Final Technical Report
TNW2007-05

Research Project Agreement No. 61-4160

Review of Private Sector Data for Roadway Monitoring

by

Mark E. Hallenbeck	Yao-Jan Wu
Director	Graduate Research Assistant
Washington State Transportation Center (TRAC)	Department of Civil and Environmental Engineering
University of Washington	University of Washington

Washington State Transportation Center (TRAC)
University of Washington, Box 354802
1107 NE 45th Street, Suite 535
Seattle, Washington 98105-4631

Sponsored by

Transportation Northwest (TransNow)
University of Washington
129 More Hall, Box 352700
Seattle, Washington 98195-2700

and in cooperation with

U.S. Department of Transportation
Federal Highway Administration

May 2007

TECHNICAL REPORT STANDARD TITLE PAGE

TNW2007-05	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
4. TITLE AND SUBTITLE Review of Private Sector Data For Roadway Monitoring		5. REPORT DATE May 2007 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Mark E. Hallenbeck and Yao-Jan Wu		8. PERFORMING ORGANIZATION REPORT NO. TNW2007-05
9. PERFORMING ORGANIZATION NAME AND ADDRESS Washington State Transportation Center (TRAC) University of Washington, Box 354802 University District Building; 1107 NE 45th Street, Suite 535 Seattle, Washington 98105-4631		10. WORK UNIT NO. 11. CONTRACT GRANT NO. DTRS99-G-0010
12. SPONSORING AGENCY NAME AND ADDRESS Transportation Northwest (TransNow) University of Washington 129 More Hall, Box 352700 Seattle, Washington 98195-2700		14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES This study was conducted in cooperation with the University of Washington and the US Department of Transportation.		

16. ABSTRACT

This document presents a summary of TRAC's examination of privately collected roadway performance data in the Puget Sound Region. The review examined the accuracy and reliability of those data for use by WSDOT on roadways in the Puget Sound metropolitan region that are not covered by existing freeway surveillance systems. This project is the first step in the analysis of the potential use of privately collected datasets for arterial performance monitoring for both monitoring of signal control systems and freight (truck movement) analysis.

These test results apply only to the Seattle metropolitan region, and only to Spring 2007. The project team's conclusions from the tests are that the private sector data are currently overly conservative estimates of roadway speed and performance. At this time, it is not recommended that the data be used for arterial performance monitoring.

It is recommended that WSDOT and other roadway agencies be open to additional testing of these data sources as these improvements are made.

17. KEY WORDS Arterial Performance, Vehicle Probes , GPS ,		18. DISTRIBUTION STATEMENT	
19. SECURITY CLASSIF. (OF THIS REPORT) None	20. SECURITY CLASSIF. (OF THIS PAGE) None	21. NO. OF PAGES 13	22. PRICE

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

REVIEW OF PRIVATE SECTOR DATA FOR ROADWAY MONITORING

The demands of subscription and advertising based, privately operated, traveler information systems combined with technology advancements in the areas of global positioning systems, cell phone location tracking, and wireless communication are driving the private sector to seek new cost effective means of collecting roadway performance information. These privately collected and reported data offer roadway agencies a potential new source of roadway performance data that can be used for a wide variety of purposes. NCHRP project 70-01, "Private Sector Provision of Congestion Data" looked nationally at the accuracy and potential cost of privately collected roadway performance data.

This document presents a summary of TRAC's examination of privately collected and marketed roadway performance data in the Puget Sound Region. Several private sector firms currently offer data sets that cover subsets of the metropolitan roadway systems. One of those firms agreed to provide data for testing to the UW at no charge.

However, in order to obtain the data, the TRAC-UW team was required to sign a non-disclosure agreement prohibiting the publishing of the detailed, firm specific, test results to agencies that have not signed non-disclosure agreements with the participating private sector companies or that are not part of the UW/WSDOT team examining the potential use of those data for performance monitoring. Therefore no company references are included in this report.

The review examined the accuracy and reliability of those data for use by WSDOT on roadways in the Puget Sound metropolitan region that are not covered by existing freeway surveillance systems. This project is the first step in the analysis of the potential use of privately collected datasets for arterial performance monitoring for both monitoring of signal control systems and freight (truck movement) analysis. If the freeway data proved to be accurate, additional data would be requested which covered major arterials in the metropolitan region.

