THE ECONOMICS AND STATISTICS OF PASSIVE TELEMATICS MONITORING AS A SOURCE OF TRAFFIC DATA [A STUDY ON THE I-90 AND I-87 CORRIDORS]

Final Report on Phase I

Prepared for

THE NEW YORK STATE
ENERGY RESEARCH AND DEVELOPMENT AUTHORITY
Albany, New York

Joseph D. Tario, PE
Senior Project Manager

THE NEW YORK STATE
DEPARTMENT OF TRANSPORTATION
Albany, New York

Richard McDonough
Transportation Specialist III

Prepared by

CALMAR TELEMATICS LLC
Liverpool, New York

Ross D. Sheckler
Project Manager

NYSERDA Agreement Number: 9110-1-0

NYSERDA
Report 20079110

July 2007

NOTICE

ABSTRACT

A study was undertaken to determine the statistics and economics associated with the use of the passive monitoring of commercial vehicle telematics systems as a source of highway traffic data. Relationships were established with a group of private sector commercial trucking firms and the New York State Motor Truck Associations. Systems were devised for extracting data from vehicle messaging systems while honoring the privacy and security concerns of the commercial carriers. A database of traffic activity was constructed, a stream of current traffic data conditions was fed to a publicly accessible web site and a stream of current traffic data conditions was made available to state engineers. It was determined that through establishing a close partnership with the local trucking industry it was possible to provide traffic monitoring on the major roads of New York State at a cost significantly below the cost of installing hard sensor systems.

TABLE OF CONTENTS

FIGURES

TABLES

SUMMARY

Calmar Telematics has completed the first year of a three year effort to build a highway information system, known as HIVIS, based upon the data mining of the millions of messages that are passed between commercial vehicles and their dispatchers each day. The first year of this project has been focused on the establishment of supply chain structures, the statistical and economic realities of this data source and the business model considerations that must be taken into account in constructing a highway information system around passive probe data.

This final report covers the work laid out in the proposal to NYSERDA as Phase1A and Phase 1B. The primary tasks of establishing the flow of data from active fleets into a prototype database, performing highway speed and traffic volume analysis, and analyzing the necessary fleet size and the resulting economics of collecting data have all been accomplished. Progress which moves the program ahead of schedule has been made in the area of user interfaces and the development of advanced statistics.

Overall, at the conclusion of Phase I, the project is ahead of schedule and on budget. With the capability to model the traffic flow on the Interstates in New York established and with a variety of clients requesting data to support 2008 projects, Calmar can assert that the further NYSERDA goals of executing a commercialization plan are well underway.

Accomplishments of Phase I Include:

- Fleet to HIVIS software completed and in service.
- Prototype GIS and database system completed.
- Security measures approved by NYSMTA and are in place.
- Fleets have joined and data is flowing through the process at a rate in excess of 100,000 points per day.
- Costs of acquiring data have been established.
- The second generation of GIS methods and databasing systems have been built
- The New York State base-map has been geometrically redefined to be applicable for probe data.
- The base highway traffic statistics are being generated and a road-segment database is being populated.
- Basic traffic statistics are being delivered to participating fleets via the MyHIVIS web site.
- The sample size necessary for traffic information has been established.

1.0 DISCUSSION

Calmar Telematics has entered into a two phase program with the New York State Energy Research and Development Authority (NYSERDA) and the New York State Department of Transportation (NYSDOT) to assess the issues associated with using commercial vehicles as probes to measure the status of traffic on the highways. The resulting Highway Visibility System (HIVIS) is actively producing traffic speed data along the Interstate 90 and 87 corridors in New York State.

The first phase of this program has been focused on the study of the statistics and economics associated with the use of the passive monitoring of commercial vehicle telematics systems as a source of traffic data. This program has included an extensive effort is to establish the data retrieval mechanism and relationships of a data supply chain. To this end, relationships have been developed between Calmar Telematics, the commercial fleet operators, and the New York State Motor Truck Associations. The relationships with these 'Pioneer Fleets' are fostered through continuous communication and feedback providing a basis for the establishment of the economics of accessing data.

The second phase of the program will increase the scope of the effort from a regional to a national program incorporating a bigger segment of the industry. The technical aspects of the second phase are focused around the integration of the resulting data stream into a series of interfaces for use by a variety of stakeholders in the process. Tools for accessing the data for standardized analysis and for special research needs will be formalized and data streams established in Phase 1 will be directed to specific traffic data interfaces.

Specifically, Phase II will see the construction of a web based user interface which merges the primary Calmar Telematics traffic speed data feed with a number of secondary feeds from other commercial sources and state sensors to provide a comprehensive view of the traffic conditions feeding the Champlain Border crossing region. Later in this second phase Calmar's unique knowledge of the commercial vehicle messaging industry will leverage the same data resources to support the NYSDOT efforts to build an Intelligent Transportation System on the Tappan Zee Corridor.

Security Measures and the Fleet to HIVIS Software

By its very nature passive probe data is generated by entities that are not primarily, or even secondarily, interested in the generation of this data, which demands that the act of collecting this data not interfere with the primary activities of the parties that are generating the data. In most cases this means that no additional risk can be incurred through this activity. Therefore, privacy must be a primary concern of the organization that is acquiring the data. In the case of data that is generated through the telematics systems of commercial vehicles the demand for privacy is written in bold.

While not all commercial fleets profess an extreme concern for privacy, for many it is such an important issue that no level of security, even measures far beyond their own in-house security, will convince them to release data. Calmar has gone to great lengths to provide multiple layers of encryption and data protection to honor these concerns as they are quite important.

Calmar began its traffic data ventures from within the trucking industry. Early conceptual discussions on an industry based solution to traffic data began with the American Trucking Association (ATA). When it became time to start acquiring data, Calmar focused again on the local level and went to the New York State Motor Truck Associations (NYSMTA) to be certain that the company's purpose and methods aligned with the best interest of the industry. The NYSMTA has taken on the role of "agent and auditor" of Calmar's processes, thoroughly testing the software which is used for primary encryption and data exchange. Furthermore, they have continual access to Calmar's servers and staff in order to regularly monitor the multiple security measures. The letter from NYSMTA announcing the pilot program is included in Figure 1.

Through the rest of 2007 Calmar expects to win endorsements from key motor truck associations (MTA's) and the ATA as efforts move from New York to the rest of the country. These relationships are not just a strategy for Calmar it is part of our identity. From the first step to the last, this operation must be an industry effort. That is the level of concern that must be brought to the issue of data privacy.

Calmar Telematics has gone to great lengths to develop multiple layers of security and identity scrubbing to keep this information anonymous. The first step comes by establishing a 'data push' from within the fleet's firewall. A small Windows application is installed on the fleet's database server which the fleet IT staff sets up to pass chosen fields out to Calmar's servers at regular (1 minute) intervals. This software has been tested by the IT staff of the NYSMTA in the future months the fleets will actually download this polling application from the NYSMTA web site. This HIVIS pushing application can be terminated at any time by the operator. Encrypting the name and latitude of the vehicle before it is sent to Calmar Telematics Services creates a second layer of protection. The third and final protection layer scrubs vehicle identities from the records and transfers the records to a separate database, and the original records are wiped from the Telematics Services database. A system flow diagram is given in Figure 2.

