Title
Components of Congestion: Delay from Incidents, Special Events, Lane Closures, Weather, Potential Ramp Metering Gain, and Excess Demand

Authors
Kwon, Jaimyoung
Mauch, Michael
Varaiya, Pravin

Publication Date
2006

Components of Congestion

Delay from Incidents, Special Events, Lane Closures, Weather, Potential Ramp Metering Gain, and Excess Demand

Jaimyoung Kwon, Michael Mauch, and Pravin Varaiya

A method is presented to divide the total congestion delay in a freeway section into six components: the delay caused by incidents, special events, lane closures, and adverse weather; the potential reduction in delay at bottlenecks that ideal ramp metering can achieve; and the remaining delay, caused mainly by excess demand. The fully automated method involves two steps. First, the components of nonrecurrent congestion are estimated by statistical regression. Second, the method locates all bottlenecks and estimates the potential reduction in delay that ideal ramp metering can achieve. The method can be applied to any site with minimum calibration. It requires data about traffic volume and speed; the time and location of incidents, special events, and lane closures; and adverse weather. Applied to a 45-mi section of I-880 in the San Francisco Bay Area in California, the method reveals that incidents, special events, rain, potential reduction by ideal ramp metering, and excess demand respectively account for 13.3%, 4.5%, 1.6%, 33.2%, and 47.4%, respectively, of the total daily delay. The delay distribution of the various components is different between the morning and evening peak periods and between the two freeway directions. Quantifying the components of congestion at individual freeway sites is essential for developing effective congestion mitigation strategies.

Congestion is caused by incidents, special events, lane closures, weather, inefficient operations, and excess demand. The impact of these can be summarized in the division of the congestion pie into its components, as in Figure 1. Knowledge of the congestion pie is essential to the selection of effective congestion mitigation strategies (*1*).

This paper presents a method for dividing the total congestion D_{total} into six components:

- D_{col}, congestion caused by incidents, which could be reduced by quicker response;
- D_{event}, congestion caused by special events, which could be reduced by public information and coordination with transit;
- D_{lane}, congestion caused by lane closures, which could be reduced by better scheduling of lane closures;
- $D_{weather}$, congestion caused by adverse weather, which could be reduced by demand management and a better weather response system;
- D_{pot}, congestion that can be eliminated by ideal ramp metering; and
- D_{excess}, residual delay, largely caused by demand that exceeds the maximum sustainable flow.

The method is applied to a 45-mi section of I-880 in the San Francisco Bay Area with data for January through June, 2004.

The method refines previous studies that group D_{pot} and D_{excess} together as recurrent congestion (*2–4*). It also refines the authors' previous work, which considers only three components (D_{col}, D_{pot}, and D_{excess}) (*5*). Transportation agencies measure recurrent congestion in various ways and find it accounts for 40% to 70% of total congestion (*6*). The availability of more comprehensive data has prompted attempts to separately estimate the contribution of different causes of congestion. Some studies divide total congestion into recurrent and nonrecurrent congestion, and some studies divide the nonrecurrent congestion into accident-induced congestion and other incident-induced congestion. There also are estimates of the congestion caused by adverse weather. These studies are reviewed in the next section.

The studies leave a large fraction (between 40% and 70%) of the total congestion unexplained. This unexplained residual is often called recurrent congestion. As Hallenbeck et al. observed, "Many large delays still occur for which incidents are not responsible, and for which no 'cause' is present in the [data]" (*2*). They suggested that one cause of these delays may be "unusual volume surges at ramps . . . that are not being effectively handled by the ramp metering program" (*2*, p. 11). The proposed method estimates this potential reduction in delay, D_{pot}.

PREVIOUS STUDIES

Transportation agencies until recently reported only recurrent congestion (*7, 6*). The availability of more comprehensive data has inspired studies to quantify the relative impact of different causes of congestion.

Several studies estimate the impact of incidents. The earliest studies relied on correlating incident data collected by using floating cars with data from loop detectors (*8*). These data provide a great deal of information about the nature of incidents, but the data collection efforts are too expensive to replicate on a large scale or on a continuing basis.

Date from California Highway Patrol computer-aided dispatch and Freeway Service Patrol (FSP) logs were used to evaluate FSP effectiveness on Los Angeles freeways (*9*) and in Oregon (*10*). These studies need much human effort, data analysis skill, and subjective

J. Kwon, Department of Statistics, California State University, East Bay, Hayward, CA 94542. M. Mauch, DKS Associates, Suite 340, 8950 Cal Center Drive, Sacramento, CA 95826-3225. P. Varaiya, Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720.

Transportation Research Record: Journal of the Transportation Research Board, No. 1959, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 84–91.

