The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors

Zhanfeng Jia, Chao Chen, Ben Coifman and Pravin Varaiya

Abstract— **This paper presents the PeMS algorithms for the accurate, adaptive, real-time estimation of the g-factor and vehicle speeds from single-loop detector data. The estimates are validated by comparison with independent, direct measurements of the g-factor and vehicle speeds from 20 double-loop detectors on I-80 over a three-month period. The algorithm is used to process data from all freeways in Caltrans District 12 (Orange County, CA) over a 20-month period beginning January 1998. Analysis of those data shows that the g-factors for *different* loops in the district differ by as much as 100 percent, and the g-factor for the *same* loop can vary up to 50 percent over a 24-hour period. Many transportation districts now post real-time speed and travel time estimates on the World Wide Web. Those estimates often are derived from single-loop detector data assuming a *common* g-factor for all detectors in the district. This study suggests that those estimates can be in error by 50 percent, and so they are of little value to travelers. The use of the PeMS algorithm will reduce those errors.**

I. Introduction

Vehicle speed is the most important measure of highway service quality. Speed influences how drivers experience a trip, its cost, and the emissions it produces. Expectations of travel time (and hence of speed) enter trip-making and route-selection decisions. Vehicle speed is a principal component of indicators of highway performance, such as congestion and mobility, that agencies use to guide operations and investment. The accurate measurement of vehicle speed can improve the decisions of both consumers and providers of highway transportation service.

Single-loop detectors are the most common sensors used in today's highways. The detectors provide the only comprehensive data about highway performance that we have. But highway agencies seem not to know how well their loop detectors work, and they do not systematically archive the data. The resulting unreliability of the detectors and data unavailability partly explain why they are not used in procedures to improve highway operations or investment planning. Since they are not so used, there is reduced incentive to monitor and maintain detectors or to archive historical data. This is unfortunate because as this paper and companion work[1] show, loop detector data contain unique information of immense value to highway managers, engineers, and the public. One such valuable information is vehicle speed.

It is deceptively easy to estimate the speed at a single-loop detector data from the relationship,

$$v(t) = g(t) \times \frac{c(t)}{o(t) \times T} \ . \tag{1}$$

Here T is the duration of the reporting period, commonly 20 or 30 seconds. At the end of period t the detector reports two numbers, $c(t)$ and $g(t)$. The count $c(t)$ is the number of vehicles that crossed the detector during period t, and the occupancy $o(t)$ is the fraction of time during this period that the detector sensed a vehicle above it. The 'g-factor' $g(t)$ is the effective vehicle length in this period. It cannot be directly measured at single loops, and its value must be assumed or estimated. If the correct value of $g(t)$ is known, then the average speed $v(t)$ can be trivially calculated from the detector data using (1).[2]

Several transportation districts now regularly post speed and travel time estimates on the World Wide Web for use by travelers. Some of these districts use equation (1) to estimate speed, assuming a common, constant, district-wide g-factor. This is also Caltrans practice. As is evident from (1), if the true $g(t)$ is different from this assumed value, the speed estimate will be in error in the same proportion. If the variability in the g-factor that we find in Orange County is typical of other districts, it follows that the assumption of a constant g-factor will *result in errors in speed estimates of more than 50 percent!* Travelers can have no confidence in the posted speeds or the travel times derived from them. If these speed estimates are used as components of performance measures, those measures will similarly suffer large errors. The PeMS algorithms eliminate these errors.

This paper makes two contributions. The first provides empirical evidence of the variability in the g-factor. The significance of this variability was just noted above. We analyze 20 months of data beginning in January 1998, from the freeway network in Caltrans District 12 (Orange County). The analysis shows that the g-factors for any two randomly-picked detectors in this network differ on average by 26 percent, and with probability 0.12 that they will dif-

Jia, Chen and Varaiya are with University of California, Berkeley. Coifman is with Ohio State University. Research supported by the National Science Foundation and by Caltrans, California State Department of Transportation, through the California PATH Program. The views expressed here are those of the authors alone.

[1]This paper is part of a series of studies, software, and services produced by the PeMS project. See [3] and the articles posted in [2].

[2]Provided, of course, the data are valid. In practice validity is determined by some simple tests: for example, the data may be declared invalid if $o(t)$ is very small. There also are random measurement errors: a detector may count a vehicle in a neighboring lane because of cross-talk, or it may miss a lane-changing vehicle. These random errors are not of interest in this discussion which is concerned with large, systematic errors.

