VARIABLE SPEED LIMIT CONTROL TO INCREASE DISCHARGE RATES OF FREEWAY INCIDENT BOTTLENECKS

Danjue Chen
Research Associate
Department of Civil and Environmental Engineering
University of Wisconsin – Madison
1415 Engineering Drive
Madison, WI 53706
Fax: 608-262-5199 (dept.)
Email: dchen24@wisc.edu

Soyoung Ahn (Corresponding Author)
Associate Professor
Department of Civil and Environmental Engineering
University of Wisconsin – Madison
1415 Engineering Drive
Madison, WI 53706
Phone: 608-265-9067
Fax: 608-262-5199 (dept.)
Email: ahn37@wisc.edu
ABSTRACT
Variable speed limit (VSL) schemes are developed based on the Kinematic Wave theory to increase discharge rates at freeway incident bottlenecks while smoothing speed transition. The main control principle is to restrict upstream demand (in free-flow) progressively to achieve three important objectives: (i) to provide gradual speed transition at the tail of an incident-induced queue, (ii) to clear the queue around the bottleneck, and (iii) to discharge traffic at the stable maximum flow that can be sustained at the incident bottleneck without breakdown. These control objectives are accomplished without imposing overly restrictive speed limits. We further provide remedies for the case of a re-emerging queue at the bottleneck due to an overestimated stable maximum flow. The results from a parameter analysis suggest that significant delay savings can be realized with the VSL control strategies.

Key words: incidents, bottlenecks, variable speed limit, capacity drop
INTRODUCTION
Variable speed limit (VSL) control seeks to improve safety by smoothing out shock waves at the tail end of a queue and freeway efficiency by deferring the onset of congestion or increasing the bottleneck (BN) discharge rate. Earlier efforts to improve freeway efficiency via VSL focused on harmonizing the speed across vehicles in different lanes to create a more homogenous, stable one-pipe flow with few lane-changes (LC) (1, 2). This can presumably lead to higher capacity and critical density, thereby deferring or preventing onset of congestion (3). Several studies have shown that VSL control indeed induces more balanced speed and utilization of lanes (1, 2, 4-6).

SPECIALIST (SPEed ControllIng Algorithm using Shock wave Theory) seeks to proactively resolve a moving jam and maximize the discharge rate by limiting the speed and density of the inflow to the moving jam via VSL (7, 8). The algorithm was tested in the field with reasonable success. Another notable scheme is the mainstream traffic flow control (MTFC) developed in the framework of discrete-time optimal control (9). The main objective is to control the free-flow traffic upstream of a BN (before a queue arises) via VSL or ramp metering to prevent BN activation. The algorithm was tested on a Dutch network via simulation (10). Later, local feedback control was incorporated into MTFC to further improve field implementation (11).

Chen et al. (12) developed different VSL schemes based on the Kinematic Wave theory (13, 14) to increase freeway BN discharge rates and manage the queue upstream (for smoother speed transition) under two scenarios: steady queue and oscillatory queue that can inevitably arise at fixed BNs. The key principle is to impose VSL control some distance upstream of a BN to starve the inflow to the BN and dissipate the queue. Once the queue near the BN vanishes, another less restrictive VSL is imposed upstream to (i) resolve the heavy queue generated by the first VSL and (ii) regulate the inflow to sustain the stable maximum BN discharge rate and prevent BN re-activation.

The strategies cited above were designed primarily to address recurrent BNs or moving jams, in which a reduction in discharge rate typically ranges from 5 to 15% (15, 16). These strategies, however, are not suitable for non-recurrent bottlenecks such as incident BNs, in which discharge rates can reduce by more than 15%. Incident BNs are usually characterized by moderate to severe congestion due to significant reductions in system throughput (in greater proportion than lanes blocked (17) and sharp transition upstream from free-flow traffic to the queue, which may cause secondary incidents. A significant reduction in system throughput is attributable to: (1) a decrease in capacity due to road blockage, (2) rubbernecking around the incident, (3) change in driver characteristics (17), and (4) disruptive LCs away from incident location. In regard to (4), speed and flow may vary significantly among lanes with lower speed and flow closer to the incident location. LCs away from the incident location are likely to create voids in other lanes and reduce the discharge rate similar to the capacity drop phenomenon of recurrent BNs (18).

