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

Report for North Carolina (#10)
Arterial Validation (Low-volume Roadways)

Data Collection: June 2021
Report Date: November 2021

Prepared by:
Zachary Vander Laan, Stanley E. Young, Sara Zahedian
University of Maryland, College Park

Acknowledgements:
The research team would like to express its gratitude for the assistance it received from North Carolina Department of Transportation during the data collection phase of the validation effort.

1. Table of Contents

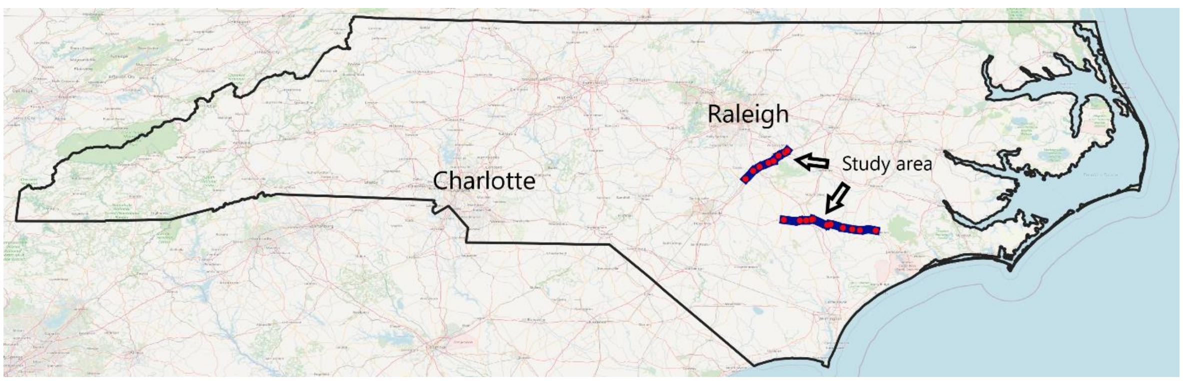
List of Tables

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. Probe data has been proven an effective representative along freeways and many high-volume arterials, and as such, the Coalition has recently commissioned multiple studies to investigate data quality on low-volume roads. This report follows two other recent low volume studies (Maryland and Vermont), focusing on two rural arterials in North Carolina.

This report uses reference data acquired in North Carolina on NC-24 and US-301 corridors (shown in ES Figure 1) from June 1-12, 2021. The road characteristics on these corridors are summarized in ES Table 1.



ES Figure 1 – Low-volume corridors used for data collection

ES Table 1 – Arterial Corridor Description in North Carolina

Corridor	# Segments	Miles (directional)	Speed Limit (MPH)	Signals / mile	Avg. AADT
NC-24	18	84	35-55	0.19	7.8k
US-301	14	50	35-55	0.86	8.9k

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 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, and the probability of observing travel time fluctuations is extremely low due to lower overall traffic volumes. Small sample sizes can impact both the number of 5-minute 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 different approach by focusing on vendors' ability to capture repeatable travel time patterns**.**

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 **there were few opportunities to evaluate low-speed data**; there were no observations in the lowest speed bin (0-15 mph), and less than 5% of all observations came from the second lowest bin (15-25 mph). Even the measurements in the second lowest bin (15-25 mph) do not reflect fluctuations in speed, but, after careful review of the data, primarily came from two segments in which traffic signal timing induced slower speeds consistently throughout the day.

Overall, each of the three vendors met contract specifications, in all speed bins (that has adequate ground truth data). The results are given below. However, it should be noted that no unusual traffic patterns – either recurring congestion, or non-recurring congestion occurred on these segments during the validation period. The few low-speed observations were all attributable to consistent signal delay on two segments. Even the signal delay was invariant by time of day.

ES Table 2 indicates that **HERE data was within contract specifications in all speed bins that had valid data for both AASE and SEB error measures**.

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	-	-	-	-	-
15-25 MPH	2.44	5.3	2.43	5.23	329
25-35 MPH	0.67	2.39	0.38	1.03	803
>35 MPH	1.32	3.55	-0.83	-1.98	6404
All Speeds	1.3	3.51	-0.55	-1.34	7536

ES Table 3 indicates that **INRIX data was within contract specifications in all speed bins that had valid data for both AASE and SEB error measures.**

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	-	-	-	-	-
15-25 MPH	1.07	3.15	1.05	2.86	325
25-35 MPH	0.82	2.7	0.31	0.5	808
>35 MPH	0.83	2.75	0.03	0.02	6467
All Speeds	0.84	2.76	0.1	0.19	7600

ES Table 4 indicates that **TomTom data was within contract specifications in all speed bins with valid data 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	-	-	-	-	-
15-25 MPH	1.31	3.23	1.27	2.6	338
25-35 MPH	0.97	2.98	0.65	1.59	847
>35 MPH	0.77	2.63	0.31	0.69	6895
All Speeds	0.81	2.69	0.39	0.86	8080

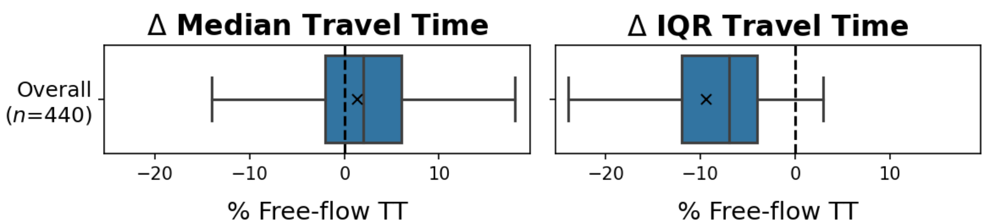
Sampled Distribution Method Results

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 periods (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.

Two simple performance measures, Median Travel Time (50th percentile) and IQR Travel Time (75th -25th percentile), are used throughout the report to characterize the central tendency and variation of travel times during each time-of-day period. 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., $\Delta\text{Median} = \text{Median}_{\text{Vendor}} - \text{Median}_{\text{WRTM}}$). To make performance measures comparable across segments, these values are normalized by the free-flow travel time for the segment, which is defined here as the 15th percentile travel time during overnight hours (10pm – 5am).

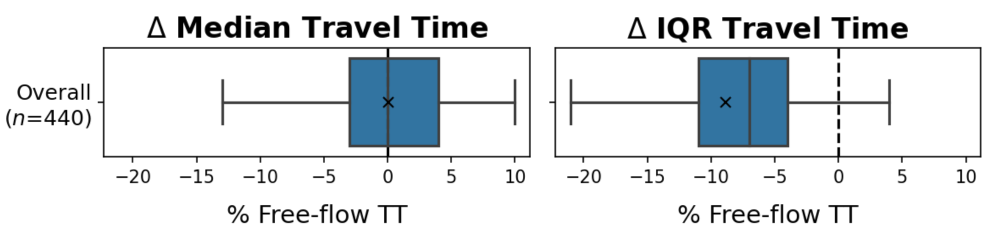
ES Figures 2-4 summarize the "deltas" (Δ) between Vendor and WRTM performance measures over 440 hourly time-of-day periods – **most of which have minimal, if any congestion**. The boxplots show the distribution of delta values, with the box containing the middle 50% of values and the whiskers representing values 1.5*IQR away from the box (i.e., lower whisker = Q1 – 1.5*IQR, upper whisker = Q3 + 1.5*IQR). Delta (Δ) values that are close to 0 show agreement between Vendor and WRTM measures, while delta values that are larger in magnitude represent discrepancies. Further details, including a breakdown of accuracy by level of congestion, can be found in the body of the report. Note that this analysis is intended to provide high level insight into vendors' ability to capture recurrent travel time patterns, *not measure compliance.*

ES Figure 2 shows that **HERE's median travel times agree with the reference data for most periods** and tend to be slightly higher than their WRTM counterparts (Δ Median is small – typically within +/- 5% of free-flow travel time and has a slight positive bias). HERE data shows a bit less variation than the reference data (Δ IQR is nearly always negative), with differences in IQR around 5-10% of the free-flow travel time.



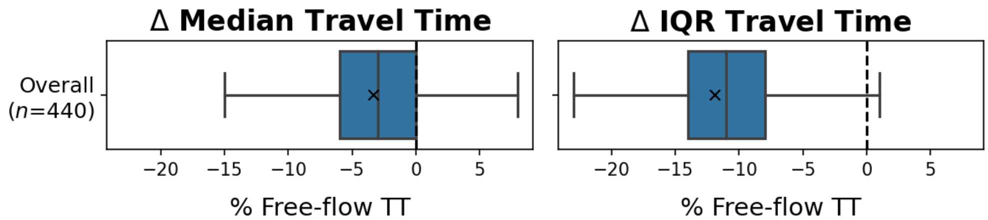
ES Figure 2 – HERE summary of Δ Median and Δ IQR values

ES Figure 3 shows that **INRIX's median travel times agree with the reference data for most periods** and do not show any noticeable bias (Δ Median is small – typically within +/- 5% of free-flow travel time and is centered around 0). INRIX data shows a bit less variation than the reference data (Δ IQR is nearly always negative), with differences in IQR around 5-10% of the free-flow travel time.



ES Figure 3 – INRIX summary of Δ Median and Δ IQR values

ES Figure 4 shows that **TomTom's median travel times agree with the reference data for most periods** and tend to be slightly lower than their WRTM counterparts (Δ Median is small – typically within 5% of free-flow travel time and has a slight negative bias). TomTom data shows a bit less variation than the reference data (Δ IQR is nearly always negative), with differences in IQR around 10% of the free-flow travel time.



ES Figure 4 – TomTom summary of Δ Median and Δ IQR values

Findings

This study reinforces the findings from recent low-volume arterial studies; probe vendors were consistently within contract specifications for the traditional analysis, but sparse data and lack of congestion made it difficult to evaluate the data quality in low-speed conditions. In this study, there were no opportunities to evaluate point-in-time accuracy in the 0-15 MPH speed bin, which typically is the most challenging for vendors to achieve. Likewise, while the sampled distribution analysis showed that vendors generally captured true time-of-day travel time patterns, the low-volume traffic conditions in this data collection did not provide meaningful opportunities to test data quality in challenging situations. As a result, **the validation team recommends significant changes in methodology for assessing data quality on low volume roadways** (discussed below).

In summary:

- This validation attempted to capture performance on rural, lower volume roads
- The test roads had low AADT (between 8000 and 9000) with sparse signal spacing
- No major congestion (as would be expected for this type of roadway) occurred during the test period
- The statistical validation results (both Traditional and Sampled Distribution) showed that vendors performed within contract specifications and reflected the normal patterns of traffic travel time experienced on the roadway.
- **No non-recurring congestion events (e.g., incident, inclement weather, special event) occurred during the data collection period to test vendors' ability to reflect non-recurring congestion on such roadways.**

Overall, this validation, as well as previous validations on lower volume roads, exposed that current validation methods, originally conceived for highly congested urban roadways, are inadequate to reflect the quality of probe data on lower volume roads. To remedy this, the validation team recommends the following:

- **Concentrate data validation on low-volume roads during periods of known traffic disruption**. Any validation activities should concentrate on fluctuations in expected speed and travel time. With low volume roads this is more aligned with weather events, road work, and/or incidents, and not predictable volume induced congestion. In contrast to a set two-week data collection, the validation should identify periods of time in which such fluctuations have occurred, and then compare how well vendors captured such fluctuations.

- **Compare travel time data between vendors, as well as against emerging connected vehicle data that can be acquired after the fact.** Because traffic disruptions are so infrequent in low-volume conditions, the current data collection approach (i.e., field collection via wireless re-identification sensors) would become prohibitively expensive, requiring sensor deployment for long periods of time until appropriate phenomena occur.

  As such, once a traffic disruption event is identified, the Coalition will request data from the vendor for roadways that would have been impacted and (1) check relative agreement among the various vendors as an indication of consistency and accuracy, and (2) procure connected vehicle data from the same roadways and compare the available trace vehicle data with from travel time vendors as another source of reference comparison.

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. Also, low-volume roadways experience much less congestion (either recurring or non-recurring), further limiting the ability to test the dynamic range of vendors on such roadways. As such, this report conducts both the traditional point-in-time analysis and an additional sampled distribution method to characterize data accuracy from multiple perspectives.

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

[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)

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.

