AUTOMATED IDENTIFICATION OF TRAFFIC DETECTOR MALFUNCTIONS

Final Report

PROJECT SPR 837

Oregon Department of Transportation

AUTOMATED IDENTIFICATION OF TRAFFIC DETECTOR MALFUNCTIONS

Final Report

PROJECT SPR 837

by

Edward Smaglik, Ph.D., P.E., Professor, Katherine Riffle, Graduate Research Assistant
Steven Gehrke, Ph.D., Assistant Professor, Brendan Russo, Ph.D., P.E., Associate Professor
Steven Procaccio, Undergraduate Research Assistant,
Northern Arizona University

David Hurwitz, Ph.D., Professor, Yujun Liu, Graduate Research Assistant
Eileen Chai, Graduate Research Assistant
Oregon State University

Xiugang Li, Ph.D., P.E.
Oregon Department of Transportation

for

Oregon Department of Transportation
Research Section
555 13th Street NE, Suite 1
Salem OR 97301

and

Federal Highway Administration
1200 New Jersey Avenue SE
Washington, DC 20590

September 2022

Technical Report Documentation Page

1. Report No. FHWA-OR-RD-23-04	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Automated Identification of Traffic Detector Malfunctions		5. Report Date: September 2022 6. Performing Organization Code
7. Author(s) Edward Smaglik, Ph.D., P.E., 0000-0002-7034-6619 Katherine Riffle, 0000-0001-8302-0223 Steven Gehrke, Ph.D., 0000-0001-9355-5571 Brendan Russo, Ph.D., P.E., 0000-0002-0606-7973 Steven Procaccio, 0000-0003-0562-1503 David Hurwitz, Ph.D., 0000-0001-8450-6516 Yujun Liu, 0000-0001-6120-2907 Eileen Chai, 0000-0002-8093-0990 Xiugang Li, Ph.D., P.E., 0000-0002-5966-8158		8. Performing Organization Report No.
9. Performing Organization Name and Address Oregon Department of Transportation Research Section 555 13th Street NE, Suite 1 Salem, OR 97301		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Oregon Dept. of Transportation Research Section and Federal Highway Admin. 555 13th Street NE, Suite 1 1200 New Jersey Avenue SE Salem, OR 97301 Washington, DC 20590		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract: There is a need to improve signalized intersection operations through identifying malfunctioning detectors, as recent work has shown that errors in data quality and accuracy may be widespread due to issues with aging equipment and unmet maintenance needs. Accordingly, there is a desire for policies, procedures, and techniques to identify malfunctioning detection equipment and evaluate the quality of data produced by detectors. Current tools, including those available through newer Advanced Traffic Controller (ATC) standards, are able to detect major detector failures by examining the presence, absence, or frequency of data being sent by a detector, but these tools are not able to assess the quality of the information sent; therefore, the health of the detector is commonly unmonitored. To address this issue, using event-based data, researchers isolated saturated flow from individual presence detectors at signalized intersections, and using detector activations and occupancy, were able to calculate volume and density metrics for individual green intervals. These data points were then used to approximate the undersaturated portion of a Volume vs. Density fundamental diagram for detectors at various sites in Oregon. Using the mathematical concept of percent difference between integrals, several performance datasets were developed for this work for use in algorithms also developed as part of this work. These algorithms assess detector health, both initially and over time. Finally, a system design and implementation plan was developed to aid in the deployment of this system.

17. Key Words: Traffic Detector; Detector Health	18. Distribution Statement: Copies available from NTIS, and online at www.oregon.gov/ODOT/TD/TP_RES/		
19. Security Classification (of this report): Unclassified	20. Security Classification (of this page): Unclassified	21. No. of Pages: 125	22. Price

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	millimeters squared	mm^2
ft^2	square feet	0.093	meters squared	m^2
yd^2	square yards	0.836	meters squared	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	kilometers squared	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	ml
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	meters cubed	m^3
yd^3	cubic yards	0.765	meters cubed	m^3
~NOTE: Volumes greater than 1000 L shall be shown in m^3.				
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	Fahrenheit	(F-32)/1.8	Celsius	°C

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	millimeters squared	0.0016	square inches	in^2
m^2	meters squared	10.764	square feet	ft^2
m^2	meters squared	1.196	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	kilometers squared	0.386	square miles	mi^2
		VOLUME		
ml	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	meters cubed	35.315	cubic feet	ft^3
m^3	meters cubed	1.308	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams	1.102	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celsius	1.8C+32	Fahrenheit	°F

*SI is the symbol for the International System of Measurement

ACKNOWLEDGEMENTS

The authors would like to thank Oregon State University Undergraduate Research Assistants Alden Sova, Brandy Quach, and Elsa Moreno Rangel and Graduate Research Assistant Mckenna Milacek for their collective efforts transcribing drone videos collected in the field, Nakai McAddis for his work on the System Implementation and Design, and the members of the Technical Advisory Committee, as this project could not have been completed without their guidance and assistance. We especially thank Shawn Strasser of ODOT for the support.

DISCLAIMER

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

ODOT desires to improve signalized intersection operations through identifying malfunctioning detectors. Past research (Smaglik et al. 2017) has shown declining operational performance from invasive and non-invasive detection units at intersections across Oregon. Specifically, errors in data quality and accuracy showed widespread issues with aging equipment and unmet maintenance needs. Accordingly, there is a need for policies, procedures, and techniques to identify malfunctioning detection equipment and evaluate the quality of data produced by detectors.

Current tools, including those available through newer Advanced Traffic Controller (ATC) standards, are able to detect major detector failures by examining the presence, absence, or frequency of data being sent by a detector, but these tools are not able to assess the quality of the information sent; therefore, the health of the detector is commonly unmonitored. For example, detrimental detector behaviors at signalized intersections such as a loop that fails for 3 minutes and works for 1 minute may not send a phase into recall. This partial failure could go unnoticed leading to poor performance and potentially encourage unsafe driver behaviors such as disobedience of signal indications. Complete failure of a detection zone is identified, but if the detector is operating, it can be hard to discern the quality of the data provided. To address this issue, this project developed a technology agnostic detector health monitoring procedure that can be deployed to identify detection performance issues beyond complete detector failure. This research provides guidance for the action and implementation of detector health analysis as a low-cost option for identifying faulty infrastructure.

2.0 LITERATURE REVIEW

The objective of this literature review is to explore previous research relevant to the areas of detector performance, detector health monitoring, and traffic flow theory as it applies to detector operations.

2.1 DETECTION TECHNOLOGY

Outside of downtown grid networks, signalized intersections are typically operated with some type of actuation. The complexity of the actuated control algorithm is directly related to the vehicle detection required to effectively operate the control. With control algorithms ranging from legacy call and extend operation to complex traffic responsive and adaptive operations, detection requirements can vary from as simple as a presence detection zone to call a side street phase for service to an array of sensors covering a network tasked with delivering presence, count, and occupancy information.

Vehicle detection falls into two general categories, invasive technologies, those which are within the pavement, and non-invasive technologies, located outside of the roadway surface. Invasive sensors are commonly based upon inductive detection, taking the form of an in-pavement wire loop, preformed loop, small form factor loop (micro-loop), or wireless magnetometer. Non-invasive sensors vary in technology, including video, both visible and infrared, radar, and recently to the market, combination video and radar units. In-pavement wired loops have been deployed in vehicle sensing operations for fifty years, with wireless magnetometer units entering the marketplace a little more than a decade ago. Various non-invasive sources have been employed in assorted vehicle detection operations for more than twenty years. It is noted that, per the direction of the SPR 837 Technical Advisory Committee (TAC), inductive loop and radar technologies will be used to develop the algorithms in this work; as such, little focus will be given to other detection sources.

2.1.1 Inductive Loop Detector

Historically, inductive loop detection has been the most widely used sensor for vehicle detection (Day et al. 2011) and, when functioning properly, have been purported to be the most accurate detection technology available. Loop detectors are installed in the pavement at various points leading up to an intersection. Figure 2.1 shows an example schematic of a typical loop installation.

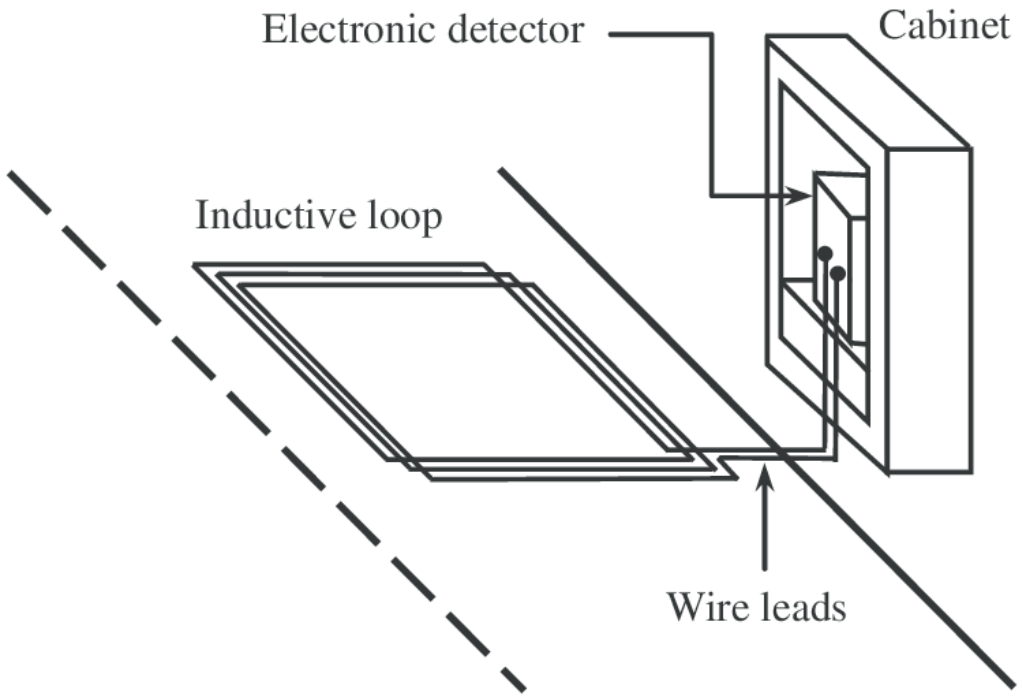


Figure 2.1: Wire inductive loop setup (Lamas-Seco et al. 2016)

Inductive loop detection has been used as a ground truth in a number of other detection performance evaluations (Day et al. 2010) (Rhodes, Bullock, and Sturdevant 2006) (Grossman et al. 2012), and using the performance characteristics of an inductive loop, the Indiana Department of Transportation (INDOT) developed detection performance specification (INDOT 2015) (Middleton et al. 2009) to address the issue of detector latency and other performance issues identified with non-invasive detection devices. Inductive loops are not without their challenges, however. Placing loops directly into the pavement can exacerbate pavement distress. While preformed loops placed under the surface course do not have this drawback, both types of installations are susceptible being compromised due to common in-ground hazards, including freeze/thaw cycling, vermin, and wayward construction equipment, all of which can cause performance degradation and impact detector health.

2.1.2 Radar Detection

Radar technology has been in use for the development of vehicle performance measures on freeway facilities for a number of years, however, only recently have products been brought to market to employ this technology at signalized intersections. Earlier units focused on advance detection only, avoiding the inherent challenge of detection vehicles at the stop line with a technology that uses object motion to operate. Researchers at the Texas A&M Transportation Institute (TTI) tested a unit in 2008 and found that the unit accounted for a 23-48% increase in phase termination over video detection (Middleton, Charara, and Longmire 2009). Research personnel at Purdue University noted that the use of this type of technology for advance detection has the potential to increase both efficiency and safety of dilemma zone protection

since it tracks the vehicle all the way through the detection zone as opposed to extrapolating from an advance speed trap (Sharma et al. 2008). These results were supported by (Hurwitz et al. 2012) who documented a reduced frequency of drivers captured in the type two dilemma zone when a wide area radar detection system was employed as compared to in-pavement loops. Another research group noted that the units recorded speed and volume values comparable to loops during both free flow and congested conditions, although some occlusion issues were noted (Minge, Kotzenmacher, and Peterson 2010). In favorable weather conditions, false and missed calls ranged from 0.4% to 6.1% of vehicles. Investigation into the performance of these units under varying environmental conditions has been conducted, with the researchers noting that an increase in precipitation was correlated to performance degradation (Medina, Ramezani, and Benekohal 2013). Performance degradation for radar units can also come from out-of-date software, movement of the unit so that it no longer is pointing at the proper target area, and failure of the individual radio channels inside the unit.

Figure 2.2 shows a radar set up on a pole at an intersection in Florida from the brand Wavetronix. Radar detectors are most commonly positioned at a high elevation to provide a wide, unobstructed view of the intersection to minimize issues with occlusion.

Figure 2.2: Wavetronix radar detector (Huotari 2015)

2.2 DETECTOR HEALTH MONITORING

Monitoring of detector health can be generally divided into three separate methods: monitoring through traffic control products, monitoring through traffic control software / algorithm, and monitoring through the use of in-person assessments. The following subsections will detail what is available in scientific as well as vendor literature regarding these techniques.

2.2.1 Detector Health Monitoring with Traffic Control Products and Software

As was noted earlier, most traffic controllers and detection devices are able to detect major detector failures by examining the presence, absence, or frequency of data being sent by a detector, but these tools are not able to assess the quality of the information sent; therefore, the health of the detector is commonly unmonitored. For example, detrimental detector behaviors at signalized intersections such as a loop that fails for 3 minutes and works for 1 minute may not send a phase into recall, and therefore may not be observed.

Given the implementation of Q-Free/Intelight products on the ODOT system, the research team reached out to the vendor to request information regarding how their products monitor detector health (Q-Free Intelight, 2020). An excerpt from the email response from a project manager at Q-Free is summarized as follows:

> MAXTIME local control software includes three ways to identify a malfunctioning sensor. Collectively these features are called "detector diagnostics" in the software. These are an optional feature that can be programmed per detector.
>
> - No Activity – Assume a failure if no calls are received on a detector for a configurable period of time.
> - Max Presence – Assume a failure if a continuous call is placed on a detector for a configurable period of time.
> - Erratic Count – Assume a failure if a more than a specified number of calls are placed on a detector in a configurable period of time.
>
> When a detector is considered failed, a couple responses are possible.
>
> - Place a minimum or maximum recall.
>   - MAXTIME software is pretty flexible on this and lets you pick between Min 1 or Min 2 and Max 1, Max 2, or Max 3.
> - Define a "failed link" detector.
>   - This defines a detector that will be used in lieu of inputs from a failed detector.
>
> The controller has some internal storage where detector failures will be logged for a limited period of time. If a jurisdiction is using MAXVIEW atms (central system) then they can also get alarms pulled into a Traffic Management Center type program for review.

As noted from this communication, MAXVIEW does identify detector faults, but only at the ends of the performance spectrum. If performance has degraded slightly due to increased latency or some other performance issue, this would likely not be identified.

Other vendors incorporate similar capabilities in their control software. Econolite's Centracs SPM central system specifications notes that this system applies statistical data science to analyze detectors that may not be fully operational, and creates a list within the monitored corridor that may have degraded detector performance (Econolite 2020b). To accomplish this, Econolite's traffic controller can be programmed to identify a lack of activity on a certain detector by time of day as a possible failure. Additionally, their SPM tool can look historically at previous days to identify differences and use that information to flag a failure.

McCain is another manufacturer that sells controllers and intersection control software, but their published literature does not detail how their products address sensor health (McCain 2020), and attempts to acquire further information from the manufacturer were unsuccessful.

In researching detector health monitoring accomplished by detection devices, the research team reviewed various inductive loop and radar detection units and noted that the extent of health monitoring is reporting faults and logging them. Vendor websites did not provide detail on how faults were identified, however given what is known about common practices by the research team, it is presumed that faults are identified by examining the presence, absence, or frequency of data being sent by a detector. (Econolite 2020a; Iteris 2020; Wavetronix - SmartSensor V 2020)

2.3 DETECTOR HEALTH MONITORING THROUGH ALGORITHMS / POST PROCESSING

Algorithms can be used either in real time or through post-processing to identify problematic detector operation. Statistical methods can be used to identify outliers, infeasible data, and erroneous data, making it suitable to develop graphs and tables to find the location of the erroneous data within the data set. From there, it is possible to find the detector itself that was causing the poor data quality. While the work in this project is focused on interrupted flow facilities, algorithms in applied to uninterrupted flow are considered as well.

Researchers at the Washington State Transportation Center developed an algorithm to identify and correct dual-loop sensitivity problems that resulted in inaccurate reporting of truck volumes. Using individual vehicle information developed from event based high resolution data, the researchers were able to identify sensitivity discrepancies and then retune the detectors, the end result of this work being the implementation of the algorithm in a software tool for convenient usage (Nihan et al. 2006). In a study that used loop detector data from almost 15,000 Caltrans inductive loops, malfunctioning loops are identified through their volume and occupancy measurements. These measurements are compared against values at neighboring detectors as well as historical data to identify when a detector may be problematic, improving on earlier methods that only relied on data from a single detector (Chen et al. 2003). In related work, researchers at the University of Nebraska developed a methodology to identify malfunctions such as detector and communication failures that lead to erroneous data (Vanajakshi & Rilett 2006). This research focused on the conservation of vehicles principle on a system-wide level to identify locations where the principle was violated. It was then validated using a CORSIM model.

The Portland Oregon Regional Transportation Archive Listing (PORTAL) is the ITS data archive for freeway loop detector data for the Portland metropolitan region, documenting aggregated data and performance measures. Data uploaded into PORTAL is filtered to identify erroneous data through a series of data quality flags as well as comparison against plausibility thresholds. For the former technique, if a detector logs a speed as zero when the same detector logs a count greater than zero, a flag is raised. For the latter technique, data samples that have a speed above100 miles per hour, or below five miles per hour would be flagged. Data samples are then broken into four categories: Good, Suspicious (failed one or more data quality conditions), No Traffic, or Communication Failure. This information is then made known to the user when downloaded and can also be plotted to identify the scale of erroneous data by type of filter. Figure 2.3 shows a monthly report that is used to compare data samples from detectors to find failing units based on occupancy, volume, and speed thresholds (Tufte et al. 2007).

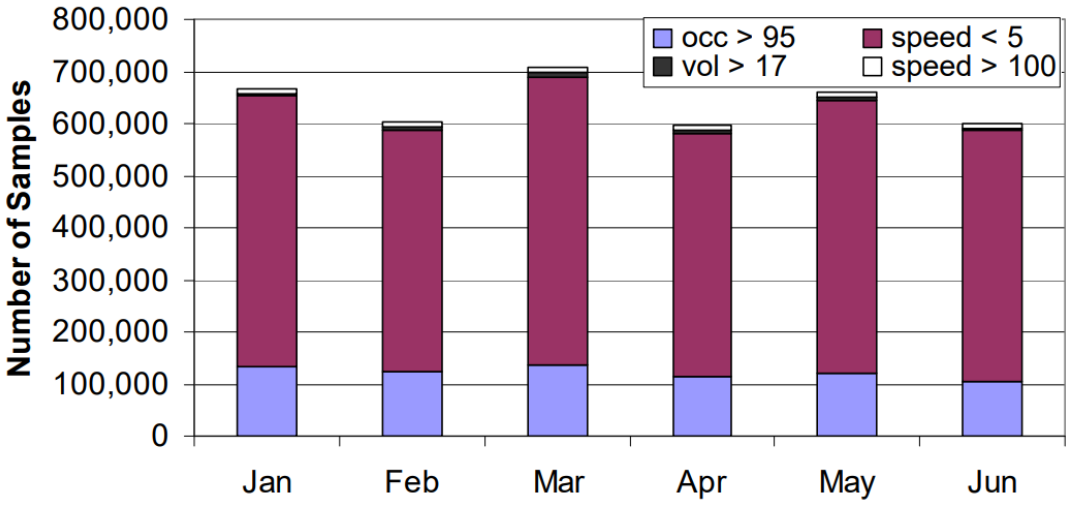


Figure 2.3: PORTAL - Number of samples failing selected conditions (Tufte et al. 2007)

Researchers in Sweden and Finland collaborated to develop a Fuzzy Intelligent Traffic Signal (FITS) control, a method which provides an inexpensive approach to improve signal control based on road infrastructure (J Jin et al. 2016). A simulation-based framework is used to evaluate different traffic control strategies based on certain criteria such as vehicle flows, pedestrian flows, priorities, and platoon management. In this methodology, stop line detectors assist in vehicle actuated timing and advance detectors play a crucial role in the decision making process (J Jin et al. 2016). In running their FITS simulations, the researchers determined that traffic states can still be properly estimated and proper decisions can be made even if a few detectors are malfunctioning, though the authors noted that there is a threshold where this falls apart (J Jin et al. 2016). Another project that related detection performance to advanced signal control was commissioned by Oregon DOT and completed in 2017. In this project, researchers at Northern Arizona University led a team that investigated the impact on non-invasive detection performance on adaptive control. As part of their site evaluation researchers noted that only 42% of the coupled detection zones (inductive loop and non-invasive technology) passed a human ground truth comparison. Additionally, the research team was able to identify other poorly performing detectors by comparing collected detector data (for example, occupancy with a video detector) with expected performance norms. One of the conclusions of this study was that

detector health monitoring is critical for sensors used for higher level control (Smaglik et al., 2017).

2.3.1.1 Automated Traffic Signal Performance Measures (ATSPMs)

Use of ATSPMs began in the mid-2000s with the collection and analysis of high-resolution event based data for traffic signal performance (Smaglik et al., 2007). Since then, researchers at Purdue University along with practitioners at the Indiana Department of Transportation and Utah Department of Transportation have evolved the use of event based data into a method of assessing and improving the performance of traffic signals, traffic signal systems, and traffic signal system business practices (Day et al., 2014). From a technical standpoint, the suite of ATSPMs can allow an agency to monitor capacity, progression, multimodal, and maintenance performance measures without the added expense of a central- or adaptive traffic signal system. These performance measures can be developed though robust communication and typical traffic signal detector information, though additional detection is required to take advantage of all the performance measures. On the topic of detection performance, detector health can be determined through identification of phases in recall over time, as this is an indication that the detector is not performing properly. These locations are aggregated and then reported to agency managers for repair prioritization.

2.4 DETECTOR HEALTH MONITORING THROUGH ON-SITE INVESTIGATION

While it is preferable to identify malfunctioning detectors through off-site means, equipment and procedures can be implemented on-site as well. Researchers in Germany developed a portable Malfunction Sniffer to identify errors in inductive loop detector outputs (Kuhnel et al 2011). Their device, shown in Figure 2.4, was effectively a portable method of ground truthing detector data. Once programmed with the exact location of the detection zones, the system would corroborate the outputs of the detectors with an audiovisual signal indicating vehicle passage so that supervisor could monitor the output. It was noted that this system did not work as well for video detectors.

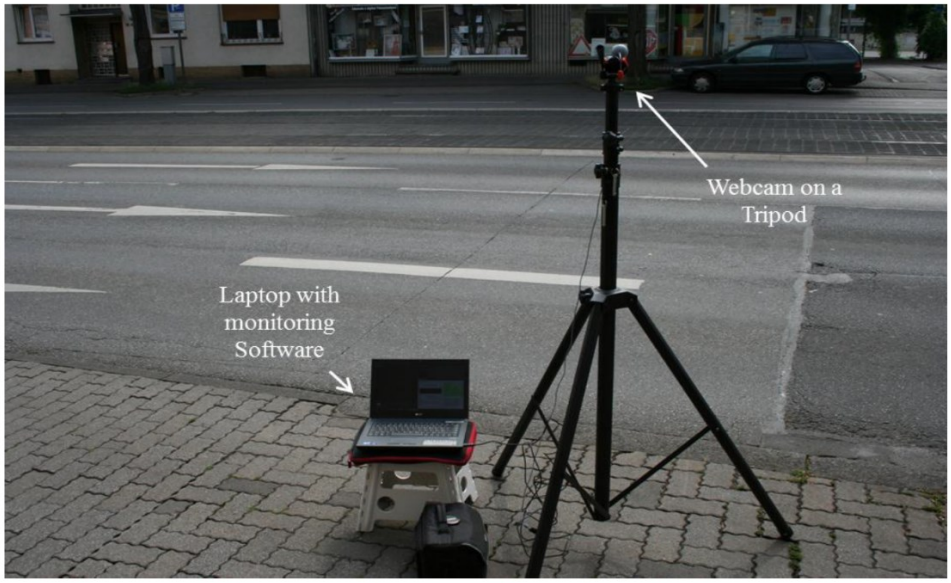


Figure 2.4: Malfunction sniffer (Kuhnel, Weisheit, and Hoyer 2011)

A project sponsored by the Federal Highway Administration (FHWA) attempted to use Ground Penetrating Radar (GPR) to identify the location of loop detectors, determine if they were functioning, and perform detailed analysis to assess the conditions of the sensor (Arnold et al. 2011). While the device developed and deployed in this work was able to accomplish all three goals to some degree, it was noted that the device was not able to detect defect and deterioration, and further work is required. Lastly, in a study performed by Purdue University, wireless magnetometers were tested against a standard loop detector to evaluate their effectiveness and accuracy at picking up calls. While wireless magnetometers are not the focus of this work, one conclusion of this study was that 8 foot spacing be observed between sensors adjacent to the stop line to minimize missed calls, indicating that design standards may have an impact on the performance of detection devices (Day et al. 2010).

2.5 TRAFFIC FLOW THEORY AND FUNDAMENTAL WORK

2.5.1 Greenshields Model

Traffic flow theory is the basis of conceptual modeling of traffic. Greenshields Model of traffic flow (Greenshields 1935) is an elegant relationship that illustrates the connected nature of volume, speed, and density within traffic operations. This relationship, shown in Equation 2-1, leads to the fundamental diagrams of the Greenshields model, shown in Figure 2.5.

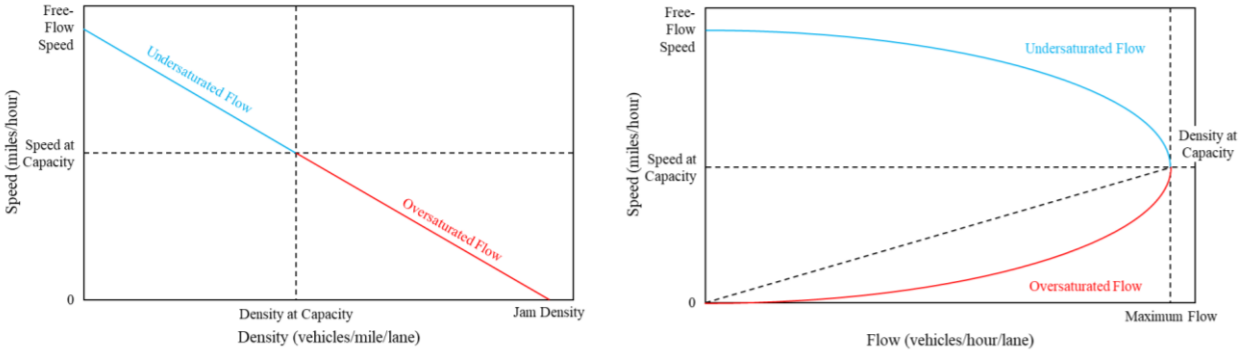
$$\boldsymbol{V = S * D} \qquad \textbf{(2-1)}$$

where:

V = Volume (vehicles/hour)

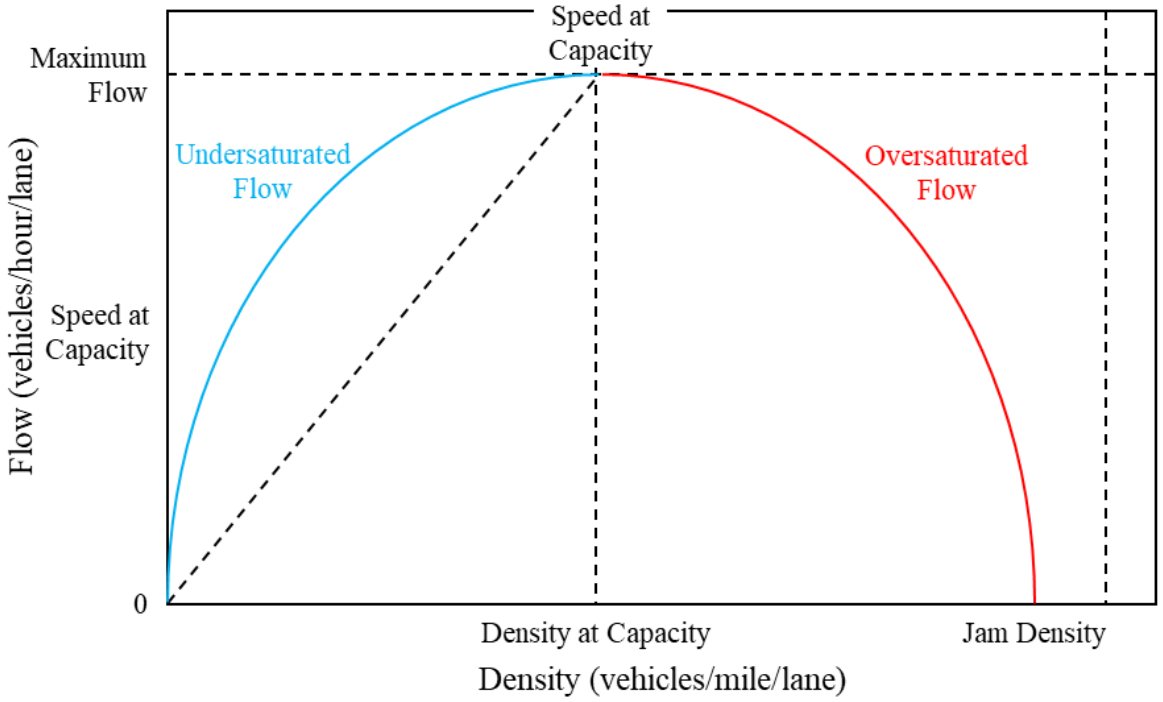
S = Speed (miles/hour)

D = Density (vehicles/mile)



(a) Speed vs Density

(b) Speed vs Flow (Volume)



(c) Flow (Volume) vs Density

Figure 2.5: Fundamental diagrams of Greenshields model

These diagrams illustrate the idealized conceptual relationships between the three macroscopic traffic stream parameters, Volume, Speed, and Density. They encompass two distinct regions of flow, undersaturated (under capacity) and oversaturated (over capacity). These diagrams are conceptual in nature, in that Volume, Speed, and Density data collected to model traffic flow at any given location when plotted would not give way to a smooth diagram as is shown in Figure 2.5, but would look more like Figure 2.6, which is a Speed / Density plot developed from real world data. The linear dashed line in Figure 2.6 represents Greenshields model, while the red points are the empirical data.

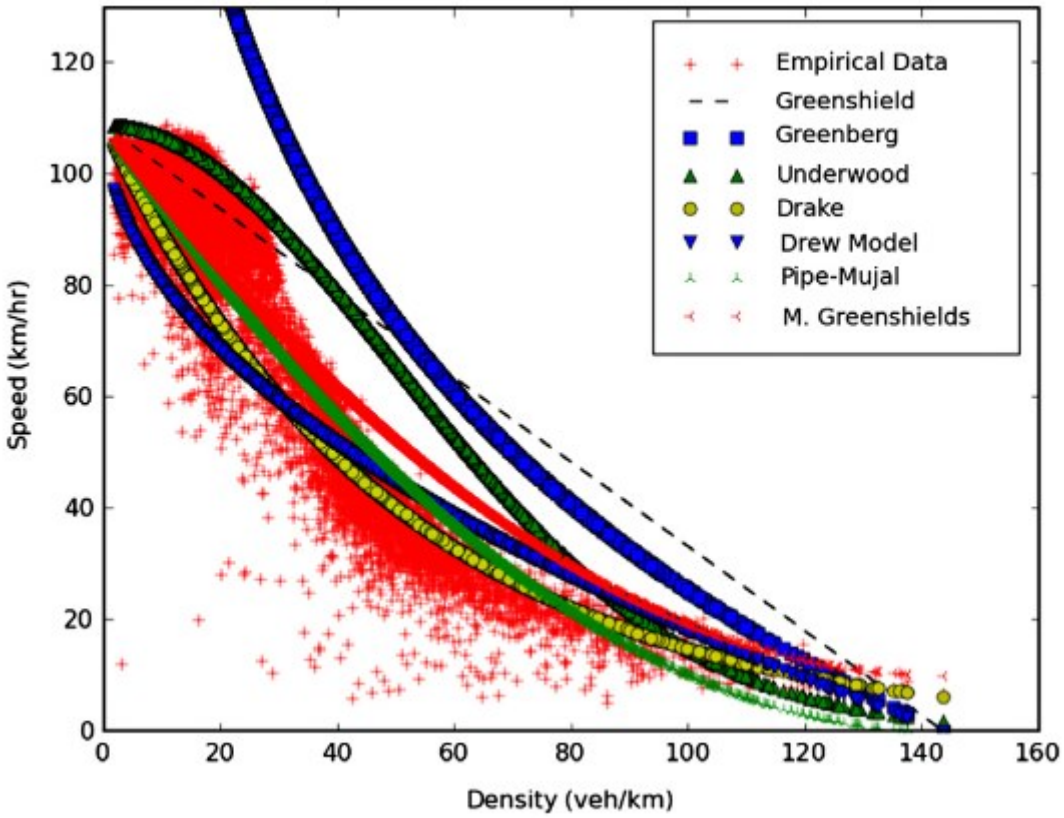


Figure 2.6: Real world speed-density plot (Wang et al. 2011)

The fundamental diagrams have been used in traffic research to assist in the investigation of incident detection (Jing Jin and Ran 2009), car-following models for simulation (Deng & Zhang 2012), the effects of weather on traffic operations (Dhaliwal et al. 2017), and variable speed limits (Bertini et al 2006), among countless other topics, but to the research team's knowledge have not been used in detector health applications.

2.5.2 Saturated Flow Rate and Headways

The departing vehicle flow rate at capacity from a signalized intersection is defined as the Saturation Flow Rate. This rate of flow occurs as vehicles in a standing queue depart, starting from the 5th vehicle in the queue as the first four vehicles in the queue depart at a lower flow rate due to time lost as the queue moves from a stopped to a moving queue (Transportation Research Board 2016). This Saturation Flow Rate can be determined in three separate ways. First, it can be calculated based upon site characteristics using methods set forth in the Highway Capacity Manual, as shown in Figure 2.7.

$$s = s_o f_w f_{HVg} f_p f_{bb} f_a f_{LU} f_{LT} f_{RT} f_{Lpb} f_{Rpb} f_{wz} f_{ms} f_{sp}$$

where

s = adjusted saturation flow rate (veh/h/ln),

s_o = base saturation flow rate (pc/h/ln),

f_w = adjustment factor for lane width,

f_{HVg} = adjustment factor for heavy vehicles and grade,

f_p = adjustment factor for existence of a parking lane and parking activity adjacent to lane group,

f_{bb} = adjustment factor for blocking effect of local buses that stop within intersection area,

f_a = adjustment factor for area type,

f_{LU} = adjustment factor for lane utilization,

f_{LT} = adjustment factor for left-turn vehicle presence in a lane group,

f_{RT} = adjustment factor for right-turn vehicle presence in a lane group,

f_{Lpb} = pedestrian adjustment factor for left-turn groups,

f_{Rpb} = pedestrian–bicycle adjustment factor for right-turn groups,

f_{wz} = adjustment factor for work zone presence at the intersection,

f_{ms} = adjustment factor for downstream lane blockage, and

f_{sp} = adjustment factor for sustained spillback.

