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

Report for Vermont (#01)
Arterial Validation (Low-volume Roadways)

Data Collection: September-October 2020
Report Date: March 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 Vermont Agency of Transportation 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 from Wireless Re-identification Traffic Monitoring (WRTM) sensors, 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, commissioning two low-volume validation studies for 2020-2021: one in Maryland and a second in Vermont (this report).

This report uses reference data acquired on various corridors in both greater Burlington and central Vermont from September 21 – October 9, 2020: US-2 (Burlington), VT-289/VT-15/VT-117 (Essex), and US-302/VT-62 (Berlin), whose road characteristics are summarized in Table 1. **The AADT values range from about 8k-14k, reflecting the lowest volume roads validated to date**.

ES Table 1 - Arterial Corridor Description in Vermont

Corridor	# Segments	Miles (directional)	Speed Limit (MPH)	Signals / mile	Avg. AADT
US-2	4	12.5	40	0.8	13k
VT-289	3	8.5	50	0.2	10k
VT-15	4	5.6	35 – 40	2.1	14k
VT-117	4	16.1	25 – 40	0.4	8k
US-302	2	1.7	40	3.5	14k
VT-62	4	7.3	50	0.8	12k

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 but can have limitations in low-volume traffic conditions when there are fewer opportunities to obtain reference travel time observations. Small sample sizes can impact both the number of intervals suitable for comparison with vendor data, and the uncertainty associated with the ground truth mean speed estimate.

Accordingly, the traditional analysis is supplemented with an additional approach to characterize probe data quality: **the sampled distribution method.** While the traditional analysis 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). Note that **the lowest speed bin represents less than 1% of the 5-minute samples used for evaluation** – most of which come from the same segment, making it difficult to draw definitive conclusions about data quality in low-speed conditions.

ES Table 2 indicates that **HERE data was within contract specifications in all speed bins for AASE, and for all but the lowest speed bin for SEB**.

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	6.10	7.96	6.10	7.96	129
15-25 MPH	2.56	5.43	2.53	5.19	2595
25-35 MPH	1.28	3.66	0.56	1.19	6565
>35 MPH	1.90	5.13	-1.18	-2.8	9400
All Speeds	1.80	4.67	0.00	-0.21	18689

ES Table 3 indicates that **INRIX data was within contract specifications in all speed bins for AASE, and for all but the lowest speed bin for SEB.**

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	5.41	7.28	5.41	7.27	128
15-25 MPH	2.11	4.61	2.03	4.19	3066
25-35 MPH	1.12	3.35	0.49	1.24	7593
>35 MPH	1.43	4.41	-0.69	-1.89	9091
All Speeds	1.44	4.06	0.22	0.3	19878

ES Table 4 indicates that **TomTom data was within contract specifications in all speed bins for both AASE and SEB error measures.**

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	2.02	3.52	1.97	3.27	129
15-25 MPH	1.99	4.43	1.88	3.76	3178
25-35 MPH	1.14	3.49	0.90	2.37	7975
>35 MPH	2.03	5	0.50	0.78	9405
All Speeds	1.68	4.32	0.88	1.86	20687

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.

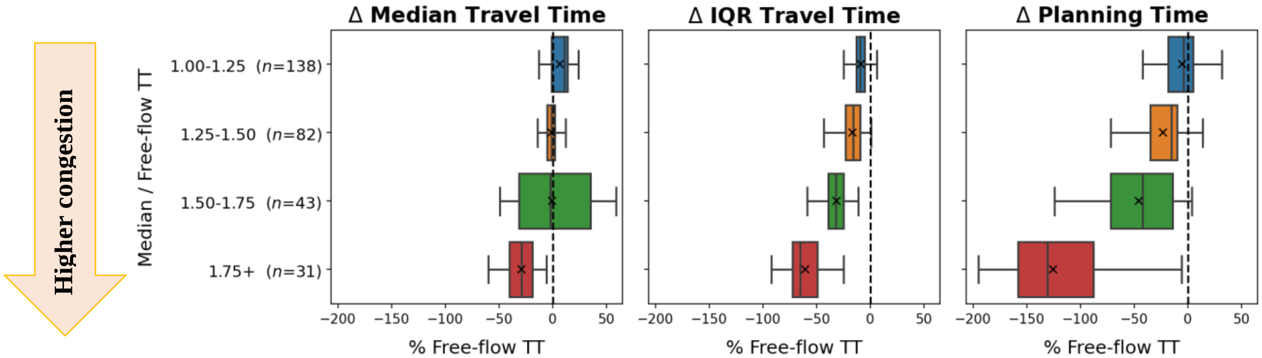
Three performance measures shown in ES Table 5 are used throughout the report to characterize the central tendency and variation of travel times during each time-of-day period and were chosen because they are intuitive and relevant to transportation planning applications. When communicating the difference 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., $\Delta\text{Median} = \text{Median}_{\text{Vendor}} - \text{Median}_{\text{WRTM}}$). Performance Measure values (as well as deltas between Vendor and WRTM PM values) are often expressed with respect to free-flow travel time (TT_{FF}) to normalize across validation segments.

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

Performance Measure (PM)	Median Travel Time	IQR Travel Time	Planning Time
Definition	TT_{50}	$(TT_{75} - TT_{25})$	TT_{95}
Type	Central Tendency	Reliability / Variation	Reliability / Variation

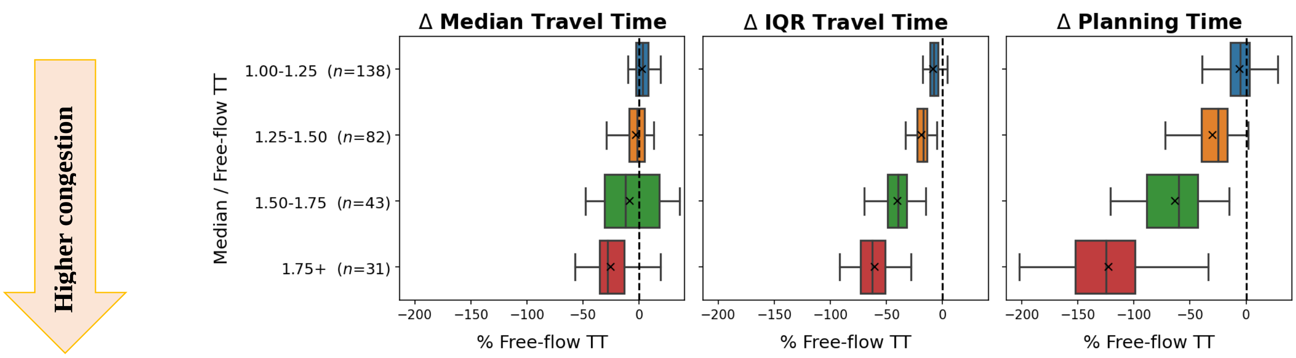
ES Figures 1-3 summarize the distribution of "delta" PM values when traffic is characterized by different levels of congestion. Note that the analysis is intended to provide high level insight into vendors' ability to capture recurrent travel time patterns, *not measure compliance.*

ES Figure 1 shows that **HERE's median travel times agree with the reference data for most periods** (small Δ Median), with the highest congestion bin showing a slight negative bias. Travel time reliability is most accurate (small Δ IQR, Δ Planning Time) at lower levels of congestion and tends to underestimate true variation at higher levels of congestion.



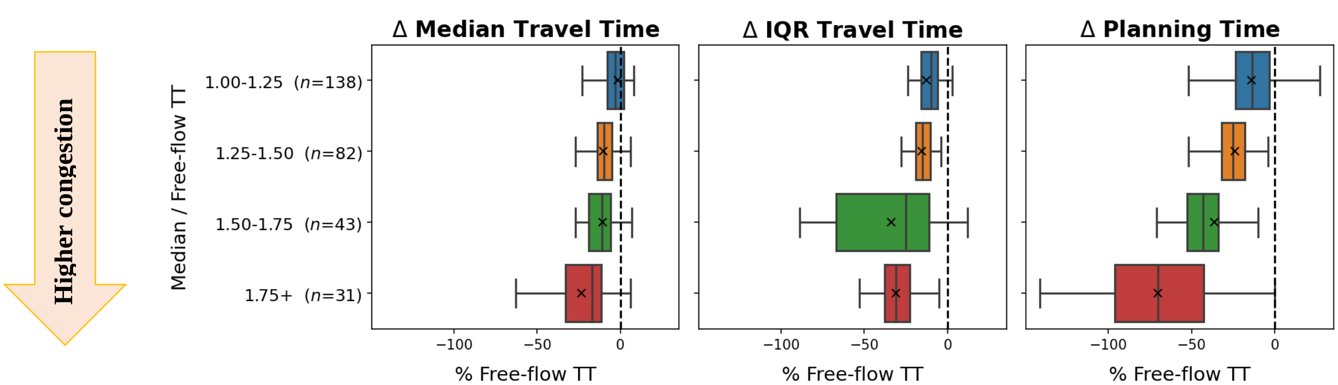
ES Figure 1 - HERE summary of delta PM values grouped by congestion level

ES Figure 2 shows that **INRIX's median travel times agree with the reference data for most periods** (small Δ Median)**,** with slight negative bias at higher congestion levels. Travel time reliability is most accurate (small Δ IQR, Δ Planning Time) 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 Median)**,** with data becoming a bit more negative biased at higher levels of congestion. Travel time reliability is most accurate (small Δ IQR, Δ Planning Time) at low levels of congestion and tends to underestimate true variation at higher levels of congestion.



