An Enhancement to Speed Estimation Using Single Loop Detectors

Wei-Hua Lin[1], Joy Dahlgren[2], and Hong Huo[3]

Abstract— **Speed estimation plays an important role in many components in Intelligent Transportation Systems (ITS). In practice, traffic data from single loop detectors are one of the dominant data sources. Speed estimation from single loop detectors is mainly based on occupancy data, a conversion factor from occupancy to density (which is potentially related to the vehicle length), and the assumed relationship between flow, speed, and density. This paper investigates the discrepancy between the speed estimated with single loops and the speed estimated with double loops. It was found that the discrepancy between the two is mainly caused by the high variance in vehicle pace, which, especially under congested traffic, is often accompanied with the high skewness. Accuracy can be improved by computing occupancy in a different way, using the median vehicle passage time over the detector as opposed to the mean vehicle passage time often used in the conventional method. The performance of the enhanced speed estimation method is very promising. The use of the median vehicle passage time reduces the skewness of pace data.**

Index Terms — **Traffic sensors, traffic flow theory, traffic management systems, traveler information systems, and traffic control.**

I. INTRODUCTION

ESTIMATION of prevailing vehicle speeds for traffic streams plays a crucial role in traffic management and traveler information systems since speed is an important indicator of traffic conditions. This is especially true for Intelligent Transportation Systems (ITS). In recent years, many algorithms and procedures developed for ITS applications, such as the variable message sign, ramp metering, and dynamic traveler information systems, have been using vehicle speed information in one way or another on a real time basis. An improvement to the accuracy in speed estimation will greatly enhance the performance of these algorithms and procedures.

For the past decade or so, traffic surveillance technologies have been improved at a rapid pace. Emerging sophisticated technologies have made it possible to capture the characteristics of vehicles in ways that were not possible in the past. Probes and other devices have been adopted for collecting traffic data relevant to traffic control. However, traffic data from loop detectors remain to be one of the primary data source used in practice. There are two types of loop detectors, double loops and single loops. Though double loops are better in capturing vehicle speed of a traffic stream, detectors for traffic counts used in many places are predominantly single loops. The focus of this paper is on speed estimation with single loops. It is expected that the use of loop detectors, especially single loops, will still be dominant for traffic data collection in the foreseeable future. Therefore, the study of how to obtain accurate speed measurement from single loops is not only of interest to the research world, but also of interest to the field application. The enhancement to the quality of loop detector data in estimating traffic states has become a challenging task that many traffic operation centers have to deal with on a daily basis.

It is often contended that the estimation of speed is very challenging for congested traffic. When traffic is congested, occupancy data become very noisy, making it difficult to capture the relationship between flow, density, and speed. To address this issue, various models were proposed in the past, including the one suggesting a discontinuous flow-density relationship [1]. For a detailed review of both recent and past empirical research on congested flow, see [2]. Unlike speed measurement using double loop detectors, speed measurement using single loop detectors is indirect. In practice, speed calculated from single loops is based on flow, occupancy, and a conversion factor, g, i.e., the estimated speed = flow / (occupancy * g), based on the assumption that occupancy is linearly proportional to density. This assumption has been challenged for the past two decades. Hall and Persaud argued that the g value is not a constant [3]. Rather, it is a function of occupancy. A better fit can be obtained when g changes in accordance with the change in the occupancy level. The data used in their research are conventional data aggregated for a pre-specified time interval. Cassidy and Coifman [4], however, contended that a constant g is reasonable if traffic parameters are measured in the manner consistent with the one defined by Edie [5]. They suggested the use of non-conventional data aggregated on a pre-specified number of vehicles, which effectively allows variable lengths for sampling intervals.

