Quality of Private Sector Travel-Time Data on Arterials

Jia Hu, Ph.D.[1]; Michael D. Fontaine, Ph.D., P.E.[2]; and Jiaqi Ma, Ph.D.[3]

Abstract: Accurate traffic state information is essential for both travelers and transportation agencies. In the past, traffic condition data were usually collected by a government agency using its own sensors. Recently, a number of private sector companies have started selling travel-time and speed data collected using probe vehicles, which provides a viable opportunity to outsource traffic data collection. Because these data sources and their related algorithms are proprietary, the reliability and accuracy of this private sector data is often an important issue for transportation agencies. Previous studies have examined the accuracy of private sector data on freeways, but arterials have not been examined extensively. Arterials represent a fundamentally more challenging environment for probe vehicle data given the larger variance in travel times created by traffic signals and other intermediate access points. In the research, the quality of private sector data on arterials is evaluated by utilizing Bluetooth travel-time data as the ground truth. The evaluation is conducted from two perspectives: the ability to track real-time conditions, and the ability to identify long-term traffic state changes. The study sites are three signalized arterials in the state of Virginia. The results indicate that the private sector data evaluated were not suitable for real-time applications, but could be used to measure long-term traffic state changes for performance measurement programs. **DOI:** 10.1061/(ASCE)TE.1943-5436.0000815. *© 2016 American Society of Civil Engineers.*

Author keywords: Private sector data; Arterial speed data; Real-time data quality; Long-term data quality.

Introduction

In urban areas, arterials play a key role in the transportation system, yet travel-time data on these routes have historically been limited. This lack of travel-time data has made it difficult to develop arterial advanced traveler information systems (ATISs) or performance measurement programs that can track arterial performance over time.

In the past, traffic condition data were usually collected by sensors owned by state or local DOTs. Recently, a number of private sector companies like INRIX, TomTom, and Here have begun selling network speed and travel-time data based on probe vehicle data. These providers typically offer extensive geographic coverage and do not need to install any in-pavement sensors because they rely on probe data. As a result, these providers may allow DOTs to save money by outsourcing travel-time data collection. For example, data from INRIX have already been adopted in a number of performance evaluation studies (Fontaine et al. 2014, p. 21; Wikander et al. 2014, p. 19; Belzowski and Ekstrom 2013, p. 38; Eisele et al. 2013, p. 21). However, because the data source and related algorithms from these companies are proprietary (Elefteriadou et al. 2014, p. 90), the reliability and accuracy of this private sector data is often a concern for DOTs.

A number of studies have investigated the accuracy of private sector data on freeways. These studies have been summarized in Table 1. These studies have generally found good agreement between private sector data and ground truth data, which is often floating car, Bluetooth reidentification, or toll tag reader data. Nevertheless, several inconsistencies were observed. Two studies noted that there was a time lag between when the private sector data showed the onset and recovery period of congestion and when it was observed in the baseline data (Chase et al. 2012, pp. 110–119; Elefteriadou et al. 2014, p. 90). Another study noted data quality issues when average speed drops below 56.3 km/h (35 mi/h) (Kim and Coifman 2014, p. 19), and also found that the data provider had a bias toward overreporting speed (Kim and Coifman 2014, p. 19).

Arterials represent a fundamentally more challenging environment for probe vehicles than freeways, and the accuracy of private sector data on arterials has not been well explored (Cambridge Systematics 2012). The question of private sector data accuracy on arterials was first raised by Young (2010). He examined the private sector data on a 1.6-km (1-mi) arterial in Virginia for the duration of 1 day, and then pointed out that the speeds on the arterial have a higher overall variance. Multiple distinct travel times can be observed on arterials because traffic signals divide traffic into pulsed flows. Therefore, in order to achieve levels of accuracy comparable to those of freeways, larger sample sizes are required. The Texas Transportation Institute (TTI) attempted to examine the accuracy of private sector arterial data in 2011 (Turner 2011) using Minnesota DOT fixed-point sensor data. Unfortunately, due to the limitations of the fixed-point sensors, the research could not conclude whether the private sector data were accurate or not. As of now, the only quantitative study of arterial data quality was conducted by Wang et al., who found a systematic bias (Wang et al. 2014, p. 122) where the probe data provider was less responsive to changes in traffic state. Wang et al.'s study is also summarized in Table 1. A limitation of Wang et al.'s study was that it was based on only 7 days of data.

The floating car method is one of the most frequently used methods to provide ground truth travel-time data. However, this method requires a significant effort to collect data and therefore usually leads to very limited sample sizes. Bluetooth travel-time

[1]Research Scholar, Turner-Fairbank Highway Research Center, Federal Highway Administration, 6300 Georgetown Pike, McLean, VA 22101 (corresponding author). E-mail: Jia.Hu@dot.gov; jh8dn@virginia.edu

[2]Associate Principal Research Scientist, Virginia Center for Transportation Innovation and Research, 530 Edgemont Rd., Charlottesville, VA 22903. E-mail: Michael.Fontaine@VDOT.Virginia.Gov

[3]Transportation Research Engineer, Leidos, Inc., 11251 Roger Bacon Dr., Reston, VA 20190. E-mail: jiaqi.ma@leidos.com

Note. This manuscript was submitted on March 29, 2015; approved on September 8, 2015; published online on January 19, 2016. Discussion period open until June 19, 2016; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering***, © ASCE, ISSN 0733-947X.

Table 1. Past Studies Investigating Private Sector Data Accuracy

Location	Road type	Length [km (mi)]	Duration	Ground truth	Measure of effectiveness	Accuracy	Reference
I-295, Florida	Freeway	12.7 (7.89)	3 h	Floating car	Percent error	5–38%	Elefteriadou et al. (2014, p. 90)
I-71, Ohio	Freeway	22.5 (14)	6 months	Loop detectors	Congestion onset	Lag of 6 min	Coifman and Kim (2013, p. 60)
I-287, New Jersey	Freeway	33.0 (20.5)	1 day	Toll tag reidentification	Average absolute speed error	3–6 mi/h	Kim and Coifman (2014, p. 19)
I-78, New Jersey	Freeway	28.2 (17.5)	1 day	Toll tag reidentification	Speed error bias	−2 to 4 mi/h	Kim and Coifman (2014, p. 19)
SR 522, Washington	Arterial	4.8 (3)	7 days	License plate readers	Mean absolute percent error (speed)	17–73%	Wang et al. (2014, p. 122)
I-90, Washington	Freeway	123.9 (77)	7 days	License plate readers	Reaction to road closure	Quick response	Wang et al. (2014, p. 122)
I-91, Massachusetts	Freeway	N/A	2 days	License plate matching	Mean absolute percent error (travel time)	1–1.5%	Jia et al. (2013, p. 16)
I-10, Florida	Freeway	32.2 (20)	4 days	Floating car	Average absolute speed error	6.27 mi/h	Lattimer and Glotzbach (2012, p. 8)
I-95 Maryland, Virginia, Delaware, and New Jersey	Freeway	148.1 (92)	4 months	Floating car	Average absolute speed error	1.96 mi/h	Haghani et al. (2008)

data have proven to be comparable to the floating car data in terms of accuracy, but can provide a significantly higher number of data points than the floating car methodology (Schneider et al. 2010, p. 309). Some studies concluded that the larger data set based on Bluetooth is superior on arterials over floating car measurements (Quayle et al. 2010, pp. 185–193).

