New Method for Estimating Freeway Incident Congestion

H. AL-DEEK, A. GARIB, AND A. E. RADWAN

Incidents are a major cause of travel delays on urban freeways. This paper describes development and application of a new method for estimating freeway incident congestion where extensive loop and incident data are available. Using shock wave analysis, a time-space domain is determined for each incident. This is used to define the congestion boundaries of an incident and to decide whether the incident should be analyzed as isolated or as a multiple-incident case. The freeway section is divided into smaller segments, each segment containing only one mainline loop station. Traffic speed and counts at freeway mainline stations and traffic counts at on/off ramp stations upstream and downstream of the incident location are used to calculate incident delay on each segment during small time slices, then cumulative incident delay is calculated. Satisfactory results were achieved when the new method was applied to a sample of isolated and multiple-incident cases collected recently as part of the Freeway Service Patrol Evaluation Project on I-880 in California.

Freeway incident congestion is viewed as a major problem in urban travel. In the short-term incident congestion causes delay to travelers, wastage of fuel, secondary accidents, wear and tear of vehicles and roadways, and environmental pollution. In the long-term congestion adversely affects the economic competitiveness of a region. National statistics indicate that more than 60% of urban freeway congestion is related to incidents (*1*).

Efficient incident management may be achieved by implementing Intelligent Transportation Systems (ITS) technologies (e.g., improved incident detection techniques using video image processing). To evaluate the efficiency of ITS in reducing incident delays, it is necessary to develop methods that can estimate accurately the magnitude of nonrecurring congestion. This paper describes the development and application of a new method for estimating congestion using incident and loop data collected simultaneously.

OVERVIEW OF THE PROBLEM

Incidents include accidents, disabled vehicles, law enforcement and emergency vehicles, and spills. Several methods have been used to estimate congestion caused by an incident. Morales (*2*) developed an analytical method that plots the cumulative arrival and departure curves and calculates the cumulative vehicle hours of incident delay. In this method the congested time period is divided into smaller time intervals during which demand and/or capacity are assumed to be constant. This results in linear arrival and departure curves at the incident bottleneck. The method assumes that initial demand is less than capacity of the freeway section. The HCM (*3*) method uses the same approach of the Morales method with an important modification: it considers cases of incidents occurring during the peak period congestion with initial demand exceeding capacity. The Morales (*2*) and HCM (*3*) methods are widely used by practitioners and researchers to estimate incident delay.

Messer et al. (*4*) used the kinematic wave theory of Lighthill and Whitham to develop a method for predicting individual travel times on the freeway during incident conditions. They divided the time-space plane during incidents into areas representing four different traffic flow conditions: normal flow, queue flow, metered flow, and capacity flow. The boundaries of these areas were defined by linear shock waves, and the speed of each shock wave was derived assuming a linear speed-density relationship developed by Greenshields (*5*). The method was applied to four incidents that occurred on the Gulf Freeway in Houston. It was found that two-thirds of the observed travel times were within 10 percent of the predicted travel times. The linear Greenshields' speed-density model results in parabolic relations for volume-speed and volume-density plots. The major problem with using the parabolic curves is that if they do not match the actual conditions in regions upstream of the incident, then significant errors can be made in calculating the wave speeds. It would be more accurate to use empirical data in calculating wave speeds.

Wirasinghe (*6*) used shock wave theory to develop formulas for calculating individual and total delays upstream of incidents. The formulas are based on areas and densities of regions representing different traffic conditions (mainly congested and capacity regions) that are formed by shock waves in the time-space plot.

Chow (*7*) compared two methods for calculating total incident delay on a highway section: shock wave analysis and queuing analysis. He assumed a *unique flow-density relationship* and derived the equations of total delay, which were found to be identical for both methods. Chow (*7*) concluded that if he had used a time-dependent flow-density relationship, which is more realistic, then the two methods would yield different results.

Wicks and Lieberman (*8*) developed INTRAS, a microscopic freeway traffic simulation model designed for freeway corridor simulations. An enhanced version of INTRAS called FRESIM (*9*) is currently under testing. Both INTRAS and FRESIM have the same fundamental structure, which is based on car-following theory. The most recent calibration of INTRAS by Cheu et al. (*10*) used data from Los Angeles freeways. They concluded that INTRAS may underestimate the occupancy during free flow conditions in the recovery periods after incidents. In addition, they indicated that the car-following theory equation used in INTRAS gave satisfactory results in general, but it failed to produce high volume and occupancy that match collected field values in their study site. INTRAS and its successor, FRESIM, can be used to estimate incident congestion by simulating the freeway with and without the incident and finding the difference in vehicle hours of travel.

Department of Civil and Environmental Engineering, University of Central Florida, P.O. Box 162450, Orlando, Fla. 32816.

The main limitations in the existing methods for estimating incident congestion are summarized as follows. The assumption of static demand is clearly unrealistic under peak hour conditions because it ignores the effects of traffic diversion from the freeway to alternate routes and/or traffic avoiding the freeway system if informed ahead about the incident. Further, assuming that the initial demand level is smaller than the capacity of the freeway is not valid under peak conditions. The theoretical shock wave models assume constant densities throughout each traffic flow region, which affects the accuracy of the models. Most importantly in estimating incident congestion is that macroscopic freeway traffic models have been used to analyze/simulate one incident at a time. In real life, multiple incidents could occur simultaneously or within short periods of time on the same stretch of highway. Consequently, incident queues may merge together, and it becomes very difficult to segregate the effect of one incident from another on the magnitude of congestion. Furthermore, incident queues may merge with other queues of recurring congestion that may be present at the time when the incident occurs. In effect, it becomes very challenging to segregate incident and nonincident congestion. None of the existing methods considers these real possibilities, which puts their accuracy of estimating incident delays in question.

