

Executive Summary: Eastern Transport Coalition's Atlanta Travel Time Study and TomTom Data Validation

Transportation Data Marketplace (TDM) Travel Time / Speed Data Validation Data Collection December 2023 & Report Published September 2024

Transportation data sold through the Eastern Transportation Coalition (ETC) Transportation Data Marketplace (TDM) is procured from private industry based on contract specifications. The fidelity and accuracy of the data is assessed via a rigorous validation process across six data categories in various contexts by a technical advisory group that sets general direction and reviews results. The TDM includes both quantitative and qualitative analysis of datasets available through the marketplace as appropriate for each data type. This report evaluates one of the products sold in the TDM: Travel Time / Speed data.

Although Travel Time and Speed data has been well-studied, this validation activity was developed for two purposes: (1) to benchmark vendor accuracy after disruptions to data supply chains during 2022-2023 (after Wejo filed for bankruptcy); and (2) to measure data latency for the first time in the TDM.

Data Vendors

All vendors in the Travel Time and Speed category (Carto, HERE, INRIX, Iteris, Timmons) participated in the validation study and submitted data for the specified validation segments. It is important to note that the underlying data for the Carto and Timmons submissions was provided by TomTom. A key requirement for this study was that data be submitted in real-time to provide an audit trail for latency calculations. Just two vendors, Carto and Iteris, had certified real-time submissions suitable for latency analysis.

Latency Analysis

Results were based on 101 slowdown and recovery events distributed across 13 validation segments. These congestion events – all of which were about one hour or more in length – were manually identified using analyst judgement and the criteria described in the Report.

The latency results clearly distinguished these vendors. Carto's TomTom-powered data showed the lowest and most consistent latency across the entire study. Slowdown-period latency values were tightly concentrated between one and two minutes, with a mean of 1.1 minutes and a median of 1.0 minute. Ninety-five percent of congestion events were detected in less than three minutes, and at least seventy-five percent of values were below two minutes across slowdown, recovery and full-event measures. This pattern reflects strong waypoint density and an efficient (every 30 seconds) processing pipeline.

In contrast, the only other certified vendor showed a wider spread of latency values, with means near two and a half minutes and occasional outliers. The uncertified vendors had a combined average slowdown latency of 1.7 minutes, which remained higher than the certified Carto's TomTom-powered latency even with the advantage of post-hoc submission.

Average Absolute Speed Error (AASE) Analysis

All the vendors assessed were within specifications established by the Coalition, but Carto and Timmons' TomTom-powered data were the most accurate overall. Carto and Timmons consistently were top performers in congested bins, where accuracy matters most for

operations. In congested conditions (0–30 mph), Carto and Timmons outperformed the next-best vendor by up to 1.3 mph. Lower AASE in congested bins means better reliability for incident detection, traveler information and performance reporting. Accuracy during congestion is critical for agencies managing dynamic traffic conditions.

This study represents the first time the TDM program measured latency using certified real-time data. Overall, TomTom-powered data demonstrated the lowest latency and the highest level of consistency among all evaluated vendors. The tight distribution of values and low mean latency indicate a system that not only detects congestion rapidly but also does so predictably across a wide range of traffic scenarios, confirming the reliability and operational maturity of TomTom's travel time and speed dataset.

The Eastern Transportation Coalition report follows in its entirety.



TDM Validation Activity: Atlanta Travel Time Study

TETC Traffic Data Marketplace Data Validation

Prepared by: TETC Validation Team

9/27/2024

TDM Validation Activity: Atlanta Travel Time Study

TETC Traffic Data Marketplace Data Validation

Travel time validation activity on a congested urban freeway, including assessment of probe data latency.

TDM-VAL-7

Data Categories: Travel Time

Data Collection: December 2023

9/27/2024

The Eastern Transportation Coalition is a partnership of 19 states and the District of Columbia focused on connecting public agencies across modes of travel to increase safety and efficiency. Additional information on the Coalition, including other project reports, can be found on the Coalition's website: www.tetcoalition.org

Executive Summary

This report describes a Transportation Data Marketplace (TDM) data validation effort conducted by The Eastern Transportation Coalition (TETC) data validation team focusing on probe-based Travel Time/Speed data – one of the six core data items sold in the TDM. This activity was conducted in Atlanta, GA and targeted an urban freeway that experiences significant congestion.

Although Travel Time/Speed probe data has been well-studied in previous Coalition validation activities, this study was developed for two purposes: (i) to benchmark vendor accuracy after disruptions to data supply chains during 2022-2023, and (ii) to measure data latency for the first time in the TDM based on an established methodology. The latency analysis is notable because it marks the first time such measurements have been possible using “certified” real-time data submitted by vendors.

The study area consisted of 18 road segments along GA-400 / US-19 between I-85 and GA-120 (Old Milton Parkway) shown in Figure 1, of which 13 were deemed suitable for analysis (others had to be dropped to deployment challenges or data quality issues). This roadway was chosen with the expectation of capturing significant congestion to test vendor performance in dynamic conditions. Reference data was collected between Dec 4-16, 2023, using wireless re-identification traffic monitoring equipment (i.e., Bluetooth detectors) and corresponding travel time vendor data was provided by each of the five vendors. Each of the vendors submitted travel time data for evaluation, with two (Carto and Iteris) successfully submitting in real-time, as requested. The other three vendors submitted data after the fact (although INRIX and Timmons attempted real-time submissions before running into technical issues). Although the timestamps from submissions delivered after-the-fact could not be certified for latency, they reflect each vendor’s real-time product.

Vendor data submissions were evaluated using two methods: (1) the traditional analysis that calculates two error metrics, Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) for five-minute period, and (2) a latency analysis, which measures the lag time between when a congestion event takes place and when it is reflected in data feeds. The following points summarize the key observations:

- **All vendors captured recurrent and non-recurrent congestion patterns with a high degree of fidelity.** Each vendor was within specification for AASE and SEB metrics in all four speed bins, including the lower speed ranges that reflect congested periods and dynamic traffic conditions. Additionally, visual inspection of the trace data corroborated the high accuracy demonstrated by the AASE and SEB error measures, with vendors closely tracking fluctuations in speed observed in the reference data.

This result was not unexpected, as probe data has been high quality for many years on high volume freeways. Nonetheless, the findings are encouraging, as it provides an important benchmark after data supply chains were disrupted in 2022-2023, including the bankruptcy of Wejo (a data supplier to at least one travel time vendor) in June 2023.

- **A major milestone was achieved by successfully benchmarking latency with real-time data.** Previous validation activities laid the foundation for this outcome, which resulted in four of the five vendors submitting travel time records using the CATTWorks Geo-referencing Protocol (CWGP) in real-time. Two of the vendors, Carto and Iteris, were ultimately successful in this endeavor (two others attempted, but had to re-submit data after-the-fact), marking the first time the validation team has measured probe data latency based on “certified” real-time submissions.

- **Care should be taken when comparing latency values.** Latency values reported in this report are based on comparisons with reference re-identification data whose records are assigned to the *trip end time*. This means that the latency values reported here cannot be blindly compared with industry claims without understanding the methodology used – particularly with respect to temporal point of reference. Regardless of the absolute value of latency reported, calculating latency the same way will provide insight into changes over time.
- **Latency results are encouraging and have improved relative to VPPII.** When calculating latency at the onset of congestion (i.e., the slowdown portion of a slowdown and recovery event), the two vendors who delivered certified real-time data had average latencies of 1.1 and 2.4 minutes (with 95% of observations under 3.0 and 4.3 minutes, respectively). The other three vendors' combined average latency was 1.7 minutes. These results show an improvement of 1-2 minutes relative to VPPII, where average latency at the onset of congestion ranged from 3.4 to 7.1 minutes.