The review of literature reveals a number of trends. First, few studies have examined the performance of private sector probe data on arterials. Second, past studies often only examined a limited time period due to the cost of collecting travel-time, speed, and delay data using traditional methods. Conclusions drawn based on 1 week of observations could be misleading because results may not represent data quality under different demand patterns, nonrecurring incidents, and/or weather conditions. Finally, the usage of private sector data for arterial performance measurement has not been explored.

Objectives and Scope

The goal of this study is to examine the accuracy of private sector travel-time data on multiple real-world arterial corridors in Virginia over an extended period of time. The research investigates the responsiveness of the private sector data to real-time traffic state change, as well as its accuracy capturing long-term traffic state change due to a systematic alteration, which in this research is the installation of an adaptive traffic signal control system.

The private sector companies selling travel-time data are constantly attempting to improve their product, so any analysis of their data quality represents performance at a discrete point in time. Thus, it is entirely possible that current generation performance may differ from what was analyzed in this paper. As a result, the specific provider used in this analysis will not be explicitly named. The findings presented in this paper should be viewed as providing high-level guidance to agencies interested in using private sector data on arterials, and agencies could use this research as a way to identify critical use cases that should be investigated in more detail with a specific provider. It is possible that performance of an individual provider could deviate substantially from what is reviewed in this paper.

Data

Study Sites and Data Collection

The Virginia DOT has recently completed a pilot test of an adaptive traffic signal control (ATSC) system on a number of corridors in Virginia. A total of three corridors were selected as study sites for this research, and their characteristics are presented in Table 2. All of these corridors where suburban multilane arterials and experienced peak period congestion. Posted speed limits were between 56.3 and 72.4 km/h (35 and 45 mi/h). All the signals within the study areas are included in the ATSC deployment. The authors collected data from both before and after the ATSC system was activated. Three months of after data were collected at all three sites. The duration of the before period was limited by when Bluetooth devices were initially deployed in the corridor.

Figs. 1–3 show maps of Traffic Message Channel (TMC) start and end locations for all three study sites. TMC could cover only one intersection when the intersection is wide. They are frequently seen for U.S. Route 17.

Table 2. Study Site Summary

Locations	Length [km (mi)]	Number of signals	Active date	Before start	Before end	After start	After end	Annual average daily traffic (AADT) [vehicle per day (vpd)]
U.S. Route 250, Staunton, Virginia	2.8 (1.75)	10	July 10, 2012	June 16, 2012	July 9, 2012	February 1, 2013	April 30, 2013	22,704
Virginia State Route 7, Winchester, Virginia, and Frederick County	2.4 (1.5)	12	May 7, 2012	April 10, 2011	May 6, 2011	February 1, 2013	April 30, 2013	24,718
U.S. Route 17, York County	13.4 (8.3)	19	June 19, 2012	May 30, 2012	June 18, 2012	February 1, 2013	April 30, 2013	49,695

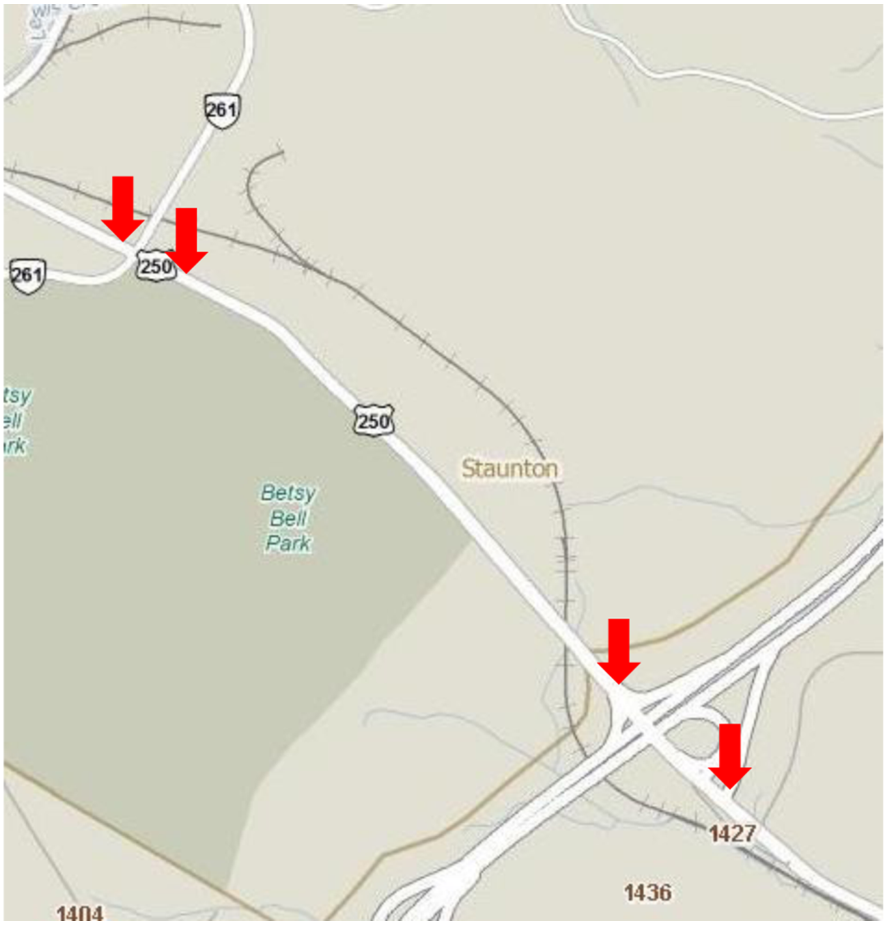

Fig. 1. TMC start and end locations at U.S. Route 250 (map image courtesy of Virginia Department of Transportation)