The system tested obtains roadway performance data from two major sources; data available through fixed sensor networks and a probe vehicle fleet. A proprietary fusing of these two data sources is then performed, and a "best estimate" of average vehicle speeds by roadway segment is provided based on the output of the data fusion process. The details of that proprietary process were not made available to the project team testing the accuracy of those speeds.

Because data are obtained from the probe vehicle fleet regardless of where those vehicles operate, roadway performance information is obtained on all roads that are traversed by participating, instrumented vehicles. As a result, traffic performance statistics are available for a number of roadways on which no fixed sensors have been placed. If the privately collected roadway performance statistics obtained in this manner are sufficiently accurate and reliable, their use would allow WSDOT, the UW, or other agencies to obtain roadway performance statistics without having to purchase, install, operate, and maintain fixed roadway sensors.

Three major factors affect the accuracy and reliability of the privately collected statistics. The first is the density of data points being obtained. The more data being collected on a given roadway segment, the more accurate and reliable the roadway performance statistic. The second factor is whether the performance statistics being reported are biased with respect to "general" roadway performance. For example, large trucks have different performance characteristics than passenger cars, particularly in congested conditions.[1] If a portion of the participating probe vehicle fleet are large commercial trucks, some bias may occur in the resulting performance measures. In other cases, for example when cell phones are being used as vehicle probes, there are concerns that people walking with cell phones may be viewed as "vehicles" and thus bias the roadway performance statistics. Similarly, where parallel transportation facilities (HOV lanes, general purpose lanes, arterials, rail transit lines) exist, errors may be induced when probes operating on one facility are incorrectly assigned to other facilities. Finally, on multi-lane roadways, speed biases often exist between lanes, with lanes affected by merging/diverging traffic streams often having very different speed characteristics than lanes not affected by these movements. Therefore, there is potential that the privately collected roadway performance statistics will be biased. The third factor is the data fusion process itself. Data from a number of different sources must be combined in order to compute a single speed statistic for a given roadway segment for a given reporting period. In some cases this fusion process involves combining data from several WSDOT sensors, as well as a number of individual probe speed and location data points, reported at different times and locations from within a defined roadway segment and reporting period.

The primary concerns of WSDOT regarding the privately collected data are 1) whether a sufficient number of probes are traveling on roadways for which data are desired when fixed sensors are not also available, and 2) whether the characteristics of those probes are biased in some way, thus providing inaccurate roadway performance information.

The private supplier of traffic data did not supply specific information on the number of data points available on test roadways, nor on the composition of the probe vehicle fleet, nor on the specific techniques used to obtain their probe fleet data. Therefore, the TRAC project team could only test the accuracy of the data against ground truth speed and travel time statistics and judge the combined effects of probe vehicle location/speed data collection methodology, data availability, and data fusion on the accuracy of reported statistics.

The accuracy of the data is reported primarily in qualitative terms, with some quantitative back-up of those qualitative descriptions. The reason for this rather "non-numerical" analysis is that the comparisons that could be performed did not result in direct measurements of "system accuracy." The privately reported statistics are "averages" of roadway conditions for the reporting period for road segments that vary in length. Travel conditions themselves vary over those time intervals and those geographic spaces. It is nearly impossible to compute a "true average speed" over a road segment

[1] See "Freight Data from Intelligent Transportation System Devices," WA-RD 566.1, 2003, by Hallenbeck, McCormack, Nee, and Wright.

that is at least a half- to one-mile long (or longer) roadway segment and over a time period of multiple minutes, without massive sensor systems. Even when an instrumented probe vehicle traverses a defined road segment, the data collected accurately describe only the "experiences" of that specific vehicle, not the average condition of the entire test section over the full reporting period.

Although this report does not wish to describe the "accuracy" of the privately collected data in definitive terms that overstate the precision of the comparison tests, these tests were sufficiently reliable to provide a relatively clear view of how the data should or should not be used at this point in time (spring 2007) in the Seattle area.

It should be noted that these results apply only to the Seattle metropolitan region, and only to Spring 2007. Where the number of participating vehicle probes or the number of data points being reported by each vehicle per unit of time is larger than existed during the test phase in the Seattle region, data collected using this methodology can be expected to be more accurate and reliable. Where the number of participating probe vehicles (or the data provided by each vehicle) is lower than in the test region, the results can be expected to be less accurate and reliable. This means that in general system accuracy and reliability will increase on higher volume roads and decrease on lower volume roads.

