An Adaptive Algorithm for Freeway Speed Estimation with Single-Loop Measurements

Yinhai Wang and Nancy L. Nihan
Department of Civil and Environmental Engineering
University of Washington
Box 352700, Seattle, WA98195-2700

ABSTRACT

Accurate, real-time traffic speed data are important inputs to successful freeway traffic management systems. Unfortunately, vehicle speeds cannot be directly measured by single-loop detectors, which are the most common detectors available in current freeway infrastructures. Algorithms are required to estimate speed using single-loop measurements. In this paper, we present a two-step speed estimation algorithm: in the first step, single loop measurements are filtered to screen out intervals containing long vehicles; and in the second step, space-mean speed is calculated using measurements for intervals containing only passenger cars. Twenty-four hour data that contain both free flow conditions and moderately congested conditions are used to test the algorithm. Speeds estimated by the proposed method are very close to the speeds observed by the corresponding dual-loop detector. Compared to the commonly adopted speed estimation algorithm with unfiltered data, the proposed method improves speed estimation accuracy significantly.

Keyword: speed, filter, vehicle length, single loop

1. INTRODUCTION

Advanced traffic management systems (ATMS) and advanced traveler information systems (ATIS) improve the efficiency of freeway networks. Both ATMS and ATIS require accurate and reliable speed data for successful operation. Although dual-loop detectors provide reliable speed data, most freeway networks have too few of these to meet the ATMS and ATIS operational needs. The current ATMS and ATIS systems rely mainly on single-loop data, as single-loop detectors are much more widely deployed in the existing freeway systems. Single-loop detectors, however, measure nothing but volume and lane occupancy directly. Traffic speed must be estimated from these volume and lane-occupancy measurements. Therefore, the ability to use single-loop measurements for accurate speed estimation is of practical significance for transportation researchers and operators.

In this paper, an adaptive algorithm for freeway speed estimation using single-loop outputs is described. First, previous studies that estimated traffic speed based on single-loop data are briefly reviewed. This is followed by a discussion on the inherent inaccuracies associated with the most commonly adopted speed estimation algorithm. Next, a methodology that can improve the accuracy of speed estimation using single-loop data is presented and evaluated. There are two steps in the proposed method: the first is the identification and separation of interval measurements with or without long vehicles (LVs, defined as vehicles longer than 26 ft or 7.92 m), and the second is the calculation of space-mean speed using only measurements for intervals without LVs. Finally, findings of this study are summarized and future studies are recommended.

2. BACKGROUND

According to the fundamental traffic flow equation, space-mean speed can be calculated by Eq. (1) if hourly vehicle volume and lane density are known [1],

$$\textit{space-mean speed} = \textit{volume} / \textit{density} \tag{1}$$

Volume is a direct output of a single loop, but density is not. However, a single-loop detector measures lane occupancy, and the lane occupancy is the product of density and the mean effective vehicle length (MEVL) under the uniform vehicle length assumption [2]. Hence, speed can be estimated using single-loop measurements as shown in Eq. (2).

$$\bar{s}(i) = \frac{n(i)}{T \cdot \frac{o(i)}{\bar{l}(i)}} \tag{2}$$

Where i = time interval index; $\bar{s}$ = space-mean speed; n = vehicles per interval; o = percentage of time loop is occupied by vehicles (lane occupancy); T = time length per interval; and l = MEVL for the interval. The MEVL is roughly equivalent to the sum of vehicle length and the length of the single-loop detector.

Athol [3] neglected the MEVL difference from interval to interval, and suggested Eq. (3) for speed estimation.

$$\bar{s}(i) = \frac{n(i)}{T \cdot o(i) \cdot g} \tag{3}$$

Where g is often referred to as speed estimation parameter and has a constant value equivalent to the reciprocal of the MEVL. Since Eq. (3) does not require any complicated calibration and uses only single-loop outputs for speed estimation, it has been commonly adopted in practice.

