Comparing INRIX speed data against concurrent loop detector stations over several months

Seoungbum Kim, PhD

Post-Doctoral Researcher
Department of Transportation and Logistics Engineering
Hanyang University at Ansan
1271, Sa-3 dong, Sangnokgu, Ansan-si
Gyeonggi-do, 426-791, Korea

Benjamin Coifman, PhD

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

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

Abstract

Many real-time traffic-monitoring applications only require speed or travel time. In recent years INRIX Traffic has started collecting and selling real-time speed data collected from "a variety of sources." The clients include direct to consumer and operating agencies alike. So far the INRIX speed data have received little independent evaluation in the literature, with only a few published studies. The current study exploits a unique juncture as the Ohio Department of Transportation transitioned from loop detectors to third party traffic data for real time management. The two traffic surveillance systems operated concurrently for about half a year in Columbus, Ohio, USA. This paper uses two months of the concurrent data to evaluate INRIX performance on 14 mi of I-71, including both recurrent and non-recurrent events.

The work compared reported speeds from INRIX against the concurrent loop detector data, as detailed herein. Three issues became apparent: First, the reported INRIX speeds tend to lag the loop detector measurements by almost 6 min. This latency appears to be within INRIX specifications, but from an operational standpoint it is important that time sensitive applications account for it, e.g., traffic responsive ramp metering. Second, although INRIX reports speed every minute, most of the time the reported speed is identical to the previous sample, suggesting that INRIX is effectively calculating the speeds over a longer period than it uses to report the speeds. This work observed an effective average sampling period of 3-5 min, with many periods of repeated reported speed lasting in excess of 10 min. Third, although INRIX reports two measures of confidence, these confidence measures do not appear to reflect the latency or the occurrence of repeated INRIX reported speeds.

Key words:
INRIX; speed measurement; Freeway traffic; performance evaluation

1. Introduction

Until recently freeway traffic data were almost exclusively collected in-house by operating agencies using fixed location detectors (e.g., loop detectors). In recent years several companies have begun collecting travel time data and selling this information to the public, traveler information providers, and even operating agencies. INRIX Traffic (2013) is among these new companies, since 2005 they have used a proprietary system to aggregate, "traffic-related information from millions of GPS-enabled vehicles and mobile devices, traditional road sensors and hundreds of other sources."

This paper exploits a unique juncture as the Ohio Department of Transportation transitioned from loop detectors to third party traffic data for real time management. The two traffic surveillance systems operated concurrently for about half a year in Columbus, Ohio, USA. This paper uses two months of the concurrent data to evaluate INRIX performance on 14 mi of I-71 in Columbus. The extended study period included potentially challenging conditions for the INRIX data collection that includes both recurrent events (rush hour congestion and late night low flow) and non-recurrent events (incidents and precipitation). This long scope of time increased the chance of observing intermittent features that could easily go unobserved in a typical validation study against manual ground truth that lasts only a few hours. The analysis undertook a broad overview of the data, comparing the reported speeds from INRIX against the concurrent aggregated loop detector data and then explicitly looked for any anomalous patterns such as latency and system outages.

Putting this work in context, some of the earliest published performance evaluation of the INRIX data was by the I-95 Corridor Coalition, where INRIX was an active partner in the consortium. Much of the evaluations consist of instantaneous comparisons offering only a "binary outcome," and the tests, "average errors over all time intervals, [so] only a few high-variance outliers are needed to invalidate the whole segment, even if the majority of the data points are within the acceptable range," (Aliari and Haghani, 2012). Instead, Aliari and Haghani contemplated a time series evaluation. Several other papers evaluated INRIX data during strictly free flow periods. In this context, Jia et al. (2013) examined a rural interstate with no congestion and found that the INRIX data had an accuracy of 80-90%, while Lattimer and Glotzbach (2012) found INRIX exhibited a 6 mph bias relative to ground truth on an uncongested freeway. Kim et al. (2011) looked at INRIX performance under recurring congestion in an urban area and found INRIX exhibited both a bias and latency, though the paper did not quantify either of these claims. Belzowski and Ekstrom (2013) compared real time data from INRIX and several other data providers against the actual travel of a vehicle. In general they found all of the data providers had difficulty reporting jams, especially short duration jams. They report that on the highway INRIX found 44% of jams lasting less than 10 min and 53% of jams lasting longer than 10 min. Kim et al. (2014) evaluated three data providers against actual tollway link travel times measured from toll tag passage times. They found the overall average speed errors to be within 10 mph throughout various levels of congestion. However, they also found that all three data providers missed a major incident lasting more than 4 hours. Kim et al. shows INRIX reported speeds 30 mph higher than the toll tag link speeds while INRIX classified the reports as being with "high confidence" during this major incident.

While the prior studies have mostly looked at direct comparison of concurrent speed measurements over short study periods, the current work uses archived data to examine continuous performance over a two-month period. Particular attention is paid to the timing of when major speed transitions were reported in the data to quantify any lags in the system. The remainder of this paper is as

follows: Section 2 presents the data and analysis area used in this study. Section 3 presents the performance of the INRIX data in three subsections. Finally, the paper closes in Section 4 with the conclusions of the study.

2. Data and Analysis Area

As noted above, this paper uses concurrent data from INRIX and loop detectors along 14 miles of I-71 in the first half of 2011 as the Ohio Department of Transportation (ODOT) transitioned from in-house data collection to third party traffic data for real time management. This study investigates 44 directional INRIX links on I-71 (20 northbound and 24 southbound), where each INRIX link is evaluated against a loop detector station located within the given link. The average link length is 0.6 mi. The INRIX data are reported every minute for each link, with each report including: timestamp, speed, and two parameters for the confidence of the reported INRIX speed, namely *Confidence Score* and *Confidence Value* (the details of these confidence measures will be discussed in Section 3.3).

The loop detector stations are spaced at roughly 1/3 mi intervals in this corridor. Figure 1a-b show the ends of the directional INRIX links and the concurrent loop detector stations along the corridor, with the southern end at the I-70/I-71/SR-315 interchange in the central business district (CBD) and the northern end outside the I-270 beltway in the northern suburbs. These loop detector stations are unique in their own right, rather than aggregating the data in the field, they send all of the vehicle actuations back to the traffic management center (TMC). Before aggregation at the TMC, the raw actuation data are archived (Coifman, 2006). With the actuation data, we can identify and correct for many errors at the loop detectors (Lee and Coifman, 2011, 2012a, 2012b), yielding very accurate measurements as have previously been verified by hundreds of GPS probe vehicle runs (Lee and Coifman, 2012b) and manually generated ground truth from concurrent video for over 63,000 vehicle actuations (Lee and Coifman, 2011).

There are several INRIX links that span multiple loop detector stations. Most of these multiple station links are north of loop detector Station 14, and from the loop data, this portion of I-71 does not see much recurring congestion, i.e., the INRIX links on I-71 tend to be shorter in the areas that see more recurring congestion. When an INRIX station spans multiple loop detector stations this study selects the best performing station of the available stations in the link. The given loop detector station could fall anywhere within the link, but across all links the average loop detector station location is at 50.2% of the distance of the link, i.e., the middle of the link (greater details of the link and loop detector station locations can be found in Coifman and Kim, 2013).

The current study investigates two months of loop detector data from April and May 2011 to evaluate the concurrent performance of the INRIX system. Meanwhile, according to ODOT, INRIX claims that the quality of their data in this region has improved significantly since the ODOT loop detectors were decommissioned in 2011. So as a result, a second batch of INRIX data from the same corridor was collected during the spring of 2013, after the loops had been shut down.

3. Evaluation of the INRIX data

The following analysis examines many features of the INRIX data by link for the 44 directional links in this study. Since the INRIX link names are long and lack any simple intuitive meaning, this paper will refer to the various links by the loop detector station that is contained in the link. Section 3.1 presents the overall performance of the INRIX data versus the loop detector data. Section 3.2 examines repeated reports by INRIX in which INRIX reports the exact same speed for several successive reporting periods.

Section 3.3 examines the Confidence Score and Confidence Value reported by INRIX to see if they reflect the repeated INRIX reports.

Since the INRIX process is proprietary, there is no way to know if the INRIX data stream includes real time data collected by ODOT, i.e., the INRIX processing might include measurements from the same sensors that this study uses for evaluation. As such, the comparisons against the loop detector data should be viewed in this context. So the differences between the two data sources in this study might be smaller than what drivers would experience on an otherwise non-instrumented freeway.

