Final Research Report
TNW2009-12

TransNow Budget No. 61-6022

DEVELOPMENT OF A STATEWIDE ONLINE SYSTEM FOR TRAFFIC DATA QUALITY CONTROL AND SHARING

by

Yinhai Wang
Associate Professor

Jonathan Corey
Graduate Research Assistant

Yunteng Lao
Graduate Research Assistant

Yao-Jan Wu
Graduate Research Assistant

A Report prepared for

Transportation Northwest (TransNow)
University of Washington
129 More Hall Box 352700
Seattle, Washington 98195-2700

and

Washington State Department of Transportation (WSDOT)
310 Maple Park Avenue SE
PO Box 47300
Olympia, WA 98504-7300

in cooperation with

U.S. Department of Transportation
Federal Highway Administration

December 2009

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
TNW2009-12		

4. TITLE AND SUBTITLE
DEVELOPMENT OF A STATEWIDE ONLINE SYSTEM FOR TRAFFIC DATA QUALITY CONTROL AND SHARING

5. REPORT DATE
December 2009

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
Yinhai Wang, Jonathan Corey, Yunteng Lao, Yao-Jan Wu

8. PERFORMING ORGANIZATION REPORT NO.
TNW2009-12

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Transportation Northwest Regional Center X (TransNow)
Box 352700, 123 More Hall
University of Washington
Seattle, WA 98195-2700

10. WORK UNIT NO.

11. CONTRACT OR GRANT NO.
DTRT07-G-0010

12. SPONSORING AGENCY NAME AND ADDRESS
United States Department of Transportation
Office of the Secretary of Transportation
1200 New Jersey Ave, SE
Washington, D.C. 20590

13. TYPE OF REPORT AND PERIOD COVERED
Final Research Report

14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES
This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. ABSTRACT

The Washington State Department of Transportation (WSDOT) operates thousands of Inductive Loop Detectors (ILDs) on the freeways and highways of Washington State. The collection and disbursement of this data is handled at the regional level, which has led to formatting differences and data fragmentation. The Datamart Project is intended to consolidate data from all regions and store it in one location easily searchable and accessible by all authorized users. Therefore a database and a web site have been created in a computer application called Datamart to accomplish these tasks and demonstrate the idea.

Additionally, this research project has been also tasked with creating software applications for traffic sensor data acquisition and developing error detection and correction methodologies for data quality control. Error detection is accomplished in a three-step test proposed by this study to identify loop detectors suffering from severe errors. The default conditions used by this test approach are general enough to be applied in other states. Meanwhile, the approach also has its flexibility to use site specific parameters to optimize its performance at a specific location. These severe errors identifiable by this approach include wrong mode setting, cross chatting, extreme under sensitivity, and incorrect sensitivity level, among others. Loop detectors suffering from incorrect sensitivity levels may be adjusted by the proposed correction methodology. The correction methodology is designed to identify the true length of the loop detector's detection zone. Properly calibrated ILDs are generally assumed to have detection zone sizes determined by the size of the loop coil buried in the roadway. The reality is that when ILD sensitivity levels drift from their correct settings, the detection zone size changes. By using the correct detection zone lengths, loop detector measurements can be corrected. This software-based approach is proven effective and can be retroactively applied to archived data as well.

The loop error identification and correction algorithms have been implemented in the prototype Datamart system, which is an online database system backed up by Microsoft SQL Server 2008 and the Google Map technologies. It is highly scalable and has the potential to add new data sources from other transportation agencies and online analysis functions for regional transportation planning, traffic management, and analysis purposes.

17. KEY WORDS
Dual-Loop Detectors, Loop Sensitivity, Speed, Data Correction, Error Detection

18. DISTRIBUTION STATEMENT
No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None	91	

DISCLAIMER

Table of Contents

Table of Figures

Table of Tables

Executive Summary

There are thousands of inductance loop detectors, including both single and dual loops, deployed in Washington State freeways. These loop detectors provide real-time measurements of certain traffic flow parameters and are important information sources for Advanced Traffic Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS). However, previous studies (Zhang et al., 2003 and Cheevarunothai et al., 2006 and 2007) found that these loop detectors are subject to various malfunctions that degrade loop detector data remarkably. Wrong sensitivity level settings are a fairly common problem that drags down the data quality for approximately 80% of the dual-loop detectors operated by the Washington State Department of Transportation (WSDOT) (Zhang et al., 2003).

Additionally, loop detector data are handled separately by regional offices in WSDOT. This system of data management is not only costly, but also inconvenient because such an isolated data storage structure and inconsistent data formats make it very challenging to retrieve statewide traffic data.

Therefore, this research intends to address the above data issues by developing an online system called Datamart for traffic data quality control and sharing. It targets improvements on traffic detector data quality, storage, and data sharing to make the WSDOT traffic sensor data across the state better quality, more accessible, and more consistent. Specifically, the objectives of this study are to investigate the causes of loop detector errors, to design an algorithm for identifying and correcting the loop sensitivity problems, to develop a computer application that implements and automates the proposed algorithm; and to establish an online database system for loop detector data management and statewide traffic data sharing.

To achieve these goals requires three distinct phases of data handling: 1) data acquisition from transportation agencies and data storage; 2) error detection and correction; and 3) online data sharing. Since data acquisition and web-based distribution

are relatively straightforward, this report focuses mostly on the second step, error detection and correction.

Traffic sensor data acquisition is performed by a computer program written in Microsoft Visual C#. This program executes on the Datamart server hosted at the Smart Transportation Applications and Research Laboratory (STAR Lab) of the University of Washington (UW). It monitors a WSDOT File Transfer Protocol (FTP) website that distributes the 20-second interval loop measurements in real time. Once a new set of data is uploaded to the FTP site, the program downloads the new data set, decodes it, and invokes proper database procedures of the Datamart system to store the date into the corresponding relations. Considering that the sensor data flow is continuous and the amount of data to process daily is huge, special attention is paid to saving storage space and enhancing database performance in the design of the traffic sensor database.

The error detection algorithms analyze loop data for three critical characteristics that determine the suitability of a loop detector for software-based error correction. First, the loop detector data must report values within a reasonable range for known vehicle types and speeds. Second, the vehicle lengths estimated by recorded loop on-times at known speed must show a complete vehicle distribution including both short and long vehicles. Third, the loop detector must be reporting data showing significant deviation from the expected values output by a correctly sensitive loop detector. When all three conditions are met the loop detector data may be corrected by the software-based approach proposed by this study.

The data correction methodology presented here seeks to determine the true length of an inductive loop detector's actual detection zone. Typically this length is assumed to be equal to the loop coil's length for a properly calibrated loop detector. The reality is that sensitivity level errors may increase or decrease the effective length of the detection zone. By determining the difference between the actual detection zone length and the expected detection zone length it is possible to generate corrected equations for single and dual loop detector data productions. These new equations generate lane occupancy, traffic speed, and vehicle length data that more accurately reflects the expected traffic conditions.

The Datamart prototype system is designed as an online data acquisition and analysis tool. It utilizes a group of the most advanced technologies for online application development, including Enterprise Edition Java (J2EE), the Google Web Toolkit (GWT), and Apache Tomcat. GWT and J2EE work in concert to make the inclusion of Google's maps and map based functionalities smoother. With the chosen implementation of GWT and J2EE, Apache Tomcat is the logical choice for hosting the website. The selected set of technologies is widely supported by different operating systems compared to competing technology sets. This definitely makes it easier to expand the Datamart system when necessary.

In addition to the software and methodology developments, the research team has the following principal findings from this research:

1) Even without high-resolution loop detector event data, several error types can still be detected through the proposed error identification method based on 20-second integrated data.

2) When a loop's sensitivity drops to a certain level, it starts to miss or split vehicles, particularly those with higher beds.

3) Malfunctions leading to wrong volume counts, such as crosstalk, pulse break up, extreme under sensitivity, etc. are very difficult, if not impossible, to correct using software-based approach. Such errors are better corrected at the hardware level.

4) Errors resulting from incorrect mode settings cannot be corrected by the proposed method because the distributions used by the correction algorithm to calculate the detection zone change are destroyed.

5) Loops with moderate sensitivity errors but reasonably accurate volume counts may be corrected by the proposed software-based method with a high degree of success.

6) Because a dual-loop detector can measure more information and have two single loops placed closely on the same lane, software-based methods may use the excess information and redundant measurements to correct more types of errors.

In summary, the proposed error identification and correction methods have been proven effective for loops with correctable sensitivity errors. If implemented at the controller level, these methods could improve single loop occupancy data, dual loop match rates, and speed and vehicle length measurements as well.

Future studies may focus on how under sensitive loops affect vehicle on-time measurements. It seems that they do not affect all kinds of vehicles equally and further research investigations are needed. Also, this study encountered a bifurcation problem of the on-times for short vehicles at some loop stations. The causes of this problem also deserve further study.

Chapter 1 Introduction

1.1 General Background

Of all of the currently available traffic detectors, the Inductive Loop Detector (ILD) is the most widely used (Klein et al., 2006). ILDs owe their popularity to their simplicity, reliability and durability. ILDs are simple in their construction and installation. An ILD is composed of a coil embedded in the roadway, wire to connect the coil with a roadside cabinet, and a controller card in the cabinet which communicates with the traffic controller. The simplicity of design eases maintenance and limits failures.

ILDs are very reliable as traffic detectors; they can function in rain, fog, and snow without appreciable affect (Klein et al., 2006). The ability to function well in adverse weather offers a clear advantage to the ILD when compared to camera based systems which suffer when visibility is limited. Similar limitations apply to many other non-intrusive alternative sensors. Also, the ILD is very durable; none of its components are exposed to wear. When properly installed, the loop coil is sealed inside the roadway and the electrical connections are protected inside pull boxes and cabinets, limiting weather and corrosion related problems to a large extent. By contrast, most other sensors, such as cameras, require lenses to be cleaned or other such ongoing maintenance (Klein et al., 2006).

Finally, ILDs tend to be relatively inexpensive in terms of hardware; though installation costs should not be neglected as a lane must be closed while crews cut the roadway to install the loop coil and lead wires. All of these factors explain why ILDs are popular with transportation agencies and also why they are so prevalent. ILDs have been in use since the early 1960's (Klein et al., 2006).

The Washington State Department of Transportation (WSDOT) maintains and operates Interstate and Washington State's freeways and highways. The WSDOT is divided into six regions: Northwest, North Central, Eastern, South Central, Southwest, and Olympic Regions. The Northwest Region, which is responsible for the Seattle metropolitan area, maintains approximately seven thousand unique ILDs. Of these 4200 are installed as single or dual ILDs on mainline freeways and highways. This large investment in sensing infrastructure allows WSDOT and researchers to monitor traffic flows around the metropolitan area.

Similar systems allow cities, such as the City of Bellevue and the City of Lynnwood, to monitor traffic within their jurisdictions. For example, the Smart Transportation Applications and Research Laboratory (STAR Lab) of the University of Washington (UW) created a traffic flow map for the City of Bellevue using the real-time traffic data collected from ILDs. The traffic flow map and analysis system is called the Google-map based Arterial Traffic Information (GATI) System (Wu et al., 2007). The GATI system and similar efforts by WSDOT are examples of Advanced Traveler Information Systems (ATIS). ATIS aims to give drivers more information about traffic conditions so that congested areas may be avoided, decreasing delay, decreasing pollution and improving traffic system performance.

In addition to the ATIS, Advanced Traffic Management Systems (ATMS) also rely on traffic sensors, mostly ILDs. ATMS intends to manage roadway traffic dynamically and optimally to reduce urban traffic congestion. For example, Active Traffic Management (ATM) is an emerging concept for congestion mitigation and safety enhancement. An ATM system monitors traffic flow in real time and detects any incidents in the area using traffic sensor data. Once traffic conditions start to deteriorate, the ATM system will take immediate actions, including speed harmonization, lane closure, shoulder driving, queue warning, etc., to improve roadway throughput and traffic safety. Such systems require accurate data at frequent intervals to function at their peak efficiency. To achieve such high levels of coverage and reliability, the inclusion of existing ILDs is virtually certain due to budgetary concerns. Therefore, the quality of ILD data is crucial for both ATIS and ATMS. As such, it is important that the ILDs produce the highest quality of data possible.

1.2 Inductive Loop Detector (ILD) Configurations

There are two installation configurations commonly used. The first is the single loop detector which is installed in the center of a lane. Single loop detectors can only register the presence or absence of a vehicle. They do this by recording the inductance drop of the coil of wire emplaced in the roadway. When inductance drops and then returns to normal the controller registers one vehicle. Controllers typically aggregate this data and report the total number of inductance drops and the percentage of time the loop coil spends below an inductance threshold. The more common terminology for these two measurements is volume and occupancy.

The second configuration is called a dual-loop or speed trap. A dual loop detector employs two distinct ILDs placed in a lane with a specified distance between the leading edge of the first loop coil and the leading edge of the second loop coil. Figure 1-1 shows a conceptual dual loop detector. The WSDOT standard is for the leading edge of the upstream ILD, referred to as the “M loop” by the WSDOT convention, to be seventeen feet from the leading edge of the downstream ILD, called the “S loop” by the WSDOT convention.

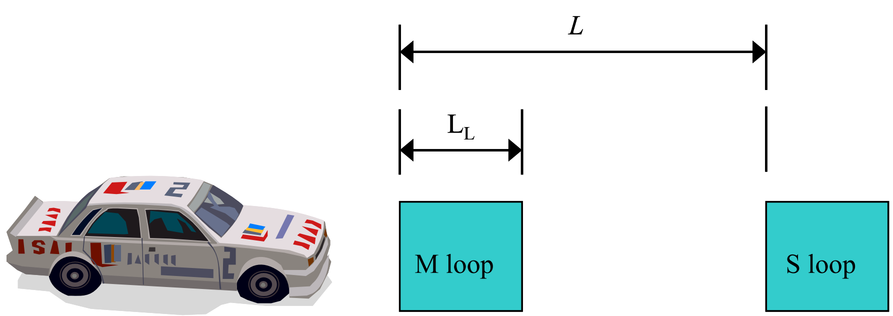


Figure 1-1 Dual Loop Detector Configuration

By measuring the difference in arrival time between the M and S loops, the controller can directly calculate the vehicle’s speed. With the vehicle speed and the on-time measurement by either the M or the S loop, vehicle lengths can be calculated as well.

1.3 ILD Error Types

ILDs can suffer from a number of errors. Some errors are related to hardware failure and others are caused by sensitivity errors. Many hardware-related errors affect the ability of the ILD to maintain a stable inductance level. Pulse break up, where a single vehicle’s inductance drop is registered as multiple inductance drops with gaps in between, is an example of such a hardware failure (Chen and May, 1987). Noisy signals from the ILD coil can cause similar results. Another hardware error, crosstalk, can cause the ILD to detect traffic in adjacent lanes or in ILDs where an electrical connection has been established, possibly by a short circuit (Chen and May, 1987). Crosstalk can also occur when adjacent ILD loop coils activate each other directly (Bhagat and Woods, 1997). Another hardware error to be concerned with is pulse mode. For freeway operations the desired mode is presence mode where an ILD will register as occupied only while a vehicle is present over the loop coil. ILDs set in pulse mode will generate a fixed length pulse regardless of the time a vehicle takes to traverse the loop coil.

The other major category of ILD errors relates to the sensitivity levels of the ILDs. ILDs operate by passing a current oscillating at a rate exceeding 10 kHz through the ILD loop coil (Klein et al., 2006). When the metal of a vehicle passes through the loop coil's electromagnetic field, it induces eddy currents in the loop coil which reduces the overall inductance. ILDs have a number of discrete sensitivity levels to control the level of inductance drop that will register as the presence of a vehicle. Incorrect sensitivity levels can affect the duration of vehicle detection. Extremely low sensitivity levels can distort vehicle detections and even completely miss vehicles. High sensitivity levels can cause detection of vehicles in adjacent lanes and may begin to detect random fluctuations.

Many loop detector controller cards such as the Reno C-1000 and EDI Model 222 use sensitivity levels that are based on powers of two. For these controller cards, sensitivity level 7 will detect a 0.01% inductance change in the attached loop. Lower sensitivity levels require larger changes in inductance to register. Level 4, for example, requires an inductance change of 0.08% to register the presence of a vehicle, while sensitivity level 1 requires a 0.64% change in inductance. (EDI, 2005; Reno A&E, 2004) This doubling of sensitivity levels can easily lead to situations where the ideal sensitivity level is simply unobtainable. If, for example, the ideal inductance change sensitivity would be such that the controller card registers a vehicle's presence at an inductance change of 0.06% this would require that the sensitivity level be set between levels 4 (0.08%) and 5 (0.04%), which is not possible.

With ILDs offering such longevity and durability, their working environments may change significantly over time and hence affect their sensitivity for vehicle detection. Another sensitivity related error can occur after an overlay. Overlays increase the depth at which the loop coil is buried in the roadway and can reduce the sensitivity correspondingly. It should be noted that until relatively recently there were no tools available to check an ILD's sensitivity on site. The Advanced Loop Event Data Analyzer (ALEDA) developed by the STAR Lab offers a great solution to detect and correct dual loop sensitivity errors at a relatively low cost (Cheevarunothai et al., 2005). Sensitivity problems in general will affect an ILD's detection capability by increasing or decreasing the actual distance from the loop coil at which vehicles are detected. See Figure 1-2 for an illustration.

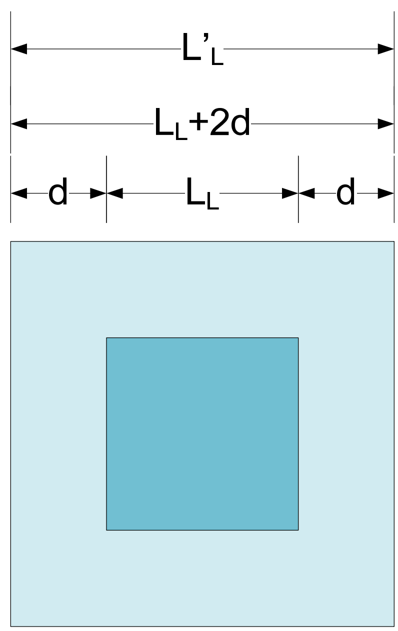


Figure 1-2 Single Loop Actual Detection Zone Change due to Sensitivity Error

Because each dual loop detector is composed of two distinct ILD's, each ILD can fail as described above. Dual loop detectors add a number of additional considerations. A dual loop detector is based on the assumption that each ILD will detect vehicles in the same manner. This assumption affects the distance L between the M and S loops, and the ILD coil detection zone length L_L. If the ILDs have different sensitivities, the distance traveled between the arrivals at the M and S loops is no longer L. Similarly, sensitivity errors affect L_L by changing the distance at which vehicles are detected. Figure 1-3 shows an example scenario that illustrates the problem. Changing the two variables, particularly in an unpredictable manner, induces errors in the length and speed calculations. To control these errors, the WSDOT dual loop algorithm throws out results from vehicle length calculations if the on-time difference between the M and S loop measurements is more than 10 percent (Cheevarunothai et al., 2005).

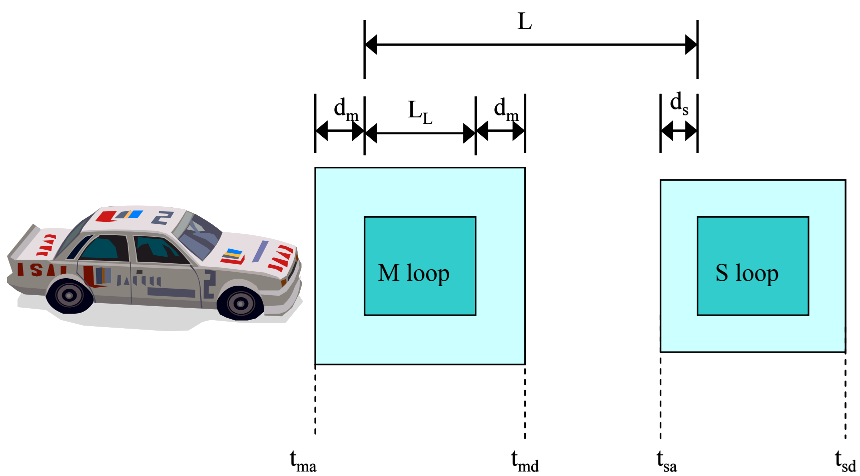


Figure 1-3: Dual Loop Detector with an Inconsistent Sensitivity Problem

1.4 Research Objectives

There are thousands of inductance loop detectors, including both single and dual loops, deployed in Washington State freeways, over 7000 in WSDOT's Northwest Region alone. These loop detectors provide real-time measurements of certain traffic flow parameters and are important information sources for Advanced Traffic Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS). However, previous studies (Zhang et al., 2003 and Cheevarunothai et al., 2006 and 2007) found that these loop detectors are subject to various malfunctions that degrade loop detector data remarkably. Wrong sensitivity level settings are a fairly common problem that drags down the data quality for approximately 80% of the dual-loop detectors operated by the WSDOT (Zhang et al., 2003).

In addition to the data quality problem, these traffic sensor data are handled separately by regional offices in WSDOT. This way of data management is not only costly, but also inconvenient because such an isolated data storage structure and inconsistent data formats make it very challenging to retrieve statewide traffic data.

Therefore, this research project targets improvements to traffic detector data, data collection, and data management that will make WSDOT traffic sensor data across the state better and more consistent. Specifically, the objectives of this study are:

- To investigate the causes of detector errors;
- To design an algorithm for identifying and correcting the loop sensitivity problems;
- To develop a computer application that implements and automates the proposed algorithm; and
- To establish an online database system for loop detector data management and statewide traffic data sharing.

Chapter 2 Literature Review

Inductive Loop Detectors (ILDs) have formed the backbone of many traffic detector networks for decades (Klein et al., 2006, Cheevarunothai, 2009). As such, their proper configuration and operation is important for traffic data quality. Unfortunately, ILDs are prone to a variety of errors (Payne and Thompson, 1997) that can dramatically affect their accuracy. The ILD data quality is increasing in importance as both ATIS and ATMS require accurate traffic sensor data to operate properly. Therefore minimizing ILD data errors by correcting inaccurate measurements when possible has become an important issue for modern traffic operations.

2.1 ILD Error Detection

ILD error detection research has spanned decades. Much of the early work focused on hardware diagnostics. Some diagnostics used specialized equipment to test ILDs and determine the quality of their wiring (Ingram 1976; Klein et al., 2006). The majority of more recent research has focused on software-based identification of errors. These techniques are broadly divided by data source, aggregated data or high-resolution event data (Cheevarunothai et al., 2009).