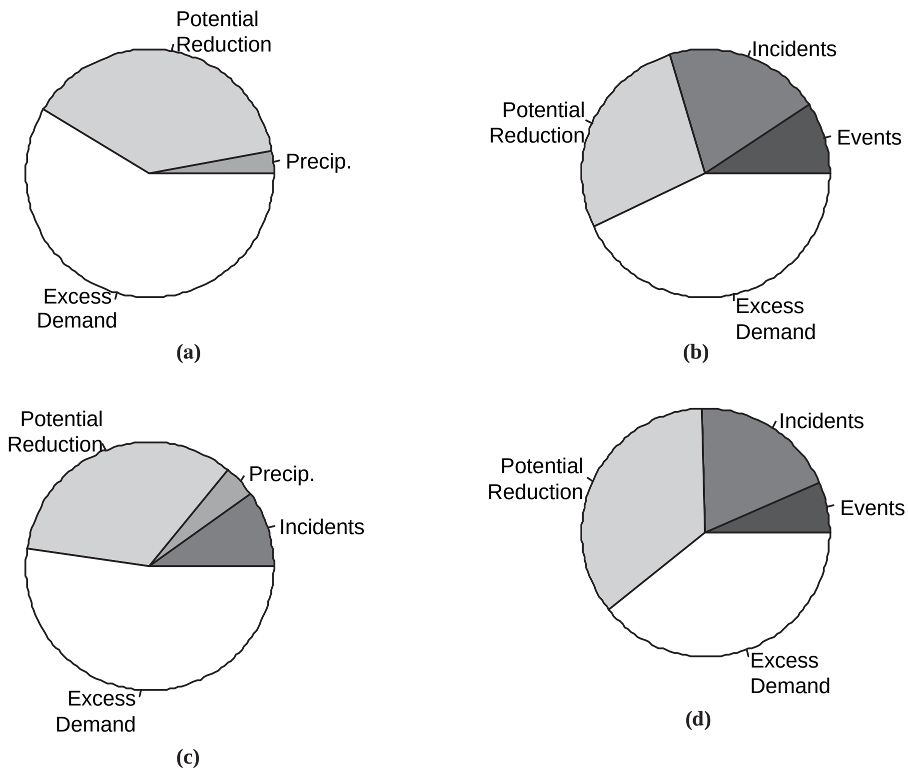

FIGURE 1 Congestion pie chart for four scenarios on I-880: (*a*) NB AM, (*b*) NB PM, (*c*) SB AM, and (*d*) SB PM.

judgment to determine the spatial and temporal region of the congestion impact of an incident. The authors' previous work developed an automated method with which to delineate an incident's impact region (*5*), but that approach requires accurate time and location of incidents, which may not be available.

Determining every individual incident's impact region can be avoided if one is willing to average out the impact of individual incidents, as done by Hallenbeck et al. and Skabardonis et al. (*2, 3*), who separated nonrecurrent and recurrent congestion, but who differed in definition and method.

Skabardonis et al. considered a freeway section during a peak period (*3*). The total congestion on each of several days is calculated as the additional vehicle hours spent driving below 60 mph (see Equation 1). Each day is classified as incident-free or incident-present. The average congestion in incident-free days is defined to be the recurrent delay. Total congestion in incident-present days is considered to be the sum of recurrent and incident-induced congestion. Subtracting average recurrent congestion from this gives an estimate of the average nonrecurrent or incident-induced congestion. On the other hand, Hallenbeck et al. took the median traffic conditions on days when a freeway section did not experience lane-blocking incidents as the "expected, recurring condition" (*2*).

A less data-intensive approach was taken by Bremmer et al. (*4*). In the absence of incident data, they simply assumed that an incident has occurred if a trip "takes twice as long as a free-flow trip for that route." The aim for this study was to forecast travel times, measure travel time reliability, and conduct cost–benefit analysis of operational improvements, rather than to measure the congestion contribution of different causes.

The impact of inclement weather on freeway congestion was addressed in the *Highway Capacity Manual* (*11,* Chapter 22) and by Chin et al. (*12*). These found that light rain or snow, heavy rain, and heavy snow reduce traffic speed by 10%, 16%, and 40%, respectively.

PROPOSED METHOD

The method applies to a contiguous section of freeway with n detectors indexed $i = 1, \ldots, n,$ whose flow (volume) and speed measurements are averaged over 5-min intervals indexed $t = 1, \ldots, T$. Days in the study period are denoted by $d = 1,2, \ldots, N$. Detector i is located at postmile x_i; $v_i(d,t) = v(x_i, d, t)$ is its speed (mph) and $q_i(d,t) = q(x_i, d, t)$ is its flow [vehicles per hour (vph)] at time t of day d.