In this study, we develop VSL strategies to proactively improve the discharge rate of incident BNs and manage the upstream queue for smoother speed transition. An increase in discharge rate is achieved by clearing the queue around the incident and then maintaining a stable
maximum flow with harmonized speed to minimize disruptive LCs. Note that other
aforementioned factors for discharge rate reductions ((1)-(3) in the previous paragraph) are not
within the scope of this paper because VSL control may be not the best option for these issues.
The strategies developed in this paper are based on the KW theory and use the logic similar to
Chen et al. (12). However, the new strategies address more effectively several critical issues for
incident BNs: (a) a restrictive speed limit (lower than the speed in queue as in Chen et al. (12)
should be avoided because incident-induced congestion is likely more severe; (b) the upstream
queue management should be more elaborate to provide smoother queue transition; and (c) it
may not be straightforward to precisely estimate the stable maximum discharge rate, which can
lead to another queue formation at the incident BN.

This study develops a theoretical framework for VSL control to improve the performance of
incident BNs. This is an important contribution given that existing strategies were primarily
designed to mitigate recurrent BNs. The VSL control strategies developed in this study are
relatively simple, yet capable of addressing two critical issues for incident BNs: (i) reducing
incident-induced total delays significantly and (ii) providing smoother speed transition for better
safety. The theoretical approach provides insights into traffic dynamics with VSL control and
the impact of control parameters on the system performance (e.g., delay savings). Moreover, it
provides a foundational framework to address more complex freeway networks and incorporate
various implementation issues (e.g., detection and control technologies).

The remaining manuscript is organized as follows. Section 2 describes the basic VSL control
strategy including the analysis of parameters on the system performance and sensitivity. Two
sequel VSL strategies are developed in Section 3 to remedy a re-emerging queue at the incident
BN due to an over-estimated stable maximum flow. Concluding remarks are provided in Section
4.

BASIC VSL CONTROL FOR INCIDENT BN

Baseline Case
We study freeway bottlenecks due to incidents that may partially block the roadways and reduce
the throughputs, as shown by Fig. 1(a). We assume that the traffic evolution can be well
approximated by the KW model with triangular fundamental diagrams (FD). The upper FD in
Fig. 1(b) describes traffic states upstream of an incident with free-flow speed $u$, wave speed $w$
and jam density $k_j$; and the lower FD describes traffic states at the incident location with lower
free-flow speed $u^{inc}$ and jam density $k_j^{inc}$. Note that we assume $u^{inc} < u$ due to rubbernecking
and other effects induced by the incident.

We assume that traffic demand is constant in state $A$, and traffic breaks down to state $H$ after the
incident. After the incident is cleared, traffic recovers the full, normal capacity of $q_{max}$ in state
$M$. State $e$ represents the stable maximum flow state that can be sustained at the incident BN
without breakdown for an extended period; i.e., BN capacity, $q_{BN} = q_e$. (The notation, $q_-$,
represents flow in the traffic state denoted in subscript.) Note that $q_H < q_{BN}$ due to LC.
disruptions; and $q_{BN} - q_H$ represents the potential gain of system throughput. States $G$ and $E$ correspond respectively to the free-flow and congested states with the same flow as $q_{BN}$ (i.e., $q_{BN} = q_E = q_G = q_e$) upstream of the incident.

The spatiotemporal traffic evolution without any control is illustrated in Fig. 1(c). After the incident, a heavy queue (in state $H$) propagates upstream, forming a shock wave, $s_{AH}$. (Hereafter, $s_-$ refers to a shock wave delineating two different traffic states denoted in subscript and represents shock wave speed when used in equations.) When the incident is cleared at time $T_M$, the normal capacity recovers and traffic evolves to state $M$, with the transition marked by $s_{HM}$. The queue ends when $s_{AH}$ collides with $s_{HM}$, at which state $A$ is resumed; this demand recovery is marked by $s_{AM}$. The queue ending time, $t_{end}^{base}$, and the total delay resulting from the incident, $W^{base}$, can be easily derived from the queuing diagram in Fig. 1(d), in which $A(t)$ and $D(t)$ denote the virtual arrival and departure curves at the bottleneck, respectively:

$$t_{end}^{base} = \frac{(q_M - q_H) T_M}{q_M - q_A},$$

$$W^{base} = \int (A(t) - D(t)) dt = \frac{1}{2} (q_A - q_H) T_M t_{end}^{base} = \frac{(q_A - q_H)(q_M - q_H) T_M^2}{2(q_M - q_A)}.$$  

Notably, if the speed in queued state $H$ is low, which is quite likely with an incident, the speed drop along $s_{AH}$ would be abrupt. In this case, it would be undesirable to impose a restrictive speed limit ($< v_H$) to clear the queue around the BN as prescribed by Chen et al. (12).

Additionally, the upstream queue management strategies may not be sufficient to provide smooth enough transition at the queue’s tail. Below we introduce a new VSL control strategy to address these problems.

