

A Methodological Approach for Estimating Temporal and Spatial Extent of Delays Caused by Freeway Accidents

Younshik Chung and Wilfred W. Recker

Abstract—Given that reliable prediction of such relatively rare—and random—events as accident occurrence will remain elusive, the most important potentially soluble factor in the development of accident management strategies is to identify and quantify the conditions affecting the nonrecurrent congestion caused by accidents once they have been known to have occurred. The objective of the research reported in this paper is to develop a method of quantifying the delay due to accidents on urban freeways, as well as to identify the causal factors affecting the total delay caused by such accidents. Binary integer programming (BIP) is applied in estimating the temporal and spatial extent of delay caused by freeway accidents, based solely on commonly available inductance loop detector data. The basic idea behind the method is to estimate the most likely temporal and spatial extent of the region of congestion caused by an accident by solving a BIP problem that is consistent with the topology of the spatio-temporal region that defines candidate speed differences between normal flow conditions and accident conditions. The procedures developed in this paper will be useful for the performance evaluation of accident management systems by quantifying accident congestion in terms of the total delay to evaluate the benefit of accident management systems accrued from efficient traffic operations. The procedures are demonstrated by a case study using accident data collected from six major freeways in Orange County, CA.

Index Terms—Accidents, binary integer programming (BIP), censored data, nonparametric survival analysis, nonrecurrent congestion, shockwave, spatio-temporal congestion region.

I. INTRODUCTION

ALTHOUGH freeways comprise only a small fraction of metropolitan transportation network mileage, they form the backbone of most urban transportation networks, carrying more than a third of all vehicular travel, and typically are the primary focus of concern of traffic congestion. Most studies of freeway congestion have focused on recurring congestion, which occurs when traffic demand exceeds capacity. Since recurrent congestion typically follows a predictable pattern, alleviation is generally amenable to such demand control policies

as high-occupancy vehicle (HOV) lanes, congestion toll pricing, transit incentives, and ramp metering, as well as expansion of capacity through construction. Such is not generally the case with nonrecurrent congestion, particularly that due to accidents, which is typified by unpredictable locations, times, types, and severity. It is estimated that nonrecurrent congestion caused by such events as accidents, disabled vehicles, and weather events accounts for one-half to three-fourths of the total congestion on metropolitan freeways. Effective management of nonrecurring congestion due to such unpredictable events as traffic accidents (one of the main sources of non recurrent congestion) heavily relies on timely response with appropriate measures (e.g., deployment of tow trucks, motorist messaging, clean-up crews, etc.); a specific field, i.e., accident management, has become an important component of freeway traffic management systems. The effectiveness of this sort of traffic management system is dependent on an accurate forecast of the potential extent of congestion delay attendant to the accident(s) to evaluate both the nature and the magnitude of the best response and, in the case of limited resources and multiple incidents, to determine appropriate system asset management strategies.

A number of methods have been proposed to estimate total nonrecurrent congestion delay. They can be classified into four groups: 1) analytical methods using deterministic queuing diagrams (e.g., [1], [2]–[9]); 2) kinematic wave (i.e., shockwave) theory (e.g., [3], [10]–[15]); 3) heuristic methods (e.g., [16], [17]–[22]); and 4) simulation methods [23]–[25]. The results from these approaches have been applied both alone and in concert with each other. However, since virtually none of these studies uses a source of incident data with descriptive variables and reporting techniques in common with the others, comparisons between different methods are difficult [26], and their spatial transferability is limited. These efforts notwithstanding, to date, there has been only limited research into models that can estimate how long, how far, and to what extent any particular incident will affect traffic.

The objective of the research reported here is to develop methods to perform the following: 1) quantify, using loop detector data, accident delay and the temporal and spatial extent of accident-related congestion and 2) identify the factors affecting the total delay caused by an accident. To demonstrate the method, we apply the model to estimate the total delay caused by a series of accidents on six major Orange County, CA, freeways using archived loop detector data and traffic accident data; we rely only on fundamental accident data such as

Manuscript received May 6, 2011; revised October 28, 2011 and February 9, 2012; accepted March 3, 2012. The Associate Editor for this paper was M. Barth.

Y. Chung is with the Department of Highway, ITS, and Safety Research, Korea Transport Institute, Goyang 411-701, Korea (e-mail: tpgist@koti.re.kr).

W. W. Recker is with the Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697 USA (e-mail: wwrecker@uci.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TITS.2012.2190282

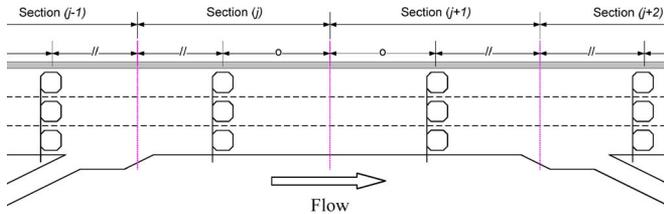


Fig. 1. Section definition and the corresponding detector location.

accident time and location and traffic flow data such as volume, speed, and occupancy data from loop detectors. We believe that the methods can provide a foundation to develop a performance measure to evaluate transportation policies and planning-level analyses associated with the design of transportation systems or as a basis for preparation of operating plans for safety, as well as for evaluation of deployed transportation projects or technologies.

II. ANALYTICAL MODEL

A. Section Definition

For purposes of analysis, we first segment the freeway of interest into sections that correspond to mainline loop detector stations. Section boundaries are defined by the middle of two detector stations, as shown in Fig. 1; we assume that estimated speeds at the station are representative of the speed for the corresponding section. Based on the sections and their corresponding detector stations, estimated speeds and densities for each section j are calculated for each subinterval (e.g., at 5-min intervals) during the analysis period, i.e., T weeks; sections downstream of section j are consecutively labeled as $(j + 1)$, $(j + 2)$, \dots , whereas those upstream are consecutively labeled as $(j - 1)$, $(j - 2)$, \dots

B. Nominal Speed Distribution

For each freeway section j and for each specific subinterval t_m on each day of the week, a nominal total of T observations¹ are available during the observation period, on which to base calculations of flow $q_j(t_m)$ and occupancy $occ_j(t_m)$ under conditions that are accident-free. For example, section i on Monday from $t_m = 08:10$ to $08:15$ (assuming 5-min subintervals) for an observation period $T = 52$ weeks is composed of 52 samples. The speed associated with the n th observation for any particular day-of-week/time interval/section combination can be estimated (either directly from dual loop detector stations or using the well-known g -factor for single loops) as $s_{jn}(t_m)$, where subscript n is used to designate the estimate of the corresponding value based on the n th observation. From these (nominally) T observations, a distribution of $s_{jn}(t_m)$ can be constructed: let $\Omega_{jm} = \Omega(\bar{s}_j(t_m), \sigma_{s_j(t_m)})$ denote the set of parameters defining the distribution of speeds $s_{jn}(t_m)$ corresponding to this “accident-free” base case. (Here, there is no attempt made to categorize these distributions among the set of well-known distributions; rather, there is an assumption

¹Exclusive of time intervals ostensibly affected by accident conditions prevailing in the vicinity of the section at the time of the observation or otherwise missing data.