Dear Member:

The New York State Motor Truck Association (NYSMTA) is partnering with Calmar Telematics to develop a program to help the efficiency and productivity of your fleet. In addition, this same program will help government make important decisions regarding our state infrastructure.

The program, called HIVIS (Highway Visibility System), uses information that is collected via on-board satellite systems. This information is used to generate maps that will give you information on traffic congestion, road use and conditions, speed of travel and more! The program works by allowing Calmar to view information that is already being collected from your on-board systems. The information is stripped of your company identifiers and then sent to Calmar. Calmar will then combine your data with the data of other carriers to generate a mapping system that shows where your trucks are as compared to areas of congestion or bad road conditions. This will allow you to be pro-active in determining which way to route your drivers. The program will help you avoid lengthy delays by seeing where the congestion is before your drivers get there.

In addition, Calmar has been contracted by the state to provide the information to various government agencies for their use in determining where their efforts and money should be concentrated to improve our highway system. This program will allow them to see the most heavily traveled routes in order to make a more informed decision regarding construction projects, road improvement and traffic design.

The NYSMTA will be taking on the role of auditor to ensure that data you provide to Calmar is indeed stripped of your company identifiers and is secure. We will monitor the flow of information to make sure that data that is given to the government cannot be tracked back to the originating carrier. We have been given copies of the software program being used and will also have access to the Calmar systems to verify that all information is stripped and completely anonymous. The NYSMTA will monitor the data flow from your system, to the Calmar system, and ultimately to the national database system. Our primary function as part of this program is to protect the security and anonymity of our members.

If you are currently using any type of on-board satellite system, or trailer tracking system, we encourage you to get involved with this program. It becomes more beneficial as more carriers participate!

For more information on the HIVIS program or to set up a meeting with Calmar to discuss your involvement please contact Bill McAnelly at 315-247-6302, or email bmcanelly@nytrucks.org. We will be happy to answer any questions and address any concerns you may have.

Sincerely,

William G. Joyce, Jr.

Figure 1: NYSMTA Support Letter

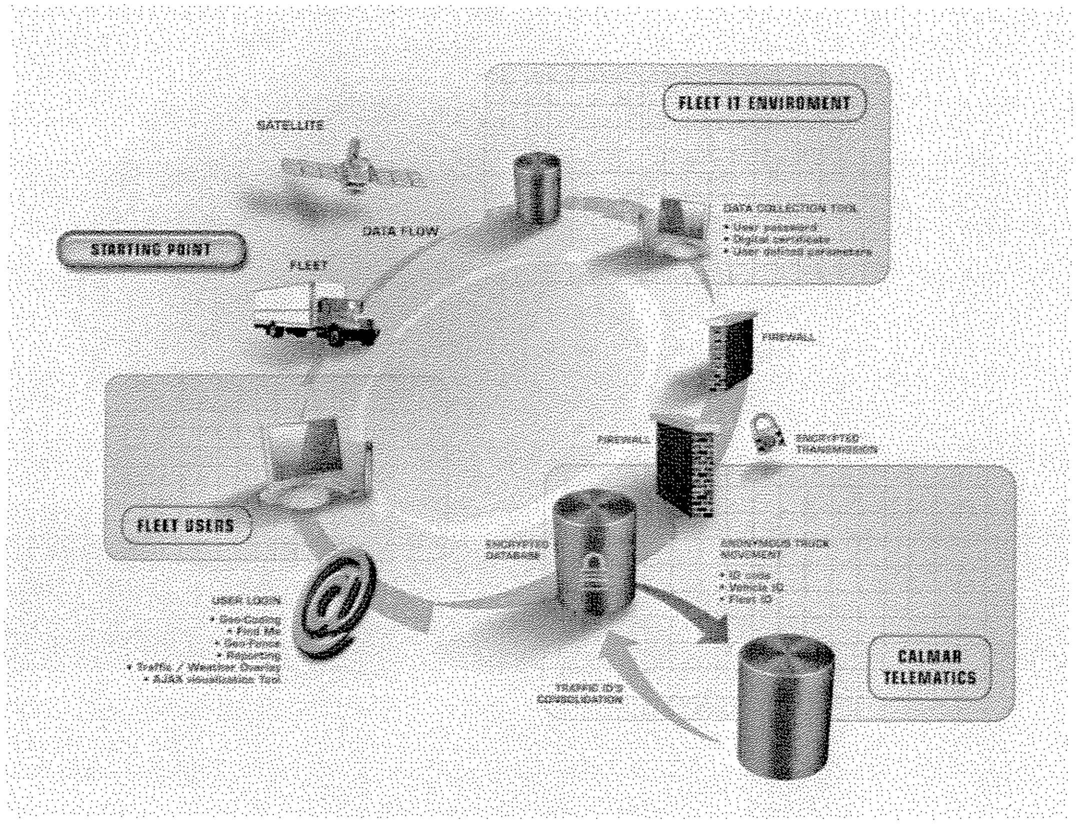

Figure 2: Telematics Data Cycle

GIS, Maps, and Roads

In order for telematics information to be useful, each individual record must be identified with a specific road segment. The process of geo-coding telematics data to road segments requires a high quality road network. A review of the map systems that are available either commercially or as public information reveals that accuracy varies from a few feet of deviation at the centerline to errors that are so large it is impossible to identify a global positioning system (GPS) location with the proper road.

Early in this first phase of the program Calmar's subcontractor, The Institute for Applied Geo-spatial Technologies (IAGT), was tasked to establish methods for geo-coding which utilized commercial off the shelf software from ESRII. While this method functioned adequately they were found to be excessively time consuming and the cost of the commercial-off-the-shelf (COTS) software was not in proportion to its function. In early 2007 the Calmar team developed its own geo-coding routine that is able to process 150 data points per second.

Over the previous six months Calmar has been actively importing millions of telematics data points and associating this data with road segments around the North East. Through this process Calmar Telematics has established a number of metrics to measure the quality of a road network and has also developed a hybrid road network based on these metrics. These metrics involve the geometric accuracy and completeness of the road network as well as the mathematical efficiency of the model. Many of these concerns are unique to the process of using probe vehicles for traffic applications and Calmar was unable to find any map set that completely met its requirements. In order to build efficiencies into the technical aspects of Phase I Calmar restructured New York State's ALIS network using a combination of program and internal funds. [Appendix A includes a discussion of road segmentation and the mathematics of road networks and how it relates to the use of probe data.]