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, 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.

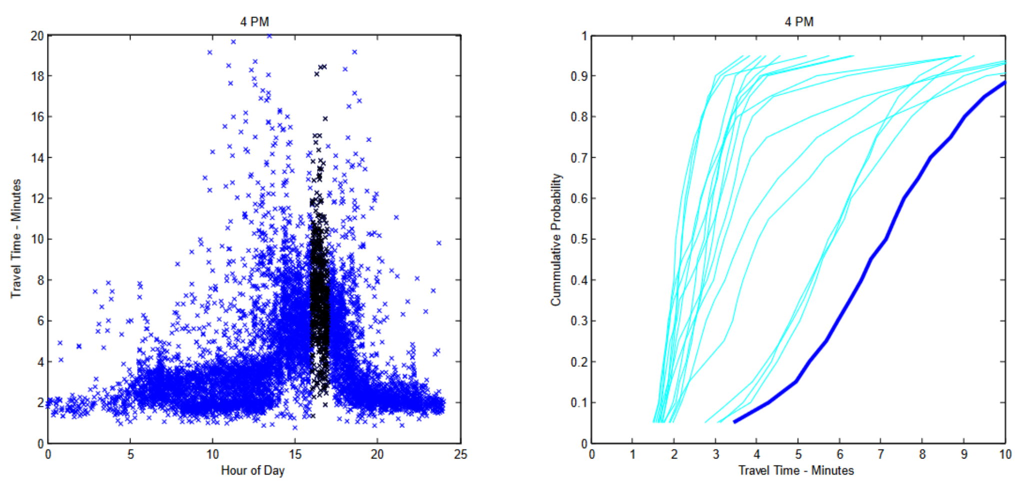
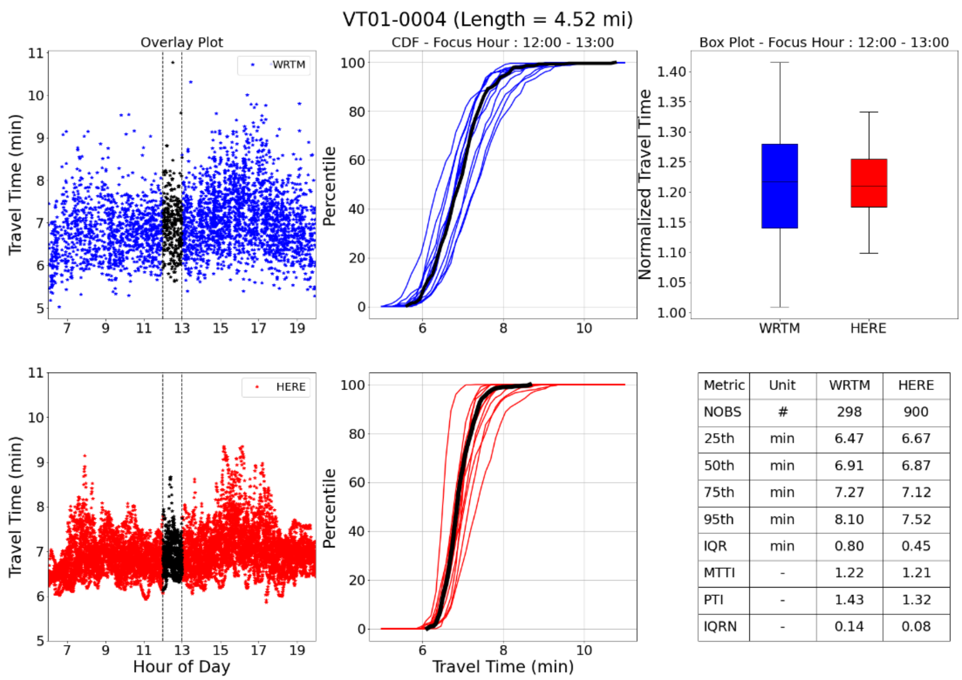


Figure 1 – Example 24-hr overlay plot and cumulative frequency diagram.

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).

Percentile-based performance measures are used throughout the analysis and are defined using the notation TT_X, where X refers to the Xth percentile travel time. Examples include:

- **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

Additionally, it is useful to summarize WRTM and Vendor 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., $\Delta\ \text{Median} = \text{Median}_{\text{Vendor}} - \text{Median}_{\text{WRTM}}$).

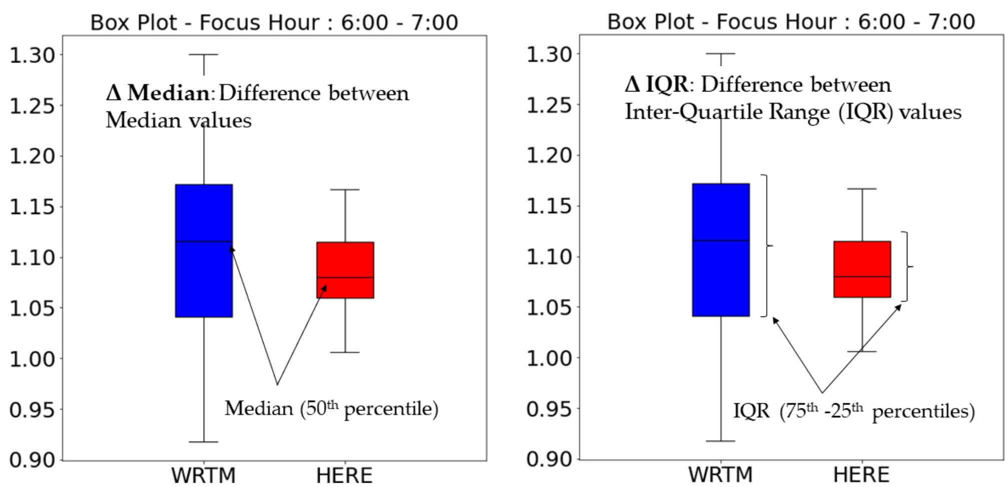


Figure 3 – Visual explanation of Delta Median and Delta IQR calculations

Because delta values will be summarized across road segments with different speed limits and lengths, it generally makes sense to express them 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.

- **By level of travel time variation (TT_{IQR}/TT_{FF})**: Each hourly period is assigned a variation bin based on the WRTM's IQRN value.

- **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) 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. Figure 4 gives an example of one such aggregation, using box and whisker plots to visualize the Δ Median and Δ IQR performance measure values separately for three congestion levels (where congestion levels are defined by the ratio of median to free-flow travel time).

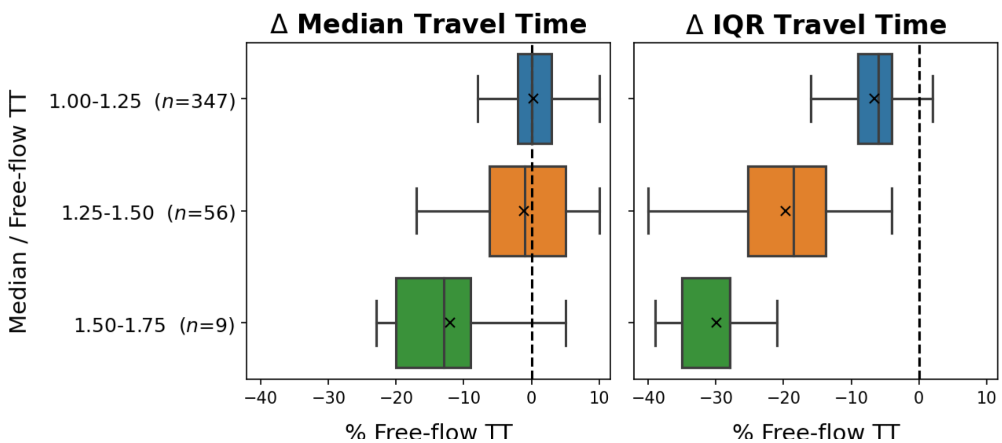


Figure 4 – Example of delta Median and delta IQR measures reported by congestion level

Data Collection

Travel time samples were collected along directional validation road segments in North Carolina from June 1-12, 2021, which are shown below in Figure 5. The data was collected with the support of North Carolina Department of Transportation (NCDOT) using temporary Bluetooth and Wi-fi-based sensors to acquire the re-identification data.

Table 1 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. There were a few minor issues with the reference data that needed to be addressed to ensure that vendor data was fairly evaluated. One sensor (S02) experienced clock issues during part of the data collection, so validation paths starting and ending at these sensors were omitted from the analysis during times when the reference data was potentially suspect.

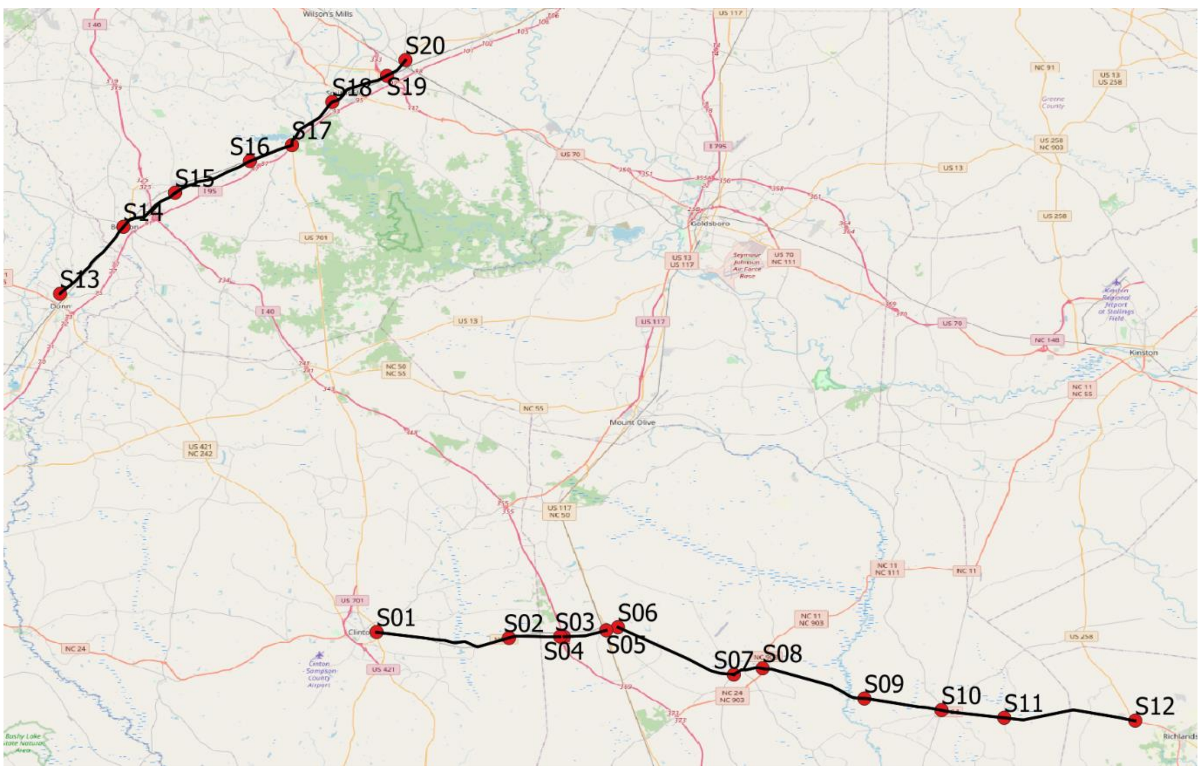


Figure 5 – WRTM Sensor locations

Table 1 – 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	NC-24	Souteast Blvd	1	1.00	6927	26	34.99761	-78.30918	7.36
NC10-0001	Eastbound	N Main st	2	0.14		55	34.99225	-78.18378	
A2	NC-24	N Main st	1	0.00	8545	11	34.99225	-78.18378	2.78
NC10-0002	Eastbound	I-40 (South)	2	0.00		55	34.99339	-78.13516	
A3	NC-24	I-40 (North)	2	1.00	7673	16	34.99345	-78.13146	2.32
NC10-0003	Eastbound	N Pine St	2	0.43		35	34.99937	-78.09149	
A4	NC-24	kenansvite Hwy	1	1.00	4074	28	35.00229	-78.08029	7.03
NC10-0004	Eastbound	Mallard St	2	0.14		55	34.95759	-77.97013	
A5	NC-24	Mallard St	1	2.00	5361	16	34.95759	-77.97013	1.67
NC10-0005	Eastbound	Routledge Rd	2	1.20		35	34.96325	-77.94319	
A6	NC-24	Routledge Rd	2	1.00	8701	14	34.96325	-77.94319	5.91
NC10-0006	Eastbound	N Williams Rd	2	0.17		55	34.93475	-77.84650	
A7	NC-24	N Williams Rd	2	0.00	9560	17	34.93475	-77.84650	4.21
NC10-0007	Eastbound	S Jackson St	2	0.00		55	34.92370	-77.77362	
A8	NC-24	S Jackson St	2	2.00	9998	15	34.92370	-77.77362	3.42
NC10-0008	Eastbound	Fountaintown Rd	2	0.58		55	34.91641	-77.71405	
A9	NC-24	Fountaintown Rd	2	0.00	9901	21	34.91641	-77.71405	7.21
NC10-0009	Eastbound	Kinston Hwy	2	0.00		55	34.91366	-77.58994	
A10	NC-24	Kinston Hwy	2	1.00	9902	21	34.91366	-77.58994	7.21
NC10-0010	Westbound	Fountaintown Rd	2	0.14		55	34.91659	-77.71405	
A11	NC-24	Fountaintown Rd	2	1.00	9999	15	34.91659	-77.71405	3.42
NC10-0011	Westbound	S Jackson St	2	0.29		55	34.92370	-77.77362	

