Probe Vehicle Runs or Loop Detectors?

Effect of Detector Spacing and Sample Size on Accuracy of Freeway Congestion Monitoring

Jaimyoung Kwon, Karl Petty, and Pravin Varaiya

Freeway congestion monitoring can be based either on sampling-based methods, such as probe vehicle runs, or on continuous data from loop detector infrastructure. Sample size, in terms of the number of days sampled, affects the accuracy of sampling-based methods; detector spacing or detector density affects the accuracy of the detector-based method. This paper presents an empirical model of the effect of the two parameters—sample size and detector spacing—on the accuracy of both methods in estimating the annual average of three congestion parameters: total delay, average duration of congestion, and average spatial extent of congestion. The model is developed with data from four urban freeway corridors in California. Among other conclusions, the model predicts that to measure the congestion parameters with 10% error, 4 to 6 days' worth of good probe vehicle data or loop detector data with half-mile spacing is needed. The proposed model facilitates comparison of the two alternatives in regard to the cost for achieving the same target accuracy. The result can also be used as a guide to determine the sample size or detector spacing in planning new congestion monitoring.

Freeway congestion monitoring can be based on either sampling-based methods, such as probe vehicle runs (so-called tach data), or continuous data from loop detector infrastructure. Study of congestion monitoring practices in different districts of California (*1*) suggests that although agencies follow broadly similar data collection and processing practices, differences in details are very large. Some agencies use only probe vehicle data; some use only detector data; still others use a mixture of the two.

Key outputs of congestion monitoring include the annual averages of the magnitude, extent, and duration of congestion—the three main elements of the highway congestion monitoring program (HICOMP) in California (*2*). Two parameters drive the error of estimates of such performance measures: sample size and detector spacing. The sample size, in regard to the number of days sampled, affects the accuracy of the probe vehicle method; detector spacing or detector density affects the accuracy of the detector-based method.

J. Kwon, Department of Statistics, California State University, 229 North Science, East Bay, Hayward, CA 94542. K. Petty, Berkeley Transportation Systems, Inc., 2030 Addison Avenue, Suite 640, Berkeley, CA 94704. P. Varaiya, Department of Electrical Engineering and Computer Science, University of California, 231 Cory Hall, Berkeley, CA 94720. Corresponding author: J. Kwon, jaimyoung.kwon@csueastbay.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2012, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 57–63.
DOI: 10.3141/2012-07

It is obvious that larger detector spacing or smaller sample size would lead to larger error, and smaller detector spacing or larger sample size gives smaller error. The effect of sample size can be studied using standard statistical sampling theory (*3*). The effect of detector spacing has been studied either analytically (*4*) or along the lines of incident detection (*5*). Still, an empirical, unifying framework under which different alternatives can be compared has not been developed.

This paper presents an empirical model of the effect of the two parameters—sample size n and detector density k (inverse of detector spacing)—on the accuracy of the probe vehicle method and the detector-based method. The effect of sample size will be analyzed using sampling theory combined with the empirical variation of daily congestion parameters. The effect of detector spacing will be analyzed using the Monte Carlo method, by removing part of the baseline detectors and measuring the increase in the error. The model is developed using data from four urban freeway corridors in California.

The paper is organized as follows. The proposed method is described in the next section, followed by a section on the empirical model along with discussions. The conclusions section completes the paper.

METHODS

Three congestion parameters for a freeway corridor are considered:

1. Magnitude of delay (vehicle hours): total delay experienced by drivers,
2. Extent of congestion (mi): length of freeway affected by congestion, and
3. Duration of congestion (h): time period affected by congestion.

The symbols $D(d)$, $L(d)$, and $T(d)$ are used to denote the true daily total delay, extent, and duration on day $d = 1, \ldots, N$, respectively. The daily delay varies randomly during the year. The true annual average D and the standard deviation σ_D of the daily delay are

$$D = \frac{1}{N}\sum_{d=1}^{N} D(d)$$

$$\sigma_D = \frac{1}{\sqrt{N}}\left[\sum_{d=1}^{N}\left(D(d) - D\right)^2\right]^{1/2} \qquad (1)$$

The true annual average daily extent L and duration T of congestion as well as their standard deviations σ_L and σ_T, respectively, are defined similarly.

