“Development of a Statewide Travel Speed Survey”

Prepared for

Delaware Department of Transportation

By

David P. Racca

September 2013

Center for Applied Demography & Survey Research
College of Arts and Sciences
University of Delaware

Newark, Delaware

The University of Delaware is committed to assuring equal opportunity to all persons and does not discriminate on the basis of race, color, gender, religion, ancestry, national origin, sexual preference, veteran status, age, or disability in its educational programs, activities, admissions, or employment practices as required by Title IX of the Educational Amendments of 1972, Title VI of the Civil Rights Act of 1964, the Rehabilitation Act of 1973, the Americans with Disabilities Act, other applicable statutes, and University policy. Inquiries concerning these statutes and information regarding campus accessibility and Title VI should be referred to the Affirmative Action Officer, 305 Hullihen Hall, 302/831-2835 (voice), 302/831-4552(TDD).

TABLE OF CONTENTS

Page

LIST OF FIGURES

Executive Summary

Vehicles owned and operated by the State of Delaware are currently monitored using GPS equipment to assist with vehicle location, status, and diagnostics. There are over 2,400 vehicles on the system, each broadcasting locations, speeds, and vehicle status. More than two million measurements are taken per month for roadways across the State, and the program has been in place since 2007. Access to this data is courtesy of the Delaware Office of Budget and Management, who direct the fleet monitoring program. State Fleet vehicles act as speed probes making it possible for the first time to collect speed measurements throughout the State for roads large and small, and produce a highly detailed view of travel speeds and travel time by directional flow, turning movement, time of day, day of week, month and year.

Data processing procedures were developed and refined in this project building on past research. The result is an apparently very extensive and consistent data set that can be used to evaluate the performance of Delaware's road network and serve many applications that include:

- Support of congestion management programs that identify specific areas of concern and focus improvements and capacity preservation efforts.
- Development of state wide routing capabilities that factor in time of day, day of week, season, and holidays.
- A better understanding of the relationship of land use and the performance of the transportation system.
- The examination of the relationship of volume, speed, and capacity for roads throughout the network.
- Evaluation of travel time reliability for corridors in Delaware.
- Measurement and study of delay at intersections.
- Before and after studies, such as the evaluation of the effects of transportation system improvements, introduction of new land uses, and road closure and other changes in road capacity.

These and other applications can be served and the research also made great progress in demonstrating effective way of referencing, managing, and accessing traffic data in general.

INTRODUCTION

The determination of speeds and travel times on Delaware's road network can provide valuable performance measures that can be used for transportation operations and planning. Existing technology in place on Delaware's roads provide good measures of traffic volumes throughout the day in many areas, but here have been considerably less measures of speed. A relatively small number of devices measure traffic speed at fixed points in the roadway network, and in general these are only deployed on major roads and generally additional speed measures are focused on corridor wide studies.

Vehicles owned and operated by the State of Delaware are currently monitored using GPS equipment to assist with vehicle location, status, and diagnostics. There are over 2,400 vehicles on the system, each broadcasting locations, speeds, and vehicle status. Each trip in the data set is defined by the time a driver turns on the ignition to the time the driver turns off the ignition. Measurements are taken every two minutes during each trip. More than two million measurements are taken per month for roadways across the State, and the program has been in place since 2007. Access to this data is courtesy of the Delaware Office of Budget and Management, who direct the fleet monitoring program, and costs for the collection of the data are covered by this program. As discussed here this is a very valuable resource to provide statewide measures.

Availability of very large data sets for traffic data is increasing rapidly. A corresponding familiarity and approach to these large data sets needs to be advanced to fully take advantage of these resources and the investments that made them possible. Challenges arising from the large amount of information and steps to make it usable are discussed.

A statewide travel speed survey was developed in this project and examples of its applications are provided. What is shown is a first look at the data. There will be further analysis of this information that will follow the initial preparation of the data and there are many potential applications to consider.

VEHICLE GPS DATA OVERVIEW

Over 2,400 State of Delaware public vehicles are equipped with GPS devices that broadcast diagnostic information every two minutes while in operation. These broadcasted signals include a time and date stamp, longitude, latitude, instantaneous speed, and the average speed since the last measurement. Each month, millions of measurements can be compiled for tens of thousands of trips. The State of Delaware works with fleet management company, Networkfleet (www.networkfleet.com) for the fleet tracking system. Using

Networkfleet software, historical and real time information can be accessed, maps of the real time location of vehicles can be viewed, and summary reports can be generated.

Figure 1 shows a sample of the main dataset used in this research. Each point on the map represents a single GPS reading from a state fleet vehicle during one month. As part of the processing, these GPS points are associated with portions of the roadway between intersections, and are identified with specific trips of each vehicle. The entire dataset has similar coverage across the State. For many roads, one month's worth of data provides hundreds of measures per month.

Figure 1 One month of GPS measurements near or on Old Baltimore Pike

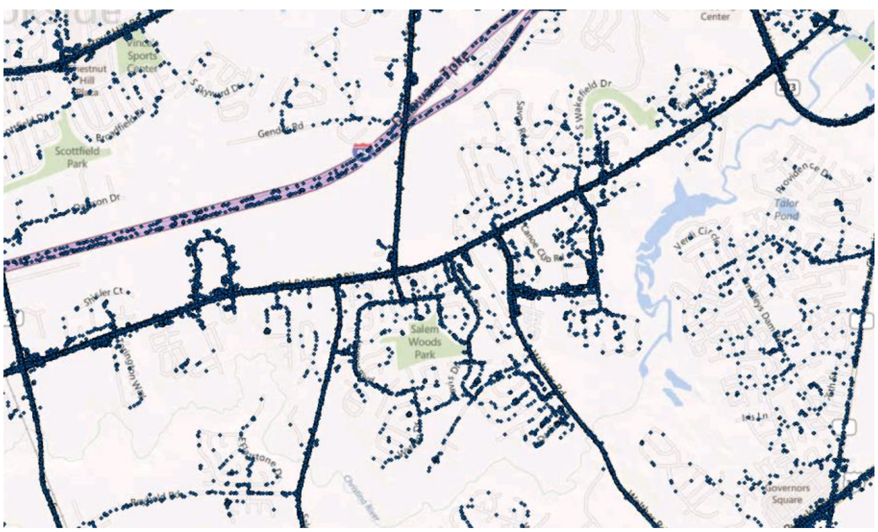
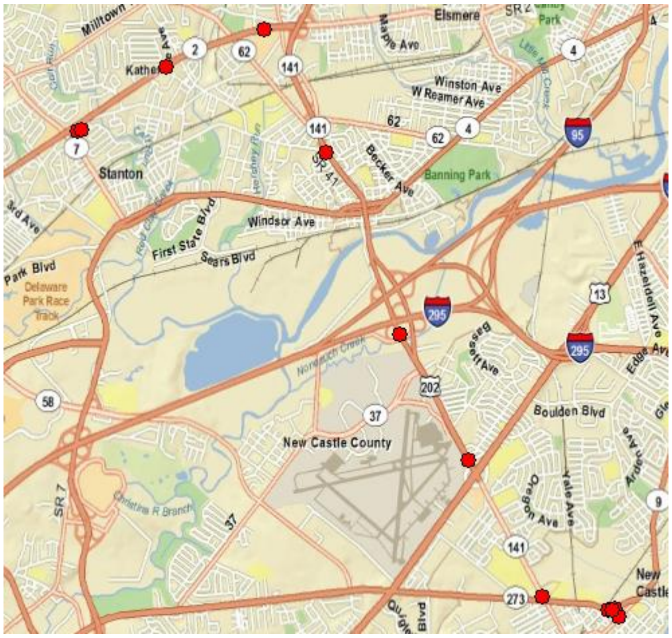


Figure 2, What GPS measurements of a single vehicle trip look like
GPS locations shown as red dots



A sample of what a trip would look like is shown above in Figure 2, where the red dots are the location of actual GPS measurements. With the current GPS equipment installed, Netfleet estimates that 90% of the reported positions are within 8 meters of their actual position. Figure 3 and 4 below provide a view of the coverage for Delaware counties. The data set stretches across all roads large and small and close to 2/3 of all roads have 1000's of measures. Numerous measures are available for even 1 hour of weekday measures as seen in Figure 5 and 6. It is important to note that this paper generally references travel speed. The measure of road link distances are tabulated and measures can be provided in terms of travel time (minutes, seconds).

Figure 3 Number of GPS Measurements (OBSERVS) for each road in New Castle County, Delaware Approximately one year of data, Delaware State Fleet GPS vehicle data 2012-2013, for road portions between intersections.

