IMPROVED SPEED ESTIMATES FROM FREEWAY TRAFFIC DETECTORS

Manuscript: TE/2003/22970

Manish Jain
Graduate Research Associate
Department of Civil and Environmental Engineering and Geodetic Science,

The Ohio State University,
Columbus, OH 43210.
Jain.67@OSU.edu

Benjamin Coifman, PhD
Assistant Professor, Department of Civil and Environmental Engineering and Geodetic Science
Assistant Professor, Department of Electrical Engineering

The Ohio State University,
470 Hitchcock Hall,
2070 Neil Ave,
Columbus, OH 43210.
Coifman.1@OSU.edu

http://www-ceg.eng.ohio-state.edu/~coifman

(614) 292-4282

Subjects: Traffic Speed, Traffic Surveillance, Traffic Congestion, Data Collection, Data Processing

ABSTRACT

This paper presents an analytical methodology to increase the accuracy of speed estimates from freeway traffic detectors by integrating information across lanes. Typically these estimates are quite noisy due to unobserved variables and measurement errors. Yet most traffic-monitoring applications employed by operating agencies only require a single measurement per direction at a detector station. Earlier efforts have focused on improving speed estimates of the individual lanes and then using a simple average to arrive at a single measure of speed at the detector station. But simply averaging the data across lanes is likely to preserve a measurement error in any one of the lanes.

This paper improves speed estimates on a lane-by-lane basis using conventional aggregated flow and occupancy data obtained from single loop detectors. It then goes further by exploiting the information in the adjacent lanes to eliminate noise. In the course of analysis, data cleaning tools have been developed to identify and exclude malfunctioning detectors and transient errors, improving the validity of traffic data before estimating speed. These data cleaning tools consist of threshold value tests and basic traffic flow theory principles. To reduce the vulnerability to common transient measurement errors in individual lanes, the median of the improved estimates across lanes is taken in each direction. The methodology is then extended to clean noisy speeds measured from dual loop detectors and other sensors that mimic single loops.

INTRODUCTION

Vehicle speed is one of the most important measures of freeway service quality. It is a principle indicator of highway performance, such as congestion and mobility, which agencies use to monitor freeways. Many Intelligent Transportation Systems (ITS) applications like traveler information and incident detection use speed as an input. Greater accuracy in speed estimates can improve decisions by agencies. Single loop detectors are the most common sensors used on today's freeways. The single loop detectors can measure flow, q, the number of vehicles that crossed the detector per unit time period, T, and occupancy, occ, the fraction of time the detector sensed vehicles above it. Most transportation agencies aggregate q and occ over fixed periods ranging from 20 seconds to 5 minutes. Mean speed is typically estimated using the following equation:

$$\hat{V} = \frac{L * q}{occ} \tag{1}$$

where,

$\hat{V}$ = Estimated mean speed over T,

L = Assumed average vehicle length.

L and $\hat{V}$ cannot be measured independently at a single loop so L is normally set to a constant, but the simple application of Equation 1 fails to account for the variation in true effective average vehicle length. As a result, small variations in q or occ measurements from sample to sample can cause large errors in $\hat{V}$. This problem is most pronounced during periods of low q and occ. During congested periods, $\hat{V}$ exhibits smaller absolute errors due to lower speeds and generally a lower frequency of long vehicles. There are several methods proposed to choose L (Coifman et al. 2003, Dailey 1999, Jia et al. 2001). These methods seem to work well when the aggregation is over longer periods, like five minutes. As shown by Coifman et al. (2003), the assumed average vehicle length becomes more representative as the sampling period increases because the number of vehicles increases. When the number of vehicles in a sample is small, a long vehicle can skew occupancy and hence the speed estimation. Recognizing that low occupancy must correspond to free flow conditions, Jia et al. (2001) and van Zwet et al. (2003) developed a methodology where a time dependent, location specific average vehicle length curve

is developed using known local free-flow speeds. The L during high occupancy periods is estimated by fitting a smooth regression curve with the assumption that L is independent of occupancy (van Zwet et al. 2003). Historical information and average vehicle length time-series are used to make real time speed estimation for subsequent days. This methodology works fine if the time dependent average vehicle length curve can be developed reliably, the free-flow speeds are known, and there is little variation in vehicle lengths within a sample. Other researchers have produced clean estimates of speed by developing algorithms that use statistical methods, like minimum mean-square error, to estimate speed and assume that flow, occupancy, speed and vehicle length follow a distribution with a mean value and variance (Mikhalkin 1972, Dailey 1999, Wang and Nihan 2000, Kwon et al. 2003).

Some researchers have focused on improving data quality using adjacent lane information for detecting bad data samples and for imputing missing or bad samples, e.g., Chen et al. (2003). Combining data across lanes increases the number of vehicles during shorter aggregation periods, since the adjacent lanes provide redundant information about the traffic state (Coifman et al. 2003). In many traffic-monitoring applications, operators are only interested in a single estimate of speed per direction at a detector station. Conventionally, this aggregation is obtained by taking the arithmetic mean of $\hat{V}$ across the individual lanes. But the accuracy of $\hat{V}$ is degraded by the fact that the true vehicle length spans a large range, some vehicles are four times longer than the average vehicle (Coifman et al. 2003). This fact affects the estimates by Equation 1 as it uses fixed L The median value is less sensitive to outliers when there are more than two lanes and is a better central measure for this application.

This paper describes a data cleaning and speed estimation methodology that exploits the redundant information from adjacent lanes and applies basic traffic flow theory principles to 30 second aggregated single loop detector data. The methodology is generic and could be easily be modified for other sampling periods, though it might require some changes to thresholds. The methodology was validated by comparing with measured speeds from dual loop detectors (a pair of detectors in each lane at the detector station, which can measure speed directly from the difference in actuation times between the two loops) to the estimates from a single loop of the pair using data from I-71 in Columbus, Ohio and I-80 in Berkeley, California (Coifman et al. 2000). Of course dual loop detector data are not without error. The largest errors arise when a

vehicle actuates only one of the loops in the dual loop detector, leading to an erroneous speed measurement for that vehicle. As such, the methodology presented in Coifman and Ergueta (2003) was used to exclude unmatched actuations at either loop of a dual loop detector when measuring speed. The excluded actuations were under one percent of all measurements, and are most frequently due to lane change maneuvers over the detectors (Coifman <in review>). After this cleaning step, the harmonic mean of these individual vehicle speed measurements were calculated for each dual loop detector (as has been shown in numerous locations, the space-mean speed can be measured by taking the harmonic mean of individual vehicles' speeds as they pass a point in space and Equation 1 is an estimate for the space-mean speed, e.g., Coifman 2001). Throughout the remainder of this paper these dual loop measurements are used as the reference speed since the intent of this work is to make single loop detectors to approach the accuracy of dual loop detectors. Finally, note that the unmatched actuations were retained when estimating speed from the single loop, since they are representative of what that detector would normally measure.

The methodology presented in this paper is transferable to other sensors that mimic single loops, e.g., the Remote Traffic Microwave Sensor (RTMS), as well as speeds measured from dual loop detectors. The paper is organized into four major sections: data cleaning tools and types of tests to identify malfunctioning detectors are given first. Next, an overview of the development of the speed estimation procedure, including detailed descriptions of the steps and application of the procedure to the data is explained. An analysis and discussion of the results is then presented, followed by conclusions.

IDENTIFYING MALFUNCTIONING DETECTORS AND ERRONEOUS DATA

Before estimating speed from q and occ, it is important to identify erroneous data reported by the detectors. The most common errors arise from a malfunctioning detector or data lost during transmission. Often the malfunctioning detector can be identified easily by visualizing the data. The time-series plots of measurements from a chronically abnormal detector are strikingly different from the typical time-series. Figure 1 compares the time-series of q and occ as reported by two adjacent detector stations on I-71 southbound in Columbus, Ohio. The malfunctioning detector station exhibits flows between 2800 vehicles per hour per lane (vphpl) to 3800 vphpl

between 13:00 and 15:00, compared to flows below 2000 vphpl at the working station just 180 m downstream. Flows as high as 3100 vphpl have been reported in the literature (e.g., Turochy, and Smith 2000), but such high flows are not sustained for long periods and in general, the maximum sustainable flow is expected to be in the range of 2500 vphpl. The shaded region of Figure 1C-D shows the periods when either the detector did not report any data or reported unusually high values of flow. It can be seen that by this criterion, more than half of the data reported from this station are erroneous on this day.

To identify if the detector is malfunctioning, several threshold value tests are applied to the data by comparing measurements against maximum and minimum acceptable values (Chen et al. 2003). For example, a test may consider q greater than 3000 vphpl as erroneous; if the frequency of such records is greater than an acceptable value then the detector is noted as malfunctioning and ignored by later steps in our speed estimation method. The thresholds for various tests described below can be set by looking at historical data and data from reliable detectors in that lane. For this paper, the threshold values were calculated using data for 5 weekdays at 10 single loop detector stations obtained via Performance Measurement System (PeMS), <http://pems.eecs.berkeley.edu/>, from the Main Line loops along 93 km (mile-marker 52 to 110) of SR-101 in the San Francisco Bay area of California. SR-101 was chosen because data for several consecutive stations and days were available and there were many detectors with noisy and erroneous data. The threshold values for the tests described below were established empirically such that only data from severely problematic detectors are discarded. These results were tested with data from other stations on SR-101 and nearby I-80 yielding similar results.

The following threshold value detector diagnostic tests are designed to be applied to each day of detector data to decide if the detector can be trusted on the following day. If a detector fails any of the threshold value tests then data from that detector are not used for speed estimation until it passes all of the tests.