Loop-Based Congestion Monitoring

Consider a freeway segment instrumented with m detectors indexed $i = 1, \ldots, m,$ whose flow (volume) and speed measurements are averaged over, say, 5-min intervals indexed $t = 1, \ldots, n_t$. Days (weekdays) in the study period are denoted by $d = 1, 2, \ldots, N$. If the study period is a year, $N = 260$ approximately. Detector i is located at postmile x_i; $v_i(d, t) = v(x_i, d, t)$ is the measured speed (miles per hour, mph), and $q_i(d, t) = q(x_i, d, t)$ is the measured flow (vehicles per hour, vph) at time t of day d.

The m detectors divide the freeway into m segments. Each segment's (congestion) delay is defined as the additional vehicle hours traveled driving below free-flow speed v_{ref}, taken to be 35 mph. So the delay in segment i in time t is

$$D_i(d,t) = l_i \times q_i(d,t) \times \max\left\{1/v_i(d,t) - 1/v_{\text{ref}}, 0\right\} \text{ vehicle hours} \tag{2}$$

where 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,

$$D_{\text{loop}}(d) = \sum_{i=1}^{n}\sum_{t=1}^{T} D_i(d,t) \tag{3}$$

The annual average daily total delay and its standard deviation are simply

$$D_{\text{loop}} = \frac{1}{N}\sum_{d=1}^{N} D_{\text{loop}}(d)$$

$$\sigma_{D,\text{loop}} = \frac{1}{\sqrt{N}}\left[\sum_{d=1}^{N}\left(D_{\text{loop}}(d) - D_{\text{loop}}\right)^2\right]^{1/2} \tag{4}$$

In the application below, the daily delay during the peak period, 5 to 10 a.m. for the morning peak and 3 to 8 p.m. for the afternoon peak, are considered separately.

The duration of congestion on day d is computed as

$$T_{\text{loop}}(d) = \#\left\{\text{5-min time periods when } v_i(d,t) < v_{\text{ref}} \text{ for any } t\right\}/12(\text{h}) \tag{5}$$

and the extent of congestion as

$$L_{\text{loop}}(d) = \text{sum}\left\{l_i : v_i(d,t) < v_{\text{ref}} \text{ for any } t\right\}(\text{mi}) \tag{6}$$

The annual average duration of congestion T_{loop} and the annual average extent of congestion L_{loop} are similarly defined as the "population average" as in Equation 4.

Probe Vehicle–Based Congestion Monitoring

Different agencies or districts follow different procedures to estimate the congestion parameters from probe vehicle runs (*1*), but the core idea is as follows: probe vehicle runs are first processed to estimate the duration $D_{\text{probe}}(d)$ and extent $L_{\text{probe}}(d)$ of congestion; the magnitude of delay is then derived using the formula

$$D_{\text{probe}}(d) = (\text{vehicles affected per hour}) \times D_{\text{probe}}(d) \times L_{\text{probe}}(d) \times \left[\frac{1}{v_{\text{congest}}} - \frac{1}{v_{\text{ref}}}\right] \tag{7}$$

Usually probe vehicle runs are performed on a few sampled days. Let n be the number of days sampled, and index sampled days as $d = 1, \ldots, n$. The annual average daily total delay is estimated by

$$D_{\text{probe},n} = \frac{1}{n}\sum_{d=1}^{n} D_{\text{probe}}(d) \tag{8}$$

The probe vehicle estimate of annual average duration of congestion $T_{\text{probe},n}$ and the annual average extent of congestion $L_{\text{probe},n}$ are similarly defined as the "sample average" as above.

Effect of Sample Size

From here on the focus is mainly on the total delay D, for simplicity, because extent L and duration T of congestion are treated similarly. Recall that the percentage error of an estimate is defined as the standard error of an estimate divided by the true value of a parameter. For the probe vehicle method, which is run on n days randomly sampled from $N = 260$ weekdays in a year, the percentage error in estimating the total delay D is given by

$$\text{percentage error of } D_{\text{probe},n} = \frac{\sigma_D}{D}\sqrt{\frac{N-n}{n(N-1)}} \tag{9}$$

from standard sampling theory in statistics (*3*). This formula assumes that a probe vehicle run on a particular day $D_{\text{probe}}(d)$ provides an accurate measurement of the true $D(d)$.