TABLE I
BASE CASE OF DISTRIBUTIONAL PROPERTIES FOR
ACCIDENT-FREE SPEEDS

Time	Freeway section (traffic flow direction ←)					
	i	$i-1$	\dots	2	1	
t_1	$\Omega_{i,1} = \Omega(\bar{s}_i(t_1), \sigma_{s_i(t_1)})$	\dots	\dots	\dots	$\Omega_{1,1} = \Omega(\bar{s}_1(t_1), \sigma_{s_1(t_1)})$	
t_2	$\Omega_{i,2} = \Omega(\bar{s}_i(t_2), \sigma_{s_i(t_2)})$	\dots	\dots	\dots	$\Omega_{1,2} = \Omega(\bar{s}_1(t_2), \sigma_{s_1(t_2)})$	
t_3	$\Omega_{i,3} = \Omega(\bar{s}_i(t_3), \sigma_{s_i(t_3)})$	\dots	\dots	\dots	$\Omega_{1,3} = \Omega(\bar{s}_1(t_3), \sigma_{s_1(t_3)})$	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
t_{M-1}	$\Omega_{i,M-1} = \Omega(\bar{s}_i(t_{M-1}), \sigma_{s_i(t_{M-1})})$	\dots	\dots	\dots	$\Omega_{1,M-1} = \Omega(\bar{s}_1(t_{M-1}), \sigma_{s_1(t_{M-1})})$	
t_M	$\Omega_{i,M} = \Omega(\bar{s}_i(t_M), \sigma_{s_i(t_M)})$	\dots	\dots	\dots	$\Omega_{1,M} = \Omega(\bar{s}_1(t_M), \sigma_{s_1(t_M)})$	

TABLE II
OBSERVED ACCIDENT SPEEDS

Time	Freeway section (traffic flow direction ←)					
	i	$i-1$	$i-2$	\dots	2	1
t_1	$\hat{s}_i(t_1)$	$\hat{s}_{i-1}(t_1)$	\dots	\dots	\dots	$\hat{s}_1(t_1)$
t_2	$\hat{s}_i(t_2)$	$\hat{s}_{i-1}(t_2)$	\dots	\dots	\dots	$\hat{s}_1(t_2)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
t_m	$\hat{s}_i(t_m)$	$\hat{s}_{i-1}(t_m)$	\dots	\dots	\dots	$\hat{s}_1(t_m)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
t_{M-1}	$\hat{s}_i(t_{M-1})$	$\hat{s}_{i-1}(t_{M-1})$	\dots	\dots	\dots	$\hat{s}_1(t_{M-1})$
t_M	$\hat{s}_i(t_M)$	$\hat{s}_{i-1}(t_M)$	\dots	\dots	\dots	$\hat{s}_1(t_M)$

made for simplicity and, with no loss in generality, that they can be adequately defined by their means and standard deviations.)

Then, associated with an accident known to have occurred on freeway section i during some time interval t_1 , we can compose a matrix of the parameters defining the distribution of base case conditions (i.e., conditions in which there is no accident) that can be expected to prevail for $t_m \geq t_1$, $m = 1, 2, \dots$ (i.e., time intervals after the accident that occur in section i at t_1), for all upstream sections that could possibly have been affected by the accident, as in Table I.

Similarly, associated with a known accident that occurs on freeway section i at time t_1 , a speed matrix can be constructed based on observed freeway traffic flow data under the accident conditions. Table II shows the speed matrix under the accident condition.

Relative to the display of information in Table II, we can schematically identify the negative effects (i.e., speed reduction) of the accident, as shown in Fig. 2. The negative effect by the accident will be propagated from the accident section to upstream sections. Such a distinct discontinuity between noncongested and congested flow is known as a shockwave [14]. If the dot-shaded area affected by the shockwave in Fig. 2 is identified, then the temporal and spatial impacts of the accident can also be determined. We approach this identification as a topology problem, in which the objective is to construct the most probable set of cells that define the accident-affected temporal/spatial region, vis-à-vis any recurrent congestion effects subject to constraints imposed by allowable shapes of the affected region. The succeeding sections describe the method

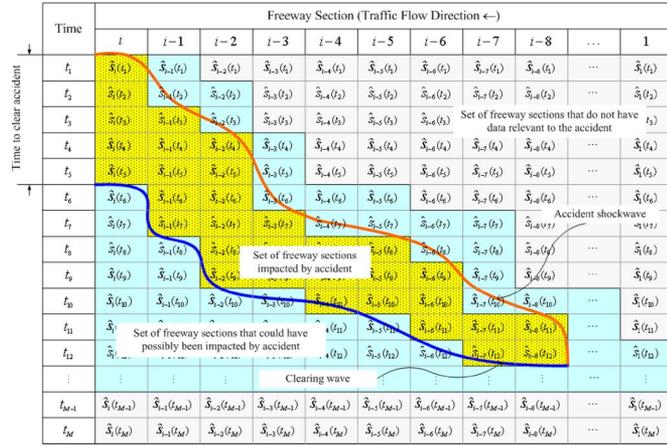


Fig. 2. Topology of freeway sections impacted by accident.