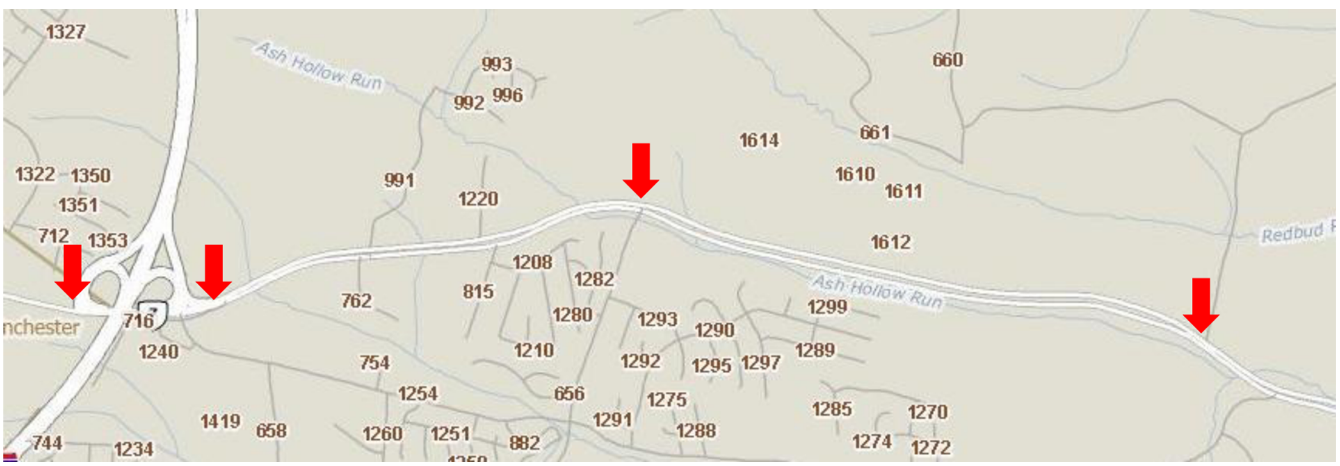

Fig. 2. TMC start and end locations at Virginia State Route 7 (map image courtesy of Virginia Department of Transportation)

Data Sources

Real-time speed data were acquired from a private sector data provider for the corridors shown in Table 2. All of the speed data are aggregated into 15-min periods. The speed data analyzed are space mean speed. To be noted, according to the vendor, this space mean speed is computed using a combination of instantaneous speeds and speeds derived from location estimates at multiple points (point-pair estimates).

Bluetooth media access control (MAC) address reidentification was used to establish the ground truth data used for comparison purposes. The system matches anonymous MAC addresses at the beginning and end of a roadway segment and calculates travel time through analysis of the time stamps of matched MAC addresses.

Data Processing

Because the private sector data rely on the presence of vehicle probes, there may be time periods in which an insufficient number of probes traverse the section to generate a reliable travel-time measure. During these periods, the data provider reported a travel time

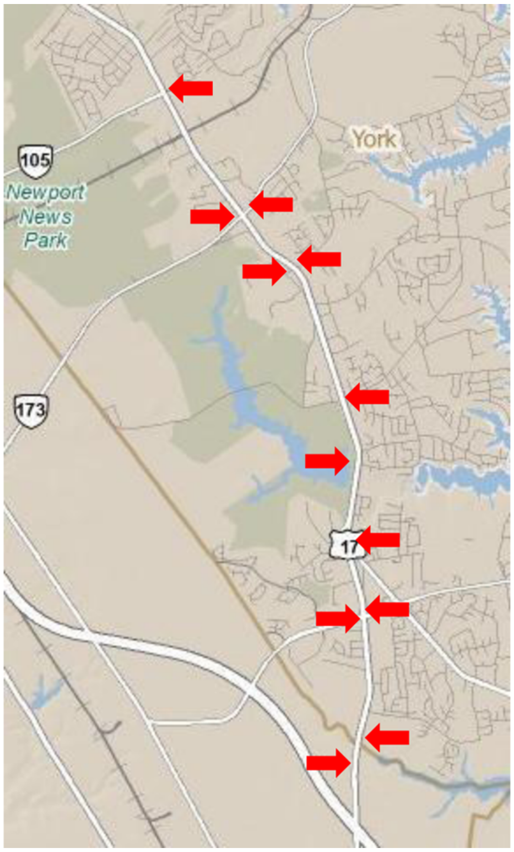

Fig. 3. TMC start and end locations at U.S. Route 17 (map image courtesy of Virginia Department of Transportation)

based off of historic data. These data are definitely not useful when assessing the real-time performance of the data stream, so these data must be screened out. The data provider reported travel times on discrete segments called TMC links. TMCs are defined by digital mapping companies like HERE (formerly NAVTEQ) and TeleAtlas and typically span roadway sections between intersections. The vendor includes a confidence score based on whether the data are derived from real time or historic data for each TMC for every time period, so periods using historic data could be easily identified. The Virginia DOT has defined that at least 85% of the TMCs composing a travel-time corridor must be reporting real-time information before the data can be used to post real-time travel-time estimates. This threshold was checked for each discrete 15-min interval studied. For example, the U.S. Route 17 corridor in York County is composed of 13 TMCs. At least $13 \times 0.85 = 11.05$, rounded up to 12, TMCs must report real-time information in a 15-min interval before the travel-time data would be included in the evaluation. The historic data from the remaining TMCs were then combined with the real-time TMCs to generate a corridor travel time. If a time interval has less than 12 TMCs reporting real-time data, that time interval would be discarded from further analysis.

For the Bluetooth data, data are aggregated in 15-min intervals. The number of Bluetooth probes in each 15 min is examined to determine if the sample size is sufficient to generate a reliable estimate of mean travel time. This process is discussed subsequently in the paper. In the event that the minimum sample size requirement is not met, that 15-min interval is removed from the analysis. The locations studied were generally in suburban areas with minimal bicycle and pedestrian traffic, so nonvehicular traffic was not expected to exert a significant bias.

Before the data sets were analyzed and compared, data harmonization was required. When possible, the researchers attempted to align the Bluetooth reader locations with the TMC endpoints so that travel times were measured over a consistent spatial basis. Sometimes this was not possible based on site characteristics, however. In order to facilitate a fair comparison, the compared travel-time links needed to be equivalent. This was accomplished by taking the private sector travel-time data and converting it to a calculated travel time based on the length of the corresponding Bluetooth segment. The calculation is based on known TMC travel times. When the end of a TMC is beyond the end of the Bluetooth segment, interpolation is utilized to find the calculated private sector travel time on the shorter Bluetooth segment. Otherwise, extrapolation is utilized. The harmonization process requires the assumption that speeds are fairly constant within each TMC segment. This is expected to be true in this study because (1) the end TMCs are typically short, averaging 0.6 km (0.4 mi), and (2) there are no major roads with high traffic volumes crossing the end TMCs. The summation of all harmonized TMC travel times gives the total travel time of the entire corridor. The average speed is the determined by dividing the length of the corridor by the associated total travel time.

Ground Truth Validation

The Bluetooth equipment provided both a smoothed 15-min average speed as well as individual Bluetooth device speeds. The Bluetooth vendor did not specify the methodology that it used to develop smoothed 15-min Bluetooth speeds, so it was necessary to examine whether processed individual vehicle speeds were comparable to the reported 15-min smoothed speeds. First, individual Bluetooth matches were retrieved for U.S. Route 17 southbound during the time period February 1 to April 30, 2013. The aggregation interval for this Bluetooth data was defined as 15 min in order to match the private sector data. The minimum required sample size was examined for each time interval, utilizing the following equation based on the central limit theorem:

$$n = \frac{t_{\frac{\alpha}{2},n-1} \times s}{d}$$

where n = minimum sample size; α = desired confidence; s = standard deviation of available data; and d = maximum allowable error.

During the examination, the following assumptions are made:

- Maximum allowable error is 8.0 km/h (5 mi/h); and
- Desired confidence is $\alpha = 0.05$.

The time intervals that fail the minimum sample size examination are excluded from the following analysis.

