5-2012

An Evaluation of Non-Intrusive Traffic Detectors at the NTC/NDOR Detector Test Bed

Benjamin W. Grone
University of Nebraska-Lincoln, ben.grone@gmail.com

AN EVALUATION OF NON-INTRUSIVE TRAFFIC DETECTORS AT THE NTC/NDOR DETECTOR TEST BED

by

Benjamin W. Grone

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Civil Engineering

Under the Supervision of Professor Laurence R. Rilett

Lincoln, Nebraska

May, 2012

AN EVALUATION OF NON-INTRUSIVE TRAFFIC DETECTORS AT THE NTC/NDOR DETECTOR TEST BED

Benjamin W. Grone, M.S.

University of Nebraska, 2012

Adviser: Laurence R. Rilett

Throughout the field of transportation engineering, decision makers require quality information. The information used in transportation operations, planning, and design is based, in part, on data from traffic detectors. The need for quality data has spurred innovations in data collection including the introduction of modern, commercially available, non-intrusive traffic detectors. As these new technologies become available, a need exists to understand their capabilities and limitations—especially limitations that are unique to a specific region.

This thesis examined the accuracy of four non-intrusive traffic detector technologies considered for potential data collection applications on Nebraska's highways. The technologies evaluated included the Solo Pro II video image processor (VIP), 3M Canoga Microloop 702 magnetic induction detector, Image Sensing Systems RTMS G4 microwave radar detector, and Wavetronix SmartSensor 105 microwave radar detector. These four detectors were installed at the NTC/NDOR non-intrusive detector test bed along Interstate 80 near the Giles Road interchange in Omaha, Nebraska. Data were collected in June, July, and August of 2011, and these detectors were analyzed based on the accuracy of their volume, speed, and length-based vehicle classification.

The analysis in this thesis utilizes numerous graphical and statistical methods to demonstrate the significance of errors in the data from the four evaluated detectors. The impacts of lighting, rain, traffic volume, and various levels of temporal aggregation on the detectors' accuracies were analyzed. Multiple regression analysis revealed that the volume accuracy of the Solo Pro II was affected by night lighting, as well as by the combined effect of dawn lighting and rain. The volume accuracies of the Microloop 702 and G4 were significantly affected by the combination of dusk lighting and rain, while the volume accuracy of the SmartSensor 105 was not found to be significantly affected by lighting or rain conditions. In addition to these results, this thesis analyzed the collected data in order to provide hypotheses pertaining to potential links between significant environmental factors and physical operating characteristics of the evaluated non-intrusive traffic detectors.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my thesis adviser, Dr. Laurence Rilett, for his direction, guidance, and input relating to this thesis, as well as for the additional opportunities he provided and encouraged me to pursue during my graduate studies. I would also like to acknowledge the other members of my advisory committee, Dr. Anuj Sharma and Dr. Aemal Khattak, who have my sincere gratitude for their instruction and recommendations, and for the time they spent critiquing my work.

I would also like to express my gratitude for the other students and researchers in the Nebraska Transportation Center (NTC) Transportation Systems Engineering office, who I had the pleasure of getting to know and working with throughout this process. Thanks, especially, to Dr. Justice Appiah for his direction relating to my questions about the appropriate ways to approach statistical issues in my research. Also thanks to Dr. Bhaven Naik for similar types of instruction. I am also thankful for the numerous friends I have made among my peers in the office. I will not attempt to name them all here for fear of leaving a name out, but they know who they are.

There were also a number of people in the NTC business center that deserve recognition for the support they gave me throughout this process. Chris LeFrois and Larissa Sazama were an invaluable resource and provided much needed encouragement when I ventured outside my comfort zone into the realm of intelligent transportation systems (ITS) communications. Valerie Lefler deserves special recognition for her ability to smile and be an encouragement even when she had a hundred different things on her mind. The rest of the NTC business center staff also deserve recognition for the great

work they do—which largely goes unnoticed by the research assistants like myself but makes the research we do possible.

Outside of this office, I would like to thank the many people at the Nebraska Department of Roads (NDOR) who made their valuable time available to me throughout the course of this thesis. In the ITS section, I would especially like to thank Sarah Tracy and Steve Olson for their involvement throughout this study. I would also like to thank Don Wood, the District 2 electronics tech leader, for his assistance with hardware at the NTC/NDOR non-intrusive detector test bed throughout the study. Lastly, I would like to thank the various people at the District 2 Traffic Operations Center for their interaction throughout the study, and their interest in my safety when test bed visits were necessary.

I would like to express my gratitude to the many contacts with whom I interacted, who represent the manufacturers and distributors of the various non-intrusive detection technologies evaluated in this thesis. Many of these people were very generous with their time in helping me shape my understanding of the operation of the detectors. They also assisted with the proper calibration of the detectors, as well as in troubleshooting communications issues.

