THE EASTERN TRANSPORTATION COALITION

CONNECTING FOR SOLUTIONS

Validation of Waze Data for Travel Time/Speed Applications

Characterizes the accuracy of travel time / speed data extracted from Waze data feeds to determine suitability for DOT operations and decision making.

August 7, 2026

Prepared by: The Eastern Transportation Coalition

Executive Summary

This report presents findings from the Waze Validation special cooperative program (SCOOP) – an initiative funded by states in The Eastern Transportation Coalition (TETC) -- that assesses the suitability of using Waze's probe-based travel time and speed data for operations and decision-making applications.

The majority of Coalition member states share data and collaborate with Waze on enhanced traveler information, enabling Waze to display timely and accurate information to their navigation users (e.g., incidents that may impede traffic flow, lane closures for roadway construction or maintenance, visibility issues, and the presence of emergency responders). However, several agencies have begun to use data provided by Waze through the Waze Partner API for other operational use cases, such as populating travel time message on overhead DMS or monitor signal timing along corridors. As such, TETC's members requested a validation study to assess the efficacy of this real-time travel time/speed data for DOT operations and decision-making.

The validation study area focuses on three corridors in Florida: two signalized arterials in the Orlando metro area (SR-436 and W Lake Mary Blvd) and one freeway corridor just outside Jacksonville (I-95). These study corridors were chosen to encompass multiple road classes (signalized arterials and freeways) and volume levels (very high volume on I-95, plus two different volume ranges on the arterials). The freeway corridor, I-95 SE of Jacksonville, has AADTs ranging from 118k-166k (bi-directional) and experiences major congestion. The arterial corridors in the Orlando area are characterized by mid-to-low volumes with high signal density; SR-436 has AADTs in the 48k-65k range (bi-directional), while W Lake Mary has lower volumes around 18k-23k (bi-directional). Reference travel time and speed records were collected using Florida DOT's permanently deployed Bluetooth sensors between November 8 - December 5, 2025, while corresponding Waze data was collected by querying Waze's real-time API during the same period.

Waze data was evaluated using two methods commonly employed in TETC validation activities: (1) a 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. The following takeaways emerged from the analysis:

- **Waze's web platform allows users to define custom analysis segments, avoiding the need to later conflate probe data to desired segment definitions.**

  The Waze platform's built-in georeferencing capability is a significant feature, placing the burden of conflation on the data provider (Waze) rather than the consumer, and lowering the barrier-to-entry for using the data quickly. This is different than most probe datasets, which are usually delivered on static TMC or vendor-specific road networks, and require agencies to perform conflation to align with their preferred road network segmentation. Note, however, that this only works for small study areas; because segments are user defined, it is well suited for monitoring specific corridors – not an entire city or state.

- **Waze's probe-based travel time and speed data closely matched reference data acquired from FDOT Bluetooth sensors on both freeways and arterials.**

  Waze data showed high accuracy (as measured using AASE and SEB error metrics) across 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 Waze data closely tracking fluctuations in speed observed in the reference data. This high level of fidelity is encouraging – particularly with Waze data being updated less frequently than the commercial offerings in the TDM (2 minutes vs 1 minute).

Note, however, that both study areas (Orlando and Jacksonville) represent major urban areas that experience regular congestion, and thus may have a higher density of Waze users than other locations. While these results are encouraging, this does not necessarily mean that the results are transferable to all locations within a state.

- **When traffic speeds were bimodal or dispersed over a range of values, Waze data tended to track the slower mode.**

  The two signalized arterial corridors in the study area were characterized by a wide range of travel time and speeds – not just due to recurrent congestion at predicable times, but also due to cycle failures and the complexity associated with traffic entering/exiting at the many access points. In such cases, Waze's reported speeds were within the range of observed reference data observations, but tended to track slower speeds, reflecting a more conservative estimate of average speed.

- **Waze data showed minimal latency, indicating an ability to quickly identify the presence of congestion.**

  When calculating latency at the onset of congestion (i.e., the slowdown portion of a slowdown and recovery event), typical latency values are slightly negative (-0.7 minutes for both mean and median), with 95% of slowdowns having latency less than 2.3 minutes. Although this sounds surprising, negative numbers (i.e., Waze data detecting the slowdown before the reference data source) are possible because, by convention, reference re-identification travel time records are assigned to the *trip end time.* The latency values would increase slightly if using a different temporal point of reference; however, in any case, the **Waze real-time feed detects the onset of congestion with low latency.**

Based on these observations, it appears that Waze's Travel Time and Speed data is well-suited for travel monitoring applications along corridors or small study areas, particularly in major urban areas such as the ones analyzed in this study. However, agencies should recognize that it is not intended for large-scale (e.g., network-wide) analysis and that Google/Waze has not made any commitment to maintaining this API and functionality in the future, nor made any service level agreements (SLAs). Further study is needed to see whether the performance observed in this study would generalize to lower volume areas or locations with less Waze application usage.

Table of Contents

1. Introduction

The Eastern Transportation Coalition has a longstanding partnership with Waze. The majority of our member states share data and collaborate with Waze on enhanced traveler information. This data from DOTs allows Waze to display timely and accurate information to their navigation users, including incidents that may impede traffic flow, lane closures for roadway construction or maintenance, visibility issues, and the presence of emergency responders in or alongside the road to mitigate risks of struck-by incidents. A few states, however, capitalize on data being provided by Waze through the Waze Partner API for other operational use cases. Rhode Island DOT, for example, utilizes travel time/speed data from Waze to populate travel time messages on their overhead DMS. Florida DOT (District 5) is utilizing travel time/speed data to analyze and monitor signal timing along corridors in rural areas where they do not have other means of detection or ATSPM capabilities. As a result of conversations about these use cases for Waze travel time/speed data, TETC's members requested a validation study to assess the efficacy of this data for DOT operations and decision-making.

Since 2008, The Eastern Transportation Coalition data validation program has regularly benchmarked the quality of probe-based travel time and speed data sold through the Transportation Data Marketplace (formerly the Vehicle Probe Project, or VPP). To assess the quality of commercial probe datasets, UMD developed a methodology using wireless re-identification traffic monitoring technology. At a high level, this 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. This reference data is used to characterize 'ground truth' speed/travel time conditions and serves as a point of comparison with commercial probe data. Common error metrics and visualizations have been developed to quantify probe vendor performance and track data quality over time.

2. Study Area

Figures 1 and 2 show the geographic scope of the study area, which includes three corridors: two signalized arterials near Orlando, FL (SR-436 and W Lake Mary Blvd) and a freeway corridor near Jacksonville, FL (I-95). In total, 46 segments are defined across both directions of SR-436, W Lake Mary Blvd, and I-95.

These study corridors were chosen to encompass multiple road classes (signalized arterials and freeways) and volume levels (very high volume on I-95, plus two different volume ranges on the arterials). The freeway corridor, I-95 SE of Jacksonville, has AADTs ranging from 118k-166k (bi-directional) and experiences major congestion. The arterial corridors in the Orlando area are characterized by mid-to-low volumes with high signal density; SR-436 has AADTs in the 48k-65k range (bi-directional), while W Lake Mary has lower volumes around 18k-23k (bi-directional).

