Speed Estimation and Length Based Vehicle Classification from Freeway Single Loop Detectors

Benjamin Coifman, PhD
Associate Professor
The Ohio State University
Joint appointment with the Department of Civil and Environmental Engineering and Geodetic Science, and the Department of Electrical and Computer Engineering
Hitchcock Hall 470
2070 Neil Ave, Columbus, OH 43210
Tel: (614) 292-4282
Fax: (614) 292-3780
E-mail: Coifman.1@OSU.edu

SeoungBum Kim
Graduate Research Assistant
The Ohio State University
Department of Civil and Environmental Engineering and Geodetic Science,
Hitchcock Hall 470
2070 Neil Ave, Columbus, OH 43210
E-mail: kim.1936@OSU.edu

1 ABSTRACT

Roadway usage, particularly by large vehicles, is one of the fundamental factors determining the lifespan of highway infrastructure. Each state typically has several dozen Weigh in Motion stations to monitor large vehicle usage. These stations are expensive to install and maintain, so they are usually supplemented with many more vehicle classification stations. Some of the classification stations employ axle counters, but the simplest of these stations use dual-loop detectors to measure vehicle length from the product of measured speed and detector on-time, and classify vehicles based on this measurement.

Meanwhile, single-loop detectors are the most common vehicle detector in use to monitor traffic, both for real-time operations and for collecting census data such as Annual Average Daily Travel (AADT). New, out-of-pavement detectors seek to replace loop detectors using wayside mounted sensors, but most of these detectors emulate the operation of single-loop detectors. In either case, collecting reliable length data from these detectors has been considered impossible due to the noisy speed estimates provided by conventional data aggregation at single-loop detectors.

This research refines unconventional techniques for estimating speed at a single-loop detector, yielding estimates that approach the accuracy of a dual-loop detector's measurements. Employing these speed estimation advances, this research brings length based vehicle classification to single-loop detectors, (and by extension, many of the emerging out-of-pavement detectors). The research promises to extend vehicle classification to single-loop detector count stations and the many single-loop detector stations already deployed for real-time traffic management. The work also offers a viable treatment in the event that one of the loops in a dual-loop detector classification station fails.

The classification methodology is evaluated against concurrent measurements from video and dual-loop detectors. To capture higher truck volumes than empirically observed, a process of generating synthetic on-times is developed.

2 INTRODUCTION

Roadway usage, particularly by large vehicles, is one of the fundamental factors determining the lifespan of highway infrastructure. The importance of road usage is evidenced by the federally mandated Highway Performance Monitoring System (HPMS) and the significance of large vehicles is reflected in the Weigh in Motion (WIM) data collected for the Long Term Pavement Performance (LTPP) program. Interest in the movement of these large vehicles has also increased from the transportation planning perspective, as freight shipments are becoming more common in the planning process.

Each state typically has several dozen WIM stations to monitor large vehicle usage. These stations are expensive to install and maintain, so they are usually supplemented with many more vehicle classification stations. Some of the classification stations employ axle counters, but the simplest of these stations use dual-loop detectors to measure vehicle length from the product of measured speed and detector on-time, and classify vehicles based on this measurement.

Meanwhile, single-loop detectors are the most common vehicle detector in use to monitor traffic, both for real-time operations and for collecting census data such as Annual Average Daily Travel (AADT). New, out-of-pavement detectors seek to replace loop detectors using wayside mounted sensors, e.g., the Remote Traffic Microwave Sensor (RTMS), but most of these detectors emulate the operation of single-loop detectors. In either case, collecting reliable length data from these detectors has been considered impossible due to the noisy speed estimates provided by conventional data aggregation at single-loop detectors.

This research refines unconventional techniques for estimating speed at a single-loop detector, yielding estimates that approach the accuracy of a dual-loop detector's measurements. Employing these speed estimation advances, this research brings length based vehicle classification to single-loop detectors, (and by extension, many of the emerging out-of-pavement detectors). The research promises to extend vehicle classification to single-loop detector count stations and the many single-loop detector stations already deployed for real-time traffic management. The work also offers a viable treatment in the event that one of the loops in a dual-loop detector classification station fails.

After reviewing the related literature, this work presents the new speed estimation techniques. Vehicle length is then estimated from the product of speed and on-time. To capture higher truck volumes than empirically observed, a process of generating synthetic on-times is developed. Following the Ohio Department of Transportation (ODOT) length based classification scheme for dual-loop detectors, the lengths are used to classify vehicles into three bins with divisions at effective vehicle lengths of 28 ft and 46 ft. This classification is evaluated against concurrent measurements from video and dual-loop detectors.

2.1 *Estimating Speed and Lengths*

This research seeks to mainstream advances in speed and length estimation from single-loop detectors and develop a vehicle classification methodology for these detectors. Benekohal and Girianna (2003) note that it is, "necessary to encourage state DOTs to include classification counts in their annual traffic-monitoring program." As noted in a draft research statement from the TRB Committee on Highway Traffic Monitoring, "Classification based solely on vehicle length is an alternative to axle-based classification but there has been no systematic study of how well it works -- or how it should work." The present research had to address many of these issues in the course of verifying the performance of single-loop detector based classification. There has been considerable research on vehicle classification leading to the conventional technologies as well as on-going work in emerging technologies. Needless to say, the body of work is broad. Limiting the scope to loop detector based speed estimation, length estimation, and vehicle classification, this section reviews the related literature.

For length-based classification from loop detectors, there are three interrelated parameters that can be measured or estimated for each passing vehicle, namely effective vehicle length (l), speed (v) and the amount of time the detector is "on", i.e., the *on-time* (on). These parameters are related by the following equation,

$$l = v \cdot on \tag{1}$$

At a single-loop detector, only the on-time can be measured directly, while a dual-loop detector can measure the speed from the quotient of the detector spacing and the difference in actuation times at the two loops. Given two of the three parameters, obviously the third is defined by Equation 1. In the absence of accurate speed estimation from single-loops, these detectors generally have not been used to estimate vehicle length or classify vehicles.

As a precursor, many researchers have sought better estimates of speed from single-loop detectors. The preceding research has emphasized techniques that use many samples of aggregate flow (q) and occupancy (occ) to reduce the estimation error, e.g., Mikhalkin et al (1972), Pushkar et al (1994), Dailey (1999), Wang and Nihan (2000), Coifman (2001). Although rarely noted, these techniques effectively seek to reduce the bias due to long vehicles in measured occupancy. Rather than manipulating aggregate data, we developed new aggregation methods to reduce the estimation errors.

Provided that vehicle lengths and vehicle speeds are uncorrelated, (see, e.g., Coifman, 2001), following conventional practice, speed (mean v) and assumed mean vehicle length (L^A) for a given sample are related by:

$$mean\ v \approx \frac{q \cdot L^A}{occ} \tag{2}$$

This equation is an extension of Equation 1, since,

$$\frac{q \cdot L^A}{occ} = \frac{L^A}{mean(on)} \tag{3}$$

and as with Equation 1, average length and average speed cannot be measured independently at a single-loop detector. Typically, an operating agency will set L^A to a constant value and use Equation 2 to estimate speed from single-loop measurements. But this approach fails to account for the fact that the percentage of long vehicles may change during the day or the simple fact that a sample may not include "typical" vehicle lengths. Particularly during low flow, when the number of vehicles in a sample is small, a long vehicle can skew occupancy simply because it takes more time for that vehicle to pass the detector. For example, at one detector station Coifman (2001) found that approximately 85 percent of the individual vehicle lengths observed were between 15 and 22 feet, but some vehicles were as long as 85 feet, or roughly four times the median length. This large range of average vehicle lengths arises due to the small number of vehicles with lengths far from the center of the skewed distribution. The median of a sample is much less sensitive to these outliers, and

$$median\ v \approx \frac{L^A}{median(on)} \tag{4}$$

provides an alternative estimate of speed. As shown in Coifman et al (2003), Equation 4 performs significantly better than Equation 2, and in fact it approaches the accuracy of dual-loop detector measurements for that study's data.

There have also been several efforts based on time-series trends in flow and occupancy to estimate the percentage of vehicles passing a single-loop detector that are long. Kwon et al (2003), developed a method employing aggregate flow and occupancy from single-loop detectors to estimate the percentage of long vehicles that passed. The work depends on two fundamental assumptions: the presence of a truck-free lane, and that the detector station exhibits high lane-to-lane speed correlation. They employed conventional detectors, used many days, using several stations from three facilities. The work only validated the results against aggregate dual-loop measurements and WIM data. The former yielded good results, while the latter had 20 percent overestimation, highlighting the importance of employing a truly independent measure of ground truth. The research studied facilities with low to moderate truck volumes (under 10 percent of the fleet) and did not explicitly single out performance in congested conditions. In fact they note that, "the estimate of truck volume is biased and unstable at the start of the congestion period."

Wang and Nihan (2003, 2004) also developed a method employing aggregate flow and occupancy from single-loop detectors to estimate the percentage of long vehicles that passed. Like Kwon et al, their work also depends on two fundamental assumptions, though slightly different, "constant average speed for each [three minute long] time period and at least two intervals containing only [short vehicles] in each period." They employed conventional detectors, used many days, using four detector stations. Also like Kwon et al, the work only validated the results against aggregate dual-loop measurements. The research studied facilities with low to moderate truck volumes (under 10 percent of the fleet) and did not explicitly single out performance in congested conditions. These authors note, "the algorithm should work better under less congested conditions." The authors also explicitly note the limitation of the small number of test sites, stating that, "future research is needed to handle the conditions when one or both of the assumptions are violated in order to reduce estimation errors.... The proposed algorithm will be more robust and accurate when the violation circumstances are properly addressed." More recently, this group has revised their methodology (Zhang et al, 2006). This recent study is subject to many of the same limitations as their earlier work, it employs aggregate flow and occupancy, was tested at only two detector stations (with approximately 10 percent truck flows), and only compared the results against aggregate dual-loop detector measurements. The final conclusion of Zhang et al states that although the method produced favorable bin volumes, further improvements to its performance are possible through optimizing its design and training, especially under heavily congested conditions.

