

Evaluation of the Accuracy and Automation of Travel Time and Delay Data Collection Methods

Robert Suarez¹, Ardeshir Faghri^{2*}, Mingxin Li²

¹Airport Design Consultants, Inc., Columbia, USA

²Department of Civil & Environmental Engineering, University of Delaware, Newark, USA

Email: *faghri@udel.edu

Received October 2, 2013; revised November 5, 2013; accepted December 1, 2013

Copyright © 2014 Robert Suarez *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for SCIRP and the owner of the intellectual property Robert Suarez *et al.* All Copyright © 2014 are guarded by law and by SCIRP as a guardian.

ABSTRACT

Travel time and delay are among the most important measures for gauging a transportation system's performance. To address the growing problem of congestion in the US, transportation planning legislation mandated the monitoring and analysis of system performance and produced a renewed interest in travel time and delay studies. The use of traditional sensors installed on major roads (e.g. inductive loops) for collecting data is necessary but not sufficient because of their limited coverage and expensive costs for setting up and maintaining the required infrastructure. The GPS-based techniques employed by the University of Delaware have evolved into an automated system, which provides more realistic experience of a traffic flow throughout the road links. However, human error and the weaknesses of using GPS devices in urban settings still have the potential to create inaccuracies. By simultaneously collecting data using three different techniques, the accuracy of the GPS positioning data and the resulting travel time and delay values could be objectively compared for automation and statistically compared for accuracy. It was found that the new technique provided the greatest automation requiring minimal attention of the data collectors and automatically processing the data sets. The data samples were statistically analyzed by using a combination of parametric and nonparametric statistical tests. This analysis greatly favored the GeoStats GPS method over the rest methods.

KEYWORDS

Travel Time and Delay; Data Collection; Accuracy and Automation; GPS

1. Introduction

The concept of travel time as a means of evaluating roadway performance has been in place for almost as long as automobiles have been in production. Travel time alone serves as a fundamental quantitative measure to compute other valuable congestion information like average speed and delay time. As highlighted by various researchers, accurate measurement of travel time and delay is essential for transport planning, modelling, mental health and assessing the impact of safety countermeasures [1-3]. It can also be used to evaluate the benefits of traveler information provision strategies in a realistic stochastic environment [4-8]. Its importance in traffic management is well documented, and as a result, travel time data collection

has been integrated into congestion management legislation for several decades. In the 1990s, travel time data collection became a mandate of the Intermodal Surface Transportation Efficiency Act (ISTEA) as part of each state's Congestion Management System (CMS) [9].

Like most of the other fields of transportation, seemingly unrelated technological advances have led to the advancement of transportation technologies. Travel time data collection is no exception. The active test vehicle technique began as a manual data collection method with the use of a test vehicle, stopwatches, and copious notes and calculations. This method relied heavily on human accuracy. The use of traditional on-road sensors (e.g. inductive loops installed on major roads) for collecting data is necessary but not sufficient because of their limited coverage and expensive costs for setting up and

*Corresponding author.

maintaining the required infrastructure. Computer technology automated the process in a variety of ways, but each new method carried its own sources of inaccuracies.

Each travel time data collection technique is capable of measuring travel times within a range of degrees. Unfortunately, truly accurate travel time measurements are infeasible and therefore some degree of error is implicit within all data collection techniques available. Furthermore, some data collection techniques exhibit small errors incrementally. These errors tend to propagate throughout the entire data set compounding the problem with each passing mile. Over what distance is a small error acceptable when error propagation is of concern? Ultimately, what level of inaccuracy is acceptable?

Some techniques are also affected by human error. Inaccuracies in how the data collector perceives a given situation or errors in the manner the data is collected can cause an underestimation or overestimation in travel time and delay values. Furthermore, with the large amount of data and extensive processing, human error can be introduced easily due to carelessness, a calculation mistake, or a weak knowledge of the task [10]. Is it possible to completely eliminate human error from the data collection process?

The GPS-based techniques employed by the University of Delaware have evolved into an automated system, which provides more realistic experience of a traffic flow throughout the road links. However, human error and the weaknesses of using GPS devices in urban settings still have the potential to create inaccuracies. Additionally, the current GPS test vehicle technique contains features that may present sources of inaccuracy. The measured routes are divided by nodes or “control points” into several segments in order to more precisely determine sources of congestion in the network. How accurately should these control points be determined to achieve more uniformity for each pass before it begins to affect accuracy too severely? Also, are these control points necessary in the analysis of the data at all?

The purpose of this paper is to conduct a thorough investigation of the methods and assumptions used in the current practice of collecting travel time and delay data. Every feature that may potentially create a source of inaccuracy will be scrutinized in an effort to make the data more representative of the true performance of the roadway network. To address these inaccuracy sources, three different methods for the active test vehicle technique will be compared. The following objectives will be accomplished:

- 1) Examine each travel time and delay data collection method to determine which provides the greatest benefit to transportation system users.
- 2) Perform a detailed analysis of the possible sources of human error in an attempt to eliminate as many

sources of error as possible for the preprocessing, data collection, and post-processing phases.

3) Evaluate the specific features, methods, and assumptions adopted for this particular application of data collection.

4) Determine the optimal method for performing an active test vehicle travel time and delay data collection study with focus placed on automation and accuracy.

The remainder of this paper will be organized as follows: Section 2 will provide a literature review into the history of travel time and delay data collection. Section 3 will discuss the travel time and delay study as it is coordinated presently. Section 4 will provide an overview for the structure of the experiment conducted to evaluate the data collection method accuracy and automation. Section 5 will describe the applications used for analysis and comparison of the data collection methods. Section 6 will provide conclusions about the experimental methods and results.