<u>Test 1:</u>

> *Samples that show zero flow and zero occupancy:* During a day there are periods when samples show zero flow and occupancy. The frequency of zero flow-occupancy measurements is a function of time of the day and we can expect to see many such

samples during late night and early morning periods. But for the rest of the day, we do not expect to see many samples with zero flow and occupancy. By counting the total number of zero flow and occupancy samples in the period between 05:00 a.m. and 10:00 p.m. and comparing with a set threshold, we can identify whether the detector is stuck or not. The threshold value for this test is set at 50 percent of the samples, if exceeded, the detector is considered to be malfunctioning.

Test 2:

Samples that show zero flow and non-zero occupancy: A sensor reporting zero flow and non-zero occupancy over a 30-second period can be viewed potentially malfunctioning. However, one such sample cannot be deemed faulty. A vehicle is counted in q at a single instant, e.g., upon entering the detector, but may contribute to occ of more than one sample if it remains over the detector at the end of a sample. As a result, a vehicle stopped above the sensor for the entire 30-second period can result in zero flow and 100 percent occupancy, or a vehicle begins the 30-second period positioned above the sensor, then moves away from the sensor without another vehicle passing the sensor in the 30-second period, resulting in zero flow and nonzero occupancy. This second scenario tends to occur in very heavy traffic or in very light traffic. If the traffic is known to be light, then we do not expect occupancy to be greater then 3 percent as typically a free flow vehicle takes less than 1 second to cross the detector. If the number of samples between 5:00 a.m. and 10:00 p.m. with zero flow and non-zero occupancy exceeds a threshold value of 25 percent, then detector is considered to be malfunctioning. When the period of aggregation is increased to five minutes these events become far less likely and this type of validity check becomes more indicative for longer aggregation intervals.

Test 3:

Samples that show non-zero flow and zero occupancy: Any vehicle that passes the sensor must register a positive value of occupancy during the sample it contributes to flow. The threshold value for this test is set at 25 percent of the samples from the whole day. If during a day, the number of samples with zero occupancy and non-zero flow is greater than the threshold value, the detector is assumed to be malfunctioning.

Test 4:

Samples that show abnormally high flow or occupancy: Chattering and pulse break up at the detector can cause the flow in the sampling period to be abnormally high. Samples with flow as high as 3100 vphpl (25 vehicles per 30 seconds) were found in the test data from SR-101. Setting a threshold at a conservative 3100 vphpl (Turochy and Smith 2000), a sample above this value is considered suspicious. Once more, the number of samples that exceed this threshold is counted and if greater than 25 percent of the samples for the day, the detector is considered to be malfunctioning. However, if there are only a few such samples, then only those samples might be erroneous. A similar test with occupancy data can also be performed to check the validity of the data. An occupancy above 35 percent is considered high, arising from congestion. We assume that any freeway should be uncongested at least 60 percent of the time (van Zwet et al. 2003), thus if number of samples with occupancy exceeding 50 percent for a day, the detector is considered to be malfunctioning.

Test 5:

Samples that show constant occupancy and flow: Sometimes the field controller gets stuck and reports the same flow and occupancy for successive samples. This error can be identified by taking the difference of successive flow measurements and again for occupancy for entire day's data. The presence of a long series of zeroes (3 hours in our case) verifies that the detector is stuck and such data should be discarded.

MACROSCOPIC SPEED ESTIMATION AND FILTER RULES

After excluding all malfunctioning detectors based on the preceding tests, a clean estimate of vehicle speeds is obtained by following the Steps 1-6 below. As previously noted, the conventional speed estimation from Equation 1 is noisy. These raw speed estimates are refined on an individual lane basis by passing them through a series of filters, Steps 2-6, which use data from adjacent lanes and apply basic traffic flow theory principles to identify erroneous speed estimates. It is important to follow Steps 1 and 2 in order followed by Steps 3-5, not necessarily in the given order, and concluded with Step 6. The reason for this ordering is because the earlier

steps affect the thresholds in the subsequent steps, as will become more apparent in the following section.

Unless otherwise stated, the data for illustrating the filtering process were obtained from detector station 4 on I-80 westbound at Berkeley Highway Laboratory (BHL), California (Coifman et al. 2000). This station has five lanes in each direction and data were aggregated to 30-second samples such that each day's data had 2880 samples per lane. The data from PeMS were not used because of the poor quality of dual loop speed measurements, as mentioned in a white paper at PeMS and verified by our empirical analysis. Meanwhile, the data from SR 101 were used to develop thresholds in the previous section, Identifying Malfunctioning Detectors and Erroneous Data, as there are only eight dual-directional detector stations in the BHL.

Step 1:

Identify erroneous samples: In this step, erroneous data samples are identified and are either discarded or replaced by an estimate. The methodology to identify erroneous data samples is similar in principal to the detector diagnostic discussed in previous section. Samples with q = 0 and occ > 3 percent, or q > 0 and occ = 0 are suspect (refer to Test 2 and 3). Also successive samples with the exact same flow, occupancy, or both are suspect. The suspect samples for individual lanes are validated by looking at the samples in their "neighborhood"; where the neighborhood consists of adjacent lanes during the same sample period and the three most recent valid measurements in the subject lane within previous five minutes, if they exist. If the suspect sample is consistent with the traffic state indicated by neighboring samples, then it is retained as a valid sample. The remaining suspect samples are discarded at this stage so that the errors introduced by imputing bad samples do not affect estimated speeds. If one wishes to eliminate holes in the data due to presence of erroneous samples, they could be filled by an estimate from the average q and occ in the adjacent lanes. If the samples in the adjacent lanes are not reliable then the median of the samples from the past 3 valid measurements can be used as an estimate. If the last three valid samples in the subject lane are older than 10 samples, then the estimation can be done by combining data from the valid samples from all lanes over the last five minutes. This data estimation methodology is applicable to

short-term malfunctions only. If a malfunction persists for several minutes, the traffic measurements may no longer apply to the current traffic situation, resulting in inaccurate estimates. Furthermore, this mechanism is applicable only to locations where individual lane data are available and there are at least two reliable lanes. If this condition is no longer valid then the estimate might be as poor as the erroneous measurement and in such case we would be better off simply flagging the sample as invalid.

The scatter plot in Figure 2A compares $\hat{V}$, for an individual lane, at the end of Step 1 plotted against measured speed. The data were collected on August 12, 1998. Notice the considerable amount of scatter.

Step 2:

Raw estimate of speed: First a target period of the day is chosen which is defined as a period when the traffic is expected to be free flowing and the flow far from zero. When the traffic is free flowing, flow and occupancy measurements are expected to be in phase, i.e., increase or decrease together through small fluctuations. The correlation among q and occ measured in successive 5-minute samples in the target period is calculated and if most of these calculations are positive, then q and occ are in phase verifying free-flow. Usually, the period between the morning and evening peaks or after the evening peak is a good choice for the target period. A ratio of measured flow over occupancy is taken. Ordinarily, this ratio would be multiplied by L to estimate $\hat{V}$ and we initially set L=6.1 m. We then estimate the "raw speed" using Equation 1. A correction factor is then calculated by taking the ratio of the median of raw speed during the free-flowing target period and the assumed free-flow speed for that station and lane. This assumed free-flow speed might come from the posted speed limit, as in this chapter, or direct measurement, e.g., from a radar gun. More precise measurements of free-flow speed are desirable but that would require more resources for an operating agency and the assumption is similar to what is being practiced with existing single loop detector data. The correction factors are calculated for all lanes at every station and in subsequent steps, the raw speed estimates are multiplied by corresponding correction factor.

If non free-flow conditions were present during the target period then the computed correction factor would be erroneous. The median of the estimated speeds after applying the correction factor to entire day's data should be close to the assumed free-flow speed thereby validating the correction factor and the presence of free-flow conditions in the target period. When the phase data indicate congested conditions, the correction factors can be computed by using recent historical data from last few days. Taking the moving median of the recent correction factors can eliminate error due to undetected congestion and reduce fluctuations due to daily changes in the traffic patterns. Using data from 15 consecutive days, over the period between 19:00 and 23:00 the daily correction factor for each lane was calculated and as expected, there was little variation from day to day (Figure 3). In this manner, these factors can be developed for every station and help correct for any error due to sensitivity of the detector. The scatter plot in Figure 2B shows $\hat{V}$ at the end of Step 2.

Step 3:

Speed-Flow filter: A typical curve representing the relationship between speed and flow is shown in the right half of Figure 4. Using this relation one can define threshold values for free flow speed corresponding to different flows. Flow-Speed pairs that are outside these threshold boundaries either fall in an infeasible region or are from congested conditions. The left hand side of Figure 4 shows the monotonically decreasing relationship between speed and occupancy. The exact shapes of these curves are not important; we will develop our thresholds to be loose enough to accommodate any reasonable approximations. In Figure 4, we do not expect a flow-speed pair to fall in shaded region 'A' unless it is from congested conditions. But such measurements should have high occupancy if valid. Given the maximum speed (V_{max}) of region 'A', the valid data belonging to region 'A' must correspond to occupancy higher than threshold (Occ_{min}) on left of region 'B'. If a measurement falls in both region 'A' and 'B' then it is erroneous. For our test data, threshold values were established such that a sample would belong to an infeasible region if q < 1000 vphpl, occ < 15 percent and $\hat{V}$ < 80 km/h. When an estimate fails this test, it is compared to the sample neighborhood and if it is similar to the other estimates it is retained; otherwise it is flagged as erroneous and re-estimated as

follows. The speed is estimated by taking the median of the speeds in the adjacent lanes. If the samples in the adjacent lanes are not reliable then the median of the samples from the past 3 valid measurements can be used as an estimate. If the last three valid samples in the subject lane are older than 10 samples, then the estimation can be done by combining data from the valid samples from all lanes over the last five minutes. If valid data are not available then the sample is marked as erroneous and no estimation is made. In this process, care was taken to use the median rather than mean to reduce sensitivity to outliers. The improvement in speed estimates at the end of this step is shown in Figure 2C.