Table of Contents

- Executive Summary 3**
- Table of Contents 5**
- Introduction 6**
- Data Vendors 6**
- Study Area 7**
- Data Collection 8**
 - Reference Data..... 8
 - Vendor Data 9
- Evaluation Methodology 10**
 - Traditional Analysis..... 10
 - Latency Analysis..... 11
 - Implications of reference data timestamp..... 13
- Results and Discussion 15**
 - Traditional Analysis..... 15
 - Latency Analysis..... 19
 - Carto 19
 - Iteris 21
 - Comparison to VPPII..... 22
 - Interpretation Notes..... 22
- Conclusions..... 23**

Introduction

Transportation data sold through the Eastern Transportation Coalition (TETC) Transportation Data Marketplace (TDM) is procured from private industry based on contract specifications. The intent of the Coalition's validation program has evolved from a highly prescriptive methodology of the original Vehicle Probe Project (VPP) validation to one with more freedom to explore the fidelity and accuracy of the data feeds across six data categories in various contexts. The TDM now has the flexibility to adjust to the needs of the Coalition members as the market evolves and data needs expand. The validation process is overseen by a technical advisory group that sets general direction and reviews results. The TDM includes both quantitative and qualitative analysis of datasets available through the marketplace as appropriate for each data type. The marketplace currently contains six core data items: Travel Time/Speed, Volume, Waypoint, Origin-Destination, Freight, and Conflation, with all but one (Travel Time/Speed) being sold through the marketplace for the first time. As such, the validation team, under the guidance of TETC's TDM Leadership Group is beginning to establish standards and methods for effectively evaluating data quality and value across these six types of data sets.

This report evaluates one of the products sold in the TDM: Travel Time/Speed data, which was the core product from VPPII. Although Travel Time and Speed data has been well-studied, this validation activity was developed for two purposes: (1) to benchmark vendor accuracy after disruptions to data supply chains during 2022-2023, and (2) to measure data latency for the first time in the TDM. One of the major disruptions came in August 2023, when Wejo, a former vendor in the TDM and supplier of connected vehicle waypoint data, filed for bankruptcy. Wejo was supplying waypoint data (the key ingredient to probe-based speed and travel time data) to at least one TDM Travel Time/Speed vendor, so this validation presents an opportunity to assess the state of data quality leading into 2024.

As part of the validation study, reference Travel Time datasets were collected using Bluetooth/Wi-Fi-based re-identification sensors and used as the basis for evaluating reported vendor data. Vendor data was assessed through several techniques, including traditional error metrics, a visual analysis of time series plots, and latency analysis. The latency analysis is notable because it marks the first time such measurements have been possible using "certified" real-time data submitted by vendors.

Data Vendors

All vendors selected through the TDM RFP process in the Travel Time category (Carto, HERE, INRIX, Iteris, Timmons) participated in the validation study and submitted data for the specified validation segments. A key requirement for this study was that data be submitted in real-time to provide an audit trail for latency calculations. Four of the five Travel Time/Speed vendors (Carto, INRIX, Iteris, Timmons) submitted data in real-time as specified, while one (HERE) provided a post-hoc submission. However, INRIX and Timmons' submissions had data issues that needed to be resolved, so both vendors ended up resubmitting data after-the-fact. As such, just two vendors, Carto and Iteris, had certified real-time submissions suitable for latency analysis.

Study Area

Figure 1 shows the geographic scope of the study area within the Atlanta, GA metro region, consisting of 18 segments defined across both directions of GA-400 / US-19 from I-85 to GA-120. This roadway was chosen in coordination with Georgia DOT based on its high volume (about 128,000 AADT) and significant congestion (both recurring and non-recurring), which were needed to support a latency analysis.



Figure 1 – Validation segments (left) with sample imagery from Google StreetView (right)

Validation segments, shown in Figure 1, were defined between interchanges (or other logical break points) approximately 1-2 miles apart, with adjustments made based on deployment considerations (e.g., availability of mounting furniture for Bluetooth sensors, safety, minimizing detections from parallel facilities). 18 segments were included in the segment definitions, with 13 ultimately used for calculating error metrics and latency. Several segments were not included in the final analysis due to a deployment issue (one sensor was not able to be deployed) and data quality considerations (reference data that was corrupted and could not be trusted as the basis for comparison with probe vendors).

Table 1 shows the key attributes used to communicate the validation segment locations to vendors via CWGP. This information was shared with vendors as a GeoJSON file with MultiPoint coordinate geometry (representing the start and end points of each segment) so that it could be readily processed with a computer or displayed via web or Geographic Information System (GIS) tools. All segments are shown here for completeness; however, segments with an asterisk were not ultimately included in the analysis.

Table 1 - CWGP attributes for travel time validation segments

ID	Start Location	Start Heading	End Location	End Heading	Road Name	Road Class	Length (Miles)
P01	-84.360051 33.831055	11	-84.367928 33.848552	330	GA-400	2	1.397
P02	-84.367928 33.848552	330	-84.362945 33.896563	30	GA-400	2	3.413
P03	-84.362945 33.896563	30	-84.357763 33.91815	0	GA-400	2	1.564
P04	-84.357763 33.91815	0	-84.357906 33.942061	357	GA-400	2	1.648
P05	-84.357906 33.942061	357	-84.34118 33.983236	26	GA-400	2	3.047
P06	-84.34118 33.983236	26	-84.326624 34.01845	36	GA-400	2	2.605
P07*	-84.326624 34.01845	36	-84.314919 34.034529	36	GA-400	2	1.296
P08*	-84.314919 34.034529	36	-84.292123 34.05245	74	GA-400	2	1.844
P09	-84.292123 34.05245	74	-84.273849 34.06472	27	GA-400	2	1.418
P10*	-84.27398 34.064952	208	-84.29154 34.052807	254	GA-400	2	1.38
P11	-84.29154 34.052807	254	-84.315121 34.034579	215	GA-400	2	1.895
P12	-84.315121 34.034579	215	-84.327261 34.017887	216	GA-400	2	1.345
P13	-84.327261 34.017887	216	-84.342032 33.981984	207	GA-400	2	2.653
P14	-84.342032 33.981984	207	-84.358124 33.942816	178	GA-400	2	2.901
P15*	-84.358124 33.942816	178	-84.357956 33.917187	178	GA-400	2	1.767
P16*	-84.357956 33.917187	178	-84.363243 33.896686	210	GA-400	2	1.49
P17	-84.363243 33.896686	210	-84.368232 33.848277	150	GA-400	2	3.441
P18	-84.368232 33.848277	150	-84.360384 33.830594	189	GA-400	2	1.411

Data Collection

Reference Data

Reference travel time samples were collected along the validation segments from December 4 to December 16, 2023. Bluetooth re-identification sensors were deployed by a contractor that specializes in field data collection, and the resulting raw data was processed by the Coalition validation team to obtain travel time measurements.

At a high level, the processing approach involves: (1) preprocessing/consolidating raw Bluetooth detections at each sensor/station location, (b) matching detections between origin and destination locations (i.e., the endpoints of each segment), and (c) filtering resulting travel time observations to identify outliers. A custom software library was developed to support these analysis routines.

Figure 2 shows example travel time patterns along a representative segment (P11), which shows the fluctuations in speed that regularly occur on the roadway over the course of the study period (upper plot). On this segment – and many others along the corridor, there are periods of congestion during weekday morning and evening peak periods, with relatively stable speed/travel time values observed during off-peak periods. The lower plot shows estimated traffic flow over the study period, under the rough assumption of a 5% penetration rate. The observed travel time patterns are reinforced by the overlay plot shown in Figure 3, which plots travel time by hour-of-day for all weekdays in the study period to identify recurrent congestion patterns. These clear congestion events provide ample opportunity to assess vendor data latency.