In this paper, we first show with the field data some interesting features regarding the variance of some of the key traffic parameters important for speed estimation at different speed levels. We found that the high variance and the skewness of the pace data under congested traffic have a profound effect on the accuracy in speed estimation. Based on

1. Wei-Hua Lin is with Department of Systems and Industrial Engineering, The University of Arizona, Tucson, AZ 85721-0020, U.S.A. (email: weilin@sie.arizona.edu);
2. Joy Dalhgren is with the Institute of Transportation Studies at University of California at Berkeley, Berkeley, CA, 94804-4698, U.S.A. (email: joy@uclink4.berkeley.edu).
3. Hong Huo is with Department of Systems and Industrial Engineering, The University of Arizona, Tucson, AZ 85721-0020, U.S.A. (email: hhuo@email.arizona.edu)

this finding, we proceed to propose a method that uses median of the vehicle passage time of the detectors instead of the mean of the vehicle passage time to represent the occupancy information for speed estimation. This simple modification reduces the skewness in data and yields more robust estimation compared with the method using occupancy calculated in the conventional way. The method appears to be promising when tested with the limited data sets available to this study. It consistently outperformed the method based on the occupancy data processed in a conventional way.

The paper is organized as follows. In the next section, we describe briefly how speeds are estimated using single loop detectors and the key factors that would affect the accuracy in estimation. Sec. 3 provides some empirical evidence showing the behavior of the variance of pace, speed, and occupancy under different traffic conditions and how it would affect the accuracy in estimating the prevailing vehicle speeds of a traffic stream. Sec. 4 describes the proposed method for speed estimation. Field data were then used to compare the performance of the proposed method and the performance of the conventional method based on the commonly used performance measures, the mean squared error and the mean relative error. The results are summarized in Sec. 5. Sec. 6 summarizes the proposed method and identifies the direction for future research.

II. Conventional Method for Speed Estimation Using Single Loop Detectors

The following is a set of notation used in the paper:

l_i : the length of the ith vehicle;

v_i : the speed of the ith vehicle;

v_s : the actual space-mean speed;

$\hat{v}_s$: the estimated space-mean speed;

τ_i : the passage time over the detector for the ith vehicle:

k : vehicle density (veh/distance unit);

q : vehicle flow (veh/time unit);

o : vehicle occupancy, percentage time detector is covered by vehicles;

p_i : the pace of ith vehicle, which is the reciprocal of the vehicle speed;

T : time interval during which traffic parameters are aggregated.

With double loops, density and occupancy can be measured independently. Vehicle density for a specific time interval T, in which there are n vehicles crossing the detector of concern, is calculated by:

$$k = \frac{\sum_{i=1}^{n} \frac{1}{v_i}}{T},$$

whereas vehicle occupancy is calculated by:

$$o = \frac{\sum_{i=1}^{n} \frac{l_i}{v_i}}{T}.$$

The relationship between occupancy and density is as follows:

$$o = \frac{\sum_{i=1}^{n} \frac{l_i}{v_i}}{T} = \frac{\left(\sum_{i=1}^{n} \frac{1}{v_i}\right)}{T} \sum_{i=1}^{n} \left(\frac{\frac{1}{v_i}}{\sum_{i=1}^{n} \frac{1}{v_i}} \right) l_i = k \sum_{i=1}^{n} \left(\frac{p_i}{\sum_{i=1}^{n} p_i} \right) l_i .$$

If we can approximate $\sum_{i=1}^{n} \left(\frac{p_i}{\sum_{i=1}^{n} p_i} \right) l_i$ by an appropriate function $g(\cdot)$, we should be able to estimate the space-mean speed for a given time interval T by:

$$\hat{v}_s = \frac{q}{k} = \frac{\frac{n}{T}}{\frac{o}{g(\cdot)}} = \frac{n}{\left(\sum_{i=1}^{n} \frac{1}{v_i}\right)} \frac{g(\cdot)}{\sum_{i=1}^{n} \left(\frac{p_i}{\sum_{i=1}^{n} p_i} \right) l_i} = v_s \frac{g(\cdot)}{\sum_{i=1}^{n} \left(\frac{p_i}{\sum_{i=1}^{n} p_i} \right) l_i} .$$