In general the GPS data from telematics devices has proven itself to have a higher degree of repeatability, and one presumes accuracy than even the accurate map systems such as ALIS. The present state of the HIVIS system is such that it can provide lane-specific speed results where the map systems are defined to the proper degree of accuracy. In the coming years Calmar will undertake efforts to define the lane-centers to within a foot of accuracy.

Database Development

Data received by Calmar goes through a series of servers and processes as is used to populate both short term and long term data systems. The hardware and software to receive and store incoming data has been purchased and successfully brought online. A T1 line currently provides sufficient bandwidth and speed for receiving telematics data as well as the fleet specific user interface communication requirements of the MyHIVIS website.

Databases of highway segment statistics are maintained on a minute-by-minute, daily, and day-of-the-week by month basis, an overall geo-coded data point archive database is maintained, and soon an intersection behavior database will be populated.

Data is first received by Calmar into a 'threshold server' where the data is maintained for one week for use by the originating vehicle's owner. Shortly after, each data point is re-labeled with its vehicle's week specific, randomly created ID and transferred to the HIVIS archive database. Thus, a string of reports from a single vehicle has been maintained but the actual identity has again been encrypted.

The highway statistics are collected for each minute and rolled up into five minute intervals for reporting to New York State and for sharing with the fleets. This five minute stream is available to New York State agencies upon request. Since early 2007 Calmar has maintained average speeds and counts on 5 minute intervals for each day-of-the-week for every month (i.e. 8:00-8:05am for Monday's in May) for roads in New York State.

Beginning in July Calmar will start to aggregate intersection statistics in a manner similar to the road segment statistics.

Establishing backup and secondary servers will be a major part of the second phase efforts, as well as communicating and maintaining a live feed service.

Data Latency

One of the primary concerns frequently associated with the use of vehicle probes as a source of traffic data is the very timeliness or 'latency' of the data system. Calmar has found that by working closely with the fleets themselves data latency can be reduced to very acceptable levels.

There are two forms of reporting systems commonly used in telematics units: real-time, which transmits data immediately upon creation, and store-and-forward, which periodically transmits bulk data dumps. Real-time units have remote communications via cellular or satellite carriers. These units are programmed to send location and vehicle state information at regular intervals and at the occurrence of specified events such as stopping. Alternatively, units that are referred to as store-and-forward are programmed to store multiple location and state data points to be transferred in larger blocks information at once.

While the real-time telematics systems are directly applicable to the production of traffic information, generally, only the last point or two from store and forward systems have any value in the production of real-time traffic information. It is worth noting that the typically more detailed route information that is generated by store-and-forward systems is highly valuable in route and commerce flow studies..

A latency test for this real-time data was performed on June 18th, 2007, calculating the time difference from a record's creation to its insertion into the HIVIS database. The results for this test demonstrated that the average latency is usually less than 2 minutes. A distribution of data latency is shown in Figure 3.

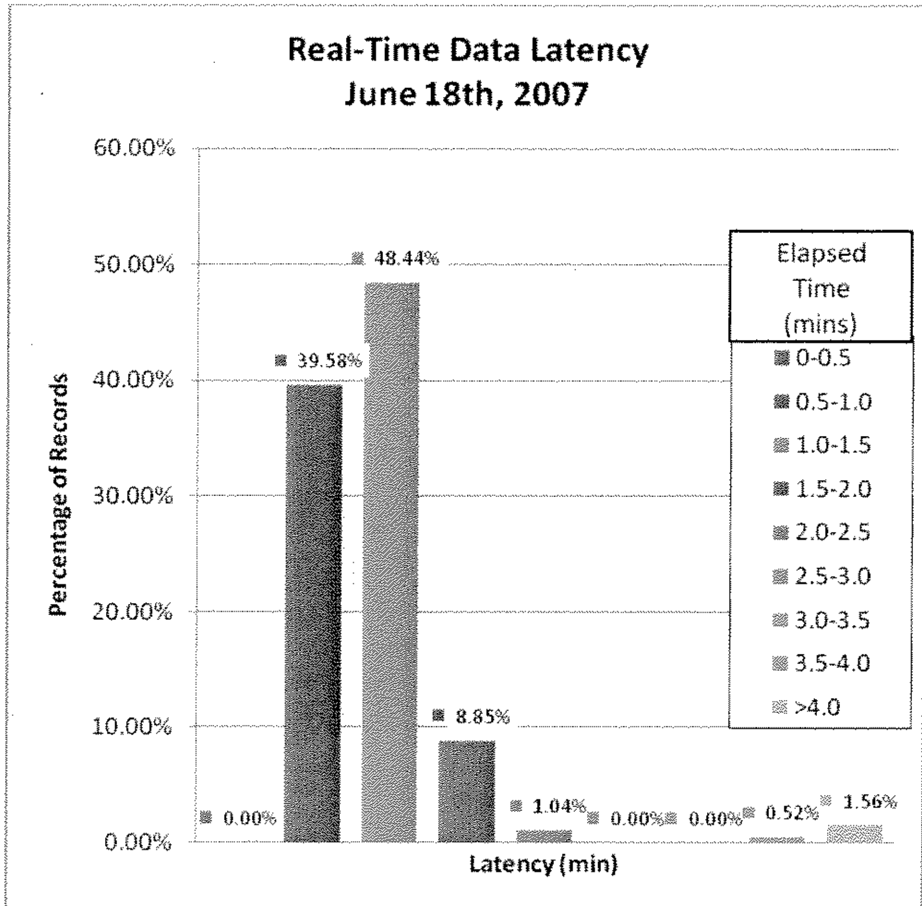

Figure 3: Typical Real-Time Data Latency Results

Mean Segment Traffic Speed

There are a couple of unique concerns with calculating mean segment speeds using telematics data each of these is associated with the source of data, the probes, moving about the subject road network independently.

The first problem that is noticed lies, once again, in the road segmentation. If the roads are not segmented in a manner that is appropriate for probe data the results will be sporadic. For instance, the ALIS base-map has many road segments that are less than 100 feet in length, with a statistical data method such as probe tracking the odds of actually having a data point land on that 100 ft. length of road is very low and the segmentation become irrelevant.

The less obvious issue in the use of probes for traffic speeds is a bias towards slower speeds. This arises from the fact that polling isn't done at specific locations, but rather at a set time interval or upon the occurrence of specific events. So, instead of one vehicle passing a single location once and therefore being counted exactly once for that road segment, each vehicle spends time on each road segment and therefore slower speeds lead to more records. This affects calculations in two main situations, when traffic speed changes greatly within the defined time interval, and traffic is really moving at two separate speeds, like on a multiple lane highway.

To understand the extent of this issue, the mean speed was compared to the median speed, while ignoring any possible bias, for 15 minute time intervals. Ideally, the median speed will exactly equal the mean speed, displaying that the 50% mark is also the average and therefore the bias mentioned above is insignificant. Figure 4 shows that the median tends to be slightly higher than the mean, with an average difference between the median and the mean of 0.7mph and a standard deviation of 2.2mph. So, while other methods of determining a more accurate average segment speed should still be explored, the current bias does nothing more than make calculated mean segment speeds conservative estimates.

