VEHICLE LENGTH MEASUREMENT AND LENGTH-BASED VEHICLE CLASSIFICATION IN CONGESTED FREEWAY TRAFFIC

Lan Wu

Graduate Research Assistant
The Ohio State University
Department of Civil, Environmental, and Geodetic Engineering

Benjamin Coifman, PhD

Associate Professor
The Ohio State University
Joint appointment with the Department of Civil, Environmental, and Geodetic Engineering, and the Department of Electrical and Computer Engineering

Hitchcock Hall 470
2070 Neil Ave, Columbus, OH 43210
Phone: (614) 292-4282
E-mail: Coifman.1@OSU.edu

ABSTRACT

Classified vehicle counts are a critical measure for forecasting the health of the roadway infrastructure and for planning future improvements to the transportation network. Balancing the cost of data collection with the fidelity of the measurements, length-based vehicle classification is one of the most common techniques used to collect classified vehicle counts. Typically the length-based vehicle classification process uses a pair of detectors in a given lane to measure effective vehicle length. While the calculation is simple and seems well defined, this paper demonstrates that small changes in the calculations can lead to large differences in performance during challenging conditions. In particular, most conventional calculations assume that acceleration can be ignored, which simply is not the case in congested traffic. As a result of this fact, many operating agencies are reluctant to deploy classification stations on roadways where traffic is frequently congested. This paper examines six variations of the conventional vehicle length calculation and develops a seventh that also estimates constant acceleration. It then highlights two of these approaches that work well in extreme conditions on freeways for speeds down to 15 mph. This range should be sufficient for most applications. Then using empirically collected data we find that the extreme events were uncommon and even the conventional method did quite well in stop-and-go traffic since the slower traffic moves, the lower the flow during that period. In any event, the key to success is the use of well-tuned detectors.

Key words:
Loop detectors; Dual-loop detector; Speed and length measurement; Vehicle classification; Freeway traffic.

INTRODUCTION

Classified vehicle counts are a critical measure for forecasting the health of the roadway infrastructure and for planning future improvements to the transportation network, e.g., the USDOT mandates that all states collect these classification data. Balancing the cost of data collection with the fidelity of the measurements, length-based vehicle classification is one of the most common techniques used to collect classified vehicle counts. Typically the length-based vehicle classification process uses a dual-loop detector (consisting of a pair of loop detectors in a given lane) to measure each vehicle's traversal time and then converts it to speed by taking the quotient of traversal time with the known distance between the detection zones. The product of this speed measurement and the dwell-time over one or both loop detectors in the dual-loop is then used to calculate the effective vehicle length. While the calculation is simple and seems well defined, we demonstrate that small changes in the calculations can lead to large differences in performance during challenging conditions. In particular, most conventional calculations assume that acceleration can be ignored, which simply is not the case in congested traffic. Thus, many operating agencies are reluctant to deploy classification stations on freeways with recurring congestion. This paper examines six variations of the conventional vehicle length calculation and develops a seventh that also estimates constant acceleration.

There have been empirical studies that have looked at the accuracy of individual vehicle measurements, e.g., [1-3], but very few have contemplated the nuances of the calculations as we do herein, e.g., [4]. This approach also applies to other sensors that use a pair of detection zones to emulate dual-loop detector operation, e.g., Wavetronix SmartSensor. Furthermore, while the focus of the current work is on length-based classification, the same speed measurement techniques are commonly employed at axle-based classification stations and this work should apply there as well.

The remainder of this paper is as follows. The first section develops the six conventional and one new length measurement methods. The second section uses the equations of motion to evaluate the performance of the seven methods. The third section uses empirically collected vehicle trajectories and actual dual-loop detector measurements to evaluate the performance of the seven methods in stop-and-go traffic. While we sought to improve general performance with a new length measurement method, at least at the locations studied we find that the extreme events were uncommon and even the conventional method did quite well in stop-and-go traffic. One key to success is the use of well-tuned detectors and we discuss these practical considerations at the end of the paper. Finally, the paper closes with conclusions.

LENGTH MEASUREMENT METHODS

Conventional Vehicle Length Measurement- Zero Acceleration Method

The conventional length measurement is not precisely defined. This ambiguity arises from the redundancy of the dual-loop detector and from the fact that the measurements occur over space. Figure 1 shows the two loop detectors in a dual-loop detector, separated by spacing S (leading-edge to leading-edge). A passing vehicle first crosses detector #1, with the resulting pulse from t_1 to t_2 in the bivalent output, Figure 1b, and then detector #2, with the resulting pulse from t_3 to t_4. The dual-loop detector yields two separate measures of traversal time, one from the rising edges of the two pulses: $TT_r = t_3 - t_1$; and one from the falling edges: $TT_f = t_4 - t_2$, which in turn yield two separate measures of speed, $V_r = S/TT_r$ and $V_f = S/TT_f$. Similarly, there are two separate measures of dwell-time for the passing vehicle, $T_u = t_2 - t_1$

from detector #1 and $T_d = t_4 - t_3$ from detector #2. When the vehicle traverses the dual-loop detector at a constant speed, then $V_r = V_f$ and $T_u = T_d$ assuming no detector errors. Because the two loop detectors in a dual-loop are separated by S and the vehicle has an *effective length*, L_e, (the sum of the physical vehicle length and the detection zone size) these temporal measurements also have a spatial component. Any change in speed will impact the measurements, which will usually cause the redundant measurements to differ in value. At free speed the impacts from acceleration are negligible, but at lower speeds the difference can become large and if the vehicle stops over the dual-loop detector it is intractable.

Conventionally speed is measured assuming acceleration is zero. Formalizing this conventional method, CM, there are two possible length measurements, defined by Equation 1. Typically an operating agency would only use one of the variants, e.g., some Caltrans engineers have expressed a preference for 1a because they feel V_r is measured more accurately than V_f by a dual-loop detector.

$$L_{CMr} = V_r * T_u \quad (1a)$$

$$L_{CMf} = V_f * T_d \quad (1b)$$

The form of these equations is important. Figure 1 shows that V_r and T_u are measured roughly concurrently, and similarly V_f and T_d are measured roughly concurrently. So in general the paired speed and dwell-time should be impacted similarly (but not identically) by any change in speed while the vehicle traverses the dual-loop detector. Swapping the pairing yields the method we term CM-, as defined in Equation 2. While similar in form to Equation 1, now the time period where speed is measured does not overlap the period that the dwell-time measured and thus, eliminates any benefits of the overlapping measurement period.

$$L_{CM-r} = V_r * T_d \quad (2a)$$

$$L_{CM-f} = V_f * T_u \quad (2b)$$

Our group has previously calculated the arithmetic average of the two terms from Equation 1 to reduce measurement noise [5], Equation 3, which in the present work we term CM+.

$$L_{CM+} = \frac{V_r*T_u+V_f*T_d}{2} = Av(L_{CM}) \quad (3)$$

WSDOT (Washington State Department of Transportation) used the product of rising edge speed V_r, and the arithmetic average of T_u and T_d to measure the effective length [6]. After that, [7] extended the WSDOT algorithm to instead use the arithmetic average of V_r and V_f via Equation 4, which we term CMO.