Fig. 4 presents the plot of the smoothed speed data against the individual Bluetooth speed data. The Bluetooth speed data were computed from the per vehicle records. The two sources are plotted against each other using the 15-min speed measurements from the same 15-min time stamp. A total of 163 data points are plotted. As shown in Fig. 4, all the dots are evenly distributed along the $y = x$ line with a mean absolute percentage error (MAPE) of 13%. The Pearson correlation is calculated for these 163 pairs of data, which is equal to 0.507. These findings validate that the smoothed data are comparable to what would be developed from processing individual vehicle data. As a result, the smoothed 15-min vehicle speeds were used as a benchmark in this study.

Methodology

The accuracy of the private sector data was evaluated based on two perspectives:

J. Transp. Eng., 2016, 142(4): -1--1

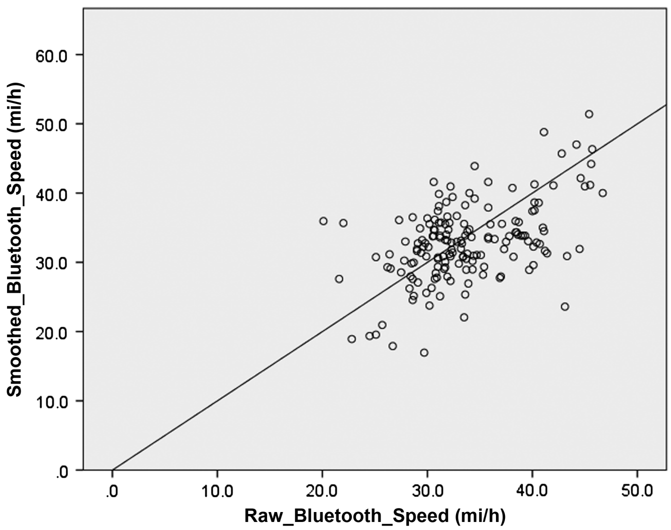

Fig. 4. Smoothed versus raw 15-min Bluetooth speeds

- Real-time quality: This evaluation examined how well the private sector data respond to the dynamic changes of traffic state in real time. It reflects the feasibility of adopting the data source to monitor arterial traffic conditions and provide real-time traveler information.
- Long-term performance measurement validity: This evaluation assesses whether long-term trends in benchmark data are reflected in the probe data. It is possible that a data provider may not be able to provide continuous real-time data, but that the data may show consistent trends with the Bluetooth data when aggregated over longer periods of time.

For the purpose of this analysis, all the data for each corridor are summarized by time interval. The dividing criteria includes day of the week [Mondays, weekdays (Tuesday to Thursday), Fridays, and weekends] and time of the day (morning, midday, afternoon, and night). In this study, weekdays only include Tuesday, Wednesday, and Thursday. Monday and Friday are separated because those two days often have distinct traffic patterns from the other weekdays. Morning is defined as from 6–10 a.m. and afternoon is defined as from 3–7 p.m. The rest of the day would be contained in the midday and overnight periods.

In order to perform a quantitative evaluation on the accuracy of private sector data, several performance measures were calculated.

The descriptive statistics and distribution of the data sources are examined and reported. Kurtosis is calculated, which is a measurement of the flatness of a distribution. A higher kurtosis value represents a more peaked distribution. Kurtosis is calculated as follows:

$$K = \frac{m_4}{m_2^2}$$

$$m_2 = \frac{\sum_{i=1}^{n}(v_i - \bar{V})^2}{n}$$

$$m_4 = \frac{\sum_{i=1}^{n}(v_i - \bar{V})^4}{n}$$

where n = number of observations; v_i = speed observation i; and $\bar{V}$ = averaged speed over all observations.

Skewness was also computed, and serves to measure the degree of asymmetry in a distribution (left or right tailed). A negative skewness value represents left-tailed distribution, and vice versa. The mathematical formulation of the skewness is defined as

$$S = \frac{m_3}{m_2\sqrt{m_2}}$$

$$m_2 = \frac{\sum_{i=1}^{n}(v_i - \bar{V})^2}{n}$$

$$m_3 = \frac{\sum_{i=1}^{n}(v_i - \bar{V})^3}{n}$$

where n = number of observations; v_i = speed observation i; and $\bar{V}$ = averaged speed over all observations.

The calculation of kurtosis and skewness helps to show whether the distribution of speeds differs between the private sector data and the Bluetooth data. Given that arterials experience fundamentally more variable traffic flow, this will help show whether there is any bias in the speed estimates generated by the two data sources.

Two metrics were utilized to quantify the error. First, speed absolute error (SAE) was calculated using the following equation:

$$\text{SAE} = |\hat{V} - v_i|$$

where $\hat{V}$ = corresponding ground truth (Bluetooth) speed.

The speed absolute error shows, on average, how much deviation there is between the Bluetooth and private sector data. A similar performance measure that was used was the MAPE. It was calculated as follows:

$$\text{MAPE} = 100 \times \frac{1}{n}\sum_{i=1}^{N}\frac{|\hat{V} - v_i|}{v_i}$$

The MAPE and SAE are used jointly to assess the level of deviation between the two data sets.

Results

Real-Time Data Quality

A total of 36,222 15-min time intervals were available when both Bluetooth and private sector data provided real-time coverage. These intervals include all three corridors for both before and after periods. The Pearson correlation between the Bluetooth and private sector data was tested by Bluetooth speed range, as shown in Table 3. Although the correlation is statistically significant, the magnitude of the correlation coefficient is less than 0.4. Therefore, no strong correlation is detected between the two data sources. The magnitude of correlation varies based on speed, ranging from −0.236 to 0.360. The error of the private sector data is within 15% when the average speed is above 48.3 km/h (30 mi/h) and it decreases as the speed increases. However, a low error rate does not indicate high correlation. The highest correlation is observed when the speed is between 64.4 and 72.4 km/h (40 and 45 mi/h).

Fig. 5 shows the probability density function (PDF) of all 36,222 data points. Based on this figure and descriptive statistics demonstrated in Table 4, it can be seen that:

- The private sector data tend to overestimate speeds by approximately 20% relative to Bluetooth;
- The distribution of the private sector data speed skews to the right, while Bluetooth skews to the left; and

Table 3. Pearson Correlation Comparison

Bluetooth speed range [km/h (mi/h)]	Count	MAPE (%)	Pearson correlation (two-tailed)
<24.1 (<15)	92	−209.2	−0.236
24.1–32.2 (15–20)	374	83.7	0.215
32.2–40.2 (20–25)	5,479	54.0	0.123
40.2–48.3 (25–30)	11,654	28.6	−0.014
48.3–56.3 (30–35)	13,235	14.7	0.064
56.3–64.4 (35–40)	4,026	9.3	0.137
64.4–72.4 (40–45)	1,170	8.1	0.360
72.4–80.5 (45–50)	192	4.1	0.102
0–80.5 (0–50)	36,222	25.6	0.247

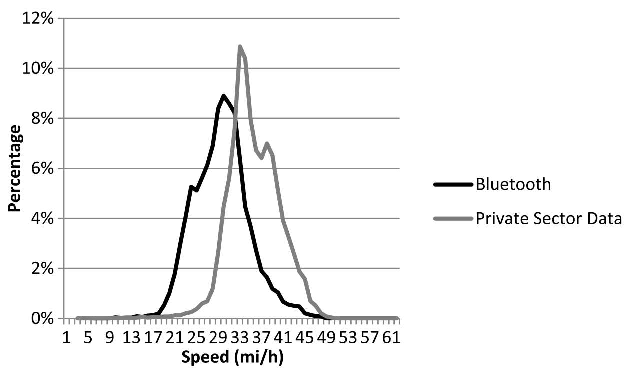

Fig. 5. Probability density function plot of the two data sources

- The private sector speed distribution is more concentrated, showing less variability.