Finally, I would also like to thank my wife, Melani, for her support and relentless encouragement throughout the ups and downs of this thesis. Without her playing the role she did, this feat would not have been possible for me. Finally, I would like to thank my God, who, among his many other blessings, blessed me with a critical mind and natural curiosity that made this study engaging for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 Background

Decisions relating to highway transportation are made at many different administrative levels. These decisions are often based on information that comes from collected data. They can only be as sound as the collected data upon which they are based. The data used in traffic engineering generally fit into one of two categories. Inventory data, which address the available highway resources, include items such as road classification, cross-sectional characteristics, pavement quality indices, and intersection characteristics; this type of data is generally taken from design documents or by direct measurement. The second type of data is demand data, which is concerned with the degree to which the stated resources are currently, have historically, or are projected to be utilized. Demand data include items such as origin-destination matrices, travel time, traffic volume, and vehicle classification. Data on the characteristics of traffic on a given roadway or network are vitally important to management decision-making. Decision-makers work under the assumption that the data are reasonably reliable, but acknowledge that there will be errors inherent in a given dataset. While it is rather difficult to improve historical data, there has been an ongoing effort by officials responsible for data collection to improve the quality of data currently being collected, or that which will be collected in the future.

Since the 1960s, inductive loop detectors have been the primary source of vehicular traffic data, e.g., volume, speed, and classification (1). However, there are a number of problems presented by loop detectors that have warranted research into alternative means of traffic data collection. Some of the problems with inductive loop detectors include their high failure rate, the intrusive nature of their installation and

maintenance (traffic disruption and danger for installers), and their undermining of the structural integrity of the surrounding pavement (2, 3, 4). Research into detector technologies has yielded six major scientific properties that allow detectors to detect vehicles: sound, opacity, geomagnetism, reflection of transmitted energy, electromagnetic induction, and vibration (5). Most of the state-of-the-art detectors on the market fit into a category with one of these detected properties, or could be considered combination detectors (i.e., those which observe multiple properties of vehicles).

The goal of this thesis was to make statistical comparisons between some of the non-intrusive technologies currently available for traffic detection for performance under various environmental conditions. Statistical analyses on comparisons ranging from disaggregate presence detection to higher parameters such as speed and classification were conducted to arrive at value judgments of the various traffic detectors under examination. The evaluation of the detectors also included an analysis of the impacts environmental conditions exert on the various detectors. It was anticipated that the statistical analysis presented in this thesis would advance the field not only by delineating the characteristics of the set of non-intrusive traffic detectors upon which it was conducted, but also by informing future research on yet undeveloped traffic detectors.

1.1 Problem Statement

While there exists a substantial body of literature reporting on the accuracy of various traffic detector technologies, the majority of such research was conducted under ideal environmental conditions (adequate lighting, low wind, and no precipitation), or without explicit acknowledgment of the impacts that environmental conditions may have on detector accuracy. Because agencies that implement these technologies for traffic data

collection purposes do so with the expectation that the data they are receiving has a reasonable accuracy across environmental characteristics, a need exists to provide quantified, empirical assessment of the factors associated with adverse environmental conditions (such as low lighting, lighting transition, and precipitation), specifically those conditions frequently encountered in the state of Nebraska.

1.2 Research Objectives

The primary objective of an currently ongoing research pursuit in the field of traffic detectors, led by the Nebraska Transportation Center (NTC), is to provide a sound methodological framework for use in analyzing the fitness of various non-intrusive traffic detection technologies—technologies which, importantly, inform policy-makers and designers. As technology rapidly evolves, this is an ongoing task. The current study is valuable to this ongoing research, as it implements a series of statistical tools and analyses to closely examine and document the responsivity of numerous traffic data technologies to various environmental conditions. Analyses were conducted on four technologies that represent alternatives to the traditional inductive loop for traffic data collection. The study assessed the accuracy of vehicular traffic volume, speed, and length-based classification data, collected by one video detector, two different radar detectors, and a magnetic induction microloop detector under fair and adverse conditions including rain and lighting conditions (i.e., dawn, dusk, and night [dark]). Review of these data informs upon which of these detector technologies are most robust against adverse environmental conditions. A primary focus of this thesis was on scientifically defensible statistical analyses of the error rates of these four technologies, conducted under the full spectrum of potentially adverse environmental conditions.

1.3 Research Program

The research presented in this thesis was carried out by following the program of tasks in the order presented in this section.

1.3.1 Literature Review

The first step was to conduct a literature review examining the existing body of knowledge pertaining to state-of-the-art traffic detectors and their various accuracies. This review provided a base of evidence upon which to construct a research program capable of furthering collective understanding of this subject. This review was conducted by examining existing publications relevant to the historical and current use of traffic detectors, industry accepted inaccuracies, and technological limitations of different traffic detectors. The literature review is outlined in chapter 2 of this thesis.