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

Findings and Recommendations

The top-level findings are summarized below for both operations and planning perspectives (based on the traditional and sampled distribution analyses, respectively):

Operations Perspective

- **Probe data was generally effective in representing ground truth traffic speed / travel times** during the periods where there was sufficient reference data for evaluation.
    - All vendors were compliant with VPP requirements for the Average Absolute Speed Error (AASE) measure specification and minimally achieved compliance in all but the lowest speed bin for the Speed Error Bias (SEB) specification.
    - Two vendors' SEB measures were slightly out of the target range in the lowest speed bin (0-15 MPH), but this scenario represented less than 1% of the observations used for evaluation and mostly came from a single segment.

- It was difficult to draw definitive conclusions about low-speed data quality, but **it is recommended that probe data be monitored in low-speed conditions.**
    - There were limited opportunities to evaluate data quality in the lowest speed bin, and in these limited samples two vendors were slightly out of target range for SEB.
    - The lowest speed bin has historically been most challenging for compliance even on higher-volume roads, so it is worth continuing to monitor in low-volume conditions that have been less studied.

Planning / Arterial Management Perspective

- Median vendor and reference travel time values matched 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.**
    - This agreement is particularly strong when there is less congestion and variation in travel times in each time-of-day period.

- Whenever there was significant variation in travel times during a time-of-day period, the vendors usually under-reported travel time variation compared to the reference data (i.e., overestimated travel time reliability). However, this may be at least partially explained by the inherent aggregation of the vendor data product (the fact that probe data is reported in one-minute averages).

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 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 limited WRTM travel time data to characterize ground truth traffic conditions, which can require discarding temporal periods from the analysis. Even if there is enough data, the low sample size can impact the confidence band used to characterize the ground truth mean speed. Accordingly, this report conducts both the traditional point-in-time analysis and an additional 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 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 Cumulative 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 non-outlier 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.

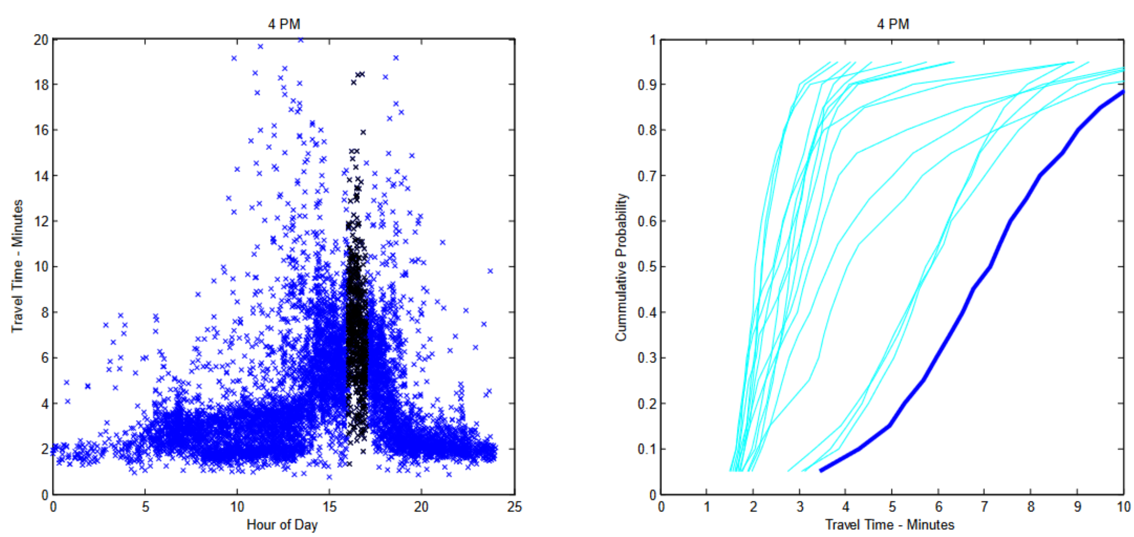
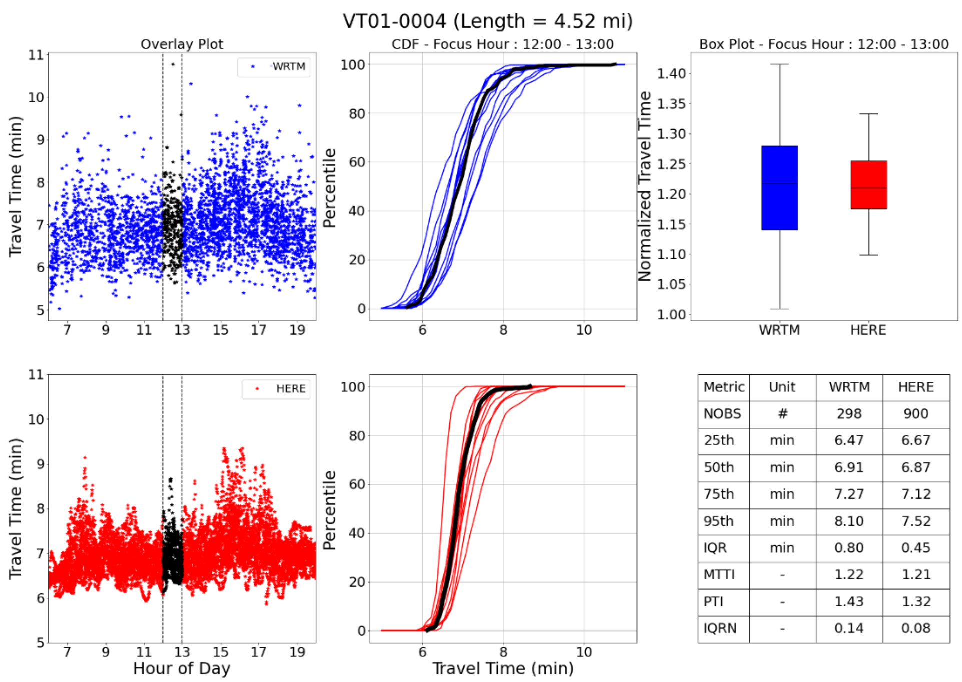


Figure 1 - Example 24-hr overlay plot and cumulative frequency diagram.

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	#	298	900
25th	min	6.47	6.67
50th	min	6.91	6.87
75th	min	7.27	7.12
95th	min	8.10	7.52
IQR	min	0.80	0.45
MTTI	-	1.22	1.21
PTI	-	1.43	1.32
IQRN	-	0.14	0.08

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

Three performance measures are used throughout the analysis and are defined as follows, with the notation TT_X referring to the Xth percentile travel time.

- **Median Travel Time**: TT_{50}
  - Measures typical travel time during period (central tendency)
- **Inter-Quartile Range (IQR):** $TT_{75} - TT_{25}$
  - Measures how travel time varies during period (reliability)
- **Planning Time (PT):** TT_{95}
  - Measures near worst-case travel time during period (reliability)

Often these travel time values are normalized by the free-flow travel time for the segment, TT_{FF}, which is the free-flow travel time obtained by taking the 15th percentile travel time during overnight hours (10pm – 5am). 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.

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. 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., $\Delta\,\text{Median} = \text{Median}_{\text{Vendor}} - \text{Median}_{\text{WRTM}}$). Because delta values will be summarized across road segments with different speed limits and lengths, they are expressed as unitless values normalized by the free-flow travel time. For example, Median Travel Time would be expressed as $\frac{TT_{50}}{TT_{FF}}$, which represents the 50th percentile travel time as a ratio (or equivalently as a percentage of) the free-flow travel time TT_{FF}. Common aggregation strategies include:

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

- **By level of congestion ($\mathbf{TT_{50}/TT_{FF}}$)**: Each hourly period is assigned a congestion bin based on the WRTM's ratio of median to free-flow travel time 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 level of travel time variation ($\mathbf{TT_{IQR}/TT_{FF}}$)**: Each hourly period is assigned a variation 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 delta values (i.e., Δ Median, Δ IQR, Δ PT) 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 congestion bin (Median / Free-flow travel time), 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 congestion level

Performance Measure (PM)	Median Travel Time	IQR Travel Time	Planning Time	
PM Definition	TT_{50}	$(TT_{75} - TT_{25})$	TT_{95}	
	$\overline{\Delta\ \text{Median}}$ ($\%\ TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{IQR}}$ ($\%\ TT_{\text{Free flow}}$)	$\overline{\Delta\ \text{PT}}$ ($\%\ TT_{\text{Free flow}}$)	Samples
Median / Free-flow TT				
1.00 – 1.25	7	-9	-5	138
1.25 – 1.50	-2	-16	-24	80
1.50 – 1.75	0	-31	-44	43
> 1.75	-31	-59	-124	33