$$L_{CMO} = \frac{V_r+V_f}{2} * \frac{T_u+T_d}{2} = Av(V) * Av(T) \quad (4)$$

For completeness we consider two different averages that we have not found in the literature: first, the product of the harmonic average speed and the arithmetic average dwell-time via Equation 5, which we term CMX. Second, we find the product of the harmonic average speed and harmonic average dwell-time via Equation 6, which we term CMY.

$$L_{CMX} = \frac{S}{TT_r+TT_f} * (T_u + T_d) = Hav(V) * Av(T) \quad (5)$$

$$L_{CMY} = \frac{2*S}{TT_r+TT_f} * \frac{2}{1/T_u+1/T_d} = Hav(V) * Hav(T) \quad (6)$$

New Method- Constant Acceleration Method

All six of the conventional method variants assume that acceleration can be ignored. As shown in the next section, if the vehicle's speed changes as it traverses the dual-loop detector, the performance of these six methods varies greatly. To lessen these impacts, the new method, NM, uses the equations of motion to calculate speed, V_0, length, L_{NM}, and constant acceleration, a, using the four *transition-times*, t_1 to t_4, from Figure 1b assuming that acceleration is constant as the vehicle traverses the dual-loop detector, yielding

Equations 7-9 (obviously, at low speeds even the assumption of constant acceleration can break down).

$$S = V_0 * TT_f + \frac{1}{2}a * TT_f^2 \quad (7)$$

$$L_{NM} = V_0 * T_u + \frac{1}{2}a * T_u^2 \quad (8)$$

$$S + L_{NM} = V_0 * (T_u + TT_f) + \frac{1}{2}a * (T_u + TT_f)^2 \quad (9)$$

Next, Equations 10-12 solve for the three unknowns in terms of the *interval-times*, T_u, T_d, TT_r, TT_f, from Figure 1.

$$a = \frac{2(V_r - V_f)}{(T_u + T_d)} \quad (10)$$

$$V_0 = \frac{(T_u + T_d)S - (V_r - V_f)TT_r^2}{(T_u + T_d)TT_r} \quad (11)$$

$$L_{NM} = S * \frac{(TTr + TTf)Td * Tu}{(Tu + Td) * TTr * TTf} = \frac{V_r + V_f}{2} * \frac{2}{1/T_u + 1/T_d} = Av(V) * Hav(T) \quad (12)$$

The constant acceleration assumption was also used in [8-9], yielding a measurement equivalent to Equation 12. Although length is equivalent, [8-9] measure speed differently and do not measure a. It appears that this constant acceleration idea has largely been forgotten, though it occasionally reappears in the literature, e.g., [10].

PERFORMANCE EVALUATION UNDER DIFFERENT VEHICLE-MOTIONS

This section models vehicle-motion using four different models: constant speed, constant acceleration, non-constant acceleration, and stop model, as discussed below. Using a given vehicle-motion model, we then synthesize the resulting dual-loop detector interval-times as a function of a passing vehicle's true effective length, L_e, initial speed, V_0, and acceleration, a (if any). These detector measurements are used to calculate L using each of the seven length measurement methods. The L measurements are compared to L_e to find the resulting error. To avoid confounding factors the detector measurements are synthesized without any measurement errors (this assumption is eventually relaxed in the evaluation section). So the results in this section represent strictly the impacts of the assumptions in the given measurement method and the resulting biases from the implementation.

Constant speed model

The simplest vehicle-motion is the constant speed model, whereby the vehicle passes a dual-loop detector with a constant speed. In this case, there is no acceleration while traversing the dual-loop detector, thus $V_r = V_f$ and $T_u = T_d$. In this case all seven of the length measurement methods will yield the same L for a given vehicle, without any errors.

Constant acceleration model

Of course vehicles do not travel at constant speeds for their entire trip. The impacts of a given acceleration rate increase as V_0 decreases since a lower speed means a vehicle will be over a dual-loop detector for a longer time. The simplest vehicle-motion with acceleration is one of constant acceleration, with no stops over the dual-loop detector. Equations 13-16 show the resulting interval-times.

$$TT_r = \frac{-V_0 + \sqrt{V_0^2 + 2a*S}}{a} \quad (13)$$

$$TT_f = \frac{\sqrt{V_0^2 + 2a*(L_e + S)} - \sqrt{V_0^2 + 2a*L_e}}{a} \quad (14)$$

$$T_u = \frac{-V_0 + \sqrt{V_0^2 + 2a*L_e}}{a} \quad (15)$$

$$T_d = \frac{\sqrt{V_0^2 + 2a*(L_e + S)} - \sqrt{V_0^2 + 2a*S}}{a} \quad (16)$$

We synthesized a range of: a, V_0, L_e, found the interval-times and then used these to measure L from all seven methods. In this case the assumptions of the conventional methods differ from the actual vehicle-motion and as expected, length measurement errors occurred, while the NM implicitly assumes a constant acceleration. Indeed, in this case we find that NM is always the most accurate L, without any error. Of greater interest is the fact that the six conventional methods exhibit a range of performance, with some markedly better than others. Figure 2 shows the relative error in length via Equation 17 from all seven methods as a function of V_0 given four different values of a (one per row of subplots) and four different values of L_e (one per column of subplots). Both CM and CM- yield two different L per vehicle, to facilitate readability we only show one measurement from the given method (i.e., from Equations 1a and 2a).

$$\text{relative error in length} = \frac{L_{measured} - L_e}{L_e} * 100\% \quad (17)$$

As one might expect, across all of the conventional methods the magnitude of the errors generally increases with a, L_e, and the inverse of V_0. Note that in many of the subplots the curve for CM- and sometimes even CM fall completely beyond the 5% boundaries used in the given subplot. As noted above NM has no error given constant a. The next best method is CM+, the only conventional method that maintained an absolute error below 5% on all but one pairwise combination of L_e and a. For the shortest L_e the next three best methods are in this order: CMX, CMY, CMO. For the longer values of L_e the order changes to CMO, CMX, CMY. In any event CM- offers the worst performance throughout Figure 2, followed by CM. The various averaging methods perform better than CM because the impacts of acceleration on the different interval-times can cancel one another out, something that does not occur with CM and the situation is exasperated with CM- where the two interval-times do not overlap at all.

Non-constant acceleration model

The more time a vehicle spends over a dual-loop detector (both due to L_e and V_0) the greater the opportunity for the driver to change their acceleration rate. While the possibilities are literally infinite, we use a piecewise constant acceleration model to represent a general case of non-constant acceleration. The vehicle enters the dual-loop detector at V_0; accelerates at constant a_i for period t_i, reaching speed V_X, and then accelerates at constant a_j for period t_j.