Step 4:

Speed-Occupancy filter: Again, using the argument presented in Step 3 and the relation between flow, occupancy and speed, one can define the threshold values for speed corresponding to different flow-occupancy values. As has been observed, both in our research and by other researchers (e.g., Coifman 2001), capacity flow corresponds to occupancy somewhere in the range of 10-25 percent and below this free-flow conditions exist. We set a conservative threshold for free-flow occupancy at 8 percent, thus samples below this threshold and $\hat{V}$ < 80 km/h are suspicious, e.g., region 'B' in Figure 4. These suspect estimates are compared to their neighborhood samples to verify they are not feasible. Also, the current traffic state (free-flowing or congested) is verified using the phase information between flow and occupancy measurements of the previous few samples. If the suspect sample's speed is found to be inconsistent with the current traffic state then the speed for the sample is re-estimated using the process outlined in Step 3. Now, however, we do not expect an occupancy-speed pair to fall in region 'B'. In this case, the measurements with $\hat{V}$ < 80 km/h, occ < 8 percent and q < 840 vphpl are replaced with the assumed local free-flow speed. The scatter in $\hat{V}$ at the end of this step is shown in Figure 2D.

Step 5:

Speed filter: In this step, samples with $\hat{V}$ > 144 km/h are identified as erroneous since we do not expect the speeds to be this high. This error could be either due to transient error from a short on time or due to the overestimation of L. If most of the vehicles in the sample are truly shorter than assumed L, then we will overestimate the speed. The phase information and information from adjacent lanes are used to determine if traffic is free-flowing or congested. If the traffic is free flowing then these samples are scaled down to the assumed free-flow speed, otherwise these samples are discarded altogether. Figure 2E shows $\hat{V}$ at the end of Step 5.

Step 6:

Moving median of three samples: Equation 1 is sensitive to small variations in q and occ. This sensitivity is inversely proportional to the number of vehicles observed. When only a few vehicles pass the detector in a given time interval, i.e., light traffic, the true average vehicle length is more likely to differ substantially from L, and $\hat{V}$ will be noisy (Coifman 2001). The errors in speed estimation might not be eliminated if the sample falls in the feasible regions of the previous steps. In order to reduce error from high variance in L we take a moving median of three samples. However, we restrict the moving median only to estimates with $\hat{V}$ >40 km/h since the true speed can change significantly from one sample to the next below this threshold while L typically exhibits smaller variance during lower speeds due to the higher proportion of passenger vehicles. The speed estimates at the end of filtering Step 6 are shown in Figure 2F.

Finally, a single accurate estimate of speed per direction at a detector station is obtained by taking the median of the improved individual lanes' speed estimates. The lane-by-lane $\hat{V}$ after filtering is shown in Figure 5A-E and median $\hat{V}$ across all lanes is shown in Figure 5F.

ANALYSIS AND DISCUSSION

Figure 6A-F shows several scatter plots of different estimates of speed versus measured speed for a day of data from BHL station 4. The left column shows conventional estimates from Equation 1 while the right column shows the new estimates from the previous section, Macroscopic Speed Estimation and Filter Rules. This figure shows (A-B) estimates from a single lane, (C-D) mean across all lanes and (E-F) median across all lanes. The mean and median come from those lanes that were not eliminated by the filters. In each row, the circles are placed at the same location in the plot pair to highlight the difference between the estimations. If the estimations were perfect a given plot would show all the points on the diagonal line inclined at 45-degrees. In all three cases the speed estimation using conventional method shows greater deviation from the diagonal line then the new methodology. The estimation for individual lanes is not as good as that across lanes, which is expected because there are fewer vehicles in the individual lane samples. This drop in quality can be observed by comparing the scatter around the diagonal line in plots B, D and F.

Figure 7 compares the time series of estimated speeds after Step 2 and those after Step 6. As can be observed from the plot, there is high scatter in the raw speeds estimated by Equation 1, especially during early morning periods. For the same data, when the speeds are passed through the filters developed above, the scatter is reduced significantly without degrading the estimates. Figure 7A compares speed for an individual lane, Figure 7B compare the mean speed across all lanes and Figure 7C compares median speed across all lanes. Observe that the mean and median across all lanes provides a cleaner estimate than a single lane for both the raw and processed data.

It is important to ensure that the speed estimates after filtering are representative of the traffic state. The estimates should not indicate congested conditions when it is actually free-flow and vice versa. Though we do not expect the estimates to perfectly match the true measurements, they should be as close as possible. To check the quality of estimates, the number of estimates with $\hat{V} > 80$ km/h and measured speed less than 56 km/h (Type-1 error) and $\hat{V} < 48$ km/h and measured speed greater than 80 km/h (Type-2 error) per day were counted. The quality of the estimates is good if the number of estimates with each error type is zero or close to zero. One

month of data from BHL station 4 were used. Figure 8A shows the cumulative distribution function (cdf) for the median error of each error type across the lanes. As can be observed from the plot, the number of samples with error type-1 is zero for almost 90 percent of the time and few samples (15 out of 2880) exhibit error type-2, which likely is acceptable for most applications.

A considerable amount of error in raw speed estimates is simply due to bias in L. The bias in L was removed by calculating the average representative vehicle length for every two hours using measured speed. Unbiased speeds were estimated using Equation 1 with new average vehicle lengths, representing the best possible performance from conventional practice. Figure 8B shows the cdf of absolute error in speed estimation for an individual lane data, Figure 8C for the median across lanes and Figure 8D for the mean across all lanes. Each plot shows the cdf of error for the cleaned estimate from the filters, raw unbiased estimates and raw estimates. Notice that simply by removing the bias, we get significant improvement in the individual lane speed estimates and similar performance improvements when combining data across lanes, though the proportional improvement is smaller than the earlier improvements in individual lanes. For the raw estimates, the median across all lanes performs better than mean, which can be observed by comparing Figure 8C-D. Also the plots in Figure 6C-F and Figure 7B-C highlight that median is less sensitive to outliers.

As observed by comparing Figure 8B with Figure 8C-D, the absolute error decreases by combining data across lanes. In all cases, the error after cleaning the data is smaller than unbiased speed estimates error. The mean absolute error of cleaned speeds is less than less than 11 km/h for individual lanes and is less than 5 km/h when data are combined across lanes.

The algorithm developed in the section on Macroscopic Speed Estimation and Filter Rules uses single loop detector data to estimate speed, but it can also be extended and applied to speeds measured from dual loop detectors that are noisy. For cleaning noisy speed measurements, only the part of the Step 2, which estimates raw speed using Equation 1, needs to be omitted and the rest of the algorithm can be applied without change. For dual loop detectors, there is no need to correct for error in L as speed is calculated by measuring the time taken to travel between the

two loop-detectors. In this case, the correction factors from Step 2 correct for any error in detector spacing (which is constant, though potentially unknown).

Analysis of RTMS data:

This algorithm was also tested with data obtained from an RTMS in side fire mode. The RTMS is functionally very similar to a single loop detector station since the unit measures flow and occupancy and then estimates speed. The RTMS incorporates the functionality of the traffic controller, but it also provides contact closure outputs that can be input to a conventional traffic controller (henceforth referred to RTMS-via-controller). For this study, concurrent RTMS, RTMS-via-controller and dual loop detector data were collected on May 26, 2000 at eastbound BHL station 7 and aggregated to 30-second sample.

Figure 9 shows several scatter plots comparing median estimated and measured speed across all lanes. The left hand column shows the raw estimates and the right hand column shows the corresponding cleaned estimates after applying the filters. The first row, A-B, show the speed estimates reported by the RTMS's internal controller. The second row, C-D, show the speed estimates after applying Equation 1 to the measured q and occ reported directly by the RTMS. The third row, E-F, show the speed estimates from Equation 1 applied to q and occ for RTMS-via-controller (the downstream loop detector was within 6 m of the RTMS detection zone in each lane). For reference, the fourth row shows the speed estimates from the loop q and occ. Quantifying these errors, Figure 10 shows the distribution of absolute and percentage error for each of the rows from Figure 9. Note that our filtering improves the performance in all four cases and similar results were observed on an individual lane basis, as shown in Table 1. Among the raw data, the direct RTMS estimates (Figure 9A) show the best performance because the RTMS applies its own filters to the estimation process while the estimates from q and occ are worst for the direct RTMS measurements (Figure 9C) because it only reports integer occupancy. Except for these rounding errors, Figure 10 shows that after applying the filter from this paper, the other three estimates exhibited similar performance (Figure 10A, E, G).

SPACE-MEAN SPEED

Although it is common to take the arithmetic mean of speed across lanes (potentially weighted by flow in each lane), such an approach does not preserve the space-mean speed. In fact, it can easily be shown that,

$$\overline{V} = \frac{n}{\sum_{i=1}^{n} \frac{1}{v_i}} = \frac{q_A}{\sum_{j=1}^{m} \frac{q_j}{\overline{V}_j}} \tag{2}$$

where,

$\overline{V}$ = space-mean speed across all lanes over the sample of duration T,

n = total number of vehicles that pass in all lanes over T,

v_i = the measured speed of the i-th vehicle,

q_A = total flow across all lanes,

m = total number of lanes at the detector station,

q_j = flow in j-th lane,

$\overline{V}_j$ = space mean speed in the j-th lane over the sample of duration T.