2. Related Works

Introduction of GPS into the public sector provided a significant advancement in the active test vehicle data collection. In recent years, a number of applications for GPS technology have led to innovative methodologies that have direct and indirect relevance to travel time data collection. Many of these experiments offer new ways of organizing and automating the data collection procedure. In one such study, Hunter developed the Travel Run Intersection Passing Time Identification (TRIPTI) algorithm [11] for the collection and analysis of GPS-based travel time data. The TRIPTI algorithm first checks each data point location against the known location of each intersection to determine which intersections were traversed. Then the algorithm determines the crossing time of the data point nearest the exiting reference line.

Another study, created by Demers [12], was developed to create a real-time probe-based traveler information system. Probe vehicles equipped with a GPS receiver, a Pocket PC, a 3G wireless card, and CoPilot route guidance software gathered real-time travel time data that was used to make path choices for them by selecting the fastest route based on the real-time data from 200 vehicles.

Similar to Demers' approach to a more wide scale probe vehicle-based study, Shladover [13] developed a data sampling process as part of vehicle-infrastructure integration (VII). Each VII-equipped vehicle on the road serves as a probe for data transmission. Snapshots are generated periodically to identify vehicle speed, vehicle stopping, vehicle starting, and other special events.

Since 1996, Delaware Department of Transportation (DelDOT) with the help of the Civil and Environmental Engineering Department at the University of Delaware, has been measuring travel time and delay along most of

Delaware's major collectors, arterials, and freeways. When the study was first established, data was collected using the manual active test vehicle technique.

Manual data collection involved the use of stop-watches to measure the passage of time as a vehicle was driven along the corridor being studied. Because the manual method relied so heavily on human accuracy, a GPS application was adapted. The GPS data has proven to be at least as accurate as the data collected by conventional methods, and is 50% more efficient in terms of manpower.

In 2002, an analysis of GPS applications in traffic management systems was performed and published by Faghri and Hamad [14]. Their analysis compared the performance of the GPS average vehicle technique against the manual average vehicle technique. In 2005, an additional evaluation was completed to compare the manual method and the GPS method to the outputs of a distance measurement instrument (DMI) [10].

To address the concern of appropriate sample size, Faghri and Hamad performed an investigation into the number of runs required to maintain a confidence level of 95% [14]. Data was collected on trip length, trip time, and delay time which was used to compute running speed. Based on the calculated values of the average range in running speed (5.0 mph) and information provided by ITE, the minimum number of runs was computed to be two [15].

3. Solution Overview

As mentioned, many travel time and delay data collection techniques can be performed in a variety of ways ranging from manual data logging to completely automated, computer-aided record keeping. The active test vehicle technique that is used by the University of Delaware has evolved over the years from a manual method, requiring a high degree of manpower, to an automated method, involving GPS utilization.

Independent of the collection technique used, a number of experiment features must be defined prior to the field work of data collection. Objectives for the experiment must be clearly defined to ensure the proper result is achieved. Variables that will determine data collection scheduling, frequency, duration, and sample size must also be clearly defined based on the intended objectives.

3.1. Data Collection Methodology

The process of gathering travel time and delay data involves three primary steps: preprocessing, data collection runs, and post-processing. Prior to the formulation of a precise course of action, a comprehensive organization of the study was arranged to provide the strong foundation needed to reach a high level of accuracy. In order to proceed, the why, where, when, and how of data collection

was established.

The "why" or the purpose of data collection can be explained by the typical trends in Delaware's traffic. The geographical location of the state places it in the unique situation of catering to the travel needs of its own residents in population centers like Wilmington, Newark, and Dover, and to the travel needs of those traveling along the Interstate 95 corridor between Philadelphia and Baltimore. These characteristics come together to create unstable traffic conditions during rush hour periods.

Next, the "where" or the location is determined. The chosen purpose has provided a broad scope for the study area, but specific roadways were selected for study. In this particular case, most of the radial and circumferential arterials and collectors surrounding major trip generating land uses are included with special emphasis in the areas of Newark, Wilmington, and Dover, Delaware.

Though it may seem simple enough to select times for data collection, this "when" feature of the design is vital. The most challenging time a roadway will face will be the peak hour of its regular week to week flow pattern. Our selection of ideal peak hour times involved analysis of automatic traffic recorder (ATR) data. Based on the ATR data, morning peak hour was determined to occur between 7:00AM and 9:00AM and afternoon peak between 4:00PM and 6:00PM.

Recommendations for this revised methodology were provided by the Delaware Center for Transportation. The "how" of the three phases of preprocessing, data collection runs, and post-processing can be explained as follows.

3.2. Preprocessing

The preprocessing stage involves the organization of much of the study. While GPS data will be collected every second along the trip, that data must be clustered in a way that makes its presentation more manageable and useful for transportation planners. For this reason, each route that is studied is subdivided into small portions called segments. Each segment can vary in length, but is always preceded and followed by a control point. Control points are used to designate significant positions in the route where vehicles are introduced to or removed from the traffic stream or where the functional classification of the route changes.

3.3. Data Collection Run

TerraSync begins collecting data second by second as soon as a connection with the GPS satellites is secured. If human or technology errors occur, causing an inaccuracy, the run must be restarted from the last accurately recorded control point. Whenever GPS devices are used for collecting traffic data, it is important to recognize the weaknesses of using GPS devices in urban settings. GPS

satellite signals are often lost when a vehicle drives next to buildings or beneath a dense tree canopy. Therefore, data collection must sometimes be conducted using a less automated, manual technique as described previously. Given a successful run using the GPS method, each run will yield a second by second record of the route’s latitude, longitude, speed, bearing, inclination, and corresponding time stamp.

3.4. Post-Processing

During the post-processing phase, the output files from the data collection run are manipulated to yield the data. The output data yield several important pieces of information that are compiled into a tabular format for presentation purposes. Those values that are calculated from the data collection outputs include:

- Mean Travel Time—The average time in seconds that was taken to travel the length of the segment during the peak hour.
- Mean Travel Speed—The average speed of the test vehicle traveling from one control point to the next during the peak hour. This value is obtained by dividing the distance of the segment by the mean travel time.
- Total Delay—The time spent in delay traveling through the given segment during the peak hour. By DeIDOT’s definition, delay is the time when vehicle’s

speed drops below 5 miles per hour.