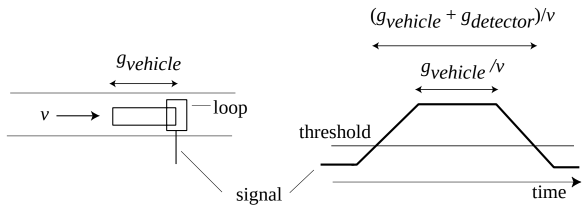

Fig. 1. A vehicle moving at speed v occupies a detector for time $(g_{vehicle} + g_{detector})/v$, where $g_{vehicle}$ is the vehicle length, and $g_{detector}$ depends on the detector electronics.

fer by 50 percent. Furthermore, the g-factor for the *same* detector can vary by as much as 50 percent over a 24-hour period.

More importantly, the paper describes the PeMS adaptive, real-time algorithms for accurately estimating g-factor $g(t)$ and speed.[3] Confidence in the algorithms' accuracy is established by comparing their estimates with direct, independent measurements of g-factors and speeds for three months at 20 specially-equipped dual-loop detectors on I-80 in Emeryville, CA. The algorithms are now implemented in PeMS to estimate in real time the g-factor and the speed at every detector in District 12.

The rest of the paper is organized as follows. In section II we discuss the causes of variation in the g-factor and we show how large this variation is for the detectors in District 12. In section IV we compare the g-factor and speed estimates obtained from the PeMS algorithms with their independently measured, true values. In section III we present the PeMS algorithms. Concluding remarks are gathered in section V. For an alternative approach to measuring speed from single loop detectors see [4].

II. The variation in g-factors

A loop detector infers the presence of a vehicle when its signal exceeds a threshold. Figure 1 shows that a vehicle 'occupies' a loop for a duration of time equal to

$$\text{on-time} = \frac{g_{vehicle}}{v} + \frac{g_{detector}}{v}\ .$$

Here v is the vehicle speed and $g_{vehicle}$ is its length. $g_{detector}$ is a function of the threshold value and the slope of the detector signal, which depend on the detector electronics. If during period t, $c(t)$ vehicles cross the detector at the same speed $v(t)$, the loop is occupied for time

$$o(t) \times T = c(t) \times (g_{traffic}(t) + g_{detector}) \times v(t)$$

where $g_{traffic}(t)$ is the average value of vehicle length, $g_{vehicle}$, during period t. Comparison with (1) shows that $g(t)$ is the sum of two components,

$$g(t) = g_{traffic}(t) + g_{detector}. \tag{2}$$

The first component, $g_{traffic}(t)$, depends on the mix of vehicle types (autos, vans, trucks) crossing the detector during period t. In urban freeways this mix exhibits a 24-hour cycle and a weekly cycle. For example, trucks account for a larger percentage of the traffic during the early morning weekday hours, while autos predominate during commute hours and weekends. On some freeways, moreover, trucks are confined to the outer lanes.

With 18-foot long autos and 60-foot long trucks, we expect $g_{traffic}(t)$ to range from 18 feet for inner, auto-only lanes to as much as 60 feet in the early morning for outer lanes over freeway links with heavy truck traffic. Suppose instead of accounting for this lane-, link- and time-dependent variation, $g_{traffic}(t)$ is assumed to be a constant, say, 20 feet. Then auto-only lane speeds will be underestimated by 10 percent and speeds in lanes with a large proportion of trucks will be overestimated by nearly 300 percent. Much more commonly adjacent lanes on many links will exhibit, over a 24-hour period, a shift in proportion of truck traffic from 0 percent (giving $g_{traffic}(t) = 18$ feet) to 20 percent (giving $g_{traffic}(t) = 26.4$ feet), or a variation of 46 percent, so that taking a constant value will lead to speed estimates that are easily off by 25 percent.

The second component of $g(t)$, $g_{detector}$, depends on the detector itself—its threshold value and age, how it is installed, and the characteristics of nearby loops. In a freeway system detectors are deployed over a period of years and they are not uniformly maintained. As a result, $g_{detector}$ may vary greatly across detectors, even though, for the same detector it will likely change only slowly over time. The evidence from Orange County suggests that the variation is large.