This distinction is important because much of traffic flow theory is built upon the fundamental equation,

$$q_A = k_A \overline{V} \tag{3}$$

where,

k_A = total density across all lanes.

The methodology presented in this paper does not preserve the space-mean speed across all lanes since the last step of the proposed estimation process is,

$$\hat{V} = median\left(\hat{V}_j\right) \tag{4}$$

where,

$\hat{V}_j$ = cleaned estimate of space mean speed in the j-th lane,

and the median operation will not generally be equivalent to the inverse of the flow-weighted harmonic mean of Equation 2. Similarly, the "measured speeds" across all lanes presented in the figures and the table are the median or arithmetic mean (as indicated) of the individual lanes' measured speeds.

The median is taken for practical reasons since even with the proposed cleaning and filtering, in each lane $\hat{V}_j$ is still noisy. This noise is evident between 10:00 and 15:00 in the "clean" estimate shown in Figure 7A. One would expect this free flow period to be characterized by relatively stable speeds around 96 km/h, but the plot clearly shows transient errors dropping down as low as 48 km/h during this period. This error is due largely to long vehicles skewing the estimate from equation 1, though noted earlier, it is clear from Figures 6A and 7A that the error within a single lane is much smaller after the proposed cleaning. Comparing the data highlighted with the large oval in Figures 6B and 6F, one can see that after taking the median across lanes the measured and estimated values tend to be much closer.

CONCLUSIONS

Many researchers have sought better estimates of speed from single loop detectors. The earlier works have emphasized statistical techniques to reduce the bias from long vehicles in mean speed estimates on a lane-by-lane basis. This paper has taken a different approach, it applies threshold value tests and traffic flow theory principles to identify erroneous speed estimates. These erroneous estimates are re-estimated using redundant information from adjacent lanes and recent history. These individual lane estimates were shown to have improved by as much as 70 percent as compared to unbiased conventional estimates and by as much as 125 percent as compared to conventional raw estimates. It was also shown that the directional estimate of speed obtained by averaging the improved individual lane estimates across all lanes is more accurate than conventional estimates.

In the course of developing the algorithm, data validation and detector diagnostic tools were developed to eliminate malfunctioning detectors and transient errors. These data validation tools can help improve the data quality, which in turn would improve speed estimates and decision-making process by agencies. The methodology was shown to be less sensitive to site-specific

characteristics of the traffic flow and also applicable to noisy traffic data collected by other methods.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contributions from the anonymous reviewers that have improved the quality of this paper, particularly in the section discussing space-mean speed.

This material is based upon work supported in part by the National Science Foundation under Grant No. 0133278 and by the California PATH (Partners for Advanced Highways and Transit) Program of the University of California, in cooperation with the State of California Business, Transportation and Housing Agency, Department of Transportation.

The Contents of this report reflect the views of the author who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification or regulation.

REFERENCES

Chen, C., Kwon, J., Skabardonis, A., and Varaiya, P. (2003). "Detecting Errors and Imputing Missing Data for Single Loop Surveillance Systems," TRB, *Transportation Research Record 1855*, pp 160-167.

Coifman, B. (2001). "Improved Velocity Estimation Using Single Loop Detectors," *Transportation Research: Part A*, 35(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*, 11(3-4), pp 211-222.

Coifman, B. and Ergueta, E. (2003). "Improved Vehicle Reidentification and Travel Time Measurement on Congested Freeways," *ASCE Journal of Transportation Engineering*, 129(5), pp 475-483.

Coifman, B., Lyddy, D., and Skabardonis, A. (2000). "The Berkeley Highway Laboratory-Building on the I-880 Field Experiment", *Proc. Third Intelligent Transportation Systems Conference*, Detroit, MI, USA, IEEE, pp 5-10.

Coifman, B. (in review). "Freeway Detector Assessment- Part I, Vehicle Level Evaluation of Loop Detectors and the RTMS," submitted for publication in *Proc. of the 84th Annual Meeting of the Transportation Research Board*.

Dailey, D. (1999). "A Statistical Algorithm for Estimating Speed from Single Loop Volume and Occupancy Measurements," *Transportation Research: Part B*, 33(5), pp 313-322.

Jia, Z., Chen, C., Coifman, B. and Varaiya, P. (2001). "The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors," *Proc. Fourth Intelligent Transportation Systems Conference*, Oakland, CA, USA, IEEE, pp 536-541.

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

Mikhalkin, B., Payne, H. J., and Isaksen, L. (1972). "Estimation of speed from presence detectors," *Highway Research Record. 388*, Highway Research Board, pp 73-83.

Turochy, R. and Smith, B. (2000). "A new procedure for Detector Data Screening In Traffic Management Systems", *Transportation Research Record 1727*, pp 127-131.

Van Zwet, E., Chen, C., Jia, Z. and Kwon, J. (2003). "A statistical method for estimating speed from single loop detectors," Freeway Performance Measurement System (PeMS), <http://pems.eecs.berkeley.edu/>, March, 2003.

Wang, Y. and Nihan, N. (2000). "Freeway Traffic Speed Estimation with Single-Loop Outputs," TRB, *Transportation Research Record 1727*, pp 120-126.

FIGURES AND TABLES

List of Tables

List of Figures

Figure 6, This figure shows (A-B) estimates from a single lane, (C-D) mean across all lanes and (E-F) median across all lanes. The left column shows conventional estimates from Equation 1 while the right column shows the new estimates from the methodology developed in this paper. In each row, the circles are placed at the same location in the plot pair to highlight the difference between the estimations. Observe that mean across lanes shows greater scatter than median across lanes for raw estimates (C and E). After cleaning the estimates using the filters, both mean and median across lanes show similar performance.

Figure 7, Time series of speeds after Step 6 (Clean) plotted over the speeds at the end of Step 2 (Raw). (A) Speeds for Lane 4, (B) mean speeds across lanes and (C) median speed across lanes. The circles are placed to illustrate the effect of outlier in a lane. The mean value is more sensitive to outliers than median

Figure 8, Cumulative Distribution Function (CDF) of (A) number of erroneous samples out of 2880 30-second samples for error types 1-2. (B) CDF of absolute mean error for Lane 4, (C) absolute mean error for median of speeds across all lanes and (D) absolute mean error for mean of speeds across all lanes for Raw and Unbiased estimates. For reference, (D) shows CDF of errors for median across lanes for clean speeds.

Figure 9, Scatter plots comparing median estimated and measured speed across all lanes. The left hand column shows the raw estimates and the right hand column shows the corresponding cleaned estimates after applying the filters. The first row, A-B, shows the speed estimates reported by the RTMS. The second row, C-D, shows the speed estimates after applying Equation 1 to the measured q and occ reported directly by the RTMS. The third row, E-F, shows the speed estimates from Equation 1 applied to q and occ for RTMS-via-controller. The fourth row shows the speed estimates from the loop q and occ.

Figure 10, Distribution of absolute and percentage error for median estimated and measured speed across all lanes. The left hand column shows the distribution of absolute mean error and the right hand column shows the distribution for absolute

percentage error. The first row, A-B, shows the speed estimates reported by the RTMS. The second row, C-D, shows the speed estimates after applying Equation 1 to the measured q and occ reported directly by the RTMS. The third row, E-F, shows the speed estimates from Equation 1 applied to q and occ for RTMS-via-controller. The fourth row shows the speed estimates from the loop q and occ.

Table 1, Absolute mean error and absolute percentage mean error for RTMS estimates.

	Lane 1		**Lane 2**		**Lane 3**		**Lane 4**		**Lane 5**		**All lanes combined**	
	Abs Error (km/h)	Abs % Error	Abs Error (km/h)	Abs % Error	Abs Error (km/h)	Abs % Error	Abs Error (km/h)	Abs % Error	Abs Error (km/h)	Abs % Error	Abs Error (km/h)	Mean Abs % Error
RTMS Measured	6.2	18.7	3.8	15.8	3.5	16.2	2.5	10.1	4.0	14.5	2.4	6.3
RTMS Measured Clean	6.4	16.5	2.8	11.7	2.1	11.2	2.0	8.8	3.2	12.5	1.6	3.4
RTMS (q, occ) Raw	14.5	39.3	7.8	23.4	6.6	22.3	7.1	21.5	20.3	55.9	6.2	15.0
RTMS (q, occ) Clean	7.5	18.4	5.2	16.1	4.7	18.2	5.1	19.0	5.4	16.4	3.9	11.4
RTMS-via-controller (q, occ) Raw	9.7	24.3	8.2	25.1	10.3	32.6	9.2	27.2	4.8	15.0	8.0	22.8
RTMS-via-controller (q, occ) Clean	5.9	13.8	3.0	9.6	3.4	14.7	4.0	14.4	3.5	11.8	2.1	5.5
d/s Loop (q, occ) Raw	3.4	11.1	1.3	5.5	4.9	16.3	6.4	19.2	7.6	22.1	4.0	10.0
d/s Loop (q, occ) Clean	1.9	6.0	1.2	5.1	3.5	12.2	3.7	11.4	3.0	8.3	1.7	3.9