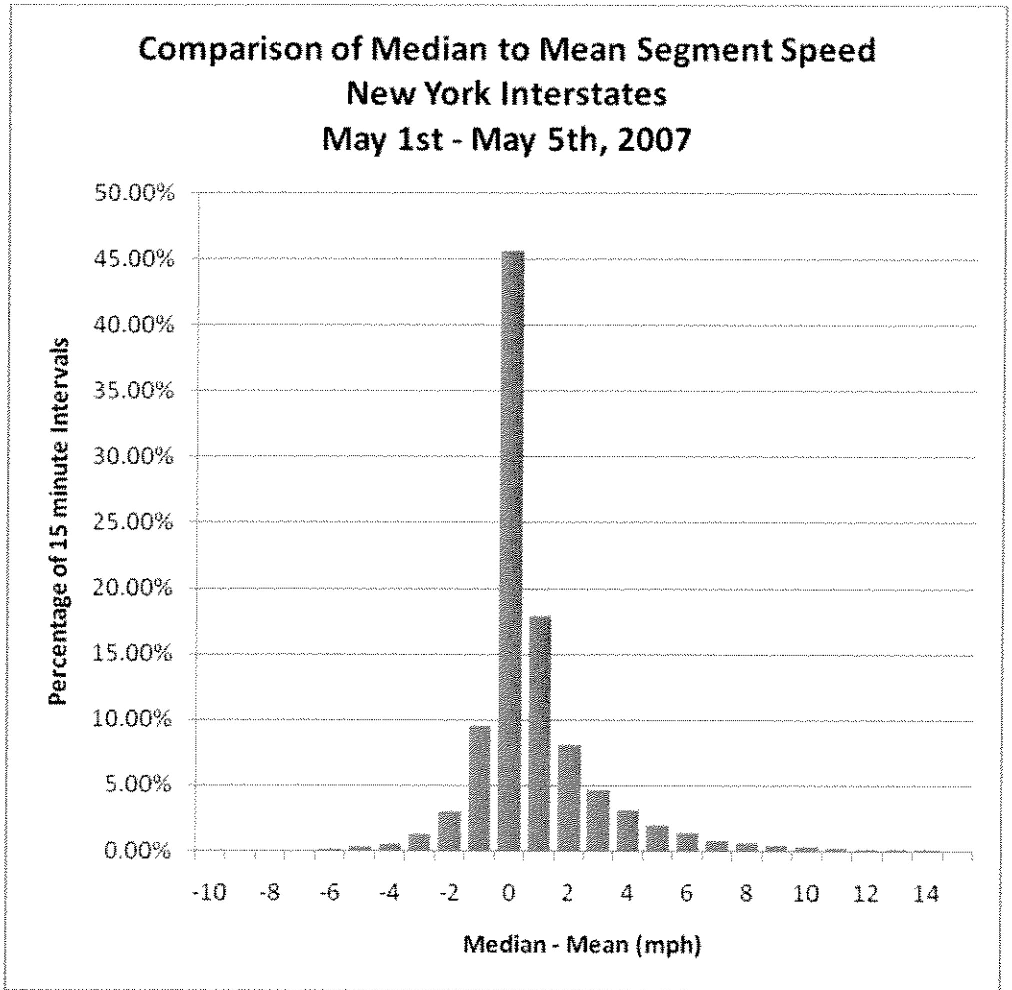

Figure 4: Median Speed to Mean Speed Comparison

Once it was determined that mean segment speeds were not unacceptably biased, their results could be validated. The New York State Thruway Authority (NYSTA) uses dual inductance-loop sensors to monitor average speeds and traffic counts. Using this data around the Albany area from May 1st through May 5th, it was found that the Calmar data has an absolute average error of 5.615 mph and a speed error bias of -4.48 mph. Some of this deviation undoubtedly reflects the tendency of commercial vehicles to travel slightly slower than passenger vehicles. The consistency of this error, approximately 5 mph slower than actual traffic, for the five day period on the segment of eastbound I-90 between exits 25 and 24 is shown in Figure 5. This sample of data also shows the important fact that Calmar's data consistently detected the Albany inbound, morning rush-hour congestion during the week and the congestion relief on Friday and Saturday.