There have also been efforts to use new loop detector sensors to measure the inductive vehicle signature for vehicle classification, e.g., Reijmers (1979) and Gajda et al (2001). While these inductive signature based efforts are promising, the published studies typically employ validation sets on the order of 100 vehicles. Conventional binary loop detector output remains by far the dominant configuration for single-loop detectors.

As noted earlier, most of the non-invasive vehicle detectors that have entered conventional practice mimic the operation of single-loop detectors, the two most prevalent examples of these detectors being the SmartSensor by Wavetronix and RTMS by EIS. Both sensors can provide length based classification data, though the specific algorithms are proprietary. While the sensors often provide reasonable counts and speed estimates in aggregate data, per-vehicle analysis has shown that the aggregate data allow over-counting errors to cancel under-counting errors and that individual vehicle on-times can be subject to large errors (see, e.g., Zwahlen et al, 2005; Coifman, 2006a). The literature is surprisingly lacking in terms of evaluating the classification performance from these sensors. Among the few available studies, Zwahlen et al, (2005) evaluated the SmartSensor in uncongested, low volume traffic, with low truck flows. While these conditions should lead to favorable performance by the sensor, after comparing the classification results against manually generated ground truth data the authors concluded that, "vehicle classification is unreliable; the fraction of trucks in a lane can be severely overestimated or underestimated." Trucks were undercounted by as much as 80 percent in the worst case and "at this time, the system does not reliably estimate the number of trucks in the traffic stream." French and French (2006) examined the performance of RTMS and SmartSensor, including vehicle classification, at four temporary locations and three fixed locations. Even though manufacturer representatives calibrated the detectors the reported truck counts from the non-invasive detectors were typically off by a factor of two and sometimes as much as ten. Almost all of the test locations were characterized by low truck flows, below 5 percent of the traffic. So while the manufacturers offer vehicle classification from these non-invasive sensors, the specific algorithms are undocumented and to the extent that they have been evaluated in the literature, the performance is poor.

Returning the focus to conventional single-loop detectors, the present research seeks to estimate vehicle lengths and classify vehicles. Assuming the loop detector is functioning properly, Equation 1 shows that a given on-time measurement is simply a function of the vehicle's length and speed. During free flow conditions the vehicle speeds typically fall in a small range and during congested conditions the difference between two successive vehicles' speeds is usually small. If one assumes that all of the vehicles in a sample are traveling near the median speed, one can use Equation 4 in conjunction with measured on-times to estimate individual vehicle lengths. Of course the number of vehicles per sample must be small enough for the speed assumption to hold and one must control for low speed conditions, when acceleration becomes non-negligible within the sample. Using samples of ten consecutive vehicles and restricting the analysis to samples with $v>20$ mph (from Equation 4), Coifman et al (2003) found the average absolute error in estimated length (via Equation 1) is less than six percent for 210,000 vehicles in the sample data set.

In the presence of heavy truck traffic, e.g., 40-60 percent of the flow, the improvements from Equation 4 degrade because of the high variability in sample median vehicle length. Using data from a detector with heavy truck traffic, Neelisetty and Coifman (2004) developed a methodology to address this problem. As demonstrated in Neelisetty and Coifman, two consecutive vehicles usually have similar speed, even during congestion, and thus, from Equation 1, the ratio of the on-times is a good approximation of the ratio of their lengths. The extension explicitly recalibrates speed estimates by looking for two consecutive vehicle measurements possessing the longest feasible vehicle length and the shortest feasible length, roughly 80 ft and 18 ft, respectively or a ratio of 4:1 in successive on-times. When this ratio is observed in the on-times, one then knows the vehicle lengths in addition to the on-times and can use Equation 1 to estimate speed. Further checks are then made to eliminate transient detection errors that would otherwise disrupt this speed estimation. The paper reported an average absolute percent error under six percent for a detector with heavy truck traffic, but the site also had little congestion.

3 IMPROVED SPEED ESTIMATION FROM SINGLE-LOOP DETECTORS

This research set out to estimate speed using a combination of the moving median method of Equation 4, Coifman et al (2003) and the sequence method of Neelisetty and Coifman (2004). Some shortcomings were encountered and a third technique was devised that examines the on-time distribution within a sample (henceforth called the distribution method). The research also considered the conventional speed estimates from Equation 2. Details of the three non-conventional speed estimation methods follow.

3.1 Moving Median Method

For this study, the median on time in Equation 4 is taken from a fixed window of 33 vehicles centered on the current vehicle. The window moves by one vehicle each sample, hence "moving median". This same window is used when applying the conventional speed estimate from Equation 2 as well as the other non-conventional methods. In any case, the fixed number of vehicles ensures that there will be many vehicles in the sample, even during periods of low flow.

3.2 Sequence Method

If the percentage of long vehicles can fluctuate from sample to sample, then the true value of L in Equation 2 will vary as well. If the fluctuation is large enough, the true value of L in Equation 4 will also vary. Following Neelisetty and Coifman (2004), the on-time ratio between two successive vehicles should be proportional to their length, even during congestion. For most pairs of successive vehicles this fact does not help; however, when the two successive vehicles are the longest and shortest vehicles, one can deduce their lengths directly from the on-times. From Figure 1A, the longest vehicles are about 70 ft and the shortest are about 20 ft, i.e., a ratio of 3.5:1. In the absence of detector errors, this length ratio can only be observed from such a pair of long (LV) and short (SV) vehicles, i.e., SV followed by LV; or LV followed by SV. To accommodate the fact that these two populations have some variability in lengths and that the speeds might not be exactly equal, the method looks for ratios between successive on-times that fall in the range of 3.0 to 4.5. When such a ratio is observed in the on-times, Equation 1 is used to estimate the speed of the two vehicles given l_{SV}^{A}=20 ft and l_{LV}^{A}=70 ft, i.e.,

$$\hat{v}_{SV} = \frac{l_{SV}^{A}}{on_{SV}}; \qquad \hat{v}_{LV} = \frac{l_{LV}^{A}}{on_{LV}} \tag{5}$$

If there are multiple sequences within sample, the algorithm keeps estimating speed for each sequence and then assigns the median speed from all of the individual estimates to all vehicles within the sample. Otherwise, when there are no such sequences within the sample, the algorithm falls back to the moving median method. After working with the sequence method it was found that it fails too frequently during congestion. The assumption that two successive vehicles have the same speed simply does not hold at low speeds when acceleration is non-negligible, typically when speeds drop below 10 mph in stop-and-go traffic. When truck volumes are low, e.g., under typical urban conditions of around 10 percent, the sequence method uses only a few vehicles in the sample to estimate speed, making it vulnerable to making large errors if these vehicles are measured incorrectly or if the pair of vehicles used to estimate speed are far removed from the subject vehicle (i.e., the pair fall at the start or end of the sample of 33 vehicles) and speed varies significantly over the sample.

3.3 Distribution Method

The limitations of the Sequence method in congestion led to the development of a new method that considers the entire distribution of on-times observed in a sample. As with the moving median, vehicles are sampled in a moving window of a fixed number of 33 vehicles, centered on the subject vehicle. The measurements are sorted into bins by every 1/6 sec and a moving average of three bins is taken. If this sample exhibits a clear bimodal distribution, e.g., as seen in Figure 1B, then the two peaks can be localized and the speeds estimated using Equation 5. If the resulting distribution is not bimodal, a series of steps are taken to estimate the speed. The details of the process are as follows.

First check to see if the sample exhibits the expected bimodal distribution, i.e., establish whether there are two peaks. If so, following the same logic used in the Sequence Method, check to see if the ratio between the two mode on-times is in the neighborhood of 3-4.5. Explicitly enumerating the steps,

1) find the dominant mode on-time, i.e., the largest peak
2) search for observations within 3 to 4.5 times larger than the dominant mode
3) search for observations within 3 to 4.5 times smaller than the dominant mode
4) compare the number of observations (2) and (3) to decide which one has more observations
5) if a clear secondary peak from (4) emerges with three (just under 10 percent of the sample) or more vehicles, the sample is considered bimodal and analysis continues to (6), otherwise, the sample is treated using one of the techniques that follow
6) assign assumed average vehicle length to the dominant mode based on the location of the secondary peak with more observations from (4) (l_{SV}^{A} or l_{LV}^{A}) and estimate speed from Equation 5

As will all steps in this method, once a mode has been identified, the exact on-time is determined by taking the median of all of the individual on-times within the mode. With the threshold of 10 percent of the vehicles having to fall in the secondary peak before the distribution is considered bimodal, one would frequently expect samples to be classified as unimodal, e.g., in practice it is not uncommon to find that all 33 vehicles within a sample are passenger vehicles yielding a unimodal on-time distribution. For these unimodal distributions, taking 45 mph as a conservative lower bound to free flow conditions, using Equation 1 one can calculate the feasible on-times for SV and LV under different traffic conditions. The on-time of 20ft vehicle at 85 mph should be 0.16 sec and at 45 mph should be 0.3 sec. Similarly the on-time of 70 ft vehicle at 85 mph should be 0.6 sec and at 45 mph should be 1.1 sec. In other words, Region 1- 0.16 < mode(on-time) < 0.3: 20 ft vehicle traveling above 45 mph (free flow)