NEW METHOD

This section describes the development of a new method for estimating freeway incident congestion. The new method has two steps:

1. Determination of the time-space domain (or the area of influence) of an incident,
2. Calculation of delays based on speed reduction caused by the incident on freeway segments located within the time-space domain of the incident determined in Step 1. Two types of incident cases are considered: single (isolated) incidents and multiple incidents.

The two steps are described below.

Time-Space Domain of Incident

Shock wave analysis is used to determine the time-space domain of an incident. This is the area that defines the time-space boundaries of congestion caused by a specific incident as shown in Figure 1 (cases A and B). The area has dimensions ($T + D$, X), where T is

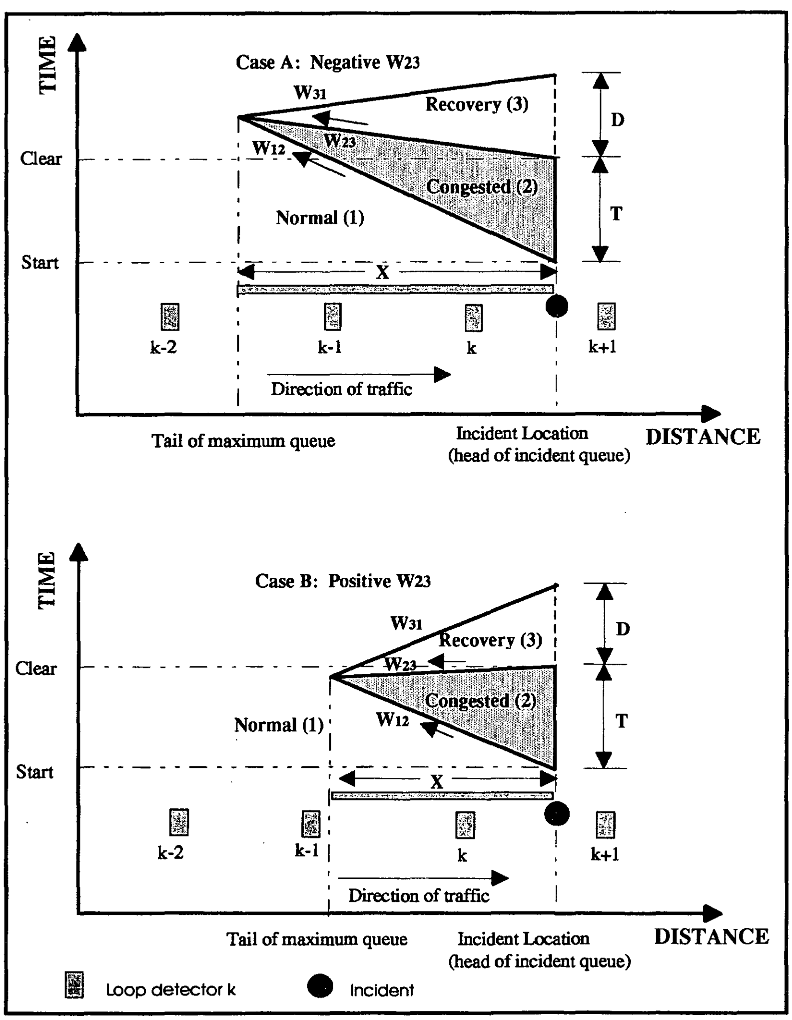

FIGURE 1 Time-space domain of incident.

the incident duration, D is the time to discharge incident queue and return to normal conditions, X is the maximum length of incident queue, and W_{ij} is the speed of shock waves forming along boundaries of traffic conditions i and j. Cases A and B of Figure 1 will be explained later. Because downstream conditions are likely to be affected by platoon dispersion, geometric bottlenecks, and other incidents, they are not considered in this analysis. We have assumed linear shock waves for simplicity. In reality shock waves may be nonlinear. Furthermore, incidents occurring upstream of the subject incident may distort the shock wave diagram by altering the queue length shown in Figure 1. Nonetheless, for the purpose of calculating the time-space domain, these effects are ignored. This is not expected to have a major impact on accuracy of congestion estimates mainly because the equations derived in this section will not be used to calculate congestion; instead, they will only be used to provide a rough approximation for the congestion boundaries of each incident. As will be explained later, incident congestion calculations are based only on reductions in loop speeds on freeway segments located within the time-space domain.

Three shock waves are shown in Figure 1 (with speeds W_{12}, W_{23}, and W_{31}) forming along boundaries of three traffic conditions (normal, congested, and recovery flow). These traffic conditions are defined as follows.

Normal Condition

This is the traffic condition at the first detector *upstream* of the incident location (i.e., detector k in Figure 1) during *the last time slice before the incident occurs,* $i - 1$. The length of time slice can be chosen arbitrarily; however, a 1-min time slice is recommended in this analysis. The traffic condition is defined in terms of three parameters: flow (F), density (K), and speed (V). The density K can be estimated by F/V, where both F and V are known from data of detector k. The three parameters will be denoted as ($F_{k,1}^{i-1}$, $K_{k,1}^{i-1}$, $V_{k,1}^{i-1}$), where the subscript 1 denotes "normal traffic condition" (*1*) and the superscript $i - 1$ refers to the last time slice before the incident occurs.

Congested Condition

Two congested traffic conditions are of interest: traffic condition at the first detector *upstream* of the incident location (i.e., detector k in Figure 1) during *time slice i, the first time slice after the incident has occurred*, and traffic condition at the same detector k during *time slice i*, the last time slice before the incident is cleared.* These two congested conditions have parameters ($F_{k,2}^{i}$, $K_{k,2}^{i}$, $V_{k,2}^{i}$) and ($F_{k,2}^{i*}$, $K_{k,2}^{i*}$, $V_{k,2}^{i*}$), respectively, where the subscript 2 denotes "Congested traffic condition" (*2*). This is indicated by the shaded area in Figure 1.

