International Journal of Traffic and Transportation Engineering 2013, 2(6): 149-158
DOI: 10.5923/j.ijtte.20130206.03

Evaluation of Travel Time Data Collection Techniques: A Statistical Analysis

Laura Berzina[1], Ardeshir Faghri[2], Morteza Tabatabaie Shourijeh[3], Mingxin Li[2,*]

[1]McCormick Taylor, Inc., Newark, Delaware, USA
[2]Department of Civil & Environmental Engineering, University of Delaware, Newark, Delaware, USA
[3]I.S. Engineers, LLC, Houston, Texas, USA

Abstract Measuring traffic congestion is a key element of Congestion Management Systems (CMS) and is utilized to document traffic information and make decisions regarding existing traffic congestion problems. With the improvement of Global Positioning Systems (GPS) technologies, amongst other Intelligent Transportation Systems (ITS) components, various data collection procedures can be performed more efficiently. This paper examines travel time data collection methods. Data collection approach is introduced and a non-parametric statistical analysis of three different data collection methods, namely GPS, DMI (Distance Measuring Instrument), and MAN (manual), is conducted. The statistical analysis shows that the GPS approach is more consistent in terms of accuracy than the other two methods. The results also indicate that for short travel distances and trips without any delay, the three methods are accepted as being equal with a 95% confidence level. However, segment with different characteristics need to be investigated further through more data collection and analysis.

Keywords GPS, Statistical analysis, Travel time and delay, Data collection, Accuracy

1. Introduction

Measuring traffic congestion provides the data that is used in decision-making towards congestion managementby documenting congestion information such as travel time, average speed, or delay time[1]. Travel time and delay data is perhaps the most important type of data used tocalibrate and validate the simulation model that supports Advanced Traveler Information Systems (ATIS) and Advanced Traffic Management Systems (ATMS) strategies for road capacity enhancement, such as traffic signal optimization and incident management on freeways and arterials[2-6]. Quiroga and Bullock[7] grouped the most commonly used techniques for collecting travel time data into two categories: roadside techniques and vehicle techniques. With improvement of GPS technologies, various travel time data collection procedures can be performed more efficiently. Data collection techniques for travel time and delay are usually organized in four general groups[8]:

- The active test vehicle technique
- The license plate matching technique
- The passive ITS probe vehicle technique
- New and non-traditional techniques

* Corresponding author:
lmx@udel.edu (Mingxin Li)
Published online at http://journal.sapub.org/ijtte

Since the 1920's, the test car technique has been used to collect travel time, and is still one of the most common methods. Using this technique the driver behaves like an average driver, without being below or exceeding the speed of an average vehicle in the stream of traffic. In this paper, the active test vehicle technique (also known as test vehicle technique or probe vehicle technique) is studied in three different levels of its own:

- Manual (MAN)- recording, manually, elapsed time at predefined control points
- The Distance Measuring Instrument (DMI)- determining travel time and distance along a corridor based upon speed
- The Global Positioning System (GPS)- provides test vehicle position, time, and speed

The use of traditional sensors installed on major roads (e.g. inductive loops, AVI sensors)[8][9] or more recent Bluetooth sensors[10] along arterials and freeways for collecting data is necessary but not sufficient because of their limited coverage and expensive costs for setting up and maintaining the required infrastructure[11]. Moreover, traditional sensors are capable of monitoring discrete points along the most congested roadway but do not provide information about conditions on the road sections between sensors. As a relatively low-cost, high accuracy solution, GPS related data collection techniques have gained acceptance among transportation engineers and practitioners [12]. Since1997, the Delaware Department of Transportation

(DelDOT), with help from the Delaware Center for Transportation (DCT) at the University of Delaware, has been using the GPS technique to collect the travel time and delay data on all major roads and highways in the State of Delaware (Figure 1). Previous studies have proven that travel time data collection using the GPS is equally accurate as the manual test vehicle method, and that it is 50% less labor-intensive[13][14]. However, there is a lack of studies to compare the three techniques. A comparative analysis of manual and GPS techniques has shown with a 95% confidence level that the difference between the manual and the GPS technique is not statistically significant[13]. Using the parametric t-test and the F-test the variances were found to be equal and the means of the methods were not significant at the two-tailed probability level.

This paper furthers this investigation and presents an extensive non-parametric statistical comparison of the data accuracies using GPS, DMI, and manual data collection techniques.

2. Experimental Data Collection Procedure

The experimental data collection method includes the design of the process and data variables. First, about a four-mile long stretch of Route 896, south of Newark, Delaware was chosen, a map of which appears in Figure 2. The main purpose of choosing this particular section was to obtain a fewconsecutive segments,each with different characteristics, to test if the segment variables (length, number of signals, curves etc.) have an impact on travel time and delay time data collection methods.

Distance on a two-lane arterial highway was divided into three segments in each direction, for a total of six segments (Table 1). Each of the three one-directional segments had slightly different characteristics: Segment 1 has a straight section with several traffic lights, Segment 2 is a stretch of road with one light in the southbound direction and with no delay in the northbound direction, and Segment 3 is a longer segment including four traffic lights and is curved at one end of the segment.

Table 1. Road segments

Segments	*Southbound*	*Northbound*
1	SR 4 – I-95 (SB1)	I-95 – SR 4 (NB1)
2	I-95 – Old Baltimore Pike (SB2)	Old Baltimore Pike – I-95 (NB2)
3	Old Baltimore Pike – US 40 (SB3)	US 40 – Old Baltimore pike (NB3)

The data collection for each segment included measuring travel time in seconds, delay time in seconds, and distance in miles. Distance measurements obtained using the GPS and manual methods were in miles with accuracy to hundredth of a mile. Although the initial segment distance values obtained by DMI were in feet, they were converted to miles to be consistent with other measurements.