Region 2- 0.3 < mode(on-time) < 0.6: 20 ft vehicle traveling below 45 mph (congestion)

Region 3- 0.6 < mode(on-time) < 1.1: Either 20 ft traveling below 45 mph or 70 ft vehicle traveling above 45 mph (uncertain)

Region 4- 1.1 < mode(on-time): Either 20 ft or 70 ft vehicle traveling below 45 mph (congestion)

For each sample the dominant mode on-time will fall in one of these four regions. If the dominant mode falls within Region 1 or 2, it can be deduced directly that the mode corresponds to a SV and speed can be estimated from Equation 5. In region 4 it is not clear what the dominant vehicle is, but it is clearly congested. The largest ambiguity arises in region 3, the mode is either due to free flowing LV or congested SV. To identify the traffic condition of a unimodal sample falling in region 3, we apply the following three tests:

Occupancy filter: Empirically, low occupancy corresponds to freely flowing traffic with low flow (Jain and Coifman, 2005). Therefore, a sample can be considered as free flowing if its occupancy is less than a certain threshold (15 percent in this study). If so, speed is estimated from Equation 5 assuming the mode corresponds to a LV. Otherwise, analysis continues with the next two steps in parallel,

On-time variance: In general speed during free flow is more stable than during congestion because a common feature of congested traffic is acceleration and deceleration waves. Furthermore, the relative impact to on-time of a given small speed fluctuation (e.g., 1 mph) is inversely proportional to speed. For the same level of speed fluctuations the variation of on-time during free flow is less than congestion. An on-time sample variance of 0.11 [sec^2] is used as the threshold between free flow and congested, as derived from empirical analysis of dual-loop data.

Estimated speed from previous sample: Two successive samples will typically have similar speeds, i.e., the transitions between free flow and congestion are only observed a few times a day (if at all). So if a unimodal distribution is found with the mode in Region 3 in one sample, the estimated speed from the preceding sample is used as a proxy for the traffic condition of the current sample.

If the sample is deemed congested by the on-time variance and this result is consistent with the previous sample, speed is estimated from Equation 5 assuming the mode corresponds to a SV. Likewise, if both tests indicate that conditions are free flowing, then speed is estimated from Equation 5 assuming the mode corresponds to a LV. If none of the above cases are met, then the sample is treated as an exception, as discussed below. (As presented in Coifman, 2007, most of the non-exception samples are assigned to the correct traffic condition).

When the mode falls in Region 4, traffic has to be congested, whether the dominant vehicle is long or short. But differentiating between the possible vehicle lengths is necessary to get an accurate speed estimate. Given a unimodal distribution, one cannot differentiate between the two situations. So the algorithm increases the sample size to 51 vehicles and examines whether the distribution has changed to a bimodal distribution or remains a unimodal distribution. If the larger sample turns out to have a bimodal distribution then the vehicle corresponding to the dominant mode (LV or SV) is assumed to apply to the single mode of the smaller sample and used to estimate speed from the smaller sample via Equation 5. Otherwise, the sample is also treated as an exception.

There are three exceptions where the above methodology is not applied to estimate speed for a given sample, namely,

- Samples whose distributions have more than two modes
- Unimodal samples falling in Region 3 that are not filtered from the three tests
- Samples falling in Region 4 that still have unimodal distribution after expanding the sample size

Since it is not likely to observe 33 successive long vehicles in a lane, the shortest on-times likely come from passenger vehicles. So for the exceptions, the second shortest on-time measurement within each sample is taken and assumed to come from a passenger vehicle. Taking the second shortest reduces sensitivity to detector errors that might cause erroneously short on-times. Speed is then estimated from this on-time using SV in Equation 5. Thereby estimating speed for one of the faster passenger vehicles in the sample and assuming it applies to all of the vehicles in the sample. Finally, note that these exceptions are relatively uncommon, comprising less than one percent of the samples in Coifman (2007).

4 PERFORMANCE EVALUATION AGAINST DUAL-LOOP DETECTORS

The four speed estimation methods were evaluated in two ways, first in terms of the actual measured on-times (upstream loop) and speeds from dual-loop detectors on I-71 in Columbus, OH (Coifman, 2006b). The monitored portion of I-71 extends from the central business district (CBD) to the northern suburbs, as highlighted in Figure 2A. The deployment covered roughly 14 miles, with dual-loop detector stations every mile and an average of two single-loop detector stations between each successive pair of dual-loop detector stations. The detector stations report individual transition data whenever a given detector becomes occupied or clears with each vehicle that passes, sampled at 240 Hz. As with the earlier speed estimation studies, the dual-loop detectors provided a ready source of ground truth for vehicle speeds and lengths. Also like the earlier studies, except for a few detectors, these urban data are characterized by relatively low truck volume. Figure 1A shows a typical distribution of individual vehicle lengths observed in this corridor over a 24 hr period. As with most stations, this bimodal distribution is characterized by a tall, narrow peak around 20 ft due to passenger vehicles and a shorter and broader peak around 70 ft due to longer vehicles. Figure 1B shows the corresponding on-times from known free flow periods.

But one of the objectives of this research is to extend single-loop based length classification to detectors with high truck volumes. Which leads to the second evaluation, data with higher truck volumes were synthesized by combining individual measured speed and arrival times from a dual-loop detector station that experiences recurring congestion with synthetic vehicle lengths for the vehicles and then calculating the new set of on-times that would result from Equation 1. Needless to say, the actual lengths and on-times are discarded. Each synthetic vehicle length was determined via a two-step process, first randomly determine whether the given vehicle was long or short based on the desired percentages of each type of vehicle (the threshold between the two groups was set at 50 ft, to fall between the two modes, e.g., Figure 1A). Then for the given vehicle type, randomly sample a synthetic vehicle length from an empirically observed distribution of either LV or SV lengths at a dual-loop detector station. This new length is assigned to the vehicle and the corresponding on-time is calculated via Equation 1. The individual vehicle speed and synthetic length are stored for validation purposes. When estimating speed for the conventional method (Equation 2) with synthetic data, the process of synthesizing the data has disrupted the occupancy and the new off-times are inaccurate. So instead, we calculate mean(on) and employ Equation 3 to estimate speed. All three of the non-conventional methods only use the on-times and they are not impacted by the disruption to the off-times.

4.1 Speed Estimation

The four speed estimation methods were applied over the same fixed sampling windows of 33 vehicles, to the measured on-times from the upstream loops at all 13 operational, northbound dual-loop detectors on I-71 for the entire month of April 2005. The cumulative distribution function (CDF) of the absolute error in speed estimation relative to the measured speed from each of the four methods was calculated over the month at each station. Figure 3 summarizes the results, tabulating the 90th percentile of the absolute difference between the each estimate and the corresponding measured speed (Figure 3A-D) and measured length (Figure 3E-H) over the entire month for each lane at each station. All three of the new estimation methods yield similar performance, and this performance is generally better than the conventional method. Among the three new methods, the sequence method yields slightly poorer results

Moving to the synthesized data from station 1 for the same month, the percentage of LV was varied from 10 to 90 percent in 10 percent increments. Performance relative to the measured speed and measured vehicle length were examined, with similar results. Figure 4 shows the CDF of the absolute error in speed estimation relative to the measured speed from each of the four methods, over the month of synthetic data. Each row corresponds to a different method and each column corresponds to a different lane. In each subplot of Figure 4A-B the CDFs forms a fan, with the 10 percent LV with the smallest errors on the far left and 90 percent LV with the largest errors on the far right. Note that throughout this figure the synthetic data with 10 percent LV is comparable to the results with measured data (see Coifman, 2007 for details). If one only examined the 10 percent LV curves, there is little difference between the three non-conventional methods, while the errors are roughly twice as large in the conventional method. At higher truck volumes, both the conventional method and the moving median method degrade due to the fact that L^A in Equations 2 and 4 is no longer representative of the vehicle fleet. As the fleet becomes more homogeneous at higher truck flows, conceivably the errors could be countered at least in part by actively selecting a new value for L^A. But when the percentage of trucks and cars vary throughout the day or if the two groups are roughly equal in numbers, even such a recalibration will fall to solve all of the problems. In contrast, the sequence method and distribution method show little change in performance as the percentage of long vehicles increases. In other words, for these two methods there is no need to recalibrate l_{SV}^A and l_{LV}^A in the presence of different percentages of trucks. Close inspection of Figure 4C-D reveals that the distribution method has smaller errors. The difference between the two methods is more apparent in Figure 5A-B, which tabulates the average absolute difference between each estimate and the corresponding measured speed over the month, across all lanes at station 1, as the percentage of trucks varies between 10 and 90 percent. The figure presents separately the results during free flow and congestion, using a measured speed of 45 mph as the threshold. The sequence method has a higher absolute average error because the method typically only uses a small number of the on-times observed in a given sample, and thus, is more sensitive to detector errors and changes in speed over the sample. The process was repeated in Figure 5C-D for the estimated length, with similar results. As a result, the distribution method will be used throughout the remainder of this work.