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

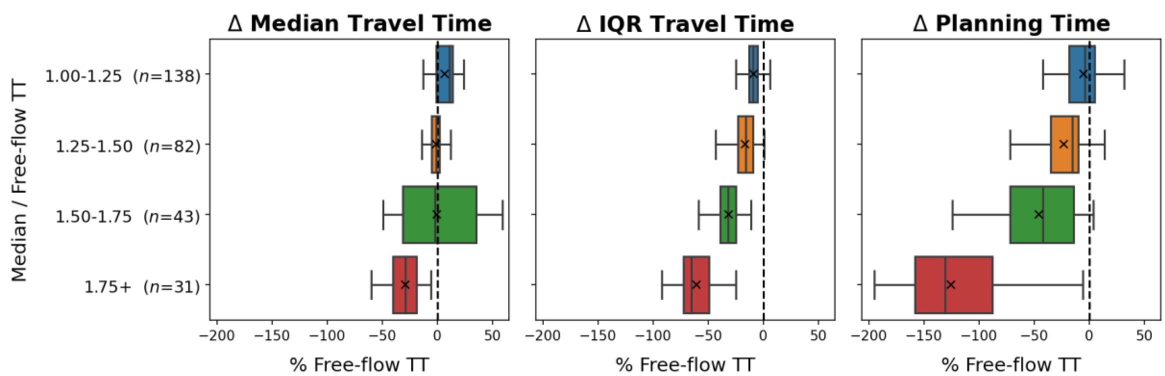


Figure 3: Sample distribution of delta values when aggregated by MTTI Bin

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 Median 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 Median 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 Planning Time 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 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 is often less variation than in WRTM observations due to the averaging inherent in the product. Thus, discrepancies in Δ IQR and Δ Planning Time, 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.

- Delta values are expressed with respect to free-flow travel time to normalize across different road segments.
    - E.g., Δ PT = -0.7 (or -70 % TT_{FF}) means that the Vendor's Planning Time value is less than the WRTM 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 directional validation road segments in Vermont from September 21 through October 9, 2020, which are shown below in Figure 4. The data collection was conducted by Vermont Agency of Transportation (VTrans), who used a combination of permanent and temporary Bluetooth sensors to acquire the re-identification data.

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.

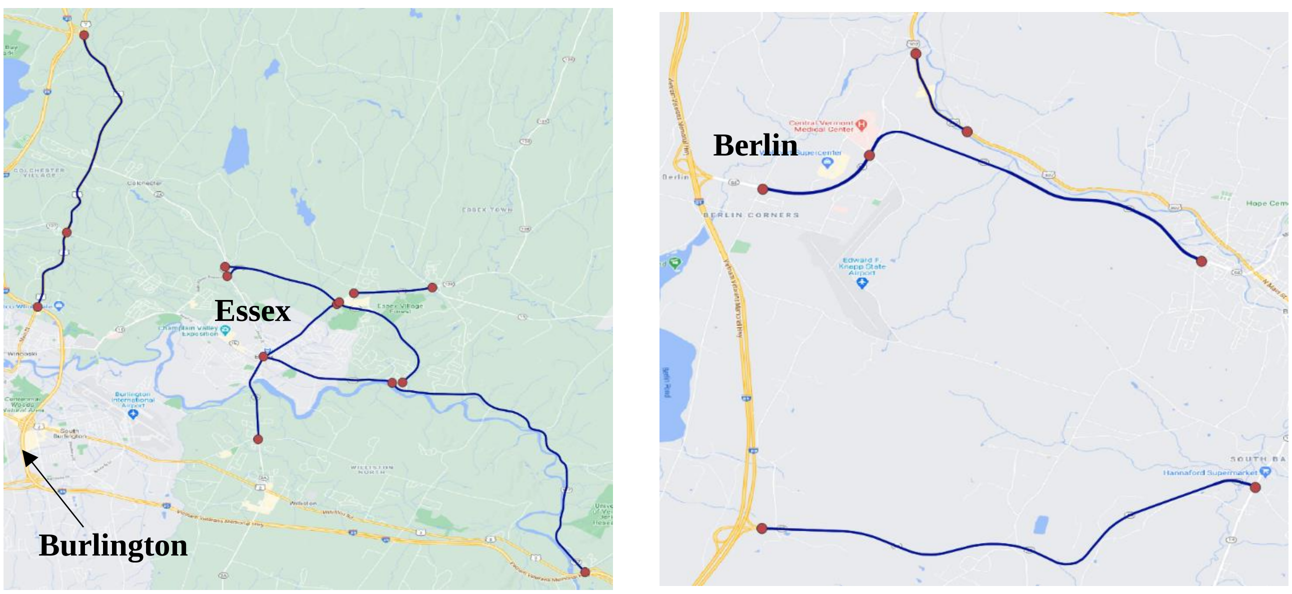


Figure 4 – WRTM Sensor locations

Table 2 - Validation Segment Attributes

Segment (Map Link)	DESCRIPTION						Deployment		
	Highway	Starting at	Lane (Min)	Signals	AADT	Access Points	Begin Lat/Lon		Length (mile)
	Direction	Ending at	Lane (Max)	Signal/mile		Speed Limit	End Lat/Lon		
Arterial									
A1	US-2	US 2 / US 7	1	2	11361	17	44.58952	-73.16645	4.52
VT01-0001	Southbound	US 2 / VT 127	1	0.44		40	44.52887	-73.17195	
A2	US-2	US 2 / VT 127	1	4	15958	9	44.52887	-73.17195	1.74
VT01-0002	Southbound	US 2 / I-89 N	2	2.29		40	44.50599	-73.18101	
A3	US-2	US 2 / I-89 N	1	4	15958	9	44.50599	-73.18101	1.74
VT01-0003	Northbound	US 2 / VT 127	2	2.29		40	44.52887	-73.17195	
A4	US-2	US 2 / VT 127	1	2	11361	17	44.52887	-73.17195	4.52
VT01-0004	Northbound	US 2 / US 7	1	0.44		40	44.58952	-73.16645	
A5	VT-289	VT 2A / VT 289	1	1	16410	1	44.51543	-73.12232	1.98
VT01-0005	Southbound	VT 15 / VT 289	2	0.50		50	44.50702	-73.08834	
A6	VT-289	VT 15 / VT 289	1	0	5170	4	44.50702	-73.08834	2.29
VT01-0006	Southbound	VT 289 / VT 117	2	0.0		50	44.48385	-73.06720	
A7	VT-289	VT 289 / VT 117	1	1	10334	6	44.48385	-73.06720	4.23
VT01-0007	Northbound	VT 2A / VT 289	2	0.24		50	44.51835	-73.12319	
A8	VT-15	VT 15 / VT 2A / VT 117	1	3	12081	13	44.49053	-73.11119	1.59
VT01-0009	Northbound	VT 15 / Commonwealth Ave	1	1.89		40	44.51016	-73.08351	
A9	VT-15	VT 15 / Commonwealth Ave	1	3	17025	14	44.51016	-73.08351	1.20
VT01-0010	Northbound	VT 15 / VT 128	1	2.50		35	44.51186	-73.05935	
A10	VT-15	VT 15 / VT 128	1	3	17025	14	44.51186	-73.05935	1.20
VT01-0011	Southbound	VT 15 / Commonwealth Ave	1	2.50		35	44.51016	-73.08351	
A11	VT-15	VT 15 / Commonwealth Ave	1	3	12081	14	44.51016	-73.08351	1.59
VT01-0012	Southbound	VT 15 / VT 2A / VT 117	1	1.89		40	44.49053	-73.11119	
A12	VT-117	VT 15 / VT 2A / VT 117	1	2	7616	20	44.49053	-73.11119	2.09
VT01-0014	Eastbound	VT 117 / VT 289	1	0.96		25	44.48243	-73.07177	

Segment (Map Link)	DESCRIPTION Highway Direction	Starting at Ending at	Lane (Min) Lane (Max)	Signals Signal/mile	AADT	Access Points Speed Limit	Deployment Begin Lat/Lon End Lat/Lon		Length (mile)
Arterial									
A13 VT01-0015	VT-117 Eastbound	VT 117 / VT 289 VT 117 / US 2	1 1	2 0.34	7768	12 40	44.48243 44.42437	-73.07177 -73.01217	5.97
A14 VT01-0016	VT-117 Westbound	VT 117 / US 2 VT 117 / VT 289	1 1	2 0.34	7768	12 40	44.42437 44.48243	-73.01217 -73.07178	5.97
A15 VT01-0017	VT-117 Westbound	VT 117 / VT 289 VT 15 / VT 2A / VT 117	1 1	2 0.96	7616	20 25	44.48243 44.49053	-73.07178 -73.11119	2.09
A16 VT01-0018	US-302 Southbound	US 302 / Berlin State Highway US 302 / Ames Dr	1 1	3 3.56	13740	14 40	44.23198 44.22146	-72.55363 -72.54658	0.84
A17 VT01-0019	US-302 Northbound	US 302 / Ames Dr US 302 / Berlin State Highway	1 1	3 3.56	13740	14 40	44.22146 44.23198	-72.54658 -72.55363	0.84
A18 VT01-0020	VT-62 Eastbound	VT 62 / Paine Turnpike N VT 62 / Airport/Fisher Rd	2 2	2 2.23	11120	2 50	44.21375 44.21825	-72.57438 -72.55980	0.92
A19 VT01-0021	VT-62 Eastbound	VT 62 / Airport/Fisher Rd VT 62 / Berlin St	1 2	2 0.73	11900	2 50	44.21825 44.20397	-72.55980 -72.51474	2.74
A20 VT01-0022	VT-62 Westbound	VT 62 / Berlin St VT 62 / Airport/Fisher Rd	1 2	2 0.73	11900	2 50	44.20408 44.21836	-72.51469 -72.55990	2.74
A21 VT01-0023	VT-62 Westbound	VT 62 / Airport/Fisher Rd VT 62 / Paine Turnpike N	2 2	2 2.25	11120	2 50	44.21836 44.21388	-72.55990 -72.57434	0.91