Fig. 6 shows all of the speed data collected, including both the before and after period at all three sites. The data from the two sources are plotted against each other using the 15-min speed measurements from the same 15-min time stamp. The line of $y = x$ is drawn on the chart as a reference point, so if the two data sources match perfectly, all the data should lie on the $y = x$ line. It again shows that the private sector data tend to overestimate speed on the arterials relative to Bluetooth. There is also a clear tendency that the plot converges as average speeds increase. This is consistent with freeway results that showed good agreement between private sector data and Bluetooth data under free-flow conditions. Differences become more pronounced as congestion becomes more severe.

Fig. 7 is a heat map showing the SAE of the private sector data relative to the Bluetooth data. Color represents the density of data points falling in the corresponding speed and error level, where black represents higher densities of observations and white represents lower densities. If the private sector data perfectly match the Bluetooth data, the chart would appear to be all white with a colored horizontal band on the bottom given the data that are available for the specific speed bins. The figure can be interpreted by reading it using each vertical speed bin on a horizontal axis. The white areas indicate no data points are available. The color demonstrates the absolute number of data points during certain conditions, and colors cannot be compared horizontally because each speed bin does

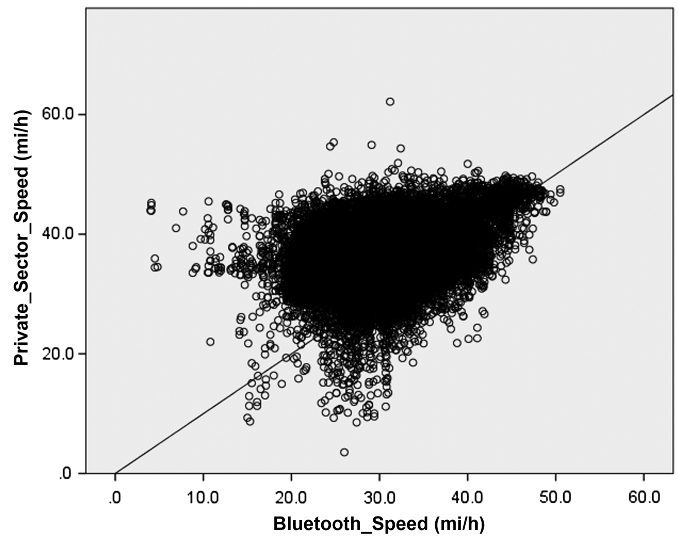

Fig. 6. Private sector data versus bluetooth speeds

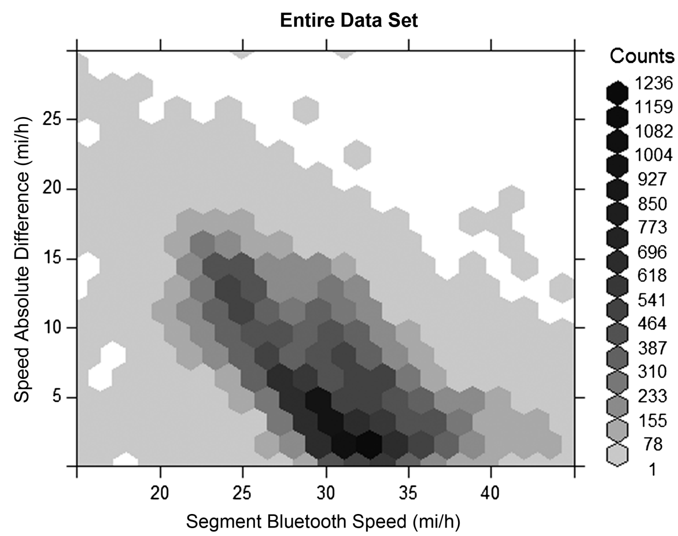

Fig. 7. Speed absolute error density hexbin plot

not have the same sample size. When the speed is above 56.3 km/h (35 mi/h), the heat maps show that most of the available data in the study have speed differences of 0–4.8 km/h (0–3 mi/h). In the 30–35 mi/h bin, private sector speed deviations mostly concentrate in the 0–4.8 km/h (0–3 mi/h) range, as demonstrated by the darkest color in Fig. 7. Because more data are available, some data are showing SAE values of 4.8–12.9 km/h (3–8 mi/h). As the speed continues to drop to 32.2–48.3 km/h (20–30 mi/h), almost no data demonstrate error smaller than 4.8 km/h (3 mi/h), so the differences appear to increase as speed declines.

Table 4. Speed Descriptive Statistics

Sources	N	Range [km/h (mi/h)]	Mean [km/h (mi/h)]	Standard deviation [km/h (mi/h)]	Skewness		Kurtosis	
					Statistic	Standard error	Statistic	Standard error
Bluetooth speed	36,222	74.8 (46.50)	48.4 (30.09)	8.2 (5.08)	0.24	0.013	0.6	0.026
Private sector speed	36,222	94.3 (58.60)	58.0 (36.04)	7.5 (4.67)	−0.16	0.013	1.3	0.026