Recovery Condition

This is the traffic condition at the first detector *upstream* of the incident location (i.e., detector k in Figure 1) during *time slice* $i^* + 1$, *the first time slice after the incident has been cleared.* This traffic condition has parameters ($F_{k,3}^{i*+1}$, $K_{k,3}^{i*+1}$, $V_{k,3}^{i*+1}$), where the subscript 3 denotes "recovery condition" (*3*).

The shock wave defines the boundary separating between two traffic conditions, and its speed equals the difference in flows divided by the difference in densities of the two conditions. Hence,

$$W_{12} = \frac{F_{k,1}^{i-1} - F_{k,2}^{i}}{K_{k,1}^{i-1} - K_{k,2}^{i}}, \tag{1}$$

$$W_{23} = \frac{F_{k,2}^{i*} - F_{k,3}^{i*+1}}{K_{k,2}^{i*} - K_{k,1}^{i*+1}}, \tag{2}$$

$$W_{31} = \frac{F_{k,3}^{i*+1} - F_{k,1}^{i-1}}{K_{k,3}^{i*+1} - K_{k,1}^{i-1}}. \tag{3}$$

Using the geometry in Figure 1, it can be shown that

$$D = \frac{TW_{12}\,(W_{31} + W_{23})}{W_{31}|W_{23} - W_{12}|} \tag{4}$$

and

$$X = \frac{TW_{12}W_{23}}{|W_{23} - W_{12}|}. \tag{5}$$

Note that a wave speed equals the reciprocal of the slope in Figure 1. Also, note that W_{12} has always a negative sign (negative slope) because it is a backward-forming shock wave, whereas W_{31} always has a positive sign because it is a forward-recovery shock wave. But W_{23} can have either a negative sign (Figure 1, case A) or a positive sign (Figure 1, case B). Equation 4 applies if W_{23} is negative, but if W_{23} is positive then Equation 4 should be modified to

$$D = \frac{TW_{12}\,(W_{23} - W_{31})}{W_{31}(W_{23} + W_{12})} \tag{4'}$$

and Equation 5 should be modified to

$$X = \frac{TW_{12}W_{23}}{(W_{23} + W_{12})} \tag{5'}$$

Finally, it is important to mention that Equations 4, 5, 4′, and 5′ use only the absolute magnitude of W_{12}, W_{23}, and W_{31} to calculate D and X.

Search Procedure

The forming (W_{12}) and recovery (W_{23}) waves meet and define the congested region (*2*). Because both waves have to meet, otherwise the queue never diminishes, the following conditions apply:

$$W_{12} < 0,$$
$$W_{31} > 0, \tag{6}$$
$$\text{If } W_{23} < 0, \text{ then } |W_{12}| < |W_{23}|.$$

If the data of detector k during time slices $i - 1$, i, i^*, and $i^* + 1$ violate one or more of these conditions, then the actual incident start and/or end times may be slightly different from what has been observed in the field and, therefore, these times should be adjusted.

More specifically, the incident may have been detected a few minutes after it occurred; also, in some cases, an incident may have no effect on speed for several minutes before the reported clearance time. For example, in the I-880 incident database observers drove tach cars on the freeway continuously during 3 hours in the morning (6:30 a.m. to 9:30 a.m.) and 3 hours in the afternoon (3:30 p.m. to 6:30 p.m.) to report start and end times of incidents (*11*). The average headway of the tach cars was 7.5 minutes. Because loops normally are spaced at 0.54 to 0.81 km (0.33–0.50 mi) in most freeway systems, the incident exact location may be as far as 0.81 km (0.50 mi) from the nearest upstream loop detector. This means that the incident effect will take some time after the incident occurs until it reflects in the data of the upstream loop. Further, it is possible in incident cases where the demand level is low that after moving the incident vehicle to the shoulder and emergency vehicles leave the scene, the effect of the vehicle's presence on loop speed is negligible. Add to this the uncertainty in incident duration that has been addressed by researchers as a real concern for inaccuracy in modeling incident congestion (*4*). In an attempt to overcome these problems, we have developed a heuristic search procedure to find the adjusted incident start and end times from loop data for the purpose of obtaining accurate wave speeds that satisfy the conditions listed in condition 6 above. The search procedure is applied to speed and flow data of the nearest detector upstream of the incident and within a few minutes before and a few minutes after the observed start and end times of the incident. It is suggested that one uses 1-min time slices in this procedure. The procedure is described in the following steps:

Step 1. Identify the observed start and end times of the incident and the nearest loop detector station upstream of the incident location from field data.

Step 2. Calculate the wave speed W_{12} according to Equation 1 and using speed flow data of detector k during the first minute before and during the first minute after the incident observed start time. The wave speed W_{12} should always be negative; if not, then the incident start time needs to be adjusted and Step 2 will be repeated with the incident start time being shifted downward by 1 min. This process of downward shifting continues until a negative W_{12} is achieved. It is recommended, however, that the maximum downward shift not exceed 3 min. The 3-min threshold is about half of the tach car headway in the FSP study (*11*). A different (maybe larger) value for the maximum threshold should be used if the incident data are collected via surveillance systems (e.g., CCTV). If the 3-min downward shift does not achieve a negative W_{12}, then the incident start time is adjusted by an upward shift (in steps of 1 min to a maximum of 3 min). The time slice that produces a negative W_{12} is considered to be the adjusted incident start time.