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). Note that **the lowest speed bin (0-15 mph) contains less than 1% of all 5-minute samples used for evaluation – most of which come from the same segment, making it difficult to draw conclusions about data quality in low-speed conditions.**

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 in all speed bins when compared to both the SEM band and Mean. The Speed Error Bias (SEB) is within specification in all but the lowest speed bin (0-15 mph) when compared to the SEM Band, and in the upper two bins (25-35 mph, 35+ mph) 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	6.1	7.96	6.1	7.96	129
15-25 MPH	2.56	5.43	2.53	5.19	2595
25-35 MPH	1.28	3.66	0.56	1.19	6565
>35 MPH	1.9	5.13	-1.18	-2.8	9400
All Speeds	1.8	4.67	0	-0.21	18689

INRIX

Table 4 reports the traditional analysis error metrics for INRIX. The Average Absolute Speed Error (AASE) is within specification in all speed bins when compared to both the SEM band and Mean. The Speed Error Bias (SEB) is within specification in all but the lowest speed bin (0-15 mph) when compared to the SEM Band and 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	5.41	7.28	5.41	7.27	128
15-25 MPH	2.11	4.61	2.03	4.19	3066
25-35 MPH	1.12	3.35	0.49	1.24	7593
>35 MPH	1.43	4.41	-0.69	-1.89	9091
All Speeds	1.44	4.06	0.22	0.3	19878

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	2.02	3.52	1.97	3.27	129
15-25 MPH	1.99	4.43	1.88	3.76	3178
25-35 MPH	1.14	3.49	0.9	2.37	7975
>35 MPH	2.03	5	0.5	0.78	9405
All Speeds	1.68	4.32	0.88	1.86	20687

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 Median, Δ IQR, Δ PT) 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 Median, Δ IQR, Δ PT boxplots for each path. 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 path: VT01-0010, which reinforces the underlying methodology for computing the performance measures. VT01-0010, shown below in Figure 5, is located along VT-10 in the Eastbound direction and carries traffic NE of downtown Essex. It is about 1.2 miles in length with an AADT of approximately 17k (bidirectional, so approximately 8-9k in the Eastbound direction), and has 3 traffic signals and 9 access points, thus providing several 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 and indicates that travel times range from 2 - 4.5 minutes over the course of most days, occasionally reaching 5-7 minutes during periods of higher congestion.

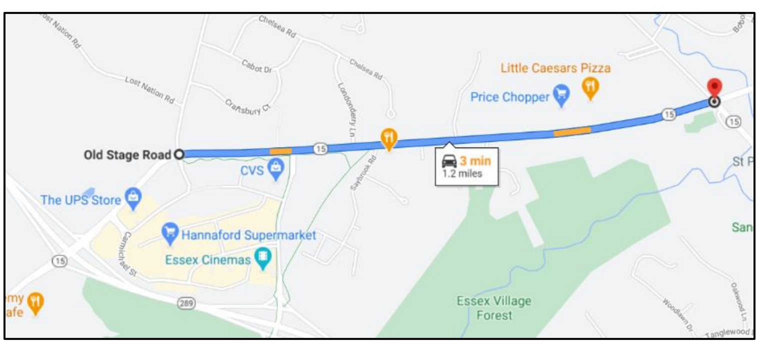


Figure 5 – Segment boundaries and screenshot for validation path VT01-0010.

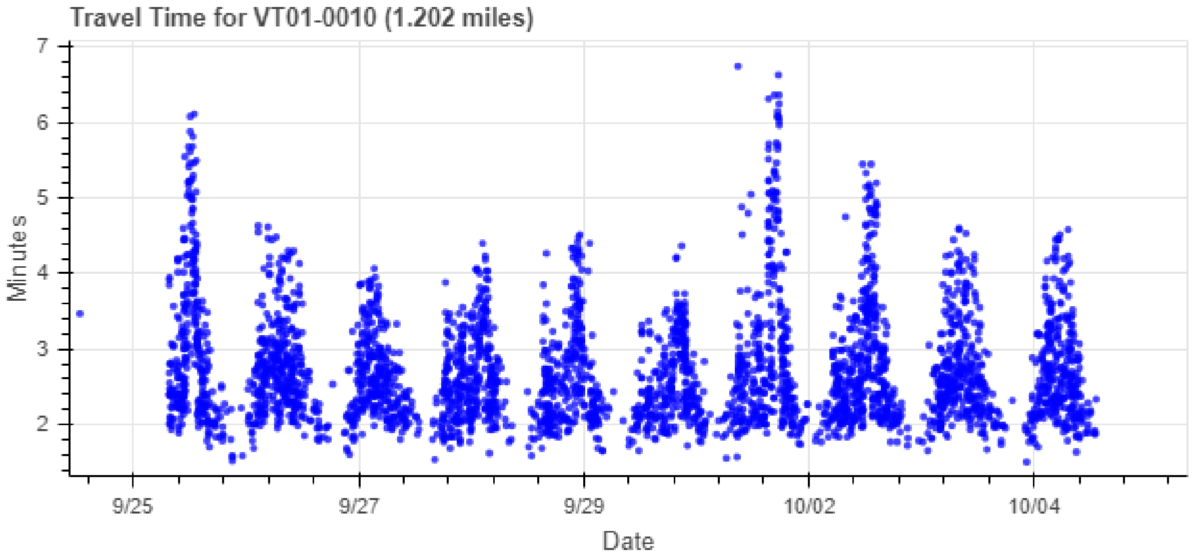


Figure 6 – Travel times along path VT01-0010 over data collection period.

Note that higher 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 automated filtering algorithms and human judgement.

When interpreting the results, it is important 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 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 for Planning Time, 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 Median Travel Time)**:
  - Δ Median values are centered close to 0 for the lowest three congestion levels in Figure 7, indicating that WRTM and vendor median travel time values match well during most periods. The vendor tends to underreport the median travel time in the highest congestion bin.
  - Median travel times show strong agreement when there is minimal travel time variation. In periods with more variation, there appears to be a slight negative bias, as well as a wider range of discrepancies in the highest bin.

- **Reliability Comparison (Δ IQR Travel Time, Δ Planning Time)**:
  - Reliability PM values (IQR, PT) are close when congestion and variation levels are low. As congestion and travel time variation increase, the Vendor increasingly underreports true travel time variation (more negative delta values).

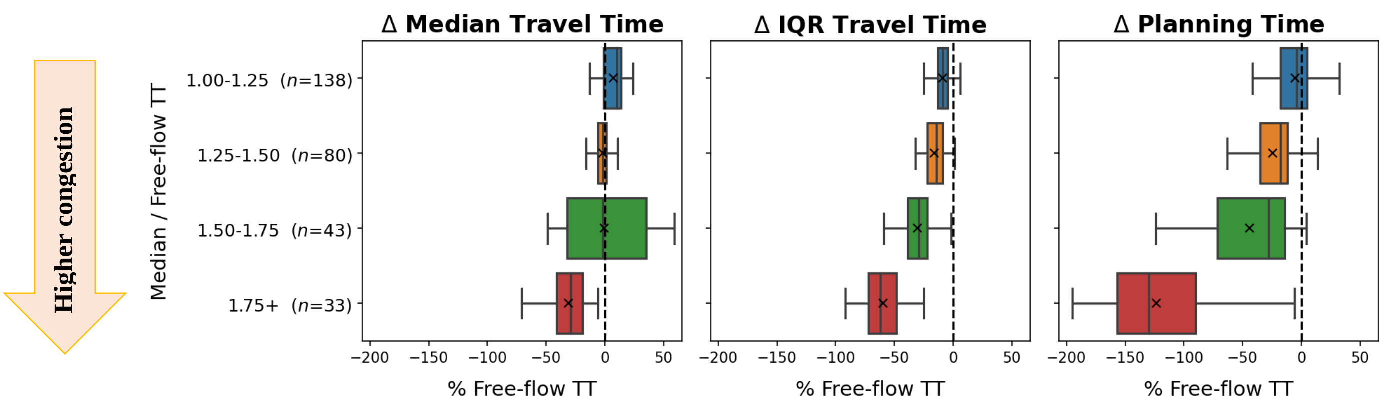


Figure 7 - HERE summary of delta PM values grouped by congestion level

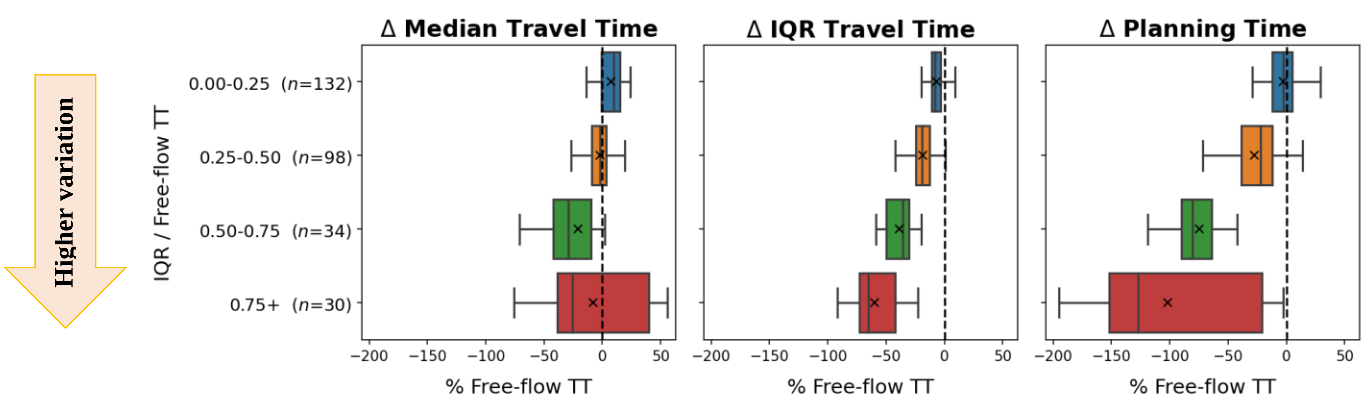


Figure 8 - HERE summary of delta PM values grouped by level of travel time variation