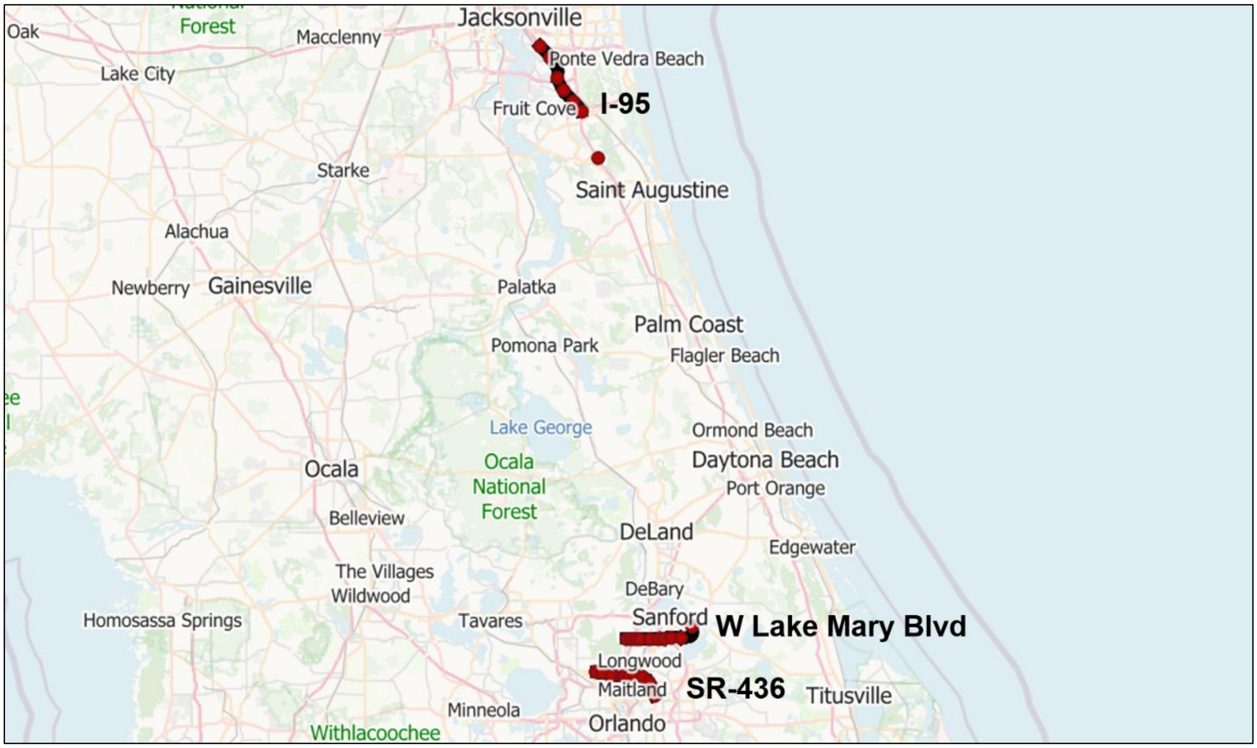

Figure 1: Summary view of study corridors

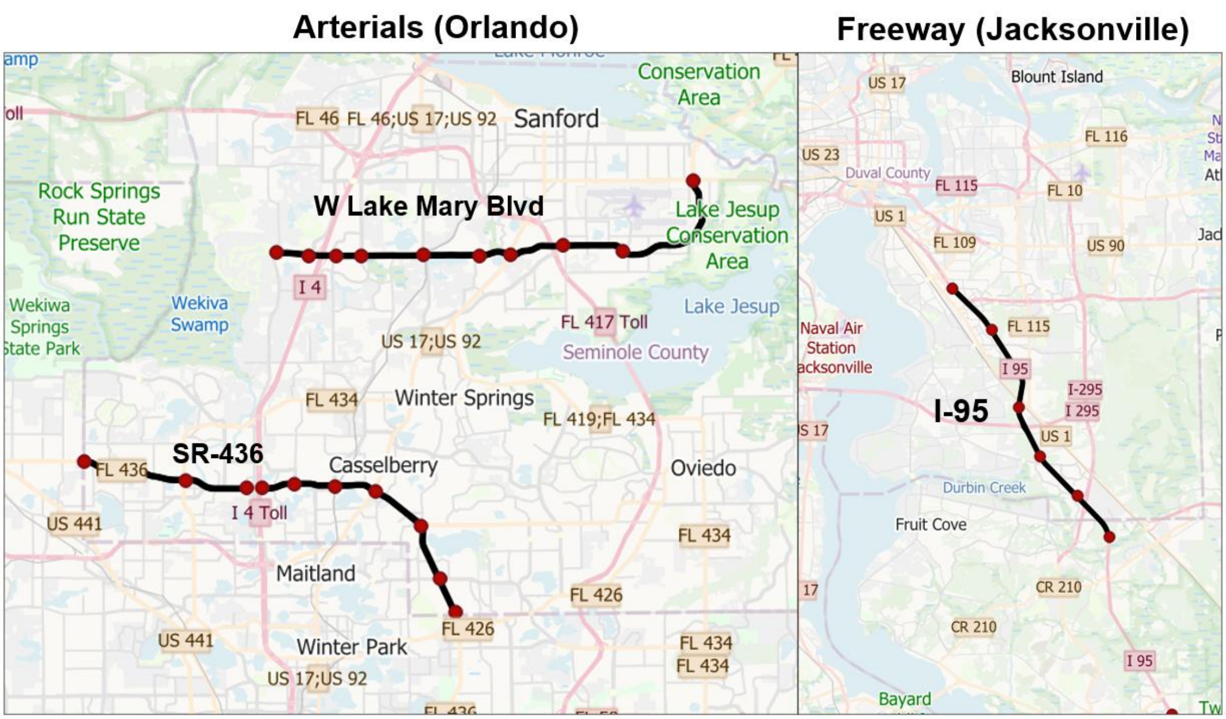

Figure 2: Validation segment definitions for two signalized arterials (left) and one freeway (right)

Validation segments, shown in Figures 1-2, were defined between pairs of Florida DOT Bluetooth devices, which typically coincide with interchanges, intersections, or other logical break points -- often about 1-2 miles apart. Because FDOT has a robust network of permanently deployed sensors with adequate spacing to support the study, Coalition staff did not deploy any supplemental counters for short-term collection.

In total, 46 segments were included in the study area, with 41 ultimately used for analysis. Several segments were not included in the final analysis due to reference data issues (e.g., sensor malfunctioned or produced reference data that could not be trusted as the basis for comparison with Waze data).

Table 1 shows the key attributes used to describe validation segment locations via the CATT Works Georeferencing Protocol (CWGP) – a lightweight georeferencing method that unambiguously describes where travel time/speed data will be evaluated. Although shown in tabular form here, this information is typically encoded as a GeoJSON file with MultiPoint coordinate geometry (representing the start and end points of each segment) so that it can be readily processed with a computer or displayed via web or Geographic Information System (GIS) tools. Given the large number of the segments included in the study, only a subset of segments is shown here for reference.