TEST METHODOLOGY

In the initial set of tests, privately collected data were compared with data from GPS-equipped floating car travel time runs performed specifically to examine the accuracy of privately reported data in locations where WSDOT surveillance data were not available. The floating car runs took place on the fringes of the Puget Sound freeway system, in corridors that do not contain WSDOT fixed sensors. The result was a direct comparison of the roadway performance data collected by a passenger car with data reported by the tested system, which are based exclusively on the performance statistics reported by its participating fleets.

The results are indicative of the "accuracy" of privately collected data for other non-instrumented freeways in the region. However, even though we can compute "average floating car speed" over the length of a test roadway segment, there is no guarantee that the floating car is itself the "true average condition" for that road segment and reporting period.

A second set of tests was performed that compared WSDOT fixed sensor (loop) data against the fused private data.

Similar to the floating car analyses, care must be taken when fixed sensor point detector data are used as the "ground truth" against which the privately collected, fused data are compared. Vehicle speed can change significantly over the length of a roadway segment, and thus the fixed sensor data may not be a perfectly "ground truth" measure of roadway conditions over the length of that segment. (For example, vehicles using that segment of roadway may experience congestion at some point in that segment that is not observed by the point detector being used for ground truth.) As the distance between fixed sensor locations increases, the ability of data from fixed point detectors to accurately reflect the "true average condition" of the roadway segment between sensors decreases. (That is, as the length of the roadway segment associated with a point measurement increases, the probability that conditions within that segment are no longer homogeneous increases. When conditions are not homogeneous, a point measure becomes a less accurate measure of average conditions within that segment.) Thus, point measurements too are at best inexact ground truth measures of "average roadway speed."

One advantage of the floating car test runs is that they collect GPS-based vehicle speed statistics every 5 seconds. This means that not only can we compare the average speed of the floating car over the length of the segment against the privately reported data, we can examine the variations in vehicle speeds that occur within a reporting segment for portions of specific reporting periods, such as the 5-minute interval used by WSDOT for reporting performance statistics. This provides considerable insight into possible reasons speeds reported by the private company might differ from the average speed on that road section reported by a floating car.

PRIVATE SECTOR DATA AVAILABILITY

The roadway network upon which privately reported traffic statistics were available during the time period these tests took place consists of the major freeways in the Puget Sound region, including King, Pierce, Snohomish and parts of Kitsap, Thurston, Mason and Grays Harbor Counties. Additional data, including some major arterials, are expected to become available on additional road segments later this calendar year (2007).

FLOATING CAR TEST RESULTS

This section describes the results of the comparison of privately collected and reported performance statistics to GPS data collected via floating car travel time surveys.

Test Locations

The floating car runs were concentrated on two major roadways, with some additional roadways covered as the floating car traveled to/from the primary highways. The intent of the floating car runs was to collect "ground truth" data on roadways where no WSDOT traffic surveillance was available in order to test the accuracy and reliability of traffic statistics based exclusively on the private probe vehicle system.

Two primary routes were selected, SR 512 in Pierce County and a small section of SR 522 in Snohomish County. Data were collected in both AM and PM peak periods to determine whether privately collected data correctly noted the onset of congestion and whether the speeds during the peak congestion periods were consistent with those observed in the floating car.

SR 512 was examined from its interchange with SR 167 to its interchange with I-5. Because private sector data currently exists on SR 522 only on the roadway section east of I-405 and west of SR 522's intersection with SR 524, floating car data were only obtained on that limited road section.

Because the floating car traversed SR 167 to reach the SR 512 roadway, a limited number of data points were also collected and compared with privately reported data for the section of SR 167 that goes from the SR 167/SR 18 interchange to the SR 167/SR 512 interchange.

SR 512 Test Results

The SR 512 test was intended to show vehicle probe data at their best because the relatively high volume should mean that probe vehicles are using the system during most of the day. SR 512 is a moderately high volume freeway, with an average annual daily traffic (AADT) of roughly 80,000 vehicles. Approximately 8 percent of those vehicles are trucks.

Floating car data were collected on the afternoon of March 28th and both the morning and afternoon of March 29, 2007. The results from comparing the floating car data for these time periods were similar. In general, the privately collected data reported considerably more congestion than was observed in the floating car runs. The private data system routinely reported speeds that were 15 to 35 miles per hour slower than those observed in the floating car runs during the peak periods. The largest differentials were during peak periods and peak directions when congestion was reported by the private system but was not noted in the floating car runs. When the floating car runs did observe heavy congestion, the private system generally reported speeds similar to those noted by the floating car data.

