# I-95 Corridor Coalition

I-95 Corridor Coalition
Vehicle Probe Project: Validation
of Arterial Probe Data 2015



*July 2015*

# I-95 CORRIDOR COALITION VEHICLE PROBE PROJECT VALIDATION OF ARTERIAL PROBE DATA
## JULY 2015

## *Summary Report*

*Prepared by:*
Stanley E. Young
Masoud Hamedi
Elham Sharifi
Reuben Morris Juster
Kartik Kaushik
Sepideh Eshragh

University of Maryland, College Park

## *Executive Summary*

At the beginning of 2013 the I-95 Corridor Coalition directed the Vehicle Probe Project (VPP) validation team to shift focus from freeways to signal controlled arterials, and to develop a comprehensive analysis of probe data quality on arterial roadways. In response, the VPP validation team collected traffic data and compared it to VPP reported data on a variety of arterial roadways within participating states from April 2013 through June of 2014. During this period nine data collection activities were carried out on 14 corridors within the states of New Jersey, Pennsylvania, Maryland, Virginia, and North Carolina covering 320 miles, and spanned roadway functional classes of *Other Principal Arterials*, *Minor Arterials*, and *Major Collectors*, based on Highway Performance Monitoring System[1] (HPMS) classifications. On all of these facilities outsourced probe data provided through the VPP was compared with field-collected Bluetooth Traffic Monitoring (BTM) data which provided a reference source for travel time on each segment. The comparison of the two data sources was conducted using four analysis methods which included:

(1) A *traditional analysis* using precision and bias metrics which compared the speed reported by the probe data with the speed reported by BTM reference data in five minute intervals. The traditional analysis was derived from the freeway validation method that has been used since 2008.

(2) A *slowdown analysis* rated how accurately probe data capture significant disruptions defined as reductions of traffic speed by at least 10 to 15 miles per hour for 30 minutes or longer. Three ratings were used for each slowdown: *fully captured*, *partially captured*, and *failed to capture*.

(3) A *sampled distribution method* assessed probe data's ability to accurately portray recurring congestion. Weekday overlay plots combined with percentiles-based *cumulative frequency diagrams* (CFDs) provided visual feedback on the fidelity of probe data.

(4) Lastly, the probe and BTM data from each day of data collection on each segment were thoroughly reviewed by a team of researchers. This provided first-hand review of the fidelity of arterial probe data in comparison to BTM, and it provided the research team a perspective to judge the effectiveness of the various analysis methods.

This extensive validation effort resulted in recommendations on the use of outsourced probe data for operations and performance measure purposes, considerations for future use, and future validation emphasis for probe data on arterial roadways. The initial top level recommendations for accuracy and use of existing VPP data for arterial roadways is summarized in Table ES-1

---

[1] Highway Performance Monitoring System

Table ES-1. Arterial Probe Data Usability

| ✓ RECOMMENDED | 🔍 SHOULD BE TESTED | ✖ NOT RECOMMENDED |
|---|---|---|
| ● <= 1 signal per mile <br>● AADT > 40,000 vpd (2-way) <br>● Limited curb cuts <br>**Principal Arterials** <br>**Likely to be accurate…** | ● 1 to 2 signals per mile <br>● AADT 20K to 40K vpd (2-way) <br>● Moderate number of curb cuts <br>**Minor Arterials** <br>**Possibly accurate, test …** | ● >= 2 signals per mile <br>● AADT < 20K (2-way) - low volume <br>● Substantial number of curb cuts <br>**Major Collectors** <br>**Unlikely to be accurate…** |

- **Probe data is recommended for operations and performance measures when the average signal density on a corridor is 1 signal per mile or less**, the Average Annual Daily Traffic[2] (AADT) is 40,000 or greater, and limited curb cuts and access points to disrupt traffic flow. The quality of probe data on such roadways was observed to approach that of freeway data quality, and is expected to accurately capture a large majority of significant slowdowns. On these roadways probe data can support a broad range of applications such as performance measures for MAP-21, planning studies, before and after analysis, traffic operations, and traveler information.

- **Probe data should be used with caution when the average signal density on a corridor is between 1 to 2 signals per mile or less**, the AADT is between 20,000 and 40,000 or greater, and the roadway has a moderate number of curb cuts and access points to disrupt traffic flow. On such roadways, the VPP data may fail to capture significant slowdowns up to 50% of the time. Probe data may be used for comparative before and after studies and as an operations support tool. However, data should be test if used for performance measures, planning, or traveler information.

- **Probe data is NOT recommended for operations and performance measures when the average signal density on a corridor is 2 signals per mile or greater**, the AADT falls below 20,000, substantial curb cuts and access points to disrupt traffic flow, and/or number of lanes fall below two per direction. On such roadways, the VPP data is expected to not capture the majority of significant slowdowns, and is not recommended for applications at this time.

Probe data performance correlated best with signal density. Although other geometric attributes are listed in Table ES-1, signal density is the foremost indicator of probe data performance. Increased volume, as measured by AADT, will increase accuracy of probe data all other factors being equal. However, volume of traffic alone does not overcome the challenges of reporting accurate speed and travel time as a result of the complex stop and go motion of vehicles on arterials with dense signal spacing.

Probe data quality is anticipated to improve in time with increased probe density and improved algorithms. Data collection in spring of 2015 on US-1 in Virginia showed evidence of improvement. Arterial data quality should continue to be periodically monitored by the Coalition's validation efforts.

Throughout all the testing the VPP data exhibited some fundamental issues related to traffic characterization on signalized roadways. These fundamental issues included:

---

[2] Average Annual Daily Traffic as defined by HPMS

- **Probe data consistently errored toward faster speeds during congested periods.** As probe data quality improves it will report the full extent of slowdown more accurately, and congestion may appear to grow worse when in actuality it is only the quality of the probe data that is improving. This scenario has been corroborated by early adopters of probe data for arterial performance measures.
- **Whenever platoons of vehicles were consistently split by a red light resulting in two distinct speed profiles, probe data invariably reported the faster of the two modes.** This phenomenon was independent of the type of roadway and their geometric attributes, and appeared any time bi-modal flow was encountered with only a handful of exceptions.
- **Complex flow patterns common on signalized roadways cannot be observed in VPP data.** Bi-modal flow resulting from a portion of the platoon progressing on green and a portion forced to stop till the next cycle is just one example of complex flow. Probe data reports only an average speed, and, unlike BTM data, provides no information on the variation of speed.

This validation effort was the most expansive to date and resulted in several findings related to appropriate validation methodology for arterials. A by-product of this validation effort was the advancement of appropriate methodologies for arterial analysis, and documenting the shortcomings of procedures originally created for freeway analysis. Future arterial validation efforts (as well as future work in the areas of appropriate arterial performance measures) should consider the advancements in the use of the various methodologies employed herein.

As a result of the findings of this study, the following recommendations are made to the I-95 Corridor Coalition and its members.
- **The I-95 Corridor Coalition should continue to monitor outsourced probe data fidelity on arterials as part of the VPPII initiative.** Probe data quality is anticipated to improve as probe data densities increase, and algorithms allowing for point pairing improve. Validation of multi-vendor probe data available in VPP II will continue to benchmark industry capability moving forward.
- **Future work on arterial performance measures and probe data validation on arterials should build on the methodologies established,** realizing that freeway measures and methods are inadequate on arterials. Future effort should engage traffic engineering community along with the planning and operations community to merge current work on arterial management with probe data initiatives.
- **The Coalition should engage probe data providers and industry researchers to explore and prototype new data items capable of fully characterizing the complex arterial travel patterns including resulting from signal control.** At the base of this discussion are issues related to 'what should be reported?'. Current practice of mean speed measurements alone fail to capture the dynamics of arterial traffic flow.

The I-95 Corridor Coalition's Vehicle Probe Project continues to lead the country as the epicenter for probe-based operations and planning performance measures. The unique nature of a common licensing agreement combined with common data formats and analysis tools provides the best foundation to continue to advance arterial performance measure practice from a multi-state, industry and research collaborative environment.

# Table of Contents

# List of Figures and Tables

# Validation Results for Arterial Probe Data

## 1.0   Introduction and Background

The I-95 Corridor Coalition's Vehicle Probe Project (VPP) began in 2008 with a vision of an East Coast wide traffic monitoring system that provided a common operations picture to all jurisdictions, as well as a basis for performance measurement.  That vision included freeways as well as major arterials which provide access to the freeway system and arterials that provide alternative paths in the event of incidents or severe congestion.   For this reason signal-controlled, interrupted-flow facilities, generally referred to as arterials within this report, were included in the original procurement for the VPP.  Such arterial facilities have taken a more prominent role now that the VPP has entered its second phase, referenced as VPPII.  The procurement specifications for VPPII began to put quality requirements on arterials based on knowledge gained during the VPPI validation program.   The first few years of the VPPI validation program were dedicated primarily to freeway data quality.  As the freeway data quality became better understood, validation resources were diverted to explore the relatively unknown arterial probe data quality.

The initial look at arterial data quality occurred from 2008 through 2010 resulting in an initial white paper to the Coalition summarizing the analysis.  In 2011 and 2012, several targeted data collections were performed that furthered the understanding of probe data on arterials, and whose results were conveyed in a webinar in 2013 to the Coalition.  The work from 2011 to 2012 also resulted in a new method to assess accuracy based on the analysis of repeatable traffic patterns. This method was based on overlaying multiple days of traffic data, typically weekdays, to reinforce recurring congestion patterns and boost the data density to observe the detail in such patterns.   Statistical summaries of the overlay method, specifically Cumulative Frequency Diagrams (CFD), provided direct measures of the statistical travel time distribution.  Starting in 2013, several targeted arterial data collections were planned, building on information learned in the prior two efforts.  This report is a result of the information gained from the targeted data collections which were conducted in 2013 and 2014.  Because the data collection and analysis methodologies used in the 2013 to 2014 effort built on the previous findings, a review of the 2010 white paper and 2013 webinar are provided as background.

*Summary of Initial Analysis – November 2010*
As opportunity allowed from 2008 through 2010, a portion of the Bluetooth traffic monitoring (BTM) sensors were placed on adjoining arterials when collecting reference data for freeway validation. This enabled the University of Maryland (UMD) to begin to explore VPP data fidelity on arterials.   The results of these activities revealed more about the contrast between arterial traffic flow and freeway traffic flow than they revealed about VPP data quality on arterials. Several issues were identified that impacted arterial data quality (both from probe data and BTM reference data) as well as its validation process. The significant findings from this initial analysis included:

- **Interrupted-flow arterials encompass a wide variety of functional road classes** spanning high-volume, multi-lane arterials, all the way down to low-volume local streets. Initial results indicated that probe data was likely to be usable for some applications on the higher class facilities, specifically those that are multi-lane, exhibit sparse signal spacing, and medium to low mid-block friction.

- **Traffic flow on arterials is more diverse than on freeways** (referring specifically to speeds and travel times), resulting in higher speed variation on arterials. Higher speed variance in turn requires a higher sampling rate to attain the same level of data quality or confidence. The primary cause of the higher variation is directly attributable to signal control. On any given segment containing signal control, traffic progresses through a cyclic pattern which includes deceleration, queuing, queue dissipation on green, acceleration and then free-flow, which repeats itself with each signal cycle. Not only do speeds vary considerably along the length of the segment, the travel times experienced by vehicles also vary depending on their position in the traffic stream relative to the signal cycle. Some may make it through the corridor on all green, while other may be forced to wait until the next cycle. Large speed variations are induced primarily by signal control, however increased turning opportunities afforded by mid-block access to goods and services contribute as well.

- In addition to higher overall variance in speeds, **traffic signals on arterials tend to divide traffic into pulsed flows**, with two or more distinct travel times, with the difference between these times equal to the prevailing signal cycle time. Several of the test segments exhibited two distinct travel times in which approximately half the platoon made it through a signalized corridor without stopping, while the other half of the platoon was forced to stop and wait for the green phase to complete the segment. These bi-modal flows are a significant technical challenge to effectively use any type of arterial traffic data, not just probe data.