Table 1 - CWGP attributes for subset of validation segments

ID	Start Location	Start Heading	End Location	End Heading	Road Name	Road Class	Length (Miles)
Seg01	-81.459672 28.672713	91	-81.443846 28.669401	90	SR-436	2	1.02
Seg02	-81.4438 28.66936	88	-81.419172 28.665256	90	SR-436	2	1.548
Seg03	-81.419172 28.665256	91	-81.393528 28.661788	89	SR-436	2	1.618
Seg04	-81.388213 28.661943	114	-81.374136 28.663723	90	SR-436	2	0.879
…							
Seg46	-81.519744 30.120035	136	-81.497977 30.093016	169	I-95	1	2.305

3. Data Collection

Reference Data

Reference travel time samples were collected along validation segments from November 8 to December 5, 2025. Raw detection records were acquired from Bluetooth re-identification sensors owned and managed by Florida DOT, 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 by the TETC validation team to support these analysis routines.

Figure 3 shows speeds along a representative arterial segment during the study period (Seg 10 on SR-436), while Figure 4 shows a representative freeway segment (Seg 41 on I-95). In Figure 3 (arterial), the average speed fluctuates from about 20-50mph, dipping down to 10mph during the slowest period during weekdays, with individual speed observations spanning an even wider range. Few, if any, major slowdown events are observed along either arterial segment, but traffic speeds vary significantly during most periods, due to interrupted traffic conditions from signals and many access points along the corridors. In contrast, Figure 4 (freeway) shows much higher speeds – often around 70-75mph, with major slowdown events present, and much less dispersion around the average speed values. These clear congestion events provide ample opportunity to assess vendor data latency.

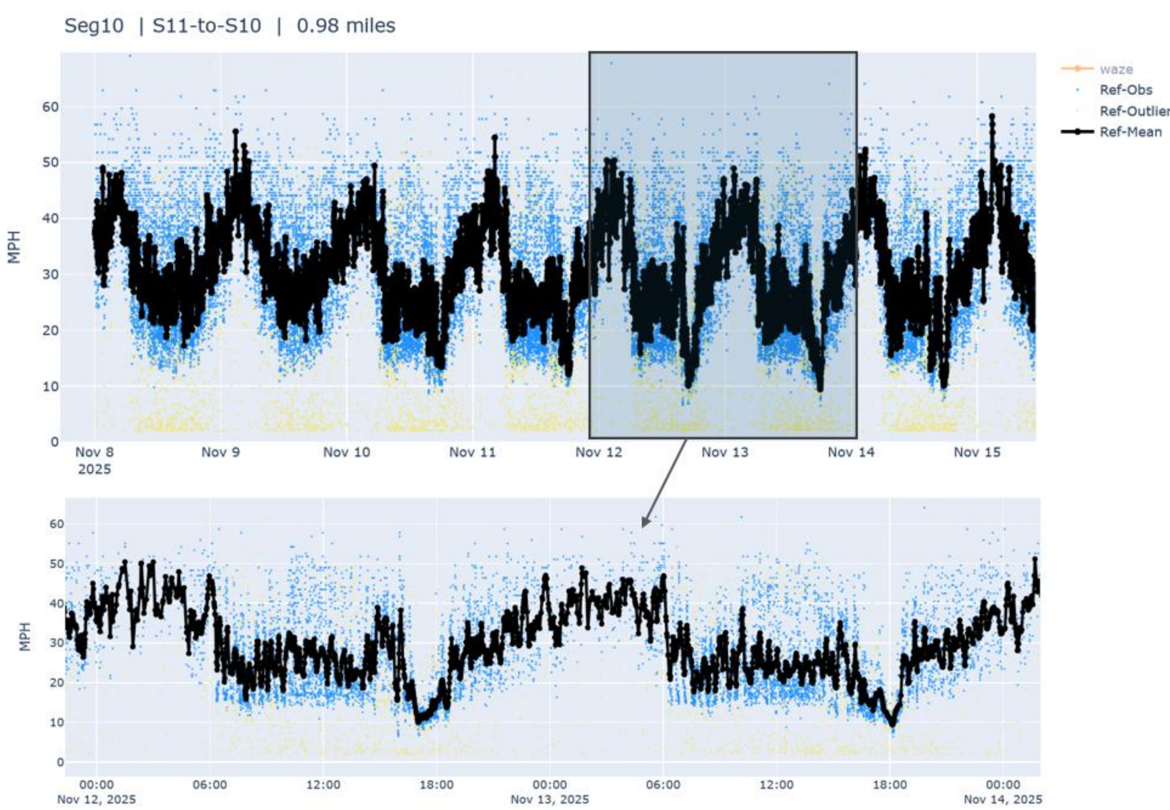

Figure 3 – Representative plot showing speed fluctuations on SR-436 (arterial)

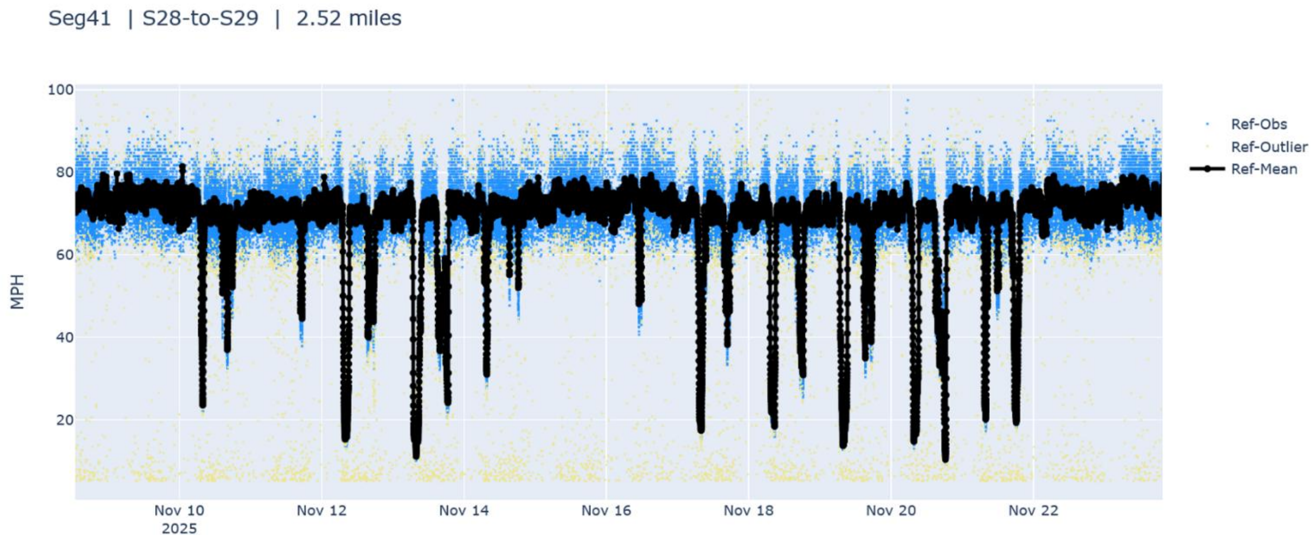

Figure 4 – Representative plot showing major slowdown events on I-95 (freeway)

Waze Data

Waze travel time data was obtained in real time through the Waze Partner Hub API. This process involves using Waze's online platform to first create a study region (custom geographic polygon) and then define a set of custom segment definitions, an example of which is shown below in Figure 5. To create a custom segment, a user clicks the desired start and end locations on an interactive web map and provides a unique ID (route name) to reference the segment. After clicking the start and end locations, the online tool 'snaps' the route to the road network. Waze segments were defined to align with the Reference segments, which were defined based on where FDOT Bluetooth segments were deployed. Once a new segment is saved, real-time traffic data can be accessed through a real-time API that is updated every 2 minutes and available in JSON (JavaScript Object Notation) or XML (Extensible Markup Language) formats.