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									
A12 NC10-0012	NC-24 Westbound	S Jackson St N Williams Rd	2 2	1 0.24	9559	17 55	34.92370 34.93498	-77.77362 -77.84649	4.21
A13 NC10-0013	NC-24 Westbound	N Williams Rd Routledge Rd	2 2	0 0.00	8701	14 55	34.93498 34.96325	-77.84649 -77.94319	5.91
A14 NC10-0014	NC-24 Westbound	Routledge Rd Mallard St	1 2	2 1.19	5294	16 35	34.96325 34.95759	-77.94319 -77.97013	1.68
A15 NC10-0015	NC-24 Westbound	Mallard St kenansvite Hwy	1 2	1 0.14	4074	28 55	34.95759 35.00229	-77.97013 -78.08029	7.03
A16 NC10-0016	NC-24 Westbound	N Pine St I-40 (North)	2 2	1 0.43	7673	16 35	34.99937 34.99345	-78.09149 -78.13146	2.32
A17 NC10-0017	NC-24 Westbound	I-40 (South) N Main st	1 2	1 0.36	8545	11 55	34.99339 34.99225	-78.13516 -78.18378	2.78
A18 NC10-0018	NC-24 Westbound	N Main st Souteast Blvd	1 2	0 0.00	6927	26 55	34.99225 34.99761	-78.18378 -78.30918	7.36
A19 NC10-0019	US-301 Northbound	W Granville St E Main St	1 1	2 0.36	6566	29 45	35.31850 35.38200	-78.60927 -78.54894	5.58
A20 NC10-0020	US-301 Northbound	E Main St Raleigh Rd	1 2	2 0.54	3832	21 55	35.38200 35.41398	-78.54894 -78.50006	3.68
A21 NC10-0021	US-301 Northbound	Raleigh Rd S Main St	1 1	0 0.00	3643	19 55	35.41398 35.44406	-78.50006 -78.42904	4.57
A22 NC10-0022	US-301 Northbound	S Main St US-701	1 1	3 1.21	4483	17 45	35.44406 35.45946	-78.42904 -78.38934	2.49

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								
A23 NC10-0023	US-301 Northbound	US-701 Brogden Rd	1 1	1 0.27	11924	31 45	35.45946 -78.38934 35.50008 -78.35105	3.66
A24 NC10-0024	US-301 Northbound	Brogden Rd US-70	2 2	9 2.58	22386	56 35	35.50008 -78.35105 35.52497 -78.29962	3.49
A25 NC10-0025	US-301 Northbound	US-70 W Richardson St	1 2	5 3.43	14063	22 35	35.52497 -78.29962 35.53963 -78.28174	1.46
A26 NC10-0026	US-301 Southbound	W Richardson St US-70	1 2	5 3.43	14063	22 35	35.53963 -78.28174 35.52497 -78.29962	1.46
A27 NC10-0027	US-301 Southbound	US-70 Brogden Rd	2 2	9 2.58	22386	56 35	35.52497 -78.29962 35.50008 -78.35105	3.49
A28 NC10-0028	US-301 Southbound	Brogden Rd US-701	1 1	2 0.55	11916	31 45	35.50008 -78.35105 35.45946 -78.38934	3.66
A29 NC10-0029	US-301 Southbound	US-701 S Main St	1 1	1 0.40	4483	17 45	35.45970 -78.38968 35.44406 -78.42904	2.48
A30 NC10-0030	US-301 Southbound	S Main St Raleigh Rd	1 1	1 0.22	3643	19 55	35.44406 -78.42904 35.41398 -78.50006	4.57
A31 NC10-0031	US-301 Southbound	Raleigh Rd E Main St	1 2	1 0.27	3832	21 55	35.41398 -78.50006 35.38200 -78.54894	3.68
A32 NC10-0032	US-301 Southbound	E Main St W Granville St	1 1	2 0.36	6566	29 45	35.38200 -78.54894 35.31850 -78.60927	5.58

Traditional Validation Results

Tables 2-4 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 given the low-volume nature of this location, **the lowest speed bin (0-15 mph) does not have any observations, while the second lowest bin (15-25%) has only about 5% of all 5-minute samples used for evaluation.**

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 2 reports the traditional analysis error metrics for HERE. The Average Absolute Speed Error (AASE) is within specification in all speed bins that have observations when compared to both the SEM band and mean. The Speed Error Bias (SEB) is within specification in all speed bins that have observations when compared to the SEM Band, and all but the 15-25 MPH bin when compared to the Mean.

Table 2 – 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	-	-	-	-	-
15-25 MPH	2.44	5.3	2.43	5.23	329
25-35 MPH	0.67	2.39	0.38	1.03	803
>35 MPH	1.32	3.55	-0.83	-1.98	6404
All Speeds	1.3	3.51	-0.55	-1.34	7536

INRIX

Table 3 reports the traditional analysis error metrics for INRIX Both the Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) are within specification in all speed bins that have observations when compared to the SEM band and Mean.

Table 3 – 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	-	-	-	-	-
15-25 MPH	1.07	3.15	1.05	2.86	325
25-35 MPH	0.82	2.7	0.31	0.5	808
>35 MPH	0.83	2.75	0.03	0.02	6467
All Speeds	0.84	2.76	0.1	0.19	7600

TomTom

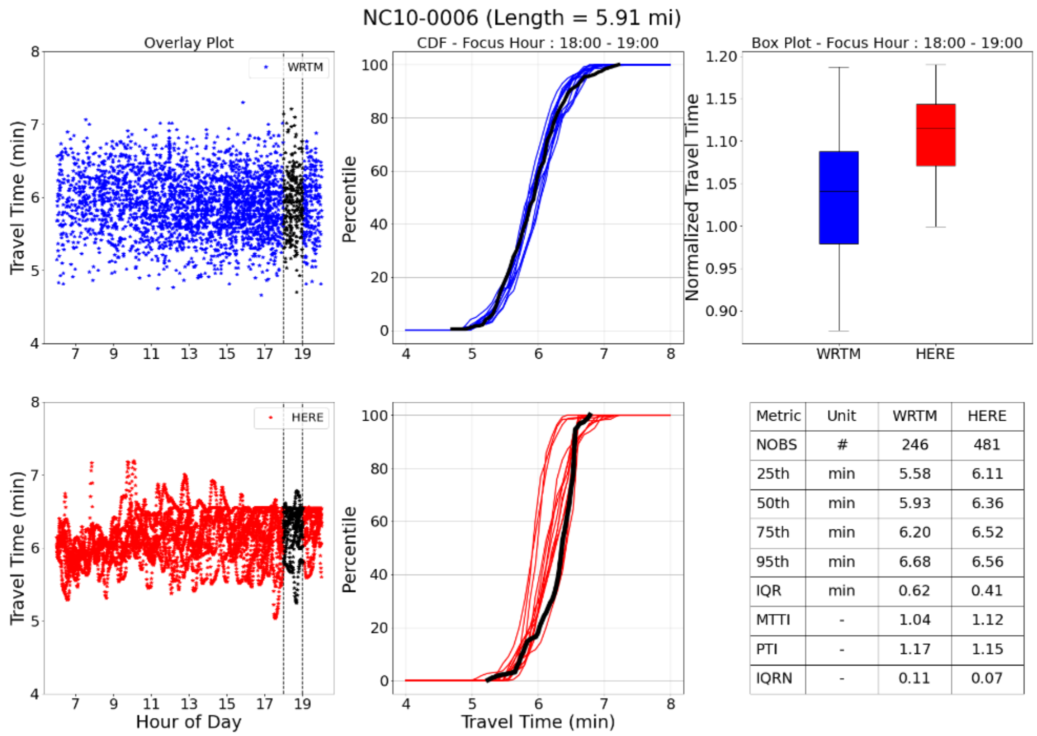
Table 4 reports the traditional analysis error metrics for TomTom. Both the Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) are within specification for all speed bins when compared to the SEM Band and Mean.

Table 4 – 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	-	-	-	-	-
15-25 MPH	1.31	3.23	1.27	2.6	338
25-35 MPH	0.97	2.98	0.65	1.59	847
>35 MPH	0.77	2.63	0.31	0.69	6895
All Speeds	0.81	2.69	0.39	0.86	8080

Sampled Distribution Method Results

24-hour overlay plots were generated for both reference and vendor data along validation paths and were subsequently used to produce time-of-day CFD curves for each weekday hour between 6AM and 8PM (440 total curves). Although this distribution-based approach is well-suited to working with sparse data (since observations from the entire dataset are organized into time-of-day bins), the observed traffic patterns along the low-volume roadways in the study area were quite uninteresting. Figure 6 shows a representative example, where CFD curves are basically identical throughout different times of day without any notable patterns (e.g., no AM/PM peak spikes, signal timing patterns, bimodality, etc.).



Metric	Unit	WRTM	HERE
NOBS	#	246	481
25th	min	5.58	6.11
50th	min	5.93	6.36
75th	min	6.20	6.52
95th	min	6.68	6.56
IQR	min	0.62	0.41
MTTI	-	1.04	1.12
PTI	-	1.17	1.15
IQRN	-	0.11	0.07

Figure 6 – Example CFD plots for a single path.

As such, in this section we do not produce overlay/CFD plots in detail, and instead summarize the reference WRTM and vendor time-of-day distributions via boxplots. For each vendor, we produce boxplots for each validation path and weekday hour, which allows many travel time distributions to be shown visually in a compact way. Figure 7 is an example of the boxplots produced for a single path; reference WRTM travel time distributions are shown in blue with vendor data in orange. Each box contains a line in the middle that represents the median value, while the width represents the IQR value, making it straightforward to visually compare simple measures of central tendency and travel time variation.

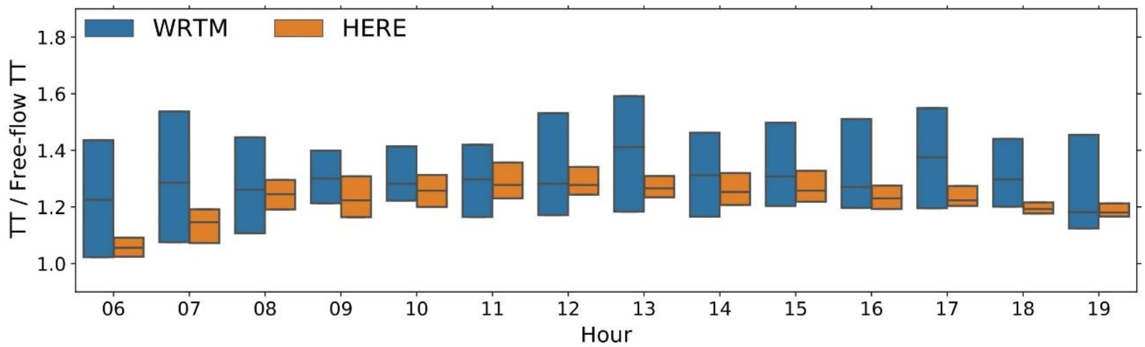


Figure 7 – Example time-of-day boxplots produced for a single path

Additionally, the WRTM and Vendor median/IQR values from each of the boxplots are extracted and compared, resulting in 440 "delta" values (i.e., Δ Median and Δ IQR values, which represent the difference in median/IQR). These values, which show how closely WRTM and Vendor measures of central tendency and variation match, are reported as overall measures and separately by level of congestion as described in the methodology section above. 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, at time the filtering strategy can play a role in the performance measure values

- 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

Figure 8 shows that HERE's median travel times agree with the reference data for most periods and tend to be slightly higher than their WRTM counterparts (Δ Median is small – typically within +/- 5% of free-flow travel time and has a slight positive bias). HERE data shows a bit less variation than the reference data (Δ IQR is nearly always negative), with differences in IQR around 5-10% of the free-flow travel time.