3.1 Time series evaluation of INRIX speeds against loop detector speeds

Figure 1c shows the summary plot of traffic speeds in the time-space plane on a typical day from the loop detector speeds (5 min median speed across all lanes at the given station) from the northbound stations selected for each INRIX link, as denoted with stars in Figure 1a. In this plot the lighter the shade the higher the speed, with the exception that white denotes no data. Thus, Figure 1c shows that a heavy queue formed around mile 4.8 at approximately 16:00, grew upstream, and eventually dissipates around 18:00. Over this same period a lighter queue formed around mile 8, it also grew upstream and overran the heavier queue around 17:00. Figure 1d shows the corresponding summary plot from the INRIX data, using 5 min averages of the native 1 min reporting period used by INRIX. Comparing Figure 1c with 1d, these two plots are consistent, showing queuing at roughly the same times and locations. Figures 1e-f repeats this comparison for the southbound traffic on the same day, with similar performance.

To investigate any biases between two systems, the time-series speed from a given loop detector station and corresponding INRIX link are compared. Figure 2a shows the time-series speed curves from Station 1 northbound and the corresponding INRIX link between 14:00 and 19:00 from May 26, 2011. Now, however, the time-series speed from the loop detectors is aggregated to 1 min, the native reporting rate of the INRIX data. Figure 2b shows the difference between the two time-series speeds and the horizontal line at 5 mph indicates the median of this difference over the day. The free flow periods show that most individual differences are above zero, indicating a measurable bias during off-peak periods between the two systems. From Figure 2b it is evident that the bias before and after the peak period is consistent with the daily median speed difference. Thus, the bias may reflect a poor calibration of one or the other system, and if so, it should be easy to correct the bias via a scale factor applied to the speed data from either of those systems. As such, the 5 mph daily bias is not taken to be an error by either system. This process was repeated for several links in the study corridor with recurring congestion and they yielded similar results (not shown).

Returning to Figure 2a, inspection shows that the INRIX speed lags the loop detector speed by many minutes; the clearest indication occurs when the speed drops around 15:00, and then the recovery back to free speed around 18:30. The time offset appears to be roughly balanced for both the onset of queuing and the subsequent dissipation. This plot shows that the loop detector data responded to the onset of queuing before the INRIX data and similarly the recovery is also seen first by the loop detectors. Assuming the specifications from INRIX reported by the I-95 Corridor Coalition also apply to these data, namely that the lag time measure is less than or equal to eight minutes (Haghani et al., 2009), then this latency is within the specifications of the INRIX deployment, but from an operational standpoint it is important to know about it for various time sensitive applications, e.g., traffic responsive ramp metering. To measure the amount of latency, this study calculates a correlation coefficient (CF) via Equation 1, while shifting the time-series loop detector with 10 sec steps. In this study the amount of time shifted from zero is called the time offset (TO). Figure 3a shows CF as a function of TO corresponding to Figure

2a. In this case the CF increases until the TO is equal to 420 sec and then decreases as the TO is increased further. Figure 2c repeats the comparison from Figure 2a except that the loop detector data are now shifted to the right by 420 sec. After shifting the loop detector data by TO both systems pick up the start and end of the congestion at roughly to the same time.

$$CF(x,y) = \frac{Cov\ (x,y)}{\sqrt{Cov\ (x,x) \times Cov(y,y)}} \tag{1}$$

where,

x : Time-series speed from loop detector after shifting by TO
y : Corresponding time-series from INRIX
cov (x,y) : Covariance of x and y

Obviously, the discrete location of the loop detector station should not measure the exact same time series speed as seen in the measures from the INRIX link over space. Careful inspection of Figure 1a reveals that Station 1 northbound is at the upstream end of the INRIX link. Since the queue grew from downstream of the INRIX link, queuing and traffic flow theory dictate that the delay should first be seen in the spatial INRIX data since the link is predominantly downstream of the loop detector station. This fact is supported by the time series in Figure 2c, where the duration of the congested period is slightly longer in the INRIX data because the closer the measurement location is to the bottleneck, the longer the period of time that it should exhibit queuing. As such, the INRIX latency is likely to be slightly larger than measured in Figure 3.

To investigate the reproducibility of the latency from day-to-day, this work repeats the comparison over 20 weekdays exhibiting congestion from April 2011 (Figure 3b) and May 2011 (Figure 3c) using the same loop and INRIX data pair. Figure 3d summarizes the results. Generally the INRIX speed is reported later than the loop detectors, with an average latency of about 340 sec (356 sec in April, 325 sec in May). Coifman and Kim (2013) repeat this process at several other links in the study corridor with recurring congestion and find similar results.

In contrast to Figure 2a, Figure 4a shows the two time series speed on an atypical day at Station 1. Notice the two speed drops in the INRIX data at roughly 18:00 and 18:30 are not evident in the loop detector data. Of the INRIX links with loop detector stations, this inconsistency was only observed at Station 1 northbound and it occurred on several days. The spacing between Station 1 northbound and the upstream detector station is 0.7 mi, there are several short INRIX links in this gap that do not have a corresponding detector station but also exhibit this pattern in a consistent manner.

Further investigation turned up 31 cases with similar displacement in speed in this link/station pair from 16 days in May 2011. Figure 5 shows a scatter plot of duration versus time of day for these speed displacements. In all cases the loop detector speed is indicative of free flowing speed while the INRIX speed drops below free flow. Investigating the adjacent INRIX links and loop detectors upstream and downstream of this location did not reveal a clear cause for these disturbances in the INRIX data, though the INRIX links without loop detectors were generally consistent with the transient speed drops. So the displacements may represent some combination of spatial and historical averaging in the INRIX processing (e.g., the two speed displacements in Figure 4a correspond to times of the day where the station is often congested, e.g., Figure 2a).

3.2 Repeated INRIX reported speeds

While INRIX reports speed for a given link every minute, INRIX typically reports the same speed for several successive reporting periods. This trend suggests that INRIX is effectively calculating the speeds over a longer period than it uses to report the speeds because one would expect to see at least a small variability from one sample to the next[1]. Figure 4b shows that although INRIX reports speeds every minute, most of the time the reported speed is simply repeated from the previous sample (for clarity only samples repeated at least four times are highlighted for the INRIX data, with the shorter durations suppressed in this plot). The two transient speed displacements in the INRIX data both follow extended periods of repeated speeds by INRIX, further supporting the hypothesis that they capture some aspect of historical averaging in the INRIX processing.

Although the reported traffic patterns at Station 1 northbound on April 15, 2011 were atypical of most of the links, the INRIX data transmission pattern was typical. Figure 4c repeats this exercise at the same station on another day that did not exhibit the speed displacements. This figure also shows the total number of repeated samples for each instance whenever the exact same INRIX speed is repeated in at least four successive samples.

For any INRIX link there should be 60 reported speeds per hour. For each sample this work tallies the duration of repeated samples, e.g., in Figure 4c the hour between 14:00 and 15:00 starts with 5 samples from a cluster of 10 repeated speeds, yielding 5 samples of "10 duration" in this hour (since the first 5 samples from the cluster fell in the previous hour) and repeating this tally for the entire hour results in the distribution in Figure 6a. Adding in the results from the same hour of the day for every day in the month at this link yields Figure 6b. Binning the results from Figure 6b in to: 2-5, 6-10, and 11+ and calculating the cumulative percentage yields Figure 6c. Then this work repeats the process for every hour in the day at the link, yielding Figure 6d (Figure 6c corresponds to the 15th bar in Figure 6d).

This performance varies by link, Figure 6e shows the corresponding results for the INRIX link from Station 15 northbound, almost 5 miles further north. Between 8:00 and 20:00 the median repetition of a time report is in the 2-5 min bin at Station 1 and it falls in the 6-10 min bin at Station 15. Figure 7a shows the results for all northbound links in May 2011. The work found similar results southbound, e.g., Figure 6f shows the corresponding results from Station 15 southbound.

As noted above, the INRIX data quality has reportedly improved since 2011, so Figure 7b repeats this analysis for the northbound stations in May 2013. Comparing a given link between 2011 and 2013 in all cases the cumulative distribution has shifted towards the shorter duration bins in 2013. These comparisons were repeated for the southbound links with similar results, e.g., compare the 2011 results from southbound Station 15 against the corresponding results from 2013 in Figures 6f-g.