Figure 2.7: Calculation of saturation flow rate (Transportation Research Board 2016)

Second, it can be directly measured in the field by counting the vehicles departing at capacity during a certain time period. Lastly, it can be determined in the field by measuring departure headways of vehicles departing at capacity, with headway defined as, "the time between successive vehicles as they pass a point on a lane or roadway, measured from the same point on each vehicle" (Transportation Research Board 2016). The relation between headway and volume is shown in Equation 2-2. If the headway measured occurs during queue discharge at capacity, the corresponding volume that will be calculated will be that of the saturation flow rate.

$$\boldsymbol{V = \frac{3600}{h}} \qquad \textbf{(2-2)}$$

where:

V = Volume (vehicles/hour)

h = Departure headway (seconds/vehicle)

The concepts of headway, saturation headway, and saturation flow rate were developed through applied research, and as part of the foundation of traffic operations theory, appear in research

endeavors covering all aspects of traffic theory, including intersection capacity (Laufer et al. 2019), the impact of automated vehicles on mixed-use lanes (Mohajerpoor & Ramezani 2019), bicycle operations (Raksuntorn & Khan 2003), geometric design (Potts et al. 2007), and weather conditions (Asamer & Van Zuylen 2011), among others, but they have not been applied to detector health.

2.6 CONCURRENT RESEARCH PROJECTS

The research team was aware of one concurrent research project that had a similar focus to this project:

- Multi-Stage Algorithm for Detection-Error Identification and Data Screening (Azin & Yang, 2020)
    - Funding Agency: Utah Department of transportation
    - Contractor: University of Utah.
    - PI: Xianfeng Terry Yang
    - Project Start Date: 01/24/2019
    - Project End Date: October 2020
    - Funding Amount: Not listed
    - The goal of this project was to develop a screening tool to identify detector errors from data within the Utah DOT detector data database. This work used statistical analysis as well as historical detector information to identify malfunctioning detectors from data within the database through a multi-stage process, using a combination of historical data, data from neighboring detectors, and the application of traffic flow theory to detector data to identify problematic detectors. The outcome of this project was a methodology that can conduct in-depth data reviews of those identified detector stations with potential detection-errors.

In this study, data is compiled from UDOT's Performance Measurement System (PeMS) from detectors along a corridor. The PeMS system receives vehicle count and occupancy data at 20 second intervals. Speed, flow, and occupancy are analyzed to find potential errors in a one-month data collection period. The primary method of detector health evaluation in this study is through comparison of adjacent detectors upstream or downstream of each other on this roadway.

While this project is related to the work described in the SPR 837 workplan, it was strictly concerned with detectors on free-flow facilities. Second, this work uses neighboring detectors in its identification algorithm; SPR 837 relies on detectors at an isolated signalized intersection. (Azin & Yang, 2020)

2.7 CONCLUDING REMARKS

This literature review has covered the basics of inductive loop and radar detection technology, the state of the practice regarding detector health monitoring, and the elements of traffic theory that will be used in monitoring detector health. Inductive loops, when functioning properly, are purported to be the most accurate detection technology, likely due to their close proximity to the traffic being detected, a consequence of being an invasive technology. But, because of their invasive nature, there are a number of issues that can compromise the performance of an inductive loop detection. Radar detection, one type of non-invasive detection, has been shown in research to be generally reliable, with environmental factors causing a minimal impact on performance, however internal components can fail without a complete failure of the unit, which can also compromise performance.

In the area of detector health, three different techniques were covered in this literature review: monitoring with traffic control products and software, monitoring with algorithms / post processing, and on-site monitoring. Traffic control products and software typically identify poorly performing detectors through monitoring for flickering, lack of a call, or a constant call from a specific detector. Most online vendor literature is vague when it comes to describing how detector health is monitored, if mentioned at all. This, combined with the lack of information in the literature focused on detector health monitoring in the field, indicates that detector health monitoring is typically accomplished with these aforementioned heuristics. If data is post processed, a number of different methods can be used to identify problems with detector health. This can be accomplished through comparing detector outputs with outputs of neighboring detectors, comparing detector outputs with historical data, or evaluating detector data with plausibility thresholds. Additionally, using ATSPMs, the health of a detector is monitored by identifying actuated phases operating in recall, an indication that the detector is not providing proper information to the controller. Lastly, on-site investigations can also be conducted to identify poorly performing detectors, if so desired.

Finally, Greenshields model and content within the Highway Capacity Manual form the theoretical basis for capacity analysis of interrupted and uninterrupted flow facilities. Each intersection approach has a unique discharge capacity that can be either calculated or measured in the field through two separate methods. These methods, along with the fundamental diagrams yielded through application of Greenshields' model, and combined with high resolution detection data, reveal an opportunity to monitor detector health through traffic flow information on a per intersection approach basis.

2.7.1 Application to the Project

There were no methods found through the literature review that are directly similar to the evaluation this study develops. This project incorporates aspects from each of the detector health monitoring topics researched in this literature review: data collection software and post-processing in this project integrates existing traffic theory, and a new method of on-site investigation is introduced to validate the project's health-monitoring algorithm.

In this project, the Fundamental Diagrams of Greenshields Model, Figure 2.5, and the associated theories are used to develop an algorithm for identifying detector malfunctions. The conceptual

quadratic relationship between Density and Volume is integral in deriving methods of detector health evaluation. Other relationships derived from the fundamental relationship between volume, speed, and density, shown in Equation 2-1, incorporate additional aspects of the detector data and the detector's location characteristics into this evaluation. Approximating uninterrupted saturated traffic flow is necessary for analyzing the data using existing traffic theory.

The remainder of this report is structured as follows. Chapter 3 will describe the analysis of the selected study sites, the process of data collection and reduction, and the manual verification of the selected detectors. Chapter 4 will be on data processing and data analysis, and the development of the health assessment algorithm. Chapter 5 will be the system design and implementation plan for incorporating the developed detector health monitoring algorithm. Chapter 6 will then explore the conclusions, lessons learned, and limitations of this project.

3.0 SITE SELECTION AND DETECTOR PERFORMANCE EVALUATION

Development of an algorithm to assess detector health is predicated on using developmental data from detectors that are performing properly. As such, a heuristic was developed to assess the performance of detectors in situ, comparing event-based data provided by ODOT to detector performance data reduced from field observations. This chapter documents this approach, and the outcome of the assessment. Table 3.1 lists all sites considered for analysis during this approach. Some sites were removed from consideration because of broken detection, unavailability of data during specific date ranges, or other reasons as noted in Table 3.1.

Table 3.1: Sites Evaluated During Site Selection Process

Number	Site	Location	Notes	Reason for Inclusion/Exclusion
1	Technology Loop	Corvallis		Some loops ground out
2	US101@N22nd	Lincoln City	Has extend/delay on detectors (removed 8/25)	Has extend/delay on detectors (removed 8/25)
3	OR34@Peoria	Corvallis	Loops and radar, has extend/delay on detectors (removed 8/24)	Loops and radar, has extend/delay on detectors (removed 8/24)
4	OR212@135th	Happy Valley	Replaces OR99W @ Tualatin-Sherwood-RD	Too much broken data
5	OR51@16th	Independence	-	Replaces OR34@I-5 N B Ramp; PreCovid data not available
6	OR99W@OR18	Dundee	-	PreCovid data not available; No detector event data
7	OR22@I-5 S B Off Ramp	Salem	Has extend/delay on detectors (removed 8/24)	Added as an option
8	US20@15th	Corvallis	Extend/delay on detectors removed	-
9	OR34@I-5 S B Ramp	Albany	Has loops and radar and no stretch or delay time on loops (verified August 17)	-
10	US26@Meinig-Pioneer	Sandy	Delay/extend removed from detectors	Replaces OR34@I-5 N B Ramp for something closer to Portland and isn't an on/off ramp
-	OR99W@Tualatin-Sherwood-RD	Sherwood	-	-
-	OR34@I-5 N B Ramp	Albany	Has loops and radar and no stretch or delay time on loops (verified August 17)	Removing to replace with oversaturated location

Notes:
Gray shading indicates the site was excluded from the study
Unshaded rows indicate the location was a study site

Event-based data for the six sites, provided by ODOT personnel, were compiled for analysis. All data for this project, collected and compiled by the project team, and collected by ODOT, are stored securely on Box.com. Additionally, all video files are held onsite at Oregon State University, and all Event Log files are stored on an NAU Research Dropbox Account.

3.1 EVENT LOG AND DRONE VIDEO DATA COLLECTION

This section covers the types of data collected or compiled as part of this task: Event Log data, drone video data, and elements obtained from drone video reduction.

3.1.1 ODOT-Provided Event Log Data

Event Log data from vetted detection devices at the six selected sites were used in subsequent tasks to develop algorithm(s) to identify poorly performing detectors. This Event Log Data reports information using Event IDs and corresponding Parameters (Day et al., 2014). While there are many different types of events contained in a typical log, the list of Event IDs and the corresponding Parameter used in this task are shown in Figure 3.1. Event IDs 1 and 8 were used to identify the start of each green and yellow phase, with timestamps attached to specific events used to determine the length of each cycle and each green and yellow/red phase. Event IDs 82 and 81 indicated the Vehicle Detector On and Vehicle Detector Off, respectively. With all radar and loop detection zones operating in presence, data from these events can be used to determine activations (which are used as a surrogate for vehicle counts in this work, as count detector outputs are not available) and occupancy, which will be used in subsequent sections to evaluate the efficacy of the detection zones at the study sites. Parameter outputs of Event IDs 81 and 82 in the Event Log are MaxTime (MT) numbers, which correlate to either inductive loop or radar detection zones as shown on intersection plans. Table 3.2 documents the correlations between the field loop / radar zones and the MaxTime numbers. Only detection zones which have valid Event ID 81/82 detector activity are listed in Table 3.2.

Event ID	Name	Description	Parameter Description
1	Phase Begin Green	Set when either solid or flashing green indication has begun.	Phase #
8	Phase Begin Yellow Clearance	Set when phase yellow indication becomes active and clearance timer begins.	Phase #
81	Vehicle Detector Off	Vehicle detector has turned off. Detector on and off events are triggered post any detector delay/extension processing.	Vehicle detector #
82	Vehicle Detector On	Vehicle detector has turned on. Detector on and off events are triggered post any detector delay/extension processing.	Vehicle detector #

Figure 3.1: Event log IDs and parameters (Day et al., 2014)

Table 3.2: MaxTime Numbers and Corresponding Detector Numbers

MaxTime Numbers and Corresponding Detector Numbers (1-14)														
Intersection	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
OR22 @ I-5SB Ramp		1	2	3	4-6			7	8	9-10	11-12	13-14		
OR34 @ I-5		2					Rad7		7	8		9	Rad13	Rad14
OR34 @ Peoria	20	1	2			3-4		10	11-12				21-22	28-29
US20 @ 15th	13	1	2					8	9-10				11-12	
US26 @ Meinig		1	2	3				4	5-6					
US101 @ 22nd	21	1	2	3-4	5-6			10	11-12	13-14			22-23	
MaxTime Numbers and Corresponding Detector Numbers (15-28)														
Intersection	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**
OR22 @ I-5SB Ramp		15	16	17-18										
OR34 @ I-5	1						Rad21		Rad23	Rad24	Rad25		Rad27	Rad28
OR34 @ Peoria		15	16					23	24	26-27		25		
US20 @ 15th	3	15	14					19	17-18				4-5	
US26 @ Meinig							7	10	11-12	13-14				8-9
US101 @ 22nd	7	15	16	17-18	19-20			24	25-26				8-9	

Two of the sites provided for this work, OR34 @ I5 and OR34 @ Peoria, are equipped with radar detection, in addition to inductive loop detection. At the outset of the analysis, it was determined that only radar count zones operating in 'Normal' mode (which is akin to a loop detector operating in presence mode) would be used in this analysis, as the outputs of the larger stop line and advance radar detection zones are manipulated by proprietary vendor software to achieve various objectives, and as such cannot be linked to traffic theory. Thus, they are excluded from analysis, and are not shown in Table 3.2. Table 3.3 lists the number of days of event-based data available for each site.

Table 3.3: Event Log Data Availability for Each of Six Intersections

Intersection	Dates	Days Available
OR22 @ I-5SB Ramp	8/2/20 – 8/8/20; 10/5/20 – 2/15/21	133
OR34 @ I-5	8/2/20 – 8/8/20; 10/5/20 – 2/15/21	140
OR34 @ Peoria	8/2/20 – 8/8/20; 10/5/20 – 2/15/21	133
US20 @ 15th	8/14/20 – 8/17/20; 10/5/20 – 2/15/21	137
US26 @ Meinig	8/26/20; 10/5/20 – 2/15/21	134
US101@ 22nd	8/2/20 – 8/8/20; 10/5/20 – 2/15/21	133

3.1.2 Drone Video Collection

Oregon State University research team members were responsible for the acquisition of field data to provide an inventory of existing infrastructure elements, to support the validation of controller logs of detector calls for service, and to support the calculation of saturation flow rates based on the current Highway Capacity Manual methodology. The following sections document how this work was conducted.

3.1.2.1 Roles and Responsibilities

Research team member roles and responsibilities were established to ensure the safe and efficient collection of field data. As this field work required the use of a small, unmanned aircraft system (i.e., drone) to collect aerial videos that would clearly show vehicles passing over in-pavement loop detectors, the following team roles were defined:

- *Remote Pilot-in-Command (PIC)*: The PIC checked local air traffic control requirements and submitted the application for controlled area if required. On site, the PIC led the team to find an appropriate place to test and set up the equipment. The PIC was responsible for operating the drone and making any needed flight adjustments to account for other users in the field or changing weather conditions.

- *Visual Observer (VO*): Once the drone was prepared to fly, the VO scanned the airspace in which the drone would operate to detect any potential collision hazards. Also, the VO maintained awareness of the position of the drone and effective communication with PIC. When data was being recorded, the VO alerted the PIC of any changes in safety relevant conditions.

- *Research Assistants*: During the drone data collection, the research assistant measured the position and length of detectors as well as other relevant measurements. After data collection, the research assistants edited the video data, and annotated the detector numbers, lengths, and positions on photos from the field.

3.1.2.2 Equipment

This experiment required the use of a drone, a distance measuring wheel, and a high-resolution camera, among other items. These tools are described in the subsections 3.1.2.2.1 and 3.1.2.2.2.

3.1.2.2.1 Drone

A DJI Mavic 2 Pro was used to collect all drone data in the field. Figure 3.2 and Figure 3.3 display the components stored in a Pelican case and Shoulder bag.

Figure 3.2: Primary drone components and registration

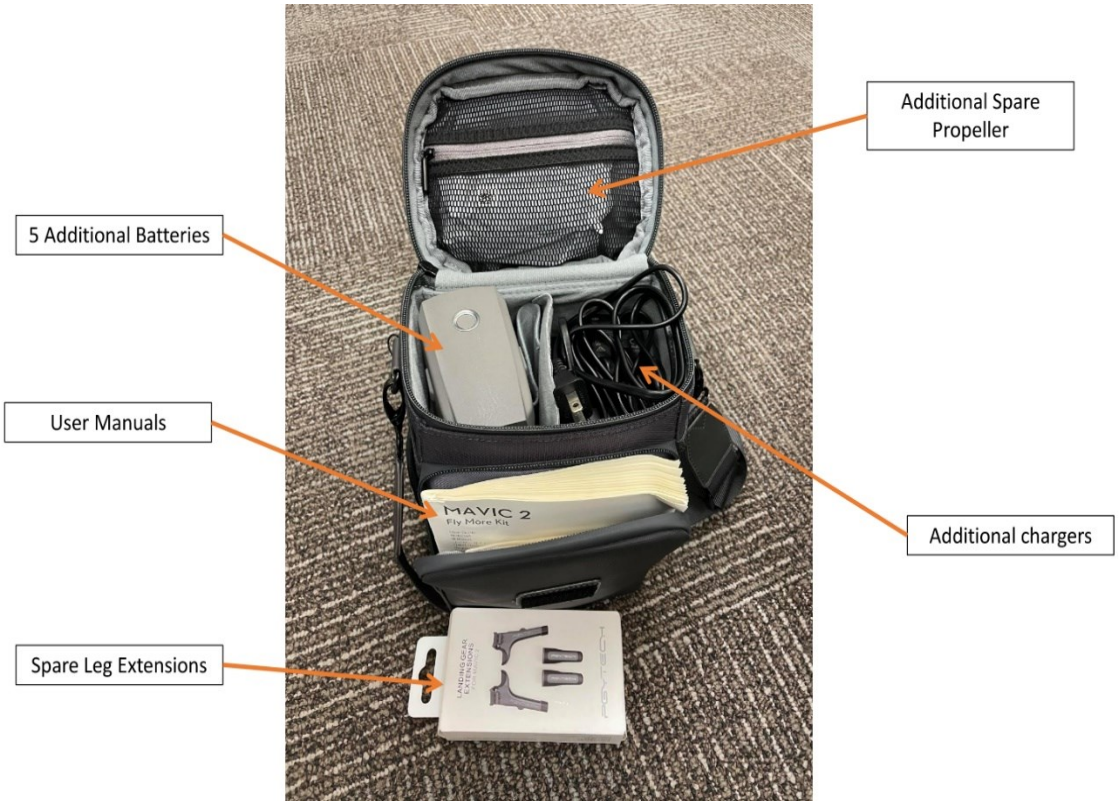


Figure 3.3: Drone batteries and charging tools

3.1.2.2.2 Supplemental Equipment

In addition to the drone equipment, the field work was supported by a landing pad, a solar powered electric generator, and a measuring wheel. The landing pad made it easier to initiate takeoffs and landings on uneven terrain, the field generator was used to recharge drone batteries in the field between flights, the measuring wheel was used to document the dimensions of detectors and their distance from the stop lines, and personal protective equipment (PPE) contributed to the safety of researchers in the field. Figure 3.4 displays images of the supplemental equipment used in the field.

(a) Landing Pad

(b) Generator

(c) Measuring Wheel

(d) PPE

Figure 3.4: Supplemental equipment including (a) Landing pad, (b) Generator, (c) Measuring wheel, and (d) PPE

3.1.2.3 ODOT UAS Contractor Requirements

Before the field work could be performed, OSU fulfilled nine ODOT UAS Contractor Requirements. Those requirements include:

1. UAS needs to be registered with the Oregon Department of Aviation and FAA and proof of registration needs to be submitted;
2. Proof of UAS Insurance;
3. FAA Part 107 Registration for PIC for all missions;
4. ODOT requires a PIC and Visual Observer for all missions;
5. Provide waiver documentation, if needed;
6. Have and follow “rules of engagement,” which includes what to do during an emergency;

7. Pilot must have flown at least 3 missions in the last 90 days;

8. Pilot must follow all Part 107 regulations, with exception of airspace with a waiver of authorization;

9. If the work is construed as surveying, photogrammetry, or engineering, the work must be performed under the responsible charge of a professional engineer, professional land surveyor or professional photogrammetrist in the State of Oregon.

OSU completed all applicable requirements within this set of requirements and was granted approval by ODOT to perform the field work.

3.1.2.4 Experimental Data Collected

Three primary types of data were collected in the field: detector position and dimensions, aerial video of vehicles driving over detectors, and inputs to calculate the HCM adjusted saturation flow rate.

3.1.2.4.1 Detector Position and Dimensions

Research assistants used signal plans provided by ODOT, photographs collected on site, and distance measuring wheels, to confirm the existence, placement, and dimensions of detectors. These details were then annotated on photos from the field and the provided signal plans. Figure 3.5 shows the road measurement details. A research assistant measured the diameter of the circular detectors, the nearest length from detector to stop line (placement), and the dimensions of the parallelogram detectors.

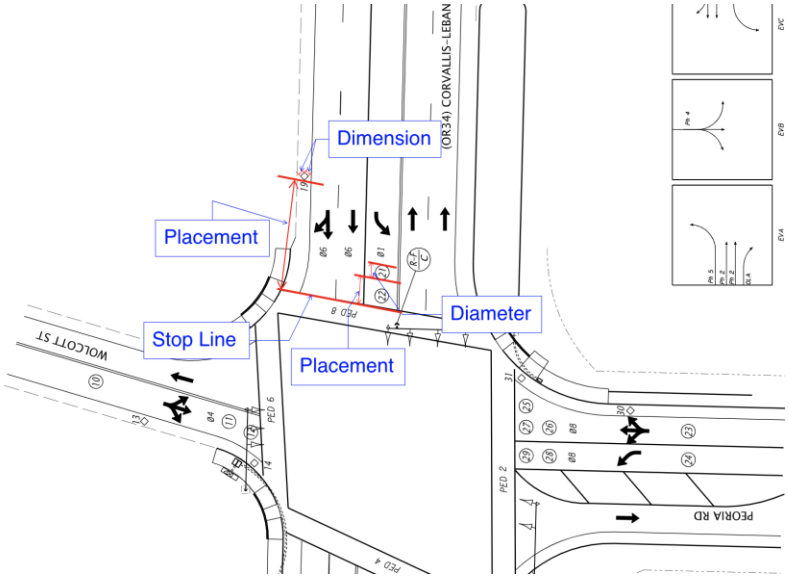


Figure 3.5: Road measurement details

3.1.2.4.2 Drone Video of Vehicles Driving over Detectors

As noted earlier, six intersections were selected as experimental sites for this research. Each site has a unique configuration (e.g., number, position, and dimensions) of detectors. To collect usable videos, i.e., stable images with good contrast of detectors against pavement, the weather conditions and the drone position had to be carefully considered. Additionally, some detectors are located hundreds of feet away from each other. This combination of factors required multiple drone flights on each approach to directly observe each individual in-pavement detector. With one available drone (DJI Mavic 2 Pro) and nine batteries, the video observations were collected one intersection per day and one video at a time with a maximum video duration of approximately 20 minutes.

To ensure the safety of the drone, the research team, and the traveling public as well as to ensure the quality of videos, research assistants found an appropriate area to set up the landing pad. This is a critical choice in avoiding collisions with power lines, span wires, tree branches, and other overhead obstacles. Once the drone was in flight, as stipulated by the FAA, the maximum flight elevation cannot exceed 400 ft, and the drone must not fly over the road or any non-research personnel. Moreover, the VO needed to continuously surveil the surrounding environment while the PIC adjusted the camera angle to ensure the detectors and traffic signal display were captured on video simultaneously. During the recording period, the PIC was responsible for attending to the drone and the controller, maintaining constant communication with the VO, and ensuring the flight occurred safely until the drone had landed. After the field data was collected from all six locations, research assistants cropped the videos, added timestamps, and designed an Excel template to transcribe the video data.

3.1.2.5 HCM Adjusted Saturated Flow Rate

The Highway Capacity Manual (HCM) describes a methodology for evaluating the capacity and quality of service provided to motorized vehicles at a signalized intersection. One aspect of this methodology includes a calculation procedure for determining an adjusted Saturation Flow Rate. The procedure includes several elements: adjustments for lane width, heavy vehicle and grade, parking, bus blockage, area type, right turns and left turns, pedestrians and bicycles, work zone presence, downstream lane blockage, and sustained spillback. In the field, research assistants collected inventory data related to each of these adjustment factors for each intersection approach. Figure 2.7, i.e., Equation 19-8 from the 2016 version of the HCM, was used to compute the adjusted Saturation Flow Rate per lane for the subject lane group. The individual factors were documented in the field in the template displayed in Figure 3.6. This information will be used for algorithm development work in subsequent tasks.

Adjustment Factors		NB					SB					EB					WB					
	Lane	1	2	3	4	Bike	1	2	3	4	Bike	1	2	3	4	Bike	1	2	3	4	Bike	
fw	Width (ft)																					HCM Exhibit 19-20 (<10:0.96, >=10-12.9: 1.00, >12.9: 1.04)
	/																					
fHVg																						Not sure how to determine
fp	Street Parking: Y/N																					
fbb	Bus stop: Y/N																					
fa																						Not sure how to determine
fLU	Lane Type	L/R/T	L/R/T	L/R/T	L/R/T		L/R/T	L/R/T	L/R/T	L/R/T		L/R/T	L/R/T	L/R/T	L/R/T		L/R/T	L/R/T	L/R/T	L/R/T		
	Marking: Y/N																					
	/																					1 if lane group has one shared lane or one exclusive lane
fLT																						Er=1.18 or El=1.05, if not, refer to Chapter 21 Section 3
fRT																						
fLpb	Ped/Bike: Y/N																					Chapter 31 Section 2, Yes for the appearance of pedestrian and bike
fRpb																						
fwz	Work Zone: Y/N																					
fms	Closure: Y/N																					
fsp	Spillback observed: Y/N																					

Figure 3.6: Adjustment factors form

3.2 DATA TRANSCRIPTION AND REDUCTION

3.2.1 Drone Video Data Transcription

The video captured by the drone was transcribed to obtain usable ground truth information about detector calls. From the video, research assistants concentrated on any vehicles that traversed an active detector. Time stamps were recorded when the front bumper of the vehicle arrived at the upstream edge of the detector zone (Figure 3.7a) and then again when the rear bumper of the vehicle departed the downstream edge of the detector zone (Figure 3.7b). Additionally, the active traffic signal display was recorded during each call for service. Transcription for an individual detection zone was performed for the either the duration of the entire video or for the first 100 vehicle incursions.

(a) (b)

Figure 3.7: Detector 19 on the NB approach of US20 and 15th Street. (a) Detector 19 call on. (b) Detector 19 call off.

Figure 3.8 shows the data transcribed for Detector 19 on the NB approach of US20 and 15th Street, the same scenario described in Figure 3.7. The white van arrived at 13:07:29:10 and left

at 13:07:30:27 on November 20, 2020. Equation 3-1 was used to calculate the duration of the call for service.

$$\left[(\boldsymbol{Departure\ Hour} * \mathbf{3600}) + (\boldsymbol{Departure\ Minute} * \mathbf{60}) + (\boldsymbol{Departure\ Second}) + \left(\frac{\boldsymbol{Frames}}{\mathbf{30}}\right)\right] - \left[(\boldsymbol{Arrival\ Hour} * \mathbf{3600}) + (\boldsymbol{Arrival\ Minute} * \mathbf{60}) + (\boldsymbol{Arrival\ Second}) + \left(\frac{\boldsymbol{Frames}}{\mathbf{30}}\right)\right] =$$

$$\left((\mathbf{13} * \mathbf{3600}) + (\mathbf{7} * \mathbf{60}) + \mathbf{29} + \left(\frac{\mathbf{10}}{\mathbf{30}}\right)\right) - \left((\mathbf{13} * \mathbf{3600}) + (\mathbf{7} * \mathbf{60}) + \mathbf{30} + \left(\frac{\mathbf{27}}{\mathbf{30}}\right)\right) = \mathbf{1.57}\boldsymbol{s}$$

(3-1)

Site		US20@15th St														
Date		11/20/2020														
Direction		NB														
Video	Vehicle	Time*											Traffic Signal Status**			Comment(s)
		Arrival	Depart	Calculation									Red	Yellow	Green	
				Arrival (Degree)				Depart (Degree)				Duration (Decimal) (second)				
				h	m	s	sss	h	m	s	sss					
NB #17-19	Detector	19														
	1	13:07:02	13:07:03	13	7	2	6	13	7	3	4	0.93	x			
	2	13:07:29	13:07:30	13	7	29	10	13	7	30	27	1.57	x			
	3	13:08:47	13:08:48	13	8	47	23	13	8	48	21	0.93	x			
	4	13:08:59	13:09:00	13	8	59	21	13	9	0	25	1.13	x			
	5	13:10:19	13:10:20	13	10	19	10	13	10	20	12	1.07	x			
	6	13:10:41	13:10:42	13	10	41	7	13	10	42	8	1.03	x			
	7	13:11:38	13:11:48	13	11	38	12	13	11	48	0	9.60	x			
	8	13:11:51	13:11:53	13	11	51	4	13	11	53	29	2.83			x	
	9	13:15:08	13:15:09	13	15	8	14	13	15	9	5	0.70	x			
	10	13:15:12	13:15:13	13	15	12	25	13	15	13	23	0.93	x			

Figure 3.8: Sample data transcription form

3.2.2 Manual vs Event Log Comparison

Detector performance was evaluated by comparing manually reduced drone-video data of vehicles entering and exiting detection zones to the corresponding Event Log detector outputs.

3.2.2.1 Event Log Data Reduction and Data Set Preparation

While the timestamps on the drone Video Log were close to the timestamps in the Event Log (within a minute, generally), the specific Event Log data which corresponded directly to the reduced drone video data needed to be identified. To accomplish this, individual vehicles were identified within both data sets, and the time between Vehicle Detector On indications in both the Video Log and Event Log were used to match up a specific Event Log vehicle activation with the corresponding Video Log vehicle activation. This process was conducted for the first and last vehicle of each manually reduced Video Data log to develop a complete list of Event Log Vehicle Detector On and Vehicle Detector Off activations that would correspond to activity in the Video Log during the same timeframe. Once this was accomplished,

the initial vehicle green indication noted in the Video Log was used to shift the timestamps so that the initial vehicle activation in both the Video Log and Event Log occurred simultaneously. This process was undertaken for each detection zone for which data were available, as was noted in Section 3.1.1. An example of this data reduction is shown in Table 3.4 below for Detector 7 of the intersection of OR22 and I-5. In this example, the detector on duration is calculated for each activation to be used for comparative analysis.

Table 3.4: Timestamps Example: Detector On and Off Indications for Detector 7 OR22 at I-5

Detector Indication (minutes:seconds.00)				Detector On Duration	
Video Log		Event Log		Video Log	Event Log
On	Off	On	Off	= Off - On	= Off - On
43:29.67	43:31.13	43:30.10	43:31.70	0:00:01.46	0:00:01.60
43:32.13	43:33.27	43:32.60	43:33.80	0:00:01.14	0:00:01.20
43:37.80	43:38.67	43:38.20	43:39.10	0:00:00.87	0:00:00.90
43:44.37	43:44.90	43:44.90	43:45.40	0:00:00.53	0:00:00.50
43:46.53	43:47.10	43:47.00	43:47.60	0:00:00.57	0:00:00.60
43:59.33	43:59.90	43:59.70	44:00.40	0:00:00.57	0:00:00.70
44:07.23	44:08.40	44:07.80	44:09.00	0:00:01.17	0:00:01.20
44:12.43	45:24.77	44:12.90	45:25.40	0:01:12.34	0:01:12.50
45:28.30	45:29.60	45:28.80	45:30.30	0:00:01.30	0:00:01.50
45:31.00	45:32.23	45:31.70	45:32.80	0:00:01.23	0:00:01.10
45:34.87	45:36.03	45:35.50	45:36.70	0:00:01.16	0:00:01.20
45:47.63	45:47.93	45:48.20	45:48.50	0:00:00.30	0:00:00.30

Finally, it should be noted that for radar count detection zones at the intersection of OR34 @ I5, the Event Log outputs were compared to the Video Log for closest neighboring inductive loop detector as the exact location of the radar count zone is not visible. For the advanced detectors 7 and 21 that span the width of the entire approach, the activations of multiple neighboring loop detectors were compiled chronologically to develop a consistent comparative set.

3.3 DETECTOR PERFORMANCE EVALUATION

3.3.1 Comparative Metrics

With the data sets prepared as shown in Table 3.4, the research team settled on two separate metrics for comparative analysis of the two logs to determine whether or not the detector would be considered suitable for the basis of this work. It should be noted that a detector noted as suitable in the context of this work does not imply the detector is not able to perform its prescribed traffic control function; it means that the detector is able to perform at a higher level of consistency for the purpose of algorithm development. The first comparative metric used was the total number of activations noted by each log during the analysis period. A difference

threshold of 10% was used to determine whether or not the radar / loop detector was healthy with regard to activations. The calculation is shown in Equation 3-2.

$$\frac{\textbf{\textit{Total Observation Event Log Activations} - \textbf{\textit{Total Observation Video Log Activations}}}{\textbf{\textit{Total Observation Video Log Observations}}} \quad \textbf{(3-2)}$$

While the threshold of 10% is a general rule of thumb when comparing counts from vehicle sources, it has been used in previous research works for this same purpose (detector performance) (Smaglik et al. 2007) (Smaglik et al. 2017). The second metric used for determining health of the detectors in this study is the Detector On Duration. This value, when combined with an analysis period duration, can be used to determine the occupancy of a detection zone. As was shown in Table 3.4, the Detector On Duration was found for each activation for both the Event Log and Video Log data sets. The distributions of Detector On Durations for both the Video Log and Event Log were compared using a paired t-Test to identify statistically significant differences (an F-Test was used to check whether each pair of distributions had Equal or Unequal Variances, and the corresponding t-test was used based upon the outcome of that test). If the t-Test indicated a Significant Difference, then the detector was determined to be unhealthy for the purpose of this analysis (Montgomery, 2018).

3.3.2 Results

Table 3.5 lists the detectors that passed the comparative analysis undertaken in this work. A total of 79 detection zones underwent the above comparative analysis (70 inductive loop and 9 radar). The comparative analysis in this chapter is to select the detectors for further data analysis and algorithm development for this project. The detectors that did not pass the comparative analysis may not have issues because the detectors' health should be verified by signal timers or electricians. On the following pages, Table 3.6 through Table 3.11 and Figure 3.9 through Figure 3.14 list the results of the comparative analysis for each of the six study intersections. The subset of detection zones that passed the comparative analysis include stop line, advanced, single lane, multiple lane, short, and long detection zones over a variety of lane usages and provided a robust basis for development of a detector health monitoring algorithm.

Table 3.5: Usable Detectors Summary Table

Usable Detectors from Each Study Intersection					
Intersection	**Det#**	**MT#**		**Lanes**	**Location**
OR22 at I-5	1	2	Loop	1	Advanced
	2	3		1	Advanced
	4-6	5		3	Advanced
	7	8		1	Advanced
	8	9		1	Advanced
	9-10	10		2	Advanced
	17-18	18		2	Advanced
OR34 at I-5	2	2	Loop	1	Advanced
	7	8		1	Advanced
	8	9		1	Advanced
	9	12		1	Advanced
	13	13	Radar	1	Stop Bar
	14	14		1	Stop Bar
	23	23		1	Stop Bar
	25	25		1	Stop Bar
	27	27		1	Stop Bar
	28	28		1	Stop Bar
OR34 at Peoria	3-4	6	Loop	2	Advanced
	10	7		1	Advanced
	16	17		1	Advanced
	20	1		1	Advanced
	21-22	13		1	Stop Bar
	23	22		1	Advanced
	24	23		1	Advanced
US20 at 15th	1	2	Loop	1	Advanced
	2	3		1	Advanced
	3	15		1	Advanced
	4-5	27		1	Stop Bar
	8	8		1	Advanced

Usable Detectors from Each Study Intersection					
Intersection	**Det#**	**MT#**		**Lanes**	**Location**
	9-10	9		1	Stop Bar
	13	1		1	Advanced
	15	16		1	Advanced
	19	22		1	Advanced
US26 at Meinig	4	8	Loop	1	Advanced
	5-6	9		1	Stop Bar
	7	21		1	Advanced
US101 at 22nd	3-4	4	Loop	1	Stop Bar
	5-6	5		1	Stop Bar
	7	15		1	Advanced
	8-9	27		1	Stop Bar
	17-18	18		1	Stop Bar
	19-20	19		1	Stop Bar
	21	1		1	Advanced
	24	22		1	Advanced
	25-26	23		1	Stop Bar

Table 3.6: Difference Summary Table: OR22 at I-5

Det	MT	Activations			Detector On Duration Mean			Usable?
		Manual	Event Log	Difference	Manual	Event Log	Difference	
1	2	100	95	-5*	00:00.34	00:00.37	00:00.03	Y
2	3	105	103	-2	00:00.34	00:00.36	00:00.02	Y
3	4	72	72	0	00:00.28	00:00.33	00:00.05**	N
4-6	5	100	90	-10	00:00.40	00:00.47	00:00.07	Y
7	8	58	58	0	00:03.69	00:03.72	00:00.03	Y
8	9	75	75	0	00:02.39	00:02.55	00:00.16	Y
9-10	10	103	98	-5	00:03.78	00:04.87	00:01.10	Y
11-12	11	59	76	17*	00:09.36	00:07.96	-00:01.40	N
13-14	12	78	59	-19*	00:07.75	00:09.31	00:01.56	N
15	16	100	48	-52*	00:00.24	00:00.58	00:00.34**	N
16	17	100	58	-42*	00:00.27	00:00.86	00:00.59**	N
17-18	18	100	93	-7	00:00.95	00:01.04	00:00.09	Y

* indicates a difference of >10% between the Manually reported and Event Log number of activations

** indicates Significant Difference at 95% CI in the Detector On Durations as reported by the t-Test

Green cell shading indicates the detector had neither a Significant Difference in the Detector On Durations as reported by the t-Test, nor a difference of >10% between the Manually reported and Event Log number of activations. Red cell shading indicates that the detector had one or both of these.