1.3.2 Identification and Setup of Test Bed

The test bed for this detector study was an area along westbound Interstate 80 (I-80) at the Giles Road interchange in Omaha. This is a permanent traffic detector test bed maintained by the Nebraska Department of Roads (NDOR) and Nebraska Transportation Center (NTC). At this location, NDOR installed three above-ground detection systems and one buried detection system, which were each analyzed in this study. The buried detector was a 3M Canoga Microloop. The three above-ground systems were the Autoscope Solo Pro II, Image Sensing Systems RTMS G4, and Wavetronix SmartSensor SS105. The current research primarily involved the logistical planning of data collection; installation of additional site apparati for electronic communications and data collection; and calibration of the detectors. The test bed setup and detector calibration are documented in chapter 3 of this thesis.

1.3.3 Collection and Reduction of Data

Time-stamped vehicle observation, speed, and length data were collected from the four detection systems over a five-month period spanning March 2011 through July 2011. To facilitate analyses involving environmental conditions, weather data were collected from the KMLE weather station located at the Millard, Nebraska Airport, approximately 0.5 miles from the test bed. In addition to the collection of these data files, video was recorded so that subsequent manual observation could be conducted in order to establish ground truth vehicle count and classification, as well as manual verification of weather conditions. A subset of the collected, data representing various environmental and traffic conditions, was selected for analysis. Data reduction involved establishing ground truth from the recorded video and aggregating the output from the various detectors for this data set. Data collection and reduction are documented in chapter 4 of this thesis.

1.3.4 Analysis of Data

Data analysis took two forms. Aggregate analysis considered the detector performance in the detection of volume, speed, and vehicle classification over temporal aggregation intervals of one, five, and fifteen minutes. Disaggregate analysis considered the per-vehicle detection performance of the various detectors relating to presence, speed, and vehicle classification. While disaggregate analysis provided a resolution of data unobtainable in the aggregate analysis, the aggregate analysis provided information on detection abilities at an aggregation level consistent with the practical application of these detectors for intelligent transportation systems (ITS) support. Therefore, both types of analyses provided valuable information on the detection performance of alternative traffic detectors. Aggregate analysis is documented in chapter 6, while disaggregate analysis is

documented in chapter 7 of this thesis. The statistical methods utilized in the analyses are detailed in chapter 5.

1.3.5 Inference of Results

The trends that arose in the analyses were documented, and to the extent that it was practical, were also tested for statistical significance. Upon documentation of the findings, attempts were made to reconcile the findings with what was previously acknowledged regarding the physical operating characteristics of the various detection technologies, in order to offer potential explanations for the deviations from ground truth. These explanations are offered alongside the analysis description in chapters 6 and 7. The most significant of these results are reiterated in the conclusions in chapter 8, as are recommendations for future research relating to the assessment of non-intrusive traffic detectors.

1.3.6 Dissemination of Findings

This thesis documents the culmination of the results of the current study, but there have been other published documents and presentations focusing on specific aspects of this study, and future documents are in their planning stages. The purpose of these documents and presentations is to make the lessons and recommendations garnered from this research available to all interested parties.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

While an extensive body of research has analyzed various traffic detector technologies, there exists a need for further research based on the rate at which manufacturers are producing new detectors or improving algorithms for previously released detector technologies. It cannot be assumed that, simply because a given technology provided the best accuracy for cost five years ago, it will still be the best technology today. To this end, this literature review begins with a basic explanation of the different technologies that are used in state-of-the-art traffic detectors. It then presents the various metrics which have been used in previous studies to compare traffic detectors. Finally, the findings of the most relevant and most recent traffic detector technology evaluations are summarized to facilitate comparison with the results of this study.

2.2 Available Detection Technologies

One of the most basic schemes for the classification of traffic detectors divides them into the following three categories: intrusive detectors, non-intrusive detectors, and off-roadway technologies (2). Intrusive detectors refer to technologies that require the installation of the detector under, in, or on the roadway. Detectors of this type are characterized by the need to intrude upon and obstruct traffic flow during their installation and maintenance. This is frequently cited in the literature as causing additional delay, as well as placing the installer in a potentially dangerous location near traffic. Non-intrusive detectors refer to technologies which do not require obstruction of traffic during their installation and maintenance. Most frequently, these detectors are installed either alongside the roadway, or overhead. Finally, off-roadway technologies

refer to non-point technologies employed in the collection of traffic information. Examples of off-roadway technologies include probe vehicles, bluetooth vehicle reidentification, automatic vehicle identification (AVI), or remote imaging (satellite or aircraft). This literature review was primarily concerned with intrusive and non-intrusive detector technologies.