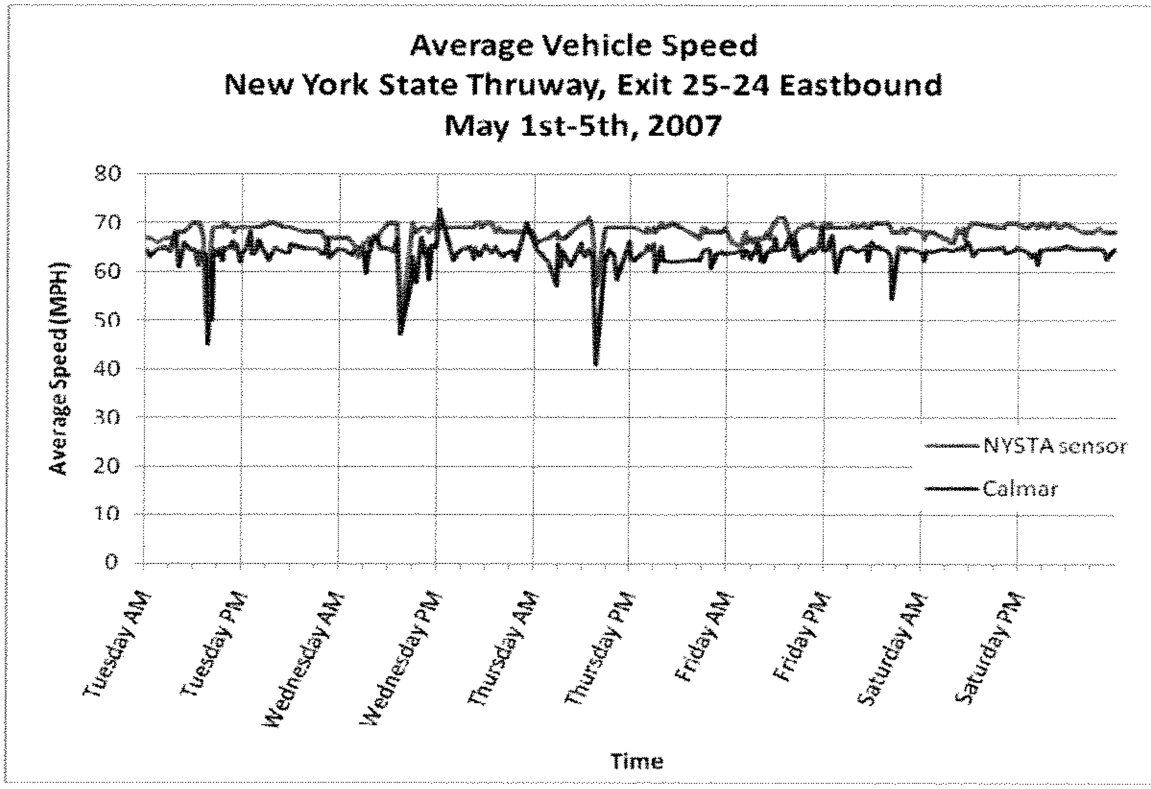

Figure 5: Average Vehicle Speed Validation

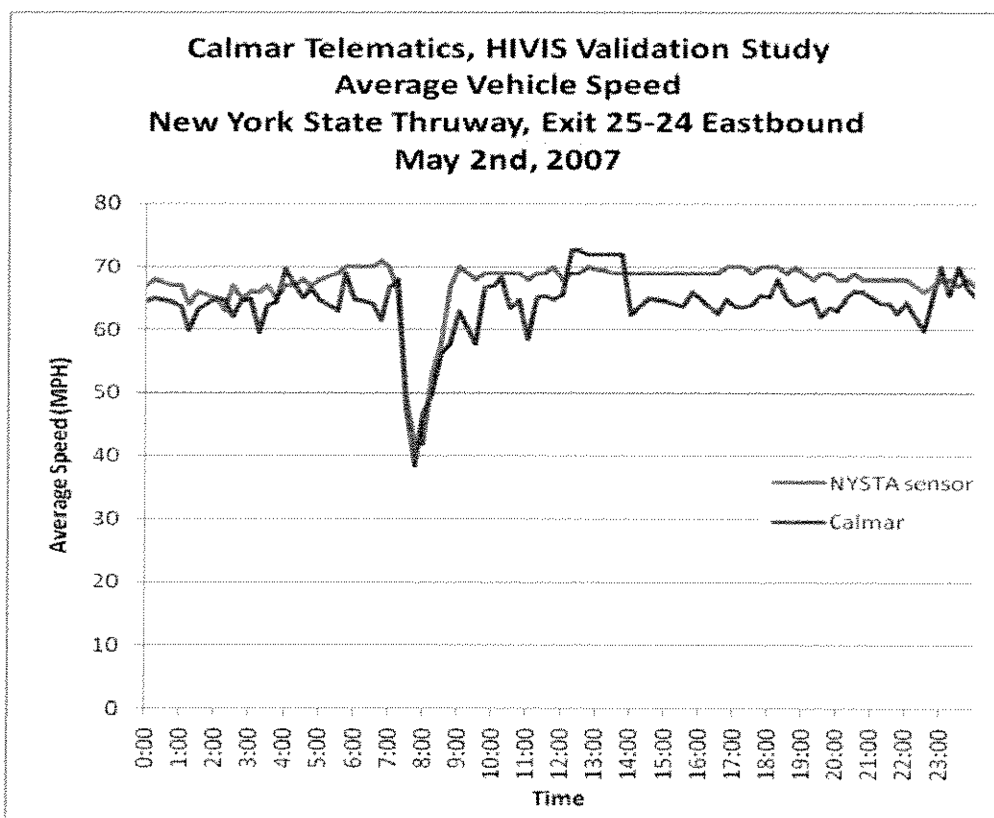

Figure 6: Average Vehicle Speed Validation - Wednesday Detail

Reliability

The issue of reliability in the field of computing and communications has been the subject of significant development work and is now at the point of common practice among thousands of firms. Dual sites, back-up power, secondary communication paths, and multiple layers of redundancies are all process that are being put into practice at Calmar Telematics in 2007.

It is also important to know that in the field of traffic data there is another level of detail that must be considered in the subject of reliability and that is the area of data supply. Traffic data efforts that have been tied to a single source of data, such as a cellular provider, or a couple of sources such as a handful of telematics service providers have had repeated interruptions due to abandonment by the cellular and/or telematics providers.

Calmar's approach has been quite different signing agreements directly with a wide variety of commercial fleets which use a wide variety of communications service providers. While fleets will certainly come and go, one or two or even dozens of defections will not shut down the HIVIS system. Furthermore, because the communications systems of these fleets are varied the failure of one mode will not bring the system down completely. In fact because some of the fleets are on satellite based networks even a complete black-out of a region of the country will not take the reporting down to zero.

Data Availability

If a highway/traffic management organization is to build their operations around any source of data it is incumbent upon that organization to understand the reliability and the vulnerability of that data supply. In the case of fixed, infrastructure intensive sensors, it is important to know the life expectancy, the probability of interrupted operations, and requirements for installation and maintenance. In the case of passive telematics observation, however, the infrastructure is spread across thousands of vehicles and any one or even a dozen failures won't bring down the system. Instead the primary concern is fleet participation and data communication.

The data that is necessary to operate a traffic monitoring system exists within the trucking industry and this program has demonstrated that it can be accessed for this use. A constituency of 10,000 trucks transmitting messages at rates approaching those seen in this program will produce 3 million reports each day on ACC Class 1-3 roads in New York. This equates to one report every 10 minutes for every mile of road. This is a fleet size that is indeed achievable and it remains the responsibility of the aggregator to get a balanced distribution of vehicle operations and vehicle types. Moreover it is the responsibility of the aggregator to create a business arrangement which provides the right combination of risk and cost to compel fleets to participate as well as reward them for doing so. Calmar Telematics feels that the groundwork laid in Phase I of this program is a strong basis for reaching this level of data flow by mid 2008 and maintaining it well into the future.

Report Rate and Highway Presence

Through the term of Phase 1 of this project Calmar received data from a variety of telematics systems with report rates ranging from an hour to less than a minute. The average report rate during Phase 1 was roughly two minutes, significantly better than the expected 15 minutes. This gives an even higher data resolution than was originally expected.

Calmar focused efforts on New York and the Albany region in particular. As a result, the fleets which participated in this phase exhibited a strong presence on the primary highways in the region as shown in Table 1.

Highway	% of All NY Records
I-90 MA to Albany	2.93%
I-90 Albany to Syracuse	12.92%
I-90 Syracuse to Buffalo	4.12%
I-87 Tappan Zee to Albany	5.34%
I-87 Albany to Canada	8.64%
I-88 Binghamton to Schenectady	1.26%
I-81 Binghamton to Syracuse	1.26%
I-81 Syracuse to Canada	1.04%

Table 1: Interstate Distribution of Data in New York

The combination of fleet size and report rate has resulted in the ability to provide reasonably complete traffic monitoring in the area of focus, around the Capital District of Albany. The following figures are reports of 15 minute intervals between 8AM and 9AM in the month of May 2007.