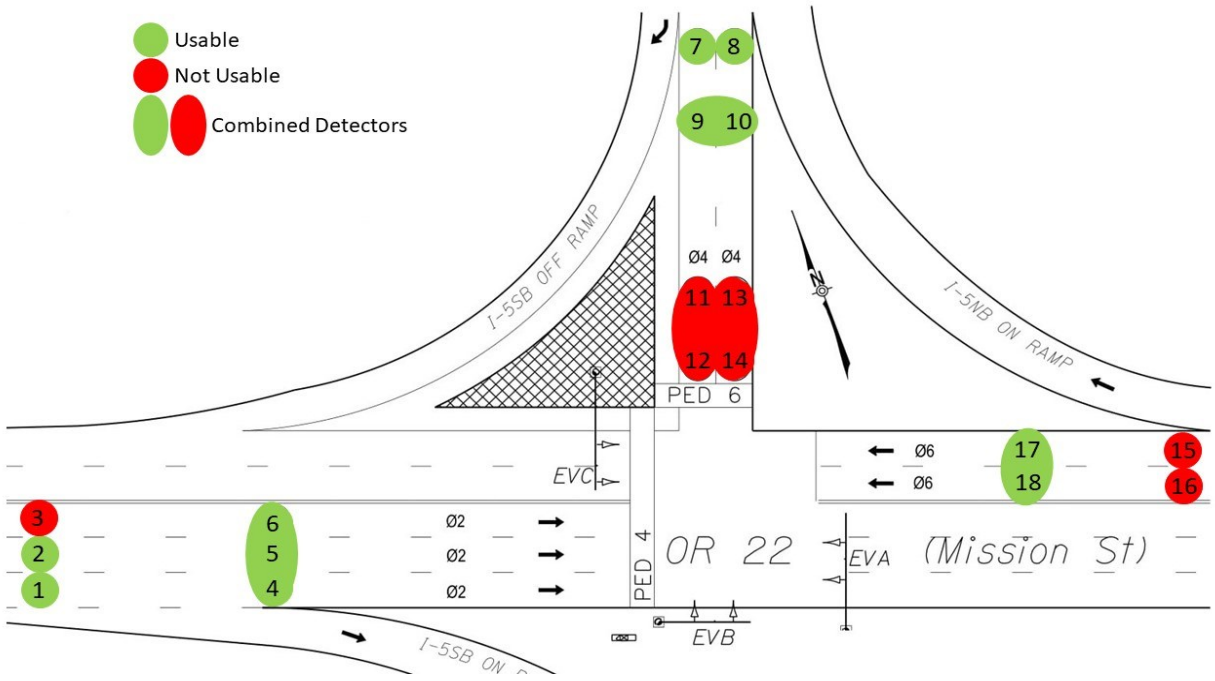


Figure 3.9: Comparative analysis results: OR22 at I-5

Table 3.7: Difference Summary Table: OR34 at I-5

Det	MT	Activations			Detector On Duration Mean			Usable?
		Manual	Event Log	Difference	Manual	Event Log	Difference	
1	15	25	35	10*	00:00.80	00:00.55	-00:00.25	N
2	2	100	110	10	00:00.41	00:00.41	00:00.00	Y
7	8	92	95	3	00:00.68	00:00.65	-00:00.03	Y
8	9	29	30	1	00:00.53	00:00.51	-00:00.02	Y
9	12	54	53	-1	00:01.41	00:01.45	-00:00.04	Y
RAD7	7	156	125	-31*				N
RAD13	13	27	26	-1				Y
RAD14	14	100	101	1				Y
RAD23	23	53	53	0				Y
RAD24	24	25	30	5*				N
RAD25	25	98	93	-5				Y
RAD27	27	64	63	-1				Y
RAD28	28	57	58	1				Y
RAD21	21	131	106	-25*				N

* indicates a difference of >10% between the Manually reported and Event Log number of activations

** indicates Significant Difference at 95% CI in the Detector On Durations as reported by the t-Test

Green cell shading indicates the detector had neither a Significant Difference in the Detector On Durations as reported by the t-Test, nor a difference of >10% between the Manually reported and Event Log number of activations. Red cell shading indicates that the detector had one or both of these.

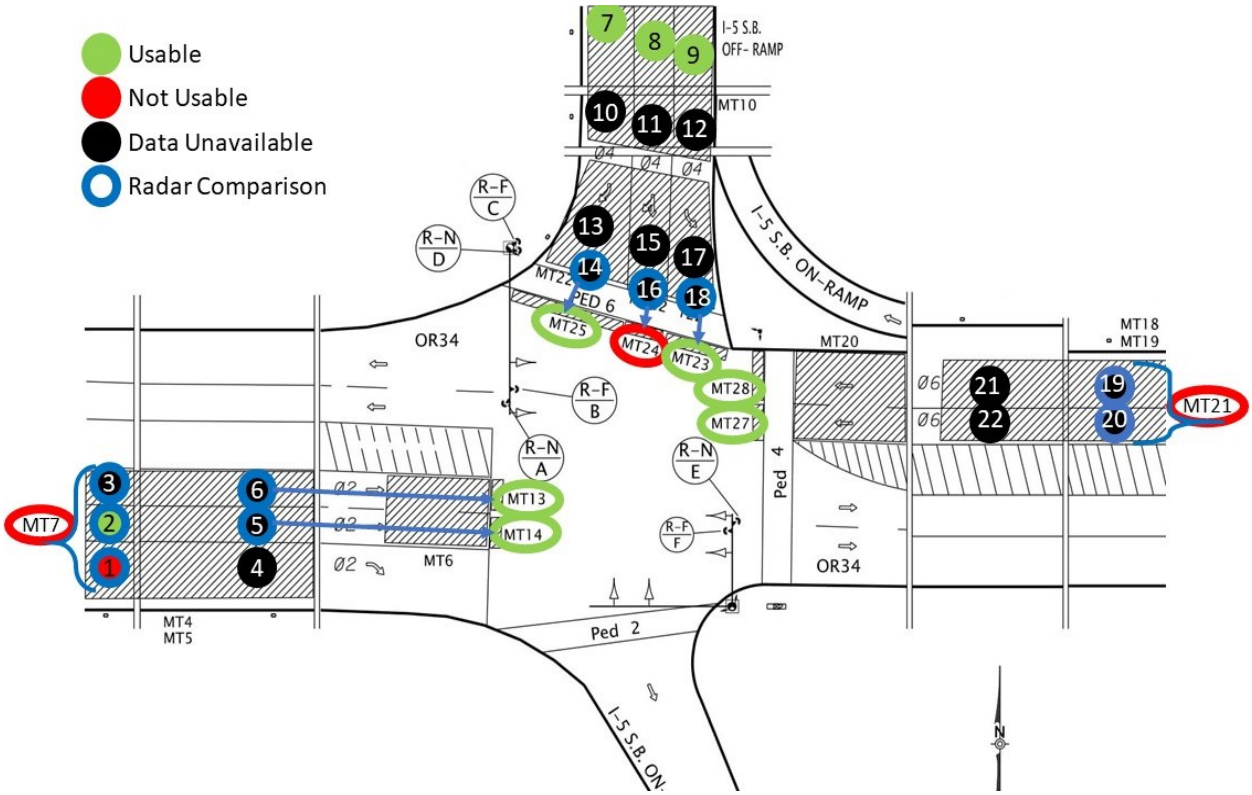


Figure 3.10: Comparative analysis results: OR34 at I-5

Table 3.8: Difference Summary Table: OR34 at Peoria

Det	MT	Activations			Detector On Duration Mean			Usable?
		Manual	Event Log	Difference	Manual	Event Log	Difference	
1	2	48	47	-1	00:00.23	00:00.40	00:00.18**	N
2	3	63	63	0	00:00.15	00:00.34	00:00.19**	N
3-4	6	100	98	-2	00:00.65	00:00.62	-00:00.03	Y
10	8	6	6	0	00:01.39	00:01.58	00:00.20	Y
11-12	9	6	5	-1*	00:10.13	00:11.76	00:01.63	N
15	16	88	89	1	00:00.42	00:00.34	-00:00.08**	N
16	17	100	100	0	00:00.80	00:00.71	-00:00.10	Y
20	1	4	4	0	00:02.08	00:01.73	-00:00.35	Y
21-22	13	4	4	0	00:12.14	00:10.67	-00:01.47	Y
23	22	24	24	0	00:02.53	00:02.50	-00:00.03	Y
24	23	35	35	0	00:05.39	00:05.56	00:00.17	Y
25	26	9	11	2*	00:16.87	00:12.69	-00:04.18	N
26-27	24	14	18	4*	00:20.72	00:18.08	-00:02.64	N
28-29	14	26	33	7*	00:23.16	00:17.76	-00:05.40	N

* indicates a difference of >10% between the Manually reported and Event Log number of activations

** indicates Significant Difference at 95% CI in the Detector On Durations as reported by the t-Test

Green cell shading indicates the detector had neither a Significant Difference in the Detector On Durations as reported by the t-Test, nor a difference of >10% between the Manually reported and Event Log number of activations. Red cell shading indicates that the detector had one or both of these.

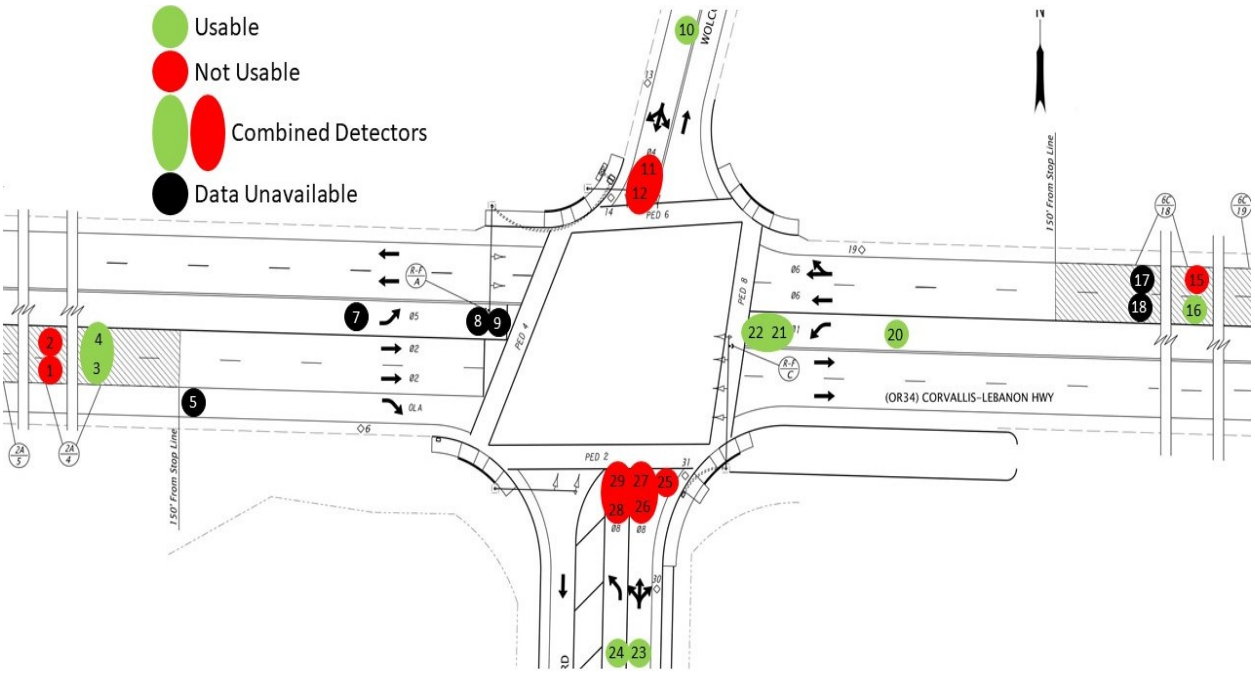


Figure 3.11: Comparative analysis results: OR34 at Peoria

Table 3.9: Difference Summary Table: US20 at 15th

Det	MT	Activations			Detector On Duration Mean			Usable?
		Manual	Event Log	Difference	Manual	Event Log	Difference	
1	2	100	102	2	00:00.44	00:00.43	-00:00.01	Y
2	3	100	100	0	00:01.02	00:01.01	-00:00.01	Y
3	15	15	15	0	00:03.99	00:03.85	-00:00.14	Y
4-5	27	13	14	1	00:28.90	00:26.57	-00:02.33	Y
8	8	32	33	1	00:01.46	00:01.55	00:00.10	Y
9-10	9	22	21	-1	00:18.84	00:19.39	00:00.55	Y
11-12	13	7	9	2*	00:38.35	00:29.62	-00:08.73	N
13	1	8	8	0	00:10.10	00:10.08	-00:00.02	Y
14	17	100	100	0	00:00.97	00:00.78	-00:00.18**	N
15	16	100	99	-1	00:00.96	00:00.94	-00:00.02	Y
17-18	23	18	22	4*	00:28.25	00:22.71	-00:05.54	N
19	22	26	26	0	00:04.28	00:04.63	00:00.34	Y

* indicates a difference of >10% between the Manually reported and Event Log number of activations

** indicates Significant Difference at 95% CI in the Detector On Durations as reported by the t-Test

Green cell shading indicates the detector had neither a Significant Difference in the Detector On Durations as reported by the t-Test, nor a difference of >10% between the Manually reported and Event Log number of activations. Red cell shading indicates that the detector had one or both of these.

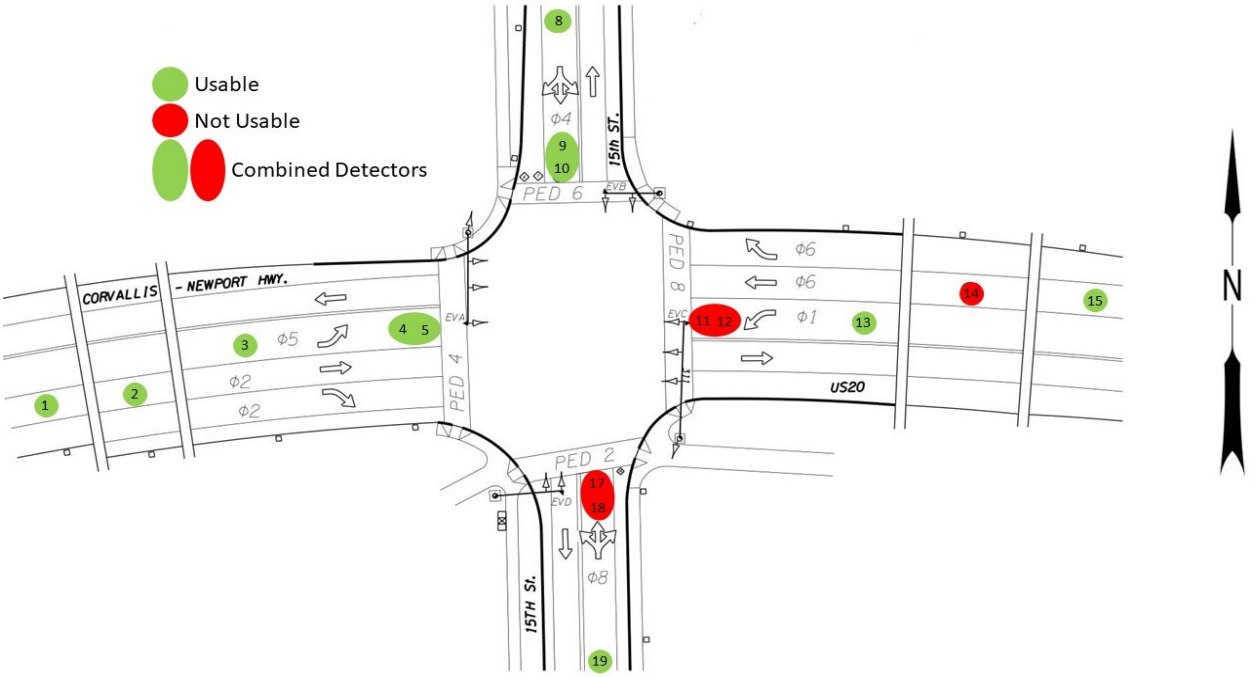


Figure 3.12: Comparative analysis results: US20 at 15th

Table 3.10: Difference Summary Table: US26 at Meinig

Det	MT	Activations			Detector On Duration Mean			Usable?
		Manual	Event Log	Difference	Manual	Event Log	Difference	
1	2	100	198	98*	00:01.33	00:01.84	00:00.52	N
2	3	100	99	-1	00:01.29	00:01.91	00:00.62**	N
3	4	10	28	18*	00:01.11	00:00.87	-00:00.23	N
4	8	35	35	0	00:04.75	00:05.70	00:00.95	Y
5-6	9	36	35	-1	00:16.65	00:16.85	00:00.20	Y
7	21	12	12	0	00:21.95	00:21.56	-00:00.39	Y
8-9	28	13	15	2*	01:10.75	01:07.01	-00:03.74	N
10	22	46	94	48*	00:11.37	00:08.23	-00:03.14**	N
11-12	23	83	48	-35*	00:18.04	00:20.86	00:02.82	N
13-14	24	45	35	-10*	00:24.41	00:24.67	00:00.26	N

* indicates a difference of >10% between the Manually reported and Event Log number of activations

** indicates Significant Difference at 95% CI in the Detector On Durations as reported by the t-Test

Green cell shading indicates the detector had neither a Significant Difference in the Detector On Durations as reported by the t-Test, nor a difference of >10% between the Manually reported and Event Log number of activations. Red cell shading indicates that the detector had one or both of these.

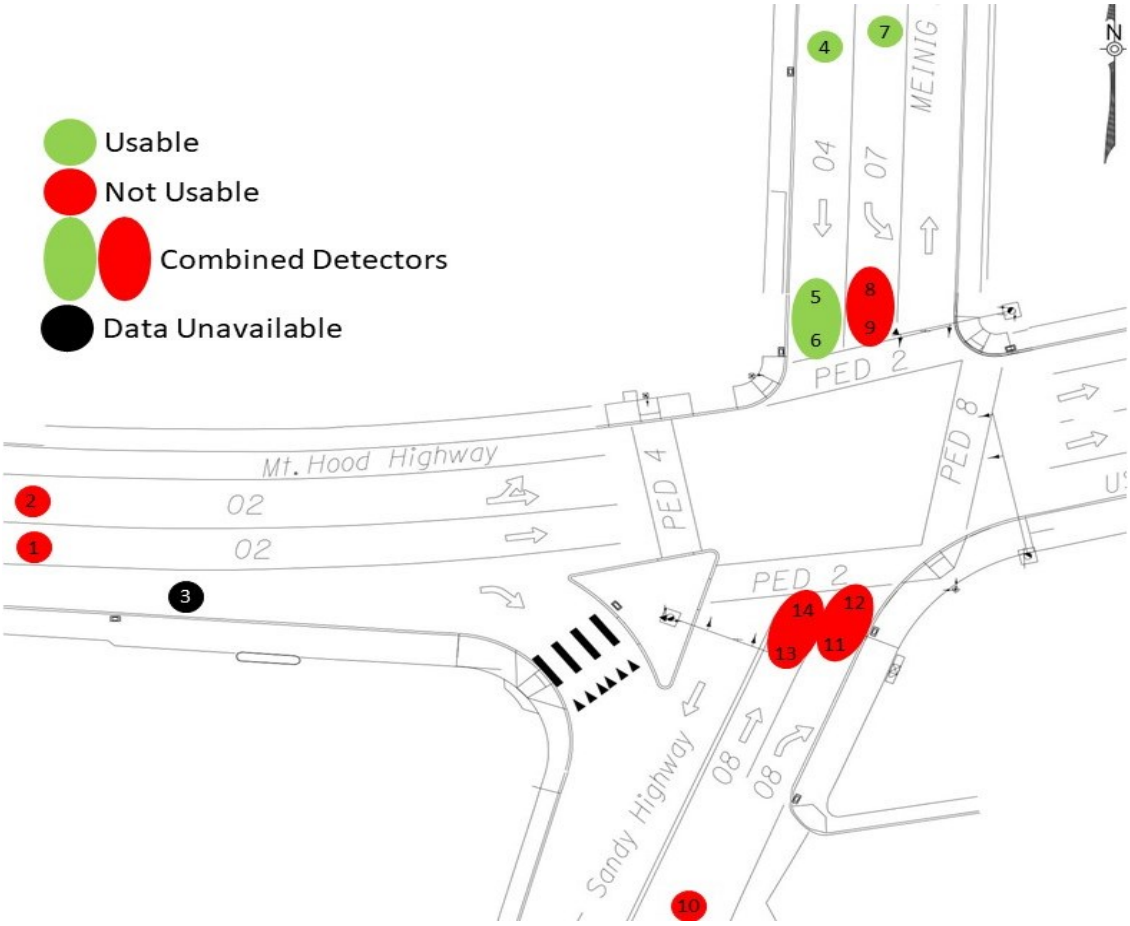


Figure 3.13: Comparative analysis results: US26 at Meinig

Table 3.11: Difference Summary Table: US101 at 22nd

Det	MT	Activations			Detector On Duration Mean			Usable?
		Manual	Event Log	Difference	Manual	Event Log	Difference	
1	2	85	94	9*	00:00.63	00:00.55	-00:00.08	N
2	3	100	101	1	00:00.75	00:00.62	-00:00.13**	N
3-4	4	100	102	2	00:02.17	00:02.02	-00:00.15	Y
5-6	5	100	100	0	00:01.85	00:01.74	-00:00.11	Y
7	15	19	20	1	00:00.90	00:00.86	-00:00.04	Y
8-9	27	19	19	0	00:05.76	00:05.64	-00:00.12	Y
10	8	26	37	11*	00:01.62	00:02.66	00:01.04**	N
11-12	9	22	17	-5*	00:07.63	00:14.12	00:06.49**	N
13-14	10	17	21	4*	00:21.25	00:06.86	-00:14.39**	N
15	16	90	139	49*	00:00.59	00:00.57	-00:00.02	N
16	17	100	64	-36*	00:00.56	00:00.56	00:00.01	N
17-18	18	86	90	4	00:03.50	00:03.13	-00:00.37	Y
19-20	19	100	101	1	00:02.08	00:01.96	-00:00.12	Y
21	1	3	3	0	00:01.01	00:00.83	-00:00.18	Y
22-23	13	7	9	2*	00:12.08	00:09.36	-00:02.72	N
24	22	13	14	1	00:01.89	00:02.08	00:00.18	Y
25-26	23	10	11	1	00:34.77	00:31.18	-00:03.59	Y

* indicates a difference of >10% between the Manually reported and Event Log number of activations

** indicates Significant Difference at 95% CI in the Detector On Durations as reported by the t-Test

Green cell shading indicates the detector had neither a Significant Difference in the Detector On Durations as reported by the t-Test, nor a difference of >10% between the Manually reported and Event Log number of activations. Red cell shading indicates that the detector had one or both of these.

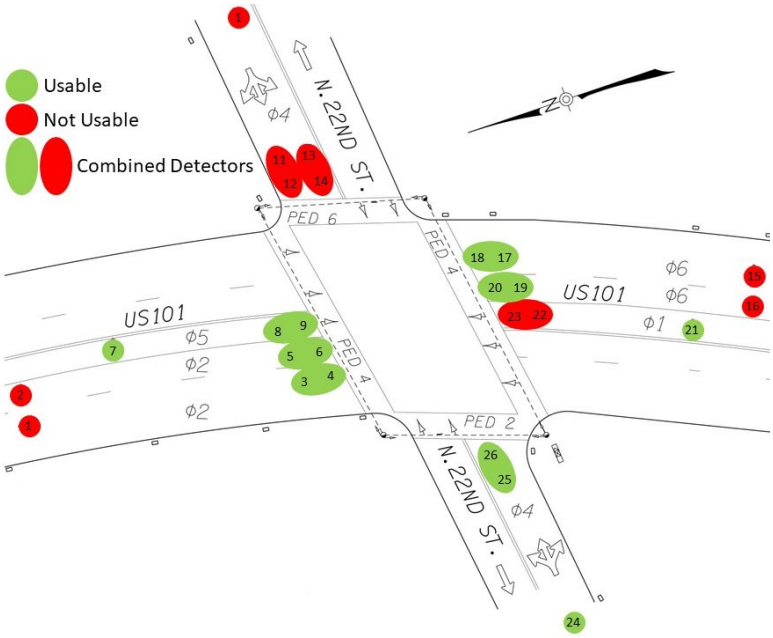


Figure 3.14: Comparative analysis results: US101 at 22nd

4.0 DATA ANALYSIS AND ALGORITHM DEVELOPMENT

4.1 INTRODUCTION

This section will describe the steps taken to develop the algorithm used to assess detector health. The general approach used was to link traffic detector data with traffic theory, using statistical rigor but with an eye for ease of implementation across the ODOT system. As part of this work, several performance datasets were developed for use in the algorithm's comparative analytics. These performance datasets will also be described in this chapter. Lastly, the algorithm itself will be presented, with step-by-step instructions tied back to the analytic work covered earlier in the chapter.

4.2 GENERAL FORM OF COMPARATIVE PROCESS

The overall comparative process for the algorithm is illustrated in Figure 4.1. There are two separate health assessments that are part of the outcomes of this work. First, as shown by item '1' in Figure 4.1, an initial detector health assessment is undertaken. Site and detector data from the detector in question are used to develop two separate comparative points for initial health assessment, both of which use the mathematical concept of integration and Greenshields's volume / density fundamental diagram (Figure 2.5c) as the basis for comparison. Comparative point one of the initial health assessment for a detector involves the development of a conceptual volume / density curve based upon site characteristics and detector data, development of a predicted volume / density curve based upon the outcome of modeling in this work (to be discussed in a successive section) and site characteristics, and then comparing the difference of integral values of those two curves to a performance dataset developed for this work (PPD, Predicted Performance Dataset). Figure 4.2 and Equation 4-1 illustrates this comparison, showing how the 'percent difference' value is generated from the predicted and the conceptual curves.

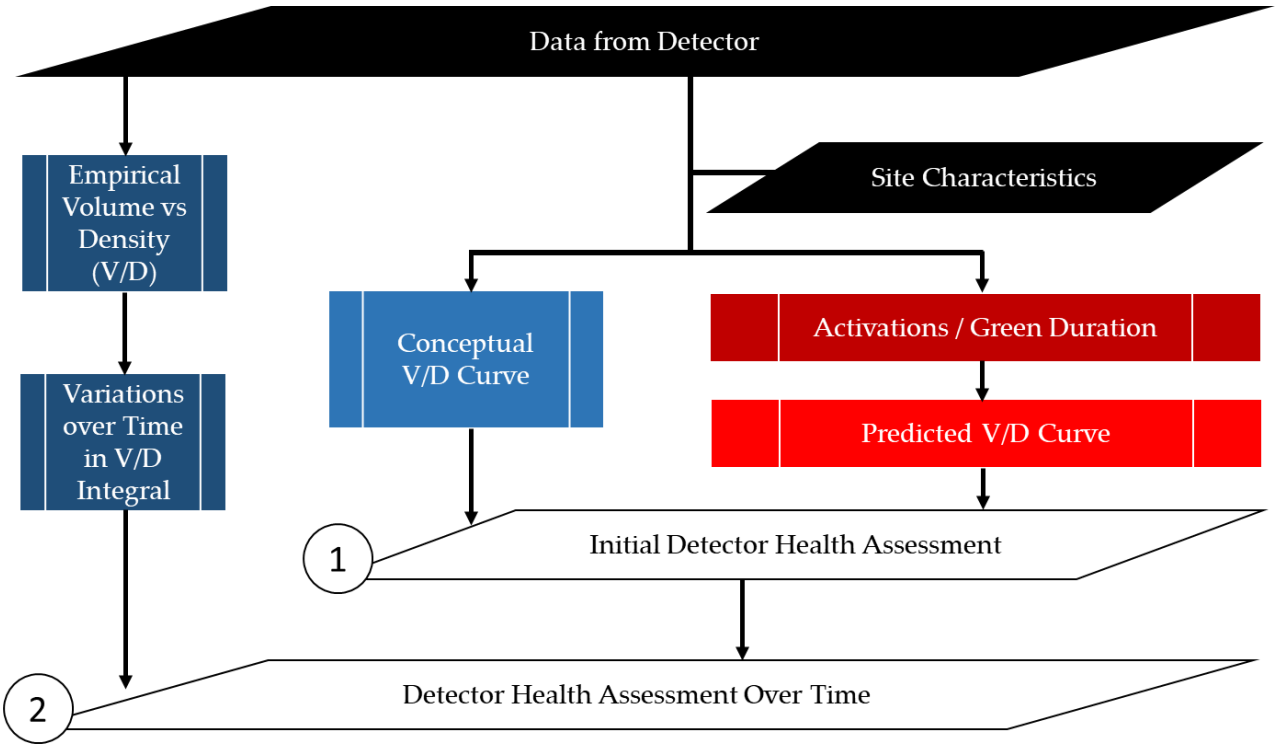


Figure 4.1: Data analysis flowchart

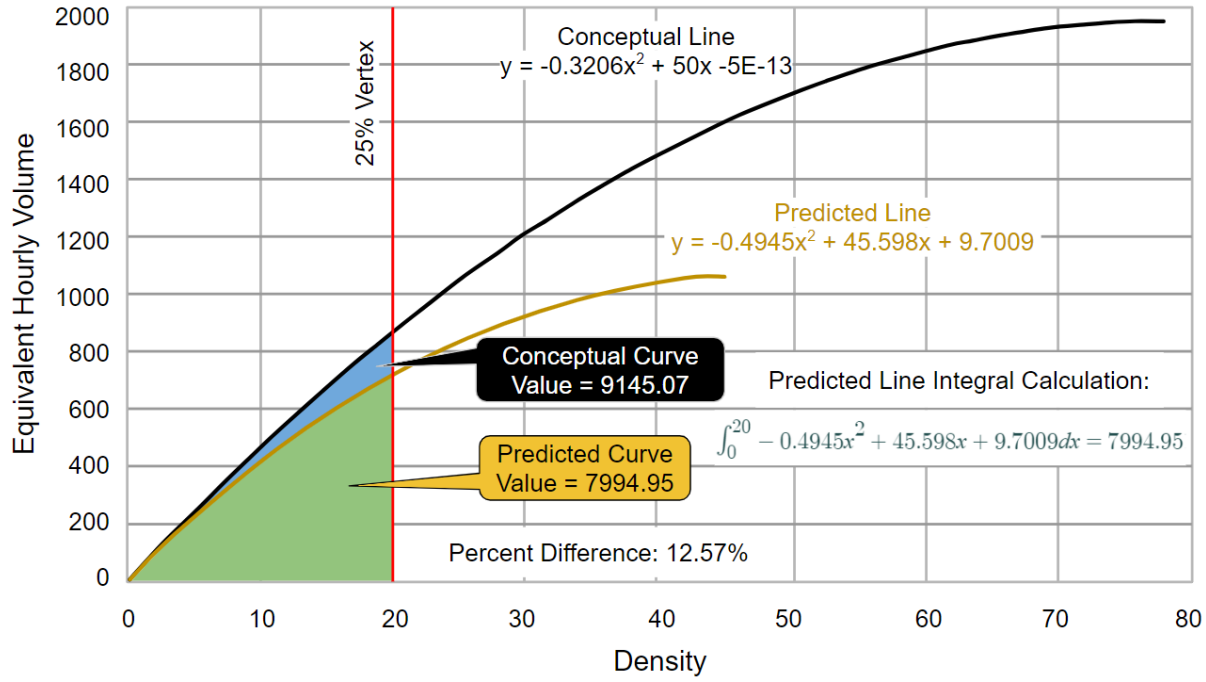


Figure 4.2: Integral percent difference illustration

$$\left|\frac{\boldsymbol{Predicted\ Integral - Conceptual\ Integral}}{\boldsymbol{Conceptual\ Integral}}\right| = \boldsymbol{Percent\ Difference}$$

(4-1)

The second comparative point of the initial health assessment involves comparing the week-to-week percent difference values between integrals of four weeks of empirical volume / density curves to a second performance data set (EPD, Empirical Performance Dataset) developed as part of this work.

Item '2' in Figure 4.1 is the second health assessment, that of detector health over time. For this assessment, percent difference values between integrals are computed from empirical data in rolling four-week increments. These values are then plotted on a control chart and compared with the EPD. Over time, the control chart limits can be adjusted based upon current data.

The crux of these comparisons is the approximation of uninterrupted flow from detector data on an interrupted flow facility. To approximate this type of flow, several filtering techniques were applied to the raw data to remove vehicle actuations that are likely occurring outside of uninterrupted flow, allowing the isolation of the saturated portion of queue discharge to be used for this approximation. While it is expected that the resultant empirical curves will be the same general shape as a conceptual Volume vs. Density curve, it is not expected that they will closely approximate the x- and y-dimensions of such a curve.

4.3 DATA CLEANSING AND APPROXIMATION OF UNINTERRUPTED FLOW

ODOT personnel provided approximately 19 weeks of raw data for the six intersections. The first step in reduction was to remove any spurious data that would hinder data processing. There are two known issues with the provided event-based data, that of repeated 'Detector On' events for the same detector, and that of repeated 'Green On' interval data events. Any repeated 'Vehicle Detector On' and 'Vehicle Detector Off' events were removed to allow for proper processing of the data, while the complete cycle was removed for any cycle that contained a repeating Green Time Start and Yellow Time Start. This was done to ensure that green durations and their related volume characteristics were consistent from cycle to cycle. These steps were completed for each detector that passed the performance heuristic applied in Section 3.3 of this work. Data from these detectors were used to build the relationships and algorithms documented in the subsequent sections.

4.3.1 Peak Period Selection

As noted earlier, a critical step in the modeling completed in this work is the approximation of uninterrupted flow from detector data, which is most likely to occur during the saturated flow period of queue discharge. To achieve this end, several filtering techniques were applied to remove data that were unlikely to reflect saturated conditions. Given that saturate discharge is most likely to occur during periods of high volumes, it was determined that the data from peak commuting periods (Tuesday, Wednesday, and Thursday, from 6:00 AM – 9:00 and 4:00 PM –

7:00 PM) would be used for analysis, as these time periods would likely provide the highest proportion of intervals exhibiting this type of discharge.