Step 3. Update the incident clearance time by adding the incident observed duration to the adjusted start time found in Step 2 above. Calculate the wave speed W_{23} according to Equation 2 and using speed flow data of detector k during the first minute before and during the first minute after the incident updated clearance time. If the calculated W_{23} is positive, or if it is negative and satisfies condition 6 above, then the calculated W_{23} is used and the clearance time is not adjusted. Otherwise, Step 3 is repeated with downward and/or upward shifts to the observed incident clearance time until condition 6 is achieved. The time slice that satisfies this condition is considered to be the adjusted clearance time of the incident.

Step 4. Calculate the wave speed W_{31} according to Equation 3 and using speed flow data of detector k during the first minute before and during the first minute after the incident adjusted start and clearance times found in Steps 2 and 3, respectively. The calculated W_{31} should be positive; if not, then the above Steps 2 and 3 should be repeated with more shifts to the observed start and/or end times of the incident until W_{31} is positive and the conditions in 6 are satisfied.

The purpose of the above shock wave analysis is to determine where to stop (at what loop detector?) and when to stop (at what time slice?) calculating delays on freeway segments.

Calculation of Delays

Single (Isolated) Incident Delay

Nonrecurrent congestion on a freeway section can be caused by one or more incidents. Isolated incidents are cases where nonrecurrent congestion is caused by only one incident. The default is to analyze all incidents as isolated cases. However, multiple incident cases are possible whenever one or more incidents occur within the time-space domain of another incident. As will be explained later, a special algorithm has been developed to deal with multiple incident cases and to separate the congestion caused by each incident.

The freeway section under study must be divided into smaller segments. Each segment should include no more than *one* mainline station of loop detectors. The segment ends are determined by the midpoint between detectors, if there are no ramps, or by the ramp termination point when ramps exist. If all loops in a station are not working (not producing valid data), then the segment to which the station belongs is eliminated by allocating half of its length to each of the segments directly upstream and downstream of it. The purpose of freeway segmentation is to maximize use of all available loop data. The time period of interest is divided into smaller time intervals (time slices), and delay is calculated for each time slice on each freeway segment. The time-space domain of an incident is divided into a certain number of freeway segments and time slices, then delay is calculated for each time slice on each freeway segment. To estimate delay caused by an isolated incident, the following assumptions are made:

- Traffic speed and volume data are determined from the loop station on the segment, and these data are homogenous throughout the segment.
- The incident delay is calculated with respect to a reference (or base) average speed that reflects normal conditions that may or may not be congested. The reference speed represents a historical speed profile that may be used to segregate incident and nonincident congestion. The historical profile can be determined for each segment using incident free loop data as follows: for each loop detector in

the freeway section and the same travel direction, the reference speed is the average of 1-min speeds for all incident-free days in the data set studied. In other words, during each minute of the incident-free day and for each loop detector within the study section, there are two reference speeds, one for each direction of travel. One-minute speed averages are considered an appropriate level of resolution in the incident delay analysis.

Within the time-space domain of an incident, delay on each freeway segment is calculated only if the speed on the segment drops below the reference speed; otherwise, delay is null. The delay formula for each segment upstream of the incident is given by

$$
\begin{aligned}
D_k^i &= L_k \frac{\Delta T}{60} F_k^i \left(\frac{1}{V_k^i} - \frac{1}{V_k^{i,r}} \right) && \text{for } 0 < V_k^i < V_k^{i,r} \\
D_k^i &= F_k^i \left(\frac{\Delta T}{60} \right)^2 && \text{for } V_k^i = 0 \qquad (7) \\
D_k^i &= 0 && \text{for } V_k^i > V_k^{i,r}
\end{aligned}
$$

where

D_k^i = delay on freeway segment k during time slice i (vehicle-hours),
L_k = length of segment k (km),
ΔT = length of time slice i (min),
F_k^i = flow (from loops) on segment k during time slice i (vehicles/hr),
V_k^i = speed (from loops) on segment k during time slice i (km/hr), and
$V_k^{i,r}$ = reference average speed on segment k during time slice i (km/hr).

The total delay on the freeway section that is caused by the incident is given by

$$TD = \sum_{i=1}^{m} \sum_{k=1}^{n} D_k^i, \qquad (8)$$

where

TD = total incident delay on freeway segments upstream of segment k affected by the incident (vehicle-hours),
n = number of freeway segments upstream of segment k (determined by Equation 5 or 5′), and
m = number of time slices with incident congestion (determined by $T + D$, where D is found by Equation 4 or 4′).

Multiple Incident Delays

The assumptions used in single incident analysis are also applicable here.

Separating Congestion of Multiple Incidents. Suppose that two incidents (1 and 2) occur at times t_1 and t_2, respectively, at locations as shown in Figure 2. The inter-arrival time (the time gap

