The Eastern Transportation Coalition Probe Data Validation: HERE, INRIX and TomTom

Report for Maryland (#13)
Arterial Validation (Low-volume Roadways)

Data Collection: July 2020
Report Date: January 2021

Prepared by:
Zachary Vander Laan, Sara Zahedian
University of Maryland, College Park

Acknowledgements:
The research team would like to express its gratitude for the assistance it received from the state highway officials in Maryland during the data collection phase of the validation effort.

1. Table of Contents

List of Tables

2. List of Figures

Executive Summary

The Eastern Transportation Coalition (TETC) data validation program regularly compares probe-based travel time and speed data with reference data collected via field measurements, seeking to benchmark the quality of probe data sold through the Vehicle Probe Project (VPP) marketplace under a variety of conditions. Previous validation studies have primarily emphasized freeways and signalized arterial facilities that carry mid-to-high traffic volumes. Probe data has been proven an effective representative along freeways and many high-volume arterials. As such, the Coalition is now moving to also understand data quality on low-volume roads.

This report uses reference data acquired on three low-volume corridors in Maryland: MD-30, US-11, and US-40, whose road characteristics are summarized in Table 1. Note that the AADT values are generally around 15k-24k, with the majority of segments and directional miles in the 15k-16k range – a volume range much lower than most recent data collections.

ES Table 1 - Arterial Corridor Description in Maryland

Corridor	# Segments	Miles (directional)	Speed Limit (MPH)	Signals / mile	Avg. AADT
MD-30	20	36	30 – 50	0.5	16k
US-11	4	4	25 – 30	3.7	15k
US-40	6	8	35 – 55	5.7	24k

The standard method for evaluating probe data (called the "traditional analysis" throughout the report) involves comparing reference and VPP data sources over 5-minute periods and summarizing the results in terms of two standard error measures. This operations-focused evaluation method serves as the basis for contract compliance and is always conducted as part of regular validation efforts, but can have limitations in low-volume traffic conditions. In particular, low traffic volumes provide fewer opportunities to obtain reference travel time observations, which in some cases can be insufficient to characterize ground truth conditions.

Accordingly, the validation team is taking this opportunity to augment the traditional analysis by re-introducing and extending a previously-proposed approach: **the sampled distribution method.** While the traditional approach measures accuracy at any point in time, the sampled distribution method takes a fundamentally different perspective by focusing on vendors' ability to capture repeatable travel time patterns – an approach oriented towards arterial management and planning applications. Coalition members have recently inquired about whether probe data on low-volume roads can be trusted for planning purposes, and the sampled distribution method provides a useful tool to help answer this question.

Traditional Analysis Results

ES Tables 2-4 summarize each vendor's error metrics using the traditional analysis. Vendor probe speeds are compared to (a) the mean WRTM speed, and (b) 1.96 SEM (Standard Error of the Mean) speed band for each 5-minute interval. These comparisons are quantified in terms of two error metrics: Average Absolute Speed Error (AASE) and Speed Error Bias (SEB), which are computed separately for four different speed bins. Columns relevant to contract specifications are outlined in red, and error metric values are colored green or red to indicate whether the value is within contract specifications (AASE <= 10 mph, SEB <= +/-5 mph). **The contract specifications are met in every speed bin for each Vendor.**

ES Table 2 – HERE Traditional Analysis Summary

Speed Bin	Average Absolute Speed Error (<10mph)		Speed Error Bias(<5mph)		Number of 5 Minute Samples
	1.96 SEM Band	Mean	1.96 SEM Band	Mean	
0-15 MPH	2.60	5.53	2.57	5.23	239
15-25 MPH	1.76	4.7	1.24	2.91	2263
25-35 MPH	1.13	4.02	0.60	1.79	3903
>35 MPH	1.50	4.21	-0.74	-1.53	16868
All Speeds	1.47	4.24	-0.29	-0.47	23273

ES Table 3 – INRIX Traditional Analysis Summary

Speed Bin	Average Absolute Speed Error (<10mph)		Speed Error Bias (<5mph)		Number of 5 Minute Samples
	1.96 SEM Band	Mean	1.96 SEM Band	Mean	
0-15 MPH	3.25	6.58	3.22	6.47	239
15-25 MPH	2.10	5.27	1.97	4.56	2263
25-35 MPH	1.55	4.71	1.16	3.07	3903
>35 MPH	1.40	4.08	0.10	0.48	16868
All Speeds	1.51	4.33	0.49	1.37	23273

ES Table 4 – TomTom Traditional Analysis Summary

Speed Bin	Average Absolute Speed Error (<10mph)		Speed Error Bias (<5mph)		Number of 5 Minute Samples
	1.96 SEM Band	Mean	1.96 SEM Band	Mean	
0-15 MPH	1.17	4.07	1.08	3.77	239
15-25 MPH	1.50	4.46	1.23	3.23	2263
25-35 MPH	1.16	4.11	0.73	2.35	3905
>35 MPH	1.16	3.73	0.37	1.11	16882
All Speeds	1.20	3.87	0.52	1.55	23289

Sampled Distribution Method Results

The sampled distribution method seeks to quantify the extent to which vendor data captures repeatable, time-of-day travel time patterns. To this end, WRTM and vendor travel time records are each organized by validation segment and time-of-day (in this case one-hour periods from 6AM-8PM on weekdays), and used to generate cumulative frequency diagrams (CFD) and compute percentile-based performance measures (PM). The resulting Vendor-based CFD diagrams and performance measures are compared to their WRTM counterparts for evaluation.

Three performance measures shown in ES Table 5 are used throughout the report to characterize the central tendency and variation of travel times during each time-of-day period, and were chosen because they are intuitive and relevant to transportation planning applications. When communicating the difference between Vendor and WRTM performance measures for each hourly period and location, the delta symbol (Δ) is used to represent the signed difference between vendor and WRTM values (e.g., $\Delta\text{MTTI} = \text{MTTI}_{\text{Vendor}} - \text{MTTI}_{\text{WRTM}}$). Performance Measure values (as well as deltas between Vendor and WRTM PM values) are expressed with respect to free-flow travel time as a way to normalize across segments.

ES Table 5 – Performance Measures used to summarize travel time central tendency and reliability

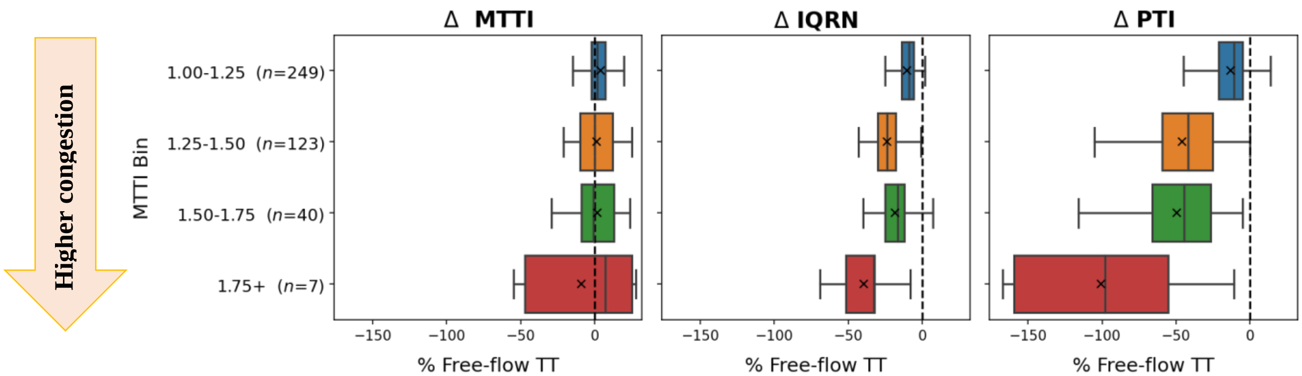
Performance Measure (PM)	Median Travel Time Index (MTTI)	IQR Normalized Travel Time (IQRN)	Planning Time Index (PTI)
PM Definition	$\frac{TT_{50}}{TT_{FF}}$	$\frac{(TT_{75} - TT_{25})}{TT_{FF}}$	$\frac{TT_{95}}{TT_{FF}}$
PM Type	Central Tendency	Reliability / Variation	Reliability / Variation

ES Figures 1-3 summarize the distribution of delta values when traffic is characterized by different levels of congestion (i.e., by MTTI bin). Note that the analysis is intended to provide high level insight into vendors' ability to capture recurrent travel time patterns, *not measure compliance.* The following bullet points provide guidance for interpretation:

- Small Δ MTTI values indicate agreement between WRTM and Vendor median travel times for each time-of-day period, while large values show a positive or negative bias.

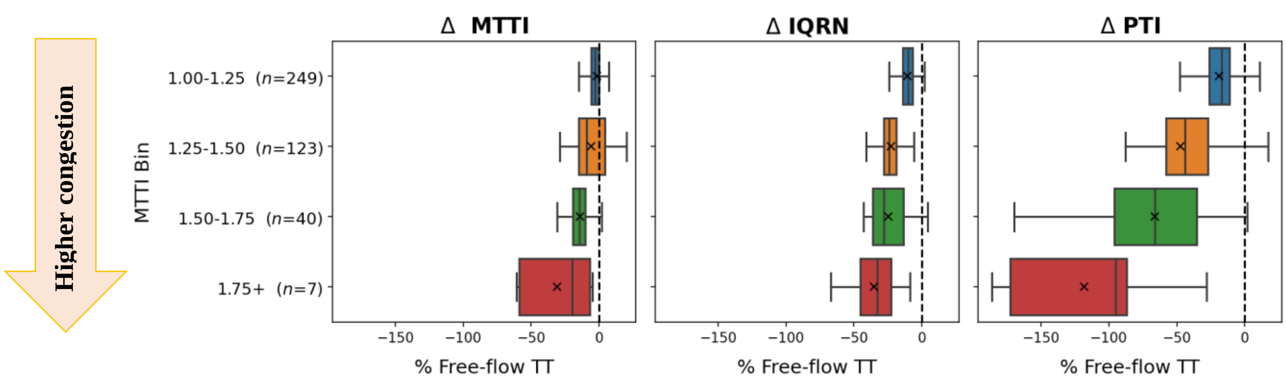
- Small Δ IQR and Δ PTI values indicate agreement between WRTM and Vendor travel time variation / reliability measures for each time-of-day period, while large values show a difference in measured variation.
    - <u>NOTE</u>: Vendors report 1-minute average travel times, and thus will naturally show less variation relative to WRTM data, which considers all observations.

ES Figure 1 shows that **HERE's median travel times agree with the reference data for most periods** (small Δ MTTI)**,** and do not show a consistent positive or negative bias. Travel time reliability is most accurate (small Δ IQR, Δ PTI) at low levels of congestion, and tends to underestimate true variation at higher levels of congestion.



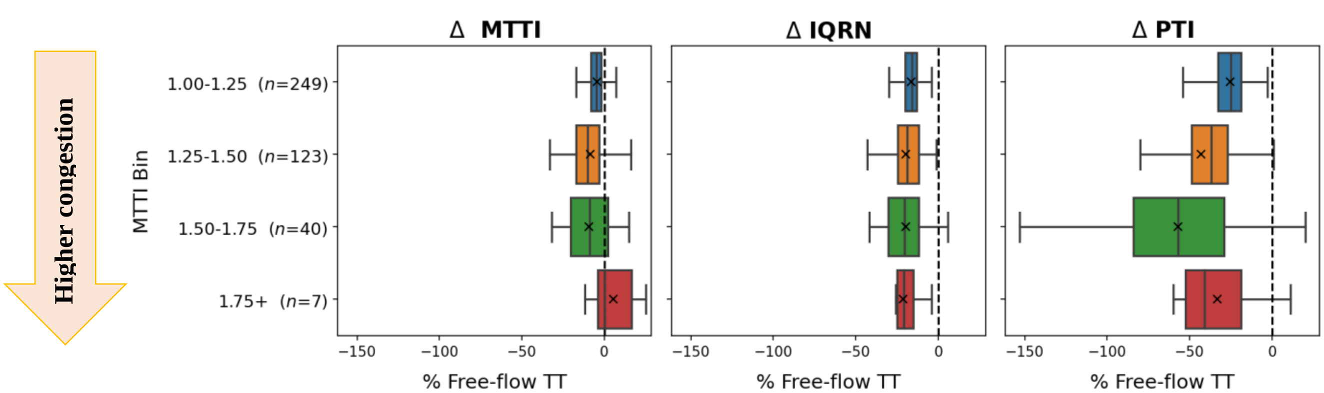
ES Figure 1 - HERE summary of delta PM values grouped by congestion level

ES Figure 2 shows that **INRIX's median travel times agree with the reference data for most periods** (small Δ MTTI)**,** with moderate negative bias in the highest congestion bin (that only represents <2% of sample periods). Travel time reliability is most accurate (small Δ IQR, Δ PTI) at low levels of congestion, and tends to underestimate true variation at higher levels of congestion.