4.3.2 Start-up Vehicles

Each time an indication changes from red to green, and a stopped queue is transitioned to a moving, saturated queue, lost time is incurred. Equation 4-2 and 4-3 show the pertinent equations from the Highway Capacity Manual for this concept, with the former illustrating where the concept of lost time falls within the context of the effective green calculation, and the latter noting how many vehicles of the queue contribute to start-up lost time. Based upon this theory, the first four vehicles of each cycle would were removed from the analysis as they exhibit headways which are larger than headways expected during saturated flow (Transportation Research Board 2016). This process was undertaken for each traffic signal cycle of data analyzed.

$$\boldsymbol{g = G + Y + R - (l_1 + l_2)} \tag{4-2}$$

where:

g = Effective Green Time (seconds)

G = Actual Green Interval (seconds)

Y = Actual Yellow Change Interval (seconds)

R = Actual Red Clearance Interval (seconds)

l_1 = Start-up Lost Time (seconds)

l_2 = Clearance Lost Time (seconds)

$$\boldsymbol{l_1 = \sum_{i=1}^{4} t_i} \tag{4-3}$$

where:

l_1 = Start-up Lost Time (seconds)

t_i = Lost time for ith vehicle in queue (seconds)

4.3.3 Saturated Headway Value

Additionally, activations with headways above a certain threshold would need to be removed from the analysis to ensure approximation of saturated flow. As part of the process of

determining a cutoff headway value (headways above this cutoff value would be removed from the analysis), headways were evaluated per vehicle position in the queue for each green-phase interval. Figure 4.3 illustrates three hours of afternoon peak period headway data for one detector, sorted by vehicle position (vehicles in positions 1-4 were removed from the visualization, per section 4.3.2).

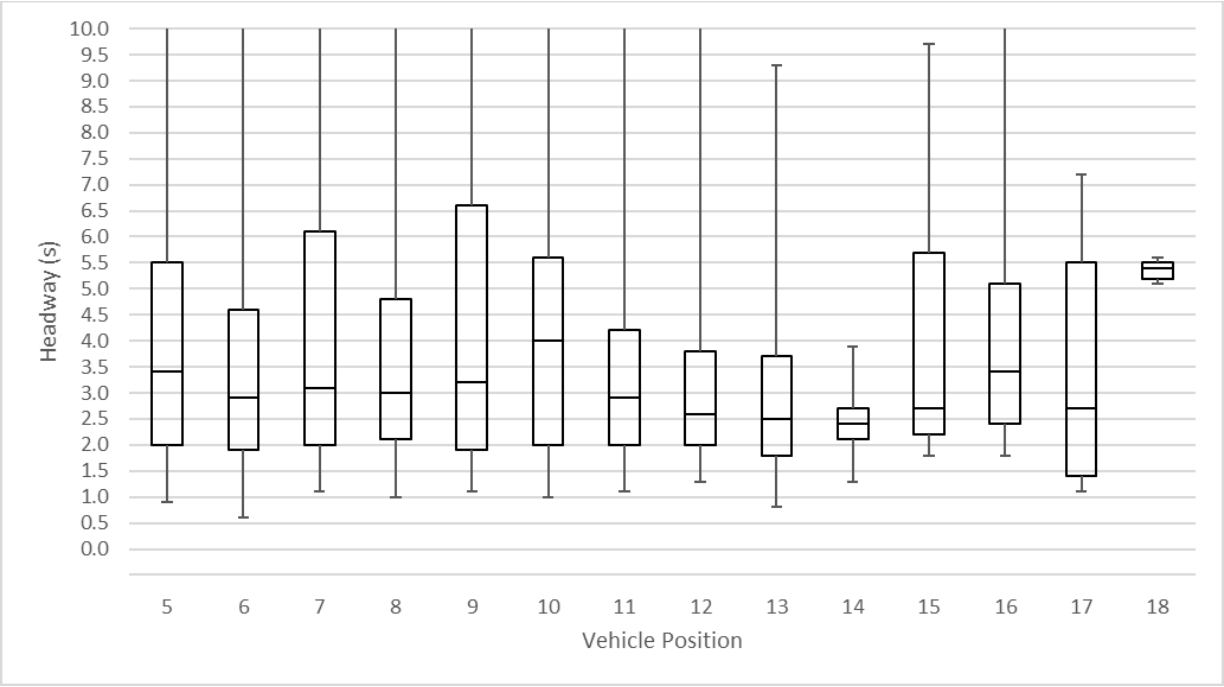


Figure 4.3: Headway data per vehicle position, no data removed, for one detector

Of noted in Figure 4.3 is that the mean headway of each vehicle position, with the exception of positions 10 and 18, is between 2.0 and 3.5 seconds, reasonable values for a saturated headway based upon field conditions. With this knowledge, several exploratory techniques were applied to the data to determine a proper headway cutoff value. First, the top quartile of headways for each position was remove, the results of this are plotted in Figure 4.4 (this is from the same data used to develop Figure 4.3).

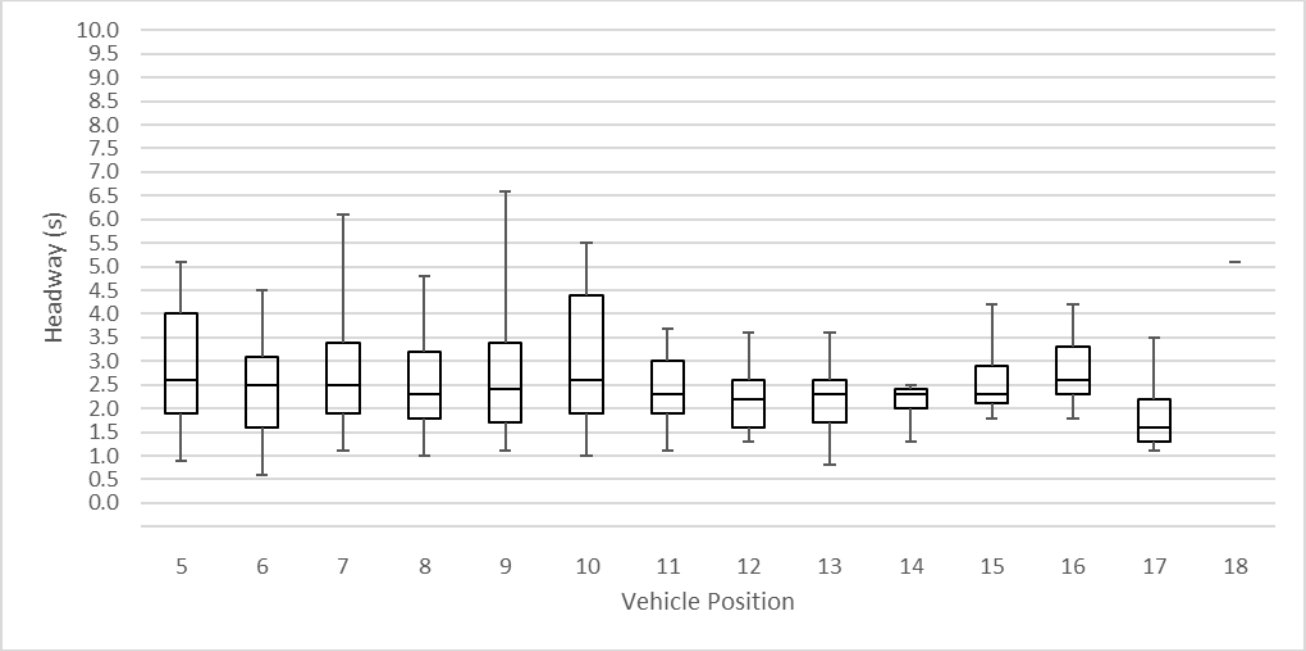


Figure 4.4: Headway data per vehicle position, top quartile of headway data removed from each vehicle position, for one detector

While the means of the vehicle positions are lower, showing that some of the larger headways were removed, the resulting data still contains many headways that are not indicative of saturated flow, which could negatively impact the data analysis. Several other approaches were applied to limit the number of headways in the data that would represent non-saturated flow, including removing headways larger than twice or three times the median headway for each position, removing activations that are detected within the last 6.0 seconds before the yellow-phase interval, and removing the data from the entire green-phase interval if the first vehicle's headway was greater than 8.0 seconds. In the end, it was determined that limiting the headways used in this work to 3.0 seconds, a common value used in setting gap timers within actuated control, would more effectively accomplish the desired outcome without the need to carry out complicated mathematical procedures. This approach has the added benefit of including activations later in the green interval as part of the data analysis, as saturated conditions can occur well into stale green. The activations remaining after filtering for Start-Up Lost Time and headways is referred to as filtered activation data.

4.4 CALCULATING EQUIVALENT HOURLY VOLUME AND DENSITY FROM FILTERED DATA

With data filtered to remove spurious data and approximate saturated flow, the next step is to calculate volume and density values for each cycle of data, which can then be plotted and used to approximate a Volume vs. Density relationship. For volume, Equivalent Hourly Volume (EHV) was used, a concept which uses volume (or counts) from a time period shorter than one hour that is then scaled to reflect an hourly volume. Equation 4-4 shows the calculation for EHV, using the number of activations during green as a surrogate for departure volume. It is recognized that several vehicles may arrive on red, and that these are not captured in departure volume. This,

along with the filtering techniques described earlier in this chapter, will likely result in volumes that are lower than actuality, however it is believed that the constant application of these methods across all data processed will limit any additional variability these lower volumes might introduce into the process. EHV was calculated for each cycle.

$$\boldsymbol{EHV = {3600}/{(3600 \times 24 \times C)(A)}}$$

(4-4)

where:

EHV = Equivalent Hourly Volume

C = Cycle Duration

A = Number of Filtered Activations per Green Duration

In determining density, occupancy on a per cycle basis was used. Equation 4-5 shows the calculation for density. Occupancy was calculated using Equation 4-6. A density value was calculated for each cycle.

$$\boldsymbol{D = {O \times 5280}/{(L_{Veh} + L_{Det})}}$$

(4-5)

where:

D = Density

O = Occupancy

L_{Veh} = Average Vehicle Length

L_{Det} = Detector Length

$$\boldsymbol{Occupancy = \frac{Filtered\ Detector\ On\ During\ Green\ \ Duration}{Green\ Duration}}$$

(4-6)

Application of Equation 4-6 required several assumptions / inputs. First, the Vehicle Length is assumed to be the design passenger vehicle length of 19 feet (American Association of State Highway and Transportation Officials 2011). Second, Detector Length is measured from the field (or approximated in the case of radar detection zones). Lastly, a constant vehicle length was presumed to present in the traffic stream. These assumptions are commonly used in the calculation of density from occupancy data.

4.5 APPROXIMATED VOLUME DENSITY RELATIONSHIP

Once the processing and calculations described in Section 4.3 through 4.4 are completed for an entire week of data (Tuesday, Wednesday, and Thursday from 6:00 AM – 9:00 AM and 4:00 PM – 7:00 PM), each individual cycle's EHV and density values are plotted and an empirical line of best fit is created by applying a quadratic best fit line to the plotted data, as shown in Figure 4.5. This amounts to one week of data, a quantity that will be used later in the algorithm description. The Coefficient of Determination (R^2) for the fit of the line to the data is also shown, as well as a conceptual Volume vs. Density line derived from site information for that detector, to be described shortly (Site 245, Detector 1).

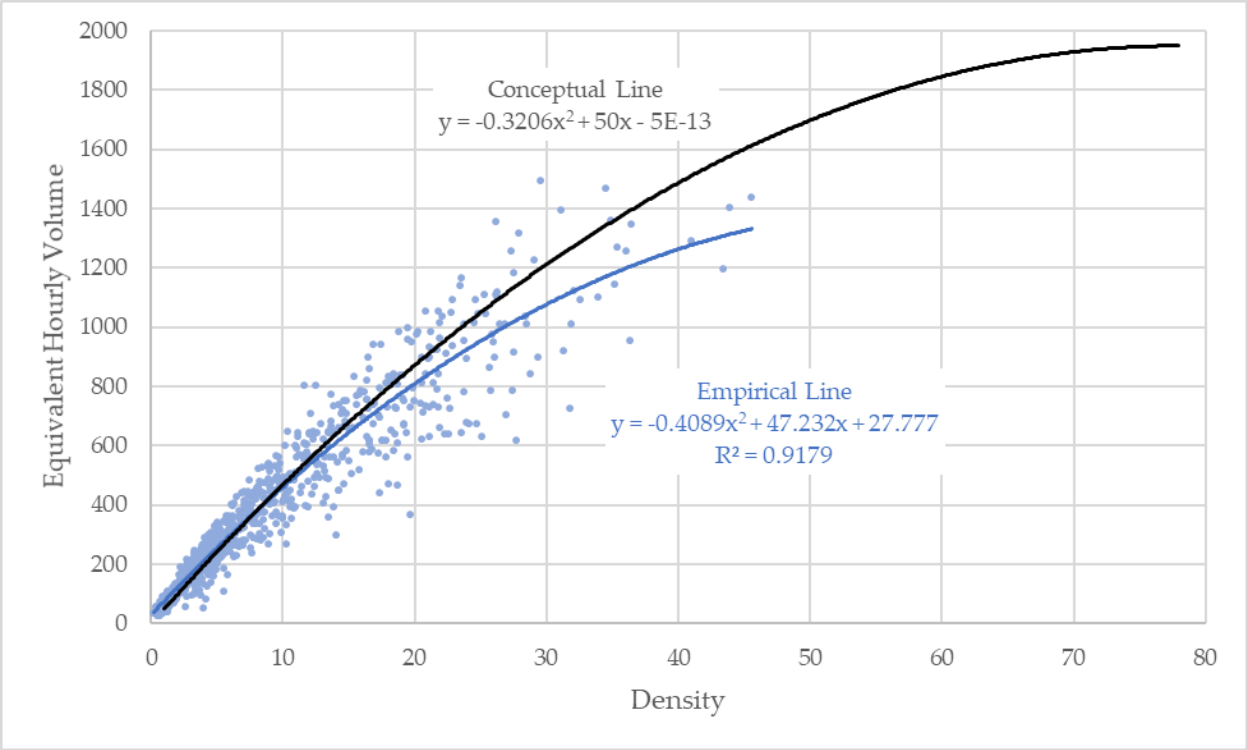


Figure 4.5: Example empirical line and empirical data, and conceptual line

It should be noted that for weeks of data with less than 50 data points, an empirical line was not developed as the number of points was insufficient to develop a relationship (this value was determined through visualization and judgment by the research team). Additionally, any empirical line with a R^2 of less than 0.70 was not used for further analysis, as an R^2 less than 0.70 generally indicates a weak or moderate fit of the trendline to the underlying data. Four weeks of data were processed for each detector, resulting in a total of 100 weekly empirical lines developed across the data set for use in further analysis.

Also shown in Figure 4.5 is a conceptual Volume vs. Density quadratic relationship for the detector from which this empirical data was collected. This relationship was developed for each detector using the Optimum Density (Equation 4-7) and Maximum Equivalent Hourly Volume (Equation 4-8), per Greenshield's relationship (Equation 2-1). Maximum Volume was

approximated by using an average headway for the approach; Optimum Speed was approximated using the speed limit for the approach.

$$D_O = \frac{V_{Max}}{S_O} \tag{4-7}$$

where:

D_O = Optimum Density

V_{Max} = Maximum Volume

S_O = Optimum Speed (½ of Posted Speed Limit)

$$V_{Max} = \frac{3600}{Average\ Headway} \tag{4-8}$$

where:

V_{Max} = Maximum Volume

While the Optimum Speed was determined to the ½ the posted speed limit, an approximation directly derived from Greenshield's work, several sets of data for a detector were analyzed to determine the most effective method of calculating the average headway to be used for calculating Maximum Volume. Table 4.1 shows the average headway for a detector for various days and time periods of analysis (each day consists of six hours of peak period data, as has been used previously in this work). Given the relatively small spread in the average headway from the various windows of data analyzed, and the fact the event-based data is collected at a 0.1s resolution, it was determined that using the average headway for the first day of data would be sufficient and allow for data to be analyzed quickly.

Table 4.1: Average Headways for Sensitivity Analysis

Date	Average Headway (seconds)
Jan 12	2.11
Jan 12-14	2.08
Jan 19-21	2.08
Feb 2-4	2.06
Jan 12-14, Jan 19-21	2.08
Jan 12-14, Jan 19-21, and Feb 2-4	2.07

4.5.1 Outlier Control Experimentation

To improve the correlation between EHV and Density and reduce variability in the processed data, an experiment with outlier control was performed. The method used for this experiment was the interquartile range method, which removes data points located at the ends of the spectrum of plotted data. In this method, the interquartile range (IQR) (i.e., the 75th percentile – the 25th percentile of the data set) is calculated, and any point falling below the 25th percentile – 1.5*IQR or above the 75th percentile + 1.5*IQR is removed as an outlier. Figure 4.6 and Figure 4.7 illustrate the effects of this interquartile range outlier control, with the former figure showing all data points for one week of collected data, and the latter showing on those remaining after the outliers were removed. Note that the conceptual line is unchanged between the two plots.

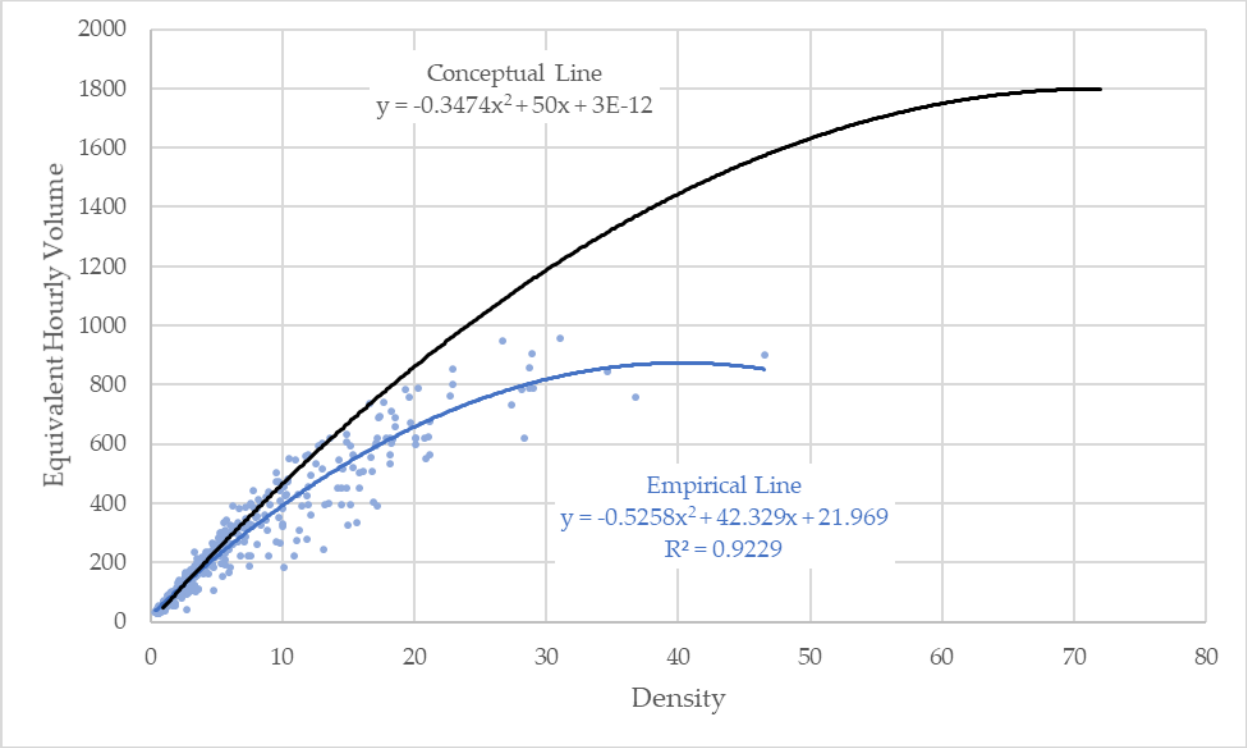


Figure 4.6: Example empirical line and empirical data without outlier removal

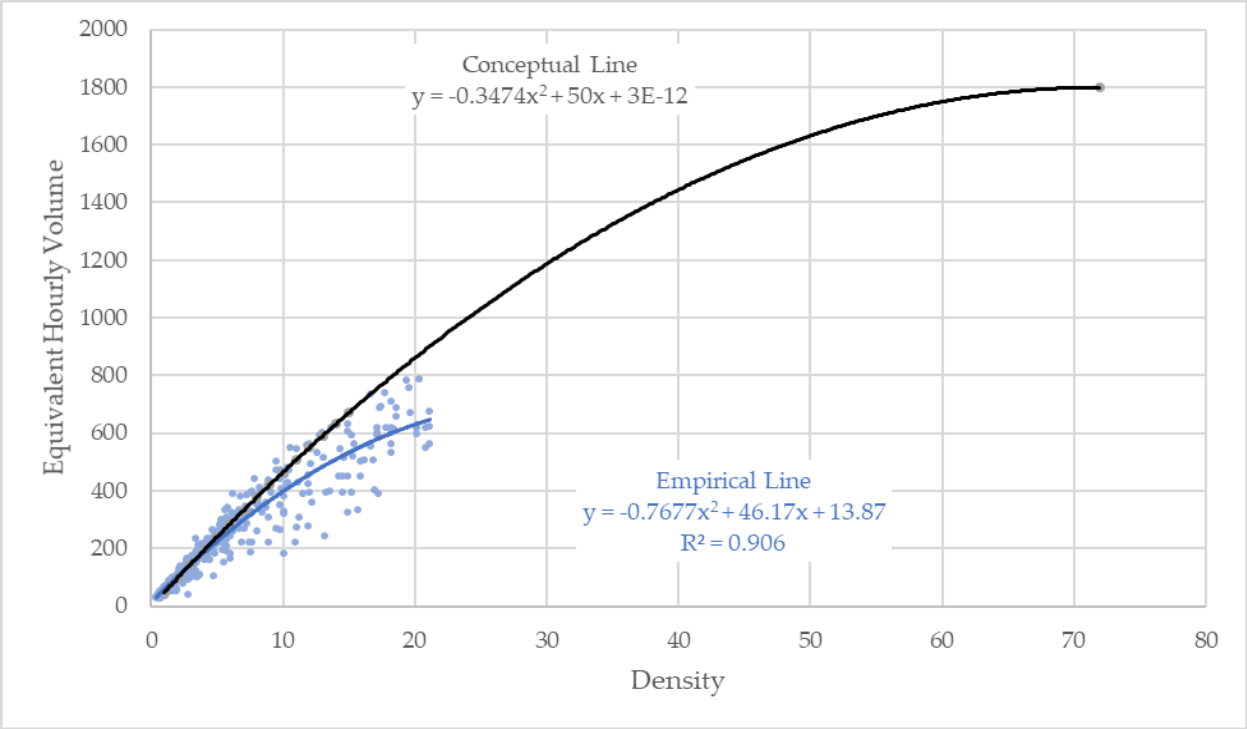


Figure 4.7: Example empirical line and empirical data with the outliers removed

As shown in Figure 4.6 and Figure 4.7, removing the Volume and Density outliers through the Interquartile Range process changes the empirical line of best fit as well as the R^2 value. The IQR method was applied to all processed empirical lines, with summary data shown in Table 4.2. Applying this method slightly reduced both the mean and standard deviation of the processed data. However, it was noted that the R^2 value of the best fit line was generally reduced, and because of this, it was determined that the benefit of the higher R^2 outweighed the reduced variability in the dataset. Because of this, the results of this outlier removal process were discarded, and not used in the overall analysis.

Table 4.2: Mean and Standard Deviation of the Equivalent Hourly Volume and Density Empirical Data with Outlier Removal

	No Outliers Removed		**EHV after EHV Outliers Removed**	**Density after Density Outliers Removed**	**Combined EHV and Density Outliers Removed**	
	EHV (veh/hr)	**Density (veh/ln-mi)**	**EHV (veh/hr)**	**Density (veh/ln-mi)**	**EHV (veh/hr)**	**Density (veh/ln-mi)**
Mean	260	6.82	248	5.99	238	5.99
Standard Deviation	212	7.05	197	5.53	184	5.53
Sample Size	402	402	395	388	387	388

4.6 GREEN ACTIVATIONS

The number of detector activations per green duration was also complied from the initial data processing. These data are derived after the raw data have been cleansed of redundancies, but do not incorporate the filtering process for approximating uninterrupted traffic flow. For simplicity's sake, we name the number of detector activation per green duration as green activation in this report.

We used four weeks' data from 6 AM to 9 AM and from 4 PM to 7 PM on Tuesdays, Wednesdays, and Thursdays. The data of green duration and green activation were plotted for each selected detector (or detector group) at all six intersections listed in Chapter 3. The regression modeling was done for each detector (or detector group) to produce the relationship between the green activation and green duration.

At the intersection of Mission Street and I-5 Southbound Off Ramp (as known as Site 245 in this report), detector 7 is an advance loop detector on the left-turn lane of the I-5 Southbound Off-Ramp, as well as detector 8 on the other left-turn lane of the same off-ramp from the freeway. Figure 4.8 through Figure 4.14 show that the number of activations increase with the green duration for detector 7. For days from January 12 to 14 (Tuesday to Thursday), 2021, the R^2 is 0.76. But when combining two weeks' data (the week of January 12 and the week of January 19), the R^2 is 0.39, which is not significant. When combing three weeks' data, the R^2 is 0.28, which is not significant either.

The detector 8 is on the adjacent left-turn lane of the same off-ramp. The regression modeling results of green activations and green durations for detector 8 are similar to the results of detector 7. For simplicity, we do not show the figures for detector 8 here.

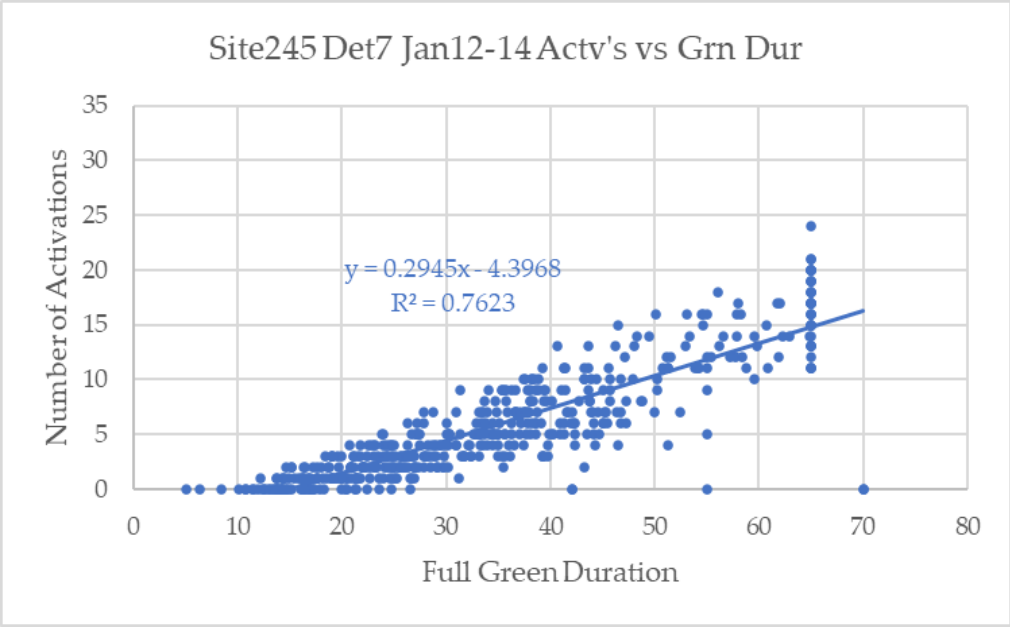


Figure 4.8: Green activation and green duration for detector 7 for the week of January 12

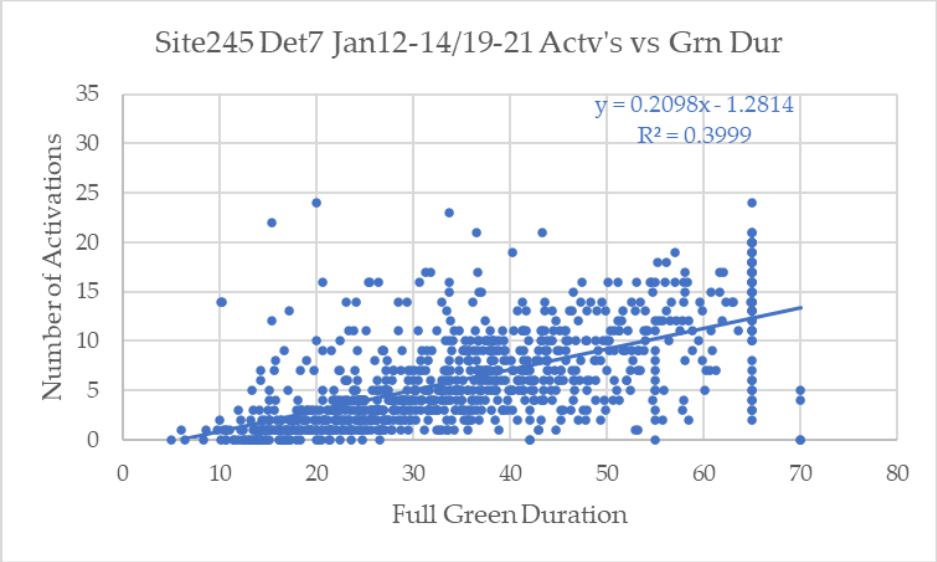


Figure 4.9: Green activation and green duration for detector 7 for two weeks

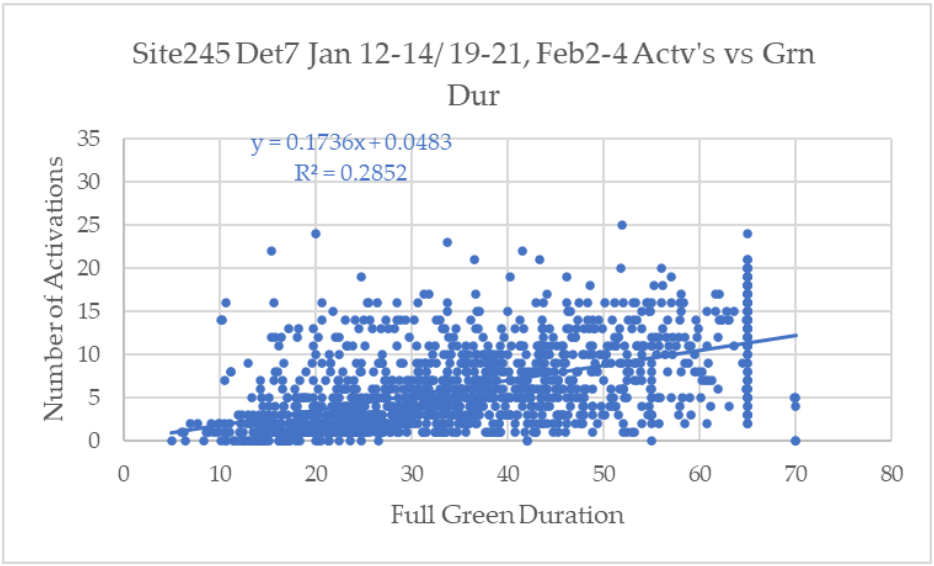


Figure 4.10: Green activation and green duration for detector 7 for three weeks

Detector group 9-10 has two advance loop detectors 9 and 10, downstream of detectors 7 and 8 on the same off-ramp. We have data for the detector group but not for detector 9 or 10 separately. The following figures show that the number of activations increase with the green duration for the detector group, but the R^2 values are close to 0.2, which is not significant.

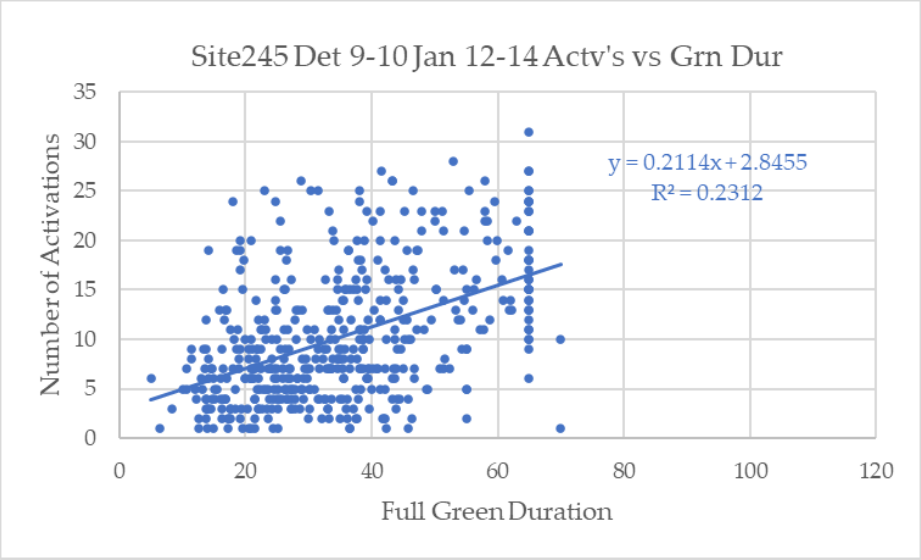


Figure 4.11: Green activation and green duration for detector Group 9-10 for the week of January 12

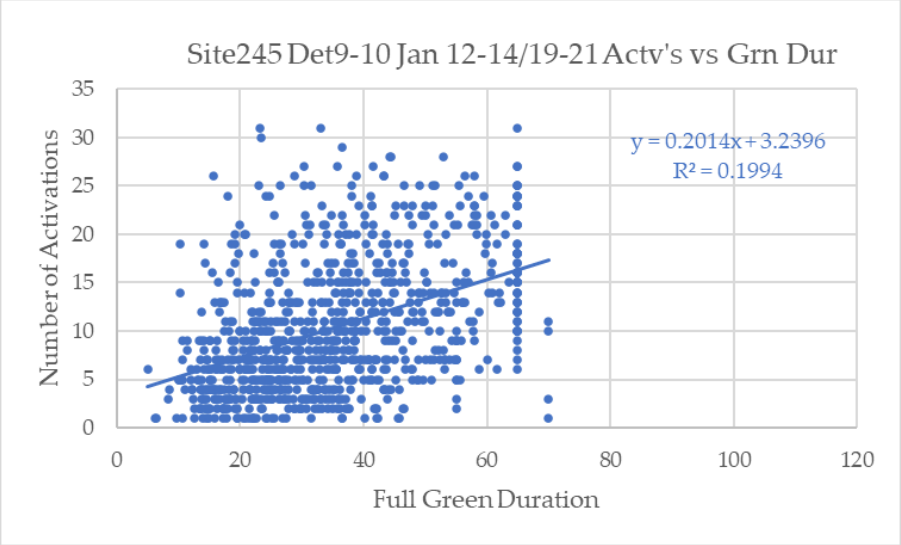


Figure 4.12: Green activation and green duration for detector group 9-10 for two weeks

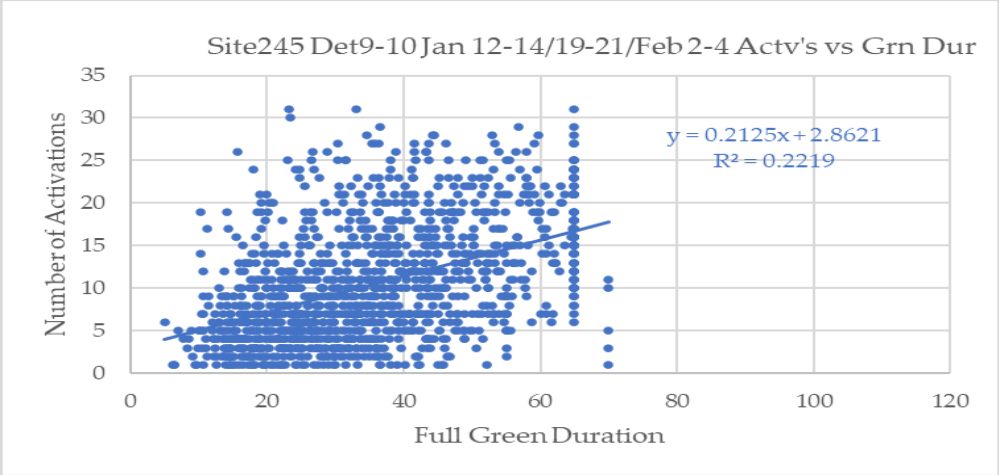


Figure 4.13: Green activation and green duration for detector group 9-10 for three weeks

The regression modeling results for the other 42 detectors (or detector groups) show that the R^2 values are less than 0.4, which is not significant. In addition, some regression models' coefficients for the variable of green duration are close to 0, which is not statistically significant. For simplicity, we do not present all the regression modeling results in this report but show the following figure as an example. Therefore, we have not found statistically significant relationships between the green activation and the green duration from the collected data for most of the detectors.