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

Performance Measure (PM)	Median Travel Time	IQR Travel Time	Planning Time	
PM Definition	TT_{50}	$(TT_{75} - TT_{25})$	TT_{95}	
	$\overline{\Delta \text{ Median}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta \text{ IQR}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta \text{ PT}}$ (% $TT_{\text{Free flow}}$)	**Samples**
Overall	-1	-20	-29	294
Higher congestion				
Median / Free-flow TT				
1.00 – 1.25	7	-9	-5	138
1.25 – 1.50	-2	-16	-24	80
1.50 – 1.75	0	-31	-44	43
> 1.75	-31	-59	-124	33
Higher variation				
IQR / Free-flow TT				
0.00 – 0.25	7	-7	-3	132
0.25 – 0.50	-2	-18	-27	98
0.50 – 0.75	-21	-39	-75	34
> 0.75	-8	-60	-102	30

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 9 shows the distribution of PM delta values across all validation paths. Many locations have small Δ Median values, indicating general agreement in median travel times across hourly periods. Δ IQR values are quite consistent across locations as well, showing that the vendor data slightly underreports IQR values. Δ PT values are much more noticeably negative - especially for a handful of paths (e.g., P10, P12, P17-P20) – an intuitive result given that these segments have multiple 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 reflected in the average vendor data (and thus impacts the 95th percentile travel time values used in the Planning Time performance measure computation).

Figure 10 shows detailed results for path VT01-0010, whose travel times were previously shown in Figure 6 for the entire validation period, and which is one of the few locations to exhibit repeatable time-of-day patterns. The overlay plot shows that WRTM travel times are typically concentrated around 2 – 3.5 minutes for most hours between 6AM and 8PM but can reach upwards of 5-7 minutes between 3-6 PM (hours 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 are lower in magnitude than was observed with WRTM. Because there is less variation in the vendor travel time records than WRTM for each hourly period, the delta Planning Time values are noticeably negative – an expected result given that vendor data is reported in one-minute averages.

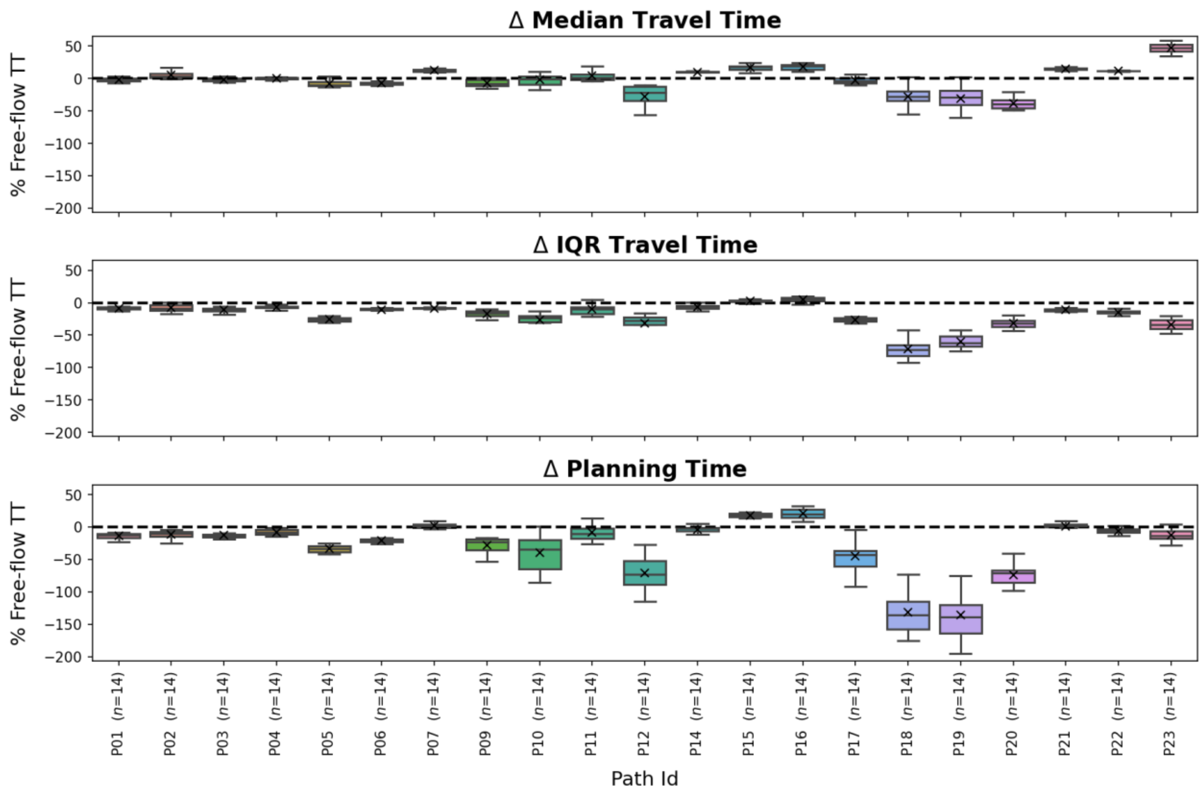
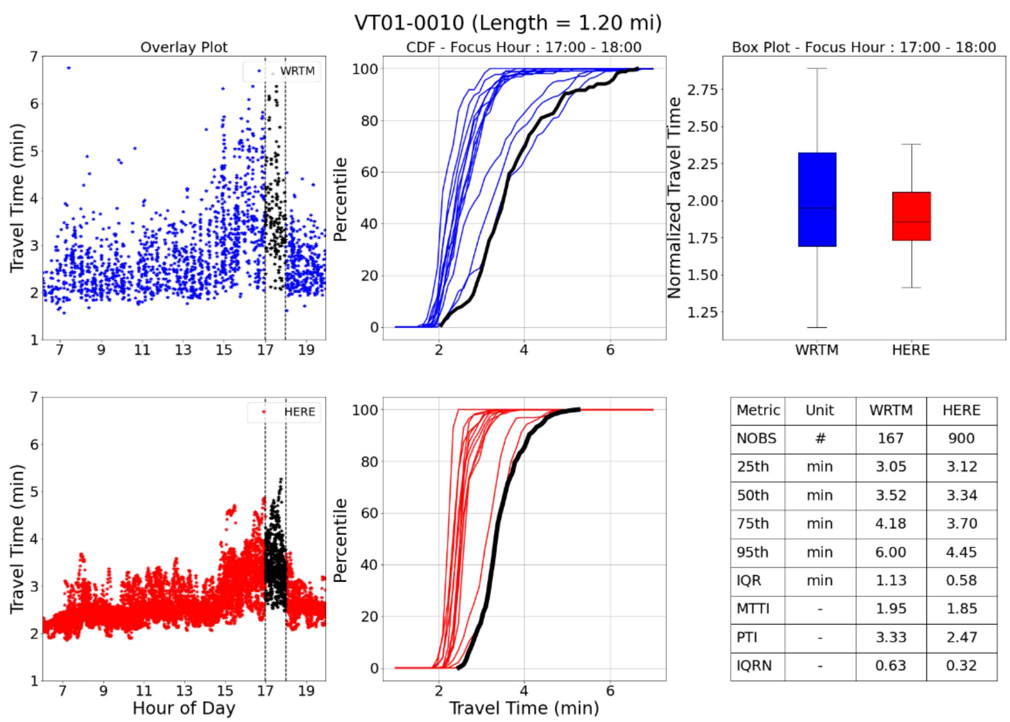


Figure 9 - HERE path-level delta PM values



Metric	Unit	WRTM	HERE
NOBS	#	167	900
25th	min	3.05	3.12
50th	min	3.52	3.34
75th	min	4.18	3.70
95th	min	6.00	4.45
IQR	min	1.13	0.58
MTTI	-	1.95	1.85
PTI	-	3.33	2.47
IQRN	-	0.63	0.32

Figure 10 – HERE overlay and CFD plots for segment VT01-0010.