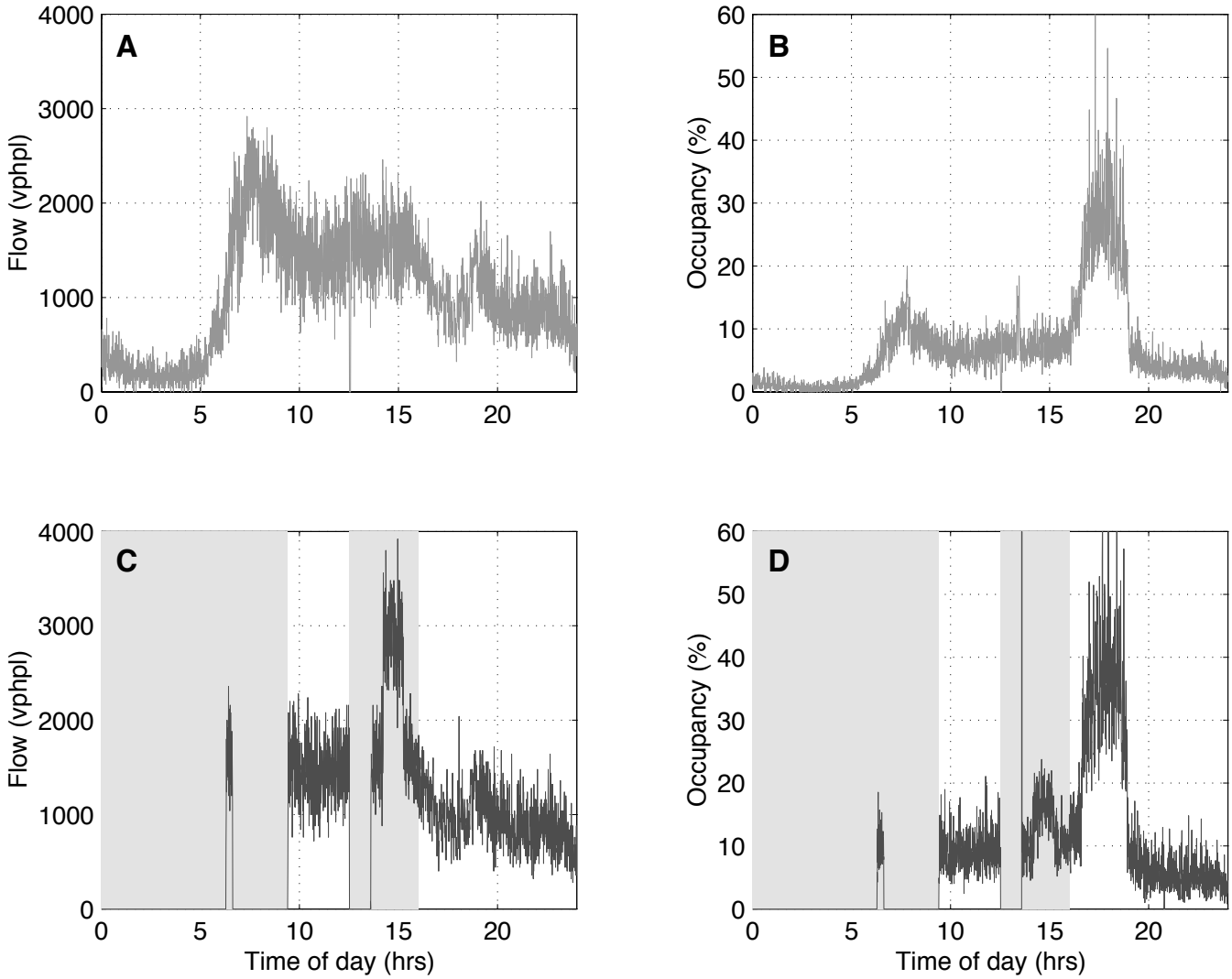

Figure 1, Stations exhibiting typical (A & B) and abnormal (C & D) flow and occupancy measurements. The plots show average flow and occupancy data across southbound lanes from two adjacent stations (St7: A & B, St8: C & D) at I-71, Columbus Ohio on 4th October 2002. The shaded region in plots C & D shows the period when either the detector did not report any data or reported unusually high values of flow.

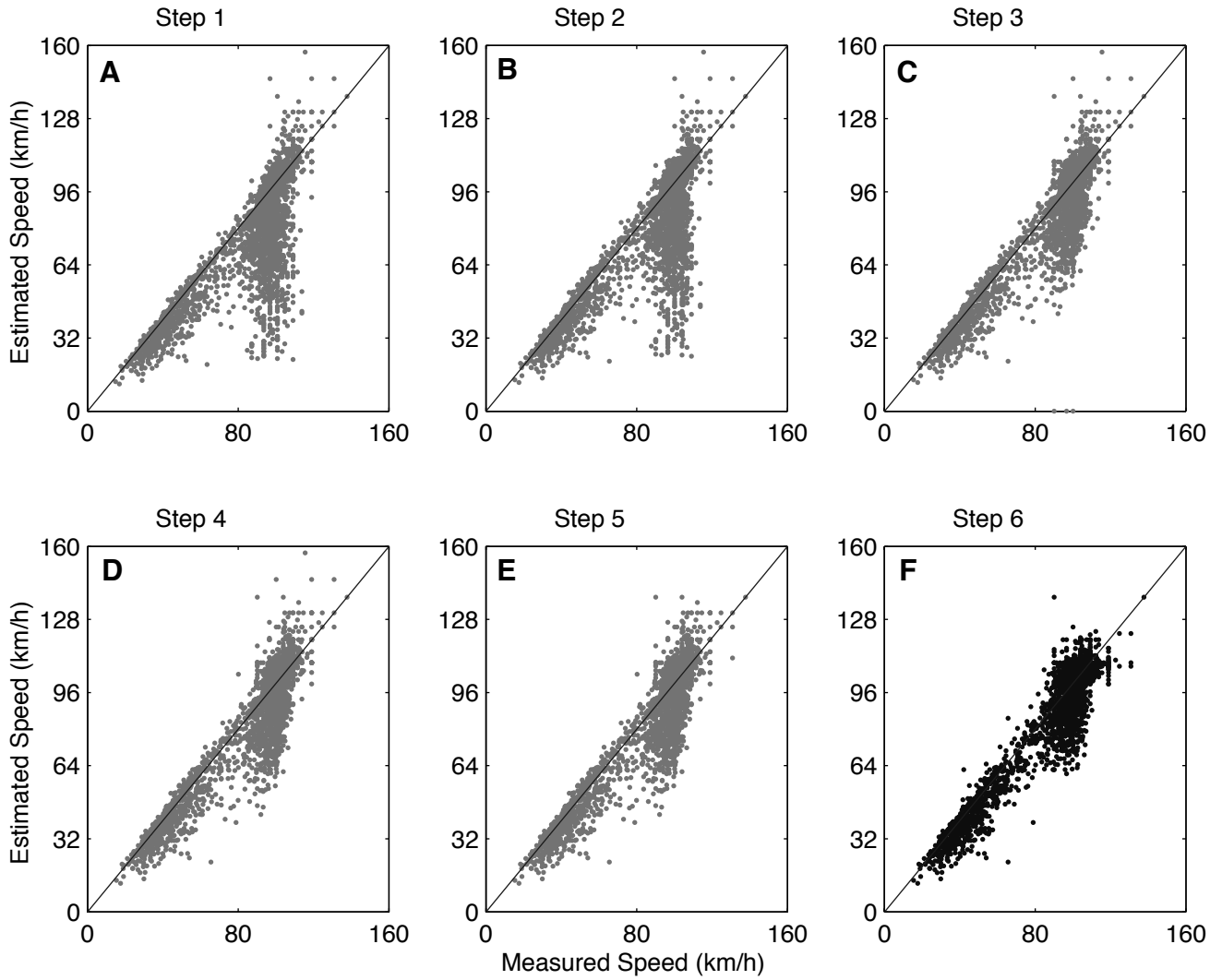

Figure 2, Plot showing progression of improvement in speed estimates for lane 4 on BHL station 4 westbound on 12th August 1998. The plot A shows the estimates at the end of Step 1, B shows the estimates at the end of Step 2, C shows the estimates at the end of Step 3, D shows the estimates at end of Step 4, E at the end of Step 5 and F at the end of Step 6.