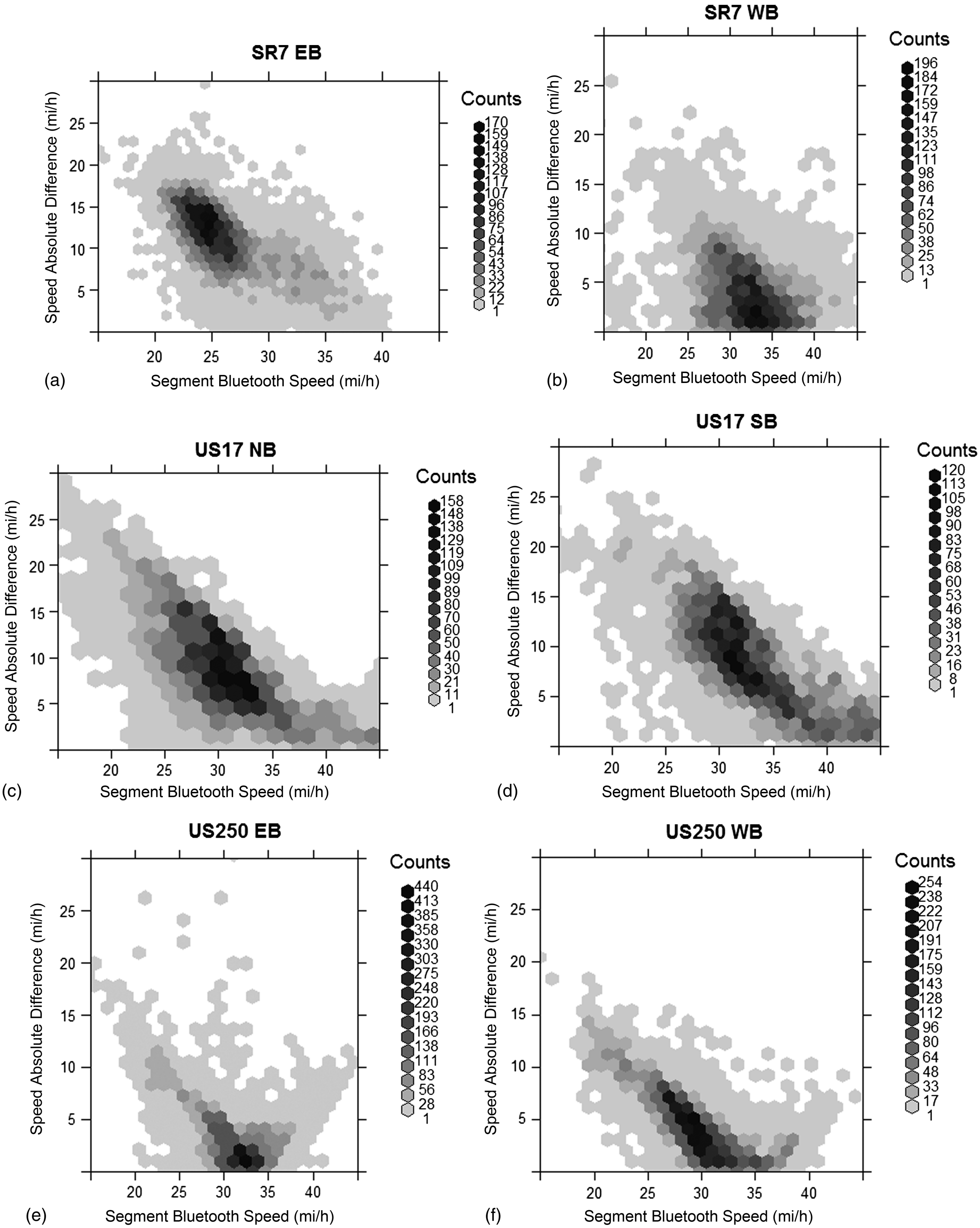

Fig. 8. Speed absolute error by site

Fig. 8 examines the speed absolute error of private sector data by site and by direction of travel, with all sites showing similar trend as in Fig. 7.

One major conclusion from these charts is that the absolute error of the private sector data is often negatively correlated to the average speed. More interestingly, the magnitude of decreases in average speed is almost equal to the amount of increase in absolute error. It implies that the reported speed of the private sector data does not always respond to traffic state changes. In other

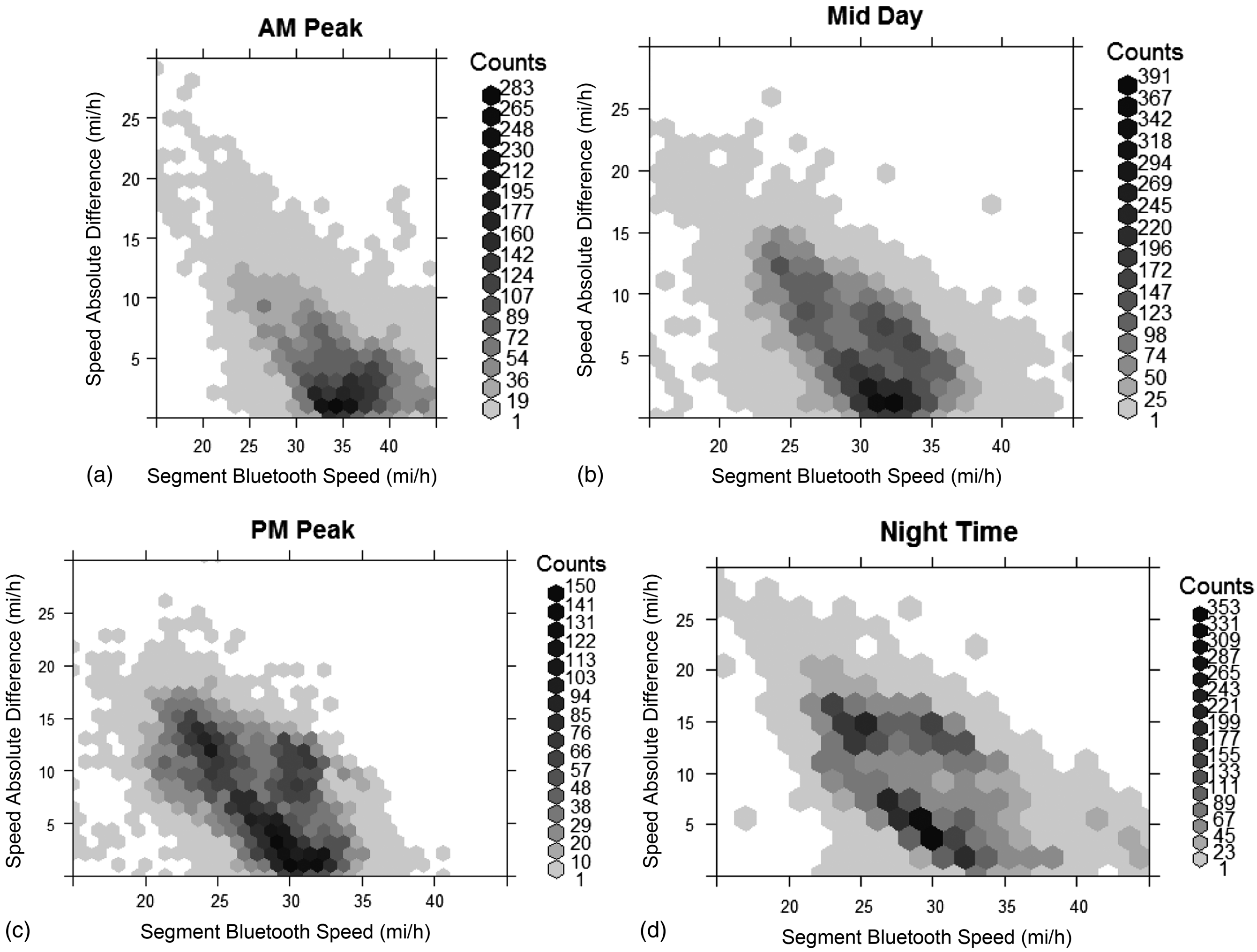

Fig. 9. Speed absolute error by time of day

words, the private sector speed may not fully report speed changes on arterials, especially during more congested conditions.