![FIGURE 1 Traffic evolution at incident BN.](image-url)
Basic VSL Control Strategy
The control principle is simple: restrict the upstream demand (in free-flow) progressively. When done strategically, this will achieve three important objectives together: (i) to induce gradual transition at the queue’s tail, (ii) to clear the queue around the BN, and (iii) to discharge traffic at stable maximum flow without breakdown. The control procedure consists of the following steps. (All traffic states resulting from the VSL control are shown on the FDs in Fig. 2(a), and the corresponding traffic evolution is shown in the time-space diagram in Fig. 2(b).)

**Step 1:** This main step is to control the upstream demand and clear the queue around the BN.

Set several intermediate speed limit values between \( u \) and \( v_E \) with an even increment. (The notation, \( v_- \), represents speed in the traffic state denoted in subscript.) The number of speed limit values will depend on the difference between \( u \) and \( v_E \) and the increment deemed acceptable to drivers. For an illustration purpose, we assume three intermediate values, denoted as \( V_1, V_2, \) and \( V_3 \).

**Step 1-1:** At the start of control \( (t_0) \), impose \( V_1 \) simultaneously over an extended segment immediately upstream of the queue; see Fig. 2(b). The length of the segment will be discussed later. This results in a zone with state \( \tilde{a}_1 \) (zone ‘1’ in the figure), which has the same density as \( A \), but the speed equals to \( V_1 \); see the FD in Fig. 2(a). Since the control is imposed simultaneously in space, the transition between \( A \) and \( \tilde{a}_1 \) forms a vertical shock, \( s_{A\tilde{a}_1} \). At the downstream end of the control, state \( \tilde{a}_1 \) meets state \( H \) and forms \( s_{\tilde{a}_1H} \). Notice that with this control, the queue propagates more slowly; i.e., \( |s_{\tilde{a}_1H}| < |s_{AH}| \).

**Step 1-2:** Switch the speed limit from \( V_1 \) to \( V_2 \) at time \( T_{V2} \) to create zone 2. Again, the new speed limit is actuated simultaneously in space immediately upstream of the queue. Similar to zone 1, zone 2 is in state \( \tilde{a}_2 \) with the same density as \( A \) but with speed \( V_2 \) and forms \( s_{\tilde{a}_2H} \) where it meets state \( H \). Notice that the queue propagates even more slowly; i.e., \( |s_{\tilde{a}_2H}| < |s_{\tilde{a}_1H}| < |s_{AH}| \).

**Step 1-3:** Similar to Step 1-2, but change speed limit from \( V_2 \) to \( V_3 \) at time \( T_{V3} \) to create zone 3.

**Step 1-4:** Similar to Step 1-2, but change speed limit from \( V_3 \) to \( v_E \) at time \( T_{v_E} \) to create zone 4 in state \( \tilde{a}_E \). Finally, the queue moves forward since \( s_{\tilde{a}_EH} > 0 \) and is resolved \( s_{\tilde{a}_EH} \) arrives at the BN at \( T_{\tilde{a}_E} \).

**Step 2:** This is to discharge traffic at the stable maximum flow without breakdown. The main idea is to have upstream traffic (in state \( A \)) gradually evolve to state \( E \), and then let them fully accelerate to state \( G \) before passing the BN.

Create an acceleration zone immediately upstream of the BN so that traffic can accelerate to \( u \). The length of this zone, \( L \), is set to be 0.35-1 km as in Chen et al. (12).
Step 2-1: When $H$ is resolved at $\tau_0$, de-activate $v_E$ at the rate of the maximum backward moving wave speed, $w$, to guide the traffic in state $\alpha_E$ to accelerate to state $\alpha'_E$ until the shock reaches the entrance of the acceleration zone; see Fig. 2(b). Note that $q_{\alpha'_E}$ may be smaller than $q_{BN}$, resulting in under-utilization of the BN capacity. Fortunately, this period is short ($= \left( \frac{L}{u} - \frac{L}{w} \right)$) and negligible compared to the incident duration. For example, it is about 4 minutes when $u = 100$ km/h, $w = -18$ km/h, and $L = 0.75$ km. Nevertheless, one possible remedy is to increase the de-activation rate to achieve a higher discharge rate than $q_{\alpha'_E}$, which may be possible given the low density in $\alpha_E$ (relative to the FD). For simplicity when calculating the delay saving, we assume $q_{\alpha'_E} = q_H$, which corresponds to a de-activation rate, $w' = \frac{q_H - (k_A + q_A/w)v_E}{q_H - (k_A + q_A/w)}$.

The trajectory of the first vehicle that crosses the whole acceleration zone at $u$ is denoted by the connected red arrows in Fig. 2(b), referred to as the first vehicle trajectory (FVT). The FVT serves as the boundary for the second set of VSL in the next step and is used to back-calculate the lengths of the vertical control at $v_E$, $v_3$, $v_2$, and $v_1$, sequentially, as pictured in Fig. 2(b).