In Equation 9, values of D and σ_D are unknown. For a well-instrumented freeway, however, $D_{\text{loop}}(d)$ is very close to the true $D(d)$ and $D_{\text{loop}}(d)$ can be used in place of $D(d)$ in Equation 1 to obtain approximates of D and σ_D. In other words, the approximation

$$\sigma_{D,\text{loop}} \approx D_{\text{loop}}$$

and

$$\sigma_D \approx \sigma_{D,\text{loop}} \tag{10}$$

will be used for a well-instrumented freeway with detector spacing less than a half mile.

Also, the constant σ_D/D is site-specific and varies from site to site. To obtain the gross error model, the percentage error curve (Equation 9) for a few sites is obtained and the curves are averaged:

$$\text{gross percentage error of } D_{\text{probe},n} = \text{average}\left\{\text{percentage error of } D_{\text{probe},n} \text{ over sites}\right\} \tag{11}$$

Effect of Detector Density

In the detector-based method, congestion parameters are calculated from measurements at discrete locations on the freeway segment. To quantify errors from such spatial sampling, proceed as follows:

1. Begin with well-instrumented corridor of length l (mi) with m detectors.
2. Take the measurements reported by the m detectors as "ground truth."
3. Consider 1 day.
4. Randomly pick m' ($<m$) of the m stations, and estimate the congestion parameters for that day from the m' selected detectors.
5. For each size m', make a different random selection a large number of times, say, $5 \times m$ times. In this way, for each sample size n, there are $5m$ measurements of the congestion parameters, which are compared with the ground truth given by measurements from all m stations. Write a particular estimate as $D_{\text{loop},k}$, where $k = m'/l$ is the detector density.
6. Compute percentage error for detector density k on that day as the standard error of the $5m$ measurements divided by the ground truth.
7. Repeat Steps 4–6 for different $m' = 1, \ldots, m - 1$.
8. Repeat Steps 3–7 for, say, 10 randomly sampled days, and average the daily percentage error curve over the 10 days. That is the percentage error versus detector density curve for the site.

Just as in Equation 11, site-specific percentage error versus the detector density curve is averaged over sites to obtain

$$\text{gross percentage error of } D_{\text{loop},k} = \text{average}\left\{\text{percentage error of } D_{\text{loop},k} \text{ over sites}\right\} \quad (12)$$

Combining the Two Empirical Models

The gross percentage error of $D_{\text{probe},n}$ is used to approximate the percentage error that can be achieved with the given sample size n in a probe-based method; the gross percentage error of $D_{\text{loop},k}$ is used to approximate the percentage error achieved with the given detector density k for a detector-based method.

Detector data may not be available for many days because of detector or communication failures. In such cases, sampling of days and detector density both contribute to the percentage error. In such cases, the simple formula

$$\text{gross percentage error of detectors with detector density } k \text{ and sample size } n = \text{gross percentage error of } D_{\text{probe},n} + \text{gross percentage error of } D_{\text{loop},k}$$

can be used to approximate the percentage error. This formula is based on the reasonable assumption that the error from sampling of days and the error from spatial sampling are independent.

EMPIRICAL MODEL

The model is developed using data from four urban freeway corridors in California. The four sites are selected using the following criteria:

1. The site is "interesting," that is, it experiences (recurrent) congestion.
2. Probe vehicle run data are available for the site.
3. The location is well instrumented, having loop detectors with detector spacing < 0.5 mi.

Table 1 lists the four sites. For those sites, probe vehicle data are obtained from Caltrans as described in detail in Kwon et al. (*1*). Loop detector data are obtained from the Freeway Performance Measurement System (PeMS) (*6*).

Speed Correlation Between Probe Vehicle Runs and Loop Data

There is qualitative agreement in the congestion measures based on the two data sets, that is, there is a large, positive correlation between PeMS-reported speeds and tach vehicle speeds (see Figure 1). But there is disagreement over the quantitative estimates because of differences in the procedures used to calculate them.

Effect of Sample Size

Table 2 lists the true average values and the standard deviations approximated by Equation 10 for the four study sites. The numbers in Table 2 have the following meaning: on roughly 68% of the days in 2002, the congestion delay on the 51N study segment was in the interval $1{,}116.6 \pm 640.1$ vehicle hours, the extent was within 3.46 ± 1.08 mi, and the duration was within 4.47 ± 1.75 h. A similar interpretation holds for the other three sites. Observe that the standard deviation is almost 50% of the average value, indicating a large variation in congestion from one day to the next.