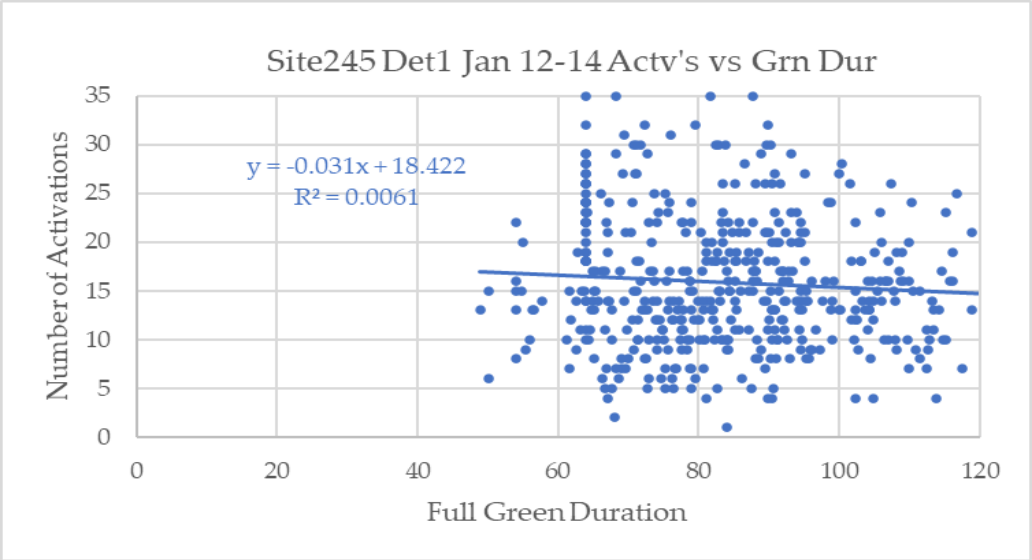


Figure 4.14: Green activation and green duration for detector 1 for the week of January 12

4.7 PREDICTED VOLUME/DENSITY RELATIONSHIP AND PERFORMANCE DATA SETS

With the processes identified for generating conceptual and empirical Volume vs. Density relationships, the next step is to develop the modeling for predicted Volume vs. Density

relationships, a Predicted Performance Dataset (PPD), both necessary for the Initial Detector Health Assessment (Item 1 in Figure 4.1), and an Empirical Performance Dataset (EPD), necessary for both the Initial Health Assessment and Detector Health Assessment Over Time (Item 2 in Figure 4.1). The process used for modeling will be described first.

4.7.1 Predicted Volume/Density Relationship Modeling

In developing a model for predicting Volume vs. Density relationships for a specific site, the choice of an appropriate modeling technique was contingent on having sufficient data observed at each location, including that these data exhibit a downward parabolic shape when fit to a quadratic curve, additional site-level data available to explore how independent factors may contribute to observed volume-density curve variation is available, and using a model structure that accounts for the interdependence between observed *a*, *b*, and *c* terms in the quadratic formula in predicting empirical lines at each location. The form of a quadratic expression is shown in Equation 4-9.

$$\boldsymbol{y = ax^2 + bx + c} \tag{4-9}$$

An initial step in the modeling process was the construction of a data set with sufficient content at each detector across the four weeks of collected data. Accordingly, a potentially complete sample of 180 records (four weeks of empirical data collected at 45 detectors) was reduced to a sample of 106 records after removing records with empirical data fewer than ten points (74 records). Of this remaining sample, five records were removed in which the *a* term exhibited a positive value, which would have produced an upward parabolic shape when plotted, resulting in a final sample of 101 records.

Using this reduced sample, a next analytic step was to specify multiple regression models where observed values of *a*, *b*, and *c* were functions of various site characteristics that may account for variation in these outcome variables. Given the inter-relationship between the three outcome variables and a desire to model a single set of predictor variables, a multivariate multiple regression modeling structure was chosen. Specification of predictors in this simultaneous model of multiple outcomes was pursued with an iterative process that first assessed statistically significant intergroup (or model outcome) differences in the effects of independent predictors and then examined the explanatory power of any selected predictor variables. Regarding the former assessment, a multiple analysis of variance (MANOVA) was conducted to test differences in mean values of *a*, *b*, and *c* terms per location across several categorical site characteristics. Table 4.3 summarizes the results of this analysis in which significant variation was found for mean values of the *a*, *b*, or *c* terms in each of the tested categorical variables except for detector length. Of note, continuous measures of green activation and detector indication were examined as binary variables in the MANOVA, with low and high values based on relationship of locational measure with mean value of variables within the full final sample.

Table 4.3: Descriptive Statistics for Four-Week Sample of Detector Summary Data

	a				*b*				*c*			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
Week 1 (n=25)	-0.391	0.580	-2.876	-0.003	32.773	15.300	1.572	60.183	18.129	33.676	-43.426	151.920
Week 2 (n=25)	-0.290	0.280	-1.203	-0.003	31.249	14.761	1.838	60.204	24.368	33.496	-4.092	168.150
Week 3 (n+26)	-0.282	0.232	-0.964	-0.005	30.074	14.492	2.447	55.122	28.731	36.060	0.192	182.680
Week 4 (n=25)	-0.370	0.417	-1.838	-0.003	32.322	19.095	0.440	76.858	26.243	23.622	-2.049	110.320
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
0: Tech loop (n=21)	0.108	0.150	-0.581	-0.006	20.941	8.520	7.723	48.414	19.510	12.389	-0.568	47.687
1: Tech Loop (n=80)	-0.392	0.420	-2.876	-0.003	34.385	16.109	0.440	76.858	25.697	35.202	-43.426	182.680
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
0: Detect Advance (n=44)	-0.157	0.300	-1.838	-0.003	21.384	9.031	0.440	48.414	18.654	22.296	-20.946	110.300
1: Detect Advance (n=57)	-0.468	0.410	-2.876	-0.003	39.467	15.430	1.572	76.858	28.855	37.235	-43.426	182.680
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
0: Main Street (n=29)	-0.246	0.526	-2.876	-0.003	21.538	11.337	1.572	53.939	31.962	51.713	-43.426	182.680
1: Main Street (n=72)	-0.367	0.328	-1.838	-0.003	35.638	15.587	0.440	76.858	21.36	18.539	-4.092	110.320
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
0: Single Lane (n=16)	-0.350	0.212	-0.821	-0.053	43.523	17.286	15.130	76.858	34.750	21.327	-20.946	77.010
1: Single Lane (n=85)	-0.329	0.423	-2.876	-0.003	29.343	14.545	0.440	60.204	22.465	33.228	-43.426	182.680
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
0: >10-ft Length (n=66)	-0.358	0.421	-2.876	-0.003	31.313	15.011	1.572	60.204	24.173	34.942	-43.426	182.680
1: > 10-ft Length (n=35)	-0.284	0.344	-1.838	-0.003	32.111	17.403	0.440	76.858	24.859	25.569	-20.946	110.320
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
0: Thru Lane Only (n=34)	-0.26	0.49	-2.88	0.00	21.40	11.38	0.44	53.94	29.67	48.83	-43.43	182.68
1: Thru Lane Only (n=67)	-0.37	0.34	-1.84	0.00	36.76	15.26	2.14	76.86	21.74	18.05	-4.09	110.32
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680

	a				*b*				*c*			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
0: High Activation (n=61) ^	-0.232	0.440	-2.876	-0.003	22.304	10.298	0.440	53.939	25.475	38.839	-43.426	182.680
1: High Activation (n=40) ^	-0.486	0.254	-1.203	-0.020	45.750	11.675	15.714	76.858	22.788	16.732	-4.092	50.243
Study Sample (n=101)*	-0.333	0.396	-2.876	-0.003	31.589	15.799	0.440	76.858	24.411	31.875	-43.426	182.680
0: High Indication (n=56) ^	-0.431	0.324	-1.838	-0.053	39.609	13.711	15.130	76.858	19.696	18.306	-20.946	77.013
1: High Indication (n=45) ^	-0.210	0.445	-2.876	-0.003	21.609	12.188	0.440	53.939	30.279	42.740	-43.426	182.680

Notes:

* Reduction of complete sample (n=180) after removing records with non-negative value for a (n=72) or less than 10 observations (n=74).

^ High represents activation/indication level above the mean value for the study sample ($\bar{x}$ activations = 304.927 and $\bar{x}$ indications = 37.476).

Cells in GREEN reflect a statistically significant difference in group means ($p<0.05$).

Cells in BLUE reflect a marginally significant difference in group means ($p<0.10$).

Having established that significant mean differences across the outcomes existed for seven independent variables, a backwards elimination model specification process was undertaken to determine a consistent set of predictors in the multivariate multiple regression model used to create the predicted empirical line. The final model specification—shown in Table 4.4—was determined once the removal of a single predictor resulted in no improvement to the adjusted R^2 value of the reduced model and that the Type II MANOVA test statistic for each remaining predictor variable was marginally statistically significant ($p<0.10$). Looking at individual model performances, the overall fit for the model of the *b* term ($R^2=0.661$) was higher than the specification for the *a* term ($R^2=0.172$) and *c* term ($R^2=0.154$). The presence of advanced detection technology and the continuous green activation metric were statistically significant in each specification, with these two variables being the lone significant predictors in the *b* term model. The former predictor as well as the presence of a loop detector were only marginally significant in the *a* term model, while the number of indications, presence of advanced detector technology, and site location within a single lane roadway were all statistically significant ($p<0.05$) in the *c* term model. Equation 4-10, Equation 4-11 and Equation 4-12 present the results of the developed models for the *a*, *b*, and *c* coefficients, respectively, of the predicted curve.

Table 4.4: Multivariate Multiple Regression Model Estimates

	a			***b***			***c***		
Predictor Variable	**Beta**	**Std. Error**	**p-value**	**Beta**	**Std. Error**	**p-value**	**Beta**	**Std. Error**	**p-value**
(intercept)	0.629	0.338	0.066	6.337	8.624	0.464	-10.341	27.478	0.707
Tech Loop	-0.267	0.136	0.052	3.773	3.472	0.280	9.171	11.062	0.409
Detect Advance	-0.180	0.100	0.074	6.754	2.542	0.009	21.385	8.098	0.010
Single Lane	-0.171	0.120	0.157	4.700	3.062	0.128	-29.725	9.758	0.003
Activations	-0.001	<0.001	0.022	0.064	0.007	<0.001	-0.047	0.024	0.051
Indications	-0.008	0.006	0.162	-0.136	0.145	0.348	1.458	0.461	0.002
Model Summary									
Adjusted R^2	0.172			0.661			0.154		

$$\hat{y}_a = \mathbf{0.629} - \mathbf{0.267}\left(x_{tech_{loop}}\right) - \mathbf{0.180}\left(x_{detect_{adv}}\right) - \mathbf{0.171}\left(x_{lane_{single}}\right) - \mathbf{0.001}\left(x_{wk_{acthour}}\right) - \mathbf{0.008}\left(x_{wk\_grnhr}\right)$$

(4-10)

$$\hat{y}_b = \mathbf{6.337} + \mathbf{3.773}\left(x_{tech\_loop}\right) + \mathbf{6.754}\left(x_{detect\_adv}\right) + \mathbf{4.700}\left(x_{lane\_single}\right) + \mathbf{0.064}\left(x_{wk\_acthour}\right) - \mathbf{0.136}\left(x_{wk\_grnhr}\right)$$

(4-11)

$$\hat{y}_c = -10.341 + 9.171(x_{tech\_loop}) + 21.385(x_{detect\_adv}) - 29.725(x_{lane\_single}) - 0.047(x_{wk\_acthour}) + 1.458(x_{wk\_grnhr})$$

(4-12)

where:

$\hat{y}_a$, $\hat{y}_b$, $\hat{y}_c$ equals the predicted values of a, b, and c

x_{tech_loop} equals the presence of a loop detector (binary)

x_{detect_adv} equals the presence of advanced detector technology (binary)

x_{lane_single} equals site location within a single lane roadway (binary)

$x_{wk\_acthour}$ equals the number of activations per hour (continuous)

x_{wk_grnhr} equals the number of indications per hour (continuous)

Using this final multivariate multiple regression model specification, the final step was to predict the value of *a*, *b*, and *c* terms for each combination of detector location and week of empirical data, which would then be used to build the PPD for health comparison. Prediction of a maximum of four empirical lines per detector location was accomplished by inserting the observed value of each predictor variable in the final model specification for all records. While predictive estimates for *a*, *b*, and *c* terms using every week of recorded detector data helps to provide a more robust assessment of detector health, a location-level aggregation of these terms across the data collection period can also be useful in investigating the predictive model's performance at sites with varying characteristics not isolated in the final specification. Figure 4.15, Figure 4.16, and Figure 4.17 offer a comparison of the average predicted and observed values of *a*, *b*, and *c*, respectively, for locations in the data set over the four-week collection period. The line numbers on the y-axis correspond to specific detectors in the data analysis. Table 4.5 links each line number to a specific site and detector number. While the residual differences are generally fairly small, there are several coefficients that have larger differences, such as Lines 9 and 20 for the *a* coefficient, and Line 11 for the *c* coefficient.

Table 4.5: Line Numbers to Site and Detector Number(s)

Line Number	Site Number	Site Name	Detector Number(s)
1	245	OR22 @ I-5SB	1
2	245	OR22 @ I-5SB	2
3	245	OR22 @ I-5SB	4-6
4	245	OR22 @ I-5SB	7
5	245	OR22 @ I-5SB	8
6	245	OR22 @ I-5SB	9-10
7	245	OR22 @ I-5SB	17-18
8	502	OR34 @ I-5SB	2
9	502	OR34 @ I-5SB	7
11	502	OR34 @ I-5SB	9
12	502	OR34 @ I-5SB	13
13	502	OR34 @ I-5SB	14
14	502	OR34 @ I-5SB	23
15	502	OR34 @ I-5SB	25
16	502	OR34 @ I-5SB	27
17	502	OR34 @ I-5SB	28
18	55	OR34 @ Peoria	3-4
20	55	OR34 @ Peoria	16
24	55	OR34 @ Peoria	24
25	503	US20 @ 15th	1
26	503	US20 @ 15th	2
32	503	US20 @ 15th	15
35	585	US26 @ Meinig	5-6
37	2	US101 @ 22nd	3-4
38	2	US101 @ 22nd	5-6
41	2	US101 @ 22nd	18
42	2	US101 @ 22nd	9

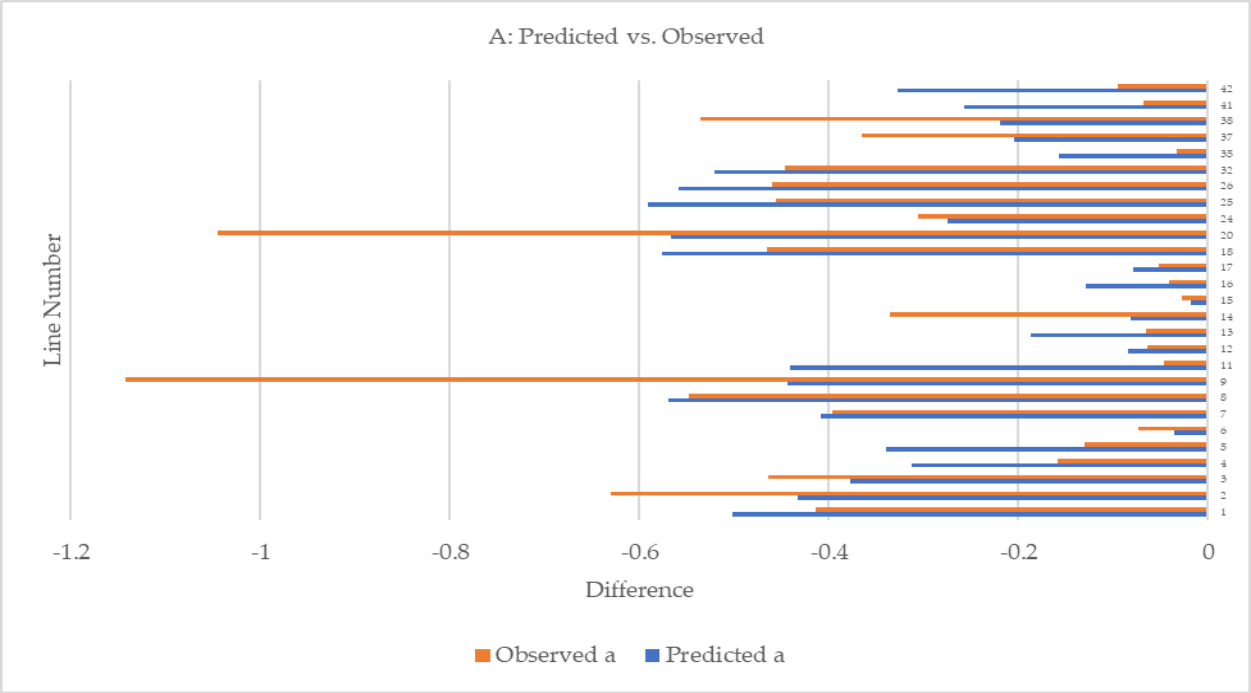


Figure 4.15: Comparison of mean observed and predicted values of *a* term over data collection period

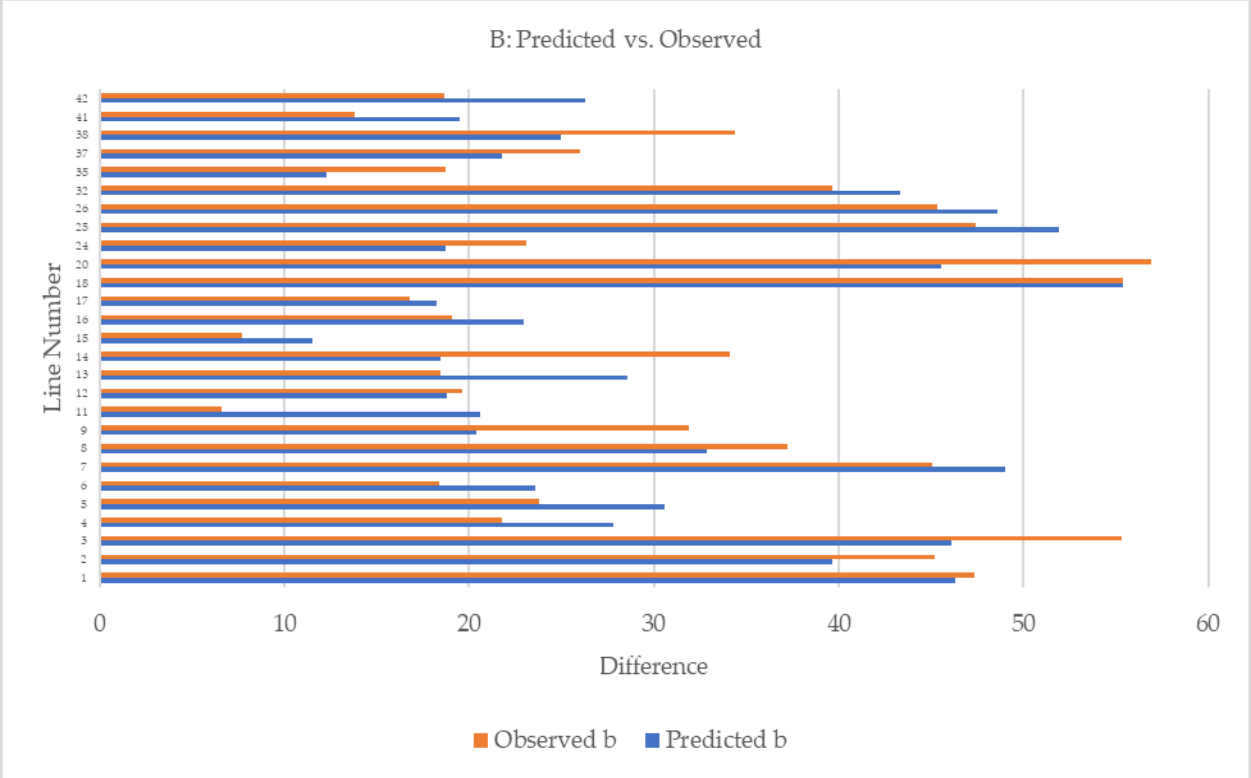


Figure 4.16: Comparison of mean observed and predicted values of *b* term over data collection period

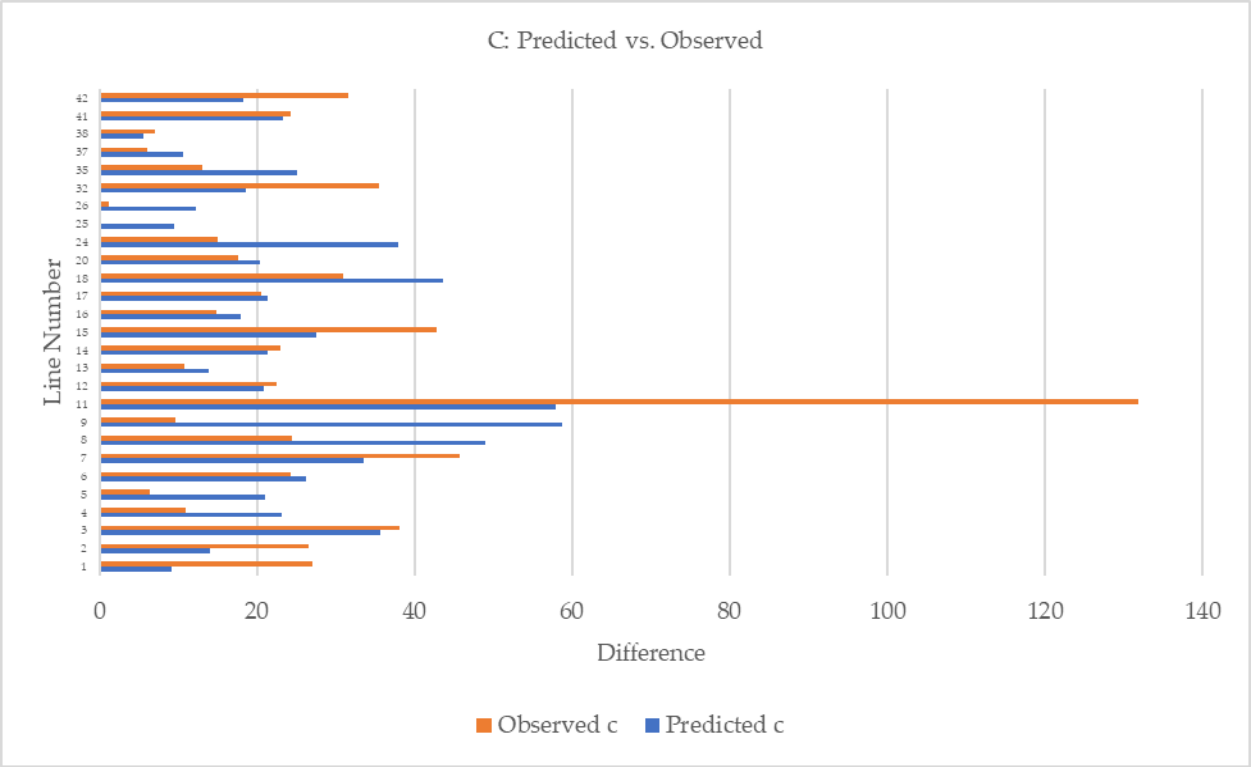


Figure 4.17: Comparison of mean observed and predicted values of *c* term over data collection period

4.7.2 Development of Empirical Performance Dataset (EPD)

As was noted in Section 4.2 and illustrated in Figure 4.2, the concept of mathematical integration was used for comparing the various iterations of Volume vs. Density curves developed in this work. Then, the percent difference in the integral value between two curves was compared and used as a metric for analysis. Figure 4.18 illustrates all three curves, Conceptual, Empirical, and Predicted for a given site.

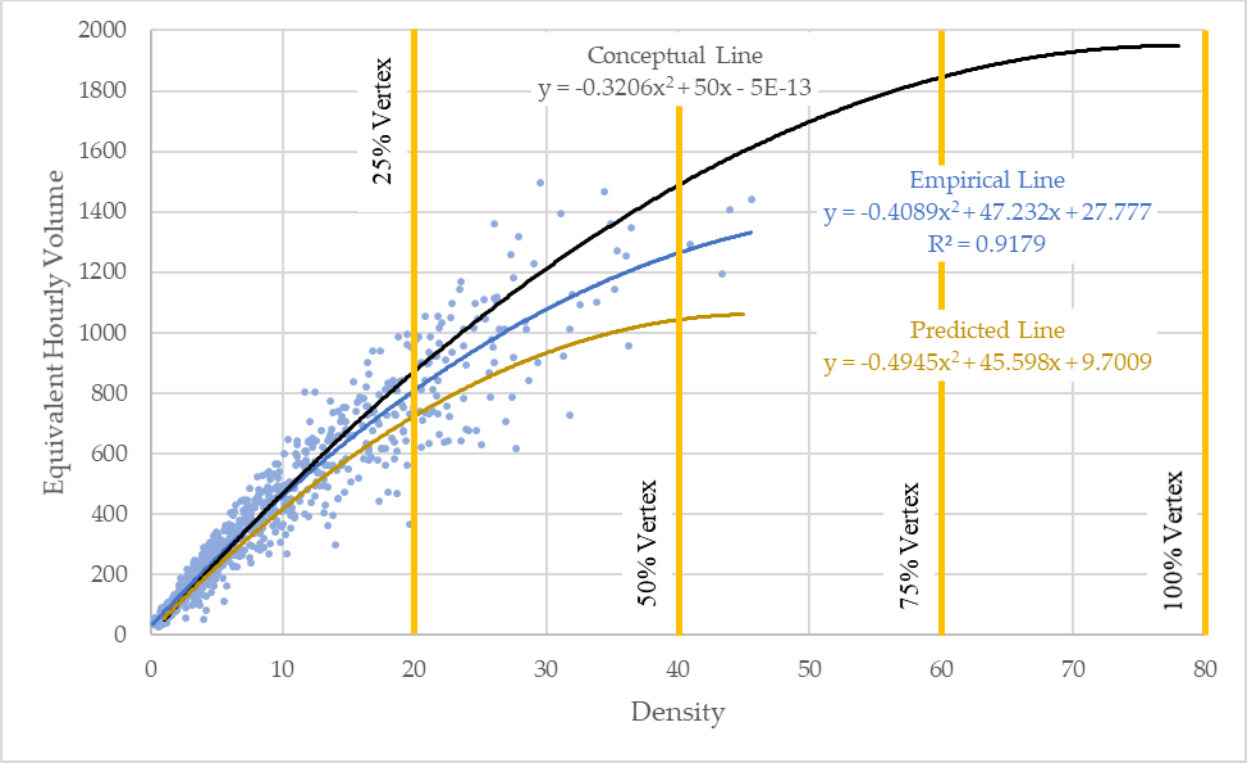


Figure 4.18: Empirical line, predicted line, and conceptual curve with integration bounds (25%, 50%, 75%, and 100% of the conceptual curve)

In using the concept of percent difference described earlier, a decision needed to be made regarding integration bounds for these comparisons. Given the typical application of field data to an uninterrupted conceptual Volume vs. Density curve, initially integration bounds for these comparisons were 0 and the x-coordinate of the vertex of the conceptual curve (100% vertex as shown in Figure 4.18). However, given that the majority of the plotted empirical data points fell within 0 to 50% of the x-axis vertex range, comparative integrals were developed for four different sets of ranges, from 0% to 100% of the x-axis vertex coordinate, in 25% increments, to determine which set of bounds would be the most appropriate for this work.

Prior to describing the results of this comparison, the specifics of the percent difference values calculated should be noted. For the EPD, percent difference calculations were made between consecutive weeks of data, using Equation 4-13.

$$\boldsymbol{Percent\ Difference} = \mathbf{100} \times \frac{(\boldsymbol{week\ n+1}) - (\boldsymbol{week\ n})}{(\boldsymbol{week\ n})} \qquad \textbf{(4-13)}$$

where:

week n = the integration value of the earlier week in the comparison

week n+1 = the integration value of the latter week in the comparison

As four weeks of data was processed for each detector, a total of six percent differences were generated, as each week was treated as an individual data point, regardless of the temporal sequence of the data. This resulted in 147 data points for the EPD.

- Week 2 compared to Week 1
- Week 3 compared to Week 1
- Week 4 compared to Week 1
- Week 3 compared to Week 2
- Week 4 compared to Week 2
- Week 4 compared to Week 3

Figure 4.19 through Figure 4.22 summarize the distributions of the percent difference calculations conducted with the various integration bounds. The percent differences when integrated to 100% of the conceptual vertex were typically the highest, as can be seen in these figures and in Table 4.6, which summarizes the mean and standard deviations of each of the distributions. Both the mean and standard deviation continued decrease as the integration bounds were reduced, with the smallest values observed at the 25% threshold.

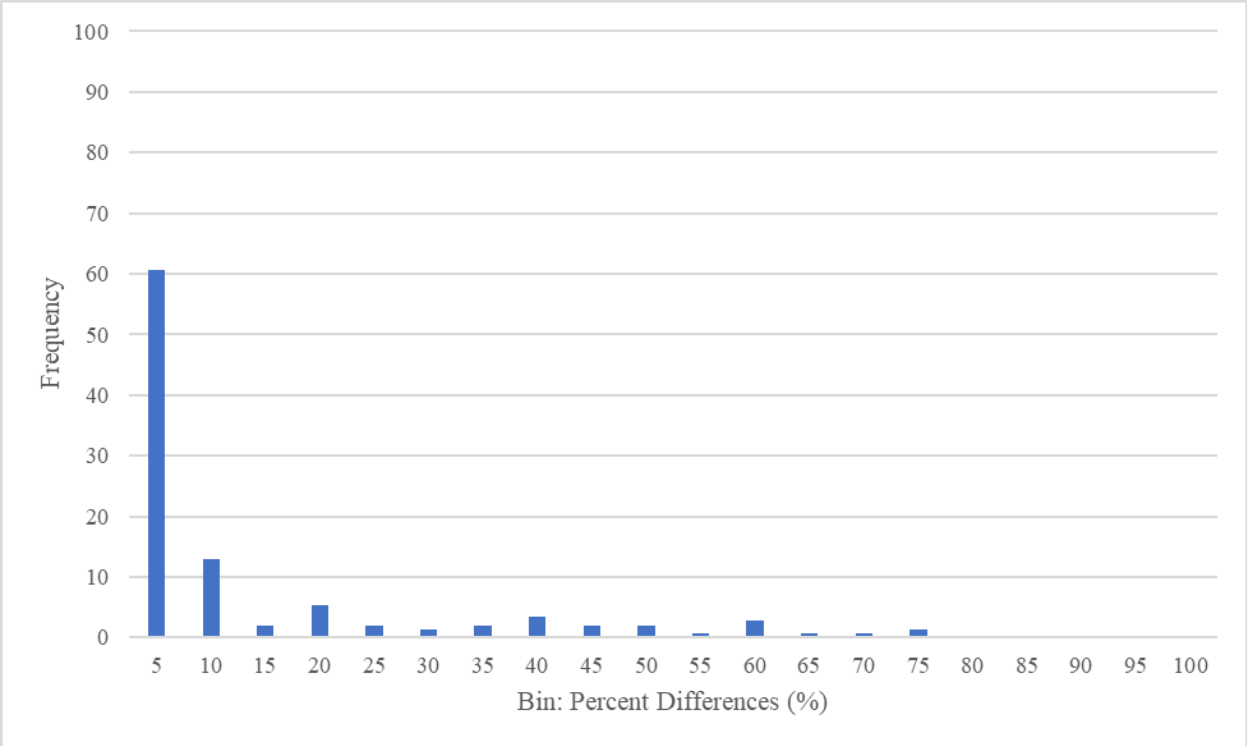


Figure 4.19: Percent difference distribution for integrating the conceptual and empirical lines to 25% of the conceptual vertex

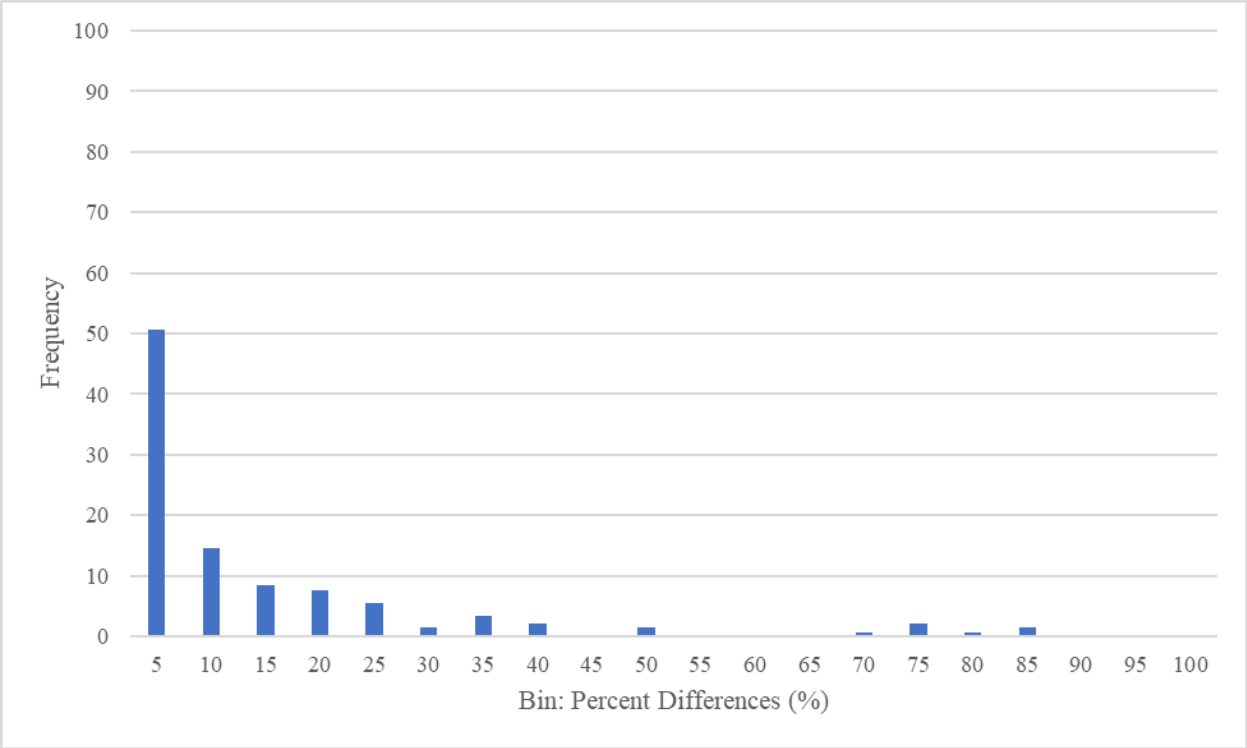


Figure 4.20: Percent difference distribution for integrating the conceptual and empirical lines to 50% of the conceptual vertex