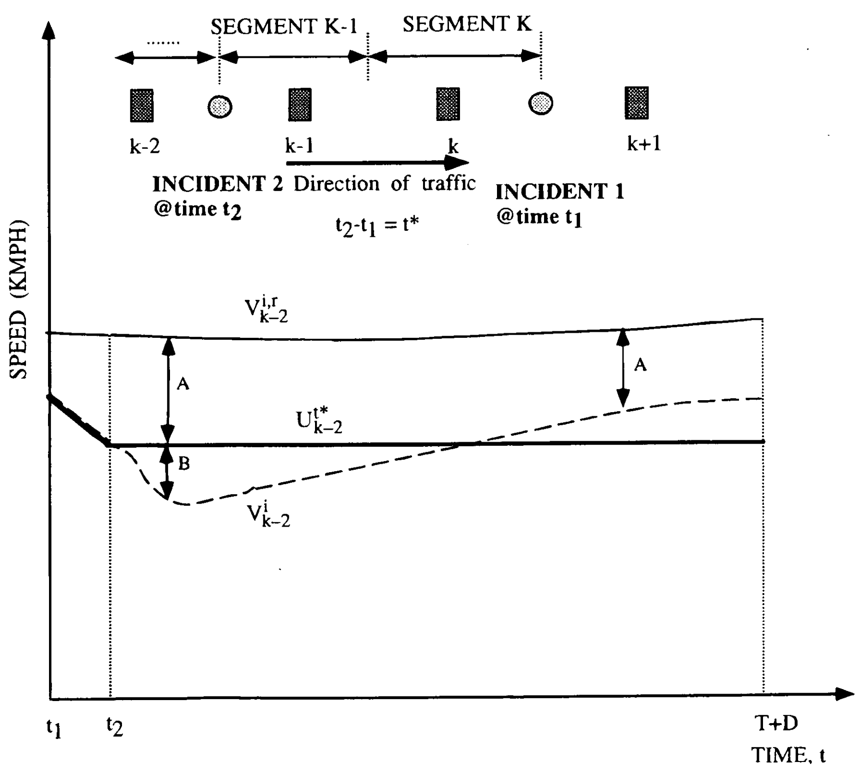

FIGURE 2 Separation of multiple incident congestion using speed profiles.

between occurrences) of the two incidents is t^*. The following illustrates an algorithm for separating the congestion caused by each incident.

Incident 1 Congestion During Time t^*. Congestion of *incident 1* during time slices before incident 2 occurs is found using Equations 7 and 8 above. Note that the two equations can be applied to find *incident 1* congestion as long as *incident 2* has not occurred, that is, these equations can be applied as long as the following condition holds:

$$m < t^*/\Delta T$$

Incident 1 Congestion on Segments Downstream of Incident 2 for Time $> t_2$. Equations 7 and 8 can be applied to find *incident 1* congestion on all segments downstream of *incident 2*. Note that the number of segments here generally will be less than n (mentioned in Equation 8 above).

Incident 1 Congestion on Segments Upstream of Incident 2 for Time $> t_2$. It will be assumed here that the congestion effect caused by *incident 1* on segments *upstream of incident 2* is captured by a drop in speed denoted by A as shown in Figure 2, where A is the difference between the reference average speed on segment $k - 2$ ($V^{i,r}_{k-2}$) and the greater of two speeds: the speed during the inter-arrival time t^* ($U^{t^*}_{k-2}$) or the actual speed during *incident 2* on segment $k - 2$ (V^{i}_{k-2}). All segments upstream of *incident 2* ($k - 2$, $k - 3$, $k - 4$, . . .) will be assessed for this type of speed drop. The delay effect of *incident 1* on segment $k - 2$ upstream of *incident 2* can be found by

$$D^{i}_{k-2} = L_{k-2}\,\frac{\Delta T}{60}\,F^{i}_{k-2}\left(\frac{1}{U^{i}_{k-2}} - \frac{1}{V^{i,r}_{k-2}}\right) \text{for } 0 < U^{i}_{k-2} < V^{i,r}_{k-2} \quad (9a)$$

$$D^{i}_{k-2} = F^{i}_{k-2}\left(\frac{\Delta T}{60}\right)^{2} \qquad \text{for } U^{i}_{k-2} = 0 \quad (9b)$$

$$D^{i}_{k-2} = 0 \qquad \text{for } U^{i}_{k-2} > V^{i,r}_{k-2}, \quad (9c)$$

where

$$U^{i}_{k-2} = \text{Maximum}\;\{U^{t^*}_{k-2}, V^{i}_{k-2}\}.$$

These equations can be replicated for segments $k - 3$, $k - 4$, and so on. Then, the total delay can be found in a way similar to Equation 8 above.

Incident 2 Congestion. *Incident 2* will be considered to have an effect only if it reduces speed below the speed level that prevails under *incident 1* conditions. It will be assumed in this analysis that the speed level under *incident 1* conditions is represented by $U^{t^*}_{k-2}$, which is the speed under *incident 1* conditions before *incident 2* occurs. This speed obviously will vary during different time slices from loop-to-loop on the freeway section studied. For *incident 2* to have an effect on congestion (in addition to the effect of *incident 1*), the speed drop below $U^{t^*}_{k-2}$, which is depicted by the difference between $U^{t^*}_{k-2}$ and V^{i}_{k-2} (shown as B in Figure 2), must be significant at the 95 percent level. The congestion effect of *incident 2* is found by

$$D^{i}_{k-2} = L_{k-2}\,\frac{\Delta T}{60}\,F^{i}_{k-2}\left(\frac{1}{V^{i}_{k-2}} - \frac{1}{U^{t^*}_{k-2}}\right) \text{ for } 0 < V^{i}_{k-2} < U^{t^*}_{k-2} \quad (10a)$$

$$D^{i}_{k-2} = F^{i}_{k-2}\left(\frac{\Delta T}{60}\right)^{2} - Z \qquad \text{for } V^{i}_{k-2} = 0 \text{ and } U^{t^*}_{k-2} > 0 \quad (10b)$$

$$D^{i}_{k-2} = 0 \qquad \text{for } V^{i}_{k-2} > U^{t^*}_{k-2} \quad (10c)$$

where

$$Z = \text{The delay, } D^{i}_{k-2}, \text{ calculated in Equation 9a above} \quad (10d)$$

These equations can be implemented into a simple spreadsheet where delay calculations can be performed for each time slice on any number of segments with incident congestion. Then delay is accumulated over all time slices and all segments affected by the incident to give the total cumulative vehicle-hours of incident delay.

INCIDENT DELAY ANALYSIS

In this section we present the results of applying the new method to cases of isolated and multiple incidents that were selected from the I-880 incident database (*11*). Incident and loop data were collected on a 11.8 km (7.3 mi) freeway section on I-880 as part of the Freeway Service Patrol Evaluation Project (FSP) in Alameda County, California. Loop stations are located approximately every 0.54 km (0.33 mi) on the study section of I-880.