 TABLE III
 EXAMPLE OF MAXIMUM POSSIBLE EXTENT OF SHOCKWAVE

Acc ID #	Acc time	Time lapse	Shockwave Length (mile)	Possibly affected section IDs	
				Upstream section	
9	2001-03-01 14:15:00	14:15:00	0.000	11	
9	2001-03-01 14:15:00	14:20:00	-0.787	11	
9	2001-03-01 14:15:00	14:25:00	-1.517	11,10,9	
9	2001-03-01 14:15:00	14:30:00	-2.179	11,10,9,8,7	
9	2001-03-01 14:15:00	14:35:00	-2.945	11,10,9,8,7	
9	2001-03-01 14:15:00	14:40:00	-3.501	11,10,9,8,7,6,5	
9	2001-03-01 14:15:00	14:45:00	-4.081	11,10,9,8,7,6,5,4	
9	2001-03-01 14:15:00	14:50:00	-4.593	11,10,9,8,7,6,5,4,3	
9	2001-03-01 14:15:00	14:55:00	-5.264	11,10,9,8,7,6,5,4,3,2	
9	2001-03-01 14:15:00	15:00:00	-5.813	11,10,9,8,7,6,5,4,3,2,1	
9	2001-03-01 14:15:00	15:05:00	-6.376	11,10,9,8,7,6,5,4,3,2,1	
9	2001-03-01 14:15:00	15:10:00	-6.750	11,10,9,8,7,6,5,4,3,2,1	

for discriminating the regions between the noncongested and the congested area due to traffic accidents.

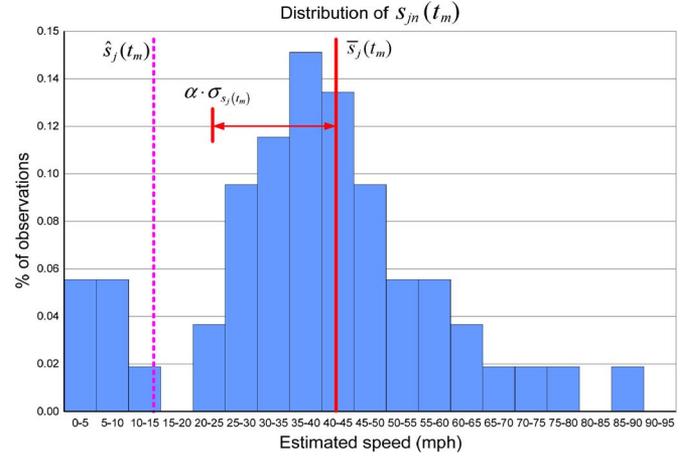
C. Determining the Maximum Extent of Accident Influence

Since data neither on the severity (such as the number of closed lanes by the accident) nor when the accident was cleared are directly obtainable from loop data, we first estimate the maximum possible extent of the shockwave by assuming the worst possible conditions, i.e., total blockage for some prespecified time period. Thus, for any given accident occurring at section i at time t_1 , we compute the maximum number of upstream sections that could be affected by the assumed persistent total blockage at section i at time t_1 . Table III provides an example of one such set of calculations for an accident (Accident ID #9) that was reported on section $i = 11$ of SR-22 in Orange County, CA, at $t_1 = 14 : 15 : 00$ on 03/01/2001.

Using this sort of data, we can schematically construct the “maximum area of interest” for any accident occurring at section i at time t_1 as the shaded (blue) area in Fig. 2. Based on this interpretation, the only data relevant to the current example of an accident occurring at section i at time t_1 can be restricted to cells in the shaded (blue) area in Fig. 2.

D. Determining Congested Region

In Fig. 2, the cells in the dot-shaded (yellow) area represent speeds that have been determined to have been reduced due to


 Fig. 3. Conclusion that cell is affected by accident: $\hat{s}_j \leq \bar{s}_j(t_m) - \alpha \cdot \sigma_{s_j(t_m)}$.

the accident (i.e., affected by accident), which is deemed to be significantly lower than would be expected based on the distribution of accident-free speeds for that particular section-time interval combination. Alternatively, the plain-shaded (blue) cells represent particular section-time interval combinations for which the observed speeds are not significantly different from accident-free conditions. That is, they are speeds that are determined from observation not to be affected by the pertinent accident but within the maximum possible affected congestible sections by an accident of greatest consequence (when all lanes are blocked, as previously assumed).

Considering that the speed of traffic in sections adversely affected by the traffic accident will be lowered, the basic idea behind discriminating between these two regions is to compare the accident speed $\hat{s}_j(t_m)$, for any particular freeway section j and subinterval t_m , to the distribution of the accident-free speeds $s_{jn}(t_m)$, $n = 1, 2, \dots, n_{\text{obs}}$, $n_{\text{obs}} \leq T$, and assign some level of confidence that any particular speed was not drawn from the distribution of $s_{jn}(t_m)$. If we limit our level of confidence that the speed was not drawn from the distribution of $s_{jn}(t_m)$ to those that compute to be at least $\alpha \cdot \sigma_{s_j(t_m)}$, where α is a positive number, below the mean speed of the distribution of $s_{jn}(t_m)$, we can identify affected versus nonaffected cases according to the difference between the mean speed and the threshold, as shown in Fig. 3.

By assigning to the cells in Fig. 2 the outcome of the probability that the speed in such cells is drawn from the distribution of speeds associated with accident-free conditions, we can construct the binary matrix shown in Fig. 4 as the representation of the affected region.

However, owing to uncertainties and other externalities, assigning an outcome based on a significance test related to the null hypothesis that the cell speed actually came from the accident-free speed distribution will generally result in less-than-accurate conclusions regarding the affected region, as shown in Fig. 5, where the true boundary of the affected region is shown by the thick line (red line).

If we define P_{jm} as follows:

$$P_{jm} = \begin{cases} 0, & \hat{s}_j(t_m) \leq \bar{s}_j(t_m) - \alpha \cdot \sigma_{s_j(t_m)} \\ 1, & \hat{s}_j(t_m) > \bar{s}_j(t_m) - \alpha \cdot \sigma_{s_j(t_m)} \end{cases} \quad (1)$$

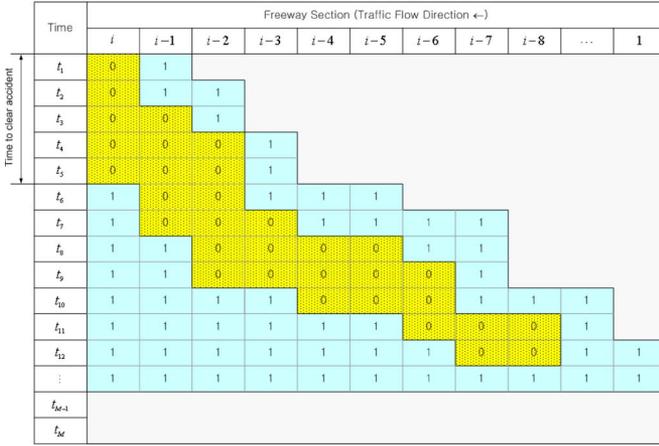


Fig. 4. Binary matrix of accident speeds.