ES Figure 2 - INRIX summary of delta PM values grouped by congestion level

ES Figure 3 shows that **TomTom median travel times agree with the reference data for most periods** (small Δ MTTI)**.** Travel time reliability is reasonably consistent across congestion levels, and tends to under-report true variation.



ES Figure 3 - TomTom summary of delta PM values grouped by congestion level

Findings and Recommendations

Overall, all three vendors met the contract compliance requirements specified for the Vehicle Probe Project, achieving the specified targets for Average Absolute Speed Error and Speed Error Bias in all required speed bins. At a high level, this indicates **that all three vendors are providing speed and travel time data in line with expectations.**

However, given the low-volume nature of the study, the traditional analysis should not be considered in isolation. The sampled distribution method is less sensitive to sample size and focuses on vendors' ability to capture repeatable travel time patterns – a perspective that is particularly relevant to planning applications. Key findings include the following points, which are not specific to any one vendor:

- Median Vendor and WRTM travel time values tend to match well for most time-of day periods, indicating that **the data would likely be useful for planning applications that require accurate "typical" travel time (or speed) values**.

- Whenever there is significant variation in travel times during a time-of-day period, the vendors usually fail to capture the full extent of variation. However, this does not necessarily reflect poor accuracy; vendors report data in 1-minute averages, and thus naturally exhibit less variation than individual WRTM travel time observations.

In the future, the validation team expects to continue using a variation of the Sampled Distribution Method to complement the point-in-time accuracy measures typically employed for validation – particularly for low-volume scenarios and along arterials with complex traffic flow patterns. This method will likely be refined over time based on feedback received from Coalition members.

Introduction

The University of Maryland (UMD), acting on behalf of The Eastern Transportation Coalition (formerly the I-95 Corridor Coalition, was given the responsibility of evaluating the quality of Vehicle Probe Project (VPP) data at the inception of the project in 2009. To assess the quality of travel time and speed data, UMD developed a methodology using wireless re-identification traffic monitoring (WRTM) technology, which is documented in detail in the previously referenced full report: I-95 Corridor Coalition Vehicle Probe Project: Validation of INRIX Data, January 2009.

At a high level, WRTM equipment is deployed at strategic locations along selected road segments and identifies – and later re-identifies – unique signals emitted by in-vehicle electronic equipment via Bluetooth, Wi-Fi and other technologies, thus allowing direct measurement of travel times from a sample of vehicles. Initial research conducted by UMD shows that this sampling approach is capable of accurately characterizing travel times (speeds), and thus WRTM data serves as the ground-truth data source against which reported probe speeds are compared.

In 2014, the project moved to a second phase (VPPII), during which a probe data marketplace was created. Currently there are three data vendors that provide travel time and speed data through this marketplace: HERE, INRIX, and TomTom. The purpose of this report, which is produced on a regular basis, is to continue to rigorously assess the accuracy of speeds reported by each vendor on various road segments from Coalition member states.

This particular validation effort focuses on a low-volume arterial facility. One of the challenges associated with validating probe data in low-volume conditions is that there is often insufficient WRTM travel time data to characterize ground truth traffic conditions. Accordingly, this report conducts both the traditional point-in-time analysis and also proposes a new analysis method for quantifying data accuracy in low volume conditions.

Probe Data Vendors

Data from Three probe data vendors are evaluated in this report: HERE, INRIX, and TomTom. Each vendor provides travel time and speed data along the road segments and time periods of interest, which are subsequently compared to ground truth WRTM observations in order to assess data accuracy.

Specifically, each vendor reports travel time and speed data in one-minute intervals either along road segments defined by the WRTM sensor locations (i.e., validation segmentation) or Traffic Message Channel (TMC) segments. In the latter case the TMC-based speeds must first be transformed to equivalent speeds on validation segments before a direct comparison can be made.

Methodology

The primary means of evaluating the vendor data is through the **traditional validation analysis**, which is documented in the original report (I-95 Corridor Coalition Vehicle Probe Project: Validation of INRIX Data July-September 2008) and summarized below. Additionally, supplemental analyses may be conducted depending on the road type being evaluated and observed data characteristics. The most common supplemental analyses are the **slowdown analysis**, which evaluates probe data quality during major congestion events on arterials, and the **sampled distribution method**, which evaluates vendors' ability to capture repeatable time-of-day patterns.

This report utilizes the traditional and sampled distribution methods, omitting the slowdown analysis because there were an insufficient number of congestion events. The sampled distribution method is useful in low-volume deployments such as this one, as it involves quantifying performance for weekday time-of-day periods, and thus does not require the same granularity of travel time observations.

Traditional validation analysis

Overview

The traditional validation analysis consists of comparing ground truth (i.e., WRTM) speeds against vendor speeds over five-minute intervals and quantifying the discrepancy in terms of two error metrics defined in the contract specifications.

Obtain vendor speed data along validation road segments

Road segments used for validation are defined based on WRTM sensor locations – often resulting in different segment definitions than those typically reported by the probe vendors. Accordingly, vendors may either report speeds directly on the validation road segmentation used for evaluation, or report speeds based on standard Traffic Message Channel (TMC) segments. In the latter case, equivalent vendor speeds must be obtained for the geometry specified by the WRTM sensors, which is accomplished via a trajectory reconstruction algorithm. This algorithm is described in another report[1] and works by (a) identifying the portions of vendor road segments that correspond to the validation segment, and (b) using the speeds reported on the vendor's segments during multiple time intervals to calculate the equivalent speed.

---

[1] Ali Haghani, Masoud Hamedi, Kaveh Farokhi Sadabadi, Estimation of Travel Times for Multiple TMC Segments, prepared for I-95 Corridor Coalition, February 2010 (link)

Filter and aggregate ground truth data

Raw travel time (speed) observations are first filtered to remove outliers. The filtering step is necessary because WRTM sensors sometimes re-identify vehicles that stop between sensors or record travel times from pedestrians or non-motorized vehicles that are not representative of actual traffic conditions. After the outlier observations are removed, the remaining representative observations are aggregated for each segment over five-minute intervals, and intervals with too few observations or excessive variation are discarded.

The remaining intervals are deemed suitable for evaluation of vendor probe data and are summarized in terms of (a) space-mean speed and (b) confidence band around the mean. The space-mean speed captures average ground truth traffic behavior, while the confidence band accounts for sample size and variation in the observed speeds.

Several statistical measures were initially evaluated to define the width of this uncertainty band, all of which are described and reported in the original report. Ultimately, the standard error of the mean (SEM) measure was selected due to its simplicity and sensitivity to both variability and number of observations used for calculations. The SEM is calculated as the standard deviation (SD) of the WRTM speeds divided by the square root of the number WRTM data points (n) taken for a given time. In other words, $\text{SEM} = \frac{SD_{WRTM}}{\sqrt{n}}$. A band based on this statistic (i.e., 1.96 times the SEM band) narrows when there is a higher degree of confidence in the ground truth data (i.e., more observations or less variation) and widens when there is less confidence, serving as a proxy for a 95% confidence interval of ground truth speeds.

Compute Error Metrics

A statistical analysis of the data is conducted for four defined speed bins, where each five-minute interval is associated with a speed bin based on its corresponding ground truth space-mean speed (0-15 mph, 15-30 mph, 30-45 mph, 45+ mph for arterials; 0-30 mph, 30-45 mph, 45-60 mph, 60+ mph for freeways). Reported probe speeds are compared to both the space-mean and SEM band ground truth speed for each five-minute time interval, and the discrepancies are quantified in terms of two error metrics: Average Absolute Speed Error (AASE) and Speed Error Bias (SEB), which are reported separately for each speed bin. According to contract specifications, AASE and SEB values must be within 10 mph and 5 mph, respectively, when compared with the SEM band.

AASE is calculated by summing up the absolute difference between probe vendor speeds (S_P) and ground truth speeds (S_{GT}) for each time interval and taking the average over n observations. That is, $\text{AASE} = \frac{1}{n}\sum_{i=1}^{n}|S_P - S_{GT}|$. Because the absolute value is used, positive and negative errors cannot cancel, and the result is always positive. Speed Error Bias is calculated similarly, with the difference that the absolute value of the errors is not taken. In other words, $\text{SEB} = \frac{1}{n}\sum_{i=1}^{n} S_P - S_{GT}$. Thus, positive and negative errors can cancel each other out, and the resulting value can provide insight into whether there is a consistent positive or negative error.

Sampled Distribution Method

In contrast to the traditional analyses that aggregate data in 5-minute bins and quantify data quality for each point in time, the sampled distribution method seeks to quantify the extent to which probe data captures *repeatable* travel time patterns. Accordingly, it aggregates the data over multiple days, grouping WRTM travel time observations and Vendor records into time-of-day periods (e.g., 7-8 AM on weekdays), using overlay plots and cumulative frequency diagrams (CFD) to summarize the travel time distributions of each analysis period. By comparing CFD curves for WRTM and VPP data – both visually and via percentile-based statistics, this method seeks to quantify the extent to which probe data captures repeatable traffic phenomena.

Overlay Plots and Cumulate Frequency Diagrams

Overlay plots are constructed by taking multiple days of observation and graphing them on a single 24-hour timeline, a process that typically only includes weekday data. This method of overlaying data on a single 24-hour plot reinforces repeatable traffic phenomena, enhancing the density of travel-time samples and thus increasing the visible detail of any recurring congestion. Figure 1 (left panel) shows an example of an overlay plot for an arterial segment, where each travel time data point collected using WRTM equipment on a weekday from an approximate two-week period is graphed on a 24-hour timeline. The relative density of the data provides a visual indication of the probability of traversing the corridor at the travel time indicated on the y-axis. A corresponding distribution analysis is constructed from the data in the overlay plot. Each curve in a distribution analysis (called an empirical cumulative frequency diagram or CFD) is constructed from the percentiles of the travel time data in the overlay plot. Figure 1 shows example CFD curves based on the overlay plot travel time records.