Once more we synthesized a range of: a, V_0, L_e, found the interval-times and then used these to measure L from all seven methods. Under the non-constant acceleration model the assumptions of all seven of the length measurement methods differ from the actual vehicle-motion and the exhaustive set of all possible combinations of parameters becomes difficult to present. Figure 3 shows a representative example, presenting the absolute relative error via Equation 18 from all seven methods as a function of V_0 given four different sets of a_i and a_j (one set per row of subplots) and four different values of L_e (one per column of subplots). In Figure 3 the magnitude of a_i and a_j are equal, but they are opposite in sign. So here the vehicle accelerates for exactly half of the time it is over the dual-loop detector and then decelerates for the second half of the time. Within a given subplot V_X, t_i and t_j are constant for a given V_0, but these values vary with V_0 and they differ

from one subplot to the next due to the independent L_e, V_0, and $|a_i|$.

$$\text{absolute relative error in length} = \frac{|L_{measured} - L_e|}{L_e} * 100\% \quad (18)$$

In this case there is no single "best" method. There is a cluster of NM, CM+, CMX and CMY offering similar performance with the lowest errors. For the shortest L_e CMY exhibits the best performance, for longer values of L_e, NM tends to do slightly better than the rest, followed by CM+. After this cluster comes CM, then CMO, and finally CM- (though sometimes CMO pulls ahead of CM).

Stop model

In heavy congestion traffic often comes to a complete stop. This situation is the most challenging for dual-loop detector length measurement. Short vehicles can stop between the two loop detectors, in which case the low speed will be reflected in the speed measurements but not the dwell-times. On the other hand, long vehicles can stop over both loop detectors in the dual-loop, so the low speed will be reflected in the dwell-times but not the measured speed since V_r is measured strictly before the stop and V_f strictly after the stop. We developed the stop model to capture this situation, as follows. A given vehicle with L_e and V_0 will traverse a dual-loop detector until coming to a stop with a constant, negative acceleration rate a_i, for a period t_i, remain stopped for some time, Δt, and then depart with a constant, positive acceleration rate a_j, for a period t_j until completely past the dual-loop detector. Once more we synthesized a range of: a, V_0, L_e, found the interval-times and then used these to measure L from all seven methods. As with the non-constant acceleration model the assumptions of all seven of the length measurement methods differ from the actual vehicle-motion.

Recognizing the fact that any non-zero stop time can only degrade performance beyond the zero stop time case, we set $\Delta t = 0$ to present the best case scenario and thus, the stop model reduces to a special case of the non-constant acceleration model, with $V_x = 0$. Figure 4 shows a representative example, presenting the absolute relative error from all seven methods as a function of V_0 given four different sets of a_i and a_j (one set per row of subplots) and four different values of L_e (one per column of subplots) for the range in which a vehicle will stop over the dual-loop detector. Again the example shows the case when the magnitudes of a_i and a_j are equal, but opposite in sign.

In Figure 4 we see that all of the scenarios exhibit a peak error at a different value of V_0 (note that the horizontal scale changes from one subplot to the next to clearly show the given feasible stop range). The peak error corresponds to the case when a short vehicle stops completely between the two loop detectors in the dual-loop or a long vehicle stops over both of the loop detectors. The distance to a stop varies based on L_e, a_i, and V_0 as the vehicle first enters the dual-loop detector, so as a_i changes for a given V_0, it should not be surprising that the worst performance seemingly moves around different values of V_0. This shifting simply reflects the interrelationship of the various parameters. When the vehicle stops in the worst possible location, i.e., the middle of the dual-loop detector, all seven of the methods have errors in excess of 40% for all 16 scenarios presented. Away from this peak error, NM usually offers the best performance followed by CM+. The remaining five methods generally exhibit errors in excess of 10% throughout the range of V_0 associated with the stopping vehicle.

PERFORMANCE EVALUATION WITH EMPIRICAL DATA

The previous section relied strictly on assumed models of motion, now we consider performance from empirical data without any presumed motion. To this end, we employ the Next Generation Simulation (NGSIM) datasets. The NGSIM program was initiated by the Federal Highway Administration to collect

high-quality, empirical vehicle trajectory data to support the development of better traffic simulation [11]. To validate our work, first we use the I-80 dataset, which includes vehicle trajectories over approximately 1/3 mi of I-80 in Emeryville, California for 45 min during rush-hour on April 13, 2005, then we use the US-101 dataset, which includes vehicle trajectories over a similar distance of US-101 in Los Angeles, California for 45 min during rush-hour on June 15, 2005.

NGSIM synthesis and validation

The NGSIM data include instantaneous speed and location for every passing vehicle as well as the vehicle's physical length. We simulated a dual-loop detector in each lane, located a fixed distance past the entry point (800 ft for I-80 data and 1,000 ft for US-101). The detection zone size was set to 6 ft, with $S = 20$ ft. The synthesized transition-times t_1 and t_3 come directly from the NGSIM trajectory data as the vehicle passes the leading edge of each simulated loop detector. The trajectories are linearly interpolated to find the exact passage time because the raw NGSIM data are sampled at 10 Hz, which is too slow to calculate accurate vehicle lengths. Transition-times t_2 and t_4 come from a given vehicle's trajectory shifted upstream in space by the vehicle's length and the size of the detection zone. These four synthetic transition-times are then used to measure L from all seven methods, while L_e, comes from the recorded NGSIM vehicle length plus the size of the detection zone.

The seven length measurement methods are first evaluated using the entire NGSIM I-80 dataset, i.e., all of the available 45 min from each of 6 lanes. The arithmetic average, $\bar{v} = \frac{1}{2}(V_r + V_f)$, is used to sort the results into bins by 5 mph from 0 to 30 mph and by 10 mph when higher than 30 mph. The top part of Table 1 shows the total number of vehicles falling into each speed bin and then for each of the seven methods, reports the number of those vehicles that had an absolute relative length measurement error below 1%. Figure 5a shows these same results as a percent of the total number of vehicles in each speed bin. NM and CM+ show the best performance overall, with CM+ dropping below NM in the highest speed bin (a difference of 2/36 vehicles). Both CM+ and NM had over 90% of the length measurements within 1% absolute relative error in every speed bin except for the lowest speed bin. This process is repeated in the bottom part of Table 1 and Figure 5b tallying those vehicles with measurement error below 5%. The results are similar to the 1% threshold, though now NM and CM+ are above 85% for even the lowest speed bin and above 99% for all speed bins above 10 mph. Figure 5a-b also show that performance is already very good using CM for this dataset, suggesting that the extreme accelerations in the lower portions of Figures 2 and 3, as well as the worst case stop locations of Figure 4 are fairly uncommon.

The length measurement results are encouraging, but it is important to also consider the specific application: length-based vehicle classification. As noted in [4], "the [length based] classification scheme is tolerant to large length estimation errors provided the true length is far from the boundary between two classes." In fact [4] found a classification error rate of 1.6% during free flow at a good dual-loop detector station and showed that most of the classification errors arose due to small length measurement errors for vehicles with lengths close to the boundary between two classes. Using the three length classes from [4], with boundaries at 28 ft and 46 ft, Table 2 shows the classification results for the NGSIM I-80 data from NM, CM+ and for reference, CM. NM had a classification error rate of $11/5{,}675 = 0.19\%$, CM+ had an error rate of $10/5{,}675 = 0.18\%$, and CM had an error rate of $16/5{,}675 = 0.28\%$.

All of this analysis was repeated on US-101 with similar results, e.g., Figure 5c-d. Table 3 shows the classification performance on US-101. In this case NM had a classification error rate of $14/6{,}098 = 0.23\%$, CM+ had an error rate of $13/6{,}098 = 0.21\%$, and CM had an error rate of $19/6{,}098 = 0.31\%$.