The estimation is accurate if $g(\cdot) = \sum_{i=1}^{n} \left(\frac{p_i}{\sum_{i=1}^{n} p_i} \right) l_i$ for every time interval. However, with single loops, neither p_i nor l_i can be measured directly. Recognizing that $g(\cdot)$ is to approximate a linear combination of vehicle length weighed by pace, it is natural to use the average vehicle length instead in hopes that $\frac{\sum_{i=1}^{n} l_i}{n} \approx \sum_{i=1}^{n} \left(\frac{p_i}{\sum_{i=1}^{n} p_i} \right) l_i$. An approximation to $\frac{\sum_{i=1}^{n} l_i}{n}$ can be obtained off-line if we assume that the average length of vehicles on the road does not vary substantially at different sampling intervals. This is how speed is measured with a conventional approach. Note that an accurate measurement of $\frac{\sum_{i=1}^{n} l_i}{n}$ would not necessarily lead to a good approximation of $\sum_{i=1}^{n}\left(\frac{p_i}{\sum_{i=1}^{n}p_i}\right)l_i$, even if we were able to update $\frac{\sum_{i=1}^{n}l_i}{n}$ at every sampling interval on a real-time basis. The variance structure of the pace would also affect the accuracy in speed estimation. It is not difficult to see that the difference between $\frac{\sum_{i=1}^{n}l_i}{n}$ and $\sum_{i=1}^{n}\left(\frac{p_i}{\sum_{i=1}^{n}p_i}\right)l_i$ should be small if the variance of pace remains to be small since $\frac{p_i}{\sum_{i=1}^{n}p_i}\approx\frac{1}{n}$. A large variance in pace would increase the discrepancy between $\frac{p_i}{\sum_{i=1}^{n}p_i}$ and $\frac{1}{n}$, which could potentially result in a large discrepancy between $\sum_{i=1}^{n}\left(\frac{p_i}{\sum_{i=1}^{n}p_i}\right)l_i$ and $\frac{\sum_{i=1}^{n}l_i}{n}$. In fact, the variance of pace could be a key contributor to the accuracy in speed estimation. Moreover, the strong skewness is usually accompanied with a large variance. We will show in the following section with empirical data that the variance of pace is indeed very large when traffic is congested. The skewness is also quite pronounced for the pace data under congested traffic.

III. Empirical Evidence of the Behavior of Pace

The data available to this study is 1/60[th] second data from a 4 mile-long stretch of freeway along Interstate 80 of California, provided by the Berkeley Highway Laboratory. They were collected from double loop detectors. The data kept at the tick level provides very detailed information about each individual vehicle as it crosses a given detector station. As indicated in Fig.1, each record in the data file contains the information of a single vehicle crossing a specific double loop, including the information about the station number, the lane number, and the four time stamps that track the times for the vehicle to cross the upstream (t_1 and t_2) and downstream (t_3 and t_4) loops with its front and rear bumpers, respectively.

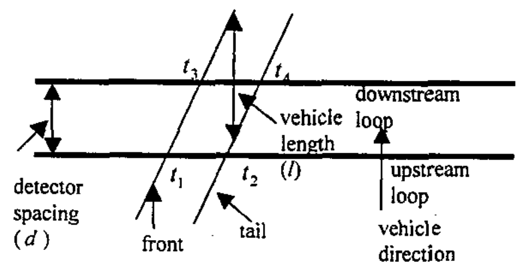

Fig. 1. Schematic representation of data collected from double loop vehicle detectors.

With data of this resolution, we can get very detailed information about each individual vehicle. Exact speed can be calculated by:

$$v_i=\frac{d}{t_3-t_1}\text{ (measured with the front bumper)}$$

or $$v_i=\frac{d}{t_4-t_2}\text{ (measured with the rear bumper)}$$

Vehicle length can be calculated by:

$$l_i=v_i(t_2-t_1)$$

or $$l_i=v_i(t_4-t_3)\,.$$

With data of this resolution, one can even obtain the acceleration and deceleration information of vehicles when they traverse the detector. A higher speed measured with the front bumper than the one measured with the rear bumper suggests a deceleration process. In addition to individual vehicle data, we can also produce aggregated data such as the commonly used flow, occupancy, density, and speed within a chosen sampling interval. The double loop data will be used as the "ground truth" data in this study.

A. Variance of speed and pace under different traffic conditions

When we compare the variance of the speed for congested traffic and the variance of the speed for uncongested traffic, we found that the difference between the two is not very large. In many situations, the variance of the speed for congested traffic tends to be only slightly higher than the variance of the speed for uncongested traffic. Occasionally, we observed that the variance of the speed for congested traffic could be smaller. This is reasonable in part because when traffic becomes congested, the room for drivers to freely control their speeds also becomes very small, whereas during the free flow traffic condition drivers tend to only loosely obey the speed limit sign and travel at their own desired speeds. However, the variance of the pace under the congested traffic condition is always much higher than that under the uncongested traffic condition.