- **Maximum traffic volume on arterials is generally half that of freeways for the same geometric configuration** (number of lanes). Combining lower volume with the impact of higher variance (which requires larger sample sizes), it was anticipated that VPP probe data would be challenged to provide quality traffic data due to its inherent sampling nature. Similarly, decreased volumes and higher variance also limit opportunities to validate with BTM reference data, as it is also a sampling technology.

- **Congested flow is more difficult to discern from free-flow conditions on arterials than on freeways**. Whereas freeway congestion can be identified with a simple speed threshold, differing travel times occur on arterials at different times of day due to the different signal timing plans in effect. The freeway validation method, which emphasizes performance during congested periods, would be difficult to adapt for arterials.

- **Traffic Message Channel (TMC) codes** (the prevailing industrial standard for roadway segmentation), though adequate for most freeway applications, **may not be adequate for the complexity of arterial networks**. On freeways, TMC segments break at each break in access, whereas on arterials, TMC segment breaks are less consistent, breaking primarily at intersections with major facilities. Signalized intersections with minor facilities are contained in a single TMC. In addition, TMC coverage can be discontinuous.

These finding were summarized in a contrast chart between freeway and arterial traffic flow as shown in Table 1 below.

Table 1. Contrast of Freeway and Arterial characteristics

| | Freeways | Arterials |
|---|---|---|
| Volume | 2200 vphpl[*] | 1400 vphpl[*] on green |
| Speed Range | 20-70 mph | 10-45 mph |
| Freeflow | 65 mph | Unknown |
| Congestion Types | Recurring / Non-recurring | Cycle Failure / Mid-Block Friction |
| Congestion Signature / Incident | Slowdowns $< 55$ mph | Difficult to recognize |
| Flow characteristic | Uniform | Higher Variance, Frequently Bi-Modal |
| Traffic Message Channel (TMC) | Adequate / Breaks at each access | Inadequate / Breaks at only at major facilities |

*Vehicle per hour per lane

With these finding in mind, UMD diverted more validation resources to arterial roadways from 2011 to 2012, targeting arterial roadways identified as *'higher-class facilities, those that are multi-lane, exhibit sparse signal spacing, and have medium to low mid-block friction.'*

## Summary of Second Phase Analysis – January 2013

Beginning in 2010, UMD collected validation data sets on US-1 in Pennsylvania and Virginia, State Route 3 in Maryland, and State Route 13 in Delaware. Using the basic tools developed for freeway validation, the results from these four arterial validations reinforced the previous findings, and began to further clarify when probe data should be considered for use. *Probe data quality on arterial facilities was found to be most correlated to traffic volume and secondarily to signal density in these exercises.* The analysis indicated that probe data began to exhibit consistent quality on arterial roadways once Average Annual Daily Traffic (AADT) surpassed 30,000. These findings were obtained by visual review of graphs that directly plotted BTM reference data against VPP data over a 24 hour period.

*It was also observed that the traditional freeway validation metrics of Average Absolute Speed Error (AASE) and Speed Error Bias (SEB) did not consistently correlate well to the results obtained from careful review of the data.*

**Average Absolute Speed Error (AASE)**
The AASE is defined as the mean absolute value of the difference between the mean
speed reported from the VPP and the reference mean speed for each five minute time
period. The AASE is the primary freeway accuracy metric, and reflects the variation of
the probe data with respect to the reference data. Contract specifications for freeways
allow a maximum AASE of 10 miles per hour (MPH) in each of four speed ranges.
Speed ranges for freeway validations are defined as 0-30 MPH, 30-45 MPH, 45-60 MPH,
and > 60 MPH.

**Speed Error Bias (SEB)**
The SEB is defined as the average speed error (not the absolute value) in each speed bin.
SEB is a measure of whether the speed reported in the VPP consistently under or over
estimates speed as compared to ground truth speed. Contract specifications for freeways
allow a maximum SEB of ± 5 MPH in each speed range as defined previously.

The reason for this discrepancy is related to the high-variance and often multi-modal nature of
arterial traffic data. Within the freeway methodology, the AASE and SEB are measured against
the 95th confidence interval of the mean of the reference data, referred to as the Standard Error of
the Mean (SEM) band. For freeways, the SEM band accounts for uncertainty in the BTM
reference data arising from either low number of samples, or high variance in the reference data.
The SEM band discounts (or de-weights) validation results whenever confidence in the mean
speed as measured by BTM data was low. On freeways the SEM band is relatively narrow much
of the time, usually about two mph on average. There are periods when the SEM band does
become significant. Such periods include:

- Overnight periods when the number of data samples are low and vehicles travel at
  differing speeds according to driver preference

- Occasional traffic incidents in which traffic progresses at distinct and differing speeds
  such as when traffic is diverted around an accident in two distinct paths

- Facilities that have special use lanes such as local and express lanes in which each set of
  lanes travel at distinctly different speeds, creating a large SEM band.

Such time periods in which the SEM becomes significant on freeways is a small fraction of the
overall observations. In contrast, much of the arterial traffic data exhibited high-variance as a
result of the basic nature of the arterial facilities (as previously discussed), as well as frequently
exhibited multi-modal flow as a result of signal control. As a consequence, the SEM band was
frequently wide, yielding measures of AASE and SEB that appeared to fall within the quality
guidelines established for freeways, but did not correlate with the quality as revealed from visual
inspections of data.

In 2012, UMD began to look at case studies from permanent installations of BTM equipment on
US-1 in Northern Virginia available as a result of a federal demonstration grant. The validation
process adapted some statistical techniques inspired from floating car runs typically employed to
assess the before/after impact of signal retiming. These techniques were initially applied to
BTM data on US-1, and led to the use of *overlay plots* and *distribution analyzes* as explained
below as a better method to quantify probe data fidelity

*Overlay plots* are constructed by taking multiple days of observation and graphing them on a
single 24 hour timeline. Typically only data from weekdays are combined, excluding data from

weekends. This method of overlaying data on a single 24-hour plot reinforces repeatable traffic phenomenon, enhancing the density of travel-time samples and thus increasing the visible detail of any recurring congestion.  Figure 1a is an example of an overlay plot for an arterial segment. Each travel time data point collected using BTM equipment on a weekday from an approximate two week period is graphed on a 24 hour timeline.  The relative density of the data provides a visual indication of the probability of traversing the corridor at the travel time indicated on the y-axis.

A corresponding *distribution analysis* is constructed from the data in the overlay plot.  Each curve in a distribution analysis (called a *cumulative frequency diagrams* or CFDs) is constructed from the percentiles of the travel time data in the overlay plot. (CFDs are also known as cumulative distribution functions, or CDFs, in statistical literature).   In the distribution analysis charts used in this validation, a CFD is constructed for each hour from the overlay plot (excluding overnight hours).   For example, the travel times from the overlay plot for the morning rush hour period from 8 AM to 9 AM are used to calculate the 5th, 10th, 15th, 20th, ….. 95th percentile travel times.  The plot of those percentiles becomes the CFD for the 8 AM to 9 AM morning rush hour.  Figure 1b is an example of a distribution analysis of the data in Figure 1a.  Multiple CFDs from multiple one hour time periods throughout the day are plotted in the distribution analysis, creating an ensemble of curves.   The CFD for the evening rush hour from 5 PM to 6 PM is plotted in black to highlight its CFD in contrast to the other hours of the day.
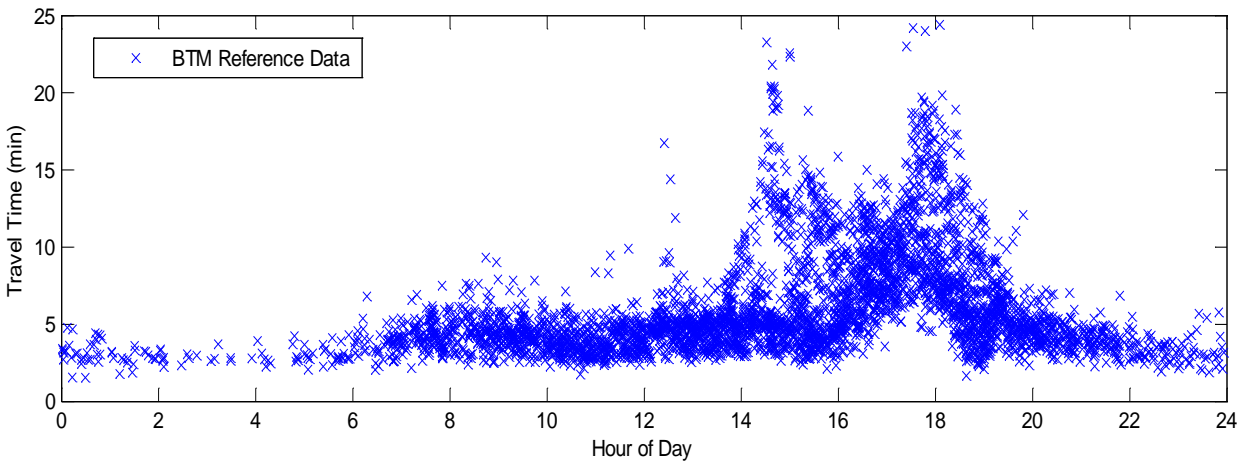


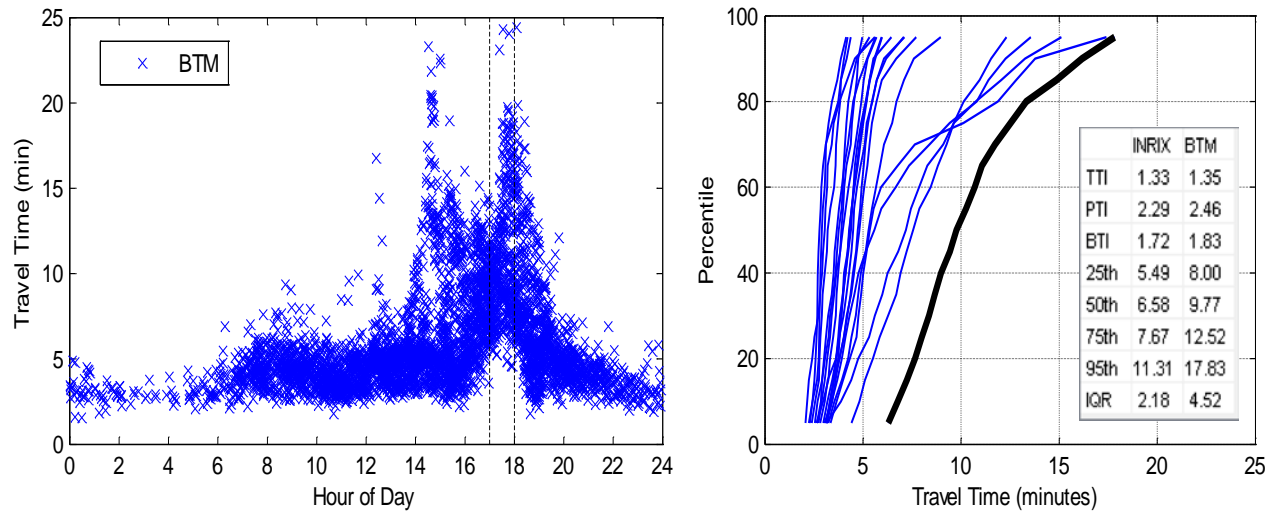Figure 1a. Sample 24 hour weekday overlay plot of BTM data.

Figure 1b. Sample CFD plot of hourly traffic distribution patterns.

This overlay approach and distribution analysis overcame many of the issues related to insufficient sample size, high-variance, and multi-modal data. *Whereas the strength of the traditional analysis is to assess point-in-time performance, the strength of the distribution method is to assess the ability of the traffic data to capture repeatable patterns of weekday traffic.*

The CFD plots also provide a method to directly calculate many common travel-time and travel-time reliability measures. As a result, common performance measures such as the travel time index (TTI), buffer time index (BTI), planning time index (PTI), medians, etc., could be directly calculated and compared for peak periods between probe data and BTM reference data.