Another way of measuring this effective sampling period that emerges in the INRIX data is to tally the total number of changes in speed per link over a given hour (which is equal to the total number of "uniquely" reported speeds minus one), and calculate percentage of the samples that change every hour. So if 20% of the hourly samples are changed (12 unique samples out of 60 reported samples), it indicates that *on average* a unique speed is reported every 5 min (equivalent to *an average* effective sampling period of 5 min during that hour if the new reports were evenly spaced), while 33% unique samples corresponds to *an average* of one unique speed every 3 min. Figure 6h shows a bar graph of the hourly statistics in May 2011 contrasted against May 2013 for southbound station 15. Not surprisingly, in both

[1] Given the competitive nature of the for profit traffic data providers, it is quite possible that in response to these findings INRIX and oher data providers will start adding a small random noise to their reported speeds to preclude such analysis in the future.

years the frequency of speed change tends to be lowest during the early morning and late evening, while the highest frequencies are typically seen during the morning or evening peak period, consistent with daily trends in traffic volume. Figure 8 repeats this exercise for all of the southbound stations with data in both years. Most stations exhibit 20%-30% of the samples exhibiting a speed change during the day, corresponding to effective sampling periods between 3 and 5 min. Between 2011 and 2013, the southern-most stations that are closest to the CBD (the first four entries in the top row) show a slight reduction in in the frequency of speed changes while most of the remaining stations showed a slight increase of roughly 5% in the frequency of speed changes. These small changes in frequency suggest that much of the improvement seen between Figure 7a in 2011 and Figure 7b in 2013 came from reducing the number of extremely long periods without a new reported speeds, e.g., compare the difference between Figure 6f and 6g against Figure 6h from the same station.[2] This finding is consistent with the fact that in Figure 7 the interface between "unique sample" and "2-5 min" bins did not change much between the two years.

3.3 Confidence Score and Confidence Values reported by INRIX.

This section focuses on the repeated INRIX reported speeds from Section 3.2 in the context of the concurrent INRIX confidence measurements; thus, these results are independent of the loop detector data. INRIX reports two measures of confidence in a given reported speed. The first measure INRIX calls the *Confidence Score*, in which INRIX assigns one of three possible outcomes:

30: High confidence, indicating that the report is based on real-time data for that specific segment.
20: Medium confidence, indicating that the report is based on real-time data across multiple segments and/or based on a combination of expected and real-time data.
10: Low confidence, indicating that the report is based primarily on historic data.

Reviewing the INRIX data from May 2011 and May 2013, this research counts the number of speed records associated with each of the Confidence Scores at each of the INRIX links in the study area. Overall, 94.0% of the 2011 data and 96.3% of the 2013 data had Confidence Score of 30.[3] Figure 9 show the binned cumulative distribution of Confidence Score by time of day (minute) across all stations in the given direction and all 31 days in the given month. Like Figure 7, the different values of the binned distribution are distinguished by shading, with a given vertical bar in any one of the four plots corresponding to the binned cumulative distribution for a single minute of the day. The darker the shading, the lower the Confidence Score. Figure 9 shows that between 10% and 30% of the Confidence Scores are 10, and the rest are 30 during early morning and late night periods. Over the remainder of the day there are no Confidence Scores of 10, instead, there are between 0% and 20% of the scores at 20 and the rest at 30. Although not shown, the individual station results are similar to these aggregated plots, with a score of 30 for over 95% of the mid-day and over 80%-90% of the early morning hours. Compared to 2011, the frequency of scores of 10 and 20 dropped in 2013 regardless of direction.

---

[2] Similar results were seen in both directions, so for brevity this paper uses the format of Figure 7 to present the northbound results, and the format of Figure 8 to present the southbound results. Both formats for both directions are presented in Coifman and Kim, 2013.

[3] The Confidence Scores were extracted from the INRIX archives at different times, and as a result, a few stations were only extracted in one year or the other. For the sake of comparison, this section only evaluates the stations for which there was extracted data in both 2011 and 2013.

Figure 10a revisits the data from Figure 2, only now highlighting the repeated speeds reported by INRIX. Figure 10b shows the corresponding Confidence Score reported by INRIX for all of the samples in this period, again highlighting those records corresponding to repeated speeds from INRIX. This example is typical of other days and stations: the repeated speeds usually have the highest Confidence Score, 30. In fact over the entire period shown in this plot the Confidence Score never deviated from 30. As such, the Confidence Score does not reflect how many times the reported speed has been repeated by INRIX. Both the name and the definition of "Confidence Score" from INRIX imply that high scores correspond to up to date data. Yet as noted above, if the data were measured every minute one would expect to see the speed fluctuate a little from sample period to sample period, so contrary to the Confidence Score, the repeated INRIX reported speeds suggest that the INRIX data are potentially out of date and may not capture the current, real-time conditions on the roadway.

INRIX reports a second measure of confidence, which they call the *Confidence Value.* Reportedly, the Confidence Value is based on a comparison against historical trends, though the details of this evaluation are proprietary. The Confidence Value only applies when the Confidence Score is 30. Given the claim that the data quality improved between 2011 and 2013, the analysis focused on the 2013 data.

Figure 11a-b repeats the analysis from Figure 10 on a new day: May 2, 2013. Again Figure 11b shows that the repeated INRIX speeds have the highest Confidence Score, 30, as do all of the other samples over the entire period shown. To investigate whether the so-called Confidence Value captures the repeated INRIX reported speeds when the reported Confidence Score does not, Figure 11c repeats the analysis using the Confidence Value. Similar to the Confidence Score, the repeated INRIX speeds usually have the highest Confidence Value, 100, though a few exceptions are evident, as shown in detail in Figure 11d-f. For the three times that the Confidence Value deviates from 100, a period of repeated INRIX speeds starts out with a lower Confidence Value, the subsequent repetitions of the same reported speed usually have progressively higher Confidence Values (though exceptions exist, e.g., the trend at 16.8 hr initially is decreasing). These results are typical of other days at the station. Investigating more groups of repeated samples with respect to Confidence Value, this work found 8,943 groups with repeated speed samples over May 2013 at Station 1. The total number of reported speeds from these groups is 40,468, which is about 91% of the INRIX reported speeds for this station in the month. Of the 8,943 groups, 7,099 had a constant Confidence Value where 84% of these were at 100. Meanwhile, 1,844 of the groups consist of multiple Confidence Values. In 1,370 groups (74% of those with multiple Confidence Values) the Confidence Values increase monotonically. In 213 groups (11% of those with multiple Confidence Values) the Confidence Values decrease monotonically, while the remaining 15% of the groups with multiple Confidence Values show mixed trends. This analysis at Station 1 is typical of other stations. The exact nature of the Confidence Value calculation is proprietary, but like the Confidence Score, it appears that the Confidence Value does not reflect how long a given speed has been reported by INRIX.

Returning to the unexplained speed displacements from the end of Section 3.1, this work also investigated Confidence Score and Confidence Value during those events. Figure 5 already shows the duration and time of day when INRIX deviated from the free flow speeds indicated by the loop detectors for this link. Figure 12a then sorts these points based on the reported Confidence Score. An “x” indicates the INRIX speed displacements that had all corresponding Confidence Scores equal to 30, while an “o” represents displacements with Confidence Scores less than 30. As can be seen, 25/31 (81%) of the displacement samples have the highest Confidence Scores. Meanwhile, as shown in Figure 12b, the majority of samples with the unexplained speed displacement with Confidence Scores equal to 30 (squares with “x”) have average Confidence Value less than 100 (12/20 or 60%). This plot shows all of

the Confidence Values that were recorded, so although the Confidence Value only applies to Confidence Scores of 30, since it was reported for lower Confidence Scores they are all shown on the plot. On the other hand, some samples were lacking the Confidence Value (denoted NaN) because they were not extracted from the archived data used for this study. Figure 12 shows evidence suggesting that most of the observed speed displacements at Station 1 northbound either had reduced Confidence Value or Confidence Score, though the number of observations is too small to provide any proscription for interpreting the Confidence Scores or Confidence Values.

4. Conclusions

Many real-time traffic-monitoring applications only require speed or travel time, e.g., traveler information or detecting the presence congestion. In recent years INRIX has started collecting and selling real-time speed data collected from "a variety of sources." The clients include direct to consumer and operating agencies alike. As discussed in Section 1, so far the INRIX speed data have received little independent evaluation. This paper presents new analysis on several aspects of the INRIX data not found in the existing literature. For much of the analysis the work used individual vehicle actuations from well-tuned loop detectors as the baseline. Unlike prior studies, this work examined an extended period, with several months of contiguous data from concurrent detectors, allowing the work to catch intermittent events as well as normal operation.

At the coarse level of 5 min aggregated speed data plotted in the time space plane, the INRIX summary plots were comparable to the corresponding plots derived from loop detector data, both showing similar patterns of congestion, queue growth, and so forth. Upon examining the time series speed data aggregated at the INRIX's native 1 min reporting period, three issues became apparent. First, the INRIX speeds tend to lag the loop detector measurements by almost 6 min. In light of the latency findings in this work, one can review the past studies that relied on individual probe vehicle runs for ground truth and observe evidence of similar latency, e.g., plotting the data from Table 1 in Aliari and Haghani (2012) and Figure 3b in Kim et al. (2011) the INRIX data appear to exhibit latency of approximately 10 min relative to the concurrent link speeds measured from Bluetooth IDs. This latency appears to be within the specifications of the sensor deployment, but from an operational standpoint it is important to know about it for time sensitive applications, e.g., traffic responsive ramp metering. Second, although INRIX reports speed every minute, most of the time the reported speed is identical to the previous sample. After tracking the changes in reported speed, this work observed an effective average sampling period of 3-5 min, with many periods of repeated INRIX reported speed lasting in excess of 10 min. The repeated INRIX reported speeds suggest that the INRIX data are potentially out of date and may not capture the current, real-time conditions on the roadway. Between 2011 and 2013 the effective sampling period dropped, indicating a positive trend. Third, although INRIX reports two measures of confidence, these confidence measures do not appear to reflect the latency or the occurrence of repeated INRIX reported speeds, i.e., most of the repeated speeds had no impact on the confidence measures, and the confidence measures do not provide any insight into how many times a given INRIX speed has been repeated. While the findings are notable, conceivably INRIX data could still be used for real time operations provided the control decisions are designed to accommodate the findings discussed above.