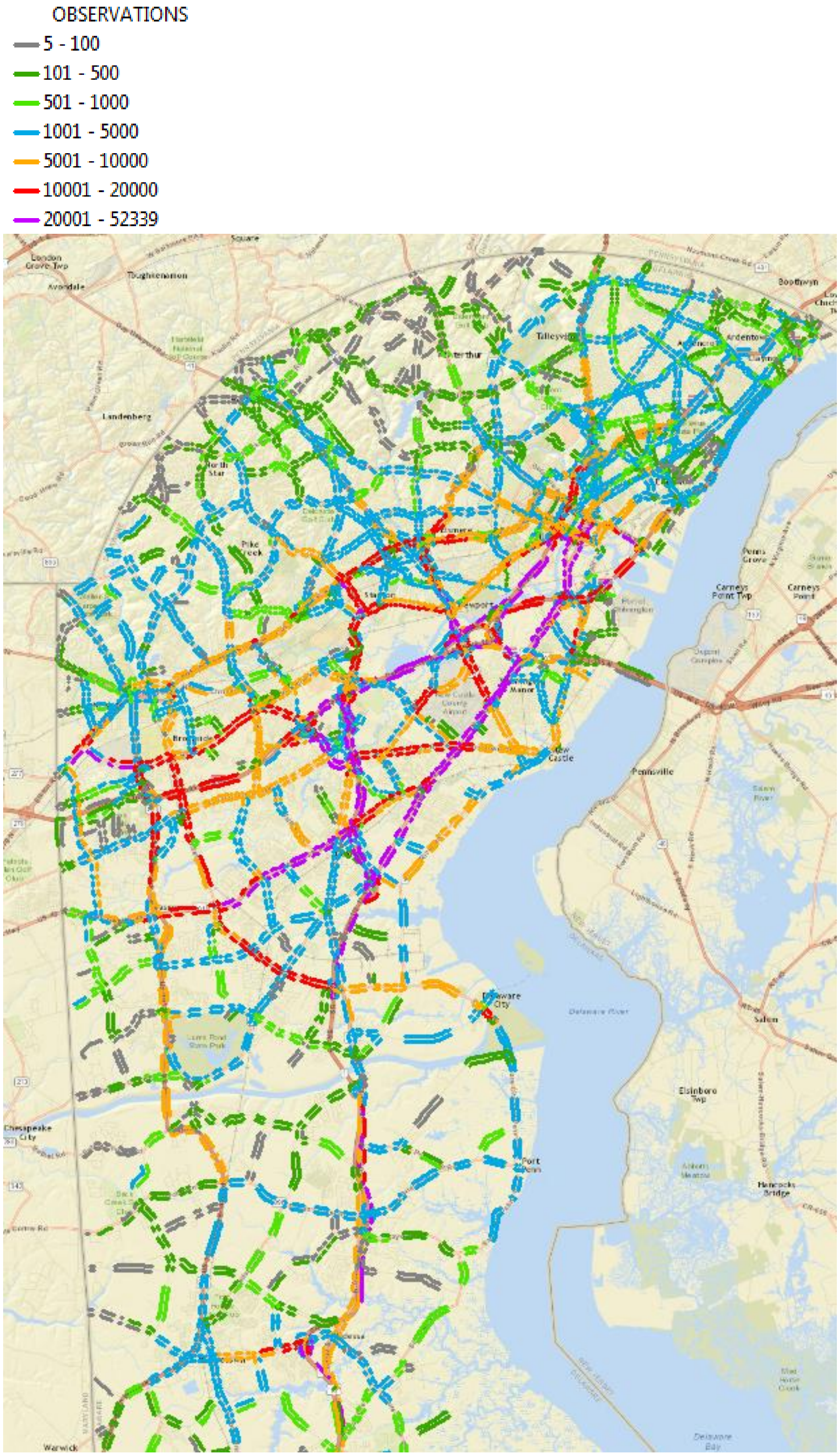


Figure 4 Number of GPS measurements for roads in Kent and Sussex Counties
Approximately one year of data, Delaware State Fleet GPS vehicle data 2012-2013, for road portions between intersections.

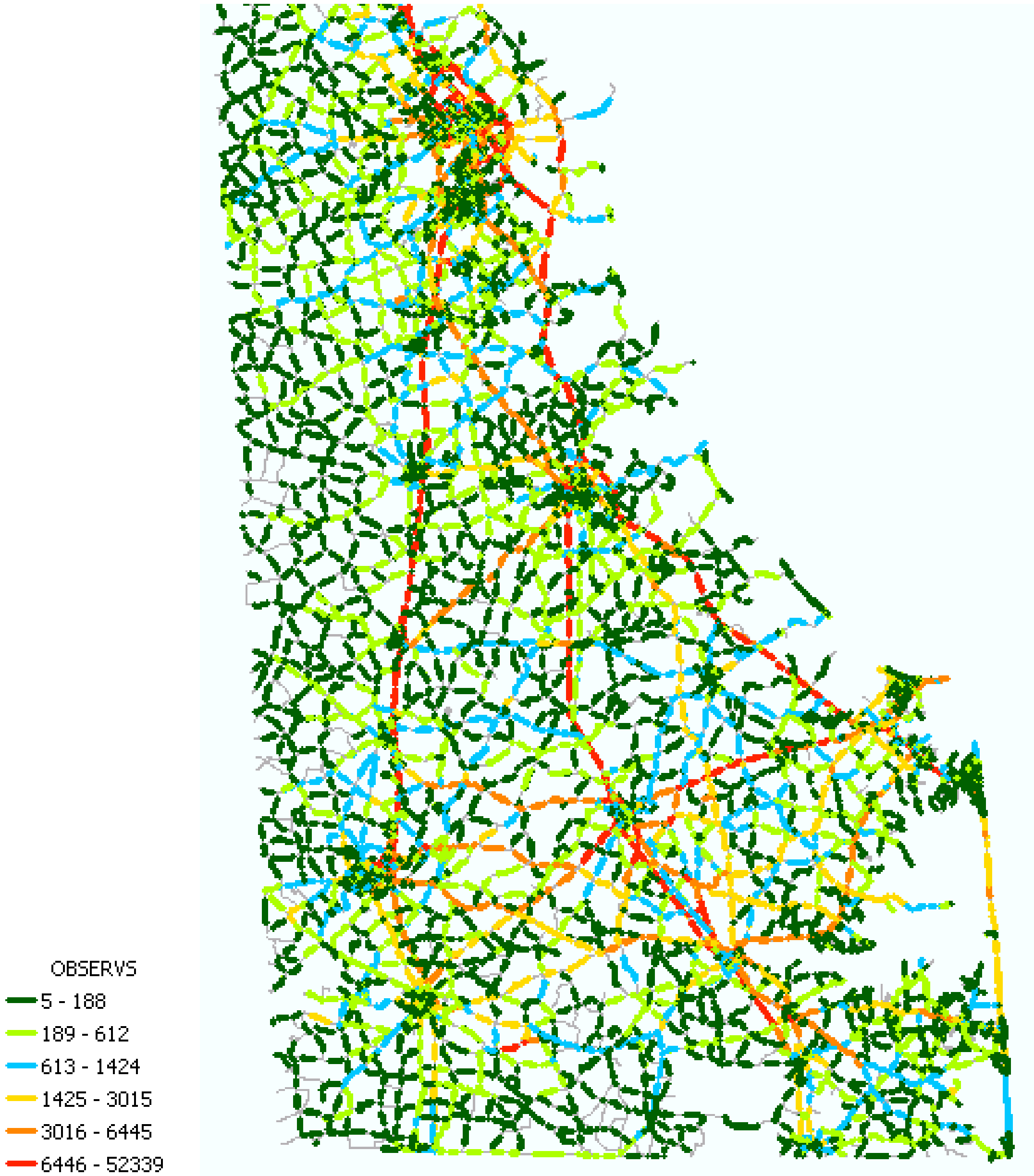


Figure 5 Number of GPS Measurements for each roads in New Castle County for the 8AM Weekday Hour Approximately one year of data, Delaware State Fleet GPS vehicle data 2012-2013, for road portions between intersections.

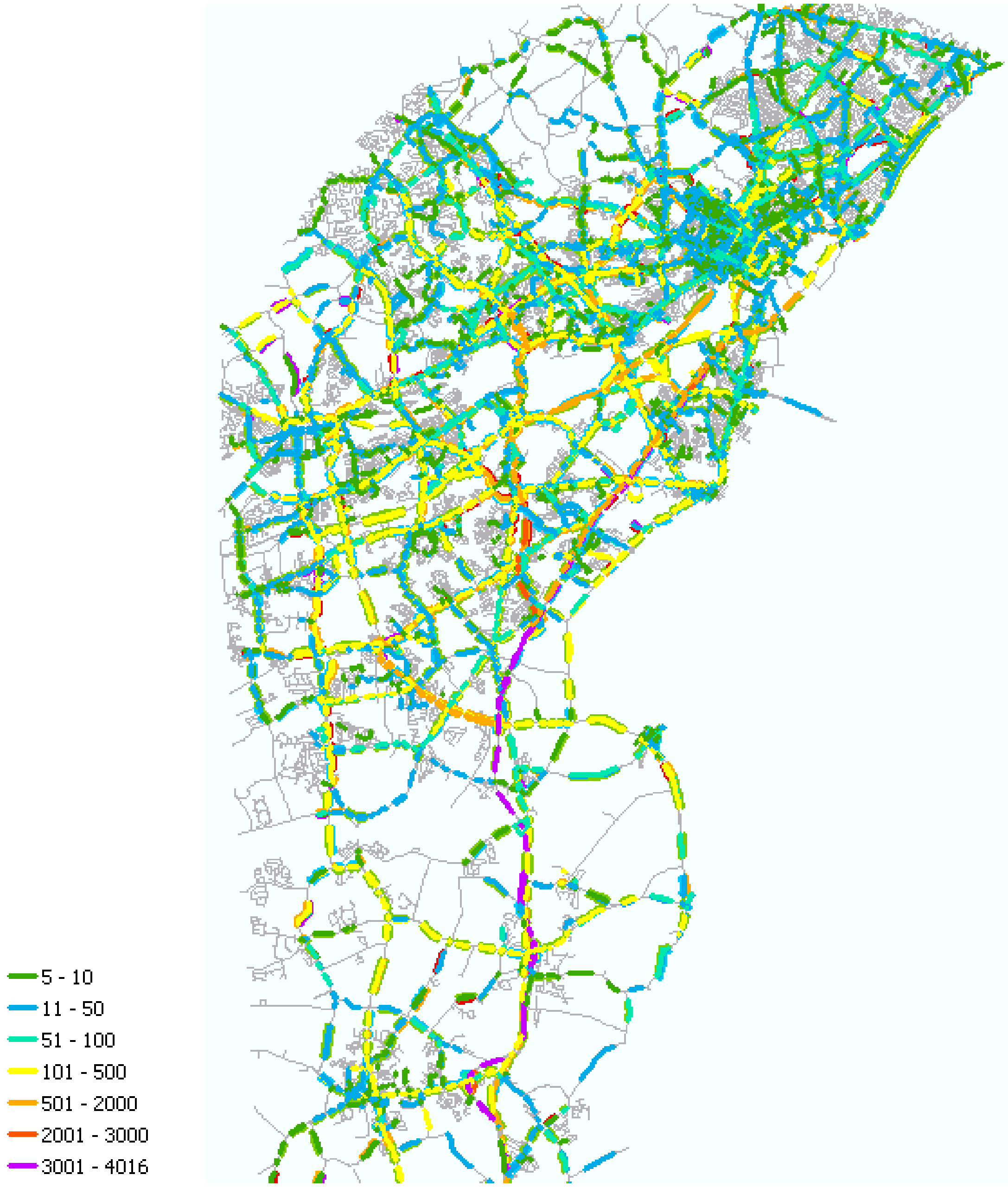
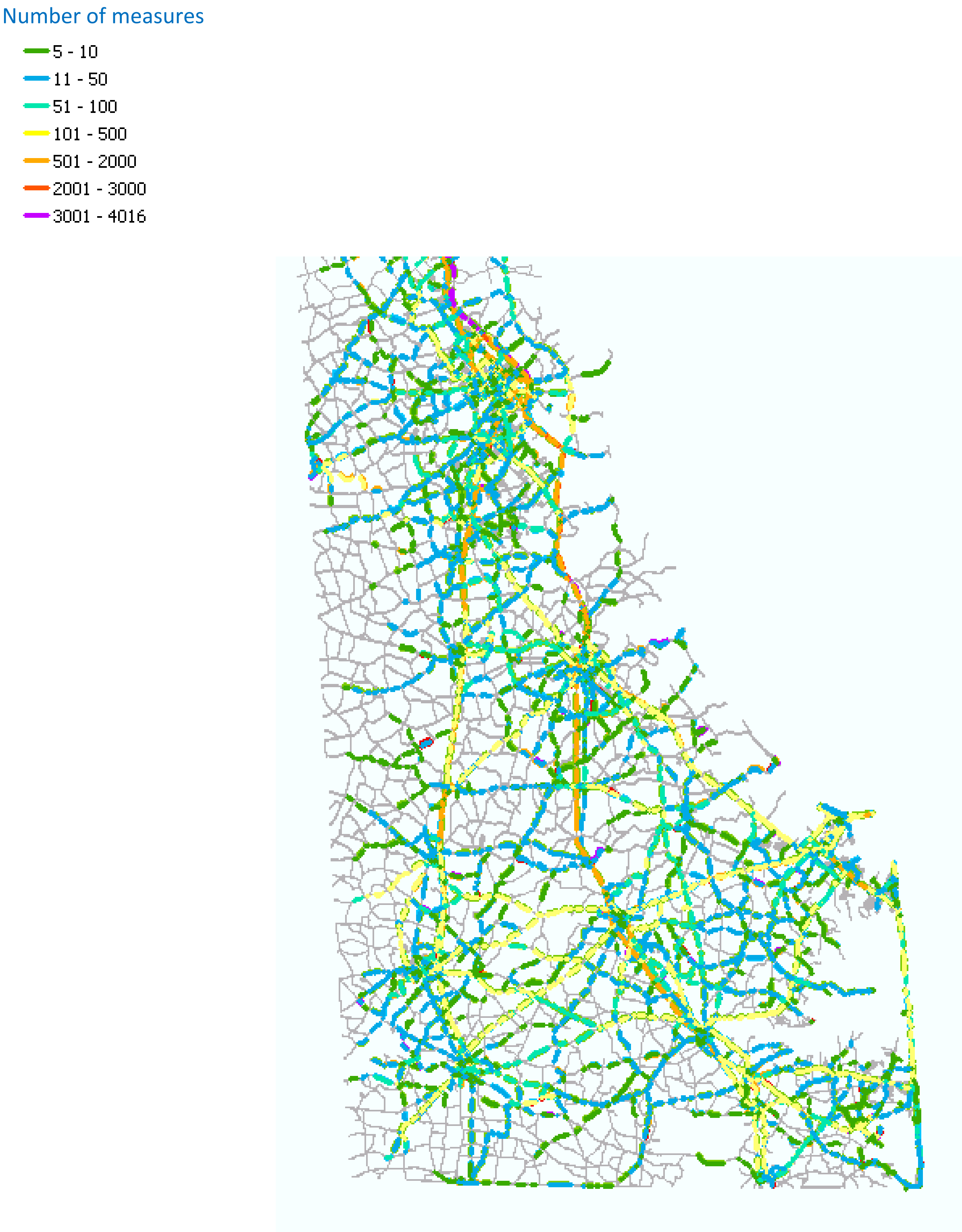