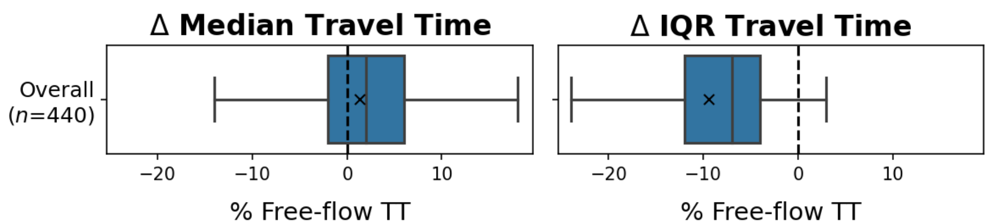


Figure 8 – HERE summary of Δ Median and Δ IQR values

Figure 9 reports the Δ Median and Δ IQR values separately for three congestion levels, where each period is assigned to a congestion bin based on the ratio of the median travel time to free flow speed. The three congestion bins are (1.00-1.25], (1.25-1.50], and (1.50-1.75], with higher values representing higher typical travel times relative to free-flow conditions. The number of observations in each bin are not evenly distributed (i.e., the lowest bin has 375 of the 440 observations) because most hourly time-of-day periods are characterized by minimal congestion. For both Median and IQR comparisons, the HERE probe data is closest to the reference source in the lowest bin and begins to underreport the Median/IQR as the congestion bin increases.

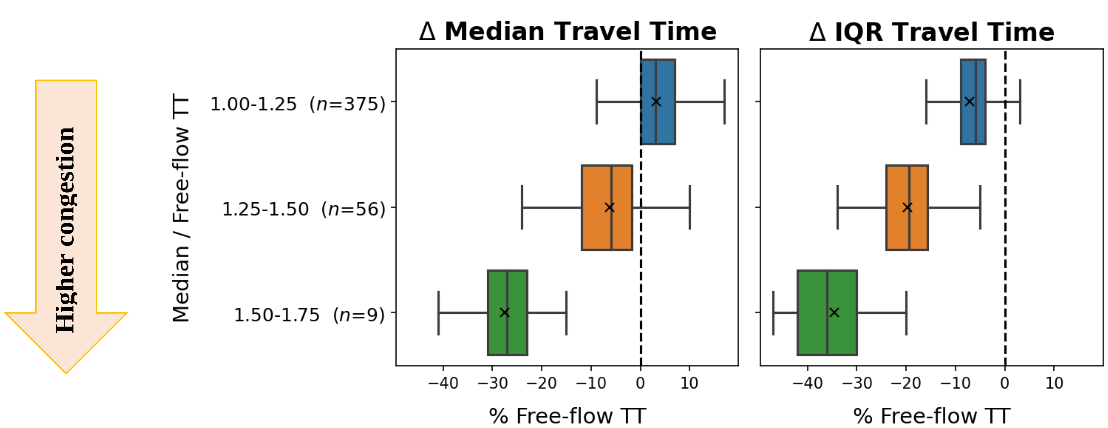


Figure 9 – HERE Δ Median and Δ IQR values summarized by congestion bin

Figures 10-13, shown on the following pages, contain boxplots summarizing the Vendor and WRTM travel time distributions for all validation paths and time-of-day periods. The summary figures on this page (Figures 8 and 9) were constructed by comparing the median and IQR values for each of the boxplots in Figures 10-13 and aggregating the results.

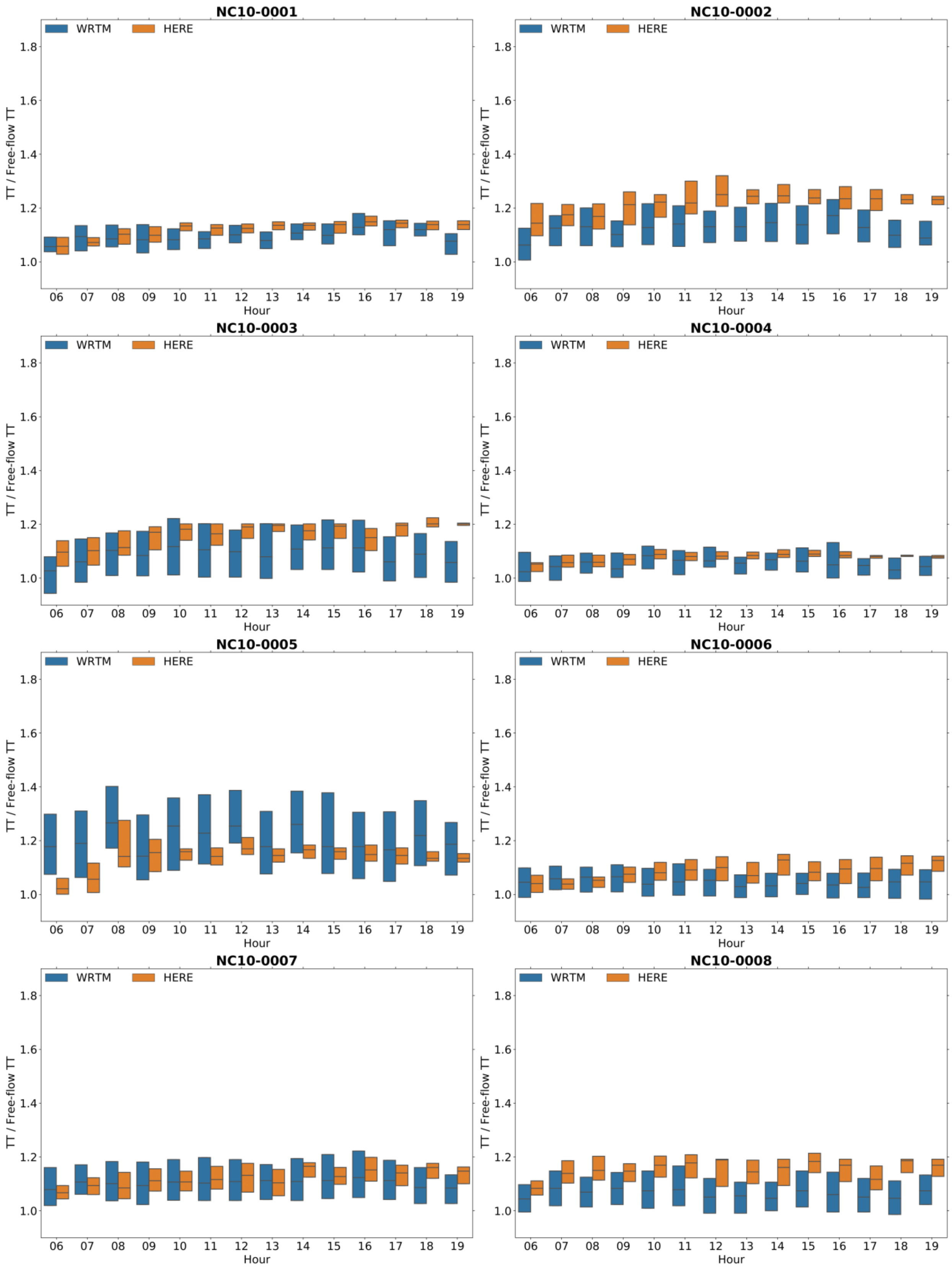


Figure 10 – HERE path-level travel time boxplots for each time-of-day period (1/4)

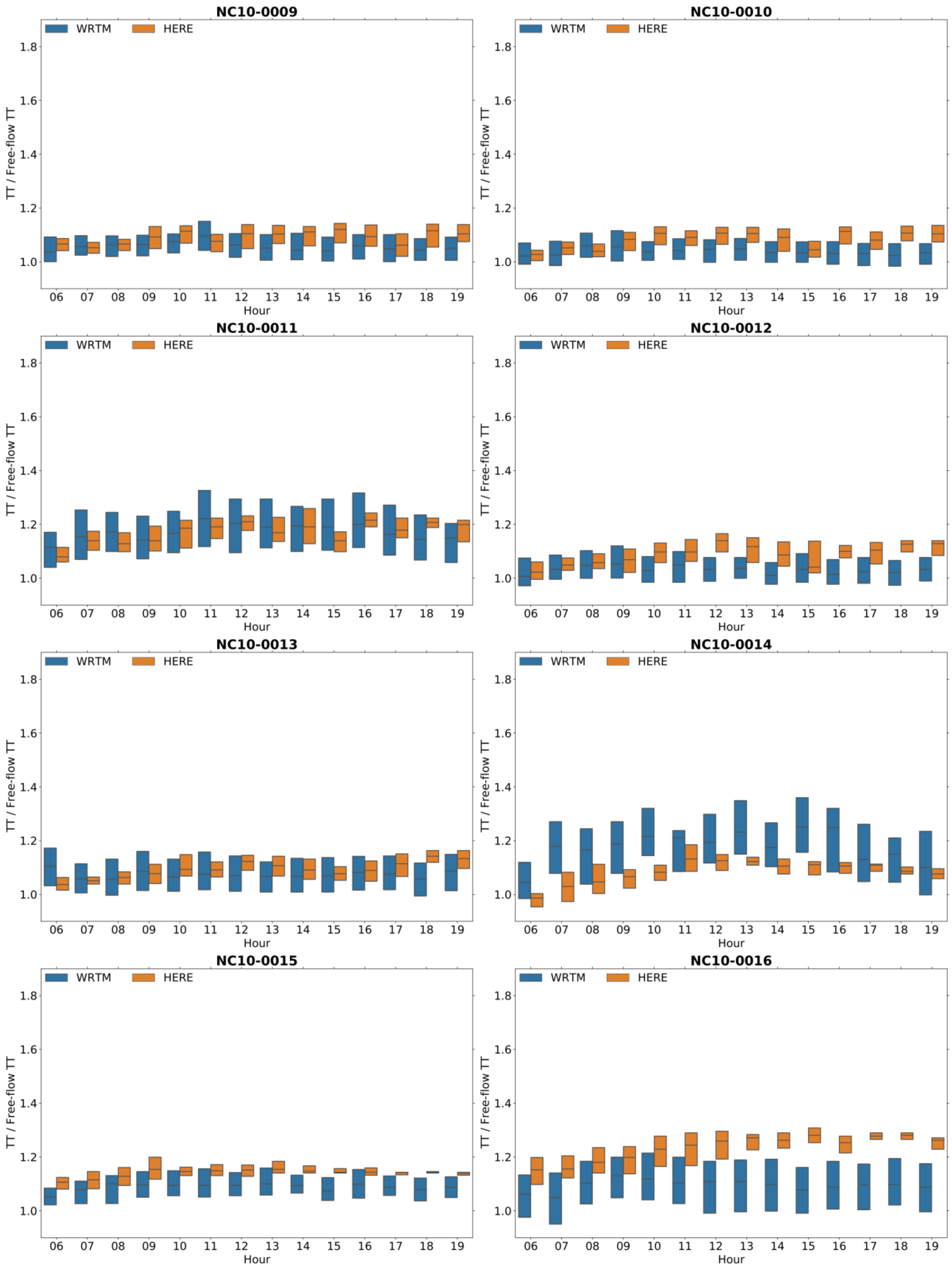


Figure 11 – HERE path-level travel time boxplots for each time-of-day period (2/4)