The second initiative concluded in early 2013 with an attempt to quantify the attributes which determine whether or not VPP probe data could be effectively used. Table 2 below was created primarily based on the results from the second initiative, and augmented with early observations from the current initiative. It summarizes expectations regarding the fidelity of probe data based on a number of geometric attributes. This table also provided an expectation of 'predicted performance' for each corridor in the case studies from 2013 to 2014 which was then contrasted against observed performance.

Table 2. Arterial Probe Data Usability

| Principal Arterials | Minor Arterials | Major Collectors |
|---|---|---|
| ● AADT > 40,000 vpd (2-way) | ● AADT 20K to 40K vpd (2-way) | ● AADT < 20K (2-way) - low volume |
| ● 2+ lanes per direction | ● 2+ lanes per direction | ● <= 2 lanes per direction |
| ● <= 1 signal per mile | ● <= 2 signals per mile | ● >= 2 signals per mile |
| ● Limited curb cuts | ● Moderate number of curb cuts | ● Substantial number of curb cuts |
| **Likely to have accurate probe data…** | **Possibly accurate probe data…** | **Unlikely probe data is accurate…** |
| ✓RECOMMENDED | 🔍SHOULD BE TESTED | ✖NOT RECOMMENDED |

Having established workable tools, as well as having exhaustively sampled and verified freeway data fidelity, the Coalition directed much of its validation efforts in 2013 and early 2014 to sampling a variety of arterial corridors to further quantify the value of probe data for various functional classes of arterial roadways.

## 2.0 Case Study Locations

From April 2013 through June 2014, the UMD deployed BTM data collection equipment on nine separate occasions, covering a variety of arterial roadways from North Carolina to New Jersey. Table 3 below lists the data collection trials and the roadways covered. These nine test locations were chosen in consultation with the Coalition member having jurisdiction over the roadway. Sites were chosen that were of interest to the Coalition member, anticipated to provide samples of congested and uncongested traffic flow, and had geometric attributes within the bounds listed in Table 2. In general, most roadways that were selected were multi-lane facilities with an AADT of 20,000 or greater. A minor portion of roadway segments within some study areas fell below this threshold. For each roadway (more specifically, each segment of each roadway), the geometric and traffic attributes were documented including:

- Traffic volume as reported by the Highway Performance Management System (HPMS)

- Posted speed limit

- Signal density

- Access density (curb cuts, driveways, and crossroads)

- Number of lanes

- Road geometrics (median barriers, turn lanes, etc.)

Each roadway corridor chosen for data collection provides a case study for the effectiveness of VPP data. This approach creates a cross reference for similar facilities. The performance of VPP data on any given arterial roadway can be estimated by the performance of similar roadways from this validation exercise with respect to volume, speed limit, signal density, and other geometric attributes.

## Table 3. Case Study Locations and Attributes

| Case Study Number | Data Set (State-ID#) | Road Number | Road Name | Validation Date Span | # of Segments | # of Through Lanes | AADT Min-Max / Weighted Average[1] (in 1000s) | Length[2] (mile) | # Signals / Density[3] | # of Access Points | Median Barrier | Speed Limit (mph) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | NC-06 | NC-55 | Williams St, Apex Hwy. | Apr 30-May 13, 2013 | 18 | 1-3 | 15-43/25 | 30.3 | 62 / 2.05 | 231 | Partial | 35-50 |
| 2 | MD-07 | MD-355 | Wisconsin Ave, Rockville Pike, Hungerford Dr, Frederick Rd | July 6-20, 2013 | 10 | 2-4 | 32-67/44 | 17.1 | 67 / 3.9 | 221 | Partial | 30-45 |
| | | MD-586 | Veirs Mill Rd | | 6 | 2-3 | 21-43/34 | 6.2 | 19 / 3.1 | 56 | Yes | 30-45 |
| 3 | NJ-11 | US-1 | Trenton Fwy, Brunswick Pike | Sep 10 - 24, 2013 | 10 | 2-4 | 33 – 90/70 | 14.2 | 10 / 0.7 | 112 | Yes | 55 |
| | | NJ-42 | Black Horse Pike | | 8 | 2 | 25-54/48 | 12.5 | 23 / 1.8 | 260 | Yes | 45-50 |
| | | US-130 | Burlington Pike | | 10 | 3 | 42-42/42 | 14.3 | 28 / 2.0 | 229 | Yes | 50 |
| 4 | NJ-12 | NJ-38 | Kaighn Ave. | Nov 5-19, 2013 | 16 | 2-4 | 32-80/46 | 24.5 | 44 / 1.8 | 235 | Yes | 50 |
| | | NJ-73 | Palmyra Bridge Rd. | | 18 | 2-4 | 33-74/52 | 23.9 | 41 / 1.7 | 236 | Yes | 45-55 |
| 5 | PA-05 | US-1 | Lincoln Highway | Dec 3 - 14, 2013 | 28 | 2 - 3+3 | 21 – 100/45 | 30.6 | 107 / 3.5 | 178 | Yes | 40 - 50 |
| | | US-322 | Conchester Highway | | 6 | 1-2 | 22 – 34/25 | 14.3 | 7 / 0.5 | 48 | No | 35 - 45 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | PA-06 | PA-611 | Easton Rd Old York Rd | Jan 9 - 22, 2014 | 14 | 1-4 | 18-32/27 | 20.5 | 68/ 3.3 | 227 | Partial | 15-45 |
| | | PA-611 | N Broad St | | 12 | 2-4 | 16-30/21 | 12.5 | 144/ 11.5 | 251 | Partial | 25-40 |
| 7 | VA-07 | VA-7 | Leesburg Pike and Harry Byrd Hwy | April 5-16, 2014 | 30 | 2-4 | 45-60/56 | 30.5 | 57 / 1.9 | 203 | Yes | 35-55 |
| | | US-29 | Lee Hwy (S Washington St) | | 4 | 2 | 14-25/21 | 4.4 | 22 / 5.0 | 114 | Partial | 30 |
| 8 | VA-08 | US-29 | Lee Hwy | May 8-19, 2014 | 26 | 2-4 | 15-45/33 | 31.9 | 115/3.6 | 287 | Partial | 35-50 |
| 9 | MD-08 | MD-140 | Reistertown Rd | Jun 5-17, 2014 | 20 | 1 - 3 | 19-44/31 | 17.4 | 68 / 3.9 | 221 | No | 30-40 |
| | | | Baltimore Blvd | | 8 | 2 - 3 | 40-53/42 | 15.5 | 18/ 1.2 | 52 | Partial | 50-55 |

[1]AADT weighted average is weighted by length of the segments

[2]Uni-directional mileage, a 2 mile roadway in which both direction of travel are analyzed will result in 4 total miles of tested roadway.

[3]Signal density is number of signals per mile

# 3.0 Individual Case Study Analysis Methodology

A full analysis was performed resulting in individual reports with a summary of findings for each case study listed in Table 3. A Web Portal has been created that contains all the analytical results for each roadway segment of each case study. The Web Portal can be accessed at [https://app.box.com/I95-ArterialValidationArchive ]. The portal is organized hierarchically in directories by case study name (i.e. NJ-11, NJ-12, etc.).

For each case study, the analysis is organized into four sections: (1) *Roadway Attributes*, (2) *Traditional Validation*, (3) *Sampled Distribution Method,* and (4) *Discussion of Results* as explained in further detail later in this section. A standard naming scheme is used for each segment in the report. A segment is referenced by the case study name (which is a state abbreviation followed by an ID#), and then the segment within that case study. For example, NJ11-03 refers to the 3rd segment from case study NJ11. This naming scheme is consistent in each case study as well as this final summary report. [Note that in some material, the segment number may be augmented with leading 0's, such as NJ11-0003, in the example given.] The assignment of segment numbers within a case study is typically sequential along the direction of travel. Exceptions occur. For example, NJ11 begins with segment 03 rather than segment 01. Such exceptions typically indicate a field data collection issue that prevented analysis of a certain segment, and thus some segments were omitted from reporting.

Each individual case study report provides representative samples from the various analysis methodologies. Along with the individual case study report available in the Web Portal, the full analysis results from each type of analysis are included. These analysis reports include:

- **Traditional Analysis Report**: This document conveys results of the initial freeway-centric validation process. This report contains segment by segment results for the AASE and SEB comparisons between the VPP data and the BTM reference data collection.

- **Traditional Graphs**: This series of folders contains the daily 24 hour graphs for each segment, for each day, contrasting the VPP data with the BTM reference data at five minute intervals. Each validation segment has a sub-folder containing a comparative graph for each day of the validation. Representative graphs in the case study report, and even in this overall summary are drawn from these archives.

- **Distribution Analysis**: This folder consists of graphs that compare BTM and VPP data using the distribution analysis methodology for each segment. The first set of graphs consists of 24 hour overlay plots for all days (weekends and weekdays combined), weekdays only, and weekends only similar in format to Figure 1a. The second set of graphs presents overlay scatter plots and cumulative frequency diagrams (CFDs) highlighting each hour of the day similar in format to Figure 1b. The distribution analysis graphs containing the CFDs are provided only for weekdays.

At any time the reader can access the Web Portal to obtain a more detailed analysis for any segment or corridor.

Each individual case study report summarizes the analysis results of each case study, and is organized into sections as follows:

**(1) Roadway Attributes:**     This is a detailed description of the individual segments comprising the case study. In contrast to freeways, the expected or nominal operation of each

arterial segment is unique and dependent on a number of variables. Understanding the segment and where it falls within the throughput / land access (mobility/ accessibility) spectrum is critical when reviewing validation results. The roadway attributes table presents information about roadway geometry, TMC codes, placement of BTM sensors and general description/notes for each segment in each case study. Each roadway attribute table includes the following data items for each segment.

- **SEGMENT / (Map Link) -** Segment number linked to Google Maps for convenience of viewing.

- **Geometric Description**
    - **Crossroads Start/End** – The primary landmark, typically an intersection, that demarks the beginning and end of the segment. The landmark may also be political or geographical in addition to crossroads.

    - **Lanes** - the minimum and maximum number of through lanes observed on the segment obtained from aerial photography.

    - **AADT** - the minimum and maximum AADT reported as obtained from the publicly released Highway Performance Monitoring System (HPMS) shapefiles.

    - **Signals: Number and Density -**the total number of signalized intersections on the segment as obtained from aerial photography. If the segment begins at a signalized intersection, that intersection is not included in the count (it will be included in the preceding segment). Density is calculated by number of signals divided by segment length.

    - **Number of Access Points** –includes intersections with other roads (other than signalized intersections) and entrances to business centers and malls. Data is obtained from aerial photography. (A driveway to a house is not considered an access point.)

    - **Speed Limit** – posted speed limit as obtained by street view photography.

    - **Median Barrier** – the existence of a median barrier on the segment that prevents left hand turning movements except at designated access areas as obtained by aerial photography.

    - **Number of Major Junctions** - the number of junctions with another roadway of significantly higher class such that the segment acts as a feeder/distributer to another roadway at that junction. Criteria for a major junction:
        - If crossroad is a freeway
        - If crossroad intersection is a grade separated interchange and the number of through lanes is the same or greater than the segment
        - If crossroad intersection is signalized, the number of through lanes of the cross road is greater than the segment

- **TMC Codes**
    - **Begin//End** – the beginning and ending Traffic Message Channel (TMC) code for the segment.

- o **Length -** the combined length of all TMCs for the segment in miles.

- o **Number (#)** - total number of TMCs that exist on the segment.

- **Bluetooth Data Sensor**

  - o **Begin//End** – the begin/end BTM sensor identifier.

  - o **Length –** the distance between the upstream and downstream BTM sensor.