- Mean Running Speed—The average speed that a vehicle would travel through the route segment if delay were not experienced. This value is obtained by the following equation:

$$MRS = \frac{\text{Distance}}{\text{Mean Travel Time} - \text{Total Delay}} \quad (1)$$

- Percent Time in Delay—The percentage of time spent in delay for that route segment. The percentage is obtained by dividing the total delay by the mean travel time, then multiplying the quantity by 100.

$$\text{Percent Time in Delay} = \frac{\text{Total Delay}}{\text{Mean Travel Time}} \times 100 \quad (2)$$

- Level of Service (LOS) —LOS is a quality measure describing operational conditions within the traffic stream. The LOS is determined based on the percent difference between the weighted average posted speed and the mean travel speed.

The resulting data is then integrated into a Geographic Information System (GIS) database (Figure 1) for the assessment of traffic systems.

3.5. Preliminary Evaluation of Alternatives

In the search for a new method for travel time and delay data collection, it was essential to find alternatives that

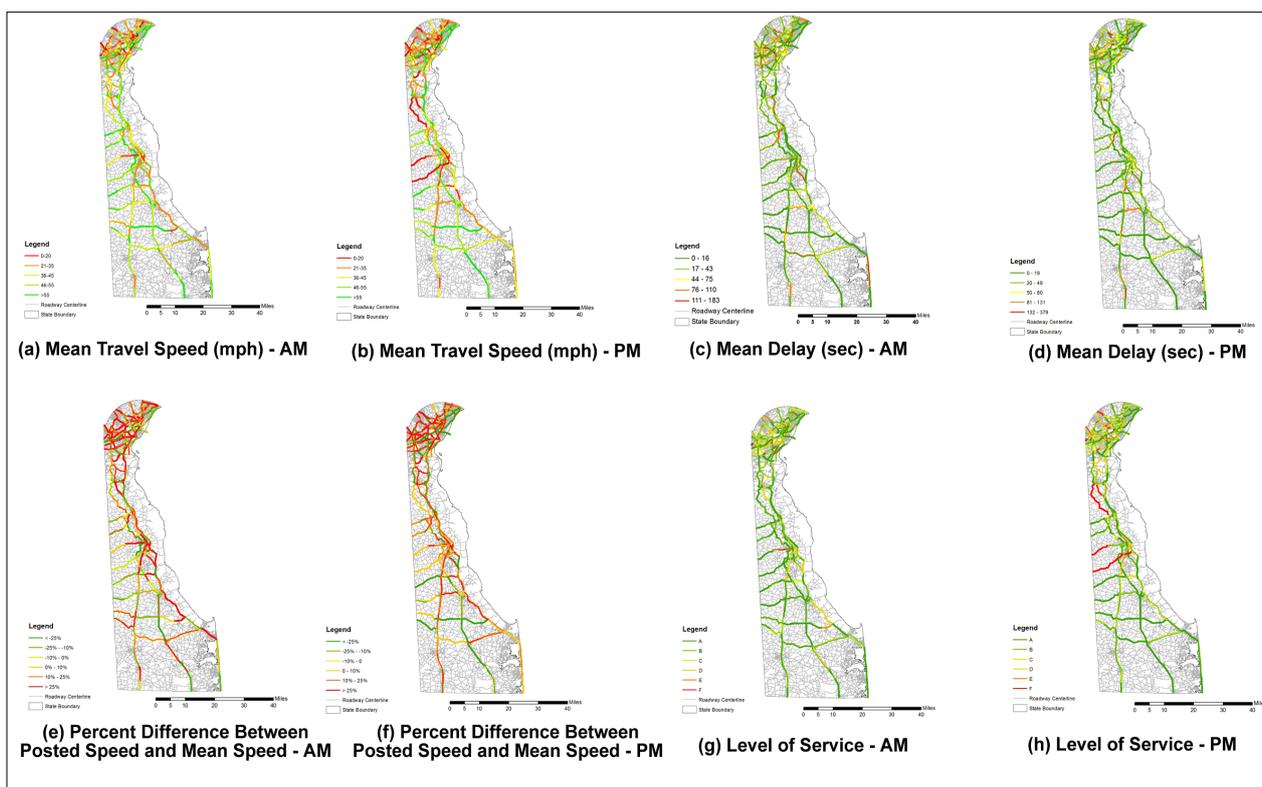


Figure 1. Travel time and delay measurements in Delaware, USA.

would address the maximum number of error sources while still delivering the data in a manner that best suits the MPO and DOT needs. Several viable alternatives were considered, but of all the travel time and delay equipment and software packages available, the following four were chosen for greater consideration as they were marketed as being useful for travel time applications:

- ESRI ArcPad 8
- Magellan Professional MobileMapper 6
- PC-Travel Software Suite 2
- GeoStat TravTime 2.0

Upon initial inspection of each alternative, it was found that while ESRI ArcPad and Magellan MobileMapper were both powerful pieces of equipment, these systems were designed with flexibility for multiple applications. Unfortunately that flexibility frequently limited the amount of automation that was possible using those systems. The remaining two systems lack the flexibility found in ArcPad and MobileMapper, but because they are so specialized, they are able to provide the highest level of automation to a travel time and delay data collection study.

PC-Travel operates in a very similar way to TravTime with a few exceptions. PC-Travel does not have a designated data storing unit and simply utilizes a laptop or PDA. This enables the data collector to view the status of the unit continuously throughout the run.

The TravTime system is able to eliminate a number of error sources related to ambiguous control point locations and missed or duplicated control points. By supplying the software with the latitude and longitude of the points of interest, all processing of the data is done by the computer. This also reduces the amount of post-processing fatigue that is experienced by the human data processor as much of the work is done by the software itself. While PC-Travel also eliminates the problem of ambiguous, missed, and duplicated control points, post-processing fatigue still persists as the start and end points of each route must be trimmed by hand.