2.2.1 Intrusive Detectors

The most common intrusive detector is the inductive loop. An inductive loop detector is a system comprised of four parts, including one or more coils of wire embedded in or under the pavement, an electronics unit which provides the circuit with power and senses a change in inductance, a lead in wire from the loop(s) to the pull box, and a lead in cable from the pull box to the electronics unit in a controller cabinet (5). When a vehicle with conductive metal passes over the loop, the inductance is reduced, thereby increasing the frequency of the oscillator. The higher frequency is registered by the detector oscillator, and the vehicle's presence is registered.

Another type of intrusive detector is the pneumatic road tube (2). The pneumatic road tube is a tube laid across the travelled lane. The tube is capped so that the passage of a vehicle's tires over the tube increases the air pressure in the tube. This pulse of higher pressure is registered by a sensor at one end of the tube, which records an axle passage. Vehicle count, speed, and classification data are calculated from axle passages. The wear that these tubes receive makes them more suited to short-term installations than long-term data collection.

Magnetometers are intrusive traffic detectors that sense the earth's magnetic field. They have two or three distinct coils around perpendicular axes, and are therefore more

properly known as two-axis or three-axis fluxgate magnetometers (6). These multiple axes allow them to detect changes in both the vertical and horizontal components of the earth's magnetic field, which in turn allows magnetometers to detect the presence of stopped vehicles as well as the passage of moving vehicles. Magnetometers have greater lane discretion than the magnetic detectors discussed in the non-intrusive detectors section below, which means that they are less likely to register false calls from magnetic spillover. However, their larger size requires an intrusive installation, while some magnetic detectors can be installed non-intrusively.

A final class of intrusive traffic detector with a specialized application is weigh-in-motion (WIM), which is achieved through one of three primary technologies (7). The first of these technologies is the piezoelectric sensor, which is installed in a saw cut across the travel lane and produces a voltage proportional to the force exerted on it by the wheels of a single axle. The dynamic load is calculated from the detected voltage. The second type of WIM detector is a bending plate. A bending plate detector consists of high-strength steel plates in each wheel path of a travel lane. The bottom of each steel plate is equipped with a strain gauge. From the reported strain in both plates, the dynamic axle load can be calculated. The third type of WIM detector is a load cell. A load cell detector consists of a single load cell with two scales (one in each wheel path). The load cell is equipped with a strain gauge which registers the dynamic axle load. For each of the three systems, the dynamic load is processed through a calibrated computation which estimates the vehicle's static load. WIM detectors are frequently paired with a different detector, such as an inductive loop, to allow other parameters such as speed and vehicle classification to be recorded.

2.2.2 Non-Intrusive Detectors

Much research over the past two decades has been conducted toward the development and analysis of various non-intrusive detectors. Six classes of non-intrusive detectors have emerged, based on the respective technologies the detectors employ for vehicle detection. These classes are: video image processor, microwave radar, magnetic, acoustic, infrared, and combined technology. Each of these detector classes has varied in its degree of use by the industry, and each thrives in different applications. Table 2.1 provides a cursory list of non-intrusive detector models with their classification by technology.

Table 2.1 Non-Intrusive Detector Models

Manufacturer	Model	Technology
Econolite	Autoscope Solo Pro II	Video Image Processor
Econolite	Autoscope Solo Terra	Video Image Processor
Iteris	Vantage	Video Image Processor
Iteris	VersiCam	Video Image Processor
Miovision	Video Analysis Service	Video Image Processor
Traficon	Detector Board VIP	Video Image Processor
Traficon	TrafiCam	Video Image Processor
ISS	RTMS G4	Microwave Radar (FMCW)
GMH Engineering	Delta DRS1000	Microwave Radar (Doppler)
IRD	TMS-SA	Microwave Radar (Doppler)
MS Sedco	Intersector	Microwave Radar (FMCW)
MS Sedco	TC26-B	Microwave Radar (Doppler)
Naztec	Accuwave 150-LX	Microwave Radar (FMCW)
Stalker	Speed Sensor	Microwave Radar (Doppler)
Wavetronix	SmartSensor 105	Microwave Radar (FMCW)
Wavetronix	SmartSensor Advance	Microwave Radar (FMCW)
Wavetronix	SmartSensor HD	Microwave Radar (FMCW)
Wavetronix	SmartSensor Matrix	Microwave Radar (FMCW)
Wavetronix	SmartSensor V	Microwave Radar (FMCW)
Xtralis	ASIM MW 334	Microwave Radar (Doppler)
GTT	Canoga Microloop 702	Magnetic
MS Sedco	TC30	Acoustic (Ultrasonic)
SmarTek Systems	SAS-1	Acoustic (Passive)
OSI LaserScan	AutoSense	Infrared (Active)
Xtralis	ASIM IR 30x	Infrared (Passive)
Xtralis	ASIM DT 351	Combined (Doppler Radar, Passive Infrared)
Xtralis	ASIM DT 372	Combined (Ultrasonic, Passive Infrared)
Xtralis	ASIM TT 29x	Combined (Doppler Radar, Ultrasonic, Passive Infrared)

