Eurasian Journal of Civil Engineering and Architecture

Volume 2 Issue 1

Evaluation of the Utilization of Smartphone Applications in Active Probe Vehicle Traffic Data Collection

A. Ozden*[1], A. Faghri[2]

[1*]Eskisehir Osmangazi University Department of Civil Engineering, Eskisehir, Turkey.
[2]University of Delaware, Civil & Environmental Engineering Department, Newark, DE, USA
(Corresponding Author's E-mail: *aozden@ogu.edu.tr*)

ABSTRACT

Probe vehicle data has been widely used as a mean of traffic monitoring, specifically for travel time, delay and speed measures. Technological developments in the last decade have increased the availability of technologies and tools used in probe vehicle data collection. One of the most common methods is obtaining necessary data from GPS equipped vehicles. Transportation agencies can utilize fleet data for continuous monitoring of a study area or assign a certain number of vehicles to perform data collection on a specific corridor/area. However, if the number of probe vehicles is low, the location accuracy becomes more critical. The purpose of this study is to evaluate the possibility of using existing smartphone applications in the market for collecting travel time and delay data in probe vehicles and compare with high-end GPS product. With this goal, the study aims to reduce the cost of data collection and test the accuracy and reliability of limited probe vehicle data. The data has been collected simultaneously on 102 segments in different speed, density and environmental conditions on major roadways in Delaware. The mean and variance of the travel time and delay measures are compared with statistical methods and the results revealed that there is no significant difference between smartphone application data and high-end GPS product data for travel time and delay measures. Therefore, it is emphasized that the smartphones are capable of collecting probe vehicle data for management and operation of the roadways even in specific data collections where the number of probe vehicles is limited.

Keywords: Travel Time, Delay, Probe Vehicle Data, Smartphone Applications, and Traffic Data

INTRODUCTION

Data has been gaining more attention and getting extremely valuable in the age of technology. Similarly, traffic data become a valuable asset not only to alleviate the current problems but also to predict and prevent from possible issues we may face. In recent years, we have been witnessing the emergence of new traffic data collection technologies due to an exponential growth in technology, and data sharing and analysis tools that facilitates the processing of large data sets. Since the non-intrusive traffic data collection technologies become available and cost-effective, the more real-time and near-time traffic data are being used for planning and operational purposes.

Traditional in-pavement sensors (inductive loop detectors) and pneumatic tubes have been widely used since the beginning of traffic data collection for providing key traffic measures such as traffic volume and speed data. These intrusive methods provide limited coverage with high initial and maintenance costs (Antoniou, Balakrishna, & Koutsopoulos, 2011; Leduc, 2008). Moreover, manual counts have extensively used in addition to traditional methods for the same purpose until the emergence of non-intrusive technologies. More traffic data collection technologies and tools become available in the early 2000s and brought opportunities for cost-effective and non-intrusive data collection. Turner (1996) introduced the emerging technologies and explained the cost, usability and strength/weaknesses while the technologies are in their infancy. Video and image processing by using cameras; monitoring the movement of vehicles via GPS enabled devices in the vehicles; multiple radar technologies for detecting the presence and length of vehicles; monitoring the vehicles via the phones and other Bluetooth capable devices are some of the well-known traffic applications. Turner et al. (1998) studied both traditional and emerging technologies and data collection methods that were utilized by many transportation departments, specifically in the U.S. With the emergence of new technologies and increasing effort to collect data, the amount and type of incoming traffic data have been increased significantly. Antoniou et al. (2011), and Leduc (2008) have successfully presented the traffic data collection technologies in detail, including coverage of the traffic measures, capital and maintenance costs and data collection capabilities.

The principal of probe vehicle data is to identify an individual vehicle equipped with a proper device (usually a GPS), and time-stamp the location of the same vehicle in certain intervals or at certain locations for the calculation of the required traffic measures such as travel time, average travel speed, and origin-destination (OD) matrices (Shawn M Turner et al., 1998). This method has been widely used for different purposes: i.e. fleet management, bus services, toll services, and travel time and speed data collection. For instance, all state- or city-owned GPS-equipped vehicles generate a network level travel time and average travel speed data. However, since the fleet data contains only certain type of vehicles (i.e. bus fleet), using the respective data for traffic management requires careful evaluation and adjustment. Combining the variety of data from different sources need applying proper data fusion methods (Choi & Chung, 2002; Faouzi & Klein, 2016; Faouzi, Leung, & Kurian, 2011). Traffic data from a certain type of vehicle fleet does not reflect the general traffic condition on the roads.