Overall, it is believed that the GeoStat TravTime system yields greater benefits than that of the PC-Travel software.

4. Experimental Procedures

Prior to the collection and comparison of the experiment's data, experimental design details must be prepared. This preparation includes instrumentation calibration and a description of the experimental data collection procedure.

4.1. Instrumentation Calibration

The instruments that would be used in the experiment

included two stopwatches, the current Trimble GPS unit and corresponding laptop, and the Garmin GPS unit and GeoStats GeoLogger. Each of these pieces of equipment was calibrated if possible to ensure that each was operating at the highest possible level of accuracy.

The Trimble GPS unit was used in the same manner previously described. The literature made no reference to any need for GPS calibration. According to literature on the composition of GPS signal data, however, accurate position data can only be determined if the almanac and ephemeris data are up-to-date [16].

The Garmin GPS unit and GeoLogger were used in the same manner previously described. First, rules to control the logging functionality of the data logger were established using the unit's download utility software. Logging rules included:

- Speed Filter—this rule would save only speeds above 1.15 mph.
- Save Speed—this rule would record a speed for each recorded GPS point.
- Save Altitude—this rule would record altitude for each recorded GPS point.
- Time Filter—this rule would log GPS points less frequently than every one second.

4.2. Experimental Data Collection Procedure

In order to test the three travel time and delay data collection methods against each other, a number of experiment features and variables were outlined to reduce the potential for biases and errors. Of the core pieces of information needed, only travel time and delay would be collected in the field by each of the three methods independently.

First, the location of the data collection test runs was chosen. The route chosen was SR 2, Kirkwood Highway, a 4-lane major arterial serving as a main artery between Wilmington, DE and Newark, DE. The route was studied between its intersections with SR 273 and SR 7 and subdivided into four segments. The control points and unique characteristics from each segment have been outlined in the [Table 1](#). A map of the route layout is depicted in [Figure 2](#).

All three data collection methods were used simultaneously using one vehicle. A team of data collectors was assembled, all with the basic knowledge and experience needed to collect travel time and delay data using all three methods. On any given day, three members of the team were selected to perform data collection runs. The first person was responsible for driving the vehicle in the same manner as the average driver on the road and operating one stopwatch used for delay time measurements. The second person was responsible for operating the other stopwatch used for travel time measurements. The third person was responsible for operating the Trimble

Table 1. State route 2 (Kirkwood Hwy) segments.

Seg.	Eastbound	Westbound	Unique Characteristics
1	SR 273 to SR 72	SR 72 to SR 273	Originates in city center. Travels through low-speed residential.
2	SR 72 to Polly Drummund Hill Rd	Polly Drummund Hill Rd to SR 72	Covers short distance surrounded by strip malls.
3	Polly Drummund Hill Rd to MeadowoodDr	MeadowoodDr to Polly Drummund Hill Rd	Several traffic lights through mixed land uses.
4	MeadowoodDr to SR 7	SR 7 to MeadowoodDr	Several traffic lights and additional lane in each direction.

**Figure 2. Experiment study area.**

GPS unit and corresponding laptop while logging control point and delay features. While in the field, the three individuals rotated responsibilities at the conclusion of each run to reduce the likelihood of imparting a single person's data collection or driving habits onto the data.

5. Data Collection Method Evaluation

It has been established that several of the procedural points for improvement have been resolved thereby accomplishing an increased automation of the travel time data collection. The increased accuracy, however, requires that mathematical analyses be applied to ensure the viability of any new data collection method. The comparison will be conducted in two parts:

- Evaluation of GPS positioning accuracy.
- Evaluation of travel time and delay measurement accuracy.

The first evaluation will focus on the data collection by the Trimble GPS receiver and the Garmin GPS receiver. The positioning accuracy of each GPS unit varies based on a number of factors. Because the two receivers were not created by the same manufacturer, it is important to verify that they are comparable in their abilities to determine their positions precisely and accurately.

The second evaluation will focus on the experimental data of all three methods: manual, Trimble GPS, and GeoStats GPS. Using a number of methods for comparison, their accuracy in determining travel time and delay will be compared. A hypothesis will be developed to make an inference about the method's viability with respect to the two accepted methods for travel time and delay data collection. Statistical analyses of the experimental data will then be used to reach a decision to accept or reject the developed hypothesis.

5.1. GPS Position Accuracy

All measurement devices have subtle yet implicit sources of inaccuracy and GPS devices are no exception. Positioning determination from a GPS unit can never be exactly accurate or precise due to error sources such as atmospheric effects, multipath effects, clock errors, and relativity. If the units are shown to have comparable accuracies, travel time and delay measurements should theoretically be unaffected by the choice of GPS receiver used. However, if the units vary significantly in their accuracies, the travel time and delay data may also exhibit this same inaccuracy due simply to the GPS triangulation methods.

To visually and numerically compare the global position calculated by the GPS receivers, each unit's data sets were post-processed and formatted for consistency between files. The key pieces of information contained in each file include the latitude, longitude, time stamp, and velocity as recorded on a second-by-second basis by each respective GPS unit. These files were then paired based on the trip number they recorded; therefore, the Trimble GPS file and the GeoStat GPS file of trip number#1 would be paired and so on for trips #2, #3, etc. Each pair of files was imported into a GIS environment and visually overlaid with a shape file containing the roadway network for the state of Delaware.

Initially a simple visual inspection of the GIS map was performed to determine if a notable variation between the GPS unit outputs was apparent. Overall it was found that much of the GPS positioning data did not vary significantly over the entirety of each run, however, in some instances along each run, the precision between units appeared to differ somewhat (Figure 3).