The required sample size was estimated using the recommendations of a study done by Faghri and Hamad[13]. From the given power of their sample size, it was decided to obtain thirty-two samples for each segment.

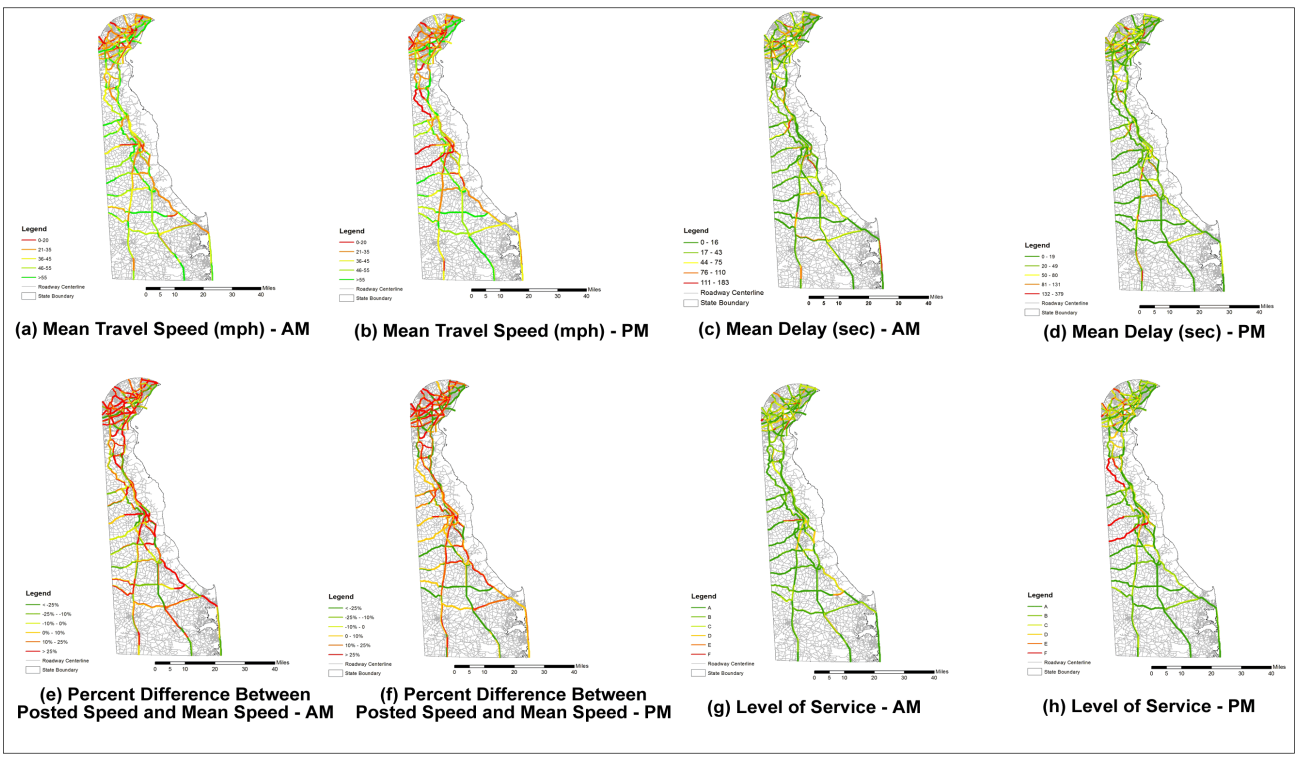

Figure 1. Example of route map output from GIS

Figure 2. The study area with three segments of the State Route 896, Newark, DE (Source: ©2013 Google Map)

The three data collection methods were used simultaneously, using one test vehicle with three operators randomly assigned to different methods to avoid human biases. In reality it was more complicated than that; the driver was using a stopwatch for delay measurements and reading the odometer at each control-point. One of the two remaining people was using the TDC-8 and a stopwatch for travel time measurements, and the second passenger was using the GPS unit. The data were collected during the off-peak hour from 9:30 AM to 11:00 AM during the weekdays. Each day a maximum number of four runs in each direction were completed. Upon the completion of the runs, the data were processed.

The hypothesis tested in this study is whether the three data collection methods are equal or different; therefore, a two-tailed hypothesis test will be applied, and as a result we test:

H_0: the methods are the same
H_A: the methods are different

3. Statistical Analysis of the Data

Before applying statistical tests, we must make sure that the correct statistical approach is used. Therefore, each data set has to be tested for normality to find if the underlying distribution of the samples is normal. The simplest normality tests usually indicate if the distribution is normal or not, but do not show the type of the distribution. Also a test of data independence will be performed.

The travel time, delay time, and distance data were collected simultaneously for all three methods, but we cannot evaluate normality for data obtained in series; therefore they will be evaluated in pairs: manual and DMI methods, DMI and GPS methods, and GPS and manual to compare and evaluate the behavior of the results.

3.1. Analysis for Normality

The purpose of the normality tests is to see whether there is detectable evidence that the normality assumptions for a given hypothesis test are being violated. Each test has its own set of assumptions to be held. For instance, the paired t-test assumes that the differences of the samples are independent and normally distributed. Most of the normality tests do not indicate the cause of the non-normality, but only detect that the normality assumption has been violated. It is also known that small sample sizes usually are in agreement with the normality assumption, and therefore, it isharder to detect the assumption violation.