The n detectors divide the freeway into n segments. Each segment's (congestion) delay is defined as the additional vehicle hours traveled driving below free-flow speed v_{ref}, taken to be 60 mph. Thus the delay in segment i in time t is

$$D_i(d,t) = l_i \times q_i(d,t) \times \max\left[1/v_i(d,t) - 1/v_{\text{ref}}, 0\right] \text{ vehicle hours} \quad (1)$$

in which l_i is the segment length in miles. The total delay in the freeway section on day d is the delay over all segments and times,

$$D_{\text{total}}(d) = \sum_{i=1}^{n} \sum_{t=1}^{T} D_i(d,t) \quad (2)$$

The average daily total delay is simply

$$D_{\text{total}} = \frac{1}{N}\sum_{d=1}^{N} D_{\text{total}}(d) \qquad (3)$$

The following application separately considers the daily delay over two peak periods: 5:00 to 10:00 a.m. for the morning peak and 3:00 to 8:00 p.m. for the afternoon peak.

Incidents are indexed $a = 1, 2, \ldots$. The time τ_a when incident a occurs and its location σ_a are approximately known. The incident clearance time and the spatial and temporal region of the incident's impact are not known.

Decomposition of Delay

The method divides the average daily total delay, Equation 3, into six components:

$$D_{\text{total}} = D_{\text{col}} + D_{\text{event}} + D_{\text{lane}} + D_{\text{weather}} + D_{\text{pot}} + D_{\text{excess}} \qquad (4)$$

It will be useful to define

$$D_{\text{non-rec}} = D_{\text{col}} + D_{\text{event}} + D_{\text{lane}} + D_{\text{weather}} \qquad (5)$$

$$D_{\text{rec}} = D_{\text{tot}} - D_{\text{non-rec}} = D_{\text{pot}} + D_{\text{excess}} \qquad (6)$$

where

D_{col} = the daily delay caused by incidents,
D_{event} = daily delay caused by special events;
D_{lane} = daily delay caused by lane closure;
D_{weather} = daily delay caused by adverse weather condition;
D_{pot} = potential reduction of D_{rec} by ramp metering;
D_{excess} = residual delay, attributed mostly to excess demand;
D_{rec} = daily recurrent delay; and
$D_{\text{non-rec}}$ = daily nonrecurrent delay.

D_{total}, calculated from flow and speed data, is the average daily total delay. D_{col}, D_{event}, D_{lane}, and D_{weather} are components of so-called nonrecurrent congestion. The difference between their sum and D_{total} is the recurrent congestion (*2, 3*). A portion of recurrent congestion due to frequently occurring bottlenecks in principle could be reduced by ramp metering. That potential reduction is estimated as D_{pot}. The remaining delay, D_{excess}, is caused by all other causes, most of which is likely due to demand in excess of the maximum sustainable flow. The delay caused by excess demand can be reduced only by changing trip patterns.

Nonrecurrent Delays

The components of nonrecurrent delay are identified with the use of the model

$$D_{\text{total}}(d) = \beta_0 + \beta_{\text{col}}X_{\text{col}}(d) + \beta_{\text{event}}X_{\text{event}}(d) + \beta_{\text{lane}}X_{\text{lane}}(d) + \beta_{\text{weather}}X_{\text{weather}}(d) + \epsilon(d) \qquad (7)$$

where

$\epsilon(d)$ = the error term with mean zero,
$X_{\text{col}}(d)$ = the number of incidents on day *d,*
$X_{\text{event}}(d)$ = the number of congestion-inducing special events such as sport games on day *d,*
$X_{\text{lane}}(d)$ = the number of lane-closures on day *d,* and
$X_{\text{weather}}(d)$ = the 0-1 indicator of adverse weather condition on day *d.*

These explanatory variables are used in the application, but the list could be augmented if additional data are available. For example, $X_{\text{event}}(d)$ could be the attendance at special events instead of the number of special events; $X_{\text{lane}}(d)$ could be the duration instead of the number of lane closures; and $X_{\text{weather}}(d)$ could be the precipitation (as in the application).

The model assumes that each incident, special event, lane closure, and adverse weather condition contributes linearly to the delay. Figure 2 illustrates that such model is reasonable for the study site. More complicated causality between explanatory variables, such as between the bad weather and the number of accidents, is not considered, to keep the number of parameters in the model small. But if one has enough data and the interaction is strong enough, such interaction terms could be included. (For the San Francisco Bay Area data, the correlation coefficient between precipitation and number of accidents is only 0.032.)

Fitting the model to the data via linear least squares gives the parameter estimates, again denoted β_0, β_{col}, β_{event}, β_{lane}, and β_{weather}. The components of the total delay then are

$$D_{\text{col}} = \beta_{\text{col}} \times \text{avg}\left[X_{\text{col}}(d)\right] \qquad (8)$$

$$D_{\text{event}} = \beta_{\text{event}} \times \text{avg}\left[X_{\text{event}}(d)\right] \qquad (9)$$

$$D_{\text{lane}} = \beta_{\text{lane}} \times \text{avg}\left[X_{\text{lane}}(d)\right] \qquad (10)$$

and

$$D_{\text{weather}} = \beta_{\text{weather}} \times \text{avg}\left[X_{\text{weather}}(d)\right] \qquad (11)$$

in which the average is taken over days, $d = 1, \ldots, N$.