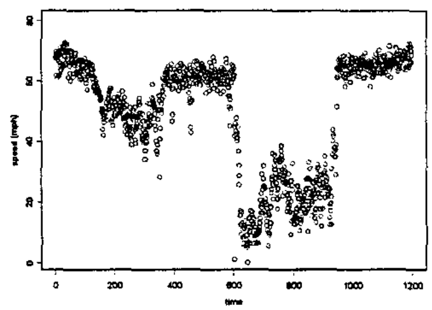

Fig. 2: Plot of speed vs. time (each time unit = 1 minute).

Fig. 2 is a plot of vehicle speeds for a time period of 1200 minutes, covering both congested and uncongested traffic. One can see that traffic has a high speed between t=410 to 510, and low speed for t = 750 to 850. The ratio of the variance of the speed for the time period (750, 850) to the variance of the speed for the time period (410,510) is 1.26, whereas the ratio of the variance of the pace for these two time periods is as high as 316.8. Even if we consider the speed difference during these two time periods and adjust the variance with respect to the mean speed, we still get 3.37 for the ratio of the variance of the speed and 103 for the ratio of the variance of the pace. We have also tested some other data sets. It appeared that in all of the cases we have tested the variance of the pace grows much more quickly than the variance of the speed as traffic becomes more congested.

For the data set we use sometimes the average vehicle length for light traffic is longer than that for congested traffic; the variance tends to be larger as well. This is reasonable since a large percentage of trucks is usually in full operation at midnight or early in the morning due to regulations that specifically prohibit trucks from using some major highways during certain time periods.

B. Variance of the vehicle passage time over detectors

The variance of the vehicle pace is closely related to the vehicle passage time over detectors, a measure that can be obtained with sing loops. By definition, vehicle passage times over the loop detector are used to compute occupancy. It is well known that the plot of flow-occupancy data often exhibits a well-clustered and smooth left regime (for uncongested traffic) but a scattered and noisy right region (for congested traffic). In other words, for the same level of flow, occupancy exhibits a very small variance for uncongesthted traffic but a large variance for congested traffic. In fact, one can easily find the answer to this phenomenon by plotting the occupancy data in a different way, showing the detailed vehicle times for each individual vehicle over the loop detector. In the past, vehicle times over loop detectors were rarely documented in the literature since they are the raw data, which are often unavailable to users.

Figs. 3 and 4 are plots for vehicle passage times over the detector. Each horizontal stripe represents the vehicle passage time over detectors. Each row is a collection of vehicle passage times recorded from a single one-minute sampling interval. To facilitate the comparison, the number of stripes at each row is identical, suggesting the same level of flow for each time interval. The number to the left is the average speed for that time interval in the unit of miles per hour. Fig. 3 is a plot for vehicles traveling at low to moderate speeds (< 45 mph), whereas Fig. 4 is a plot for vehicles traveling at high speeds (> 55 mph). The two plots exhibit very different patterns for vehicle passage time over the detectors. The vehicle passage times shown in Fig. 4 for uncongested traffic exhibit a smaller variance, whereas the vehicle passage times shown in Fig. 3 fluctuate wildly for congested traffic. Note that the vehicle passage time over detectors is calculated by l/v. A pair of similar vehicle passage time may come from a very different combination of vehicle length and speed.

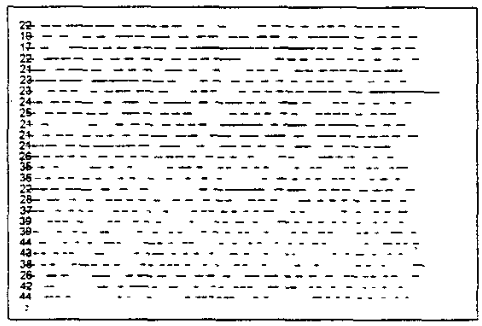

Fig. 3: Plot of times vehicles spent over loop detectors (speed < 45 mph).