On the other hand, probe vehicles are used for signal timing optimization on a corridor by driving one or more probe vehicles on a certain section of the corridor for certain repetition. However, the primary limitation of the probe vehicle is that there are only limited numbers of vehicles equipped with GPS units. Additionally, cost of the GPS units increases significantly to provide high accuracy for the location of the vehicles. A high-end commercial GPS unit used in the study costs nearly $8,000 per unit (“Trimble Store,” 2018). Since the number of probe vehicles is limited in this type of studies, the accuracy of location data becomes critical. Considering the variety of applications reveals that it is essential to define the primary objectives of the data collection for selecting the most appropriate method and tools to be used.

The major advantages of probe vehicle data collection are given as low cost per unit of data, continuous and automated data collection, no disruption of traffic, and real-time data retrieval. On the other hand, few disadvantages are mentioned as privacy issues, data processing complexity, and high implementation cost in some cases such as license plate technology (Antoniou et al., 2011; Leduc, 2008; Shawn M Turner et al., 1998).

On the other hand, Global Positioning System (GPS) has been nearly ubiquitous in the last few decades providing an accurate location, velocity and timing information. Addition of new satellite systems and technological advancements enabled precise location determination on earth surface (Djuknic & Richton, 2001; Kaplan & Hegarty, 2006). The capability of portable devices (phones, watches, tablets, etc.) with accurate location determination increased the availability of applications for this purpose.

Delaware Department of Transportation (DelDOT) and University of Delaware Center for Transportation has been utilizing active probe vehicle data collection method for the traffic monitoring program in the State of Delaware. Active probe vehicle method includes operating specially equipped vehicles during peak hours and measuring travel time, delay time, delay reasons and speed of the vehicles (Faghri & Hamad, 2002). This effort has been continuing since 1995 for determination of the state of the traffic for morning and evening peak hours for commuters and weekend peak hours for beach traffic in southern parts of Delaware. The fall coverage includes nearly 4000 km driving and the summer coverage includes nearly 3000 km driving. Following Figure 1 represents the coverage of roadways for both commuter and beach traffic data collection.