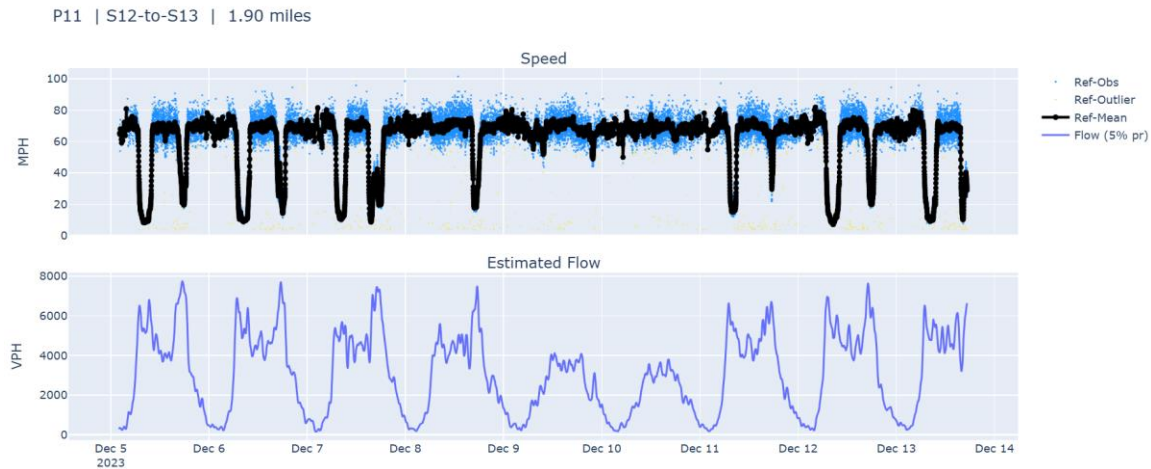


Figure 2 – Representative time series plot showing recurring congestion events

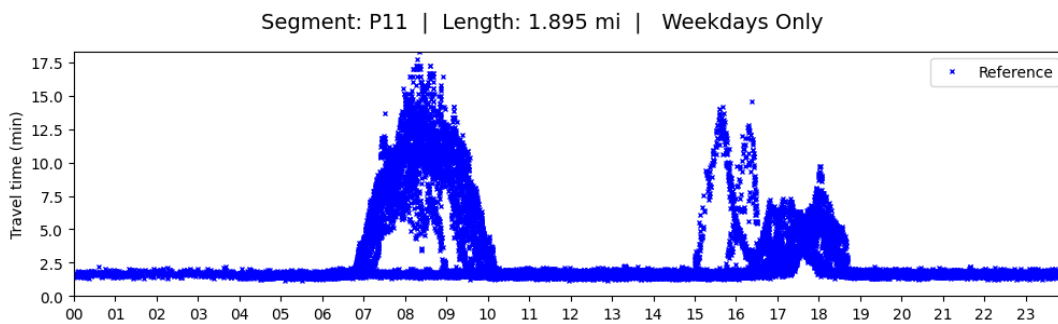


Figure 3 – Overlay plot highlighting weekday AM and PM peak period congestion.

Vendor Data

Vendors were instructed to submit travel time/speed data for all validation segments (as defined by the validation team using CWGP, not TMCs or other native segments) in real-time, which required delivering records once per minute to a SFTP server managed by the validation team. The validation team provided a reference implementation of how to submit data in real-time using the Python programming language, and each vendor was given its own server credentials for secure access to the server.

Four of the five Travel Time/Speed vendors (Carto, INRIX, Iteris, Timmons) attempted to submit data in real-time as described above, with two (Carto and Iteris) ultimately successful and the other two (INRIX and Timmons) encountering problems that resulted in needed to re-submit after-the-fact. The other vendor, HERE provided a post-hoc submission.

Figure 4 shows a sample file that is representative of what vendors delivered to the validation server each minute. Each file, delivered via SFTP and following the JSON Lines (i.e., newline-delimited JSON) format, contains one record per validation segment, with fields for location, timestamp, speed, travel time, and corresponding status flag. Note that the timestamp field in the file is populated by the vendor. However, another timestamp value was added by the validation team, which corresponds to when the file was received on the server. This server timestamp was ultimately used for analysis purposes for vendors that submitted in real time.

```
1 {"loc_id": "P01", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 65.2, "tt_minutes": 1.29, "status_flag": 30}
2 {"loc_id": "P02", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 68.4, "tt_minutes": 3.0, "status_flag": 30}
3 {"loc_id": "P03", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 62.1, "tt_minutes": 1.51, "status_flag": 30}
4 {"loc_id": "P04", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 67.7, "tt_minutes": 1.46, "status_flag": 30}
5 {"loc_id": "P05", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 69.6, "tt_minutes": 2.63, "status_flag": 30}
6 {"loc_id": "P06", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 70.7, "tt_minutes": 2.21, "status_flag": 30}
7 {"loc_id": "P07", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 70.8, "tt_minutes": 1.1, "status_flag": 30}
8 {"loc_id": "P08", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 70.1, "tt_minutes": 1.58, "status_flag": 30}
9 {"loc_id": "P09", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 69.0, "tt_minutes": 1.23, "status_flag": 30}
10 {"loc_id": "P10", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 67.7, "tt_minutes": 1.22, "status_flag": 30}
11 {"loc_id": "P11", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 70.8, "tt_minutes": 1.6, "status_flag": 30}
12 {"loc_id": "P12", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 69.6, "tt_minutes": 1.16, "status_flag": 30}
13 {"loc_id": "P13", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 69.0, "tt_minutes": 2.31, "status_flag": 30}
14 {"loc_id": "P14", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 67.4, "tt_minutes": 2.58, "status_flag": 30}
15 {"loc_id": "P15", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 67.1, "tt_minutes": 1.58, "status_flag": 30}
16 {"loc_id": "P16", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 67.1, "tt_minutes": 1.32, "status_flag": 30}
17 {"loc_id": "P17", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 67.7, "tt_minutes": 3.05, "status_flag": 30}
18 {"loc_id": "P18", "timestamp": "2023-12-16 08:00:55-0500", "speed_mph": 67.7, "tt_minutes": 1.25, "status_flag": 30}
```

Figure 4: Sample vendor file submission

Using the server timestamp associated with data delivery certifies that the corresponding travel time and speed data were produced as a real time product. To ensure that the server timestamps were valid, the submission server's clock was compared to the global Network Time Protocol (NTP) server and found to match within less than a second (0.0008 to 0.019 seconds, tested on two different dates).

Evaluation Methodology

Vendor data submissions were evaluated using two methods: (1) the traditional analysis that calculates two error metrics, Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) for five-minute periods, and (2) a latency analysis, which measures the lag time between when a congestion event takes place and when it is reflected in data feeds. Vendor data was also assessed by visually inspecting speed and travel time plots for each segment in comparison to the reference data, focusing in particular on periods where congestion was observed.

Traditional Analysis

The traditional validation analysis consists of comparing reference speeds against vendor speeds over five-minute intervals and quantifying the discrepancy in terms of two error metrics defined in the contract specifications. Prior to computing error metrics, both reference (Bluetooth) and vendor data must be aggregated to five-minute bins.

For reference data, each Bluetooth travel time observation is assigned to a 5-minute time bin based on the trip end timestamp (e.g., a trip ending on 12/14/2023 at 8:03:14 AM would be

assigned to the time bin 12/15/2023 8:00-8:05 AM). Each 5-minute bin is summarized in terms of the average (space-mean) speed, upper and lower confidence bands around the average speed (Standard Error of the Mean (SEM) band), number of observations, and coefficient of variation. Periods with insufficient Bluetooth observations or too much variation in speeds are not used in the analysis. Similarly, vendor data is aggregated to 5-minute bins by averaging the 1-minute records submitted in the delivery.

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 the space-mean speed value for the reference data (0-30 mph, 30-45 mph, 45-60 mph, 60+ mph for freeways). Reported probe speeds are compared to both the space-mean and confidence band ground truth speeds 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 the absolute value of SEB should be less than 5mph and 4mph, respectively, on freeways.

AASE is calculated by summing up the absolute difference between probe vendor speeds (S_V) and reference speeds (S_R) for each time interval and taking the average over n observations. That is, $AASE = \frac{1}{n} \sum_{i=1}^n |S_{V_i} - S_{R_i}|$. 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, $SEB = \frac{1}{n} \sum_{i=1}^n S_{V_i} - S_{R_i}$. 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.

Latency Analysis

Latency within the TDM is defined as the difference between the time the traffic flow is perturbed – often as a result of an incident, and the time that the change in speed is reported in the probe data. Figure 5 shows this visually, with the black curve representing reference data and the orange curve showing the vendor data that lags the reference curve.