Figure 6 Number of GPS Measurements for roads in Sussex County and Kent County for 8AM Weekday Hour, Delaware State Fleet GPS vehicle data 2012-2013, for road portions between intersections.



The Public Vehicle Sample

There are approximately 2400 State vehicles that are part of the fleet that is tracked. About 40% of the vehicles are 2 or 4 person passenger cars, about 34% are various types of passenger vans, 13% are pickups, and about 10% are SUV's. The fleet is predominantly focused on moving people rather than for maintenance and construction, with only about 8 dump trucks and 7 trucks with trailers. The vehicles studied include para-transit vehicles but not transit buses. Public safety vehicles are not included. This of course is not a random sample of vehicles on the road but the assumption is that vehicles progress with the predominant traffic speed particularly at speeds below the free flow speed which are of most interest from a planning standpoint. With exception of the case where vehicles are hovering around origins or destinations on minor roads or driveways often within land parcels, there does not appear to be a number of trips that involve vehicles significantly deviating from expected traffic flow. While one trip that involved a vehicle that stopped on the side of a major or minor arterial with the ignition on has been seen in a review of thousands of trips, this would seem to be a rare event in the 10's of millions of measurements (highway maintenance vehicles and transit buses are not included in this set). Also no trips were observed in preliminary review as yet that would indicate a door to door type of delivery vehicle. As data from other sources of travel time and speed are available, comparisons can be made which might shed light on any particularly different characteristics of the state fleet data from the general vehicle population, and perhaps adjustments can be made to distributions if appropriate. There are also a number of cleaning steps that could be conducted to identify and remove particularly slow or otherwise uncharacteristic behavior when compared to what is expected from the general vehicle population.

Where average travel speed on roads are known to consistently flow above speed limits, predominantly with Interstate I-95, initial indications are that State vehicle speeds in the data seem to not deviate largely with traffic flow. For the case of Interstate 95 for instance with a posted speed of 55mph, free flow speeds for the state fleet vehicle sample, calculated from a 75 percentile, are between 60 and 65 mph for most portions which is expected to be slightly less than would be calculated from a random sample but at first look it doesn't appear that the difference is large. As another example, for Delaware Route 1 in areas where the posted speed is 65mph, free flow speeds of the state vehicle sample are typically close to 70mph. How the State vehicle GPS data performs in measuring speeds on highways well above speed limits would need to be studied further. Fortunately, the access restricted highways where this would be most an issue are few and existing devices for measuring speed are already in place.

PROCESSING STATE FLEET VEHICLE GPS DATA

The primary goal of this research requires viewing the point measurements, made up primarily of a longitude/latitude coordinate and a time stamp, as trips over defined links covering particular origins and destinations. In order to make this point-to-link relationship, the data had to first be processed. The main processing steps were:

- Capture historical GPS data by querying Networkfleet for Delaware vehicles.
- Process GPS XML response data
- Create GPS point databases and GIS files
- Extract and associate GPS points with particular trips taken through time
- Build a trip and link based version of the GPS data
    - Estimate the path taken between GPS readings
    - Associate particular point measures with a particular road link, direction, and tuning movement
- Where portions of a particular trip include road links with no actual measurement, interpolate speeds estimates and travel times between GPS measures
- Screen the data for errors and anomalies.

The end result of this process maps each observation to a particular road link and direction. The specific links and directions which a particular vehicle travels, constitute its trip.

The processing was done using tools that included SPPS statistical software, Excel, ARCGIS Desktop (ESRI), and java and Python programming. Processing the most likely path between GPS readings and building trip and road link sequences was done using ARC Network Analyst. These initial steps took considerable amounts of computer processing time. Just one month of data takes a powerful University of Delaware computing cluster (MILS) almost 2 days to process. The development of the process, quality control, and error checking of the referencing network involved repeated processing for each of several months.

Imperfect GPS positional accuracy, roughly +/- 8 meters, was not considered to be a substantial source of error. This type of error is most difficult when the measurement is very near the location of an intersection or where a road is represented as a dual lane highway and the inaccuracy produces a proximity to the wrong side of the road. Steps in programs were developed to ensure that each GPS reading was associated with the correct road link and direction by examining subsequent links and nodes mapped during that trip. With errors in positional accuracy and for other reasons it is possible that the optimum path generated by Arc Network Analyst is not the actual path taken by the vehicle between the 2 minute reading. Fortunately the GPS reading from the device also includes an average speed from the last measurement. The average speed from the GPS unit was compared to the average speed that was calculated from the time stamp and the distance calculated from the optimum path algorithm. Where the reported device average speed differed considerably from the calculated speed, the measurement was discarded. Discarding these would remove many of the placement errors that resulted from inaccuracies of location reported by the GPS device.

Time and location were broadcasted from GPS devices every two minutes, and speed fluctuations between those measurements are unknown. Average speeds were estimated by dividing the distance travelled between two subsequent GPS readings by the change in time between those measures (i.e two minutes). That average speed was apportioned to any links that were completely traversed between adjacent GPS readings. For links that were partially traversed between adjacent GPS readings, average speeds were calculated as the sum of any interpolated distances on each link divided by the sum of the corresponding interpolated times on that link. For each link, we expect that the average of interpolated speeds taken over random placement of measurements will approach the average of actual speeds taken over many observations.

A large amount of time was spent reviewing and mapping thousands of paths generated between GPS measurements to identify sources of error and issues with the processing or the data. The first few months of processing involved repeated re-processing of the GPS data until all major network and processing errors were addressed. Errors generally resulted because the GPS reading was associated with the incorrect road, or the routing network used had incorrect impedances or connectivity, or the actual path taken between readings did not correspond to the expected optimum path between readings. Where problems were due to the routing network, the network would be corrected and the processing was repeated. .

The entire process was structured assuming that the data would be reprocessed at a later date. For instance at some point if an error in the road network routing model is found, estimates would need to be developed again for any trips crossing that location. As the data is further examined additional error checking and cleaning steps and adjustments would be expected. For instance, it is possible that particular vehicles for some reason might generally behave unusually, and these could be identified and removed from estimates.

REFERENCING TRAFFIC DATA, THE ADVANTAGES OF LINEAR REFERENCING

Perhaps the most important issue in processing and managing traffic data is a standard means for referencing data to the transportation network. An immense amount of traffic data is being collected by field devices, traffic impact studies, permanent counts, vehicle probes, and other means. Moreover, that information comes in different formats, and addressing different portions and lengths of the roadway, so integrating all of this data requires a consistent manner of referencing.

Typically, data points reference a GIS-based model of the road network that uniquely identifies each road segment or collection of segments using some arbitrary numeric or text identifier. There are a number of road representations typically available in any locale and there are a number of commercial products. Typically each representation includes its own identification scheme, cartography, and the manner in which the road network is broken down into portions of road that data is reported on (segmentation). The use of various identifiers and segmentation among data sets is a well known impediment for consistently locating and relating transportation data. Depending on the application, transportation data can address large corridors or the smallest portions of road, so the segmentation, can cause some of the biggest problems in relating data from various sources.

DelDOT has historically (30 years at least) used a linear referencing system (LRS) which groups continuous road segments into a single route and then assigns each of those segments a beginning and end mile marker and measures in between. Linear referencing provides the ability to reference small and large portions of the roadway.

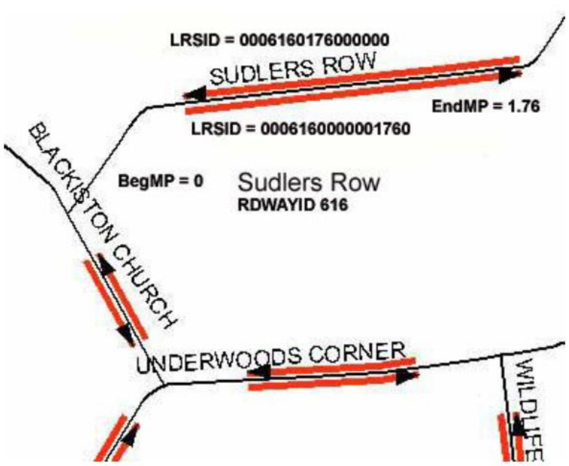
In developing a method of referencing traffic data CADSR desired a method that:

- is related to an established standard
- can reference the smallest portions of road as well as the largest
- is not dependent on a particular cartographic source
- can relate data from various sources and measurement schemes
- can be generalized to relate information about small and large road segments
- can capture the direction of traffic flow. Traffic data for a particular portion of road is directional.
- Provide a fixed identifier for use by those who cannot work with linear referencing systems (route and mile point).