There are several normality tests known, such as Lilliefor's, Kolmogorov-Smirnov (KS), Anderson-Darling (AD), andD'Agostino-Pearson. The most common goodness-of-fit test is the KS test, which helps determine if the sample population follows a specific distribution, and it is based on an empirical distribution function. The KS test hypothesis is defined as:

H_o: the data comes from a specific distribution
H_A: the data does not follow a specific distribution

The more appropriate normality test is the AD test. It is a modification of the KS test and has several advantages over that, including the fact that it is more sensitive on tails. The analysis in this section involves only two tests for normality: the AD and KS tests, which both are based on a cumulative distribution function[15][16].

Results of the normality tests are displayed in Table 2, Table 3, and Table 4. There we find that the majority of our data samples do not follow a normal distribution at the $\alpha = 0.05$ level, either with the AD or KS test as indicated by a small *p-value* indicating that normal distribution is highly unlikely.

Table 2. AD and KS test for Normality for Travel Time differences

Dir.	Seg.	Paired Diff.	AD Coef.	AD p-value	KS Coef.	KS p-value
Southbound	1	GPS-DMI	2.6998	<0.0001	1.5751	<0.01
		MAN-GPS	6.2891	<0.0001	2.3365	<0.01
		MAN-DMI	6.1506	<0.0001	2.4096	<0.01
	2	GPS-DMI	2.7138	<0.0001	1.8052	<0.01
		MAN-GPS	2.0869	<0.0001	1.6677	<0.01
		MAN-DMI	2.1009	<0.0001	1.1973	<0.01
	3	GPS-DMI	1.5411	0.0006	1.2787	<0.01
		MAN-GPS	2.0188	<0.0001	1.3799	<0.01
		MAN-DMI	1.8487	0.0001	1.2995	<0.01
Northbound	1	GPS-DMI	2.27183	<0.0001	1.5183	<0.01
		MAN-GPS	3.9349	<0.0001	2.0876	<0.01
		MAN-DMI	2.1098	<0.0001	1.2485	<0.01
	2	GPS-DMI	2.3039	<0.0001	1.4855	<0.01
		MAN-GPS	3.6470	<0.0001	2.0447	<0.01
		MAN-DMI	4.2600	<0.0001	2.1946	<0.01
	3	GPS-DMI	3.8668	<0.0001	1.9759	<0.01
		MAN-GPS	4.9326	<0.0001	2.1270	<0.01
		MAN-DMI	4.6805	<0.0001	2.1358	<0.01

According to the *p-value* in Table 2, the travel time does not follow a normal distribution; therefore we reject the H_o hypothesis in favor of H_A. Also the delay time data for paired differences in the Table 3 verify that they do not meet the normality criteria. Note that Segment 2 northbound has been excluded from the analysis, as it has zero delay.

Differences of the distance measurements have weak evidence of a normal distribution too, mostly between GPS and DMI paired differences using the AD test. In this case, KS test is not used to draw conclusions about normality, using the fact that the AD test is a modification of the KS test.

Table 3. AD and KS test for Normality for Delay Time differences

Dir.	Seg.	Paired Diff.	AD Coef.	AD p-value	KS Coef.	KS p-value
Southbound	1	GPS-DMI	1.4970	0.0007	1.3469	<0.01
		MAN-GPS	1.9539	<0.0001	1.5214	<0.01
		MAN-DMI	1.3656	0.0016	1.3289	<0.01
	2	GPS-DMI	3.0255	<0.0001	1.8627	<0.01
		MAN-GPS	5.5581	<0.0001	2.2064	<0.01
		MAN-DMI	5.8098	<0.0001	2.2670	<0.01
	3	GPS-DMI	4.6508	<0.0001	1.6757	<0.01
		MAN-GPS	1.0824	0.0077	0.9470	0.0370
		MAN-DMI	2.0696	0.0001	1.0406	<0.01
Northbound	1	GPS-DMI	0.9856	0.0134	1.1088	<0.01
		MAN-GPS	4.2785	<0.0001	1.7158	<0.01
		MAN-DMI	2.9618	<0.0001	1.4013	<0.01
	2	GPS-DMI	2.6382	<0.0001	1.8185	<0.01
		MAN-GPS	1.1959	0.0041	0.9247	0.0426
		MAN-DMI	1.4492	0.0010	1.1937	<0.01
	3	GPS-DMI	1.4970	0.0007	1.3469	<0.01
		MAN-GPS	1.9539	<0.0001	1.5214	<0.01
		MAN-DMI	1.3656	0.0016	1.3289	<0.01

Therefore, we conclude that travel time, delay time, and distance data for paired differences between GPS-DMI, MAN-GPS, and MAN-DMI are not considered normal and nonparametric statistics are more appropriate to analyze the given paired data set.