A. Variation in g-factor between detectors

We estimate the inter-detector g-factor variation in Orange County as follows. For each loop detector i we aggregate its count $c_i(k)$ and occupancy $o_i(k)$ over a ten-minute period beginning at 1.00 am for seven days, indexed $k = 1, \cdots, 7$. We assume that at this hour in the morning, vehicles are traveling at the free flow speed of 65 mph. Knowing v, we use (1) to obtain the average g-factor for this detector as

$$g_i(k) = \sum_t \frac{o_i(k)}{c_i(k)} \times v_{65} \times T.$$

The sum is over all 30-second periods between 1.00 am and 1.10 am, excluding those periods for which the occupancy is zero, v_{65} is the assumed speed of 65 mph expressed in feet/sec, and $T = 30$ sec. For each detector i, this procedure yields up to seven estimates $g_i(k)$ for the seven different days k. We compute the mean $\bar{g}_i$ of these seven estimates and the standard deviation, σ_i.

We classify each detector by its lane from 1 to 6, with lane 1 being the innermost and lane 6 being the outermost.

[3] PeMS stands for the freeway Performance Measurement System, developed in part under contract from Caltrans. It is now fully implemented for Districts 7,8, 12. PeMS is being extended to other Caltrans districts. The PeMS system comprises (1) a large database that stores historical and real-time data; (2) a set of algorithms including the g-factor and speed estimators that process these data to generate several performance measures (including congestion and travel time); and (3) a server that delivers this information to managers, engineers, and the traveling public via their web browsers.

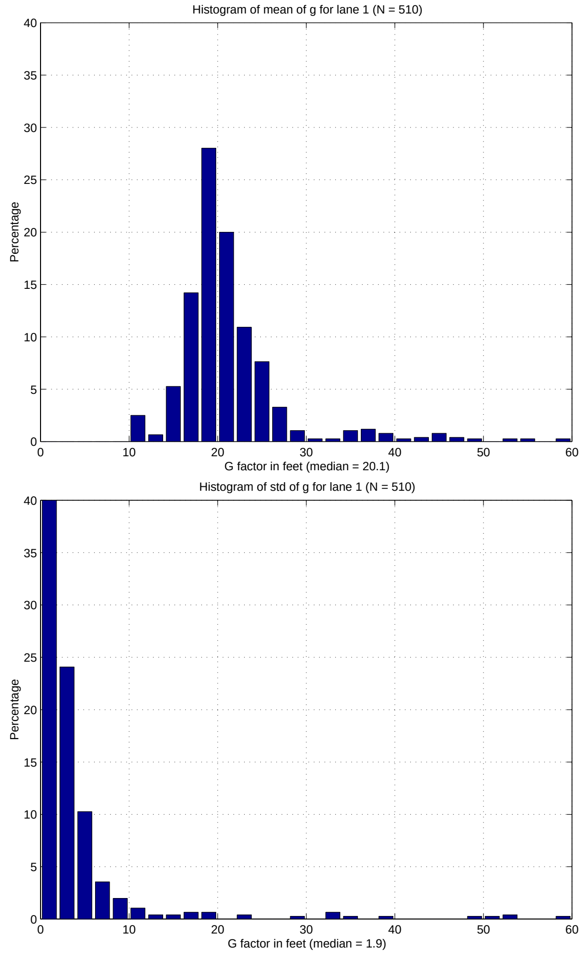

Fig. 2. Histograms of the mean $\bar{g}_i$ and standard deviation σ_i for detectors in lane 1.