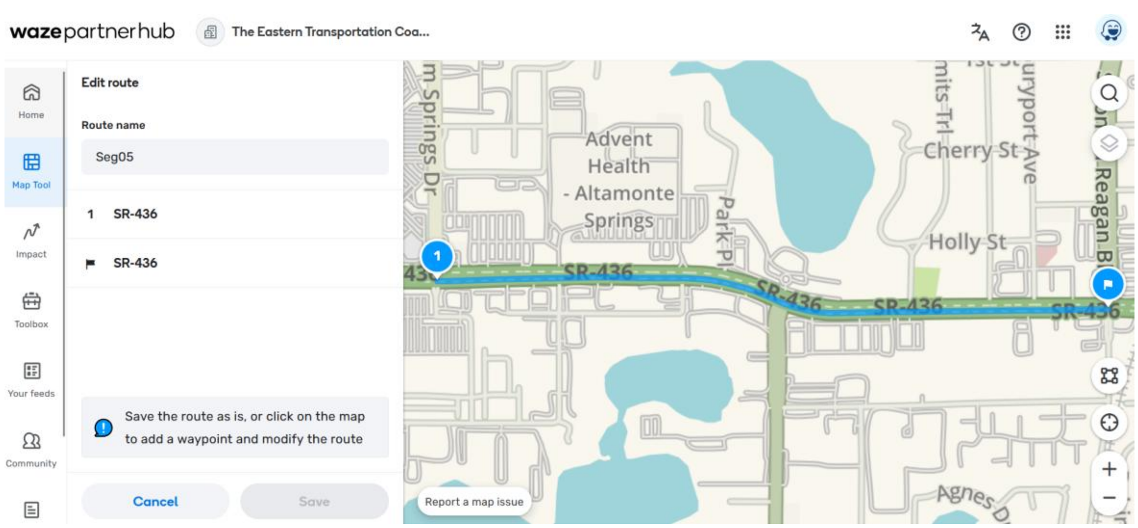

Figure 5 – Screenshot of Waze's online tool used to define custom segments.

The real-time API response contains information about both the overall study area (polygon) and user-defined segment locations. Across the entire study polygon, it flags sections of the road with irregular traffic conditions and reports the number of Waze users and cumulative length of different types of traffic jams. Additionally, at user-defined segment locations it reports the current travel time, historical travel time, and a flag indicating the level of traffic jam existing on the segment. Although information about potential traffic jams is included, there is no data density/sample size field that indicates whether the reported travel time value is based on real-time information (i.e., observed Waze users) or estimated another way.

To obtain a historical travel time/speed dataset for validation analysis, the project team wrote a script to query the custom API every minute and archived the results. Note that Waze's real-time API is updated every two minutes, but because it is unclear exactly when the update occurs, the project team decided to query every minute, even if some traffic data metrics remain the same for successive minutes.

Benefits of Using Georeferencing Protocol

The fact that Waze allows users to define their own segments is a key finding in and of itself; it makes it much more straightforward for users to obtain travel time/speed data for the exact study area of interest, without needed to figure out how to properly aggregate TMC-based data or handle cases where predefined segments do not align. That said, because Waze's web platform requires manual configuration of segments, it is only suitable for small study areas, not a full transportation network.

The ability to define custom segment definitions is analogous to the Coalition's Catt Works Georeferencing Protocol (CWGP), which is employed in travel time validation activities for the Eastern Transportation Coalition's data validation program.

4. Evaluation Methodology

Waze data was 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. Waze 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 probe speeds from Waze 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 11/20/2025 at 8:03:14 AM would be assigned to the time bin 11/20/2025 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.

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, 0-15 mph, 15-25 mph, 25-35 mph, 35+ mph for arterials). 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 5 mph and 4 mph, 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, $\text{AASE} = \sum_{i=1}^{n} \left| S_{V_i} - S_{R_i} \right|$. Because the absolute value is used, positive and negative errors cannot cancel, and the result is always positive. Speed Error Bias is calculated similarly, with the difference that the absolute value of the errors is not taken. In other words, $\text{SEB} = \sum_{i=1}^{n} S_{V_i} - S_{R_i}$. 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 6 shows this visually, with the black curve representing reference data and the orange curve showing the vendor data that lags the reference curve.

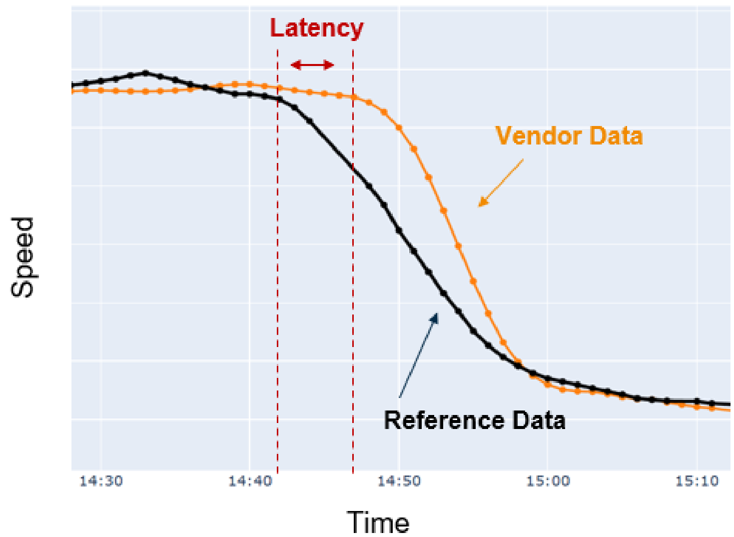

Figure 6 – Illustration of latency

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]

  Because Waze's dataset has a 2-minute update frequency, this process will result in a smoothed 1-minute dataset that can be aligned with the smoothed reference data.