A GIS tool was used to compute the distance between GPS features with corresponding time stamps. The resulting distances would measure the offsets between the positioning readings of both GPS units. From this analysis, the largest offset experienced was approximately 26 feet. Positioning inaccuracy is most likely to occur following a stop at a red traffic signal or stop sign. Because most control points are located at stop signs and traffic signals that may require the vehicle to stop, travel times have the potential to be overestimated or underestimated due to this observed offset between GPS units.

To determine if these larger offsets could impose significant travel time inaccuracy, acceptable bounds for this type of error must be determined. The shortest segment currently studied is approximately 820 feet in length with a speed limit of 35 mph. Assuming a speed of 35 mph and an underestimation and overestimation of 26 feet at each end of the segment, the travel time ranges from 15.48 to 16.49 seconds. This represents a 3% error which is within the current 5% error bounds observed by the study outline.

Based on this analysis, it can be assumed that the new Garmin GPS unit is acceptably accurate with respect to the current Trimble GPS unit. Therefore, any dramatic difference between travel time and delay readings must be attributed to the data collection procedures rather than the GPS unit triangulation.

5.2. Travel Time and Delay Accuracy

Following the post-processing of the collected data, statistical tests must be applied to determine whether the new data collection method is comparable in accuracy to the two methods already employed for travel time and delay studies. Determining which statistical tests to apply

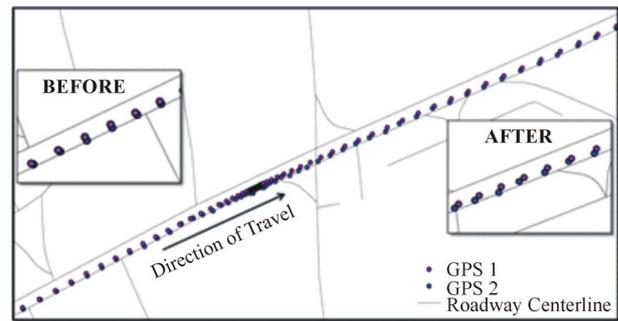


Figure 3. GPS positioning comparison.

depends on the characteristics of the data being considered. Each statistical test is based on a set of assumptions about randomness, distribution, independence, etc. In the event that a data set does not meet all of the assumption criteria, it is best to apply several different statistical tests. For analysis of the travel time and delay data, this approach of using a number of tests will be applied to achieve a clearer picture of each data collection method's relationship among the others.

Initially, a common hypothesis is developed to which each statistical test will be applied. The data will then to be analyzed using three methods:

- 1) Analysis of Means and Variances
- 2) Wilcoxon Signed Rank Test
- 3) Correlation Analysis

5.3. Hypothesis Testing

A statistical hypothesis (Table 2) is a claim about the value of a population characteristic [17]. The null hypothesis (H_0) represents the prior belief, or the claim initially assumed to be true. The contradictory claim is known as the alternative hypothesis (H_a).

The null hypothesis is understood to be true unless the sample evidence suggests that H_0 is false, in which case H_0 is rejected in favor of H_a . Hypothesis testing is used to reach a decision to reject H_0 or fail to reject H_0 .

To test the mean of a sample, there are three basic hypothesis tests available: Tests 1 and 2 represent one-tailed hypothesis tests essentially making a decision about which value is higher or lower than the other. Test 3, however, represents a two-tailed hypothesis test in which neither value is decidedly higher or lower, but simply that the values are the same or different. Because of the implicit errors present in data measurement, the absolute truth with regards to travel time and delay can never be known. Without an absolute truth to compare all other measures to, a decision about which data collection method is superior cannot be made. In this case, the two-tailed hypothesis will be adopted to show whether the three data collection methods are equal or different. Therefore, the hypothesis to be tested is as follows:

H_0 : the methods are the same

H_a: the methods are different

5.4. Analysis of Means and Variances

Two samples of measurements, even when taken from the same population, are unlikely to have exactly the same mean. By testing for significant differences between means of two samples of measurements, it can be determined whether these differences are due to chance or if they are statistically significant. To perform such a comparison, the numerical differences in the means and the variability of the measurements in the two samples are evaluated. The variability of the measurements is characterized as the standard deviation of the difference of the means (\hat{s}) and can be calculated as (3):

$$\hat{s} = \sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} \tag{3}$$

s_1, n_1 = standard deviation and number of observations in Sample 1; s_2, n_2 = standard deviation and number of observations in Sample 2. Over a normal distribution of the difference in means, $\hat{s}, 2\hat{s}$, and $3\hat{s}$ represent the 68.26, 95.46, and 99.73 percent of cases respectively. If the numerical difference in means falls outside of $\pm 3\hat{s}$, that value would be considered highly suspect and unlikely to be due to chance alone [18]. Therefore, for values that fall between $\pm \hat{s}$, H_0 will be accepted. After applying the formulas to the collected travel time and delay data, the resulting comparisons indicate that H_0 is accepted in every case (Table 3). Therefore, with regards to the sample means of each data collection method, all three methods generate comparable values and are essentially the same. These results are also displayed in Figures 4 and 5.

In the same fashion, two samples of measurements, even when taken from the same population, are unlikely to have exactly the same variance. For this comparison, the F test is used to compare the ratio of the two sample variances with the values taken from the F distribution at the 0.05 level [19]. A value of F close to 1 provides evidence that the underlying population variances are equal. If $F < 1$, F_{crit} gives the critical value less than 1 for $\alpha = 0.05$. If $F > 1$, F_{crit} gives the critical value greater than 1 for $\alpha = 0.05$. If the value of F lies between 1 and F_{crit} , H_0 will be accepted.

After applying the F -test to the collected travel time and delay data, the resulting comparisons indicate that H_0 is accepted in every case (Table 4). Therefore, with regards to the sample variances of each data collection method, all three methods generate comparable values and are essentially the same.

5.5. Wilcoxon Signed Rank Test

Nonparametric or distribution-free procedures are used in

Table 2. Hypothesis tests.