In one of the 44 directional links studied, the INRIX reported speeds exhibited intermittent speed drops while the concurrent loop detector speeds remained representative of free flowing conditions. This error occurred 31 times in May 2011 and each occurrence lasted less than 15 min. While the cause of the

error is unknown, it is suspected to be related to historical averaging. In most occurrences one or both of the confidence measures appears to have been impacted, suggesting that the confidence measures are performing as intended in this case. The number of observations is too small to provide any proscription for interpreting the Confidence Scores or Confidence Values and so it remains a topic for future research.

5. Acknowledgements

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

We are particularly grateful for the assistance and input from George Saylor, Bryan Comer, Merih Ocbazghi, Jason Yeray, and John MacAdam at the Ohio Department of Transportation, and Zhuojun Jiang at the Mid-Ohio Regional Planning Commission (MORPC).

The authors would also like to thank the anonymous reviewers for their constructive feedback that helped make this a better paper.

6. References

Aliari, Y., Haghani, A. (2012). Bluetooth Sensor Data and Ground Truth Testing of Reported Travel Times. *Transportation Research Record: Journal of the Transportation Research Board, Issue 2308*, pp 167-172

Belzowski, B., Ekstrom, A., (2013). *Stuck in Traffic: Analyzing Real Time Traffic Capabilities of Personal Navigation Devices and Traffic Phone Applications*, University of Michigan Transportation Research Institute, UMTRI-2013-36, 38p

Coifman, B., (2006). *The Columbus Metropolitan Freeway Management System (CMFMS) Effectiveness Study: Part 2 - The After Study*, Ohio Department of Transportation, FHWA/OH-2006/19, 372p

Coifman, B., Kim, S., (2013). *Assessing the Performance of the SpeedInfo Sensor*, Ohio Department of Transportation, FHWA/OH-2013/21, 60p

Haghani, A., Hamedi, M., Sadabadi, K., (2009). *I-95 Corridor Coalition Vehicle Probe Project: Validation Of Inrix Data July-September 2008, Final Report*, I-95 Corridor Coalition, 129p

INRIX Traffic (2013). http://www.inrix.com/ (accessed September 12, 2013)

Jia, C., Li, Q., Oppong, S., Ni, D., Collura, J., Shuldiner, P. (2013). Evaluation of Alternative Technologies to Estimate Travel Time on Rural Interstates. *Proc. of the Transportation Research Board 92nd Annual Meeting*, Transportation Research Board, 16p.

Kim, K., Chien, S., Spasovic, L. (2011). Evaluation of Technologies for Freeway Travel Time Estimation: Case Study of I-287 in New Jersey. *Proc. of the Transportation Research Board 90th Annual Meeting*, Transportation Research Board, 17p

Kim, K., Motiani, D., Spasovic, L., Dimitrijevic, B., Chien, S., (2014). Assessment of Speed Information Based on Probe Vehicle Data: A Case Study in New Jersey, *Proc. of the Transportation Research Board 93rd Annual Meeting*, 19p

Lattimer, C., Glotzbach, G. (2012). Evaluation of Third Party Travel Time Data. *Proc. of the ITS America 22nd Annual Meeting & Exposition*, ITS America, 8p

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

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

Lee, H., Coifman, B., (2012b). Quantifying Loop Detector Sensitivity and Correcting Detection Problems on Freeways, *Journal of Transportation Engineering, ASCE*, vol. 138, no. 8, pp 1051-1061

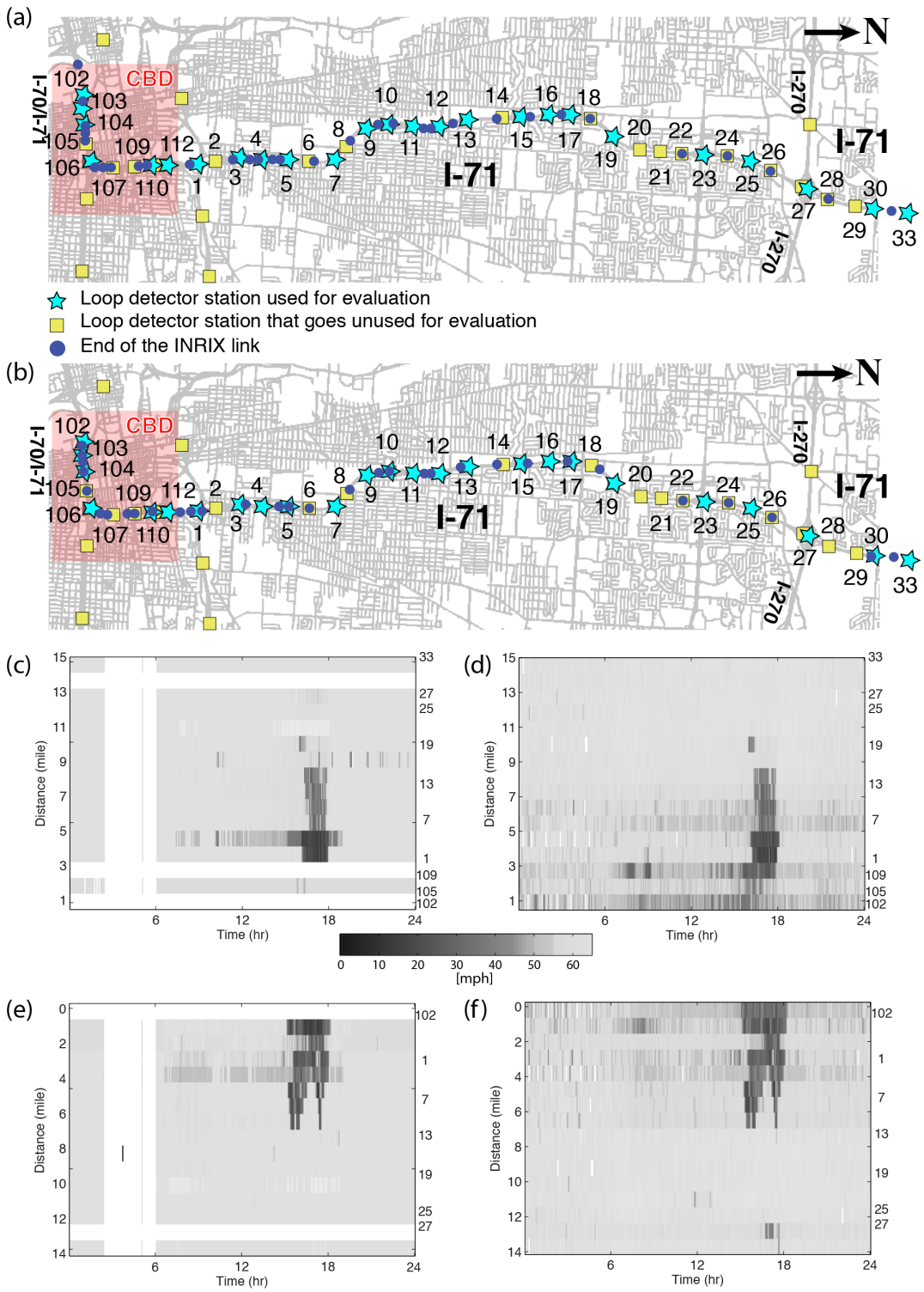

Figure 1, The I-71 corridor and the ends of the (a) northbound and (b) southbound INRIX links used in this study. Summary plots of 5 min median speeds along the I-71 corridor on a typical day, May 17, 2011, from (c) the northbound loop detectors, (d) the corresponding northbound INRIX data, (e) the southbound loop detectors, and (f) the corresponding southbound INRIX data. In all summary plots, traffic flows from bottom to top.