Table 4. AD and KS test for Normality for Distance differences

Dir.	Seg.	Paired Diff.	AD Coef.	AD p-value	KS Coef.	KS p-value
Southbound	1	GPS-DMI	0.7033	0.0398	0.7297	>0.15
		MAN-GPS	1.0854	0.0076	0.9903	0.0262
		MAN-DMI	1.2264	0.0018	1.0670	<0.01
	2	GPS-DMI	0.4127	0.3385	0.7340	>0.15
		MAN-GPS	0.7004	0.0675	0.9138	0.0453
		MAN-DMI	1.1163	0.0064	0.8820	0.0586
	3	GPS-DMI	0.6869	0.0729	0.8430	0.0842
		MAN-GPS	1.0386	0.0099	0.8616	0.0720
		MAN-DMI	1.1628	0.0049	0.9627	0.0331
Northbound	1	GPS-DMI	1.7960	0.0277	1.7960	0.0001
		MAN-GPS	1.2199	0.0035	1.0889	<0.01
		MAN-DMI	1.8099	0.0001	1.2183	<0.01
	2	GPS-DMI	1.8082	<0.0001	0.9960	0.0246
		MAN-GPS	1.3521	0.0017	0.9887	0.0266
		MAN-DMI	1.5244	0.0006	1.1686	<0.01
	3	GPS-DMI	0.3669	0.4326	0.6642	>0.15
		MAN-GPS	3.2831	<0.0001	1.5904	<0.01
		MAN-DMI	3.9921	<0.0001	1.7582	<0.01

3.2. Analysis for Independence

The degree to which two variables are related to each other can be measured using the Pearson Product Moment Sample Correlation Coefficient[17][18].

However, previously we have determined that our data does not follow a normal distribution; therefore, we must use the Spearman test, which is used for non-normal data. This is a test of independence based on ranks, where r_s is the correlation coefficient. The correlation value ranges between -1 to +1, where +1 indicates a perfect positive correlation between the n variables, but - 1 indicates a perfect negative relationship. There is no relationship between the variables if the correlation coefficient is zero or very close to zero. The correlation test has been applied to each data set.

From the correlation results of the travel time and delay time data in Table 5, one can clearly see that the given continuous data are closely correlated, therefore, dependent. As a result, only non-parametric or distribution-free statistics can be applied for further data analysis. Although distance measurements in Table 5 indicate little correlation between the variables, showing either weak positive or negative relationship between the data sets, non-parametric tests will also be applied to distance data.

Table 5. Spearman test for Independence

Dir.	Seg.	Paired Diff.	r_s statistic		
			Travel Time diff.	Delay Time diff.	Distance diff.
Southbound	1	GPS-DMI	0.99	0.98	0.47
		MAN-GPS	0.97	0.93	0.27
		MAN-DMI	0.97	0.95	-0.03
	2	GPS-DMI	0.99	1.00	0.44
		MAN-GPS	0.99	0.99	-0.04
		MAN-DMI	1.00	0.99	-0.25
	3	GPS-DMI	1.00	1.00	0.73
		MAN-GPS	1.00	0.99	0.21
		MAN-DMI	1.00	0.98	0.11
Northbound	1	GPS-DMI	1.00	1.00	-0.05
		MAN-GPS	1.00	0.99	-0.12
		MAN-DMI	1.00	0.99	-0.04
	2	GPS-DMI	0.92	-	-0.10
		MAN-GPS	0.96	-	0.10
		MAN-DMI	0.97	-	0.18
	3	GPS-DMI	1.00	1.00	0.25
		MAN-GPS	1.00	0.98	-0.13
		MAN-DMI	1.00	0.99	0.10

3.3. Non-parametric Statistical Tests

Parametric statistics assume that a given population follows a normal distribution with a mean μ, and variance σ^2. On the other hand, non-parametric statistics have no assumption about the underlying distribution; therefore, it is often called a distribution-free statistic[17][19]. In previous sections we determined that travel time, delay, and most of distance data are not normal and are dependent. Therefore, travel time data collection method paired data will be evaluated using the three non-parametric tests: Wilcoxon Signed Rank test, Sign test, and Minimum Chi-square test.

Wilcoxon Signed Rank Test

Wilcoxon Signed Rank test is known as the nonparametric alternative to the paired t-test. The Wilcoxon Signed Rank Test is designed to test the hypothesis about a shift in location of a median Θ in the distribution. However, there are a few assumptions to be considered:

- the paired differences are assumed to be independent,
- each paired difference comes from a continuous distribution that issymmetric, and
- the paired differences have the same median.

The null hypothesis assumes that there is a zero shift of each distribution for the differences. First, the observations of the study are taken in pairs, and then the differences between each pair are calculated. The next step includes rank R_i of the differences from the smallest to the biggest, ordered by absolute value $|Z_1|$, ..., $|Z_n|$ and summing the positive values to obtain T^+[17][17].

$$T^+ = \sum_{i=1}^{n} R_i \psi_i \tag{1}$$

where

$$\psi_i = \begin{cases} 1, & \text{if } Z_i > 0 \\ 0, & \text{if } Z_i < 0 \end{cases} \tag{2}$$

Table 6. Wilcoxon Signed Rank test for Travel Time Differences

Dir.	Seg.	Paired Diff.	*Wilcoxon W*	*p - value*	H_0
Southbound	1	GPS-DMI	36.5	0.01	reject
		MAN-GPS	22.5	0.3296	accept
		MAN-DMI	14.0	0.0029	reject
	2	GPS-DMI	20	0.0655	accept
		MAN-GPS	101	0.0686	accept
		MAN-DMI	105	0.6831	accept
	3	GPS-DMI	435	<0.0001	reject
		MAN-GPS	63.5	0.4893	accept
		MAN-DMI	496	<0.0001	reject
Northbound	1	GPS-DMI	437.5	<0.0001	reject
		MAN-GPS	50.5	0.3343	accept
		MAN-DMI	496	<0.0001	reject
	2	GPS-DMI	109	0.5375	accept
		MAN-GPS	36	0.7963	accept
		MAN-DMI	35	0.4054	accept
	3	GPS-DMI	6.5	0.0027	reject
		MAN-GPS	20	0.7389	accept
		MAN-DMI	0	0.0008	reject