Test 1	Test 2	Test 3
$H_0: \mu \geq \mu_0$	$H_0: \mu \leq \mu_0$	$H_0: \mu = \mu_0$
$H_a: \mu < \mu_0$	$H_a: \mu > \mu_0$	$H_a: \mu \neq \mu_0$

Table 3. Analysis of means for Segment 1.

Comparison		$\Delta\mu$	\hat{s}	$2\hat{s}$	$3\hat{s}$	H_0
Travel Time EB	Manual_Trimble	0.15	9.77	19.53	29.30	accept
	Trimble_GeoStats	7.77	9.29	18.58	27.87	accept
	Manual_GeoStats	7.62	9.22	18.44	27.66	accept
Travel Time EB	Manual_Trimble	1.44	7.55	15.10	22.65	accept
	Trimble_GeoStats	1.79	7.52	15.03	22.55	accept
	Manual_GeoStats	3.23	7.53	15.07	22.60	accept
Delay Time EB	Manual_Trimble	1.36	8.41	16.82	25.23	accept
	Trimble_GeoStats	5.58	7.76	15.51	23.27	accept
	Manual_GeoStats	6.94	7.93	15.87	23.80	accept
Delay Time WB	Manual_Trimble	0.56	6.44	12.88	19.32	accept
	Trimble_GeoStats	0.10	6.46	12.92	19.38	accept
	Manual_GeoStats	0.45	6.46	12.92	19.38	accept

Table 4. Analysis of variances for Segment 1.

Comparison		F	F_{crit}	H_0
Travel Time EB	Manual_Trimble	0.97	0.52	accept
	Trimble_GeoStats	1.49	2.05	accept
	Manual_GeoStats	1.45	2.05	accept
Travel Time WB	Manual_Trimble	0.97	0.51	accept
	Trimble_GeoStats	1.01	1.93	accept
	Manual_GeoStats	0.98	0.51	accept
Delay Time EB	Manual_Trimble	1.04	1.94	accept
	Trimble_GeoStats	1.52	2.01	accept
	Manual_GeoStats	1.59	2.02	accept
Delay Time WB	Manual_Trimble	1.00	0.52	accept
	Trimble_GeoStats	0.99	0.52	accept
	Manual_GeoStats	0.99	0.52	accept

cases when the distributional assumption of normality is invalid. Past studies of travel time and delay data have determined that the data does not follow a normal distribution because of the variability in the frequency and duration of signalized intersection interruptions [10,19]. Based on the assumption of non-normal data, the nonparametric Wilcoxon Signed-Rank Test will be used. This test determines the magnitude of departures from the hypothetical median among the sample population [20].

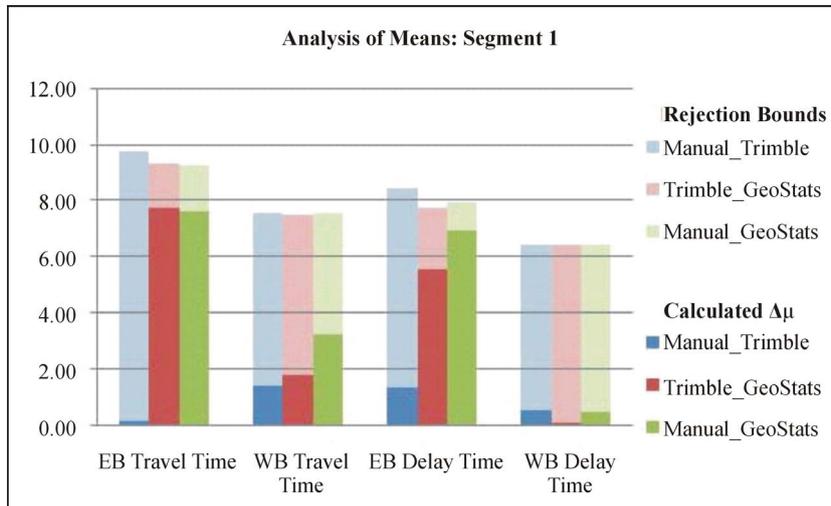


Figure 4. Analysis of means for Segment 1.

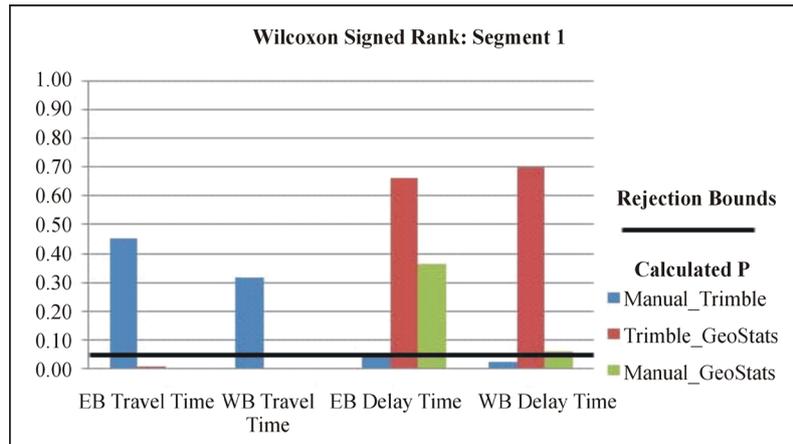


Figure 5. Wilcoxon signed rank test for Segment 1.

The measured travel times and delay times are first paired for direct comparison. If these data collection methods were all identical, their differences would always reach a median value of zero. Therefore, from each pair, a difference is calculated and applied using a two-tailed approach. These absolute values of these differences are arranged in order of magnitude and assigned ranks in ascending order. Finally, the ranks of the nonnegative differences are summed which yield the Wilcoxon W value. W values that are relatively high or relatively low suggest a large number of values are shifted above or below the hypothetical median. Additionally, each comparison generates a P-value, which represents the exact probability of obtaining a value of $|t|$ equal to or greater than that observed when H_0 is true [20,21]. At the $\alpha = 0.05$ level, if the P-value ≥ 0.05 , H_0 will be accepted.