The intercept β_0 in Equation 7 is the delay when there are no incidents, special events, lane closures, or adverse weather. Thus, consistent with convention, it may be identified with recurrent congestion, since it equals total delay minus the nonrecurrent delay $D_{\text{non-rec}}$,

$$\beta_0 = D_{\text{rec}} = D_{\text{total}} - D_{\text{non-rec}} \qquad (12)$$

Recurrent Delay Algorithm: Separating Recurrent and Nonrecurrent Congestion

The next step is to divide the recurrent delay into the delay that can be eliminated by ramp metering and the delay caused by excess demand. For this, the method identifies recurrent bottlenecks on the freeway section by using the automatic bottleneck identification algorithm proposed by Chen et al. (*13*). Then the ideal ramp metering (IRM) is run on those recurrent bottlenecks that are activated on more than 20% of the weekdays considered (*14, 5*).

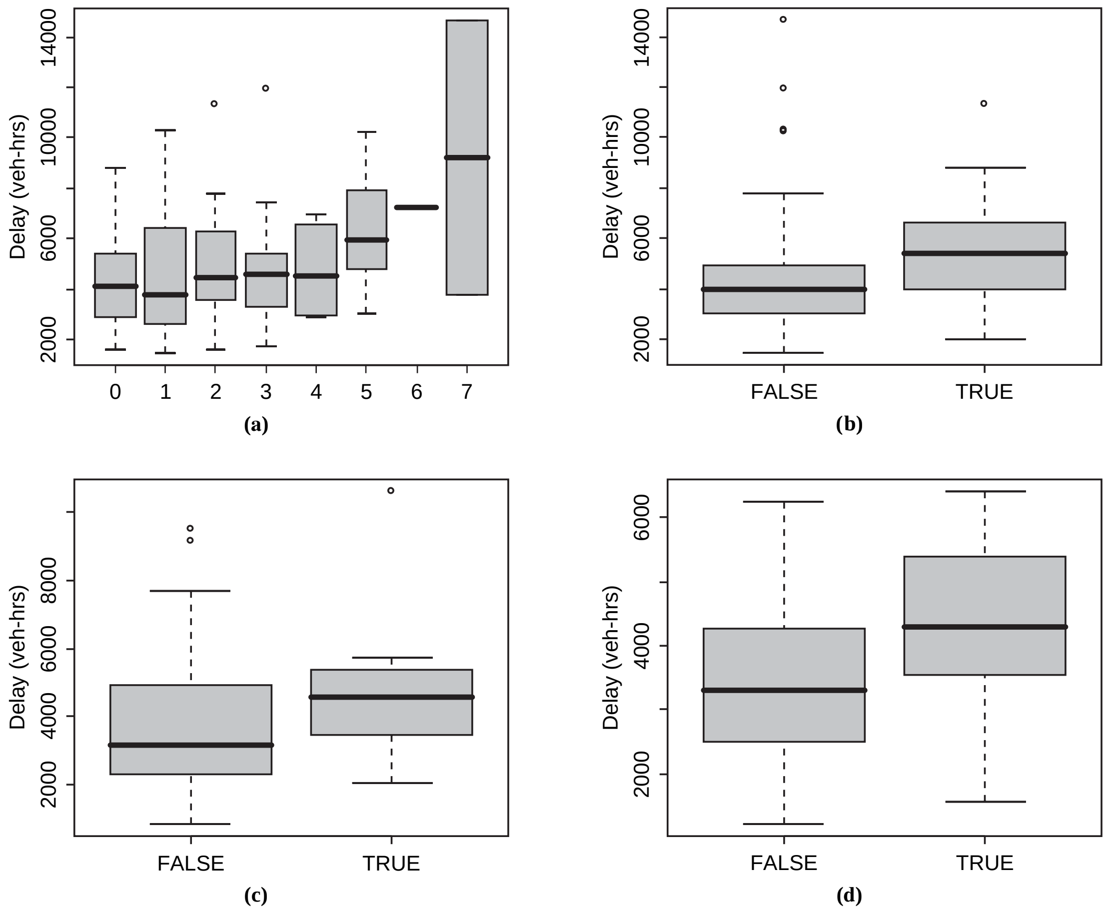

FIGURE 2 Relationship between delay and selected factors. Distribution of average daily total delay $D_{total}(d)$, summarized as box-and-whisker plot, is shown for each level of (*a*) number of incidents, (*b*) special event occurrence, and (*c*), (*d*) adverse weather condition.