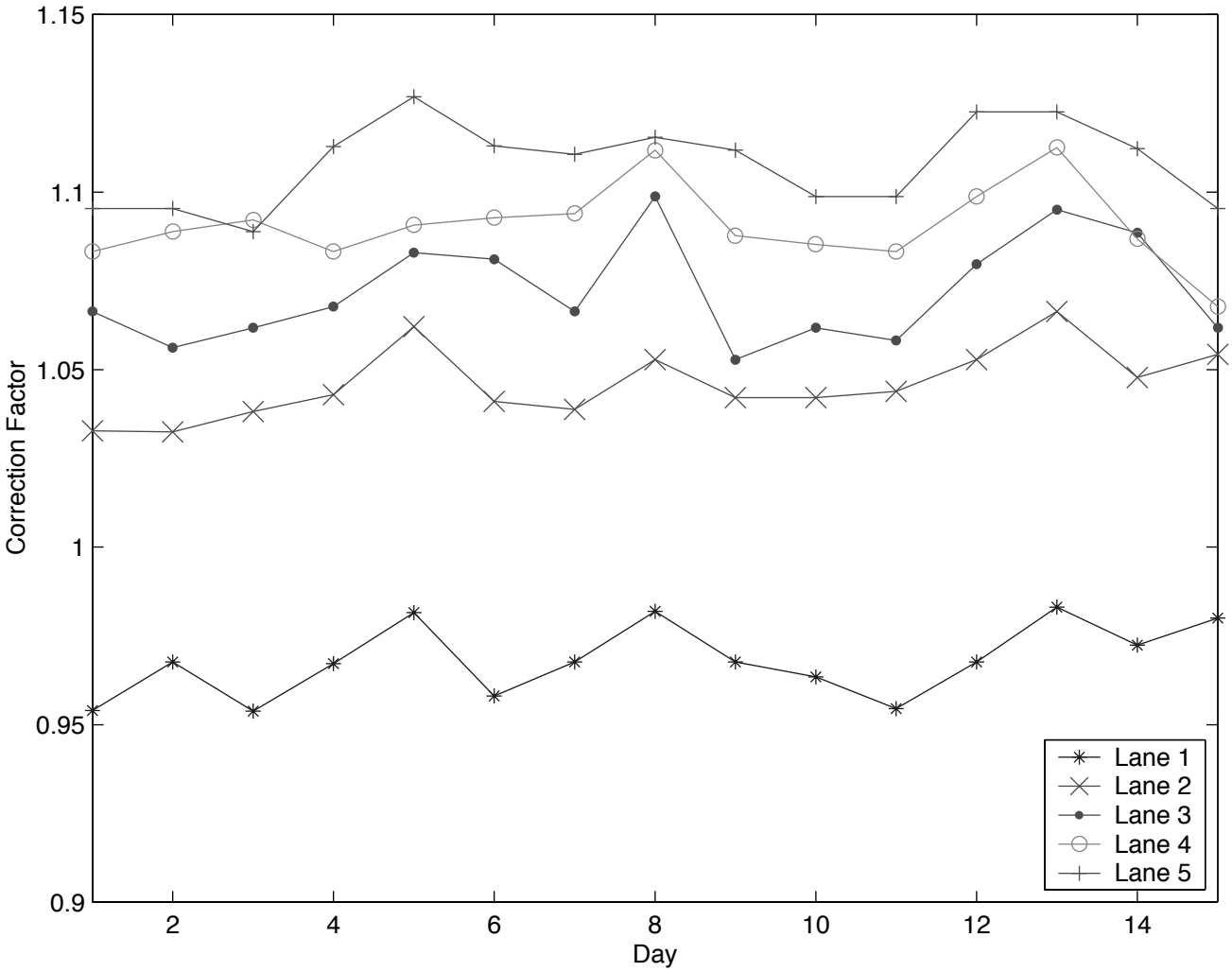

Figure 3, Variation in correction factors across days for 15 days data collected at BHL station 4 westbound.

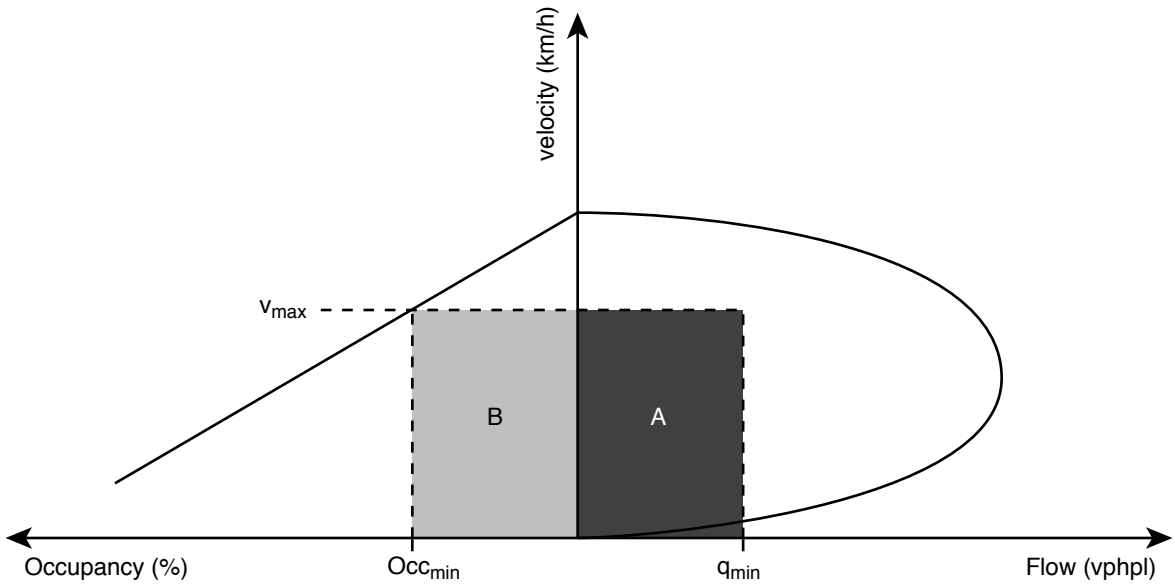

Figure 4, Plot showing relationship between flow-speed and occupancy-speed

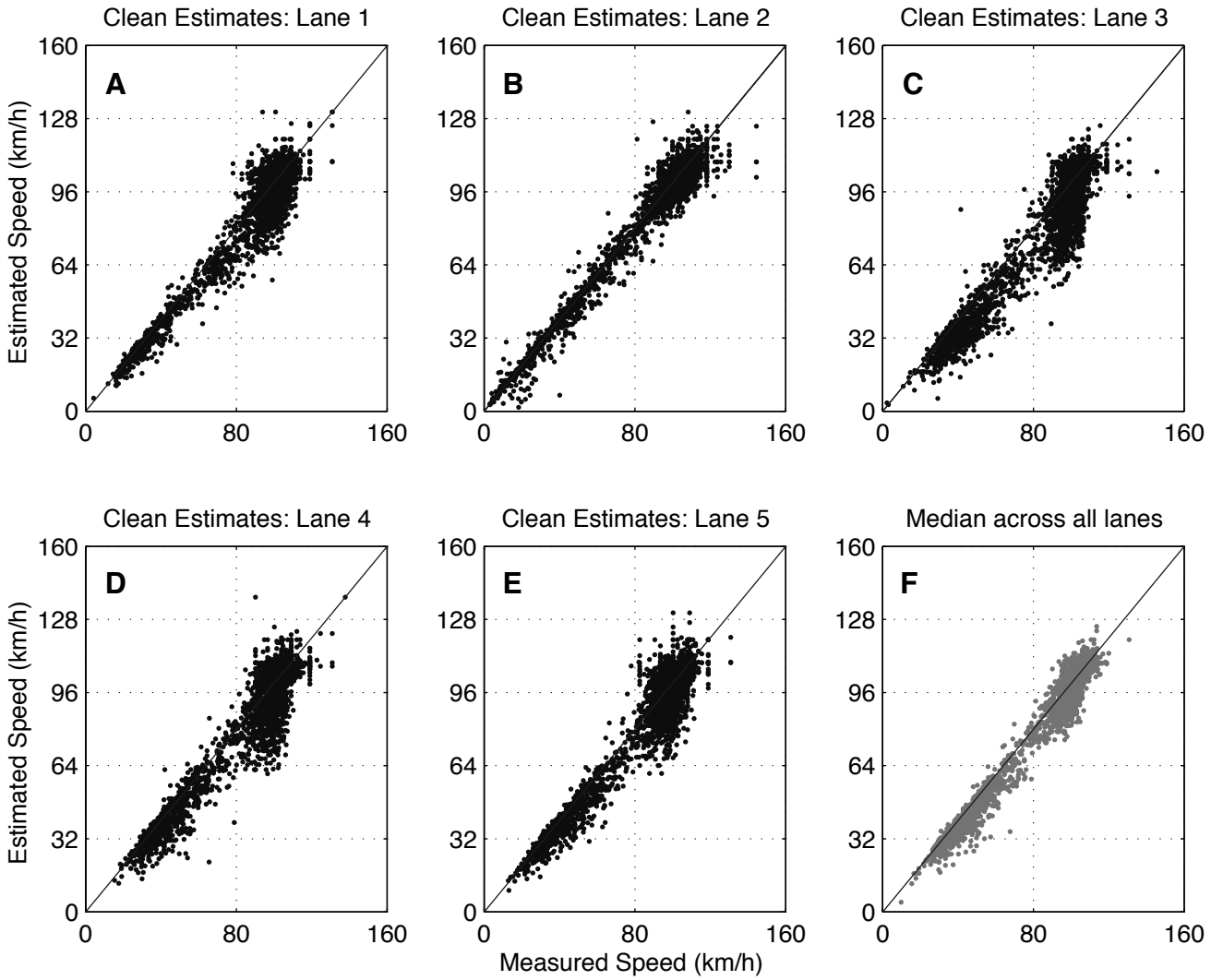

Figure 5, Plot showing improved speed estimates at the end of filtering for individual lanes (A-E: lanes 1-5) and improved median estimate across all lanes (F). Note that lane 1 is restricted to High Occupancy Vehicles (HOV) and we do not use adjacent lanes data to estimate speeds for this lane because its time-series is different from the other lanes.