The detailed results of shock wave analysis applied to an isolated incident case (incident #1456) are shown in Table 1, and delay is depicted in Figure 3. Note that the incident duration has been adjusted (it has been reduced by 4 min) using the search procedure described earlier. The calculated maximum incident queue length is 7.1 km (4.4 mi), and the duration of incident congested conditions ($T + D$) is about 55 min. It has been verified through the incident database that no other incident occurred within the time-space domain of incident #1456 (i.e., no other incident was observed along a 7.1 km (4.4 mi) freeway segment upstream of incident #1456 for a period of 55 min from the start of this incident). The estimated maximum incident queue length using loop speeds is 3.4 km (2.1 mi), which indicates that the method has overestimated the incident congestion boundaries (X and $T + D$).

Incidents #651 and #655 of the FSP database represent a multiple-incident case. Results of shock wave analysis are shown in Table 2. The methodology for separating incident delay was applied to segregate the delay for each incident case, and the delay results are shown in Figure 4. A smaller value for the maximum queue length was estimated from loop speeds (8.1 km (5 mi)), which again indicates that the method has overestimated X and, consequently, the incident congestion boundaries. Actually, it has been found that the new method overestimates the incident congestion boundaries in most of the 231 cases analyzed (see Table 3).

Table 3 shows the average and *SD* of delays for each category of incidents during morning and evening shifts. It is clear that "in-lane accidents" have the lion's share in terms of delay in this sample, whereas "right shoulder breakdowns" come second in the list (but these have the largest frequency). There are large variations in

TABLE 1 Isolated Incident Case (Incident #1456, NB I-880)

	Observed		Adjusted		Traffic Condition			
Incident # 1456	Start Time	Duration	Start Time	Duration	Normal	Congestion1[a]	Congestion2[b]	Recovery
Time(hh:mm)	5:04 PM	0:46	5:07 PM	0:42	5:06 PM	5:08 PM	5:48 PM	5:50 PM
Loop					loop11	loop11	loop11	loop11
Flow(vph)					1512	1272	1320	1266
Density(vpkm)[c]					18.5	29.9	16.4	13.7
Speed(kmph)					81.6	42.5	80.3	92.6

Incident #	Shock Wave Speed (kmph)			Time-Space Domain	
	W_{12}	W_{23}	W_{31}	D(min)	X(km)
1456	-21.1	19.5	50.7	13.4	7.1

[a]*Traffic condition at the first detector upstream of the incident location during the first time slice after the incident has occurred.*

[b]*Traffic condition at the first detector upstream of the incident location during the last time slice before the incident is cleared.*

[c]*One kilometer (km) = 0.6 mile.*

delays (e.g., *SD* is more than twice the average delay for most categories). Most studies indicated that incident duration has a large *SD*. Because the incident delay is very sensitive to incident duration, it is not surprising to see large variations in the incident delay estimates.

METHOD DISCUSSION

The following provides possible causes of over-predicting the size of the incident congestion boundaries in the new method. For simplicity, the wave speeds were based on point values of flows and densities obtained from one detector station located immediately upstream of the incident. This implied using constant and linear wave speeds over the freeway segments upstream of the incident as shown in Figure 1. In real life, wave speeds are nonlinear and dynamic. But the estimated wave speeds are likely to be large because of the sharp differences in densities immediately upstream of the incident during the one minute before and the one minute after the occurrence of the incident. Later on this difference in densities will be smaller and, consequently, the actual wave speeds will be smaller than the estimated ones. This translates into smaller congestion envelopes. That is, the congestion envelope produced by linear shock waves is expected to contain envelopes of the more realistic nonlinear waves. It can be demonstrated, using Figure 1, that for the same incident duration the maximum incident queue (X) is larger when the waves are faster than when they are slower.

Overestimation of the incident congestion boundaries does not necessarily result in overestimation of delays using the new method. The formulas for incident congestion do not use the magnitude of

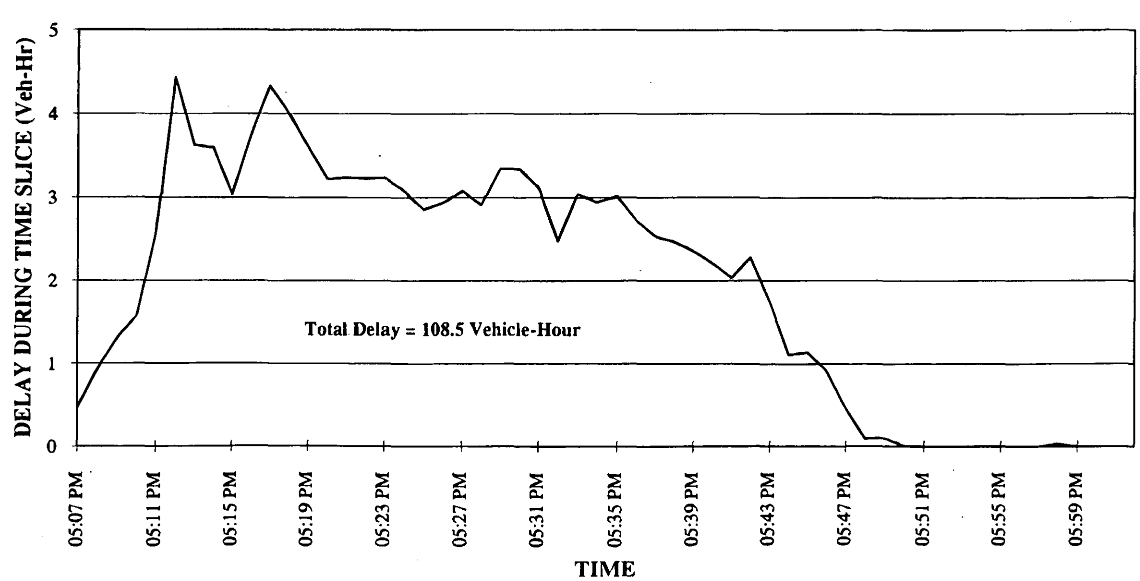

FIGURE 3 Delay for an isolated incident (#1456, NB I-880).