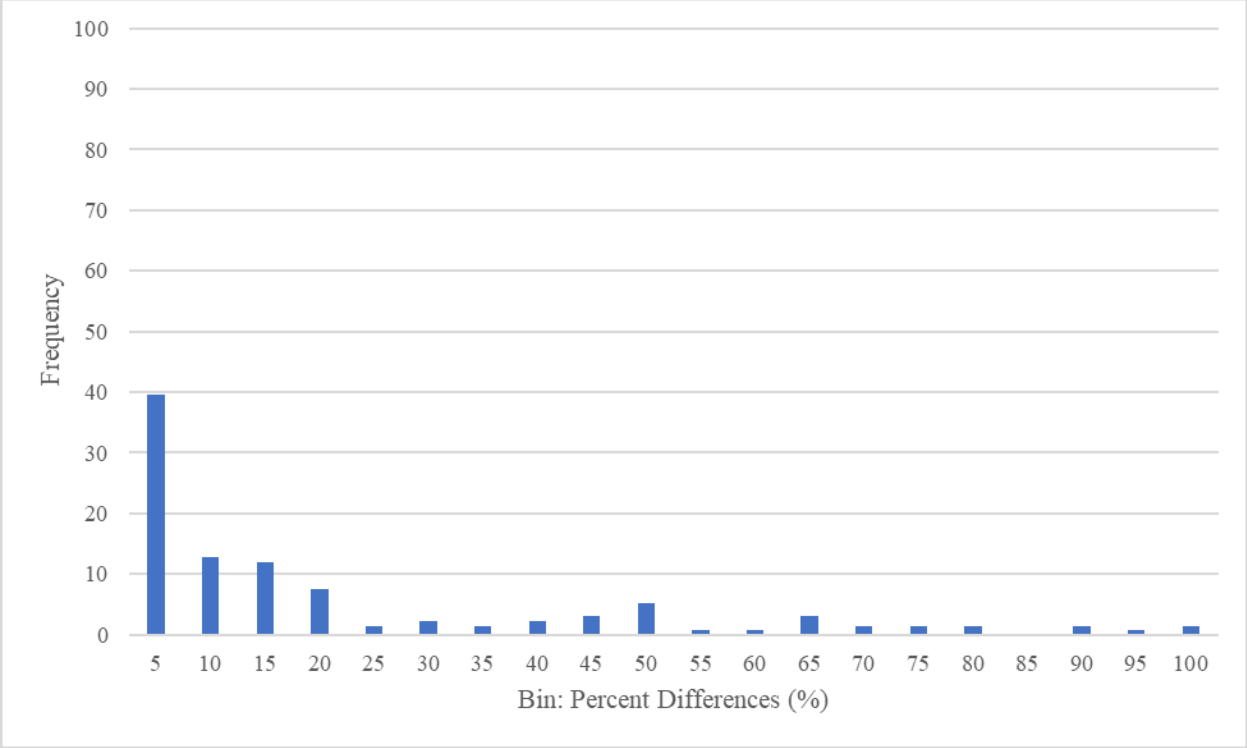


Figure 4.21: Percent difference distribution for integrating the conceptual and empirical lines to 75% of the conceptual vertex

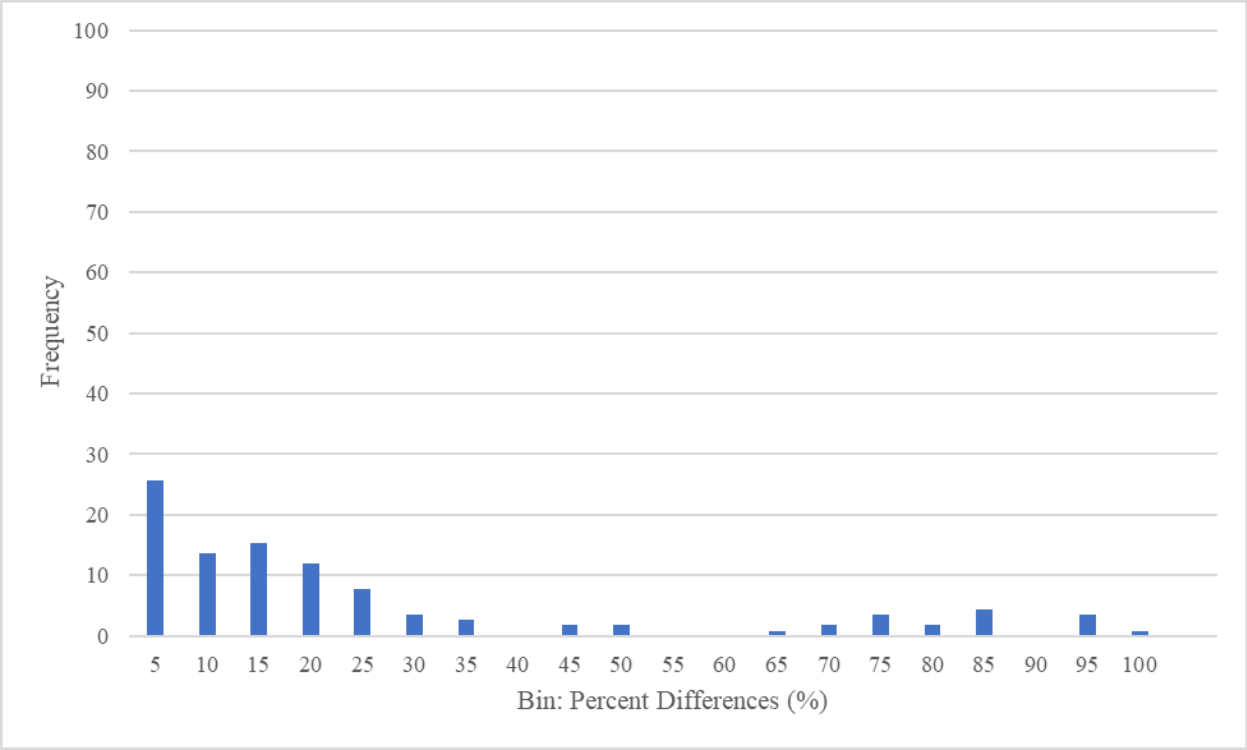


Figure 4.22: Percent difference distribution for integrating the conceptual and empirical lines to 100% of the conceptual vertex

Table 4.6: Mean and Standard Deviation of Integration Values at Different Integration Thresholds

	25%	50%	75%	100%
Mean	10.64	13.51	24.54	142.55
Std Dev	16.69	24.81	36.36	386.77

Figure 4.23 shows the cumulative percent differences for each integration threshold. A line is drawn at the 20% difference bin in Figure 4.23 to allow for a comparison between the four trace lines. This line illustrates that, at this point on the plot, roughly 80% of the data points in both the 25% and 50% threshold have values of 20% percent difference or lower. At the 20% bin, only 70% of the data is encompassed for the 75% threshold, and 65% for the 100% threshold. More data points below 20% are an indicator of less week-to-week variability in the processed data, something that is desirable in developing a dataset for performance comparisons.

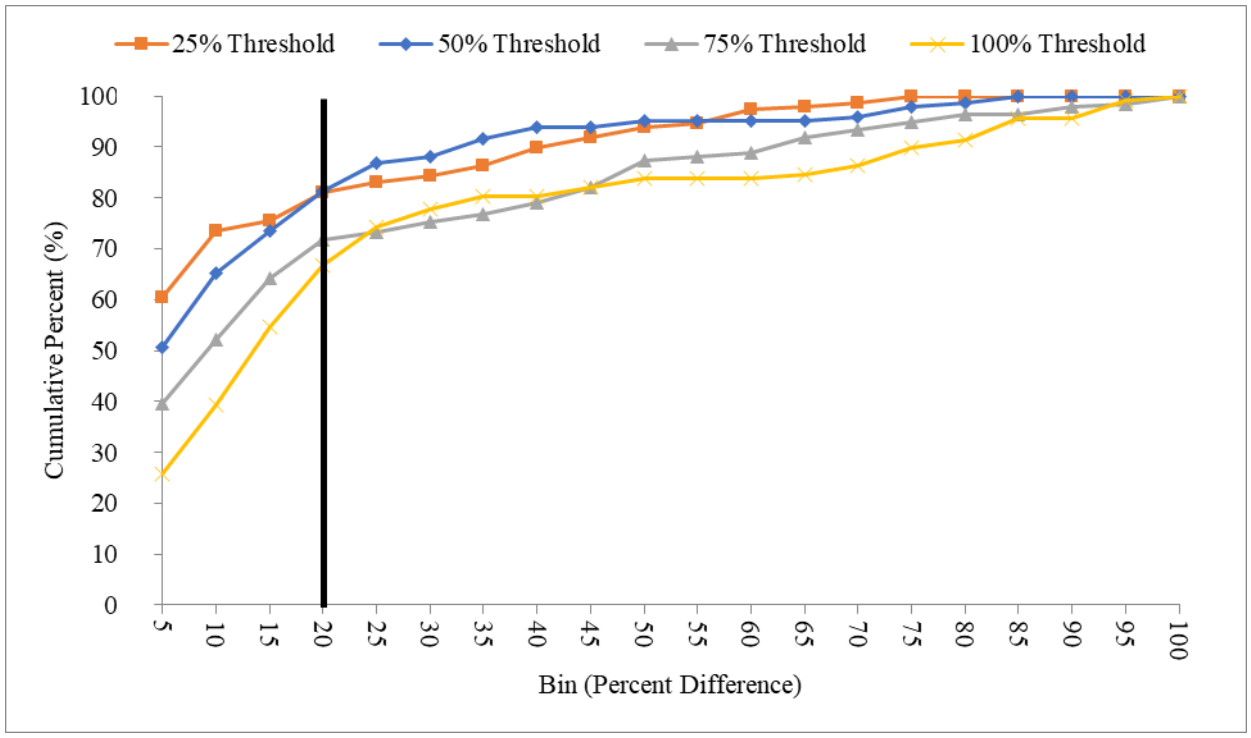


Figure 4.23: Cumulative percent difference comparison of the integration thresholds

Because of this, the desire to have roughly 80% of the percent difference values at or below 20% (black line shown in Figure 4.23), and the fact that the majority of data points developed to create the empirical lines were in this section of the plots, it was determined that using bounds of integration from 0 to 25% of the conceptual vertex would provide the most predictable performance assessment, as the mean and standard deviation of this distribution (red boxed data in Table 4.6) will be used and termed the Empirical Performance Dataset (EPD).

4.7.3 Development of Predicted Performance Dataset (PPD)

The final step in the development of data for detector health comparison is the creation of the Predicted Performance Dataset (PPD). To do that, a second percent difference calculation was made, this one between the conceptual line for each detector and each the predicted line for each week for that respective detector, as was calculated in Section 4.7.1 (this calculation was illustrated in figure 4.2). This was completed using the same process as described in Section 4.7.2, using integrals carried out between the 0% and 25% range. The distribution of these percent differences between the conceptual and predicted lines is shown in figure 4.24 with the mean of this data as 2.8 percent difference and the standard deviation 5.5 percent difference. These values will also be used in detector health analysis and termed the Empirical Performance Dataset (EPD).

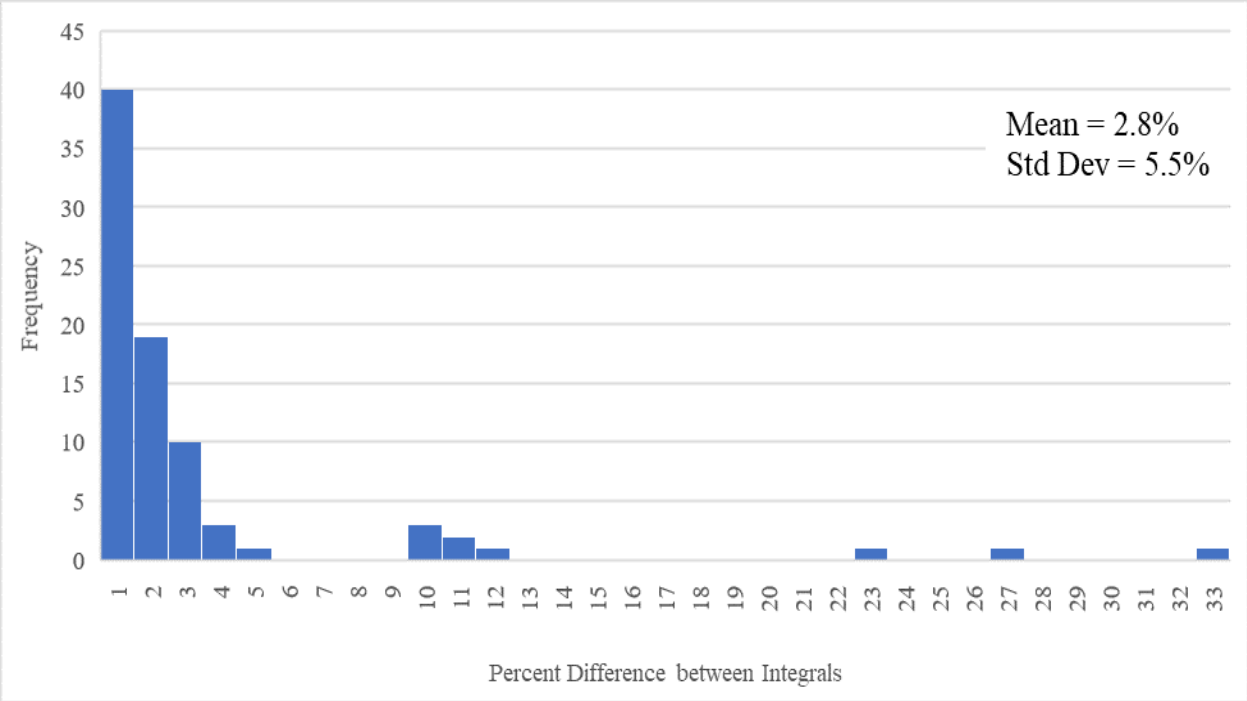


Figure 4.24: Percent difference between conceptual and predicted empirical line integrals for all detectors in predicted performance dataset

4.8 PROCESSED DATA PERFORMANCE COMPARISONS

To investigate whether or not the calculated means and standard deviations for the two performance datasets represent thresholds that differ from detectors that may not be performing properly, two different exercises were carried out.

4.8.1 Comparison of EPD to Underperforming Detectors

First, the same steps of data analysis and processing (as described in Section 4.7.2) were undertaken to determine the distribution of percent difference values from those detectors that did not pass the heuristic assessment, termed underperforming detectors. Figure 4.25 and Figure 4.26 show the distribution and cumulative distribution for these detectors (11 detectors, resulting in 60 percent difference data points). In comparison to EPD, the mean percent difference of the 0-25% integration bounds of the underperforming detectors is 42.2% (an increase of 31.6%), and the standard deviation is 117.9% (an increase of 101.2%). The increase in both of these values represents an increase in variability of the integral values from week to week of the underperforming detectors compared with those in the EPD, indicating that the use of percent difference values should be effective in identifying underperforming detectors. The 50% bound CDF line is shown for comparison, illustrating that the wider integration bounds greatly increase the percent difference variability in the underperforming detector data set. The same line is drawn on Figure 4.26 at the 20% bin. For the EPD, both of the 25% and 50% traces encompassed roughly 80% of the data points. For these underperforming detectors, roughly 70% of the data points are 20% difference or less with the 25% threshold, while roughly 45% of the data points are 20% or lower for the 50% threshold.

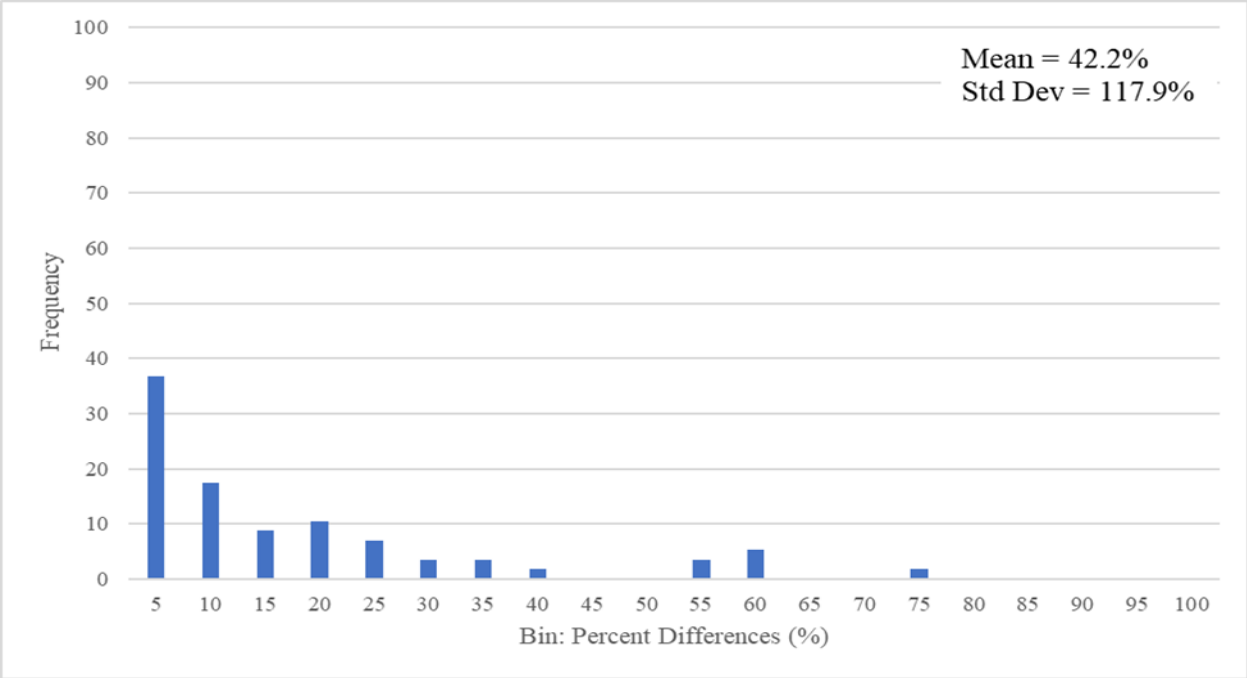


Figure 4.25: Percent difference distribution for integrating the conceptual and empirical lines to 25% of the conceptual vertex for underperforming detectors

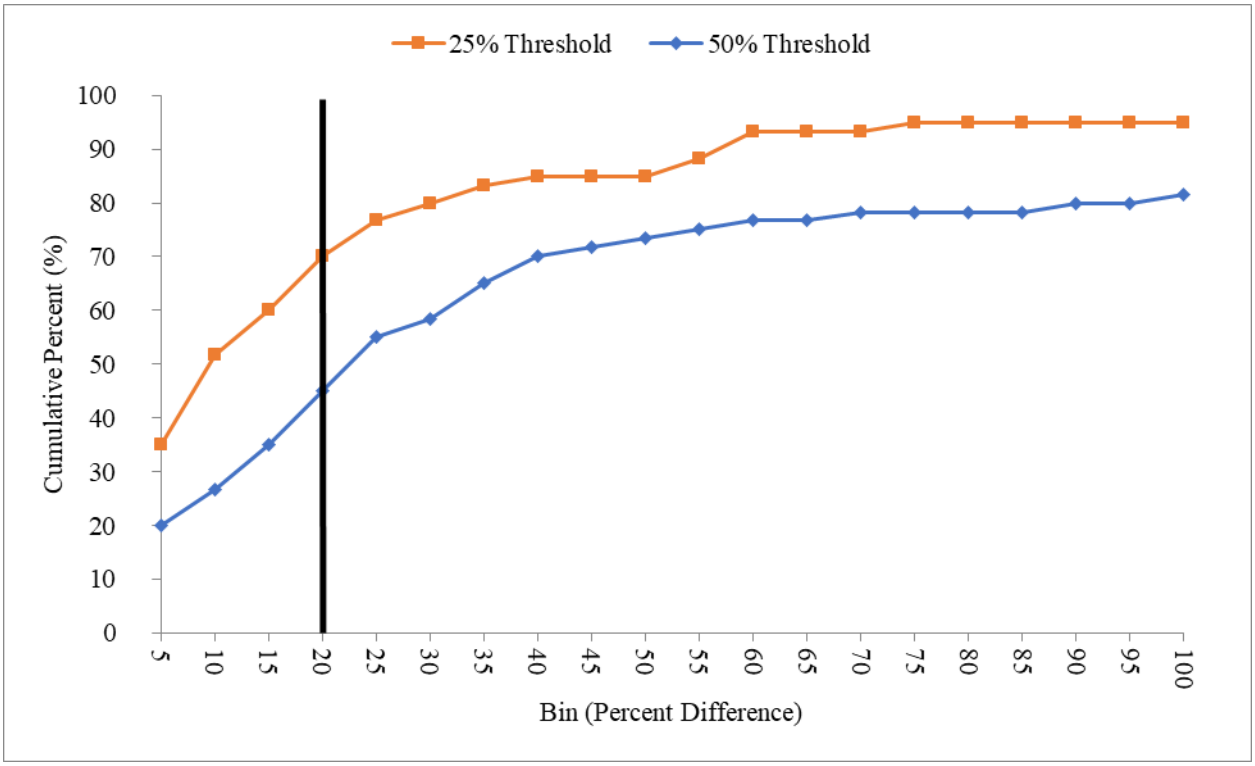


Figure 4.26: Cumulative percent difference comparison of the integration thresholds for underperforming detectors

4.8.2 Integral Value Changes For 'Bad' Detector Data

Second, in an attempt to gain an understanding of the impact that degraded detector performance has on the difference in integral values, the calculated Volume and Density data points for four weeks of data for one detector were artificially increased and decreased by 10%, 20%, and 30% (a review of literature did not identify any works where improperly performing, but still functioning, detector data was logged or profiled in this manner). Figure 4.27, Figure 4.28, Figure 4.29, and Figure 4.30 show the impact of these increases on the location of the empirical line for an increase in Density, increase in Volume, decrease in Density, and decrease in Volume, respectively.

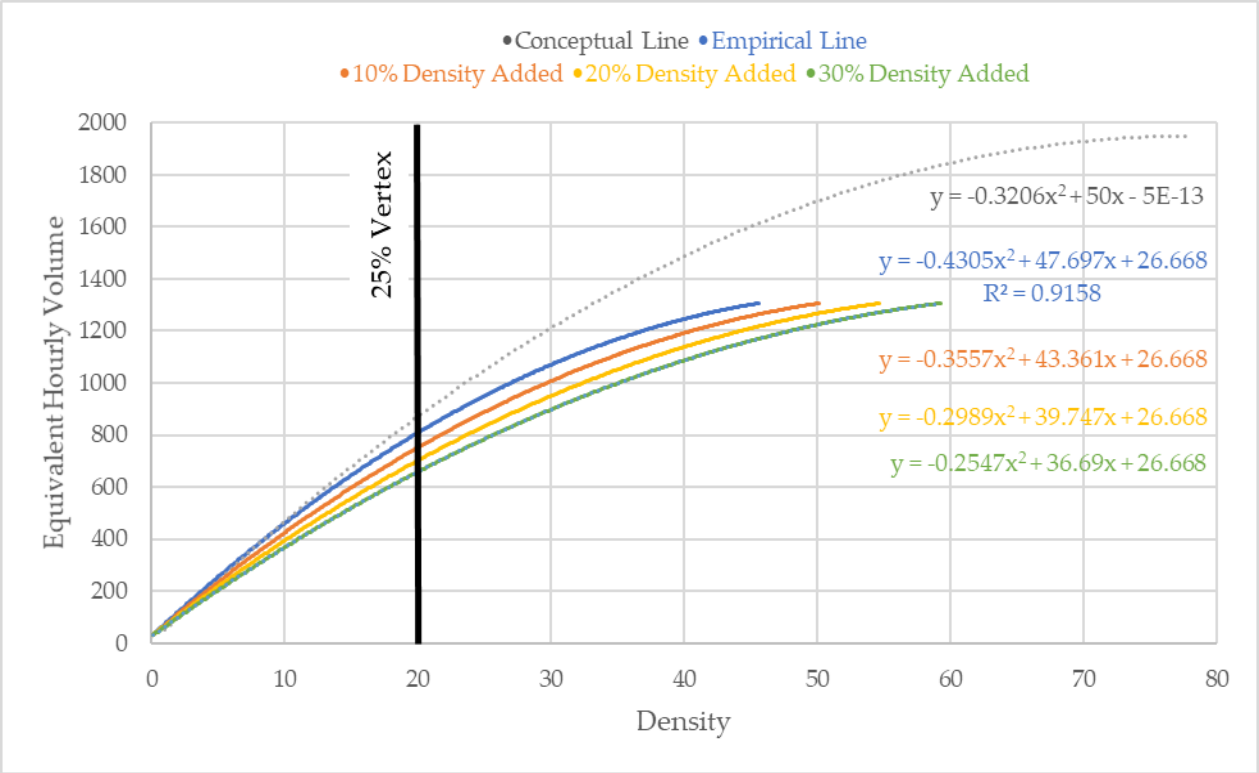


Figure 4.27: Site 245 detector 1, 4 weeks of data, density increased

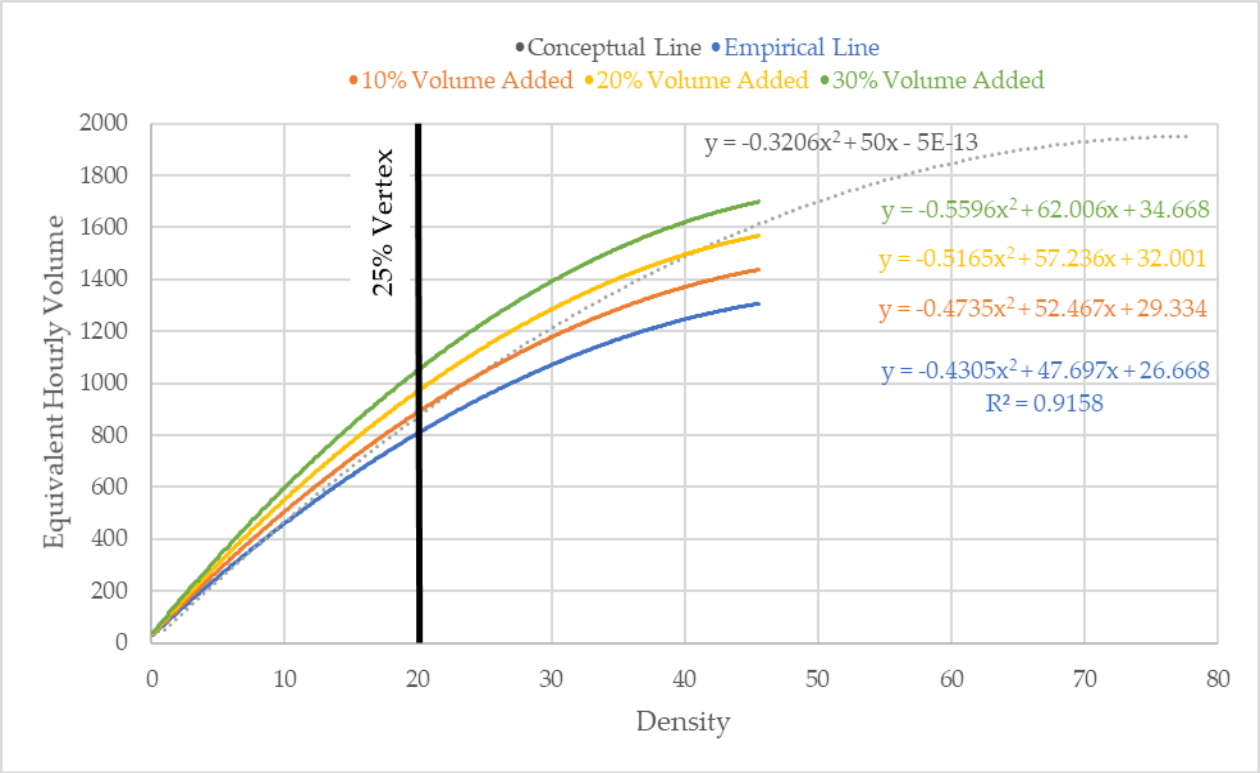


Figure 4.28: Site 245 detector 1, 4 weeks of data, volume increased

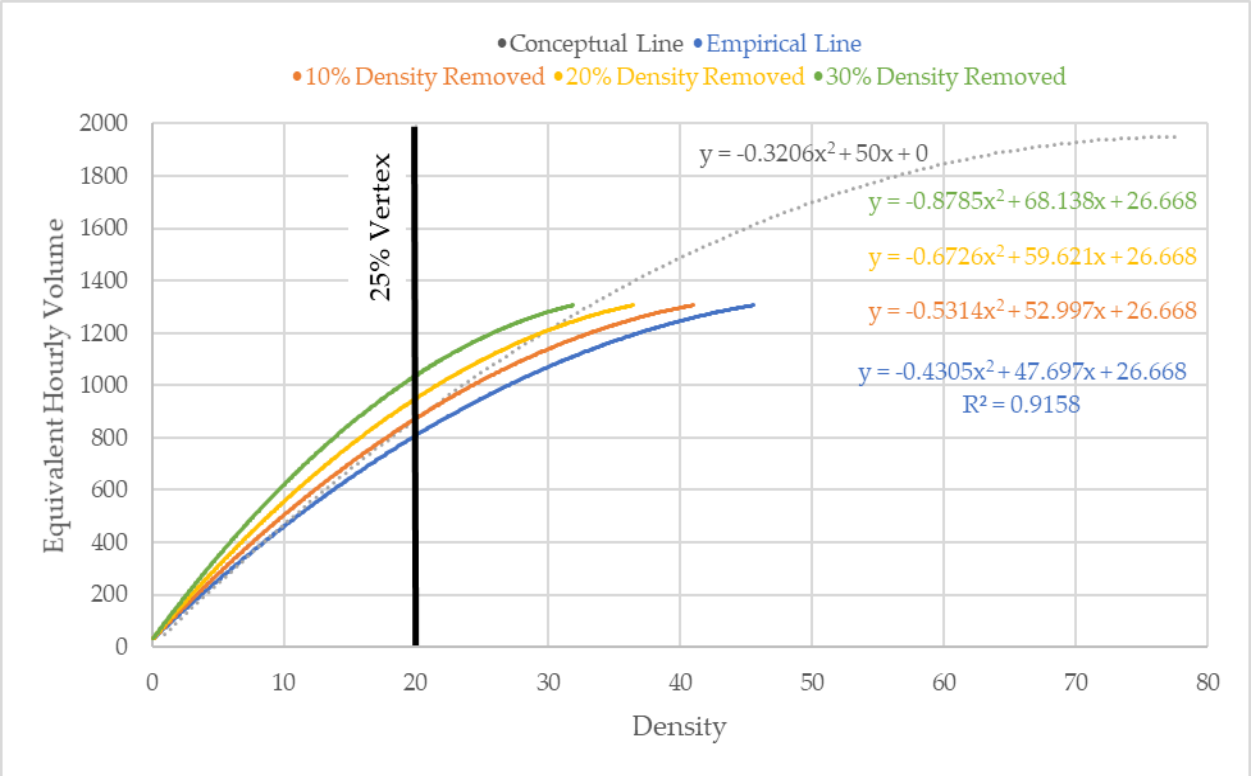


Figure 4.29: Site 245 detector 1, 4 weeks of data density decreased

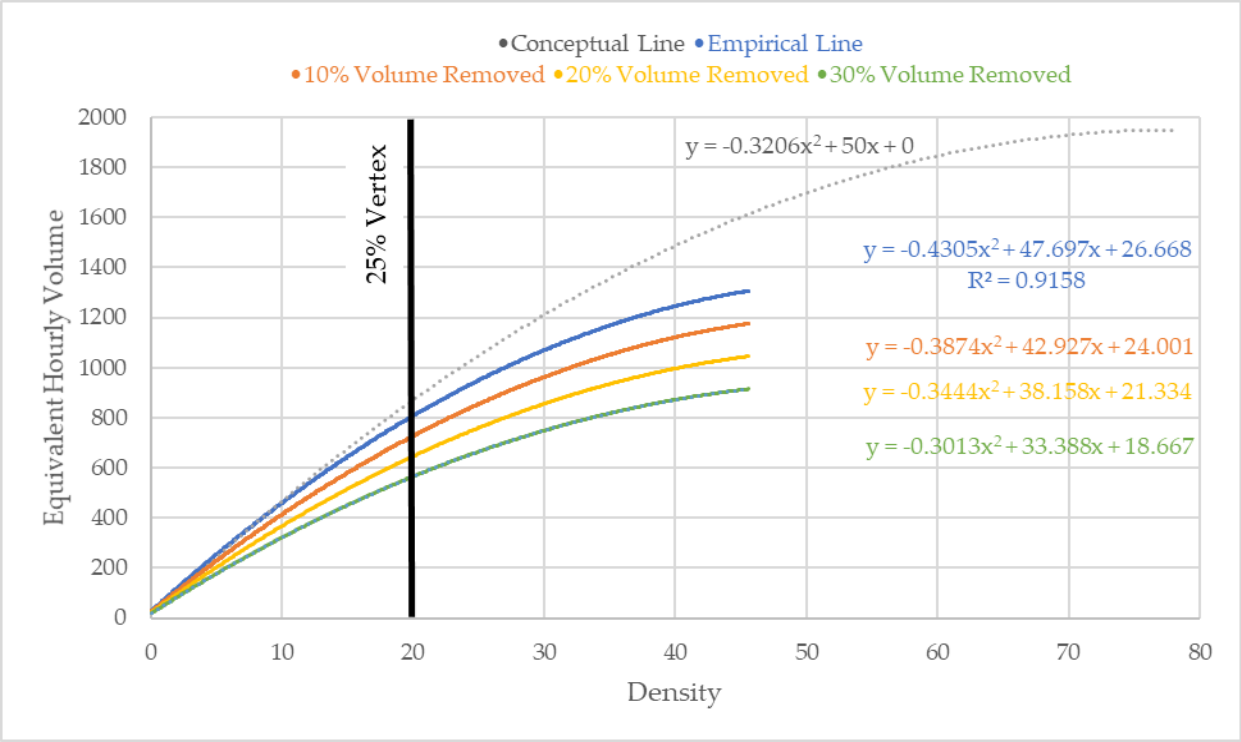


Figure 4.30: Site 245 detector 1, 4 weeks of data, volume decreased

The adjusted lines illustrated in the above figures were compared to the detector's conceptual line by integrating from 0 to 25% of the conceptual line's vertex and calculating the percent difference between these lines. The integration bound is also shown in the figures. The results of these integration comparisons are shown in Table 4.7.

Table 4.7: Site 245 Detector 1, 4 Weeks of Data: Integration Percent Differences from the Conceptual Integral

% Difference from Conceptual Integral (Conceptual – x) / Conceptual Integral from 0 to 25% of Conceptual Vertex		
Empirical		**2.2%**
Density Added	10% Added	9.5%
	20% Added	15.8%
	30% Added	21.2%
Volume Added	10% Added	7.6%
	20% Added	17.4%
	30% Added	27.2%
Density Removed	10% Removed	6.5%
	20% Removed	17.0%
	30% Removed	29.7%
Volume Removed	10% Removed	12.0%
	20% Removed	21.7%
	30% Removed	31.5%

In the EPD, the mean and standard deviation of the percent differences between the empirical and conceptual lines is 10.64% and 16.69% respectively. Staying within 1.5 standard deviations from the mean, a comparative value from a statistical standpoint and a threshold used in the algorithms presented in successive sections (in a normal distribution, ~87% of the data fall with +/- 1.5 std deviations), would create a threshold of 35.68% percent difference between the conceptual and empirical lines. When 10%, 20%, and 30% was increased or decreased from the Density and Volume datasets, all of the percent differences between the conceptual and the empirical lines were still within that 35.68% threshold. This comparison may indicate that using 1.5 standard deviations from the mean may not be sensitive enough to identify all detector malfunctions or poor data quality.