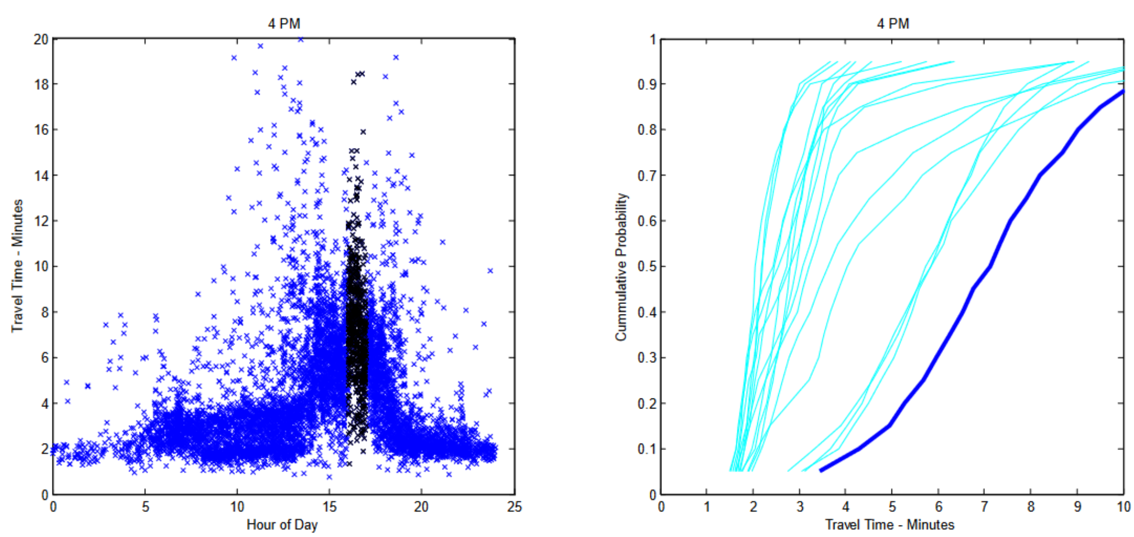
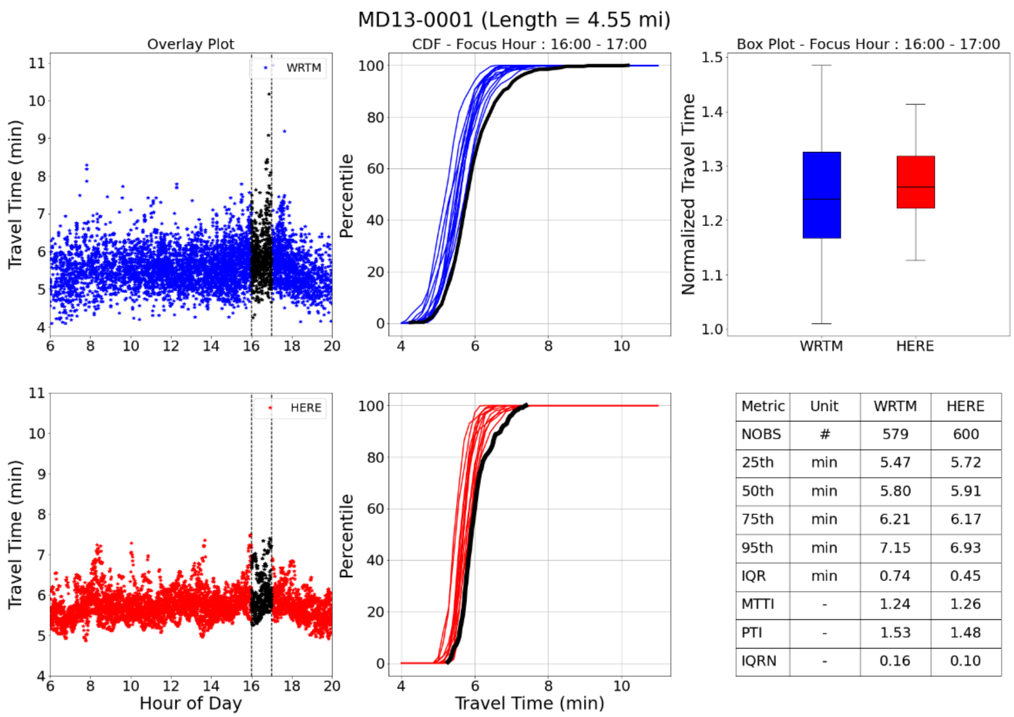


Figure 1 - Example 24-hr overlay plot and cumulative frequency diagram.

Comparing WRTM and Vendor Distributions

CFD curves and percentile-based performance measures are generated separately for both WRTM and Vendor data, and then compared for each validation segment and time of day period. Figure 2 provides an example of the visuals that are produced for each hourly period to facilitate this comparison. Starting from the left, the first column of figures shows 24-hour travel time overlay plots for WRTM (top left) and Vendor (bottom left), with the target hour bounded by dotted vertical lines. The CFD curves associated with each hourly period are summarized in the middle column, with the target hour's curve (which corresponds to travel time observations between the dotted vertical lines in the overlay plot) highlighted in black. The right column of figures contains boxplots of the travel time distribution for the target hour (upper right plot) and a summary table of relevant travel time percentiles and percentile-based performance measures (bottom right).



Metric	Unit	WRTM	HERE
NOBS	#	579	600
25th	min	5.47	5.72
50th	min	5.80	5.91
75th	min	6.21	6.17
95th	min	7.15	6.93
IQR	min	0.74	0.45
MTTI	-	1.24	1.26
PTI	-	1.53	1.48
IQRN	-	0.16	0.10

Figure 2 - Overlay plots (left), cumulative frequency curves (middle), and boxplots/stats (right).

Three unitless performance measures (MTTI, IQRN, PTI) are included in the summary table and are based on the median travel time, the IQR of travel time and the Planning Time (the 95th percentile "worst case" travel time) – all normalized by the free-flow travel time. The idea behind using normalized performance measure values is that they can be easily compared across validation segments that have different lengths and allowable traffic speeds. These performance measures are defined as follows, where TT_X is the Xth percentile travel time, and TT_{FF} is the free-flow travel time obtained by taking the 15th percentile travel time during overnight hours (10pm – 5am):

- **Median Travel Time Index (MTTI)**: $\mathrm{MTTI} = \frac{\mathrm{TT}_{50}}{\mathrm{TT}_{\mathrm{FF}}}$
  - Measures typical travel time during period (central tendency)
- **Inter-Quartile Range Normalized (IQRN)**: $\mathrm{IQRN} = \frac{\mathrm{TT}_{75} - \mathrm{TT}_{25}}{\mathrm{TT}_{\mathrm{FF}}}$
  - Measures how travel time varies during period (reliability)
- **Planning Time Index (PTI):** $\mathrm{PTI} = \frac{\mathrm{TT}_{95}}{\mathrm{TT}_{\mathrm{FF}}}$
  - Measures near worst-case travel time during period (reliability)

Summarizing Differences between WRTM and Vendor Performance Measures

In addition to comparing WRTM and vendor cumulative frequency curves and their performance measures for individual time periods, it is useful to summarize their differences more broadly across all locations and times of day to draw high level insights. In order to do so, we compute Vendor and WRTM performance measures during each period, calculate the difference between the measures, and then aggregate the results under different traffic conditions, times of day, and locations.

Specifically, the delta symbol (Δ) is used to represent the signed difference between vendor and WRTM performance measure values (e.g., $\Delta\,\mathrm{MTTI} = \mathrm{MTTI}_{\mathrm{Vendor}} - \mathrm{MTTI}_{\mathrm{WRTM}}$). Because each PM (MTTI, IQRN, PTI) is normalized with respect to free-flow travel time, the delta values are also unitless and represent the fraction of the free-flow travel time (or equivalently a percentage of free flow travel time), making it easy to compare delta values across road segments with different speed limits and lengths. Common aggregation strategies include:

- **Overall:** Aggregated over all hours and validation paths

- **By MTTI Bin (level of congestion)**: Each hourly period is assigned a MTTI bin based on the WRTM's MTTI value. Suggested ranges include: 1.0 – 1.25, 1.25 – 1.5, 1.5 – 1.75, 1.75+, with 1.0 – 1.25 representing minimal congestion and > 1.75 representing higher levels of congestion.

- **By IQRN Bin (level of variation)**: Each hourly period is assigned a IQRN bin based on the WRTM's IQRN value. Suggested ranges include: 0 – 0.25, 0.25 – 0.5, 0.5 – 0.75, > 0.75, with 0 – 0.25 representing minimal travel time variation and > 0.75 representing high levels of variation.

- **By Time of Day:** Grouped by hour of day or pre-defined periods (e.g., peak/off-peak)

- **By Validation Path**: Grouped by validation segment

Using any of these grouping/aggregation strategies, the resulting Δ PM values (i.e., Δ MTTI, Δ IQRN, Δ PTI) can be summarized in terms of a single value (e.g., the average or median deltas) or visuals that show the distribution of values in each group. For example, Table 1 reports average delta values for all performance measures when aggregated by MTTI bin, while Figure 3 shows the same aggregation but provides more information about the delta values by showing their distributions via boxplots.

Table 1: Sample summary of average delta values aggregated by MTTI Bin

Performance Measure (PM)	Median Travel Time Index (MTTI)	IQR Normalized Travel Time (IQRN)	Planning Time Index (PTI)	
PM Definition	$\frac{TT_{50}}{TT_{FF}}$	$\frac{(TT_{75} - TT_{25})}{TT_{FF}}$	$\frac{TT_{95}}{TT_{FF}}$	
	$\overline{\Delta\ \text{MTTI}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{IQRN}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{PTI}}$ (% $TT_{\text{Free flow}}$)	Samples
MTTI Bin				
1.00 – 1.25	-1	-1	-1	267
1.25 – 1.50	-15	-15	-15	69
1.50 – 1.75	-9	-9	-9	50
1.75 +	-19	-19	-19	30

Delta (*Δ*): Difference in hourly performance measure values (Vendor -WRTM), expressed as percent of free-flow TT
Avg Delta ($\bar{\Delta}$): Average delta value computed using relevant 1-hr time-of-day samples, rounded to the nearest percent

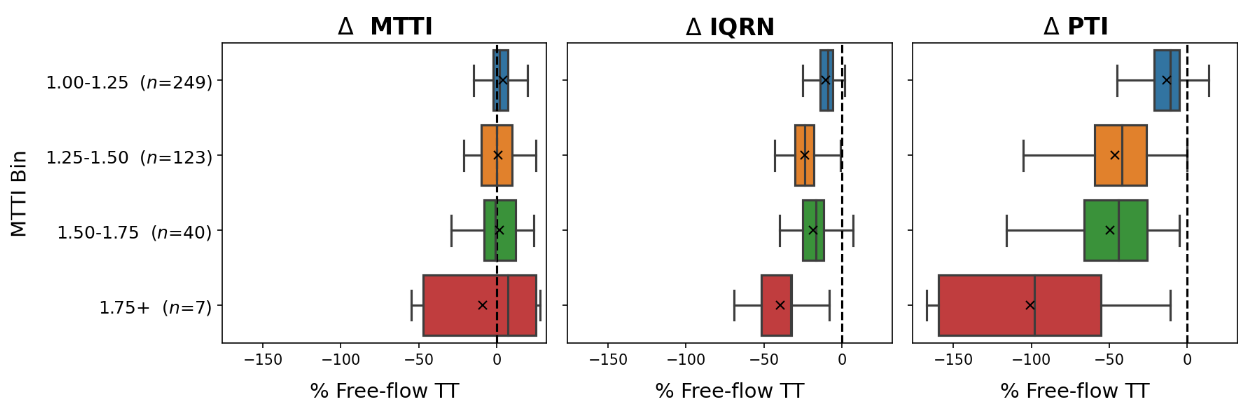


Figure 3: Sample distribution of delta values when aggregated by MTTI Bin

The following comments provide guidance on interpreting the resulting delta PM measures:

- This analysis is intended to provide high level insight into vendors' ability to capture recurrent travel time patterns, *not measure compliance.* Currently there are no agreed-upon specifications for this analysis that determine "good" or "bad" performance.

- Small Δ MTTI values indicate agreement between WRTM and Vendor median travel times for each time-of-day period, while large values show a positive or negative bias. Achieving a small Δ MTTI tends to be easier when traffic is close to free-flow conditions (minimal congestion) and more difficult when there is congestion and/or variation in travel times.

- Small Δ IQR and Δ PTI values indicate agreement between WRTM and Vendor travel time variation / reliability measures for each time-of-day period, while large values show a difference in measured variation. Achieving small Δ IQR and Δ PTI values is easier when there is minimal congestion or variation in travel times (because there is little variation to measure), and more difficult when there is higher variation.
    - NOTE: WRTM captures the travel times of individual vehicles, while vendors report average travel times for each 1-minute (without any indication of variation during the period). Although Vendor data *does* exhibit variation in the 1-minute average values over time-of-day periods, there often less variation than in WRTM observations due to the averaging inherent in the product. Thus, discrepancies in Δ IQR and Δ PTI, while useful for understanding how well Vendor data captures true travel time variation, may be partially caused by Vendor data aggregation rather than poor performance.

- Performance Measure (PM) values are expressed with respect to free-flow travel time as a way to normalize across different road segments.
    - E.g., MTTI = 1.3 implies that the median travel time is 1.3 times (or 130% of) the time to traverse a segment at free-flow speeds.
    - E.g., Δ PTI = -0.7 (or -70 % TT_{FF}) means that the Vendor's PTI value is less than the WRTM PTI value by 70% of free-flow travel time. If the free-flow travel time is 1 minute, then this implies that the Vendor 95^{th} percentile travel time is 0.7 minutes (42 seconds) less than the WRTM 95^{th} percentile travel time.