  - o **% Diff –**the percentage difference between the TMC length and BTM sensor length.

- **General Description –** this section provides a general description and notes about the segment.

**(2) Traditional Validation:** The second section of each case study report provides a summary of the traditional validation process originally established for freeways and adapted for arterials based on AASE and SEB by speed category. The methodology used on arterials follows the process developed for freeways as documented in the Corridor Coalition's archive of freeway validations. (Available at http://www.i95coalition.org/wp-content/uploads/2015/02/I-95-CC-Final-Report-Jan-28-2009.pdf?5a9c76 ) The speed categories used for arterial analysis differ from those used for freeways. Speed bins are adjusted to appropriate levels for the facility. Typical speed ranges for arterials are 0-15 mph, 15-30 mph, 30-45 mph, and >45 mph, though the actual speed ranges will differ based on the facility. Apart from the speed bins, the analysis methodology remains the same as that for freeways. A full traditional validation report is available for each case study providing segment specific as well as overall corridor analysis. The traditional report for each case study is made available to the reader via the Web Portal.

A by-product of the traditional analysis are daily 24-hour graphs contrasting the VPP data with the BTM reference data and its five minute mean and standard error of the mean (SEM) limits. Representative samples of these graphs are frequently exhibited in the individual case study reports to exemplify issues observed with data quality. A sample 24-hour data plot is shown in Figure 2 from a principal arterial facility. Each individual BTM reference datum is depicted as a blue 'x', and reflects the speed of a single vehicle based on its measured travel time. In aggregate, these observations form a cloud of data indicating the ranges of speeds directly observed on the test segment for any given time. The BTM reference data are shown in contrast to the average speed reported by the VPP data averaged into five-minute intervals. The VPP five-minute speed data are depicted with red diamonds. The agreement between the BTM reference data shown in blue and the VPP data in red forms the basis of comparison.

Other attributes visible on the 24-hour graph include BTM outliers, and the mean and SEM band of the BTM reference data. If a BTM data point is determined to be an outlier based on a statistical test, it is demarked with a black '**.**' overlaid on the blue 'x'. The majority of BTM outliers are created by vehicles that take significantly longer time to traverse the test segment than other vehicles. These BTM data points typically appear as speed measures between zero and five miles an hour, indicating that the vehicle stopped for services within the test segment such as fuel or food.

The mean of the BTM reference data (excluding outliers) is depicted as a solid black line, and the standard error of the mean (SEM) band is plotted using black dashed lines, '---', about the mean. This SEM band reflects the 95th percent confidence interval about the mean of the BTM

reference data. A narrow SEM band reflects low variance in the BTM speed data, whereas a wide SEM band reflects significant variance in the BTM reference data.

Each graph is for a single day (generally 24 hours in length) within the data collection period. The plot reflecting the first and last day of data collection may be less than 24 hours depending on the time when the BTM sensor placed and collected.
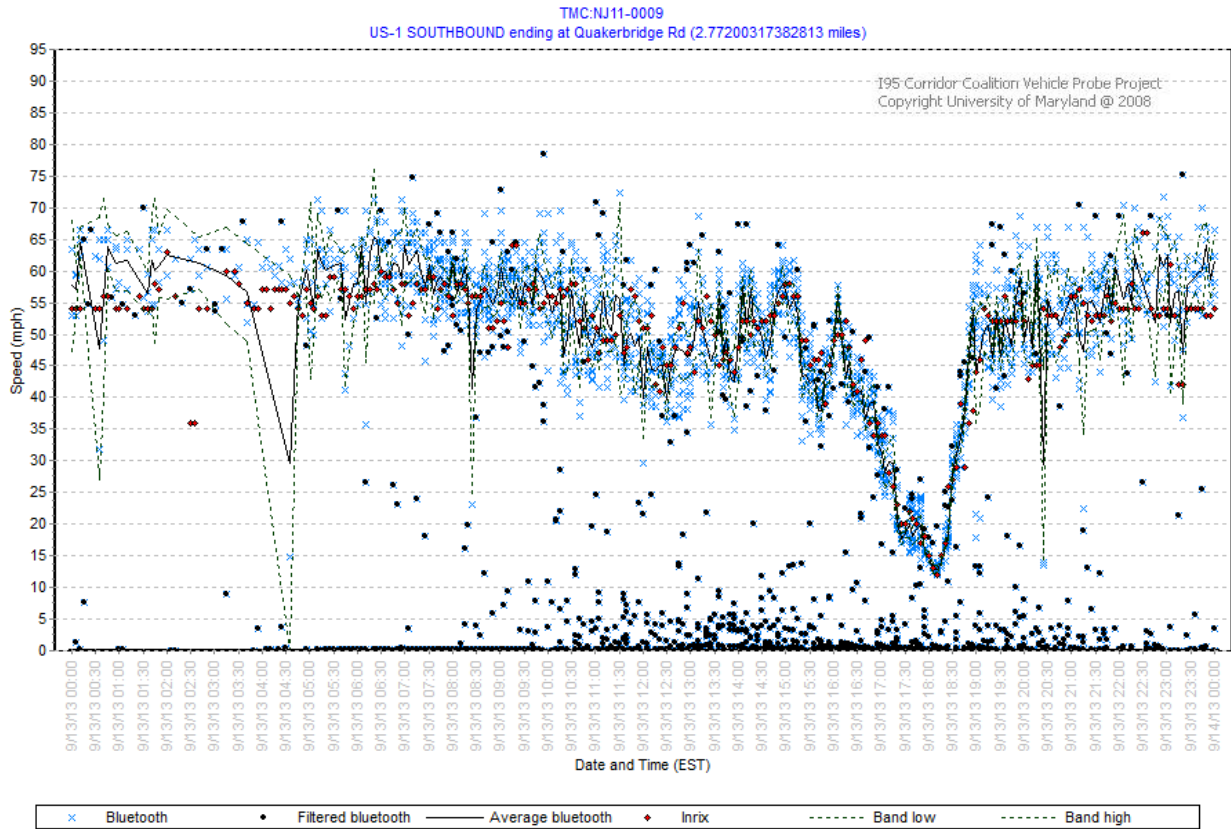


Figure 2. Sample daily plot of the traditional analysis for a freeway highlighting plotting convention

A sample 24-hour graph from a lesser arterial roadway is shown in Figure 3. In contrast to the principal arterial shown in Figure 2, the lesser arterial has fewer reference BTM data points. Also, the SEM band is significantly wider owning to a larger variance of observed reference speed which is typical as a roadway provides increased accessibility. Such graphs are used extensively to illustrate data quality issues in the individual case study reports. A complete archive of 24-graphs for each segment for each case study is available through the Web Portal.
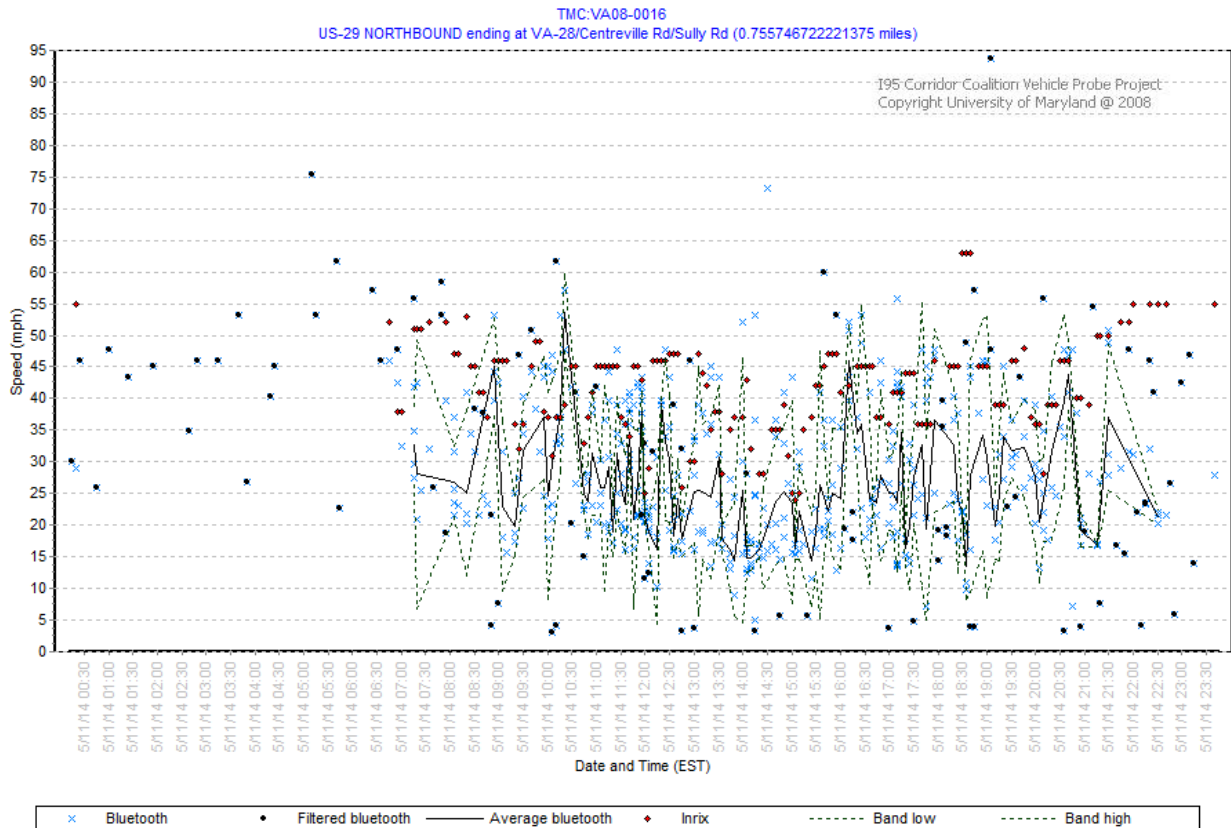
Figure 3. Sample daily plot of the traditional analysis on a minor arterial

The culmination of the traditional analysis is tabulated measures of AASE and SEB for individual segments as well as for the corridor as a whole.  A sample of the corridor level tabulation is shown in Table 4 for the US-1 principal arterial illustrated in Figure 2.  Table 5 provides the traditional results for the lesser arterial from Figure 3.  Corridor results are included in each case study report.  Individual segment results are available in the traditional report accessible via the Web Portal.

Table 4. Sample of the AASE and SEB Results on a Principal Arterial Corridor

| Speed Bin | Average Absolute Speed Error (<10mph) | | Speed Error Bias (<5mph) | | Number of 5 Minute Samples | Hours of Data Collection |
|---|---|---|---|---|---|---|
| | Comparison with SEM Band | Comparison with Mean | Comparison with SEM Band | Comparison with Mean | | |
| 0-15 MPH | 2.9 | 4.4 | 2.8 | 3.8 | 224 | 18.7 |
| 15-25 MPH | 5.3 | 7.3 | 5.2 | 6.9 | 1742 | 145.2 |
| 25-35 MPH | 5.4 | 9.6 | 5.2 | 8.8 | 3155 | 262.9 |
| >35 MPH | 2.3 | 6.5 | -1.3 | -2.9 | 21276 | 1773.0 |
| All Speeds | 2.9 | 6.9 | -0.1 | -0.8 | 26397 | 2199.8 |

Table 5. Sample of the AASE and SEB Results on a lesser Arterial Corridor

| Speed Bin | Average Absolute Speed Error (<10mph) | | Speed Error Bias (<5mph) | | Number of 5 Minute Samples | Hours of Data Collection |
|---|---|---|---|---|---|---|
| | Comparison with SEM Band | Comparison with Mean | Comparison with SEM Band | Comparison with Mean | | |
| 0-15 MPH | 7.2 | 11.9 | 7.1 | 10.7 | 1540 | 128.3 |
| 15-25 MPH | 3.8 | 7.4 | 3.7 | 7.4 | 4124 | 343.7 |
| 25-35 MPH | 1.3 | 6.8 | 0.5 | 1.8 | 5060 | 421.7 |
| >35 MPH | 2.5 | 6.2 | -1.2 | -3.9 | 6828 | 569.0 |
| All Speeds | 2.5 | 6.2 | -1.1 | -3.7 | 17552 | 1462.7 |

**(2) Slowdown Analysis Method**: The slowdown analysis is an offshoot of the traditional analysis, developed to provide a more intuitive measure of probe data's ability to capture congestion events. The slowdown analysis is effective in quantifying the ability of probe data to capture significant disruptions in traffic. The definition of a slowdown is when traffic speed reduces by at least 15 mph for a period of one hour or more. On slower speed arterials, the threshold may be reduced to a reduction in speed of 10 mph, and the duration of 30 minutes or greater. For each observed slowdown in each 24-hour data plot, the analyst rates the

performance of the probe data base on the reported speed reduction and duration on slowdown. Each slowdown is classified as either 'Fully Captured', 'Partially Captured', or 'Failed to Capture' as explained below.