While reviewing Figure 5C-D, note that the average absolute errors should be expected to increase as the truck flows increase because the average vehicle length increases. Whether looking at speed or length, the estimation errors increase significantly during congestion. Although not readily apparent in these plots, all four methods exhibited degraded performance during heavy congestion, overestimating vehicle lengths for passenger vehicles when speeds were below 20 mph (see Coifman, 2007 for details). This problem arose for several reasons; first, even with samples of just 33 vehicles, the chosen speed for the sample may not be representative of a specific vehicle's speed. Second, at these low speeds, acceleration becomes non-negligible, impacting both the measurements and the estimates. When measured speeds are above 20 mph, both the sequence method and distribution method yield good results.

4.2 Length Based Classification

Using the same month of data, the measured and estimated lengths are calculated (via the dual-loop detectors and distribution method, respectively) for the 13 operational, northbound dual-loop detectors on I-71. These lengths are then used to classify the vehicle following the ODOT dual-loop detector, length based classification scheme: three bins are used with divisions at effective vehicle lengths of 28 ft and 46 ft. Using the estimated length and repeating with the measured length, each vehicle is sorted into one of the three classes. Then the two classifications per vehicle are compared one with another. If the two classes are identical, it is considered a correctly classified vehicle. Otherwise, it is considered as either an over-classified vehicle (estimated class is higher than the measured class) or an under-classified vehicle (estimated class is lower than the measured class). Performance during free flow and congestion are examined separately (again, using measured speed of 45 mph as the threshold). The total of correctly classified, over-classified and under-classified vehicles are found in each lane for free flow and congestion, the totals are presented in Figure 6. Each station is presented in a different column in the subplots, while the same column is used for the given station in all six subplots. Each column has 60-120 points, one point per lane per day in the month. Figure 6A shows the percentage of free flow vehicles correctly classified each day, Figure 6B shows the percentage over-classified each day, and Figure 6C plot shows the percentage under-classified each day. Over 97 percent of the vehicles are correctly classified at each station. Between the two errors, over-classification is dominant because the vast majority of vehicles passing through the I-71 corridor are passenger vehicles, falling in class 1 and cannot be under-classified. Figure 6D-F repeat this analysis for congested conditions.

Turning to the month of synthetic data at station 1 to emulate detector measurements under different percentages of long vehicles, Figure 7 shows a box plot of the monthly data for each percentage of long vehicles. As with Figure 6, one point is generated per lane per day, for a total of 120 points underlying each column at this station with four lanes. In each box plot, the top and bottom edge of a box show the first and third quartiles and the horizontal line within the box shows the median value of the observations. The top and bottom edges denoted with "T" are boundaries of the maximum and minimum values of the observations, while a plus shape denotes an outlier (defined to be $1.5\times$interquartile beyond the nearest quartile). Figure 6A shows that during free flow the median performance is roughly constant across the different percentages of long vehicles, falling above 99 percent in each case, indicating that the classification methodology is not sensitive to the percentage of long vehicles. Figure 6D shows that during congestion the monthly median of the correctly classified vehicles falls between 80 and 90 percent. Although the performance is worse than under free flow conditions, it is still roughly consistent across the different percentages of long vehicles. Of course the percentage of over-classified vehicles drops and under-classified vehicles increases as the percentage of long vehicles increase, simply because class 1 vehicles can not be under-classified and class 3 vehicles can not be over-classified.

5 Performance Evaluation Against Manually Extracted Data

While performance against dual-loop detector data is good, the fact remains that dual-loop detectors are also capable of making errors, e.g., if they measure the on-time incorrectly, then length calculated from Equation 1 may agree with the dual-loop measurement while both the measurement and estimate are equally incorrect. To control for the possibility of such errors that may impact the dual-loop measurements and single-loop estimates in the same way, this research also collected concurrent detector data and video data at two locations to test the performance of single-loop detector estimated length based vehicle classification. Both locations are shown on the map of Figure 2A, while a sample frame from each site is shown in Figures 2B-C.

The first test location is a classification station on I-70, just east of Brice Rd. The station is equipped with dual-loop detectors and a piezo electric axle detector in each of the three eastbound lanes, normally used to bin vehicles into the 13 FHWA vehicle classes (see, e.g., FHWA, 2001). The station was observed midday, under clear weather and free flow conditions from 10:13 to 14:00, on June 20, 2006. All told, almost four hours of data were recorded and, 9,746 vehicles were seen. The measured speed and length were also recorded for these vehicles. A software tool was developed to semi-automate the extraction of ground truth data from the video (inspired in part by VideoSync, Caltrans, 2007). The tool allows the user to manually measure vehicle length after synchronizing the detector and digitized video data. See Coifman (2007) for details of this tool, for the purposes of this paper, it is sufficient to note that care was taken in the selection of the camera angles to ensure a view angle perpendicular to the roadway (Figures 2B-C) so as to reduce the impacts of projection errors on the video based vehicle lengths. Likewise, the video based vehicle lengths were measured as close to the base of the vehicle as possible.

Figure 8A compares the estimated length from on-times versus manually measured length across all lanes. Most of the points fall close to the diagonal, indicating the estimates are generally close to the measurements. Figure 8B clusters these points based on the resulting length class from the estimated and measured length. The correctly classified vehicles fall in the three cells on the diagonal, while the other six cells tally the various misclassifications. Figure 8C-D repeat the exercise using the measured lengths reported from the dual-loop detectors against the manually measured lengths. Comparing Figure 8A and C, the plots show the reported lengths are closer to the manually measured lengths than the estimated lengths are. However, Table 1 summarizes the classification results, and as evident, the classification performance is very similar whether using estimated or reported length. This result arises from the fact that the classification scheme is tolerant to large length estimation errors provided the true length is far from the boundary between two classes.

The second test location is station 9 on I-71, a single-loop detector station just south of Hudson St, used for real-time traffic management, with single-loop detectors in each of the three lanes, in each direction. This station was chosen because it is closest to an existing close circuit television camera. Detector data and concurrent video were collected for two hours between 12:20 and 14:20, on June 5, 2006, under clear weather and free flow conditions, with a total of 15,251 vehicles recorded by the detectors. Once more the ground truth software was used to manually measure every vehicle length. Of these actuations, 441 do not correspond uniquely to a passing vehicle in the video, and instead were due to the detector "dropping-out" in the middle of a long vehicle and causing "pulse-breakup," i.e., semi-trailer trucks frequently resulted in two or more pulses when these trucks should have only been recorded as a single pulse. These errors highlight the importance of having a validation measure independent of the loop detectors, because the pulse break-ups would degrade performance of dual-loop length measurements as well.

After accounting for the pulse-breakups, 6,998 southbound and 6,648 northbound vehicles were ground truthed. Figures 8E and G compare the estimated length against the manually measured length in the two directions. In both plots the pulses found to be due to breakup are shown in a lighter shade. Figures 8F and H cluster these points based on the resulting length class from the estimated and measured length and again, show the results including those vehicles impacted by pulse-breakup and in parentheses repeats the statistics without the vehicles impacted by pulse-breakup. As before, the correctly classified vehicles fall on the diagonal and the totals are summarized in Table 1.

From Table 1, after excluding pulse-breakups at the I-71 test-site, the methodology had an accuracy of over 99 percent for class 1 and over 93 percent for class 3, while performance was over 74 percent accurate for class 2. Of course these results are mid-day, without congestion. The lower performance in class 2 appears to be due in part to the fact that most of the class 2 vehicles are close to the lower boundary and are frequently misclassified as class 1. A similar error rate was observed for class 2 when using the reported vehicle length measured by the dual-loop detectors (last column of Table 1). Most misclassified long vehicles at the I-71 test-site were due to pulse-breakup. When the pulse-breakups are included, the on-times for long vehicles are too short and many class 3 vehicles are misclassified, as evident in Table 1. Even including these errors, from Figure 8F and H, very few of the errors were more than one class away from true.

6 CONCLUSIONS

Roadway usage, particularly by large vehicles, is one of the fundamental factors determining the lifespan of highway infrastructure. Each state typically has several dozen vehicle-classification stations to monitor large vehicle usage, the simplest of these stations use dual-loop detectors to measure vehicle length. Meanwhile, single-loop detectors are the most common vehicle detector in use to monitor traffic, both for real-time operations and for collecting census data such as AADT. Collecting reliable length data from these detectors has been considered impossible due to the noisy speed estimates provided by conventional data aggregation at single-loop detectors. This research has questioned those assumptions, demonstrating length based vehicle classification on freeways from single-loop detectors under a wide range of traffic conditions, yielding estimates that approach the accuracy of a dual-loop detector's measurements. The research promises to provide a lower cost means of collecting vehicle classification data, provide a software based solution when one of the detectors in more sophisticated classification station fails, and extend classification to traffic monitoring stations already deployed for real-time traffic management in urban areas. In fact the classification work could allow these urban traffic management systems to better monitor freight traffic within the metropolitan areas.

The present study developed a length based classification methodology from single-loop detectors. In the process, it lead to improved speed and length estimates from single-loop detectors. This work is equally applicable to the non-invasive detectors such as the RTMS and SmartSensor that seek to replace loop detectors using wayside mounted sensors by emulating the operation of single-loop detectors.

The work started by refining our existing speed estimation algorithms to accurately estimate speed under a wide range of traffic conditions: free flow to congested, as well as ranging from low to high truck volumes. An important innovation of this work was the synthetic data used to capture higher truck volumes than empirically observed. Following the ODOT length based classification scheme for dual-loop detectors, the lengths are then used to classify vehicles into three bins. This classification is evaluated against concurrent measurements from video and dual-loop detectors.