Data Collection

Travel time samples were collected along 30 directional validation road segments in Maryland from July 20 through July 31, 2020. These segments stretch from Old Hanover Rd to Mt Ventus Rd Number 2 on US-30, from Western Maryland Pkwy to N Edgewood Dr on US-40, and from E Wilson Blvd to Club Rd on US-11, and are defined based on WRTM sensor locations, which are shown in Figure 4.

Table 2 contains the summary information for each data collection segment, including WRTM sensor latitude/longitudes and an active map link, which can be followed to view each data collection segment in detail. Please note that the configuration of the test segments is often such that the endpoint of one segment coincides with the start point of the next segment, so that one WRTM sensor covers both data collection segments.

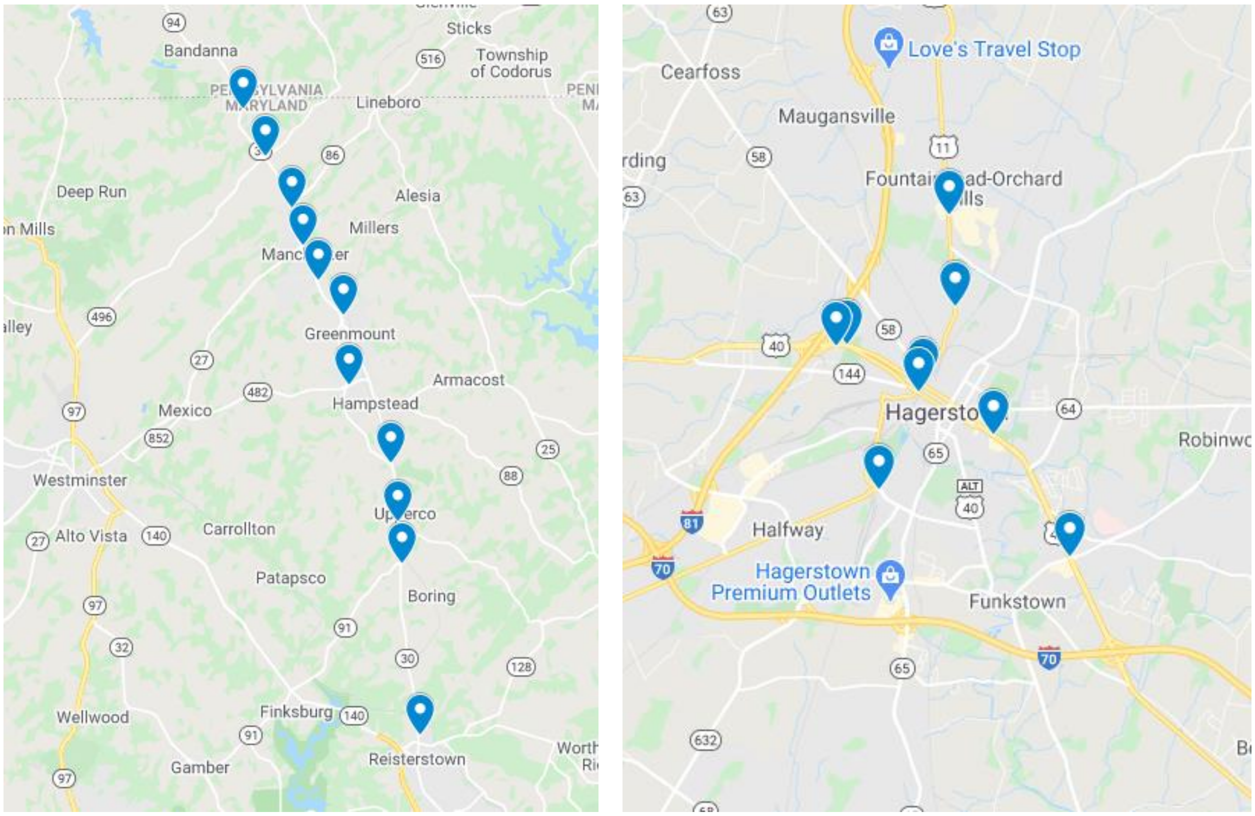


Figure 4 – WRTM Sensor locations

Table 2 - Validation Segment Attributes

Segment (Map Link)	DESCRIPTION						Deployment		
	Highway	Starting at	Lane (Min)	Highway	AADT	Access Points	Begin Lat/Lon		Length (mile)
	Direction	Ending at	Lane (Max)	Direction		Speed Limit	End Lat/Lon		
Arterial									
A1	MD-30	Old Hanover Rd	2	0	17947	27	39.47860	-76.82915	4.55
MD13-0001	Northbound	Emory Rd	2	0		40	39.54399	-76.83765	
A2	MD-30	Emory Rd	2	2	19159	5	39.54399	-76.83765	1.09
MD13-0002	Northbound	Dover Rd	2	1.83		50	39.55958	-76.83973	
A3	MD-30	Dover Rd	2	0	19004	10	39.55958	-76.83973	1.62
MD13-0003	Northbound	S Main St	4	0		50	39.58211	-76.84312	
A4	MD-30	S Main St	2	0	12384	5	39.58211	-76.84312	2.58
MD13-0004	Northbound	Hampstead Mexico Rd	4	0		55	39.61141	-76.86432	
A5	MD-30	Hampstead Mexico Rd	2	1	13751	3	39.61141	-76.86432	2.00
MD13-0005	Northbound	Eagle ridge Ct	4	0.50		55	39.63767	-76.86689	
A6	MD-30	Eagle ridge Ct	2	1	20401	11	39.63767	-76.86689	1.17
MD13-0006	Northbound	Maple Grove Rd	2	0.86		45	39.65122	-76.87925	
A7	MD-30	Maple Grove Rd	2	2	20365	14	39.65122	-76.87925	0.99
MD13-0007	Northbound	Manchester Rd	2	2.02		30	39.66442	-76.88656	
A8	MD-30	Manchester Rd	2	1	17446	13	39.66442	-76.88656	1.05
MD13-0008	Northbound	Hallie Ave	2	0.95		40	39.67866	-76.89176	
A9	MD-30	Hallie Ave	2	1	14314	14	39.67866	-76.89176	1.56
MD13-0009	Northbound	Wentz Rd	2	0.64		40	39.69814	-76.90478	
A10	MD-30	Wentz Rd	2	0	13001	12	39.69814	-76.90478	1.51
MD13-0010	Northbound	Mt ventus Rd Number 2	2	0		50	39.71626	-76.91577	
A11	MD-30	Mt ventus Rd Number 2	2	0	13001	12	39.71626	-76.91577	1.51
MD13-0011	Southbound	Wentz Rd	2	0		50	39.69814	-76.90478	

Segment (Map Link)	DESCRIPTION						Deployment		
	Highway Direction	Starting at Ending at	Lane (Min) Lane (Max)	Signals Signal/mile	AADT	Access Points Speed Limit	Begin Lat/Lon End Lat/Lon		Length (mile)
Arterial									
A12 MD13-0012	MD-30 Southbound	Wentz Rd Hallie Ave	2 2	1 0.64	14314	14 40	39.69814 39.67866	-76.90478 -76.89176	1.56
A13 MD13-0013	MD-30 Southbound	Hallie Ave Manchester Rd	2 2	1 0.95	17446	13 45	39.67866 39.66442	-76.89176 -76.88656	1.05
A14 MD13-0014	MD-30 Southbound	Manchester Rd Maple Grove Rd	2 2	2 2.02	20365	14 45	39.66442 39.65122	-76.88656 -76.87925	0.99
A15 MD13-0015	MD-30 Southbound	Maple Grove Rd Eagle ridge Ct	2 4	1 0.86	20401	11 40	39.65122 39.63767	-76.87925 -76.86689	1.17
A16 MD13-0016	MD-30 Southbound	Eagle ridge Ct Hampstead Mexico Rd	2 4	1 0.50	13685	3 55	39.63767 39.61137	-76.86689 -76.86461	2.00
A17 MD13-0017	MD-30 Southbound	Hampstead Mexico Rd S Main St	2 4	0 0	12384	5 55	39.61137 39.58195	-76.86461 -76.84319	2.57
A18 MD13-0018	MD-30 Southbound	S Main St Dover Rd	2 4	0 0	18966	10 50	39.58193 39.55958	-76.84315 -76.83973	1.60
A19 MD13-0019	MD-30 Southbound	Dover Rd Emory Rd	2 3	2 1.83	19159	5 50	39.55958 39.54399	-76.83973 -76.83765	1.09
A20 MD13-0020	MD-30 Southbound	Emory Rd Old Hanover Rd	2 3	0 0	17947	27 50	39.54399 39.47860	-76.83765 -76.82915	4.55
A21 MD13-0021	US-40 Westbound	N Edgewood Dr S Cleveland Ave	4 4	4 2.28	38234	15 45	39.61673 39.63777	-77.69248 -77.70981	1.76
A22 MD13-0022	US-40 Westbound	S Cleveland Ave N Burhans Blvd	2 4	9 8.41	13227	21 40	39.63777 39.64659	-77.70981 -77.72606	1.07
A23 MD13-0023	US-40 Westbound	N Burhans Blvd Western Maryland Pkwy	2 4	6 5.85	11732	20 55	39.64659 39.65341	-77.72606 -77.74305	1.03

Segment (Map Link)	DESCRIPTION						Deployment		
	Highway Direction	Starting at Ending at	Lane (Min) Lane (Max)	Signals Signal/mile	AADT	Access Points Speed Limit	Begin Lat/Lon End Lat/Lon		Length (mile)
Arterial									
A24 MD13-0024	US-40 Eastbound	Western Maryland Pkwy W Washington St	2 4	7 6.04	11682	20 55	39.65314 39.64518	-77.74542 -77.72684	1.16
A25 MD13-0025	US-40 Eastbound	W Washington St S Cleveland Ave	2 4	6 5.71	13227	17 35	39.64518 39.63758	-77.72684 -77.70994	1.05
A26 MD13-0026	US-40 Eastbound	S Cleveland Ave N Edgewood Dr	4 4	4 2.28	38275	17 45	39.63758 39.61659	-77.70994 -77.69274	1.75
A27* MD13-0027	US-11 Northbound	Virginia Ave W Washington St	2 2	6 3.64	12371	13 30	39.62847 39.64522	-77.73566 -77.72694	1.65
A28 MD13-0028	US-11 Northbound	W Washington St W Irvin Ave	2 4	4 3.55	15336	10 25	39.64522 39.65982	-77.72694 -77.71842	1.13
A29 MD13-0029	US-11 Northbound	W Irvin Ave Club Rd	2 2	3 2.70	15634	8 30	39.65982 39.67588	-77.71842 -77.71989	1.11
A30 MD13-0030	US-11 Southbound	Club Rd W Irvin Ave	2 2	3 2.70	15634	8 30	39.67588 39.65982	-77.71989 -77.71842	1.11
A31 MD13-0031	US-11 Southbound	W Irvin Ave W Washington St	2 4	4 3.55	15336	10 25	39.65982 39.64523	-77.71842 -77.72694	1.13
A32* MD13-0032	US-11 Southbound	W Washington St Virginia Ave	2 2	6 3.62	12371	13 30	39.64522 39.62849	-77.72694 -77.73589	1.66

* *Segment omitted due to a sensor placement issue*

Traditional Validation Results

Tables 3-5 summarize the standard error metrics obtained by comparing Vendor speeds to reference (WRTM) speeds over many 5-minute periods. Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) metrics are computed in four different speed bins (0-15, 15-25, 25-35, 35+ MPH) and reported separately based on whether reference speeds are represented by (i) the WRTM mean or (ii) the standard error of the mean (SEM) band in the calculations. The columns relevant to contract compliance (i.e., comparisons with respect to the SEM band) are outlined in red, and error metrics in these columns colored green or red to indicate whether the value is within contract specifications (AASE <= 10 mph, SEB <= +/-5 mph).

In previous reports, these AASE and SEB error metrics have also been reported separately for each validation segment. For brevity – and to focus on the sampled distribution method due to the low-volume nature of the study – these tables are omitted from this report, but can be made available upon request.