The classification error rates observed herein during congestion are almost a full order of magnitude better than the empirical results from [4] collected under free flow conditions. This difference likely reflects

several factors. First, [4] used sampled data at 240 Hz while the interpolated NGSIM data are continuous time. The discretized empirical data will exhibit errors due to the sampling and the subsample impacts are more pronounced at higher speeds. Second, over 70% of the observed accelerations had a magnitude below 2 mphps, falling between the first two rows of Figures 2, 3, and 4. Third, although the NGSIM data include stopped vehicles that should exhibit large errors, their numbers are relatively small compared to the total flow. The simple fact that flow goes to zero as speed drops means that relatively few stopping vehicles are actually observed in the data. A total of $110/5{,}675 = 1.94\%$ of I-80 vehicles stopped over the dual-loop detectors, while only $10/6{,}098 = 0.16\%$ of US-101 vehicles stopped over the dual-loop detectors. For all speed bin below 30 mph the non-constant acceleration model proved to be the dominant vehicle-motion in the NGSIM data, while constant acceleration model dominated for 30-50 mph [12]. Finally fourth, the synthetic data do not have any detector errors. Even a healthy detector is likely to have the occasional measurement error when recording the transition-times.

BHL validation

It is uncommon that individual vehicle actuations are recorded and it is even less common to have ground truth vehicle length measurements. The previous section used empirical speeds and accelerations, but synthetic detector data. This section uses real dual-loop detector data. The NGSIM I-80 dataset was collected in the Berkeley Highway Lab (BHL) [15], and BHL dual-loop detector Station 8 was within the NGSIM field of view (see [13-14] for details). Station 8 had problems on the day of collection: it was off-line most of the time that the NGSIM data were collected and approximately 1% of the loop detector actuations were non-vehicle actuations due to splashover (the non-vehicle pulses were manually identified and excluded), there were still about 12 minutes of concurrent data that had actual dual-loop detector measurements with independent ground truth vehicle lengths from NGSIM.

The left-hand column of Figure 6 shows the scatter plot (all six lanes combined) of measured effective length from the dual-loop detectors versus the corresponding NGSIM reported physical length plus a 7 ft detection zone, with one subplot for each of the three methods: NM, CM+, and CM. The horizontal and vertical lines in the plots show the divisions between length classes. Although not shown, the average absolute length error from each of the three methods is 2.5 ft.

The left-hand side of Table 4 shows the performance of NM, CM+ and CM using the actual BHL dual-loop detector data. The results are very good considering the average speed was 21 mph across these vehicles (median speed was 19 mph). Consistent with the NGSIM validation above, CM did as well as the other two methods, which suggests the acceleration impacts were very small (the average absolute acceleration for these vehicles was 2 mphps). The classification error rate from the actual dual-loop detector data was 1.5%, which is comparable to the free flow results of 1.6% from [4].

For reference, the right-hand side of Table 4 and Figure 6 show the results for the same vehicles, except now the dual-loop detector data were synthesized strictly from the NGSIM trajectories in the same manner as Table 2. There are roughly 10 fewer errors in the synthetic data, presumably reflecting the impacts of using real loop detectors. The synthetic data error rate is 0.57%, and is similar to the NGSIM analysis presented above.

Practical considerations

In practice performance could be worse than the synthetic results presented in the NGSIM section due to sampling issues and detector errors. Typically dual-loop detectors are sampled anywhere from 60 Hz to more than 1 kHz, e.g., the dual-loop detectors in the BHL section are sampled at 60 Hz. At lower sampling frequencies short interval-time measurements will be very noisy due to the measurement granularity. For

example a dual-loop detector sampling at 60 Hz with $S = 20$ ft might only be able to resolve speeds to 5 mph in free flow traffic since the traversal time is so short. These sampling issues diminish greatly at lower speeds since the traversal and dwell-times become much longer.

To avoid having to address the impacts of detector errors, this paper has been very specific about the need to use well-tuned detectors. Dual-loop detectors are prone to errors including pulse-breakup, splashover, sensitivity, etc. [16-20]. These errors will degrade performance and the resulting length measurement errors would be much larger at any speed for all seven of the methodologies discussed herein. For example, the size of a loop's detection zone typically is unknown. Although the physical loops in the BHL are 6 ft across, the 7 ft detection zone size used above for BHL station 8 resulted in an unbiased length between NGSIM and BHL. Given the true 7 ft detection zone, had we guessed a 6 ft zone, analytically we would expect length measurement errors in excess of 14% regardless of the length measurement method used. In this case we were able to find the unbiased length via the NGSIM vehicle measurements, in practice, one would need some other means of measuring the vehicle's physical length or the detection zone, e.g., [21].

To achieve well-tuned detectors it is important for an operating agency to follow an established protocol for calibration and to quantify the reliability of the classification system. It is equally important to have an ongoing performance monitoring to ensure the detectors remain well-tuned. To this end, our group has produced suite of tools that can be run in real-time to verify that a given detector is well-tuned [16-19]. If a detector fails these tests then the corresponding data are of questionable quality and the detector is in need of re-tuning.

CONCLUSIONS

This paper examined length measurement for vehicle classification at dual-loop detectors on a freeway during congested conditions where speed is low enough that acceleration cannot be neglected. We consider six variations of the conventional length measurement method (CM•), all of which assume acceleration is zero. We developed a new method for measuring length (NM) that instead assumes acceleration is constant, but might be non-zero. We then evaluated all seven of the length measurement methods under different vehicle-motions (first using strictly defined motions, then empirically observed trajectories, and finally using actual dual-loop detector data).

The six CM variants exhibited markedly different performance under the strictly defined motions, with CM+ showing the best results among the six conventional methods. Meanwhile, NM was slightly better than CM+ across all seven methods, but only by the smallest of margins. All of the methods worked well given zero acceleration. Under constant acceleration, consider Figure 2 when $L_e = 50$ ft and $a = 3$ mphps, the absolute relative error from CM exceeds 5% for $V_0 < 23$ mph, illustrating why operating agencies are reluctant to use conventional methods in congestion. Recall that across all seven methods, the absolute error increases with L_e and the reciprocal of V_0, so in contrast to CM, the error from CM+ remains below 5% under more challenging conditions: down to $V_0 = 6$ mph with $L_e = 70$ ft. Suggesting that length measurements and classification could be extended to these lower speeds provided care is taken to ensure that the detectors are well-tuned. Considering typical accelerations, $|a| \leq 3$ mphps we see that NM, CM+, CMX and CMY in Figure 3 all have errors below 5% for speeds down to 20 mph for $L_e = 70$ ft and similar performance at lower speeds for shorter values of L_e. We suspect that it is possible to realize further gains by considering the differences between the various length measurement methods, e.g., comparing Equations 2a and 2b to determine if just one if them is impacted by acceleration.