Step 2-2: Starting at the position of the FVT at $t_0$, impose $V_1$ at the rate of $s_{AA_1}$ ($s_{AA_1} = s_{AE}$). At the same rate, impose $V_2$, $V_3$, and $v_E$ when the FVT intersects $T_{V_2}$, $T_{V_3}$, and $T_{v_E}$, respectively, such that all shocks are parallel. Speed limit $v_E$ is extended to the entrance of the acceleration zone, at which drivers are informed to resume free-flow speed $u$ (or $v_e$).

This step results in four zones: zones I-IV, characterized by states $A_1-A_3$ and $E$, respectively. Particularly, state $A_1$ in zone I has the same speed as state $\alpha_1$ but a higher density; see Fig 2(a). Notice that in zones I through IV, vehicles gradually decrease their speed from $u$ (free-flow speed) to $v_E$ and then maintain $v_E$ until reaching the entrance of acceleration zone. Thereafter, they resume $u$ and traffic evolves to state $G$. Notice that state $G$ eventually evolves to state $e$ at the BN without any flow change, as prescribed by the lower FD. (Recall that the lower free-flow speed at the BN is to capture rubbernecking and other effects of the incident.) In anticipation of this, traffic may alternatively be controlled to reach state $e$ in the acceleration region. Regardless, the BN starts to discharge traffic at $q_{BN}$ ($q_{BN} = q_E = q_G = q_e$) after the arrival of the FVT.

Step 2-3: When the incident is cleared at $T_M$, de-activate VSL to restore the full capacity of $q_M$. To restore $q_M$ as soon as possible, de-activation at the entrance of the acceleration zone can be timed so that the forward moving shock, $s_{GM}$, will reach the BN at $T_M$. However, the gain is only about 1 minute (the trip time of $s_{GM}$), and thus, one may safely start the de-activation at $T_M$. Concurrently, VSL is de-activated at the rate of $w$, forming $s_{EM}$, which eventually collides with $s_{A3E}$ and terminates zone IV. Notably, the trajectory of the last vehicle that experiences state $E$ is denoted by the three connected green arrows in Fig. 2(b), referred to as the Last Vehicle Trajectory (LVT). As described later, the LVT will serve as the boundary for VSL de-activation.
Step 2-4: De-activate $V_1$, $V_2$, and $V_3$ along the LVT to end the VSL control. As a result, three new zones, zones i-iii, form naturally (i.e., no VSL control is needed) upstream of the LVT; see Fig. 2(b). States $A_1$-$A_3$ evolve to congested states $A_1^m$-$A_3^m$, respectively, due to the bounded speeds, $V_1$-$V_3$ downstream. When these states interact with the upstream demand, shocks $s_{AA_1^m}$-$s_{AA_3^m}$ form, which mark diminishing queue. As traffic emerges from state $A_3^m$, it evolves to the full capacity state, $M$, forming $s_{A_3^m}M$. When it collides with $s_{AA_3^m}$, the queue completely vanishes.

(a)  

(b)  

(c)  

(d)  

---

1 On Fig. 2(b) dotted line was used to indicate that the temporal duration of state $G$ was shorten for presentation purpose to show the complete evolution. Similar way was used in Fig. 3(c) and Fig. 4(a).
FIGURE 2 Basic VSL control strategy.
\(\Delta t \in (20, 140, 260)\)
(\(\Delta t\) is in sec, \(q_A = 0.8q_M, q_E = 0.75q_M, L_q = 2\text{km}, L = 0.75\text{km}, T_M = 2\text{h}\))

Parameter Analysis

The new queuing diagram with VSL control is shown in Fig. 2(c). Notice that the BN discharge rate increases to \(q_{BN}\) at \(\tau_1\), resulting in quicker queue dissipation at \(t_{\text{end}}^{VSL}\). \(\tau_1\) can be derived based on \(u\) and \(w'\) from the time-space diagram in Fig. 2(b), and \(t_{\text{end}}^{VSL}\) can be derived from the queuing diagram in Fig. 2(c).

\[
\begin{align*}
\tau_1 &= \tau_0 + L\left(\frac{1}{u} - \frac{1}{w'}\right) = -\frac{L_q}{s_{AH}} + \Delta t_{\text{crit}} + L\left(\frac{1}{u} - \frac{1}{w'}\right), \quad (3) \\
t_{\text{end}}^{VSL} &= ((q_E - q_H)\tau_1 + (q_M - q_E)T_M)/(q_M - q_A), \quad (4)
\end{align*}
\]

where \(\Delta t_{\text{crit}}\) is the duration of queue clearance at the BN after VSL (total duration of zones 1-4), as labeled in Fig. 2(b). \(\Delta t_{\text{crit}}\) is derived based on \(L_q, s_{\alpha zH}, s_{\alpha H}, s_{\alpha i H},\) and \(s_{\alpha E H}\) and expressed as:

\[
\Delta t_{\text{crit}} = -\frac{3q_Aw\Delta t(q_E(u-w)+k_jw)u+2L_q(q_E+k_jw)(q_Hu-q_Aw+k_jw)}{2w(-q_Eq_Hu+q_Aq_Ew-q_Hk_jw)}. \quad (5)
\]