HERE

Table 3 reports the traditional analysis error metrics for HERE. The Average Absolute Speed Error (AASE) is within specification for all speed bins when compared to both SEM Band and Mean. Speed Error Bias (SEB) is within specification in all speed bins when compared to the SEM Band and in the upper three speed bins when compared to the Mean.

Table 3 – HERE traditional analysis data quality measures

Speed Bin	Average Absolute Speed Error (<10mph)		Speed Error Bias (<5mph)		Number of 5 Minute Samples
	Compared to 1.96 SEM Band	Compared to Mean	Compared to 1.96 SEM Band	Compared to Mean	
0-15 MPH	2.60	5.53	2.57	5.23	239
15-25 MPH	1.76	4.7	1.24	2.91	2263
25-35 MPH	1.13	4.02	0.60	1.79	3903
>35 MPH	1.50	4.21	-0.74	-1.53	16868
All Speeds	1.47	4.24	-0.29	-0.47	23273

INRIX

Table 4 reports the traditional analysis error metrics for INRIX. The Average Absolute Speed Error (AASE) is within specification for all speed bins when compared to both SEM Band and Mean. Speed Error Bias (SEB) is within specification in all speed bins when compared to the SEM Band and in the upper three speed bins when compared to the Mean.

Table 4 - INRIX traditional analysis data quality measures

Speed Bin	Average Absolute Speed Error (<10mph)		Speed Error Bias (<5mph)		Number of 5 Minute Samples
	Compared to 1.96 SEM Band	Compared to Mean	Compared to 1.96 SEM Band	Compared to Mean	
0-15 MPH	3.25	6.58	3.22	6.47	239
15-25 MPH	2.10	5.27	1.97	4.56	2263
25-35 MPH	1.55	4.71	1.16	3.07	3903
>35 MPH	1.40	4.08	0.10	0.48	16868
All Speeds	1.51	4.33	0.49	1.37	23273

TomTom

Table 5 reports the traditional analysis error metrics for TomTom. Both the Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) are within specification for all speed bins when compared to the SEM Band and Mean.

Table 5 - TomTom traditional analysis data quality measures

Speed Bin	Average Absolute Speed Error (<10mph)		Speed Error Bias (<5mph)		Number of 5 Minute Samples
	Compared to 1.96 SEM Band	Compared to Mean	Compared to 1.96 SEM Band	Compared to Mean	
0-15 MPH	1.17	4.07	1.08	3.77	239
15-25 MPH	1.50	4.46	1.23	3.23	2263
25-35 MPH	1.16	4.11	0.73	2.35	3905
>35 MPH	1.16	3.73	0.37	1.11	16882
All Speeds	1.20	3.87	0.52	1.55	23289

Sampled Distribution Method Results

The following strategy is used to analyze the sampled distribution results for each vendor. First, the distribution of delta performance measure values (i.e., Δ MTTI, Δ IQRN, Δ PTI) is summarized across all validation paths and time-of day periods, and additionally partitioned into different levels of congestion and travel time variation to gain additional insight. For each scenario, average delta values are reported in tables, while the entire distribution is summarized visually via boxplots. In each boxplot, the line in the middle of the box represents the median value, while the 'x' represents the mean value, and the sample size (representing the number of hourly periods) is indicated by the text on the label (e.g., n=200).

Afterwards, the delta values are separated by validation path, resulting in Δ MTTI, Δ IQRN, and Δ PTI boxplots for each path. The first 20 paths (MD13-0001 through MD13-0020, labeled P01-P20 in the plots) correspond to a single corridor (MD-30), while the remaining ones correspond to US-40 (P21-P26) and US-11 (P28-P31). Each path-level boxplot summarizes the performance measure comparison over all hourly periods with valid data, and gives a quick visual indication of how closely vendor data matches WRTM in terms of central tendency and variation / reliability in each location.

Finally, the detailed overlay plots and CFD curves are shown for a single representative path: MD13-0007, which reinforces the underlying methodology for computing the performance measures. MD13-0007 – shown below in Figure 5 -- is located along MD-30 in the Northbound direction, and passes through the town of Manchester. It is 1 mile in length with an AADT of about 20k (bidirectional, so approximately 10k in the Northbound direction), and has 2 traffic signals and 14 access points, thus providing many opportunities for vehicles to enter and exit the road. Figure 6 shows a brief overview of the WRTM travel times across the data collection period. It indicates that travel times range from 2-4 minutes over the course of each day, and occasionally reach 6-10 minutes during periods of congestion.

Figure 5 – Segment boundaries and screenshot for representative validation path MD13-0007.

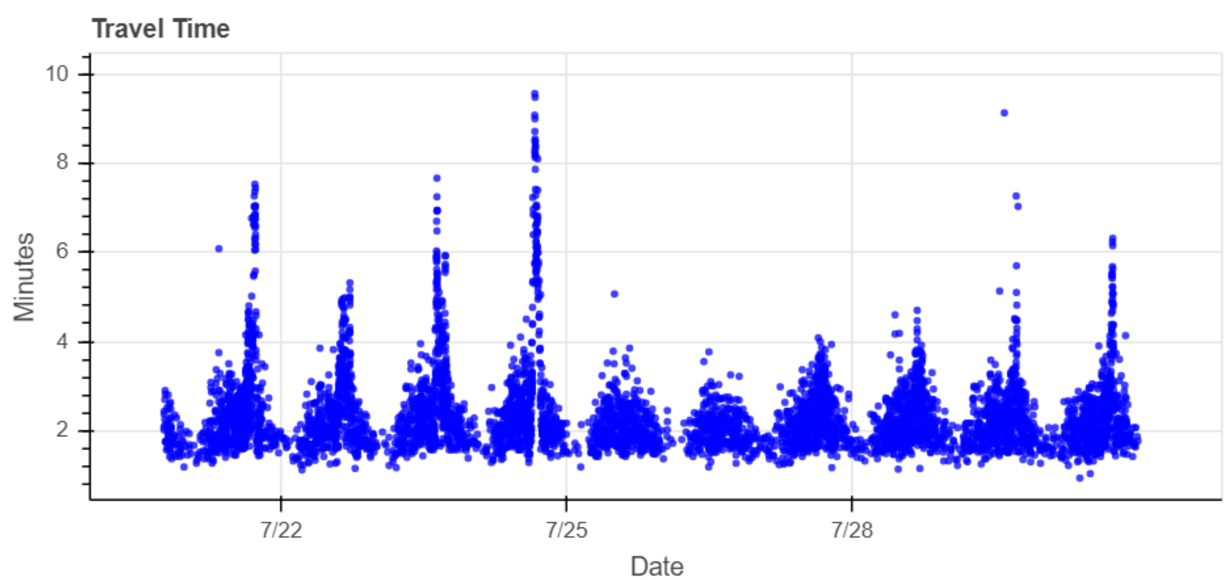


Figure 6 – Travel times along path MD13-0007 over data collection period.

Note that high travel times can be legitimately attributed to a variety of causes, including cycle failures at traffic signals, friction from vehicles entering/exiting at various access points or trying to parallel park, pedestrians crossing the street, etc. Likewise, because it is not a limited access facility, vehicles can also stop briefly (e.g., pull off the road for coffee, gas) and re-enter in time to traverse the segment in a reasonable amount of time. Distinguishing between these scenarios can be difficult, and currently requires a combination of an automated filtering algorithm and human judgement.

When interpreting the results, it is import to remember the following:

- Vendors report data in one-minute average values while WRTM data is based on individual observations, meaning that in most cases the WRTM data will show exhibit more variation.

- While care was taken to filter the WRTM data in a reasonable manner, the filtering strategy can play a role in the performance measure values on roadways with complex traffic flow – particularly PTI, which uses the 95th percentile travel time value.

- This analysis is intended to provide high level insight into vendors' ability to capture recurrent travel time patterns, *not measure compliance.* Currently there are no agreed-upon specifications for this analysis that determine "good" or "bad" performance.

HERE

Figures 7-8 summarize the distribution of delta values for HERE under a variety of congestion and travel time reliability scenarios, while Table 6 reports the average values for the same scenarios. Observations from these figures and tables include the following points:

- **Central Tendency Comparison (Δ MTTI):**
  - Δ MTTI values are centered close to 0 for all congestion levels (MTTI bins), indicating that WRTM and vendor median travel time values match well during most periods. While all congestion levels have strong MTTI agreement on average, there are a wider range of delta values during periods of high congestion.
  - Δ MTTI values are centered around 0 for all but the highest level of travel time variation (IQRN bins), indicating strong agreement in WRTM and Vendor median travel times during these periods. However, the Vendor data noticeably underreports the median travel time in periods with high travel time variation.

- **Reliability Comparison (Δ IQRN, Δ PRI):**
  - Reliability PM values are close (i.e., small Δ IQRN and Δ PTI) when congestion and variation levels are low (low MTTI and IQRN bins). As congestion and travel time variation increase, the Vendor increasingly underreports true travel time variation (more negative delta values).

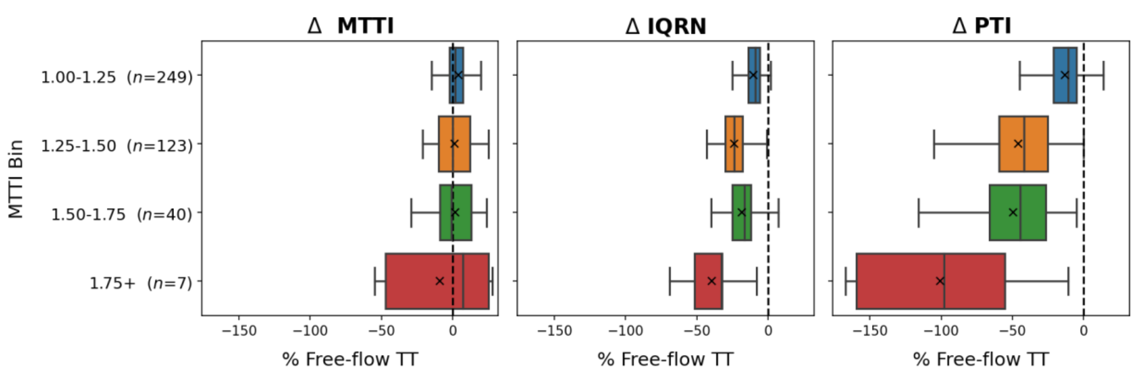


Figure 7 - HERE summary of delta PM values grouped by congestion level

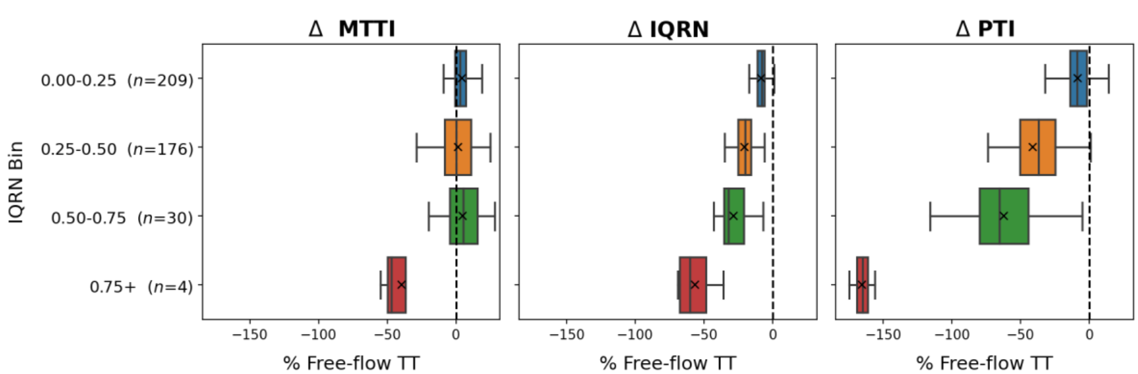


Figure 8 - HERE summary of delta PM values grouped by level of travel time variation

Table 6 - HERE summary of average delta PM values under various scenarios.