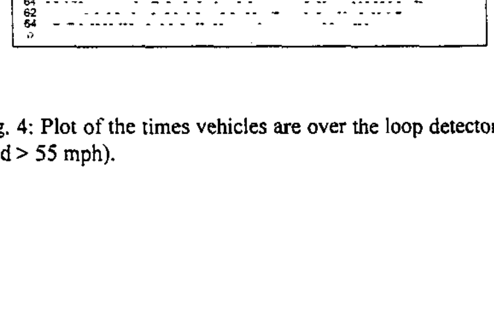

Fig. 4: Plot of the times vehicles are over the loop detectors (speed > 55 mph).

C. The relationship between the variance in pace and the accuracy in speed estimation

Our data show that the variance in pace is closely related to the accuracy in speed estimation. The variance of pace can be very large under stop-and-go traffic in which some vehicles may travel at a speed much lower than the prevailing speed of the traffic stream at the moment they cross the loop detectors. The poor estimation of the speed usually occurs when the variance in pace is very high, especially when the paces of a small portion of vehicles within the sampling interval are exceedingly large. A large fluctuation in pace would in turn lead to a large fluctuation in the vehicle passage time over the detector, as shown in Fig. 3. This observation is the key motivation for the development of an enhanced method for speed estimation proposed in the next section.

IV. Proposed Method for Estimating Vehicle Speed with the Median Vehicle Passage Time over Detectors

Let $\tau_i = \frac{l_i}{v_i}$ be the passage time over the detector for vehicle i, and $\bar{\tau}$ be the average vehicle passage time. We rewrite occupancy as follows:

$$o = \frac{\sum_{i=1}^{n} \frac{l_i}{v_i}}{T} = \frac{n \frac{\sum_{i=1}^{n} \tau_i}{n}}{T} = \frac{n\bar{\tau}}{T}.$$

The speed estimation is thus:

$$\hat{v}_s = \frac{\frac{n}{T}}{\frac{o}{g(\cdot)}} = \frac{g(\cdot)}{\bar{\tau}}.$$

As shown in Fig. 4, τ_i can be very noisy under congested traffic conditions. If we replace $\bar{\tau}$ by the median of the vehicle passage times over detectors, τ_M, we then have a new estimator for the space mean speed:

$$\hat{v}_s' = \frac{g(\cdot)}{\tau_M},$$

where τ_M is the median of τ_i, i.e., $P\{\tau_i \le \tau_M\} \ge 1/2$ and $P\{\tau_i \ge \tau_M\} \ge 1/2$. If $\tau_M = (l/v)_M$ is a good approximation of l_M/v_M, then the estimation of the space mean speed becomes:

$$\hat{v}_s' = v_M \frac{g(\cdot)}{l_M}.$$

The above estimation would perform well if the following two conditions are satisfied:

(a) v_M varies with traffic conditions and reflects the prevailing traffic speed; and

(b) l_M remains relatively stable and is invariant of traffic conditions.

Based on condition (b), we can simply use a constant g as a representation for l_M. The use of the median vehicle passage time over the detectors would effectively reduce the skewness of pace that would arise under stop-and-go traffic and the presence of long vehicles. A variation of this scheme is to use the average vehicle passage times for a range of τ_i chosen from the middle 30% percentile. We have tested this alternative with field data. For the limited data sets we use, the gain to this variation over the median method seems to be very limited.

In Fig. 5, we compare the result from the proposed method with the result from the conventional method based on the one-minute data. The plots were generated with data sets from lane 3 at station 3. As shown in the Figure, our method (labeled as the median method) performs significantly better than the conventional method (labeled as the mean method).

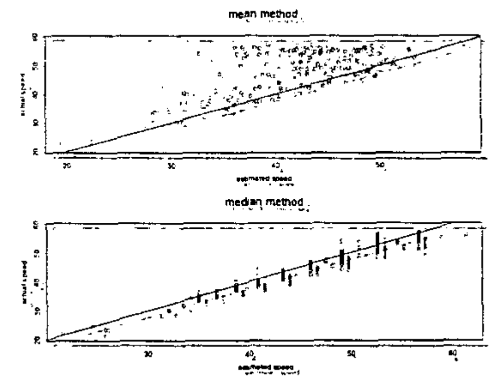

Fig. 5: Comparison of the proposed method (labeled as the median method) and the conventional method (labeled as the mean method).