Linear referencing provides support for these features and in previous research CADSR has developed a routing road centerline that employs Delaware's linear referencing system in the identifiers of road segments. A fixed segment identifier is established, called the LRSID, for each road link/segment in the network and it is comprised of the route identifier, the beginning mile point, and the end mile point. It is a text item of length 16 where the first 6 places hold the route identifier, the next 5 characters holds the beginning mile point to 3 decimal places (1000s of a mile), and the next 5 characters hold the end mile point. The direction of the traffic data assigned to the segment is determined by the order of the mile points. Figure 7 on the next page illustrates this.

Sudlers Row is a portion of road that is assigned to Route 616, begins at 0 mile point and ends at 1.76 mile point. In a text format 6 places for the route would be ‘000616’, 5 places for the beginning mile point would be ‘00000’ and for the end mile point ‘01760’. Concatenating these leaves and LRSID = ‘000616-00000-01760’ = ‘0006160000001760’ . The same portion of road referencing data in the other direction would be LRSID = ‘000616-01760-00000’ = ‘0006160176000000’. The result is a segment/link identifier that incorporates the segmentation and direction and can stand alone in databases as a meaningful identifier. (In more sophisticated databases management systems the identifier could be comprised of the three fields, route, beginning mile point, and end mile point, and joining and relates to other tables could be done on multiple fields. In general many users do not have linear referencing system tools or the ability to join on multiple fields). A focus was on being able to provide data and locations to accommodate basic GIS capabilities and analysis capabilities. Any representation of a road centerline file that includes the Delaware linear referencing system can use standalone data tables of traffic data and can reference the data without concern of different cartography or segmentation. All of the locational information about the road segment is in the identifier.

Figure 7, Example of the identification scheme used to reference GPS speed measurements



In this way CADSR created a standard referencing scheme based on the Delaware LRS which identifies each link, mile point information, and direction of traffic flow. The red arrows shown in the figure 4 are a cartographic feature that can provide a means to visualize the segment and direction.

Identification is further augmented to be able to reference turning movements. A typical trip in the vehicle GPS data set would not only proceed in a straight thru manner, with no left or right turns. At an intersection some vehicles will go straight thru, some will turn left, some will go right, and some will make a "U" turn. Of course each turning movement will usually always involve a different time (speed) to make the turn and this is very clearly shown in the processed data sets. Often at a signalized intersection, right turns will take less time due to yields and right turn on red, and left turns will take more time. The vehicle GPS data set yields different times/speeds to cross the segment and tables include an "S","R","L", or "U" at the end of the LRSID, e.g.

LRSID='0006160176000000R' (In this example the east bound right turn from Sudlers Row)

Just as there are different speeds and travel times for making that turning movement there are also different traffic volumes. Traffic volumes are referenced to the transportation network in the same way and turning movement volumes can then be easily related to speed measurements for that turning movement.

ADDITIONAL NOTES

Aggregations

Over 100 million measurements each year available for very detailed road segments throughout the State creates a large data set that requires aggregations to examine conditions with respect to various factors of interest. These factors include:

- Time of day intervals, i.e. 15 minute, 30 minute, 60 minute intervals
- Periods of the day, AM Peak, Midday, PM Peak, Evening etc.
- Day of week
- Season, i.e Summer or Non-Summer
- Holiday, non-Holiday
- Year

Such aggregations are necessary when looking at any large traffic data set whether it is for volume, speed, or capacity, in order to distill the large data into meaningful portions to support planning applications. It is also practically necessary to address limitations of computer speed. It is possible to create the aggregations on the fly using SQL statements in a robust database system such as Oracle and this has been demonstrated by the author with internet based mapping systems, but as the amount of measurements gets in the several 10's of millions this may take large amounts of time. Therefore to examine the data more easily as well as provide it to planners in useable sizes, aggregations are created such as, "Average speed by road link for weekdays (Monday thru Friday) by 30 minute intervals (epochs) through the day, for non-Summer months". These aggregations are much more manageable and useful.

Of particular interest are time intervals during the day. While the State vehicle data is voluminous, when one is trying to examine data at more detailed time intervals during the day there are less and less observations. Fifty or more observations are desired for an estimate of a particular section of road for a particular day and time of day. 5 minute intervals throughout the day for weekdays would only produce 50 or more observations for only a few of the most traveled roads with this data set. However if data is aggregated into 30 minute intervals through the day, many more observations are available for each time interval for many more roads. Looking at 60 minute time of day interval aggregations provides 100's to 1000's of observations for major roads across the state and many minor roads. Where multiple years of data are used, detail and reliability of estimates is increased. The question arises "How much data is needed and how detailed a time interval over what number of months or years is needed to serve the application?" At the planning scale and to derive the average functioning of the statewide network, 30 minute and 60 minute intervals seem to provide a very good statewide picture.

Aggregations by road segment

In a previous section the method of referencing the travel time and speed data using the Delaware linear referencing system was described. The benefits include having the data referenced in line with a State standard, the ability to reference any portion of road (dynamic segmentation), independence from a particular cartographic representation, and a very straightforward way of referencing directional data. Another very important feature of this method is the ability to aggregate data to more general representations. Data can be aggregated by time of day, day of week, month, and other factors as discussed above but it can also be aggregated into measures for larger portions of roadways. Most available measures of traffic data relate to much larger, more generalized portions of roadway than what the State vehicle GPS data provides. The road network used to process the vehicle GPS data is very detailed, with link/segment breaks at every intersection for all roads large and small. This provides very useful detail, but often, particularly when relating measures to other traffic volume or speed data, there is a need to generalize or address larger portions of the roadway. A simple example can be provided.

Figure 8 below shows a particular portion of road, Thompson Station Road, that is broken down in the routing network into two segments between Hopkins and Smith Mill Road. In the figure, the LRS identifiers are shown. Recall that the identifier is constructed by concatenating the route id with the beginning and end mile points with the mile points in the order that would indicate direction of flow.

The top segment going north, 0000820104001280	Rt 820 with begmp = 1.04 and endmp = 1.28. len = .24
The bottom segment going north,0000820044001040	Rt 820 with begmp = 0.44 and endmp = 1.04, len = 0.6
The top segment going south, 0000820128001040	Rt 820 with begmp = 1.28 and endmp = 1.04, len = .24
The bottom segment going south,0000820104000440	Rt 820 with begmp = 1.04 and endmp = 0.44, len = 0.6

From this it is seen that Thompson Station Road is measured from South to North, and it ranges from 0.44 miles to 1.28 miles on Route 820. Suppose though that it was desirable to have a measure that combined these segments for one measure between Smith Mill Road and Hopkins Road as seen in Figure 9. The identification method is consistent where

South to North Road Section 0000820044001280 = 0000820104001280 + 0000820044001040

Where for instance it is found that the average speed for 0000820104001280 is 30 mph and the average speed for 0000820044001040 is 40 mph. The mile points could be used to weight the measurement to get a speed estimate for the one segment so that,

Speed of 0000820044001280 = (length 0000820104001280 * avg speed 0000820104001280

\+ length 0000820044001040 * avg speed 0000820044001040)

/ length of 0000820044001280

= (0.24 * 40 + 0.6 * 30) / .84 = 32.8 mph

This is a simple example using a portion of road broken into two segments that is aggregated into one, but in other cases the aggregations can involve many segments aggregated to one larger segment. The use of the route and mile points in identifiers and tables facilitates the automation of aggregations and a consistent and meaningful identification that can handle various resolutions.

Figure 8 Portion of Thompson Station Road represented by two segments

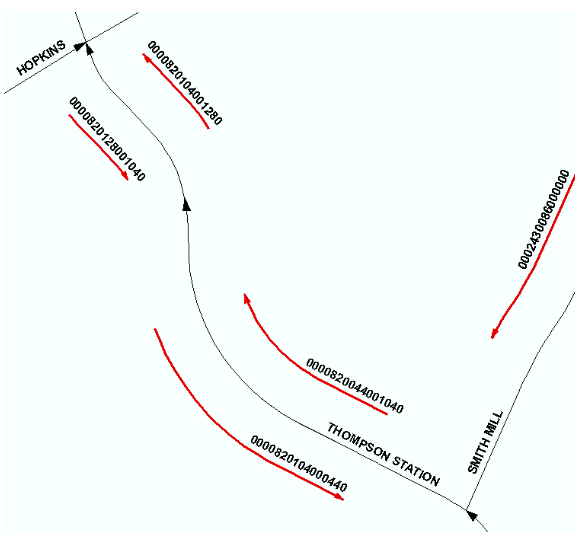
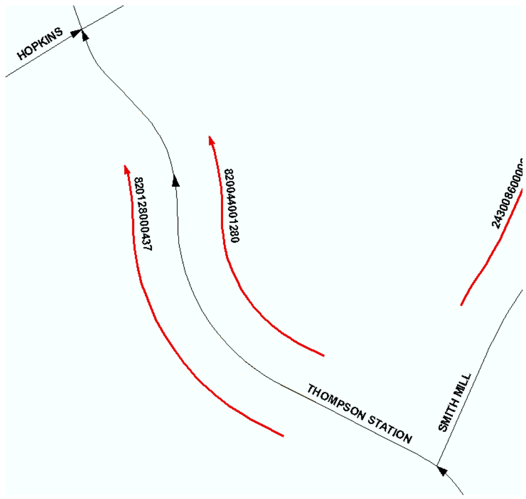


Figure 9 Portion of Thompson Station Road represented by one segment



In some cases of course the segments to be aggregated do not belong to the same linear referencing route. In this case look up tables are required to assign a route identifier to a series of road segments to be aggregated.