Unlike earlier efforts to classify vehicles from single-loop detectors, this work does not employ aggregate data, instead, it uses the individual vehicle actuations and explicitly classifies each and every vehicle. This point is important, because the earlier efforts that relied on aggregate measurements from dual-loop detectors allow over-counting errors to cancel undercounting errors, so the reported results in the earlier studies may be overly optimistic. Unlike the earlier studies, this work considered truck volumes over 10 percent of the fleet, explicitly generating synthetic detector data to simulate truck volumes up to 90 percent. Furthermore, this work did not rely strictly on dual-loop detectors for validation, we manually generated ground truth vehicle length data from concurrent video for approximately 25,000 vehicles. The manual verification from video ensured that any detector errors that might impact the dual-loop measurements would not bias the results. As it stands, in the process of generating these ground truth length data, we found loop detectors were dropping-out in the middle of semi-trailer trucks, a problem that impacted both dual and single-loop detector classifications alike.

Performance of the methodology degrades during congestion due to the fact that we estimate a "typical" speed within a sample of many vehicles and a given vehicle may have a speed that is far from typical within a congested sample. The speed estimation methods can be used to reliably detect congested conditions, so results during such periods can at least be identified by the current methodology and weighted appropriately. There is likely room for further improvement in estimating individual vehicle speed from single-loop detectors during heavy congestion.

Table 1 shows that the length based single-loop detector estimation classification results are very close to the dual-loop detector measured length based classification results for the I-70 test-site and, after excluding pulse-breakup, the I-71 test-site. But one cannot summarily exclude pulse-breakups based on the ground truth data and the fact remains that stations installed to measure speed might not count vehicles as accurately as a station deployed and tuned primarily to classify vehicles. So work remains to investigate the feasibility to catch detector errors.

7 ACKNOWLEDGEMENTS

We are indebted to ODOT for in-kind help, particularly in regard to collecting field data. At various points in this research invaluable help came from the following ODOT employees: Dave Gardner, Tony Manch, David Stewart, Kevin Calovini, Farouk Aboukar, Matt Graf, and Nick Hegemier.

8 REFERENCES

Benekohal, R., Girianna, M. (2003) "Technologies for Truck Classification and Methodologies for Estimating Truck Vehicle Miles Traveled," *Transportation Research Record 1855*, pp. 1–13.

Caltrans (2007), *VideoSync*, http://www.dot.ca.gov/research/operations/videosync/index.htm, accessed on June 5, 2007.

Coifman, B., (2001) "Improved Velocity Estimation Using Single Loop Detectors", *Transportation Research: Part A*, vol 35, no 10, pp. 863-880.

Coifman, B., Dhoorjaty, S., and Lee, Z., (2003) "Estimating Median Velocity Instead of Mean Velocity at Single Loop Detectors", *Transportation Research: Part C*, vol 11, no 3-4, pp 211-222.

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

Coifman, B., (2006b) *The Columbus Metropolitan Freeway Management System (CMFMS) Effectiveness Study: Part 2 - The After Study*, Ohio Department of Transportation.

Coifman, B., (2007) *Vehicle Classification from Single Loop Detectors*, Project 05-02, Midwest Regional University Transportation Center, University of Wisconsin, Madison.

Dailey, D. (1999) "A Statistical Algorithm for Estimating Speed from Single Loop Volume and Occupancy Measurements." *Transportation Research-B*, Vol 33B, No 5, June 1999, pp 313-322. Elsevier Science, London.

FHWA, (2001) *Traffic Monitoring Guide*, USDOT, FHWA-PL-01-021.

French, J., French, M., (2006) *Traffic Data Collection Methodologies*, Pennsylvania Department of Transportation, Contract 04-02 (C19).

Gajda, J., Sroka, R., Stencel, M., Wajda, A., Zeglen, T. (2001), "A Vehicle Classification Based on Inductive Loop Detectors." *Proc. of the 18th IEEE, Instrumentation and Measurement Technology Conference*, Vol. 1, pp 460-464.

Jain, M., Coifman, B., (2005) "Improved Speed Estimates from Freeway Traffic Detectors" *ASCE Journal of Transportation Engineering*, Vol 131, No 7, pp 483-495

Kwon, J., Varaiya, P., Skabardonis, A. (2003), "Estimation of Truck Traffic Volume from Single Loop Detectors with Lane-to-Lane Speed Correlation." *Transportation Research Record 1856*, TRB, National Research Council, Washington, D.C., 2003, pp 106-117

Neelisetty, S., Coifman, B., (2004) "Improved Single Loop Velocity Estimation in the Presence of Heavy Truck Traffic" *Proc. of the 83rd Annual Meeting of the Transportation Research Board.*

Mikhalkin, B., Payne, H., Isaksen, L. (1972) "Estimation of Speed from Presence Detectors." *Highway Research Record 388*, pp 73-83. Highway Research Board, Washington, DC.

Pushkar, A., Hall, F., Acha-Daza, J. (1994) "Estimation of Speeds from Single-Loop Freeway Flow and Occupancy Data Using Cusp Catastrophe Theory Model." *Transportation Research Record 1457*, pp 149-157. Transportation Research Board, Washington, DC.

Reijmers, J., (1979) "On-Line Vehicle Classification," *Proc. of the International Symposium on Traffic Control Systems*, Vol 2B, Institute of Transportation Studies, University of California at Berkeley, pp 87-102.

Wang, Y., Nihan, N. (2000) "Freeway Traffic Speed Estimation With Single Loop Outputs," *Transportation Research Record 1727*, pp120-126, Transportation Research Board, Washington, DC.

Wang, Y., Nihan, N. (2003) "Can Single-Loop Detectors Do the Work of Dual-Loop Detectors?" *Journal of Transportation Engineering*, Vol. 129, No. 2. ASCE, pp 169-176.

Wang, Y., Nihan, N.L. (2004), "Dynamic Estimation of Freeway Large-Truck Volumes Based on Single Loop Measurements". *Journal of Intelligent Transportation Systems*, Vol. 8, No. 3, 2004, pp. 133-141.

Zhang, G., Wang, Y., Wei, H., (2006) "Artificial Neural Network Method for Length-Based Vehicle Classification Using Single-Loop Outputs," *Transportation Research Record 1945*, pp. 100–108.

Zwahlen, H. Russ, A., Oner, E., Parthasarathy. M. (2005) "Evaluation of Microwave Radar Trailers for Non-intrusive Traffic Measurements." *Transportation Research Record 1917*, pp127-140.

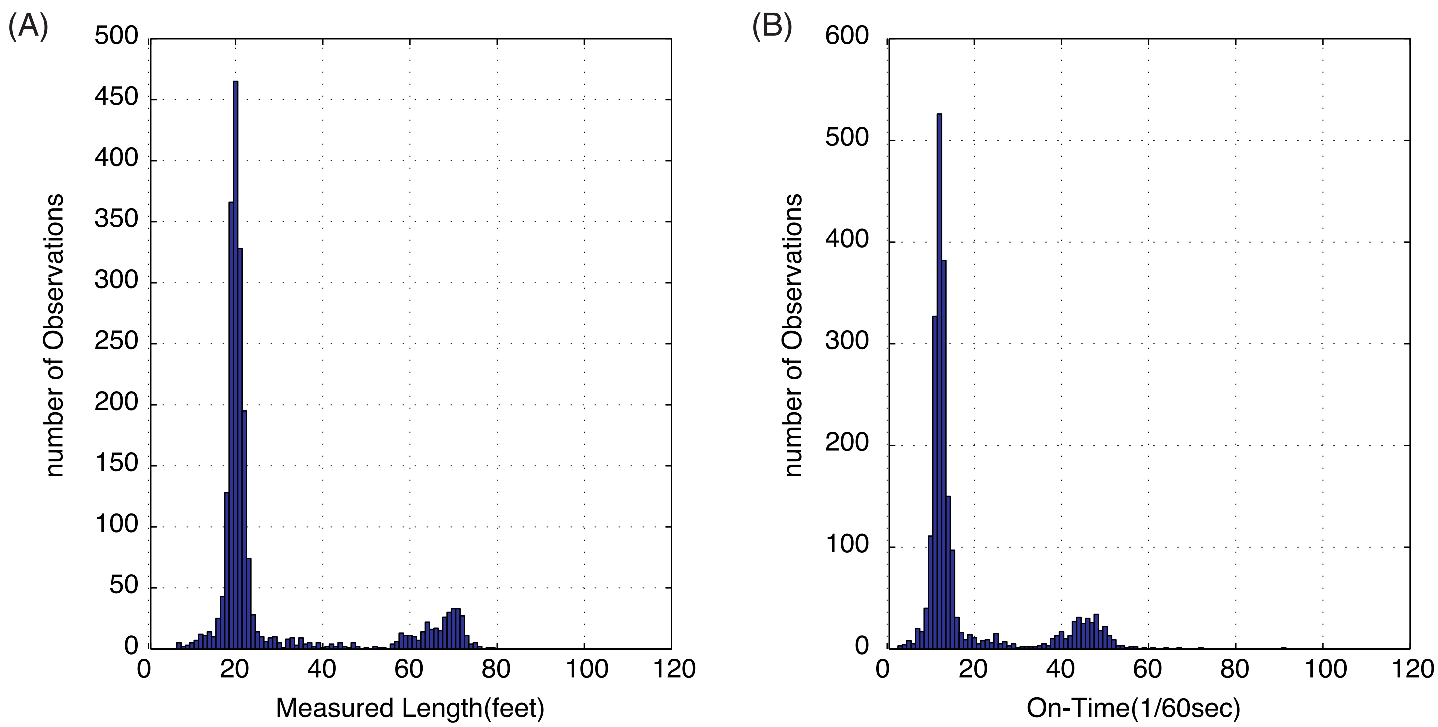

Figure 1, (A) Example of bimodal distribution of measured length, and (B) the corresponding bimodal distribution of on-times.