Performance Measure (PM)	Median Travel Time Index (MTTI)	IQR Normalized Travel Time (IQRN)	Planning Time Index (PTI)	
PM Definition	$\frac{TT_{50}}{TT_{FF}}$	$\frac{(TT_{75} - TT_{25})}{TT_{FF}}$	$\frac{TT_{95}}{TT_{FF}}$	
	$\overline{\Delta\ \text{MTTI}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{IQRN}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{PTI}}$ (% $TT_{\text{Free flow}}$)	Samples
Overall	3	-15	-28	419
MTTI Bin (Higher congestion)				
1.00 – 1.25	4	-10	-13	249
1.25 – 1.50	1	-24	-46	123
1.50 – 1.75	2	-18	-50	40
> 1.75	-9	-40	-101	7
IQRN Bin (Higher variation)				
0.00 – 0.25	4	-8	-9	209
0.25 – 0.50	1	-21	-41	176
0.50 – 0.75	5	-29	-62	30
> 0.75	-40	-56	-165	4

Delta (*Δ*): Difference in hourly performance measure values (Vendor -WRTM), expressed as percent of free-flow TT
Avg Delta ($\bar{\Delta}$): Average delta value computed using relevant 1-hr time-of-day samples, rounded to the nearest percent

Figure 9 shows the distribution of PM delta values across all validation paths. Most locations have small Δ MTTI values, indicating general agreement in median travel times across hourly periods. Δ IQRN values are quite consistent across locations as well, showing that the vendor data slightly underreports IQR values. Δ PTI values are much more noticeably negative - especially for paths P21-P31 – an intuitive result given that US-40 and US-11 have many traffic signals and access points. Because WRTM data is based on individual travel time observations while vendor data uses one-minute averages, it makes sense that WRTM data captures some higher travel times that cannot be captured in the average vendor data (and thus impacts the 95^{th} percentile travel time values used in the PTI performance measure computation).

Figure 10 shows detailed results for path MD13-0007, whose travel times were previously shown in Figure 6 for the entire validation period. The overlay plot shows that WRTM travel times are typically concentrated around 2-3 minutes for most hours between 6 AM and 8PM, but can reach upwards of 7-10 minutes between 3-6 PM (15-18 on the graph). The same pattern is reflected in the vendor's overlay plot, but there is less variation in travel time values and the high travel time values during 3-6 PM are lower in magnitude than was observed with WRTM (e.g., 4-5 minutes rather than 7 minutes from 5-6 PM). Because there is less variation in the vendor travel time records than WRTM for each hourly period, the delta IQRN and PTI values are noticeably negative – an expected result given that vendor data is reported in one-minute averages.

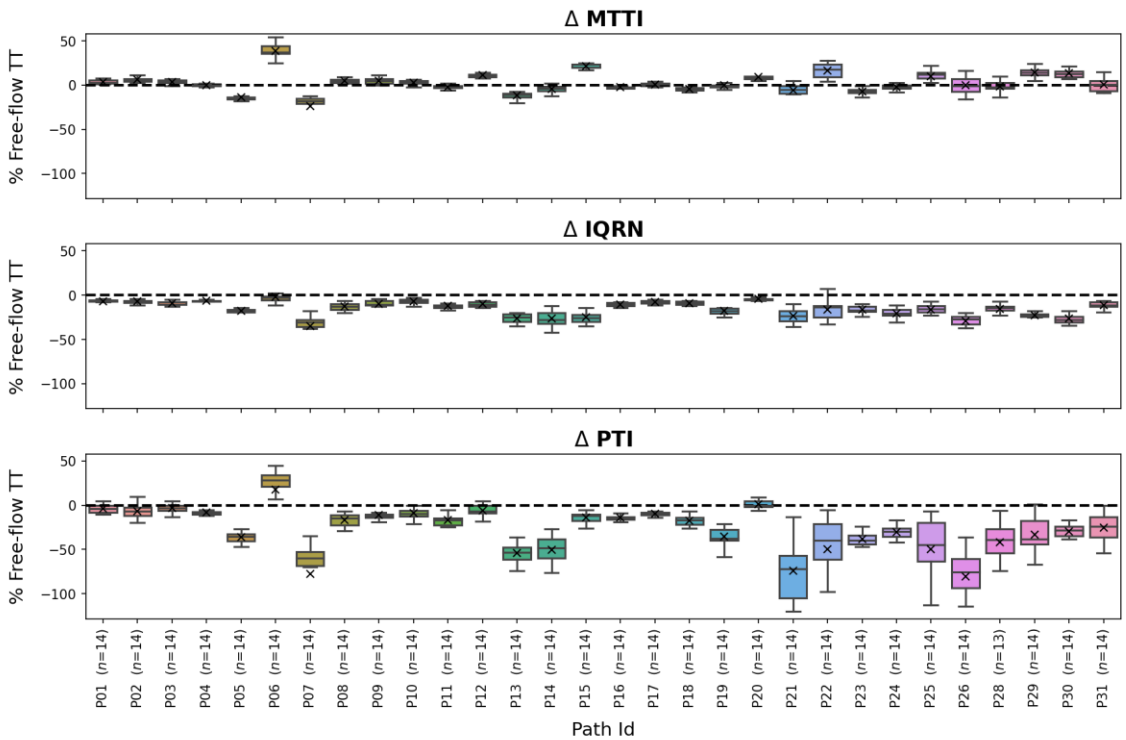
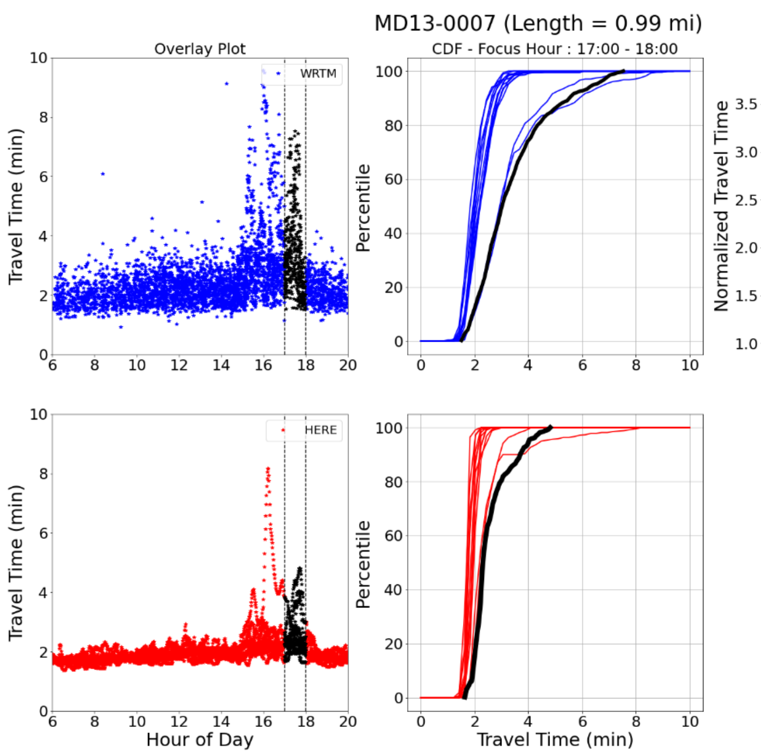


Figure 9 - HERE path-level delta PM values



Metric	Unit	WRTM	HERE
NOBS	#	418	600
25th	min	2.40	2.12
50th	min	3.02	2.31
75th	min	4.06	2.79
95th	min	6.42	4.01
IQR	min	1.66	0.67
MTTI	-	2.04	1.56
PTI	-	4.33	2.71
IQRN	-	1.12	0.45

Figure 10 – HERE overlay and CFD plots for segment MD13-0007.

INRIX

Figures 11-12 summarize the distribution of INRIX delta values under a variety of congestion and travel time reliability scenarios, while Table 7 reports the average values for the same scenarios. Observations from these figures and tables include the following points:

- **Central Tendency Comparison (Δ MTTI)**:
  - Δ MTTI values are close to 0 at the lowest congestion level (MTTI bin), and slowly become more negative at higher levels (around -25% to -30% of free-flow travel time in the highest bin). The highest bin also exhibits a greater level of Δ MTTI variation (wider box plot).
  - Δ MTTI values are close to 0 at the lowest level of travel time variation (IQRN bin), indicating strong agreement in WRTM and Vendor median travel times when travel times are stable for the analysis period. However, as more variation is observed, the delta values become more negative – particularly in the highest bin.

- **Reliability Comparison (Δ IQRN, Δ PRI)**:
  - Reliability PM values are close (i.e., small Δ IQRN and Δ PTI) when congestion is minimal (low MTTI bin) and there is minimal variation in in travel times (low IQRN bin). As congestion and travel time variation increase, the Vendor increasingly underreports true travel time variation (more negative delta values).

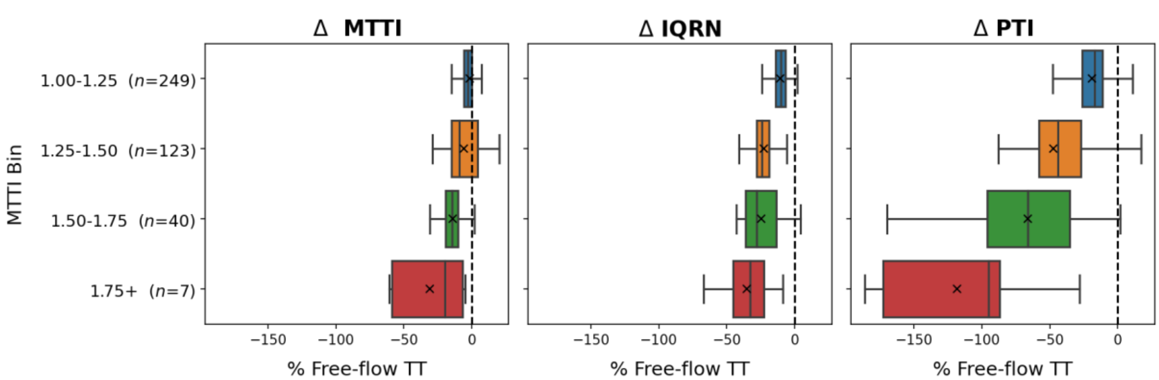


Figure 11 - INRIX summary of delta PM values grouped by congestion level

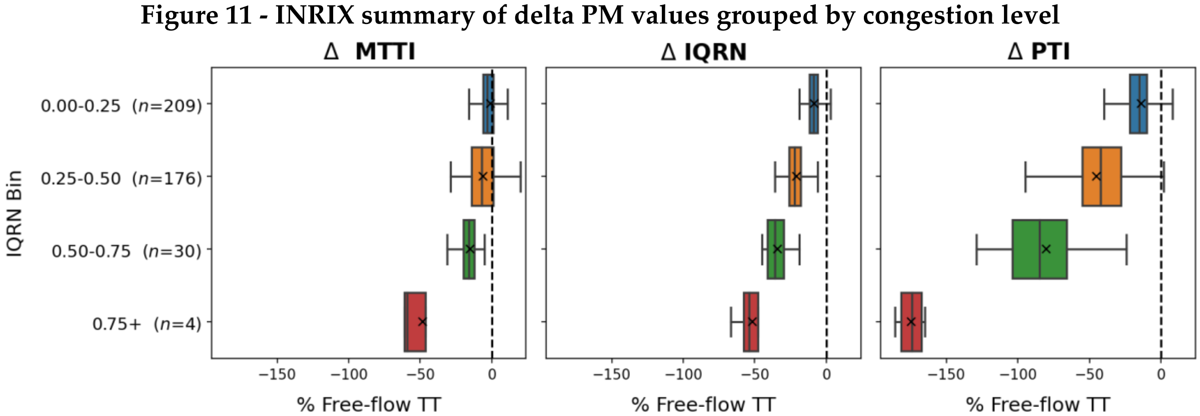


Figure 12 - INRIX summary of delta PM values grouped by level of travel time variation

Table 7 - INRIX summary of average delta PM values under various scenarios.

