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 solvable 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

A LTHOUGH 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 recurring 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...
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 $(j+1), (j+2), \ldots$, whereas those upstream are consecutively labeled as $(j-1), (j-2), \ldots$.

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 $T$ observations 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 $\Omega_{jn} = \Omega(s_{jn}(t_m), \sigma_{s_{jn}(t_m)})$ 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 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 $t_m \geq t_1$, $m = 1, 2, \ldots$ (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

---

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.
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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.

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 shown in Fig. 2. Based on this interpretation, the only data relevant to the current example 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 extent of the shockwave by assuming the accident speed behind discriminating between these two regions is to compare the accident-free speed distribution with the observed speeds.
the problem of determining the “best” set of “dot-shaded (or yellow)” cells can be formulated as the following statement:

\[
\sum_{\forall \text{dot-shaded cells}} P_{jm} + \sum_{\forall \text{shaded cells}} (1 - P_{jm}) = \text{Minimum.}
\]  

(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

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

(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 \( j \) 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:

\[
\delta_{j+k,m} \leq \left[ 1 - (\delta_{j,m} - \delta_{j+1,m}) \right] \cdot R \quad \forall j, m \forall k \leq J - j \quad (a)
\]

\[
\delta_{j,m+r} \leq \left[ 1 - (\delta_{j,m} - \delta_{j,m+1}) \right] \cdot R \quad \forall j, m \forall r \leq M - m \quad (b)
\]

\[
\delta_{j,m+k} \leq \left[ 1 - (\delta_{j,m} - \delta_{j+1,m}) \right] \cdot R \quad \forall j, m \forall k \leq M - m \quad (c)
\]

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:

\[
\text{Min } Z = \sum_{j,m} \left[ P_{jm} \cdot \delta_{jm} + (1 - P_{jm}) \cdot (1 - \delta_{jm}) \right]
\]

s.t.:

\[
\delta_{j+k,m} \leq \left[ 1 - (\delta_{j,m} - \delta_{j+1,m}) \right] \cdot R \quad \forall j, m \forall k \leq J - j
\]

\[
\delta_{j,m+r} \leq \left[ 1 - (\delta_{j,m} - \delta_{j,m+1}) \right] \cdot R \quad \forall j, m \forall r \leq M - m
\]

\[
\delta_{j,m+k} \leq \left[ 1 - (\delta_{j,m} - \delta_{j+1,m}) \right] \cdot R \quad \forall j, m \forall k \leq M - m
\]

\[
\delta_{jm} = \begin{cases} 
0 & \text{if } \dot{\delta}_{jm} \leq 0.5; \\
1 & \text{otherwise.}
\end{cases}
\]

(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

\[
TD = \sum_{\forall j, m \in \text{dot-shaded cells}} \max \left\{ L_j \cdot \left[ \frac{1}{\hat{s}_j(t_m)} - \frac{1}{s_j(t_m)} \right] \cdot V_{jm}, 0 \right\} 
\]

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 \( m \);
- \( s_j(t_m) \): 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., \( t_m = 5 \text{ min} \)) 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 \( s_{jn}(t_m) \), \( n = 1, 2, \ldots, n_{\text{obs}}, n_{\text{obs}} \leq T \), 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, sidewipe, 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

\[
\bar{s} = \frac{q}{5280 \times occ}
\]

(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

\[
\hat{g}_i(t+1) = p \cdot \hat{g}_i(t) + (1 - p) \cdot g_i(t + 1)
\]

(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.,

\[
s_{jn}(t_m) = g_{jn}(t_m) \frac{q_{jn}(t_m)}{5280 \times occ_{jn}(t_m)}, \quad n = 1, 2, \ldots, 52.
\]

(9)

In the definition of \( P_{jn} \), the threshold value \( \alpha \) 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 \( n_{\text{min obs}} \). Since 30 is commonly used as the minimum number of observations required for the law of large numbers to apply, it was determined that \( n_{\text{min obs}} = 30 \). 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), 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

---

\(^2\)Although different \( \alpha \) 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 \( \alpha \) 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 \( \alpha \) 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.

\(^3\)Such 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. 10. 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 1379.49 + 4 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 including 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

### TABLE V

RESULTS OF NONPARAMETRIC ANALYSIS AT 5% LEVEL

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

| † 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). |

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 $\alpha$ 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


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.