The methods using time series 20- or 30-second aggregated data include studies developed by Chen and May (1987), Nihan et al. (1990), Jacobson et al. (1990), Cleghorn et al. (1991), Nihan (1997), Turner (2000), Turochy and Smith (2000), and Wall et al. (2003). In these methods, multiple thresholds from empirical studies for loop malfunction detection were compared with the fundamental traffic variables extracted from aggregated loop data. For example, aggregated 30-second loop measurements with an occupancy value higher than a chosen threshold, typically values between 90 and 99 percent (Turochy and Smith, 2000) would be considered erroneous. The major drawback of using aggregated data is that a large sample variance must be accepted and certain types of loop detector errors are potentially missed because aggregated data do not contain individual vehicle actuations that are important for error identification (Coifman et al, 2004; Cheevarunothai et al, 2009). Furthermore, different types of errors may cancel each other making them very difficult to be detected in aggregated data.

As computer technology advanced, event data level analysis became possible. Chen and May (1987) placed a computer inside a traffic control cabinet with the traffic controller and ILD controller cards (a.k.a. electronic units). They were able to record the ILD actuations actually seen by the controller for data processing and aggregation. Such detector actuation data provided details on when a vehicle arrives at and departs from the ILD, how long it stays visible, and the accurate time between inductance drops. More sophisticated diagnostics have been enabled by these high-resolution event data analyses allowing the identification of which component of the system was causing the problem. This research (Chen and May, 1987) laid the groundwork for identification of pulse break up and some other errors. The researchers were able to identify the errors by examining the actuations of the ILDs.

Recent research has been attacking the problem of ILD error detection from both ends with some researchers using event data to identify errors directly and other researchers trying to isolate errors from aggregated data. Event data based error detection has been expanding upon the work done by Chen and May (1987). May et al. (2004) focused their error identification efforts on a heavily instrumented section of I-80 called the Berkeley Highway Laboratory (BHL). Event data can be collected from any of the detectors in the BHL through controller reconfigurations. A PATH research report by May et al. (2004) summarizes many of their findings with regards to event data based ILD error detection at the BHL. Their tests include several occupancy, speed, length and volume based logical tests with alternative methods of calculation used to double check results. Coifman and Dhoorjaty (2004) also used BHL data to investigate the relationship between headway and on-time and the difference between individual vehicle speeds and average vehicle speeds.

A series of event data based investigations (Zhang et al., 2003; Cheevarunothai et al., 2005) led to the creation of the Advance Loop Event Data Analyzer (ALEDA) system. The ALEDA system is a portable system that connects a laptop computer directly to the input file of a traffic control cabinet. It directly reads loop controller card actuations before they reach the cabinet's controller and hence does not affect controller's regular operations during the data collection process. The ALEDA system has been improved over time and proven effective in event data collection, loop error identification, and sensitivity problem tune up (Cheevarunothai et al., 2006; Cheevarunothai et al., 2007; Cheevarunothai et al., 2009). ALEDA is also capable of recording and replaying event data as well as simulating freeway loop actuations. A more detailed examination of the ALEDA system may be found in Nihan et al. (2006).

Aggregated ILD data have been widely used in ATIS and ATMS. Dailey (1999) and Wang and Nihan (2000) extracted speed data from 20-second aggregated data. By filtering out long vehicles for speed estimations, Wang and Nihan (2003) were able to apply a uniform vehicle length assumption to the speed equations developed by Athol (1965) with reasonable accuracy. With improved quality speed data, truck volumes could be estimated with acceptable certainty. However, if a loop's sensitivity is off its correct level, nothing extracted from its direct measurements can be reliable. Many studies have been conducted on aggregated ILD data error checking. For example, Nihan et al. (2002) directly analyzed dual loop detector errors in vehicle classification. By using a surveillance camera that overlooked the dual loop detector site, they were able to classify vehicles by length and compare the results to those reported by the detector. The results showed a general undercounting of vehicles and some incorrect binning by length. More relevant studies on ILD error identification were summarized in Nihan et al. (2002).

2.2 ILD Error Correction

The ultimate goal of error correction is to repair flawed data so that data used for traffic analysis are more reliable. Different methods exist to capture flawed data and discard them from important analysis. The simplest method is to throw out data that are identified as flawed. Data may be identified and discarded as being flawed when they exceed user-defined and flow theory based thresholds (Turochy and Smith, 2000). Typical aggregated data tests include maximum volume and occupancy levels. Event data tests can include checks for signal length, signal continuity and alternate calculations for speed and length (May et al., 2004). Additional techniques include using the two ILDs of a dual loop detector to check each other and generate corrected values based on the double observations (Coifman, 1999; Cheevarunothai, 2006). Single loop detectors have also begun receiving more correction attention with sensitivity adjustments

(Cheevarunothai, 2009). More advanced techniques can attempt to replace missing and repair flawed data through the use of upstream and downstream sensor data (Wall and Dailey, 2003).

Chapter 3 Study Data

This study uses two primary loop detector data sources: high-resolution event data and 20-second integrated data. Event data was collected at selected cabinets using the ALEDA system developed by the STAR Lab (Cheevarunothai et al., 2005). The event data record the individual loop detector status, occupied or not, at 100 Hz. Event data was collected to facilitate the development and calibration of the correction algorithms to be applied to the 20-second data. The aggregation of loop detector data into 20-second intervals is standard at WSDOT Traffic System Management Centers (TSMCs). Data files are downloaded from WSDOT's Northwest Region via a File Transfer Protocol (FTP) service website from WSDOT. Details of two datasets will be explained in the following sections.

3.1 20-second Data Collection

WSDOT's onsite traffic controllers collect single and dual loop detector event data and extract volume and occupancy measurements for each 20-second interval. For a dual-loop detector, the controller will also calculate mean vehicle lengths and speed for the same 20-second interval. Aggregated 20-second measurements are sent to the TSMC periodically for further processing and archiving. These 20-second data are posted to a FTP website for third party applications. Then they are aggregated into five-minute interval data for long term storage and analysis at TSMC. Five minute intervals are the standard long term storage interval used by WSDOT for planning and analysis.

The STAR Lab downloads the 20-second data in real-time and archive them in a Microsoft SQL Server 2008 database for research and education. A computer program was developed by the research team to automatically download, unpack, and store the 20-second data periodically. This program was written in Microsoft Visual C#. Data from all of the ILDs in the central Puget Sound region have been received and archived. As of the writing of this report, the dual loop detector data are not included in the 20-second data package on the FTP website, but WSDOT is working on it. So dual-loop data should be included in the FTP service soon.

Single loop detector data include nine attributes: *detector name*, *timestamp*, *volume*, *occupancy*, *speed*, *calculated speed*, *lane count*, *color*, and *flag*. *Speed* is the single loop speed estimated by using volume and occupancy measurements. *Calculated speed* is the same as *speed* except that it is bounded. When *speed* exceeds 60 mph, the *calculated speed* will be 60 mph. *Lane count* is "1" for individual loop detectors, but for the station data described below it will equal the number of lanes at the station. *Color* is used by the WSDOT traffic flow map to indicate traffic congestion level and is determined by the occupancy. *Flag* indicates whether WSDOT's algorithms have identified the detector as having a significant error. Each cabinet also reports the data observed for that location as a whole.

20-second aggregated data are used in the analyses to follow. To avoid segmentation error (Yu et al., 2009) and ensure that only one vehicle's on-time is being considered at a time the aggregated data must be screened to isolate individual vehicles in an interval. The criteria for selection of a 20-second aggregated interval were as follows: the selected (i^{th}) interval must have a volume exactly equal to one; the previous (i-1) interval must have both zero volume and zero occupancy; and the following (i+1) interval must also have zero occupancy and zero volume. Effectively this means that only one vehicle can pass over the detector per minute. Because of the extremely low flow rate this implies all complying intervals are assumed to be under free flow conditions. For the locations chosen for event data collection there were at least 400 such intervals in a one-month period of August 2009.

3.2 Event Data

Since event data collection requires site visits and cabinet wiring, it is not possible to collect event data from all Inductive Loop Detector (ILD) stations. Using detailed individual vehicle data, such as vehicle's arrival and departure times, an ILD's malfunction can be diagnosed. Therefore, in this study, event data collection targeted cabinets displaying signs of data error. To identify erroneous ILDs, historical 20-second loop measurements were sampled from the STAR Lab database. Dual loop detectors were the primary focus for event data collection in this study because of the following four reasons. First, dual loop detectors have two ILDs which doubles the number of available ILDs to sample. Second, dual loops can provide extra data valuable for error identifications. Third, dual loop correction is also a function of this project. Finally, one

ILD of the detector pair can be compared to the other ILD as well as to the dual loop generated speed data for verification purposes.

The WSDOT's cabinet naming system follows the following convention. The first three characters represent route number, followed by a station locator formed by "es" and the station's mile post. For example, a station may have its name as 005es15652. The first three characters, 005, indicate that the cabinet is located on Interstate 5. The final five numbers, 15652, indicate the milepost, 156.52, at which the cabinet is located.

Considering the time and budget constraints, only a couple of dual-loop stations could be selected for event data collection. The selection was based on average volumes and occupancies. The three dual loop stations selected for this study are 005es15652, 005es15996, and 005es16302. All these selected cabinets are located on I-5 south of Seattle. Both directions of I-5, north and south, have four general purpose lanes plus one High Occupancy Vehicle (HOV) lane on the inside for stations 005es15652 and 005es15996. Station 005es16302 has three general purpose lanes and a HOV lane in each direction. Both north and south bound I-5 had a directional Annual Average Daily Traffic (AADT) of approximately 120,000 in 2005 (WSDOT, 2008) for the study segment. The locations of the three detectors may be found in Figure 3-1 with the inset pictures showing close ups of each cabinet location and the approximate locations of the dual loop detector coils.

Cabinet 005es15652 was selected because one of two single ILDs that form the dual loop detector on lane 5 of southbound I-5 was reporting average volumes over 20 vehicles per interval. 20 vehicles in twenty seconds would equate to 3600 vehicles per lane per hour which far exceeds the levels expected by the Highway Capacity Manual (2000). Thus, it is suspicious to frequently report such high interval volumes. Additionally, of the five southbound dual loops at this station, two have failed ILDs and are not functioning; three dual loops are producing data, including the one on lane 5. The dual loop on lane 4 appears to be functioning well with regards to the on-times of each ILD but both appeared to be under sensitive. There are two important geometric details for this location, the loop detectors are at the crest of a small rise that blocks line of sight and the freeway is curving at this location. Both factors would be expected to reduce drivers' speeds at this location. Data collection for this cabinet began on September 24th and continued until September 29th, 2009. The lane 4 southbound dual loop detector at this cabinet is Study Site 1.

Cabinet 005es15996 was selected by another criterion. A major step toward error correction was to identify and understand ILDs with various errors such as crosstalk and pulse mode. Few ILDs could be identified from the available aggregated data as potentially suffering from crosstalk, but several ILDs were identified as potentially being set in pulse mode. The criterion to identify pulse mode looked at the range of on-times produced by the ILD. ILDs with narrow on-time ranges and near constant on-times were selected as candidates for pulse mode operations. 005es15996 had two such loops and also appeared from 20-second single loop data to have reasonably functioning dual loop detectors in adjacent lanes. This location's geometry is less interesting with only a small grade and straight line travel. Data collection for this cabinet began on September 29th and continued until October 5th, 2009. The northbound lane 4 dual loop detector at this cabinet has been selected as Study Site 2.

Cabinet 005es16302 was selected for data collection of over sensitive ILD data. An ILD of the lane 1 dual loop was showing signs of oversensitivity with occupancy values exceeding expectations for the reported volumes. Other dual loops at this station do not have malfunctions easily detected from 20-second aggregate data. As Figure 3-1 shows this location (upper right inset) is more geometrically complex. There is an on ramp after the dual loop stations (indicated by the line) and an exit ramp before them. The whole segment has a small uphill grade and is relatively straight. Due to time constraints the event data collection at this location only spanned one day from the afternoon of October 5th until the afternoon of October 6th, 2009.

Figure 3-1 Study Sites

Event data at the three locations will be used for the analyses regarding error detection and correction in the following chapters. The 4th lane from the outside (HOV) of station 005es16302 is missing from these sections because it did not record a large enough sample of vehicles during free flow conditions for comparison to other locations. For stations 005es15652 and 005es15996 event data from midnight to four in the morning were used. Station 005es16302 was analyzed from midnight to five in the morning to increase the sample size since only one day's worth of data was collected. From historical data traffic flow does not begin to drop out of free flow conditions on this segment until after 5:30 am.

Chapter 4 Error Detection

This chapter will be presented in the following manner. The chapter will begin with an introductory section on error detection which will be followed by a section discussing the observed on-time distributions. A Gaussian Mixture Model will be developed in the Section 4.3. Three error types will be examined in the Sections 4.4 through 4.6. A brief examination of 20-second data versus event data will follow the error discussions. The chapter will close with a possible explanation for one of the error types.

4.1 Introduction

There are many factors which can influence ILD detection accuracy. These may include sensitivity level setting, electrical connection quality, signal stability and many more. Detecting all kinds of errors is not always feasible because different errors may cancel each other or cause similar results in the data. For example, pulse break up and noise can both produce high volume and low occupancy readings. However, there are still some ILD errors that can be identified and corrected to improve ILD data. Zhang et al. (2003) studied the WSDOT dual-loop detectors and concluded that approximately 80% of them have remarkable sensitivity problems. Simply identifying and correcting sensitivity caused problems will largely enhance the quality of loop detector data. Therefore, this study focuses on ILD sensitivity problems. For this purpose, a series of tests have been chosen to determine the fitness of a single loop detector for correction. For dual loop detectors each of their ILDs must pass these tests before the dual loop detector can be checked.

The data required for these tests may be either formed by aggregated interval measurements at free flow speeds with only one vehicle recorded or high-resolution ILD event data. For intervals with only one vehicle recorded, the lane occupancy can be converted directly to a single vehicle's on-time by simply multiplying it by the aggregation interval length (i.e. 20 seconds for WSDOT data). This method of using aggregated data can be tedious due to the number of records required to generate a reasonable distribution, but has the benefit of not requiring the direct efforts that event data collection can. Researchers using single vehicle aggregated data should be wary of segmentation error. Segmentation error occurs when a vehicle is in the process of traversing the ILD loop coil at the moment the aggregation intervals change. Segmentation error can lower the average occupancy values observed for the ILD (Yu et al., 2009).

Since an ILD's sensitivity level ties directly to the distance at which the detectors can recognize the presence of a vehicle, an incorrect sensitivity setting can result in data errors. In this study, we assume that an ILD's detection zone is symmetrical about the loop coil center and uniform across vehicle types. Ideally, the detection zone should overlap exactly with the loop coil when sensitivity is properly set. When the ILD sensitivity is off its correct level, a non-zero distance offset *d*, exists as shown in Figure 1-2. If the ILD is over-sensitive, *d* is larger than zero. It takes a negative value when the ILD is under-sensitive.

This change in ILD detection zone affects all of the calculations for single and dual loop detectors. Figure 1-3 illustrates the detection zone problem for dual loop detectors. For dual loop detectors, vehicle speed can be accurately measured when both single ILDs are consistent with the sensitivity, i.e. $d_m = d_s$, where d_m and d_s represent the distance offset for the M and S loops, respectively. Since the actual distance traveled by a vehicle between the arrival time recorded by the M loop (t_{ma}) and the arrival time to the S loop (t_{sa}) is still *L*, the speed measurement is accurate even though both may be over-sensitive or under-sensitive. However, to obtain accurate vehicle length measurement, the two single ILDs must be adjusted to the correct sensitivity level. A hardware-based approach for dual-loop sensitivity tune-up is available in Cheevarunothai (2006). This study focuses only on software-based methods for both single and dual loop sensitivity problem identification and correction.

4.2 On-time Distribution

As mentioned earlier, an ILD's sensitivity ties directly to each vehicle's on-time measurements. Vehicles' on-time data serve as an important information source for the diagnostics. Therefore, the quality and nature of single loop detector on-time distributions must be examined first. These examinations may be carried out using aggregated data, such as 20-second aggregated data, or event data. Event data typically require special effort to collect, but support more thorough error investigations. Aggregated data have the benefit of being more easily accessible but lack individual vehicle details in general. However, when traffic volume is extremely light, a 20-second interval may observe just one vehicle. This enables us to extract individual vehicle information from the aggregated data, although aggregated data limit the resolution of on-time data for analysis. For example, 20-second aggregated data with occupancy reported in tenths of a percent would have an on-time resolution of 20 ms (50 Hz). Event data on the other hand is limited by the hardware used. In this study, event data were collected using ALEDA at 10 ms resolution (100 Hz). A comparison of results obtained through the use of event and 20-second data will be shown later in the chapter.

4.2.1 On-time calculation and research

If an ILD's detection zone contains a non-zero distance offset *d* as shown in Figure 1-2, then the effective vehicle length (L_E) changes to Equation 4-1.

$$L_E = L_V + L_L + 2d = v * (t_d - t_a) = v * OT \quad (4\text{-}1)$$

L_E is the sum of the actual vehicle length (L_V) and the length of the ILD's detection zone, which equals the loop (L_L) and twice the distance offset, d, as shown in Figure 4-1. The effective length is a function of vehicle speed and the on-time (OT).

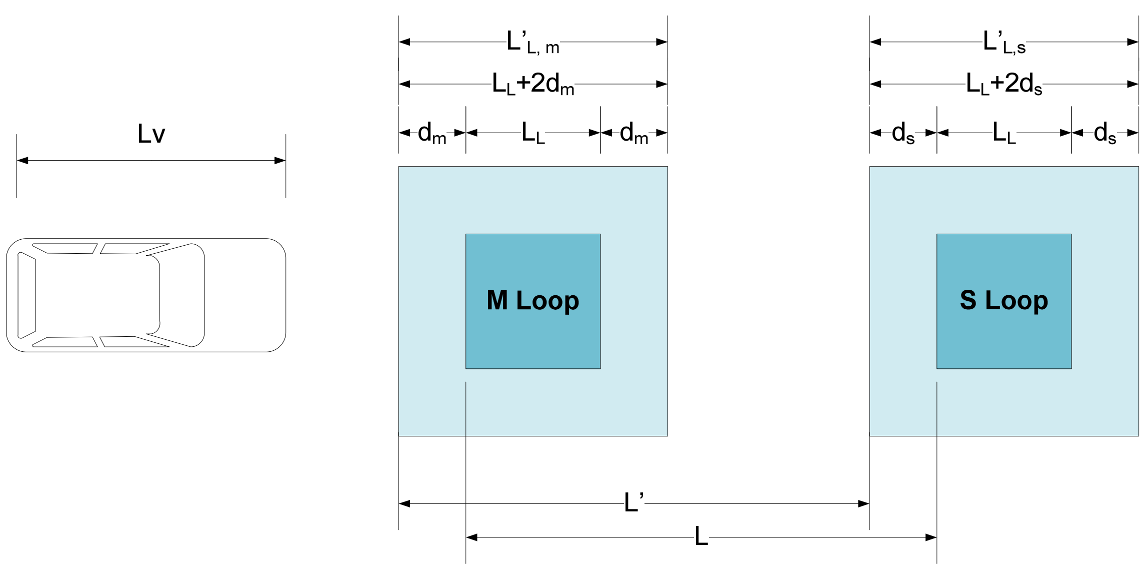


Figure 4-1 Equivalent Length Diagram

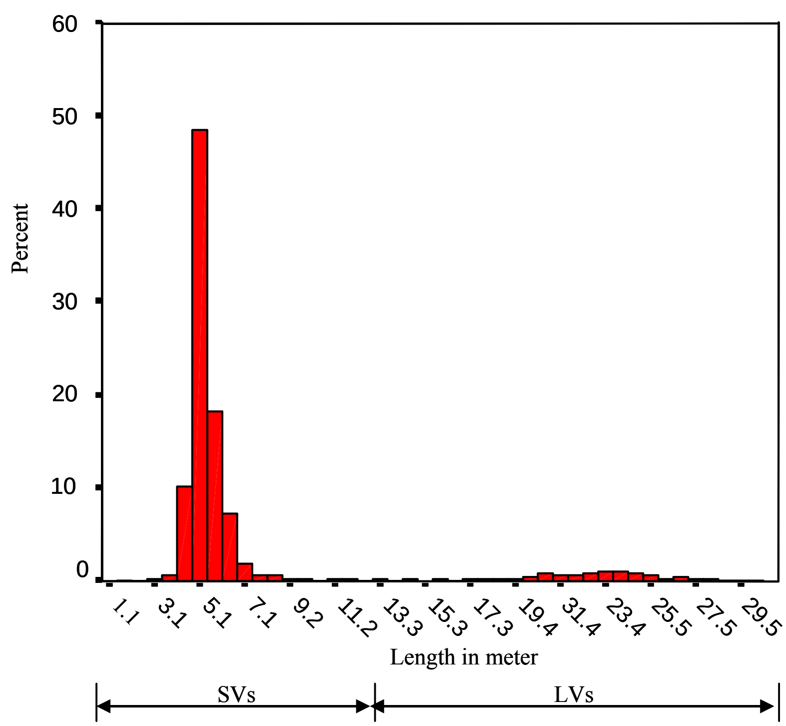
Rearranging Equation 4-1, we can get the on-time for a vehicle as:

$$OT = \frac{L_V + L_L + 2d}{v} \tag{4-2}$$

For a correctly functioning loop detector, d is negligible and L_L is a constant. Thus, with a known speed, such as the free flow speed (v_f), the on time is directly related to the distribution of vehicle lengths (L_V).

$$OT_f = \frac{L_V + L_L}{v_f} \tag{4-3}$$

An earlier study by Wang and Nihan (2000) found the mean short vehicle[1] length to be 15.18 feet and the standard deviation to be 1.31 feet. Other research has also found the vehicle length distribution to strongly favor a narrow short vehicle distribution such as findings by May et al. (2003). Figures 4-2 and 4-3, below, illustrate what would be expected of correctly functioning loop detectors from two previous studies. As shown in Figure 4-4, our research also indicates similar results with Wang and Nihan (2000, 2003) and May et al. (2004).



[1] In this study, a vehicle shorter than 26 feet was considered a short vehicle.