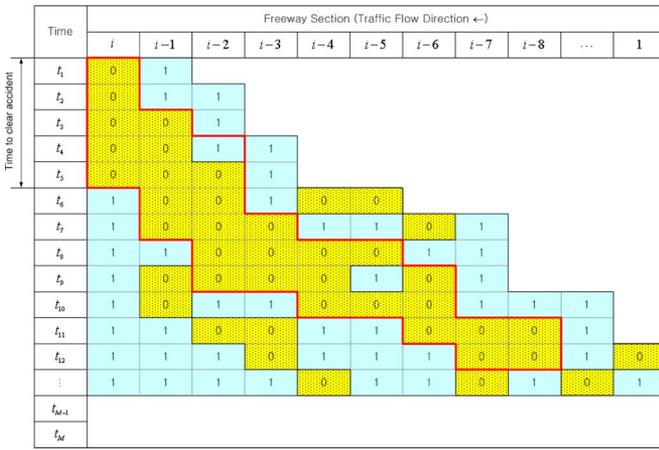


Fig. 5. Empirical binary matrix of accident speeds.

the problem of determining the “best” set of “dot-shaded (or yellow)” cells can be formulated as the following statement:

$$\sum_{\forall \text{dot-shaded cells}} P_{jm} + \sum_{\forall \text{shaded cells}} (1 - P_{jm}) = \text{Minimum.} \quad (2)$$

Or, defining

$$\delta_{jm} = \begin{cases} 1, & \text{if cell is affected by the accident} \\ 0, & \text{if cell is not affected by the accident.} \end{cases}$$

Equation (2) can be written as

$$\sum_{\forall j,m} [P_{jm} \cdot \delta_{jm} + (1 - P_{jm}) \cdot (1 - \delta_{jm})] = \text{Minimum.} \quad (3)$$

The subset of cells for which the accident speeds are significantly different from the accident-free speeds comprises a region that theoretically must obey certain topological properties. Specifically, there are three impossible local shape configurations for the subset of spatio-temporal cells congested by the accident. The first such case is a region that contains any holes (Fig. 6), i.e., the spatio-temporal progression of the accident shockwave must be uninterrupted. In addition, the vertical position (t) of any dot-shaded (“yellow”) section j

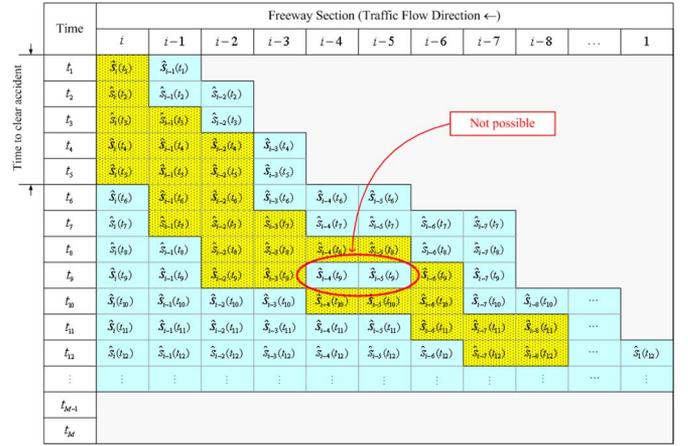


Fig. 6. Impossible shape I of the congested region.

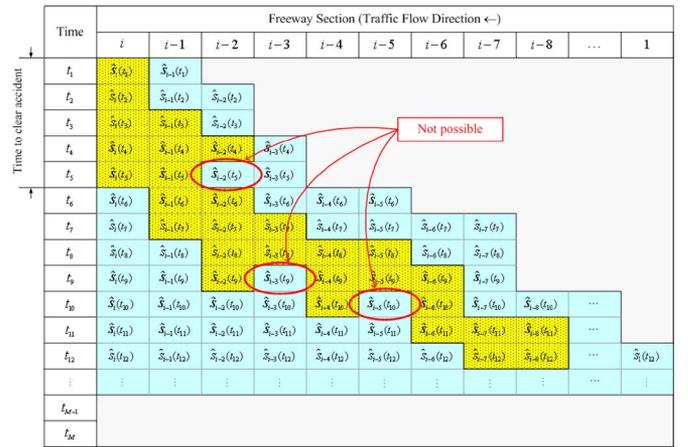


Fig. 7. Impossible shape II for the shape of the congested region.

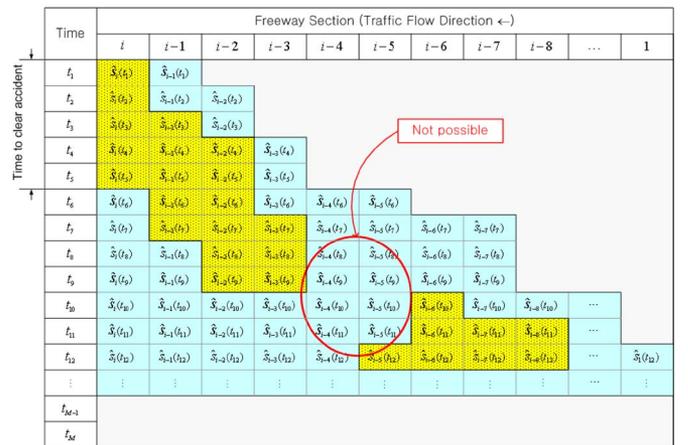


Fig. 8. Impossible shape III for the shape of the congested region.

along the boundary of the region must be either lower or the same (i.e., \leq) vertical position of the neighboring shaded (“yellow”) section $j - n$ (Fig. 7), i.e., the boundary of the spatio-temporal progression of the accident shockwave must be upstream. Likewise, the entire boundary of the affected region must be contiguous (Fig. 8).