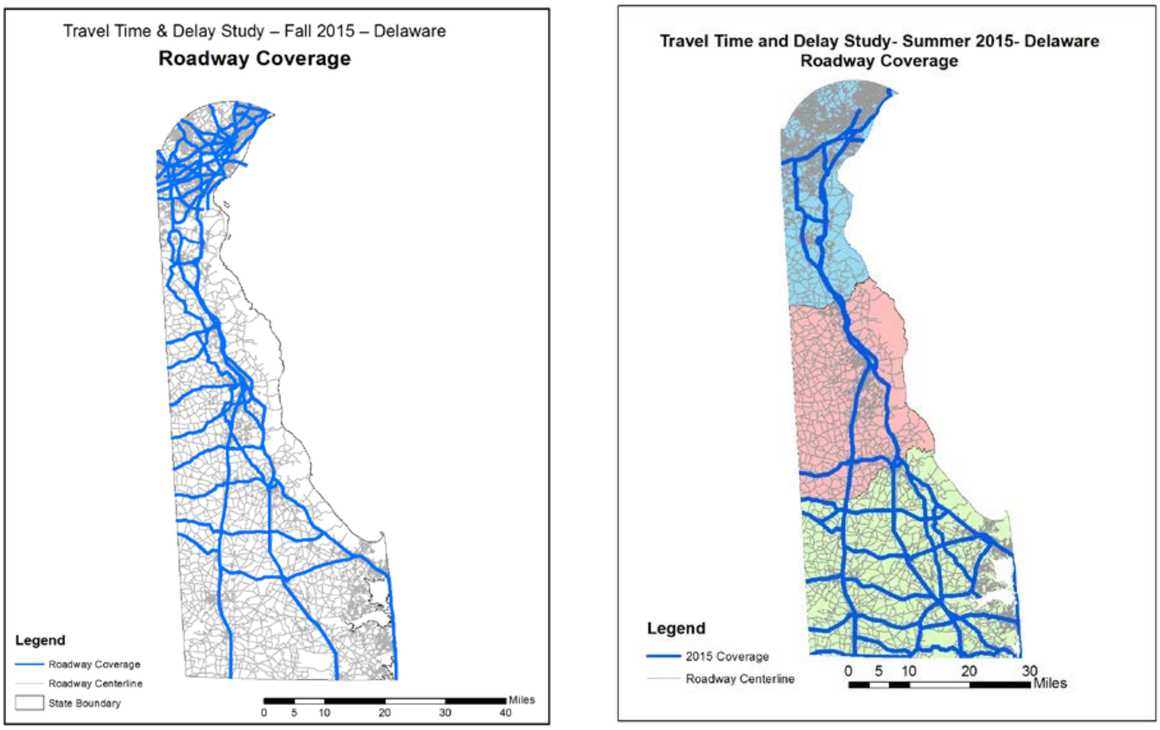

Figure 1. Roadway coverage for travel time and delay data collection

Within the scope of DelDOT traffic monitoring program and current data collection method, it is important to reduce the cost of data collection without compromising the accuracy and reliability of collected data. The primary purpose of the study is to evaluate the possibility of using smartphone applications for collecting travel time and delay data in probe vehicles to reduce the cost of data collection. If so, the need for the high-end GPS units is expected to be eliminated and it will be possible to operate more vehicles simultaneously.

MATERIALS AND METHODS

DelDOT and University of Delaware have been continuously collecting Travel Time (TT) and Delay Time (DT) data in every summer and fall for illustrating the peak hour traffic in Delaware. Data collection has been administered by using active probe vehicle method, where two GPS-equipped vehicles travel on certain road segments during peak hours. The vehicles are equipped with Trimble Geoexplorer 6000 XP portable units for data collection. Additionally, two smartphones with dedicated travel footprint applications running in it were used simultaneously. These smartphone apps, Orange and GPS Tracks, were operated in two different phones: Orange in an iPhone 5s and Tracks in a Galaxy S4.

The use of the smartphone applications has given the chance to evaluate the portable devices (smartphones, tablets, etc.), specifically GPS-enabled phones, in collecting traffic data. However, it is important to note that smartphone data has been widely used by Google, Apple, and TomTom for monitoring the traffic conditions on roadways. In this project, it is evaluated to see if the limited smartphone data is accurate and reliable enough to be used in specific traffic data collection projects that require high location accuracy.

Trimble Geoexplorer 6000 XP is a portable unit for collecting and storing geographical information. This unit was supported with an external antenna to increase the location accuracy of the operated vehicle. The horizontal and vertical accuracy is given as 30 cm and 45 cm respectively.

Orange (formally GPS Track) and GPS Tracks are travel footprint applications that are designed for tracking and recording the paths traveled, usually biked or walked. The primary audience of the applications is the ones who are exercising and jogging, not for collecting traffic data. However, the technical features provided by the application can easily be used for determination of traffic measures required.

Orange provides approximately 10-meter vertical and horizontal accuracy for determination of the location. The scan time is about one second meaning that the location of the device is recorded in every second. Both applications can provide data to be exported in .csv and .kml formats to be used in well-known data analysis and mapping applications. However, Orange provides slightly better visual features such as following the route traveled, waypoint features, and distance and speed calculations.

The data outputs of applications include latitude, longitude, timestamp, and altitude. The first three were used for the calculation of travel times and plotting the vehicle location on maps for visual evaluation. Altitude output was only used to check if the location accuracy stays within limits provided by the app developers. Additionally, Orange application provides vertical and horizontal accuracy measures for each recorded point.

The data was collected on 16 different roads and a total of 102 segments selected among the project roadway coverage presented in Figure 1. The collection of 102 segments includes combination of rural, urban, high-speed, and low-speed segments in both signalized and unsignalized corridors. The purpose of this variation is to evaluate the tested smartphone applications in various environmental and traffic conditions.

The data from each source include continuous data for a set of segments which later broken down into predetermined segments according to DelDOT segment aggregation. Then the travel time and delay measures were calculated by subtracting the time stamp values from consequent segment endpoints. Following Figure 2 presents the procedure followed for processing of the collected data. The travel time and delay definitions were taken from DelDOT project requirement for consistency.