Figure 4-2 Length Distribution of Vehicles (L_V) on Southbound I-5 (Wang and Nihan, 2003)

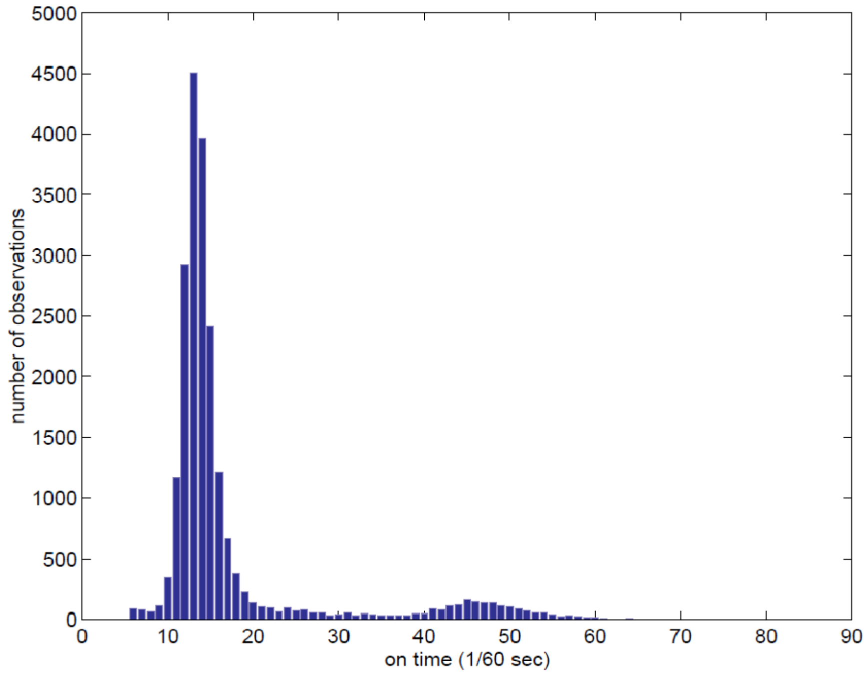
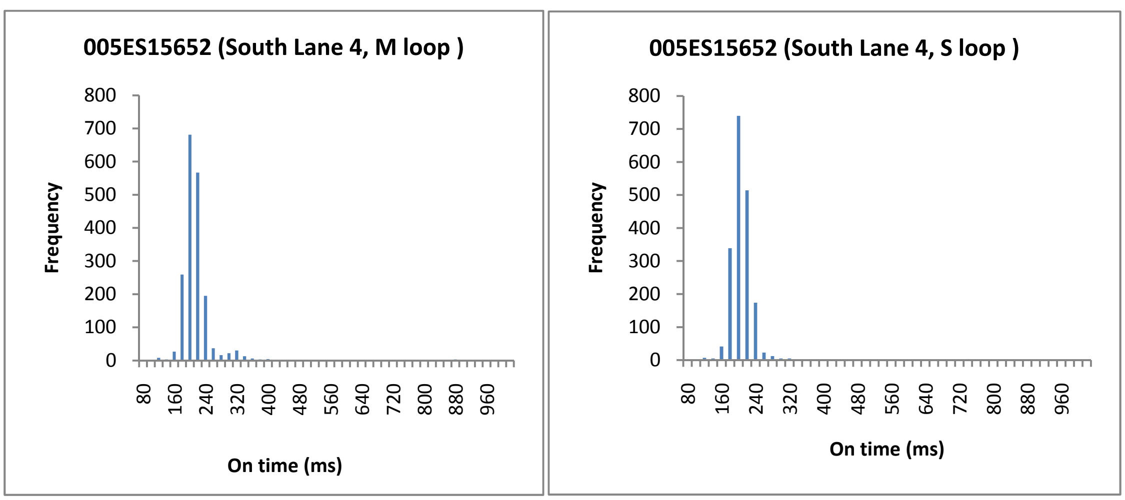
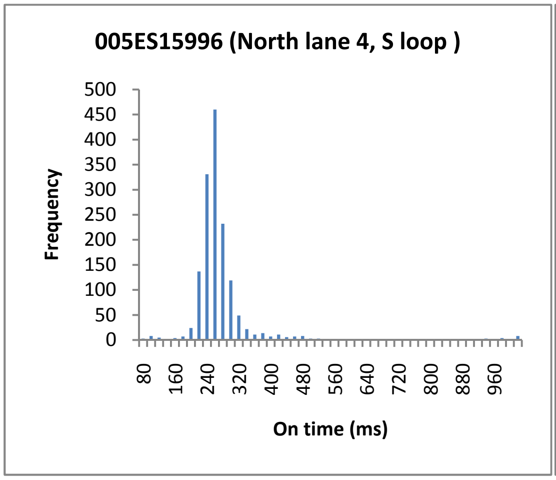


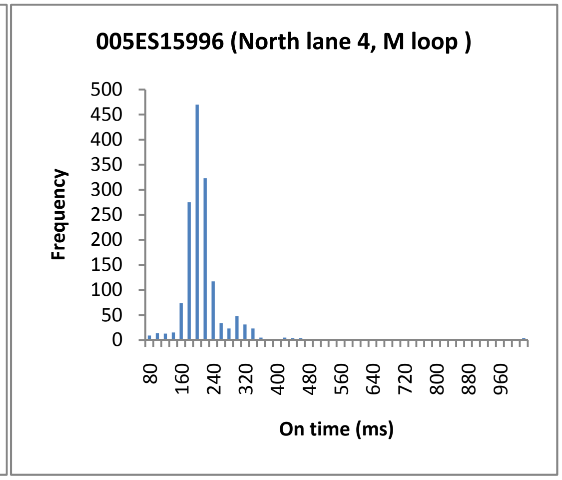
Figure4-3 Free Flow Dual Loop Detector On-Time (OT) Distribution (May et al. 2003)



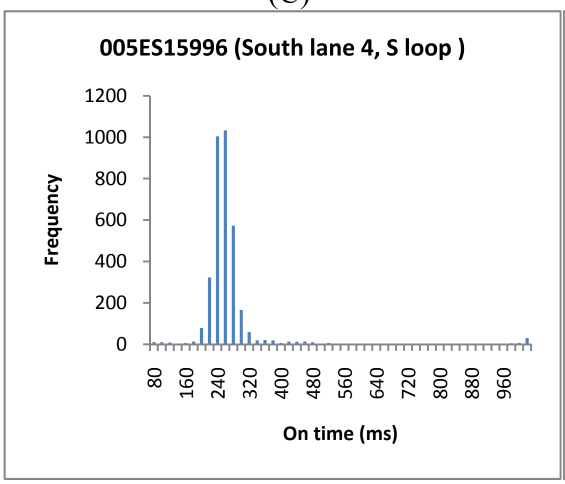
(A) (B)



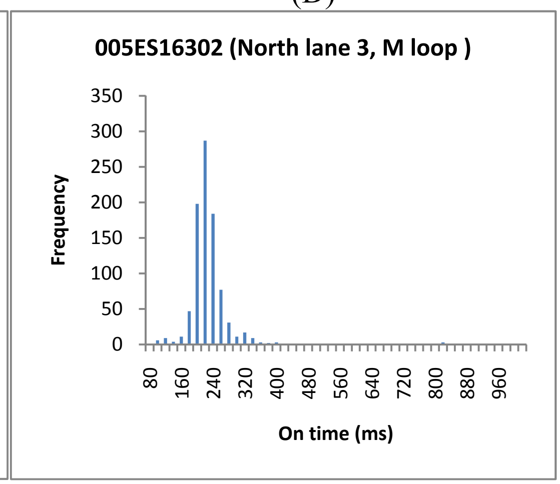
(C)



(D)



(E)



(F)

Figure 4-4 On-time Distributions for Six Loop Detectors

4.2.2 Modeling On-time distribution

Traffic is composed of many vehicle types. These vehicles may be short vehicles or any of a variety of sizes of heavy trucks. Looking at Figures 4-2 and 4-3, the on-time distributions are expected to be bi-modal. The first peak which is very high and narrow represents the short vehicle length distribution. The second peak is broader and shallower, indicating the large variation of long vehicle lengths. With these expectations in mind, a Gaussian Mixture Model (GMM) (McLachlan and Basford, 1988) was applied to the on-time distribution data in order to weed out unsuitable ILDs. The distribution traits can be modeled as a mixture of K Gaussian distributions, which would make the probability of measuring the i^{th} on-time as

$$P(L_{Vi}) = \sum_{j=1}^{K} \omega_j f(L_{Vi}, \mu_j, \sigma_j^2) \qquad (4\text{-}4)$$

where L_{Vi} is the length of the i[th] vehicle observation; K is the number of vehicle categories based on vehicle length, K=1 to 4 depending on the number of categories observable at the site; ω_j is the weighting factor of the j^{th} Gaussian distribution $f(L_{Vi}, \mu_j, \sigma_j^2)$ with a mean of μ_j and a variance of σ_j^2. The Gaussian distribution of length used here is only one-dimensional and is defined as:

$$f(L_{Vi}, \mu_j, \sigma_j^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(L_{Vi}-\mu_j)^2/2\sigma^2}, \qquad L_{Vi} > 0 \qquad (4\text{-}5)$$

The first constraint on the correction method is that the distance offset, *d*, should not be less than $-L_L/2$ so that $L_L + 2d > 0$. Thus, *d* should be bounded by

$$d > -L_L/2 \qquad (4\text{-}6)$$

If we define the first vehicle category as the short vehicle category, the average on-time for Category 1 is

$$\mu_1 = \overline{OT} = \frac{\overline{L_{V1}} + L_L + 2d}{v_f} \qquad (4\text{-}7)$$

with which we can rewrite the first constraint condition for use with the GMM as

$$\mu_1 = \overline{OT} = \frac{\overline{L_{V1}} + L_L + 2d}{v_f} \geq \frac{\overline{L_{V1}}}{v_f} = \alpha \qquad (4\text{-}8)$$

where α is the lower boundary value for the average on-time of short vehicles; $\overline{L_{V1}}$ is the average vehicle length for vehicle category 1. $\overline{L_{V1}}$ may change by location and lane. The average length for short vehicles used in this project is 15.2ft from the work conducted by Wang and Nihan (2000). The free flow speed is assumed to be smaller than 70 mph at free flow for freeways in the central Puget Sound region where the speed limit is typically 60 mph or lower. Combining the free flow speed and short vehicle lengths yields an α of

$$\mu_1 \geq \frac{\overline{L_V}}{v_f} \geq \frac{15.2\, ft}{70\, mph\, (102.67\, fps)} = .148\,\text{sec} = 148\, ms \qquad (4\text{-}9)\ (C1)$$

Another constraint condition is that the proportion of short vehicles should be much higher than any other type of vehicle on the freeways in the central Puget Sound region. This is a reasonable assumption considering the relatively low volume levels of various kinds of trucks compared to cars. With the expected vehicle length distributions strongly favoring short vehicles, a second correction suitability constraint condition is enforced through the Gaussian Mixture Model.

$$\omega_1 > \beta \qquad (4\text{-}10)\ (C2)$$

The weight factor for vehicle category 1, ω_1, must be greater than β, a lower boundary value, in the GMM. The value of β will depend on local conditions and will be most strongly influenced by the percentage of trucks in the traffic flow. For locations where trucks compose large portions of the traffic flow a lower threshold should be used. For ideal loop detectors the threshold would be equal to the percentage of small vehicle in the traffic flow. Since loop detectors are not ideal and random errors and sensitivity errors occur the threshold must be lower than the expected value to account for real life factors such as lane changing and speed variation over the ILD. For our study sites, β = 80% was used.

The third constraint corresponds to whether sensitivity correction is necessary for an ILD. If the calculated value of *d* is small enough, then the sensitivity of the ILD is at the correct level. Otherwise, sensitivity adjustment is needed. Equation 4-11 is used to judge if the sensitivity adjustment is needed.

$$|d| < \gamma \qquad (4\text{-}11)$$

Here is γ the upper boundary of allowable ILD sensitivity error in the loop detector system. By substituting Equation 4-11 into Equation 4-7, the third constraint condition can be rewritten in terms of the GMM.

$$\frac{\overline{L_V}+L_L-2\gamma}{v_f} < \mu_1 < \frac{\overline{L_V}+L_L+2\gamma}{v_f} \qquad (4\text{-}12)\ (C3)$$

For a plus and minus ten percent on-time range γ should equal 1.06 feet, given the loop length is 6 feet and the average vehicle length is 15.2 feet. The selection of free flow speed v_f is crucial. A free flow speed of 60 mph would produce a $\pm 10\%$ on-time range of 217 ms to 265 ms and for 65 mph the range moves to 200 ms to 245 ms. The correct choice of free speed at a location will be the difference between correctly identifying a loop as having a sensitivity problem or not. Where possible, good dual loop detector data should be used to establish a free flow speed. If acceptable dual loop detector data is not available other speed collection methods may be used. When other data is not available the Highway Capacity Manual (2000) free flow speed estimation methods may be used. Methods such as the HCM should be used with care because they are based on average conditions which may not be applicable to the specific location being examined.

4.3 Error Detection via Gaussian Mixture Model (GMM)

The Gaussian Mixture Model (GMM) uses multiple normal distributions to fit a data set. The model weights each distribution based upon its influence in fitting the data set as well as finding each normal distribution's descriptors such as mean and variance. The GMM has found many uses, particularly in clustering operations that are time variation insensitive (Hasan and Gan, 2009). Speech recognition fields such as speaker identification have used the GMM to identify specific speakers (Reynolds and Rose, 1995). Several subsets of robotics have used the GMM as a means to train robotic behavior such as hand grasp control (Ju et al., 2008). The GMM has been applied across a wide variety of fields with success in large part due to its simplicity and adaptability.

The GMM model provides an effective means to detect ILD sensitivity errors. Traffic is composed of vehicle categories with different lengths. Each vehicle length category has its own mean and standard deviation which can be used in analysis. Most of this research focuses on Category 1 because short vehicles make up the vast majority of traffic and short vehicles have the least length variation. The three constraints shown in Equations 4-9, 4-10, and 4-12 define characteristics of the detected vehicle lengths that may be tested by the GMM. These tests are intended to remove ILDs that have severe hardware problems and sensitivity errors severe enough to lose information from the correction process. ILDs that fail constraint 1, 2, or 3 are defined as having that type of failure. For example, an ILD that failed constraint 1 would suffer a type 1 error.

4.4 Type 1 Error

Several loop detectors examined for this project continually recorded values significantly lower than the threshold used in constraint 1. Table 4-1 and Figures 4-5, 4-6 and 4-7 show the mean on-times recorded at three loop detectors.

Table 4-1 Type 1 ILD Error Table

Station	Lane	Loop	Average on time μ_1 (ms)
005ES15652	South Lane 5(HOV)	S loop	128
005ES15996	North lane 3	M loop	122
	South lane 3	M loop	120

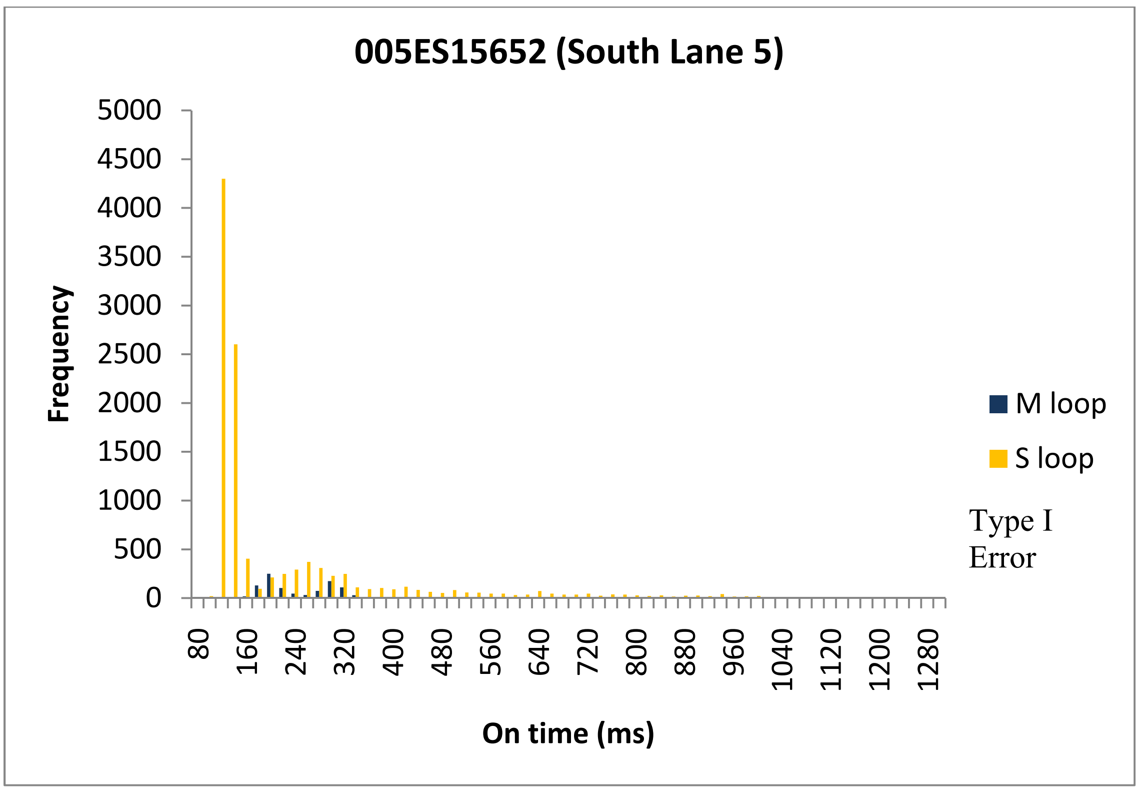


Figure 4-5 Single Loop Detector Distributions for Dual Loop Detector 005es15652 Lane 5 South

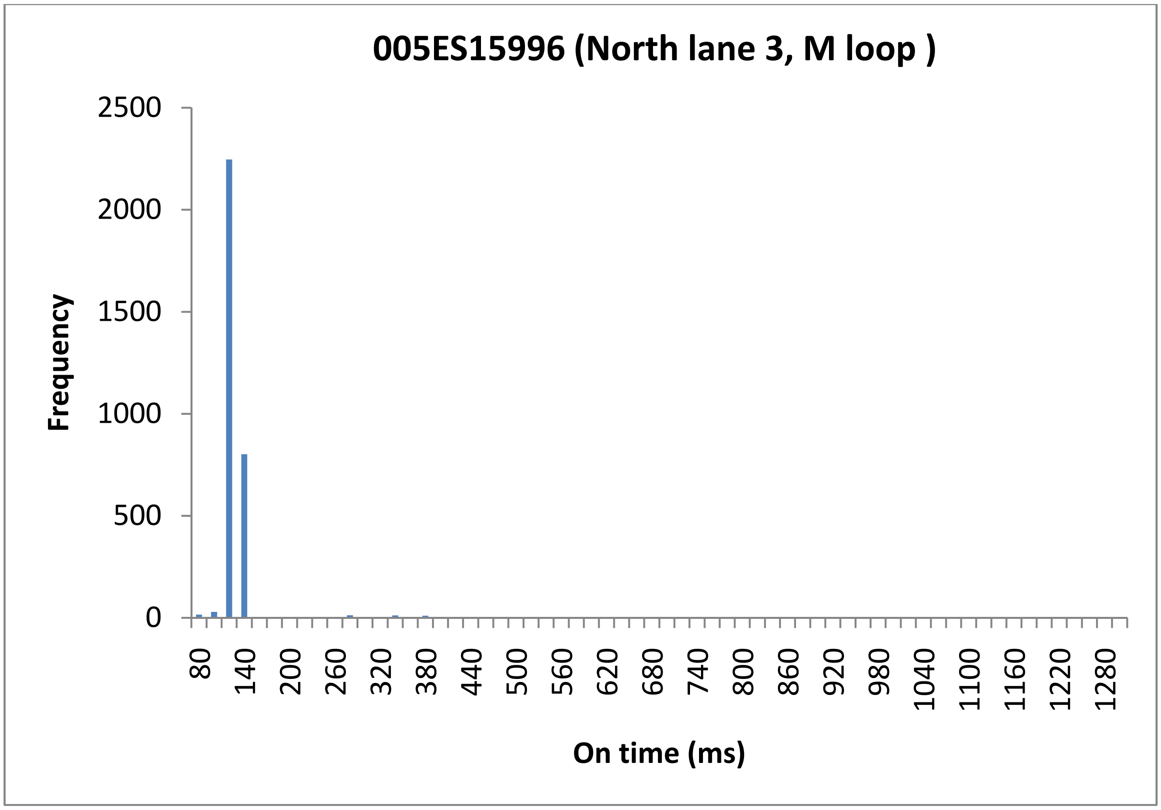


Figure 4-6 On-time Distribution for 005es15996 North Lane 3 M Loop

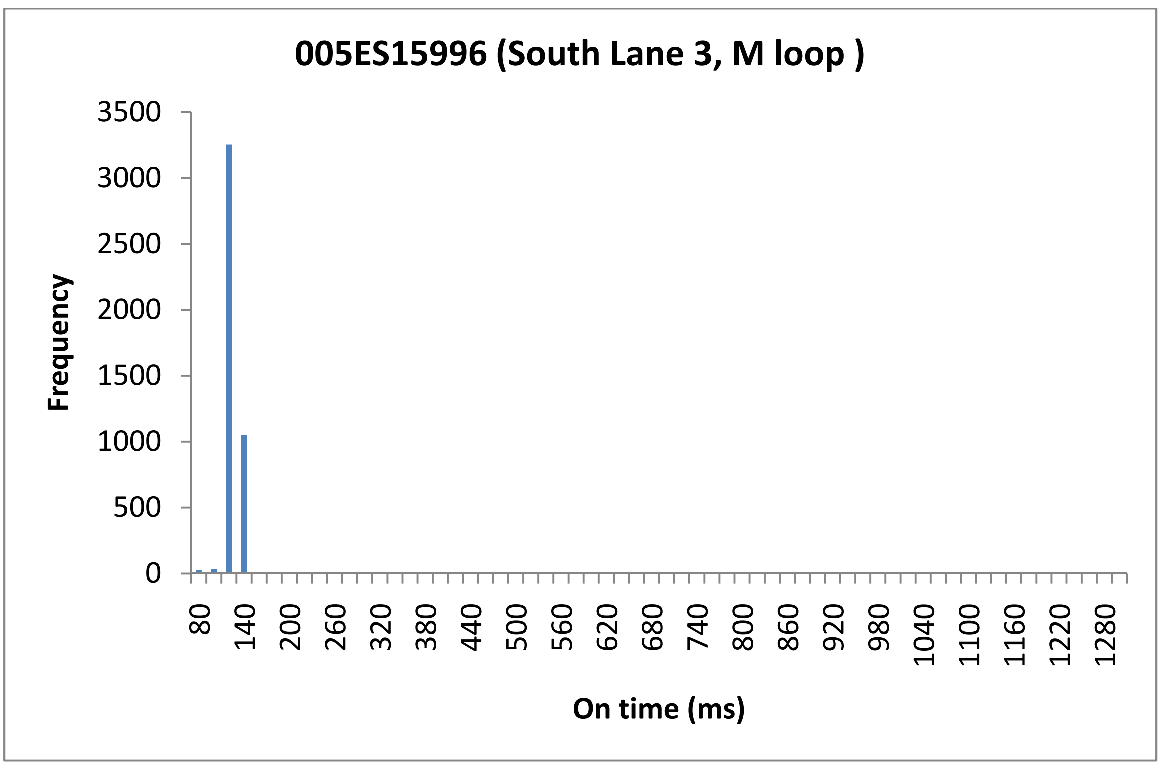


Figure 4-7 005es15996 South Lane 3 M loop

These ILDs have produced data with little to no distribution to adjust. Their values are also too short to be reasonable for vehicles at reasonable speeds. Other methods will be required to correct ILDs outputting data of this quality, but they are beyond the scope for this study. Note Figure 4-5 in particular, the two ILDs of the dual loop detector produce drastically different data with the S loop producing many times the records of the M loop. The S loop was switching on and off at random intervals producing average vehicle counts over twenty vehicles every twenty seconds or 3,600 vehicles per hour.