These conditions can be enforced by the following relationships:

$$\begin{aligned} \delta_{j+k,m} &\leq [1 - (\delta_{j,m} - \delta_{j+1,m})] \cdot R \quad \forall j, m \quad \forall k \leq J - j \quad (a) \\ \delta_{j,m+r} &\leq [1 - (\delta_{j,m} - \delta_{j,m+1})] \cdot R \quad \forall j, m \quad \forall r \leq M - m \quad (b) \\ \delta_{j,m+k} &\leq [1 - (\delta_{j,m} - \delta_{j+1,m})] \cdot R \quad \forall j, m \quad \forall k \leq M - m \quad (c) \end{aligned} \quad (4)$$

where R is a large number, J is the maximum number of upstream sections, and M is the maximum number of subinterval time periods that define the maximum duration assumed for congestion caused by the accident (e.g., for 5-min subintervals and a maximum analysis time period of 4 h, i.e., $M = 48$).

Equations (3) and (4) are in the form of the objective function and constraint, respectively, of a “binary integer programming (BIP)” problem. The determination problem previously described can be represented in the form of the following BIP problem:

$$\begin{aligned} \text{Min } Z &= \sum_{\forall j,m} [P_{jm} \cdot \delta_{jm} + (1 - P_{jm}) \cdot (1 - \delta_{jm})] \\ \text{s.t. :} & \\ \delta_{j+k,m} &\leq [1 - (\delta_{j,m} - \delta_{j+1,m})] \cdot R \quad \forall j, m \quad \forall k \leq J - j \\ \delta_{j,m+r} &\leq [1 - (\delta_{j,m} - \delta_{j,m+1})] \cdot R \quad \forall j, m \quad \forall r \leq M - m \\ \delta_{j,m+k} &\leq [1 - (\delta_{j,m} - \delta_{j+1,m})] \cdot R \quad \forall j, m \quad \forall k \leq M - m \\ \delta_{jm} &= \begin{cases} 0 \\ 1. \end{cases} \end{aligned} \quad (5)$$

E. Total Delay Estimation

To estimate the total delay (TD) caused by the accident, this study defines the nonrecurrent delay as the additional delay produced by the reduction in section speed caused by an accident over the average annual section speed (or accident-free condition). Then, having completed the preceding steps that determine the region (in time and space) that is negatively affected by any particular accident, we can calculate TD as

$$TD = \sum_{\forall j,m \in \text{dot-shaded cells}} \max \{ L_j \cdot [1/\hat{s}_j(t_m) - 1/\bar{s}_j(t_m)] \cdot V_{jm}, 0 \} \quad (6)$$

where

L_j	length of freeway segment j ;
V_{jm}	volume (count) of vehicles in segment j during time m ;
$\hat{s}_j(t_m)$	speed affected by accident in segment j at time m ;
$\bar{s}_j(t_m)$	annual average speed in segment j at time m .

III. CASE STUDY

To demonstrate the analytical procedures developed in the preceding sections, we apply the model to a case study of congestion-related effects of accidents occurring on the freeway system in Orange County, CA.

A. Traffic Flow Data

The case study uses one-year historical inductive loop detector data from 1 March 2001 to 28 February 2002 (i.e., $T = 52$) on six major freeways in Orange County, CA: Interstate 405 (I-405), Interstate 5 (I-5), State Route 22 (SR-22), State Route 55 (SR-55), State Route 57 (SR-57), and State Route 91 (SR-91). The original data include traffic counts and occupancy for each lane at each detector station every 30 s. There are 499 mainline loop detector stations, excluding stations on HOV lane(s) in the study area. Due to freeway construction and communication problems, about a quarter of the detectors were not functioning during the study period, and only a 61.3% of loop stations were able to provide traffic data coverage for more than 50% of the study period. In this application, lane-by-lane traffic data are aggregated into 5-min intervals (i.e., $|t_m| = 5$ min) at each detector station to obtain stable traffic data from each point.

Prior to determining the congested region, it is required to separate any particular accident-induced speed $\hat{s}_j(t_m)$ from the distribution of accident-free speeds $s_{jn}(t_m)$. In this paper, the speed observations included in the maximum extent of accident influence were regarded as accident-induced speed. On the other hand, the distribution of the base case condition (i.e., accident-free speed) in Table I was constructed from days, regardless of accident influence. With this separating judgment, an accident-induced speed and its associated accident-free speed distribution can be determined. Additionally, when calculating the parameters of the distribution of the accident-free speeds $s_{jn}(t_m)$, $n = 1, 2, \dots, n_{\text{obs}}$, $n_{\text{obs}} \leq T$, and assigning some level of confidence that any particular $\hat{s}_j(t_m)$ was not drawn from the distribution of $s_{jn}(t_m)$, for cells in Fig. 2 for which $n_{\text{obs}} < n_{\text{min obs}}$, we set $P_{jm} = 0.5$.

B. Accident Data

Accident data were obtained from the Traffic Accident Surveillance and Analysis System (TASAS) maintained by the California Department of Transportation (Caltrans). Approximately 6200 crashes were recorded during the study period. The accident data include basic information related to accident time and accident location in terms of freeway milepost. In addition, data for each crash include three primary crash characteristics: 1) crash type, based on the type of collision (rear end, side-swipe, or hit object), the number of vehicles involved, and the movement of these vehicles prior to the crash; 2) crash location, based on the location of the primary collision (left lane, interior lanes, right lane, right shoulder area, or off-road beyond right shoulder); and 3) crash severity, in terms injuries and fatalities per vehicle.

C. Speed Estimation Using Single-Loop Data

Loop detectors in Orange County predominantly are single-loop detectors that provide only traffic counts and occupancies (occ) for each lane at each detector station every 30 s; travel speeds need to be estimated from these measures, assuming a so-called “ g -factor,” which is the summation of the average

vehicle length and effective detection length [27]–[29]. The average speed \bar{s} can be calculated as

$$\bar{s} = g \frac{q}{5280 \times occ} \quad (7)$$

where q = flow rate(veh/h), and $g = g - \text{factor}$.