We will use a two-sided test for the obtained travel time data to test H_o: $\theta = 0$ versus H_A:$\theta \neq 0$. The three pairs are grouped: GPS-DMI, MAN-GPS, and MAN-DMI. The results in Table 6 show that MAN-GPS pair has the closest travel time values for Segment 1 and Segment 3. On the other hand, Segment 2 indicates a close agreement between

MAN-DMI methods, showing that, in fact, the methods are equal at the 0.6831 and 0.4054 probability level. The delay time data in Table 7 shows that the GPS-DMI are the two closest related methods with probability levels of 0.7233, 0.3767, 0.1939, 0.1735, and 0.1754.

Table 7. Wilcoxon Signed Rank test for Delay Time Differences

Dir.	Seg.	Paired Diff.	*Wilcoxon W*	*p - value*	H_0
Southbound	1	GPS-DMI	61.5	0.7233	accept
		MAN-GPS	101	0.0853	accept
		MAN-DMI	216	0.0520	accept
	2	GPS-DMI	42.5	0.3767	accept
		MAN-GPS	80	0.2498	accept
		MAN-DMI	105.5	0.0489	reject
	3	GPS-DMI	152	0.1939	accept
		MAN-GPS	378.5	<0.0001	reject
		MAN-DMI	288	<0.0001	reject
Northbound	1	GPS-DMI	153.5	0.1735	accept
		MAN-GPS	101	0.0001	reject
		MAN-DMI	308.5	<0.0001	reject
	2	GPS-DMI	32.5	0.1754	accept
		MAN-GPS	245.5	<0.0001	reject
		MAN-DMI	217	0.0004	reject
	3	GPS-DMI	61.5	0.7233	accept
		MAN-GPS	101	0.0853	accept
		MAN-DMI	216	0.0520	accept

Table 8. Wilcoxon Signed Rank test for Distance Differences

Dir.	Seg.	Paired Diff.	*Wilcoxon W*	*p - value*	H_0
Southbound	1	GPS-DMI	31	<0.0001	reject
		MAN-GPS	212	0.3307	accept
		MAN-DMI	82.5	0.0007	reject
	2	GPS-DMI	254	0.9064	accept
		MAN-GPS	190.5	0.2597	accept
		MAN-DMI	166	0.1081	accept
	3	GPS-DMI	526	<0.0001	reject
		MAN-GPS	116	0.0165	reject
		MAN-DMI	309.5	0.3949	accept
Northbound	1	GPS-DMI	496	<0.0001	reject
		MAN-GPS	224.5	0.4600	accept
		MAN-DMI	412.5	0.0055	reject
	2	GPS-DMI	152	0.0362	reject
		MAN-GPS	209.5	0.4504	accept
		MAN-DMI	151.5	0.0354	reject
	3	GPS-DMI	11	<0.0001	reject
		MAN-GPS	92	0.0013	reject
		MAN-DMI	469	0.0001	reject

Distance measurements in Table 8, however, show that if the distance is shorter then it is more likely the methods will perform equally, but as soon as the distance increases, the value fluctuates. We can conclude that measuring delay time using the GPS and the DMI methods is not significantly different; therefore, the manual method is the one that performs different.

Table 9. Sign test for Travel Time Differences

Dir.	Seg.	Paired Diff.	*Sign statistic*	*p - value*	*Pos.*	*Neg.*	*Zero*
Southbound	1	GPS-DMI	3	0.0044	3	16	13
		MAN-GPS	5	1.0000	5	6	21
		MAN-DMI	1	0.0005	1	15	16
	2	GPS-DMI	3	0.0923	3	10	18
		MAN-GPS	11	0.2101	11	5	16
		MAN-DMI	9	0.6636	9	12	10
	3	GPS-DMI	29	<0.0001	29	0	3
		MAN-GPS	7	0.6291	7	10	15
		MAN-DMI	31	<0.0001	31	0	1
Northbound	1	GPS-DMI	29	<0.0001	29	1	2
		MAN-GPS	7	0.7744	7	5	20
		MAN-DMI	31	<0.0001	31	0	1
	2	GPS-DMI	10	1.0000	10	9	13
		MAN-GPS	6	1.0000	6	6	20
		MAN-DMI	6	0.7539	6	4	22
	3	GPS-DMI	1	0.0034	1	12	19
		MAN-GPS	4	1.0000	4	5	23
		MAN-DMI	0	0.0005	0	12	20

Table 10. Sign test for Delay Time Differences

Dir.	Seg.	Paired Diff.	*Sign statistic*	*p - value*	*Pos.*	*Neg.*	*Zero*
Southbound	1	GPS-DMI	7	0.8036	7	9	14
		MAN-GPS	12	0.0768	12	4	14
		MAN-DMI	18	0.0227	18	6	8
	2	GPS-DMI	7	0.5488	7	4	19
		MAN-GPS	11	0.1185	11	4	16
		MAN-DMI	13	0.0213	13	3	15
	3	GPS-DMI	12	0.6636	12	9	11
		MAN-GPS	25	<0.0001	25	3	4
		MAN-DMI	22	<0.0001	22	2	8
Northbound	1	GPS-DMI	12	0.3833	13	8	11
		MAN-GPS	22	<0.0001	22	2	8
		MAN-DMI	23	<0.0001	23	2	7
	2	GPS-DMI	5	0.4240	5	9	18
		MAN-GPS	21	<0.0001	21	1	10
		MAN-DMI	20	<0.0001	20	1	11
	3	GPS-DMI	7	0.8036	7	9	14
		MAN-GPS	12	0.0768	12	4	14
		MAN-DMI	18	0.0227	18	6	8