4.5 Type 2 Error

Another problem that became apparent as the research progressed is that some loop detectors output data with a fragmented distribution. Instead of the expected single large peak for short category 1 vehicles and wide distribution of large vehicles, these loop detectors reported on-times which resulted in three or more peaks. The exact cause of this behavior is unknown. Until the cause is determined, software correction may not be feasible. Examples of this kind of problem can be found in Figures 4-8, 4-9, 4-10 and 4-11 below. The split distributions violate the second constraint because the short vehicle category 1 distribution does not comprise over 80% of the distribution mixture. Tables 4-2, 4-3, 4-4 and 4-5 show the respective Gaussian distribution mixtures seen in the figures for each studied ILD.

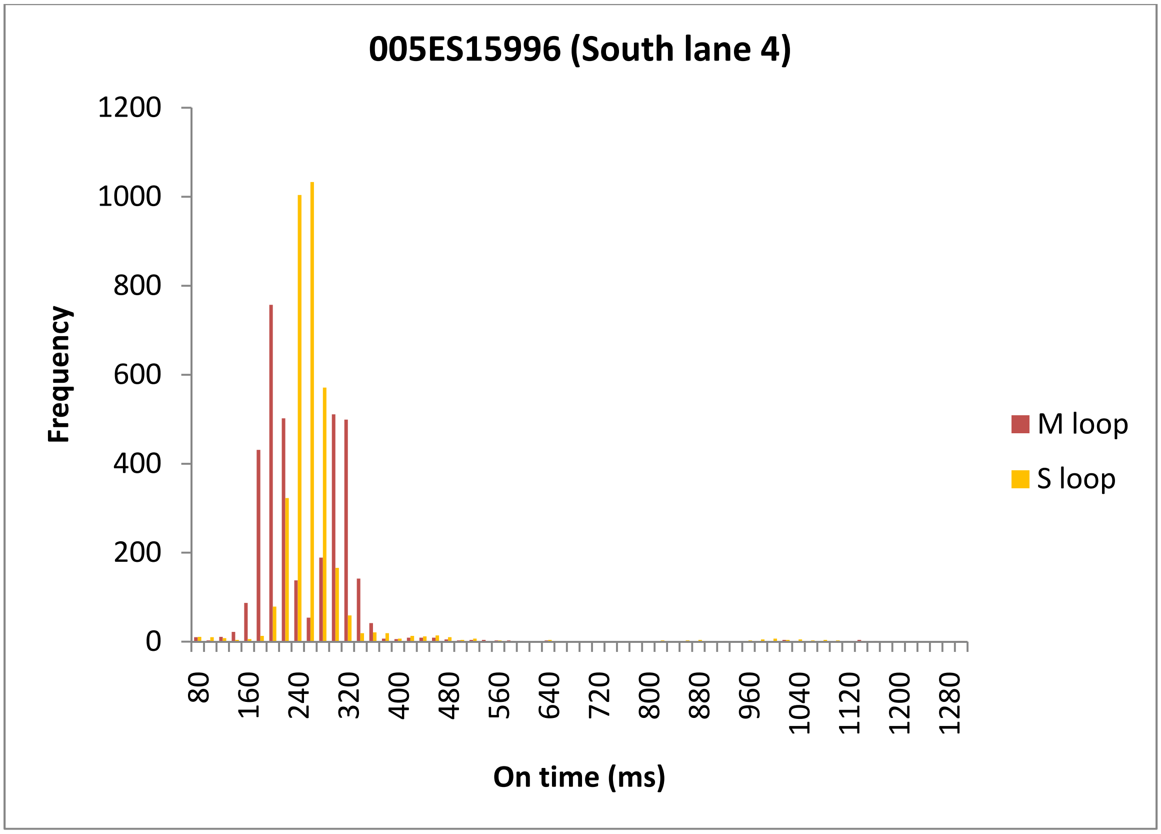


Figure 4-8 Single Loop Detector On-time Distributions from 005es15996 Lane 4 South

Table 4-2 Single Loop Detector Gaussian Distribution from 005es15996 Lane 4 South

	M loop			S loop	
	Category 1	Category 2	Category 3	Category 1	Category 2
ω_j	0.55	0.40	0.05	0.92	0.08
μ_j	194.37	300.11	432.60	246.57	478.29
σ_j^2	394.77	352.26	83750.63	508.66	96675.00

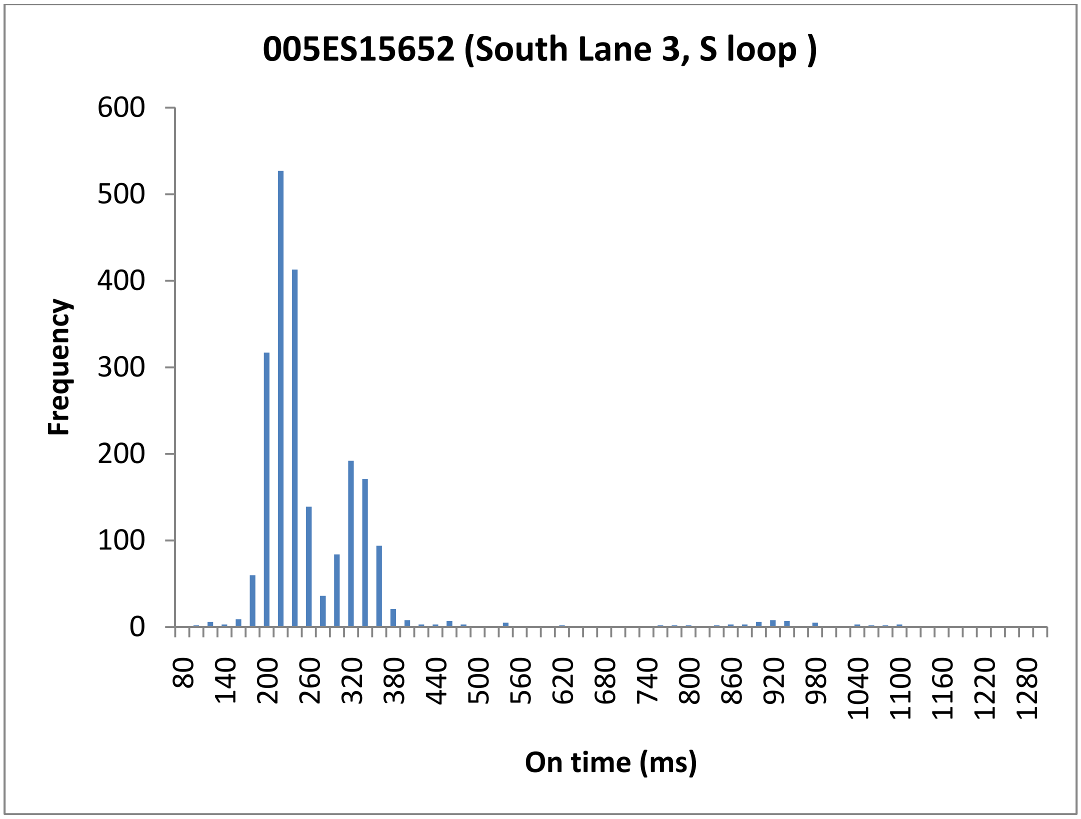


Figure 4-9 On-time Distribution for 005es15652 South Lane 3 S Loop

Table 4-3 Gaussian Distributions for 005es15652 South Lane 3 S Loop

	Category 1	Category 2
ω_j	0.61	0.39
μ_j	215.66	346.97
σ_j^2	327.96	30805.48

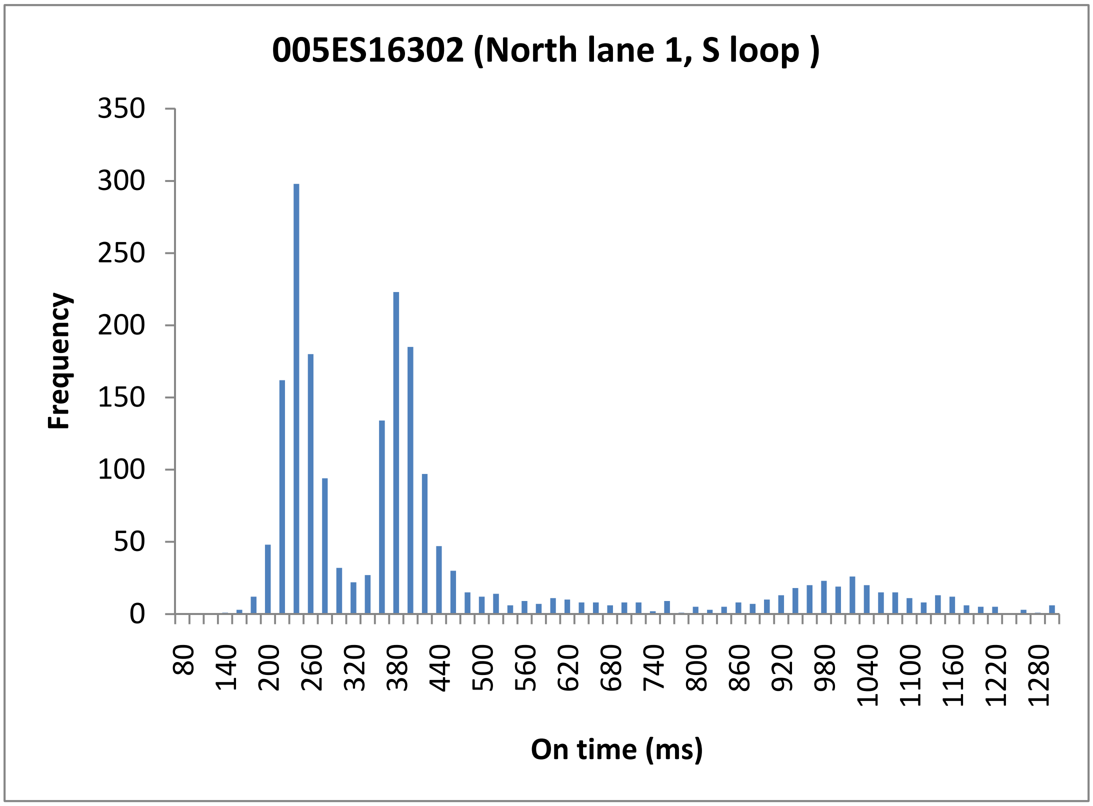


Figure 4-10 On-time Distribution for 005es16302 North Lane 1 S Loop

Table 4-4 Gaussian Distributions for 005es16302 North Lane 1 S Loop

	Category 1	Category 2	Category 3	Category 4
ω_j	0.40	0.33	0.13	0.14
μ_j	236.12	378.52	480.19	1005.93
σ_j^2	570.04	615.30	20243.78	13787.89

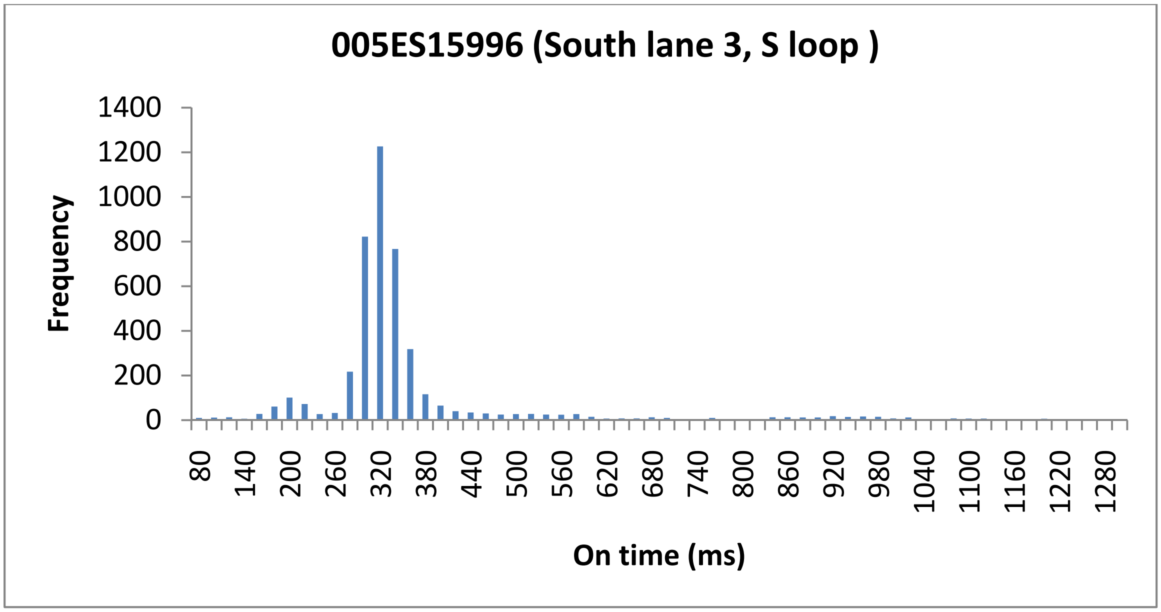


Figure 4-11 On-time Distribution for 005es15996 South Lane 3 S Loop

Table 4-5 Gaussian Distributions for 005es15996 South Lane 3 S Loop

	Category 1	Category 2	Category 3	Category 4
ω_j	0.07	0.75	0.12	0.06
μ_j	185.86	312.37	421.90	889.85
σ_j^2	1627.50	469.44	9953.74	36995.46

Comparing to the bimodal distributions shown in Figures 4-2 and 4-3 from healthy ILDs, the Category 1, short vehicle, distribution has been the split into two or more distributions. The cause of the on-time split is unknown at this time. Further investigations are needed to determine the cause of this type of error before an effective correction algorithm can be proposed.

4.6 Type 3 Error

ILDs may be free of type 1 and type 2 errors and still not generate error free data. Sensitivity errors may still exist such that the ILDs are under reporting or over reporting vehicle occupancy. With a known or chosen free flow speed, constraint 3 will define an acceptable range of on-times. Because type 3 errors will rely on details specific to each location they will be included in the case studies in the next chapter.

4.7 20-Second Aggregated Data Versus Event Data

As was stated at the beginning of the chapter either event data or 20-second aggregated data may be used. Thus far event data have been used due to both the number of samples and the additional information available to diagnose unexpected errors. Event data is not absolutely necessary to conduct the sensitivity error correction; however, these proposed analyses may be conducted with properly screened aggregated data, such as 20-second data. The 20-second data, collected in August 2009, used in this section consists of intervals with only one vehicle and empty intervals before and after, this way only that one vehicle's occupancy can be extracted and considered. By converting the occupancy percentages into on-time values the exact same analysis can be conducted as has been done with event data. As examples, Tables 4-6 through 4-8 provide the Gaussian distributions for three stations using event data and 20-second data respectively.

Table 4-6 Gaussian Distributions for 005es16302 North Lane 3 M Loop from Event and 20-Second Data

	Event Data		20-Second Data	
	Category 1	Category 2	Category 1	Category 2
ω_j	0.90	0.10	0.94	0.06
μ_j	215.17	294.88	216.95	352.50
σ_j^2	526.12	31235.57	868.92	62404.58

Table 4-7 Gaussian Distributions for 005es15996 South Lane 4 M and S Loops from Event and 20-Second Data

Event Data					
	M loop			S loop	
	Category 1	Category 2	Category 3	Category 1	Category 2
ω_j	0.55	0.40	0.05	0.92	0.08
μ_j	194.37	300.11	432.60	246.57	478.29
σ_j^2	394.77	352.26	83750.63	508.66	96675.00
20-Second Data					
	M loop			S loop	
	Category 1	Category 2	Category 3	Category 1	Category 2
ω_j	0.54	0.45	0.01	0.96	0.04
μ_j	192.17	299.43	378.02	261.52	102.54
σ_j^2	461.93	838.66	100387.20	2673.90	730.31

Table 4-8 Gaussian Distributions for 005es15652 South Lane 4 M and S Loops From Event and 20-Second Data

Event Data				
	M loop		S loop	
	Category 1	Category 2	Category 1	Category 2
ω_j	0.92	0.08	0.96	0.04
μ_j	200.80	335.89	198.01	383.53
σ_j^2	326.02	33482.96	347.05	62083.38
20-Second Data				
	M loop		S loop	
	Category 1	Category 2	Category 1	Category 2
ω_j	0.93	0.07	0.97	0.03
μ_j	203.57	453.26	200.43	555.29
σ_j^2	628.31	102129.71	627.87	65558.20

Note that in Tables 4-6 through 4-8 even though the data sets are from different sources the center points for the major distributions are relatively constant with only a few milliseconds difference between the results from the event data and the 20-second data. The minor distributions, however, ($\omega_j \leq 0.10$) can change rather substantially between 20-second data and event data with the Category 2 distribution in Table 4-6 decreasing in weight from 0.10 to 0.06 and shifting mean by ~58 ms. For the smaller distributions this is not very surprising because long vehicles are typically slower than short vehicles and the chance of capturing just one long vehicle in a 20 second interval is much lower than capturing just a short vehicle. Additionally, many of the minor distributions observed so far have had fairly large standard deviations ranging from 100 ms to nearly 317 ms. The large standard deviation tends to indicate that these distributions are sensitive to outliers because the model is attempting to include every point in the distributions.

Table 4-8 is of particular interest because it shows the changes from event data to 20-second data for one of the sites chosen for correction in the next chapter. Both of the primary distributions are practically unchanged with only ~1% difference between the event data means and the 20-second aggregated data means. There are also only small changes to the distribution weightings indicating that the results would be similar between the event data derived parameters and the 20-second aggregated data derived values.

For ILDs with available 20- or 30-second aggregated historical data the analyses performed previously may be executed using the aggregated data with little loss of accuracy. Event data remains the more flexible choice for error detection despite the additional effort that must be undertaken to gather it. Event data allows researchers to observe more varied errors that are not detectable in aggregated data. Event data may also be gathered for any time of day allowing analyses to include intervals with heavier traffic that may not easily be analyzed with aggregated data. The choice between aggregated and event data based analyses will likely come down to opportunity and selection processes with aggregated data used to select targets for event data collection.

4.8 Error Detection Discussion

When error correction was first proposed, the research team expected known loop errors such as crosstalk and pulse mode to be problematic but manageable. What were unexpected were the effect of extreme low sensitivity and the appearance of the bifurcated distributions seen earlier in the chapter such as Figure 4-9. After a close examination of the event data and research into the actual function of the loop controller card, a possible cause was found. To explain the possible cause, the working principles of ILDs are briefly introduced.

ILD loop controller cards use an electrical system that reverses the current flow through the ILD loop coil at rates above 10 kHz (Klein et al., 2006). This oscillation is tied to the inductance of the loop coil. If the inductance drops, the oscillations increase and the voltage of the system drops. The voltage drop is the actual measured quantity used by the loop controller card to determine if the loop coil is occupied or not.

The inductance change seen by the ILD is driven by eddy currents induced in the ILD loop coil by the passing of metal through the electrical field created in the loop coil by the loop controller card's oscillating current. There are two effects on loop coil inductance. The first is an increase in the inductance caused by the presence of metal in the loop coil. The second is a much larger reduction of inductance caused by eddy currents induced in the metal (Klein et al. 2006).

Very little research has looked at the actual inductive pattern developed by vehicles traveling over ILDs. A team led by Stephen Ritchie in collaboration with Carlos Sun has been examining inductive patterns and their research may provide a possible answer to several questions about the behavior seen previously. To help explain the behavior the figures below are reproduced from their 2001 PATH report.