Performance Measure (PM)	Median Travel Time Index (MTTI)	IQR Normalized Travel Time (IQRN)	Planning Time Index (PTI)	
PM Definition	$\frac{TT_{50}}{TT_{FF}}$	$\frac{(TT_{75} - TT_{25})}{TT_{FF}}$	$\frac{TT_{95}}{TT_{FF}}$	
	$\overline{\Delta\ \text{MTTI}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{IQRN}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{PTI}}$ (% $TT_{\text{Free flow}}$)	Samples
Overall	-5	-16	-33	419
MTTI Bin (Higher congestion)				
1.00 – 1.25	-2	-11	-19	249
1.25 – 1.50	-6	-23	-47	123
1.50 – 1.75	-14	-25	-66	40
> 1.75	-31	-35	-118	7
IQRN Bin (Higher variation)				
0.00 – 0.25	-1	-9	-14	209
0.25 – 0.50	-6	-21	-45	176
0.50 – 0.75	-15	-34	-80	30
> 0.75	-49	-52	-175	4

Delta (Δ): Difference in hourly performance measure values (Vendor -WRTM), expressed as percent of free-flow TT
Avg Delta ($\overline{\Delta}$): Average delta value computed using relevant 1-hr time-of-day samples, rounded to the nearest percent

Figure 13 shows the distribution of PM delta values across all validation paths. Many locations have small Δ MTTI values, indicating general agreement in median travel times across hourly periods. Δ IQRN values are quite consistent across locations as well, showing that the vendor data slightly underreports IQR values. Δ PTI values are much more noticeably negative - especially for paths P21-P31 – an intuitive result given that US-40 and US-11 have many traffic signals and access points. Because WRTM data is based on individual travel time observations while vendor data uses one-minute averages, it makes sense that WRTM data captures some higher travel times that cannot be captured in the average vendor data (and thus impacts the 95th percentile travel time values used in the PTI performance measure computation).

Figure 14 shows detailed results for path MD13-0007, whose travel times were previously shown in Figure 6 for the entire validation period. The overlay plot shows that WRTM travel times are typically concentrated around 2-3 minutes for most hours between 6 AM and 8PM, but can reach upwards of 7-10 minutes between 3-6 PM (15-18 on the graph). The same pattern is reflected in the vendor's overlay plot, but there is less variation in travel time values and the high travel time values during 3-6 PM are lower in magnitude than was observed with WRTM (e.g., 4 minutes rather than 7 during the 5-6 PM focus hour). Because there is less variation in the vendor travel time records than WRTM for each hourly period, the delta IQRN and PTI values are noticeably negative – an expected result given that vendor data is reported in one-minute averages.

.

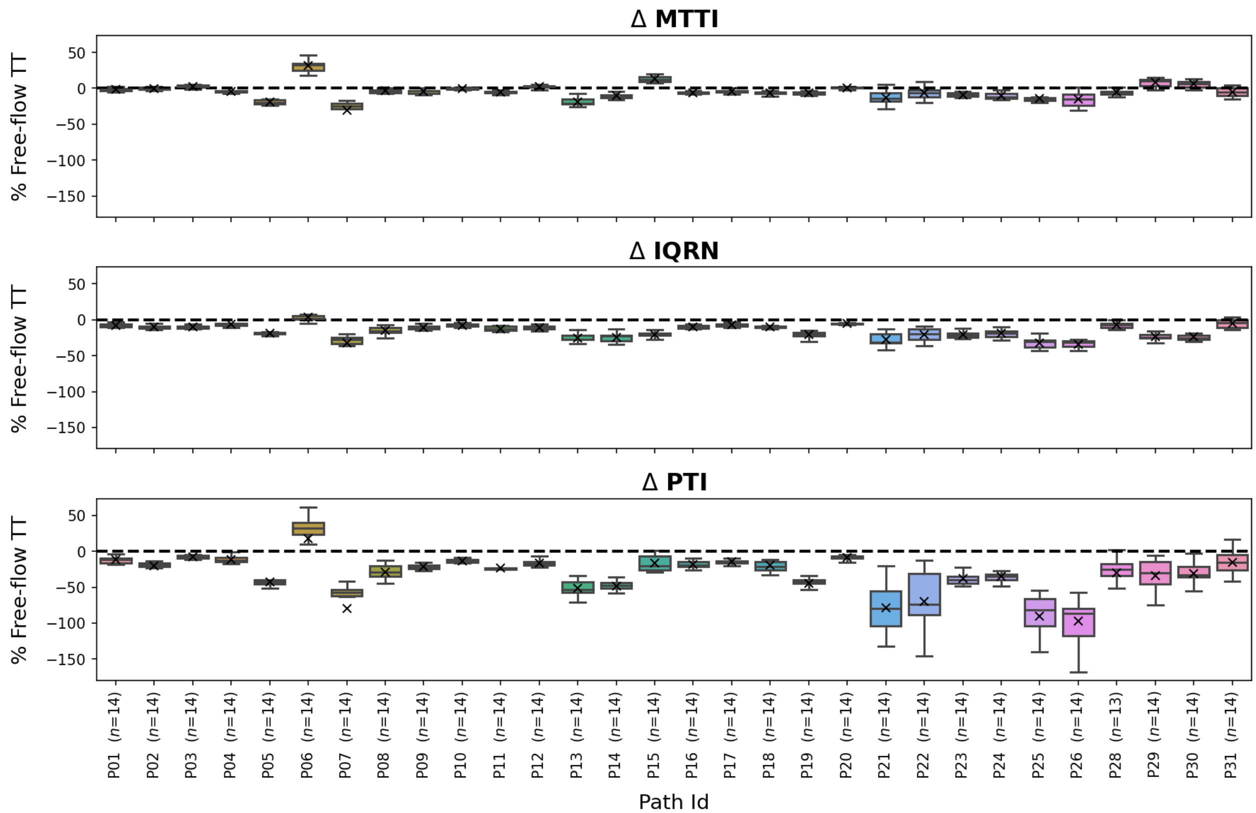
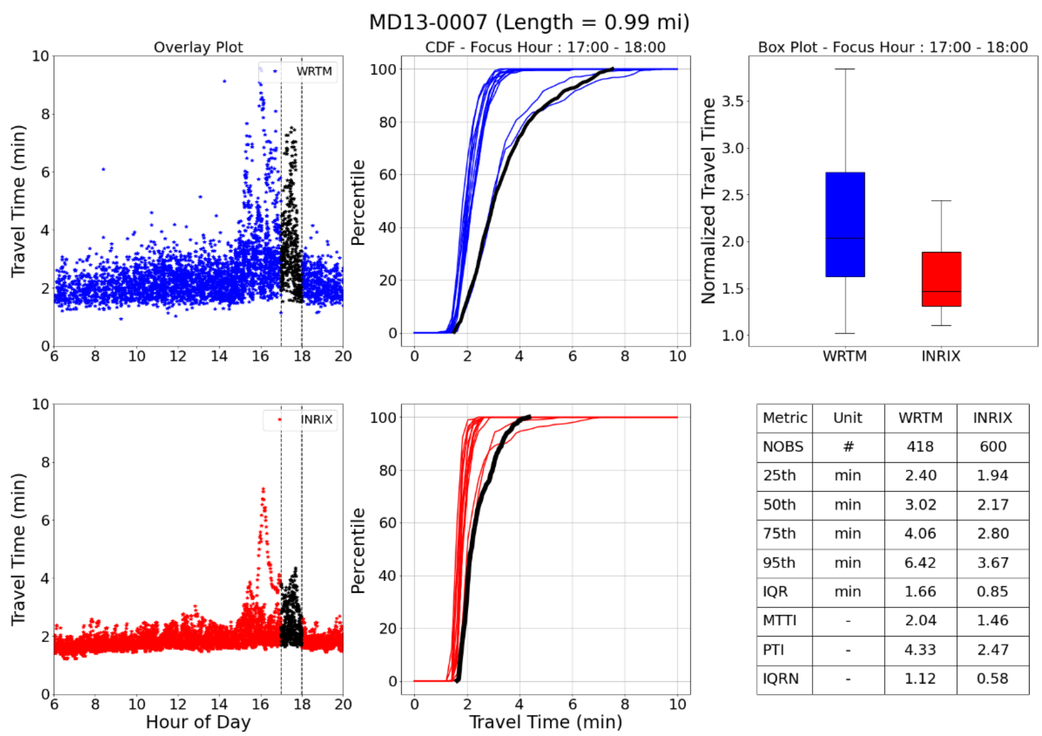


Figure 13- INRIX path-level delta PM values



Metric	Unit	WRTM	INRIX
NOBS	#	418	600
25th	min	2.40	1.94
50th	min	3.02	2.17
75th	min	4.06	2.80
95th	min	6.42	3.67
IQR	min	1.66	0.85
MTTI	-	2.04	1.46
PTI	-	4.33	2.47
IQRN	-	1.12	0.58

Figure 14 – INRIX overlay and CFD plots for segment MD13-0007.

TomTom

Figures 15-16 summarize the distribution of delta values under a variety of congestion and travel time reliability scenarios, while Table 8 reports the average values for the same scenarios. Observations from these figures and tables include the following points:

- **Central Tendency Comparison (Δ MTTI)**:
  - Δ MTTI values are centered around 0 for all congestion levels (MTTI bins), indicating that WRTM and vendor median travel time values match well during most periods. While all congestion levels have strong MTTI agreement on average, there are a wider range of delta values during periods of high congestion.
  - Δ MTTI values are close to 0 for all levels of travel time variation, indicating strong agreement in WRTM and Vendor median travel times during most periods.

- **Reliability Comparison (Δ IQRN, Δ PRI)**:
  - Reliability PM values do not follow an obvious pattern when quantified by level of congestion or variation, although in general the delta values are negative, indicating that variation is underreported. Δ IQRN remains nearly constant for different congestion levels, but becomes more negative as travel time variation increases. Δ PTI values become more negative with higher congestion and variation until reaching the highest bins, at which point it becomes more negative.

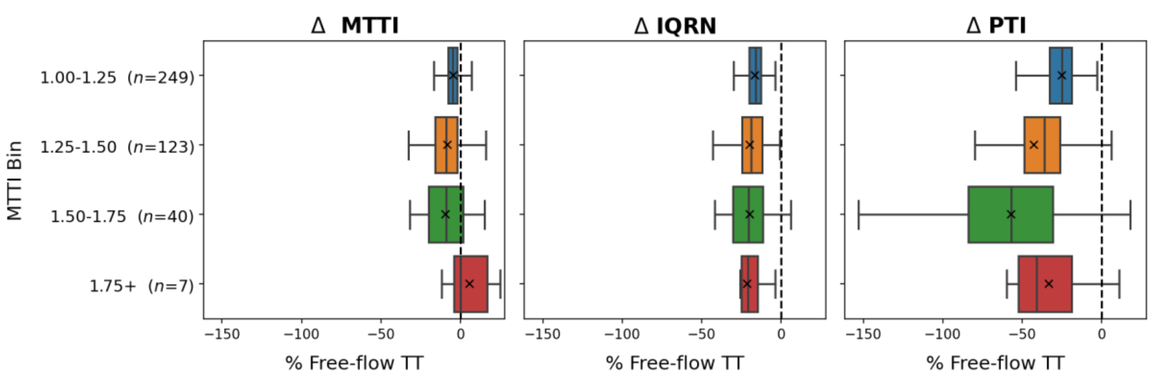


Figure 15 - TomTom summary of delta PM values grouped by congestion level

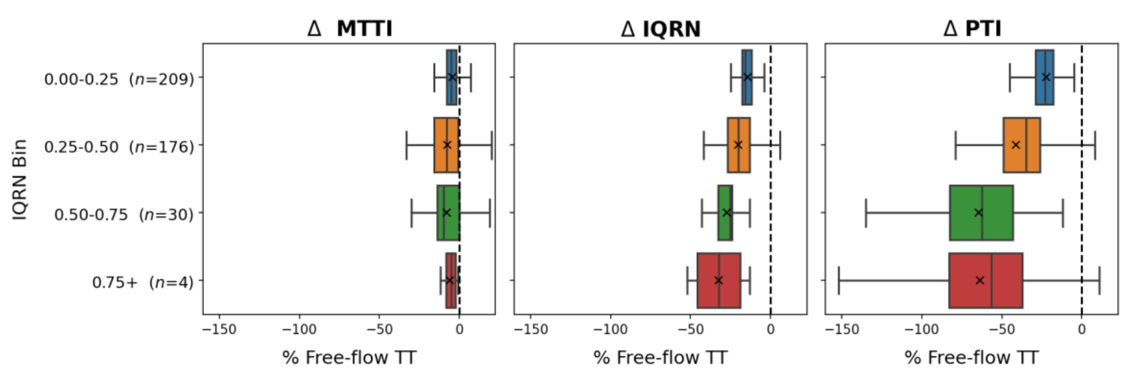


Figure 16 - TomTom summary of delta PM values grouped by level of travel time variation

Table 8 - TomTom summary of average delta PM values under various scenarios.