- A *Fully Captured* slowdown indicates that the probe data accurately characterized both the reduction in speed, and duration of the slowdown. The error in speed reduction or duration cannot exceed 20%. An example of a fully captured slowdown is shown below.
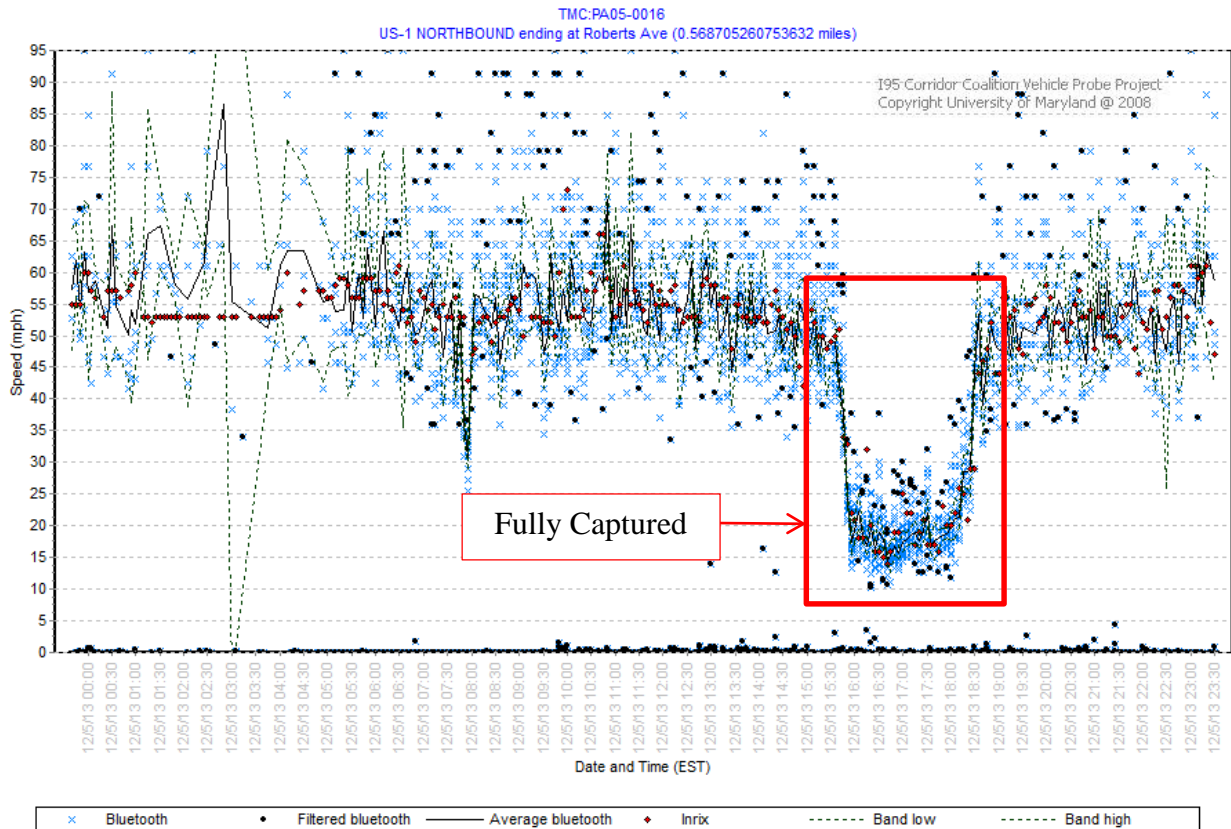


Figure 4. An example of a fully captured slowdown

- A *Partially Captured* slowdown indicates that the probe data reported a significant disruption to traffic, but the extent of speed reduction or duration of time were in error by more than 20%. An example of partially captured slowdown is shown in Figure 5.
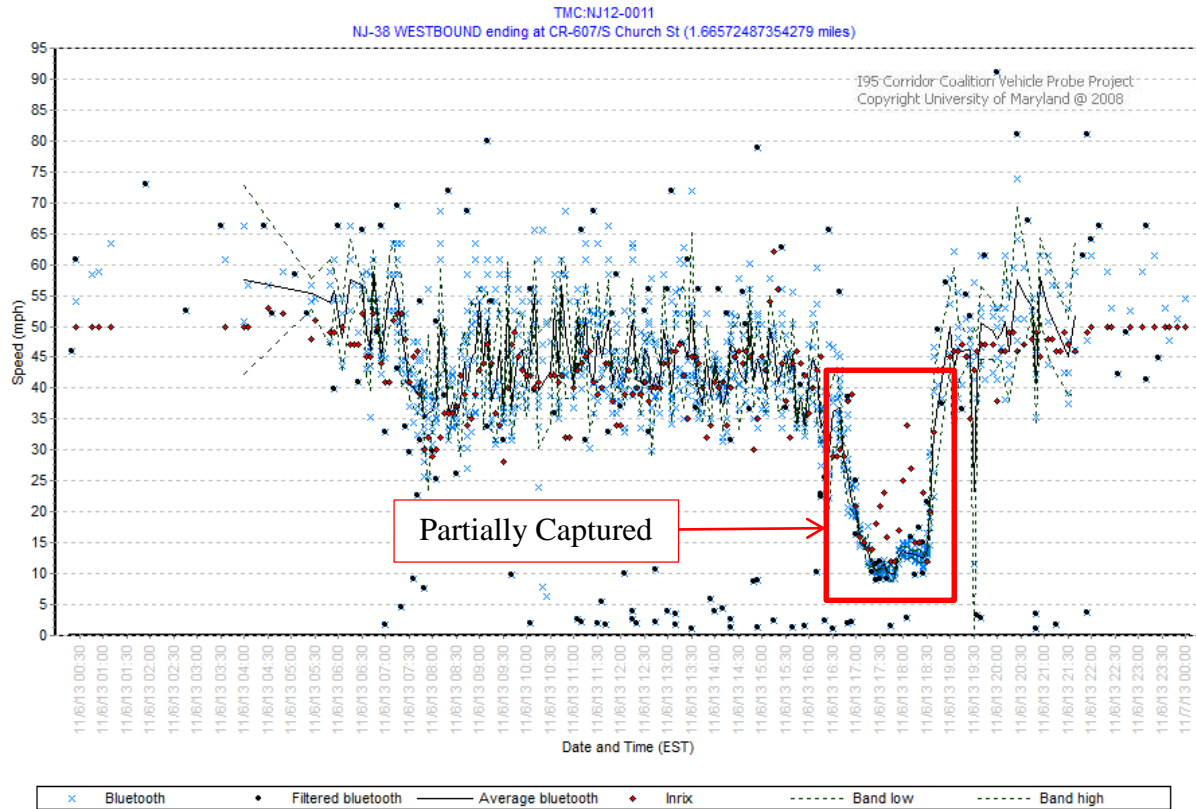
Figure 5. An example of a partially captured slowdown

- *Failed to Capture* indicates that the probe data either completely missed the slowdown, or the extent of speed reduction or duration of the event were significant in error such that the slowdown would not be interpreted as a significant disruption to traffic. An example of failed to captured slowdowns are shown in Figure 6.
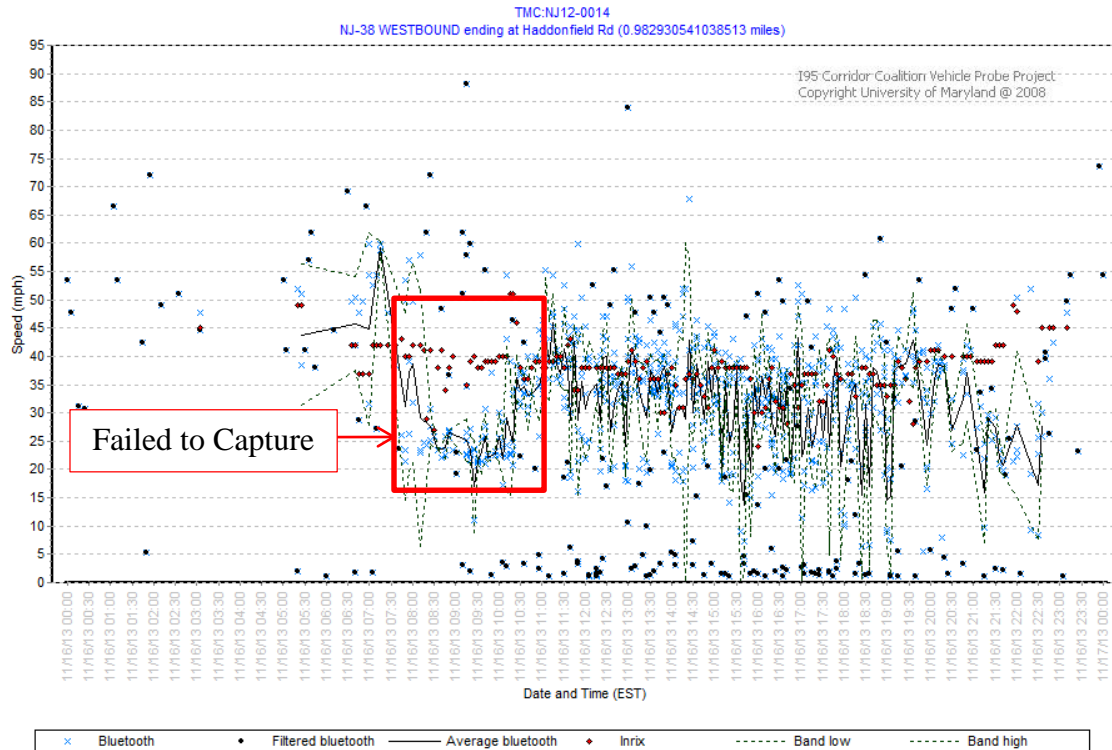
Figure 6. An example of a failed to capture slowdown