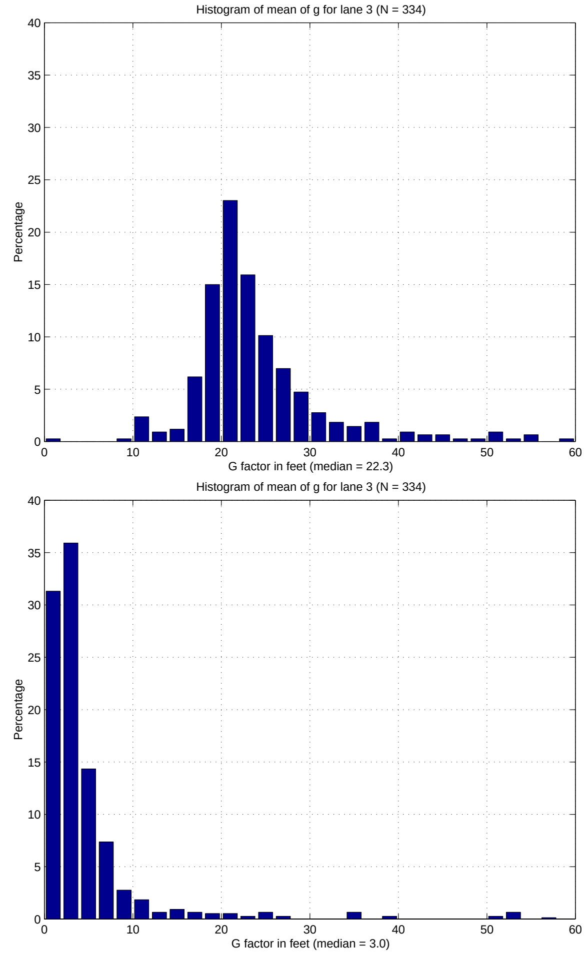

Fig. 3. Histograms of the mean $\bar{g}_i$ and standard deviation σ_i for detectors in lane 3.

In all there are 510 lane-1, 365 lane-2, 334 lane-3, 231 lane-4, 69 lane-5, and 18 lane-6 detectors. Figures 2, 3 give the histograms of $\bar{g}_i$ and σ_i for detectors in lanes 1 and 3.

The obvious feature of these histograms is that the median value of the mean increases from 20.1 feet for lane-1 detectors to 22.3 feet for lane-3 detectors (to 23.4 feet for lane-5 detectors, not shown). This increase is likely due to the increasing prevalence of vans and trucks and a true speed less than the assumed 65 mph in the outer lanes.

The surprising feature is the large variation among the mean values. Among lane-1 detectors, 30 percent have a mean below 18 feet, 47 percent have a mean between 18 and 22 feet, and 23 percent have a mean larger than 22 feet. The standard deviation of the means has a median of 1.9 feet (and a larger mean) so that two randomly picked lane-1 detectors will differ by $\sqrt{2} \times 1.9 = 2.66$ feet or 13 percent.

Among lane-3 detectors, 30 percent have a mean below 20 feet, 33 percent have a mean between 20 and 24 feet, and 37 percent have a mean larger than 24 feet. The median standard deviation is 3.0 feet, so two randomly picked lane-2 detectors differ by $\sqrt{2} \times 3.0 = 4.2$ feet or 19 percent.

Among lane-5 detectors, 23 percent have a mean below 20 feet, 36 percent have a mean between 20 and 24 feet, and 41 percent have a mean larger than 24 feet. The median standard deviation of 3.7 feet implies that two randomly picked lane-5 detectors differ by $\sqrt{2} \times 3.7 = 7.4$ feet or 31 percent.

Thus a constant $g_{detector}$ assumption will typically lead to differences in speed estimates from detectors in the *same* lane by 13 to 31 percent. The speed estimates from different lanes will be larger: randomly pickd lane-1 and lane-5 detectors will differ by 26 percent.

These are typical differences. For confidence in the estimates sufficient to alter trip-making or route-choice decisions, a much higher statistical standard must be met.

B. *Variation over time in g-factors at the same detector*

The g-factor for the *same* loop varies over the course of a week. This variation reflects the daily and weekly cycle of change in the traffic mix, $g_{traffic}(t)$. We present evidence of this variation in the form of PeMS estimates of $g_{traffic}(t)$. As noted previously, PeMS estimates are carried out for all 1,500 loop detectors in Orange County, of which only a tiny sample is presented here.

III. PeMS g-factor and speed estimation algorithms

The basic assumption underlying the algorithms is that when there are few vehicles on the road, they move at the free flow speed. On typical urban freeways, this is about 60 mph. Under this assumption, for each detector, equation (1) can be rewritten as

$$g(t) = \frac{o(t) \times T}{c(t)} \times v_{free}, \tag{3}$$

where v_{free} is the free flow speed, and t indicates the tth measurement.

An occupancy threshold is selected to determine whether the traffic is in free flow.

Using equation (3), the g-factor for every measurement can be found. However, in congestion, equation (3) does not hold and we cannot use it to calculate the g-factor. Moreover, if there is no vehicle driving through a certain loop during the tth measurement, $c(t)$ is zero, and (3) does not make sense.