Here is a brief description of the IRM algorithm. For a specific recurrent bottleneck, let segment *i* and *j* be the upstream and downstream boundaries of the bottleneck, respectively. For the upstream boundary *j,* the median queue length of the bottleneck is used. Then the total peak period volume at the two locations is computed. The difference between the two would be the difference between the total number of cars entering or exiting the freeway between the two segments. It is assumed that all those cars contributing to the difference are arriving (or leaving) at a virtual on ramp (off ramp) at the upstream segment *i.* Also, the time-series profile of that extra traffic is assumed identical to the average of those at segment *i* and *j.* This allows one to compute the modified total input volume profile at the segment *i.* The capacity of the whole section is the maximum sustainable (over 15 min) throughput at location *j,* and this is computed from the empirical data. The virtual input volume is monitored at segment *i* at 90% of C_j to prevent the breakdown of the system, assuming that the metered traffic will be free flow (60 mph) throughout the freeway section and that the upstream meter has infinite capacity.

Thus, under IRM, the delay occurs only at the meters. The potential savings from IRM at these bottlenecks for each day *d* is then computed as

$$D_{pot}(d) = D_{BN,\ before\ IRM}(d) - D_{BN,\ after\ IRM}(d) \tag{13}$$

Here $D_{BN,\ before\ IRM}(d)$ and $D_{BN,\ after\ IRM}(d)$ is the delay at the bottlenecks before and after IRM is run. The average daily potential saving is

$$D_{pot} = \min\ \left\{\text{median}\left[D_{pot}(d),\ d = 1, \ldots\right],\ D_{rec}\right\} \tag{14}$$

In Equation 14, the median instead of the mean is used to ensure that the influence of incidents, special events, and so forth, is minimized in the computation. Also, the potential saving cannot be larger than the total recurrent delay D_{rec}.

Because of the ideal nature of IRM, D_{pot} needs to be interpreted with caution. The assumption of a large, though not infinite, capacity at the meter is not realistic for many urban freeways, and metering at certain locations can lead to breakdown of arterial traffic nearby. Thus, it is recommended that D_{pot} be viewed as the maximum possible saving in the recurrent delay by metering.

Congestion Pie

The described method divides the average daily total delay D_{total} into six components, which are summarized in easily understood pie charts, like those in Figure 1.

CASE STUDY

The method is applied to a 45.33-mi (Postmile .39 to 45.72) section of southbound (SB) and northbound (NB) I-880 in the San Francisco Bay Area. Two periods are considered: morning peak (AM), 5:00 to 10:00 a.m., and evening peak (PM), 3:00 to 8:00 p.m. Data cover 110 weekdays from January 5 to June 30, 2004. There are four scenarios, distinguished by peak period and freeway direction: SB AM, SB PM, NB AM, and NB PM.

Data Sources

Traffic Speed and Volume Data

The 90 (NB) and 94 (SB) loop detector stations in the section provide 5-min lane-aggregated volume and speed data, available at the PeMS website (*15*).

FSP Incidents

Incident data are for FSP-assisted incidents. On an average nonholiday weekday, the FSP assists upward of 80 motorists on I-880 during the 6:00–10:00 a.m. and 3:00–7:00 p.m. periods. FSP peaks are an hour shorter than peaks used for computing total delay (5:00 to 10:00 a.m. and 3:00 to 8:00 p.m.), but the effect is not expected to be substantial. On weekends and holidays, FSP assistance is not provided. FSP drivers record the date and time, duration, freeway name and direction, incident description (e.g., traffic accident, flat tire, out-of-gas), and location (e.g., on or off ramp, left shoulder, right shoulder, in-lane). Here, only in-lane incidents (as opposed to those on the left or right shoulder or on a ramp) during peak hours are considered. There were 829 such incidents during the study period.

Special Events

On 45 of 110 weekdays, there were special events in the Oakland Coliseum, near Postmile 36 of I-880, including baseball (the Oakland A's) and basketball (the Golden State Warriors) games and show performances, most starting at 7:00 p.m. Data were provided by Networks Associates Coliseum and The Arena in Oakland.

Weather

Weather data were collected from California Department of Water Resources for the Oakland North station (*16*). The station reports daily precipitation, temperature, wind speed and direction, and so forth; only precipitation was considered in the analysis.

Lane Closure

Lane closure data were obtained from the lane closure system (LCS) managed by the California Department of Transportation (*17*). LCS records include for each lane closure

- Location—freeway, direction, county, and postmile;
- Begin and end date and time;
- Facility and lanes—on and off ramp, number of lanes, which lanes; and
- Type of work—sweeping, construction, and so forth.

For the first half of 2004, for NB I-880, there were 224 lane closures, 126 of them in the traffic lanes. All daytime closures were "sweeping" or "call box remove/repair" involving a moving closure of at most one lane, and have negligible impact on congestion. All congestion-inducing lane closures (repair, striping, and paving) occurred at night (after 10:00 p.m. and before 5:00 a.m.) or on weekends outside the morning and evening peaks. This was also the case for SB I-880. Thus $D_{\text{lane}} = 0$ is assigned for all scenarios.