Sign Test

We will use the Signtest to see how much our travel time, delay time, and distance values fluctuate and to which side they tend to go. Here we test the H_o hypothesis that $\theta = \theta_0$ versus H_A: $\theta \neq \theta_0$, where θ is the median of the population differences. Generally, Sign test answers *how often*, where the Wilcoxon Signed Rank test answers *how much,* our data is skewed. Sign test counts the positive differences between the paired differences. The continuous data obtained for the Sign test indicate that the closest agreement for travel time data is reached between manual and GPS techniques as indicated in Table 9.

On the other hand, delay time data indicates that manual technique has the poorest performance (Table 10), but GPS-DMI technique agrees at the 0.8036, 0.5488, 0.6636, 0.3833, and 0.4240 probability levels, which also correspond with the observations using Wilcoxon Signed Rank test. Distance measurements show a better agreement between manual and DMI techniques (Table 11).

Using the Sign test we have determined that the medians of the paired differences do not provide sufficient evidence that the differences between all the methods are zero. The data has shown that each segment and each pair of differences fluctuates for undetermined reasons. The Sign test works only on the positive and negative signs of the data, but our data represents a vast amount of zero values for travel time and delay time paired differences. As a result, in the next section we will propose an additional approach to deal with zero values.

Table 11. Sign test for Distance Differences

Dir.	Seg.	Paired Diff.	*Sign statistic*	*p - value*	*Pos.*	*Neg.*	*Zero*
Southbound	1	GPS-DMI	1	<0.0001	1	31	0
		MAN-GPS	9	0.0201	9	23	0
		MAN-DMI	10	0.0501	10	22	0
	2	GPS-DMI	16	1.000	16	16	0
		MAN-GPS	17	0.7201	17	14	1
		MAN-DMI	15	1.000	15	16	0
	3	GPS-DMI	31	<0.0001	31	1	0
		MAN-GPS	11	0.2005	11	19	2
		MAN-DMI	22	0.0501	22	10	0
Northbound	1	GPS-DMI	31	<0.0001	31	1	0
		MAN-GPS	18	0.5966	18	14	0
		MAN-DMI	19	0.3771	19	13	0
	2	GPS-DMI	11	0.1102	11	21	0
		MAN-GPS	10	0.0708	10	21	1
		MAN-DMI	11	0.1102	11	21	0
	3	GPS-DMI	2	<0.0001	2	30	0
		MAN-GPS	5	0.0001	5	27	0
		MAN-DMI	29	<0.0001	29	3	0

Dealing with Ties and Zeros (Minimum Chi-square test)

Special procedures need to be considered if the data contains ties or zero values. The ranks of the tied values need to be averaged before assigning the sign. This procedure is known as the *average rank procedure*[20], and it has been applied throughout the tests performed on our data.

The data in the previous sections have showed that almost every data set of the paired differences contains at least one or several zero values. Wilcoxon testcan be advised to drop the zero values before ranking the data and apply the test to the remaining values; this procedure is known as *reduced sample procedure*[20]. The procedure was applied to the Wilcoxon Signed Rank test and the Sign test. However, some of our data for differences between methods contain a large amount of zeros. In this case, it is suggested to apply a different procedure[17].

The procedure proposed is based on the *minimum Chi-square statistic*[21]. If there is no difference in the methods, the probability of a positive difference $P_+(\theta)$ is equal to the probability of a negative difference $P_-(\theta) = (1 - P_0(\theta))/2$, where $P_0(\theta) = \theta$, for some $0 \le \theta \le 1$, is the probability of no difference. The Chi-square statistic for any $0 \le \theta \le 1$ is:

$$\chi^2(\theta)=\frac{(O_- - nP_-(\theta))^2}{nP_-(\theta)}+\frac{(O_0 - nP_0(\theta))^2}{nP_0(\theta)}+\frac{(O_+ - nP_+(\theta))^2}{nP_+(\theta)} \quad (3)$$

Plugging $\hat{\theta}=O_0/n$, the maximum likelihood estimator of θ, into $\chi^2(\theta)$, the minimum Chi-square statistic is equal to:

$$\chi^2(\hat{\theta})=\frac{(O_+ - O_-)^2}{O_+ + O_-} \quad (4)$$

For large sample sizes *n*this statistic has an approximate Chi-square distribution with one degree of freedom, when the hypothesis of no difference is true. It follows that in this case,

$$|T|=\frac{|O_+ - O_-|}{\sqrt{O_+ + O_-}} \quad (5)$$

has approximately the distribution of $|Z|$, where Z is a standard normal random variable. To solve Equation4,first the numbers of positive and negative values for each pair between the data methods are obtained. O_+ is the number of positive signs and O_- is the number of negative signs between the paired differences.