In this paper, we account for spatial and temporal variations of the g -factors by calculating a g -factor representing each hour interval over the 52 weeks for each loop detector station as

$$\hat{g}_i(t+1) = p \cdot \hat{g}_i(t) + (1-p) \cdot g_i(t+1) \quad (8)$$

where $\hat{g}_i(t)$ denotes g -factors for section i for time step t , $g_i(t+1)$ denotes initial average g -factors for section i for time step $t+1$ obtained by assuming free-flow speed (75 mi/h) when the occupancy is lower than 0.06, and p denotes smoothing parameter (0.9 in this study).

By applying the aforementioned procedure, representative g -factors for each hour of the day are calculated for all loop detector stations. These g -factors provide basic information for speed calculations in this study. Based on these g -factors, speeds are calculated every 5 min for all 52 weeks, i.e.,

$$s_{jn}(t_m) = g_{jn}(t_m) \cdot \frac{q_{jn}(t_m)}{5280 \times occ_{jn}(t_m)}, \quad n=1, 2, \dots, 52. \quad (9)$$

In the definition of P_{jm} , the threshold value α is very critical to distinguish between speed affected by accident and the other speeds. In this paper, various values were applied to find the best threshold. The best value was empirically found from the relation between traffic data and traffic accident as $\alpha = 0.25$.²

As previously noted, we first need to set a threshold regarding the minimum number of observations that we require to have confidence in the statistical calculations for mean and standard deviation of the nominal (accident-free) speed distribution $n_{\min obs}$. Since 30 is commonly used as the minimum number of observations required for the law of large numbers to apply, it was determined that $n_{\min obs} = 30$. Only 61.3% of loop stations were able to provide more than 50% of traffic data required for the analysis. We also arbitrarily set the maximum time duration of accident effects M to 4 h. On the Orange County freeway system, there are many cases of missing data traced to reasons related to loop detectors temporarily not providing valid data, resulting in the estimation of delay for only 2232 (36.10%) of the 6182 total accidents.

D. Treatment of Censored Observations

Of the 2232 accident cases analyzed, congested regions from 232 accidents were censored by the time boundary condition (i.e., cases where congestion was still noted beyond $M = 4$ h),

²Although different α values with respect to each accident data are used for (1), the best value is employed to simplify the estimation of nonrecurrent congestion. Thus, the α value in this study was empirically found by using the relationship between randomly selected 50 accident samples and their associated traffic data, meaning that various values were applied to find the best α value with a 0.05 increment from 0.00. When the value was 0.25, then the congested region (i.e., dot-shaded region) in Fig. 2 was best discriminated from the noncongested regions.

TABLE IV
SUMMARY ON DELAY ESTIMATION RESULTS

Description	Freeway						Total
	1-405	1-5	SR-22	SR-55	SR-57	SR-91	
Number of observations	364	662	307	227	437	235	2,232
2-dimensional censoring	3	28	6	10	52	8	107
Censored by time	57	38	47	46	24	20	232
Censored by space	11	46	5	12	34	26	134
No censoring	293	550	249	159	327	181	1,759

spatially censored cases caused by either the missing data problem or a spatial boundary condition accounted for 134 observations (i.e., cases where congestion was still noted beyond J freeway sections), and 107 accidents were related to 2-D censoring by both of time and space boundaries or missing data. Table IV shows the distribution of cases for each freeway for which we were able to apply the procedures described here to estimate delay caused by accidents—about 21% of the total of 2232 accidents estimated are censored by time and/or space boundary conditions.

Among the censored results are a significant number of “serious” accidents, e.g., those related to injuries and/or fatalities, secondary accidents that occurred before the first accident was cleared, hazardous material spills, etc. We include their impact in the statistical analysis through standard survival functions. To form formal tests of hypothesis for the equality of survivor functions across groups, such nonparametric tests as the log-rank, Wilcoxon, Tarone-Ware, Peto-Prentice, and generalized Fleming-Harrington test, can be considered.³ These tests do not test the equality of the survivor functions at a specific point but compare the overall survival functions. The log-rank test is most powerful when the hazards are not equal but instead are proportional to one another [30]. If two different survival curves cross each other, the Wilcoxon test is preferred to the log-rank test. However, due to the weighting scheme of the Wilcoxon test, it has limitations if the censoring patterns differ over the test groups.

In this paper, the log-rank test is employed for univariate analysis using $p < 0.05$. Since such analysis gives a statistical result only for the difference of survivor functions across groups, another statistical analysis is required to identify which factor in a group contributes to total delay. To graphically illustrate the difference, the Kaplan-Meier (KM) estimate (or product limit estimate) is employed, which is one of the most common nonparametric survival analysis methods. An example of the KM estimates for total delay as a function of Midday versus other time periods is shown in Fig. 9.

E. Case Study Results

The median total delay for the 2232 accidents, including censored results, is 22.27 vehicle hours, and the minimum total

³Such statistical methods as Chi-square tests and t-tests are not appropriate for censored data. Moreover, ordinary least squares or other general regression models cannot properly handle such censored data.

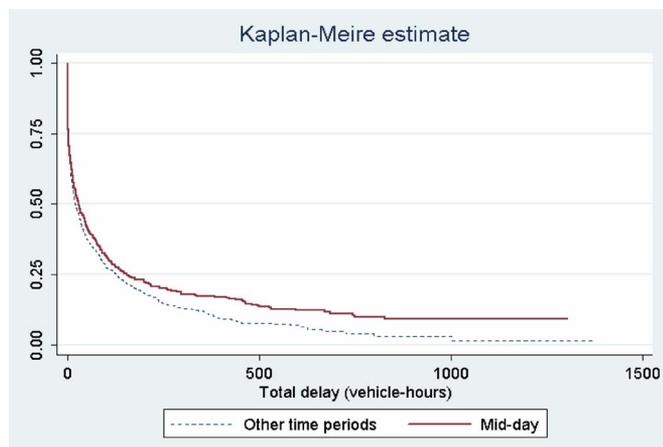


Fig. 9. Example of KM estimates for midday versus other time periods.