A deeper appreciation of the variation in congestion is gained from Figure 2, which shows the histograms of daily delay, duration, and extent for the four sites. Thick, solid vertical lines are the annual averages, also given in Table 2. Dashed vertical lines correspond to the delay, duration, and extent of congestion measured on the day of the probe vehicle run. The difference between the solid and dashed lines is the error that would be made by taking the measurements of the tach run to be the true averages. Of the 12 estimates (three parameters for each of the four sites) in Figure 2, eight have an error smaller than their corresponding standard deviation and four have an error exceeding the standard deviation.

Figure 3 shows the relationship between the percentage error and n for the four sites. The thick lines represent the average of the four curves.

TABLE 1 Four Study Sites

District and Freeway	Segment Length (mi)	Number of Days with Good Loop Data	Probe Vehicle Run Date	Average Number of Detectors per Mile
D03 51N	6.738	241	2002-11-6	1.95
D12 91E	9.891	261	2004-12-9	1.57
D12 5N	13.587	261	2002-11-5	2.23
D12 55N	11.547	261	2002-10-29	2.20

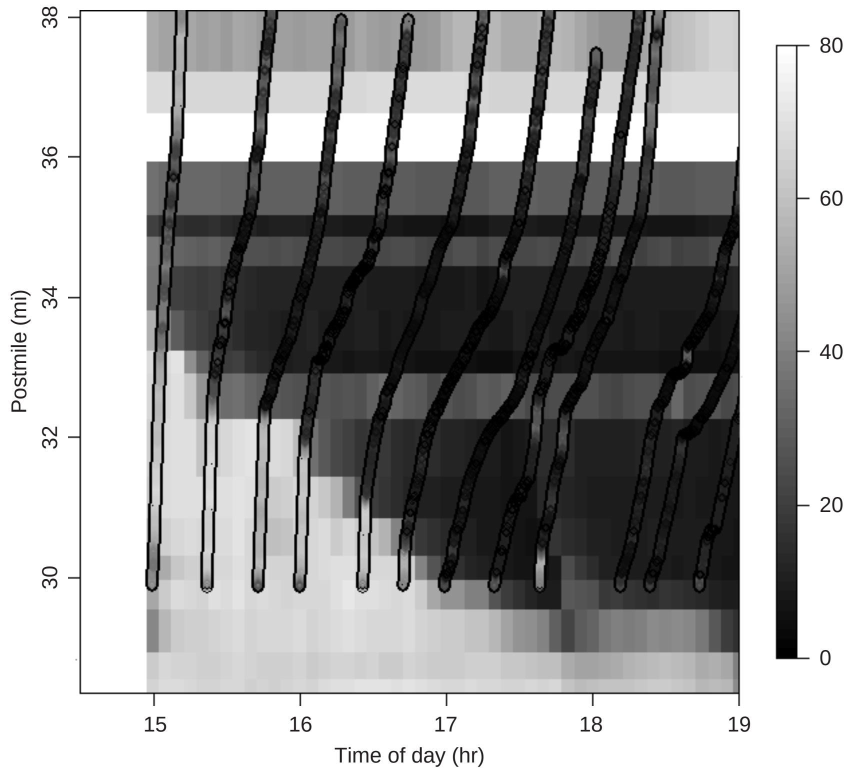

FIGURE 1 Probe vehicle runs (12 trajectories) superimposed on speed contour plot from loop data, for eastbound SR-91 study site on December 9, 2004. Background contour plot represents traffic flow speed measured from loops (mph) at given location (postmile) and time of day; the colors and angles of a trajectory represent traffic flow speed as "seen" by a particular probe vehicle.

Figure 3 implies that using only a single day's worth of data leads to 50% error in estimation of average daily delay *D*. The effect of such a small sample size is less severe for parameters *L* and *T*, with percentage error of 30% and 40%, respectively. For most parameters, achieving 10% error requires at least 4 to 6 days' worth of data. Also notable is the relatively small site-to-site variation in percentage error, except for duration.

Effect of Detector Density

The three "box and whiskers" plots in Figure 4 summarize the distribution of the 65 values of the congestion parameters for each number of detectors m'. The speed contour plot is the ground truth from all 13 detectors on November 6, 2002.