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

[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)

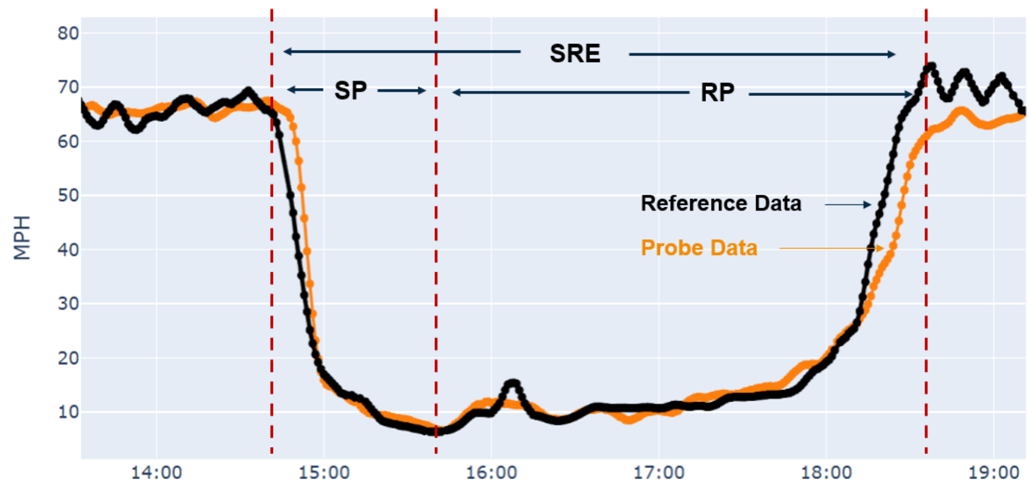

Figure 7 – Parts of a slowdown (SP, RP, SRE)

- **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 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.

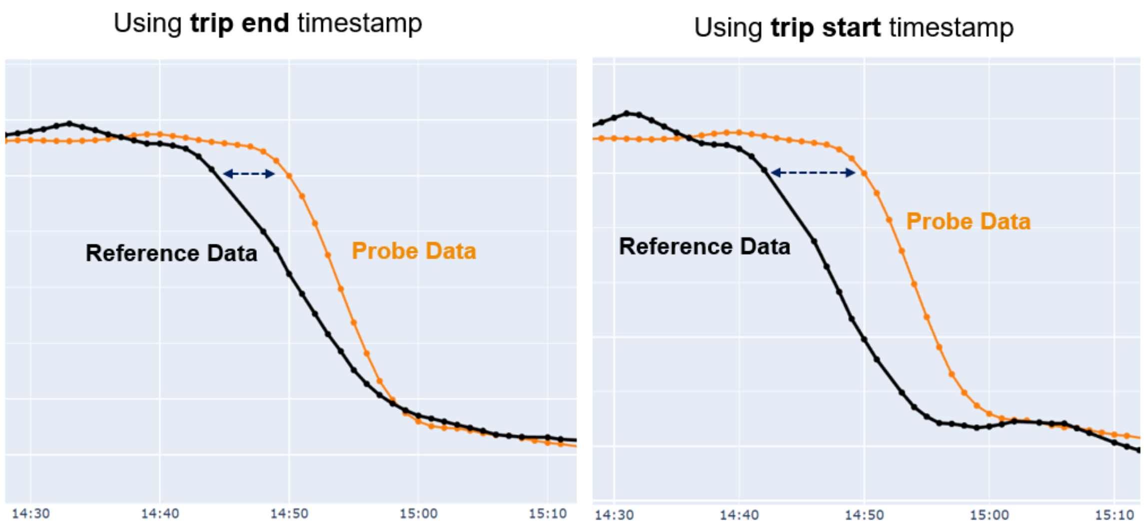

Figure 8 – Impact of using the start versus end timestamp for Reference data

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 TDM validation program*. 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.

5. Results

The following subsections report results for traditional and latency analyses separately. The traditional analysis focuses on summary accuracy metrics and is produced for both arterials and freeways (albeit with different speed bins for reporting). The latency analysis focuses specifically on freeway segments, consistent with how the method has been used in the Transportation Data Marketplace validation program.

Traditional Analysis

Tables 2 and 3 summarize the error metrics computed between reference and vendor speeds along arterial and freeway segments, respectively, while Figures 9-13 provide a representative visual example for each corridor.

For both arterials and freeways, Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) measures are well within Coalition TDM targets for all speed bins – whether measured in comparison to the 1.96 SEM band (as has historically been done for contract specifications) or the Mean speed itself (a more difficult comparison). Overall, these tables communicate that on both arterial (SR-436 and W Lake Mary Blvd) and freeway (I-95) validation segments, Waze reported speeds closely matched the observed Reference speeds.

Table 2: Error metrics for traditional analysis (Arterial)

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	
0-15	4.0	2.5	3.8	2.4	3422
15-25	3.2	1.2	1.7	0.9	38754
25-35	3.7	1.1	-1.4	-0.5	65139
35+	4.7	2.1	-4.2	-2.1	69404
All speeds	4.0	1.6	-1.7	-0.8	176719

Table 2: Error metrics for traditional analysis (Freeway)

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	1.9	1.4	0.4	0.3	776
30-45	2.9	2.0	-1.0	-0.7	1044
45-60	3.6	2.2	-2.5	-1.7	1064
60+	3.1	1.3	-2.2	-1.1	64428
All speeds	3.1	1.3	-2.2	-1.1	67312

Figures 9-12 serve as a visual example to complement the summary statistics above. Figures 9 and 10 focus on SR-436, with Figure 9 showing a representative example of Waze data tracking reference values closely over a one-to-two-day period, and Figure 10 highlighting how Waze data behaves when traffic speeds are bi-modal. Figure 9 highlights that although reference speeds are dispersed (wide band of blue speed observations, likely due to a complex arterial environment with traffic signals and many access points), the mean reference speed (black) agrees closely with what Waze reports (orange) – even during speed fluctuations. Figure 10 (upper plot) shows an example of bi-modal traffic speeds, where the blue dots (individual speed observations) are clustered into two groups, most likely due to a traffic cycle failure where some vehicles make a green light and others have to wait until the next cycle. In such cases, Waze data tends to track the lower group of speed observations, reflecting a conservative estimate of average speeds. Figure 10 (lower plot) adds the reference mean speed in solid black, showing what the mean value would be based on the observed Bluetooth records. It should be noted that Waze's lower speeds are legitimate observed speeds, even if the average disagrees slightly with the average Reference value.