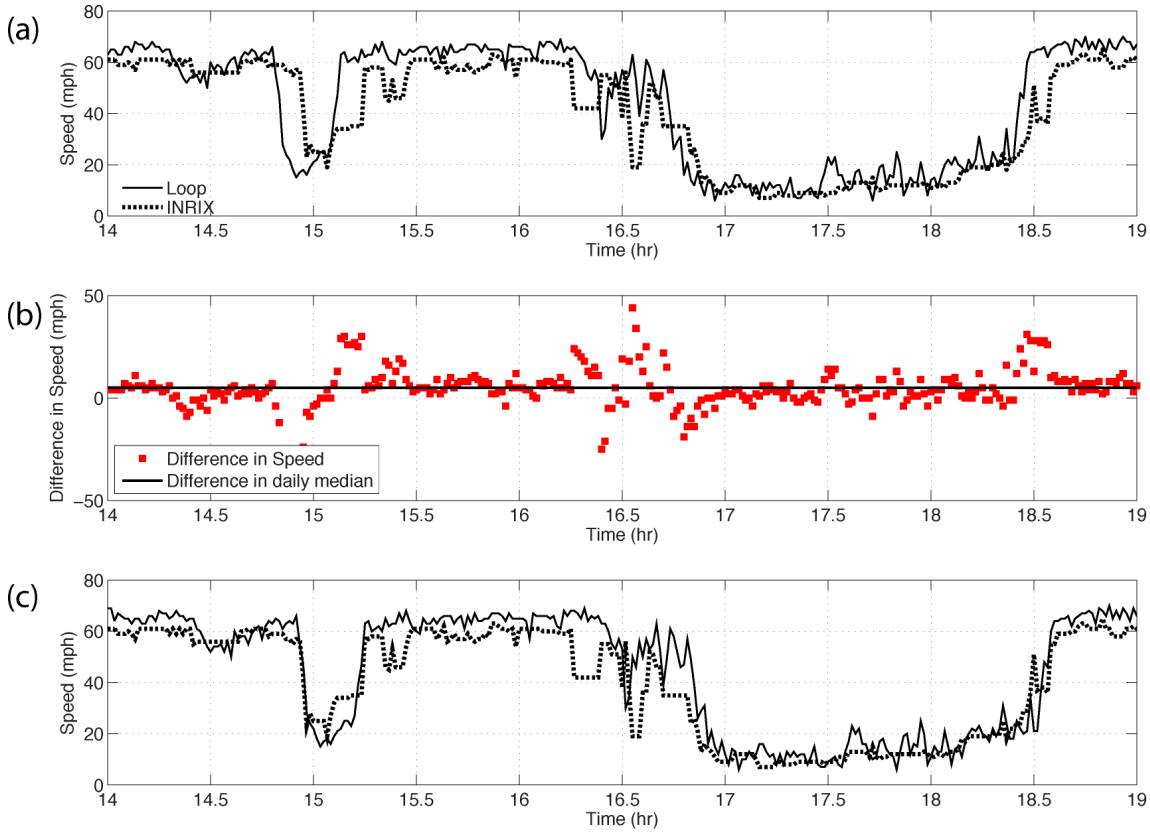

Figure 2, (a) 1 min aggregated speed as a function of time from the loop detector and INRIX systems at Station 1 northbound on a typical day, May 26, 2011, (b) the difference between the two concurrent time-series speeds in and the median bias over the day shown with a horizontal line at 5 mph, (c) time-series speed after shifting the loop detector time-series by 420 sec.

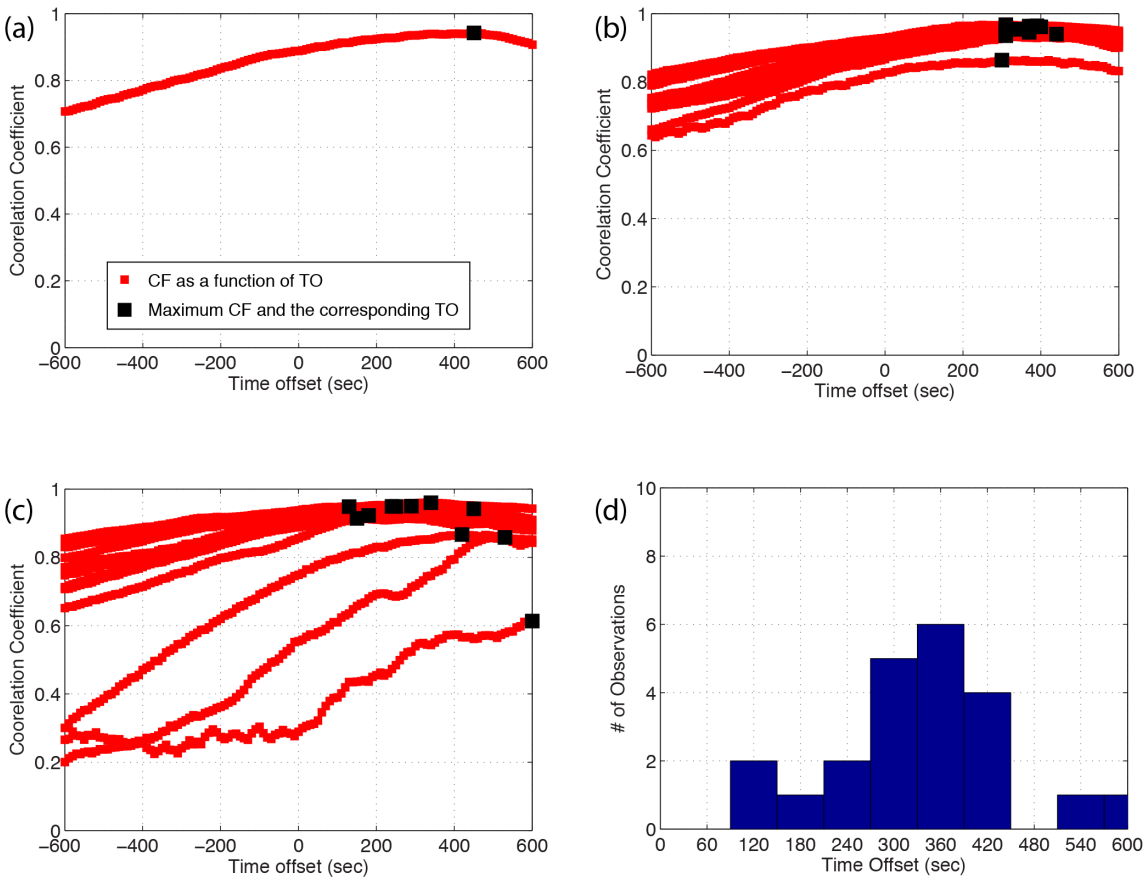

Figure 3, Correlation coefficient as a function of the time offset, highlighting the TO at the maximum CF for the INRIX data at Station 1 northbound, (a) May 26, 2011, (b) ten weekdays in April, 2011, (c) ten weekdays in May, 2011, (d) distribution of the TO at maximum CF from parts b and c.

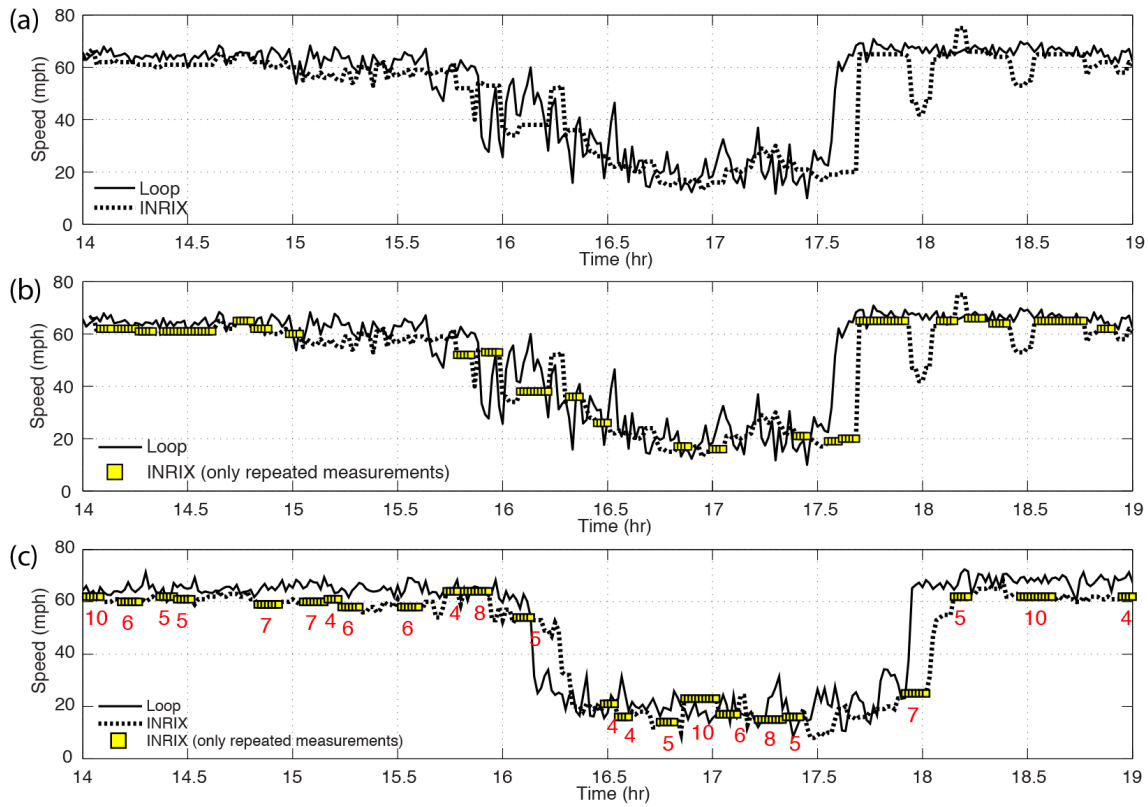

Figure 4, (a) 1 min aggregated speed as a function of time from the loop detector and INRIX systems at Station 1 northbound on a non-typical day, April 15, 2011, (b) replotting the data from part a, this time highlighting the INRIX data that repeat the speed from the previous reporting period (for clarity only repeated samples persisting at least 4 min are shown for the INRIX data, with the shorter durations suppressed in this plot), (c) repeating this comparison at the same station on April 29, 2011, denoting the duration of repeated speeds from INRIX (again only repeated samples persisting at least 4 min are shown for the INRIX data).