One type of non-intrusive traffic detector is the video image processor (VIP). This type of detector consists of a camera which captures video of the traffic stream, and a computer programmed with an algorithm to processes the recorded video. The computer recognizes changes between successive frames and extracts parameters about vehicles

that pass through the image (5). Two primary types of algorithms exist in VIP detectors: trip-line and tracking. Trip-line detection allows a user to program virtual detectors onto certain areas within the image. When a group of pixels near that area changes hue or lightness, vehicle presence at that location is registered. By defining the geometry of the image and placing multiple virtual detectors along a travel lane, a speed trap configuration is able to extract vehicle count, speed, and length parameters for vehicles in that lane. Tracking algorithms in VIPs are less fully developed and are generally considered to be more complex. While trip-line algorithms only monitor specific areas of the image for changes, a tracking algorithm monitors the entire image, thereby recognizing a vehicle as it enters the frame, tracking it through the image. Based on calibration of image geometry, this type of algorithm is able to extract parameters such as vehicle count, speed, and length. VIPs with tracking algorithms are also useful for their ability to register turning movement counts at intersections. One example of a trip-line VIP detector is the Autoscope Solo Pro II, evaluated in this study.

Another type of non-intrusive detector is microwave radar. Microwave radar functions by emitting an electromagnetic wave toward the roadway (6). When a vehicle passes through the electromagnetic wave, it reflects a portion of the wave back to the detector. There are two types of microwave radar that differ in the way this reflected wave is processed. A continuous wave (CW) Doppler radar unit senses the shift in frequency between the transmitted signal and the detected return signal. This frequency shift is used to sense vehicle presence and calculate speed based on the Doppler principle. CW Doppler radar units are unable to detect stationary objects. A frequency modulated continuous wave (FMCW) radar unit transmits an electromagnetic wave, the frequency of

which is continuously being adjusted with time. Because of this modulated frequency, it is possible to determine the range (distance) to the vehicle. Successive range readings are used to determine the vehicle speed. A FMCW radar unit is able to detect stopped vehicles. Microwave radar units are either installed in an overhead (over one lane of traffic) or side-fire (transmitting perpendicular to the direction of traffic and across multiple lanes) configuration. Examples of microwave radar units include the Wavetronix SmartSensor 105 and ISS RTMS G4, evaluated in the current study.

A magnetic detector can fall into either the intrusive or non-intrusive category, depending on the model selected. This form of detector has been included under non-intrusive detectors in this thesis, due to the fact that the one magnetic detector assessed in this study was considered non-intrusive because it was installed in a conduit bored under the travel lanes from the side of the roadway. Other magnetic detectors are placed in saw cuts, or in holes cored into the pavement. Magnetic detectors function by passively sensing the vertical component of the earth's magnetic field (6). When the earth's magnetic field at the location of the detector is perturbed by the nearby passage of a ferrous object, a vehicle detection is registered. When two magnetic detectors are placed along a travel lane in a speed trap configuration, vehicle speed and length can be reported. Examples of magnetic detectors include the 3M Canoga Microloop 702, evaluated in this study.

The two types of acoustic traffic detectors are ultrasonic and passive acoustic (2). Ultrasonic detectors employ an active acoustic technology. They function by a) transmitting ultrasonic electromagnetic pulses and measuring the time it takes each pulse to be reflected back to the detector, or b) transmitting a continuous ultrasonic wave and

using the Doppler principle to detect vehicle presence. Passive acoustic detectors sense the different sources of sound associated with a vehicle, such as engine noise and tire/road interface noise, rather than transmitting an electromagnetic wave like the ultrasonic detector. They use an array of microphones, along with an algorithm capable of locating vehicles in the detection area. Both types of acoustic detectors are capable of collecting volume, speed, and classification data.

There are three classes of infrared traffic detectors on the market: active infrared, passive infrared, and infrared axle detectors. An active infrared detector is mounted over the roadway or in a crossfire configuration at the side of the road, and emits infrared beams toward the road surface, which are reflected to the detector. Passive infrared detectors function in a similar manner, except that they rely on electromagnetic energy emitted by the vehicle, or solar and atmospheric energy reflected off of the vehicle. In both cases, the infrared energy enters the detector through an optical system that directs it to an infrared-sensitive material, which generates an electrical signal that can be processed to determine vehicle presence (6). An infrared axle detector is mounted at ground level on one or both shoulders, depending on the model. It transmits an infrared laser across the travel lanes a few inches above the road surface. An axle is detected when the infrared signal is reflected off a wheel back to the unit (for single shoulder models), or when the infrared signal between the transmitter and receiver is disrupted by a wheel (for paired, i.e., two-shoulder units). The axle counts are aggregated into vehicle counts, speeds, and classifications based on axle spacing (8).