**(3) Sampled Distribution Method:** The third section of each case study report contrasts VPP and BTM data using sampled distribution methods based on 24-hour overlay plots and CFDs as previously described. Overlay and CFD plots for representative segments are provided for each case study to illustrate findings. A full listing of the overlay and CFD charts for each segment of each case study are provided via the Web Portal. Figure 7 and Figure 8 provide samples of the overlay and CFD plots respectively comparing VPP data with BTM reference data. The CFD plot in Figure 8 highlights the 8 AM to 9 AM peak travel hour.
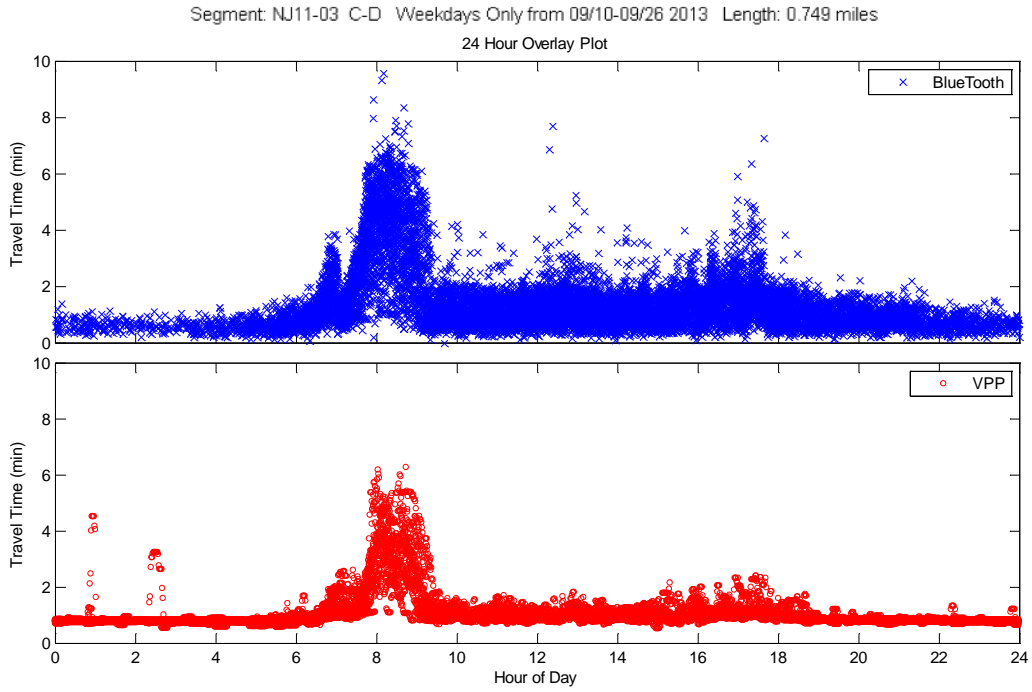
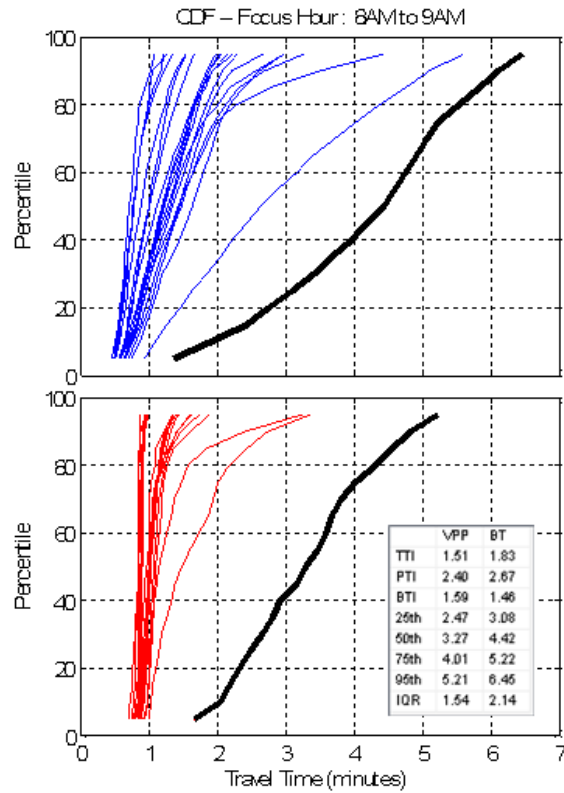Figure 7. Sample BTM (Top) and VPP (Bottom) 24-hour overlay plot



Figure 8. Sample BTM (Top) and VPP (Bottom) CFD Diagrams

In addition to visual comparisons, standard performance measures are calculated for the highlighted hour as shown in Figure 8. These hourly performance measures include:

- **Travel Time Index (TTI)** - ratio of the travel time central tendency (the median travel time is used) to the free flow travel time (estimated using the 15th percentile travel time)

- **Planning Time Index (PTI)** - ratio of the 95th percentile travel time to the free flow travel time (15th percentile travel time)

- **Buffer Time Index (BTI)** - difference between the 95th percentile travel time and travel time central tendency (the median travel time is used) divided by central tendency (the median travel time is used)

- **25th, 50th, 75th and 95th Percentiles** – percentiles directly calculated from the travel time distribution

- **Interquartile Range (IQR)** - difference between 75th and 25th percentile travel time

**(4) Discussion of Results:** The final section of each case study report provides a summary of the significant findings, and an assessment of the usability of the probe data for various performance measure applications.

# 4.0  Summary of Case Study Results

The analysis methods described in the previous section were applied to all the arterial corridors within the nine case studies listed in Table 3 to assess the effectiveness of VPP data across a broad sample of arterial roadways.  This section provides an overall summary and emphasizes correlations between arterial attributes and anticipated probe data quality.  The archive of graphs, data, tables, and individual reports available through the Web Portal provide the supporting data.

Although based on an analyst review rather than statistical compilations, the slowdown analysis provided the best insight on the fidelity of outsourced probe data for operations and performance measures.  The traditional method based on AASE and SEB provided meaningful insights, and the distribution method was effective if the corridor exhibited recurring congestion patterns.  However, the slowdown analysis provided the best overall gage of quality provided that there were a sufficient number of traffic disruptions over the time period of the case study.

Table 6 shows the percent of *fully captured*, *partially captured* and *failed to capture* slowdowns for each corridor within the nine case studies.  Table 6 also contains the roadway attributes which correlated best with the results of the analysis.  Note that the results from some case studies were omitted in the summary table if deemed insignificant. For example, freeway portions of PA-05, designated corridor code 5b, were excluded from the results as it is not an arterial roadway.

Table 6. Slowdown Analysis Results

| Case # | Data Set ID | Corridor Code | Road Name | Roadway Attributes | | | Slowdown Analysis | | | |
| | | | | AADT (in 1000s) | Signal Density | Speed Limit (mph) | Number of Slowdowns | Percent (%) | | |
| | | | | | | | | FULLY CAPTURED | PARTIALLY CAPTURED | FAILED TO CAPTURE |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | NC-06 | 1 | NC-55 | 25 | 2.05 | 35-50 | 54 | 15.0% | 44.0% | 41.0% |
| 2 | MD-07 | 2a | MD-355 | 44 | 3.9 | 30-45 | 9 | 44.4% | 22.2% | 33.3% |
| | | 2b | MD-586 | 34 | 3.1 | 30-45 | 8 | 0.0% | 37.5% | 62.5% |
| 3 | NJ-11 | 3a | US-1 | 70 | 0.7 | 55 | 101 | 63.4% | 36.6% | 0.0% |
| | | 3b | NJ-42 | 48 | 1.8 | 45-50 | 4 | 0.0% | 100.0% | 0.0% |
| | | 3c | US-130 | 42 | 2.0 | 50 | 4 | 25.0% | 50.0% | 25.0% |
| 4 | NJ-12 | 4a | NJ-38 | 46 | 1.8 | 50 | 57 | 40.4% | 38.6% | 21.1% |
| | | 4b | NJ-73 | 52 | 1.7 | 45-55 | 89 | 41.6% | 46.1% | 12.4% |
| 5 | PA-05 | 5a | US-1(a) | 45 | 3.5 | 40 - 50 | 78 | 28.2% | 48.7% | 23.1% |
| | | 5c | US-322 | 25 | 0.5 | 35 - 45 | 58 | 50.0% | 41.4% | 8.6% |
| 6 | PA-06 | 6a | PA-611 | 27 | 3.3 | 15-45 | 18 | 22.2% | 33.3% | 44.4% |
| | | 6b | PA-611 | 21 | 11.5 | 25-40 | 5 | 0.0% | 20.0% | 80.0% |
| 7 | VA-07 | 7a | VA-7(a) | 56 | 1.9 | 35-55 | 75 | 24.0% | 42.7% | 33.3% |
| | | 7b | VA-7(b) | 55 | 1.6 | 45-55 | 22 | 4.5% | 22.7% | 72.7% |
| | | 7c | US-29 | 21 | 5.0 | 30 | 1 | 0.0% | 0.0% | 100.0% |
| 8 | VA-08 | 8 | US-29 | 33 | 3.6 | 35-50 | 49 | 8.2% | 42.9% | 49.0% |
| 9 | MD-08 | 9a | MD-140(a) | 31 | 3.9 | 30-40 | 20 | 0.0% | 35.0% | 65.0% |
| | | 9b | MD-140(b) | 42 | 1.2 | 50-55 | 18 | 22.2% | 66.7% | 11.1% |

Table 7 shows results of traditional analysis yielding average AASE and SEB for each corridor under investigation.  Each AASE and SEB score is comprised of two numbers separated by a

slash ('/').  The number preceding the slash is the metric assessed against the mean of the BTM data.  The number after the slash is the same metric assessed against the SEB band.  Much of the arterial traffic data had high-variance resulting in a wide SEM band, which in turn tended to mask actual performance.  For this reason, assessing AASE and SEB against the mean rather than the SEM band is recommended on arterials.  The Vehicle Probe Project I contract specified a maximum of 10 MPH AASE and within ±5 MPH for SEB.  These specifications were based on the application needs for freeways, and are presented as reference benchmarks as no arterial specifications had been established.

Table 7. AASE and SEB (Comparison with Mean / Comparison with SEM Band)

| Corridor Code | Comparison with Mean / Comparison with SEM Band | | | | | | | |
| | Average Absolute Speed Error< 10 MPH - VPPI Spec. | | | | Speed Error Bias Within ±5 MPH - VPPI Spec | | | |
| | Speed Bin | | | | Speed Bin | | | |
| | 0-15 MPH | 15-25 MPH | 25-35 MPH | >35 MPH | 0-15 MPH | 15-25 MPH | 25-35 MPH | >35 MPH |
| 1 | 13.9/9.6 | 10.9/3.4 | 5.6/1 | 5.1/1.3 | 13.7/9.6 | 10.5/3.3 | 3.8/0.7 | -3.4/-1.1 |
| 2a | 12.7/6.4 | 7.3/3 | 3.7/0.9 | 12.5/6.3 | 12.7/6.4 | 7/2.9 | 1.3/0.3 | -12.4/-6.2 |
| 2b | 13.5/7.5 | 8.6/3.4 | 4.7/1.2 | 7.8/2.9 | 13.4/7.5 | 8.2/3.3 | 2.3/0.8 | -7.3/-2.8 |
| 3a | 4.4/2.9 | 7.3/5.3 | 9.6/5.4 | 6.5/2.3 | 3.8/2.8 | 6.9/5.2 | 8.8/5.2 | -2.9/-1.3 |
| 3b | 13.4/7.4 | 10.3/3.6 | 6.4/1.3 | 5.6/1.8 | 13.4/7.4 | 10.1/3.6 | 5.3/1.1 | -4/-1.6 |
| 3c | 19.9/12.2 | 13.8/5.1 | 7.1/2.8 | 4.9/1.5 | 19.9/12.2 | 13.7/5.1 | 6.5/2.6 | -1.4/-0.6 |
| 4a | 12.8/9.5 | 11.8/7.5 | 7.7/3.1 | 4.9/1.5 | 12.8/9.5 | 11.7/7.5 | 7.2/3 | -1.2/-0.6 |
| 4b | 7/4.7 | 9/4.1 | 7.5/3.5 | 5.2/1.8 | 7/4.7 | 8.8/4 | 6.4/3.2 | 0/-0.1 |
| 5a | 11.5/7.5 | 8.4/4.7 | 5.7/1.9 | 4.9/1.3 | 11.3/7.5 | 8.2/4.7 | 4.6/1.7 | -2.7/-0.9 |
| 5c | 8.2/5.9 | 8.3/4.3 | 6.7/2.9 | 3.1/1.3 | 8.1/5.9 | 8.2/4.3 | 6.3/2.7 | -0.1/0 |
| 6a | 9.1/4.9 | 6.1/2.9 | 4.1/1.1 | 6.1/2.3 | 9.1/4.9 | 5.7/2.8 | 1.7/0.6 | -4.4/-1.7 |
| 6b | 6.5/3.3 | 3.4/1.3 | 5.3/2.1 | 12.9/5.5 | 6.4/3.3 | 1.5/0.8 | -4.9/-2 | -12.9/-5.5 |
| 7a | 10.8/7.5 | 7.8/4.5 | 6.7/2.6 | 6.3/2.2 | 10.7/7.5 | 7.3/4.3 | 4.8/2.1 | -3.5/-1.6 |
| 7b | 19.3/18.3 | 13.8/10.8 | 12.4/6.6 | 5.2/1.7 | 19.1/18.2 | 13.4/10.6 | 12/6.5 | -0.9/-0.6 |
| 7c | 8.1/3.9 | 3.7/1.2 | 2.5/0.6 | 8.7/5.2 | 8/3.8 | 2.8/1 | -1.1/-0.5 | -8.7/-5.2 |
| 8 | 11.9/7.2 | 7.4/3.8 | 6.8/1.3 | 6.2/2.5 | 10.7/7.1 | 7.4/3.7 | 1.8/0.5 | -3.9/-1.2 |
| 9a | 9.7/5.8 | 5.1/1.9 | 4.1/1.1 | 9.2/3.8 | 9.6/5.8 | 4.2/1.8 | -1.5/-0.4 | -9/-3.8 |
| 9b | 18.9/16.9 | 13.1/10.1 | 13.5/6.8 | 6.6/2.9 | 18.9/16.9 | 12.9/10.1 | 12.5/6.2 | -2.9/-1.6 |