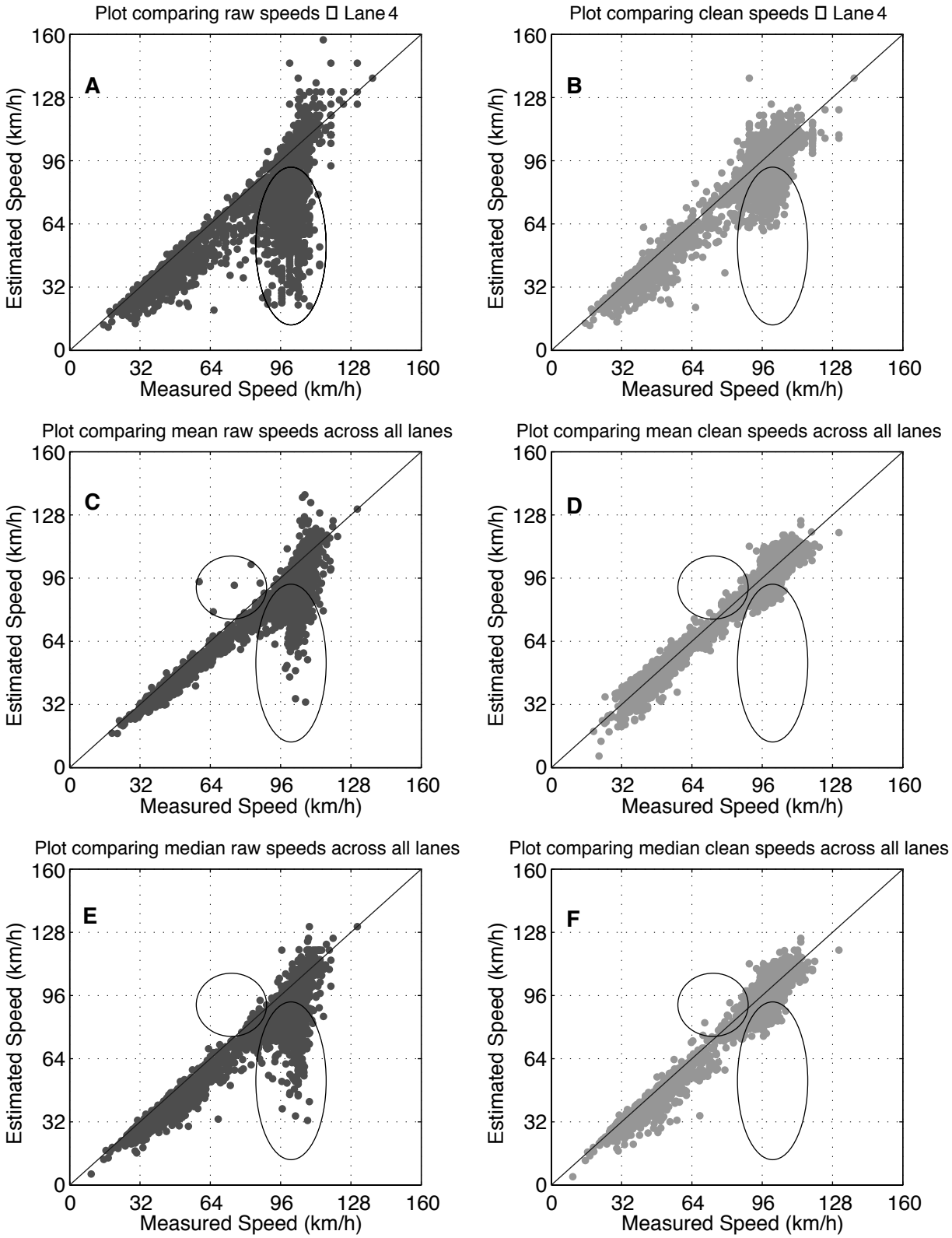

Figure 6, This figure shows (A-B) estimates from a single lane, (C-D) mean across all lanes and (E-F) median across all lanes. The left column shows conventional estimates from Equation 1 while the right column shows the new estimates from the methodology developed in this paper. In each row, the circles are placed at the same location in the plot pair to highlight the difference between the estimations. Observe that mean across lanes shows greater scatter than median across lanes for raw estimates (C and E). After cleaning the estimates using the filters, both mean and median across lanes show similar performance.

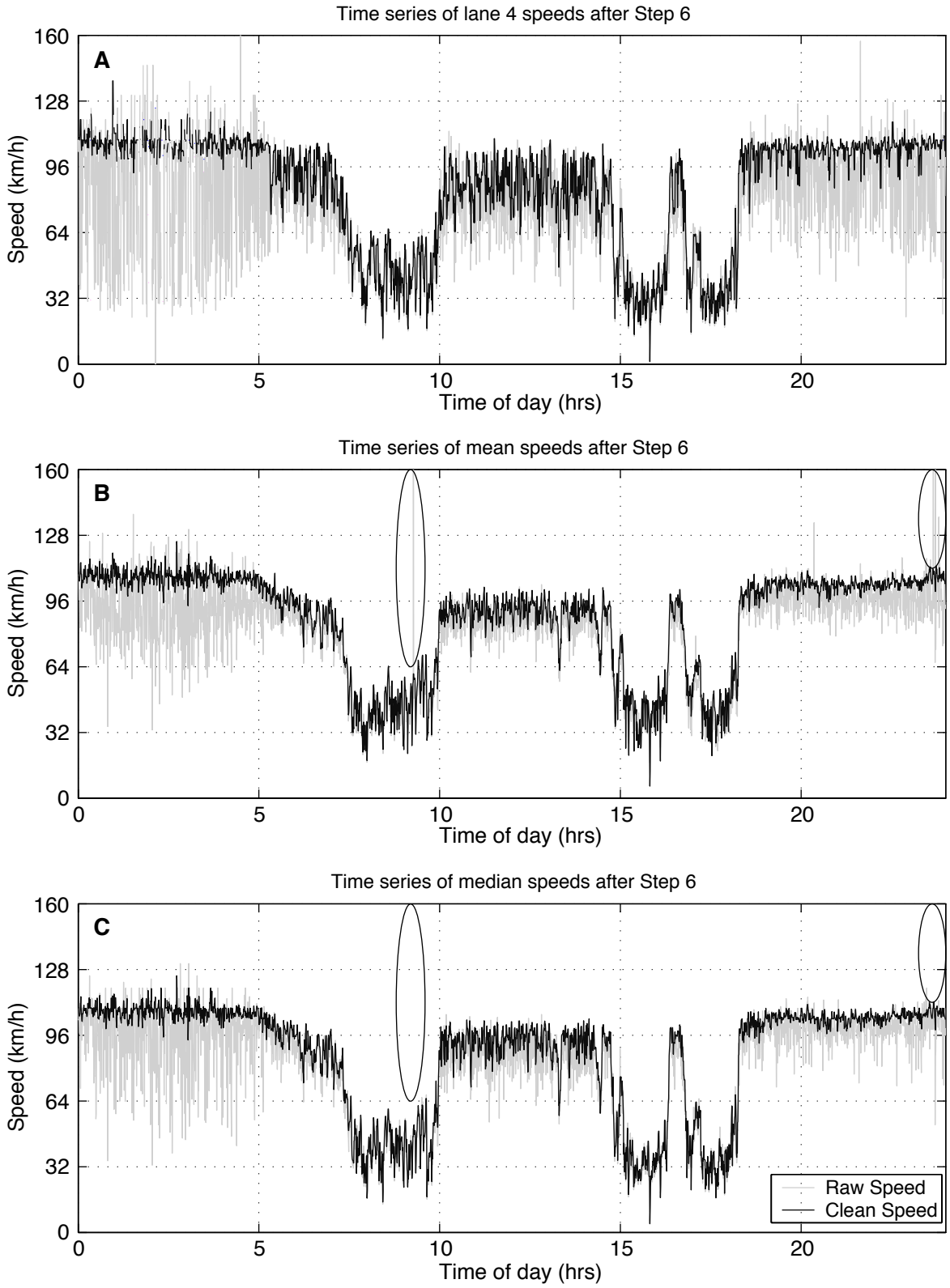

Figure 7, Time series of speeds after Step 6 (Clean) plotted over the speeds at the end of Step 2 (Raw). (A) Speeds for Lane 4, (B) mean speeds across lanes and (C) median speed across lanes. The circles are placed to illustrate the effect of outlier in a lane. The mean value is more sensitive to outliers than median

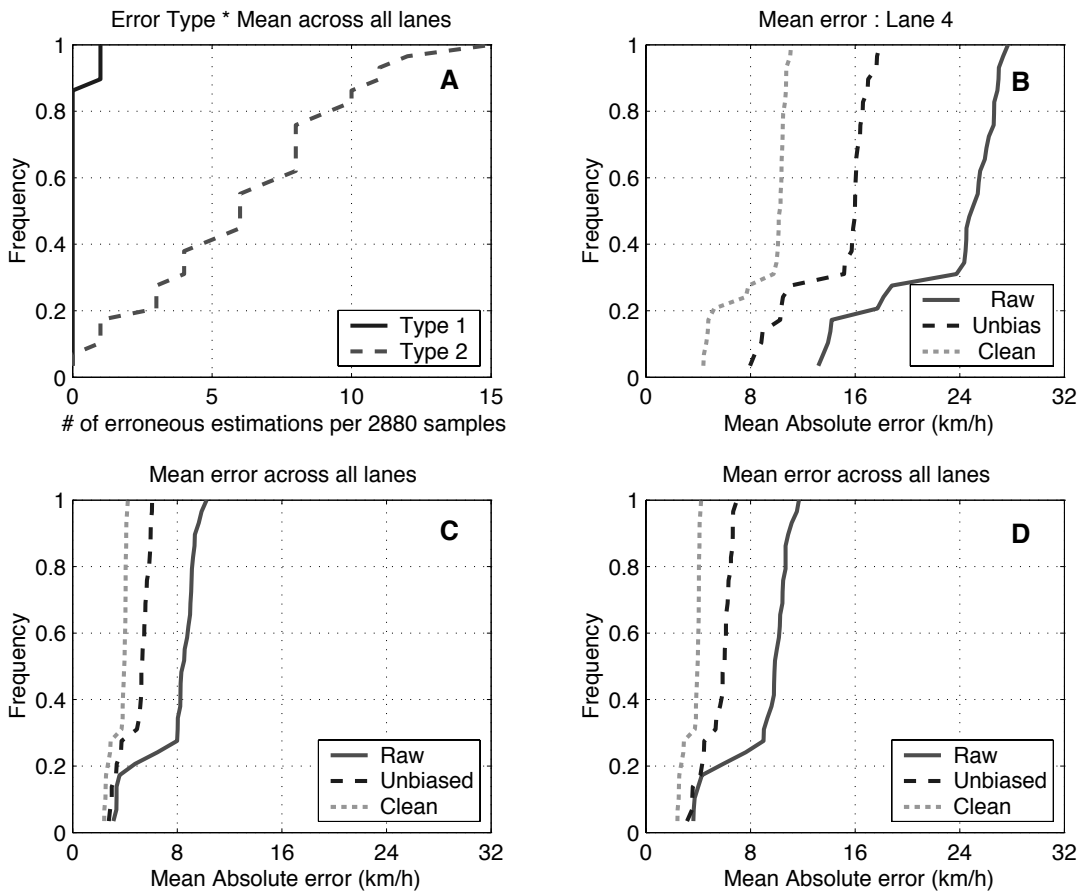

Figure 8, Cumulative Distribution Function (CDF) of (A) number of erroneous samples out of 2880 30-second samples for error types 1-2. (B) CDF of absolute mean error for Lane 4, (C) absolute mean error for median of speeds across all lanes and (D) absolute mean error for mean of speeds across all lanes for Raw and Unbiased estimates. For reference, (D) shows CDF of errors for median across lanes for clean speeds.