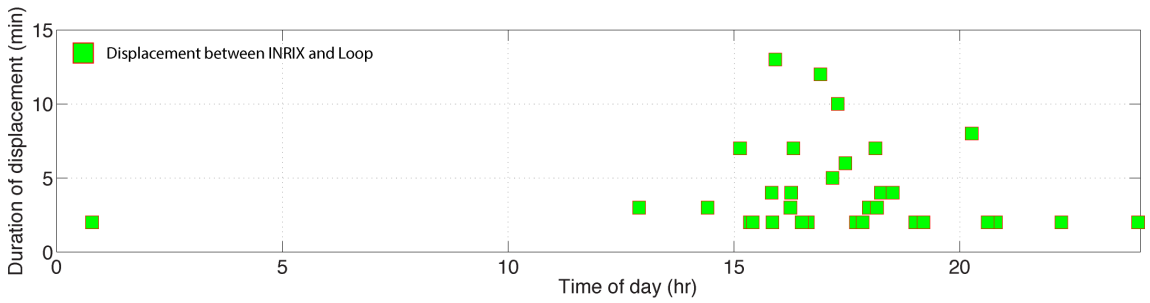

Figure 5, Duration of displacement between INRIX and loop detector speed.

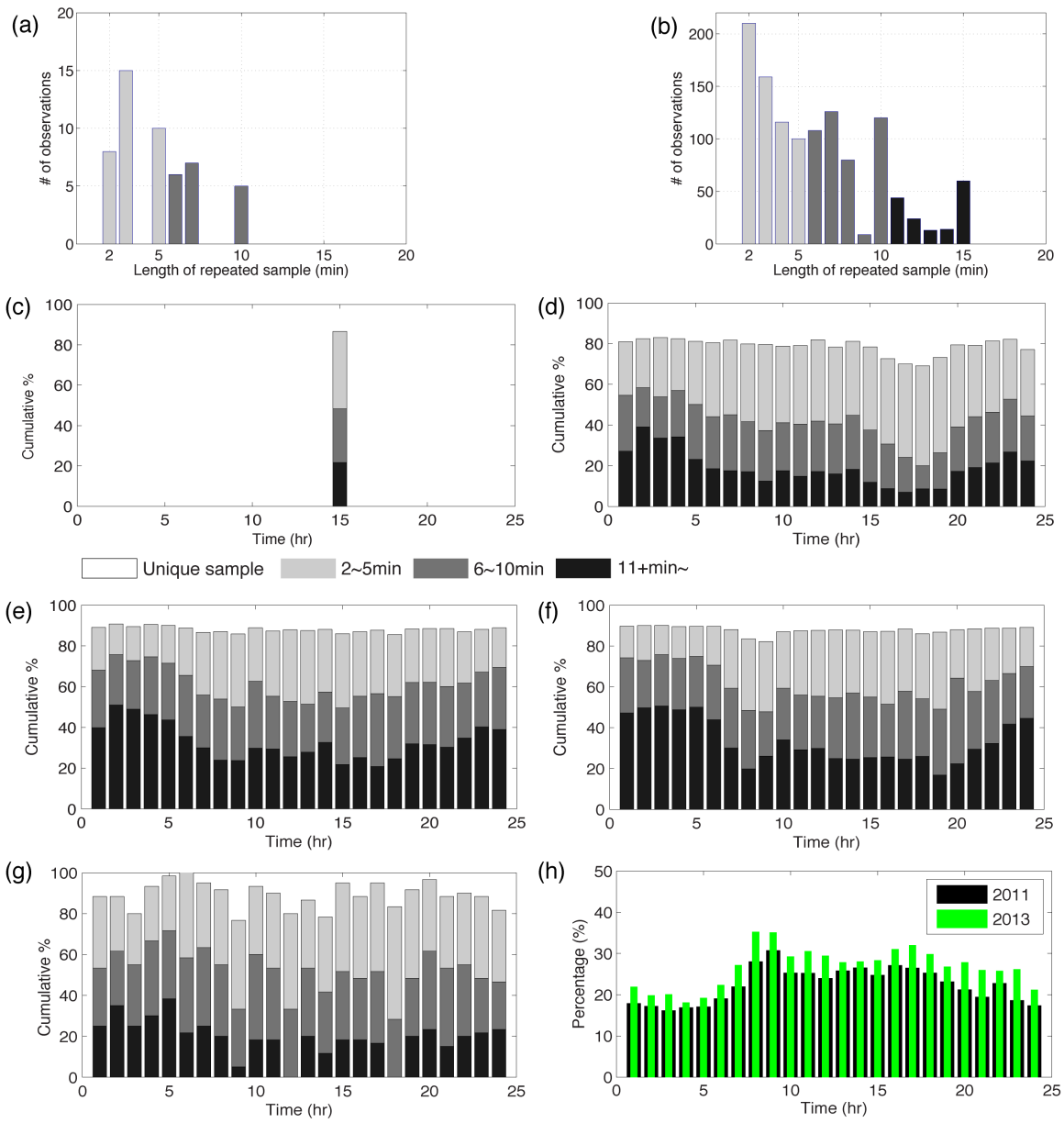

Figure 6, (a) Distribution of length of time a report is repeated (by number of successive samples) between 14:00 and 15:00 on April 29, 2011 at Station 1 northbound, (b) bar chart showing the distribution between 14:00 and 15:00 from all days in April, 2011, (c) binned cumulative distribution from the histogram in part b. Binned cumulative distribution of monthly median durations with non-unique samples by hour for (d) Station 1 northbound in April 2011, (e) Station 15 northbound in April 2011, (f) Station 15 southbound in April 2011, (g) Station 15 southbound in May 2013. (h) Percentage of samples with a change in speed by hour for Station 15 southbound in May 2011 and May 2013