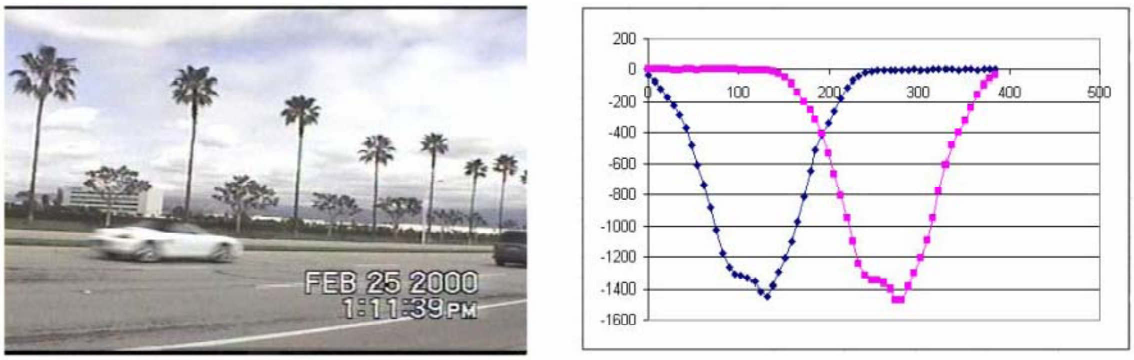


Figure 4-18 Car Inductive Signature (Ritchie et al., 2001)

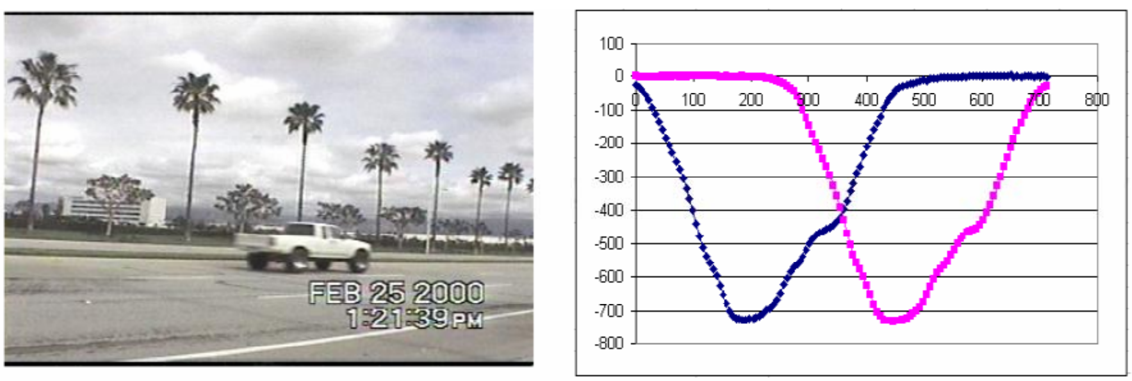


Figure 4-19 Pickup Truck Inductive Signature (Ritchie et al., 2001)

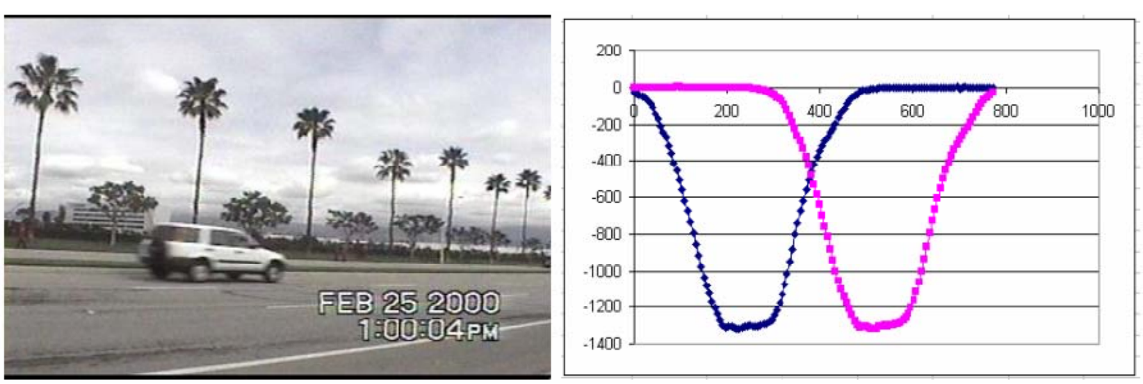


Figure 4-20 SUV Inductive Signature (Ritchie et al., 2001)

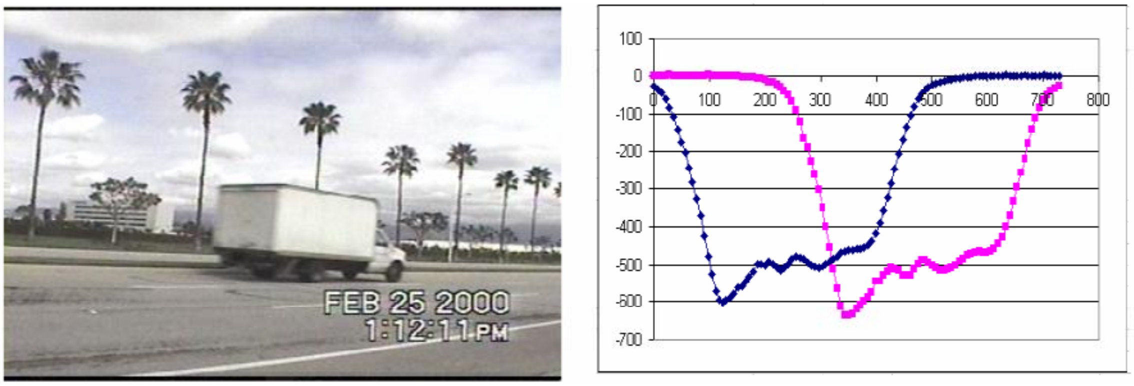


Figure 4-21 Unit Truck Inductive Signature (Ritchie et al., 2001)

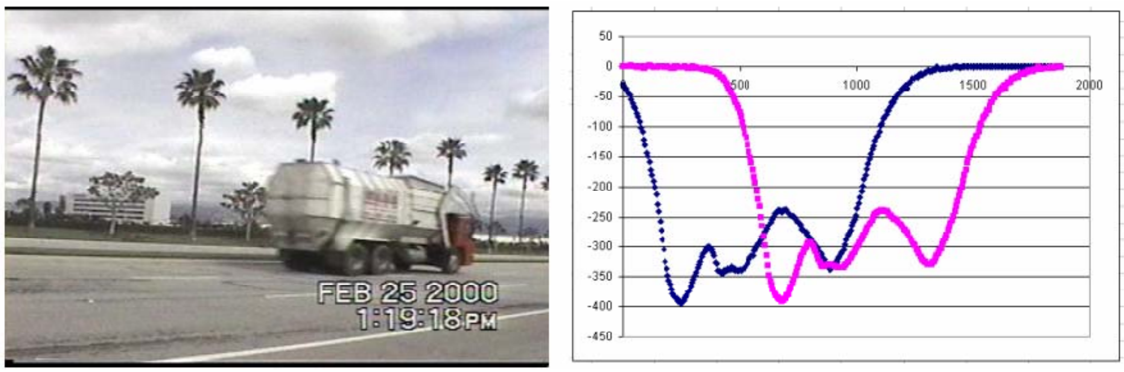


Figure 4-22 Trash Truck Inductive Signature (Ritchie et al., 2001)

Figures 4-18 through 4-22 show the inductive signatures produced by each vehicle type. For a clearer comparison and discussion of how the inductive signatures may be affecting the results seen in Figures such as Figure 4-9, the inductive signature data from Ritchie et al. (2001) have been superposed in Figure 4-23 on a common scale with three threshold values added for the discussion.

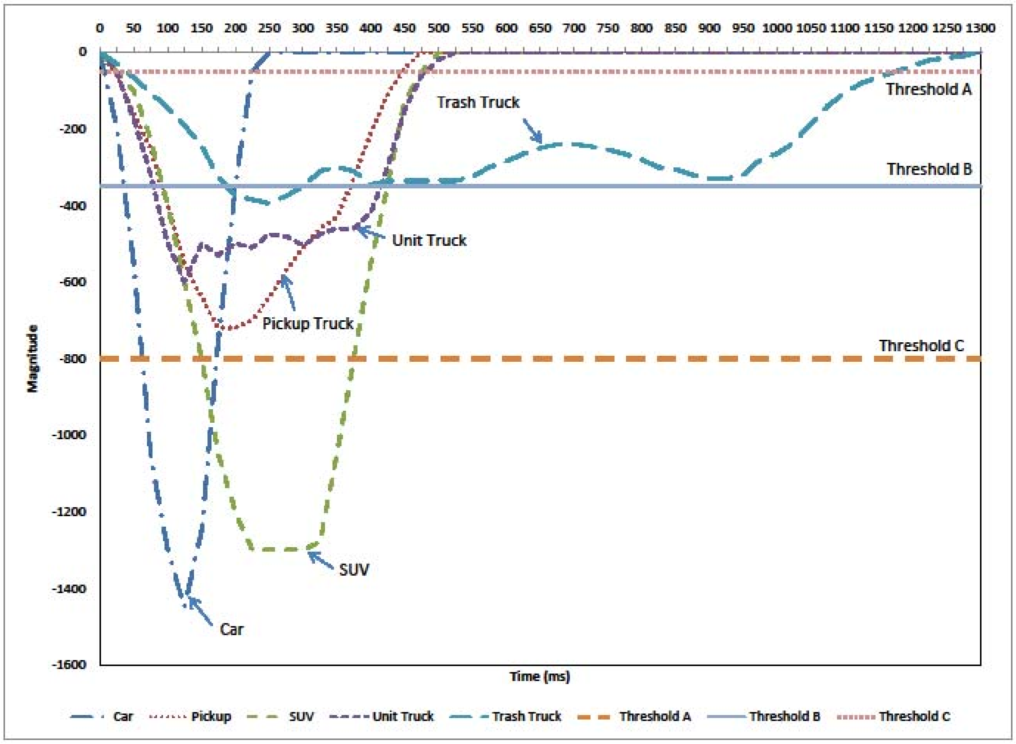


Figure 4-23 Inductive Signatures Plotted on Common Scale

Figure 4-23 shows what may be a counterintuitive result for most people, cars and SUVs generate stronger responses than large trucks. The reason for this is that the electrical field of the loop weakens according to distance squared. Large trucks may have more metal to affect the ILD loop coil but they also typically have much higher under carriages and are much larger than the loop coil which puts that metal further from the ILD loop coil.

Consider threshold A in the figure. At this sensitivity level, all vehicles are detected quickly and should have on-time values that very closely represent their true length and speed. The threshold is high enough to minimize random noise impact and yet low enough to quickly and correctly detect vehicles.

Threshold C, on the other hand, is drastically under sensitive and completely misses the trucks. SUVs and cars are detected cleanly but not for their entire lengths. This threshold may not raise failure concerns because it will still be outputting data, and much of that data will not even be flagged by most of the thresholds commonly used to screen data such as those used in (Turochy and Smith, 2000). The failure of the thresholds to detect this failure will be tied to the number of SUVs and longer cars in the traffic stream because these vehicles may produce on-time values large enough to be reasonable for small cars. At this level the SUV would produce an occupancy value of approximately 225 ms at freeway speeds, which is reasonable for a small car. However, with trucks missing entirely this data is critically flawed.

Threshold B is the most interesting case. At this level all vehicles are detected though not necessarily detected correctly. Notice that the trash truck is only detected for about one hundred milliseconds. Also, note that a small difference in threshold sensitivity or truck inductive signature could result in the trash truck generating multiple short detections. This behavior may be one cause of pulse break up. At threshold B, cars and SUVs are little affected by the sensitivity change and will be correctly reported for the majority of the expected on-time. Pickup trucks and unit trucks will be cleanly and clearly detected but their lengths will be underestimated by a large percentage compared to cars and SUVs and also in comparison to sensitivity threshold A.

This sensitivity behavior is very likely to be the cause of single distribution data that is narrowly distributed and lacking truck distributions. When the sensitivity is too low only cars and SUVs will be detected and even then only for a shorter period. See Figures 4-6 and 4-7 for possible examples. As sensitivity increases, more vehicles are detected but not necessarily completely. This incomplete detection may be responsible for the peaks at very short on-times seen in some split distributions. Figures 4-9 through 4-11 may be affected by this phenomena, however there are almost certainly additional factors affecting these distributions. Note that for trucks, including pickup trucks, their signal weakens toward the end of the truck. This could seriously affect the assumption of symmetrical d when sensitivity levels are too low.

Chapter 5 Error Correction Methodology and Application

The next step after the identification of errors is to correct those errors when possible. This correction revolves around correcting the detection zone length of an ILD. By adjusting the detection zone length used in calculations to reflect the actual ILD detection zone length, the outputted data will improve in accuracy. The adjustment of the detection zone length to its correct level is accomplished by adjusting the detection zone distance offset, *d*. Three correction scenarios will be described in the following sections. First, single loop detectors will be corrected, followed by the correction of dual loop detectors. The third section will explore one of the borderline cases. Error correction results will be discussed and summarized in the final section.

Before moving into the direct application of the error correction methodology a quick discussion of the available data is in order. Event data from eleven dual loop detectors was collected for this analysis. One lane at Study Site 3, 005es16302 northbound lane 4 (HOV), had such low traffic counts that it was not analyzed, leaving ten dual loop detectors for error detection and correction analysis. The Category 1 mean on-time and weighting factor for these detectors are summarized in Table 5-1 below.

Table 5-1 Loop Detector Characteristics Summary

Station	Lane	M loop on time		S loop on time	
		μ_1(ms)	ω_1	μ_1(ms)	ω_1
005ES15652	South 3	203	0.82	215	0.61
	South 4	201	0.92	198	0.96
	South 5	274	0.60	123	--
005ES15996	North 3	119	--	308	0.66
	North 4	195	0.82	251	0.90
	South 3	120	--	312	0.75
	South 4	194	0.55	247	0.92
005ES16302	North 1	217	0.58	236	0.40
	North 2	230	0.88	230	0.63
	North 3	215	0.90	323	0.73

In Table 5-1 the ILDs that are suffering type 1 sensitivity errors are filled with dark gray. These ILDs are reporting data with no distribution to correct and so are thrown out. Similarly, the ILDs suffering from type 2 error have their cells filled with light gray. These ILDs show split distributions in their on-time values that cannot currently be satisfactorily explained or corrected. Because the cause of this error is unknown and its effects render the correction method ineffective, these ILDs are also thrown out. Hardware correction is required to fix the type 1 and type 2 errors seen at nine of the eleven event data collection sites. With all of the type 1 and type 2 error loop detectors ineligible for correction, there are only two dual loop detector candidates remaining where one or both ILDs suffer from type 3 error, lane 4 southbound at Study Site 1 (cabinet 005es15652) and lane 4 northbound at Study Site 2 (cabinet 005es15996).

The remainder of this chapter is composed of a more detailed examination of the correction methodology, followed by two case studies and then a summary of the lessons learned. The correction methodology will be presented for two different dual loop detectors. These two dual loop detectors are used for the case studies to follow because their correction method includes correcting the individual loop detectors.

5.1 Error Correction Methodology

As mentioned in Chapter 4, Type 1 errors reduce the expected breadth of the ILD's on-time distribution to a very narrow range centered on unreasonably low values. Type 2 errors result in split on-time distributions which are beyond the scope of this study for error correction. Type 3 errors are errors in sensitivity, the focus of this research. ILDs that do not suffer from any of the three types of error do not require correction because they are already reporting data that is within the selected margin of error.

Correcting each error type requires a different strategy. Correcting type 1 errors can only be done in hardware. Type 1 errors are believed to be the results of extreme under sensitivity as discussed in the previous chapter. Correcting that under sensitivity can only be done by adjusting the loop amplifier card's sensitivity setting. An additional cause of type 1 error may be incorrect mode settings such as pulse mode, which would also require a manual correction of the ILD's settings. Software correction of Type 2 errors has not been developed because the cause of type 2 errors is still unknown and will require further research before effective correction methods can be developed. Therefore, the correction of type 2 errors can only be accomplished through hardware correction at this time. Type 3 errors indicate ILDs suffering from sensitivity level errors that rarely miss vehicles or substantially alter their on-time characteristics. These ILDs are not operating at the correct sensitivity level but can capture sufficient data about passing vehicles without impacts from adjacent loop detectors.

The determination of type 3 errors requires knowledge of either vehicle speeds at the particular location or each vehicle's effective length. Considering that the effective length of each vehicle is very difficult to get, this study focuses on the approach that uses known speeds. Typically, the mean free flow speed of a facility is very predictable. Hence, we propose collecting data during the free flow condition for checking Type 3 errors. If the free flow speed is unknown, researchers will need to estimate the free flow speed by other methods such as looking at local dual loop detector data. As a last resort, free flow speeds may be estimated by the methods presented in the Highway Capacity Manual (2000). The HCM method in particular should be applied with caution because it is based on average values which may or may not be applicable to the location being studied. For this research, the free flow speeds were either generated by the dual loop detectors themselves or from other nearby dual loop detectors.

After a free flow speed has been selected for the location, Type 3 error can be checked according to Equation 4-12, reproduced below as Equation 5-1.

$$\frac{\overline{L_V}+L_L-2\gamma}{v_f} < \mu_1 < \frac{\overline{L_V}+L_L+2\gamma}{v_f} \qquad \text{(5-1) (C3)}$$

Equation 5-1 will generate upper and lower bounds on the on-time distribution. When the category 1 mean is outside of the range described by this equation, the ILD is suffering from Type 3 error. ILDs that suffer from Type 3 error are eligible for correction.

The correction process begins with Equation 5-2 which is a reorganization of Equation 4-7 solving for the detection zone offset distance.

$$d = \frac{1}{2}\left(\mu_1 \cdot v_f - \overline{L_V} - L_L\right) \qquad \text{(5-2)}$$

In Equation 5-2, μ_1 is the category 1 short vehicle on-time distribution mean, v_f is the free flow speed, $\overline{L_V}$ is the short vehicle average length, and L_L is the loop length. Once d is calculated for an ILD it can be applied to the detection zone length L_L as in Equation 5-3.

$$L'_L = L_L + 2d \tag{5-3}$$

For dual loop detectors, each ILD has a d value which will change the distance between the M and S loops as in Equation 5-4.

$$L' = L + d_m - d_s \tag{5-4}$$

The change in the detection zone length will result in improved speed calculations for single loop detectors using Equation 5-5

$$v_{sl} = \frac{L_V + L'_L}{OT} \tag{5-5}$$

where L'_L is described in Equation 5-3, L_V is a vehicle's length, *OT* is the on-time for the vehicle, and v_{sl} is the calculated vehicle speed using single loop measurements. Since *OT* is a direct measurement of single loops and L'_L can be calculated using Equation 5-3, vehicle speed can be calculated if L_V is known. Unfortunately, vehicle length varies from vehicle to vehicle and L_V cannot be measured from single loops. However, as indicated in the vehicle length distribution shown in Figure 4-2, short vehicle lengths change only narrowly around their mean of 15.2 ft, with a standard deviation of σ=2.2 ft. Therefore, using the mean length of short vehicles $\overline{L_V}$ to replace L_V in Equation 5-5 may not introduce significant errors. This is particularly true when using the average *OT* for several short vehicles detected in a short time interval, e.g. 20 seconds or 30 seconds. Typically, traffic speed is relatively constant over such a short time interval. Assuming *n* short vehicles are detected in the interval, the sample vehicle length of these *n* short vehicles will be much closer to the mean short vehicle length, as indicated by the reduced standard deviation of $\frac{\sigma}{\sqrt{n}}$.

Rearranging Equation 5-5 yields the following equation for vehicle length calculation.

$$L_V = v_{sl} * OT - L'_L \tag{5-6}$$

When v_{sl} is known, vehicle length can be calculated using single loop measured *OT*. Although vehicle speeds change rapidly during congested periods, they are fairly predictable under free flow conditions. Therefore, we assume free flow speed is known at a particular location and propose using single loop measurements under free flow conditions for vehicle length calculation in this study.

The changes in detection zone lengths similarly affect the dual loop speed and vehicle length calculations as shown in Equations 5-7 and 5-8 below.

$$v_{dl} = \frac{L^{'}}{\Delta t_a} \tag{5-7}$$

$$L_V = v_{dl} * OT - L'_L \tag{5-8}$$

where *L'* is described by Equation 5-3, Δt_a is the difference in arrival times between the M and S loops and v_{dl} is the dual-loop calculated vehicle speed, By adjusting for the actual detection zone length, the speed and vehicle length values reported by the loop detectors can be corrected.

5.2 Case Study One

This dual loop detector is located on lane 4 of southbound I-5 at milepost 156.52 at cabinet 005es15652. This lane is the innermost general purpose lane. This location has a directional AADT of approximately 120,000 vehicles. The dual loop detectors are located at the top of a small hill. The segment is also curved which reduces the line of sight at this location.

This dual loop shows virtually the same sensitivity level between ILDs with the M loop reporting 201 ms on average for its category 1 Gaussian distribution and the S loop reports a very close 198 ms for category 1. These category 1 mean values are above the threshold set in constraint 1. Both show very strong weighting factors, ω_1, with each passing constraint 2 by more than 10%. According to Equation 5-3, with the two ILDs at nearly the same sensitivity level, *L* will approximately equal *L'* because d_m will be nearly equal to d_s. If *L* equals L' there is no difference in the speed calculated for the dual loop detector before and after correction. The assumption for this dual loop detector is that the speed measurement will not appreciably change with sensitivity adjustment because the *d* values for each ILD should be nearly equal.

The paired on-times for the two ILDs are shown below in Figure 5-1.

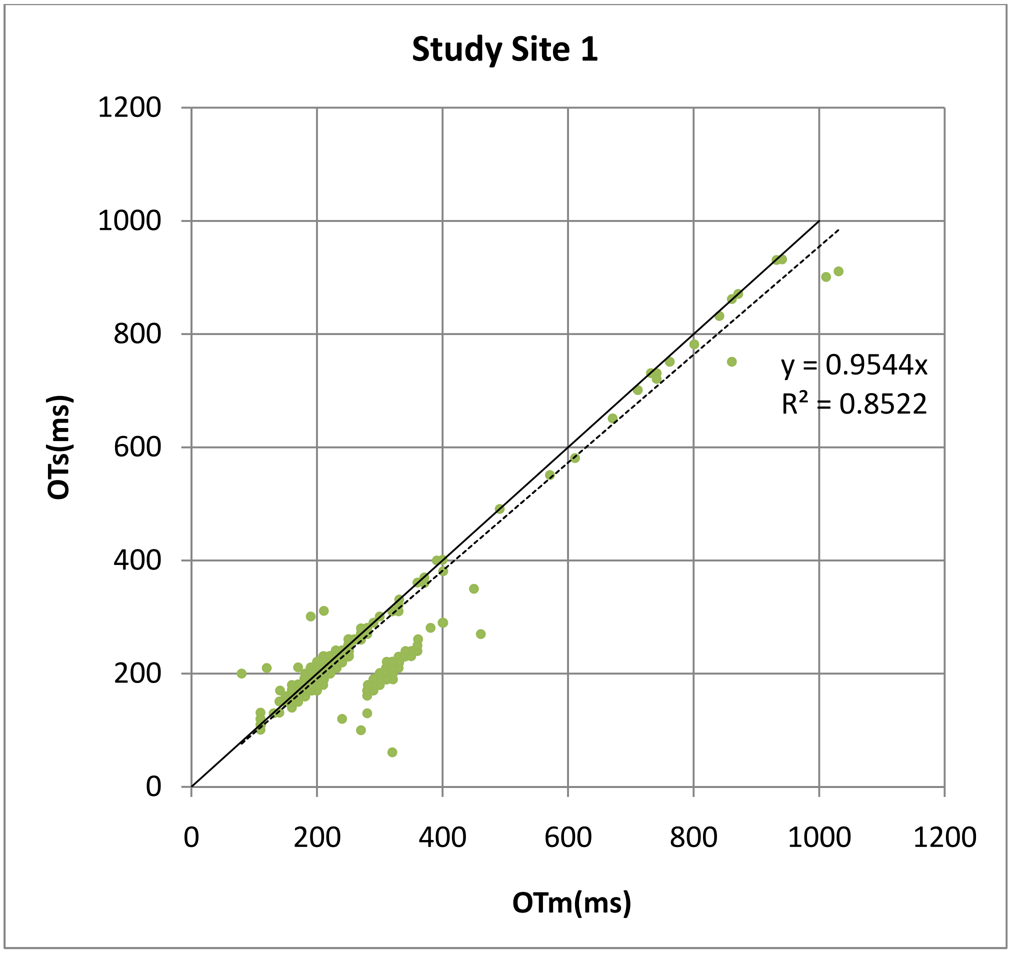


Figure 5-1 Paired On-times for South Lane 4 M and S Loops at Study Site 1

The solid 45° diagonal line in Figure 5-1 shows where the on-times of the M and S loops are equal. When the S loop reports a longer on-time than the M loop for a particular vehicle, the dot for the vehicle will be placed above the line. Otherwise, it will be below the line. For this dual loop detector, the two ILDs report substantively equal on-times as indicated by the dashed linear fit line and its R^2 value of 0.85. Note that the long vehicles with their higher on-times are behaving very similarly to the shorter vehicles.