Figure 7: Albany Area Traffic Speeds 8:00AM – 8:15AM

Figure 8: Albany Area Traffic Speeds 8:15AM – 8:30AM

Figure 9: Albany Area Traffic Speeds 8:30AM – 8:45AM

Figure 10: Albany Area Traffic Speeds 8:45AM – 9:00AM

COST ELEMENTS

Computing costs

The complete cycle of telematics data, as depicted in Figure 2, has been successfully implemented by Calmar Telematics. The cycle begins with the collection of "raw" telematics data from a number of fleets and is completed by returning this data back to the fleet user in an aggregated scrubbed format. Hardware requirements for this cycle are broken into two groups: fleet hardware and Calmar Telematics hardware. Any fleet that stores their vehicle Telematics data electronically and has access to the World Wide Web can contribute to the HIVIS project.

Calmar Telematics requires a large reliable broadband connection as well as reliable and powerful servers so that fleet data can be transmitted and received to and from the telematics servers. A T1 connection at the Calmar Telematics facility has provided the needed reliability and speed. The yearly cost of a T1 connection is around $4,800. Because the servers are currently housed at the Calmar Telematics facility there are no additional housing costs for hosting this HIVIS server. If, however, these servers were to be rented at a commercial facility the average cost can be expected to be around $7,200 a year which includes a broadband connection.

A single high powered server, residing at Calmar Telematics, has handled the current load of incoming fleet data, as well as the data dissemination loads for end-users. The one time purchasing cost of this machine was $9,338 with an additional $1,321 for software. It is expected that this server can handle around twice the current load of incoming fleet data. As more fleets participate in the HIVIS project more servers will be brought online at similar purchasing costs. The system that will eventually handle the data needs for a nationwide effort consists of approximately 25 servers and significant associated storage.

System down time can be practically eliminated with redundant servers at a remote site. The hardware and connection costs for these redundant servers will be similar to the previously mentioned prices. Because data is not only disseminated but also consumed a mirrored sight must communicate with the other servers if a complete data set is to be assembled. A mirrored as well as striped server layout is not a new problem but the size and the number of records consumed in the HIVIS project will be at an equal or greater level than existing systems.

The net computing infrastructure cost for a system that can handle a single region such as New York State is estimated to be $40,000 annually. The computing infrastructure required to operate on a national level is estimated at $130,000 annually. The associated labor cost to operate the system on a 24x7 basis is 8 full time employees and associated management at a total labor cost of $1.3M.

Cost of data

The data costs associated with the passive monitoring of telematics data as a means of monitoring the traffic flow on major highways have four key factors; 1) the rate at which the vehicle's telematics systems report to the dispatcher, 2) the presence of the participating fleet on the roads of interest, 3) the compensation to the vehicle owner, and 4) the secondary costs of compensation.

The majority of commercial fleets in operation today have built their systems around long term relationships with key clients establishing a network of depots along carefully chosen routes. This allows one to pick fleets that focus their operations on roads of interest. In Phase I Calmar focused on fleets operating in New York and the Northeast in order to establish a strong presence on Interstates 90, 87, 88, and 81. As Phase II is launched Calmar is working to bring key east coast fleets into its system in order to build a presence along I-95 and associated highways.

Calmar has engaged in two years of discussions with a wide variety of members of the trucking industry including large, medium, and small fleets, several trade associations, and suppliers of telematics services. The discussion with each entity has been different and each brings a unique spin on the issue of releasing operational messaging data in order to achieve purposes that lie outside the primary operations of the companies.

It is essential to understand that at no time is it possible to place a value on the data from vehicle messaging that even begins to approach the value that operating the vehicle generates every day. That is to say, the truck generates more than $2.00 for each mile that it is driven or an average of $20 of revenue in the time that one message is sent. It is hard to conceive any way in which the traffic value of the data in a message can approach the significance of the freight revenue. Just the same, every commercial organization is interested in increasing revenue, provided it comes with no negative impact to the primary operation of the company.

A significant segment of the trucking industry has demonstrated a willingness to provide access to their vehicle messaging data provided there is reasonable compensation in both cash and information and provided that the data aggregator is trusted partner responsive to the needs of the fleets.

Cash Compensation

Based upon the experiences of the previous two years, Calmar has found that if an adequately secure environment is assured, a compensation scheme which gives the fleets an opportunity at a meaningful secondary revenue stream. The fleets that participated in this phase of the program were extremely hesitant set any firm price on data. Generally, the fleets expressed a real interest in participating in order to position themselves for what is viewed as significant potential upside.

The participating fleets understood that one of the primary goals of this study was to establish the basic economics of operating this system. So, when pressed, a consensus was reached that an annual monetary return beginning in the neighborhood of $100 per vehicle has been generally accepted as goal to work toward as the first satisfactory level of compensation.

In Phase II Calmar will be working with fleets, the NYSMTA, and the ATA to establish a revenue sharing scheme which provides due compensation to the fleets and encourages increases in the rate of data flow.

Secondary Costs – Non-monetary Compensation

As stated earlier, the general consensus of the trucking industry is that cash compensation will need to be accompanied with complementary traffic information products for participating fleets. This can take a number of forms including real-time traffic, historical congestion, route studies, or a variety of other metrics. A key to the successful operation of this telematics monitoring venture will be striking a balance between handing the fleets a standard set of valuable tools and charging for non-standard requests for information.

In this first phase of the program information was fed back to the fleets through a web site known as MyHIVIS, which provided participating fleets with a tool to access the traffic information. The MyHIVIS process resides on the threshold server allowing the fleets to log-in with the decryption keys necessary to identify their own vehicles among all of the other vehicles in the system. While all fleets with telematics systems have interfaces to view the location of their vehicles, the MyHIVIS site allows the fleet to overlay the location of their vehicles with both current and historical traffic conditions generated by dozens of fleets.

In the follow-on phases the MyHIVIS web site will grow to include more basic data forms and may even include weather and road construction layers. The system will necessarily be expanded to service thousands of users and direct XML feeds will go to some of the larger commercial fleets.

In order to do a complete accounting of the costs associated with the use of telematics data one must understand that the market for traffic information to the commercial operator must be serviced without compensation.

The cost of operating these complimentary data services is relatively small compared to the lost opportunity to sell this data to the commercial carriers. The opportunity costs associated with using commercial vehicles as data probes is at least as large as the cash compensation mechanism. States and other clients for real-time traffic data should understand these secondary costs of data as the removal of some of the secondary markets for the data and the loss of the opportunity to share the system cost. Incidentally, this same statement is true for cellular tracking as the successful organizations that are operating with this technology are gifting back the traffic information to the cellular carrier which then offers it to their clients.

Fleet Support

Part and parcel with compensating the fleets that choose to participate in the program is the need to maintain the fleets' confidence in the program as a whole. Calmar Telematics has established a fleet relations department staffed with trucking industry professionals. This group is tasked with maintaining a high degree of industry confidence in Calmar. This includes face-to-face meetings, ready phone support, participation in state and national trade associations, and contributions to industry conferences.

In the Northeast, Calmar has been able to operate with one full time and two half time employees in fleet relations. When overhead, travel and conference costs are added to salary the fleet support function for a system operating in the Northeast adds $400,000 to the annual cost of data.