There are several factors that affect \(\Delta t_{\text{crit}}\) and thus \(\tau_1\), namely the durations of zones 1-3, number of intermediate VSL values, and the most restrictive VSL, \(v_E\). For simplicity, we assume that the durations of zones 1-3 all equal to \(\Delta t\). Clearly, the sooner we start \(v_E\) (i.e., \(\Delta t\) is smaller), the sooner we can clear the queue at the incident BN. Fig. 2(d) illustrates the impact, a linear positive relationship between \(\Delta t_{\text{crit}}\) and \(\Delta t\), which increases faster when the congestion becomes more severe (i.e., \(q_H\) decreases as the capacity drop \((1 - \frac{q_H}{q_M})\) increases). Note that \(\Delta t\) should be sufficiently long so that drivers can adapt and transition smoothly, which indicates the trade-off between the delay saving and speed transition. In the field implementation of Hegyi et al. (19), such intermediate speed was set to last 20 seconds. Also notice that the number of intermediate VSL values between \(u\) and \(v_E\) can affect \(\Delta t_{\text{crit}}\): given \(\Delta t\), it will take longer to start \(v_E\) with more intermediate VSL values. This again highlights the trade-off between the delay saving and speed transition. Finally, the effect of \(v_E\) on \(\Delta t_{\text{crit}}\) is intuitive: increasing \(v_E\) increases \(\Delta t_{\text{crit}}\) and \(\tau_1\) since VSL becomes less restrictive. Conversely, increasing \(v_E\) can have a positive effect on \(t_{\text{end}}^{VSL}\). As evident from Fig. 2(c), the BN would discharge traffic at a higher rate with higher \(v_E\) (recall that \(q_E = q_{BN}\)), increasing the rate of queue dissipation. This trade-off will be investigated in more detail shortly.
The delay saving as a result of VSL, $\Delta W_{VSL}^0$, equals to the area of the shaded region in Fig. 2(c).

$$\Delta W_{VSL}^0 = 0.5 \cdot (q_{BN} - q_H)(T_M - \tau_1)((t_{end}^{base} - \tau_1) + t_{end}^{VSL} - T_M).$$

Fig. 2(e) illustrates the impact of $q_{BN}$ on $\Delta W_{VSL}^0$. Note that they are respectively expressed as the fractions relative to the road capacity and baseline total delay (i.e., $q_{BN}/q_M$ vs. $\Delta W_{VSL}^0/W^{base}$) to better examine the relative impact. It is interesting to note, however, that in severe congestion (e.g., $q_H < 0.5q_M$), the fractional delay saving increases with $q_{BN}/q_M$ at decreasing rates up to a certain point and then decreases markedly. This trend is attributable to setting the most restrictive VSL at $v_E$ that varies with $q_{BN}$, resulting in the trade-off between $\tau_1$ and the rate of queue dissipation thereafter, as mentioned in the previous paragraph. More specifically, increasing $v_E$ and $q_E$ results in faster increases in $\tau_1$ and thus decreases in the fractional delay saving. Also notice that the impact of $\Delta t$ is quite significant. For example, when $\Delta t$ increases from 20 seconds to 140 seconds, the delay saving can drop by more than 10%.

**VSL CONTROL WITH RE-EMERGENCE OF QUEUE AT INCIDENT BN**

Uncertainty in estimation of the stable maximum flow, $q_{BN}$, is a valid concern due to the non-recurrent, wide-varying nature of incidents. Moreover, difficulty in predicting rubbernecking behavior adds to the challenge. In this section, we develop a VSL control strategy for the case of a re-emerging queue at the incident BN, as a result of an over-estimated $q_{BN}$.

We assume that the actual stable maximum flow, $q_{BN^*}$, is smaller than $q_{BN}$ (and thus $q_e$, $q_E$ and $q_G$); see Fig. 3(a). As a result, a queue forms again at the BN at $\tau_1$ (when the BN is supposed to start discharging at $q_{BN}$)\(^2\), and traffic reverts to state $H$. This new queue propagates as $s_{GH}$, as depicted in Fig. 3(b). If no action is taken, the new queue would continue to travel upstream, albeit more slowly than the initial queue ($s_{GH} < s_{AH}$), and eventually terminate transition zones I-IV. Moreover, the BN would discharge at $q_H$, meaning little to no savings in total delay.