In reality, however, since the MEVL may vary dramatically from time to time, neglecting the variation in MEVLs can result in biased speed estimation [4]. Therefore, instead of using a constant g value in Eq. (3), Wang and Nihan [5] suggested that the g value should be updated periodically in response to the changing traffic composition. They proposed a log-linear model for estimating the MEVL of each estimation period using the statistical moments of occupancy and volume. Then, the estimated MEVL was applied to calculate the g value for the period. This proposed methodology reduced estimation bias significantly. Coifman *et al* [6] suggested using Eq. (3) for median speed estimation rather than mean speed estimation, for a group of consecutive intervals or vehicles, to avoid the bias. Since most vehicles are passenger cars under typically traffic conditions and passenger car lengths do not vary much, a constant g value works much better for estimating median speeds than mean speeds.

In addition to studies that use Eq. (3) for speed estimation, other methodologies have been developed. Pushkar *et al* [7] developed a cusp catastrophe theory model to estimate speed. Comparison of the estimation results between using their model and Eq. (3) found that the cusp catastrophe theory model gave more accurate results. Dailey [8] considered random errors in the measurements and used a Kalman filter to estimate speed. The estimated average speeds per interval were basically consistent with the observed average speeds, but the estimated variance over the entire study phase was significantly smaller than the observed variance. Sun and Ritchie [9] proposed a linear model to estimate individual vehicle speeds with slew rates of single-loop inductive waveforms. They concluded that their proposed algorithm performed better than conventional methods with single-loop measurements, and was robust under different traffic conditions.

Though the aforementioned methods have various advantages, they are not yet well accepted at present. Eq. (3) is still widely adopted in practice.

3. SPEED ESTIMATION WITH Eq. (3)

Several studies [4, 10] have addressed the applicability of Eq. (3) for speed estimation, but the results have been inconsistent. To further examine the accuracy of Eq. (3), we selected two 30-min data sets, including both single and dual loop measurements of a dual-loop station (we use the first single loop of a dual loop detector as our single-loop data source), from the Washington State Department of Transportation (WSDOT) loop detection system on May 13 (Thursday), 1999. One data set (a) was collected from the second-lane (from the right) loop detectors of station ES-516R on Eastbound SR-520 from 7:00 to 7:30 am. The other data set (b) was collected from the second-lane (from right) loop detectors of station ES-167D on Southbound I-5 from 3:30 to 4:00 pm.

Based on Eq. (3), the speed estimation parameter, g, can be calculated by

$$g = \frac{1}{T}\frac{\sum_{i=1}^{m}\frac{n(i)}{o(i)}}{\sum_{i=1}^{m}\bar{s}(i)} \tag{4}$$

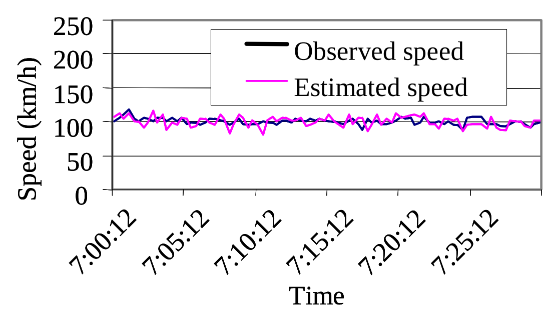
Where m is the number of samples and, for a 30-min data set, m = 90. Eq. (4) is actually the OLS (Ordinary Least Square) estimator of g. By applying this estimated g to Eq. (3), speed is calculated for each 20-second interval. The estimated speeds and dual-loop observed speeds are plotted in Figure 1. We can see that although the same equation was applied for speed estimation to both sites, the goodness of fit for each is quite different.

If we define estimation error as

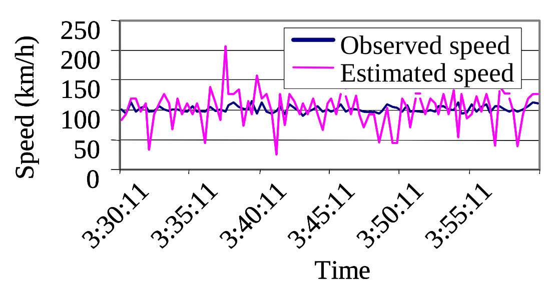
$$\varepsilon(i) = \textit{estimated speed} - \textit{observed speed} \tag{5}$$

then we can use this statistical variable to measure the accuracies of speed estimates for the two selected sites.