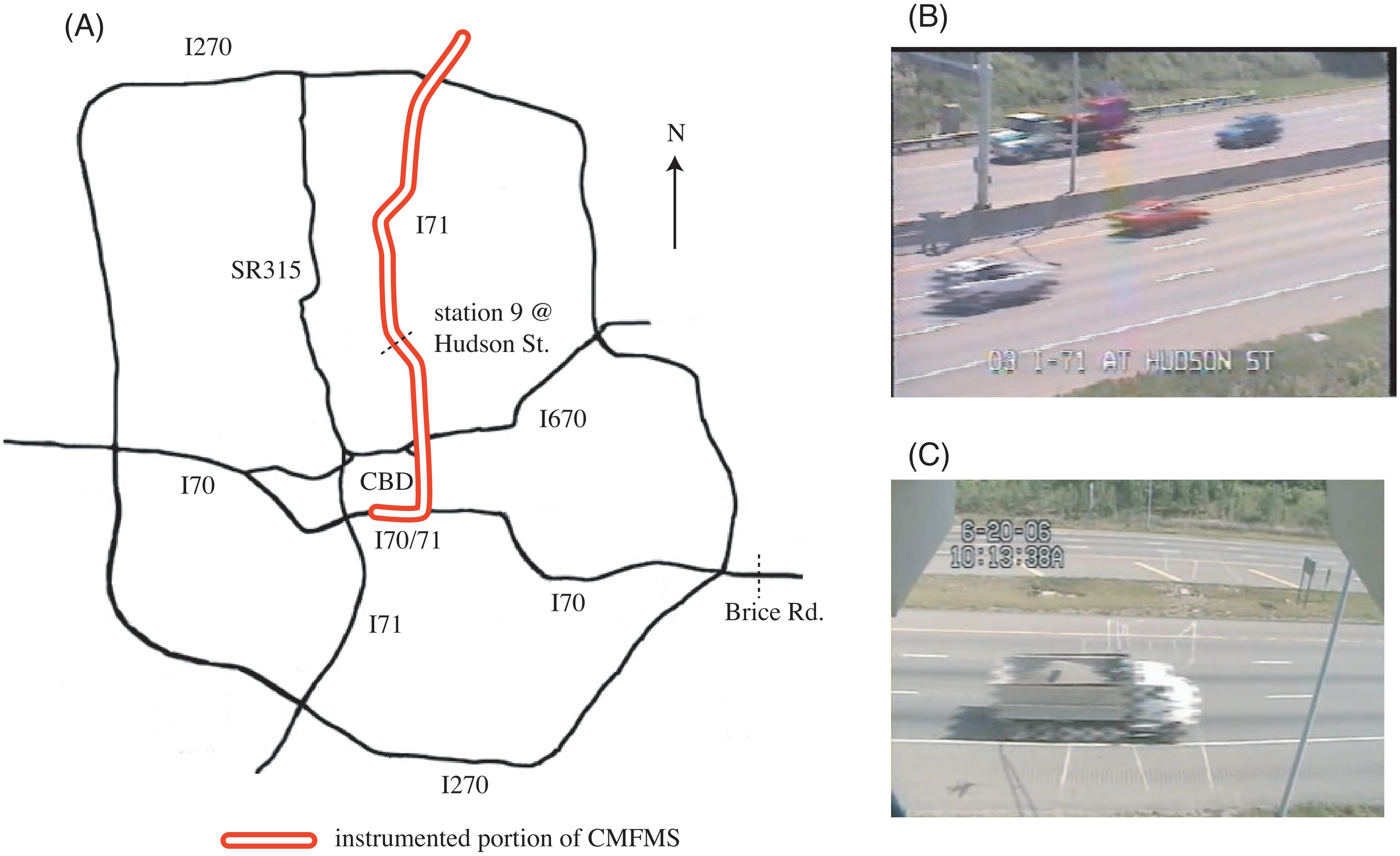

Figure 2, (A) Freeway network in Columbus Ohio, highlighting the instrumented portion of I-71 and the two test sites. Sample frames from (B) I-71 test site, and (C) I-70 test site.

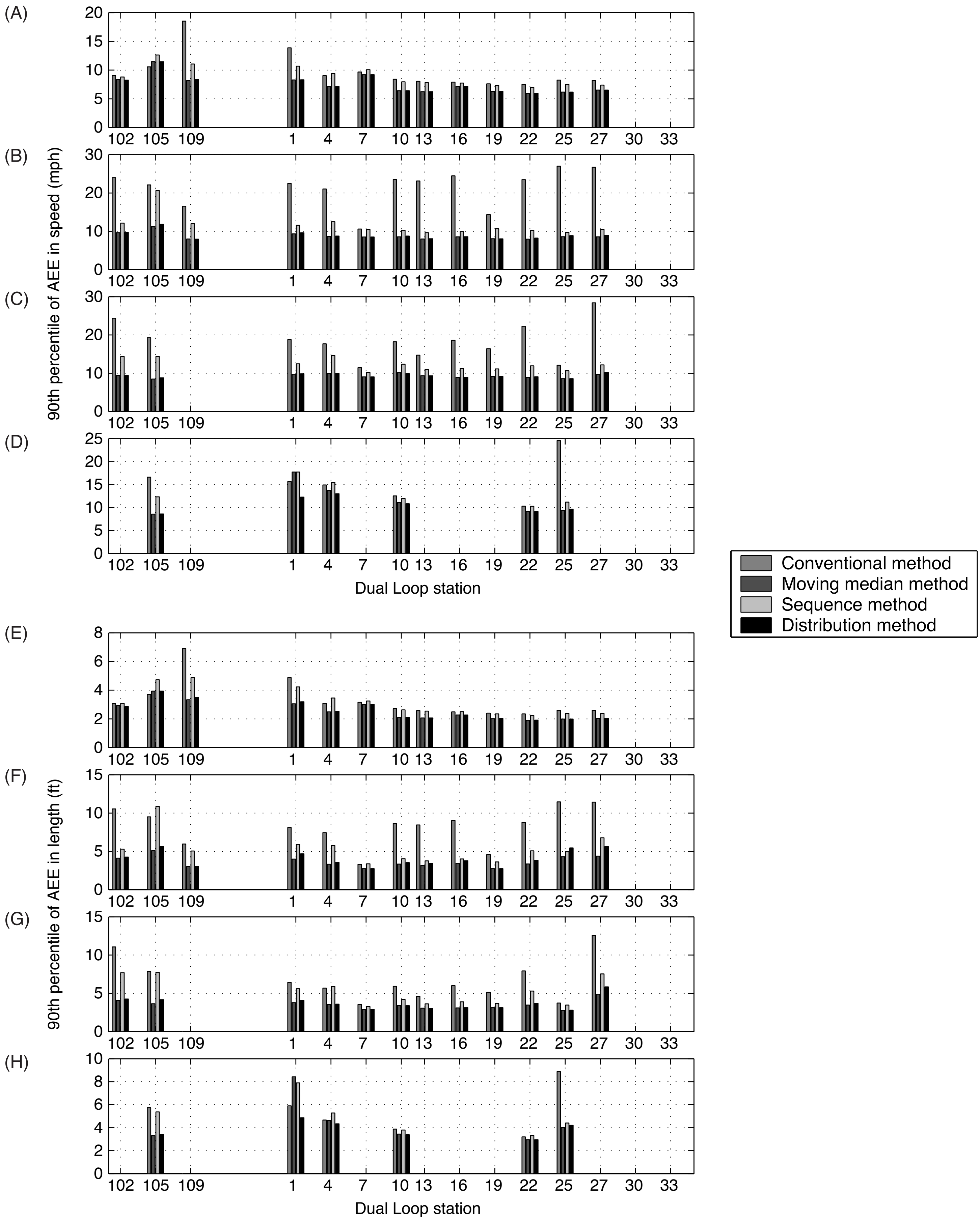

Figure 3, Absolute estimation error at 90th percentile in speed: (A) lane 1 (median), (B) lane 2, (C) lane 3, (D) lane 4 (shoulder) and in length (E) lane 1, (F) lane 2, (G) lane 3, and (H) lane 4.