TABLE 2 Multiple Incidents Case (Incidents #651 and #655, SB I-880)

	Observed		Adjusted		Traffic Condition			
Incident #651	Start Time	Duration	Start Time	Duration	Normal	Congestion 1[a]	Congestion 2[b]	Recovery
Time(hh:mm)	7:04 AM	2:24	7:00 AM	2:24	6:59 AM	7:01 AM	9:23 AM	9:25 AM
Loop					loop17	loop17	loop17	loop17
Flow(vph)					1376	1264	1369	1230
Density(vpkm)[c]					14.2	28	14.4	13
Speed(kmph)					96.6	45.1	95	95

	Observed		Adjusted		Traffic Condition			
Incident #655	Start Time	Duration	Start Time	Duration	Normal	Congestion 1	Congestion 2	Recovery
Time(hh:mm)	7:31 AM	0:20	7:31 AM	0:18	7:30 AM	7:32 AM	7:48 AM	7:50 AM
Loop					loop2	loop2	loop2	loop2
Flow(vph)					1542	1374	1656	1296
Density(vpkm)					21.3	28.4	28.6	17.1
Speed(kmph)					72.5	48.3	58	75.7

Incident #	Shock Wave Speed (kmph)			Time-Space Domain	
	W_{12}	W_{23}	W_{31}	D(min)	X(km)
651	-8.1	93.4	111.1	1.8	18
655	-22.5	32.2	64.4	3.7	4

[a] *Traffic condition at the first detector upstream of the incident location during the first time slice after the incident has occurred.*

[b] *Traffic condition at the first detector upstream of the incident location during the last time slice before the incident is cleared.*

[c] *One kilometer (km) = 0.6 mile.*

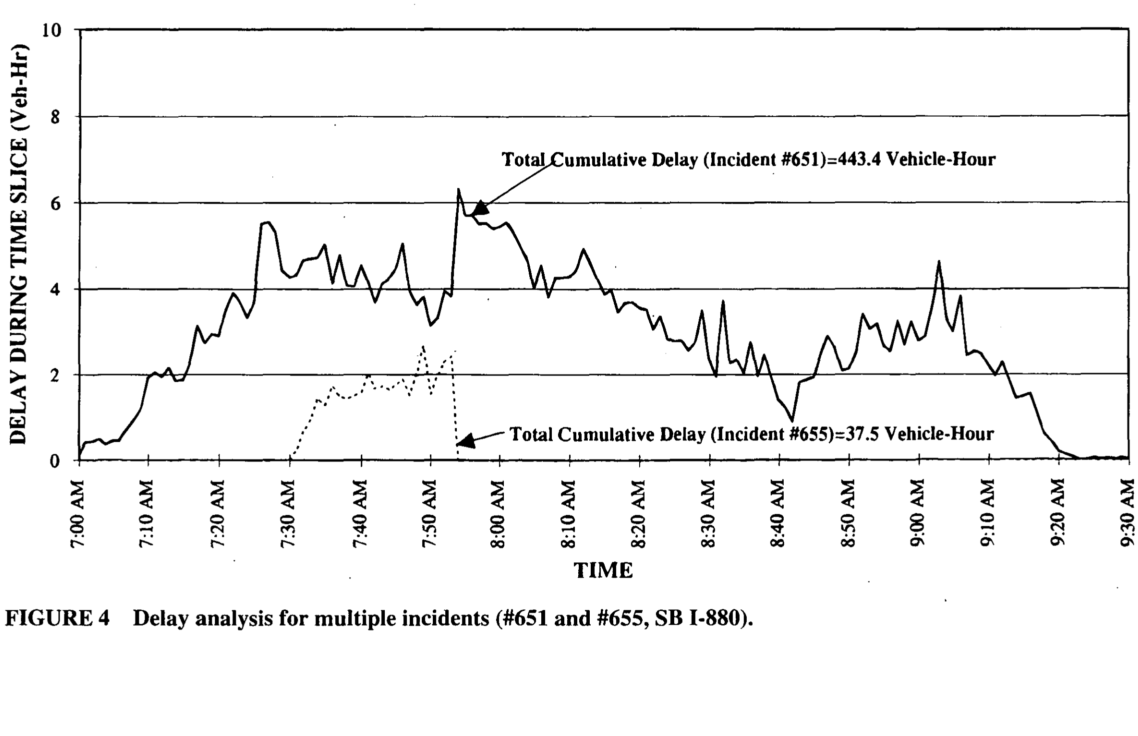

FIGURE 4 Delay analysis for multiple incidents (#651 and #655, SB I-880).