Figure 1(a) shows that the estimated speeds at station ES-516R on SR-520 are very close to the observed speeds. The absolute values of the estimation errors for 20-second intervals are no larger than 19.45 km/h, or 19.4 percent of the space-mean speed. The standard deviation of the estimation errors is 7.50 km/h. Considering the dynamic features of traffic flow and the possible data errors, the estimates are not bad. On the other hand, Figure 1(b) shows that the estimation errors at station ES-167D on I-5 are much larger. The maximum absolute value for the estimation errors is 110.36 km/h, or 109.0 percent of the space-mean speed, and the corresponding standard deviation of the estimation errors is 28.99 km/h.



(a) Data from loop detectors at SR-520EB & 84th St.



(b) Data from loop detectors at I-5SB & 145th St

FIGURE 1 Comparison of speed trap observed speeds and Equation (3) estimated speeds

The key factor that caused the difference in the estimation accuracy at stations ES-516R and ES-167D is believed to be the vehicle composition. As an example, consider an extreme case. In one 20-second interval, only one 3-m-long car is observed, and in the next 20-second interval, a 21-m-long truck passed over the loop. Even if the two observed vehicles were traveling at the same speed, the estimation obtained with a constant *g* value in Eq. (3) can give a car speed six times greater than the truck speed! In data set (a), of the 1180 observed vehicles, only 4 are LVs. But in data set (b), 18 out of 172 detected vehicles are LVs. The difference in the LV percentage between the two data sets, 0.34 percent for data set (a) and 10.47 percent for data set (b), is substantial.

Since LV lengths are very different from those for passenger cars (PCs, defined as vehicles shorter than 26 ft or 7.92 m, corresponding to Bin1 in the WSDOT loop detection system), the mixed rate of these two categories determines the MEVL for an interval. A constant *g* value in Eq. (3), however, indicates that the MEVL should be constant across all time intervals, which correspondingly requires the traffic composition to be consistent. If LV percentage changes significantly from interval to interval, the MEVL may exceeds reasonable fluctuation range and cause Eq. (3) to produce erroneous speed estimates.

Based on the above example and analysis, Eq. (3) does not fit well when the LV percentage changes significantly from interval to interval. Therefore, the speed estimation accuracy for Eq. (3) depends on the variation of traffic composition over time intervals.

4. METHODOLOGY

Due to the randomness of LV arrivals, it is almost impossible to have LVs uniformly presented at each interval. Typically, LV percentage various significantly over time. Obviously, directly applying single-loop measurements to Eq. (3) will probably cause serious estimation errors. Our solution to this particular problem is to filter single-loop outputs before inputting to Eq. (3) so that the MEVL can be consistent across intervals. The proposed filtering algorithm is based on the length difference between LVs and PCs. Its effectiveness depends largely on the features of vehicle-length distributions.

Features of Vehicle Length Distributions

Dual-loop detector data are used to examine the characteristics of typical vehicle length distributions on the freeway system. Since the WSDOT loop detection system provides aggregated measurements for 20-second intervals, only the average vehicle length is available for each interval. This indicates that single-vehicle length is not available when two or more vehicles present in one interval. Single-vehicle length data can be extracted only from intervals with exactly one vehicle presented over the entire interval duration. A total of 4703 valid single-vehicle lengths (those with nonzero error flags, which indicating some kind of data error, are excluded) were picked up from a fourteen-day (from May 3 through May 16, 1999) dual-loop data set collected by the second-lane (from right) dual-loop detector (_M___T2) at Station ES-167 on Southbound I-5. Of these observed single vehicles, 580 are LVs, which accounted for 12.33 percent of the total. The lengths of these LVs, however, ranged widely from 8.23 m to 28.35 m. Descriptive statistics of the observed LVs and PCs are given in Table 1.