4.9 GREEN ACTIVATION ANALYSIS

As mentioned in Section 4.6, we did not find statistically significant regression models between the number of detector activations per green duration (green activations) and the green duration from the collected data for 43 detectors among all 45 detectors. Therefore, we could not use the regression models, rather we used the green activation to assess the detector health. We used four weeks' data from 6 AM to 9 AM and from 4 PM to 7 PM on Tuesdays, Wednesdays, and Thursdays. R scripts were coded to produce the weekly mean and standard deviation of the green activation for each detector. We also compared the change of weekly mean to figure out how the green activation changed over the four weeks using the following equation. We also computed the sample size (number of green durations) for each week and each detector using Equation 4-14.

Current week's change of mean of green activation = (current week's mean - previous week's mean) / current week's mean

(4-14)

The weekly mean and standard deviation of green activation are listed in the following table. The sample size and change of weekly mean are shown in the following Figure. From the following table and figure, most of the weekly changes of mean of green activation are less than 15% when the sample size is more than 100. Therefore, signal timers may need to check whether there are detector issues: (1) when the sample size is less than 20; (2) when the sample size is more than 100 and the weekly change of mean is more than 20%.

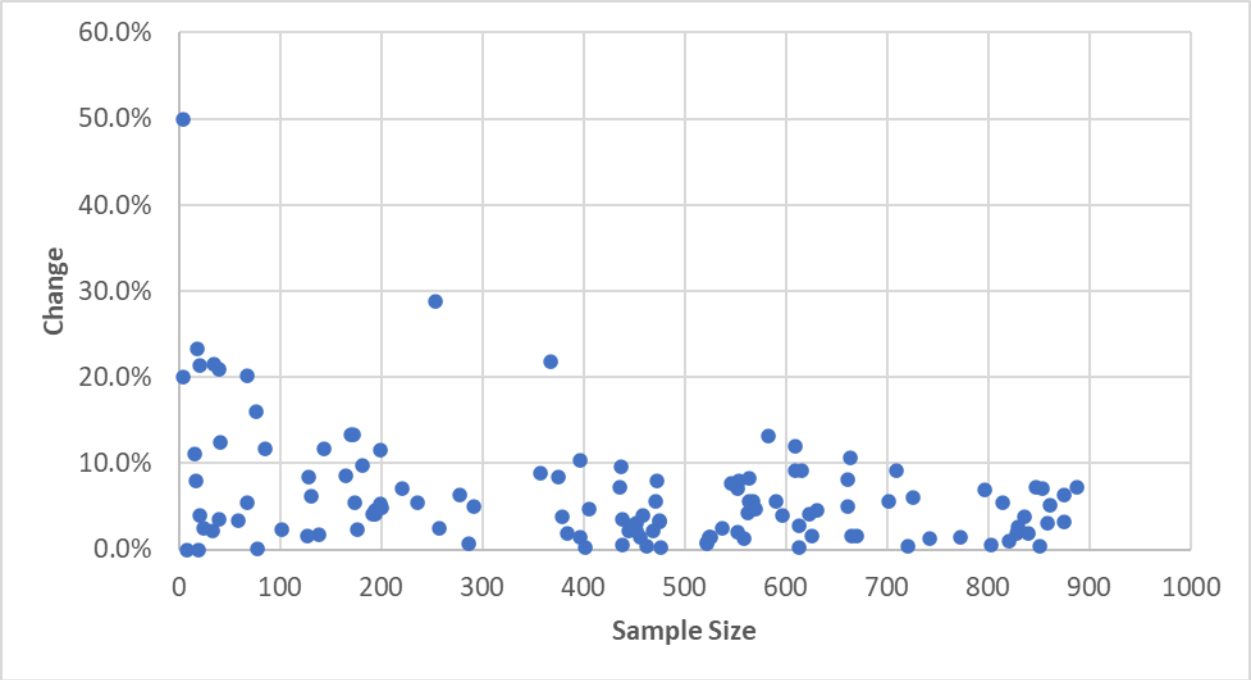


Figure 4.31: Sample size and change of weekly mean of green activation

Table 4.8: Weekly Mean and Standard Deviation of Green Activation

Mean week1	SD week1	Mean week2	SD week2	Mean week3	SD week3	Mean week4	SD week4	Change of mean week2	Change of mean week3	Change of mean week4
5.47	3.41	5.85	3.69	5.39	3.35	6.04	3.82	7.0%	7.9%	12.0%
6.92	4.47	7.49	5.18	6.92	4.73	7.30	4.78	8.2%	7.6%	5.5%
1.13	0.34	1.00	0.00	1.00	0.00	1.23	0.50	11.1%	0.0%	23.3%
1.30	0.56	1.03	0.17	1.25	0.59	1.21	0.52	20.9%	21.5%	3.4%
4.44	3.01	4.65	3.18	4.46	2.98	4.68	3.27	4.5%	4.1%	5.0%
6.79	4.65	7.51	5.59	6.90	5.02	7.29	5.19	10.6%	8.1%	5.6%
1.10	0.36	1.06	0.30	1.24	0.60	1.17	0.41	3.3%	16.0%	5.4%
1.25	0.51	1.40	0.69	1.31	0.54	1.33	0.58	11.8%	6.1%	1.7%
16.86	12.76	18.41	13.31	18.69	13.10	18.38	12.79	9.2%	1.5%	1.6%
1.00	0.00	1.00	0.00	1.50	1.00	1.20	0.45	0.0%	50.0%	20.0%
12.16	7.29	13.28	7.68	13.31	7.67	13.67	8.30	9.2%	0.2%	2.7%
1.57	1.03	1.24	0.56	1.33	0.48	1.28	0.54	21.4%	7.9%	4.0%
1.36	0.65	1.21	0.53	1.45	0.83	1.45	0.80	11.6%	20.1%	0.0%
1.60	0.96	1.68	1.08	1.45	0.83	1.51	0.88	4.6%	13.3%	4.0%
1.84	1.19	1.97	1.35	1.86	1.19	1.91	1.31	7.0%	5.4%	2.4%
15.80	6.67	16.33	7.36	16.85	7.09	17.80	8.12	3.4%	3.2%	5.6%
11.87	6.16	13.09	6.40	12.35	6.52	13.33	6.61	10.3%	5.6%	7.9%
22.19	7.64	22.15	7.37	21.83	7.35	28.10	11.75	0.2%	1.5%	28.7%
6.49	4.89	6.26	4.70	6.39	4.63	6.36	4.65	3.5%	2.1%	0.5%
7.81	5.85	7.71	5.57	8.01	5.64	7.76	5.72	1.4%	4.0%	3.1%
10.23	6.60	10.25	6.67	10.30	6.68	10.07	6.36	0.2%	0.4%	2.2%
20.86	6.85	21.30	7.17	21.88	6.64	23.97	9.32	2.1%	2.7%	9.6%
5.85	4.22	6.27	4.18	6.21	4.07	5.98	3.97	7.1%	0.9%	3.8%
2.45	1.60	2.67	1.83	2.57	1.77	2.53	1.78	9.1%	4.0%	1.6%
1.68	1.08	1.76	1.15	1.69	1.13	1.57	0.93	4.7%	3.8%	7.1%
2.82	2.22	2.66	2.11	3.02	2.50	2.89	2.40	5.6%	13.2%	4.2%
4.96	4.33	5.30	4.29	5.22	4.26	5.19	4.52	6.9%	1.5%	0.6%
7.59	5.07	8.14	4.91	7.89	4.84	7.63	4.87	7.2%	3.1%	3.3%

Mean week1	SD week1	Mean week2	SD week2	Mean week3	SD week3	Mean week4	SD week4	Change of mean week2	Change of mean week3	Change of mean week4
3.44	2.48	4.20	2.85	3.97	2.95	3.89	2.83	21.8%	5.5%	1.8%
2.60	1.63	2.76	1.69	2.77	1.71	2.73	1.79	6.1%	0.4%	1.2%
5.92	3.74	6.30	3.66	6.18	3.83	6.20	3.80	6.4%	1.8%	0.3%
4.45	2.65	4.68	2.57	4.55	2.50	4.89	3.10	5.1%	2.7%	7.3%
18.12	9.22	18.37	9.58	17.92	8.79	18.18	10.13	1.4%	2.5%	1.5%
16.63	8.42	16.40	7.83	16.52	7.75	16.66	9.10	1.3%	0.7%	0.8%
1.12	0.33	1.15	0.36	1.12	0.33	1.26	0.57	2.4%	2.2%	12.4%
1.32	0.58	1.35	0.60	1.38	0.56	1.49	0.74	2.3%	1.5%	8.4%
1.84	1.22	1.95	1.32	1.78	1.17	1.85	1.29	5.5%	8.6%	4.1%
1.69	0.99	1.88	1.12	1.79	1.18	1.87	1.28	11.6%	5.2%	4.9%
12.65	7.52	13.23	7.01	12.97	6.79	13.14	7.78	4.6%	2.0%	1.3%
1.99	1.20	1.72	1.11	1.76	1.08	1.93	1.24	13.3%	2.3%	9.8%
1.74	1.08	1.82	1.24	1.71	1.15	1.70	1.11	5.0%	6.4%	0.7%
2.17	1.47	2.35	1.72	2.14	1.54	2.10	1.53	8.4%	8.8%	1.8%

4.10 ALGORITHM FOR DETECTOR INITIAL HEALTH ASSESSMENT

Figure 4.1 described an overview of the detector health comparison processes developed for this work. The detailed algorithm for Detector Initial Health Assessment is shown in Figure 4.32. The process for each step is discussed in the successive text. Pseudocode for all algorithms is listed in Appendix A.

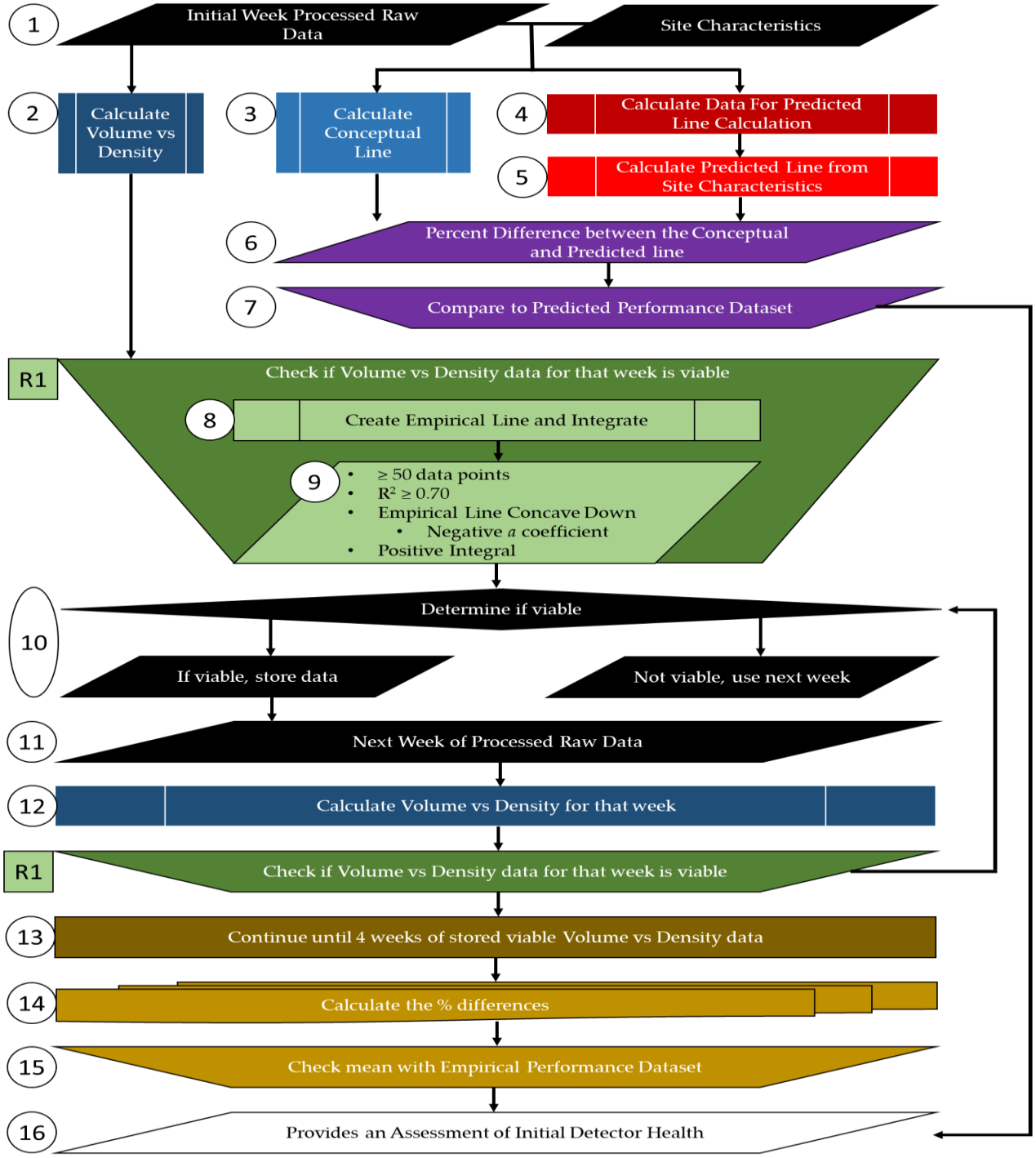


Figure 4.32: Initial health assessment flowchart

1. Input: Initial Week of Processed Raw Data. One week of filtered data (Section 4.3 through 4.3.3) is used in the algorithm for the Initial Health Assessment.
   Input: Site Characteristics. The site characteristics include the movement's speed limit, which is used in the process of calculating the conceptual line, as well as aspects of the detector's location used to model the predicted line (Section 4.7.1).

2. Process: Calculate Volume vs Density. The raw data are processed and the volume versus density relationships in the empirical data are calculated (Section 4.4).

3. Process: Calculate Conceptual Line. The conceptual line is calculated (Section 4.5)

4. Process: Calculate Data for Predicted Line Calculation. Calculate the average number of activations during green per hour, and average number of green indications per hour, both averaged across the week of data.

5. Process: Calculate Predicted Line from Site Characteristics. The process for determining the predicted line is described in Section 4.7.1. Site characteristics, activation data, and existing output models are used to calculate the predicted line.

6. Output: Percent Difference between Conceptual and Predicted Lines. The determined Conceptual and Predicted lines are compared by finding the percent differences between their integrals.

7. Compare to PPD. The percent difference between the conceptual and predicted lines' integrals is compared to the mean and standard deviation in the PPD.

8. (Routine 1) Processes: Check if Volume vs Density for that week is viable. Create an empirical line of best fit to the initial week of volume vs density data.

9. (Routine 1) Output: There are multiple checks for this data and its empirical line. The data set for the initial week must have 50 or more data points; the *Coefficient of Determination* (R^2) for the empirical line fit to the data set must be greater than or equal to 0.70; the empirical line must be concave down; and the empirical line must have a positive integral when integrated from 0 to 25% of the Conceptual line's vertex. These checks determine if the initial week of empirical data is viable for assessing the detector's health.

10. Decision: Determine if viable. The checks described above determine if the initial week of data is viable for this assessment. If it is not viable, the following week of data should be instead analyzed, and steps 1-9 repeated with the new week of data. If the initial week of data is viable, then the empirical line's integral from 0 to 25% of the conceptual line's vertex should be stored.

11. Input: The next week of processed raw data is now used in the algorithm for the initial health assessment.

12. Process: The volume vs density data should be analyzed for this next week.

13. (Routine 1) Determine if the volume vs density for that week is viable, similarly to how it was determined for the initial week, creating an empirical line and integrating it as part of the process. The same bounds are used for integration as were for the first week of data.

14. Process: If this data is viable, the integral is stored, and the process moves onto the next week of processed data. This algorithm continues until there are 4 weeks of viable data.

15. Calculation: Percent difference values are calculated from the 4 weeks of empirical line integrals. The mean of these percent differences are calculated.

16. Compare to EPD. The mean of the percent differences between the empirical lines over multiple weeks is compared is compared to the mean and standard deviation in the EPD.

17. Output: The percent differences calculated above must be within 1.5 standard deviations of the respective performance datasets (noted in steps 7 and 15). If one or both are outside this value, it is an indicator of possible poor detector health.

Note that if the detector passes both tests of detector health, the first four weeks of viable data are then used within the detector's health assessment over time.

4.11 ALGORITHM FOR HEALTH ASSESSMENT OVER TIME

The algorithm for assessing detector health over time uses control charts and the sliding window technique to compare the detector's functionality EPD. Control charts are an analysis tool used to detect variations in a dataset over time, detecting causes of shifts in the control state of a process (Montgomery & Runger 2018). It is understood that over time, there will be some acceptable level of variation in a dataset, however the goal of a control chart is to identify when variation is outside of an acceptable level. Typically, a control chart has three lines on it, one to show the mean of the dataset, and two that identify +/- 1.5 standard deviations of the sample from the mean. Points that fall outside the +/- 1.5 standard deviation lines identify excessive variation. Figure 4.33 illustrates this by using data from the 0-25% integration percent difference dataset. The values from the EPD are used in this figure, with the mean of 10.64% shown as the thick solid line, and the upper bound of 35.64 (mean + 1.5*16.69) is shown by the thin dashed line (the lower bound is below zero).

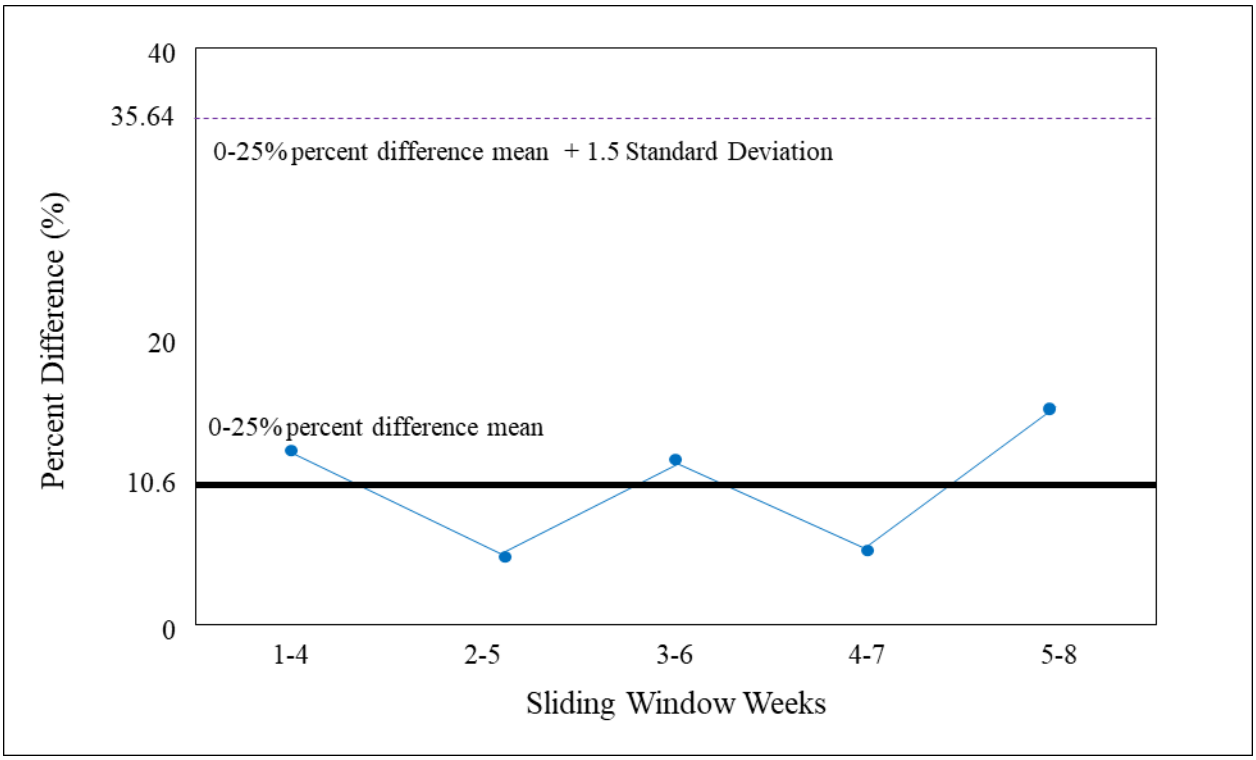


Figure 4.33: Control chart plotting the average integral percent differences from the sliding window technique

The sliding window method is derived from a technique used for evaluating safety (i.e., network screening to identify potential safety hot spots) on roadway segments incrementally (Herbel, Laing, and McGovern 2010). For example, in a hypothetical roadway safety analysis, five 0.1-mile length segments (segments 1-5) would be evaluated for a specific performance metric. Then, the next analysis would 'slide' down the facility, comprising of segments 2-6 for the second 0.5-mile-long segment. This same technique can be used to evaluate detector health over time, but instead of sliding the window over different segments of roadway, the window slides over different weeks of data. With this method, the mean of the percent differences is calculated for over four weeks of simultaneous data (as shown in Section 4.7.2). The weeks included in the calculation are then incremented by one week and calculated for the next four weeks of simultaneous viable data. That next window includes three of the four previous weeks of data, plus one new one incrementally. A point plotted outside of the upper control bound is an indicator that a detector may not be performing properly.

Figure 4.34 illustrates the algorithm for assessing detector health over time, with descriptive text following the figure.

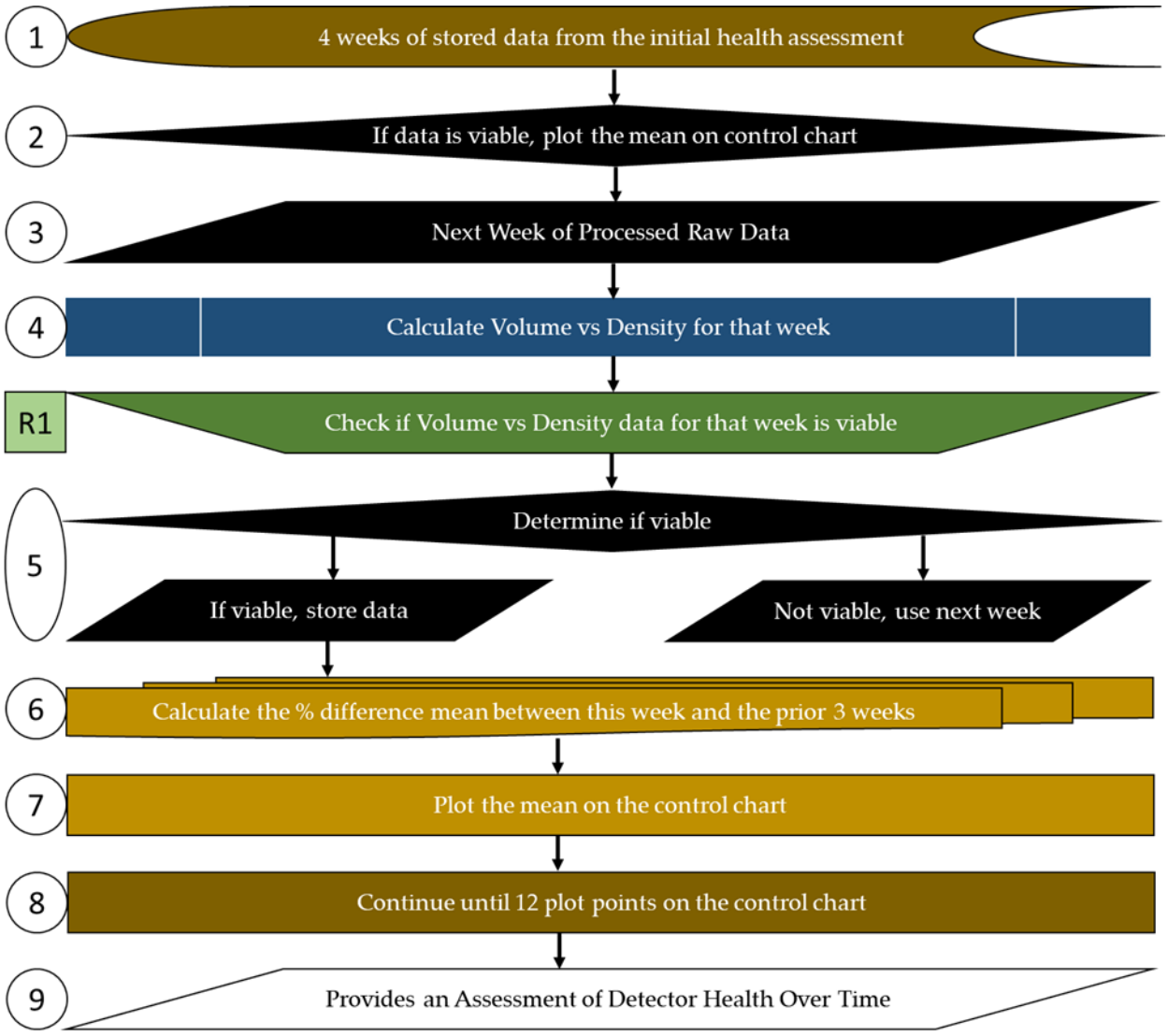


Figure 4.34: Health assessment over time flowchart

1. Stored Data: Four weeks of percent difference data are stored from the Initial Detector Health Assessment.

2. Decision: If four weeks from the initial health assessment can be considered viable, plot the percent difference mean on the control chart. If not, data are continually collected and processed until four weeks of viable data has been stored and the resulting percent difference calculation plotted on the control chart.

3. Input: The next week of processed raw data is now used in the algorithm for the health assessment over time.

4. Process: The volume vs density data should be analyzed for this next week. (Routine 1) Determine if the volume vs density for that week is viable, using the same process

as described in the initial health assessment, creating an empirical line and integrating it as part of the process.

5. Decision: If this data is viable, the integral is stored. If the data is not viable, the next week of data is used in this analysis instead.

6. Calculation: After this next week of data is processed for its integral, this integral and the integral from the prior three weeks of data are used to calculate a set of percent differences (the oldest week is dropped from the calculation, per the sliding window technique). The mean of the percent differences between these four weeks' integrals is calculated.

7. Process: The mean of the percent differences between these integrals is plotted on the control chart. If it is outside the bounds of the control chart, it may be an indicator of poor detector health. It is suggested that the mean and standard deviation of the EPD be used for the control chart.

8. Process: This process of analyzing the next week of data and comparing the results to the prior three weeks of data is continued until there are 12 plot points on the control chart.

9. Output: This control chart provides an assessment of the detector's health over time. It is suggested that the mean and standard deviation of those 12 weeks of data is used for the new control chart bounds, which allows the analysis to slowly adjust over time, and be based upon the data from the specific detector, as opposed to the entire EPD.

4.12 ALGORITHM USING GREEN ACTIVATIONS

From the analysis in Section 4.6, we propose the following algorithm to monitor detectors' health weekly using the detector activation per green duration (green activation). The algorithm could be run once per week to get warnings of potential detector issues.

Algorithm:

For the current week and each detector

- Calculate the green activation and duration.
- Calculate the sample size of green duration; calculate the mean m and standard deviation d of green activation.
- Calculate the current week's change of mean of green activation = (current week's mean - previous week's mean) / current week's mean.
- If the sample size is less than 20, produce a warning message that the detector may have potential issues.

- If the sample size is more than 100 and the current week's change of mean of green activation is more than 20%, produce a warning message that the detector may have potential issues.

- If the current week's mean of green activation is more than the previous week's green activation $m + 1.5*d$, or less than the last week's green activation $m – 1.5*d$, produce a warning message that the detector may have potential issues.

4.13 CLOSING THOUGHTS

The data driven algorithms presented in this chapter reflect a robust set of processes that can be implemented to assess the health of varied types of vehicle traffic detectors, both immediately upon integration into agency workflows, and over time. It should be noted that the selection of the value of 1.5 standard deviations from the mean was done so in line with customary control chart theory; ODOT staff may choose to adjust this selection based upon a desire to either have the algorithm function in a more aggressive manner (reduce the value) or conservative manner (increase the value). Additionally, while the algorithm developed in this work checks values with the PPD and EPD, it is recommended that as the algorithm is deployed across the system, datasets be collected and segregated for various types and locations of detectors (technology, location, length, etc.), and those datasets then be used for tracking health.

5.0 SYSTEM DESIGN AND IMPLEMENTATION PLAN

This chapter presents the design for a software system that would provide key information about the health of traffic detectors to ODOT staff along with an implementation plan outlining a process for deploying the system. The system presented makes use of the novel algorithm for evaluating traffic signal detector health developed within this project to discover traffic detectors with degraded and declining performance and will be referred to as "the algorithm" within this chapter. The research team held two Zoom meetings with the SPR 837 TAC to request input on this task deliverable and solicit feedback on a draft version. Additionally, the team software engineer worked directly with one member of the TAC in development of this final version.

5.1 SYSTEM DESIGN

This design is not intended to be a detailed prescription which ought to be built exactly as specified and should not be viewed as such. It is practically impossible create such prescriptive designs which actually meet the needs of the system's users without extensive work to determine all their requirements for the system. Instead, this design should be viewed as part of the raw materials from which the ultimately produced software will be derived. Thus, alterations by ODOT or future implementers of the system should be expected and welcomed in so far as they improve the fit of the system to ODOT's needs.

What follows is a high-level description of the system which details the overall "shape" of the system along with its key components and the relationships between them. This description remains purposefully agnostic concerning the particular tools, technologies, and data formats which may be used in the implementation of the system. Suggestions for and discussion of these implementation details, can be found in the implementation plan.

5.1.1 High Level Description

The system centrally consists of a server which is capable of executing the algorithm for a given traffic detector and reporting the results of the algorithm's execution. The server supports two modes of interaction: in the first, the server automatically runs the algorithm periodically for all known traffic detectors and makes the results available via a web dashboard which can be accessed directly by the users; in the second, a user first parameterizes a report request via a web form which is then submitted to the server and processed accordingly, yielding a report according to the specified parameters. These modes of interaction are shown in Figure 5.1.

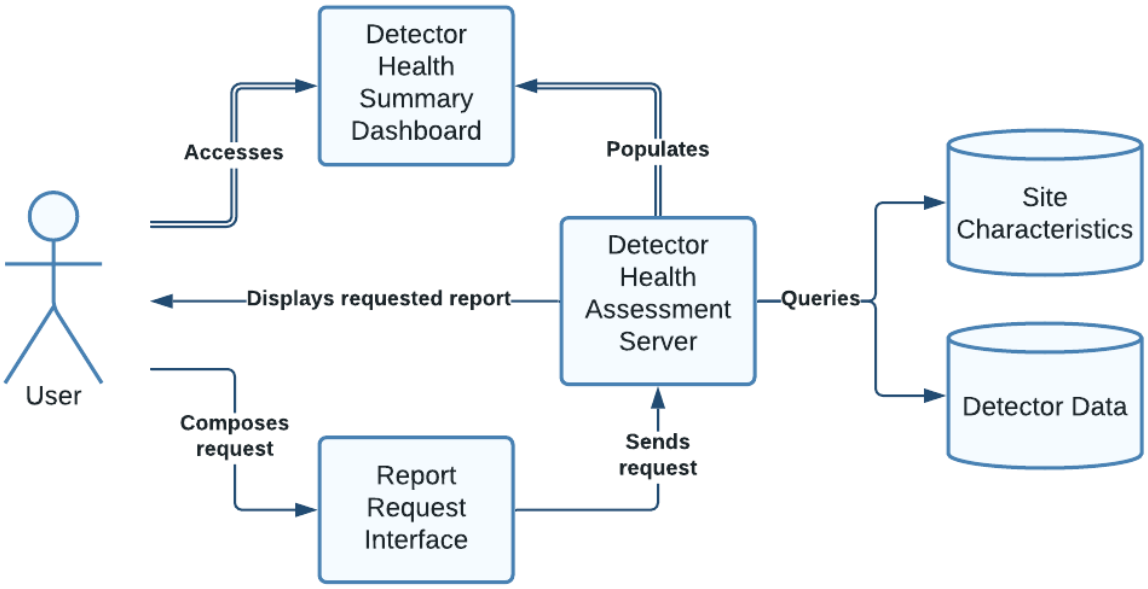


Figure 5.1: An overview of the detector health assessment system

To execute the algorithm for a given detector and therefore to provide both of these modes of interaction, the server must access certain data related to the detector to be assessed. Namely, it must obtain actuation data for the detector and also data which describes various aspects of the detector itself and the circumstances of its deployment which are referred to collectively as "site characteristics". These two data types are depicted in Figure 5.1 as housed in two distinct databases, but this is done solely to support easy visualization of the two types of data necessary for the algorithm's execution. In fact, the two types of data could be stored in a single database or across several databases with little real impact on the system's design or implementation.

5.1.2 Data Description

As stated above, the algorithm for assessing traffic detector health requires two types of data related to the detector to be assessed, as shown in Figure 5.1. The first type of data is the detector data itself which consists of activations per hour and green indications per hour. These numbers are not directly available for access in a datastore, but the raw actuation and green indication data are. As such, for a given time window the system can query the database(s) for the individual activation and green indication records occurring within the time window of interest and then perform some simple arithmetic to arrive at the hourly rates.

The second type of data needed by the algorithm is a specific set of attributes related to the detector and its deployment, namely:

- Whether the detector is a loop or radar detector
- Whether the detector is a single or multi-lane detector
- The speed limit at the detector's location

- The detector's distance from stop bar

This data is not presently available in a database system, and as such it must be collected and stored as a prerequisite to implementing the system. Suggestions for collecting and storing this data, and what other characteristics may be desirable to catalogue, are given in the implementation plan.

5.1.3 Interface Descriptions

This system provides two modes of interactions via two distinct interfaces, as shown in Figure 5.1. The first interface is the detector health summary dashboard which is automatically populated by the server using the latest data from all of the detectors known to the system. This dashboard would be similarly constructed to the department's existing detector health assessment dashboard, showing a list of all detectors ranked by their current health status and providing access to a per detector detailed view when a particular detector is selected.

The second interface is the report request interface which consists of a web form for parameterizing a detector health report. This form would allow the user to select, at minimum, which detectors should be assessed for the report and what time window should be used for the assessment. If desirable, the system could also be made to allow the user to select detectors based on combinations of their site characteristics too, e.g., a report "for all loop detectors deployed where the speed limit is less than 45mph". In any case, the resulting report could take the form of a web page returned to the user or in a downloadable file such as a PDF.

The primary purpose of the dashboard interface is to facilitate assessment of the current "state of the state" with regard to detector health, while the report request interface allows the user to investigate possible trends across site characteristics or to investigate the historical health status of certain detectors. The report request interface also allows for a detector health assessment to be easily disseminated as a static document as well.

5.1.4 Deployment Considerations

Where and how this system is deployed may have an effect on the overall design of the system as well. For example, if the system is to be deployed in such a way that it would only be accessible via a secured department intranet or Virtual Private Network (VPN), then regulating access to the system and tracking its usage may not be necessary. However, if the system were to be accessible via the public internet, then a login subsystem may need to be added to facilitate both access management and user tracking. The overall security needs of the system would need to be increased in this case. There is no way that all such design modifications dependent on the system's deployment can be predicted or exhaustively described and, as such, it is highly recommended that at the time the system is implemented these items are discussed between ODOT and the future implementers of the system.

5.2 IMPLEMENTATION PLAN

The main suggestion which forms the backbone of this implementation plan is to grow this system incrementally. It is recommended that ODOT begin by building a minimum viable

product (MVP) which is a technique for developing a product to be sufficient for early adopters, and the final features are designed after considering the initial users' feedback. (Ries 2009) In this case, the "early adopters" would actually be a select group of ODOT employees. This MVP will constitute the "seed" of the system which may be iteratively improved to arrive at the final system. Finally, suggestions are provided for the gathering and cataloguing of the detector site characteristics.