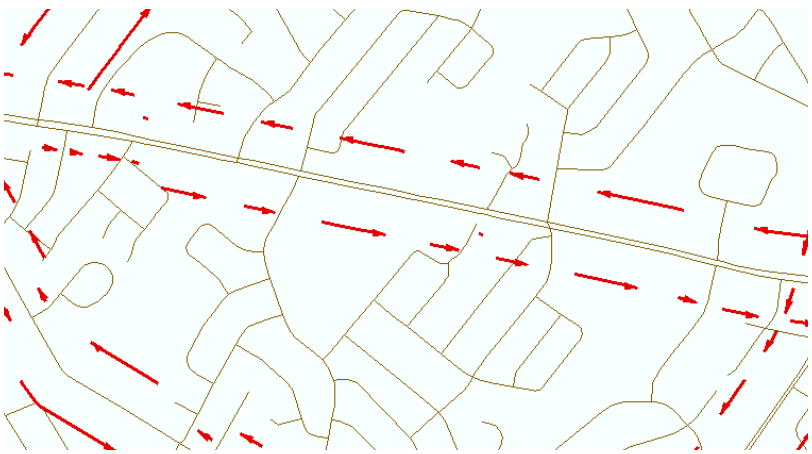
Suitability of the vehicle GPS data set as compared with other sources of speed data

Other sources of speed and travel time data are sometimes available and it is of interest to note the features of these sources and how they might compare to the State vehicle GPS data. Sources of speed or travel time data usually provide average estimates by time intervals throughout the day, for a set of road portions. Where the source for the data is sensors located in or by the road, the number of sensors, the placement and spacing, and the sample size determine the usefulness and statistical significance provided for a particular application. Such sensors often can log numerous measurements, sometimes all vehicles that pass the sensor, in aggregate time intervals as small as 5 minutes throughout the day. This would be the type of data available from permanent sensors or Blue Tooth. Another recent source of average speed data that is very promising is the National Performance Management Research Data Set (NPMRDS) that covers roads that are part of the National Highway Monitoring System. The NPMRDS produces numerous measurements for most 5 minute intervals throughout the day and is derived from commercial entities who provide summations of GPS readings as available from cell phones. Data available from permanent or temporary sensors and the NPMRDS are typically available for longer stretches of road and are especially good for the evaluation of travel time between ends of a specific corridor.

The State vehicle GPS data processed in this project differs from other emerging sources in a few ways. First, measurements are available for all segments between intersections of all roads traveled, large and small, enabling a more detailed look at flows within corridors and a picture of the entire statewide network. With measurements for smaller road portions, bottlenecks and specific candidates for improvement within a corridor may be more easily identified. The State vehicle data provides measurement coverage for minor as well as major roads throughout Delaware. For compilation of speed measures for the roadway between two sensors, the State vehicle GPS data cannot come close to the huge number of measurements that can be available with fixed sensors at defined intervals on included roads, or for identifying a particular condition at a particular time on a particular day. But for major roads there can still be thousands of measurements over a year for a particular 1 hour time interval in the vehicle GPS data, and at this time it can provide much greater coverage of the road network as a whole, particularly when several months or more of data are compiled. Fixed sensors could provide detailed information for a particular incident on a particular corridor at a particular time on a particular day. The State Fleet GPS data numerous measurements for many roads which taken in aggregate can describe the expected normal functioning throughout the day of the road network as a whole.

Figure 10, Example comparison of road segment length of State vehicle GPS versus NPMRDS data. Each red arrow represents a tabulated measurement.

State Vehicle GPS Data



NPMRDS Data

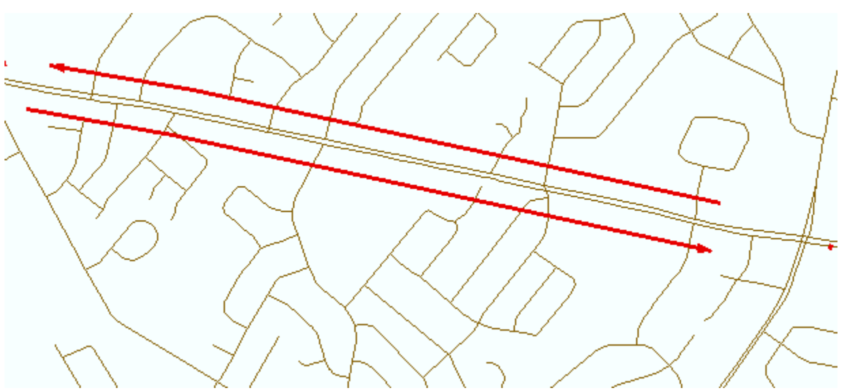
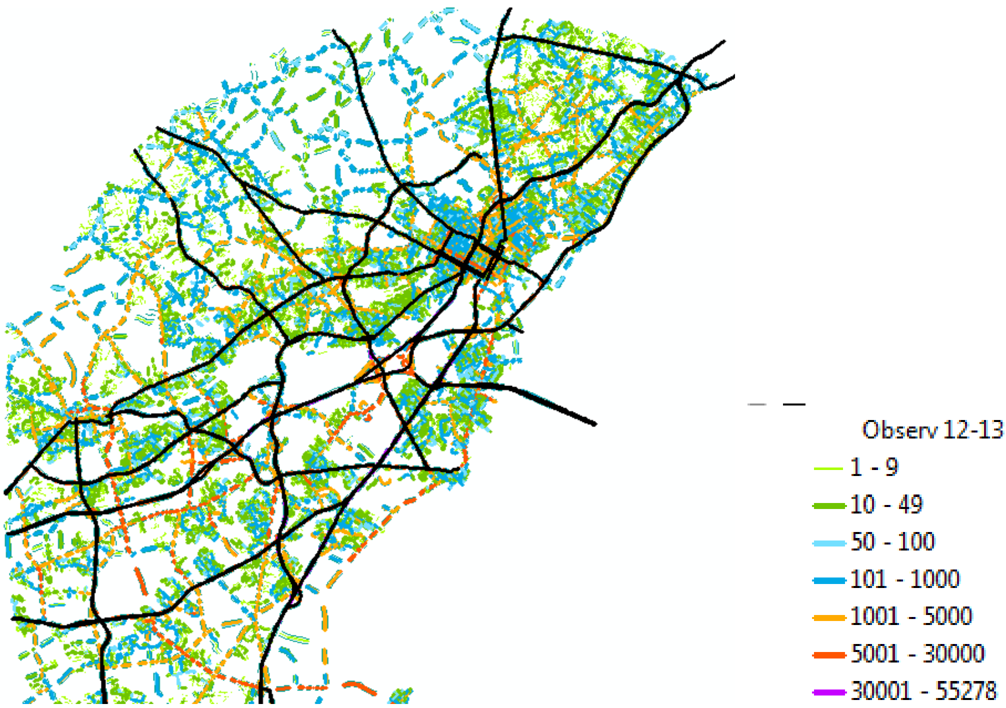
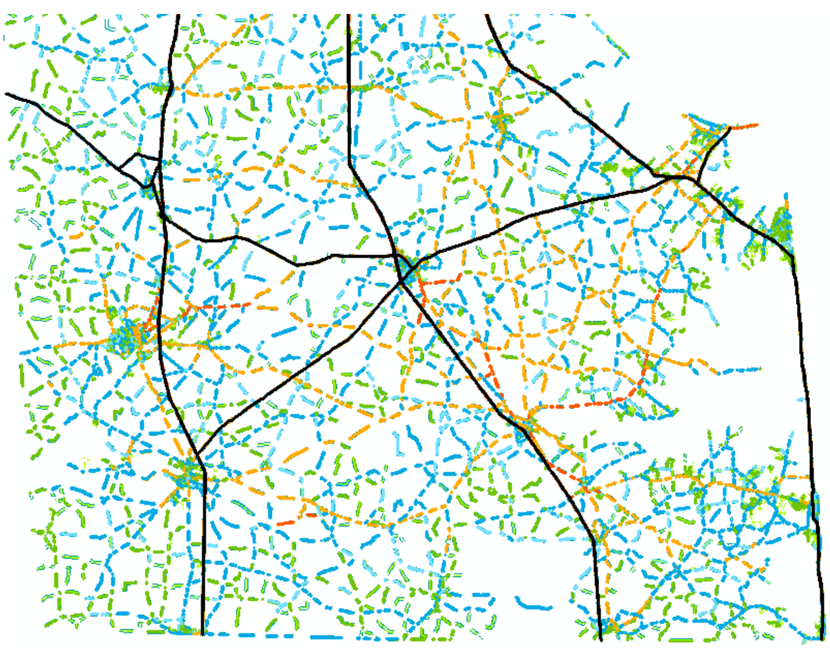


Figure 11, Comparison of Coverage, State Vehicle GPS Weekday Year 2012 vs NPMRDS, (NPMRDS in black)

NORTHERN NEW CASTLE COUNTY



SUSSEX COUNTY



Another very different and significant feature of the State vehicle GPS data is that it captures traffic speeds relative to turning movement. The amount of time it would take to traverse an intersection certainly depends on whether a vehicle is going straight thru or taking a left, right, or U turn. Because the speeds are derived from millions of actual trips throughout the network taken at totally random points in the network and because speeds for turning movements are tabulated separately, it enables the examination of performance relative to turning movement. The State vehicle GPS data is then able to examine performance at intersections. Trip based data needed to generate estimates for turning movements is typically not available often due to privacy issues. Where measurements are taken at stationary points there is typically no different measures in regards to turning movement[*].

A question arises as to how many measurements are enough. Certainly 1000s of measures throughout the day for a road at a particular time interval provides more data, but does it provide more useful information than a 100 measures? Does the application need 5 minute interval data on a specific day or would 30 minute or hour interval data averaged of numerous days suffice? The ability to compile immense data holdings does not necessarily mean it's a good idea. Practically, at what point does the significant amount of processing and management necessary for more and more data provide sufficient returns? These are questions that need additional consideration.

The vehicle GPS data and other available and emerging sources of travel time and speed data such as the NPMRDS or Bluetooth each have great value depending on the application. Speeds and distributions from one source can be compared and joined with other sources to derive a better understanding of the real performance picture. The biggest challenge (outside of getting access to the data) is seen as the ability to process and make use of and relate these various sources with each other, and with other types of traffic data.

[*] This begs the question then when looking at other sources of speed, like the NPMDS where an average speed is reported for a segment approaching an intersection, as to what the average speed measurement references. Where the estimate does not reference a turning movement, then it would necessarily seem to be an aggregation of speeds for lefts, rights, and thru traffic. Another question that arises with speeds determined from fix sensors is how delay at intersections is factored in and how placement affects the measure. If a sensor is placed at an intersection, is the time stamp set when the vehicle reaches the intersection? When it goes thru the intersection? With the vehicle GPS data, delay at intersections is incorporated in the measure through the overlay of thousands of trips capturing the flow in regard to turning movement.