The results of the Wilcoxon Signed Rank Test (Table 5) show that in every case, the manual method and Trimble method are considered comparable with regards to travel time measurements. Additionally, in every case,

the Trimble method and GeoStats method are considered comparable with regards to delay time measurements. These results are expected based on the manner in which the data is collected. The manual method involves human intervention for both travel time and delay data. The Trimble method involves human intervention for travel time data, but delay data is automated. The GeoStats method is automated for both travel time and delay data.

5.6. Correlation Analysis

The correlation coefficient is a measure of the extent to which two measurement variables are related to each other. The Pearson Product Moment Correlation Coefficient estimates the degree of linear association yielding a Pearson correlation coefficient, r . Values of r close to ± 1 represent strong positive or negative correlations, however, values of r close to zero represent independence between the two variables [20]. The Pearson correlation coefficient is determined by the following equation:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad (4)$$

The results of the Pearson correlation coefficient (Table 6) show a strong association between the three data collection methods for both travel time and delay time measurements. This is evident by the fact that most r -values are 0.99 and up, with the lowest r -value being 0.9222.

In a nonparametric context, the assumption of normality is no longer required. As stated previously, past studies of travel time and delay data have determined that the data does not follow a normal distribution; therefore, a nonparametric correlation will be applied as well. The Spearman Rank Correlation Coefficient is computationally equivalent to the Pearson coefficient calculated for ranks [20]. The Spearman correlation coefficient is de-

termined by the following equation:

$$r_s = 1 - \frac{6 \sum_i (r_i - s_i)^2}{n(n^2 - 1)} \quad (5)$$

The results of the Spearman correlation coefficient (Table 7) also show a strong association between the data collection methods. Again, most r -values are 0.99 and up, with the lowest r -value being 0.9430. Therefore, H_0 is accepted as the data sets from each data collection method do not vary significantly from each other.

5.7. Results of Comparison

To test the hypothesis of whether the manual method, the Trimble GPS method, and the GeoStats GPS method are statistically the same or different, the data sets were paired and analyzed using the Analysis of Means, Analy-

Table 5. Wilcoxon signed rank test for Segment 1.

	Comparison	Wilcoxon W	P	Est Median	H_0
Travel Time EB	Manual_Trimble	53.0	0.453	0.00	accept
	Trimble_GeoStats	227.0	0.007	0.80	reject
	Manual_GeoStats	221.5	0.002	0.60	reject
Travel Time WB	Manual_Trimble	69.0	0.315	0.00	accept
	Trimble_GeoStats	366.5	0.000	1.90	reject
	Manual_GeoStats	351.0	0.000	2.00	reject
Delay Time EB	Manual_Trimble	64.5	0.050	0.50	accept
	Trimble_GeoStats	45.0	0.660	0.00	accept
	Manual_GeoStats	76.5	0.363	0.10	accept
Delay Time WB	Manual_Trimble	165.5	0.025	1.00	reject
	Trimble_GeoStats	60.0	0.698	0.00	accept
	Manual_GeoStats	185.0	0.060	0.75	accept

Table 6. Pearson correlation analysis for Segment 1.

		Manual	Trimble	GeoStats
East bound	Travel Time	Manual	1	
		Trimble	0.9998	1
		GeoStats	0.9820	0.9829
	Delay Time	Manual	1	
		Trimble	0.9986	1
		GeoStats	0.9794	0.9777
West bound	Travel Time	Manual	1	
		Trimble	0.9994	1
		GeoStats	0.9996	0.9993
	Delay Time	Manual	1	
		Trimble	0.9924	1
		GeoStats	0.9924	0.9995

Table 7. Spearman correlation analysis for Segment 1.

			Manual	Trimble	GeoStats
East bound	Travel Time	Manual	1		
		Trimble	0.9980	1	
		GeoStats	0.9430	0.9430	1
	Delay Time	Manual	1		
		Trimble	0.9960	1	
		GeoStats	0.9800	0.9810	1
West bound	Travel Time	Manual	1		
		Trimble	0.9980	1	
		GeoStats	1.0000	0.9990	1
	Delay Time	Manual	1		
		Trimble	0.9840	1	
		GeoStats	0.9830	0.9990	1

sis of Variances, Wilcoxon Signed Rank Test, Pearson Correlation Analysis, and Spearman Correlation Analysis. All but the Wilcoxon Signed Rank Test suggested that the three data collection methods perform well when measuring both travel time and delay time.

The Wilcoxon Signed Rank Test appears to be the most restrictive in terms of classifying the data collection methods similarly. At a confidence level of 95 percent, H_0 was accepted for the following number of data pairs:

- Travel time - 14 out of 24 pairs
- Delay time - 11 out of 24 pairs

As only about half of the data pairs accepted H_0 , this may not be enough to fail to reject the hypothesis that all methods perform equally. As stipulated previously, it cannot be known which data collection method is superior; however, inferences can be made in an attempt to explain the results.

From the Wilcoxon Signed Rank Test, a strong relationship can be found between the manual method and the Trimble GPS method with regards to travel time measurement. Additionally, a strong relationship can be found between the Trimble GPS method and the GeoStats GPS method with regards to delay time measurement. Based on the manner in which data is collected through each method, it is expected that these relationships exist. The manual method involves human intervention for both travel time and delay data. The Trimble method involves human intervention for travel time data, but delay data is automated. The GeoStats method is automated for both travel time and delay data. The commonality between the Wilcoxon relationships and the manner of measurement coincide perfectly. From this it may be inferred that the manual method is inferior to the GPS methods with regards to delay measurement because of its dependence on human precision and accuracy. By the same argument it may be inferred that the manual and

Trimble GPS methods are inferior with regards to travel time measurement, also because of their dependence on human precision and accuracy.

6. Conclusions

This paper tested and evaluated three methods used for travel time and delay data collection. A 27-run sample set of manual, Trimble GPS, and GeoStats GPS data was collected for evaluation.