Although both ILDs agree with their sensitivity level, we are not sure if they are at the correct sensitivity level. This implies that both ILDs need to be adjusted toward the same direction to fix any sensitivity problem it may have. As stated previously, the free flow speed used to check constraint 3 and calculate *d* will be the speed measurement before correction. Once the individual ILDs are corrected, the dual loop detector corrections can be made and the speed and length measurements can be improved.

To determine if the ILDs need correction, constraint 3 shown in Equation 5-1 will be checked. The range for constraint 3 is defined by the selection of a free flow speed, average short vehicle length, and detection zone length. The free flow speed used for the single loops will be the median speed recorded by the dual loop detector with the two ILDs at the same sensitivity level. If the two ILDs do not have the same sensitivity level, correction must be made to adjust the S loop sensitivity to the level of the M loop. For this particular dual loop, the M and S loops already agree with the sensitivity level, so the step to make the two ILDs agree with their sensitivity levels is skipped. The free flow speed used for checking constraint 3 is 64 mph, the average short vehicle length is 15.2 feet (Wang and Nihan, 2000); and the detection zone length is 6 feet. The acceptable speed range is chosen as ±10% so $\gamma = \pm 1.06$ feet. Taking these values in Equation 5-1 results in

$$\frac{15.2\,ft + 6ft - 2(1.06\,ft)}{64\,mph} < \mu_1 < \frac{15.2\,ft + 6ft + 2(1.06\,ft)}{64\,mph} \quad (5\text{-}9)$$

$$203\,ms < \mu_1 < 248\,ms \quad (5\text{-}10)$$

Since μ_1 for the M loop is 201 ms which is beyond the acceptable range of 203 ms to 248 ms, this ILD is clearly under sensitive. Similarly the S loop with a μ_1 of 198 ms is also under sensitive. Therefore, both the M and S loops need sensitivity correction. The procedure to correct these ILDs will be described in the following two subsections.

5.2.1 Single Loop and Dual Loop Corrections

M Loop Correction

Correcting the M loop begins with calculating the detection zone offset distance, *d*, using Equation 5-2. The average short vehicle length used here is 15.2 feet and the loop length is 6 feet. For the M loop μ_1 is 201 ms or .201 seconds. The average and median speeds measured by the dual loop detector are 59.7 mph (87.56 feet per second) and 60.3 mph (88.44 feet per second), respectively. As stated previously, looking at the two ILDs, there is a reasonable expectation that this particular dual loop detector is functioning correctly with regards to speed measurement and therefore the median speed collected during the free flow period will be used as the free flow speed. The median is used because it is less sensitive to outliers (Coifman et al., 2003). The *d* value of this M loop resulting from Equation 5-2 is

$$d_m = \frac{1}{2}\left(\mu_1 \cdot v_f - \overline{L_V} - L_L\right) = \frac{1}{2}(.201\ s * 93.97\ fps - 15.2\ ft - 6\ ft) = \ -1.16\ ft \tag{5-11}$$

With this d_m value, lane occupancy measured by this loop can be corrected as

$$O'_m = O_m \times \frac{L_V+L_L}{L_V+L_L+d_m} = \frac{\sum_{i=1}^{N} OT_i}{T} \times \frac{L_V+L_L}{L_V+L_L+d_m} \tag{5-12}$$

where *T* is length of data integration interval (20 seconds for WSDOT), O'_m and O_m are the original and corrected lane occupancies, respectively, measured by the M loop, *N* is the total number of vehicles detected in *T*.

S loop Correction

The S loop's μ_1 value is 198 ms which is also clearly under sensitive with regards to constraint 3. The S loop correction proceeds in an identical manner to the M loop correction. First *d* is calculated by Equation 5-2, using the median speed from the dual loop detector, the short vehicle average length, the detection zone length, and the S loop Category 1 mean on-time:

$$d_s = \frac{1}{2}\left(\mu_1 \cdot v_f - \overline{L_V} - L_L\right) = \frac{1}{2}(.198\ s * 93.97\ fps - 15.2\ ft - 6\ ft) = \ -1.30\ ft \tag{5-13}$$

The d_s value for the S loop is larger than for the M loop because the S loop is more under sensitive. Following Equation (5-14), the original lane occupancy (O_s) measured by the S loop can also be corrected to O'_s.

$$O'_s = O_s \times \frac{L_V+L_L}{L_V+L_L+d_s} = \frac{\sum_{i=1}^{N} OT_i}{T} \times \frac{L_V+L_L}{L_V+L_L+d_s} \tag{5-14}$$

Dual Loop Correction

The M and S loop were both under sensitive according to constraint 3. When their *d* values were calculated, d_m and d_s respectively were -1.16 feet and -1.30 feet. Applying these values to Equation 5-4 and then to Equation 5-7 and Equation 5-8 yields Equations 5-15 and 5-16, respectively, for the corrected dual loop speed and vehicle length calculations.

$$v_{dl} = \frac{L + d_m - d_s}{\Delta t} = \frac{17 - 1.16 + 1.30}{\Delta t} = \frac{17.14}{\Delta t} \quad (5\text{-}15)$$

$$L_V = \frac{17.14}{\Delta t} * OT - L'_L \quad (5\text{-}16)$$

L' is only 0.14 feet larger than *L*, which represents a difference of less than 1% in speed calculations between the corrected and uncorrected data.

5.2.3 Study Site 1 Correction Validation

The validation dual loop detector correction will depend on the improvements seen in vehicle length distribution. The length distributions obtained from the dual loop before and after the correction will be compared to the length distribution data generated by Wang and Nihan (2000), which is used as the ground-truth vehicle length distribution in this study for the I-5 corridor. If the vehicle length distribution derived from the corrected dual-loop is similar to the ground-truth distribution, then the dual-loop correction is considered successful.

For single loop detectors, volume and occupancy are the only two measurements. Considering that this proposed approach is applicable to only loops with correct vehicle counts, occupancy (or on-time) is the only data to correct and validate. However, no occupancy ground-truth is available for comparison. Alternatively, speed data produced by the corrected dual loop can be used the ground-truth data for validating single loop derived vehicle speed data. Considering that a vehicle's speed ties directly to its on-time (or occupancy) measurement, if the calculated speed data can be validated, then we have a reason to believe that the sensitivity correction measure applied to the single loop is also effective.

5.2.3.1 Dual-Loop Vehicle Length Validation

Once the sensitivity problem is corrected, the dual loop should provide accurate speed measurements and does not require any validation. However, its length data involves single loop on-time measurements in calculation and should be checked with the ground-truth vehicle length distribution.

The length improvement will be based on an extended application of the GMM to the length data calculated using either the M loop or S loop measured on-times for before and after correction. The Wang and Nihan (2000) observed short vehicle length distribution and the Category 1 GMM distribution data are shown in Table 5-2.

Table 5-2 Statistics for Dual Loop Length Distributions for Category 1 (short vehicles) at Study Site 1

	Wang and Nihan (2000)	Using M loop On-times		Using S loop On-times	
		Before	After	Before	After
ω_j	NA	0.93	0.93	0.97	0.97
μ_j	15.18	12.75	15.11	12.49	15.11
σ_j^2	1.72	1.63	1.63	2.33	2.33

We can see that the similarities between the corrected short vehicle length distributions are much closer to the ground-truth short vehicle length distribution. Table 5-3 shows several t-test results for comparing the vehicle length distributions. We can see that the vehicle length data calculated using the M loop or the S loop on-times before the sensitivity correction is significantly different from the ground-truth short vehicle length distribution at the p=0.05 level. In contrast, the vehicle lengths data produced with the corrected single loop on-times from either the M loop or the S loop do not show significant difference with the ground-truth data. The "Using the M Loop OTs vs. Using the S Loop OTs" column shows the t-test results for the vehicle length differences between the M and S loop on-time calculated vehicle lengths before and after the correction. They are identical after the correction, but were significantly different before the sensitivity corrections were conducted.

Table 5-3 Hypothesis Test for Dual Loop Length Corrections at Study Site 1

	Using the M Loop OTs vs. the S Loop OTs		Using the M Loop OTs vs. Ground-truth		Using the S Loop OTs vs. Ground-truth	
	Before	After	Before	After	Before	After
t value	5.68	0	55.77	1.61	56.48	1.47
p value	0.0001	1	0.0001	0.1074	0.0001	0.1455
Significant*	Y	N	Y	N	Y	N

* Significant at the p=0.05 significant level

5.2.3.2 Speed Correction

As mentioned earlier, dual-loop speed data are considered accurate after the sensitivity correction and are used to validate speeds calculated from single-loop measurements. A summary of the speed data calculated for this study site can be found in Table 5-4. The minor difference between the dual loop speed measurements before and after correction indicates that the speed calculation of a dual loop detector is not significantly degraded when the sensitivity level is incorrect as long as both ILDs have nearly the same sensitivity level. The single loop speed estimations, however, are greatly affected by sensitivity level errors as evidenced by the before correction speeds in Table 5-4. After correction, the speed estimates from the single loop detectors are much closer to the dual loop measured speeds.

Table 5-4 Study Site 1 Speed Data Summary

	Dual Speed (mph)		M Loop Speed (mph)		S Loop Speed (mph)	
	Average	Median	Average	Median	Average	Median
Before	63.4	64.0	71.0	72.3	73.1	72.3
After	63.9	64.5	63.1	64.2	64.1	63.3

Tables 5-5, 5-6, and 5-7 show another application of the GMM, this time to the speed data from the dual loop detector, M loop single loop speed estimation, and S loop single loop speed estimation, respectively. The GMM separates the data into different distributions (Categories) and can effectively aid in the removal of outliers. Lane changing behavior and long vehicles, among a number of other factors can heavily influence single loop speed estimations. The formula in Equation 5-5 was used to calculate single loop speeds.

The single loop speed estimation method only uses on-time to estimate speed. Lane changing behavior can lead to partial on-time detections, resulting in shorter on-times than a complete detection. Long vehicles, by contrast would generate longer on-times at the same speed as a short vehicle because the long vehicle will require more time to cross the ILD's detection zone. The longer on-times of trucks appear as slow speed in the single loop speed estimations using Athol's approach (Athol 1965) and the short on-times seen for lane changing would correspond to higher speed estimates. Category A in Tables 5-5, 5-6 and 5-7 corresponds to the slow speed measurements whereas Category C corresponds to the higher speed measurements. Both Categories have few samples (lower ω_j) and the variance in Category C is generally higher than those in other cases. Therefore, Category B is adopted as the primary speed distribution and the majority of the data will be used as the reported speeds for the dual loop, M loop, and S loop speed analyses.

Table 5-5 Study Site 1 Dual Loop Speed Distributions

	Dual Loop Before			Dual Loop After		
	Category A	Category B	Category C	Category A	Category B	Category C
ω_j	0.003	0.985	0.012	0.003	0.987	0.010
μ_j	47.84	63.19	85.30	46.29	63.67	88.92
σ_j^2	13.86	18.92	1336.89	10.83	19.51	1468.40

Table 5-6 Study Site 1 M Loop Speed Distributions

	M Loop Before			M Loop After		
	Category A	Category B	Category C	Category A	Category B	Category C
ω_j	0.04	0.91	0.05	0.08	0.91	0.02
μ_j	46.67	72.34	64.72	43.00	64.45	83.77
σ_j^2	9.91	42.61	1190.01	164.97	33.35	878.40

Table 5-7 Study Site 1 S Loop Speed Distributions

	S Loop Before			S Loop After		
	Category A	Category B	Category C	Category A	Category B	Category C
ω_j	0.01	0.95	0.04	0.02	0.96	0.02
μ_j	18.30	73.37	75.54	33.81	64.37	87.59
σ_j^2	3.53	46.40	1178.75	234.42	36.79	1224.99

Table 5-8 shows the Category B mean speeds from Tables 5-5, 5-6 and 5-7 and the error percentage, defined as the percent difference between those mean speeds and ground truth data (the corrected dual loop speed mean was used). As indicated in the "Dual loop" column, the low difference in sensitivity levels between the M and S loops was having a minimal effect on the dual loop speed calculation as evidenced by the 0.75% change before and after correction. The M and S loop speed estimations reported 13.62% and 15.23% differences before correction. The error dropped to just over 1% for both loops after correction. This represents a clear improvement and successful correction accomplished by the proposed method.

Table 5-8 Study Site 1 Speed comparison

	Dual Loop		M Loop		S Loop	
	Before	After	Before	After	Before	After
Short Vehicle Speed (mph)	63.19	63.67	72.34	64.45	73.37	64.37
Error (%)	0.75	--	13.62	1.23	15.23	1.10

Error: the difference between the (corrected) measurement and the ground truth data (Corrected dual loop speed data)

5.2.4 Study Site 1 Discussion

The changes to speed and length measurements are not the only benefits brought about by sensitivity correction. The WSDOT dual loop algorithm compares the on-times of the two ILDs to decide if the measurements of a vehicle can proceed to length calculation. When the on-time difference is over 10%, the algorithm assumes that the discrepancy is the result of a bad reading or a lane-changing movement and throws out the measurements from vehicle length calculation and length-based vehicle classification (Cheevarunothai et al. 2006). This has been identified as a major cause of the low bin volume sum when compared with single loop counted volume at dual-loop stations in Washington (Cheevarunothai et al. 2007). With minor changes to either the control logic or the algorithms, the on-time discrepancies can be diminished through length-based correction. Therefore, the classified bin volume data of a dual-loop could also benefit from the sensitivity corrections.

5.3 Case Study Two

Study site 2, Cabinet 005es15996, is located on a very straight section of I-5 with a small upward incline in the northbound direction. The directional AADT at this location is also approximately 120,000. The lane 4 detector is located in the inner most general purpose lane.

As shown in Figure 5-1, the dual loop detector at Study Site 1 showed a very good agreement between the M and S loop sensitivity levels. However, this is not the case at Study Site 2. Figure 5-2 shows the on-times of each dual loop data pair for this site. This dual loop detector here has a far greater sensitivity problem than Study Site 1 as evidenced by the best fit line y=0.4746x + 327.2 and R^2 value of 0.1345. Notice that few points lie on the 45° diagonal line, which indicates equal on-times measured by the M and S loops. This implies that the two ILDs are operating at different sensitivity levels. The expectation for this dual loop detector is that with the M loop under sensitive and the S loop over sensitive, the true distance between the detection zones, L' in Equation 5-4, will be less than *L* because d_m would be negative and d_s would be positive. With *L* being greater than *L'*, the dual loop detector would be expected to overestimate vehicle speeds.

The next important detail to take note of is the cluster of points where the M loop is registering on-times of near 200 ms while the S loop is recording value around 1000 ms. This would tend to indicate that the M loop sensitivity level is low enough that the ILD is only partially detecting trucks as discussed at the end of Chapter 4. This ILD's sensitivity level would seem to be near threshold B as described in Figure 4-23. This dual loop detector can be corrected only to a certain extent, not completely due to the lost information.

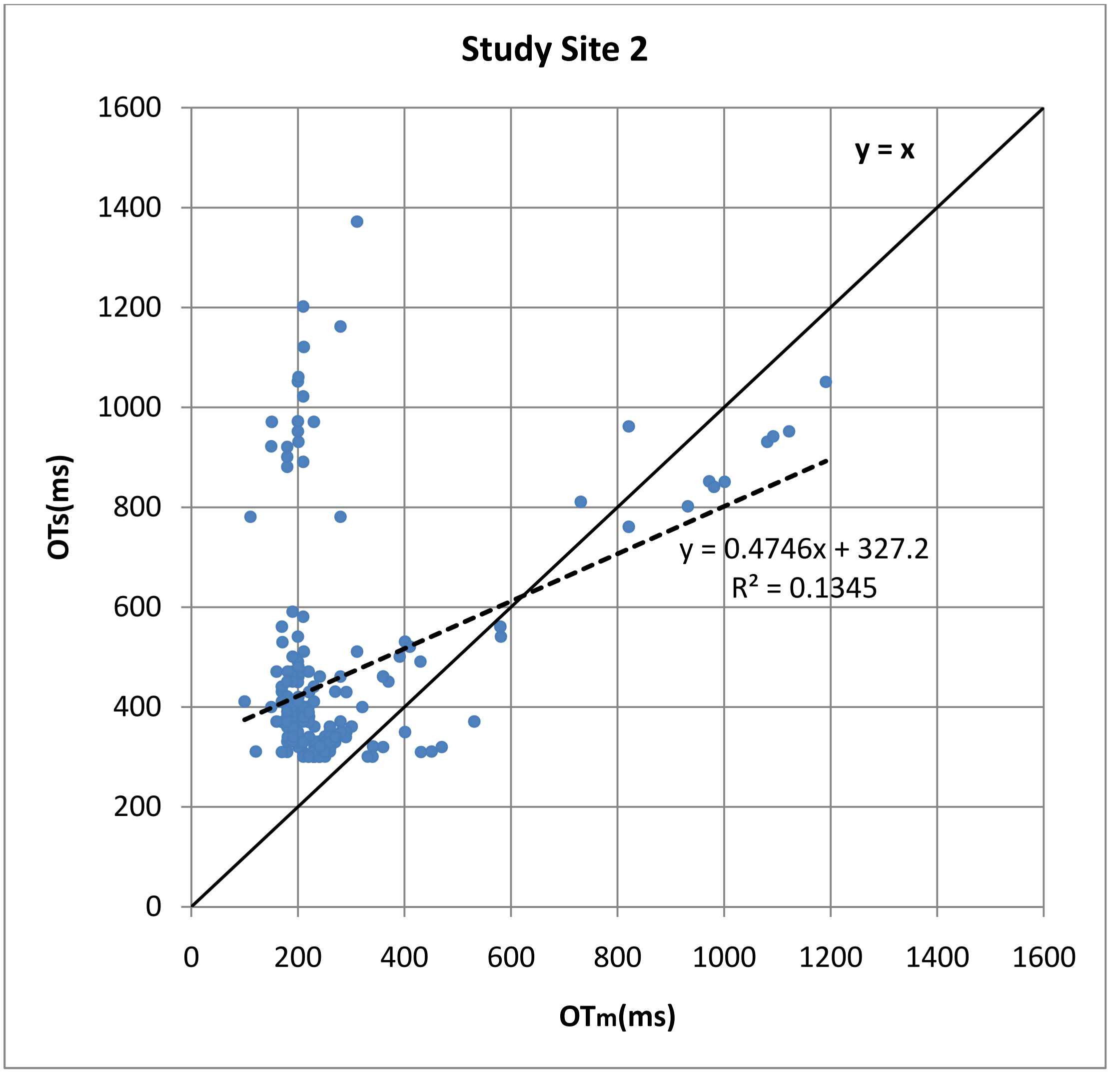


Figure 5-2 Paired On-times for North Lane 4 M and S Loops at Cabinet 005es15996

5.3.1 Free Flow Speed Estimation

Continuing with the correction will require the selection of a free flow speed. Compared to Case Study 1, this dual loop detector is likely not reporting accurate speeds and there is no historical data to reference, an appropriate free flow speed estimation for this location requires a re-examination of Equation 5-7. When the equation is considered with respect to the detection zone distance offset, d, the arrival time difference will be proportional to the difference between d_m and d_s and to the difference in on-times between the M and S loops. After combining Equation 5-2 and 5-7, this dual loop speed can be calculated as:

$$v_{dl} = \frac{L}{\Delta t_a + \frac{\Delta \mu_1}{2}} \tag{5-17}$$

where $\Delta\mu_1$ is the difference in Category 1 mean on-times between the M and S loops. Specifically, $\Delta\mu_1$, is defined as the S loop mean on-time minus the M loop mean on-time. This equation emulates the situation where the ILDs are at the same sensitivity level which would correct the speed measurement of the dual loop. When this technique is applied to Study Site 2, the resulting median speed is 65.1 mph. With this free flow speed, constraint 3 can be calculated as in Equations 5-18 and 5-19.

$$\frac{15.2\,ft+6ft-2(1.06\,ft)}{65\,mph} < \mu_1 < \frac{15.2\,ft+6ft+2(1.06\,ft)}{65\,mph} \tag{5-18}$$

$$200\,ms\ <\ \mu_1 < 245\,ms \tag{5-19}$$

In this case, the M loop has the μ_1 equal to 195 ms, which is under sensitive as defined by constraint 3. The S loop is also in error with its μ_1 of 251 ms indicating that it is over sensitive according to constraint 3.

5.3.2 Single and Dual Loop Correction

M Loop Correction

The M loop is under sensitive according to Equation 5-19. The calculation of *d* proceeds in a manner identical to the calculations for Study Site 1. To correct the M loop the *d* value must be found, as in Equation 5-20.

$$d_m = \frac{1}{2}(\mu_1 \cdot v_f - \overline{L_V} - L_L) = -1.30\text{ft} \tag{5-20}$$

The negative value for *d* indicates that this ILD is under sensitive.

S Loop Correction

Correction of the S loop proceeds in the same manner as the M loop with calculation of the *d* value as seen in Equation 5-21.

$$d_s = \frac{1}{2}(\mu_1 \cdot v_f - \overline{L_V} - L_L) = 1.36\text{ ft} \tag{5-21}$$

The positive value of *d* indicates that the S loop is over sensitive.

Dual loop Correction

With d_m and d_s calculated, the new distance between the M and S loops can be calculated as in Equation 5-22.

$$\frac{L + d_m - d_s}{\Delta t} = \frac{17\,ft - 1.30\,ft - 1.36\,ft}{\Delta t} = \frac{14.34\,ft}{\Delta t} = v \quad (5\text{-}22)$$

This time *L'* is significantly different from L. The new distance between the loop detection zones is 14.34 feet as compared to the assumption of 17 feet. This change in length represents a 16% reduction compared to the installation distance. The change in effective distance between the ILDs is directly proportional to the change in average speed.