5.2.1 The MVP Prototype System

The MVP system centrally consists of a program which will read in data and relevant site characteristics for a single traffic detector from data files, execute the algorithm to evaluate the health of a detector, and then output the results of the algorithm. There is no configuration to be performed for this system since there is only one detector to be evaluated and there is no interaction with the databases since the detector data and site characteristics are located in static files. Figure 5.2 shows a summary view of the system.

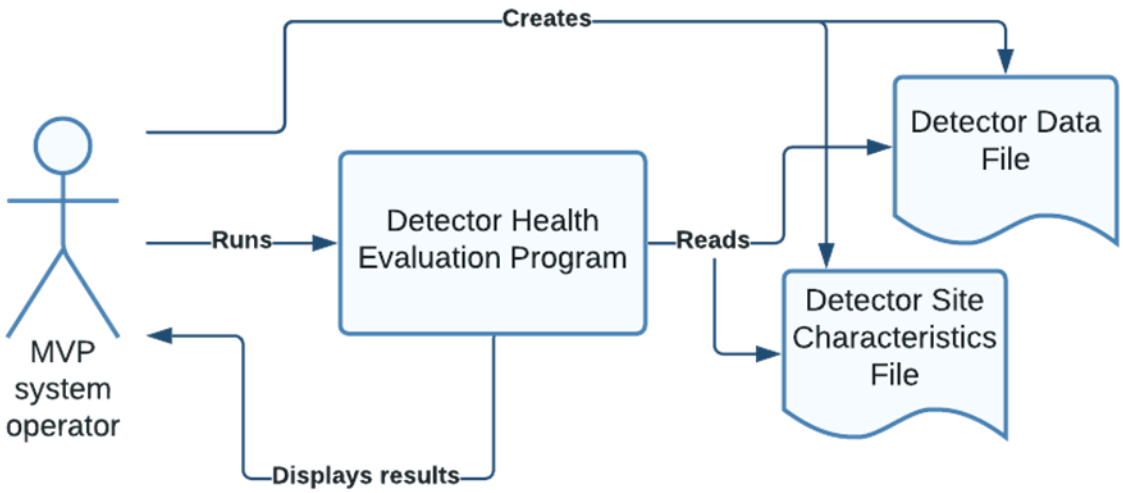


Figure 5.2: The MVP prototype system

Since the scope of this system is intentionally limited, it is suggested that it be built in-house using whatever programming language and tools the implementers desire. This will eliminate the overhead of interaction with external personnel and any learning curve associated with unfamiliar technologies. However, some consideration should be given to the available level of support for mathematical processing within the implementation tools selected since the algorithm requires some calculus. The Python and R programming languages should have sufficient support for this, via scientific computing libraries for Python and natively for R, but these are far from the only viable choices.

This system will allow a basic level of interaction with the algorithm and the opportunity to fine tune it, allow for ad hoc assessment of individual detectors if desired, but most importantly,

provide a reference implementation of the algorithm which could be used by other implementers in the future even if they are not utilizing the same technologies in their work.

5.2.2 Iterative Improvement

While the MVP provides distinct advantages for starting system implementation off well, it is not the complete system described in the system design. Nonetheless, it provides a good starting point for iteratively adding features in the next phase of development. Whether that development continues with the same implementers who built the MVP or not, prior to further build out, the choice of technologies to be used should be revisited.

At this stage, it is recommended to make use of stable, mainstream technologies to ensure the long-term viability of the system. Python is an extremely mature language with an impressive ecosystem of libraries and frameworks which will facilitate implementing the system. In particular, the NumPy library is suggested for the math involved in the algorithm and the Flask web server framework for the server itself.

At this stage, ODOT will need to confer internally or with the implementers to come up with a prioritized list of features which are desired to be implemented. Specific consideration should be paid to whether the dashboard interaction or the report generation interaction should be prioritized first. As the databases with which the system must interface come online, both of these features will become more useful, but there is no need to wait for a complete database to be in place before implementation as whatever data is available at the moment can be used, even data in static files.

5.2.3 Cataloguing Site Characteristics

In closing this chapter, it is deemed important to present a few words about the cataloging of site characteristics, as without data, even a well-engineered implementation of the system will not be very useful. As such, the cataloguing of site characteristics for all detectors of interest must be a high priority alongside the implementation of the system.

A simple table in an SQL database should be sufficient for recording detector site characteristics in the most basic way. Sufficient fields within the table for the necessary site characteristics and an ID field which can be used to correlate site characteristics with detector data is all that is required as shown in Table 5.1.

Table 5.1: Example Relevant Site Characteristics in an SQL Database

Detector ID	Detector Type	Speed Limit	Stop Bar Distance	Lanes
348	loop	55	20	2
972	radar	40	10	1

Since the primary function of the “Detector ID” field is to correlate the site characteristics with the activation data for that detector, it is suggested that it be the same as the ID of detectors in the existing activation database, or, if changed, thought be given to the choice of ID such that is

complementary with other ODOT databases, if present. It would also be helpful if the detectors were associated with a human readable name if the ID column is numeric as depicted here.

Of course, this is only the data strictly necessary for the algorithm to function, but additional site characteristics such as which lane the detector is in, the type of road the detector is deployed on, etc. may be added to this table as well and may be of interest for future exploration of trends in detector health degradation. Additionally, more complex arrangements of the data are possible such as having tables which describe roads and intersections independently of the detector table and linking them together via foreign keys or bridge tables, but it is not clear that such an arrangement would have any clear advantages over the simple one presented at this time.

Getting the site characteristic data into the database could be accomplished via the system itself via additional functionality. A feature could be added which would allow a user to interact with a web form which would allow them to select a detector and then view or edit its site characteristics, filling in some or all of them, and then save the form. Such a feature would only be worth implementing if it seems likely that detectors will be described by department staff over a longer period of time while the system is live, rather than by some third-party all at once, perhaps in a spreadsheet which is then loaded directly into the database.

6.0 CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

In this work, event-based detector data were processed for variety of detectors at select signalized intersections in Oregon to first identify which detectors were performing sufficiently to be used in an algorithm, and then for development of the algorithm. A total of 79 detection zones underwent the initial comparative analysis (70 inductive loop and 9 radar). Of the inductive loop detection zones, 39 passed the analysis, while 6 of the radar zones passed the analysis, for a total of 45 valid detection zones. The subset of detection zones that passed the comparative analysis include stop line, advanced, single lane, multiple lane, short, and long detection zones over a variety of lane usages and provided a robust basis for development of a detector health monitoring algorithm.

After cleansing the data and prepping it for analysis as well as selecting peak periods for data aggregation, empirical lines of best fit for the relationship of EHV and Density were developed for each week, with these relationships then being used to build two datasets for algorithm development as well as for initial health assessments and health assessments over time. For the former, a predicted empirical Volume vs. Density curve compared to a conceptual Volume vs. Density curve along with week-to-week percent differences between empirical curves are compared to datasets developed in this work to provide an assessment of initial health. For the latter, week-to-week percent differences are also compared to data developed in this work using the sliding window and control chart techniques. A value more than 1.5 standard deviations from the mean of the dataset was proposed as a starting point for health assessment, but this may be adjusted as desired by ODOT staff, as this value may not be sensitive enough to identify underperforming detectors, as noted in Section 4.8.2.

Moving forward, as noted in the paragraph above, it would be advisable to investigate different bounds for control chart limits (presumably statistically based). Because existing literature provided no guidance for applying this method to detector data, commonly used bounds were applied, which may not be the most suitable for vehicle detector health monitoring. Additionally, as this algorithm is implemented, it is advised to develop percent difference datasets for detectors of various technologies and configurations so that comparisons can be made between field detectors and datasets developed from detectors with similar characteristics. This can be done by segregating percent difference data from various detectors as additional sites are brought online. This should allow for tighter control limits for determining sensor health, as variation in the comparative data set would be limited by the homogenous categorization of detectors.

There are several limitations of this work that should be noted. First, the developed algorithm and datasets are modeled from a finite number of detectors. As such, the datasets developed for health assessment are based upon this set of analyzed detectors, which may not be a universally representative sample of detectors across the ODOT system. Second, due to time constraints, the algorithm developed in this work was not subjected to long term testing and validation. Based upon comparisons between the healthy and unhealthy detector datasets in this work, the methods proposed will identify a variety of unhealthy detector operations, however the thresholds chosen can be tightened up with further testing, as noted above. Both of these items can be addressed

after deployment of the algorithm on the ODOT system, as the detector set will be much more diverse, and additional data could then be used to further tune the detector health thresholds.

7.0 REFERENCES

American Association of State Highway and Transportation Officials. (2011). *A policy on geometric design of highways and streets* (6th ed.). Washington, D.C.: American Association of State Highway and Transportation Officials (AASHTO).

Arnold, J., Gibson, D., Mills, M., Scott, M., & Youtcheff, J. (2011). Using GPR to unearth sensor malfunctions. *Public Roads, 74*(4), 24-29.

Asamer, J., & van Zuylen, H. J. (2011). Saturation flow under adverse weather conditions. *Transportation Research Record: Journal of the Transportation Research Board*, *2258*(1), 103–109. https://doi.org/10.3141/2258-13

Azin, B. & Yang, X.T. (2020). *Multi-stage algorithm for detection-error identification and data screening* (Report No. UT-20.15). Salt Lake, UT: Utah Department of Transportation. Retrieved from https://rosap.ntl.bts.gov/view/dot/60813

Bertini, R. L., Boice, S., & Bogenberger, K. (2006). Dynamics of variable speed limit system surrounding bottleneck on German Autobahn. *Transportation Research Record: Journal of the Transportation Research Board*, *1978*(1), 149–159. https://doi.org/10.1177/0361198106197800119

Chen, C., Kwon, J., Rice, J., Skabardonis, A., & Varaiya, P. (2003). Detecting errors and imputing missing data for single-loop surveillance systems. *Transportation Research Record: Journal of the Transportation Research Board*, *1855*(1), 160–167. https://doi.org/10.3141/1855-20

Day, C. M., Brennan, T. M., Sturdevant, J. R., & Bullock, D. M,. (2011) *Performance evaluation of traffic sensing and control devices*. (Report No. FHWA/IN/JTRP-2011/17). West Lafayette, IN: Joint Transportation Research Program, Indiana Department of Transportation and Purdue University. https://doi.org/10.5703/1288284314641

Day, C., Bullock, D., Li, H., Remias, S., Hainen, A., Freije, R., . . . Brennan, T. (2014) *Performance measures for traffic signal systems: An outcome-oriented approach*. West Lafayette, IN: Purdue University. doi: 10.5703/1288284315333.

Day, C. M., Premachandra, H., Brennan, T. M., Sturdevant, J. R., & Bullock, D. M. (2010). Operational evaluation of wireless magnetometer vehicle detectors at signalized intersection. *Transportation Research Record: Journal of the Transportation Research Board*, *2192*(1), 11–23. https://doi.org/10.3141/2192-02

Deng, H., & Zhang, H. M. (2012). Driver anticipation in car following. *Transportation Research Record: Journal of the Transportation Research Board*, *2316*(1), 31–37. https://doi.org/10.3141/2316-04

Dhaliwal, S. S., Wu, X., Thai, J., & Jia, X. (2017). Effects of rain on freeway traffic in Southern California. *Transportation Research Record: Journal of the Transportation Research Board*, *2616*(1), 69–80. https://doi.org/10.3141/2616-08

Econolite. (2020a). AccuSense - Control. Retrieved from https://www.econolite.com/solutions/sensors/accusense/

Econolite. (2020b). *Centracs SPM Specifications*. Retrieved from https://www.econolite.com/support-page/technical-documentation/systems/centracs-spm/

Greenshields, B. D. (1935). A study in highway capacity. In *Highway Research Board Proceedings of the Fourteenth annual meeting: Held at Washington, D.C., December 6-7, 1934* (pp. 448-447). Washington, D.C.: Highway Research Board. Retrieved from https://onlinepubs.trb.org/Onlinepubs/hrbproceedings/14/14P1-023.pdf

Grossman, J., Hainen, A. M., Remias, S. M., & Bullock, D. M. (2012). Evaluation of thermal image video sensors for stop bar detection at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, *2308*(1), 184–198. https://doi.org/10.3141/2308-20

Huotari, J. (2015). *Radar-based traffic detectors* [Photograph]. Retrieved from https://oakridgetoday.com/wp-content/uploads/2015/12/Stoplight-Radar-Detection-Dec-11-2015.jpg

Hurwitz, D. S., Knodler, M. A., Nyquist, B., Moore, D., & Tuss, H. (2012). Evaluating the potential of advanced vehicle detection systems in mitigating dilemma zone safety conflicts. *ITE Journal 82*(3), 24-28. https://trid.trb.org/view/1135967.

Indiana Department of Transportation (INDOT). (2015). *Procedure for evaluating vehicle detection performance* (Report No. ITM 934-15). Indianapolis, IN: Indiana Department of Transportation. Retrieved from https://www.in.gov/indot/div/mt/itm/pubs/934_testing.pdf.

Iteris. (2020). Iteris vantage vector. Retrieved from https://www.iteris.com/oursolutions/traffic-detection/vantage-vector

Jin, J., Ma, X., Koskinen, K., Rychlik, M., & Kosonen, I. (2016). Evaluation of fuzzy intelligent traffic signal control (FITS) system using traffic simulation. In *Transportation Research Board 95th annual meeting January 10-14, 2016, Washington, D.C.* Washington, D.C.: Transportation Research Board. Retrieved from https://trid.trb.org/view/1393640

Jin, J., & Ran, B. (2009). Automatic freeway incident detection based on fundamental diagrams of traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, *2099*(1), 65–75. https://doi.org/10.3141/2099-08

Khnel, C., Weisheit, T., & Hoyer, R. (2011). Malfunction sniffing – A new approach for on-site quality evaluations of traffic data acquisition infrastructure. In *18th World Congress on Intelligent Transport Systems*. Retrieved from

https://www.researchgate.net/publication/268097738_Malfunction_Sniffing_-_A_New_Approach_for_On-Site_Quality_Evaluations_of_Traffic_Data_Acquisition_Infrastructure

Lamas-Seco, J., Castro, P., Dapena, A., & Vazquez-Araujo, F. (2016). SiDIVS: Simple detection of inductive vehicle signatures with a multiplex resonant sensor. *Sensors*, *16*(8), 1309. https://doi.org/10.3390/s16081309

Laufer, J., Yue, W.L., Shahhoseini, Z., & Javanshour, F. (2019). Determination of saturation flows in Melbourne. In *Proceedings Australasian Transport Research Forum; September 30-October 2, Canberra, Australia.* Retrieved from https://australasiantransportresearchforum.org.au/wp-content/uploads/2022/03/ATRF2019_resubmission_29.pdf

McCain. (2020). Traffic signal controllers & modules | McCain. Retrieved from https://www.mccain-inc.com/products/controllers

Medina, J. C., Ramezani, H., & Benekohal, R. R. F. (2013). Evaluation of microwave radar vehicle detectors at a signalized intersection under adverse weather conditions. *Transportation Research Record: Journal of the Transportation Research Board*, *2366*(1), 100–108. https://doi.org/10.3141/2356-12

Middleton, D., Charara, H., & Longmire, R. (2009). *Alternative vehicle detection technologies for traffic signal systems: Technical report* (Technical Report No. 0-5845-1). Austin, TX: Texas Department of Transportation. Retrieved from https://static.tti.tamu.edu/tti.tamu.edu/documents/0-5845-1.pdf.

Middleton, D., Longmire, R., Bullock, D. M., & Sturdevant, J. R. (2009). Proposed concept for specifying vehicle detection performance. *Transportation Research Record: Journal of the Transportation Research Board*, *2128*(1), 161–172. https://doi.org/10.3141/2128-17

Minge, E., Kotzenmacher, J., & Peterson, S. (2010). *Evaluation of non-intrusive technologies for traffic detection* (Report No. MN/RC 2010-36). St. Paul, MN: Minnesota Department of Transportation. Retrieved from https://www.lrrb.org/pdf/201036.pdf.

Mohajerpoor, R., & Ramezani, M. (2019). Mixed flow of autonomous and human-driven vehicles: Analytical headway modeling and optimal lane management. *Transportation Research Part C: Emerging Technologies*, *109*, 194–210. https://doi.org/10.1016/j.trc.2019.10.009

Montgomery, D. C., & Runger, G. C. (2018). *Applied statistics and probability for engineers* (7th ed.). Hoboken, NJ: Wiley.

Nihan, N., Wang, Y., & Cheevarunothai, P. (2006). *Improving Dual-Loop Truck (and Speed) Data: Quick Detection of Malfunctioning Loops and Calculation of Required Adjustments* (Report No. TNW2005). Seattle, WA: Transportation Northwest Regional Center X. Retrieved from *https://rosap.ntl.bts.gov/view/dot/38617*

Potts, I. B., Ringert, J. F., Bauer, K. M., Zegeer, J. D., Harwood, D. W., & Gilmore, D. K. (2007). Relationship of lane width to daturation flow rate on urban and suburban signalized intersection approaches. *Transportation Research Record: Journal of the Transportation Research Board*, *2027*(1), 45–51. https://doi.org/10.3141/2027-06

Q-Free Intelight. (2020). MAXVIEW Atms | Intelight is now Q-Free | learn more. Retrieved from https://www.q-free.com/intelight/

Raksuntorn, W., & Khan, S. I. (2003). Saturation flow rate, start-up lost time, and capacity for bicycles at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, *1852*(1), 105–113. https://doi.org/10.3141/1852-14

Rhodes, A., Bullock, D., & Sturdevant, J. (2006). *Evaluation of stop bar video detection accuracy at signalized intersections* (Report No. FHWA/IN/JTRP-2005/28). West Lafayette, IN: Joint Transportation Research Program, Indiana Department of Transportation and Purdue University. https://doi.org/10.5703/1288284313401.

Ries, E. (2009). Minimum viable product: A guide. Retrieved from http://www.startuplessonslearned.com/2009/08/minimum-viable-product-guide.html

Sharma, A., Harding, M., Giles, B., Bullock, D., Sturdevant, J., & Peeta, S. (2008). Performance requirements and evaluation procedures for advance wide area detectors. In *Transportation Research Board annual meeting: 87th: 2008: Washington, DC: Compendium of papers*. Washington, D.C.: TRB.

Smaglik, E. J., Sharma, A., Bullock, D. M., Sturdevant, J. R., & Duncan, G. (2007). Event-based data collection for generating actuated controller performance measures. *Transportation Research Record: Journal of the Transportation Research Board*, *2035*(1), 97–106. https://doi.org/10.3141/2035-11

Smaglik, E. J., Sharma, A., Liu, C., & Kothuri, S. (2017). *Improving adaptive/responsive signal control performance: implications of non-invasive detection and legacy timing practices* (Report No. FHWA-OR-RD-17-07). Salem, OR: Oregon Department of Transportation. Retrieved from https://www.oregon.gov/odot/Programs/ResearchDocuments /SPR781_Adaptive_Responsive_Signal.pdf.

Transportation Research Board. (2016). *Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis (Nchrp Report)*. National Academies Press. https://doi.org/10.17226/24798

Tufte, K., Bertini, A., Auffray, B., & Rucker, J. (2007). Toward systematic improvement of data quality in Portland, Oregon, regional transportation archive listing. In *TRB 86th annual meeting: Compendium of papers CD-ROM: January 21-25, 2007, Washington, D.C.* Washington, D.C.: Transportation Research Board.

Vanajakshi, L., & Rilett, L. R. (2006). System wide data quality control of inductance loop data using nonlinear optimization. *Journal of Computing in Civil Engineering*, *20*(3), 187–196. https://doi.org/10.1061/(asce)0887-3801(2006)20:3(187)

Wang, H., Li, J., Chen, Q. Y., & Ni, D. (2011). Logistic modeling of the equilibrium speed–density relationship. *Transportation Research Part A: Policy and Practice*, *45*(6), 554–566. https://doi.org/10.1016/j.tra.2011.03.010

Wavetronix - SmartSensor V. (2020). Retrieved from https://www.wavetronix.com/products/smartsensor-v

APPENDIX A

PSEUDOCODE

This pseudocode describes how to process the raw data outputs from a signalized intersection. This section does not contain any new information but is meant to assist in programming development.

INITIAL PROCESSING

This section describes the initial filtering, and the process for collecting the number of vehicle indications per green interval, and green interval duration information.

From the raw data, filter the data to get the corresponding site number, the corresponding *MaxTime* number data, the corresponding§ intersection phase number data, and the timeframes and date(s) that are being analyzed. Only detector and its corresponding phase should be analyzed at once.

> The timeframes and dates used for analysis are the morning and evening peak hours (6:00 – 9:00; 16:00 – 19:00) on Tuesdays, Wednesdays, and Thursdays. A detector's combined data analysis for a Tuesday, Wednesday, and Thursday during those hours is considered one week of data analysis for that detector.

Organize the isolated data chronologically.

For the Parameters of each event, identify repeating 1's and 8's. Remove the cycle this occurs during, and the chronologically adjacent cycles. Remove the 81 and 82 indications that occur during those cycles as well.

> From the intersection outputs, parameter 1 represents the Green Start indication for the corresponding phase; parameter 8 represents the Yellow Start indication. It is necessary to remove any repeats in these indications (two 1's in a row, or two 8's in a row) because they introduce misrepresentations of the cycle duration and the Green Interval duration.

For the Parameters of each event, identify repeating 82's and 81's. Remove the 82's and 81's that are chronologically adjacent to the repeat as well.

> From the intersection outputs, parameter 82 represents the Vehicle Detector On indication for the corresponding detector. It is necessary to remove any repeats in these indications (two 82's in a row, or two 81's in a row) because they lead to misrepresentation of number of Vehicle Detector indications, Detector On durations, and because the cause of these repeats in each instance is indeterminable from the data.

Identify the number of Vehicle Detector On indications that occur chronologically during each Green Interval, and during each Not-Green Interval (i.e. the Yellow/Red Interval).

For each separate Green Interval – chronologically between the Green Start Time (1) and the following Yellow Start Time (8) – count the number of Vehicle Detector On (82) indications. Do the same for the Not-Green Interval – chronologically between the Yellow Start Time (8) and the following Green Start Time (1).

The purpose of this is to find the number of times the Vehicle Detector indicated a vehicle during the Green Interval. Finding the number of activations during the Green Interval only is so that this information could be used to evaluate the detector's health; the Green Interval has a more consistent flow that could be better analyzed using existing traffic theory than the Not-Green interval has. The activations found during the Not-Green Interval is necessary so that this activations information is available for entire cycles in addition to just the Green Intervals.

Identify the duration of each Cycle, each Green Interval, and each Not-Green Interval.

The duration of each cycle would be found through the Green Start (1) times. The duration of each Green Interval would be from a Green Start (1) time to the following Yellow Start (8) time. The duration of each Not-Green Interval would be from a Yellow Start (8) time to the following Green Start (1) time.

Organize the Number of Activations that occurred during a Green Interval and the corresponding Green Interval Duration into a dataset that lists the number of activations during each green interval.

REDUCTION FOR EQUIVALENT HOURLY VOLUME AND DENSITY

This section describes the process for appropriately filtering the data to better estimate continuous saturated traffic flow for later analysis and reducing the data to EHV and Density for each cycle. Approximating uninterrupted saturated traffic flow is necessary for analyzing the data using existing traffic theory. This section uses the Detector On (82) indications to represent vehicles and uses only activations which occurred during a green interval.

Using the data from the previous section's process, identify the initial four vehicles (Vehicle Detector On indications) of each green interval. Exclude those vehicles from the following data analysis process.

Toward the goal of approximating uninterrupted saturated traffic flow, the first four vehicles indicated during each Green Interval should be disregarded from the EHV and Density estimates, because the first four vehicle's headways are large than the headways in traffic free flow due to the start-up lost time. These larger headways are not reflective of a detector's health but are instead detrimental to an analysis that uses EHV and Density.

Determine the headways for each vehicle that occurs during a Green Interval.

Headways can be found by using the Vehicle Detector On (82) indication time and the previous vehicle Detector On (82) indication time.

Exclude each vehicle that has a headway above 3.0 seconds from the following data analysis process.

> Toward the goal of approximating uninterrupted saturated traffic flow, it is necessary to remove any vehicles with headways greater than 3.0 seconds from the data analysis process.

After these two exclusions (vehicles in positions 1 through 4 per green interval, and vehicles with headways larger than 3.0 seconds), this data is referred to as Filtered Activations.'

Determine each of these vehicles' Detector On Duration.

> Detector On Duration is the interval that the detector remained on for each vehicle. This can be found using the Vehicle Detector On (82) indication's chronologically adjacent Vehicle Detector Off (81) indication. Detector On Duration is used to find Occupancy and Density.

For each cycle, sum the Detector On Durations for all vehicles within that cycle's Green Interval.

For each cycle, sum the number of Vehicle Detector On indications during that cycle's Green Interval.

For each cycle, use the Green Start (1) time and the following Green Start Time (1) to determine that cycle's total duration.

DETERMINE LOCATION OF THE EMPIRICAL LINE

Using this information, find the Occupancy for each cycle.

For each cycle Occupancy can be found using the equation shown below:

$$\boldsymbol{Occupancy} = \frac{\boldsymbol{Filtered\ Detector\ On\ During\ Green\ \ Duration}}{\boldsymbol{Green\ Duration}}$$

(A-1)

Use the detector location's known Speed Limit and the detector's length, as well as an average vehicle length of 19.0 ft, to find the EHV for each cycle, using the equation A-2 shown below:

$$\boldsymbol{EHV} = {\boldsymbol{3600}}/{(\boldsymbol{3600} \times \boldsymbol{24} \times \boldsymbol{C})(\boldsymbol{A})}$$

(A-2)

where:

EHV = Equivalent Hourly Volume

C = Cycle Duration

A = Number of Activations per Green Duration

Density should be found for each cycle, using the equation shown below:

$$\boldsymbol{D = O \times 5280 / (L_{Veh} + L_{Det})} \quad \textbf{(A-3)}$$

where:

D = Density

O = Occupancy

L_{Veh} = Average Vehicle Length

L_{Det} = Detector Length

Organize the EHV and Density information on a graph *Equivalent Hourly Volume vs Density*, where each point represents one cycle, and fit a quadratic line to the data (LINEST in Excel). Each graph should represent data from one week of analysis (6:00 – 9:00 and 16:00 – 19:00, Tuesday through Thursday). Using the quadratic equation best fit line, determine the *Coefficient of Determination* (R^2) for that day of data.

DETERMINE ORIGINAL PERFORMANCE OF DETECTOR

Calculate Conceptual Volume vs. Density Curve and Use the Model to Predict the Location of the Empirical Curve

Collect EHV versus Density data during the peak hours (6:00 – 9:00) of Tuesday, Wednesday, and Thursday for multiple weeks. Calculate the quadratic equation for each week.

Collect the Average Headway for each detector. Use filtered data from the first day of data collection during the peak hour.

Using that Average Headway (seconds) and the Speed Limit (mph) of that detector's lane, calculate the Maximum Volume and Optimum Density using the following equations. Optimum Speed is based off of the posted speed limit.

$$\boldsymbol{S_O = \frac{1}{2}\, Speed\ Limit} \quad \textbf{(A-4)}$$

$$\boldsymbol{Maximum\ Volume = V_{MAX} = 3600 / Average\ Headway} \quad \textbf{(A-5)}$$

$$Optimum\ Density = D_O = {}^{V_{MAX}}/_{S_O}$$

(A-6)

Use the Maximum Volume and Optimum Density to create a quadratic equation with the vertex as (Optimum Density, Maximum Volume) = (x, y) and the intercept at (0,0). This is the conceptual Volume versus Density line for all days analyzed for the detector.

Using the equations below calculate the coefficients of the predicted line. Activations per hour is all activations during green, averaged across the week. Green indications per hour is averaged across the week. Other variables come from site data.

$$\hat{y}_a = \mathbf{0.629} - \mathbf{0.267}\left(x_{tech_{loop}}\right) - \mathbf{0.180}\left(x_{detect_{adv}}\right) - \mathbf{0.171}\left(x_{lane_{single}}\right) - \mathbf{0.001}\left(x_{wk_{acthour}}\right) - \mathbf{0.008}\left(x_{wk\_grnhr}\right)$$

(A-6)

$$\hat{y}_b = \mathbf{6.337} + \mathbf{3.773}\left(x_{tech\_loop}\right) + \mathbf{6.754}\left(x_{detect\_adv}\right) + \mathbf{4.700}\left(x_{lane\_single}\right) + \mathbf{0.064}\left(x_{wk\_acthour}\right) - \mathbf{0.136}\left(x_{wk\_grnhr}\right)$$

(A-7)

$$\hat{y}_c = -\mathbf{10.341} + \mathbf{9.171}\left(x_{tech\_loop}\right) + \mathbf{21.385}\left(x_{detect\_adv}\right) - \mathbf{29.725}\left(x_{lane\_single}\right) - \mathbf{0.047}\left(x_{wk\_acthour}\right) + \mathbf{1.458}\left(x_{wk\_grnhr}\right)$$

(A-8)

where:

$\hat{y}_a$, $\hat{y}_b$, $\hat{y}_c$ equals the predicted values of a, b, and c

x_{tech_loop} equals the presence of a loop detector (binary)

x_{detect_adv} equals the presence of advanced detector technology (binary)

x_{lane_single} equals site location within a single lane roadway (binary)

$x_{wk\_acthour}$ equals the number of activations per hour (continuous)

x_{wk_grnhr} equals the number of indications per hour (continuous)

Check if the Dataset is Viable for that Week

Verify that

- The *Coefficient of Determination* R^2 is equal to or greater than 0.70,

- The total number of data points for that week is at least 50 data points,
- The *Empirical Equivalent Hourly Volume vs Density* relationship curve is concave up, or the *a* value of the quadratic equation is negative, and
- The *Empirical Equivalent Hourly Volume vs Density* relationship curve results in a positive value when integrated from 0 to 25% of the conceptual vertex

If any of the above are not true, use data from a successive week. Repeat all processes above.

COMPARE INTEGRALS TO DETERMINE DETECTOR HEALTH

Determine Initial Detector Performance

The Conceptual Line is compared to the Predicted Line for that detector as a metric for initial detector health.

The conceptual line and predicted lines are integrated to 25% of the conceptual line's vertex.

The integration value for the conceptual line is compared to the predicted line through the percent difference. This percent difference is used as a metric for initial detector health.

The percent difference of the Conceptual Line and the Predicted Line integration values is compared to the PPD. An initial detector performance assessment is provided based on whether the percent difference is within 1.5 standard deviations of the mean of what is provided in the PPD.

Next, the first four viable weeks of Empirical data are compared to each other to determine initial detector health.

The EHV versus Density quadratic line for the first four viable weeks of detector data are each integrated to 25% of the conceptual line's vertex.

The percent differences are found comparing each of the four empirical lines' integration values, so that there should be six data points of percent difference information.

The mean is found of these percent difference values. This mean is used as a metric for initial detector health.

The mean of these percent differences is compared to the EPD. An initial detector performance assessment is provided based on whether the mean is within 1.5 standard deviations of the mean of what is provided in the EPD.

Calculate Empirical Lines Over Time / Determine Detector Performance Over Time

Four weeks of viable data and the integral values when integrated from 0 to 25% of the conceptual vertex are stored from the initial health assessment. The percent differences comparing each of these four weeks of integral values are also stored. If four weeks of viable data cannot be obtained from the initial health assessment, data is continually collected and processed until four weeks of viable data has been stored and plotted on the control chart. These need not be temporally successive weeks.

The mean of these percent differences is plotted on the control chart, with the mean and standard deviation of the EPD used in the creation of the control chart bounds.

The next week of raw data is processed and analyzed for its empirical curve.

Determine if the volume versus density for that week is viable.

Verify that

- The *Coefficient of Determination* R^2 is equal to or greater than 0.70,
- The total number of data points for that week is at least 50 data points,
- The *Empirical Equivalent Hourly Volume vs Density* relationship curve is concave up, or the *a* value of the quadratic equation is negative, and
- The *Empirical Equivalent Hourly Volume vs Density* relationship curve results in a positive value when integrated from 0 to 25% of the conceptual vertex

If any of the above are not true, use data from a successive week. Process and determine if the following week of data is valid.

If the data is viable, develop an empirical curve for the data, and calculate the integral of that line from 0 to 25% of the conceptual line's vertex. Then, calculate the percent differences between this integral and the previous three weeks, and average the result.

Plot the mean of these percent differences on the control chart. A point outside of the control chart bounds (based upon the EPD) indicates a possible detector health issue.

Continue to compare four weeks of sequential viable data at a time until twelve data points of means are plotted on the control chart.

The mean and standard deviation of those 12 weeks of data is used for the new control chart bounds, which allows the analysis to slowly adjust over time

COMPARE ACTIVATIONS DURING GREEN

The following are the R scripts to compare the activations during green. The script "Main.R" processes the data for each detector. The script "Main.R" sources the script "Code.R" for each detector to calculate the mean, standard deviation, and sample size and output the results.

Main.R

```
#install.packages("tidyverse")
library(tidyverse); library(readxl)
setwd("c:/WORK/R")
w1 <- file("Output1.txt","w")
w2 <- file("Output2.txt","w")
cat("mean sd week1m week1sd week2m week2sd week3m week3sd week4m week4sd
week2mChange week3mChange week4mChange\n", file=w1)
cat("SampleSize   WeeklyChange\n", file=w2)
for (i in 1:8) {
d <- read_xlsx("Site2_GreenDuration_vs_GreenActivations_Data.xlsx", sheet=i)
source("code.R")
}
for (i in 1:7) {
  d <- read_xlsx("Site55_GreenDuration_vs_GreenActivations_Data.xlsx", sheet=i)
  source("code.R")
}
for (i in 1:7) {
  d <- read_xlsx("Site245_GreenDuration_vs_GreenActivations_Data.xlsx", sheet=i)
  source("code.R")
}
for (i in 1:10) {
  d <- read_xlsx("Site502_GreenDuration_vs_GreenActivations_Data.xlsx", sheet=i)
  source("code.R")
}
for (i in 1:8) {
  d <- read_xlsx("Site503_GreenDuration_vs_GreenActivations_Data.xlsx", sheet=i)
  source("code.R")
}
for (i in 1:2) {
  d <- read_xlsx("Site585_GreenDuration_vs_GreenActivations_Data.xlsx", sheet=i)
  source("code.R")
}

close(w1);close(w2)
```

Code.R

```
d <- d[d$Activations!=0,]; d <- d[!is.na(d$Activations),]
d1 <- d[d$Week==1,]; d2 <- d[d$Week==2,]
d3 <- d[d$Week==3,]; d4 <- d[d$Week==4,]
m <- mean(d$Activations); sd <- sd(d$Activations)
m1 <-mean(d1$Activations); sd1 <- sd(d1$Activations)
m2 <-mean(d2$Activations); sd2 <- sd(d2$Activations)
```

```
m3 <-mean(d3$Activations); sd3 <- sd(d3$Activations)
m4 <-mean(d4$Activations); sd4 <- sd(d4$Activations)
c1 <- abs(m2-m1)/m1; c2 <- abs(m3-m2)/m2;
c3 <- abs(m4-m3)/m3;
s1 <- nrow(d1); s2 <- nrow(d2); s3 <- nrow(d3)
cat(m,sd,m1,sd1,m2,sd2,m3,sd3,m4,sd4,c1,c2,c3,"\n",file=w1,sep=" ")
cat(s1,c1,"\n",file=w2,sep=" ")
cat(s2,c2,"\n",file=w2,sep=" ")
cat(s3,c3,"\n",file=w2,sep=" ")
```