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 Median Travel Time)**:
    - Δ Median values are centered close to 0 for the lowest three congestion levels in Figure 11 (particularly the lowest two), indicating that WRTM and vendor median travel time values match well during many periods. The vendor tends to underreport the median travel time in the highest congestion bin.
    - Median travel times show strong agreement when there is minimal travel time variation. In periods with more variation, there appears to be a slight negative bias, as well as a wider range of discrepancies in the highest bin.

- **Reliability Comparison (Δ IQR Travel Time, Δ Planning Time)**:
    - Reliability PM values (IQR, PT) are close when congestion and variation levels are low. As congestion and travel time variation increase, the Vendor increasingly underreports true travel time variation (more negative delta values).

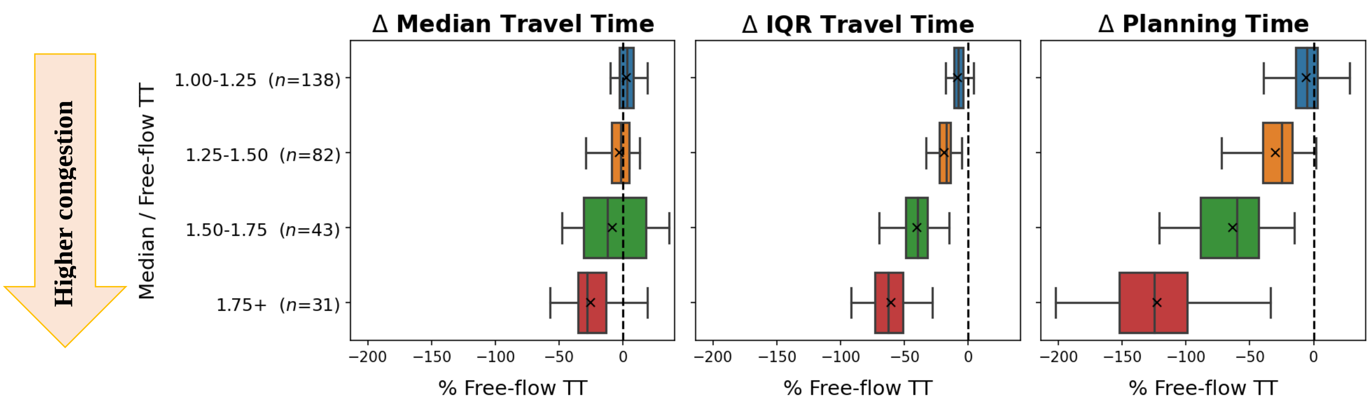


Figure 11 - INRIX summary of delta PM values grouped by congestion level

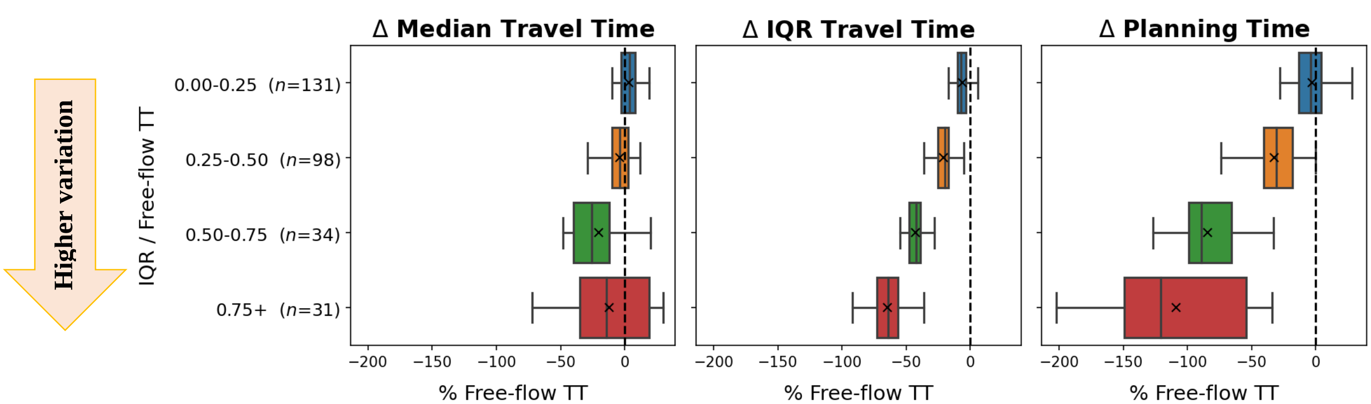


Figure 12 - INRIX summary of delta PM values grouped by level of travel time variation

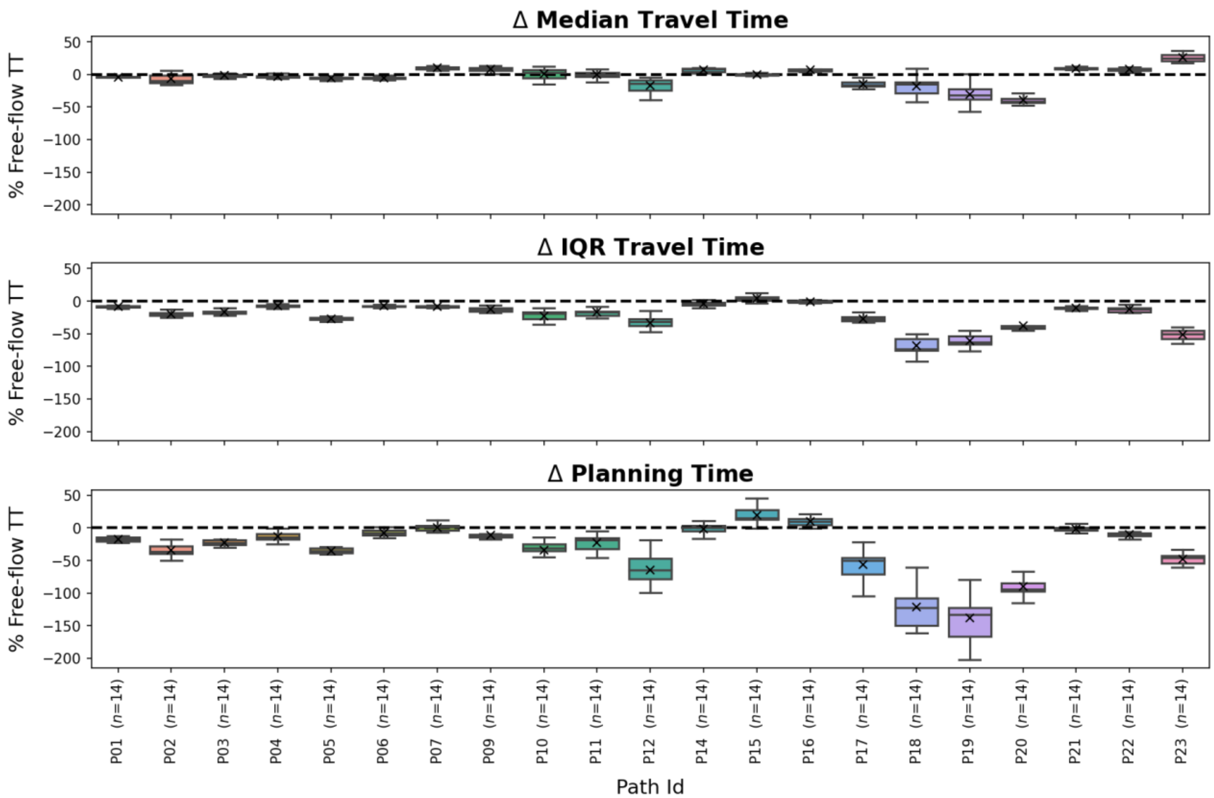
Table 7 - INRIX summary of average delta PM values under various scenarios.

Performance Measure (PM)	**Median Travel Time**	**IQR Travel Time**	**Planning Time**	
PM Definition	TT_{50}	$(TT_{75} - TT_{25})$	TT_{95}	
	$\overline{\Delta\ \text{Median}}$ $(\%\ TT_{\text{Free flow}})$	$\overline{\Delta\ \text{IQR}}$ $(\%\ TT_{\text{Free flow}})$	$\overline{\Delta\ \text{PT}}$ $(\%\ TT_{\text{Free flow}})$	**Samples**
Overall	-4	-22	-33	294
Higher congestion				
Median / Free-flow TT				
1.00 – 1.25	3	-9	-6	138
1.25 – 1.50	-3	-19	-30	82
1.50 – 1.75	-8	-40	-64	43
> 1.75	-25	-60	-122	31
Higher variation				
IQR / Free-flow TT				
0.00 – 0.25	3	-6	-3	131
0.25 – 0.50	-4	-21	-32	98
0.50 – 0.75	-20	-43	-85	34
> 0.75	-12	-65	-109	31