Results

Table 1 summarizes the regression results for nonrecurrent congestion. The last column shows the multiple *R*-squared values for each sce-

TABLE 1 Regression Results for Nonrecurrent Delay

Scenario	Factor	Estimate	Std. Error	*t*-Value	Pr(> |t|)[a]	Multiple R^2
NB AM	(Intercept)	3,301.1	191.1	17.28	0.000***	0.12
	Event	−221.5	216.2	−1.03	0.308	
	Incident	115.8	74.2	1.56	0.122	
	Weather	1,305.7	384.4	3.40	0.001***	
NB PM	(Intercept)	3,419.7	408.1	8.38	0.000***	0.14
	Event	1,084.6	416.0	2.61	0.010*	
	Incident	486.1	133.9	3.63	0.000***	
	Weather	75.4	732.7	0.10	0.918	
SB AM	(Intercept)	3,402.6	339.6	10.02	0.000***	0.17
	Event	−482.0	342.2	−1.41	0.162	
	Incident	221.1	127.6	1.73	0.086	
	Weather	2,125.6	598.5	3.55	0.001***	
SB PM	(Intercept)	3,311.1	374.8	8.83	0.000***	0.12
	Event	705.5	419.9	1.68	0.096	
	Incident	383.8	116.9	3.28	0.001**	
	Weather	28.7	751.3	0.04	0.970	

[a]Significance codes ***, **, * and "." mean the *P*-value is between 0 and .001, between .001 and .01, between .01 and .05, and between .05 and .1, respectively.

nario, which is the ratio of the sum of squares of the delay explained by the regression model and the total sum of squares around the mean. The F-statistic for testing whether the fit of the model is valid is significant with practically zero P-value for all four scenarios; this suggests that the linear regression model successfully explains the delay variation. The following also were observed:

- β_{event} is statistically significant (P-value $< .10$) only for evening shifts. This is to be expected since most special events occur in the afternoon or evening. Each special event, on the average, contributes a delay of 1,084 and 705.5 vehicle hours for NB and SB, respectively.
- β_{col} is statistically significant (P-value $< .001$) only for evening shifts. This suggests that congestion in the morning peak hours is more recurrent than in the afternoon or evening. In evening shifts, each incident contributes a delay of 486.13 (NB) and 383.75 (SB) vehicle hours on average.
- $\beta_{weather}$ is statistically significant (P-value $< .001$) only during morning shifts. On average, 1 in. rain adds 1305.7 (NB) and 2125.6 (SB) vehicle hours of delay. It rained on 29 of 110 weekdays; the median precipitation was 0.13 in. and the maximum was 2.44 in.

Figure 2 shows the relationship between D_{total} and some explanatory variables illustrating the correlation between the total delay and those variables.

Next, Formulas 8 through 11 are used to compute the delay components shown in Table 2. Before the formula is applied, those regression coefficients that are not statistically significant at significance level 0.1 are set to zero.

The automatic bottleneck detection algorithm is applied to speed data of the kind whose contour plot is shown in Figure 3. Clearly visible in the figure are a morning bottleneck near Postmile 10 and a larger evening bottleneck near Postmile 27. D_{pot} and D_{excess} are computed from the IRM algorithm and are shown in the right columns of Table 2. About 44% of recurrent delay potentially could be eliminated by ideal ramp metering (D_{pot} and D_{excess} are extrapolated from district-wide quantities; freeway-specific computation is under way in PeMS Version 6.0.)

From the charts in Figure 1 one can conclude the following:

- One-third of the congestion delay occurs at recurrent bottlenecks and can be potentially eliminated by ideal ramp metering.
- One-half of the delay is caused by excess demand in both directions and can be reduced only by changing trip patterns.
- Incidents and special events contribute 18% of the delay. The former can be reduced by more rapid detection and response; impact of special events may be reduced by information on changeable message signs.

The 486.13 (NB) and 383.75 (SB) vehicle hours of delay per incident for the evening shift are in rough agreement with other estimates. A regression of total daily delay versus number of accidents for all of Los Angeles yields a slope of 560 vehicle hours per accident (*18,* p. 20). For southbound I-5 in Seattle, Hallenbeck et al. find that a lane-blocking incident causes between 318 (conservative estimate) and 591 (liberal estimate) vehicle hours of delay (*2,* p. 15).

The average daily delay caused by incidents, D_{col}, is 986 and 837 vehicle hours, which is 20.3% and 18.8% of total evening delay for NB and SB, respectively. By way of comparison, Hallenbeck et al. found that "for the urban freeways examined [in the Central Puget Sound region of Washington State] lane-blocking incidents are responsible for between 2 and 20 percent of total daily delay" (*2,* p. 8). These average numbers must be used with caution because the delay impact of incidents varies considerably from freeway to freeway and during different times of day. For example, in the present study, during the morning peak (5:00 to 10:00), the average incident-induced delay is 0 (because β_{col} is not significantly different from 0) for NB and 9.9% of the total peak hour delay for SB.