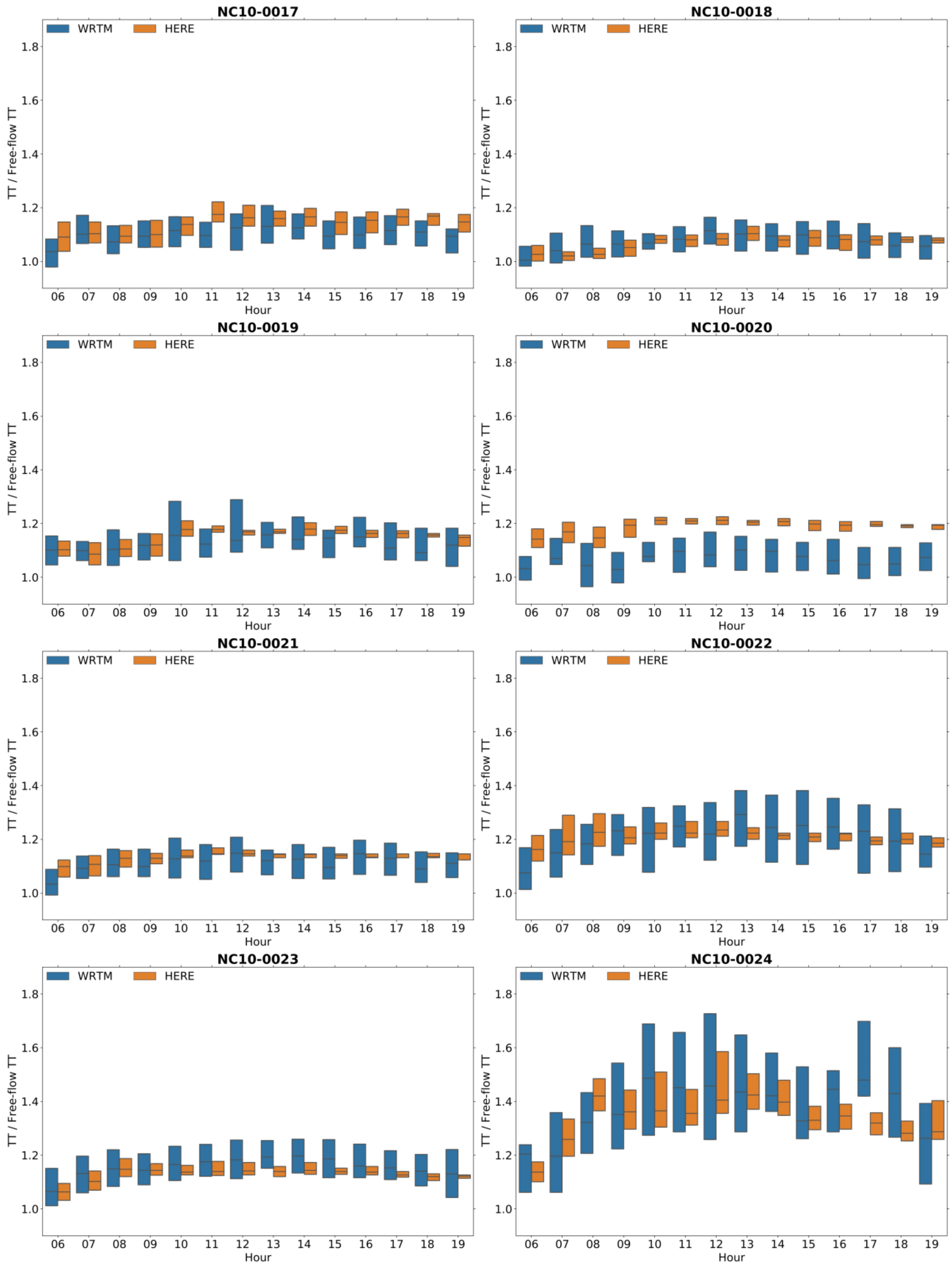


Figure 12 – HERE path-level travel time boxplots for each time-of-day period (3/4)

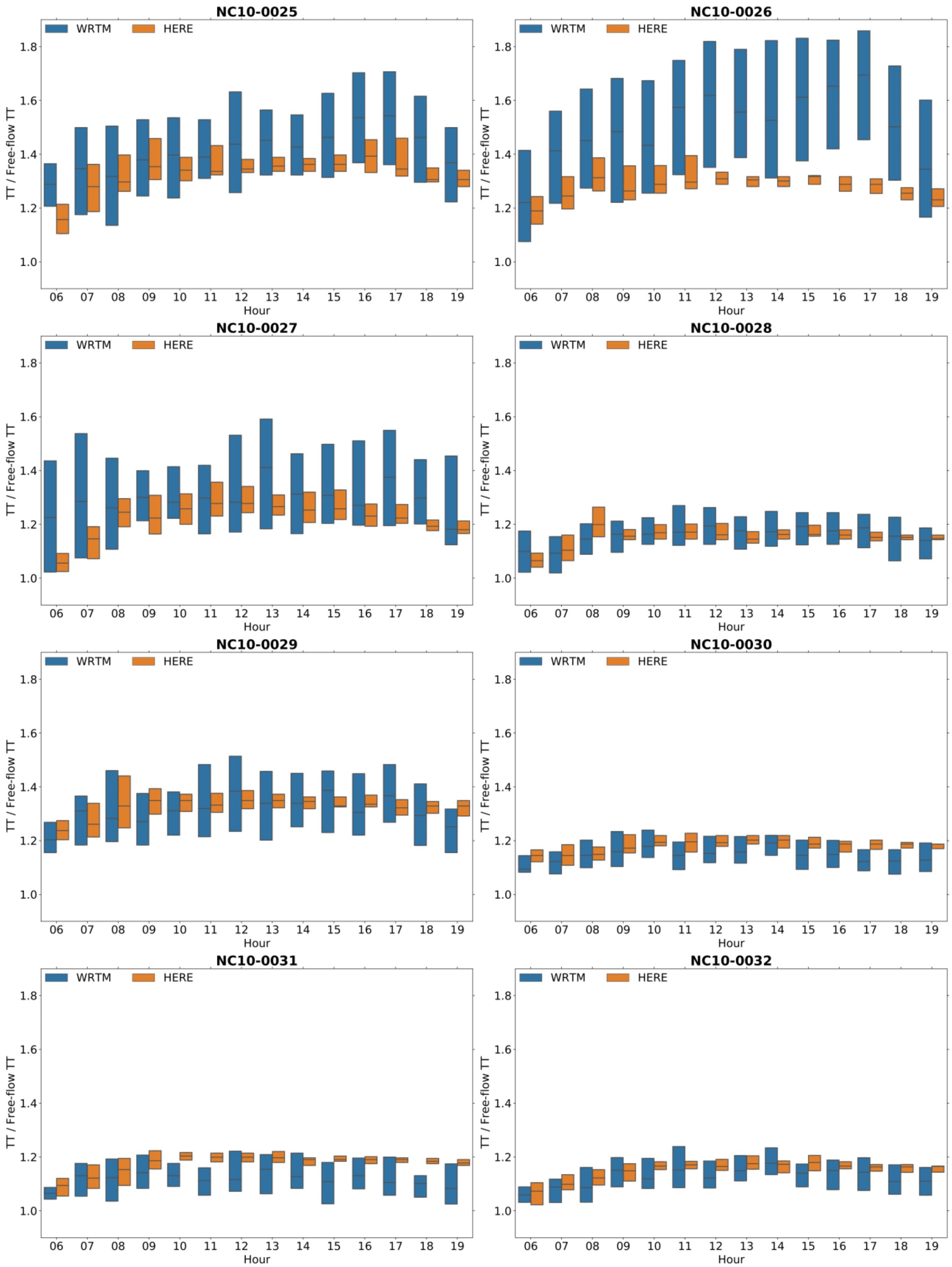


Figure 13 – HERE path-level travel time boxplots for each time-of-day period (4/4)

INRIX

Figure 14 shows that INRIX's median travel times agree with the reference data for most periods and do not show any noticeable bias (Δ Median is small – typically within +/- 5% of free-flow travel time and is centered around 0). INRIX data shows a bit less variation than the reference data (Δ IQR is nearly always negative), with differences in IQR around 5-10% of the free-flow travel time.

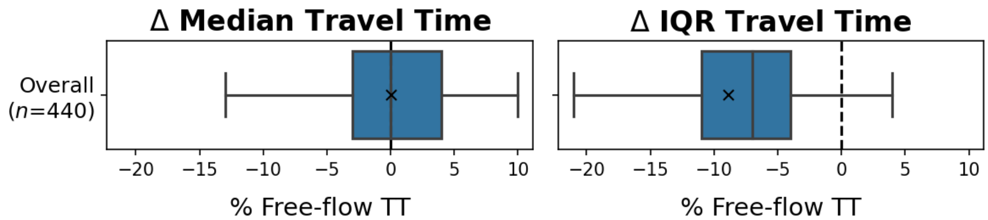


Figure 14 – INRIX summary of Δ Median and Δ IQR values

Figure 15 reports the Δ Median and Δ IQR values separately for three congestion levels, where each period is assigned to a congestion bin based on the ratio of the median travel time to free flow speed. The three congestion bins are (1.00-1.25], (1.25-1.50], and (1.50-1.75], with higher values representing higher typical travel times relative to free-flow conditions. The number of observations in each bin are not evenly distributed (i.e., the lowest bin has 375 of the 440 observations) because most hourly time-of-day periods are characterized by minimal congestion. INRIX Median travel times were very close to the reference values in the two lowest congestion bins and underreported by about 10-20% in the highest bin. INRIX IQR values were closest to the reference source in the lowest bin and increasingly began to underreport variation for higher levels of congestion.

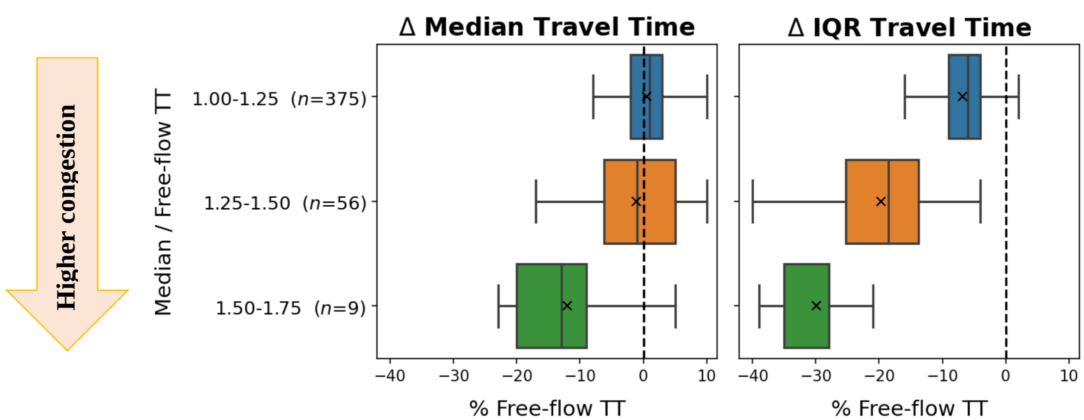


Figure 15 – INRIX Δ Median and Δ IQR values summarized by congestion bin

Figures 16-19, shown on the following pages, contain boxplots summarizing the Vendor and WRTM travel time distributions for all validation paths and time-of-day periods. The summary figures on this page (Figures 14 and 15) were constructed by comparing the median and IQR values for each of the boxplots in Figures 16-19 and aggregating the results.