V. Performance Evaluation

We compare the proposed model with the conventional method based on the two commonly used measures of performance, the mean relative error (MRE) and the mean squared error (MSE). The data used for comparison were collected at one-minute sampling interval. Seven data sets were used. The results for MRE and MSE are displayed in Tables 1 and 2, respectively. For the mean relative error measurement, the proposed method seems to be consistently better than the conventional method. The differences between the two in some cases are more significant than others. For the mean squared error measurement, the effectiveness of the proposed method seems to be more pronounced. This is intuitive since the proposed method should be more robust than the conventional method. By choosing the median of the vehicle passage time, the proposed method eliminates the extreme cases in which detector errors, acceleration or deceleration of vehicles, or excessively long vehicles may introduce excessive long passage time (long pace as well). Yet the use of median preserves the prevailing passage time.

Table 1: Comparison of the results from the proposed model and the conventional model: Mean relative errors (MRE).

Data	The proposed method	The conventional method
Station 5, lane 3	0.133	0.199
Station 5, lane 7	0.110	0.228
Station 3, lane 7	0.068	0.202
Station 3, lane 3	0.036	0.107
Station 6, lane 2	0.158	0.169
Station 6, lane 4	0.106	0.141
Station 1, lane 3	0.0981	0.151

Table 2: Comparison of the results from the proposed model and the conventional model: mean squared error (MSE).

Data	The proposed method	The conventional method
Station 5, lane 3	31.233	83.713
Station 5, lane 7	16.226	94.458
Station 3, lane 7	5.683	86.116
Station 3, lane 3	3.844	52.576
Station 6, lane 2	27.019	28.534
Station 6, lane 4	22.842	49.111
Station 1, lane 3	18.687	65.799

VI. Conclusion and Future Research

This paper proposes an enhancement to the existing methods for estimating speeds using single loop detectors. Like most existing methods, it computes the speed by converting occupancy to density first and then making use of the fundamental relationship between flow, density, and speed. The proposed method departs from the existing methods in that: instead of using the average vehicle passage time over detectors for computing occupancy, it uses the median of the vehicle passage time over detectors. This simple modification yields speed estimation with smaller mean relative errors and mean squared errors compared with the estimation generated from conventional methods. For the test cases shown in the paper, the proposed method consistently outperforms the conventional method. The proposed method requires the raw data to be processed in a non-conventional way. The requirement, however, can easily be incorporated into the microprocessors in the controller box for processing loop detector data.

In the paper, we use a constant $g(=18)$ as a conversion factor to represent the median vehicle length. The results from our preliminary studies (not shown in this paper) indicate that the choice of g might be related to the detector spacing, which could lead to a bias in the "ground truth" speed data from double loops, which subsequently, would affect the results displayed in Tables 1 and 2. Studies are ongoing to explore this issue. For future research, we will explore the potential use of g as a function of occupancy as suggested in [3]. We will also explore the use of variable sampling length as suggested in [4]. Though the proposed method outperforms the conventional method, its performance seems to vary quite substantially with different data sets obtained at different locations. Studies are ongoing to investigate this large variation. We also need to explore further the situation in which the use of median information to represent the prevailing traffic is not satisfactory, i.e., the situation in which $(l/v)_M$ is not separable.

Acknowledgment

The research from this paper was partially supported under a contract with the PATH program, University of California at Berkeley.

References

[1] Koshi, M., M. Iwaski, and I. Ohkura. (1983) Some findings and an overview on vehicular flow characteristics. In Proceeding of 8th International Symposium on Transportation and Traffic Theory. Pp. 8 – 20.

[2] Banks, J. H. (1999) Investigation of some characteristics of congested traffic flow, In *Transportation Research Record*. no.1678, pp.128-134. National Research Council.

[3] Hall, F. L. and B. N. Persaud (1989) Evaluation of speed estimates made with single-detector data from freeway traffic management systems, IN: *Transportation Research Record*. no. 1232, pp.9 – 16. National Research Council.

[4] Cassidy, M. and B. Coifman (1997). Relation among average speed, flow, and density and analogous relation between density and occupancy, IN: *Transportation Research Record*. no. 1591, pp. 1-6, National Research Council.

[5] Edie, L. C. (1963) Discussion of traffic stream measurements and definitions. IN: Proceedings of the Second International Symposium on the Theory of Traffic Flow, pp. 139 – 154. Paris.