TABLE 1 Descriptive statistics of speed trap measured vehicle lengths

	PCs only	LVs only	All vehicles
No of cases	4123	580	4703
Minimum (m)	2.13	8.23	2.13
Maximum (m)	7.92	28.35	28.35
Range (m)	5.79	20.12	26.21
Median (m)	4.57	20.27	4.57
Mean (m)	4.64	19.44	6.46
Standard Deviation (m)	0.67	4.29	2.09

The frequency distribution of the observed single-vehicle lengths is shown in Figure 2. Two peaks are obvious in the plot: one at about 5.0m, representing the length concentration for PCs, and the other at about 20.0m, representing that for LVs. The fact that the first peak is much higher than the second peak indicates that PC lengths vary much narrower than LV lengths. In Figure 3, the frequency distribution of PC lengths is shown with the corresponding normal distribution curve. The normal distribution curve fits the count histogram very well. The Kolmogrov-Smirnov Z statistic for the PC lengths is 11.42 (corresponding to $p < 0.01$), which strongly indicates that PC lengths are normally distributed.

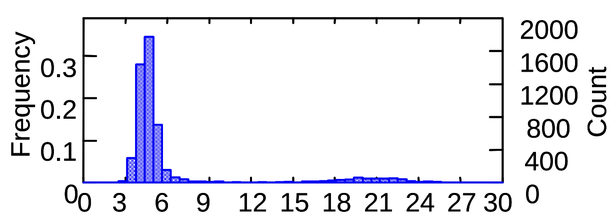


FIGURE 2 The vehicle length distribution observed by a dual loop detector at I-5SB & 145th St.

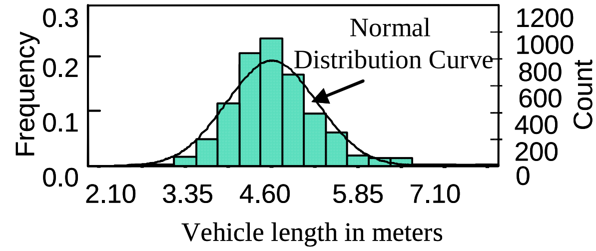


FIGURE 3 PC length distribution with the normal distribution curve

The same analysis procedure has been applied to several other sites in Puget Sound region. The length distribution of PCs is fairly consistent, while that of the LVs varies very much from site to site. The consistent distribution of PCs provides a good foundation for the proposed single-loop data-filtering algorithm to be described later.

Choosing the value for *g*

In this study, we choose $g=1/(\mu_{pc}+l_{loop})$, where μ_{pc} is the mean vehicle length for PCs (4.64 m for this study as shown in Table 1) and l_{loop} is the loop length (1.83m for this study). Such a *g* value corresponds to traffic stream with only PCs. There are two reasons for calculating *g* in this way: First, since PC lengths drift narrowly from their mean, the MEVLs for intervals with only PCs should be very close to each other and, therefore, *g* is close to a constant. Second, LV percentage is generally less than 20% for most highway routes in Washington State and such a *g* value should be true for majority of 20-second intervals. When such a g value is used in Eq. (3), only measurements from LV-free intervals can be used for speed calculation. This implies that intervals with LVs must be screened out from speed estimation.

Removing the Measurements for Intervals Containing LVs

Intervals containing LVs have longer MEVLs. However, since single-loop detectors measure only volume and occupancy for each interval, we cannot tell whether the interval contains LVs by looking only at the interval's measurements. To solve the problem, we need to look at the measurements for all the consecutive intervals, in a period, simultaneously. The relative relationships for these consecutive interval measurements provide critical information for separating the intervals with LVs. In this study, the terms "period" and "interval" are used with significant distinction. A period represents a longer time duration than an interval. An interval is 20 seconds long determined by the WSDOT loop detection system. A period contains several intervals and is chosen by the requirements of the proposed algorithm. For our data, a period of 5 minutes is selected. This is equivalent to a time length of 15 intervals. This period length has been shown to work reasonably well by our previous studies [5,11] using the same data set.

For any period *j*, there are 15 intervals. We need to assign each interval to one of the three categories: (1) intervals containing no vehicles; (2) intervals containing only PCs; and (3) intervals containing LVs. Only measurements for Category (2) intervals will be used for speed estimation for period *j*. Measurements for intervals of Categories (1) and (3) will be discarded.