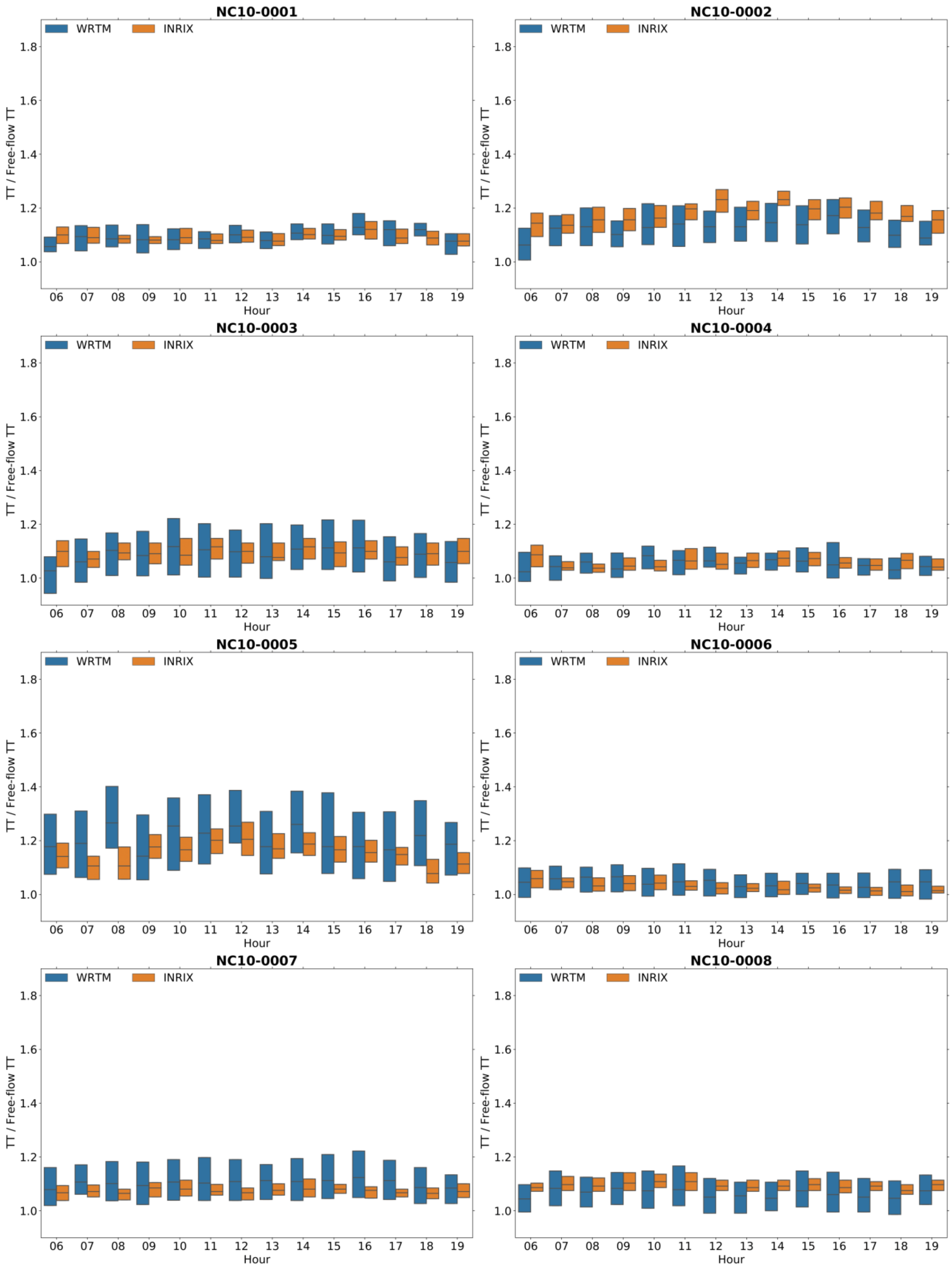


Figure 16 – INRIX path-level travel time boxplots for each time-of-day period (1/4)

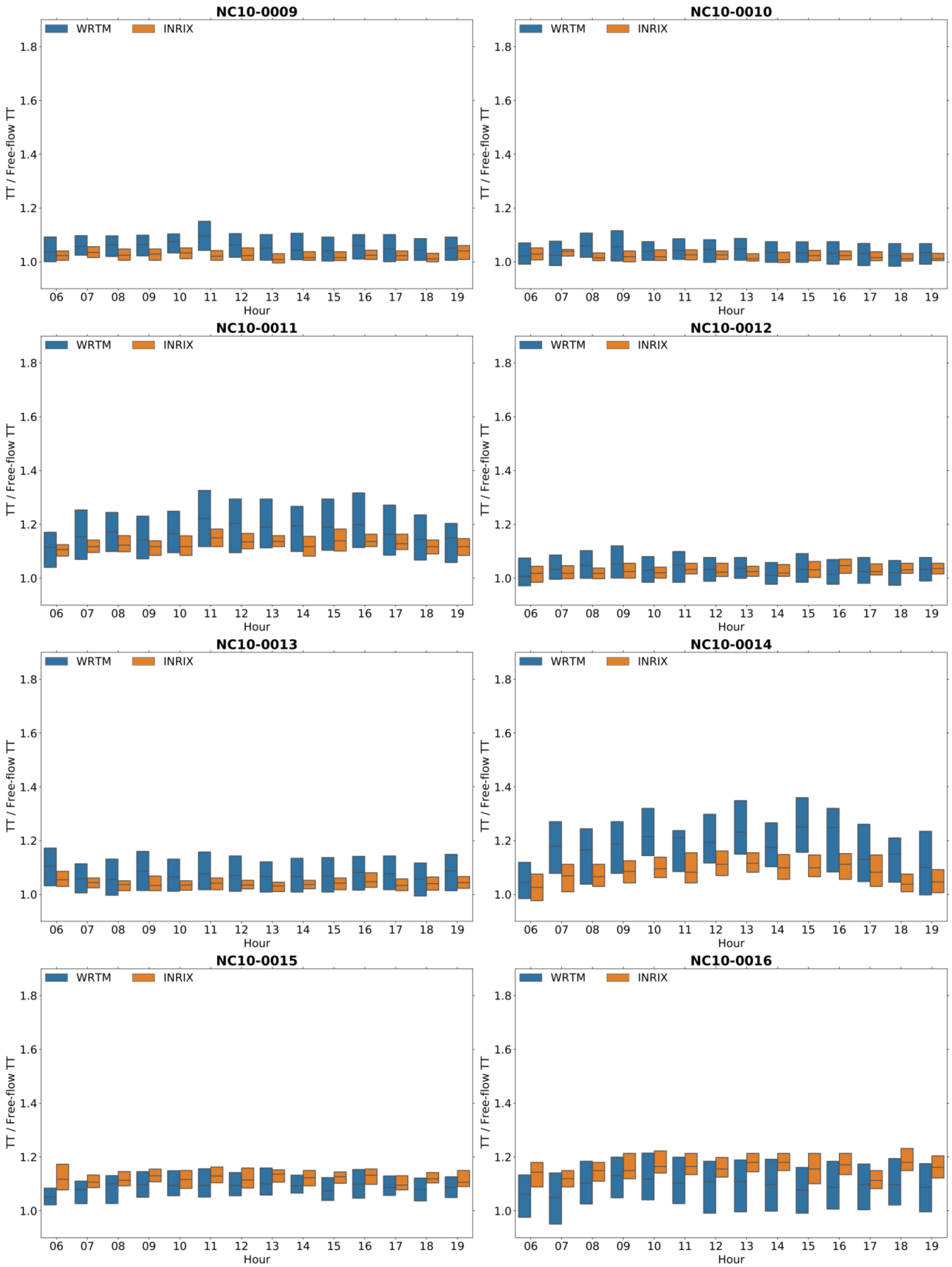


Figure 17 – INRIX path-level travel time boxplots for each time-of-day period (2/4)

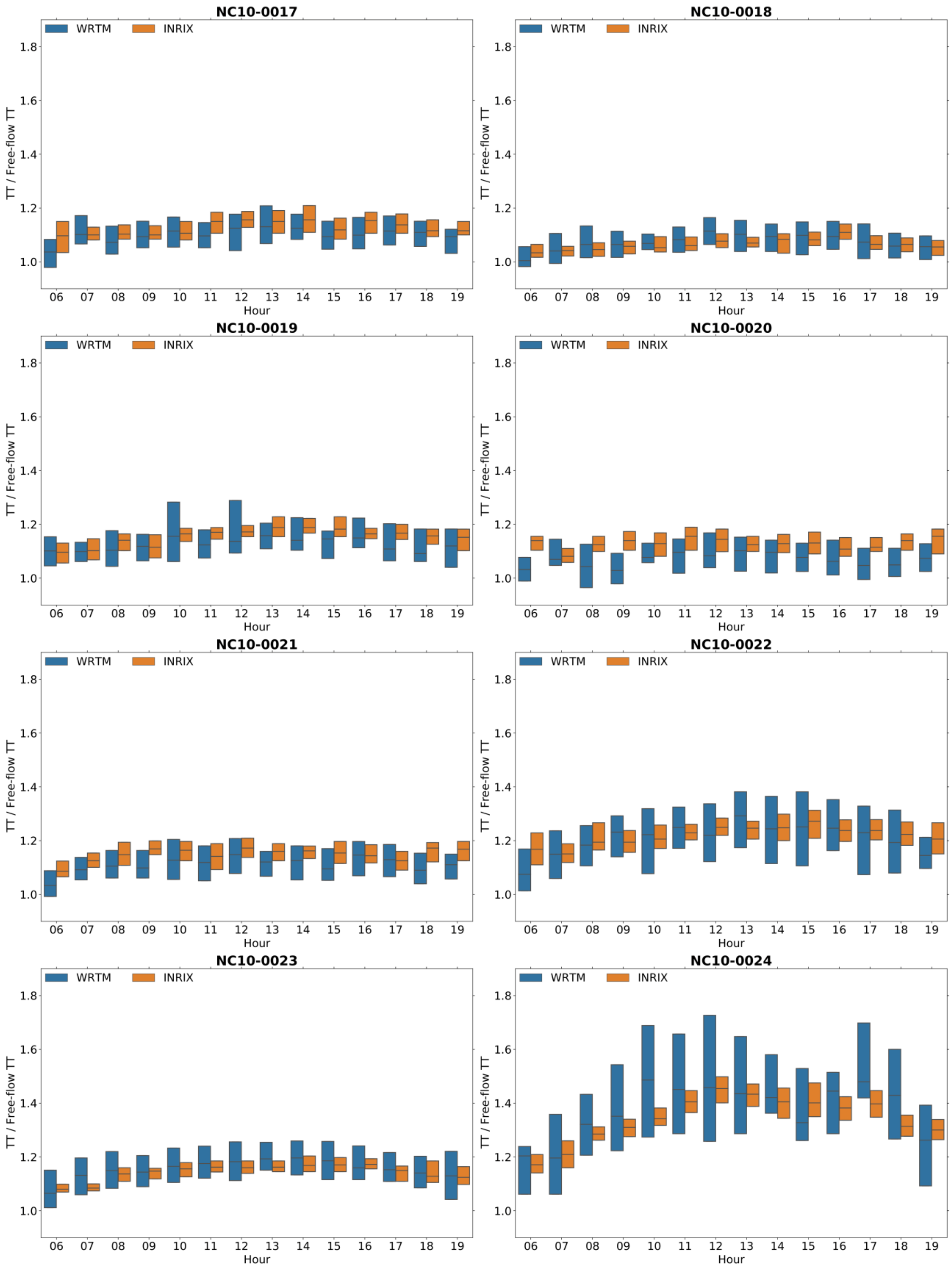


Figure 18 – INRIX path-level travel time boxplots for each time-of-day period (3/4)