The differences between the private sector data and the Bluetooth data appear to be correlated with the observed speed variability in the corridor. Generally speaking, the larger the Bluetooth variability, the larger the differences between the private sector data and the Bluetooth data. For example, Virginia State Route 7 westbound and U.S. Route 250 eastbound show minor deviations between the two data sources because the majority of their Bluetooth speed observations are clustered in a fairly narrow speed range. The opposite is true at Virginia State Route 7 eastbound and U.S. Route 250 westbound. Sites that cover longer distances (U.S. Route 17) show moderate speed deviation, while sites with shorter lengths (Virginia State Route 7 and U.S. Route 250) may experience extreme cases. In this case, longer distances serve to dampen the traffic variability experienced at signals as they are aggregated over multiple signals. Thus, larger variability appears to occur on shorter corridors due to how vehicles progress along the corridor.

The private sector data accuracy was also investigated by time of day. In this case, time of day was examined as a surrogate measure for the volume at the site and the time of day could impact the number of probes available for the private sector data, along with the level of congestion and speed variability at the site. The results are presented in Fig. 9. At first sight, morning and midday data are clustered in a smaller area at the bottom of the heat map than the afternoon or night data, thus indicating closer agreement between the private sector data and the Bluetooth data. However, closer investigation reveals that the level of congestion is the real driver of these results, not time of day. The morning and midday data show closer agreement between the Bluetooth and private sector data because the majority of the speed observations are clustered above 48.3 km/h (30 mi/h), indicating less congested traffic. The afternoon and night data show speed readings that are around or less than 48.3 km/h (30 mi/h), representing more congested traffic.

Performance Measurement Applications

While the prior analysis focused on the real-time performance of the private sector data, long-term performance measurement is also a major interest for transportation agencies. It is possible that individual 15-min intervals may have some errors, but when aggregated over time, those errors may cancel or average out. In that case, the private sector data may have applications for long-term arterial performance measurement applications even if real-time performance is not deemed acceptable.

Fig. 10 demonstrates, in the form of radar map, the performance of the private sector data relative to the Bluetooth data when aggregated over a longer time period. The comparison was conducted separately for all 16 different time intervals (4 days of the week × 4 times of day). In this case, the data are used to compare traffic conditions before and after the installation of an ATSC system on the three test corridors. The corresponding statistics are presented

J. Transp. Eng., 2016, 142(4): -1--1

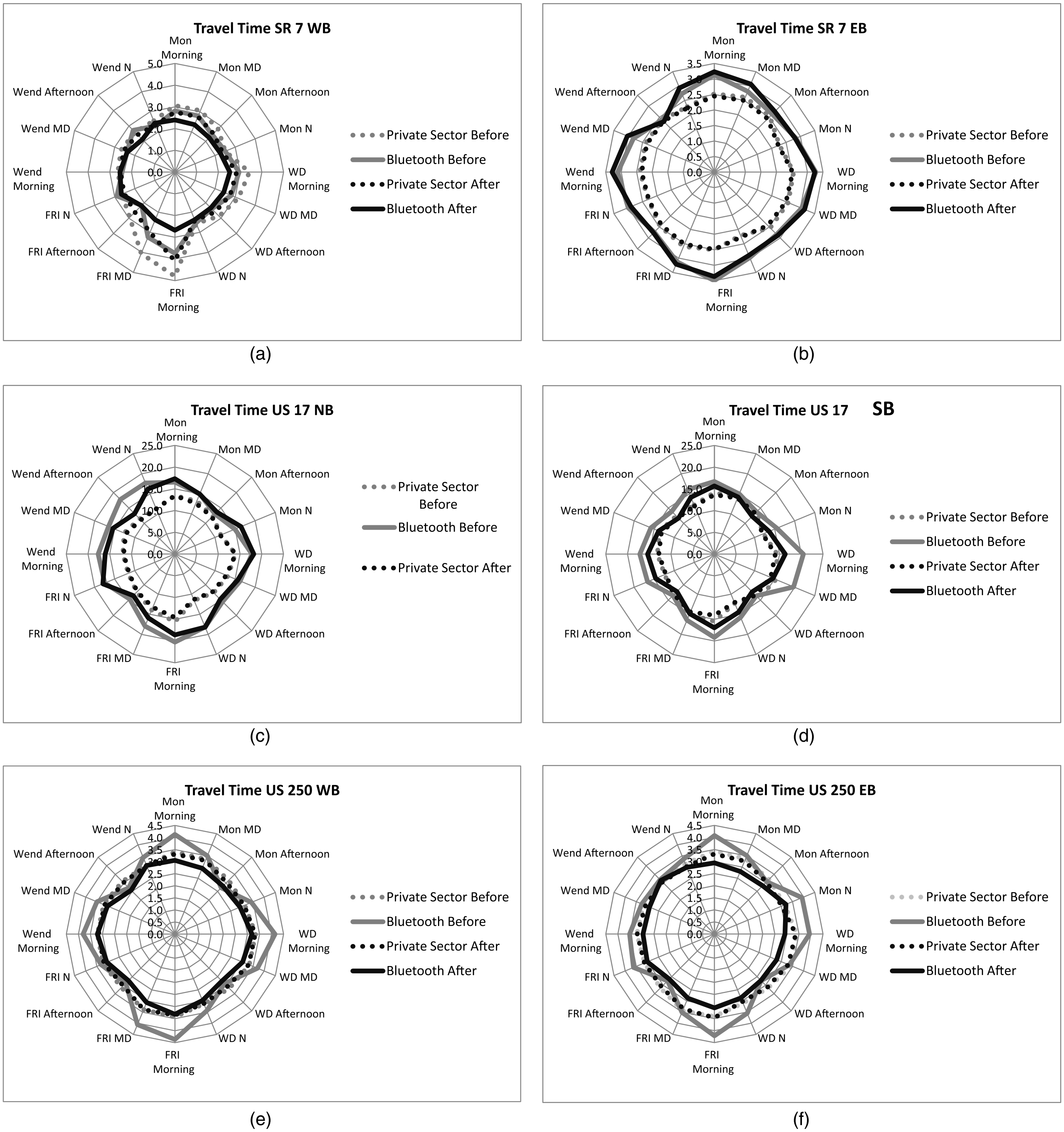

Fig. 10. Aggregate data trends using data from the before and after study

in Table 5. The goal of this evaluation is to investigate the feasibility of using private sector data to conduct before and after or wperformance measurement studies over a prolonged data evaluation period. It is possible that, although the real-time accuracy of the private sector data is not ideal, it could capture the change in traffic state when averaged over an extended period of time. Each travel direction from three corridors is reported in a separate chart. Within each chart, each data point represents the average travel time of a time period [by time of day (TOD) and day of week (DOW)] over approximately 4 months. The findings are summarized in the following:

- Private sector travel times tended to be lower than Bluetooth data. This matches the observations of the real-time data quality.
- In two of the three sites, the private sector data closely match the Bluetooth data. However, there are significant differences observed for one site: U.S. Route 250. After examining the data more closely, the two data sources are showing similar observations for the after period, but there are significant differences during the before period. Considering the fact that the before period is quite short (3 weeks) compared to the after period (3 months), it is possible that the reliability of the private sector

Table 5. Descriptive Statistics

Measurements	U.S. Route 17	Virginia State Route 7	U.S. Route 250
Average error (%)	6.2	7.6	13.2
Error % standard deviation	4.4	5.3	8.2
Difference (%)	7.0	32.8	−1.4
Bluetooth standard deviation (min)	2.6	0.3	0.4
Private sector data standard deviation (min)	0.8	0.3	0.2
Length [km (mi)]	13.4 (8.3)	2.4 (1.5)	2.8 (1.75)

data is only guaranteed when longer periods of data are collected.