To identify Category (1) intervals is easy (because both measurements should be 0), but separating Category (3) intervals from Category (2) is difficult. Here, we introduce a single-loop data-filtering algorithm to screen intervals with LVs from those without. Single-loop data filtering is the first step of this proposed adaptive speed estimation method. The filtering algorithm starts with sorting all the 15 intervals in a period in <u>ascending</u> order of average occupancy per vehicle:

$$\frac{o(1,j)}{n(1,j)} \le \frac{o(2,j)}{n(2,j)} \le \dots \le \frac{o(i,j)}{n(i,j)} \le \dots \le \frac{o(15,j)}{n(15,j)} \quad \text{for } 1 \le i \le 15 \qquad (6)$$

If an interval *b* contains no vehicles, then $o(b, j)/n(b, j)$ is defined as 0. Suppose there are *p-1* intervals that contain no vehicles, then these intervals are assigned to Category (1). Dropping these Category (1) intervals from further analysis yields

$$\frac{o(p,j)}{n(p,j)} \le \frac{o(p+1,j)}{n(p+1,j)} \le \dots \le \frac{o(i,j)}{n(i,j)} \le \dots \le \frac{o(15,j)}{n(15,j)} \quad \text{for p} \ge 1 \text{ and } p \le i \le 15 \qquad (7)$$

Because of the relatively low LV volumes and the randomness of LV arrivals, it is fairly safe to assume that there is at least one interval containing only PCs in each period. To check the violation probability for this assumption, we examine 24-hour data collected by dual-loop detector ES-167D:_M___T2 on May 13, 1999. For all 288 periods, none violates this assumption. Thus, from the sorting result shown in Eq. (7), we can conclude that interval *p* should contain only PCs. Since the PC lengths vary very narrowly, the MEVL for PCs can be treated as a constant $\bar{l}_{pc}$ (the sum of μ_{pc} and l_{loop}) without introducing significant errors. Therefore, $\bar{l}_{pc}$ can be used to approximate the MEVL for interval *p*.

Assume that both LVs and PCs travel at the same mean speed during each interval and that the mean speed is consistent over each time period. Then for any two intervals *i* and *k*, we have

$$\bar{s}(i,j) = \bar{s}(k,j) \qquad (8)$$

Since the space-mean speed for an interval *i* can be expressed as

$$\bar{s}(i,j) = \frac{1}{T} \cdot \frac{n(i,j) \cdot \bar{l}(i,j)}{o(i,j)} \qquad (9)$$

expanding Eq. (8) and rearranging terms yields the relationship of MEVLs between interval *k* and interval *i* as

$$\bar{l}(i,j) = \frac{n(k,j)}{n(i,j)} \cdot \frac{o(i,j)}{o(k,j)} \cdot \bar{l}(k,j) \qquad (10)$$

Eq. (10) indicates that when the MEVL for interval *k* is known, the MEVL for any interval *i* in period *j* can be easily calculated. So far, we have known that the MEVL for interval *p* is $\bar{l}_{pc}$. Then letting $k = p$, the MEVL for any interval *i* can be calculated by

$$\bar{l}(i,j) = \frac{n(p,j)}{n(i,j)} \cdot \frac{o(i,j)}{o(p,j)} \cdot \bar{l}_{pc} \qquad (11)$$

The critical value used for judging whether an interval contains LVs is 5.98 m, which is equivalent to the sum of the mean and twice the standard deviation of PC lengths as given in Table 1. Though such a boundary cannot guarantee all intervals with LVs are removed, it does screen out most of the LV-presented intervals while keeping the majority of PC-only intervals in our analysis.

If $\bar{l}(i, j)$ - loop length ≤ 5.98 m, then interval *i* is identified as having no LVs, and assigned to Category (2). Otherwise, interval *i* is identified as an LV-presented interval, and all the remaining intervals (from *i* to 15) are assigned to Category (3).