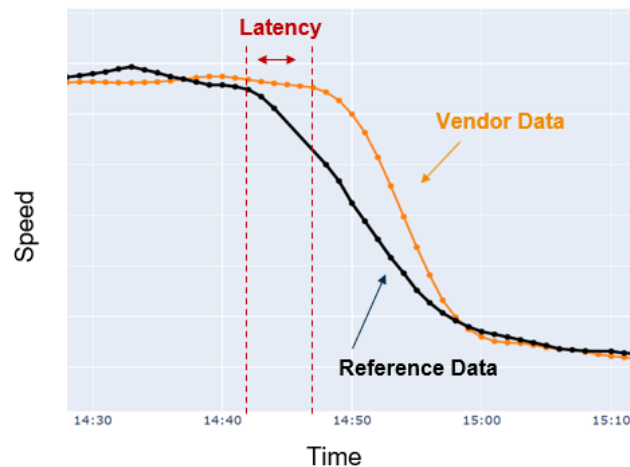


Figure 5 – Illustration of latency

Figure 6 summarizes the elements of probe data latency, starting with vehicle waypoints being recorded (far left) and ending with a user receiving probe data (far right), with various delays introduced during different steps in the process. For example, there is communication delay between when an on-board GPS unit records a vehicle location and when it is transmitted to the vendor, processing delay for the vendor to aggregate waypoints and generate the probe data product, conflation delay to translate the native product to different road segment definitions, and further communication delay to expose the data via API or transmit to the customer. Our validation is measuring the entire delay from field event to showing up in the vendor probe data. However, **the overall latency is dominated by the processing delay**, which is related to data density and is typically measured in minutes, whereas the other delays are comparatively minor and measured in terms of seconds.

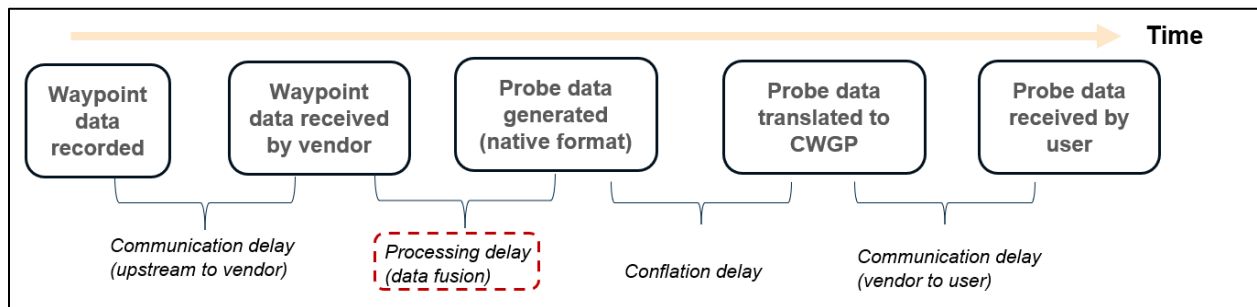


Figure 6 – Elements of probe data latency.

At a high level, the methodology used to calculate latency is based on the following steps, based on prior Coalition research and publications ^{1,2}

- **Pre-process reference and vendor data to one-minute average speed values.** Both datasets are cleaned by interpolating isolated missing values and applying a simple smoothing filter. This minimizes the impact of noise and other minor fluctuations in speed and helps facilitate the comparison of major trends in the curves. The details can be found in Wang, et. al (2017) ¹
- **Identify major slowdown events in the reference dataset.** Candidate criteria for slowdowns is a decrease in traffic speed of 20 mph or 40 percent of free-flow speed, with duration of at least an hour, although this may be relaxed in certain conditions.

For each slowdown used in the analysis, the start time, end time, and transition time are recorded, with the transition time representing the time associated with the minimum speed value during the event. Figure 7 shows how these times are used to define different parts of a slowdown – each of which can be measured separately.

- *Slowdown Period (SP):* The slowdown portion of a congestion event, measured from the onset of congestion until the transition time (time associated with minimum speed during the slowdown).
- *Recovery Period (RP):* The portion of the congestion event measured from the transition time to the end of the congestion event.

¹ A Methodology for Calculating the Latency of GPS Probe Data ([link](#))

² A Cross Vendor and Cross State Analysis of GPS Probe Data Latency ([link](#))

- *Slowdown and Recovery Event (SRE)*. The entirety of a congestion event, measured from the onset of congestion until the end of the recovery.

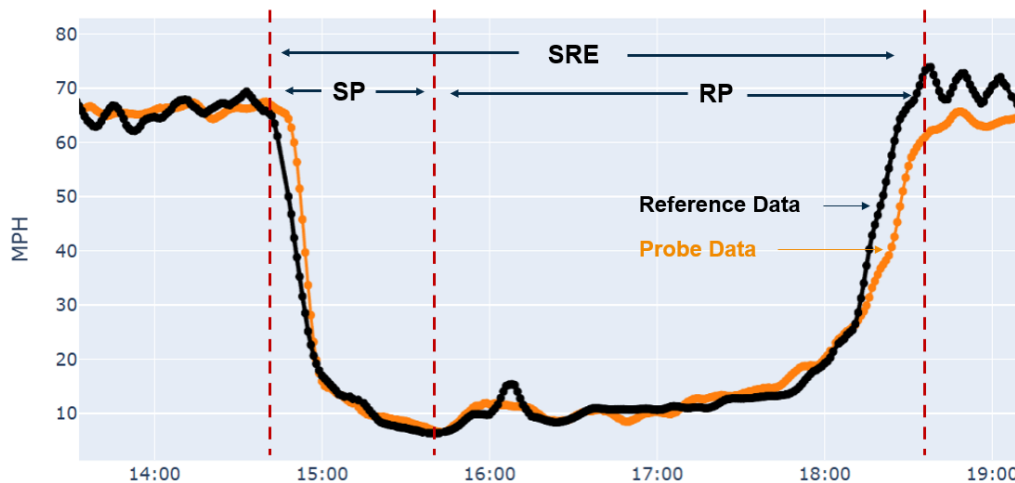


Figure 7 – Parts of a slowdown (SP, RP, SRE)

- **For each slowdown event, correlate the vendor data to the reference data with a variable time offset.** The time offset that maximizes the correlation during the slowdown between the reference and GPS probe data determines the latency for the slowdown event. Note that there are different ways to measure correlation between two time series curves. As such, we separately calculate the latency (offset) that optimizes three different fitness functions (Absolute Vertical Distance, Squared Vertical Distance, Pearson Correlation), and then average the values to obtain an overall latency for each event.

This process is repeated three times for each slowdown event to separately measure latency for the Slowdown Period (SP), Recovery Period (RP), and entire Slowdown and Recovery Event (SRE). ***SP is particularly important, and usually the value we care about most when assessing latency.***

- **Summarize latency values across the entire dataset.** Calculate the average latency and summarize the distribution of values across all slowdown events.

Implications of reference data timestamp

By definition, travel time records collected from Bluetooth sensors (i.e., reference data) are defined between a pair of sensors. Vehicles are first detected when passing an upstream sensor, and then later re-identified at a downstream sensor, with the difference between timestamps representing the travel time. However, each record must ultimately be associated with a single timestamp so that data can be aggregated and compared to vendor data during different time periods. This timestamp could theoretically be the trip start time, the end time, or some value in between; however, **the TDM (and previously VPP) validation process has always associated each record with the trip end time.** This approach is intuitive, as timestamps reflect *realized* travel time conditions at a point in time, rather than travel times that cannot be known until later when a trip is completed.

However, note that the choice of *end time* for the reference data timestamp can have implications on the latency values that are produced. Using the end time is the most generous to vendors, as it allows them more time to respond to events. For example, at the onset of congestion, slower speeds (higher travel times) will not show up in the reference data until the vehicles exit the segment, giving vendors more time to reflect the speed drop in their feeds. If start time or an intermediate time were used, the reference data curve would shift back in time and congestion would be reflected earlier. Figure 8 illustrates this concept, with the left plot showing reference data using the trip end timestamp, and the right plot showing the same data, but with reference data using the trip start timestamp. The vendor data (orange curve) is the same in both figures, but the reference data shifts to the left when the start timestamp is used, resulting in a greater latency value.