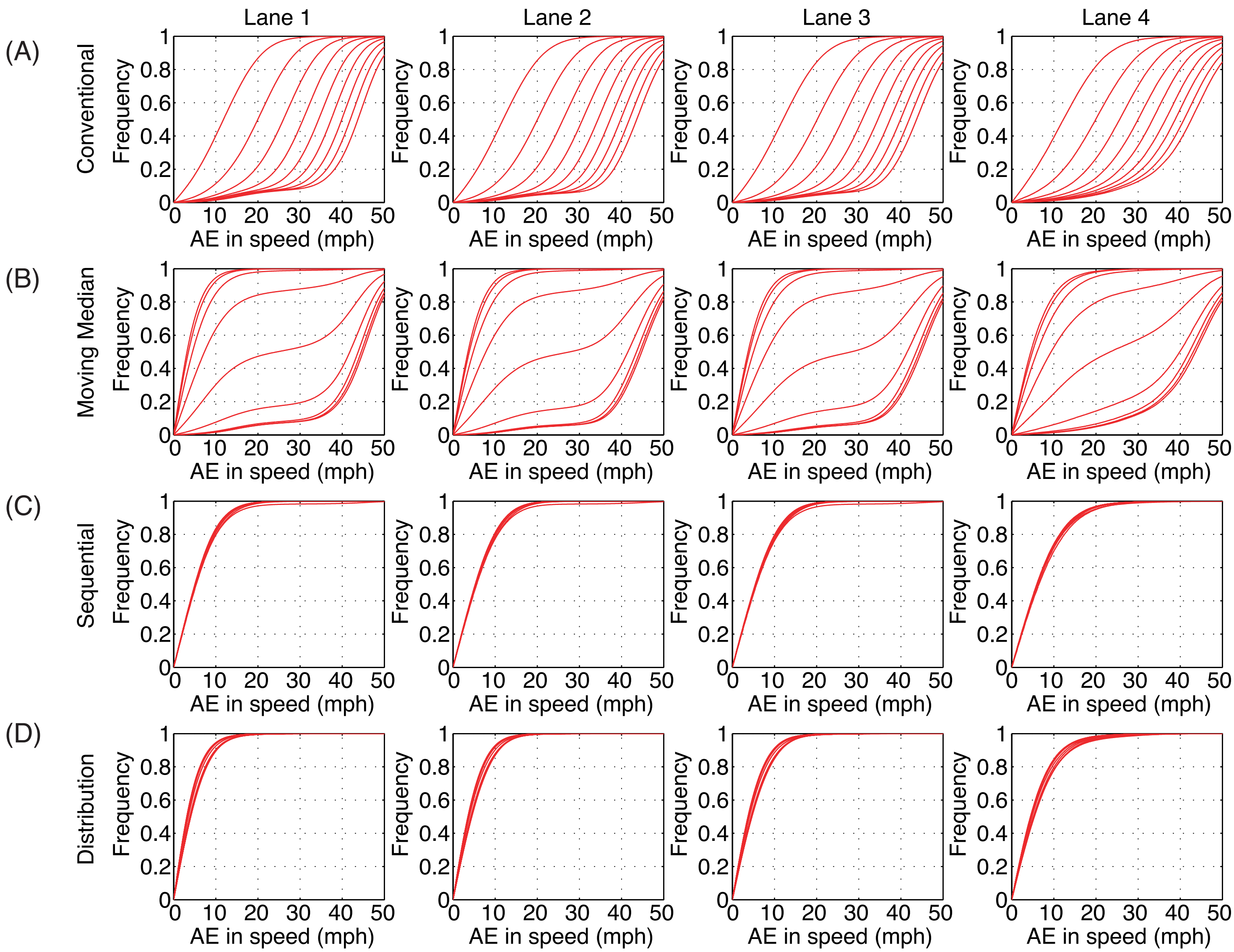

Figure 4, CDF of the absolute error (AE) from the speed estimation over one month (April, 2005) for station 1 when the percentage of trucks varies between 10% and 90%, (A) Conventional Method, (B) Moving Median Method, (C) Sequence Method, (D) Distribution Method. Each column corresponds to a different lane, as indicated.

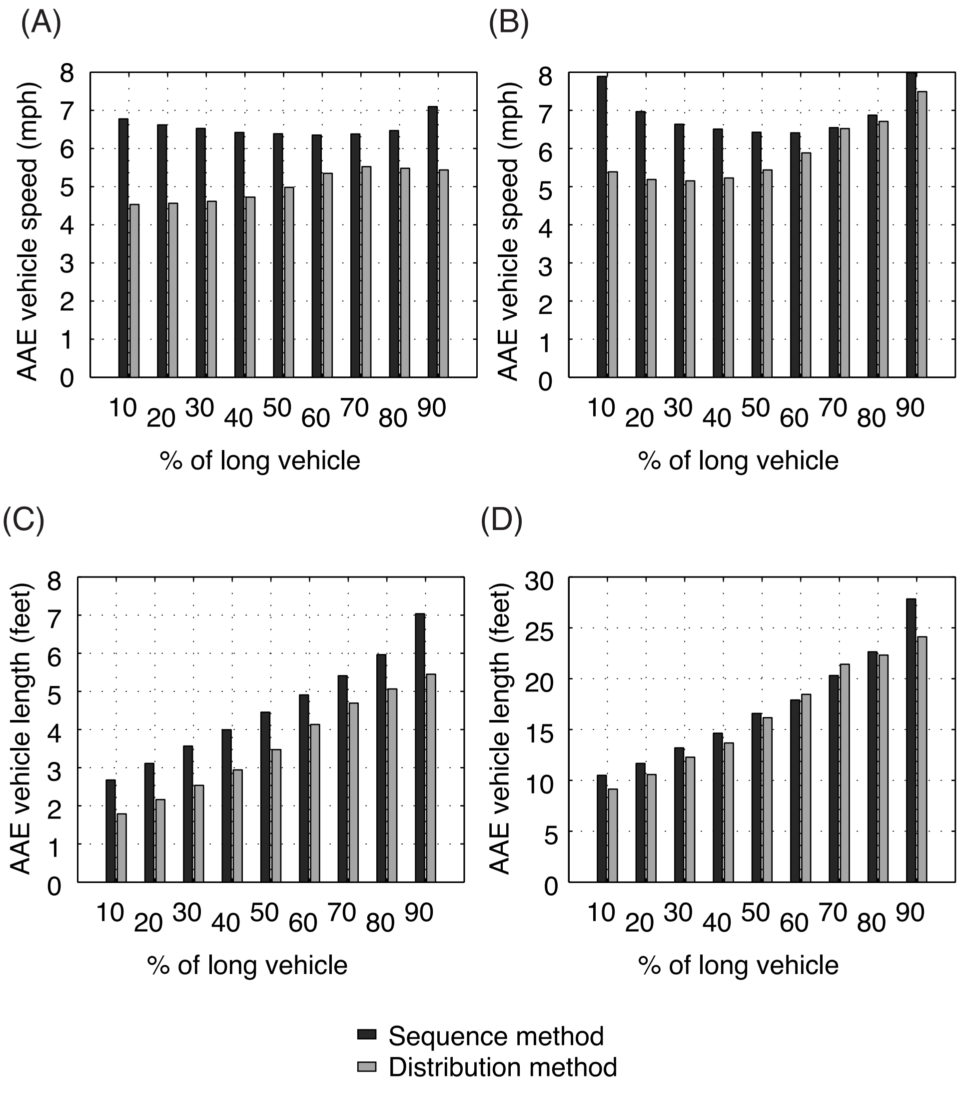

Figure 5, Average absolute error (AAE) in estimated speed across all lanes for the Sequence and Distribution methods when measured speeds are (A) above 45 mph, (B) below 45 mph, and the corresponding measures for estimated length (C) and (D). Note that the vertical scale is larger in (D).

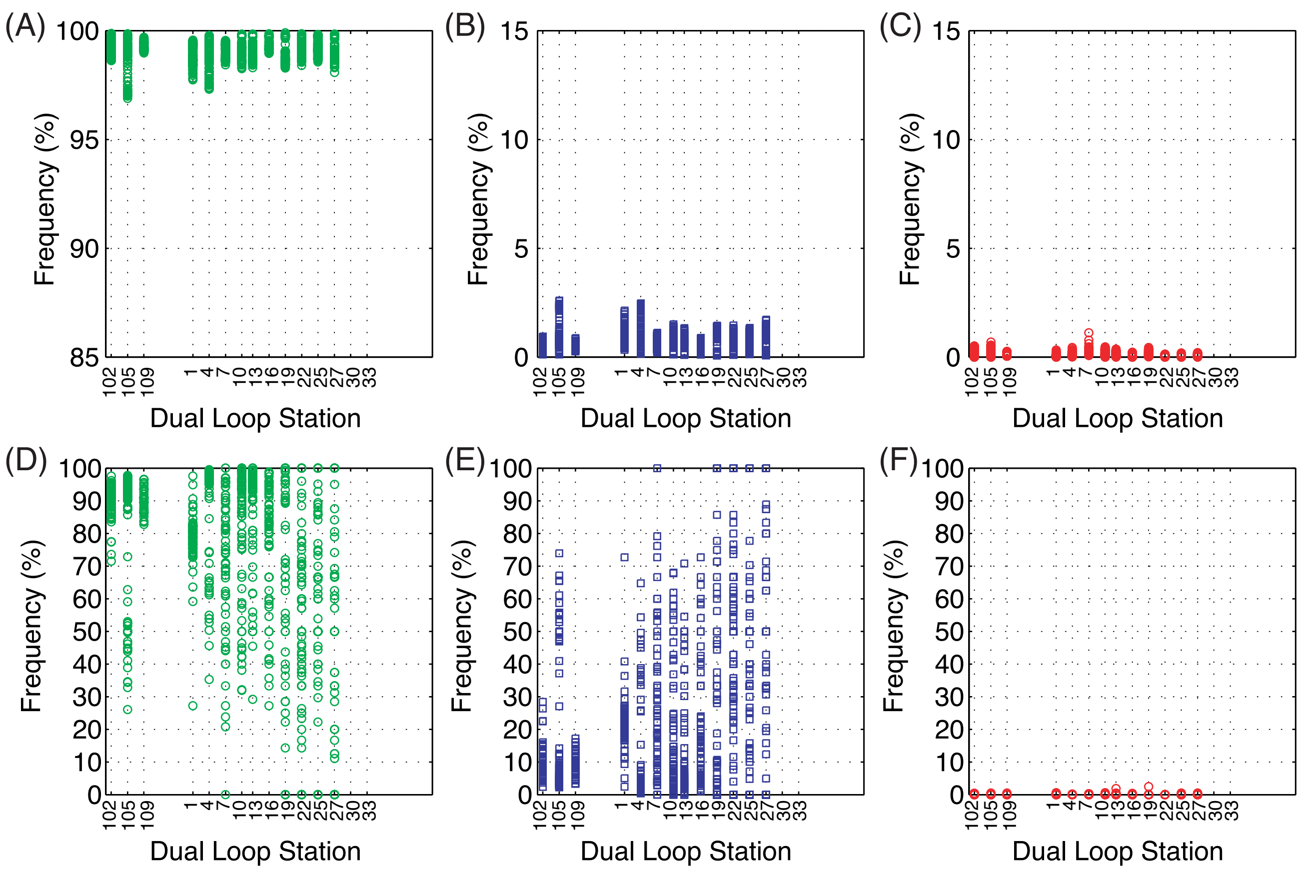

Figure 6, Monthly summary plot for vehicle classification during free flow conditions (A)-(C) and congestion (D)-(F), all lanes, all stations (one point per lane per day), (A) & (D) % of correctly classified vehicles, (B) & (E) % of over-classified vehicles, and (C) & (F) % of under-classified vehicles.