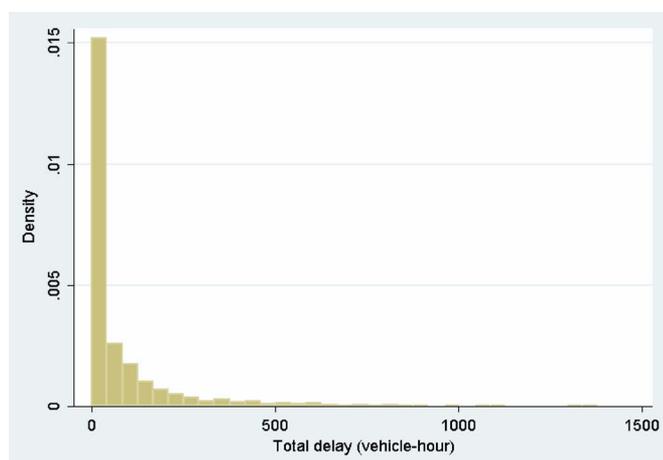


Fig. 10. Histogram of total delay including censoring observations.

delay and the maximum total delay are 0 and 1379.49⁴ vehicle hours, respectively. Fig. 10 shows the histogram of estimated total delay for the 2232 accidents, including censored results. Its shape follows a typical survivor function.

To better understand the relationships between the various factors involving an accident and its resultant congestion effects, the variation in estimated delay for the set of accidents was analyzed relative to the set of available accident characteristics. For this analysis, candidate variables are classified into three groups based on accident data from TASAS: 1) accident characteristics including accident type, accident causal factor, accident location, accident severity, number of vehicles involved, number killed, number injured, time accident occurred, and whether a truck was involved; 2) geometric characteristics in terms of number of lanes; and 3) environmental condition (i.e., whether road surface was wet).

The nominal variables that were hypothesized to affect total delay caused by accident are classified into binary variables. The accident time variable (in terms of time of day) is divided into four time intervals: 06:01 ~ 09:00, 09:01 ~ 15:30, 15:31 ~ 18:00, and 18:01 ~ 06:00. The accident time variables includ-

TABLE V
RESULTS OF NONPARAMETRIC ANALYSIS AT 5% LEVEL

Category	Variable	Samples	p-value*
Accident characteristics	Collision type	1 veh hit obj./overturn	248 0.000 (-)
		2+ veh hit obj./overturn	100 0.027 (-)
		2 veh weaving	405 0.001 (-)
		3+ veh rear end	542 0.000 (+)
Causal factor	Alcohol	68	0.001 (-)
	Speeding	1386	0.000 (+)
	Other violations	505	0.020 (-)
	Other than driver	63	0.000 (-)
Truck involved	No significant variable		N.A
Collision location	Left lane	642	0.023(+)
Severity	No significant variable		N.A
Number injured	No significant variable		N.A
Number vehicles involved	1	248	0.000(-)
	2	1302	0.004(-)
	3	489	0.000(+)
Number killed**	Not analyzed		N.A
Time of day	Night (18:01~06:00)	506	0.000(-)
	Midday (09:01~15:30)	741	0.002(+)
	PM peak (15:31~18:00)	523	0.000(+)
Day of week	Monday	383	0.002(-)
	Tuesday	442	0.005(+)
Geometric characteristics	Number lanes	6 lanes	67 0.041(-)
Environmental characteristics	Road surface	No significant variable	

* The sign in the parenthesis means the effect on delay; + implies that the covariate contributes to more delay and vice versa.

** The variable Number killed is not analyzed due to insufficient data (3 observations).

ing time of day and day of week reflect the general traffic pattern that is present.

Table V shows the result of a univariate analysis using the nonparametric survival analysis based on the log-rank tests and the KM estimates, at the level of 5%. Based on this result, we conclude that factors positively influential (i.e., producing more delay) on nonrecurrent total delay include the midday and P.M. peak periods as a function of time periods, Tuesday as a function of the day of week, 3 veh involved as a function of number vehicles involved, 3+ veh weaving and 3+ veh rear-end collision as a function of collision types, left lane as a function of collision locations, and speeding as a function of causal factors.

IV. CONCLUSION AND FUTURE STUDIES

This paper has proposed a method for the quantification of the total delay caused by a freeway accident based on applying loop detector data in the identification of the temporal and spatial extent of the congestion region. Unlike previous studies, the method used only the time of the accident and its location, together with loop detector data. As such, the method can be applied to any freeway system for which accident data are collected and that is instrumented with common inductance loop detectors. Using the procedures developed here, it is possible to develop a performance measure to evaluate transportation policies and planning level analyses associated with the design of

⁴“+” indicates right censoring.

transportation systems or the preparation of operating plans for safety, as well as for the evaluation of deployed transportation projects or technologies.

Estimated total delay using the procedure, including the effects of censoring, has been statistically analyzed for a case study involving accidents on a freeway system located in Orange County, CA. The log-rank test and KM method have been applied to investigate the variation of nonrecurrent total delay relative to a set of potential causal factors, as well as constructed (KM) survival curves. The results indicate that accidents with certain characteristics tend to lead to significantly more delay than do others.

This study employed an empirical method to find the best determinant threshold value discriminating two regions: congested regions affected by an accident and uncongested regions. Although the method utilized seems to be reasonable, more scientific approaches to set the best value of α need to be further developed in future research. In addition, since specially equipped facilities such as an aerial video camera will be needed to monitor the spatio-temporally congested area by the accident, it is very difficult and expensive to verify the result of the proposed methodology. Thus, future study is required to verify the methodology used in this study. For instance, a well-calibrated simulation model will be a good alternative method. Finally, although an accident can have an influence on other secondary accidents, the proposed method cannot directly distinguish secondary accident impact from primary accident impact. Therefore, another future research will be required to distinguish secondary accident impact from primary accident impact.