An example of paired differences, that contains 23 zero values, is given below. We test H_o: $P_+ = P_-$ and H_A: $P_+ \neq P_-$.An example is given for MAN-GPS travel time of northbound Segment 3:

$$|T|=\frac{|4-5|}{\sqrt{4+5}}=\frac{|-1|}{3}=0.333$$

and

$$\text{p-value} \approx P(|Z|>0.333)=2P(Z>0.333)$$
$$=2(1-P(Z \le 0.333)=(2(1-0.6304))=0.739$$

The results of the minimum Chi-square test are displayed in Table 12. According to the calculated values, manual and GPS methods for travel time measurements are not significantly different. In fact, the probability of them being equal is quite high. Comparing GPS and DMI methods and MAN-DMI methods, the results indicate that differences between the methods are not significant for Segment 2 only. For other segments, we reject the hypothesis that these data collection methods are equal. Differences between delay time measurements confirm that GPS and DMI methods are not significantly different when measuring delay times. However, comparisons between manual andGPS, as well as between manual and DMI, demonstrate that the pairs of these methods are not equal.

Table 12. Minimum Chi-Square test for Positive and Negative Differences

Dir.	Seg.	Paired Diff.	Travel Time	Delay	Distance
			Probability		
Southbound	1	GPS-DMI	0.0028	0.6170	0.0000
		MAN-GPS	0.7630	0.0456	0.0134
		MAN-DMI	0.0004	0.0144	0.0338
	2	GPS-DMI	0.0522	0.3658	0.8574
		MAN-GPS	0.1336	0.0706	0.0832
		MAN-DMI	0.5126	0.0124	0.8574
	3	GPS-DMI	0.0000	0.5126	0.0000
		MAN-GPS	0.4668	0.0000	0.1426
		MAN-DMI	0.0000	0.0000	0.0338
Northbound	1	GPS-DMI	0.0000	0.2752	0.0000
		MAN-GPS	0.5638	0.0000	0.4794
		MAN-DMI	0.0000	0.0000	0.2888
	2	GPS-DMI	0.8186	-	0.8348
		MAN-GPS	1.0000	-	0.0482
		MAN-DMI	0.5270	-	0.0770
	3	GPS-DMI	0.0022	0.2850	0.0000
		MAN-GPS	0.7390	0.0000	0.0002
		MAN-DMI	216	0.0520	accept

Distance data in the Table 12 suggests that manual and GPS method pairs, and manual and DMI method pairs, perform equally. Once again, Segment 2 has no significant difference between all methods.

3.4. Additional Segment Analysis

The statistical analysis has shown that for certain roadway segments of the data collection the methods follow a pattern. It has been observed that Segment 2 in both directions has less difference between the methods.

The tests possibly indicate that the methods vary according to the segment characteristics that have been described previously. The results might be a sign of correlation between the number of traffic lights on each segment and the impact on three data collection methods, as they become more apparent where segments with more traffic lights are present.

Therefore, additional linear relationships measuring travel time and delay time per mile versus number of signals per mile have been created. Table 13 represents for each method, the linear regression of travel time and delay on the number of signals per mile. Results show that R^2 values are fairly close to unity, which evidences the relationship between both the travel time and delay and the number of signals per mile.

Moreover, a multiple regression analysis was conducted considering average travel time and average delay time as dependent values, and the average length of each segment and the number of lights per segment as independent values. *L* is the average length of segment in miles, and *s* represents the average number of lights per segment. Results are presented in Table 14.

The result confirms that a relationship between the segment characteristics and travel time or delay time exists. However, one must be careful to draw conclusions, because the sample size of six is rather small to be sure about the results.

Table 13. Relationship between per mile Travel Time and Delay (y) and Signal Density (Number of signals per mile −x) for different methods

	Travel Time	Delay Time
GPS	$y=20.437x + 66.05$ $R^2=0.85$	$y=10.619x - 1.5731$ $R^2 = 0.6865$
DMI	$y=20.49x + 65.885$ $R^2 = 0.8542$	$y=10.752x - 1.7301$ $R^2 = 0.694$
MAN	$y=20.241x + 66.805$ $R^2 = 0.8438$	$y=11.519x - 2.161$ $R^2 = 0.7127$

Table 14. Multiple Regression Analysis of Average Travel Time and Delay (y) vs. Number of Signals per Segment (s) and Average Segment Length in miles (L)

	Travel Time	Delay Time
GPS	$y = 21.85s + 46.39L + 17.76$ $R^2 = 0.954$	$y = 11.12s - 6.50L + 4.14$ $R^2 = 0.515$
DMI	$y = 21.93s + 45.36L + 18.59$ $R^2 = 0.956$	$y = 11.34s - 7.50L + 4.78$ $R^2 = 0.528$
MAN	$y = 21.79s + 46.58L + 17.97$ $R^2 = 0.953$	$y = 12.09s - 7.28L + 4.09$ $R^2 = 0.546$

4. Conclusions and Recommendations

The Wilcoxon Signed Rank Test, with a confidence level of 95%, has accepted H_o hypothesis for following number of data pairs:

- Travel time - 10 out of 18 data pairs
- Delay time - 8 out of 15 data pairs, and
- Distance - 7 out of 18 data pairs.

On the other hand, the Sign test provides following results:

- Travel time - 10 out of 18 data pairs
- Delay time - 7 out of 15 data pairs, and
- Distance - 11 out of 18 data pairs.

As a result, we conclude that at least a half of the data pairs rejects the H_o hypothesis and accepts H_A hypothesis that the methods are not equal. The results show that the three travel data collection methods perform different for different data sets. We have reached a conclusion that manual and GPS methods perform equally well when collecting travel times on any of the segments. When conducting delay time studies, the GPS and DMI methods perform equally well. The manual method, assumingly, includes the difficulty of precisely reading the car's odometer, which is needed when estimating a speed of 5 mph and below.