Essentially, during the test runs, the private system routinely reported congestion prior to its occurrence both temporally and geographically. For example, westbound on the afternoon of March 29th, speeds of 47 to 49 mph were reported on 11 of the 18 private sector company's reporting sections between 3:05 and 3:15 PM (the other sections were reported to have 62 mph speeds). The floating car run that occurred in that time period reported only two segments with speeds slower than 60 mph, and no average segment speed slower than 55 mph.

By 4:40 in the afternoon (still westbound), speeds of 29 mph for the first five reporting sections were reported by the test system. The floating car run observed average section speeds of 27 and 24 mph for the third and fourth of these sections, an excellent correlation with the private data. However, the floating car measured average section speeds of 59 and 56 mph for the first two sections and 42 mph for the fifth. The private system reported 49 mph speeds for the next six road segments, while the floating car run reported speeds of 55 mph for the first two of these and then speeds near 65 mph for all remaining SR 512 road sections.

This basic geographic and temporal pattern was repeated throughout the travel time test runs. The test system reported moderate congestion (speeds of around 49 mph) as early as 4:00 AM westbound and 5:00 AM eastbound. Eastbound speeds returned to 60+ mph throughout most of the corridor by 7:00 AM. Westbound speeds dropped to 30 mph at the eastern end of the corridor a little after 7:00 AM. Test system speeds then gradually returned to 60+ mph conditions by 8:40 AM. Floating car runs during this same time period (in both directions) simply did not experience this congestion.

SR 167 Test Results

While driving to and from SR 512, the floating car vehicle used SR 167. For all but one of these trips, both the reported floating car speeds and the test system reported speeds indicted free flow travel conditions. However, on the AM return trip, made northbound on March 29th between 7:30 and 7:45 (for the portion of the trip on SR 167 between SR 512 and SR 18), the floating car experienced congestion on SR 167.

Like most other comparisons, the test system data reported for those time periods and locations generally under-estimated the traffic speeds experienced by the floating car. On five of seven matched sections, reported data were slower than the average speed experienced by the floating car by an average of 25 mph. On the remaining two sections, the reported conditions over-estimated actual average speed by 15 mph. However, significantly better matches occurred if the reported travel speeds for the reporting period immediately following the floating car's passage were used as a comparison, rather than the reporting period that included the actual passage of the floating car. (This would be the outcome if the reporting lag inherent in use of probe vehicles meant that "current" probe vehicle speed was used to estimate the "next" reporting interval's roadway performance.) In this case, the average error was only 4 mph. Note that a similar improvement did not occur in the SR 512 test runs.

Another interesting finding from the SR 167 analysis is that when the floating car test vehicle traversed a "congested" roadway segment, considerable variation was present in the instantaneous floating car speeds observed within that road segment. For example, in one segment, the floating car reported an average speed of 25.2 mph, but the standard deviation of the 20 data points reported within that section (speed was reported every 5 seconds by the GPS devices used in the floating car study) was 10.7 mph, with a minimum reported speed of 5 mph and a maximum reported speed of 36.2 mph. This same high level of variation was present on all of the examined "congested" road segments on both SR 512 and SR 167. In one roadway segment on SR 512, a differential of over 60 mph was noted between the fastest and slowest reported speeds within a single roadway segment during a single vehicle passage of that segment.

If the probe vehicle data are obtained by the private sector infrequently (either because data are only gathered when a limited number of cell phones cross cell site boundaries, or because probe vehicles only report data periodically) the private sector data collector may only have a limited number of individual vehicle speed data points from vehicles operating along any given roadway segment during a reporting period. With limited data points available for analysis, and having those data points spread geographically along a given roadway segment, it is quite understandable that a considerable differential can exist between "reported speed" as observed by a private sector data fusion engine and "actual average speed" as reported by a floating car. Unless a reasonably large sample size is present on a given segment, the presence of one or two "extreme" (but legitimate) vehicle speed reports will significantly skew the reported average speed value.

It is unclear whether this variability is the primary cause for the significant differences frequently observed between the test data and floating car data or whether one of the other potential causes of bias is the cause.