- Mean Peak Travel Time (Seconds) – the average time in seconds that was taken to travel the length of the segment. Again, the mean peak travel time of all segments are summed in the last line of each route.
- Total Peak Delay (Seconds) – the time spent in delay traveling through the given segment. *By definition, delay is the time when vehicle speed drops below 5 miles per hour.*

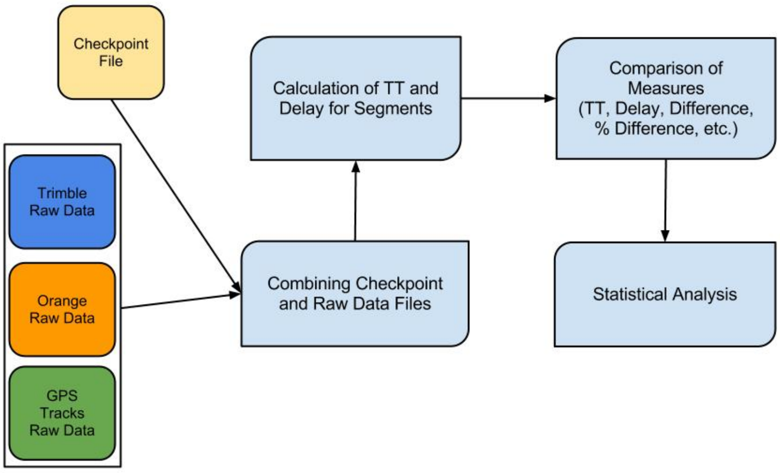

Figure 2. Data processing steps

Each one of the 102 segments is provided with three sets of travel time and delay data from Trimble, Orange and Tracks devices. The obtained data from different sources is compared with statistical analysis, specifically Analysis of Variance (ANOVA). The result of the statistical analysis provides a systematic conclusion regarding if there is a difference between data collection technologies based on given travel time and delay data.

RESULTS AND DISCUSSION

Descriptive statistics related to travel time and delay data from three different sources are presented in Table 1 and Table 2. The mean of the travel time and delay are very close in all three data sources 163.35, 163.00 and 163.08 seconds for Trimble, Orange and Tracks respectively. The standard deviation values are notably high due to the variation in the data. Length variation in segments creates variation in the range of both travel time and delay. Length range of the segments varies between 500 meters to 10 km. Thus, the travel time and delay time are between 21 and 673 seconds, and 0 and 243 seconds respectively.

Table 1. Descriptive statistics of travel time data from Trimble, Orange and Tracks (seconds)

	TRIMBLE	ORANGE	TRACKS
Mean	163.35	163.00	163.08
Standard Error	10.44	10.49	10.46
Median	135.50	136.00	136.00
Mode	125.00	125.00	125.00
Standard Deviation	105.48	105.91	105.66
Count	102	102	102

Table 2. Descriptive statistics of delay data from Trimble, Orange and Tracks (seconds)

	TRIMBLE	ORANGE	TRACKS
Mean	25.72	26.59	25.75
Standard Error	4.18	4.30	4.20
Median	0	0	0
Mode	0	0	0
Standard Deviation	42.19	43.49	42.47
Count	102	102	102

Travel time and delay time are compared in 16 roadways as presented in Figure 3 and Figure 4. It is clearly visible that both traffic measures present close values for all three data collection technologies. Additionally, it is observed that slight differences between Trimble, Orange, and Tracks are more visible in delay time compared to travel time. The one reason that may cause this variation is the smartphones location accuracy range is not well enough as compared to Trimble, 10-meter vs. 30 centimeters. Therefore, if traffic stops (mostly due to traffic lights) close to the segment endpoint, the waiting time may fall into the outside of the segment section.