Distance measurements, using the three methods, indicate that manual-GPS and manual-DMI methods perform equally well. However, because we have rounded distance values for DMI to match the decimal places of odometer readings, we may get insufficient results to correctly conclude which method is performing better. The initial distance measurements collected by DMI represented rather close values in feet, sometimes ranging for only a couple of feet.

On the other hand, the GPS collects the position data based on latitude and longitude, therefore a flat surface. Every great change in roadway grade introduces a small error in GPS distance measurements. Nevertheless, according to the statistical analysis, we would suggest the GPS method as the primary travel data collection method because we are interested in travel time and delay time consistent measurements.

The comparison of the travel time data methods evaluated in this research, provided information about differences between the methods. It also indicated that the differences could be related to the road segment characteristics. However the sample size was not sufficient to make the final conclusions in this regard.

An additional research could be conducted to investigate if the travel time method variations are based on road characteristics. This analysis could also be carried out comparing method performance during peak hour and compare with the off-peak hour performance. Also some odd results obtained could be attributed to rounding of the data and thus, it would be useful to perform more intensivedistance measurement comparison, and use the obtained values without data rounding.

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] H. Mahmassani, R. Mudge, T. Hou and J. Kim , "Use of Mobile Data for Weather- Responsive Traffic Management Models," U.S. Department of Transportation, 2012.

[2] A. Jia, X. Zhou, M. Li, N. Rouphail 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, pp. 112-121, 2011.

[3] M. Li, X. Zhou and N. Rouphail, "Planning-Level Methodology for Evaluating Traveler Information Provision Strategies under Stochastic Capacity Conditions," in Transportation Research Board's , Washington D.C., 2011.

[4] M. Li, X. Zhou and N. Rouphail, "Quantifying Benefits of Traffic Information Provision under Stochastic Demand and Capacity Conditions: A Multi-day Traffic Equilibrium Approach," in IEEE Intelligent Transportation Systems Conference (ITSC), 14th international IEEE conference on (pp. 2118-2123). IEEE., 2011.

[5] X. Zhou, N. Rouphail and M. Li, "Analytical Models for Quantifying Travel Time Variability Based on Stochastic Capacity and Demand Distributions," in Transportation Research Board 90th Annual Meeting (No. 11-3603), 2011.

[6] G. Comert, "Simple analytical models for estimating the queue lengths from probe vehicles at traffic signals," Transportation Research Part B: Methodological, vol. 55, pp. 59-74, 2013.

[7] C. A. Quiroga and D. Bullock, "Travel time studies with global positioning and geographic information systems: an integrated methodology," Transportation Research Part C: Emerging Technologies, vol. 6, no. 1, pp. 101-127, 1998.

[8] S. Robinson and J. Polak, "Modeling urban link travel time with inductive loop detector data by using the k-NN method," Transportation Research Record: Journal of the Transportation Research Board, vol. 1935, no. 1, pp. 47-56, 2005.

[9] F. Dion and H. Rakha, "stimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates," Transportation Research Part B: Methodological, vol. 40, no. 9, pp. 745-766, 2006.

[10] A. Haghani, M. Hamedi, K. Sadabadi, S. Young and P. Tarnoff, "Data collection of freeway travel time ground truth with bluetooth sensors," Transportation Research Record: Journal of the Transportation Research Board, vol. 2160, no. 1, pp. 60-68, 2010.

[11] S. Turner, W. L. Eisele, R. J. Benz and D. J. Holdener, "Travel time data collection handbook," Federal Highway Administration, Report FHWA-PL-98-035, 1998.

[12] B. Coifman and S. Krishnamurthy, "Vehicle reidentification and travel time measurement across freeway junctions using the existing detector infrastructure," Transportation Research Part C: Emerging Technologies, vol. 15, no. 3, pp. 135-153, 2007.

[13] A. Faghri and K. Hamad, "Application of GPS in Traffic Management Systems," GPS Solutions, vol. 5, no. 3, pp. 52-60, 2002.

[14] D. Papinski and D. Scott, "A GIS-based toolkit for route choice analysis," Journal of Transport Geography, vol. 19, no. 3, pp. 434-442, 2011.

[15] J. Romeu, "Anderson–darling: a goodness-of-fit test for small sample assumptions," START, vol. 10, no. 5, pp. 1-6, 2004.

[16] S. Dubey, B. Ponnu and S. Arkatkar, "Time Gap Modeling under Mixed Traffic Condition: A Statistical Analysis," Journal of Transportation Systems Engineering and Information Technology, vol. 12, no. 6, pp. 72-84, 2012.

[17] M. Hollander and D. Wolfe, Nonparametric Statistical Methods, 2nd ed., New York: John Wiley & Sons, Inc., 1999.

[18] S. Washington, M. G. Karlaftis and F. L. Mannering,

Statistical and econometric methods for transportation data analysis, CRC press, 2011.

[19] D. Anderson, D. Sweeney and T. Williams, Introduction to Statistics: An Applications Approach, St.Paul, Minnesota: West Publishing Company, 1981, pp. 405-406.

[20] J. Pratt, "Remarks on zeros and ties in the Wilcoxon signed rank procedures," Journal of the American Statistical Association, vol. 54, no. 287, pp. 655-667, 1959.

[21] A. Mood, F. Graybill and C. Duane, Introduction to the Theory of Statistics, New York: McGraw-Hill, Inc., 1974, pp. 146-147.