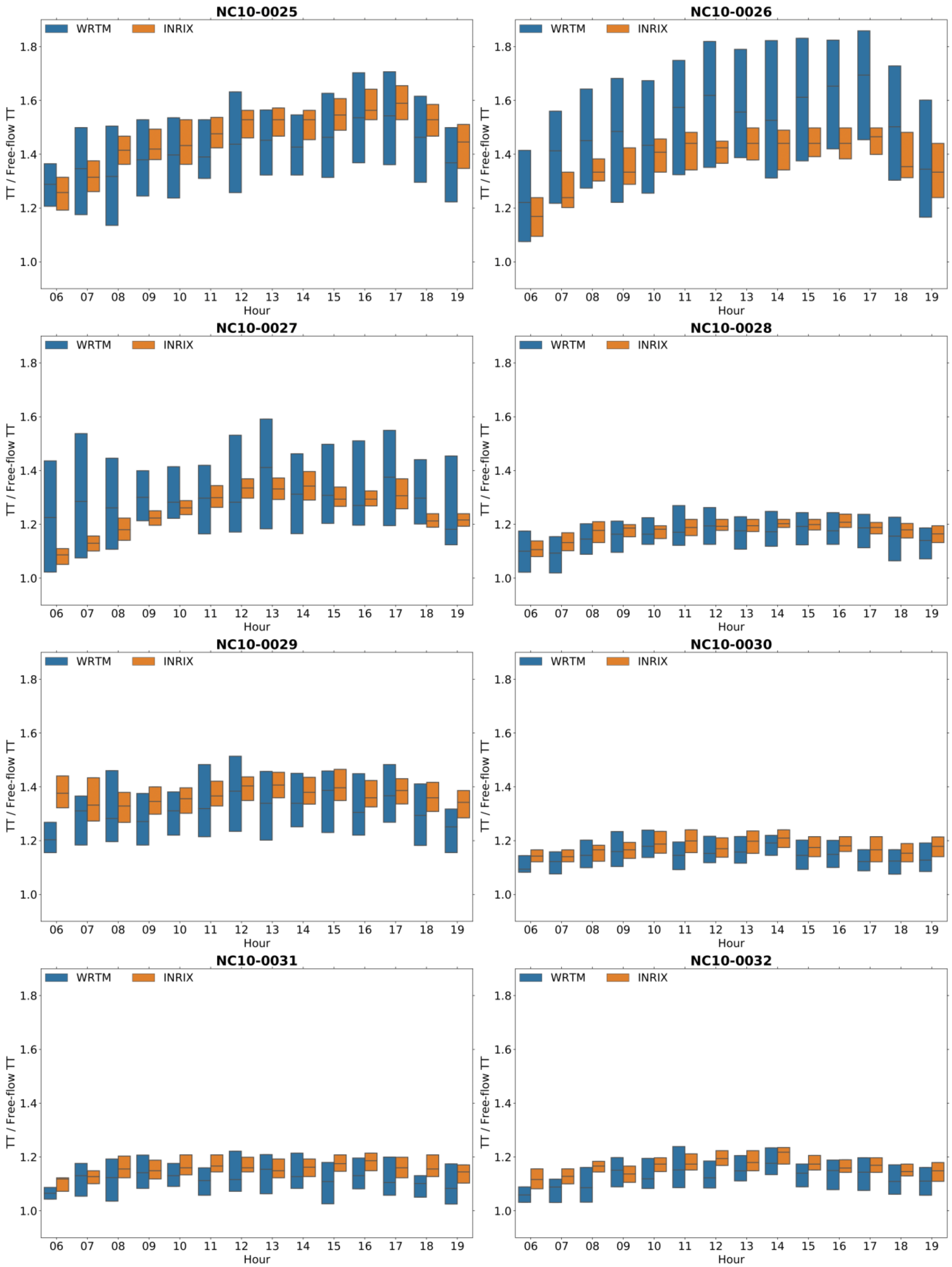


Figure 19 – INRIX path-level travel time boxplots for each time-of-day period (4/4)

TomTom

Figure 20 shows that **TomTom's median travel times agree with the reference data for most periods** and tend to be slightly lower than their WRTM counterparts (Δ Median is small – typically within 5% of free-flow travel time and has a slight negative bias). TomTom data shows a bit less variation than the reference data (Δ IQR is nearly always negative), with differences in IQR around 10% of the free-flow travel time.

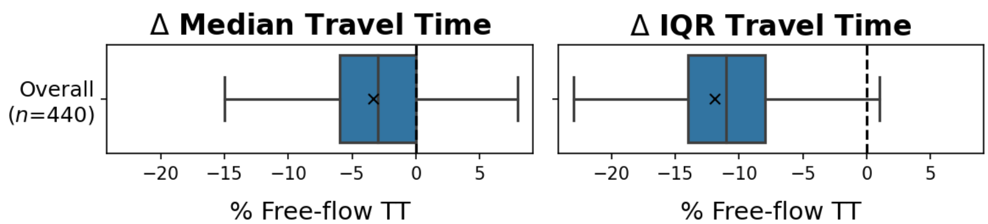


Figure 20 – TomTom summary of Δ Median and Δ IQR values

Figure 21 reports the Δ Median and Δ IQR values separately for three congestion levels, where each period is assigned to a congestion bin based on the ratio of the median travel time to free flow speed. The three congestion bins are (1.00-1.25], (1.25-1.50], and (1.50-1.75], with higher values representing higher typical travel times relative to free-flow conditions. The number of observations in each bin are not evenly distributed (i.e., the lowest bin has 375 of the 440 observations) because most hourly time-of-day periods are characterized by minimal congestion. For both Median and IQR comparisons, the TomTom probe data is closest to the reference source in the lowest bin and begins to underreport the Median/IQR as the congestion bin increases.

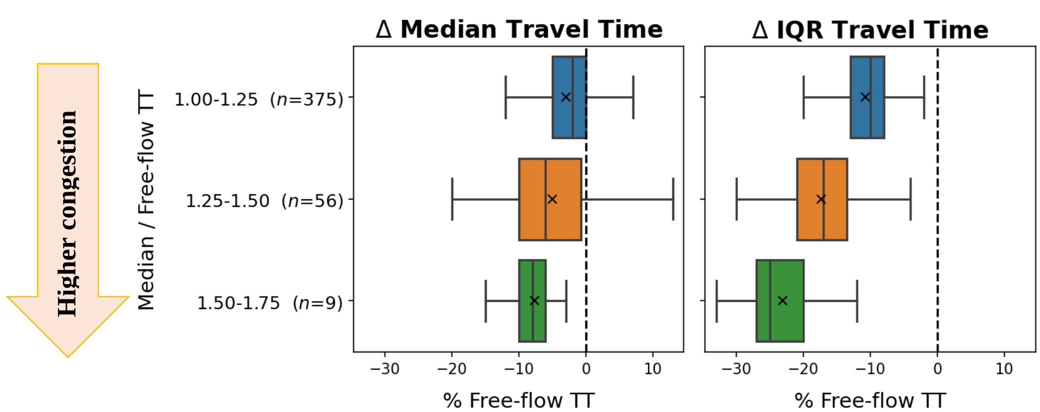


Figure 21 – TomTom Δ Median and Δ IQR values summarized by congestion bin

Figures 22-25, shown on the following two pages, contain boxplots summarizing the Vendor and WRTM travel time distributions for all validation paths and time-of-day periods. The summary figures on this page (Figures 20 and 21) were constructed by comparing the median and IQR values for each of the boxplots in Figures 22-25 and aggregating the results.

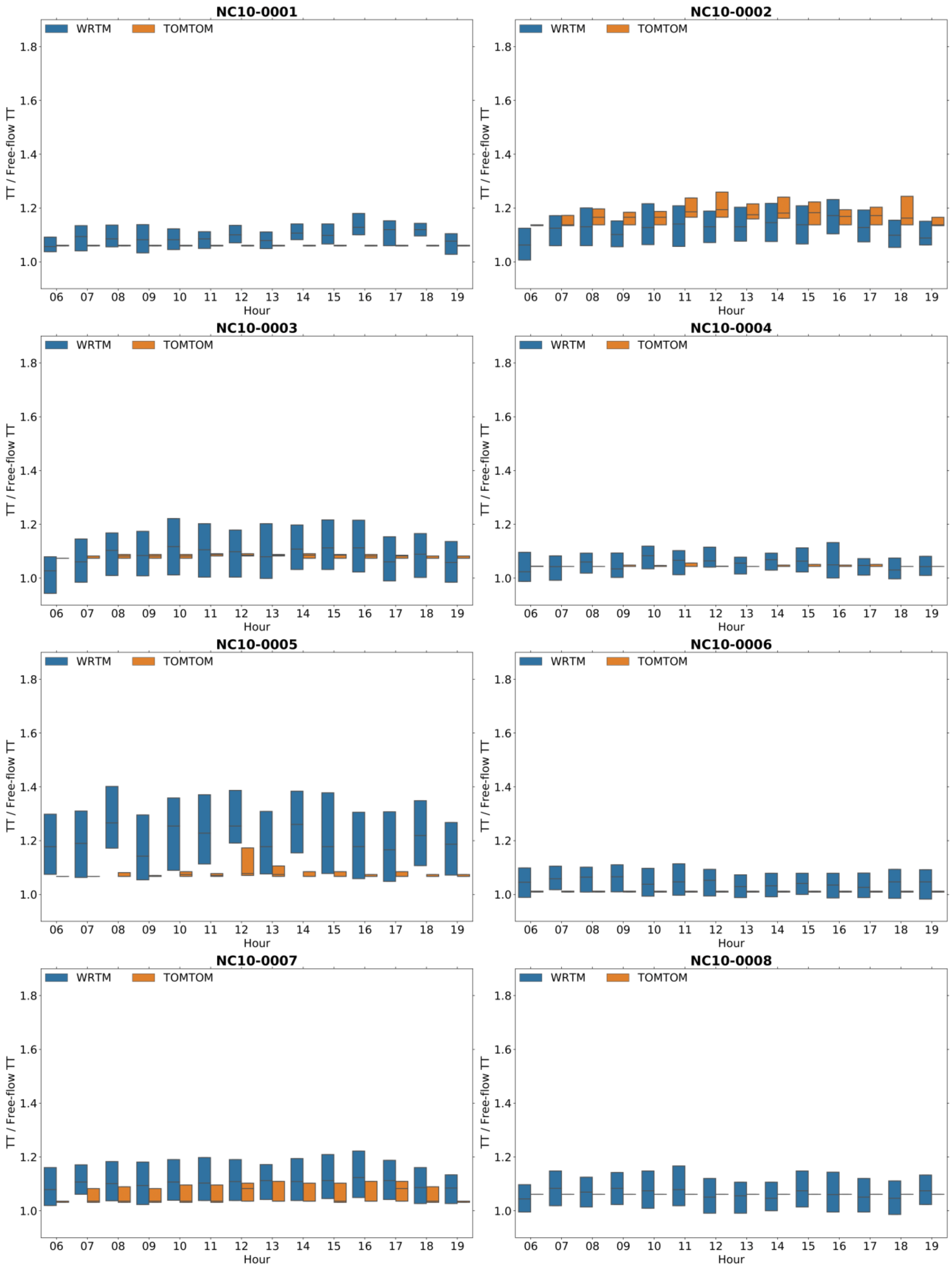


Figure 22 – TomTom path-level travel time boxplots for each time-of-day period (1/4)