Quality of probe data correlated best with density of signals.  Figure 9 is a regression plot of the percentage of *failed to capture* slowdowns versus signal density.  The coefficient of determination (the percent of variation that is attributed to the least square regression on the independent variable), R2, is 0.40 (regression line a).  If corridor 6b is omitted from the regression, because it exerts significant leverage in the regression analysis, the resulting R2 is 0.46 (regression line b).  Figure 10 is a regression plot of percentage of *failed to capture* slowdowns versus AADT, with an R2 of 0.31.  Although quality improves with increasing AADT (as expected), the correlation is not as strong as signal density.
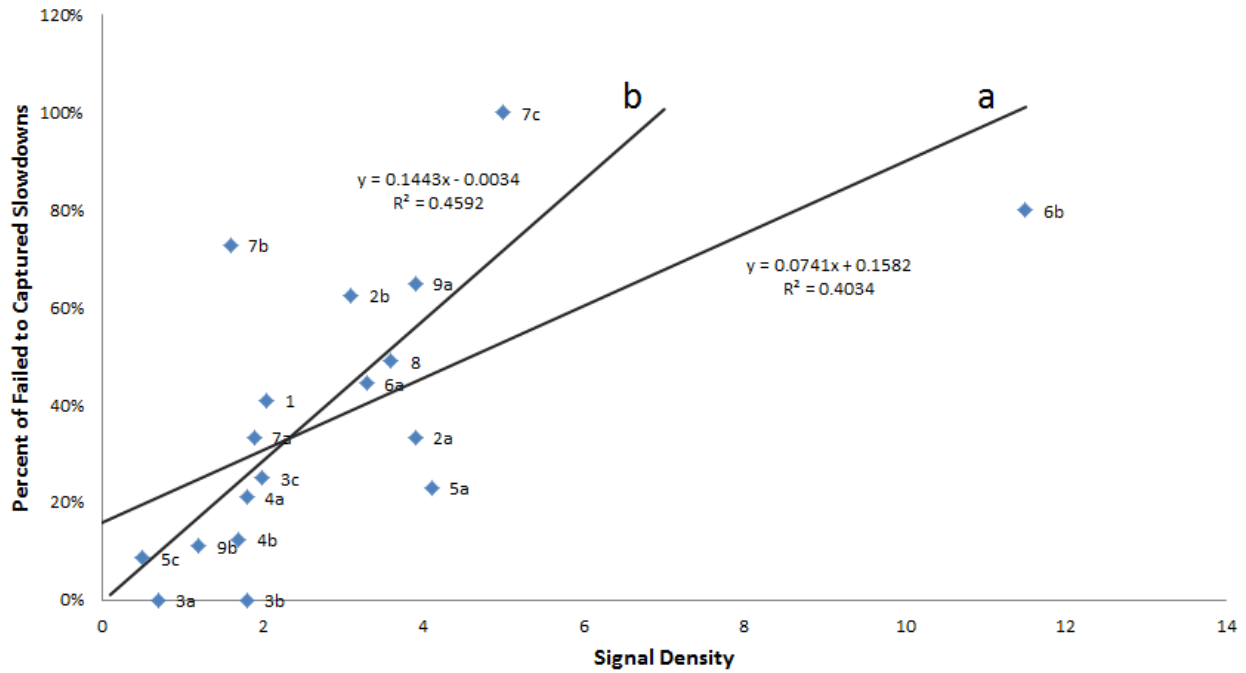


Figure 9.  Percent of failed to capture slowdowns versus signal density.
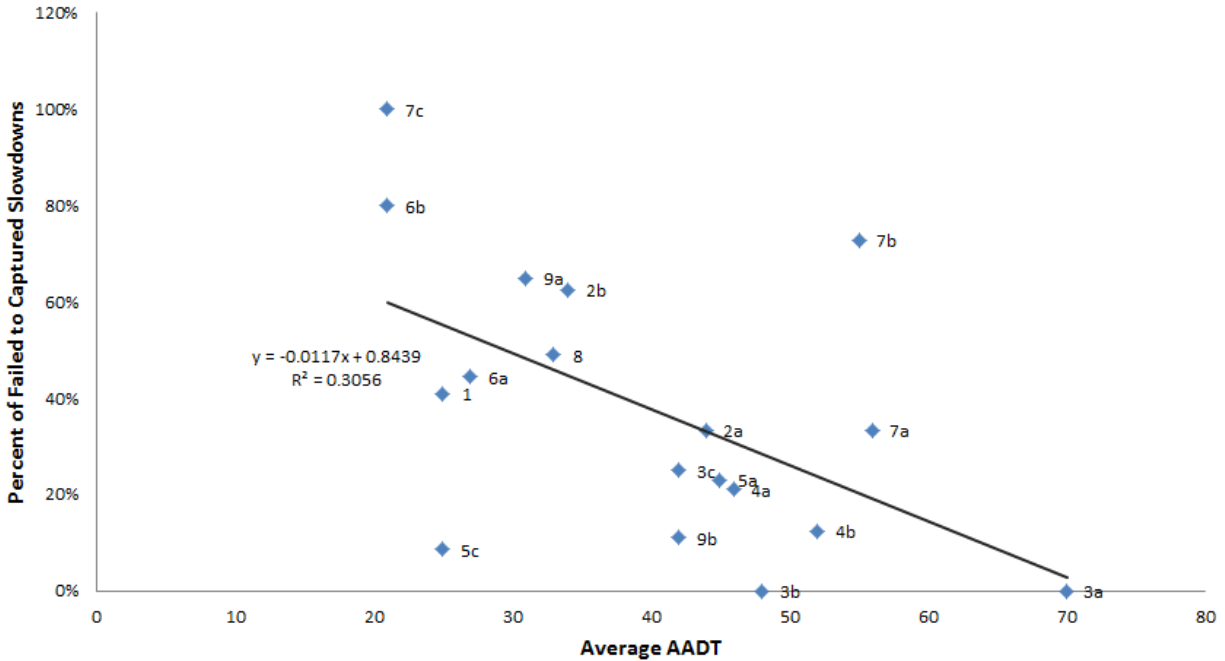
Figure 10. Percent of failed to capture slowdowns versus average AADT.

The sampled distribution method provided additional insight when corridors exhibited recurring congestion. On some segments, the VPP data reflected the patterns in the BTM reference data, an example of which is displayed in Figure 11a for segment VA08-14. The VPP travel time captures the recurrent congestion patterns during peak hours, but underestimates its magnitude. The median travel time during the AM peak period (8 AM to 9 AM) is 2.47 minutes as measured by VPP versus 3.17 minutes as measured by BTM. Underestimating the magnitude of delay during congested periods was a recurring phenomenon in all case studies. The resulting speed bias has implications if probe data is used for long term performance measures.

On some segments VPP failed to capture the recurring congestion trend. A representative example is shown in Figure 11b from segment MD08-03 in which VPP fails to capture the recurring evening delay.
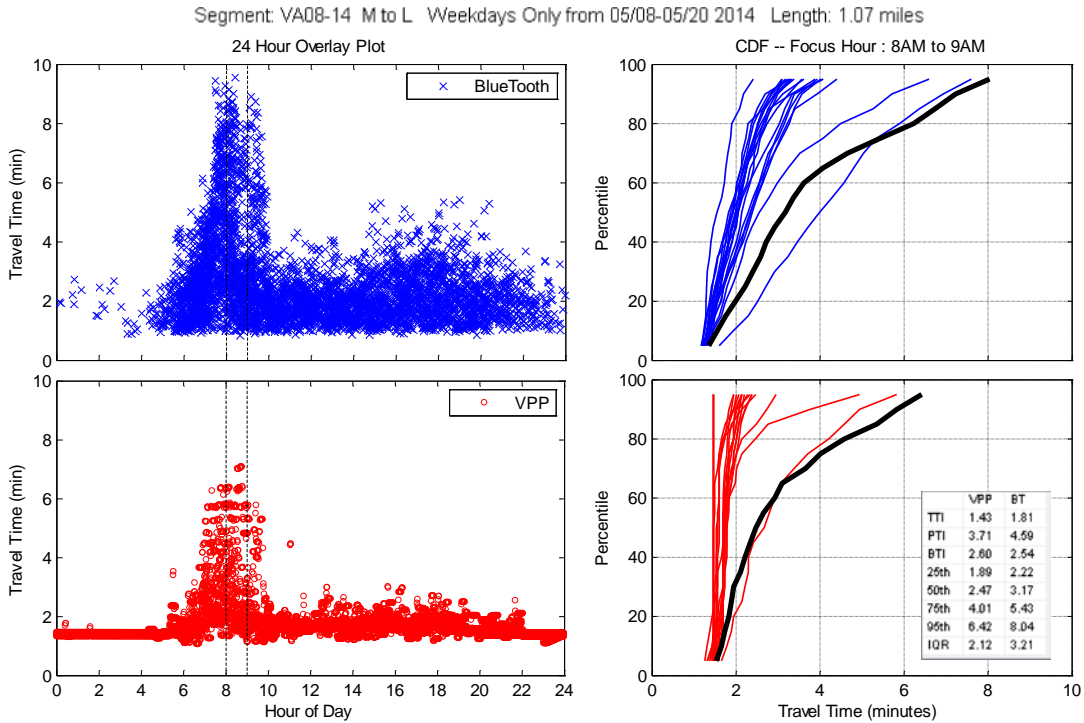
Figure 11a. 24 hour overlay plot and CDF graph from 8AM to 9AM on segment VA08-14.
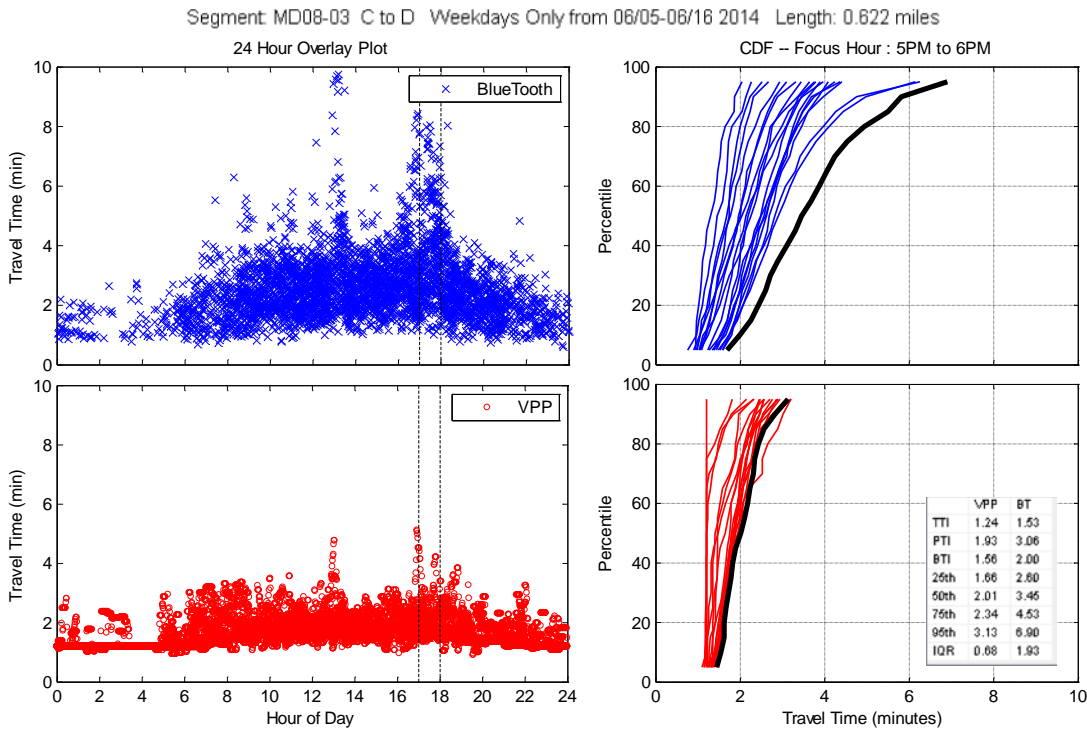


Figure 11b. 24 hour overlay plot and CDF graph from 5PM to 6PM on segment MD08-03.