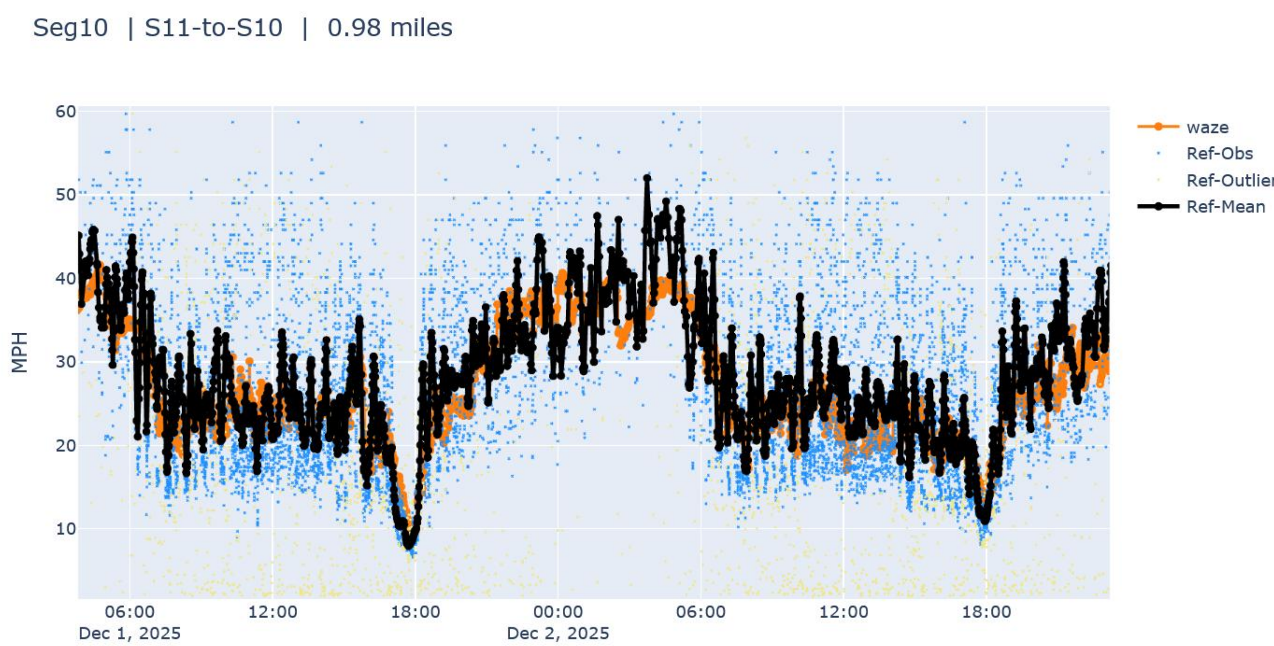

Figure 9 – Representative example on SR-436 (arterial)

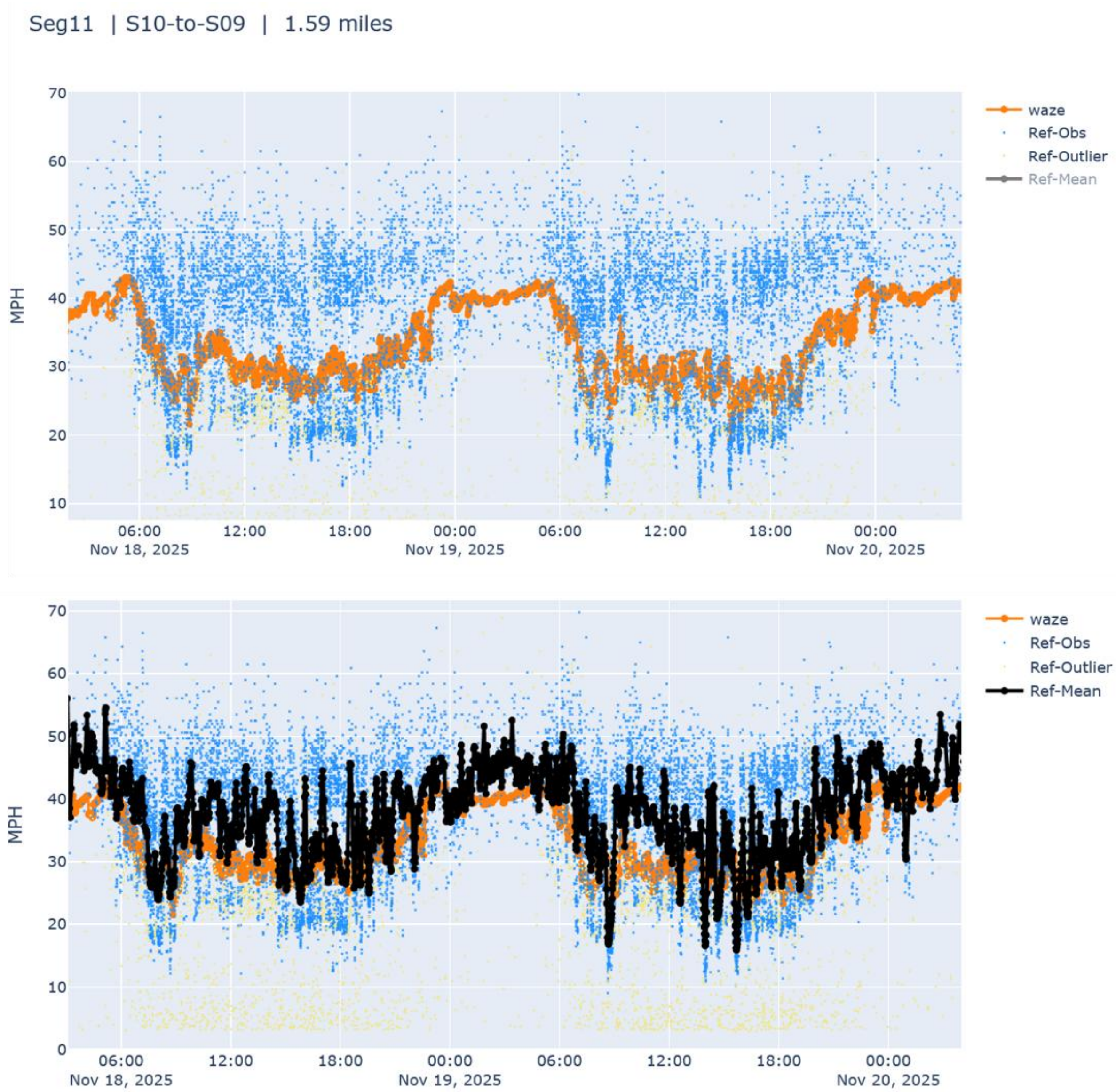

Figure 10 – Bimodal speeds on SR-436, with Waze tracking lower speed

Figure 11 is another example an arterial segment, this time on the lower-volume W Lake Mary Blvd. As was the case on the other arterial (SR-436), Waze data visually shows strong agreement. In this case, Waze speeds tend to be several mph lower during overnight and low volume periods or when speeds are highly dispersed but converges to the reference value when traffic slows and reference speeds are more uniform. These periods often coincide with higher volumes and more confidence in the reference values.