The following is an explanation of how to interpret the box plots. Consider total delay on the study segment displayed in the top left plot. Pick $m' = 3$. The lower limit of the gray box is the 25th percentile, the bold black horizontal line toward the middle of the box is the median or 50th percentile, and the upper limit of the box is the 75th percentile of the 65 values. The two "whiskers" extend from the box to the smallest and largest points that are within 1.5 × the interquartile range from the box boundaries. Finally, the small circles denote "outliers." As can be seen for $n = 3$, the median value is 1,100, the 25th percentile is 700, and the 75th percentile is 1,800 vehicle hours. As expected, as *n* increases, the variability

TABLE 2 Average and Standard Deviation of Congestion Parameters Estimated from Loop Data

	Site			
	51N	91E	5N	55N
Average				
D (veh h)	1,116.6	3,744.7	2,072.8	1,299.3
L (mi)	3.46	4.74	3.47	3.98
T (h)	4.47	6.21	5.54	3.64
Standard deviation				
σ_D (veh h)	640.1	2,283.6	1,235.3	630.0
σ_L (mi)	1.08	1.29	1.58	1.11
σ_T (h)	1.75	2.43	2.31	1.40

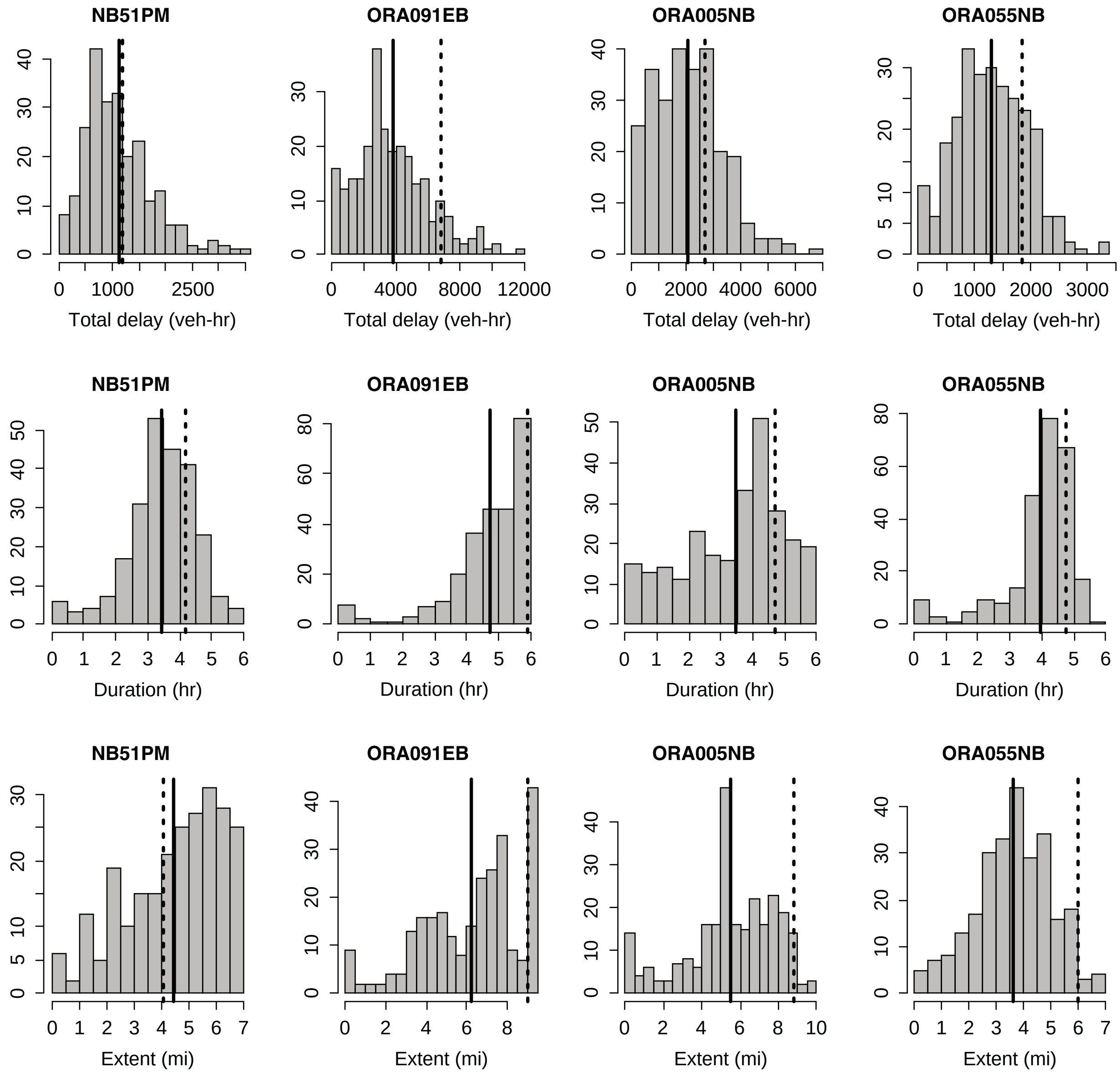

FIGURE 2 Histogram of delay, duration, and extent for the four sites. Annual averages (thick, solid vertical lines) and measurements on the day of the probe vehicle run (dashed vertical lines) are also shown.

reduces, until for $n = 13$ the "ground truth,"—1,300 vehicle hours—is obtained.