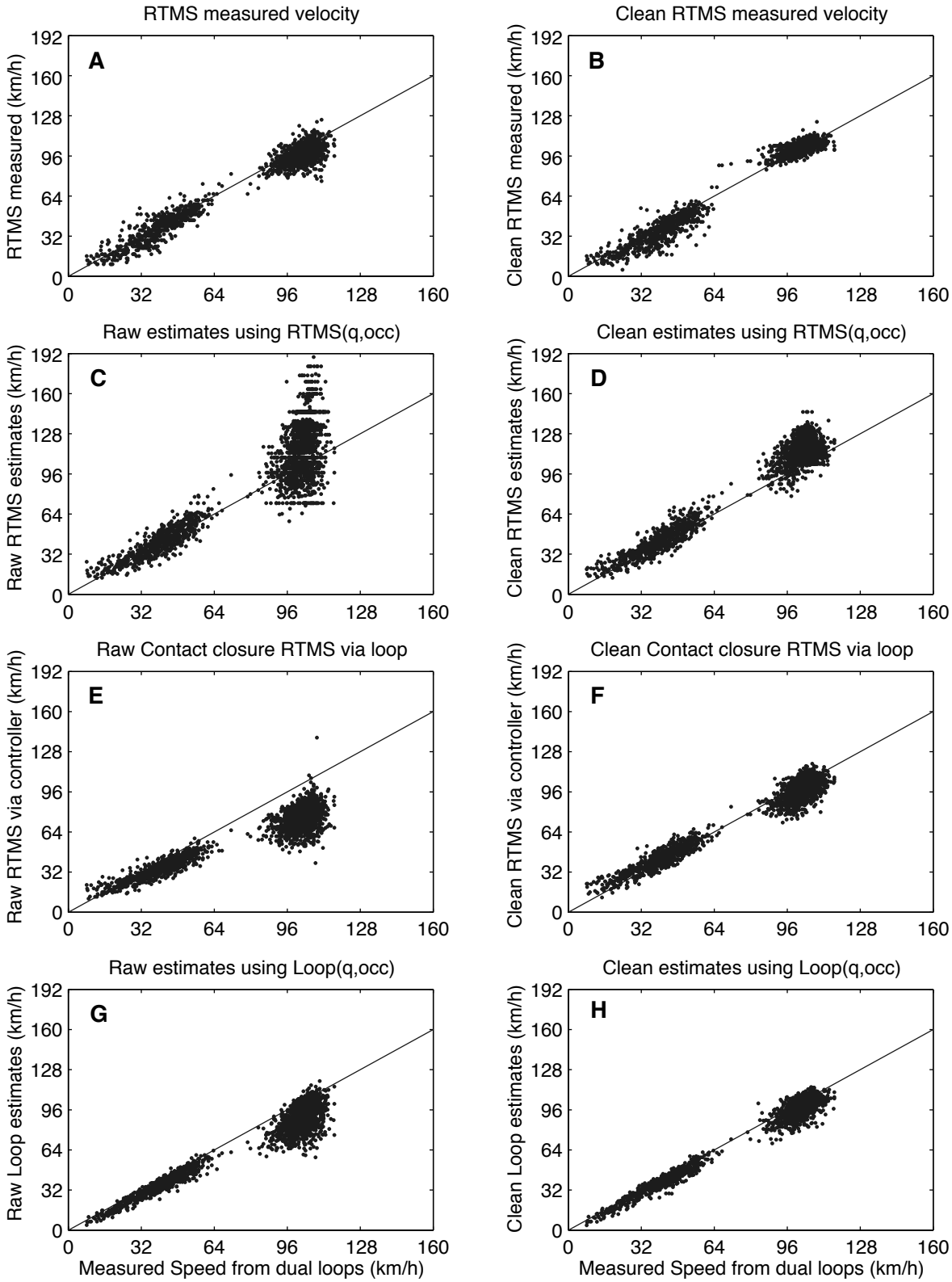

Figure 9, Scatter plots comparing median estimated and measured speed across all lanes. The left hand column shows the raw estimates and the right hand column shows the corresponding cleaned estimates after applying the filters. The first row, A-B, shows the speed estimates reported by the RTMS. The second row, C-D, shows the speed estimates after applying Equation 1 to the measured q and occ reported directly by the RTMS. The third row, E-F, shows the speed estimates from Equation 1 applied to q and occ for RTMS-via-controller. The fourth row shows the speed estimates from the loop q and occ.

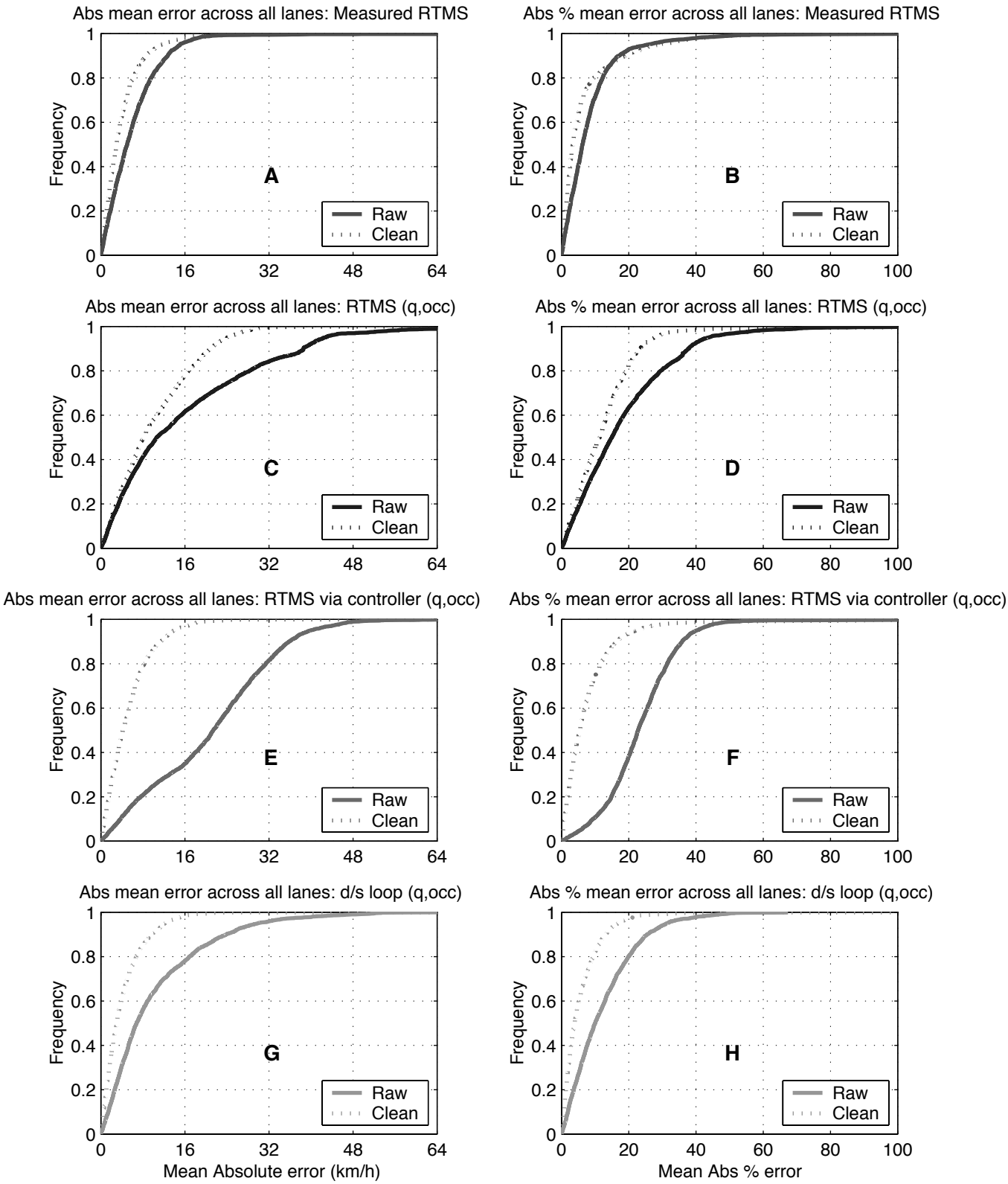

Figure 10, Distribution of absolute and percentage error for median estimated and measured speed across all lanes. The left hand column shows the distribution of absolute mean error and the right hand column shows the distribution for absolute percentage error. The first row, A-B, shows the speed estimates reported by the RTMS. The second row, C-D, shows the speed estimates after applying Equation 1 to the measured q and occ reported directly by the RTMS. The third row, E-F, shows the speed estimates from Equation 1 applied to q and occ for RTMS-via-controller. The fourth row shows the speed estimates from the loop q and occ.