With the goal of identifying the accurate and automated method for data collection, an experimental data collection procedure was applied to gather numerous samples of data. Advancement in automation was determined subjectively by evaluating the available features in each method and considering the opportunities for error to be imparted on the data. This analysis greatly favored the GeoStats GPS method over the rest.

Accuracy, however, was determined objectively by using a number of avenues. First, each GPS unit was evaluated for positioning accuracy during data collection conditions. Secondly, the data samples were statistically analyzed by using a combination of parametric and non-parametric statistical tests. Using the Analysis of Means, Analysis of Variances, Wilcoxon Signed Rank Test, Pearson Correlation Analysis, and Spearman Correlation Analysis, overall results showed that all data collection methods perform equally well for both travel time and delay time measurements. The Wilcoxon Signed Rank Test deviated slightly verifying only that manual and Trimble GPS performed well for travel time data, and Trimble GPS and GeoStats GPS performed well for delay time data.

This evaluation involved the consideration of data measured on a signalized arterial. Additional research could be conducted to investigate the relationships between data collection methods in freeway environments.

This analysis could compare the effects of different road types and different levels of service performance on the variability between data collection methods.

Acknowledgements

This work was partially supported by the Delaware Center for Transportation (DCT) co-sponsored by the University of Delaware and the Delaware Department of Transportation. The authors express their sincere gratitude to DCT for providing their extensive data sources.

REFERENCES

- [1] L. Zhang, W. Xu and M. Li, "Co-Evolution of Transportation and Land Use: Modeling Historical Dependencies in Land Use and Transportation Decision-Making," No. OTREC-RR-09-08, 2009.
- [2] S. Humphrey, A. Faghri and M. Li, "Health and Transportation: The Dangers and Prevalence of Road Rage within the Transportation System," *American Journal of Civil Engineering and Architecture*, Vol. 1, No. 6, 2013, pp. 156-163.
- [3] R. Frey, A. Faghri and M. Li, "The Development of an Expert System for Effective Countermeasure Identification at Rural Unsignalized Intersections," *International Journal of Information Science and Intelligent System*, Vol. 3, No. 1, 2014, pp. 23-40.
- [4] M. Li, X. Zhou and N. Roupail, "Planning-Level Methodology for Evaluating Traveler Information Provision Strategies under Stochastic Capacity Conditions," *Transportation Research Board's 90th Annual Meeting*, No. 11-3002, Washington DC, 2011.
- [5] M. Li, X. Zhou and N. Roupail, "Quantifying Benefits of Traffic Information Provision under Stochastic Demand and Capacity Conditions: A Multi-Day Traffic Equilibrium Approach," *14th International IEEE Conference on Intelligent Transportation Systems Conference (ITSC)*, Washington DC, 5-7 October 2011, pp. 2118-2123.
- [6] X. Zhou, N. Roupail and M. Li, "Analytical Models for Quantifying Travel Time Variability Based on Stochastic Capacity and Demand Distributions," *Transportation Research Board 90th Annual Meeting*, No. 11-3603, Washington DC, 2011.
- [7] A. Jia, X. Zhou, M. Li, N. Roupail and B. Williams, "Incorporating Stochastic Road Capacity into a Day-to-Day Traffic Simulation and Traveler Learning Framework: Model Development and Case Study," *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2254, No. 1, 2011, pp. 112-121. <http://dx.doi.org/10.3141/2254-12>
- [8] G. Comert, "Simple Analytical Models for Estimating the Queue Lengths from Probe Vehicles at Traffic Signals," *Transportation Research Part B: Methodological*, Vol. 55, 2013, pp. 59-74. <http://dx.doi.org/10.1016/j.trb.2013.05.001>
- [9] US Department of Transportation, "A Guide to Metropolitan Transportation Planning Under ISTEA: How the Pieces Fit Together," US Department of Transportation, Washington DC, 1991.
- [10] L. Berzina, "Evaluation of Travel Time Data Collection Techniques and GPS Method Post-Processing Automation," University of Delaware, Newark, 2005.
- [11] M. P. Hunter, S. K. Wu and H. K. Kim, "Practical Procedure to Collect Arterial Travel Time Data Using GPS-Instrumented Test Vehicles," *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1978, 2006, pp. 160-168.
- [12] A. Demers, G. F. List, W. A. Wallace, E. E. Lee and J. M. Wojtowicz, "Probes as Path Seekers: A New Paradigm," *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1944, 2007, pp. 107-114.
- [13] S. E. Shladover and T. M. Kuhn, "Traffic Probe Data Processing for Full-Scale Deployment of Vehicle-Infrastructure Integration," *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2086, 2008, pp. 115-123.
- [14] A. Faghri and K. Hamad, "Application of GPS in Traffic Management Systems," *GPS Solutions*, Vol. 5, No. 3, 2002, pp. 52-60. <http://dx.doi.org/10.1007/PL00012899>
- [15] H. D. Robertson, J. E. Hummer and D. C. Nelson, "Manual of Transportation Engineering Studies," Prentice-Hall, Inc., Englewood Cliffs, 1994.
- [16] Kowoma, "GPS System," 2009. <http://www.kowoma.de/en/gps>
- [17] J. L. Devore, "Probability and Statistics for Engineering and the Sciences," Brooks/Cole Publishing, Pacific Grove, 2000.
- [18] A. D. May, "Traffic Flow Fundamentals," Prentice-Hall, Inc., Englewood Cliffs, 1990.
- [19] A. Faghri, "Application of Global Positioning System (GPS) for Monitoring Congestion," University of Delaware, Newark, 1998.
- [20] P. Sprent and N. C. Smeeton, "Applied Nonparametric Statistical Methods," Chapman & Hall/CRC, New York, 2001.
- [21] M. Hollander and D. A. Wolfe, "Nonparametric Statistical Methods," John Wiley & Sons, New York, 1973.