Comparison with Delaware Blue Tooth Data

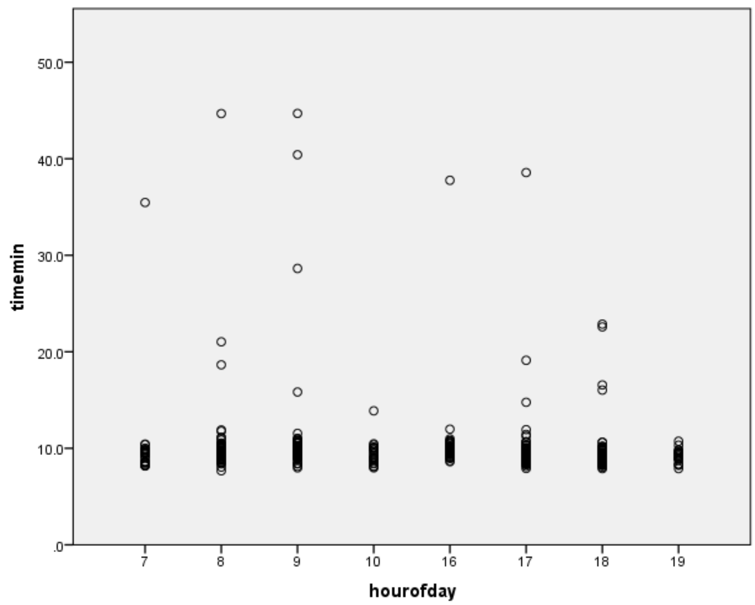
Over the next years DelDOT will be involved in the implementation of Blue Tooth technology to determine speeds and provide real time condition on major roadways. Included in this study was a brief comparison of State Fleet GPS data with data available from Blue Tooth sensors . The Bluetooth data is a small sample (portions of September and October 2013) and is preliminary but the following discussion provides a view of how the raw Blue Tooth data can be compared with vehicle GPS data. The State Fleet comparison data included about 9 months of weekday, non-summer months in the year 2012.

Bluetooth sensors are placed at intersections that can detect vehicles as they come within proximity. Figure 12 below shows the location of Blue Tooth Sensors in the sample. The basic idea is to match the Blue Tooth Identifier from devices in vehicles at one location and match that device identifier with detection at a Blue Tooth sensor at another location. Based on the time stamps of detection the travel time or speed of a vehicle traveling between Blue Tooth sensors can be estimated.

Figure 12 Blue Tooth Sensor Locations in Sussex County Delaware

The blue tooth data was processed for these 5 blue tooth stations and travel time was calculated for all station to station pairs. Travel time for a particular vehicle is calculated as the difference between the time of the last detection from one station and the last detection time from the previous station. In this sample data for these 5 stations, a little over 2500 pairs were derived. Travel time for theses origins and destinations was aggregated by station to station pairs, and hour of the day, producing a median travel time. The median is used rather than the mean since there is some noise and random error in the blue tooth data. Figure 13 below showing a scatter plot for a particular station of travel times (minutes) by hour of the day illustrates this. An approach using data within certain percentiles can also be employed to screen outliers and noise in the data.

Figure 13, Blue Tooth Travel Times by Hour of the Day
Station 3 to Station 4 southbound



The summary table on the next pages show comparisons of travel time estimates between stations from the Blue Tooth and State Fleet data. Most of the larger differences are where Blue Tooth sample size is small. Where the Blue Tooth sample is greater than 5 measures, the difference on average is 7.5%. For Blue Tooth samples greater than 10, the percentage difference on average is 6%. For Blue Tooth samples greater than 20, the percentage difference on average is 3%, and with this data set is less than 3% once Blue Tooth samples are greater than 30. A 3% percentage difference for his data is on average only around 20 seconds.

Figure 14, Comparison of Blue Tooth and State Fleet GPS Data (travel time in minutes)

(For Blue Tooth data #Observ is the number of origin and destination travel time pairs available to make the estimate. For Fleet GPS data #Observ is the average number of link speed measures used to develop the estimate.)

Orig and Destination Identifier	Hour of the Day	MPH Bluetooth	Travel Times Bluetooth	Travel times State Fleet	Percent Difference	Bluetooth #Observ	Fleet #Observ
S1N-S3N	7	62.7	15.6	16.0	2.6	17	245
S1N-S3N	8	64.1	15.3	16.1	5.3	33	262
S1N-S3N	9	59.7	16.4	16.6	1.4	29	289
S1N-S3N	16	61.0	16.0	17.2	6.9	31	597
S1N-S3N	17	62.6	15.6	16.8	7.1	41	330
S1N-S3N	18	63.9	15.3	16.7	8.5	27	161
S1N-S3N	19	59.9	16.3	17.1	4.7	7	95
S1N-S4N	7	59.6	6.7	6.8	1.0	29	296
S1N-S4N	8	58.1	6.9	6.9	.7	49	296
S1N-S4N	9	57.2	7.0	7.1	1.3	58	264
S1N-S4N	16	58.2	6.9	7.3	5.6	50	587
S1N-S4N	17	57.0	7.0	7.1	.9	93	317
S1N-S4N	18	58.9	6.8	7.1	4.0	70	168
S1N-S4N	19	57.7	7.0	7.4	6.1	17	107
S1S-S2S	7	26.3	16.6	11.6	42.7	6	144
S1S-S2S	8	32.9	13.2	12.3	7.2	20	378
S1S-S2S	9	25.5	17.0	12.6	35.2	11	422
S1S-S2S	16	32.6	13.3	14.0	4.9	8	145
S1S-S2S	17	30.1	14.4	13.6	5.9	10	120
S1S-S2S	18	34.5	12.6	12.8	1.8	17	197
S1S-S2S	19	31.8	13.6	13.1	4.0	6	145
S1S-S5S	16	42.0	25.1	26.5	5.2	1	114
S1S-S5S	17	45.0	23.5	26.0	9.7	1	91
S1S-S5S	18	43.3	24.4	24.2	.8	3	149
S1S-S5S	19	43.2	24.5	24.6	.5	1	21
S1S-S6S	17	43.7	32.5	35.3	8.0	1	61
S2N-S1N	7	34.4	12.6	11.2	12.5	5	84
S2N-S1N	8	29.7	14.8	12.2	21.4	6	100
S2N-S1N	9	17.5	24.7	13.1	88.7	11	265
S2N-S1N	16	29.9	14.7	14.2	3.2	2	425
S2N-S1N	17	30.1	14.4	12.9	11.4	16	210
S2N-S1N	18	30.2	14.4	12.3	16.7	9	108
S2N-S1N	19	32.3	13.4	13.7	2.3	3	89
S2N-S3N	7	52.3	26.9	27.0	.2	3	155
S2N-S3N	8	49.0	28.8	28.1	2.5	5	215
S2N-S3N	9	43.0	32.8	29.6	10.9	5	277
S2N-S3N	16	48.8	28.9	31.2	7.5	1	503

Figure 14 Comparison of Blue Tooth and State Fleet GPS Data (continued)

Orig and Destination Identifier	Hour of the Day	MPH Bluetooth	Travel Times Bluetooth	Travel times State Fleet	Percent Difference	Bluetooth #Observ	Fleet #Observ
S2N-S3N	17	48.8	28.9	29.5	2.1	5	264
S2N-S3N	18	49.5	28.5	28.8	1.1	5	132
S2N-S4N	7	44.1	18.9	17.9	5.8	3	148
S2N-S4N	8	34.2	24.4	18.9	29.1	9	160
S2N-S4N	9	33.4	25.8	20.0	28.8	8	266
S2N-S4N	16	37.6	22.4	21.3	5.2	2	477
S2N-S4N	17	40.9	20.4	19.8	2.9	9	244
S2N-S4N	18	42.1	19.8	19.2	3.1	5	127
S2N-S4N	19	45.0	18.5	20.8	11.0	1	96
S2S-S5S	16	53.0	11.8	12.6	6.4	6	58
S2S-S5S	17	53.2	11.7	12.5	6.1	10	38
S2S-S5S	18	58.7	10.6	11.4	6.6	8	62
S2S-S5S	19	58.2	10.7	11.5	6.7	5	41
S2S-S6S	16	44.2	22.3	22.7	1.7	3	42
S2S-S6S	17	49.7	19.9	21.8	8.9	5	22
S2S-S6S	18	50.0	19.8	19.6	.8	6	30
S2S-S6S	19	49.0	20.1	20.0	.7	1	26
S3S-S1S	7	58.5	16.7	16.6	.6	19	48
S3S-S1S	8	56.5	17.3	16.9	2.5	49	223
S3S-S1S	9	56.9	17.2	17.1	.5	40	742
S3S-S1S	16	57.6	17.0	17.1	.8	33	345
S3S-S1S	17	59.0	16.6	16.7	.8	48	369
S3S-S1S	18	60.0	16.3	16.6	1.9	62	505
S3S-S1S	19	62.4	15.7	16.6	5.6	14	263
S3S-S2S	7	37.6	45.7	27.8	64.4	2	106
S3S-S2S	8	47.9	29.4	28.7	2.4	11	314
S3S-S2S	9	48.0	29.4	29.2	.6	5	543
S3S-S2S	16	42.8	33.0	30.6	7.7	6	219
S3S-S2S	17	46.0	30.7	29.8	2.9	7	214
S3S-S2S	18	46.1	30.6	28.9	5.7	18	315
S3S-S4S	7	60.5	9.5	9.0	5.7	25	56
S3S-S4S	8	59.8	9.6	9.3	3.6	60	271
S3S-S4S	9	60.0	9.6	9.4	2.1	52	879
S3S-S4S	16	58.7	9.8	9.3	5.6	39	400
S3S-S4S	17	62.5	9.2	9.1	1.2	70	443
S3S-S4S	18	62.7	9.2	9.0	2.0	86	595
S3S-S4S	19	62.6	9.2	9.0	2.2	20	284
S3S-S5S	17	52.2	39.0	42.3	7.9	1	171
S3S-S5S	18	49.7	41.0	40.3	1.7	3	2523
S3S-S6S	17	50.0	48.0	42.3	13.4	1	171
S4N-S3N	7	65.7	8.8	9.2	4.7	18	183
S4N-S3N	8	66.4	8.7	9.2	5.8	55	445
S4N-S3N	9	62.1	9.3	9.5	2.5	46	322
S4N-S3N	16	62.0	9.3	9.9	6.2	43	611
S4N-S3N	17	65.7	8.8	9.7	9.6	64	347

Figure 14 Comparison of Blue Tooth and State Fleet GPS Data (continued)