To remedy this problem, we impose new VSL control to clear the newly formed queue and then adjust the discharge flow. We propose two strategies, A and B, with different requirements of implementation.

**Strategy A**

This is a simpler strategy with two steps, aiming at achieving $q_{BN^*}$ as soon as possible. (All traffic states resulting from strategy A are shown on the FDs in Fig. 3(a), and the corresponding traffic evolution is shown in the time-space diagram in Fig. 3(b).)

*Step A-1:* When a new queue is confirmed (at $T_2$), impose new VSL control with speed limit $v_H$ simultaneously over a segment immediately upstream of the acceleration zone to clear the queue; see Fig. 3(b). By flow conservation, a new state, $\tilde{E}$, is created in zone V with the same density as state $E$. The transition between $E$ and $\tilde{E}$ forms a vertical shock, $s_{EE}$. State $\tilde{E}$ evolves to state $I$

\(^{2}\) It is possible that a queue may form earlier if $q_{AE}$ is sufficiently high. Note, however, that a VSL strategy to remedy this case would be similar.
\( q_I = q_E \) in the acceleration zone and then resolves the queue, forming a forward moving shock, \( s_{IH} \). The arrival of \( s_{IH} \) at the BN marks the clearance of the new queue, which indicates that we can resume \( q_{BN^*} \) at the BN. The spatial extent of this VSL control at \( T_2 \), denoted by \( L_{\tilde{E}} \), should be designed so that state \( I \) does not persist at the BN (as pictured).

Upstream of zone V, state \( E \) naturally evolves to state \( H \) (zone VI) and forms \( s_{EH} \). Note that state \( H \) in zone VI is bounded by \( v_H \) due to state \( \tilde{E} \) downstream. Notably, \( s_{EH} \) is terminated as it travels upstream and collides with \( s_{A3E} \). Thereafter, state \( A_3 \) (zone III) interacts with \( H \), forming \( s_{A3H} \). Similar transitions occur to \( A_2 \) (zone II) and \( A_1 \) (zone I) sequentially. In this process, the four transition zones (I-IV) are terminated by the queue in \( H \) gradually, and the speed transition between the intermediate states and \( H \) becomes gradually more abrupt. Eventually, state \( H \) interacts with state \( A \) directly.

**Step A-2:** When the queue at the BN is cleared, impose a less restrictive VSL, \( v_{E^*} \), upstream of the acceleration zone to regulate the BN discharge rate at \( q_{E^*} (= q_{BN^*}) \). Upstream of the acceleration zone, traffic evolves from state \( H \) to \( E^* \), forming \( s_{HE^*} \), and then to state \( G^* \) in the acceleration zone. When the incident is cleared, this VSL control should be de-activated similar to Step 2-3.

State \( E^* \) propagates upstream until it is finally terminated by free-flow traffic \( A \), marked by \( s_{AE^*} \). After the VSL control is de-activated, traffic emerging from state \( E^* \) evolves to state \( M \) and resumes its full capacity.

Notably, this strategy is able to resolve the new queue and attain a higher BN discharge flow than \( q_H \), yet at the expense of sharper speed transition upstream. Notably, transitions from \( A \) to \( E^* \) and from \( A \) to \( H \) are likely very abrupt in the absence of transition layers. Therefore, we introduce strategy B to overcome this problem.
Strategy B
This strategy is built on Strategy A, but we add additional control upstream for smoother speed transition. Detailed steps follow. (All traffic states resulting from strategy B are shown on the FDs in Fig. 4(a), and the corresponding traffic evolution is shown in Fig. 4(b).)

Step B-1: Same as Step A-1.

Step B-2: Same as Step A-2.

Step B-3: Notice in strategy A (Fig. 3(b)) that upstream speed transition is compromised because the heavy queue in state $H$ (zone VI) becomes widespread over time. In this step, the heavy queue will be contained to a shorter distance and resolved sooner by imposing VSL control on states $A$ and $A_1$-$A_3$ (transition zones I-III). Specifically, when the new queue is confirmed at $T_2$, impose $V_1$ immediately upstream of state $A_1$ (zone I) over the same spatial extent as $A_1$.

Following $V_1$, impose $V_2$, $V_3$ and $v_{E^*}$ sequentially with time increment of $\Delta t$; see Fig. 4(a)-(b).

This creates low density and low flow states relative to the FD, $\tilde{a}_1$-$\tilde{a}_3$ and $\tilde{a}_E^*$. The transition between $\tilde{a}_E^*$ and $H$ forms a forward shock, $s_{\tilde{a}_E^*-H}$, which will contain and terminate state $H$ when colliding with $s_{H^*}$. The spatial extents of $V_1$-$V_3$ and $v_{E^*}$ at the start should be determined such that $\tilde{a}_E^*$ is terminated simultaneously with shock $s_{H^*}$.