While each detection technology has its own strengths and weaknesses, manufacturers have learned to leverage the strengths of multiple technologies by creating

combined detectors. These detectors aggregate data from multiple sensors to create a more robust system. For example, there are detectors that combine an infrared sensor with either an ultrasonic or microwave radar sensor. In a combined passive infrared-Doppler radar detector the passive infrared sensor is able to register slow-moving (or stopped) vehicles that a Doppler radar sensor may miss, while the Doppler radar sensor is able to provide more accurate speed readings for faster moving vehicles than is the passive infrared sensor (2).

2.3 Standards for Evaluating Traffic Detectors

Committee E17.52 of ASTM International, a leader in the development of voluntary consensus standards, is responsible for the development of standards related to traffic monitoring. This committee is currently responsible for ten active standards (9). The most pertinent of these standards is the *Standard Test Methods for Evaluating Performance of Highway Traffic Monitoring Devices* (10). This standard provides guidance for two unique test methods that can be applied to a traffic monitoring device (TMD). The first method is a "type-approval test" and the second is an "on-site verification test," the outcome of either method being an *accept* or *reject* decision for the given detector. A type-approval test is to be applied to an untested brand and model of detector in order to determine its performance in a variety of potential installation scenarios. An on-site verification test is to be conducted at each installation location on a brand and model of detector that has already passed a type-approval test.

The standard is written in such a way that it could be referenced in purchase specifications. It outlines the responsibilities of the user and the seller in the testing process. The general process includes the following steps: the user must outline the traffic

parameters to be detected and the tolerance with which each parameter is to be reported; the user and seller must agree on the source of baseline data and the accuracy of the baseline data collection method; a type-approval test should include a minimum of three hours of data collection, while for most parameters, an on-site verification test only requires a minimum of 50 vehicle observations; the device is installed and calibrated by the seller and confirmed by the user; after data is collected by the device and the agreed upon reference mechanism, the errors are calculated and compared to the pre-defined tolerance specified by the user; if the error for any parameter exceeds the tolerance, the device is rejected.

As the test provides a simple *accept* or *reject* decision, the standard explicitly states that "no information is presented about either the precision or bias of the test method for measuring the performance of a TMD since the test result is non-quantitative" (10).

Another standard from ASTM International, which is closely tied to the above standard, is the *Standard Specification for Highway Traffic Monitoring Devices* (11). While the above standard is used to define the testing method in order to confirm that tolerances set in the purchase specifications are met, this specification provides guidance for the preparation of the purchase specifications. In doing so, it defines different traffic parameters that a detector could be required to measure; and also defines measures of tolerance to be used in testing, including percent difference, single-interval absolute value difference, and multiple-interval absolute value difference. Together, these two standards assist agencies in purchasing and installing traffic detectors that are capable of reporting traffic parameters within an expected error tolerance.

2.4 Previous Traffic Detection Evaluation Studies

Over the past two decades, researchers at a number of different agencies and institutions have conducted studies to assess various traffic detection technologies. The following synopsis of the most relevant of these studies summarizes the metrics that have been considered in assessing traffic detectors, as well as the different methodologies employed and relevant qualitative and quantitative findings. An emphasis is placed specifically on performance metrics relating to detection accuracy.

2.4.1 California PATH Studies

Since 1992, the California PATH coalition has sponsored a number of studies on various traffic detection technologies. These studies have addressed a broad range of research, including accuracy assessment of different video detection models at freeway and intersection locations; prototyping new wireless magnetic detection networks; developing automated data validation algorithms for loop detectors; and developing a system to automate "ground truth" data collection for future highway detector assessments. Relevant methods and findings from these studies are presented below.

The first independent assessment of VIP technology was conducted in 1992 by California PATH. The study compared three commercially available systems and five prototype systems, and involved the processing of 280 minutes of recorded video separately, using the different VIPs under examination (12). The set of video used was selected to include numerous scenarios with different characteristics, such as more or fewer lanes, various traffic volumes, approaching and departing traffic, steep to shallow camera angles, overhead versus side mounting, varying lighting conditions, and disparate weather conditions. Ground truth for count and speed was found by manual analysis of

the recorded video (including frame-by-frame analysis for true speed). The study differentiated the video detectors into two classes based on their detection algorithm: trip-line or tracking; the study reported average absolute percent error for each class of detector under each test condition. It was determined that under optimum conditions, trip-line detectors had greater count accuracy, while tracking detectors had greater speed accuracy. Conditions that were found to degrade performance were non-optimal camera placement, transition from day to night (dusk lighting), headlight reflections on wet pavement, shadows of adjacent vehicles or objects, fog, and heavy rain. In various conditions, trip-line detectors were found to have lower error rates in count and speed data than tracking detectors. However, the authors noted that all tracking detectors analyzed were prototypes at the time of testing.