TABLE 3 Incident Delays

	Total			Morning Shifts			Evening Shifts		
Incident Type	**N[a]**	**Average**	**St-Dev[b]**	**N**	**Average**	**St-Dev**	**N**	**Average**	**St-Dev**
Right Shoulder Breakdown	176	2.5	7.1	82	0.8	2.5	94	3.9	9.2
Left Shoulder Breakdown	6	28.8	45.1	2	6.4	9.1	4	40.1	53.5
In Lane Breakdown	4	25.4	41.8	1	0.0	—[c]	3	33.8	46.8
Right Shoulder Accident	21	2.6	4.3	8	4.1	6.4	13	1.7	2.2
Left Shoulder Accident	13	11.4	28.6	6	4.5	9.1	7	17.3	38.5
In Lane Accident	11	55.0	63.8	6	74.6	72.7	5	31.6	48.0
Total Number of Incidents	231			105			126		

[a]*N* = *Number of incidents for each category.*
[b]*St-Dev* = *Standard Deviation.*
[c]*–* = *Standard deviation can not be calculated (one case only).*

the congested area itself. This area serves as a guideline for computations of the actual drop in speeds attributable to the incident over time and space. If there is no drop in speed on a specific segment located within the time-space domain of an incident, then delay is zero for that segment. Hence, although the area of search for a speed drop is larger than the actual one, incident congestion is not overestimated because zero delays are assigned to those segments not affected by the incident. Because over-prediction of the time-space domain is more likely to capture all segments with incident congestion, over-prediction is preferred over under-prediction. The only problem with over-prediction is more computational effort and time spent in checking for speed drops on what will turn out to be zero-delay segments.

CONCLUSION

Most of the conventional methods for estimating incident congestion (except for INTRAS and FRESIM) are incapable of using the detailed loop and incident data that recently became available in several surveillance systems and freeway traffic operations projects around the country. This is either because these methods are designed to deal with summarized types of data and make numerous assumptions or because they are too theoretical and have never been validated with real-life data. This makes them of limited use for practitioners. Moreover, it is not possible to use the conventional *macroscopic* methods for analyzing cases of multiple incidents that occur on the same stretch of highway resulting in multiple queues that merge together. On the other hand, *microscopic* traffic analysis tools such as INTRAS and FRESIM require extensive calibration with loop data and making assumptions about car-following theory.

This paper has presented a new macroscopic method for estimating freeway incident congestion. The method is based on shock wave analysis where the area of influence of a specific incident is demarked. If the time-space domain of incidents overlap, they result in a case of multiple incidents. An algorithm for separating congestion of each incident has been described in this paper. Also, the time-space domain of an incident is used to distinguish between isolated and multiple incident cases. In the new method, incident detection and clearance times collected simultaneously with speeds and traffic counts from mainline loop stations and on- and off-ramp stations are used to calculate incident delay on each freeway segment and for each time slice during the congested time-space region and also to obtain cumulative incident congestion. The method is applied to a sample of incident data from the FSP database of I-880 in Alameda County, California. The sample includes both isolated and multiple incident cases, and the application results are reasonable. Generally, the method overestimates the maximum incident queue length and, consequently, the incident congestion boundaries. However, this does not necessarily overestimate the incident delay. Incident delay is not calculated using the congested areas confined by the time-space domain; rather incident delay is based on the actual drop in speeds on the freeway segments upstream of the incident.

Future research will focus on two main issues:

1. Refinement of this method by calibrating the estimated time-space domain with tach car data. These data include tach vehicle travel times and speeds, which is another source of independent field data collected in the FSP project, and
2. Comparison of incident delay results with those of FRESIM for the same sample of the FSP incident database used in this paper.

Also, the authors will seek applications of the new method in other sites, where similar data have already been collected, such as I-4 in Orlando, Florida (Al-Deek, unpublished data). Although the new method needs further refinements, the authors hope that this paper has accomplished an important step toward bridging the gap between theory and practice in the field of freeway traffic operations.

ACKNOWLEDGMENTS

The authors acknowledge the help of Professor Pravin Varaiya, Dr. Alex Skabardonis, Karl Petty, Hisham Noeimi, and Dan Rydzewski of the University of California at Berkeley in providing the research team with data and comments on this paper.

REFERENCES

1. Lindley, J. Urban Freeway Congestion: Quantification of the Problem and Effectiveness of Potential Solutions. *ITE Journal*, Vol. 57, No. 1, January 1987, pp. 27–32.
2. Morales, J. M. Analytical Procedures for Estimating Freeway Traffic Congestion. *Public Road*, Vol. 50, No. 2, September 1986, pp. 55–61.
3. *Special Report 209: Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 1994, pp. 7–17.
4. Messer, C. J., C. L. Dudek, and J. D. Friebele. Method for Predicting Travel Time and Other Operational Measures in Real-Time During Freeway Incident Conditions. In *Highway Research Record 461*, TRB, National Research Council, Washington, D.C., 1973, pp. 1–16.
5. Greenshields, B. A Study of Traffic Capacity. *HRB Proceedings*, Vol. 14, 1934, pp. 448–477.
6. Wirrasinghe, S. Determination of Traffic Delays from Shock Wave Analysis. *Transportation Research*, Vol. 12, 1978, pp. 343–348.
7. Chow, W. A Study of Traffic Performance Models under Incident Conditions. In *Highway Research Record 567*, HRB, National Research Council, Washington D.C., 1974, pp. 31–36.
8. Wicks, D., and E. Lieberman. *Development and Testing of INTRAS, a Microscopic Freeway Simulation Model: Program Design, Parameter Calibration and Freeway Dynamics Component Development*. Report No. FHWA/RD-80/106, FHWA, U.S. Department of Transportation, 1980.
9. FHWA, Office of Research and Development. *FRESIM User Guide*. Beta Version 3.1, 1992.
10. Cheu, R., W. Recker, and S. Ritchie. Calibration of INTRAS for Simulation of 30-Second Loop Detector Output. *Transportation Research Board 72nd Annual Meeting*, Washington D.C., January 1993.
11. Skabardonis, A., H. Noeimi, K. Petty, D. Rydzewski, P. Varaiya, and H. Al-Deek. Freeway Service Patrol Evaluation, California PATH Research Report, UCB-ITS-PRR-95-5, Institute of Transportation Studies, University of California Berkeley, 1995.

Publication of this paper sponsored by Committee on Freeway Operations.