On the other hand, we expect the g-factor to be a continuous, smooth curve over time. This suggests using the following autoregressive filter to track the variation of the g-factor over time: If the measurement is valid and there is no congestion,

$$g_{filt}(t) = (1-p) \times g_{filt}(t-1) + p \times g_{inst}(t), \tag{4}$$

$$g_{inst}(t) = \frac{o(t) \times T}{c(t)} \times v_{free}, \tag{5}$$

otherwise,

$$g_{filt}(t) = g_{filt}(t-1). \tag{6}$$

Equation (5) is the same as (3). We call g_{inst} instantaneous g-factor since it is calculated from the instantaneous volume and occupancy measurements. This is an IIR filter. The parameter p determines the time constant τ, which is chosen to be two hours.

The problem with the filtered g-factor g_{filt} is that it introduces delay of about two hours, equal to the time constant τ. In fact, if we shifted the g-factor curve by τ, the estimate would be perfect. However, shifting will make the algorithm non-causal, so that we cannot get the g-factor in real time.

We introduce a corrector to cancel the effect of delay. The idea is to use the periodic nature of the g-factor. What we want to know is how the g-factor changes two hours into the future. Although we don't know this change in real time, we can use historical data. Statistics show that the g-factor repeats itself during weekdays.

Denote the historic g-factor as g_{hist}. The g-factor after correction is

$$g(t) = g_{filt}(t) + [g_{hist}(t+\tau) - g_{hist}(t)], \tag{7}$$

and thus the speed is

$$v(t) = \frac{o(t) \times T}{c(t)} \times g(t). \tag{8}$$

It may seem strange that we estimate speed by specifying the value of the free flow speed. In fact, we cannot determine speed from single-loop measurements without some prior knowledge. Constant g-factor algorithms assume the g-factor is known; in our adaptive g-factor algorithm the free flow speed is assumed known. The free flow speed is much easier to measure and less detector-sensitive. The next section makes this comparison in detail.

IV. Validation of the PeMS algorithms

A. *Empirical results*

In this section we present empirical results for our speed estimator and compare the results with measurements. The measurements of traffic are taken on Interstate Highway 80 (I-80) near Berkeley. There are double-loop detectors so that we can measure the speed of individual vehicles as well as the volume and occupancy. The data processing software averages the occupancy and speed, and sums the volume over a five-minute interval, that is, $T = 5\,\text{min}$. We can then calculate the real g-factor, by using equation (1), directly from measurements. The real g-factor acts as a reference in Figure 4.

Figure 4 shows how the real g-factor (dashed line) changes over a day, varying from 20 feet to 40 feet. It also shows large instantaneous variations due to bursts of heavy vehicles, especially in the early morning hours. The estimates of g-factor (solid line) follow the trend of the real values quite well. The delayed dotted line is the g-factor after the IIR filter and before the corrector. The occupancy threshold is 15%. Figure 5 shows congestion in the afternoon, between 3 and 6 pm.. According to our algorithm, the filtered g-factor is flat (unchanged) between 3-6 pm. This explains why g-factor estimates are greater than the real values since the real g-factor keeps going down during this period.

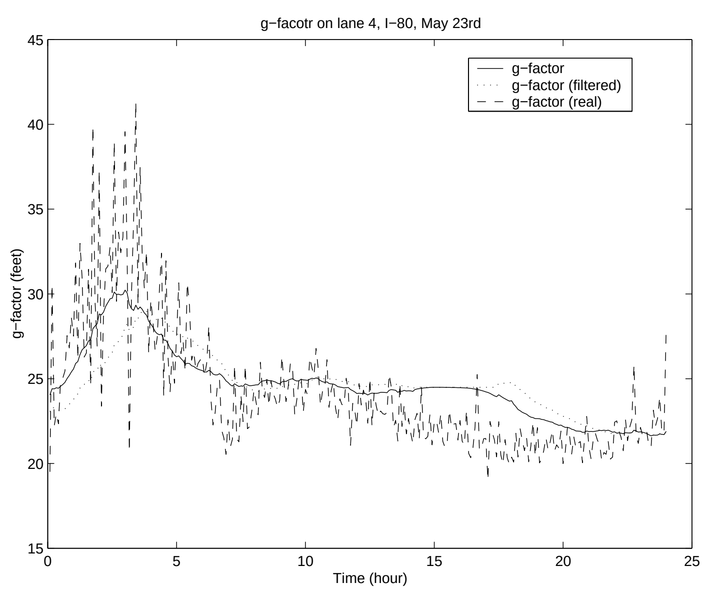

Fig. 4. g-factor estimates and observations.