5.3.3 Study Site 2 Correction Validation

5.3.3.1 Dual-Loop Vehicle Length Validation

As in Case Study 1, the GMM method was also applied to vehicle lengths data. Table 5-9 shows the Category 1 mean lengths and GMM parameters calculated using the M and S loop on-times. The M loop shows its under sensitivity with a vehicle length mean of 12.87 feet before the correction. The S loop similarly shows its over sensitivity with its vehicle length mean of 18.32 feet. Both values are remarkably different than the general short vehicle length (15.18 feet).

Table 5-9 Statistics for Dual Length Distributions for Category 1 (short vehicles) at Study Site 2

	Wang and Nihan (2000)	Using M loop On-times		Using S loop On-times	
		Before	After	Before	After
ω_j	NA	0.86	0.86	0.88	0.88
μ_j	15.18	12.87	15.49	18.32	15.60
σ_j^2	1.72	2.57	2.57	5.72	5.72

Table 5-10 has the t-statistics on the length distribution analyses. At this study site, vehicle lengths estimated from the M and S loop measured on-times showed significant difference before correction. The after correction differences are not statistically significant at the P= 0.05 significance level. The M and S loop length distributions are not brought into line with the Wang and Nihan (2000) distribution though. Even after correction, the M and S loop length measurements show statistically significant differences. The differences are not as significant as before correction though. This result indicates that vehicle length data from this dual loop detector can be improved by the correction methodology but that measurement errors will still be present.

Table 5-10 Hypothesis Test for Dual Loop Length Correction at Study Site 2

	Using the M Loop OTs vs. the S Loop OTs		Using the M Loop OTs vs. Ground-truth		Using the S Loop OTs vs. Ground-truth	
	Before	After	Before	After	Before	After
t value	83.59	1.66	47.23	6.34	49.31	6.60
p value	0.0001	0.0961	0.0001	0.0001	0.0001	0.001
Significant*	Y	N	Y	Y	Y	Y

* Significant at the p=0.05 significant level

5.3.3.2 Speed Correction

Table 5-11 shows the summary of speeds for the dual loop and single loop speed estimations. The affect of *L'* on the dual loop speed measurement is clear in the before and after for the dual loop detector data. The under and over sensitivity of the M and S loops is also clear in the table.

Table 5-11 Study Site 2 Speed Summary

	Dual Speed (mph)		M Loop Speed (mph)		S Loop Speed (mph)	
	Average	Median	Average	Median	Average	Median
Before	82.7	77.27	73.3	72.3	57.1	57.6
After	69.7	65.1	64.2	63.3	64.5	65.0

Tables 5-12, 5-13, and 5-14 show GMM parameters for a speed distribution analysis conducted on the dual loop detector, M loop, and S loop, respectively. The Category A values represent GMM parameters for the slower speed distribution and Category C values represent the high speed distribution. As discussed in Section 5.2.3 the Category A and Category C values are assumed to represent outlying speed values caused by vehicles which break the assumed conditions for speed and length calculations, i.e. the vehicle is significantly longer than expected, or the vehicle is changing lanes. Only the Category B values are used for the speed correction analysis.

Table 5-12 Study Site 2 Dual Loop Speed Distributions

	Before			After		
	Category A	Category B	Category C	Category A	Category B	Category C
ω_j	0.01	0.95	0.04	0.01	0.95	0.04
μ_j	37.88	78.44	183.35	33.21	66.14	161.56
σ_j^2	4.98	51.15	19643.55	11.92	36.54	14124.20

Table 5-13 Study Site 2 M Loop Speed Distributions

	Before			After		
	Category A	Category B	Category C	Category A	Category B	Category C
ω_j	0.07	0.82	0.10	0.07	0.83	0.10
μ_j	47.60	73.98	85.69	41.70	64.86	75.68
σ_j^2	8.77	59.97	1791.24	6.85	46.71	1429.82

Table 5-14 Study Site 2 S Loop Speed Distributions

	Before			After		
	Category A	Category B	Category C	Category A	Category B	Category C
ω_j	0.16	0.83	0.01	0.07	0.91	0.02
μ_j	48.45	57.85	148.99	38.22	65.30	108.86
σ_j^2	327.26	29.75	1214.49	187.26	44.73	2974.89

Table 5-15 summarizes the speed correction analysis for this study site. The before and after correction results are compared to the corrected dual loop speed measurement. For example, the dual loop speed before correction is 18.6% higher than after correction. The M and S loop speed estimates improve from about 12% difference to under 2% after correction. This would indicate that correction improved and even successfully corrected the speed measurements and estimations at this dual loop station, even though the length correction was not very satisfactory.

Table 5-15 Study Site 2 Speed comparison

	Dual Loop		M Loop		S Loop	
	Before	After	Before	After	Before	After
Bin1 Speed (mph)	78.44	66.14	73.98	64.86	57.85	65.30
Error (%)	18.60	--	11.85	1.94	12.53	1.27

Error: the difference between the (corrected) measurement and the ground truth data (Corrected dual loop data)

5.3.4 Study Site 2 Discussion

This dual loop detector is deeply affected by its sensitivity errors. Correction can improve flawed detectors such as these but there are definite limits to its effectiveness. Because the M loop is only partially detecting trucks, the correction will be limited at best. For this dual loop station length measurements should only be made at the S loop because the M loop is not correctly measuring truck occupancies. Even with the limited success of the correction, progress was made.

This dual loop detector is being corrected to improve its speed and length measurements. It is also being corrected as an example of the problems that will be encountered when the assumptions underlying this correction method begin to break down. This dual loop detector is obviously suffering from an M loop under sensitivity problem as described in Section 4.8 which will affect its ability to properly measure truck lengths and speeds.

5.4 Lessons Learned

Several factors influence the effectiveness of this correction process. The choice of thresholds for acceptable sensitivity errors may have the strongest influence on the success of this method. The goals of the correction should also be appraised. Speed measurement accuracy and length measurement accuracy may not change equally with correction. The results from study site 2 would seem to indicate that length measurement accuracy is lost before speed measurement accuracy. This is corroborated by the sensitivity analysis conducted in section 4.8. Therefore, the correction goals should be examined carefully with regards to threshold setting. If length estimation is the goal, tighter thresholds would be recommended. Speed estimation, by contrast, would appear to be far more robust for dual loop detectors. Also, for sites without dual loops, free flow speed must be carefully selected to ensure the best performance of the proposed approach.

Chapter 6 System Design

Currently, each of the transportation agencies independently formats and stores their traffic data. Even within a large transportation agency, offices at different regions may also have their own formats for data management and storage. This has led to a number of incompatibilities and institutional difficulties in data management. These distribution and format challenges pose significant barriers to data sharing between regions. The data management issues are particularly evident for analyses crossing regional borders where two or more regional data formats and data sources must be dealt with. These difficulties were driving forces for the development of Datamart, an online application for transportation data sharing, in this study. Therefore, the major goal for Datamart design and development is to facilitate online data distribution among transportation agencies, regional WSDOT offices, and universities.

Data acquisition and management is fundamental for Datamart. With thousands of traffic detectors deployed in the central Puget Sound region, a huge amount of traffic data flows in daily for Datamart to handle. To facilitate data quality control and support potential web-based applications and cross-agency data sharing, traffic data from different sources must be standardized in storage format. This can be enhanced in the traffic sensor database design.

Data quality is another issue for Datamart to address. As mentioned earlier, ILDs may subject to different types of malfunctions, particularly incorrect sensitivity level problems. Zhang et al. (2003) have found that large numbers of the WSDOT's ILDs have sensitivity problems. This systemic data quality problem has driven the interest in ILD error detection and correction for this research. The ILD sensitivity error identification and correction methods developed in this study will be implemented in Datamart to ensure the quality of ILD data. Because this method is software-based, both the original sensor raw data and the corrected data can be preserved. The original sensor raw data can still be used for trend analysis with the previously stored data without sensitivity correction.

Finally, Datamart will offer an online application for data downloading. All traffic sensor stations will be presented on a regional map. Users can simply click the sensor icon to activate a data download interface. Compared to the old sensor data acquisition systems that require sensor code and location to retrieve data, Datamart provides a much better interface for users to access and download data.

The overall Datamart architecture is shown in Figure 6-1. It is composed of three distinct subsystems: data acquisition, error detection and correction, and Web based distribution. The three components form a complete chain from data providers (currently only the WSDOT is included) to the online data sharing website. Details of each component are described in the following subsections.

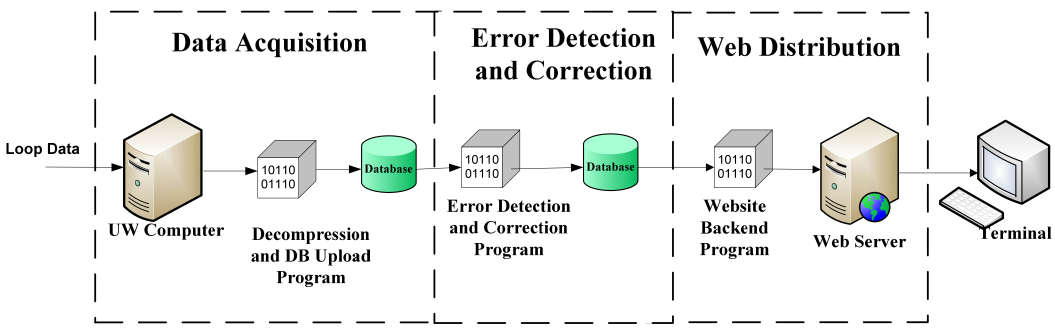


Figure 6-1 Datamart System Diagram

6.1 Data Acquisition

Datamart was designed to operate a data warehouse for the collection, correction, and distribution of traffic sensor data. At the current stage, only WSDOT 20-second ILD data are included. The data acquisition subsystem includes two components: 20-second data retrieval and storage.

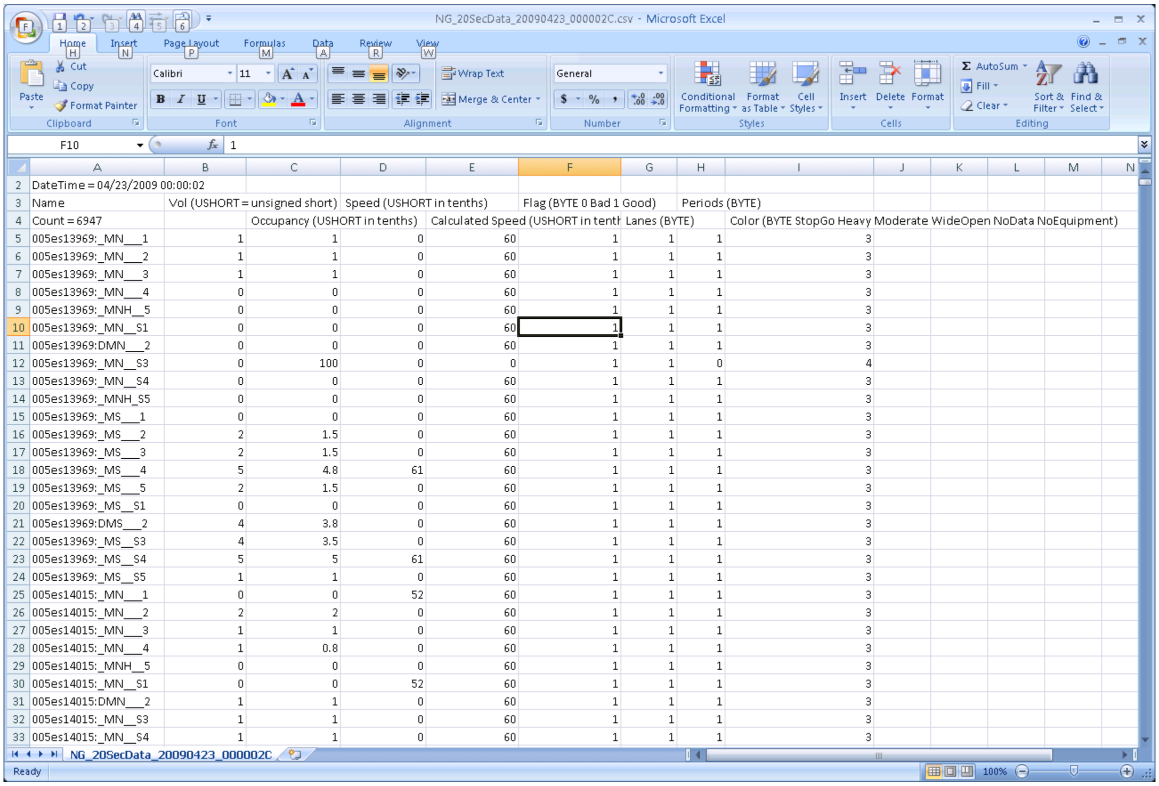
6.1.1 20-Second Data Retrieval

Aggregated data, either 20-second or 30-second, is the highest resolution data commonly archived by transportation agencies for transportation researchers and practitioners. At WSDOT the 20-second data is only held long enough to aggregate into the longer five minute periods used by planners. With increasing interest in operational projects such as

ATIS and ATMS has come an increase in WSDOT's interest in storing 20-second data for long term use.

In order to retrieve and archive 20-second loop data before they are discarded, a computer program dedicated to downloading these data was developed by the STAR Lab using Microsoft Visual C#. This program is hosted in a STAR Lab server computer and scans the WSDOT FTP site for loop data dissemination periodically for downloading any files that have not been generated by the WSDOT data server.

There are two separate files to download for the twenty second data. One file is a compressed information file that contains the actual loop information. The file has a data extension, .dat, and is generally around eighty kilobytes in size. The .dat files are available for download from the ftp site for approximately two days. The second file is in the Comma Separated Variable (CSV) format, file extension .csv, and contains the repeating data associated with the loops. Repeating data include loop name, loop location name, milepost, loop latitude, loop longitude and the number of lanes associated with the loop or station entry. The CSV file is generally about 450 kilobytes. The CSV files are replaced or updated, as necessary, whenever something changes in the system. The acquisition of 20-second data is an ongoing process and each day approximately 30 million rows are added to the database. The download program then decodes the .dat file and creates a new CSV file that combines the data from the .dat and .csv files. An example of the new file is shown in Figure 6-2



	A	B	C	D	E	F	G	H	I
2	DateTime = 04/23/2009 00:00:02								
3	Name	Vol (USHORT = unsigned short)		Speed (USHORT in tenths)		Flag (BYTE 0 Bad 1 Good)		Periods (BYTE)	
4	Count = 6947		Occupancy (USHORT in tenths)		Calculated Speed (USHORT in tenth	Lanes (BYTE)		Color (BYTE StopGo Heavy Moderate WideOpen NoData NoEquipment)	
5	005es13969:_MN___1	1	1	0	60	1	1	1	3
6	005es13969:_MN___2	1	1	0	60	1	1	1	3
7	005es13969:_MN___3	1	1	0	60	1	1	1	3
8	005es13969:_MN___4	0	0	0	60	1	1	1	3
9	005es13969:_MNH__5	0	0	0	60	1	1	1	3
10	005es13969:_MN__S1	0	0	0	60	1	1	1	3
11	005es13969:DMN___2	0	0	0	60	1	1	1	3
12	005es13969:_MN__S3	0	100	0	0	1	1	0	4
13	005es13969:_MN__S4	0	0	0	60	1	1	1	3
14	005es13969:_MNH_S5	0	0	0	60	1	1	1	3
15	005es13969:_MS___1	0	0	0	60	1	1	1	3
16	005es13969:_MS___2	2	1.5	0	60	1	1	1	3
17	005es13969:_MS___3	2	1.5	0	60	1	1	1	3
18	005es13969:_MS___4	5	4.8	61	60	1	1	1	3
19	005es13969:_MS___5	2	1.5	0	60	1	1	1	3
20	005es13969:_MS__S1	0	0	0	60	1	1	1	3
21	005es13969:DMS___2	4	3.8	0	60	1	1	1	3
22	005es13969:_MS__S3	4	3.5	0	60	1	1	1	3
23	005es13969:_MS__S4	5	5	61	60	1	1	1	3
24	005es13969:_MS__S5	1	1	0	60	1	1	1	3
25	005es14015:_MN___1	0	0	52	60	1	1	1	3
26	005es14015:_MN___2	2	2	0	60	1	1	1	3
27	005es14015:_MN___3	1	1	0	60	1	1	1	3
28	005es14015:_MN___4	1	0.8	0	60	1	1	1	3
29	005es14015:_MNH__5	0	0	0	60	1	1	1	3
30	005es14015:_MN__S1	0	0	52	60	1	1	1	3
31	005es14015:DMN___2	1	1	0	60	1	1	1	3
32	005es14015:_MN__S3	1	1	0	60	1	1	1	3
33	005es14015:_MN__S4	1	1	0	60	1	1	1	3

Figure 6-2 Decompressed .dat File

6.1.2 20-Second Data Storage

The database for the traffic sensor data storage required flexibility in design and a great deal of attention to space saving features. The database design must support the multiple data formats used by the various traffic data providers or transportation agencies. The sheer volume of data was also a major design criterion. WSDOT's Northwest Region alone has over 7000 ILDs reporting to it. When all of the regional data is combined the data volume is immense. Therefore very space efficient formats and designs were required to handle the data.

Source Data and Hardware

The data entry process begins after downloading both files described in the 20-second data description. A program runs an executable file provided by WSDOT to create a new CSV file which combines the repeating data from the CSV file and the per interval data from the data file. This CSV file forms a complete record for each loop that reported in during that particular twenty-second interval. Typically, there are about 7000 loop reports during any given interval of twenty-second for Northwest Region. Combined with the fact that twenty-second intervals equate to 4320 intervals per day, a total of approximately thirty million data entries are created every day for Northwest Region alone. After the CSV file is created the program uploads the data into the database.

The server computer hosting the database is equipped with two Intel Xeon 5520 (4C/8T) Central Processing Units (CPUs). Its operating system is Microsoft Windows Server 2008. Other details of the server hardware can be found in Table 6-1. We can see that there are 3 RAID arrays included in the server. The reason for this 3 RAID array configuration will be explained in Subsection 6.1.2.4.

The communication between the STAR Lab server and the WSDOT FTP website is via the broad band Internet service. The database system used for this study is Microsoft SQL Server 2008 Enterprise Edition.

Table 6-1 Server Specifications

Component	Number	Type
CPU	2	Intel Xeon 5520 (4C/8T)
Hard Drive Array 1	2	250 GB RAID 1
Hard Drive Array 2	4	250 GB RAID 1 + 0
Hard Drive Array 3	8+1	1 TB RAID 5 + Hot Spare
OS	---	MS Windows Server 2008

Database Design Issues

For functionality reasons, it is desirable that data be uniquely identified in the database. If data is not uniquely identifiable, confusion can result from multiple returns in place of a single answer. Fortunately, the loop data has a unique set of identifying information, or keys. The loop's name and interval timestamp uniquely identify a single record.

Another consideration in database design is how to minimize the size of the database. A given row of the loop data consumes sixty-one bytes using the default values seen in Tables 6-2 and 6-3. The loop name is eighteen characters long, and, therefore, eighteen bytes of space when stored in the char format. When speed, calculated speed, and occupancy are stored in decimal format, they consume five bytes each. Volume, lane count, flag, interval periods and color default to integer types which consume four bytes each. The total of sixty-one bytes is reached when the eight bytes needed for the timestamp's datetime format are included. When thirty million rows are added each day the per row space requirements become very important. Sixty-one bytes per row will end up requiring approximately 1.7 GB per day. It is important to note that the total would be 1.7 GB per day for Northwest Region's single loop detector data alone. There are several other regions to gather data from in addition to the planned inclusion of Northwest Region's dual loop detector data.

Table 6-2 MS SQL Data Type Space Requirements

Format	Size (byte)
integer (int)	4
smallint	2
tinyint	1
decimal (5,1)	5
decimal(8,4)	5
char(X)	X
datetime	8

Table 6-3 Default Data Table Size

Attribute	Format	Total Size	Key	Nullable
Loop Name	char(18)	18	X	
Timestamp	datetime	26	X	
Volume	int	30		
Occupancy	decimal(5,1)	35		
Speed	decimal(5,1)	40		X
Calcspeed	decimal(5,1)	45		X
Lane Count	int	49		X
Flag	int	53		X
Periods	int	57		X
Color	int	61		X

Space Saving Measures

To reduce space requirements, a number of steps were taken. The first was to create a lookup table to replace the loop name with a small integer which consumes two bytes instead of eighteen. Reducing the integer values from default four-byte integers to the smaller smallint and tinyint values results in a total savings of fourteen out of twenty-four bytes. Similarly, the decimal values can be multiplied by ten and then converted to smallints saving three bytes each. The total savings is thirty-nine bytes per row. The final total storage required is twenty-two bytes per row as can be seen in Table 6-4, or approximately six hundred megabytes per day.

Table 6-4 Final Data Table Design

Attribute	Format	Total Size	Key	Nullable
Loop ID	smallint	2	X	
YYYYMMDD	int	6	X	
HHMMSS	int	10	X	
Volume	smallint	12		
Occupancy	smallint	14		
Speed	smallint	16		X
Calcspeed	smallint	18		X
Flag	tinyint	19		X
Periods	tinyint	20		X
Lane Count	tinyint	21		X
Color	tinyint	22		X

Table 6-5 Lookup Table

Attribute	Format	Total Size	Key	Nullable
Loop Name	char(18)	18	X	
Loop ID	smallint	20		
Location	char(20)	40		X
Latitude	decimal(8,4)	45		X
Longitude	decimal(8,4)	50		X
Cabinet	char(10)	60		
Loop Type	char(6)	66		X
Milepost	decimal(8,2)	71		
Route	char(3)	75		

The lookup table, shown in Table 6-5, stores the loop name in connection with the integer replacing it in the bulk storage database. Once the lookup table was created, it made sense to move the metadata from the CSV file into the lookup table. Because the lookup table only stores one entry per loop, it has no pressing need to be particularly thrifty with space and can accommodate additional data such as the latitude and longitude far more efficiently than the bulk database. Finally, the detection zone distance offset factor, *d*, must also be stored in a table. With the d value, the original loop measurements can be easily corrected when needed to generate the corrected loop measurements. Because *d* may vary over time it is necessary for the entries in the table to contain both a beginning and ending time. The *d* table schema is

d_value_table(<u>Loop_ID, Ben_YYYYMMDD</u>, d, End_YYYYMMDD)

Following the relational database design convention in the Entity/Relationship Diagram approach, the three attributes that jointly serve as the key for this table are underlined. Data types for all attributes are listed in Table 6-6.