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

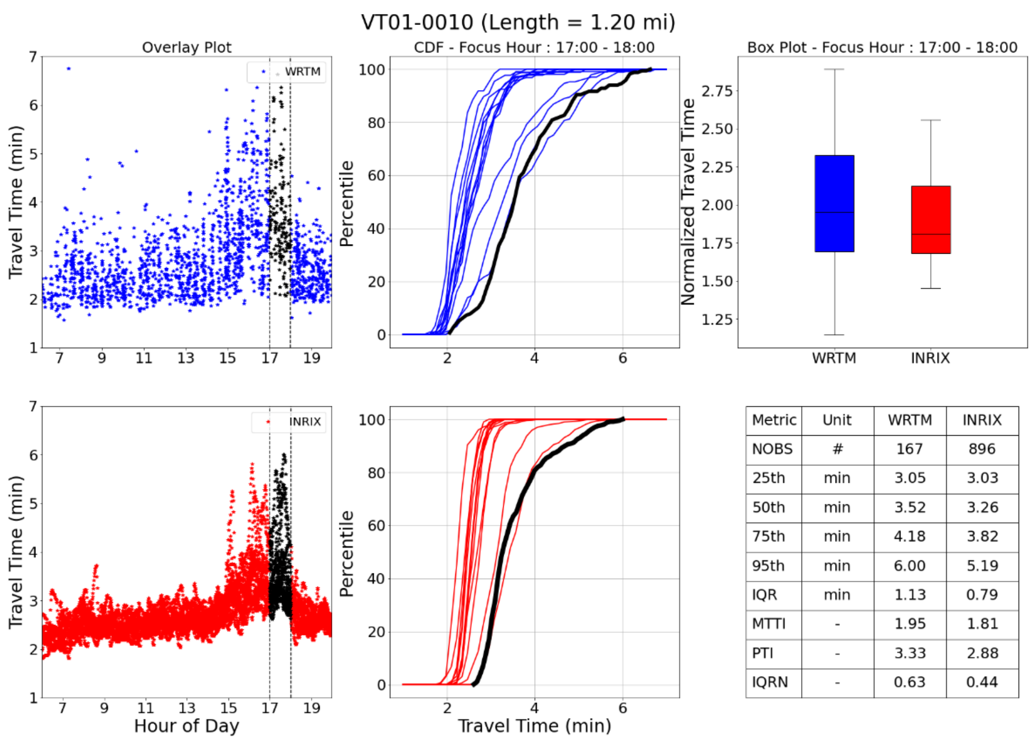
Figure 13 shows the distribution of PM delta values across all validation paths. Many locations have small Δ Median values, indicating general agreement in median travel times across hourly periods. Δ IQR values are quite consistent across locations as well, showing that the vendor data slightly underreports IQR values. Δ PT values are much more noticeably negative - especially for a handful of paths (e.g., P10, P12, P17-P20) – an intuitive result given that these segments have multiple 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 reflected in the average vendor data (and thus impacts the 95th percentile travel time values used in the Planning Time performance measure computation).

Figure 14 shows detailed results for path VT01-0010, whose travel times were previously shown in Figure 6 for the entire validation period, and which is one of the few locations to exhibit repeatable time-of-day patterns. The overlay plot shows that WRTM travel times are typically concentrated around 2 – 3.5 minutes for most hours between 6 AM and 8PM, but can reach upwards of 5-7 minutes between 3-6 PM (hours 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 are lower in magnitude than was observed with. Because there is less variation in the vendor travel time records than WRTM for each hourly period, the delta Planning Time value is 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	#	167	896
25th	min	3.05	3.03
50th	min	3.52	3.26
75th	min	4.18	3.82
95th	min	6.00	5.19
IQR	min	1.13	0.79
MTTI	-	1.95	1.81
PTI	-	3.33	2.88
IQRN	-	0.63	0.44

Figure 14 – INRIX overlay and CFD plots for segment VT01-0010.

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 Median Travel Time)**:
  - Δ Median values are centered close to 0 for the lowest congestion level in Figure 15, and then increasingly become slightly negatively biased for higher congestion levels. Overall, however, WRTM and vendor median travel time values match well during many periods.
  - Median travel times show strong agreement when there is minimal travel time variation, and a slight negative bias for higher observed levels of variation.

- **Reliability Comparison (Δ IQR Travel Time, Δ Planning Time)**:
  - Reliability PM values (IQR, PT) are close when congestion and variation levels are low. As congestion and travel time variation increase, the Vendor increasingly underreports true travel time variation (more negative delta values).

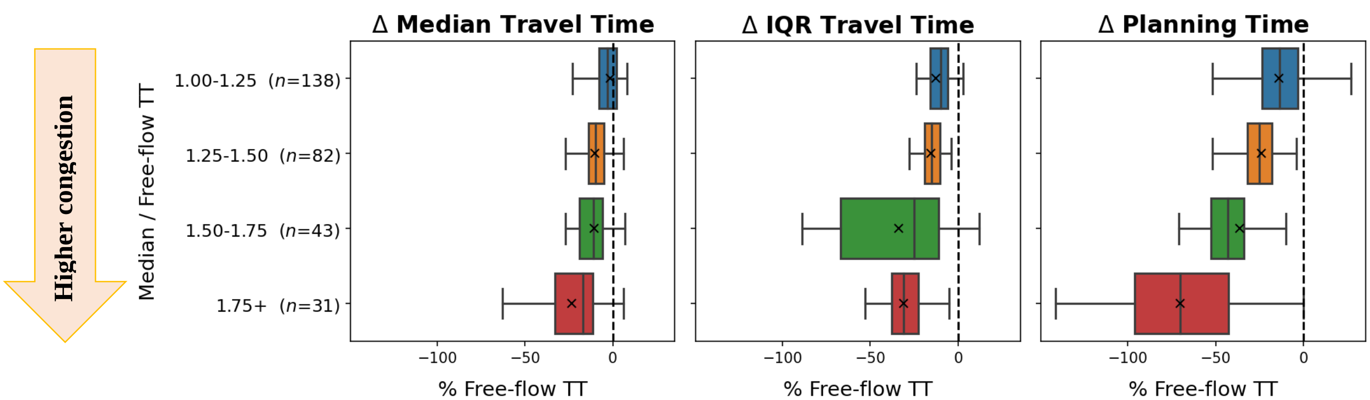


Figure 15 - TomTom summary of delta PM values grouped by congestion level

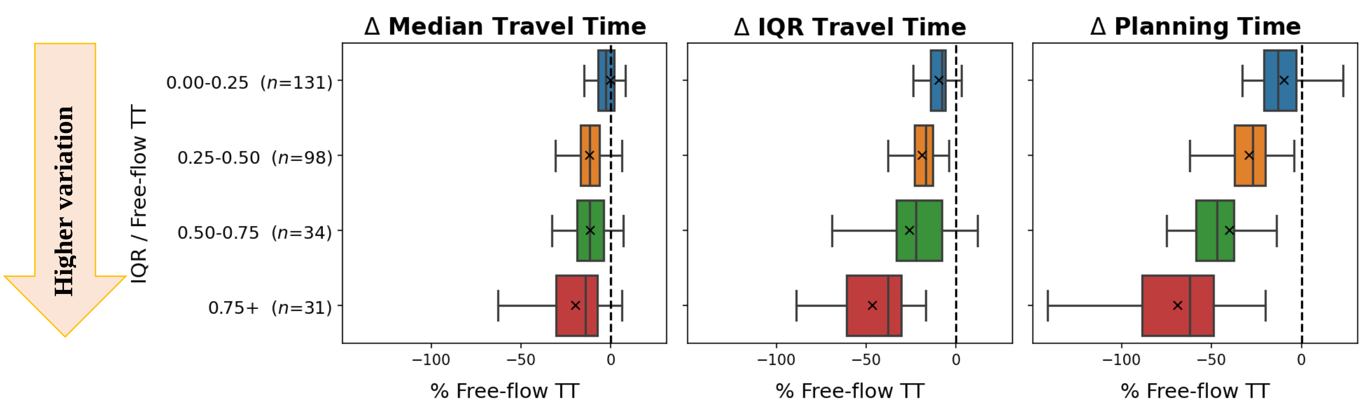


Figure 16 - TomTom summary of delta PM values grouped by level of travel time variation

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

Performance Measure (PM)	Median Travel Time	IQR Travel Time	Planning Time	
PM Definition	TT_{50}	$(TT_{75} - TT_{25})$	TT_{95}	
	$\overline{\Delta \text{ Median}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta \text{ IQR}}$ (% $TT_{\text{Free flow}}$)	$\overline{\Delta \text{ PT}}$ (% $TT_{\text{Free flow}}$)	**Samples**
Overall	-8	-18	-26	294
Higher congestion — **Median / Free-flow TT**				
1.00 – 1.25	-1	-13	-14	138
1.25 – 1.50	-10	-16	-24	82
1.50 – 1.75	-11	-34	-36	43
> 1.75	-24	-31	-70	31
Higher variation — **IQR / Free-flow TT**				
0.00 – 0.25	0	-9	-10	131
0.25 – 0.50	-12	-19	-29	98
0.50 – 0.75	-12	-26	-40	34
> 0.75	-20	-47	-69	31

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. Many locations have small Δ Median values, indicating general agreement in median travel times across hourly periods. Δ IQR values are quite consistent across locations as well, showing that the vendor data slightly underreports IQR values. Δ PT values are much more noticeably negative - especially for a handful of paths (e.g., P10, P12, P17-P20) – an intuitive result given that these segments have multiple 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 reflected in the average vendor data (and thus impacts the 95th percentile travel time values used in the Planning Time performance measure computation).