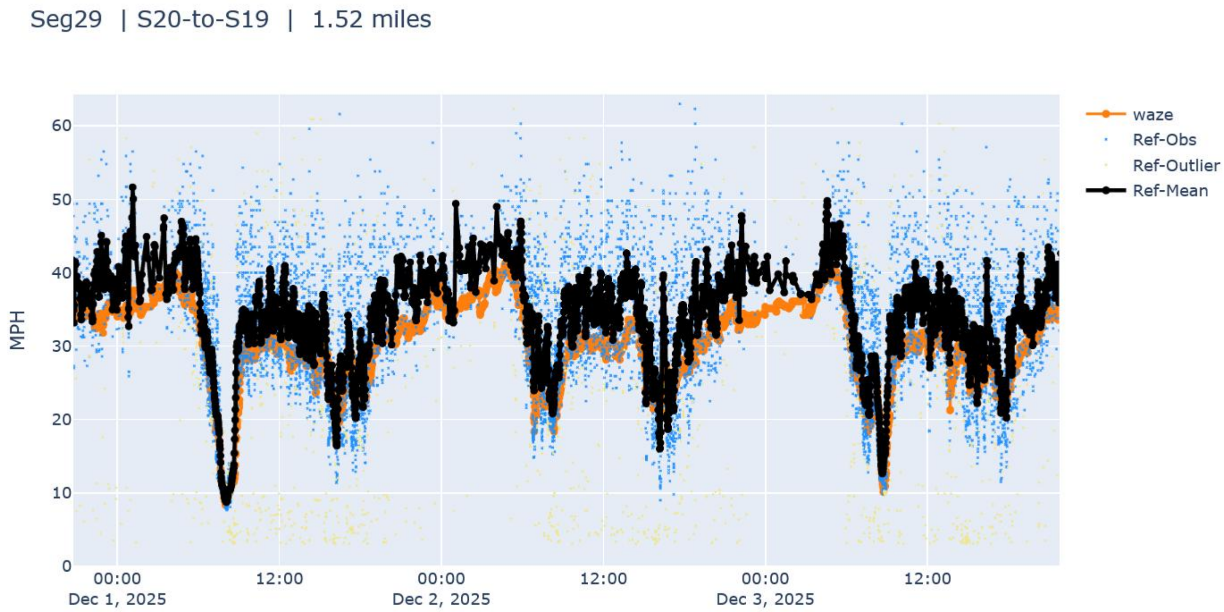

Figure 11 – Representative example W Lake Mary Blvd (Arterial)

Figure 12 shows an example segment along I-95 over the course of a day when major congestion is observed – both in the AM and PM peak periods. Free flow traffic speeds are around 65-70mph, but the average speed drops to about 15mph during the morning and 20-40mph during the afternoon peak. Visually, it is clear that the Waze data feed tracks the reference data nearly perfectly during congested periods, which is most important. However, as was observed on arterials, during periods when reference speeds are more dispersed (i.e., no congestion, close to free-flow speed), Waze reports slightly lower average speeds.

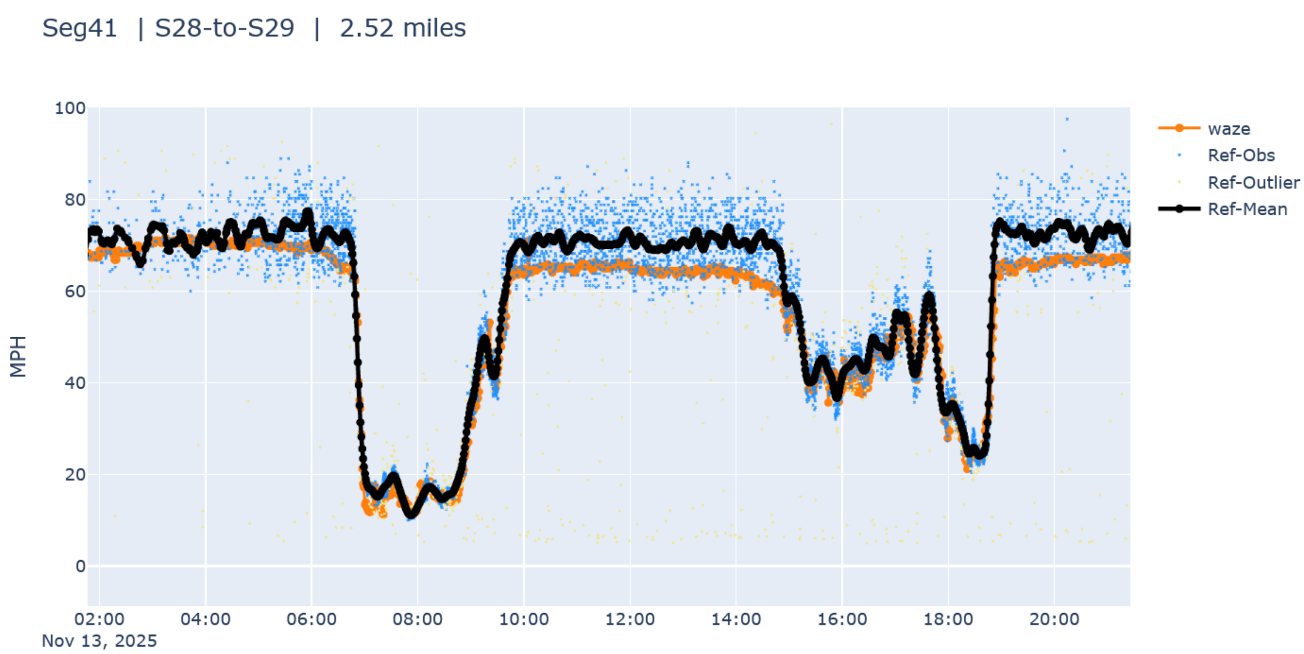

Figure 12 – Representative example from I-95

Latency Analysis

The following subsection summarizes the latency analysis results based on Waze's real-time data feed. Results are based on 62 slowdown and recovery events distributed across 10 freeway segments on I-95, which were compiled by identifying up to 10 suitable slowdown events on each segment. 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.

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).

Summary Results

Figure 13 shows a histogram summarizing the distribution of Waze latency values across all congestion events, *focusing specifically on the slowdown period (SP)* when congestion is first identified. Waze latency values typically fall in the range of -3 to 1 minute range, with both mean and median latency values of -0.7 minutes. The negative latency value is an artifact of the fact that by convention, we use the end timestamp associated with Bluetooth travel time observations. In any case, these results indicate that the **Waze travel time data has extremely low latency**.

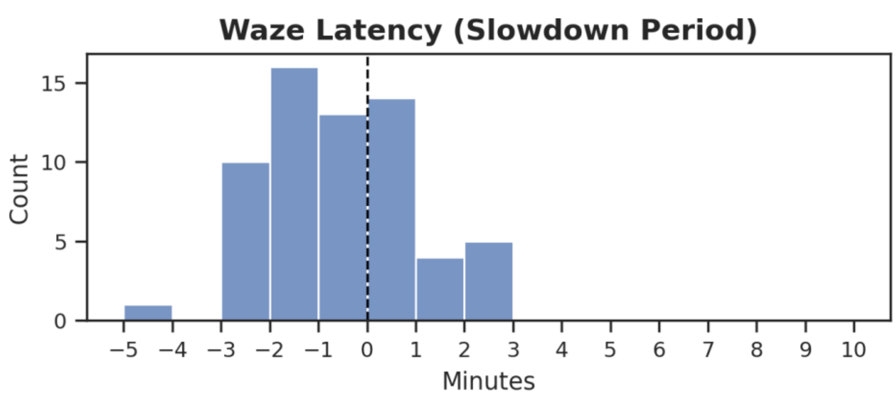

Figure 13 – Waze latency during slowdown period (SP)

Figure 14 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 Waze SP latency values were less than 1.3 minutes, with 90% less than 2.3 minutes**.