All seven of the length measurement methods will yield poor performance when a vehicle stops over

the dual-loop detector. Recognizing the severity of the measurement error when a vehicle stops over the dual-loop detector, we have ongoing work to use the dual-loop detector measurements to identify all of the vehicles that may have stopped over a dual-loop detector, e.g., [13]. In the meantime, as preliminary guidance, assuming reasonable acceleration rates, those vehicles that stop should have a measured speed below 10 mph. Fortunately, the simple fact that flow goes to zero as speed drops to zero means that relatively few stopping vehicles are actually observed in the data. From the NGSIM I-80 data, out of 5,675 vehicles we found that all of the 110 stopped vehicles had a measured speed below 10 mph, while only 10.4% of the non-stop vehicles had a speed below 10 mph even though conditions were stop-and-go.

In the empirical validation CM did almost as well as CM+ and NM in congestion, suggesting that the regular method is already doing quite well for the evaluation datasets. The extreme accelerations and the worst-case stop locations from the vhicle-motion analysis were fairly uncommon in the empirical data. Furthermore, in stop-and-go traffic few vehicle measurements will be impacted since the slower traffic moves, the lower the flow during that period. The good performance is also due in part to the fact that, "the [length based] classification scheme is tolerant to large length estimation errors provided the true length is far from the boundary between two classes," [4]. Indeed, when using real loop detector data, Figure 6a-c show a lot of scatter from length measurement errors (some up to 20 ft), but only a few of these errors place a given vehicle in a different class. Whether these results are typical of other locations will require further data collection. In any event, the empirical results also underscore the importance of using well-tuned detectors; otherwise the length measurement errors would be much larger at any speed for all seven of the methodologies discussed herein.

ACKNOWLEDGEMENTS

This material is based upon work supported in part by NEXTRANS the USDOT Region V Regional University Transportation Center. The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration. This report does not constitute a standard, specification or regulation.

REFERENCES

[1] Banks, J. (2008). *Evaluation of Portable Automated Data Collection Technologies: Final Report*. California PATH Research Report, UCB-ITSPRR-2008-15.

[2] Minge, E. (2010). *Evaluation of Non-Intrusive Technologies for Traffic Detection*. Minnesota Department of Transportation, Final Report #2010-36.

[3] Kim, S., Coifman, B. (2013). Evaluation of Axle-Based and Length-Based Vehicle Classification Stations, *Transportation Research Record 2339*, 2013, pp 1-12.

[4] Coifman, B., Kim, S. (2009). Speed Estimation and Length Based Vehicle Classification from Freeway Single Loop Detectors. *Transportation Research Part-C*, 17(4), pp 349-364.

[5] Coifman, B., Cassidy, M. (2002). Vehicle Reidentification and Travel Time Measurement on Congested Freeways. *Transportation Research- Part A*, 36(10), 2002, pp. 899–917. Washington DC.

[6] Cheevarunothai, P., Wang, Y., Nihan, N.L. (2006). Identification and Correction of Dual-Loop Sensitivity Problems. *Transportation Research Record 1945*, pp. 73 - 81.

[7] Zhang, X., Wang, Y., Nihan, N.L. (2006). Robust algorithm for improved dual-loop detection on freeways. *Proc. of the 85th Transportation Research Board Annual Meeting*, Washington DC.

[8] Gazis, D., Foote, R. (1969). Surveillance and Control of Tunnel Traffic by an On-Line Digital Computer, *Transportation Science*, Vol 3, No 3, pp 256-275.

[9] Gazis, D., Foote, R. (1970). Calibration and Correction of Magnetic Loop Detectors, *Highway Research Record 325*, pp 30-35.

[10] Huft, D. (1986). The South Dakota Bridge Weigh-In-Motion System, *Transportation Research Record 1060*, pp 111-120.

[11] Kovvali, V., Alexiadis, V., Zhang, L. (2007). "Video-Based Vehicle Trajectory Data Collection," *Proc. of the 86th Annual TRB Meeting*, TRB.

[12] Wu, L., (2013). *Improved Vehicle Length Measurement and Classification From Freeway Dual-Loop Detectors in Congested Traffic*, Masters Thesis, The Ohio State University.

[13] Wu, L., Coifman, B., (in review) "Improved Vehicle Classification from Dual-Loop Detectors in Congested Traffic"

[14] Lee, H., Coifman, B., (in review) "Using LIDAR to Validate the Performance of Vehicle Classification Stations"

[15] Coifman, B., Lyddy, D., Skabardonis, A. (2000). "The Berkeley Highway Laboratory- Building on the I-880 Field Experiment", *Proc. IEEE ITS Council Annual Meeting*, pp 5-10.

[16] Lee, H., Coifman, B. (2011). Identifying and Correcting Pulse-Breakup Errors from Freeway Loop Detectors, *Transportation Research Record 2256*, pp 68-78.

[17] Lee, H., Coifman, B. (2012). Quantifying Loop Detector Sensitivity and Correcting Detection Problems on Freeways, *ASCE Journal of Transportation Engineering*, Vol. 138, No. 7, pp 871-881.

[18] Lee, H., Coifman, B. (2012). An Algorithm to Identify Chronic Splashover Errors at Freeway Loop Detectors, *Transportation Research-Part C*. Vol 24, pp 141-156.

[19] Coifman, B. (1999). Using Dual Loop Speed Traps to Identify Detector Errors, *Transportation Research Record 1683*, Transportation Research Board, pp 47-58.

[20] Coifman, B. (2006). Vehicle Level Evaluation of Loop Detectors and the Remote Traffic Microwave Sensor, *ASCE Journal of Transportation Engineering*. Vol 132, No 3, pp 213-226.

[21] Lee, H., Coifman, B. (2012). "Side-Fire LIDAR Based Vehicle Classification," *Transportation Research Record 2308*, pp 173–183.

LIST OF TABLES

LIST OF FIGURES

Table 1, Number of vehicles in each measured average speed bin by method and across the entire NGSIM I-80 population with absolute relative error in length ≤1%, ≤5%, and correctly classified.

$\bar{v}$(mph) / #Vehicle		0~5	5~10	10~15	15~20	20~25	25~30	30~40	40~50	>50	sum
Total		176	639	1,403	1,423	887	449	473	189	36	5,675
Absolute Relative Error in Length < 1%	CM	69	416	1,162	1,238	808	405	421	171	33	4,723
	CM-	7	50	317	370	383	202	197	87	21	1,634
	CM+	88	489	1,284	1,352	868	439	459	185	34	5,198
	CMO	18	272	1,040	1,209	793	398	416	171	31	4,348
	CMX	78	451	1,238	1,309	840	425	446	180	33	5,000
	CMY	31	296	1,100	1,238	830	416	436	177	33	4,557
	NM	89	491	1,290	1,356	868	440	460	185	36	5,215
Absolute Relative Error in Length < 5%	CM	136	568	1,313	1,340	839	424	449	179	35	5,283
	CM-	15	133	748	828	671	359	387	166	31	3,338
	CM+	154	617	1,392	1,417	887	449	473	189	36	5,614
	CMO	56	478	1,264	1,302	817	411	433	175	32	4,968
	CMX	146	589	1,351	1,376	861	436	459	184	35	5,437
	CMY	75	509	1,331	1,372	861	431	456	184	33	5,252
	NM	157	616	1,392	1,418	887	449	473	189	36	5,617
Correct Class	CM	174	632	1,399	1,420	887	449	473	189	36	5,659
	CM+	174	637	1,400	1,421	887	448	473	189	36	5,665
	NM	174	636	1,400	1,421	887	448	473	189	36	5,664

Table 2, Comparison of three methods: NM, CM+ and CM for vehicle classification accuracy verification using the NGSIM I-80 dataset. Where, 1, 2, 3 denote the vehicle length class: Class 1: $L_e \leq 28$ ft, Class 2: 28 ft $< L_e \leq 46$ ft, Class 3: $L_e > 46$ ft.