Orig and Destination Identifier	Hour of the Day	MPH Bluetooth	Travel Times Bluetooth	Travel times State Fleet	Percent Difference	Bluetooth #Observ	Fleet #Observ
S4N-S3N	18	65.4	8.8	9.5	7.4	47	153
S4N-S3N	19	63.9	9.0	9.7	7.1	14	80
S4S-S1S	7	58.7	6.9	7.2	4.9	35	39
S4S-S1S	8	55.6	7.2	7.1	1.8	102	171
S4S-S1S	9	56.2	7.2	7.2	.7	81	602
S4S-S1S	16	54.5	7.4	7.3	.9	64	277
S4S-S1S	17	56.6	7.1	7.2	1.4	82	287
S4S-S1S	18	57.9	6.9	7.2	3.7	99	416
S4S-S1S	19	56.1	7.2	7.2	.5	32	239
S4S-S2S	7	24.4	34.2	18.8	81.9	1	116
S4S-S2S	8	42.0	19.9	19.4	2.4	16	324
S4S-S2S	9	39.0	21.4	19.9	7.5	7	469
S4S-S2S	16	39.7	21.1	21.3	1.0	6	179
S4S-S2S	17	39.3	21.3	20.7	2.7	13	164
S4S-S2S	18	43.1	19.4	19.9	2.6	17	254
S4S-S2S	19	38.8	21.5	20.3	6.0	4	169
S4S-S5S	16	45.7	31.9	33.9	5.9	1	143
S4S-S5S	17	48.9	29.9	33.2	10.1	1	126
S4S-S5S	18	46.0	31.7	31.3	1.2	4	196
S4S-S6S	17	46.9	38.9	42.5	8.6	1	88
S5N-S1N	16	37.7	28.0	26.9	4.2	1	300
S5N-S1N	17	25.3	41.9	25.5	64.2	3	140
S5N-S1N	18	37.2	29.0	24.4	18.8	4	75
S5N-S2N	16	55.9	11.4	12.7	10.0	7	158
S5N-S2N	17	59.9	10.7	12.7	16.0	11	62
S5N-S2N	18	60.9	10.5	12.1	13.4	8	3
S5N-S2N	19	58.5	10.9	11.9	8.2	2	8
S5N-S3N	16	45.1	45.1	44.1	2.3	1	383
S5N-S3N	18	53.4	38.1	41.0	7.0	1	99
S5N-S4N	16	41.4	35.3	34.2	3.1	1	34
S5N-S4N	18	40.3	38.0	31.5	20.6	2	91
S5S-S6S	16	39.0	9.3	10.1	8.1	9	32
S5S-S6S	17	34.6	10.5	9.3	12.9	18	14
S5S-S6S	18	39.7	9.1	8.3	9.8	13	10
S5S-S6S	19	41.5	8.7	8.5	2.5	1	17
S6N-S1N	17	37.5	38.3	34.9	9.8	2	94
S6N-S2N	16	34.6	28.9	22.5	28.5	1	92
S6N-S2N	17	49.1	20.4	22.1	7.8	6	39
S6N-S2N	18	53.5	18.7	20.2	7.4	1	27
S6N-S5N	16	26.7	14.6	9.8	48.5	2	45
S6N-S5N	17	38.6	9.4	9.4	.2	14	22
S6N-S5N	18	41.7	8.7	8.1	7.6	2	17
S6N-S5N	19	35.2	11.2	8.9	26.2	2	9

EXAMPLE RESULTS AND APPLICATIONS

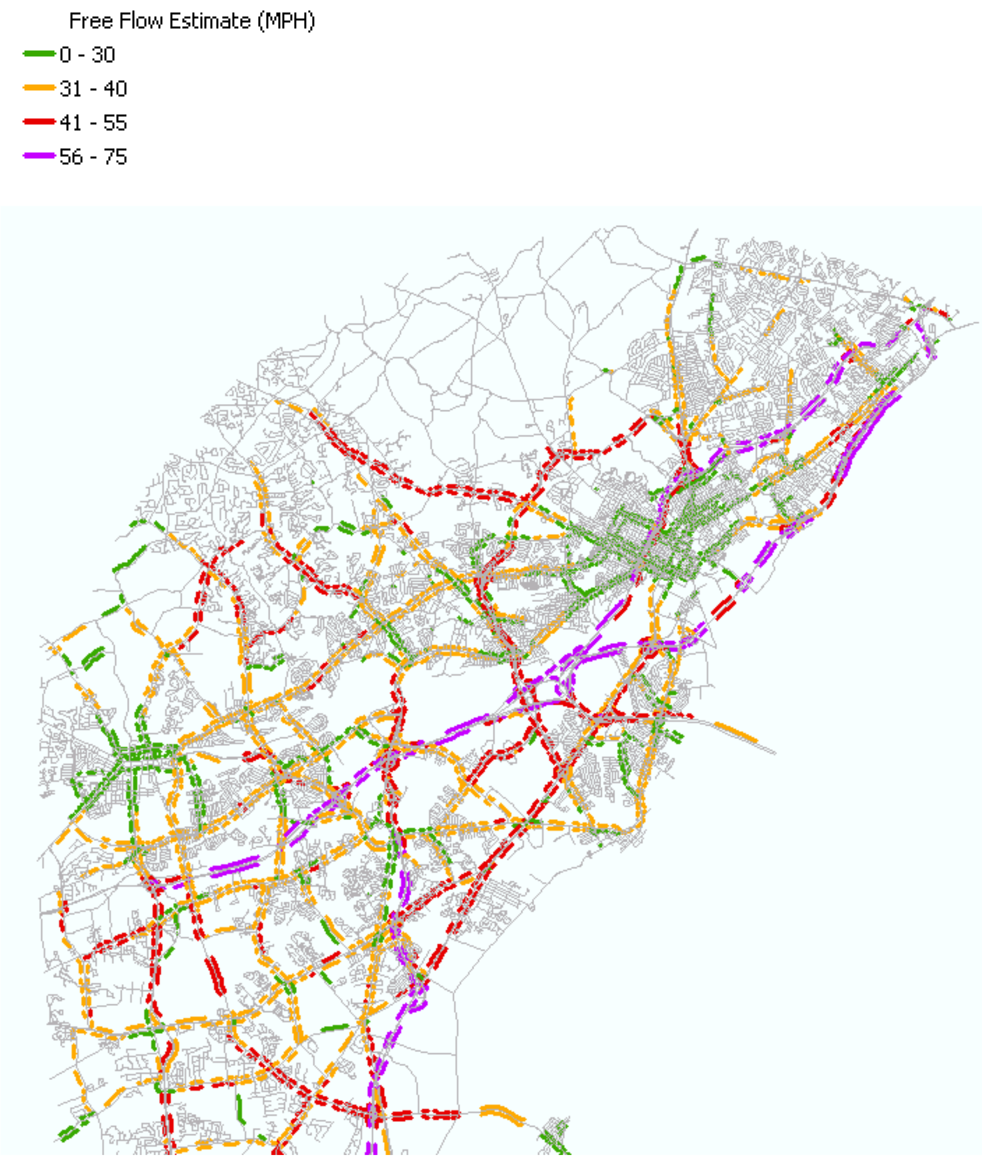
Free flow Speed

The free flow speed of a road is a useful measure that provides the expected speed where there is no congestion or impediment to flow. For each road it can provide a base line of operation. There are many approaches and modeling methods that have been investigated for free flow. It would be desirable to be able to determine free flow speed from the data and one approach would be to examine speeds through the day particularly late evening and other "off" times. For instance some type of average speed could be calculated for after dinner hours or some very early hours of the day. This might work well in the case where a corridor was studied using fixed sensors where all vehicle speeds were measured 24 hours a day, but where a network wide estimate is of interest, and for this State vehicle GPS data where there sometimes is not a great deal of data for each segment during off hours, another approach was taken. First an hourly average was developed for each direction and turning movement for every segment in the routing network . Then a 75 percent percentile was calculated from this set of hourly averages. Figure 15 below shows an example for a particular portion of road where this is done.   Notice that there are considerably less observations in the early morning and late evening hours therefore making the average hourly speed less reliable. Also the speeds are considerably higher between 1am and 5am which was seen often in the data for many segments suggesting they may not be the best hours to establish a baseline. For this section the 75% speed was calculated at 41 mph. It is important to note that this free flow speed incorporates delay from intersections and other sources and is different depending on turning movement. This approach could be refined and include additional observations.

Figure 15, Sample hourly speed averages

lrsid	hour	speedavg	observations
0001420320003730S	1	37	3
0001420320003730S	2	51	1
0001420320003730S	3	53	7
0001420320003730S	4	54	3
0001420320003730S	5	52	2
0001420320003730S	7	32	9
0001420320003730S	8	33	42
0001420320003730S	9	36	68
0001420320003730S	10	37	53
0001420320003730S	11	38	222
0001420320003730S	12	39	147
0001420320003730S	13	39	126
0001420320003730S	14	40	153
0001420320003730S	15	40	172
0001420320003730S	16	39	120
0001420320003730S	17	37	52
0001420320003730S	18	38	53
0001420320003730S	19	38	59
0001420320003730S	20	40	36
0001420320003730S	21	40	18
0001420320003730S	22	46	13
0001420320003730S	23	46	5
0001420320003730S	24	43	2

Figure 16 An example of the estimated free flow speed for major roads using the State vehicle GPS data and the hourly 75%.. Northern New Castle County Delaware, just major roads shown, Speeds incorporate delay due to intersections and other factors



View of Congestion, Percent Degradation

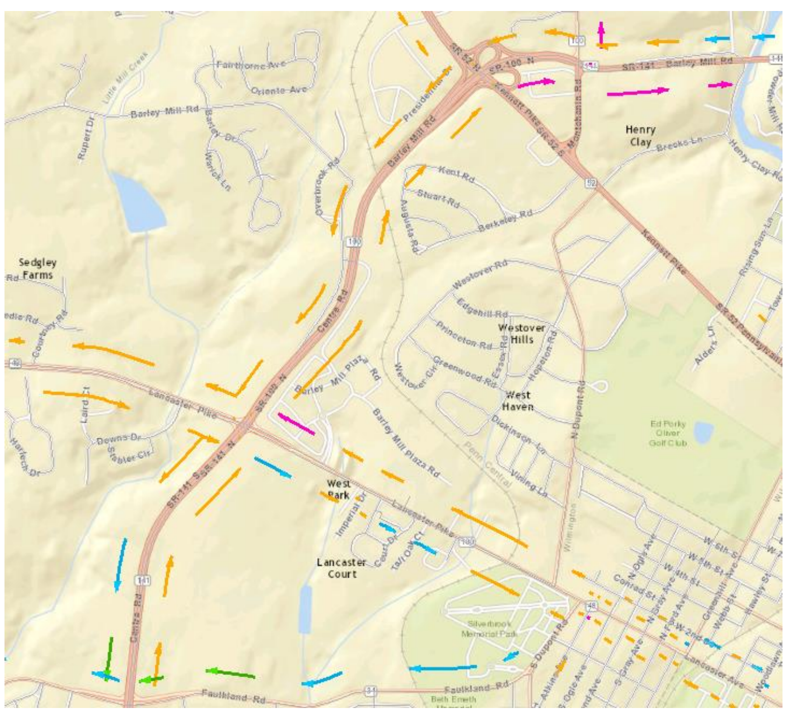
The free flow estimate for each road segment in each direction can be used to examine congestion. A percent degradation in speed can be calculated for a time interval by looking at the difference in average speed and the free flow speed (100 * (free flow – average speed) / free flow speed). Figure 17 below shows the result for the time interval between 7 and 8am. This type of view relative to a roadway's free flow baseline contains more information that just viewing average speeds.