On several segments, the BTM data revealed a bi-modal distribution of travel time. In no case study did the VPP data report any bi-modal pattern, only an average or central tendency. When bi-modal travel times were encountered, the VPP typically favored the faster mode. A representative sample is shown in Figure 12 from segment NJ11-21 in which the VPP reports the faster of the two observed modes. This was a recurring phenomenon across all data sets.
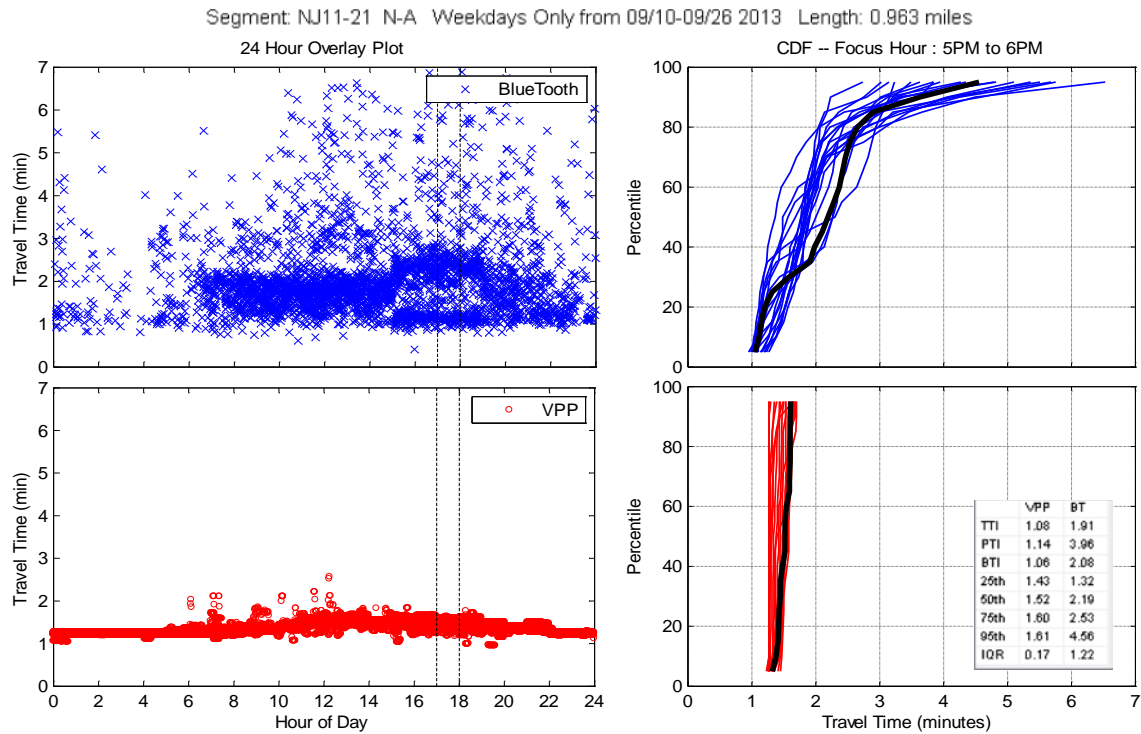


Figure 12. 24 hour overlay plot and CDF graph from 5PM to 6PM on segment NJ11-21.

# 5.0 Conclusions and Recommendations

This validation effort resulted in several recommendations with respect to the use of outsourced probe data for operations and performance measure purposes, as well as considerations for future validation, emphasis and research within the Coalition. Recommendations for the use of probe data on arterials are summarized in Table 8. Although other geometric attributes are listed in Table 8, signal density was found to be the best predictor of probe data quality. Increased volume will improve accuracy if all the other factors stay the same. However, greater volume does not overcome the challenges of reporting accurate traffic data during complex stop and go traffic on arterials with dense signal spacing.

Table 8. Arterial Probe Data Usability

| ✓ RECOMMENDED | 🔍 SHOULD BE TESTED | ✖ NOT RECOMMENDED |
|---|---|---|
| ● <= 1 signal per mile | ● 1 to 2 signals per mile | ● >= 2 signals per mile |
| ● AADT > 40,000 vpd (2-way) | ● AADT 20K to 40K vpd (2-way) | ● AADT < 20K (2-way) - low volume |
| ● Limited curb cuts | ● Moderate number of curb cuts | ● Substantial number of curb cuts |
| Principal Arterials | Minor Arterials | Major Collectors |
| Likely to be accurate… | Possibly accurate, test … | Unlikely to be accurate… |

- **Probe data is recommended for operations and performance measures when the average signal density on a corridor is one signal per mile or less**, and the AADT is 40,000 or greater. The quality of probe data for such roadways was observed to approach that of freeways. Probe data on such roadways is anticipated to *fully capture* significant slowdowns the majority of the time, and *fail to capture* less than 10% of significant slowdowns. The probe data can support a broad range of applications such as performance measures for MAP-21, planning studies, before and after analysis, traffic operations, and traveler information.

- **Outsourced probe data should be used with caution when the average signal density on a corridor is between one and two signals per mile.** Increased AADT on such roadways will increase accuracy, but increased volume does not overcome issues when bi-modal flow is encountered. Probe data is expected to *fully* or *partially capture* the majority of significant slowdowns, but *fail to capture* up to 50% significant slowdowns. If probe data is used for performance measurement, planning or traveler information, it should be tested. Probe data on such roadways may be used for comparative before and after studies, realizing there may error in the data, but it will likely be common to both before and after results.

- **Probe data is NOT recommended when signal density is above two signals per mile.** On such roadways, the VPP data is expected to *fail to capture* the majority of significant slowdowns. Probe data is not recommended for any applications above the two signal per mile threshold at this time.

Probe data quality is anticipated to improve in time with increased probe density and improved algorithms. Data collection in the spring of 2015 on US-1 in Virginia showed evidence of improvement. Arterial data quality should continue to be periodically monitored by the Coalition's validation efforts.

Throughout all the case studies, the VPP data exhibited some fundamental issues related to traffic characterization on signalized roadways. These fundamental issues included:

- **Probe data consistently errored toward faster speeds during congested periods.** The extent of slowdown measured in terms of reduction in speed was consistently underestimated as evidence by SEB measurements as well as by the distribution analysis. Even for events classified as fully captured, any error in the extent of slowdown was biased toward faster speeds. This systematic bias towards higher speeds will have programmatic significance if probe data is used in long term performance monitoring. As probe data quality improves, the data will more accurately report the full extent of slowdowns. As a result congestion may appear to grow worse when in actuality, it is only the quality of the probe data that is improving. This scenario has been corroborated by early adopters of probe data for arterial performance measures.

- **Whenever traffic progresses at two distinct travel times or speeds as a result of signal operation, probe data invariably reports the faster of the two modes.** This issue, sometimes termed 'optimistic bias in the presence of bi-modal flow', was independent of geometric attributes, and evident whenever bi-modal flow was encountered. Over the entire course of the validation there were only a handful of exceptions during which the probe data was able to successfully report a speed that was between the two modes. This behavior was evident in both the traditional plots as well as the 24-hour overlay plots.

- **Complex flow patterns common on signalized roadways cannot be observed in VPP data.** Bi-modal flow induced by signal operations is one example of complex flow. Other examples include large variation in speed due to multi-cycle failures, significant mid-block friction due to high density of access points and curb cuts, and changes in vehicle flow patterns due to signal timing changes. Because current probe data feeds report only an average speed, no information on the variation of speed is available. Complex flows were not captured either with the traditional data plots, or with the 24-hour overlay plots.

This validation effort on signal controlled arterial roadways was the most expansive to date and resulted in several findings related to appropriate validation methodology. Future arterial validation efforts (as well as future work on arterial performance measures) should consider the following issues.

- **The *Slowdown Analysis* provides the most insight into VPP data's ability to accurately capture traffic conditions.** Although the process is currently based on expert review of data as opposed to a systematic statistical computer compilation (similar to AASE and SEB), the results of the slowdown analysis provided the most intuitive and usable feedback on the fidelity of probe data. Once the slowdown analysis indicated that

probe data could adequately reflect disruptions in traffic flow, the traditional analysis and distribution analysis could begin to quantify the degree of fidelity. This methodology, currently in its infancy, should be formalized moving forward in collaboration with industry.

- *Traditional Analysis* **methods for measuring precision and bias (using AASE and SEB) should be assessed against the mean, not the standard error of the mean (SEM) band.** The SEM band, which was introduced to account for uncertainty in freeway reference data, masks the true performance of arterial probe data due to the high levels of speed variation inherent in arterial traffic flow.

- *Distribution analyzes* **provides a fundamentally better perspective to observe repeatable traffic conditions.** Overlay plots enhance repeatable traffic patterns, and the corresponding percentile-based cumulative frequency diagrams (CFDs) provide valuable visual perspective and tabular data for objective measurements. These procedures are capable of producing existing roadway performance measures, and can serve as a robust basis to compare the full breadth of performance, rather than a select metric. The distribution analysis should be continued in future validation, and considered further as the basis for standard arterial performance measure tools.

As a result of the findings of this study, the following recommendations are made to the I-95 Corridor Coalition and its members.

- **The I-95 Corridor Coalition should continue to monitor outsourced probe data fidelity on arterials as part of the VPPII initiative.** Probe data quality is anticipated to improve as probe data density increases, and algorithms allowing for point pairing improve. Validation of multi-vendor probe data available in VPPII will continue to benchmark industry capability moving forward.

- **Future work on arterial performance measures and probe data validation on arterials should build on the methodologies established in this report.** Freeway measures and methods are inadequate on arterials. Future efforts should engage the traffic engineering community along with planning and operations to merge current work on arterial management with probe data measures.

- **The Coalition should engage probe data providers and industry researchers to explore and prototype new data items capable of fully characterizing the complex arterial travel patterns including resulting from signal control.** At the base of this discussion are issues related to 'what should be reported in addition to mean speed?'. Current practice of mean speed measurements alone fail to capture the dynamics of arterial traffic flow.

The I-95 Corridor Coalition's Vehicle Probe Project continues to lead the country as the epicenter for probe-based operations and planning performance measures. The unique nature of a common licensing agreement combined with common data formats and analysis tools provides the best foundation to continue to advance arterial performance measure practice from a multi-state, industry and research collaborative environment.