From Figure 4, it is also observed that the duration T is severely underestimated and the total delay D moderately underestimated for low detector density, whereas such systematic "bias" is not present in the estimation of extent L. In all three cases, though, the "variance" is large for low detector density.

The empirical relationship between the percentage error and the detector density for the four sites is presented in Figure 5. Each line corresponds to the four study sites, and thick lines represent the average of the four curves. Because detector density is irregularly spaced, a smooth curve was fitted via LOWESS (*7*) to obtain the average curve.

For all three congestion parameters, the percentage error versus detector density curve varies more from site to site, compared with the percentage error versus sample size curve shown in Figure 3. This suggests that traffic characteristics at the site need to be considered more carefully when detector spacing is an issue.

For all three plots in Figure 5, the percentage error decreases sharply until a density of 0.5 detector/mi and decreases almost linearly from that point on. To achieve the target percentage error of 10%, the detector density needs to be at least 2 detectors/mi, that is, detector spacing needs to be less than 0.5 mi.

CONCLUSIONS

An empirical model is presented relating sample size n and detector spacing k to the accuracy of two popular congestion monitoring methods: probe vehicle runs and the detector-based method. Among other conclusions, the model predicts that to measure the congestion parameters with 10% error, 4 to 6 days' worth of good probe vehicle data or loop detector data with half-mile spacing is needed. However, sampling only a single day (however accurate that day's measurement is) or having 0.5 detector per mile, that is, a detector spacing of 2 mi (however long one measures) leads to a percentage error of as much as 50%.

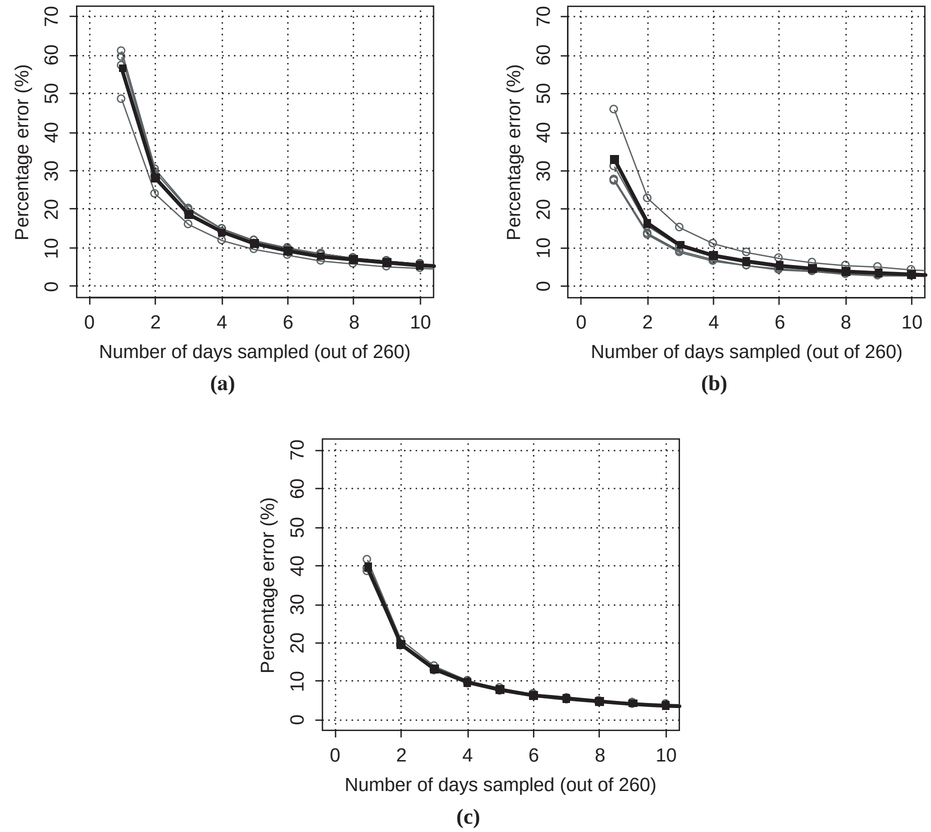

FIGURE 3 Percentage error for congestion parameters versus number of days sampled (of 260 weekdays in a year). Gray curves represent the four study sites, and thick lines represent average of the four curves: (*a*) total delay (vehicle hours), (*b*) duration (h), and (*c*) extent (mi).