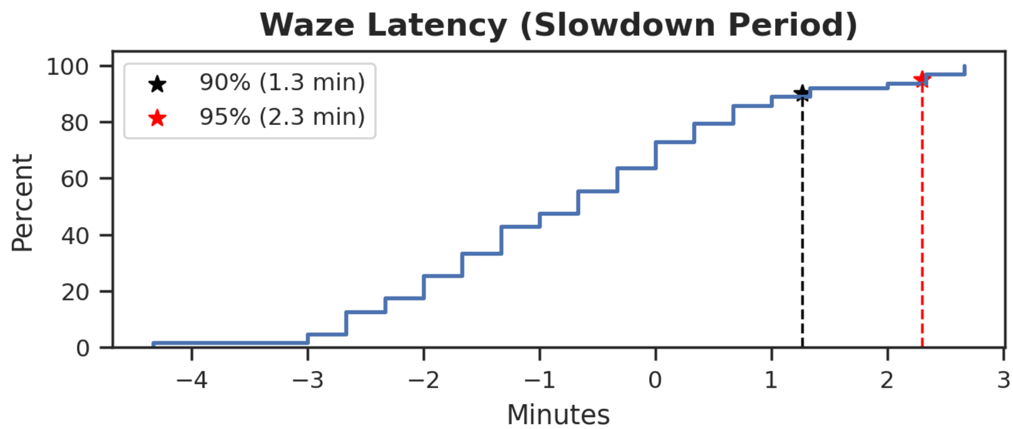

Figure 14 – CFD plot of Waze latency during slowdown period (SP)

Figure 15 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 Waze's latency values are low (boxes are centered around small numbers), reasonably consistent (boxes are similar in size and not very wide). In comparison to the slowdown period (SP), the recovery period (RP) has higher latency, although the mean and median values are still less than 2 minutes.

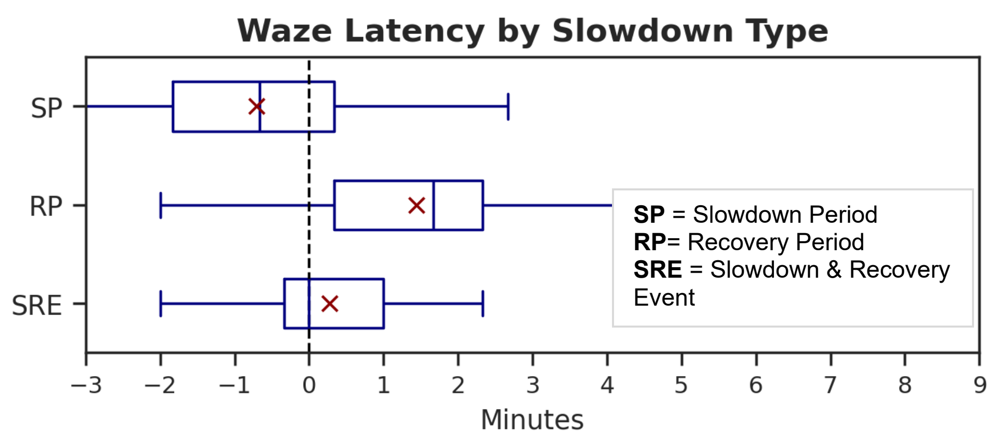

Figure 15 – Distribution of Waze latency during different parts of slowdown

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 the average latency increased by about 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 typically 1-2 miles long.

6. Conclusions

This report explores the fidelity of probe-based travel time and speed data extracted from Waze's real time API. The validation analysis focuses on three corridors in Florida: two signalized arterials in the Orlando metro area (SR-436 and W Lake Mary Blvd) and the I-95 freeway corridor just outside Jacksonville, with a study period of November 8 - December 5, 2025. Reference travel time and speed records were collected using Florida DOT's permanently deployed Bluetooth sensors, while corresponding Waze data was collected and archived by querying Waze's real-time API.

The Waze data was evaluated using two methods often employed in Coalition travel time validation activities: (1) a traditional analysis that involves calculating 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 vendor data feeds. The following conclusions emerged from the analysis:

- **Waze's web platform allows users to define custom analysis segments, avoiding the need to later conflate probe data to desired segment definitions.**

  The Waze platform's built-in georeferencing capability is a significant feature, placing the burden of conflation on the data provider (Waze) rather than the consumer, and lowering the barrier-to-entry for using the data quickly. This is different than most probe datasets, which are usually delivered on static TMC or vendor-specific road networks, and require agencies to perform conflation to align with their preferred road network segmentation. Note, however, that this only works for small study areas; because segments are user defined, it is well suited for monitoring specific corridors – not an entire city or state.

- **Waze's probe-based travel time and speed data closely matched reference data acquired from FDOT Bluetooth sensors on both freeways and arterials.**

  Waze data showed high accuracy (as measured using AASE and SEB error metrics) across 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 Waze data closely tracking fluctuations in speed observed in the reference data. This high level of fidelity is encouraging – particularly with Waze data being updated less frequently than the commercial offerings in the TDM (2 minutes vs 1 minute).

  Note, however, that both study areas (Orlando and Jacksonville) represent major urban areas that experience regular congestion, and thus may have a higher density of Waze users than other locations. While these results are encouraging, this does not necessarily mean that the results are transferable to all locations within a state.

- **When traffic speeds were bimodal or dispersed over a range of values, Waze data tended to track the slower mode.**

  The two signalized arterial corridors in the study area were characterized by a wide range of travel time and speeds – not just due to recurrent congestion at predicable times, but also due to cycle failures and the complexity associated with traffic entering/exiting at the many access points. In such cases, Waze's reported speeds were within the range of observed reference data observations, but tended to track slower speeds, reflecting a more conservative estimate of average speed.

- **Waze data showed minimal latency, indicating an ability to quickly identify the presence of congestion.**

  When calculating latency at the onset of congestion (i.e., the slowdown portion of a slowdown and recovery event), typical latency values are slightly negative (-0.7 minutes for both mean and median), with 95% of slowdowns having latency less than 2.3 minutes. Although this sounds surprising, negative numbers (i.e., Waze data detecting the slowdown before the reference data source) are possible because, by convention, reference re-identification travel time records are assigned to the *trip end time.* The latency values would increase slightly if using a different temporal point of reference; however, in any case, the **Waze real-time feed detects the onset of congestion with low latency.**

Based on these observations, it appears that Waze's Travel Time and Speed data is well-suited for travel monitoring applications along corridors or small study areas, particularly in major urban areas such as the ones analyzed in this study. However, agencies should recognize that it is not intended for large-scale (e.g., network-wide) analysis and that Google/Waze has not made any commitment to maintaining this API and functionality in the future, nor made any service level agreements (SLAs). Further study is needed to see whether the performance observed in this study would generalize to lower volume areas or locations with less Waze application usage.