Speed Estimation Using Measurements of Intervals without LVs

The second step of this proposed adaptive speed estimation method is to calculate the period speed using only measurements of Category (2) intervals. Suppose interval *q+1* is identified to have LVs, then the number of effective measurements in the period is reduced from 15-p+1 to *r* (where *r* = *q-p+1*) sets, i.e. there are *r* intervals belong to Category (2). These *r* sets of measurements are used to calculate the total applicable volume and corresponding occupancy as follows,

$$N_{pc}(j) = \sum_{i=p}^{q} n(i, j) \tag{12}$$

$$O_{pc}(j) = \sum_{i=p}^{q} o(i, j) \tag{13}$$

Then space-mean speed for period *j* is calculated by

$$\bar{s}(j) = \frac{1}{T} \cdot \frac{\sum_{i=1}^{r} n(i, j)}{\sum_{i=1}^{r} o(i, j) \cdot g} \tag{14}$$

We can see that only the measurements for Category (2) intervals are used for the speed estimation. This estimated speed, though based solely on Category (2) data, is regarded as the representative speed for the entire period.

5. ESTIMATION RESULTS

Based on the methodology presented, a computer program is developed to implement the entire procedure of the proposed method from loading data to printing speed estimation results. If real-time single-loop data can be provided as inputs to the program, traffic speed estimates can be produced in real time.

Twenty-four hour data collected by single-loop detector _MS___2 at station ES-167D on May 13, 1999 are used for testing the program. Observed speed data from the dual-loop detector formed by single loops _MS___2 and _MS __S2 are used to verify the estimation results. Since the proposed algorithm produces speed estimation for every 5-min period, dual-loop observed 20-second interval speeds are used to calculate space mean speed for each 5-min period. Dual-loop observed speeds, speeds estimated using Eq. (3) with unfiltered single-loop data, and speeds estimated by the proposed method (using filtered data) are compared and illustrated in Figure 4.

The correlation coefficient for the observed speeds and the speeds estimated by the proposed method is 0.810, which is significantly higher than that of 0.637 for the observed speeds and speeds estimated using unfiltered data. The standard deviation of estimation error (ε) for the proposed method is 5.58 km/h, while that for commonly adopted speed estimation method (using unfiltered single-loop measurements) is 9.87 km/h. This indicates that the proposed algorithm improved speed estimation accuracy over the commonly adopted method.

6. CONCLUSIONS

Eq. (3) is widely adopted for speed estimation using single-loop outputs. It uses a speed estimation parameter *g* in calculation. Though g is regarded as a constant in practice, researchers disagree on whether *g* should be treated as a constant. We addressed this issue first with our test data sets. We found that when the MEVL changes widely from interval to interval, estimation results with Eq. (3) are poor. However, this does not mean that Eq. (3) is unsuitable for speed estimation. Actually, Eq. (3) has many advantages, such as simplicity and transferability, over many of the newly developed methods. Under certain traffic circumstances, such as a consistent traffic composition over time, Eq. (3) can produce favorable speed estimates.

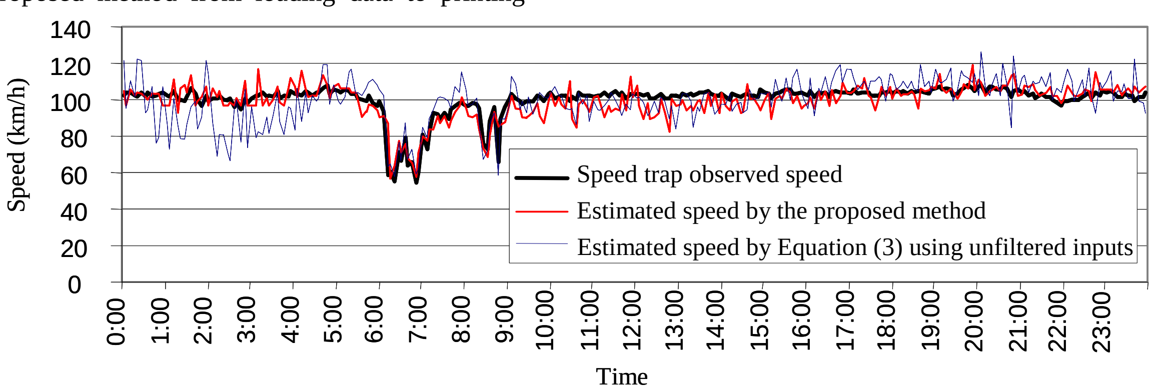


FIGURE 4 Comparison of speed trap observed speeds, Equation (3) estimated speeds using unfiltered inputs, and the proposed method estimated speeds.