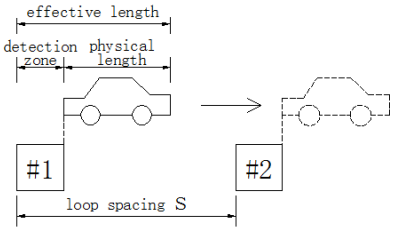
NM		True			Error Ratio
		1	2	3	
Measured	1	5,461	3	0	0.05%
	2	5	73	0	6.41%
	3	0	3	130	2.26%
Error Ratio		0.09%	7.59%	0.00%	5,664
CM+		True			Error Ratio
		1	2	3	
Measured	1	5,461	3	0	0.05%
	2	5	74	0	6.33%
	3	0	2	130	1.52%
Error Ratio		0.09%	6.33%	0.00%	5,665
CM		True			Error Ratio
		1	2	3	
Measured	1	5,459	4	0	0.07%
	2	6	73	3	10.98%
	3	1	2	127	2.31%
Error Ratio		0.13%	7.59%	2.31%	5,659

Table 3, Comparison of three methods: NM, CM+ and CM for vehicle classification accuracy verification using the NGSIM US-101 dataset. Where, 1, 2, 3 denote the vehicle length class: Class 1: $L_e \leq 28$ ft, Class 2: 28 ft $< L_e \leq 46$ ft, Class 3: $L_e > 46$ ft.

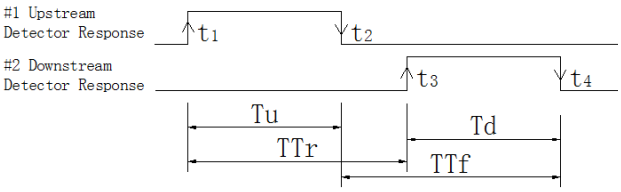
NM		True			Error Ratio
		1	2	3	
Measured	1	5,952	0	0	0.00%
	2	10	92	2	11.54%
	3	1	1	40	4.76%
Error Ratio		0.18%	1.08%	4.76%	6,084
CM+		True			Error Ratio
		1	2	3	
Measured	1	5,953	0	0	0.00%
	2	9	92	2	10.68%
	3	1	1	40	4.76%
Error Ratio		0.17%	1.08%	4.76%	6,085
CM		True			Error Ratio
		1	2	3	
Measured	1	5,951	1	0	0.02%
	2	11	89	3	13.59%
	3	1	3	39	9.30%
Error Ratio		0.20%	4.30%	7.14%	6,079

Table 4, Comparison of three methods: NM, CM+ and CM for vehicle classification accuracy verification using BHL dual-loop data (left) and for the exact same vehicles the synthetic dual-loop data from NGSIM (right).

		Actual BHL Data				Synthetic NGSIM Data			
NM		True			Error Ratio	True			Error Ratio
		1	2	3		1	2	3	
Measured	1	986	4	1	0.50%	991	0	0	0.00%
	2	8	17	0	32.00%	3	22	1	15.38%
	3	0	3	32	8.57%	0	2	32	5.88%
Error Ratio		0.80%	29.17%	3.03%	1,035	0.10%	0.30%	8.33%	1,045
CM+		True			Error Ratio	True			Error Ratio
		1	2	3		1	2	3	
Measured	1	986	4	1	0.50%	991	0	0	0.00%
	2	8	17	0	32.00%	3	22	1	15.38%
	3	0	3	32	8.57%	0	2	32	5.88%
Error Ratio		0.80%	29.17%	3.03%	1,035	0.10%	0.30%	8.33%	1,045
CM		True			Error Ratio	True			Error Ratio
		1	2	3		1	2	3	
Measured	1	987	4	1	0.50%	990	0	0	0.00%
	2	7	18	0	28.00%	4	22	1	18.52%
	3	0	2	32	5.88%	0	2	32	5.88%
Error Ratio		0.70%	25.00%	3.03%	1,037	0.10%	0.40%	8.33%	1,044



(a)



(b)

Figure 1, (a) Schematic of a vehicle passing over the two loop detectors in a dual-loop detector, (b) the time series response of the two loop detectors and the resulting measurements used to calculate speed and length.

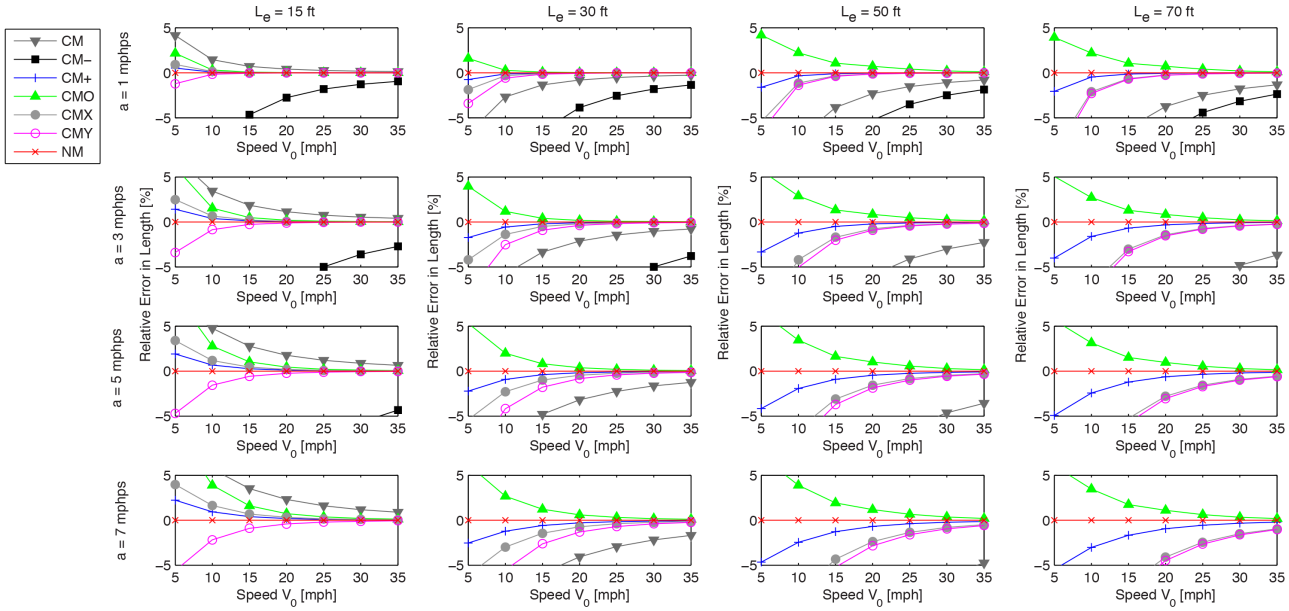


Figure 2, Family diagram of relative error in length, L_e, versus initial speed, V_0, for the seven methods under the constant acceleration model with a>0. L_e is constant in a given column of subplots, but increases from left to right, while V_0 is constant in a given row of subplots, increasing from top to bottom.