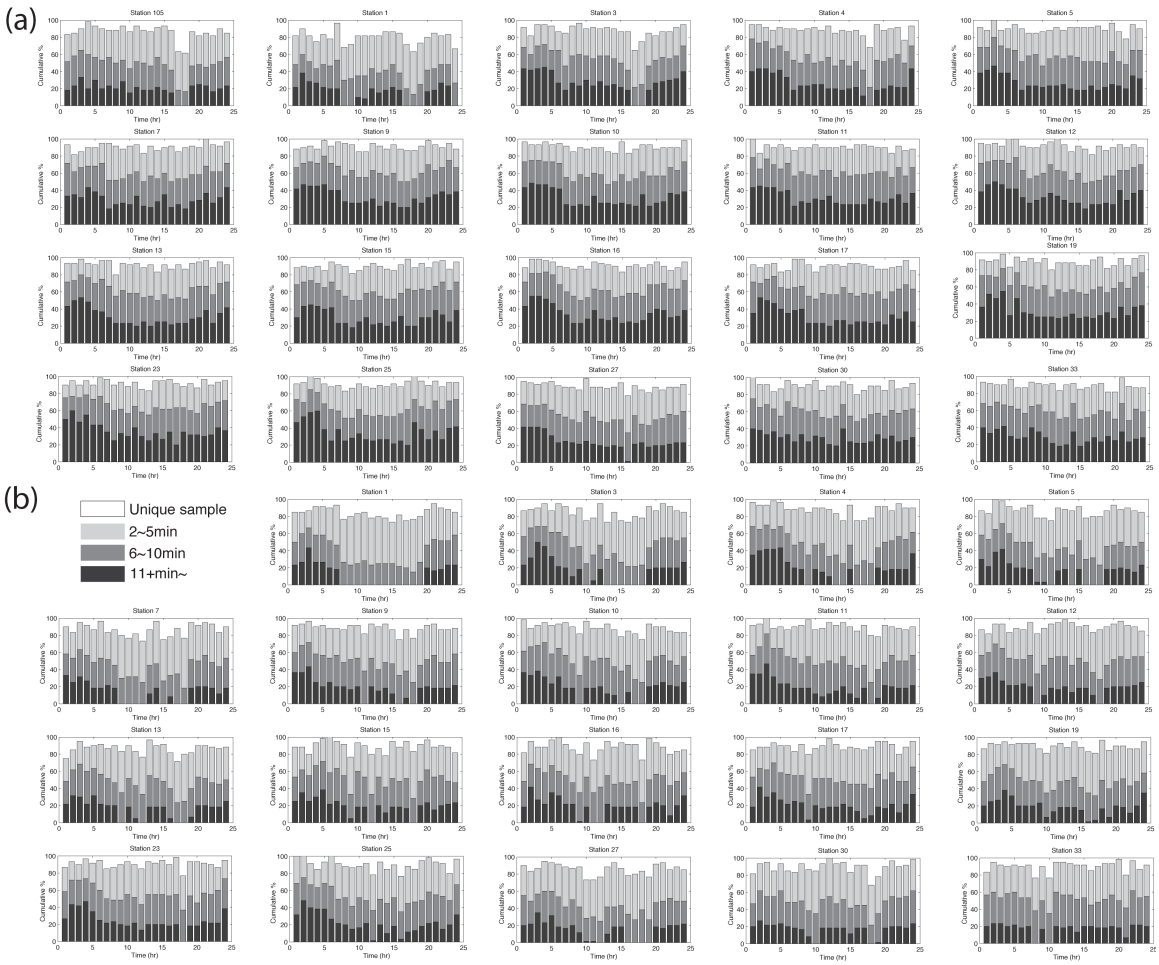

Figure 7, Binned cumulative percentage of monthly median durations with non-unique samples by hour for each northbound station, (a) May 2011, (b) May 2013. The given station number is shown above each plot. Note that the INRIX data corresponding to Station 105 simply were not extracted in 2013, hence, part b has one less plot than part a.

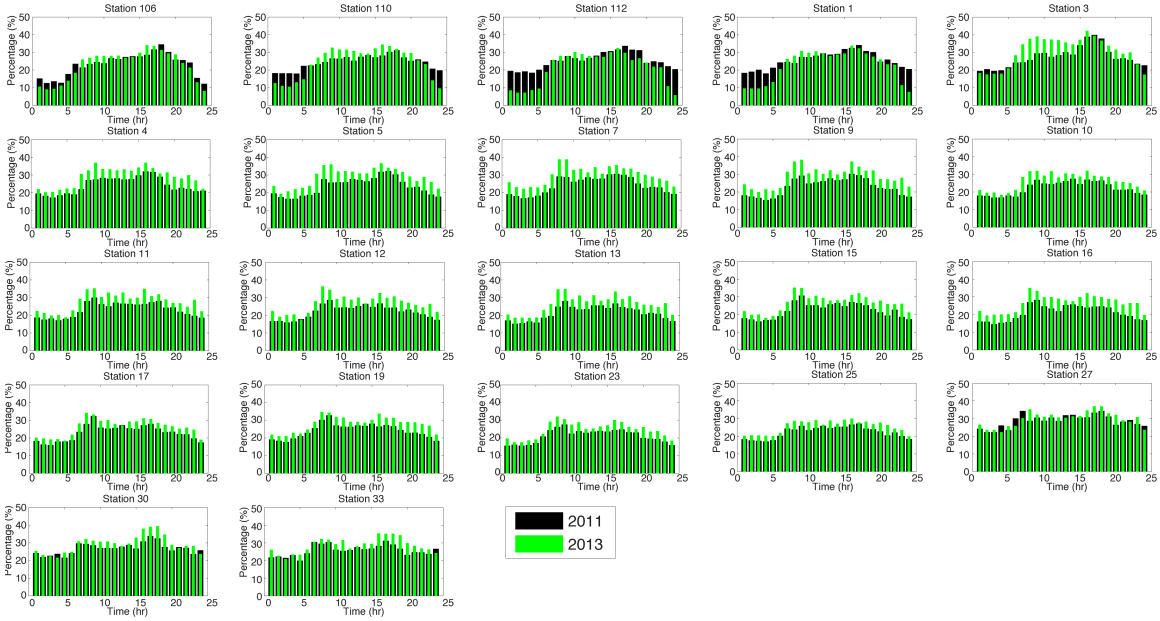

Figure 8, Percentage of samples with a change in speed by hour for all southbound stations with data in both May 2011 and May 2013. The given station number is shown above each plot.

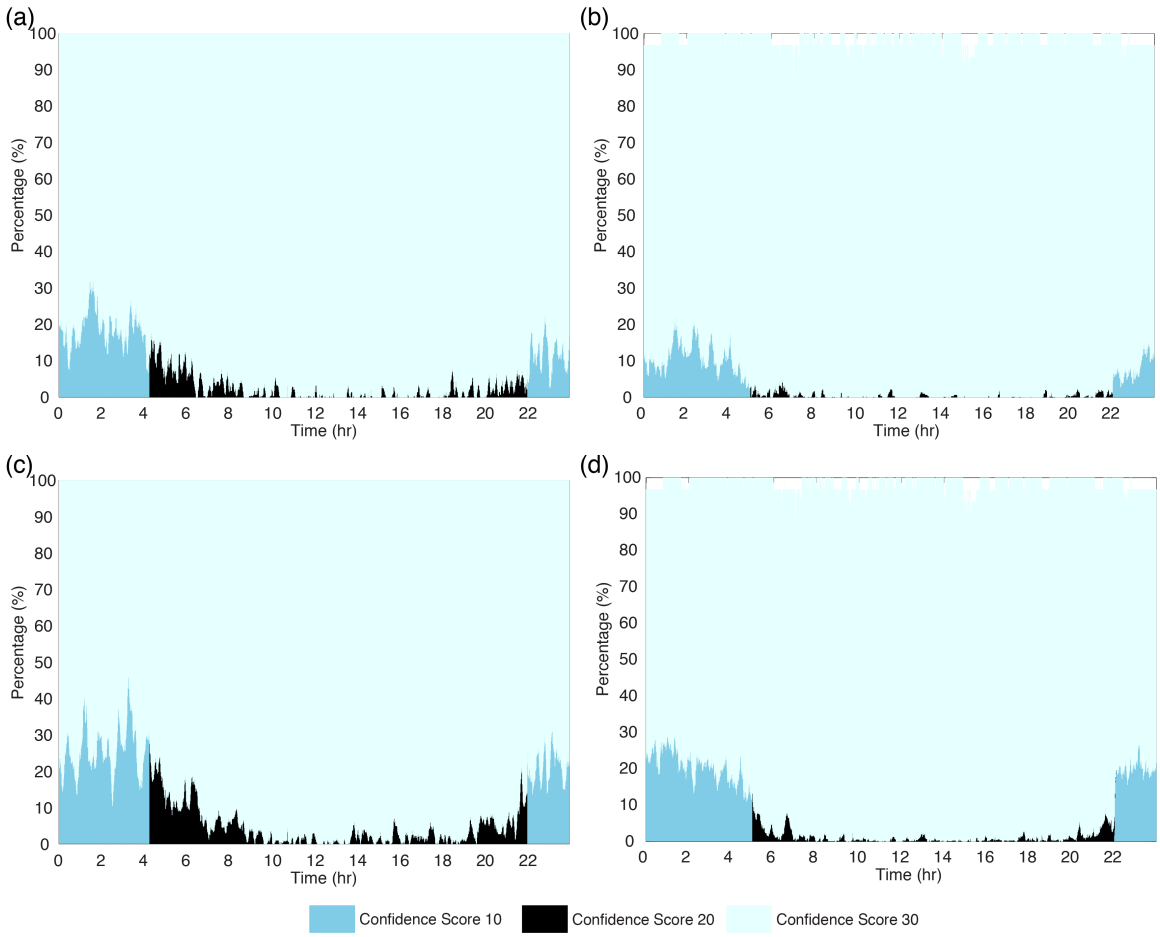

Figure 9, Binned cumulative percentage of samples showing the distribution of Confidence Scores across stations and all days in the month as a function of time of day, i.e., one vertical bar per minute of the day for: (a) northbound stations in May 2011, (b) northbound stations in May 2013, (c) southbound stations in May 2011, (d) southbound stations in May 2013.