Our proposed speed estimation method was intended to take the advantage of this commonly adopted equation, and overcome its problems by screening out intervals containing LVs. The fact that PC lengths vary narrowly around their mean provided a sound foundation for our single-loop data-filtering algorithm. By utilizing the relative relationships among interval measurements, intervals were classified into three categories: (1) intervals containing no vehicles; (2) intervals containing only PCs; and (3) intervals containing LVs. Only measurements for Category (2) intervals were used for speed estimation. Measurements for Category (1) and Category (3) intervals were discarded. Such estimated speeds were actually the space-mean speed for PCs because LV-presented interval measurements were discarded from speed calculation. However, considering that the speed difference between PCs and LVs are reasonably small compared to the speed estimation error, we can regard the estimated speed as the space-mean speed for all vehicles in each period.

Comparisons between dual-loop observed speeds, speeds estimated by the commonly adopted method, and speeds estimated using our proposed method showed that the proposed method provided better speed estimation. Its estimation accuracy was significantly higher than that using unfiltered data. Also, our 24-hour estimation results indicated that the method worked reasonably well for both free-flow and moderately congested conditions. A computer application implementing the proposed method has been developed. This application is capable of providing instant speed estimates when real-time single-loop inputs are available.

Further studies are needed to examine whether the extracted vehicle length distributions are transferable. If yes, the calibration for the proposed algorithm can be largely simplified when applying to a different site (only loop size is needed in this case). Otherwise, the mean and standard deviation for PC lengths are also required before executing the proposed algorithm.

7. REFERENCES

[1] J.G. Wardrop, "Some Theoretical Aspects of Road Traffic Research", **Proc. Instn. Engrs**., Vol 1, No. 2, 1952, pp. 325 - 363.

[2] M. Cassidy and B. Coifman, "The relation between average speed, flow and occupancy and the analogous relation between density and occupancy", **Transportation Research Record** 1591, TRB, National Research Council, Washington, DC, 1997, pp. 1-6

[3] P. Athol, "Interdependence of Certain Operational Characteristics within a Moving Traffic Stream", **Highway Research Record** 72, 1965, pp. 58-87.

[4] F.L. Hall and B.N. Persaud. "Evaluation of Speed Estimates Made with Single-detector Data from Freeway Traffic Management Systems", **Transportation Research Record** 1232, TRB, National Research Council, Washington, D.C., 1989, pp. 9-16.

[5] Y. Wang and N.L. Nihan, "Freeway Traffic Speed Estimation Using Single Loop Outputs", **Transportation Research Record** 1727, TRB, National Research Council, Washington, D.C., 2000, pp. 120-126.

[6] B. Coifman, S. Dhoorjaty, and Z. Lee, "Estimating Median Velocity Instead of Mean Velocity at Single Loop Detectors", **Transportation Research: Part C**, Vol. 11, No. 3-4, 2003, pp. 211-222.

[7] A. Pushkar, F.L. Hall, and J.A. Acha-Daza, "Estimation of Speeds from Single-loop Freeway Flow and Occupancy Data Using Cusp Catastrophe Theory Model", **Transportation Research Record** 1457, TRB, National Research Council, Washington, D.C., 1994, pp. 149-157.

[8] D.J. Dailey, "A Statistical Algorithm for Estimating Speed from Single Loop Volume and Occupancy Measurements", **Transportation Research: Part B**, Vol. 33, No. 5, 1999, pp. 313-322.

[9] C. Sun and S.G. Ritchie. "Individual Vehicle Speed Estimation Using Single Loop Inductive Waveforms", **ASCE Journal of Transportation Engineering**, Vol. 125, No. 6**,** 1999, pp. 531 - 538.

[10] B. Coifman, "A. New Methodology for Smoothing Freeway Loop Detector Data: Introduction to Digital Filtering", **Transportation Research Record** 1554, TRB, National Research Council, Washington, D.C., 1996, pp. 142-152.

[11] Y. Wang and N.L. Nihan, "Dynamic Estimation of Freeway Large Truck Volume Based on Single-Loop Measurements", **Transportation Research Record**, preprint CD-ROM, paper 01- 2853.