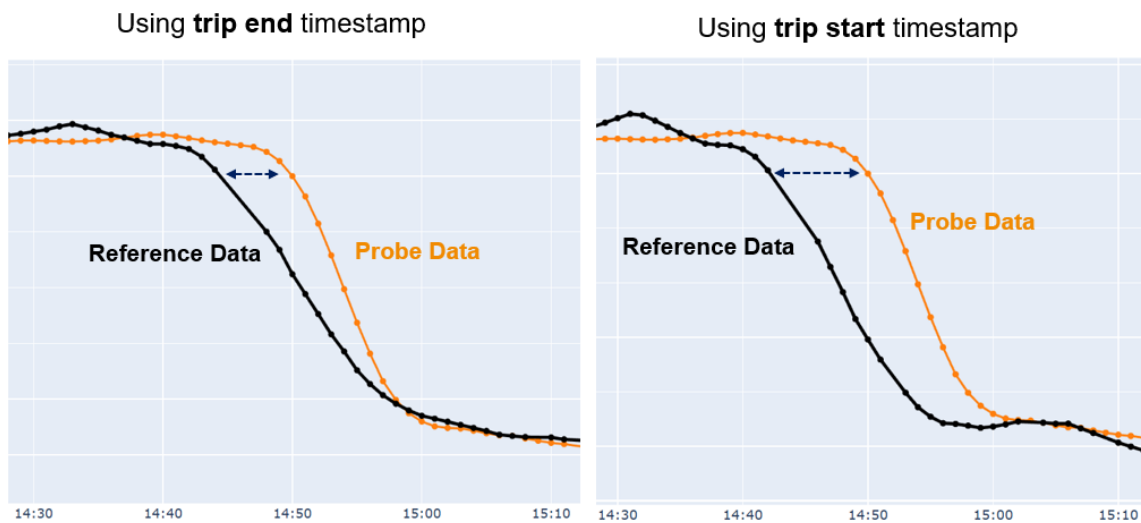


Figure 8 – Impact of using the start versus end timestamp for Reference data

As a result, **it is possible to have negative latency values**, where vendors detect the onset of congestion before it is reflected in the reference data. Although rare, this is an artifact of how reference data timestamps are defined. While other ways of defining the timestamp could be valid and may have advantages in cases such as this, *the end time will continue to be used so that results are consistent with and comparable to findings from the past 15+ years of validation.* If the latency were based on the time the vehicle enters the upstream station, the values for latency would increase by an amount corresponding to the segment travel time.

Results and Discussion

The following subsections report results for traditional and latency analyses separately. All vendors are included in the traditional analysis results, with latency analysis focusing on the two vendors who successfully submitted data in real time (Carto and Iteris).

Traditional Analysis

Tables 2-6 summarize the error metrics computed between reference and vendor speeds for all five vendors, while Figures 8-13 provide a representative visual example for each vendor on the same segment during a day with significant congestion.

For all vendors, Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) measures are within specification for all speed bins, as measured comparing to the 1.96 SEM band of the reference mean speed. The lower speed ranges are particularly accurate for all vendors, reflecting strong performance during periods of congestion. Slightly higher values are observed in the 45-60 range for most vendors; however, this is well within spec – particularly when looking at the 1.96 SEM band, which is relevant for higher speed ranges when more variation is observed in the traffic stream. Overall, these tables communicate that on average, all vendor products reported speeds that closely matched the observed Reference speeds.

Table 2: Carto error metrics for traditional analysis

Speed Bin	Average Absolute Speed Error (AASE)		Speed Error Bias (SEB)		Number of 5 Minute Samples
	Mean	1.96 SEM Band	Mean	1.96 SEM Band	
0-30	2.1	0.9	1.2	0.6	2323
30-45	2.6	0.9	1.2	0.5	1867
45-60	4.5	1.7	3.7	1.5	3759
60+	2.7	0.5	0.9	0.2	20107
All speeds	2.9	0.7	1.3	0.5	28056

Table 3: HERE error metrics for traditional analysis

Speed Bin	Average Absolute Speed Error (AASE)		Speed Error Bias (SEB)		Number of 5 Minute Samples
	Mean	1.96 SEM Band	Mean	1.96 SEM Band	
0-30	2.4	1.2	1.8	1.0	2323
30-45	3.0	1.3	2.1	1.0	1867
45-60	4.6	1.8	3.5	1.4	3761
60+	2.9	0.7	1.1	0.3	20160
All speeds	3.1	0.9	1.6	0.5	28111

Table 4: INRIX error metrics for traditional analysis

Speed Bin	Average Absolute Speed Error (AASE)		Speed Error Bias (SEB)		Number of 5 Minute Samples
	Mean	1.96 SEM Band	Mean	1.96 SEM Band	
0-30	2.4	1.1	0.9	0.6	2323
30-45	3.0	1.2	-0.2	-0.1	1867
45-60	4.1	1.4	2.3	0.8	3761
60+	2.7	0.5	0.0	-0.1	20051
All speeds	2.9	0.8	0.3	0.1	28002

Table 5: Iteris error metrics for traditional analysis

Speed Bin	Average Absolute Speed Error (AASE)		Speed Error Bias (SEB)		Number of 5 Minute Samples
	Mean	1.96 SEM Band	Mean	1.96 SEM Band	
0-30	3.4	2.0	-0.7	-0.4	2300
30-45	3.7	1.9	-0.6	-0.4	1856
45-60	4.4	1.8	0.7	0.1	3753
60+	3.4	1.0	-0.8	-0.4	20039
All speeds	3.5	1.2	-0.6	-0.3	27948

Table 6: Timmons error metrics for traditional analysis

Speed Bin	Average Absolute Speed Error (AASE)		Speed Error Bias (SEB)		Number of 5 Minute Samples
	Mean	1.96 SEM Band	Mean	1.96 SEM Band	
0-30	2.1	0.9	1.2	0.6	2323
30-45	2.7	1.0	1.3	0.6	1867
45-60	4.6	1.7	3.7	1.5	3761
60+	2.7	0.5	0.9	0.2	20111
All speeds	2.9	0.8	1.3	0.5	28062

Figures 9-13 serve as a visual example to complement the summary statistics above. These plots focus on segment P18, a segment that sees regular and sustained congestion during weekdays – particularly during AM and PM peak periods. On December 6 (Wednesday), traffic speeds averaged approximately 65-70 mph in the early morning, but dropped significantly between 7-9am – reaching as low as 20-30 mph. During mid-day, speeds recovered to about 50 mph, before experiencing major PM peak congestion from just after 3pm to about 7:30pm. The afternoon congestion pattern consisted of multiple slowdowns and partial recoveries, providing ample opportunities to test vendor performance. These figures show that all vendors captured these congestion patterns well – a result that was expected (but not guaranteed) given the high overall accuracy.