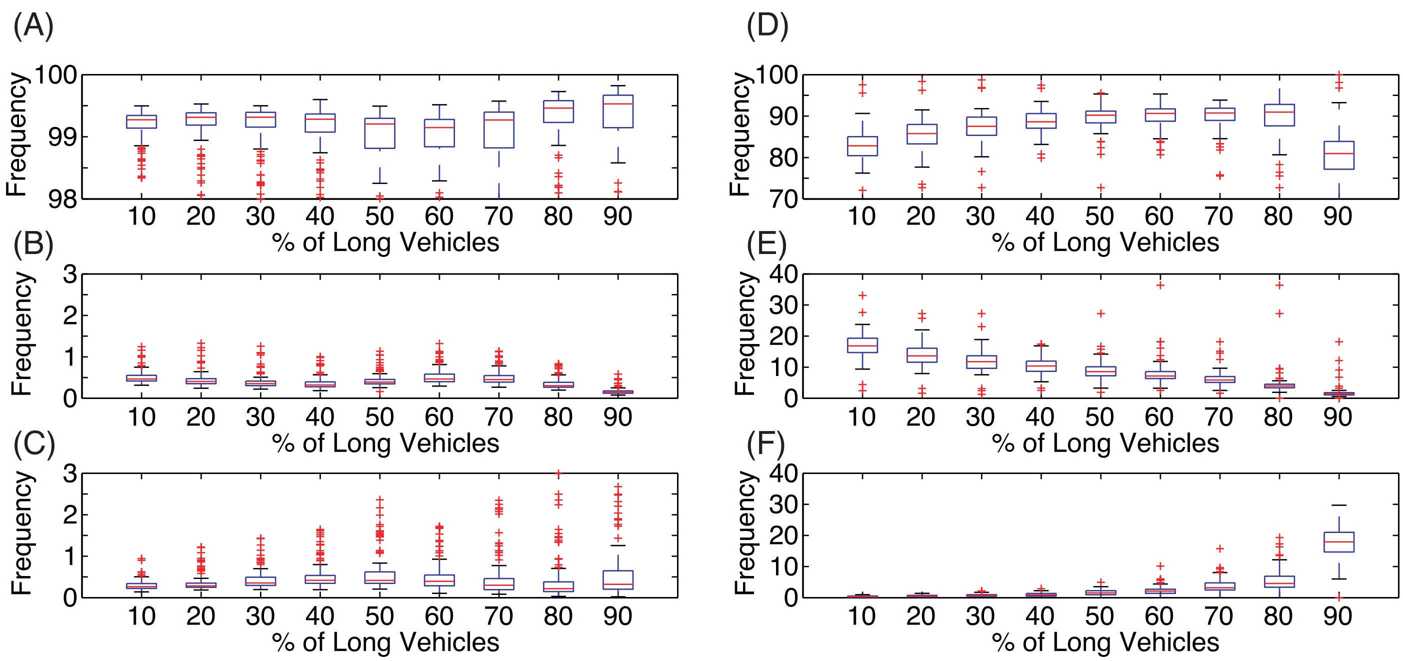

Figure 7, Box plot for vehicle classification during free flow conditions (A)-(C) and congestion (D)-(F) at station 1 northbound over one month when the percentage of trucks varies between 10% and 90%. (A) & (D) % of correctly classified vehicles, (B) & (E) % of over-classified vehicles, and (C) & (F) % of under-classified vehicles.

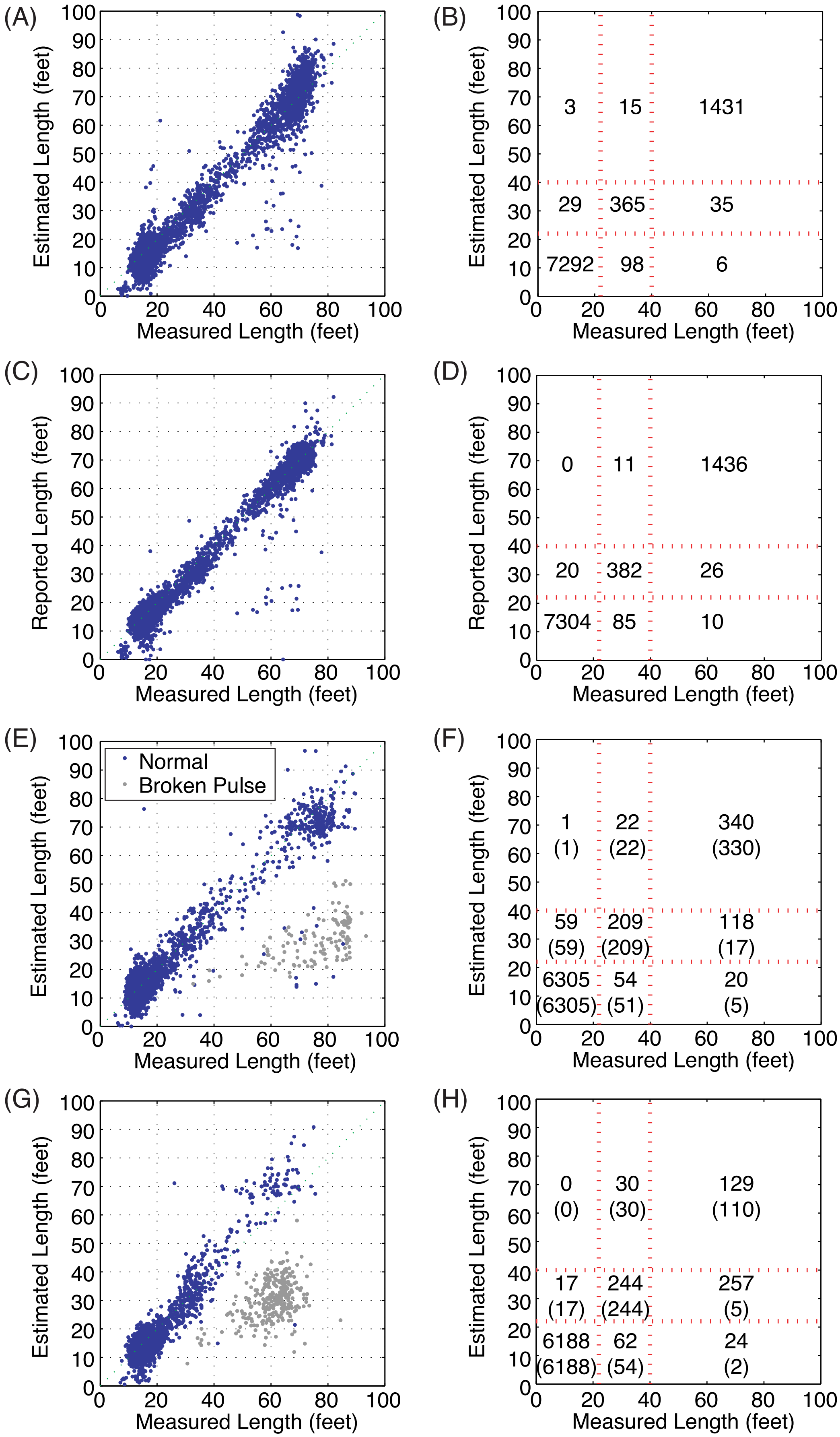

Figure 8, (A) Estimated length versus measured length at I-70 test site, (B) corresponding length based classifications, (C) reported length versus measured length at I-70 test site, (D) corresponding classifications, estimated length versus measured length at I-71 test site (E) southbound, (G) northbound, and corresponding classifications, (F) southbound, (H) northbound.

TABLE 1: Percent of correctly classified vehicles in each class at I70 and I71 test sites

I70 at Brice	Estimated length compared to measured length		Reported length compared to measured length
Class 1	99.6%		99.7%
Class 2	76.4%		79.9%
Class 3	97.2%		97.6%
I71 at Hudson	Southbound with pulse break-up	Northbound with pulse break-up	
Class 1	99.1%	99.7%	
Class 2	73.3%	72.6%	
Class 3	71.1%	31.5%	
I71 at Hudson	Southbound excluding pulse break-up	Northbound excluding pulse break-up	
Class 1	99.1%	99.7%	
Class 2	74.1%	74.4%	
Class 3	93.8%	94.0%	