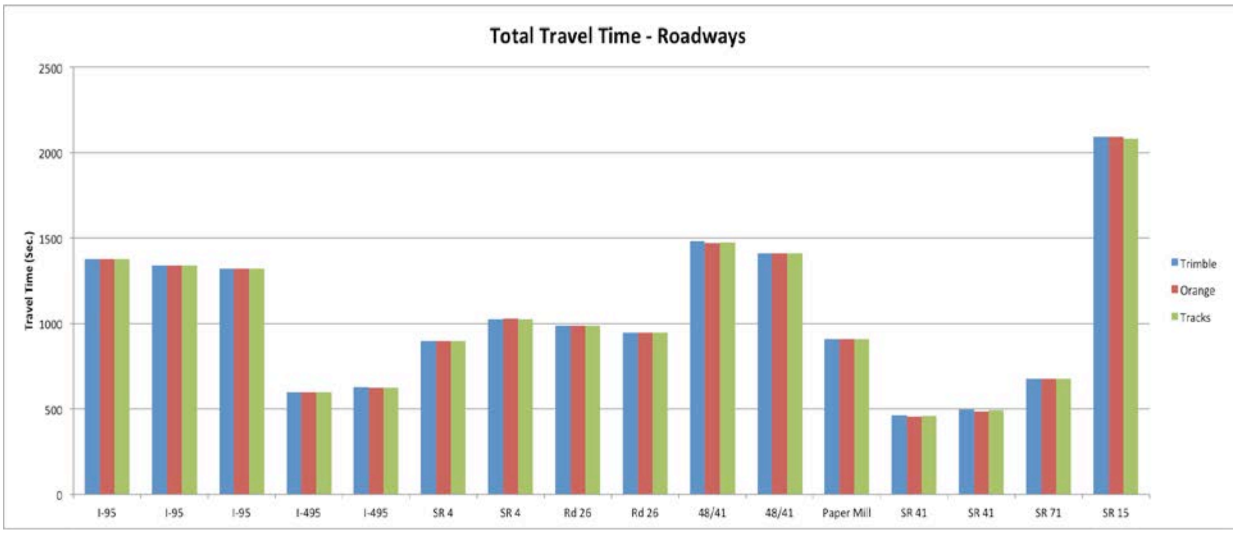

Figure 3. Comparison of travel time in aggregated road segments

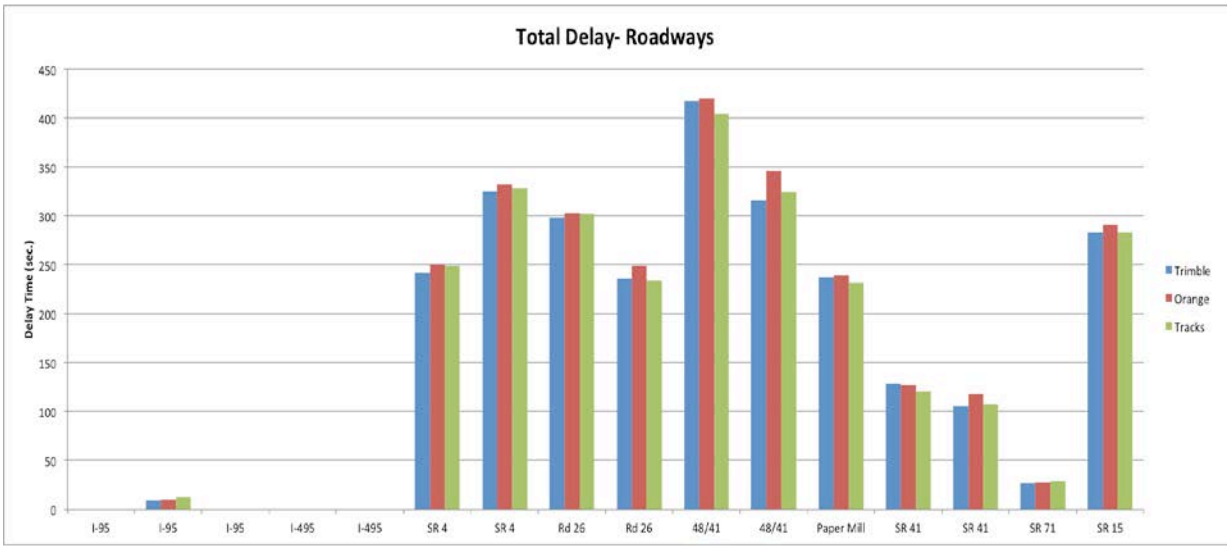

Figure 4. Comparison of delay time in aggregated road segments

ANOVA was applied to travel time and delay data of the 102 roadway segments for the comparison of the three data collection technology. The results of the statistical analysis revealed that there is no significant difference between travel time and delay time data obtained from three different sources (Trimble, Orange, and Tracks). Calculated F values are lower than the F critical value with P-values are 0.99 and 0.98 for travel time and delay time analysis respectively (Table 3 and Table 4).

Table 3. ANOVA results for travel time variable

ANOVA (Travel Time)

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	7.0065	2	3.503	0.000313	0.99968	3.0255
Within Groups	3384194	303	11168.95			
Total	3384201	305				

Table 4. ANOVA results for delay time variable

ANOVA (Delay Time)

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	50.0849	2	25.042	0.013720	0.98637	3.0255
Within Groups	553023	303	1825.16			
Total	553073	305				

The graphical examination was also carried out for the visual evaluation of the data obtained from smartphone applications. Figure 5 presents three visuals where the exact path traveled by the active probe vehicle and the location data provided by smartphone applications.