Table 6-6 d Value Table

Attribute	Format	Total Size	Key	Nullable
Loop ID	smallint	2	X	
Beg YYYYMMDD	int	10	X	
d	decimal(5,1)	15		
End YYYYMMDD	int	23		X

Changing the datetime value to two integer values was a key design choice. The datetime data type is only precise to 3.333 milliseconds which means that a particular time interval recorded in seconds with perfect precision would be stored in datetime format as the time interval ±3.333 milliseconds. The range of actual input values and lack of precision makes querying for a specific record based on time very difficult as the desired record might be at 14:22:02 but the recorded datetime format value in the table could be between 14:22:01.997 and 14:22:02.003 and might not exactly equal 14:22:02 so the query would return no result. By using an integer value for the date and another separate integer value for the time, there is no net savings in space but there is a savings in the complexity of the query required to retrieve the data. With two integers replacing the datetime value there is no more uncertainty as to what the exact value recorded in the table is. Particularly helpful is the simplification of queries for the same time period over multiple days.

Performance Increasing Measures

The next stage in database design involves selecting a design that will impose the fewest performance penalties and simplify access to the greatest degree possible. As loop data is keyed by loop ID and timestamp, these attributes are jointly used as the primary key. The next challenge is the sheer span of time that the database is expected to serve. Over time, the database will grow and consume more space and have a larger index. The increase in database size will decrease database performance and should be controlled to the extent possible. Therefore, the primary loop data table is partitioned by month so that users can quickly access the data regardless of the current size of the database. Each month's data is stored in a separate file group as required by the partitioning system and each file group is composed of two files. By creating a second file in each group, performance is increased during write operations. The performance increase is grounded in the ability to expand one file while adding data to empty pages in the other. This way the system does not have to stop and expand a file before continuing with the write operation. Partitioning also splits the index file and controls its growth.

The computer specifications listed in Table 6-1 include 3 RAID arrays. The first is a mirrored array for the operating system. This isolates all of the operating system hard drive access traffic from the drives used by the database. The second array is used for log files created by the database software. The logic is the same as for isolating the operating system, when log files are created during database operation they are created on separate disks from the database preventing the situation where the database disks must constantly switch between data and log files, which can substantially decrease performance. Finally, the large disk array for the database is composed of multiple disks to increase the reading and writing speeds. As the number of hard drives increases there are more individual disks capable of sending data to the processors, increasing throughput.

6.2 Error Detection and Correction

The methodology for error detection and correction are detailed in chapters four and five. This subsystem implements this method to automate the error detection and correction process. The error detection portion of the system looks at the distributions and properties of the ILDs to determine if the data being output is reasonable or in error. If the data is flawed then, depending on the error type the ILD will be subject to correction or marked as requiring a hardware tune up. More detailed descriptions of the functions may be found in chapters four and five.

The error detection and correction subsystem interfaces with the database in the course of executing its functions. This subsystem is developed using Microsoft Visual C#. It interfaces with the database via an Open Database Connection (ODBC). ODBC allows C# to pass queries to Microsoft SQL Server 2008 and read the query result back into the program. ODBC allows the program to operate on one computer and access the Microsoft SQL Server across a network. This functionality allows multiple copies of the program to run on several computers and access the same server for data. Given the large amount of data, parallelization can greatly increase the speed of processing. A similar program was used to analyze the event data from logged files for error diagnostics in this study.

6.3 Web Data Distribution

The Datamart Project is intended to distribute data throughout world to whatever researchers and agencies who are granted access by the data providing agencies and the STAR Lab. The simplest and most cost effective method of accomplishing this task is to create a website and tie it into the database. The web site should also allow users to select individual ILDs and time periods for which to download the data so that large downloads of undesired data are not forced on end users. Targeted downloading also reduces the strain placed on the server resources hosting the database and web site.

With these considerations in mind the web site utilizes Enterprise Edition Java (J2EE) (http://java.sun.com/javaee/), the Google Web Toolkit (GWT) (http://code.google.com/webtoolkit/), and Apache Tomcat (http://tomcat.apache.org/) to select and distribute data from the database. The primary component is the J2EE program residing on the server which handles the querying and packing of the data for the web site. This program takes input from the website to convert into queries for the database as shown in Figure 6-3.

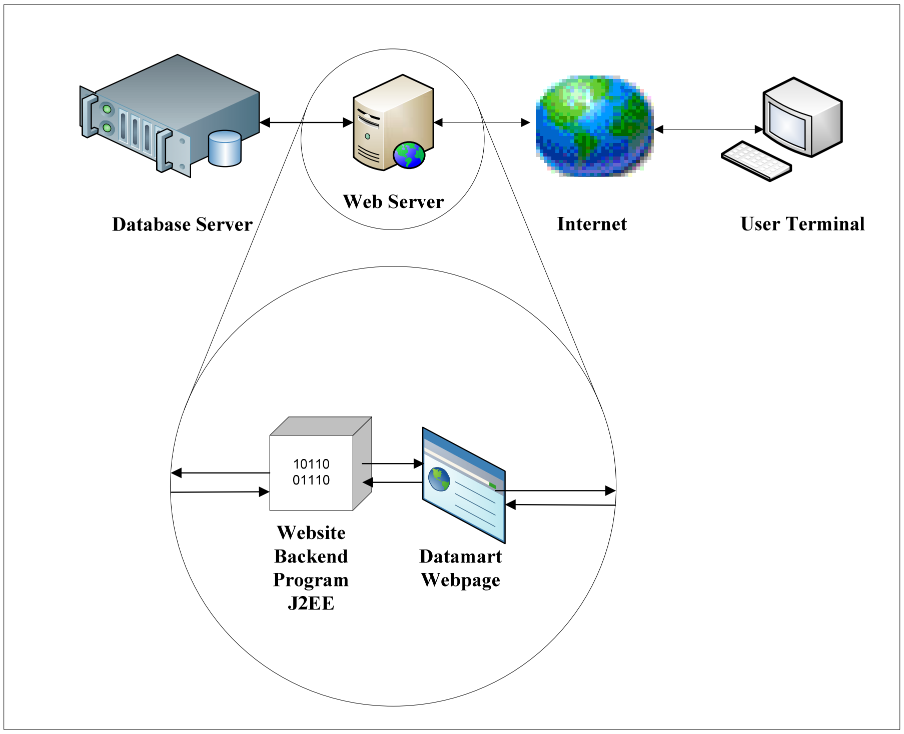


Figure 6-3 Web Data Distribution Detail Schematic

The GWT is designed to help programmers create Javascript web pages while the programming is done in J2EE or another Java variant. This allows a programmer to develop their Javascript web pages while working within a development environment that supports debugging. GWT and J2EE work in concert to make the inclusion of Google's maps and map based functionalities smoother. Apache Tomcat hosts the web page created with GWT and J2EE. Tomcat is a Java based HTTP server roughly analogous to Microsoft's Internet Information Services (IIS) and Apache HTTP Server. Tomcat was designed to be a purely Java implementation of a web server. It specializes in handling Java, Java Servlet and JavaServer Pages. With the chosen implementation of GWT and J2EE, Apache Tomcat is the logical choice for hosting the website.

Figure 6-4 shows a screen shot of the web site download function. Currently the web site is designed for use with Mozilla's Firefox (http://www.mozilla.org/) but additional browser support, particularly the newer versions of Microsoft's Internet Explorer, is in progress. The download function is behind password protection as our current data provider, WSDOT, stipulated. The WSDOT stipulation that the data be behind password protection is so that WSDOT can monitor data usage.

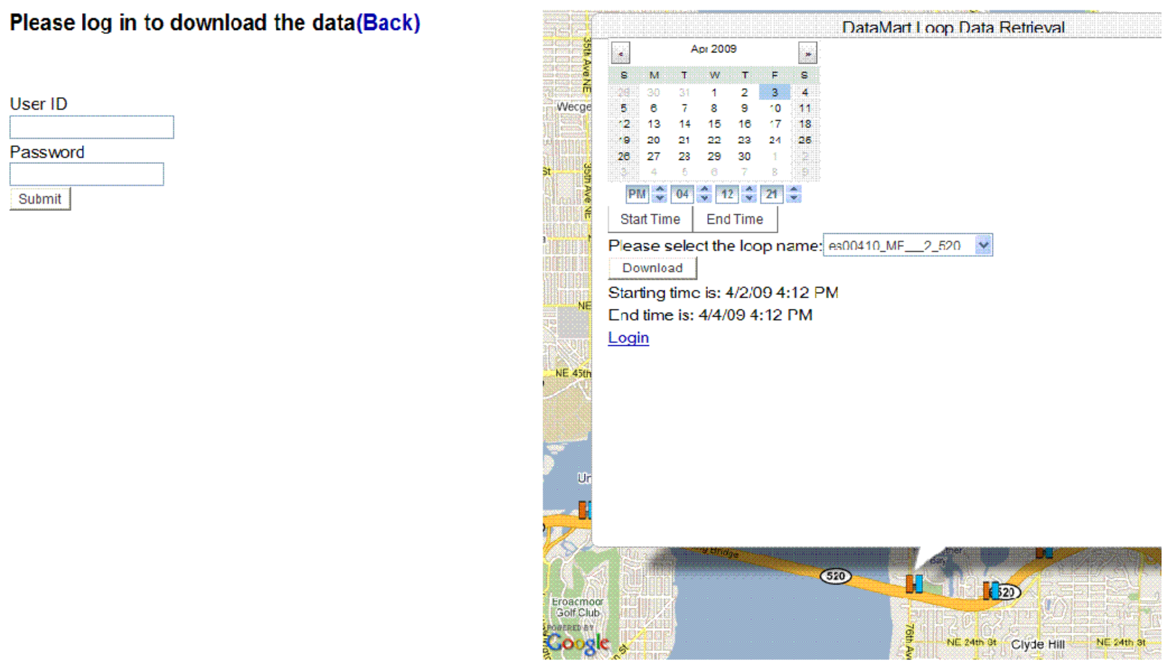


Figure 6-4 Datamart Project Website

Using the map based functionality, cabinets can be selected and data queried from within the web page. Data is downloaded as a CSV file. Users will decide on the best method to handle the data after download so a common and non-proprietary format such as CSV, which can be opened by most analysis programs, is very platform independent and allows various users to access the data without the need to support multiple data formats.

The Datamart website has been in internal use for multiple months to test the functions and collect user feedback for further improvement. The application works fine and no severe bugs have been reported.

Chapter 7 Conclusions and Recommendations

7.1 Conclusions

This research targets improvements in traffic detector data quality, storage, and data sharing to make the WSDOT traffic sensor network gather statewide data across of higher quality and make that data more accessible and consistent. Specifically, algorithms for identifying and correcting the loop sensitivity problems have been developed and implemented in the Datamart system, an online data quality control and sharing system created by this study. Datamart utilizes a group of the most advanced technologies for online application development, including Enterprise Edition Java (J2EE), the Google Web Toolkit (GWT), and Apache Tomcat. It is highly scalable and has the potential to add new data sources from other transportation agencies and online analysis functions for regional transportation planning, traffic management, and analysis purposes.

20-aggregated data has been shown to produce results comparable to those obtained using event data. The ability to use aggregated data instead of event data allows practitioners and researchers to examine ILDs remotely instead of requiring the expenditures of effort associated with event data collection.

The error detection algorithms look for characteristics in the loop detector data distributions that correspond to three types of errors. The first type of error may be caused by extreme under sensitivity or pulse mode setting, among other causes. This error is characterized by a very narrow distribution at an unreasonably low on-time without the two peaks expected for trucks and short vehicles. The second error type occurs when short vehicles do not compose the expected single large majority of the length distribution observed by the ILD. The short vehicle distribution appears split into two separate distributions in this error type. This second type of error has an unknown cause and should be the subject of further research. Finally, ILDs outputting data which is reasonable according to the calculated threshold values need not be corrected so only ILDs reporting data beyond the expected bounds suffer from Type 3 errors. ILDs suffering from Type 1 or Type 2 errors should be corrected at the source hardware.

While the proposed correction methodology can attempt to correct these errors it will meet with only varying levels of success.

ILDs with only Type 3 errors are suitable candidates for correction by the detection zone offset method used in this research. The error correction algorithm focuses on ILD detection zone length correction. It performed well when applied to suitable ILDs. ILDs suffering from Type 3 error are producing incorrect data but not data that is so flawed as to actually lose detail needed to correct the data. This research applied the correction methodology to two dual loop stations.

Study Site 1 had a dual loop detector with two under sensitive ILDs. Both ILDs were under sensitive to nearly the same degree and showed strong agreement in their data. The speed correction was minimal as expected by the similar sensitivity levels. The length correction was much more meaningful and brought the length measurements from each ILD up from a Category 1 mean of ~12.5 feet to a much more reasonable 15.11 feet after correction. Length measurements made by each ILD differed significantly prior to correction but the differences after correction were not statistically significant.

Study Site 2 contained a dual loop with an ILD that was under sensitive and an ILD that was over sensitive. The under sensitive leading loop and over sensitive trailing loop changed the true distance from the leading edge of the M loop detection zone to the leading edge of the S loop detection zone significantly. This change shrank the true distance between detection zones by 16% and correspondingly increased the dual loop speed measurements calculated using the longer expected length. This dual loop detector had another major flaw, the M loop's under sensitivity was severe enough to begin incorrectly measuring truck on-times. While the proposed correction methodology did improve the quality of the data output by this dual loop detector, it shows the limits at which the proposed correction methodology will begin to break down.

The correction algorithm adequately improves the length and speed distributions as measured by the ILDs compared to the original data. The correction of the length discrepancy between ILDs of a dual loop detector is particularly useful as it will improve dual loop data quality. If the correction method can be implemented at the controller level, it may improve data quantity for some controller logics by reducing the number of records discarded due to length or on-time discrepancy. The improved data quality seen from single and dual loop detectors will aid future projects such as ATIS and ATMS. Finally, as long as *d* values can be calculated from the available data, it is possible to retroactively apply this correction method to historical data.

7.2 Recommendations

Based on the study results, the research team would like to make the following recommendations:

- The first and most important recommendation is that transportation agencies may consider implementing the error testing and correction algorithms as completely as possible throughout their data systems, since the proposed methodologies have proven effective at the two study sites where significant data quality improvements were seen.
- The sensitivity issues with regards to vehicle detection, particularly truck detection, are in need of further research. If trucks are being missed at detector stations due to sensitivity errors, that loss of data can have repercussions beyond traffic. Accurate detection of trucks is important for interstate commerce and federal regulations. Knowledge of sensitivity's affect on vehicle detection may also open the way to new methods of error detection and correction.
- The cause for Type 2 errors seen in Chapter 4 is not fully understood at present. This situation needs to be remedied if possible because Type 2 errors were relatively common over the course of this research.

References

(1) Athol, P. "Interdependence of Certain Operational Characteristics Within a Moving Traffic Stream." Highway Research Record 72. Pg. 58-87. HRB. National Research Council. Washington, D.C. 1961.

(2) Bhagat, V. and D. Woods. "Loop Detector Crosstalk." *ITE Journal* Vol. 67, No. 2, Institute of Transportation Engineers, 1997, pp. 36-37, 47-49.

(3) Cheevarunothai, P., Y. Wang, and N. L. Nihan. "Development of Advanced Loop Event Data Analyzer (ALEDA) for Investigations of Dual-Loop Detector Malfunctions." Proceedings of the 12th World Congress of Intelligent Transportation Systems, paper 000662, November 2005.

(4) Cheevarunothai, P., Y. Wang, and N. L. Nihan. "Identification and Correction of Dual-Loop Sensitivity Problems". In *Transportation Research Record: Journal of the Transportation Research Board, No. 1945"*, TRB, National Research Council, Washington, D.C., 2006, pp. 73-81.

(5) Cheevarunothai, P., Y. Wang, and N. L. Nihan. "Using Dual-Loop Event Data to Enhance Truck Data Accuracy". In *Transportation Research Record: Journal of the Transportation Research Board, No. 1993*, TRB, National Research Council, Washington, D.C., 2007, pp. 131-137.

(6) Cheevarunothai, P. and Y. Wang. "A Practical Algorithm for Identifying and Correcting Single-Loop Sensitivity Problems". CD-ROM. Transportation Research Board of the National Academies, Washington, D.C., 2009

(7) Chen, L., and A. May. "Traffic detector errors and diagnostics". In *Transportation Research Record: Journal of the Transportation Research Board, No. 1132*, TRB, National Research Council, Washington, D.C., 1987, pp. 82–93.

(8) Cleghorn, D., F. L. Hall, and D. Garbuio. "Improved Data Screening Techniques for Freeway Traffic Management Systems", In *Transportation Research Record: Journal of the Transportation Research Board, No. 1320*, TRB, National Research Council, Washington, D.C., 1991, pp. 17-23.

(9) Coifman, B. "Using Dual Loop Speed Traps to Identify Detector Errors", Transportation Research Record No. 1683, Transportation Research Board, 1999, pp 47-58.

(10) Coifman, B., S. Dhoorjaty, , and Z. Lee, "Estimating Median Velocity Instead of Mean Velocity at Single Loop Detectors", Transportation Research: Part C, vol 11, no 3-4, 2003, pp 211-222.

(11) Coifman, B. and S.Dhoorjaty , "Event Data Based Traffic Detector Validation Tests", ASCE Journal of Transportation Engineering, Vol 130, No 3, 2004, pp 313-321.

(12) Dailey, D., "A statistical algorithm for estimating speed from single loop volume and occupancy measurements". *Transportation Research B* **33**, 1999, pp. 313–322.

(13) Hasan, B. A. S. and J. Q. Gan. *Sequential EM for Unsupervised Adaptive Gaussian Mixture Model Based Classifier.* Machine Learning and Data Mining in Pattern Recognition: 6th International Conference. Leipzig, Germany. 2009.

(14) Highway Capacity Manual (HCM). TRB, National Research Council, Washington, D.C., 2000.

(15) Ingram, J. *The inductive loop vehicle detector: Installation acceptance criteria and maintenance techniques*, California Dept. of Transportation, Sacramento, CA. 1976.

(16) Jacobson, L. N., N. L. Nihan, and J. D. Bender. “Detecting Erroneous Loop Detector Data in a Freeway Traffic Management System”. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1287*, TRB, National Research Council, Washington, D.C., 1990, pp. 151-166.

(17) Ju, Z., H. Liu, X. Zhu, Y. Xiong. *Dynamic Grasp Recognition Using Real Time Clustering, Gaussian Mixture Models and Hidden Markov Models*. Intelligent Robotics and Applications. Wuhan, China. 2008.

(18) Klein, L., M K. Mills and D. R.P. Gibson. *Traffic Detector Handbook: Third Edition*. Publication FHWA-HRT-06-108. FHWA, U.S. Department of Transportation, 2006.

(19) Payne, H.J. and S. Thompson, “Malfunction Detection and Data Repair for Induction-Loop Sensors Using I-880 Data Base”, TRB, *Transportation Research Record, 1570*, 1997, p.191-201

(20) May, A. D., B. Coifman, R. Cayford and G. Merritt. *Automatic Diagnostics of Loop Detectors and the Data Collection System in the Berkeley Highway Lab*, California PATH Research Report, 2004.

(21) McLachlan, G.J. and K.E. Basford. *Mixture Models: Inference and Applications to Clustering*, Marcel Dekker, 1988.

(22) Model C-1000 Series Operation Manual. Reno A&E, Reno, Nevada, 2004

(23) Nihan, N. L., L.N. Jacobson, J.D. Bender, and G.A. Davis. *Detector Data Validity*. Washington State Transportation Center – TRAC/WSDOT, 1990. TransNow/WSDOT Joint Final Research Report No. WA-RD 208.1, 80 pages

(24) Nihan, N. L. ‘‘Aid to determining freeway metering rates and detecting loop errors.’’ *J. Transp. Eng.,* 123(6), 1997, pp. 454–458.

(25) Nihan, N. L., X. Zhang, and Y. Wang, “Evaluation of dual-loop data accuracy using video ground truth data.” *Research Report TNW2002-02*, University of Washington, Seattle, WA. 2002.

(26) Nihan, N.L., Y. Wang, and P. Cheevarunothai. “Improving Dual-Loop Truck (and Speed) Data: Quick Detection of Malfunctioning Loops and Calculation of Required Adjustments.” Research Report for Washington State Transportation Center (TRAC) and Washington State Department of Transportation (WSDOT), Agreement T2695, Task 57. 2006.

(27) Reynolds, D. A., and R. C. Rose. “Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models.” *IEEE Transactions On Speech and Audio Processing,* Vol. 3, No. 1, 1995.

(28) Ritchie, S.G., S. Park, C. Oh and C. Sun, “Field investigation of advanced vehicle reidentification techniques and detector technologies”. MOU 3008. Partners for Advanced Transit and Highways: Institute of Transportation Studies, University of California, Irvine. 2001.

(29) State of California Model 222 Inductive Loop Detector Sensor Unit Operations Manual. Eberle Design Inc., Phoenix, AZ, 2005.

(30) Turner, S., L. Albert, B. Gajewski, and W. Eisele. ”Archived Intelligent Transportation System Data Quality: Preliminary Analysis of San Antonio Transguide Data”. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1719*, TRB, National Research Council, Washington, D.C., 2000. pp. 77-84.

(31) Turochy, R.E. and B.L. Smith. “A New Procedure for Detector Data Screening in Traffic Management Systems”. In *Transportation Research Record* 1727, Transportation Research Board, Washington, D.C., 2000, pp. 127-131.

(32) Wall, Z.R. and D.J. Dailey. “Algorithm for Detecting and Correcting Errors in Archived Traffic Data”. In *Transportation Research Record 1855,* Transportation Research Board, 2003, pp. 183-190.

(33) Wang, Y., and N. L Nihan. "Freeway traffic speed estimation using single loop outputs." *Transportation Research Record 1727*, Transportation Research Board, Washington, D.C., 2000 pp.120–126.

(34) Wang, Y., and N. L. Nihan. “Can Single-Loop Detectors Do the Work of Dual-Loop Detectors?” ASCE Journal of Transportation Engineering, 129(2), 2003, pp. 169 - 176.

(35) Washington State Department of Transportation (WSDOT). 2008 Annual Traffic Report. 2008

(36) Wu, Y. J., Y. Wang, D. Qian. “A Google-Map-Based Arterial Traffic Information System”. Presented at 10th International IEEE Conference on International Transportation Systems, Seattle, Washington, USA, 2007.

(37) Yu, Runze, G. Zhang and Y. Wang. “Loop Detector Segmentation Error and Its Impacts on Traffic Speed Estimation.” Transportation Research Record, No. 2099, 2009. pp.50-57.

(38) Zhang, X., Y. Wang, and N. L. Nihan. “Investigating Dual-Loop Errors Using Video Ground-Truth Data”. *Proc., ITS America 2003 Annual Meeting*, Minneapolis, Minn., 2003.