A subsequent study developed a video vehicle tracking algorithm to detect traffic parameters by the processing of video images (13). This study focused primarily on the technical composition of the video processing algorithm, but is relevant to the current research; the functional specifications for the system under development in the study, which are provided in table 2.2, provide insight into the desired data quality for use in ITS applications. While some of these parameters, such as flow rate, average speed, and classification, could potentially be obtained from a single detector, other parameters listed in the table, such as link travel time and origin/destination tracking, require vehicle re-identification at multiple detector stations. Analysis of the tracking algorithm utilized by the study under review found it to be very effective for velocity measurement, but less effective for measuring flow, density, and spacing—the result of missed or false detections.

Table 2.2 Recovered Parameters (13)

Parameter	Units	Range	Reporting Rate	Error
Vehicle Flow Rate	veh/h/lane	0-2500	variable	± 2.5%
Average Speed	mph	0-90	variable	± 1 mph
Link Travel Time	min	0-60	variable	± 5%
Vehicle Classification	type count	0-2400	variable	± 5%
Lane Changes	changes by lane	as measured	variable	± 5%
Queue Length	veh/type/lane	as measured	variable	± 5%
Spatial Headway	ft/veh	as measured	variable	± 5%
Acceleration	mph/sec	as measured	variable	± 5%
Origin/Destin. Tracking	enter/exit location	0-500 veh/h/loc	tracked vehicle	± 10%

Another study under the California PATH program assessed issues relating to the implementation of a new advanced traffic control system in Anaheim, California (14). The new control system was to implement SCOOT (a 1.5 generation control approach) and a video traffic detection system (VTDS). The portion of this study relevant to the current research was the assessment of the VTDS under different operating conditions at signalized intersections. At the outset of the study, it was anticipated that the VTDS, manufactured by Odetics Inc. (now Iteris), would be capable of providing presence detection for signal actuation, as well as traffic data such as count, speed, volume, and density. As the study progressed, the traffic data requirement was lowered, and the detector was assessed only for its presence detection ability. The study found that 65% of vehicles were accurately detected individually, while 81% were adequately detected for proper signal actuation. Further analysis revealed the effects of various test conditions, as outlined in table 2.3. The results of this study indicate that the performance of this early generation VIP was greatly affected by inclement environmental conditions.

Table 2.3 VTDS Detection Results (14)

Test Condition	Correct Detection
Clear, Overhead Sun, LOS A-B	81.3%
Clear, Overhead Sun, LOS C-D	82.4%
Clear, Transverse Sun, LOS B-E	74.9%
Clear, Into Sun, LOS B-E	85.2%
Clear, Low Light, LOS B-E	45.4%
Clear, Night, LOS B-E	55.9%
Rain, Day, LOS B-E	48.8%
Rain, Night, LOS B-E	61.0%
Clear, Overhead Sun, LOS B-E, Wind Vibration	61.1%
Clear, Overhead Sun, LOS B-E, EM Noise	83.4%
Clear, Overhead Sun, LOS B-E, Overhead Wires in View	43.1%
Clear, Overhead Sun, LOS A-B, Color Camera	84.6%

A study conducted in 2005 assessed the accuracy of a remote traffic microwave sensor (RTMS) along a California freeway (3). The researcher responsible for the study compared the RTMS output to the output of adjacent loop pairs in order to calculate lane-by-lane RMSE (root mean-square error) bias and MAPE (mean absolute percent error) for flow, occupancy, and speed at 30-second and 5-minute aggregation levels. Data was collected for the five eastbound lanes of a divided highway with a median barrier. The RTMS was installed in a side-fire configuration on the south side of the freeway, near the eastbound lanes. Results indicated that the RTMS overestimated flow and occupancy, underestimated velocity in lanes near the median, underestimated occupancy in lanes near the shoulder, and overestimated velocity in lanes near the shoulder. The MAPE values also demonstrated that a more aggregate sampling interval generally produced a smaller percent error than did a more disaggregate sampling interval. This study noted that excessive over-counting in the lane nearest to the median could be explained by "echoes off the concrete barrier" (3). The findings of this report also revealed extreme occupancy error in the lane nearest the detector. This appears to indicate that the detector provided

the best detection for lanes in the middle of the detection area, while having greater error rates in the nearest and farthest detection zones.

Subsequent analyses examined loop detectors and RTMS accuracy at the disaggregate, per-vehicle level, based on the same method of data collection utilized in the previous study (15). Results indicated that, across four lanes of traffic, for analysis periods including both free flow and congested traffic conditions, the count accuracy of the RTMS detector was characterized by 4.8% missed vehicles and 5.6% false detections. These two types of count errors nearly offset one another, resulting in strong count accuracy. This study also reported that the RTMS detection on-time varied lane to lane, creating a lane bias for occupancy. The larger detection zone of the RTMS led to higher occupancy measurements in comparison to the loop detectors.