Note that state $E$ may be terminated by state $H$ early if $s_{A_3E}$ and $s_{EH}$ collide before $v_{E^*}$ is supposed to go in effect, as depicted in Fig. 4(b). In this case, speed limit $v_{E^*}$ is turned on sooner than $\Delta t$, more precisely, when $s_{A_3\tilde{a}_3}$ collides with $s_{A_3H}$. By contrast, if state $E$ still remains, speed limit $v_{E^*}$ is imposed as planned (i.e., $\Delta t$ later). This would create a void (state $O$) between $\tilde{a}_E^*$ and $E$ (and later between $\tilde{a}_E^*$ and $H$) because traffic in $E$ travels faster; see Fig. 4(c).

---

3 This is because the vehicle trajectory passing these three zones is parallel to those passing zones 1-3, such as the FVT.
which zooms in the region upstream of zone VI in Fig. 4(b) to illustrate more detailed traffic evolution. If the speed difference between \( v_{E^*} \) and \( v_E \) is small, the void will eventually be resolved by \( a_{E^*} \). Thereafter, state \( a_{E^*} \) proceeds to resolve zone VI, marked by \( s_{a_{E^*H}} \). However, if the speed difference is significant, the void may pass through zone VI and proceed to the BN, resulting in a discharge flow rate of 0, which is highly undesired; see Fig. 4(d). For this case, we can delay the upstream control (Step A-1) until \( T_2' \), so that the three shocks, \( s_{EH}, s_{A_3E^*}, \) and \( s_{A_3a_{ij}} \), converge; see Fig. 4(e). With this setting, we can actuate speed limit \( v_{E^*} \) following \( V_3 \).

**Step B-4**: Upstream of \( \overline{a_1-a_3} \) and \( \overline{a_{E^*}} \), impose another set of \( V_1-V_3 \) and \( v_{E^*} \) at the rate of \( s_{AE^*} \) to transition traffic gradually from \( A \) to \( E^* \) (\( A \rightarrow A_{1*} \rightarrow A_{2*} \rightarrow A_{3*} \rightarrow E^* \)). Finally, de-activate the control when the incident is cleared, similar to Step 2.
FIGURE 4 Strategy A to address queue re-emergence.

Parameter Analysis
As expected, the delay saving would decrease if a queue re-emerges at the incident BN; see Fig. 5(a) for the queuing diagram. Note, however, that strategies A and B have the same delay savings because the additional VSL control in Strategy B only manages the queue transition while leaving the BN discharge flow the same as in Strategy A. The delay saving (compared to the base case) is $\Delta W_{VSL}^R$ and can be derived based on the queuing diagram:

$$\Delta W_{VSL}^R = 0.5(q_{BN*} - q_H)(T_M - \tau_1^R)((t_{end}^{base} - \tau_1^R) + t_{end}^R - T_M).$$  \hspace{1cm} (8)

$\tau_1^R$ represents the time when $q_{BN*}$ is resumed and equals to $\tau_1 + \Delta t_{cle}$, where $\Delta t_{cle}$ is the new queue clearance time; see the labels in Fig. 3(b).

(a) (b) (c)

FIGURE 5 Parameter analysis of Strategies A and B.
(In (b-c), $q_A = 0.8q_M, q_H = 0.5q_M, L_q = 2km, L = 0.75km, T_M = 2h$)
Notice that $\Delta t_{clc}$ depends on how responsive the system is. We assume that the new queue is detected after some buffer time, $\Delta t_{buf}$, or when the queue reaches the entrance of the acceleration zone, whichever happens first. Then, we find that $\Delta t_{clc}$ is given by:

$$\Delta t_{clc} = \min \left( -\frac{k_j \omega (L + u \Delta t_{buf})}{q_H u}, \frac{L(q_H(u-w)+u_k \omega)k_j}{(q_E-q_H)q_H u} \right),$$

(9)

in which the former expression corresponds to early detection (i.e., after $\Delta t_{buf}$) and the latter for detection at the entrance. Notice that in the former case, a linear positive relationship exists between $\Delta t_{clc}$ and $\Delta t_{buf}$. This is expected because the earlier we detect the queue, the sooner we can actuate the queue clearance process; see Fig. 5(b). In the latter case, $\Delta t_{clc}$ depends on $q_E$, the initial estimation of the stable maximum flow. The higher the $q_E$ is, the smaller the $\Delta t_{clc}$ is. This is because a higher $q_E$ (larger $|s_{GH}|$) indicates an earlier queue detection and faster queue dissipation (larger $s_{IH}$), both of which decrease $\Delta t_{clc}$. Notice that, the critical value of $\Delta t_{buf}$ when the two cases equals, denoted by $\Delta t_{buf}^{crit}$, decreases as $q_E$ increases. Notably, this indicates that with a higher $q_E$, we can detect the queue (at the entrance) earlier and thus actuate control earlier. However, it also indicates shorter time available to set up the remedy action, which could pose a different challenge.