Figure 18 shows detailed results for path VT01-0010, whose travel times were previously shown in Figure 6 for the entire validation period, and which is one of the few locations to exhibit repeatable time-of-day patterns. The overlay plot shows that WRTM travel times are typically concentrated around 2 – 3.5 minutes for most hours between 6 AM and 8PM, but can reach upwards of 5-7 minutes between 3-6 PM (hours 15-18 on the graph). The Vendor data shows a similar pattern, albeit with slightly lower travel time values on average, and a few higher observations during the PM peak hours. Because the Vendor exhibits a similar level of variation as the WRTM data, the Vendor Planning Time values are reasonably close to their WRTM counterparts, resulting in small negative delta values.

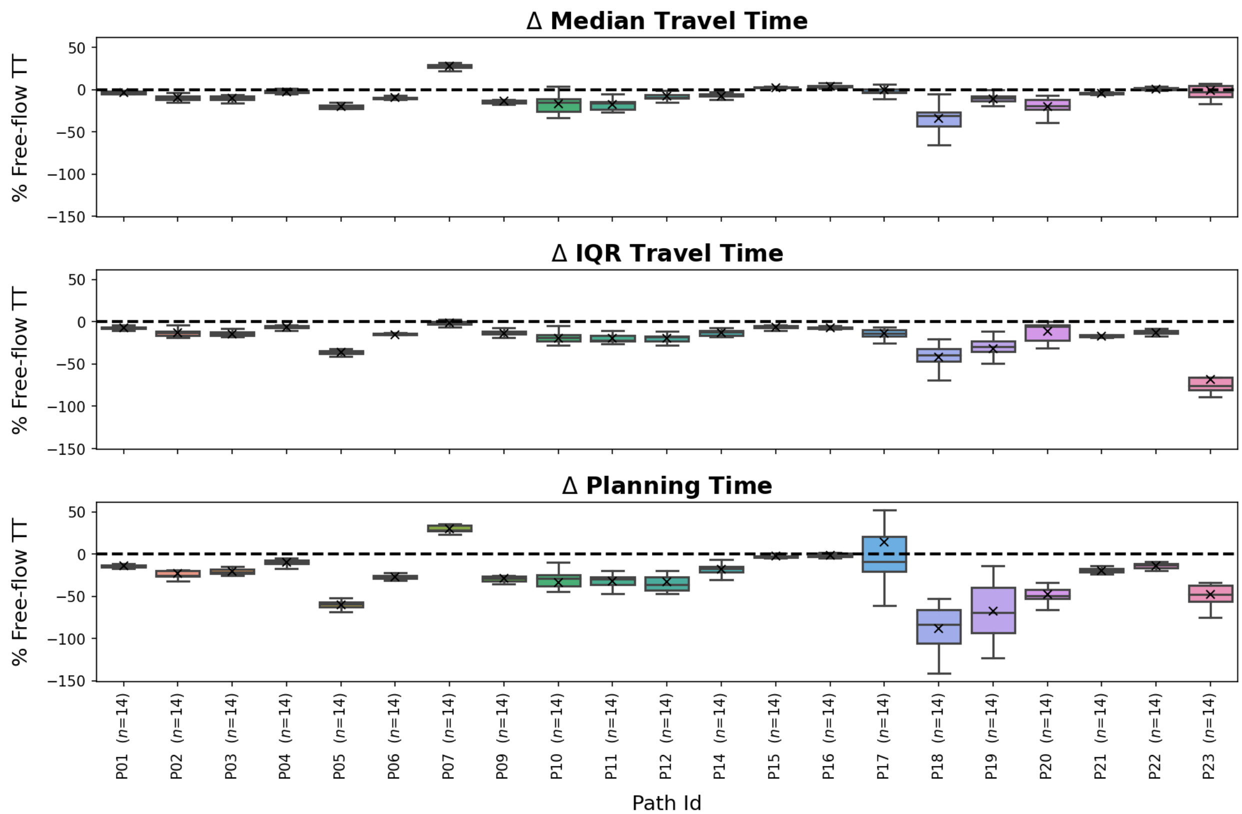
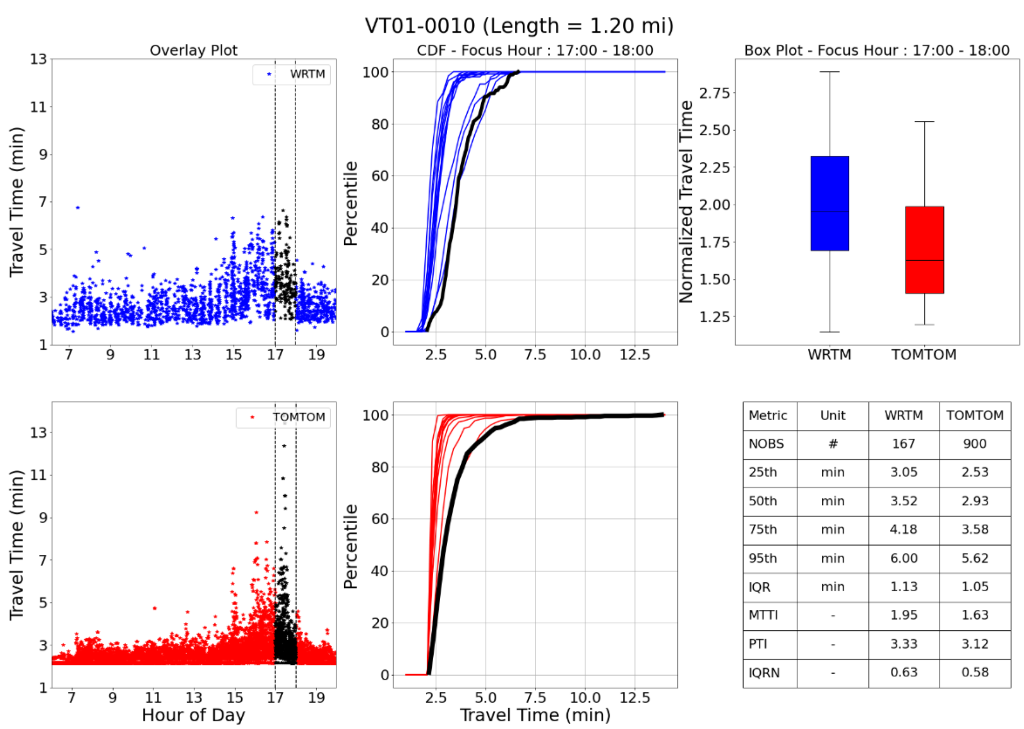


Figure 17 – TomTom path-level delta PM values



Metric	Unit	WRTM	TOMTOM
NOBS	#	167	900
25th	min	3.05	2.53
50th	min	3.52	2.93
75th	min	4.18	3.58
95th	min	6.00	5.62
IQR	min	1.13	1.05
MTTI	-	1.95	1.63
PTI	-	3.33	3.12
IQRN	-	0.63	0.58

Figure 18 – TomTom overlay and CFD plots for segment VT01-0010.

Summary and Discussion

This report supplements the traditional analysis with an additional approach to characterize probe data quality: the *sampled distribution method*. While the traditional analysis measures accuracy at any point in time, it can have limitations under low-volume conditions when there are minimal reference data samples to use for evaluation (e.g., analysis intervals that need to be dropped due to sufficient data, large confidence bands reflecting high uncertainty). Accordingly, 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. The top-level findings are summarized below for both operations and planning perspectives (based on the traditional and sampled distribution analyses, respectively):

Operations Perspective

- **Probe data was generally effective in representing ground truth traffic speed / travel times** during the periods where there was sufficient reference data for evaluation.
    - All vendors were compliant with VPP requirements for the Average Absolute Speed Error (AASE) measure specification and minimally achieved compliance in all but the lowest speed bin for the Speed Error Bias (SEB) specification.
    - Two vendors' SEB measures were slightly out of the target range in the lowest speed bin (0-15 MPH), but this scenario represented less than 1% of the observations used for evaluation and mostly came from a single segment.

- It was difficult to draw definitive conclusions about low-speed data quality, but **it is recommended that probe data be monitored in low-speed conditions.**
    - There were limited opportunities to evaluate data quality in the lowest speed bin, and in these limited samples two vendors were slightly out of target range for SEB.
    - The lowest speed bin has historically been most challenging for compliance even on higher-volume roads, so it is worth continuing to monitor in low-volume conditions that have been less studied.

Planning / Arterial Management Perspective

- Median vendor and reference travel time values matched 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.**
    - This agreement is particularly strong when there is less congestion and variation in travel times in each time-of-day period.

- Whenever there was significant variation in travel times during a time-of-day period, the vendors usually under-reported travel time variation compared to the reference data (i.e., overestimated travel time reliability). However, this may be at least partially explained by the inherent aggregation of the vendor data product (fact that probe data is reported in one-minute averages).

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 a combination of automated methods and human judgement, but it is nonetheless a challenging task that impacts the performance values used in the analysis.

The validation team plans to continue using a combination of the traditional analysis and sampled distribution method to evaluate probe data's ability to accurately report speed/travel time at each point in time, and additionally capture recurrent travel patterns that are useful for planning purposes. The sampled distribution method and related performance measures will likely be refined be refined over time based on feedback received from Coalition members.