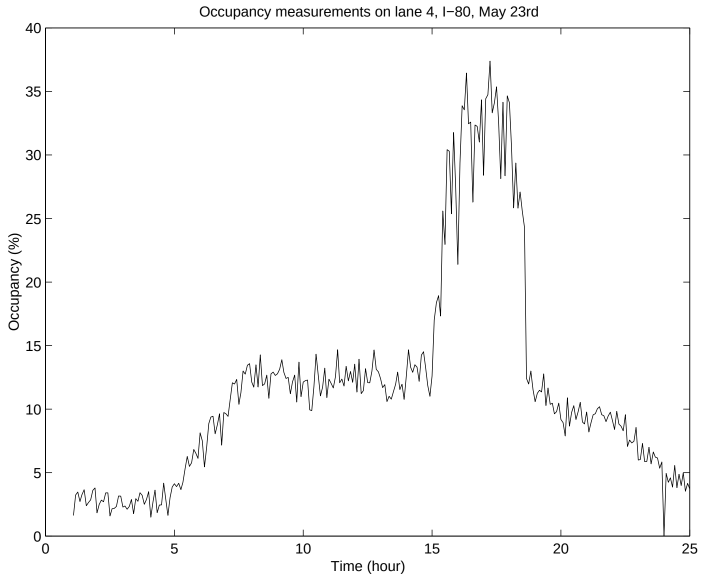

Fig. 5. Occupancy measurements.

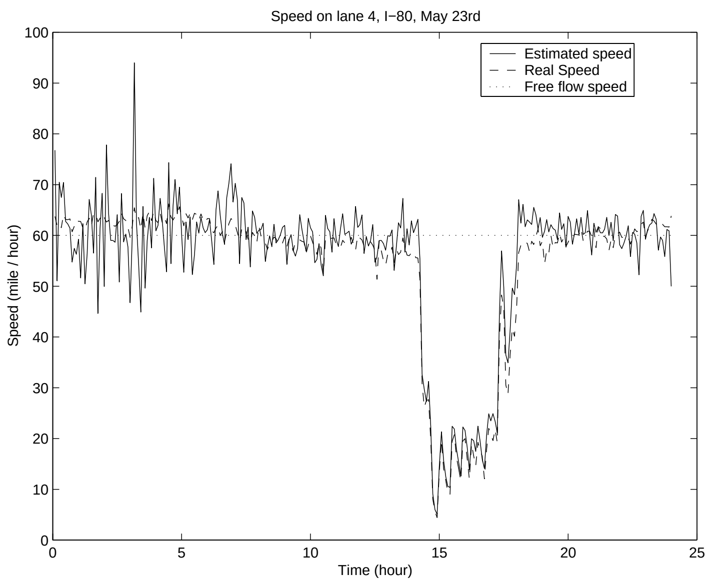

Fig. 6. Speed estimates and observations.

Figure 6 shows the speed estimates on the same day, in which the dotted line indicates the assumed free-flow speed, 60 mph. When there is no congestion, the speed estimates vibrate around 60, of course, because we calculate the g-factor based on this value. In order to estimate speed precisely, the free flow speed turns out to be an important parameter. It varies during the day. Observing the real speed in Figure 6, the average speed is 63 mph in the early morning and goes down to 57 mph just before congestion in the afternoon. Correspondingly in Figure 4, the g-factor estimates in the early morning are slightly smaller than the real values, while in the afternoon they are slightly larger.

However, what interests us most is the speed estimates when traffic is not in free flow. In particular, during congestion, the speed estimates follow the true values very well.

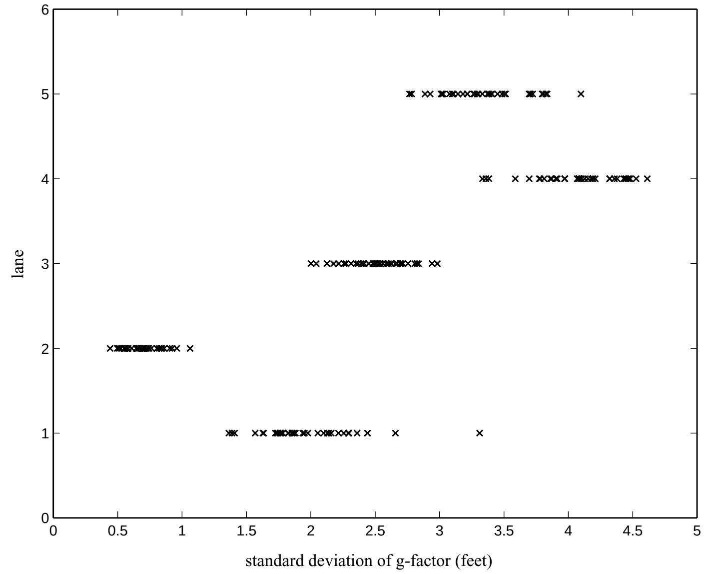

Fig. 7. Standard deviation of the g-factor. Lane 1 is the left-most lane and lane 5 is the right-most. There are 37 days of data. Each ‘x’ represents the standard deviation for one day.