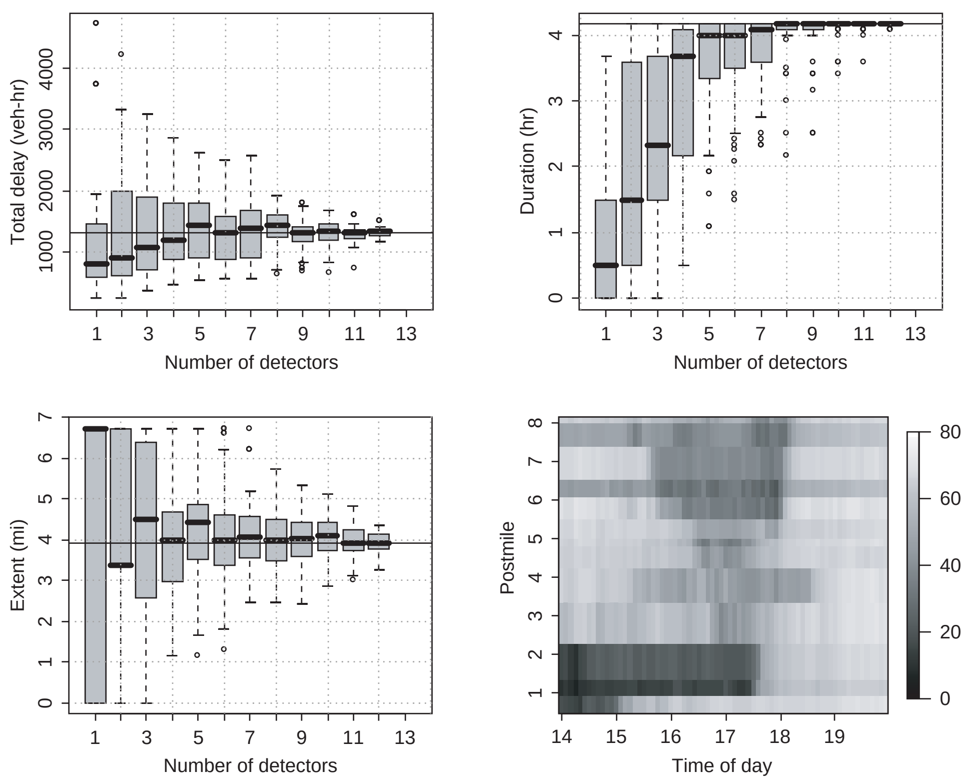

FIGURE 4 Simulation of congestion parameters with $n = 1, 2, \ldots, 13$ detectors over 9-mi section on 51N on November 6, 2002.