Performance Measure (PM)	Median Travel Time Index (MTTI)	IQR Normalized Travel Time (IQRN)	Planning Time Index (PTI)	
PM Definition	$\frac{TT_{50}}{TT_{FF}}$	$\frac{(TT_{75} - TT_{25})}{TT_{FF}}$	$\frac{TT_{95}}{TT_{FF}}$	
	$\overline{\Delta\ \text{MTTI}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{IQRN}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{PTI}}$ (% $TT_{\text{Free flow}}$)	**Samples**
Overall	-6	-18	-34	419
MTTI Bin (Higher congestion)				
1.00 – 1.25	-5	-17	-25	249
1.25 – 1.50	-8	-20	-43	123
1.50 – 1.75	-9	-20	-57	40
> 1.75	5	-21	-33	7
IQRN Bin (Higher variation)				
0.00 – 0.25	-4	-14	-22	209
0.25 – 0.50	-8	-20	-41	176
0.50 – 0.75	-8	-27	-65	30
> 0.75	-6	-33	-64	4

Delta (Δ): Difference in hourly performance measure values (Vendor -WRTM), expressed as percent of free-flow TT
Avg Delta ($\bar{\Delta}$): Average delta value computed using relevant 1-hr time-of-day samples, rounded to the nearest percent

Figure 17 shows the distribution of PM delta values across all validation paths. Most locations have small Δ MTTI values, indicating general agreement in median travel times across hourly periods. Δ IQRN values are quite consistent across locations as well, showing that the vendor data slightly underreports IQR values. Δ PTI values are much more noticeably negative - especially for paths P21-P31 – an intuitive result given that US-40 and US-11 have many traffic signals and access points. Because WRTM data is based on individual travel time observations while vendor data uses one-minute averages, it makes sense that WRTM data captures some higher travel times that cannot be captured in the average vendor data (and thus impacts the 95th percentile travel time values used in the PTI performance measure computation).

Figure 18 shows detailed results for path MD13-0007, whose travel times are shown in Figure 6 for the entire validation period. The overlay plot shows that WRTM travel times are typically concentrated around 2-3 minutes for most hours between 6 AM and 8PM, but can reach upwards of 7-10 minutes between 3-6 PM (15-18 on the graph). A similar pattern is reflected in the vendor's overlay plot, but there is less variation in travel time values and the high travel time values during 3-6 PM differ slightly in magnitude with respect to WRTM (e.g., the vendor has a higher peak during 4-5 PM, but slightly lower peak during 3-4 PM and 5-6 PM). Because there is less variation in the vendor travel time records than WRTM for each hourly period, the delta IQRN and PTI values are noticeably negative – an expected result given that vendor data is reported in one-minute averages.

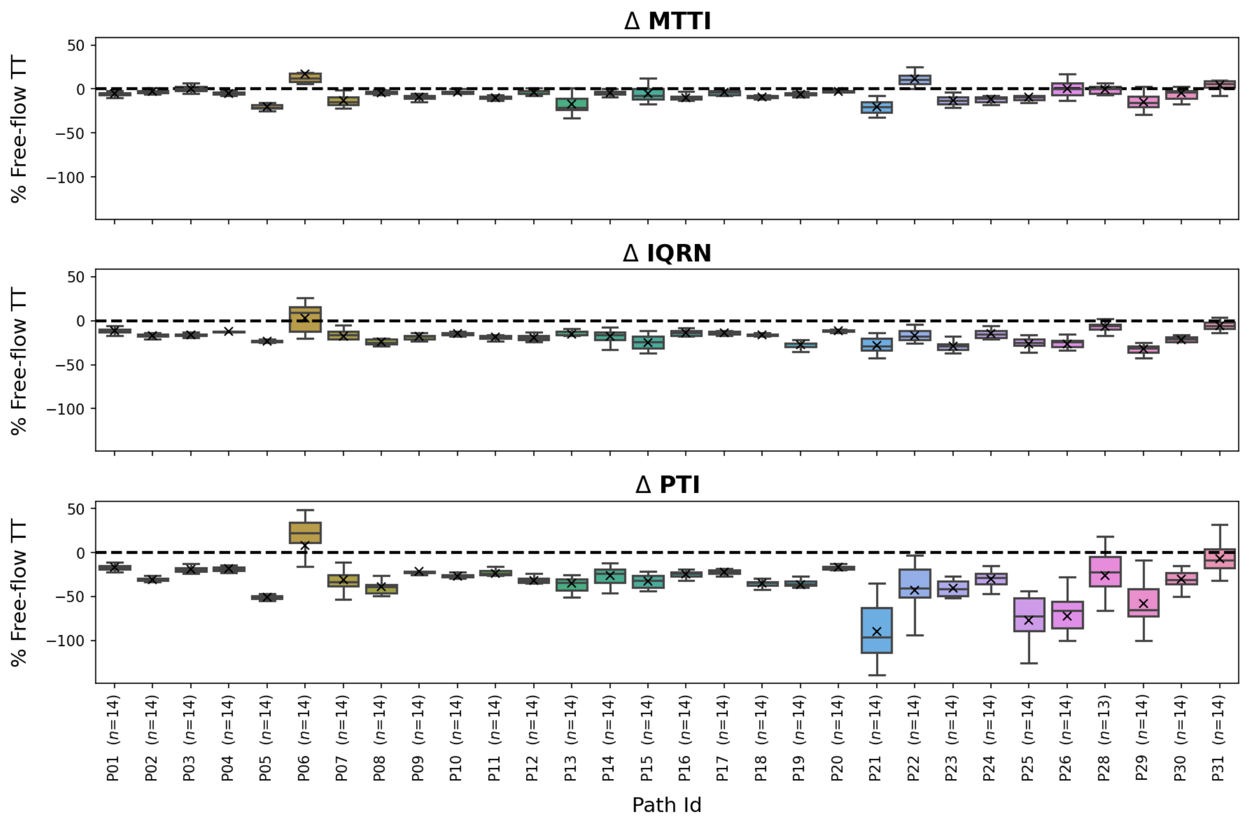
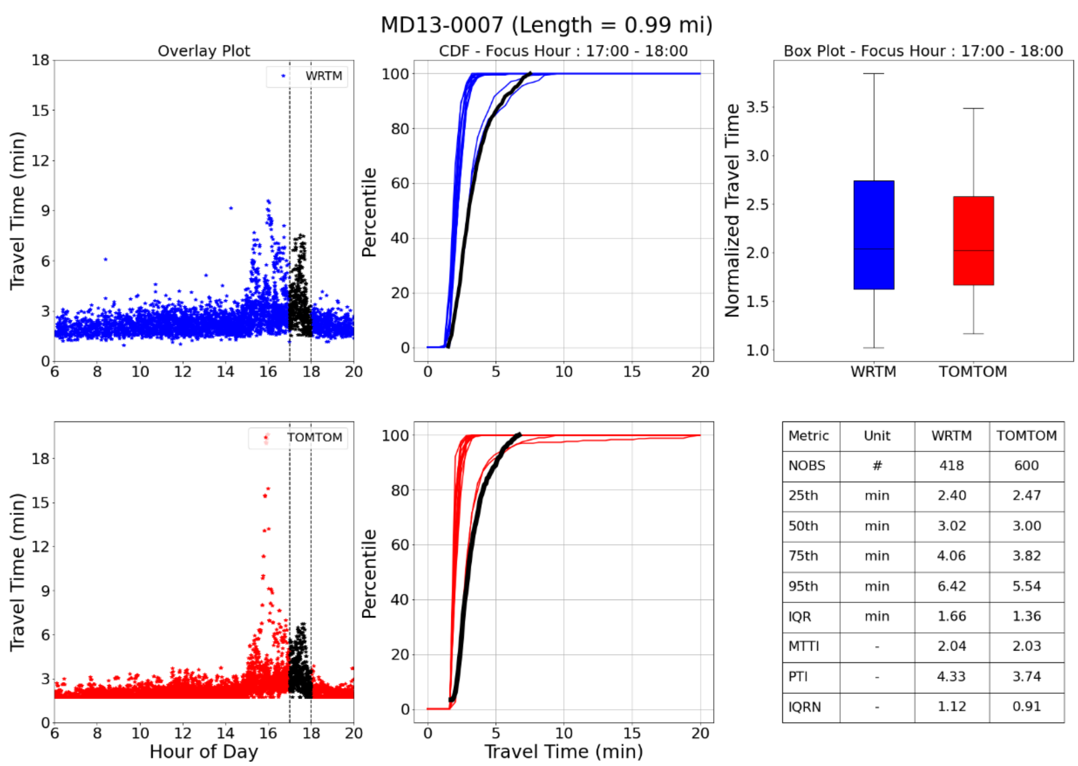


Figure 17 – TomTom path-level delta PM values



Metric	Unit	WRTM	TOMTOM
NOBS	#	418	600
25th	min	2.40	2.47
50th	min	3.02	3.00
75th	min	4.06	3.82
95th	min	6.42	5.54
IQR	min	1.66	1.36
MTTI	-	2.04	2.03
PTI	-	4.33	3.74
IQRN	-	1.12	0.91

Figure 18 – TomTom overlay and CFD plots for segment MD13-0007.

Summary and Discussion

Overall, all three vendors met the contract compliance requirements specified for the Vehicle Probe Project, achieving the specified targets for Average Absolute Speed Error and Speed Error Bias in all required speed bins. At a high level, this indicates **that all three vendors are providing speed and travel time data in line with expectations.**

However, given the low-volume nature of the study, the traditional analysis should not be considered in isolation – in particular because there are often not enough reference data samples to characterize ground truth conditions with confidence for point-in-time analysis. As a result of small sample sizes, analysis intervals sometimes have to be dropped, and even when they are not, the confidence band based on the Standard Error of the Mean is quite large and can obscure the resulting error measures.

Accordingly, this report re-introduces and extends a previously-proposed approach: the sampled distribution method, which is less sensitive to sample size and focuses on vendors' ability to capture repeatable travel time patterns. This method has been used in the past to characterize travel time patterns on signalized intersections that exhibit high variance, and provides a framework for computing performance measures based on time-of-day periods. This report extends the method to summarize differences (referred to as "deltas") between Vendor and WRTM performance measures under a variety of conditions. At the current time, the goal of this analysis is not to evaluate vendor compliance by setting benchmarks for "good" and "bad" delta values, but rather to draw high level insights from the results. Over time, this process may evolve into a more formal analysis with target error values, but at this point it simply seeks to convey how closely vendor data captures the central tendency and variation of repeatable travel time patterns.

Key findings include the following points, which are not specific to any one vendor:

- Median Vendor and WRTM travel time values tend to match well for most time-of day periods, indicating that **the data would likely be useful for planning applications that require accurate "typical" travel time (or speed) values**.

- Whenever there is significant variation in travel times during a time-of-day period, the vendors usually fail to capture the full extent of variation. However, this does not necessarily reflect poor accuracy; vendors report data in 1-minute averages, and thus naturally exhibit less variation than individual WRTM travel time observations.

While processing the data, the validation team noted that filtering WRTM travel time observations is a challenging task on roads that are not access controlled. Higher travel times can be attributed to (a) cycle failures at traffic signals and other legitimate congestion patterns, or (b) vehicles briefly leaving the road (e.g., for gas, coffee, etc.) – both of which can produce similar travel times, thus making it difficult to determine which observations are outliers. Care was taken to filter WRTM travel time operations by applying automated methods and using human judgement, but it is nonetheless a challenging task that impacts the performance values used in the analysis.

In the future, the validation team expects to continue using a variation of the Sampled Distribution Method to complement the point-in-time accuracy measures typically employed for validation – particularly for low-volume scenarios and along arterials with complex traffic flow patterns. This method will likely be refined over time based on feedback received from Coalition members.