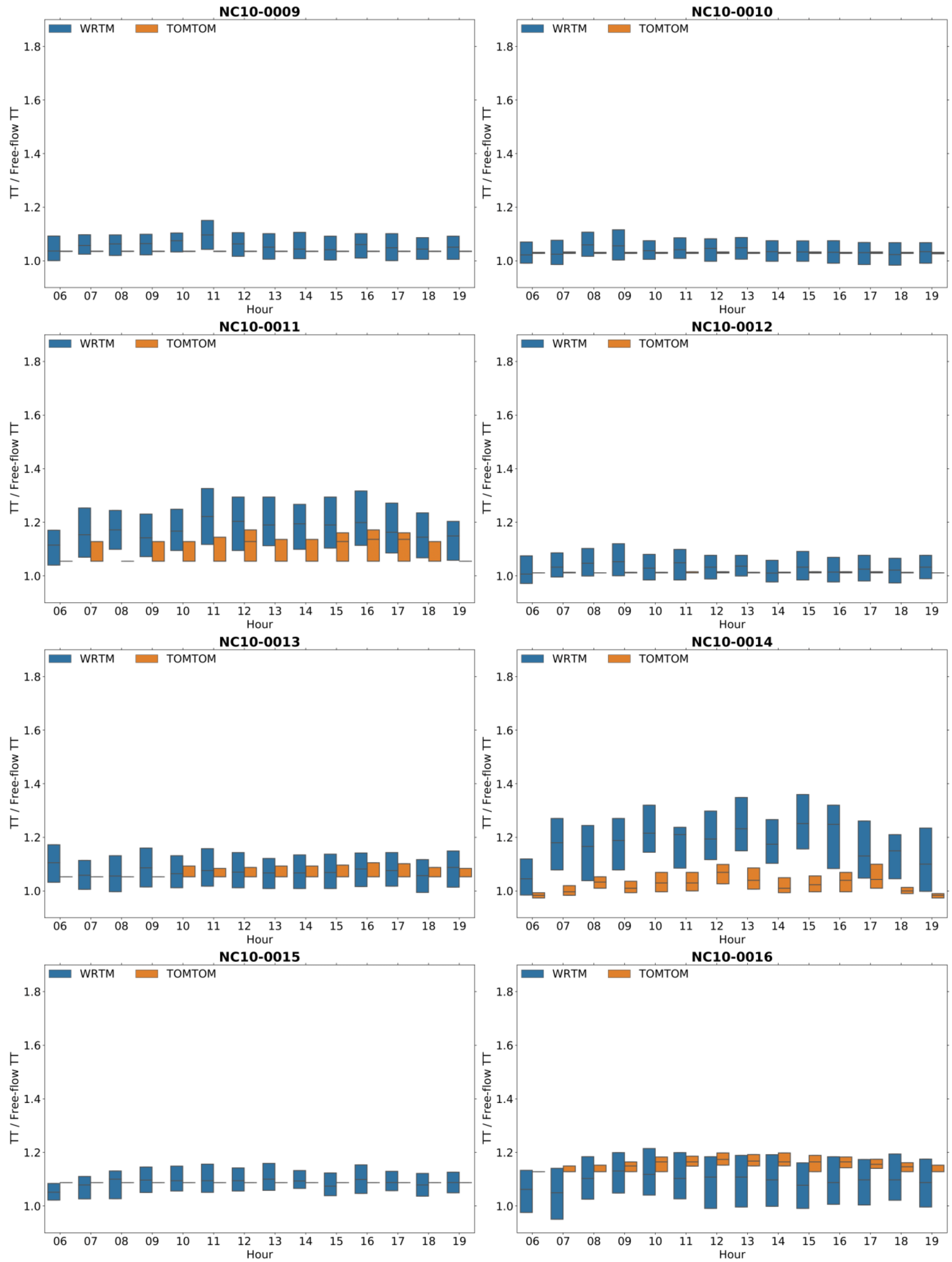


Figure 23 – TomTom path-level travel time boxplots for each time-of-day period (2/4)

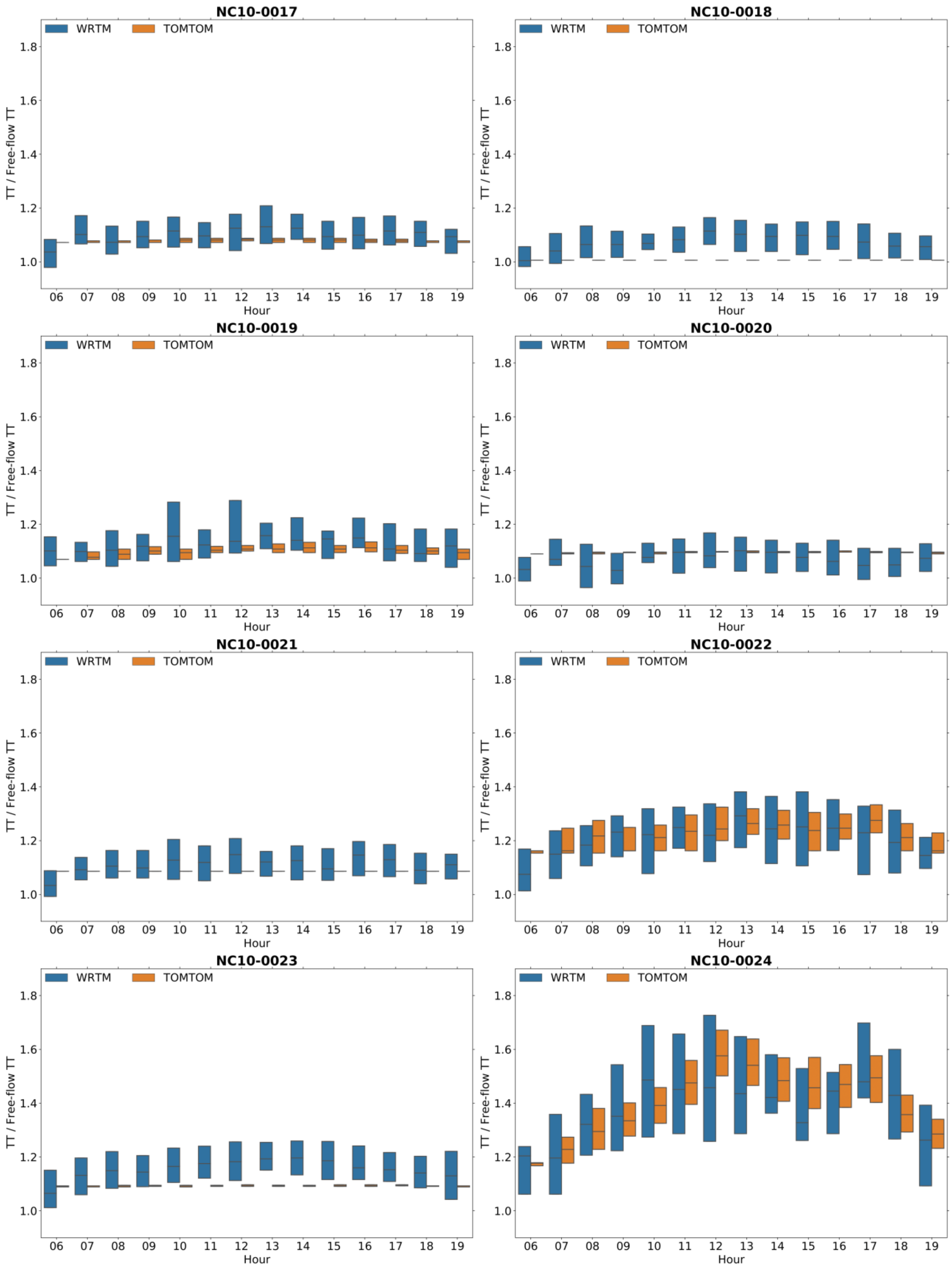


Figure 24 – TomTom path-level travel time boxplots for each time-of-day period (3/4)

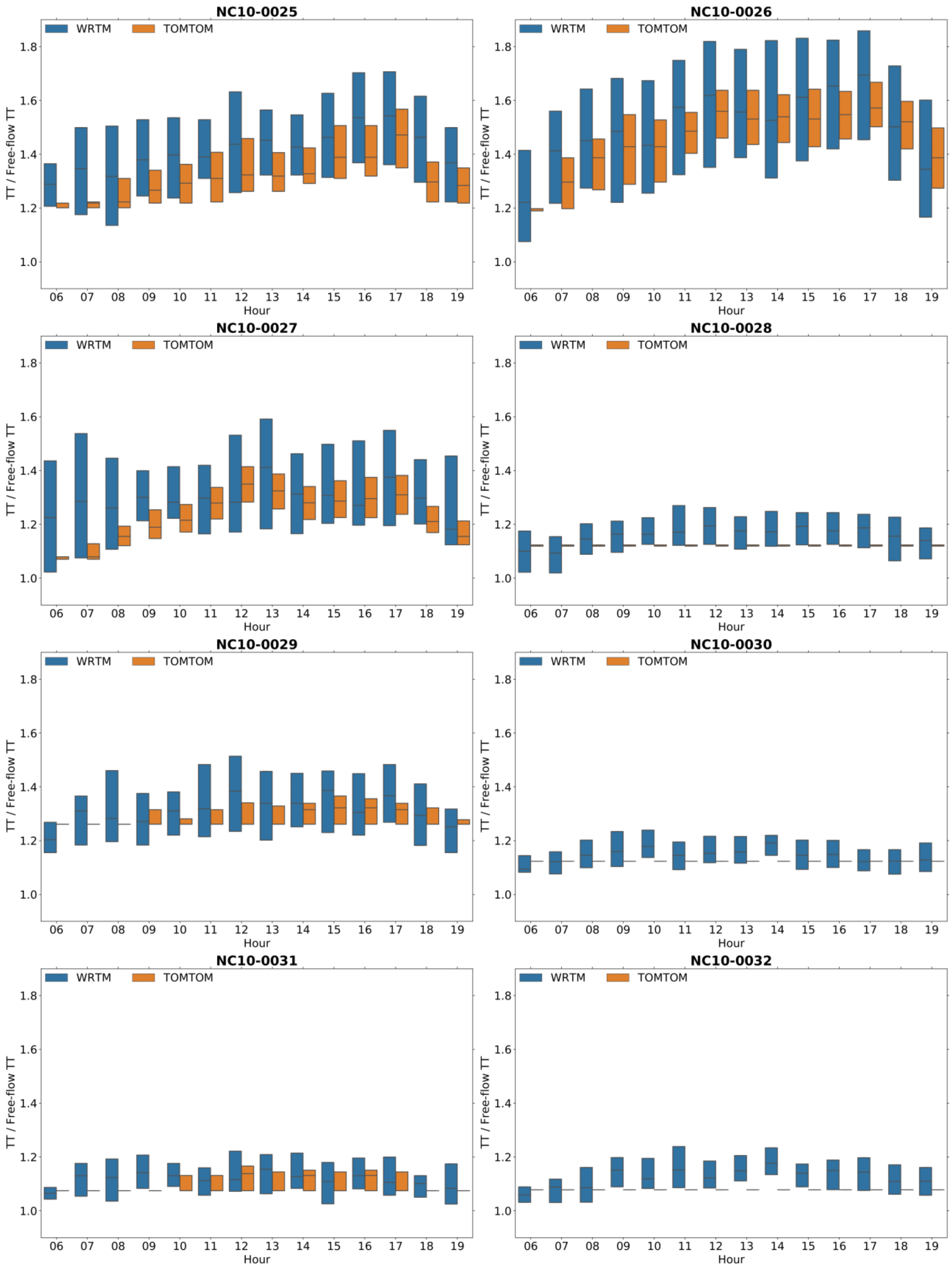


Figure 25 – TomTom path-level travel time boxplots for each time-of-day period (4/4)

Summary and Findings

This study reinforces the findings from recent low-volume arterial studies; probe vendors were consistently within contract specifications for the traditional analysis, but sparse data and lack of congestion made it difficult to evaluate the data quality in low-speed conditions. In this study, there were no opportunities to evaluate point-in-time accuracy in the 0-15 MPH speed bin, which typically is the most challenging for vendors to achieve. Likewise, while the sampled distribution analysis showed that vendors generally captured true time-of-day travel time patterns, the low-volume traffic conditions in this data collection did not provide meaningful opportunities to test data quality in challenging situations. As a result, **the validation team recommends significant changes in methodology for assessing data quality on low volume roadways** (discussed below).

In summary:

- This validation attempted to capture performance on rural, lower volume roads
- The test roads had low AADT (between 8000 and 9000) with sparse signal spacing
- No major congestion (as would be expected for this type of roadway) occurred during the test period
- The statistical validation results (both Traditional and Sampled Distribution) showed that vendors performed within contract specifications and reflected the normal patterns of traffic travel time experienced on the roadway.
- **No non-recurring congestion events (e.g., incident, inclement weather, special event) occurred during the data collection period to test vendors' ability to reflect non-recurring congestion on such roadways.**

Overall, this validation, as well as previous validations on lower volume roads, exposed that current validation methods, originally conceived for highly congested urban roadways, are inadequate to reflect the quality of probe data on lower volume roads. To remedy this, the validation team recommends the following:

- **Concentrate data validation on low-volume roads during periods of known traffic disruption**. Any validation activities should concentrate on fluctuations in expected speed and travel time. With low volume roads this is more aligned with weather events, road work, and/or incidents, and not predictable volume induced congestion. In contrast to a set two-week data collection, the validation should identify periods of time in which such fluctuations have occurred, and then compare how well vendors captured such fluctuations.

- **Compare travel time data between vendors, as well as against emerging connected vehicle data that can be acquired after the fact.** Because traffic disruptions are so infrequent in low-volume conditions, the current data collection approach (i.e., field collection via wireless re-identification sensors) would become prohibitively expensive, requiring sensor deployment for long periods of time until appropriate phenomena occur.

As such, once a traffic disruption event is identified, the Coalition will request data from the vendor for roadways that would have been impacted and (1) check relative agreement among the various vendors as an indication of consistency and accuracy, and (2) procure connected vehicle data from the same roadways and compare the available trace vehicle data with from travel time vendors as another source of reference comparison.

Such an approach would ensure meaningful feedback to the Coalition on the performance of probe data on low volume roadways.