P18 | S19-to-S20 | 1.41 miles

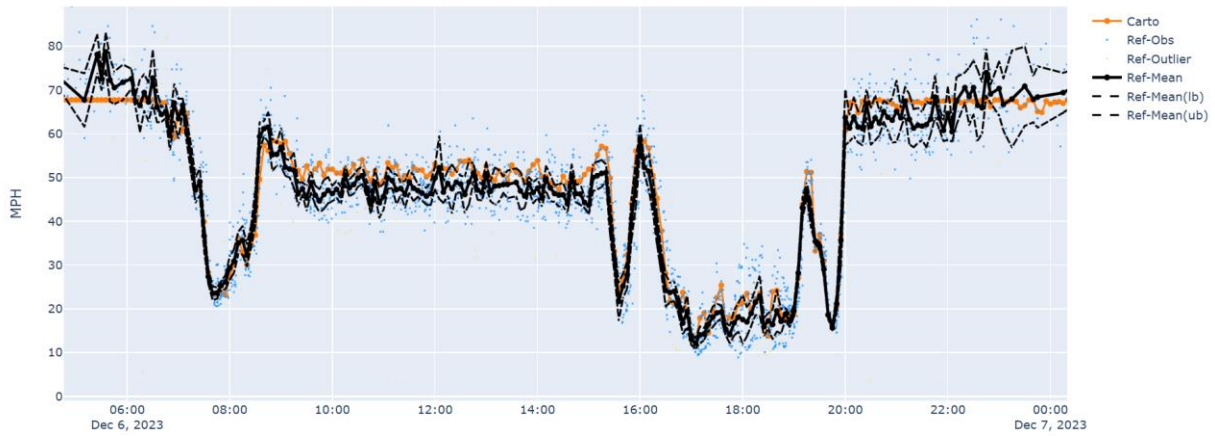


Figure 9 – Representative example from Carto

P18 | S19-to-S20 | 1.41 miles

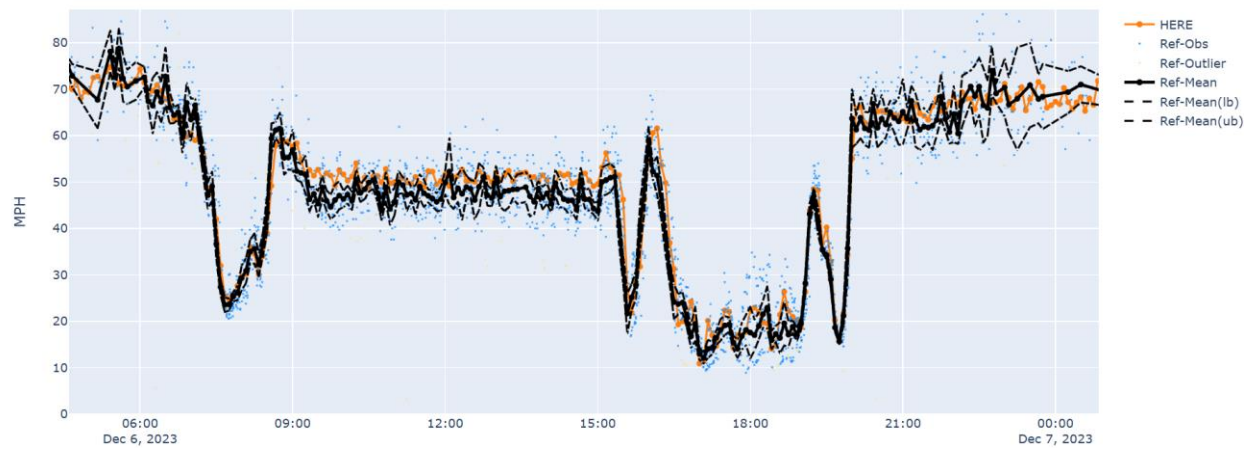


Figure 10 – Representative example from HERE

P18 | S19-to-S20 | 1.41 miles

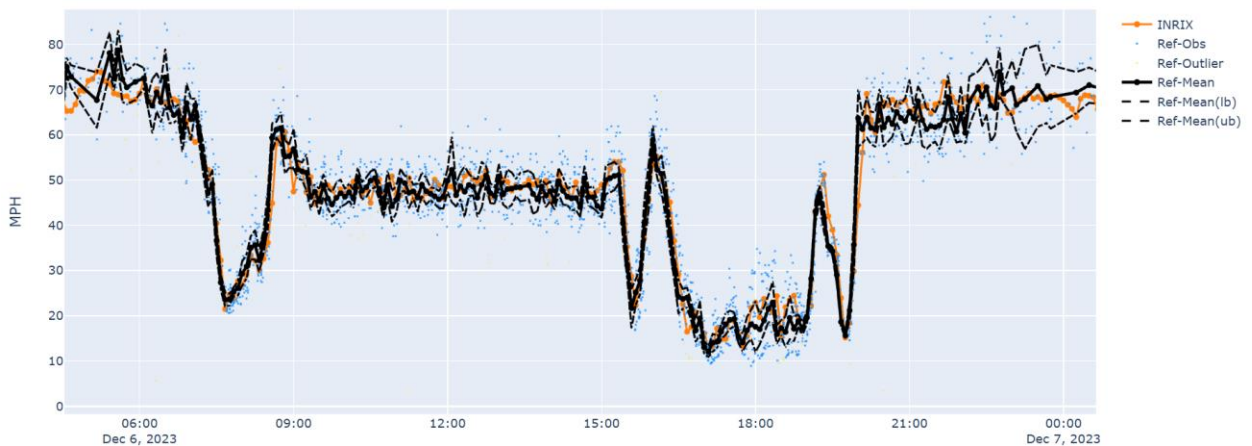


Figure 11 – Representative example from INRIX

P18 | S19-to-S20 | 1.41 miles

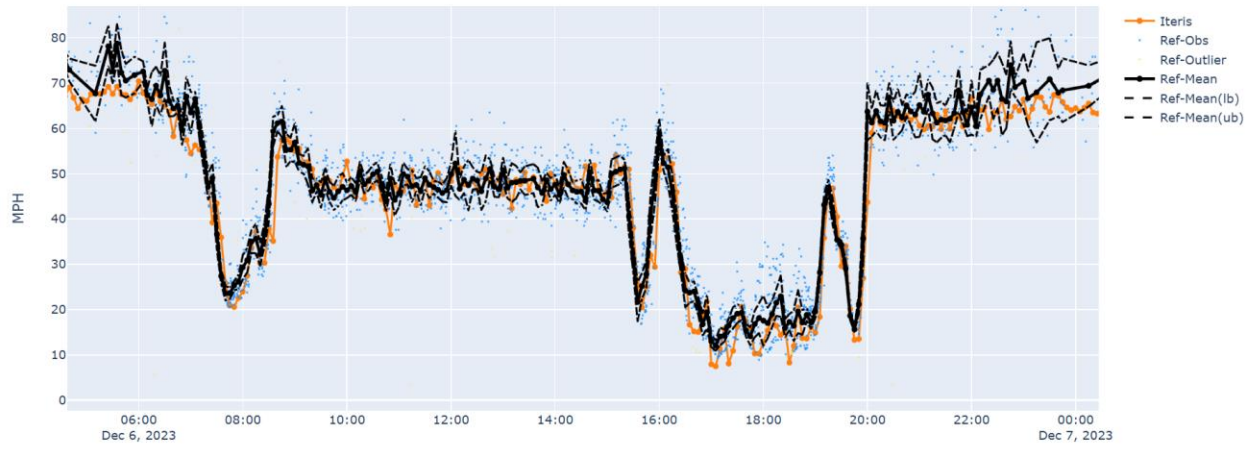


Figure 12 – Representative example from Iteris

P18 | S19-to-S20 | 1.41 miles

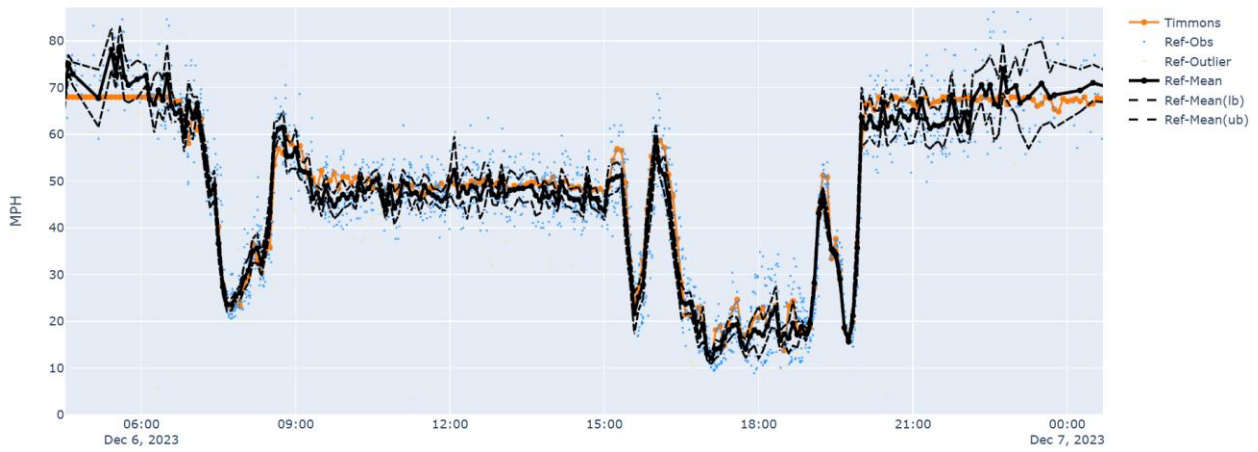


Figure 13 – Representative example from Timmons

Latency Analysis

The following two subsections summarize latency analysis results for the two vendors who submitted data in real-time, as the timestamps associated with their submissions can be certified. Aggregate summary results from all vendors who delivered data post-hoc are included at the end.

As mentioned in the methodology section above, three versions of latency were calculated for each Slowdown and Recovery Event (SRE). One measures latency for the full event (SRE), another focuses on just the slowdown period (SP), and third on just the recovery period (RP). However, **SP latency is generally the most relevant, and the focus of this analysis.** SP latency corresponds most closely to how latency is defined in the TMD RFP, is particularly important for operations applications (i.e., identifying the onset of congestion), and is most reliable to calculate (since data is typically cleaner than the recovery period, which can sometimes have lanes moving at different speeds).

Results are based on 101 slowdown and recovery events distributed across 13 validation segments. These congestion events – all of which were about one hour or more in length -- were manually identified using analyst judgement and the criteria described in the previous section.