TABLE 2 Delay Contributions from Each Cause and Congestion Pie

Scenario	Factor	β	Mean Weekday Occurrences	Delay Contributions (veh hrs)	Factor, After Bottleneck Analysis	Delay Contributions (veh hrs)	Total Delay (%)
NB AM	Recurrent	3,301	NA	3,301	Pot	1,307	38.4
		NA	NA	NA	Excess	1,994	58.6
	Event	0	0.42	0	Event	0	0.0
	Incident	0	1.55	0	Incident	0	0.0
	Weather	1,306	0.08	102	Weather	102	3.0
NB PM	Recurrent	3,420	NA	3,420	Pot	1,336	27.5
		NA	NA	NA	Excess	2,084	42.9
	Event	1,085	0.42	454	Event	454	9.3
	Incident	486	2.03	986	Incident	986	20.3
	Weather	0	0.08	0	Weather	0	0.0
SB AM	Recurrent	3,403	NA	3,403	Pot	1,327	33.5
		NA	NA	NA	Excess	2,076	52.4
	Event	0	0.42	0	Event	0	0.0
	Incident	221	1.78	394	Incident	394	9.9
	Weather	2,126	0.08	166	Weather	166	4.2
SB PM	Recurrent	3,311	NA	3,311	Pot	1,565	35.2
		NA	NA	NA	Excess	1,746	39.3
	Event	705	0.42	295	Event	295	6.6
	Incident	384	2.18	837	Incident	837	18.8
	Weather	0	0.08	0	Weather	0	0.0

NA means the number is not needed.

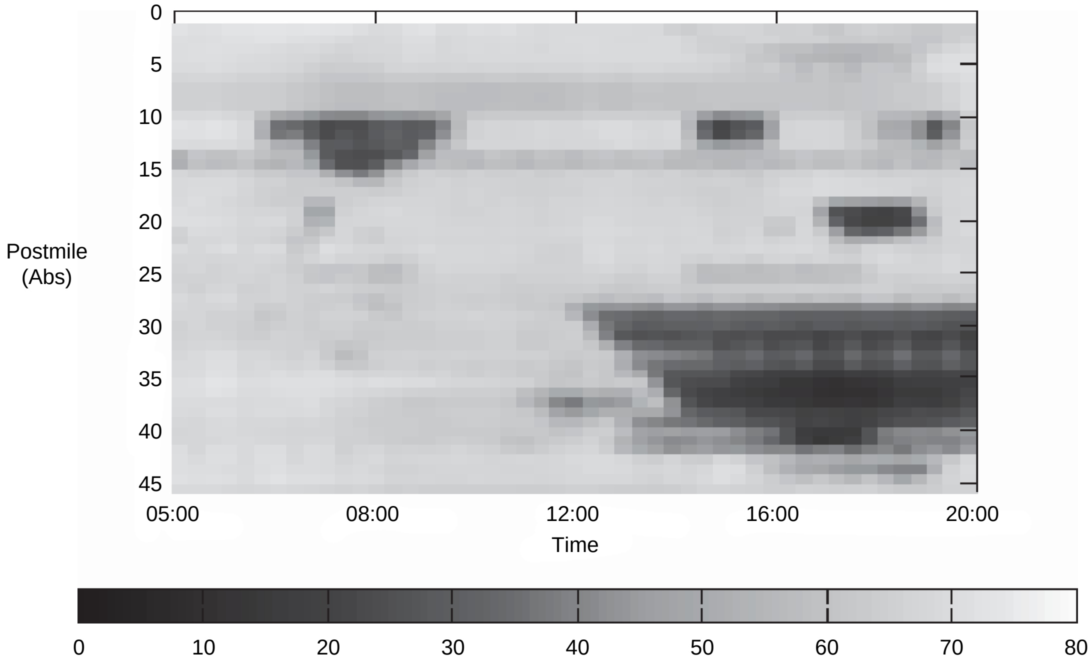

FIGURE 3 Lane-aggregated speed by postmile and time of day for I-880 SB on April 2, 2004. Traffic flows from bottom to top.

Aggregating over both peaks and both directions, the delay components are 13.3%, 4.5%, 1.6%, 33.2%, and 47.4% for incidents, special events, rain, potential reduction, and excess demand, respectively.

CONCLUSION

Between 1980 and 1999, highway route miles increased 1.5% whereas vehicle miles of travel increased 76% (*1*). In 2000, the 75 largest metropolitan areas experienced 3.6 billion hours of delay, resulting in $67.5 billion in lost productivity, according to the Texas Transportation Institute. Mitigating congestion through more efficient operations is a priority of transportation agencies. The first step in designing an effective mitigation strategy is to know how much each cause contributes to congestion. One can then design a set of action plans, each aimed at reducing the contribution of a particular cause. The more detailed the set of causes that are considered, the more effective the strategy that can be devised.