SR 522 Results

The last of the floating car tests was performed on data from SR 522 between Woodinville and Monroe. The portion of SR 522 that is currently covered by private sector data is roughly 4 miles long. WSDOT surveillance data are available at the western end of this segment but not the eastern end. Consequently, the floating car tests were performed on the eastern end of the test section, from roughly the SR 9/SR 522 interchange east to Paradise Lake Road. This section of SR 522 carries roughly 26,000 vehicles per day, with 10 percent of these vehicles being trucks.

Floating car runs done during the morning and evening peak periods observed relatively little congestion. This limited the effectiveness of the test.

Conditions along SR 522 have recently improved as a result of significant right-of-way improvements WSDOT made on the eastern end of the roadway corridor. These improvements resulted in relatively little delay observed by the floating car test vehicle in the test sections. (Most delay on SR 522 is currently occurring on roadway sections that are east of the roadway segments reported in the private sector data sets.)

No significant delays were measured westbound in the floating car runs. This matched the private sector data, which did not report congestion on those roadways.

In the eastbound direction, some minor vehicle slowing was noted by the floating car run on the eastern most section of the test road. Much of this slowing was due to the traffic light at the intersection of SR 522 and Paradise Lake Road. However, for the last segment of the reported roadway, the congestion caused by this signal resulted in an average segment speed of between 47 and 55 mph for many of the afternoon peak floating car runs. The private sector data continued to report 65 mph speeds throughout the day. This is an accurate representation of the majority of the roadway segment's performance, but is not reflective of the signal delay at the eastern end of the roadway segment.

COMPARISON WITH WSDOT SURVEILLANCE DATA

The final set of tests of privately provided traffic performance data involved a comparison of the provided data against WSDOT inductance loop statistics. Comparisons were made against both 20-second-based, lane specific, dual loop statistics and with 5-minute single loop statistics used by WSDOT for roadway performance reporting. (The 5-minute, single loop statistics cap estimated speeds at the speed limit, whereas the private data reports actual average speeds, as does the dual loop speed computation process, so these sets of statistics are not always directly comparable.)

The primary review location was southbound on I-5 at NE 145th St (WSDOT data station ES 167D.) This private sector reporting segment corresponds well with the location of the WSDOT loop location, so that differences caused by different geographic segmentation were minimized. Only one WSDOT loop location exists within this reporting section, so the private sector data can be directly correlated to the WSDOT loop data. The WSDOT site had well functioning loops, including well tuned dual loop speed traps.

Not surprisingly, in most cases the fused, private sector reported speed data were very similar to those reported by WSDOT. However, there were differences in the data sets.

There were modest differences in “reported speed” within all three data sets during free flow conditions. Not surprisingly, one difference was that average speed was different in different freeway lanes, with the left-hand HOV lanes moving more quickly than the right-most GP lane. In general, the private sector data tracked free flow speeds in lanes 2 through 4, while lane 1 (the right-most lane) speeds were slowest of all the data sets. The private sector data were routinely faster than the speeds reported by the WSDOT freeway performance reporting statistics (which use average 5-minute volume and occupancy statistics to compute vehicle speed) in part because the WSDOT reported speeds are capped at the speed limit, and (likely) in part because the WSDOT reporting statistics are weighted to include the slower lane 1 speeds.

There were also differences in when these different systems reported congested conditions. In general, whenever the WSDOT surveillance system reported congested conditions, the private sector data also reported those slow conditions. The actual speeds reported by the private sector generally closely matched the dual loop speeds, and were slightly faster than the single loop calculations. There was often a difference, however, in the duration of the time period for which slow speeds were reported. The private data often reported that congested conditions lasted longer than the period reported by either WSDOT surveillance system. An example of this is shown in Figure 1. In Figure 1, the private sector data tracked lanes 2, 3, and 4 dual loop speeds until just before noon, when slower speeds were reported. By noon, all three measurement systems had detected and reported slower roadway speeds. However, by 2:00 PM (14:00) both WSDOT surveillance systems reported speeds returning to normal, whereas the private sector data continues to indicate slow speeds for almost another 2 hours. (An examination of traffic volume statistics shows that it is unlikely that congestion still existed at this location after 2:00 PM, as volumes had returned to normal not only at this location but also at sites both up- and down-stream from this location. Furthermore, volumes were below 1,400 vehicles per lane for the four-lane roadway, indicating that sufficient capacity was available to allow recovery to occur reasonably quickly after the incident that had occurred was cleared.)