REFERENCES

- [1] *Highway Capacity Manual: Special Report 209*, 3rd ed. Washington, DC: TRB, NRC, 1994.
- [2] M. E. Goolsby and W. Smith, "Influence of incidents on freeway quality of service," in *Highway Research Record*. Washington, DC: HRB, NRC, 1971, pp. 41–46.
- [3] W.-M. Chow, "A study of traffic performance models under an incident condition," in *Highway Research Record*. Washington, DC: HRB, NRC, 1974, pp. 31–36.
- [4] J. M. Morales, "Analytical procedures for estimating freeway traffic congestion," *ITE J.*, vol. 57, no. 1, pp. 45–49, Jan. 1987.
- [5] T. W. Lawson, D. J. Lovell, and C. F. Daganzo, "Using input-output diagram to determine spatial and temporal extents of a queue upstream of a bottleneck," *Transp. Res. Rec.*, vol. 1572, pp. 140–147, 1997.
- [6] A. L. Erera, T. Lawson, and C. Daganzo, "Simple, generalized method for analysis of traffic queue upstream of a bottleneck," *Transp. Res. Rec.*, vol. 1646, pp. 132–140, 1998.
- [7] A. Skabardonis and N. Geroliminis, "Development and application of methodologies to estimate incident impacts," in *Proc. 2nd Int. Congr. Transp. Res.*, Athens, Greece, 2004.
- [8] C. S. Edward, "New model for predicting freeway incidents and incident delays," *J. Transp. Eng.*, vol. 123, no. 4, pp. 267–275, Jul./Aug. 1997.
- [9] K. Moskowitz and L. Newman, "Notes on freeway capacity," in *Highway Research Record*. Washington, DC: HRB, NRC, 1963, pp. 44–68.
- [10] S. C. Wirasinghe, "Determination of traffic delays from shock-wave analysis," *Transp. Res.*, vol. 12, no. 5, pp. 343–348, Oct. 1978.
- [11] B. Heydecker, "Incidents and Intervention on Freeways," California PATH, Berkeley, CA, UCB-ITS-PRR-94-5, 1994.
- [12] H. Al-Deek, A. Garib, and A. E. Radwan, "New method for estimating freeway incident congestion," *Transp. Res. Rec.*, vol. 1494, pp. 30–39, 1995.
- [13] M. Wohl and B. V. Martin, *Traffic System Analysis for Engineers and Planners*. New York: McGraw-Hill, 1967.
- [14] M. J. Lighthill and G. B. Whitham, "On kinematic waves. II. A theory of traffic flow on long crowded roads," *Proc. Roy. Soc. London Ser. A, Math. Phys. Sci.*, vol. 229, no. 1178, pp. 317–345, May 10, 1955.
- [15] A. Garib, A. E. Radwan, and H. Al-Deek, "Estimating magnitude and duration of incident delays," *J. Transp. Eng.*, vol. 123, no. 6, pp. 459–466, Nov./Dec. 1997.
- [16] A. Skabardonis, K. Petty, H. Noeimi, D. Ryzewski, and P. P. Varaiya, "I-880 field experiment: Data-base development and incident delay estimation procedures," *Transp. Res. Rec.*, vol. 1554, pp. 204–212, 1996.
- [17] A. Skabardonis, P. Varaiya, and K. Petty, "Measuring recurrent and non-recurrent traffic congestion," *Transp. Res. Rec.*, vol. 1856, pp. 118–124, 2003.
- [18] Y. Chung, "Quantification of nonrecurrent congestion delay caused by freeway accidents and analysis of causal factors," *Transp. Res. Rec.*, vol. 2229, pp. 8–18, 2011.
- [19] Y. Chung, "Assessment of non-recurrent traffic congestion caused by freeway work zones and its statistical analysis with unobserved heterogeneity," *Transp. Policy*, vol. 18, no. 4, pp. 587–594, Aug. 2011.
- [20] Y. Chung, H. Kim, and M. Park. (2011). Quantifying non-recurrent traffic congestion caused by freeway work zones using archived work zone and ITS traffic data. *Transportmetrica*. [Online]. Available: <http://dx.doi.org/10.1080/18128602.2010.522609>
- [21] D. Schrank and T. Lomax, "The 2003 Annual Urban Mobility Report," Texas Transp. Inst., Texas A&M Univ., College Station, TX, Tech. Rep., 2003.
- [22] Y. Chung, "Assessment of non-recurrent congestion caused by precipitation using archived weather and traffic flow data," *Transp. Policy*, vol. 19, no. 1, pp. 167–173, Jan. 2012.
- [23] R. Pal and K. C. Sinha, "Simulation model for evaluating and improving effectiveness of freeway service patrol programs," *J. Transp. Eng.*, vol. 128, no. 4, pp. 355–365, Jul./Aug. 2002.
- [24] K. Ozbay and B. Bartin, "Incident management simulation," *Simulation*, vol. 79, no. 2, pp. 69–82, Feb. 2003.
- [25] Y. Ma, M. Chowdhury, R. Fries, and K. Ozbay, "Harnessing the power of microscopic simulation to evaluate freeway service patrols," *J. Transp. Eng.*, vol. 135, no. 7, pp. 427–439, Jul. 2009.
- [26] K. Smith and B. L. Smith, "Forecasting the Clearance Time of Freeway Accidents," Center Transp. Studies, Univ. Virginia, Charlottesville, VA, Rep. STL-2001-01, 2001.
- [27] F. Hall and B. Persaud, "Evaluation of speed estimates made with single-detector data from freeway traffic management systems," *Transp. Res. Rec.*, vol. 1232, pp. 9–16, 1989.
- [28] Z. Jia, C. Chen, B. Coifman, and P. Varaiya, "The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors," in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2001, pp. 536–541.
- [29] Y. Wang and N. L. Nihan, "Can single-loop detectors do the work of dual-loop detectors?" *J. Transp. Eng.*, vol. 129, no. 2, pp. 169–176, Mar./Apr. 2003.
- [30] M. A. Cleves, W. Gould, and R. Gutierrez, *An Introduction to Survival Analysis Using Stata*, Rev. ed. College Station, TX: Stata Press, 2004.



Younshik Chung received the M.S. degree from Ajou University, Suwon, Korea, in 1999 and the Ph.D. degree in civil and environmental engineering from the University of California, Irvine, in 2007.

He has been with the Department of Highway, ITS, and Transport Research, Korea Transport Institute, Goyang, Korea, as a Research Associate, since 2007. His current research interests include traffic safety, intelligent transport systems, traffic control and operation, and sustainable transportation.



Wilfred W. Recker received the B.S., M.S., and Ph.D. degrees from Carnegie Mellon University, Pittsburgh, PA, in 1964, 1966, and 1968, respectively.

He is currently a Professor with the Department of Civil Engineering, University of California, Irvine. From 1983 to 2006, he served as Director of the Institute of Transportation Studies, University of California, Irvine. His research interests include travel demand models, disaggregate modeling methodologies, and transportation system analysis.