- There is a discrepancy observed for site Virginia State Route 7 eastbound. While the Bluetooth data show a negative effect, the private sector data show an improvement. However, the magnitudes of the changes are quite small. A t-test was performed using both data sources to examine whether there was a statistically significant change in average travel time between the before and after period. The results from two data sources agree with each other and show no statistically significant change occurred. As a result, this discrepancy between the two data sources would not result in different conclusions with respect to the effectiveness of the ATSC.
- Travel-time improvements based on the private sector data are more conservative. Therefore, it is safe to draw conclusions based on private sector data when the data are showing positive improvements. However, the magnitude of the negative effects may be underrepresented by the data.
- When looking at the magnitude of the error, sometimes sites have large percentage differences between data sets even though the absolute magnitude of the difference is not large.
- Variation in travel time captured by the private sector data is also smaller. This is also consistent with the real-time data analysis.

Conclusions

This research examined the data quality of private sector travel-time data on arterials, which has not yet been evaluated extensively. Bluetooth travel-time data was utilized as the ground truth, and the potential usage of the data for both real-time applications and performance measurement was reviewed. The key findings of this analysis were

- The private sector data sometimes do not detect speed reductions on the corridors. Differences between Bluetooth speeds and private sector speeds generally increased as congestion increased at the sites. In addition, this difference also grows with the variation in speed observations. The reported data tend to stay at or near free-flow speed, and they generally tend to overestimate speeds relative to Bluetooth. The variability captured by the private sector data is also underestimated. Therefore, the private sector data tested were not suitable for providing real-time traveler information on arterials.
- TOD does not affect the private sector data accuracy, but level of congestion and speed variability do.
- While the private sector data tested may not be able to provide consistent real-time condition estimates, they may be able to be used for long-term performance measurement. In this case, relative differences between before and after periods generally aligned with the Bluetooth data, even if the magnitudes of changes did not agree. The private sector data tended to show less variability, so they could be construed as providing a conservative view of the effect of treatments. Thus, they could be used for trend analysis on specific corridors, provided the limitations of the data are documented.

These results represent the performance of a specific private sector data stream at a moment in time, and subsequent changes to the data stream may have improved quality. This paper does imply, however, that transportation agencies should seek to verify the performance of any private sector data stream on arterials. Those evaluations should determine whether the data can be used for real-time monitoring, performance measurement, or other applications being considered by the DOTs.

References

Belzowski, B. M., and Ekstrom, A. (2013). "Stuck in traffic: Analyzing real time traffic capabilities of personal navigation devices and traffic phone applications." ⟨http://deepblue.lib.umich.edu/bitstream/handle/2027.42/100187/102967.pdf?sequence=1⟩ (Jan. 20, 2015).

Cambridge Systematics. (2012). *Travel time data collection white paper*, Florida Dept. of Transportation, Tallahassee, FL.

Chase, R. T., Williams, B. M., Rouphail, N. M., and Kim, S. (2012). "Comparative evaluation of reported speeds from corresponding fixed-point and probe-based detection systems." *Transportation Research Record*, 2308, 110–119.

Coifman, B., and Kim, S. (2013). "Assessing the performance of the SpeedInfo sensor." ⟨http://www.dot.state.oh.us/Divisions/Planning/SPR/Research/reportsandplans/Reports/2014/Planning/134723_FR.pdf⟩ (Jan. 20, 2015).

Eisele, W. L., Schrank, D. L., Schuman, R., and Lomax, T. J. (2013). "Estimating urban freight congestion costs: Methodologies, measures, and applications." ⟨http://docs.trb.org/prp/13-1344.pdf⟩ (Jan. 21, 2015).

Elefteriadou, L., Kondyli, A., and George, B. S. (2014). "Comparison of methods for measuring travel time at Florida freeways and arterials." ⟨http://www.dot.state.fl.us/research-center/Completed_Proj/Summary_PL/FDOT-BDV32-977-02-rpt.pdf⟩ (Jan. 22, 2015).

Fontaine, M. D., Chun, P., and Cottrell, B. H. (2014). "Using private sector travel time data for project-level work zone mobility performance measurement." ⟨http://docs.trb.org/prp/14-2186.pdf⟩ (Jan. 23, 2015).

Haghani, A., Hamedi, M., and Sadabadi, K. F. (2008). "Corridor coalition." ⟨http://www.i95coalition.org/i95/Portals/0/Public_Files/uploaded/Vehicle-Probe/I-95%20CC%20VP%20EX%20SUMM%20FINAL%20JAN%2022%2009%20W%20COVER.PDF⟩ (Jan. 27, 2015).

Jia, C., Li, Q., Oppong, S., Ni, D., Collura, J., and Shuldiner, P. W. (2013). "Evaluation of alternative technologies to estimate travel time on rural interstates." *Transportation Research Board 92nd Annual Meeting*, *No. 13-3892*.

Kim, S., and Coifman, B. (2014). "Comparing INRIX speed data against concurrent loop detector stations over several months." *Transp. Res. Part C*, 49, 59–72.

Lattimer, C., and Glotzbach, G. (2012). "Evaluation of third party travel time data." ⟨http://itswc.confex.com/itswc/AM2012/webprogram/Paper10870.html⟩ (Jan. 23, 2015).

Quayle, S. M., Koonce, P., DePencier, D., and Bullock, D. M. (2010). "Arterial performance measures with media access control readers: Portland, Oregon, pilot study." *Transportation Research Record*, 2192, 185–193.

Schneider, W. H., Turner, S. M., Roth, J., and Wikander, J. (2010). "Statistical validation of speeds and travel times provided by a data service vendor." ⟨http://www.dot.state.oh.us/Divisions/TransSysDev/Research/reportsandplans/Reports/2010/Traffic/134416_FR.pdf⟩ (Jan. 23, 2015).

Turner, S. N. (2011). *Technical memorandum*, Minnesota Dept. of Transportation, St. Paul, MN.

Wang, Y., Araghi, B. N., Malinovskiy, Y., Corey, J., and Cheng, T. (2014). "Error assessment for emerging traffic data collection devices." ⟨http://www.wsdot.wa.gov/research/reports/fullreports/810.1.pdf⟩ (Jan. 21, 2015).

Wikander, J., Eisele, W. L., and Schrank, D. L. (2014). "Automated congestion bottleneck identification for MAP-21 performance measure reporting using large statewide speed data sets." ⟨http://docs.trb.org/prp/14-4356.pdf⟩ (Jan. 25, 2015).

Young, S. (2010). "Memorandum to vehicle probe project arterial data quality committee." ⟨http://i95coalition.org/wp-content/uploads/2015/02/2-November-2010-Arterial-Validation-White-Paper-final.pdf?dd650d⟩ (Jan. 23, 2015).