The relationship between the fractional delay saving ($\Delta W_{VSL}/W_{base}$) and $\Delta t_{buf}$ is illustrated in Fig. 5(c). With early detection, it decreases as the buffer time increases, which is expected because $\Delta t_{clc}$ is larger. Also notice that the delay saving depends on the new BN discharge rate, $q_{BN*}$ ($q_{BN*} = q_{E*}$): obviously, a larger $q_{BN*}$ leads to a higher delay saving.

Note that in our strategies A and B, we use speed limit $v_H$ to clear the newly formed queue at the BN. This speed limit can be set to a lower value to clear the new queue faster. However, the minimum value, $v_{min}$, is bounded by the speed that yields $q_I = q_H$:

$$v_{min} = \frac{q_H \omega}{q_E + k_j \omega},$$

(10)

A speed limit lower than $v_{min}$ would result in $s_{IH} > 0$ and it would not be possible to clear the newly formed queue.

Also note that in both strategies, the duration of zone $V$ (state $\bar{E}$), $\Delta t_{\bar{E}}$, is critical for the experiment set-up, which is given by

$$\Delta t_{\bar{E}} = \Delta t_{clc} - \Delta t_{buf} - \frac{L}{u} = \min \left( \frac{(L+u \Delta t_{buf})(q_H+k_j \omega)}{q_H u}, \frac{L(q_H-k_j \omega)(q_H(u-w)+u_k \omega)}{(q_E-q_H)q_H u} \right).$$

(11-1)

Thereafter, the control distance at $T_2$, $L_{\bar{E}}$, can be derived accordingly:

$$L_{\bar{E}} = \Delta t_{\bar{E}} v_H = \min \left( \frac{(L+u \Delta t_{buf})w}{u}, \frac{L(q_H-k_j \omega)(q_H(u-w)+u_k \omega)}{(q_E-q_H)u(q_H+k_j \omega)} \right).$$

(11-2)

As revealed by Equation (11-1), in the case of early detection, $\Delta t_{\bar{E}}$ (as well as $L_{\bar{E}}$) increases with $\Delta t_{buf}$. This is because shock $s_{IH}$ has to travel longer. For the late detection, both extensions increase as $q_E$ decreases because the shock $s_{IH}$ travels more slowly.
CONCLUSIONS AND DISCUSSIONS
In this paper, variable speed limit (VSL) strategies were developed based on the Kinematic Wave theory to increase discharge rates at freeway incident bottlenecks and provide smoother speed transition upstream. Our main logic is to impose VSL control gradually on upstream demand to dissipate the queue around the incident bottleneck while inducing smoother speed transition at the queue’s tail. After the queue clearance, VSL control continues to regulate inflow to the bottleneck so that the bottleneck can discharge traffic at the stable maximum rate in free-flow without breakdown. This is accomplished without imposing overly restrictive speed limits – an important feature considering incident situations. Our findings from the parameter analysis suggest that significant delay savings can be realized with our strategy.

We further developed two sequel VSL strategies to remedy a re-emerging queue at the incident bottleneck due to an over-estimated stable maximum flow. This is a likely scenario since incidents are non-recurrent and wide-varying in nature. The first, and the simplest, strategy is designed to clear the new queue and discharge traffic at the adjusted (lower) stable maximum rate, albeit with less desirable speed transition. The other two strategies were built on the first one to better manage the upstream queue for smoother transition. The two strategies differed by the magnitude of error in estimating the stable maximum flow. Not surprisingly, we found that delay savings decreased when the stable maximum flow was over-estimated; however, they were still substantial.

Building on the theoretical framework presented in this paper, ongoing research includes formulation of discrete schemes to (i) accommodate more complex scenarios, such as time varying demand and realistic freeway networks (with on and off-ramps), (ii) shed light on spatiotemporal features of vehicle delay and speed variation, and (iii) evaluate system robustness with respect to driver compliance and traffic detection/measurement errors, delays and resolution. Finally, the proposed VSL strategies should be tested in the field and further refined to address various practical issues, such as detection of traffic states and shock waves.

ACKNOWLEDGEMENT
This research was sponsored by the National Science Foundation (Award# 1150137) and the University of Wisconsin-Madison.

REFERENCES