Carto

Figure 14 shows a histogram summarizing the distribution of Carto latency values across all congestion events, *focusing specifically on the slowdown period (SP)* when congestion is first identified. Carto latency values most commonly fall in the 1–2-minute bin, and are distributed mostly symmetrically, with some observations less than 0 (likely an artifact of the fact that Bluetooth observations are associated with the trip end time). The mean and median latency are 1.1 and 1.0 minutes, respectively.

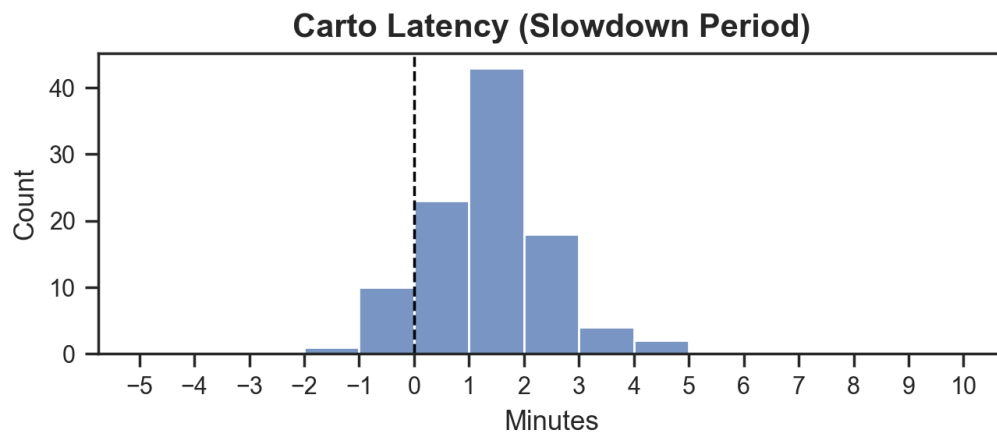


Figure 14 – Carto latency during slowdown period (SP)

Figure 15 shows the same distribution using a cumulative frequency diagram (CFD plot), which helps communicate the proportion of observations that meet different latency value thresholds. It shows that 95% of Carto SP latency values were less than 3.0 minutes, with 90% less than 2.3 minutes.

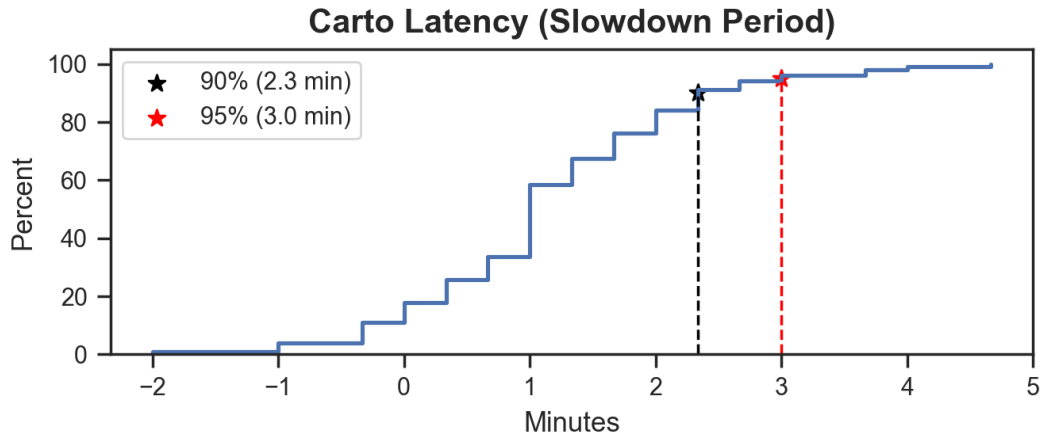


Figure 15 – CFDP plot of Carto latency during slowdown period (SP)

Figure 16 expands the analysis to consider not just latency associated with the slowdown period, but also the recovery period (RP) and the entire event (SRE). Each type of latency is summarized via boxplot, with the box portion showing the three quartiles of the distribution (Q1=25%, Q2=50%, Q3=75%), the whiskers showing the range of points that are not considered outliers (within +/- 1.5*IQR of the box), and the red X showing the mean value. Overall, these results visually communicate that Carto’s latency values are low (boxes are centered around small numbers), consistent (boxes are similar in size and not very wide), and for all three slowdown definitions, at least 75% of have latency less than 2 minutes (right edge of box, which is 75th percentile, is less than 2).

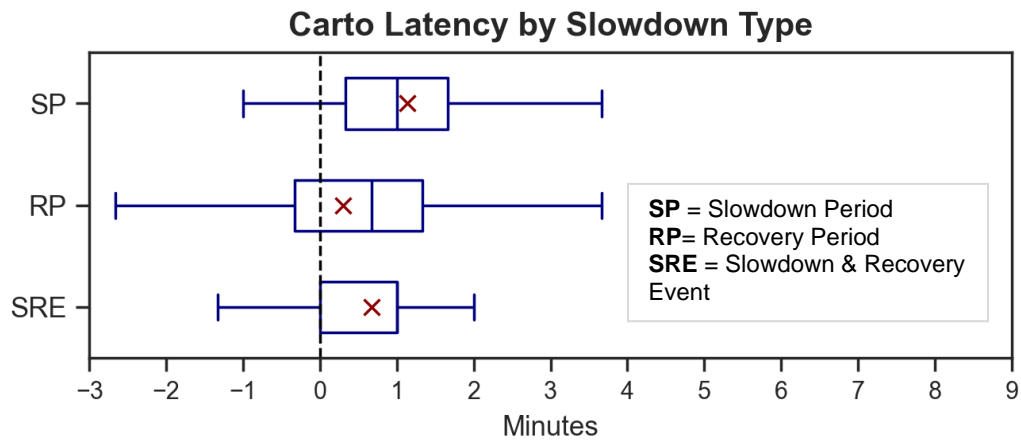


Figure 16 – Distribution of Carto latency during different parts of slowdown

Iteris

Figure 17 shows a histogram summarizing the distribution of Iteris latency values across all congestion events, *focusing specifically on the slowdown period (SP)* when congestion is first identified. The majority of Iteris latency values are roughly evenly distributed between 1-4 minutes, with a handful of larger positive and negative values in both directions. The mean and median values are 2.4 and 2.3 minutes, respectively.

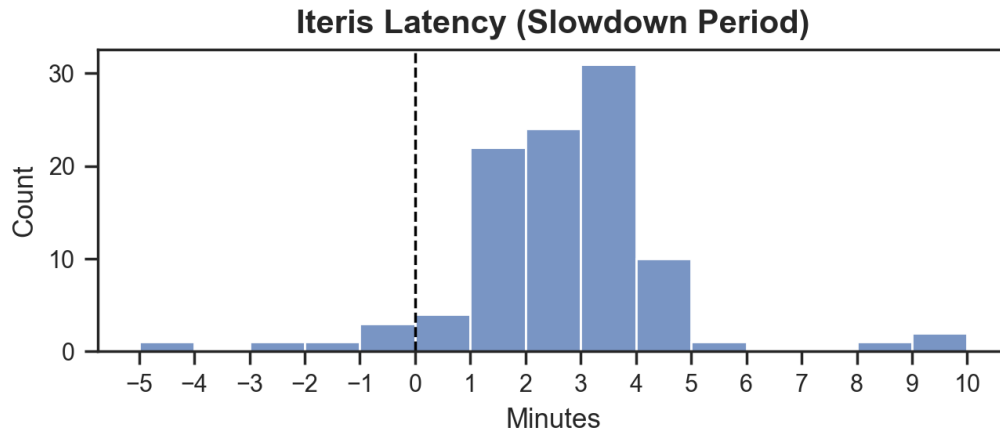


Figure 17 – Iteris latency during slowdown period (SP)

Figure 18 shows the same distribution using a cumulative frequency diagram, which helps communicate the proportion of observations that meet different latency value thresholds. It shows that 95% of Iteris SP latency values were less than 4.3 minutes, with 90% less than 4.0 minutes.