As the system expands to the entire east coast the fleet relations staff will be increased to a staff of five. Forecasts for maintaining the fleet support task across the nation as a whole will likely require a staff of ten to fifteen.

SUMMARY AND CONCLUSIONS

Calmar Telematics has successfully built and operated a state-wide traffic data system based upon the passive and anonymous monitoring of commercial vehicle telematics systems. The system known as HIVIS is currently converting more than two million reports a month into traffic speed data on the Class 1-3 roads in New York. Accurate reporting Interstate highway speeds has been verified. This data is actively being streamed to web sites and is available to NYSDOT at their request.

This first phase of Calmar's program was intended to study the economics and statistics of using the passive monitoring of telematics systems for the specific purpose of generating highway congestion and vehicle count data.

This study has made the use of commercial vehicle probes for traffic speed data on major roads a reality in New York State. Furthermore, this study has established, that provided that a proper relationship with the trucking industry is maintained, the process can be a very affordable alternative to fixed sensors.

The operation of a data system such as HIVIS for the State of New York will require approximately 10,000 vehicles at a cost of $1M - $1.3M. Personnel, computing and facilities costs will add approximately $1.2M. The estimated annual cost of operating a telematics monitoring system as a source of traffic information, such as HIVIS, is estimated at approximately $2.5M.

The vehicles of the fleets participating in the generation of highway information for New York spread regionally throughout the Northeast. This regional nature of trucking creates a system with cost benefits that can be gained through the coordination of neighboring states. Calmar's study has indicated that a total fleet of approximately 50,000 vehicles are needed to gather traffic speed data for the eastern seaboard states resulting in an annual system cost of approximately $7M. As a point of contrast, over the past decade well over $100M has been invested in the installation of fixed sensors for traffic monitoring in the eastern seaboard states.

During the one year period of this phase, the study was not able to build a large enough sample of the total population to definitively state that vehicle counts can be estimated to the 95% confidence level required by the FHWA. Qualitative measures of volume have been successful and will be continued in the second phase of the study.

The team at Calmar Telematics is excited by the strong success of Phase I and will be continuing to expand the population of the contributing fleet in coming months. Phase II of this program will be focused on building the interfaces that are necessary for our New York State clients to access this data for real-time monitoring as well as research into operational trends in traffic on the state highways.

APPENDIX A

Why a good road network is required
(Static versus point based observation)

Road statistics are traditionally obtained through geographically static devices, such as inductance loops or toll way gates. The information gathered from these devices is assumed to apply along an adjacent stretch of road. For example the Average Annual Daily Traffic calculated by an inductance loop will be applied along that stretch of road where there are no intersections.

The major difference between the traditional methods for gathering traffic information and Calmar Telematics' point based approach is that the monitoring devices are no longer static. The GPS units used to collect information on the present as well as historical traffic patterns are permanently attached to the vehicles. However, the telematics information gathered from these devices is not continuous along the vehicle route. Instead the telematics information is collected in a point-wise manner at certain intervals. These intervals are set by the local GPS unit and can be defined by time, speed, location or any number of other factors. Ideally all the units would collect information at the same location, allowing traffic information to be acquired in a manner similar to the traditional method, but this is not the case. Therefore, information gathered from telematics units must be aggregated across continuous sections of road to obtain traffic patterns (i.e. patterns of traffic that consist of more than one vehicle) for the road segment of interest.

Road segment lengths

The length of road that the telematics information is accumulated over will have a significant effect on the volume of data that is collected. To illustrate this point consider two different road segments. The first segment consists of one hundred feet of road and the second segment consists of one mile of road and encompasses the first segment. Assume the telematics data gathered for these roads have an average density of one count per sixty feet. The first segment may only contain one reading and will not provide an accurate representation of the traffic. Aggregating this same data across a one mile segment will provide around ninety readings and therefore a much better picture of the average traffic conditions.

As this example illustrates, the longer the road segment the more readings there will be to form traffic statistics. But how long is too long? An interstate traversing the length of the state is obviously too far. The ideal length for any given road would be the length over which traffic remains relatively consistent. At a minimum this would be from one intersection to the next because vehicles do not typically leave a road until an intersection. It should be noted that there are possible anomalies to this behavior such as drive ways and unmarked intersections. The former is of little influence to the traffic pattern and the latter should be monitored by identifying those vehicles that have left the known road network.

Properties of an ideal road network

A road network should have the following properties to ensure accurate traffic patterns: 1) Geospatially accurate; GPS points will not be attributed to the correct roads if the road network is not geospatially accurate, 2) road segments broken into useful lengths; too small and not enough information will be collected, too large and the details in the data may be washed out, 3) geometrically efficient (i.e. representing the road segments with the minimum amount of "shape" points while still retaining the same level of accuracy); the speed of geo-coding (the process of "snapping" GPS points to road addresses) will be greatly increased if the road network does not contain repetitive or spatially insignificant points.

Existing road network (ALIS)

Calmar began this project using the ALIS road network. This network was used to aggregate individual GPS readings to the nearest road segment. However, ALIS does not measure up against the three previous "ideal" road network properties. The accuracy of ALIS is by far its strongest area. It is possible that the current level of accuracy will allow for the proper "snapping" of GPS points but there is always room for improvement. The segmentation of ALIS is the weakest aspect of the road network. For example, in ALIS the representation of Interstate 90 East in New York is broken into over 1800 segments. Collecting 1800 traffic patterns along I-90 is not practical, efficient, or even useful. The geometric efficiency of ALIS should be improved, helping to speed up the geo-coding process. Removing some of the unnecessary shape points can also remove some of the unwanted "wiggles" in the roads. These wiggles most likely occurred when the digitizer mapped the road based on an aerial photograph.

Segment consolidation of the ALIS road network

Calmar has developed a program to consolidate the ALIS road network into a more functional segmentation. This routine is broken into three main functions: consolidating roads, segmenting new roads, and re-defining road shape points.

A FORTRAN program was developed to consolidate the road segments. The consolidated roads are then re-segmented. The philosophy behind segmenting the roads was to maintain consolidated roads where traffic patterns remain constant. Taking this to the limit means that road segments should always break where traffic has the possibility to change, i.e. road intersections. However there are two types of intersections: True road to road intersections where vehicles can turn off their current road onto any other road and overpasses where vehicles must remain on their current road and therefore traffic patterns will remain the same before and after these intersections. This process took approximately 7.75 hours to run.

The results from the consolidation/segmentation routine are then run through an Aero-Grid and Paneling System (AGPS) script to minimize the number of shape points for each road segment based on a chord-height tolerance function. The result is a more geometrically efficient network. This routine takes around forty hours to run. At this point the road consolidation is complete and the data is converted to a road network format that is more conducive to SQL and geo-coding.

Table A-1 provides the original number of road segments and shape points as well as the resultant decrease in the number of road segments and shape points after running through the road consolidation and the geometry clean up (chord-height tolerance). Using the more consolidated road network a decrease of approximately 30 percent was found in the geo-coding time.