This paper proposed a fully automated method that calculates six components of congestion: delay attributed to incidents, special events, lane closures, and weather; delay that can be eliminated by ramp metering; and the remaining delay, mostly caused by excess demand.

The method was applied to a 45-mi section of I-880 in the San Francisco Bay Area for morning and evening peaks and for both directions. Incidents and special events together account for 17.8% of total delay. Lane closures caused no delay because delay-causing closures were not scheduled during peak hours. Rain caused 1.6% of total delay. A surprisingly large 33% of all delay could be eliminated by ideal ramp metering. Finally, 47% of the delay is caused by excess demand. Certainly, as discussed in the text, the 33% potential reduction due to metering needs to be interpreted with caution, as the maximum possible reduction. Even with such precaution, if these estimates are supported in more detailed studies, it is likely that most congestion mitigation strategies would harvest large potential gains from ramp metering.

ACKNOWLEDGMENTS

The authors appreciate the comments and criticism of John Wolf, Fred Dial, Jose D. Perez, and Lisa Davies of the California Department of Transportation; Tarek Hatata of System Metrics Group; and Alex Skabardonis and Karl Petty of the PeMS Development Group. This study was supported by grants from the California Department of Transportation to the California PATH program.

REFERENCES

1. Congestion Mitigation. FHWA, U.S. Department of Transportation. www.fhwa.dot.gov/congestion/congest2.htm. Accessed July 1, 2005.
2. Hallenbeck, M. E., J. M. Ishimaru, and J. Nee. *Measurement of Recurring Versus Non-Recurring Congestion.* Washington State Transportation Center, Olympia, Oct. 2003.
3. Skabardonis, A., P. Varaiya, and K. F. Petty. Measuring Recurrent and Nonrecurrent Traffic Congestion. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1856,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 118–124.
4. Bremmer, D., K. C. Cotton, D. Cotey, C. E. Prestrud, and G. Westby. Measuring Congestion: Learning from Operational Data. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1895,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 188–196.
5. Kwon, J., and P. Varaiya. The Congestion Pie: Delay from Collisions, Potential Ramp Metering Gain, and Excess Demand. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2005.

6. Dowling Associates, Berkeley Transportation Systems, and System Metrics Group. *Measuring Non-Recurrent Traffic Congestion: Final Report.* California Department of Transportation, Sacramento, 2002.
7. *2002 HICOMP Report.* State Highway Congestion Monitoring Program, California Department of Transportation, Sacramento, Nov. 2003.
8. Petty, K., H. Noeimi, K. Sanwal, D. Rydzewski, A. Skabardonis, P. Varaiya, and H. Al-Deek. The Freeway Service Patrol Evaluation Project: Database Support Programs, and Accessibility. *Transportation Research Part C,* Vol. 4, No. 2, 1996, pp. 71–85.
9. Skabardonis, A., K. Petty, P. Varaiya, and R. Bertini. *Evaluation of the Freeway Service Patrol (FSP) in Los Angeles.* Report UCB-ITS-PRR-98-31. California PATH, University of California, Berkeley, 1998.
10. Bertini, R., S. Tantiyanugulchai, E. Anderson, R. Lindgren, and M. Leal. *Evaluation of Region 2 Incident Response Program Using Archived Data.* Transportation Research Group, Portland State University, Portland, Ore., 2001.
11. *Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 2000.
12. Chin, S. M., O. Franzese, D. L. Greene, H. L. Hwang, and R. C. Gibson. *Temporary Losses of Highway Capacity and Impacts on Performance.* Technical Report ORNL/TM-2002/3. National Transportation Research Center, Oak Ridge National Laboratory, Knoxville, Tenn., 2002.
13. Chen, C., A. Skabardonis, and P. Varaiya. Systematic Identification of Freeway Bottlenecks. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1867,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 46–52.
14. Jia, Z., P. Varaiya, C. Chen, K. Petty, and A. Skabardonis. *Congestion, Excess Demand and Effective Capacity in California Freeways.* Dec. 2000. pems.eecs.berkeley.edu. Accessed Jan. 1, 2006.
15. PeMS. pems.eecs.berkeley.edu. Accessed Jan. 1, 2006.
16. California Department of Water Resources. cdec.water.ca.gov/intro.html. Accessed Nov. 11, 2004.
17. *Lane Closure System.* California Department of Transportation. www.lcs.dot.ca.gov/. Accessed Nov. 3, 2004.
18. System Metrics Group. *Freeway Performance Report.* California Department of Transportation, Sacramento, 2003.

The Freeway Operations Committee sponsored publication of this paper.