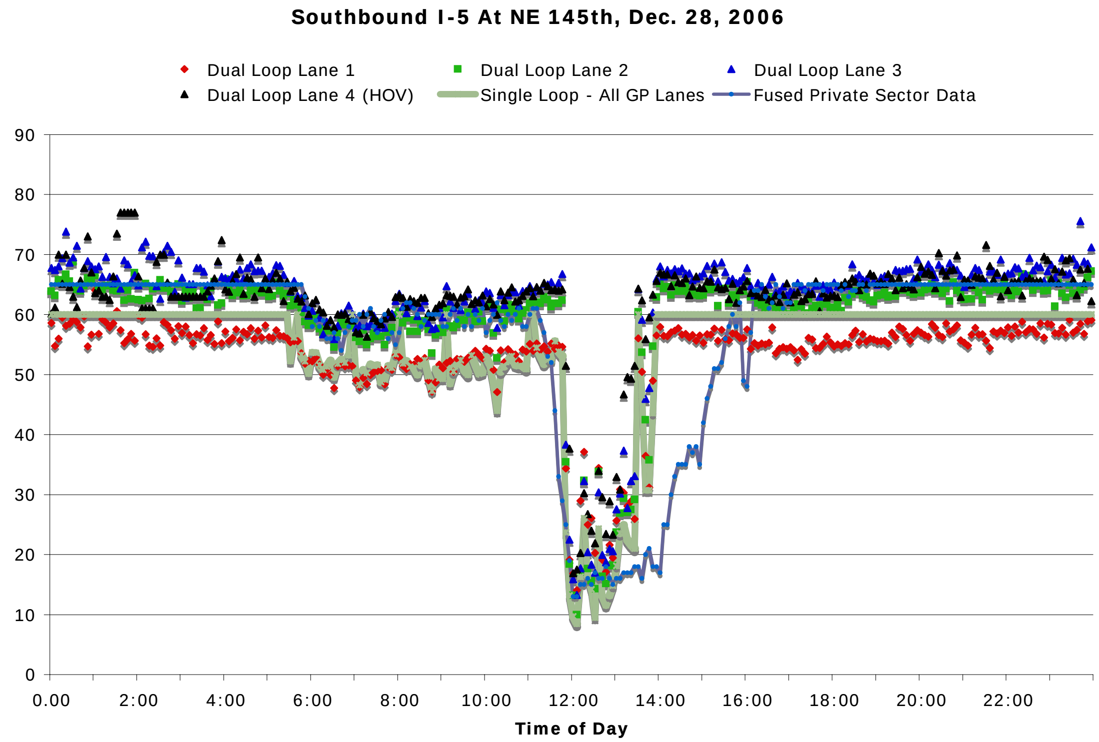

Figure 1: Speed Reporting Comparison December 28, 2007, NE 145th St.

The project team's conclusion is that the phenomenon apparent in the floating car data is also present in the roadways covered by WSDOT's surveillance system, but to a lesser degree. That is, at least on some days and in some locations, the private sector data reports more congestion than appears to exist. These differences in congestion duration are not always present. Without considerably more analysis, it is not possible to determine the extent of this condition. Without access to the data fusion algorithm and/or some of the raw data used, it is not possible to determine the cause for these "extra" segments of congestion.

CONCLUSIONS AND RECOMMENDATIONS

These results apply only to the Seattle metropolitan region, and only to Spring 2007.

The project team's conclusions from the tests are that the private sector data are currently overly conservative estimates of roadway speed and performance. They appear to under-estimate vehicle speed in some conditions and locations.

However, if WSDOT is interested in basic traveler information, the private data tested are most likely able to provide basic indications of congestion (green/yellow/red indications that can be displayed on internet web sites.) That is, the private sector data can be used to report free flow traffic conditions versus full roadways but moving conditions versus heavily congested conditions on the major freeway systems in the metropolitan area.

At this time, it is not recommended that the data be used for arterial performance monitoring, unless further evidence is presented that indicates both 1) that sufficient probe vehicle data are available to provide reliable speed estimates on the roads in question and 2) that the data fusion algorithms used by the vendor can handle the increased variability in vehicle speeds that is present on arterials as a result of stop lights and other traffic disruptions. However, these improvements are expected to occur in the near future as new traffic data collection resources continue to come to market. The result of those improvements should be continued improvement in the quality and availability of private sector data.

It is recommended that WSDOT and other roadway agencies be open to additional testing of these data sources as these improvements are made.