Figure 17, Estimated percentage difference (decrease) between free flow speed and average morning speed between 7 and 8am. Northern New Castle County. (MPH)

Percentage Degradation 8am

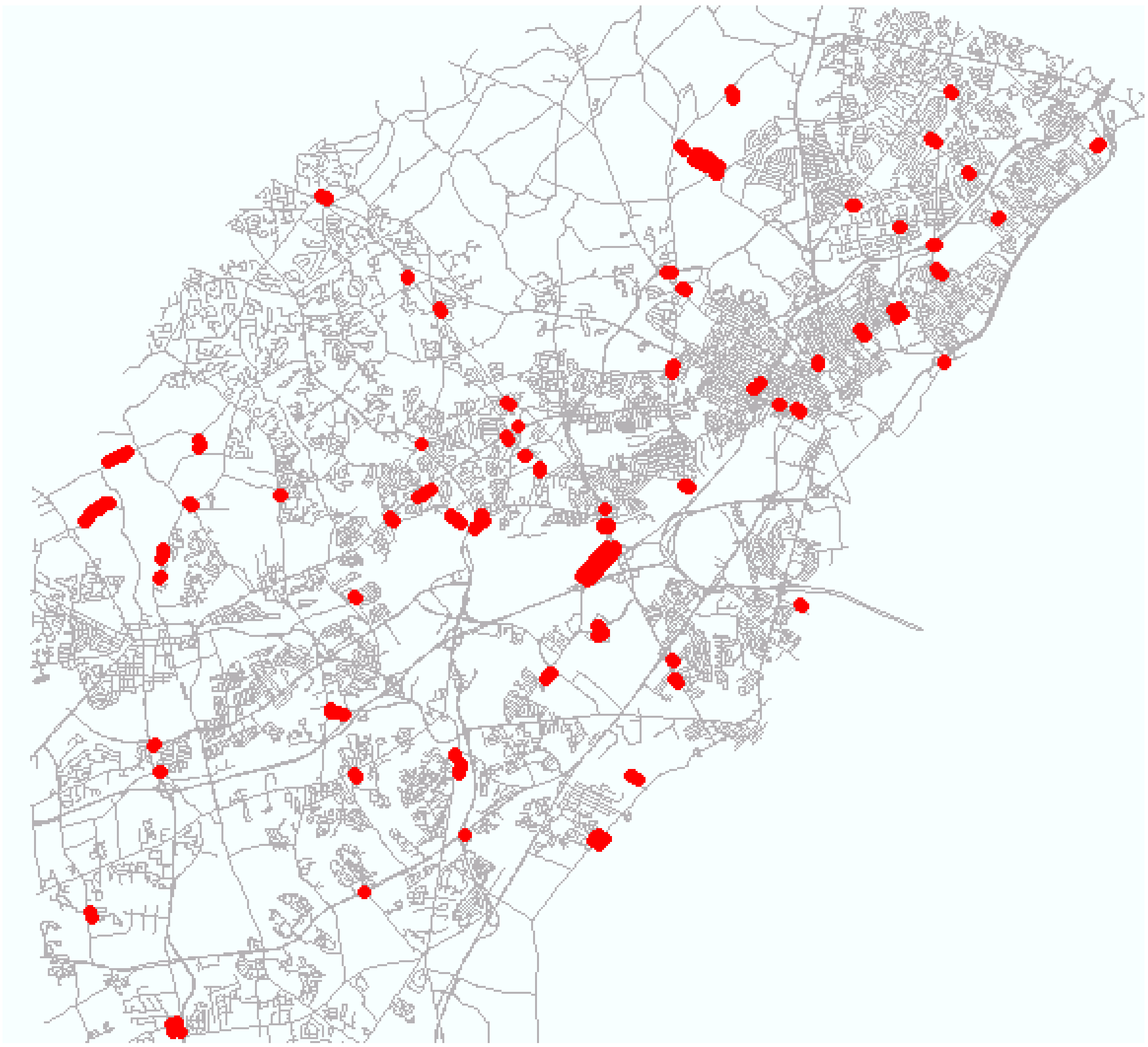
- 1 - 10
- 11 - 20
- 21 - 30
- 31 - 40
- 41 - 80

Sample exploded view



Previous example maps of the percentage degradation at the 8AM hour reference the speed estimated for road segments where travel to the next road portion is straight thru. Separate estimates of travel for segments are created where the next road would involve a right or left turn. A similar difference could be studied between a free flow condition and a congested time period. For instance, the figure below shows the location of those left turns that are most effected (> 40% degradation) by morning (8am) congestion. Such results could be investigated further. The data set could say a great deal about performance at intersections.

Figure 18, Left turns that are most effected (> 40% degradation) by morning (8am) congestions



Development of Impedances for a Statewide Routing Network

One of the most promising aspects of the State vehicle GPS data is the potential to be used to develop impedances that can be used for detailed routing applications. Optimum path calculations, and accessibility studies require routing networks that include an impedance for each road segment typically expressed in the amount of time it takes to traverse the segment. Almost always these impedances have been provided by figures estimated roughly from posted speeds or functional class or other road categories. More rare is a statewide routing network that includes directional impedances based on time of day or day of week or season that has been developed from a speed survey. Speed measurements for GPS are processed where the output is a road segment based set of speed estimates. The resulting output of processing the State vehicle GPS data is a network wide source for impedance (travel time) that is directional and can be expressed for various times of the day, season, or other factors. Because speeds for turning movements are captured, turning movement impedances as well as road link impedances are a byproduct. Each year more and more data is available to refine impedance estimates. It is hard to imagine the number of fixed sensors that would need to be deployed to achieve a similar time of day sensitive network wide impedance.

Travel Time Reliability

Beyond average speeds under various conditions, State vehicle GPS data has been shown to be useful in calculation of travel time reliability. Figure 19 provides an example output table from a study on a major corridor in Delaware. That data was drawn from a number of actual trips that traversed the full corridor and it was shown that there is sufficient data to estimate reliability figures for major corridors in Delaware.

Figure 19, Travel Times and Reliability Measures for Eastbound Trips on Old Baltimore Pike: Rt 72 to Rt 273, selections of 2010 and 2011 State vehicle GPS data, travel time in minutes

Trip Category	Mean Time	95%	Buffer Index	Planning Time Index	Avg Hourly Volume	#Trips
All Trips	6.4	8.1	.27	1.3	334	712
Peak (AM/PM)	6.7	8.5	.27	1.4	664	149
AM Peak (7-9)	**6.4**	7.8	.22	1.3	632	76
PM Peak (16-18)	7.0	9.8	.40	1.6	697	73
Midday (9-16)	6.4	7.9	.23	1.3	506	366
Weekend (all day)	6.5	8.3	.28	1.4	278	143
Off Time (<7 & >17)	6.0	7.8	.30	1.3	152	54
All Trips 2010+Jan11	6.6	8.4	.27	1.4	328	316
All Trips 2011 -Jan11	6.3	7.8	.24	1.3	336	396
Peak 2010+	7.1	9.8	.38	1.6	647	71
Peak 2011-	6.4	8.0	.25	1.3	674	78
Midday 2010+	6.5	8.3	.28	1.4	507	138
Midday 2011-	6.3	7.8	.24	1.3	508	228

- Links = (0000420333003780S to 0000420681006990S)
- The free flow time used in the Planning Time Index was the mean value of Off Time trips.

A problem for calculation of travel time reliability that is presented even with very larger data sets is that in some cases the number of trips that traversed a corridor or particular path of interest may be small if the path under study is long or circuitous. In that case, a data set such as the State vehicle GPS data which compiles data on detailed road segments is useful. Average travel times of a particular path can be estimated by a sum of averages for links, but estimating the variability requires another approach. An approach has been developed by CADSR to estimate travel time variability in the case where the number of trips actually traversing a given path is small but there is sufficient data to estimate travel time for the smaller portions of road that would make up the path. It is a method that involves the synthetic generation of travel times for non-normal distributions based on vehicle GPS data paths while accounting for correlations, and is beyond the scope of this report. It can be applied to any path or corridor and is under further study. Publication of this approach using Delaware State vehicle GPS day by the Center for Applied Demography is expected by Fall of 2014.

Reference to other potential applications

There are several other applications for a statewide speed / travel time survey. Below are references to some of these.

Before and After Studies

When some major or minor change has occurred in the transportation system or some improvement has been made, the speed and travel time data could be reviewed and mapped to see the effects of the change in the proximity and neighborhood of the change. For instance, if a light or series of signals go through a timing change, the effects at the intersection(s) could be studied as well as the neighboring area around the intersections. Sometimes when a corridor is optimized there may be some effect on travel on roads near the change.

Examining Delay at Intersections

In the State vehicle GPS data set delay at intersections is incorporated into the estimated road link speed and travel time estimate. There would seem to be some way to analyze this delay relative to all turning movements and the delay at intersections could be described.

Estimation of Capacity and Studies of the Volume and Speed Relationship

Estimations of speed at specific times for roads all across the state road network can be combined with available measures of traffic volume to support estimation of capacity and studies of the volume and speed relationship for specific roads.

Relating Traffic Flow To Land Use and Travel Demand

A speed survey captures the most fundamental performance measure for the transportation network. Planning focus is continually on the effects of land use development and trip generators and attractors and studies could be performed that shed light on how to address specific locales of interest. State vehicle GPS data has been collected as far back as year 2007. Once this data is processed, changes to the transportation system over time could be studied and related to detailed historic land use data.

Multimodal Studies

Understanding travel speeds in the vicinity of multimodal improvements would be very useful. Issues of safety could be addressed when examining travel speeds in relation to pedestrian and bicycle traffic.

CONCLUSIONS

For the first time a huge amount of detailed travel speed and time data is available for roads throughout Delaware. This data continues to be collected and there is the potential for having measures as far back as year 2007. Indications are that the State vehicle GPS data can provide valuable performance measures and serve several applications. Methods demonstrated in this project for referencing speed data could be extended to support relation and integration with a wide range of traffic and planning data. The value of the data is in its ability to describe travel characteristics of small and large roads and to capture measures at a very detailed level that includes directional flow, tuning movement, time of day, day of week, month and year.