	ALIS (Original)	Consolidated	Segmented	Chord-Height
# of Road Segments	1,141,084	324,220	824,484	824,484
% of Original Road Segments	100%	28.4%	72.2%	72.2%
# of Shape points	6,059,708	5,242,850	5,743,106	4,409,996
% of Original Shape Points	100%	86.52%	94.77%	72.77%
Time to Complete	-	54 Minutes	7 ¾ Hours	≈40 Hours

Table A-1: Performance of Road Consolidation Algorithm for ALIS Dataset

The following two figures illustrate the amount of segmentation of the road network before and after the routine. Figure A-1 shows I-87 just south of I-90 in Albany, NY before the consolidation. Figure A-2 shows this same intersection after the consolidation. Finally, Figure A-3 shows an aerial photo of the actual intersection.

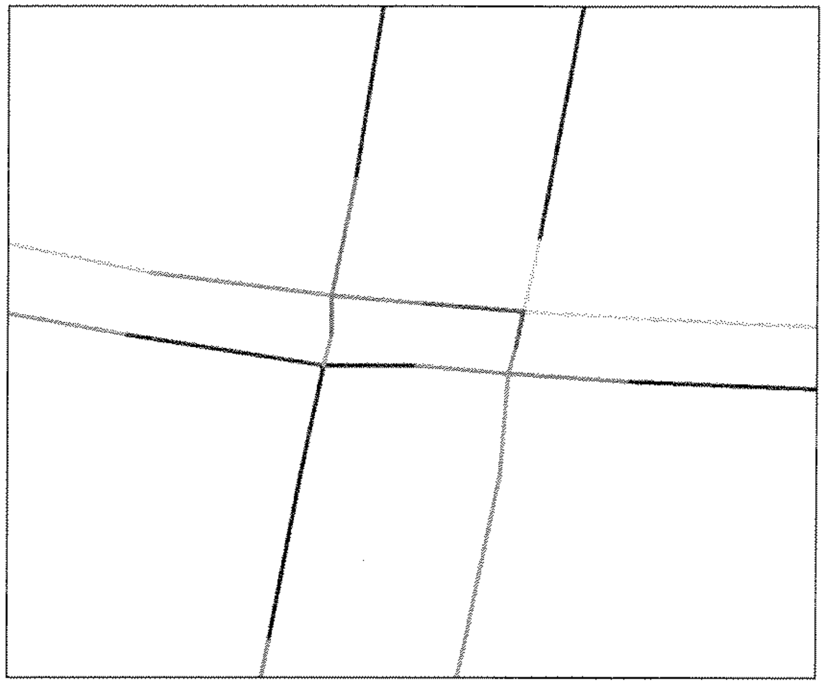

Figure A-1: Original Intersection of I-87 and Washington Ave Ext.

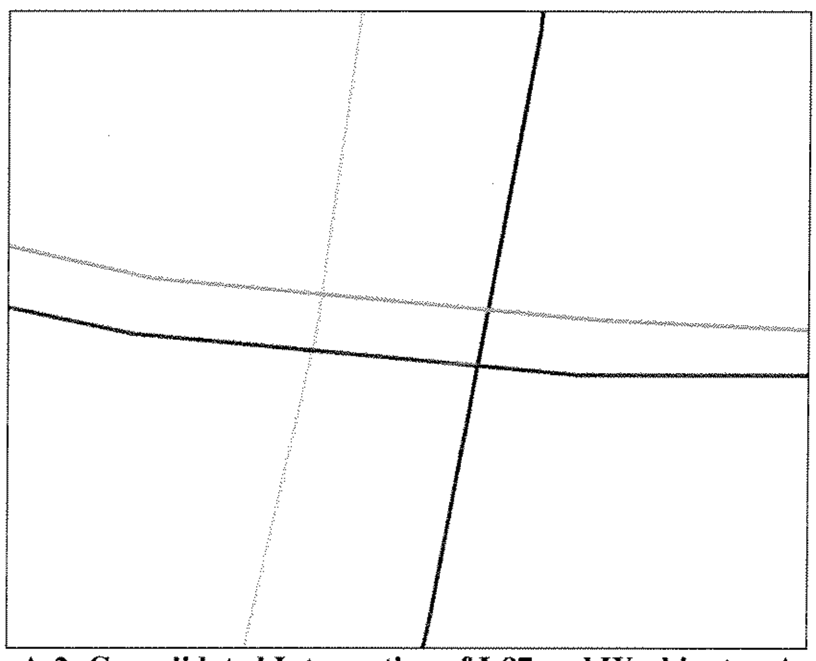

Figure A-2: Consolidated Intersection of I-87 and Washington Ave Ext.

Note the waves in the original network and how the geometry smoothing has inherently made the new road network more accurate. More research is needed to determine if smoothing will always increase the accuracy or if there are cases that will actually decrease in accuracy when smoothed.

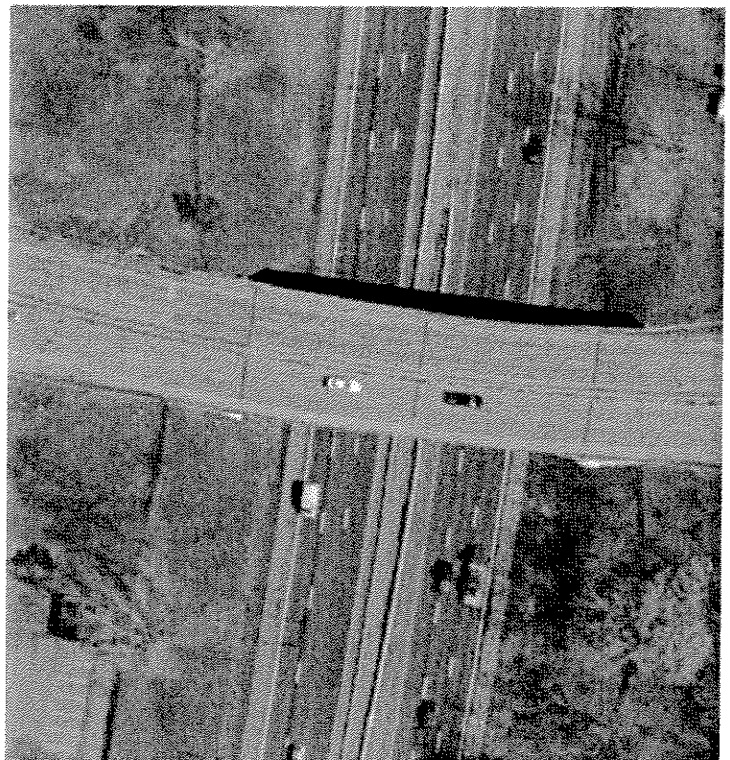

Figure A-3: Aerial Photo of I-87 and Washington Ave Ext Intersection.