B. Comparison with constant g-factor algorithm

The g-factor does not change in the same way on different lanes of freeway. The test I-80 section has five lanes. Measurements show that the lanes on the left have a flat g-factor curve while those for right lanes are peaked because of heavy trucks. Figure 7 displays the difference, looking at the variation over a day.

To evaluate the performance of a speed estimator, we define relative mean error between estimated speed and real speed as

$$Err = \sqrt{\sum_t |\frac{v_{est}(t) - v_{real}(t)}{v_{real}(t)}|^2}. \tag{9}$$

The sum is over all measurements over a day. We calculate the errors for both adaptive and constant g-factor speed estimators and show the results in Figure 8. The dataset covers 37 days of measurements on 5 lanes of I-80.

Figure 8 shows that our adaptive g-factor algorithm has better performance than a constant g-factor algorithm, especially when the g-factor variation is large, as in lanes to the right. In fact, when using a constant g-factor, the error in speed is proportional to the standard deviation of the g-factor; the effect of large g-factor variation, however, is much more limited in the adaptive g-factor algorithm. For related empirical work see [5].

V. Conclusion

The paper presents an algorithms for adaptive calculations of the g-factor and speeds in single-loop detectors, using measurements of occupancy and flow alone. Comparison tests against speed measurements from double-loop detectors show that the estimates are very accurate.

Data from District 12 in Orange County suggest that estimates based on the assumption of a constant g-factor can

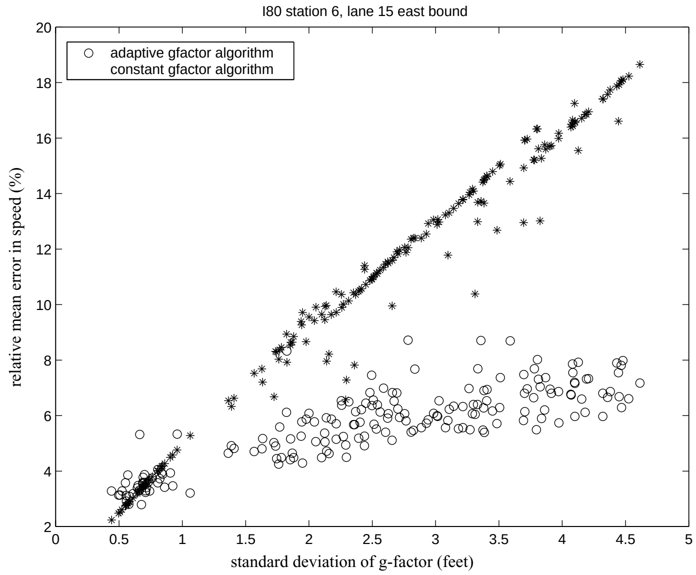

Fig. 8. Comparison of relative mean error in speed between adaptive g-factor algorithm and constant g-factor algorithm. The dataset is 37 days measurements on 5 lanes of I-80.

be easily off by 50 percent and so, such estimates cannot be trusted.

The study is part of an on-going empirical project on performance measurement of California freeways.

References

[1] D.J. Dailey, "A statistical algorithm for estimating speed from single loop volume and occupancy measurements," *Transportation Research B*, 33B(5): 313-22, June 1999.

[2] transacct.eecs.berkeley.edu

[3] C. Chen, K. Petty, A. Skabardonis, P. Varaiya and Z. Jia, "Freeway performance measurement system: mining loop detector data," 80th Annual meeting of the Transportation Research Board, Washington, D.C., January 2001.

[4] B. Coifman, "Improved velocity estimation using single loop detectors," *Transportation Research A*, 2001, forthcoming.

[5] B. Coifman, D. Lyddy, A. Skabardonis, "The Berkeley Highway Laboratory—Building on the I-880 field experiment," *Proc. IEEE ITS Council Annual Meeting*, 5-10, 2000.