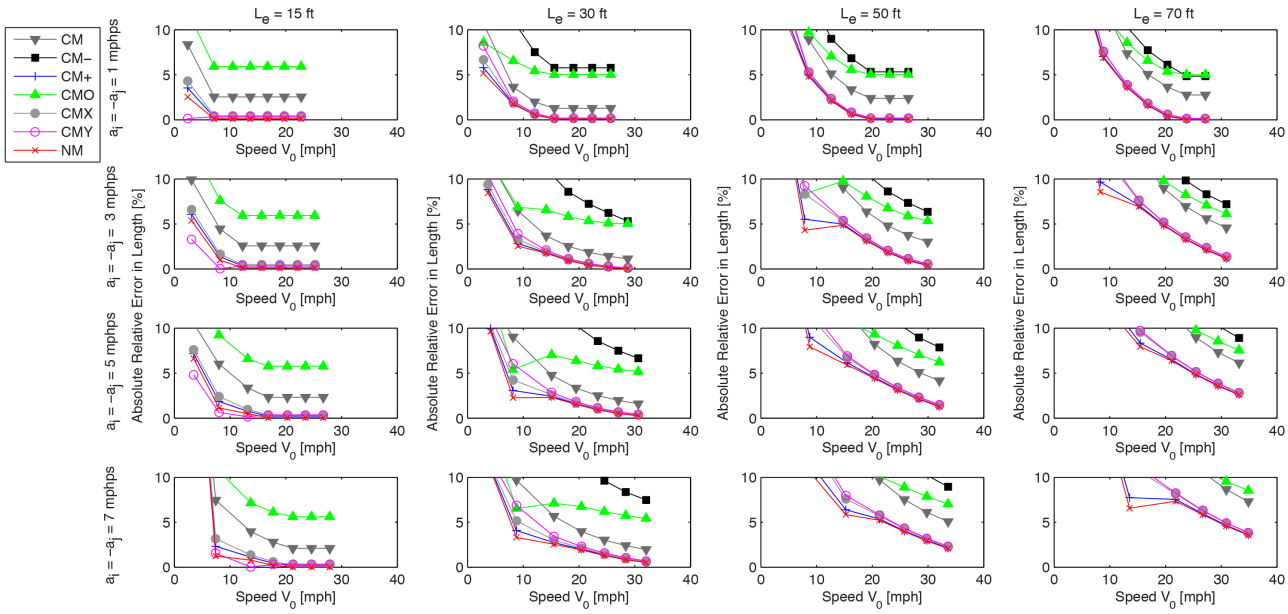


Figure 3, Family diagram of absolute relative error in length, L_e, versus initial speed, V_0, for the seven methods under the non-constant acceleration model with $a_i>0$, and $a_j<0$. L_e is constant in a given column of subplots, but increases from left to right, while V_0 is constant in a given row of subplots, increasing from top to bottom.

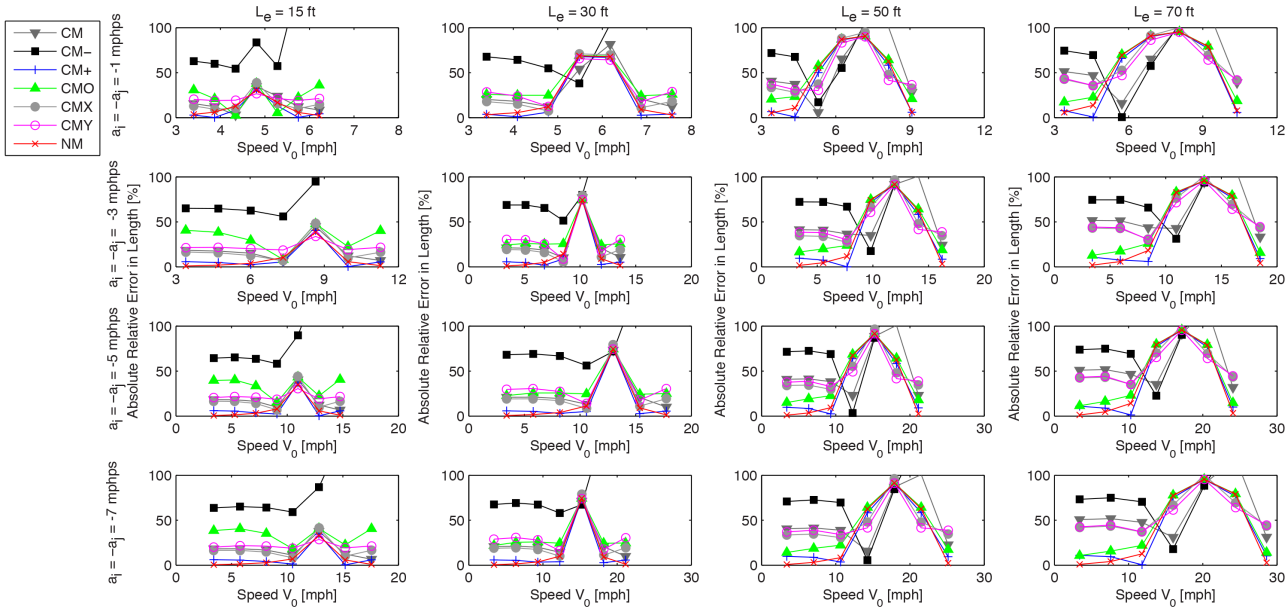


Figure 4, Family diagram of absolute relative error in length, L_e, versus initial speed, V_0, for the seven methods under the stop model with $a_i<0$, $\Delta t=0$, and $a_j>0$. L_e is constant in a given column of subplots, but increases from left to right, while V_0 is constant in a given row of subplots, increasing from top to bottom. Note that the horizontal scale changes from one subplot to the next to clearly show the given feasible stop range.