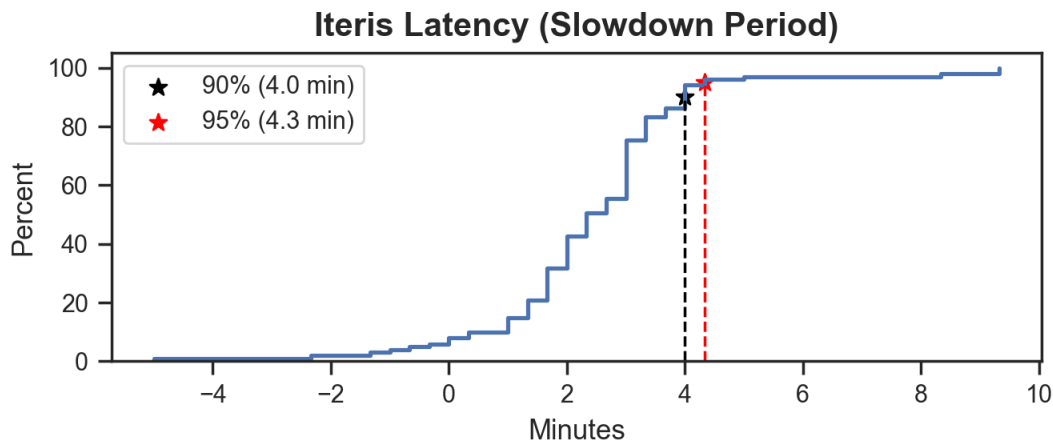


Figure 18 – CFD plot of Iteris latency during slowdown period (SP)

Figure 19 summarizes Iteris latency associated with the slowdown period (SP), recovery period (RP) and the entire event (SRE) using boxplots. These results visually communicate that Iteris' latency values are lower for SP than both RP and SRE by about 1-2 minutes (boxes are centered around small numbers). Typical (middle 50%) latency for SP is around 1.5-3.5 minutes, with RP around 3-5 minutes, and SRE around 3-4 minutes.

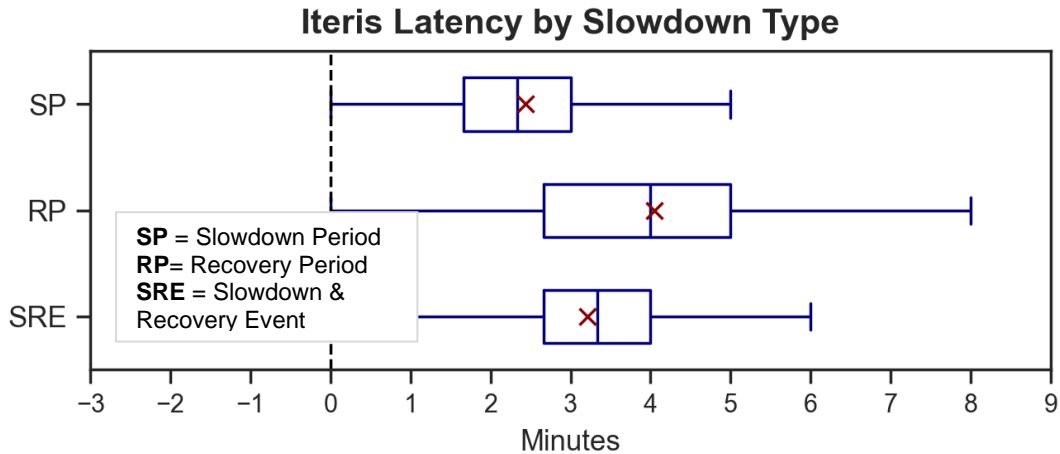


Figure 19 – Distribution of Iteris latency during different parts of slowdown

Comparison to VPPII

Table 7 summarizes the results for Carto, Iteris, and the average of all other vendors, and compares it to the values obtained during Coalition validation work in VPPII. Note that although the other vendors from TDM did not deliver data in real time, neither did all vendors in VPPII, which means that the results are generally comparable. This table shows that probe data latency has improved relative to VPPII, particularly when focusing on the most critical element of slowdowns – the onset of congestion measured by SP (slowdown period). The best SP latency values from VPPII (3.4 and 3.7 minutes for Vendors 3 and 1, respectively) were improved by 1-2 minutes in the TDM – including vendors whose timestamps could be certified via real-time delivery.

Table 7 – Comparison of average latency between VPPII and TDM

Average Latency (minutes)						
	VPP II			TDM		
	Vendor 1	Vendor 2	Vendor 3	Carto*	Iteris*	Other vendors (avg)**
SP	3.7	7.1	3.4	1.1	2.4	1.7
RP	4.1	5.2	2.9	0.3	4.0	1.0
SRE	3.9	6.6	2.8	0.7	3.2	1.4

* Certified real-time delivery ** Uncertified

Interpretation Notes

To re-iterate, the latency values reported in this section are based on comparisons with reference re-identification data whose records are assigned to the *trip end time*. This context is important and means that the latency values reported here cannot be blindly compared with industry claims without knowing the nuances of other methodologies.

To test the impact of using trip start versus end time, the validation team re-ran the analysis using the *trip start time* instead and found that for all vendors the average latency increased by about 3-3.5 minutes. However, it should be stressed that the increase in latency is not generalizable, but rather is specific to the study area and segment characteristics. The change in latency is related to segment travel times, which would likely be different if repeated in other locations. Nonetheless, this provides a ballpark figure for understanding the bounds of possible latency values at freeway speeds when segments are about 2 miles long.

Finally, one vendor noted they needed to develop custom processes to translate their native data feed to the CATT Works georeferencing format, which was less efficient than their production environment, and thus may have contributed to part of the latency. The validation team acknowledges this concern and notes that the measured latency does involve several components, including conflation to CWGP (see Figure 6). However, it should be noted that the translating data between road segment representations is an unavoidable aspect of using probe data – whether it happens on the data vendor or customer side. For example, state DOTs sometimes want to use commercial probe data on their own road network representation, which requires conflating vendor-supplied data to their own network (even if they consume real time data from a production API).

Conclusions

This report explores the fidelity of probe-based travel time and speed data from five vendors (Carto, HERE, INRIX, Iteris, and Timmons) on an urban freeway in Atlanta, GA using data collected during December 2023. Vendor data submissions were evaluated using two methods: (1) the traditional analysis that involves calculating two error metrics, Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) for five-minute period, and (2) a latency analysis, which measures the lag time between when a congestion event takes place and when it is reflected in vendor data feeds. The following conclusions emerge from the analysis:

- **All vendors captured recurrent and non-recurrent congestion patterns with a high degree of fidelity.** Each vendor was within specification for AASE and SEB metrics in all four speed bins, including the lower speed ranges that reflect congested periods and dynamic traffic conditions. Additionally, visual inspection of the trace data corroborated the high accuracy demonstrated by the AASE and SEB error measures, with vendors closely tracking fluctuations in speed observed in the reference data.

This result was not unexpected, as probe data has been high quality for many years on high volume freeways. Nonetheless, the findings are encouraging, as it provides an important benchmark after data supply chains were disrupted in 2022-2023, including the bankruptcy of Wejo (a data supplier to at least one travel time vendor) in June 2023.

- **A major milestone was achieved by successfully benchmarking latency with real-time data.** Previous validation activities laid the foundation for this outcome, which resulted in four of the five vendors submitting travel time records using the CATTWorks Geo-referencing Protocol (CWGP) in real-time. Two of the vendors, Carto and Iteris, were ultimately successful in this endeavor (two others attempted, but had to re-submit data post-hoc), marking the first time the validation team has measured probe data latency based on “certified” real-time submissions.

- **Care should be taken when comparing latency values.** Latency values reported in this report are based on comparisons with reference re-identification data whose records are assigned to the *trip end time*. This means that the latency values reported here cannot be blindly compared with industry claims without understanding the methodology used – particularly with respect to temporal point of reference. Regardless of the absolute value of latency reported, calculating latency the same way will provide insight into changes over time.
- **Latency results are encouraging and have improved relative to VPPII.** When calculating latency at the onset of congestion (i.e., the slowdown portion of a slowdown and recovery event), the two vendors who delivered certified real-time data had average latencies of 1.1 and 2.4 minutes (with 95% of observations under 3.0 and 4.3 minutes, respectively). The other three vendors' combined uncertified average latency was 1.7 minutes. These results show an improvement of 1-2 minutes relative to VPPII, where average latency at the onset of congestion ranged from 3.4 to 7.1 minutes.