Figure 5. Orange and Tracks data plotted on maps

It is seen that smartphone data fits well on roadway maps in a variety of settings. The visual on the left is from a high-speed highway section where the speed of the vehicle is around 100-120 km/h. In the middle, the plotted map is from a populated area where there are residential and commercial buildings that may interrupt the signal to and from the smartphones. However, the location accuracy is high enough to show a well-fitted line. The visual on the right is from a rural area where the cell reception is not well and causing low accuracy in triangulation. Yet, the location data plotted on the map produces an approximate line without being smooth. The application adjusts the received location based on your previous locations, calculating your direction and speed. Therefore, the location data usually comes out smoothly as the device, probe vehicle in our case, has a continuous movement. In the map on the right, since the vehicle slows down, waits for a gap for the left turn and merges to the divided highway; the app’s location determination algorithm had some difficulties for fixing the exact location of the vehicle. However, the presented line does not encounter any issue in our study for calculating the travel time and delay time. If the incoming data increases, multiple vehicles providing data at the same time, for instance, the location data can easily be smoothed with proper algorithms as known in Google, Apple and Yandex data.

CONCLUSION

Collection and processing of traffic data is becoming extremely complex due to variety of technologies and methods, and increasing amount and type of incoming data. This also causes the necessity of understanding and processing different data formats. On the other hand, technological developments brought an opportunity for reducing the cost of data collection devices. In this project, GPS enabled smartphones are used for traffic data collection. The accuracy and reliability of the smartphone data is evaluated by comparing with high-end commercial GPS units. The data was collected simultaneously in variety of geographic and traffic density conditions simulating a real condition.

The comparative analysis revealed that there is no significant difference between Trimble GPS units and GPS enabled smartphones in terms of travel time and delay time. Smartphone data is also very reliable providing consistent data in urban/rural areas, high-speed roadways, and congested corridors. One of the most important aspects is that the smartphone applications are highly cost-effective and did not cost any money for using the required application where high end Trimble GPS Unit costs nearly $8,000 per unit with external antenna. Moreover, transportation departments can invest in developing their own application for specifically collecting and processing traffic data for creating an efficient data collection and processing method. Therefore, it is emphasized that the smartphones are capable of collecting probe vehicle data for management and operation of the roadways.

References

Antoniou, C., Balakrishna, R., & Koutsopoulos, H. N. (2011). A Synthesis of emerging data collection technologies and their impact on traffic management applications. *European Transport Research Review, 3*(3), 139–148. http://doi.org/10.1007/s12544-011-0058-1

Choi, K., & Chung, Y. (2002). A Data Fusion Algorithm for Estimating Link Travel Time. *Journal of Intelligent Transportation Systems*, *7*(3–4), 235–260. http://doi.org/10.1080/714040818

Djuknic, G. M., & Richton, R. E. (2001). Geolocation and assisted GPS. *Computers*, *34*(2), 123–125. http://doi.org/10.1109/2.901174

Faghri, A., & Hamad, K. (2002). Application of GPS in Traffic Management Systems. *GPS Solutions*, *5*(3), 52–60. http://doi.org/10.1007/PL00012899

Faouzi, N. E. El, & Klein, L. A. (2016). Data Fusion for ITS: Techniques and Research Needs. In *Transportation Research Procedia* (Vol. 15, pp. 495–512). Elsevier B.V. http://doi.org/10.1016/j.trpro.2016.06.042

Faouzi, N. E. El, Leung, H., & Kurian, A. (2011). Data fusion in intelligent transportation systems: Progress and challenges - A survey. *Information Fusion*, *12*(1), 4–10. http://doi.org/10.1016/j.inffus.2010.06.001

Kaplan, E. D., & Hegarty, C. J. (2006). *Understanding GPS: Principles and Applications* (2nd ed.). Norwood, MA: Artech House Inc.

Leduc, G. (2008). *Road Traffic Data: Collection Methods and Applications*. *JRC Technical Notes* (Vol. JRC 47967). Luxembourg. Retrieved from http://ftp.jrc.es/EURdoc/EURdoc/JRC47967.TN.pdf

Trimble Store. (2018). Retrieved March 3, 2018, from https://store.trimble.com/OA_HTML/ibeCCtdMinisites.jsp?language=US

Turner, S. M. (1996). Advanced techniques for travel time data collection. *Transportation Research Record: Journal of the Transportation Research Board*, *1551*, 51–58. http://doi.org/10.1109/VNIS.1995.518816

Turner, S. M., Eisele, W. L., Benz, R. J., & Douglas, J. (1998). *Travel Time Data Collection Handbook.*