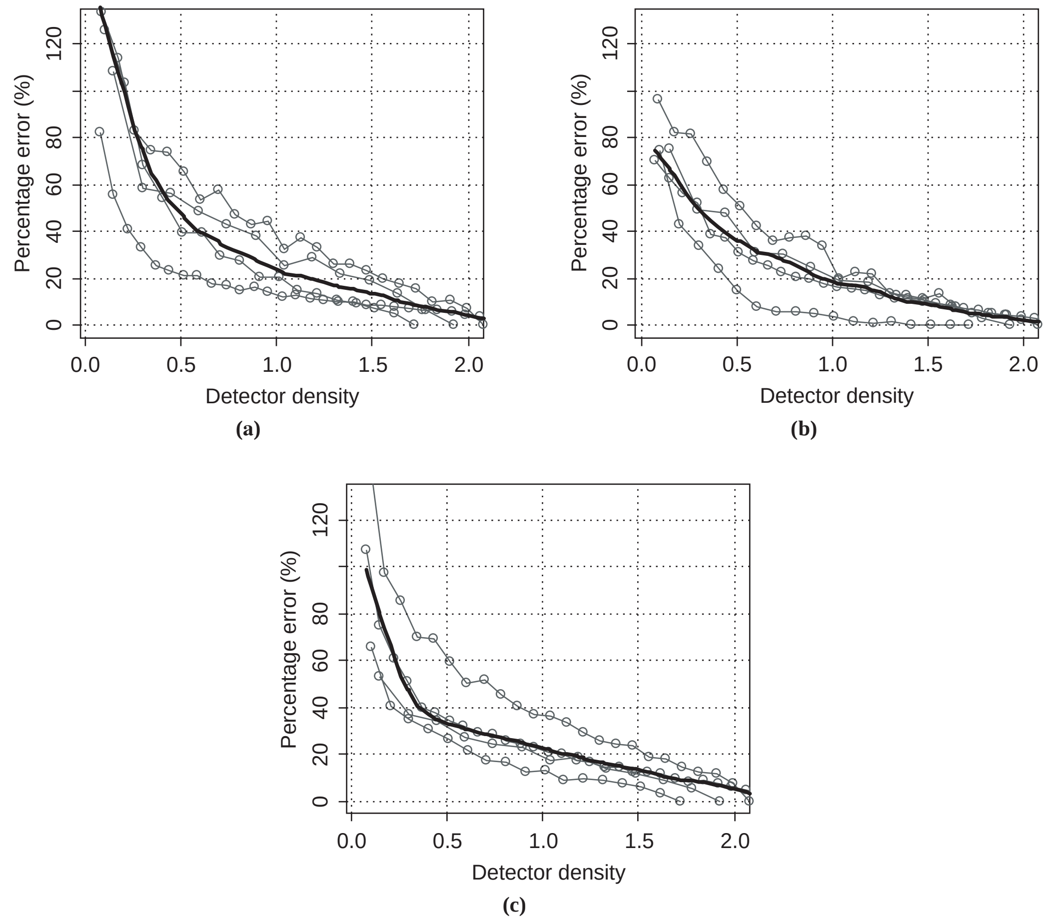

FIGURE 5 Percentage error for congestion parameters versus detector density (detectors/mi): (*a*) total delay (vehicle hours), (*b*) duration (h), and (*c*) extent (mi). Gray curves represent the four study sites, and thick lines represent the average of the four curves.

Readers need to be aware, however, of the following limitations of the model: it is based on a selection of urban freeways and cannot be applied to, for example, rural freeways with little recurrent congestion; also, even among urban freeways, there could be a sizable site-to-site variation, especially in cases in which detector density k is involved.

Nevertheless, the proposed model, being a unified framework, facilitates the comparison of the two alternatives in regard to the cost for achieving the same target accuracy, because the parameters n and k above can be easily translated into dollar costs. The result can also be used as a guide to decide the sample size or detector spacing in planning new congestion monitoring. Finally, it is easy to build a similar model for other performance measures, such as travel time.

ACKNOWLEDGMENTS

The authors are grateful for comments and criticism from John Wolf and Fred Dial of Caltrans and from Bill McCullough of the System Metrics Group. This study is supported by grants from Caltrans to the California PATH Program.

REFERENCES

1. Kwon, J., B. McCullough, K. Petty, and P. Varaiya. *Evaluation of PeMS to Improve the Congestion Monitoring Program.* Final report for PATH TO 5319. California PATH, July 2006.
2. *2002 HICOMP Report.* State Highway Congestion Monitoring Program. California Department of Transportation, Sacramento, Nov. 2003.
3. Cochran, W. G. *Sampling Techniques.* John Wiley and Sons, Inc., New York, 1977.
4. Fujito, I., R. Margiolta, W. Huang, and W. A. Perez. Effect of Sensor Spacing on Performance Measure Calculations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1945,*Transportation Research Board of the National Academies, Washington, D.C., 2006.
5. Martin, P. T., H. J. Perrin, and B. G. Hansen. Incident Detection Algorithm Evaluation. UTL-0700-31. July 2000. www.mountain-plains.org/pubs/html/mpc-01-122/. Accessed Nov. 5, 2007.
6. PeMS. PeMS website. pems.eecs.berkeley.edu. Accessed July 31, 2006.
7. Cleveland, W. S. LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression. *American Statistician,* Vol. 35, No. 54, 1981, pp. 270–280.

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented. The contents do not necessarily reflect the official views or policy of the California Department of Transportation. This paper does not constitute a standard, specification, or regulation.

The Freeway Operations Committee sponsored publication of this paper.