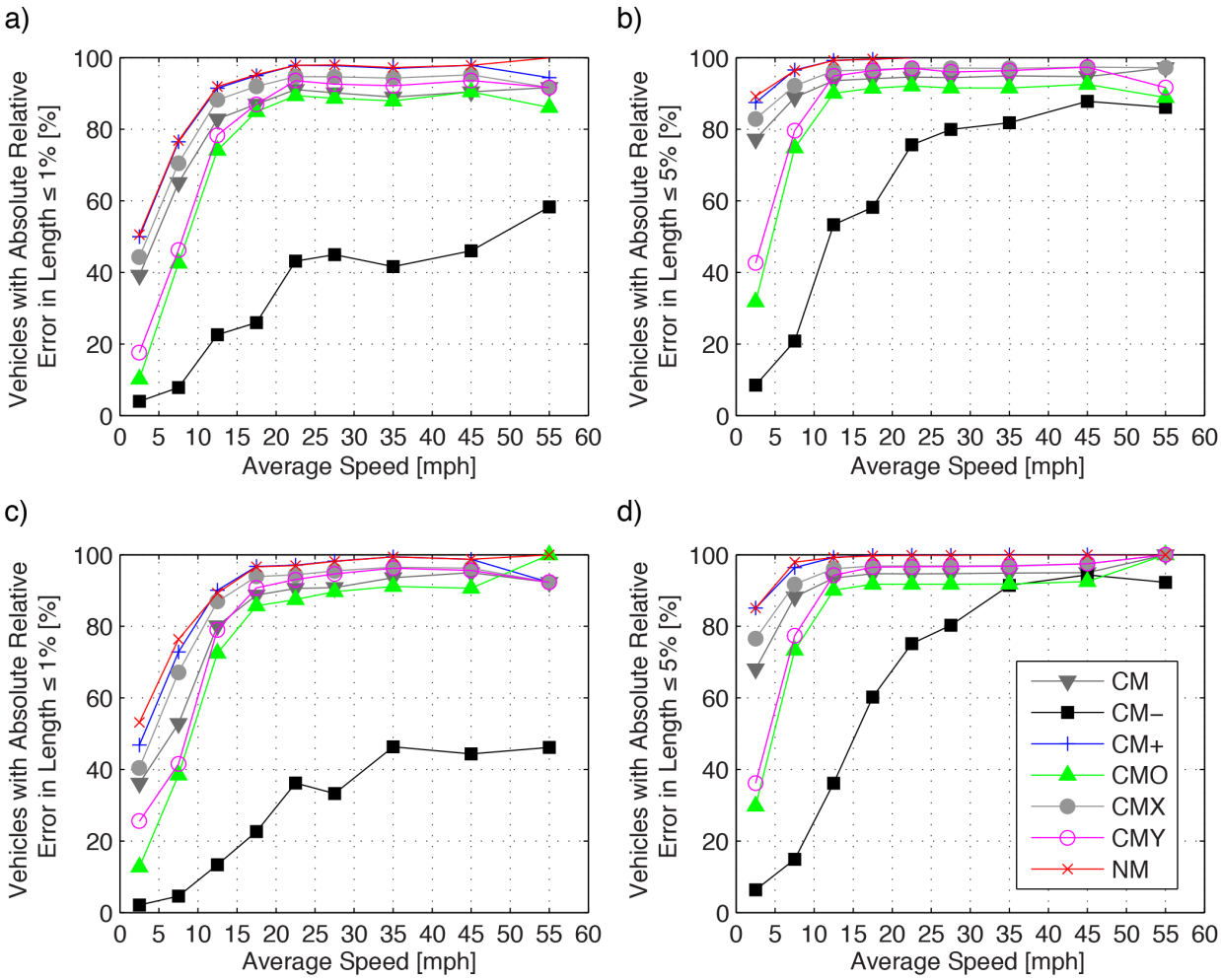


Figure 5, Percentage of vehicles in each measured average speed bin by method with absolute relative error in length for I-80 (a) ≤1%, (b) ≤5%. Repeating for US-101 (c) ≤1%, (d) ≤5%.

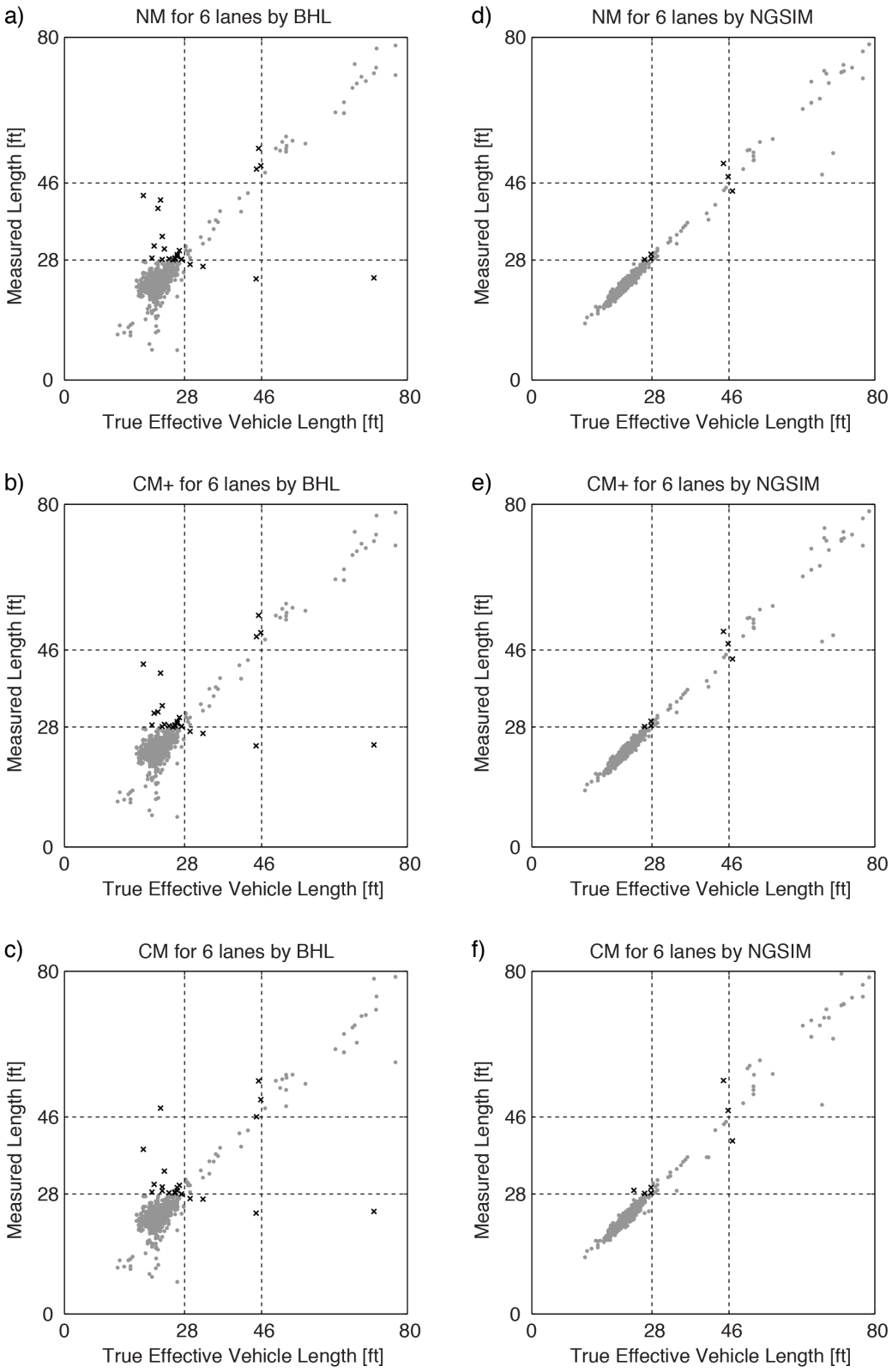


Figure 6, Measured effective length versus true effective vehicle length from (a) NM on the BHL dual-loop detector data, (b) CM+ on the BHL dual-loop detector data, (c) CM on the BHL dual-loop detector data, (d) NM on the NGSIM synthetic dual-loop detector data, (e) CM+ on the NGSIM synthetic dual-loop detector data, and (f) CM on the NGSIM synthetic dual-loop detector data. All six subplots use the same exact set of passing vehicles.