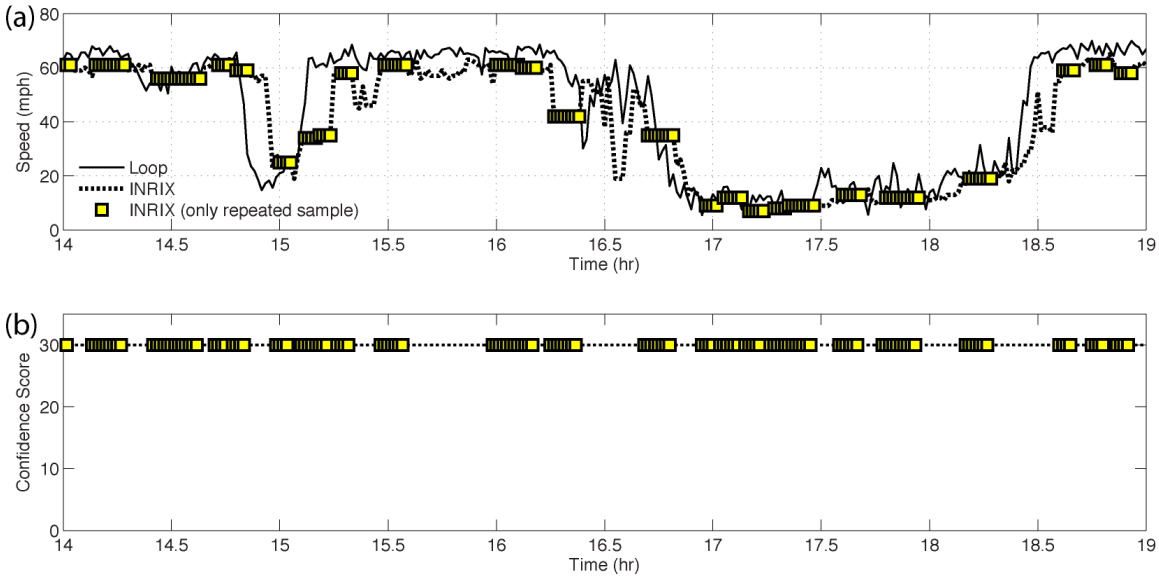

Figure 10, (a) The 1 min reported speed as a function of time from the loop detector and INRIX system at Station 1 northbound on May 26, 2011. For clarity only the repeated speeds persisting at least 4 min from INRIX are highlighted, (b) the corresponding INRIX Confidence Score.

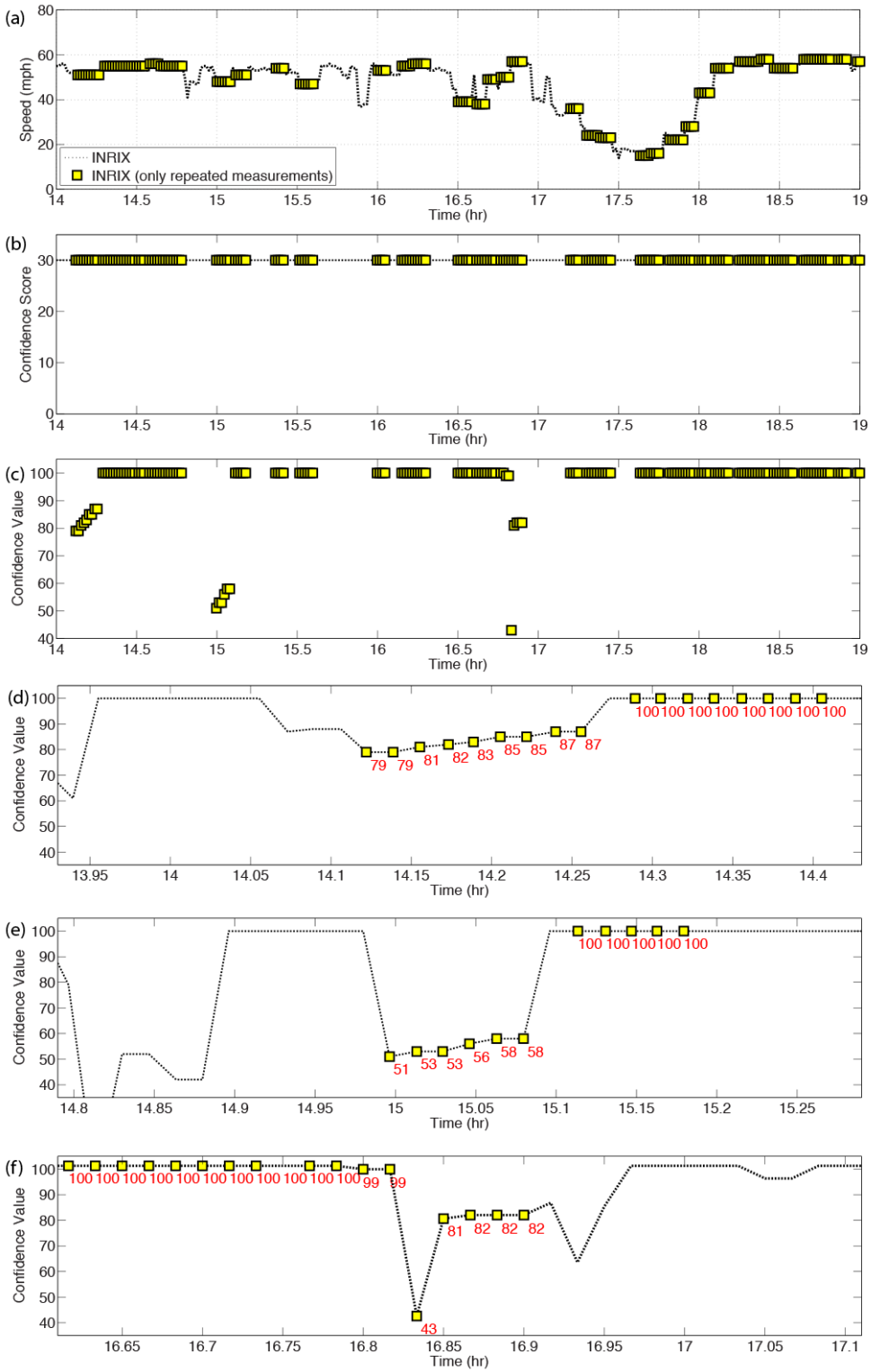

Figure 11, (a) The 1 min reported speed as a function of time from the INRIX system at Station 1 northbound on May 2, 2013. For clarity, only the repeated INRIX speeds persisting at least 4 min are highlighted. (b) The corresponding INRIX Confidence Score, (c) the corresponding INRIX Confidence Value. Details of part c: (d) around 14:10, (e) around 15:00, and (f) around 16:50.

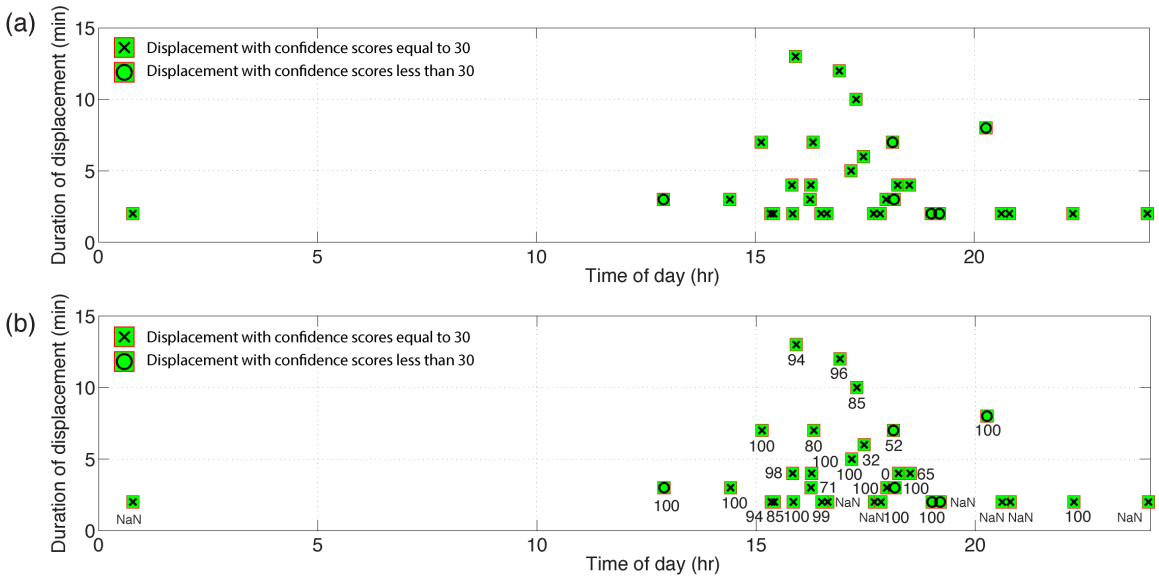

Figure 12, Duration of displacement between INRIX and loop detector speed showing the corresponding a) Confidence Score, and (b) the average Confidence Value for each of the displacement events.