The most recent research completed under California PATH relating to non-intrusive detector assessment involved efforts to develop an automated system for collecting ground truth data (16, 17). Traditionally, ground truth data for detector assessment has been collected manually via human analysis of recorded video. However, as Caltrans developed a detector test bed on Route 405 near Irvine, California, it was determined that it would be valuable to develop an automated ground truth system, which, unlike the manual collection process, would be capable of assessing large data sets. The resulting automated system was the Video Vehicle Detector Verification System (V2DVS). This system consisted of six downward-pointing video cameras (one over each lane) mounted on an overpass, each camera having a dedicated field computer that conducts video image processing, as well as a central server on which data are recorded. Under various lighting conditions, the cameras provide detection rates with

accuracies ranging between 98.3% and 99.7%, and correct velocity calculation for 96.5%-99.7% of vehicles (16). Initial testing of alternative detection technologies at this site found that missed detections were most commonly due to ambiguous vehicle lane position, non-ideal image processing conditions (shadow or occlusion) for VIPs, or reflection and occlusion problems in distant lanes for crossfire detectors. It was also concluded that frequent false detection could typically be reduced by additional calibration.

2.4.2 Detection Technology for IVHS Study

Further analysis of various traffic detection technologies was conducted under the FHWA sponsored Detector Technology for IVHS (Intelligent Vehicle-Highway Systems) study. The objectives of this program were to determine traffic parameters to be measured for IVHS applications and associated accuracy specifications; to perform laboratory and field tests of available technologies for the determination of their ability to measure these traffic parameters with acceptable accuracy; and to determine the feasibility of establishing a permanent vehicle detector test bed (18). The required accuracies for freeway data were found for two potential IVHS applications (i.e., incident management and ramp metering). The accuracy of various parameters was further divided by data aggregation intervals into tactical, strategic, and historic parameters.

Tactical data is used in applications that require data immediately at relatively short aggregation intervals (e.g., 20 seconds). Strategic traffic parameters have a greater aggregation interval (e.g., 5 minutes), thereby diminishing the noise in the data that results from the randomness of vehicle arrivals and driver behavior. Lastly, historic data is used to maintain databases and for future planning purposes. It is generally collected at

a greater aggregation interval (e.g., 15 minutes or 1 hour). Table 2.4 shows parameter specifications for freeway incident management, while table 2.5 shows parameter specifications for freeway ramp metering.

Table 2.4 Freeway Incident Detection and Management Traffic Parameter Specifications (18)

Tactical Parameters (Detection)

Parameter	Units	Range	Collection Interval	Allowable Error
Mainline Flow Rate	veh/h/lane	0-2500	20 s	± 2.5% *
Mainline Occupancy	% (by lane)	0-100	20 s	± 1%
Mainline Speed	mph (by lane)	0-80	20 s	± 1 mph
Mainline Travel Time	min		20 s	± 5%

Strategic Parameters (Incident Management)

Parameter	Units	Range	Collection Interval	Allowable Error
Mainline Flow Rate	veh/h/lane	0-2500	5 min	± 2.5% *
Mainline Occupancy	%	0-100	5 min	± 2%
Mainline Speed	mph	0-80	5 min	± 1 mph
On-Ramp Flow Rate	veh/h/lane	0-1800	5 min	± 2.5% *
Off-Ramp Flow Rate	veh/h/lane	0-1800	5 min	± 2.5% *
Link Travel Time	seconds		5 min	± 5%
Current O-D Patterns	veh/h		5 min	± 5%

Historic Parameters (Planning)

Parameter	Units	Range	Collection Interval	Allowable Error
Mainline Flow Rate	veh/h/lane	0-2500	15 min or 1 hour	± 2.5% *
Mainline Occupancy	%	0-100	15 min or 1 hour	± 2%
Mainline Speed	mph	0-80	15 min or 1 hour	± 1 mph
On-Ramp Flow Rate	veh/h/lane	0-1800	15 min or 1 hour	± 2.5% *
Off-Ramp Flow Rate	veh/h/lane	0-1800	15 min or 1 hour	± 2.5% *
Link Travel Time	seconds		15 min or 1 hour	± 5%
Current O-D Patterns	veh/h		15 min or 1 hour	± 5%

* @ 500 veh/h/lane

Table 2.5 Freeway Metering Control Traffic Parameter Specifications (18)

Tactical Parameters (Local Responsive Control)

Parameter	Units	Range	Collection Interval	Allowable Error
Ramp Demand	Yes/No		0.1 s	0% (No misses)
Ramp Passage	Yes/No		0.1 s	0% (No misses)
Ramp Queue Length	vehicles	0-40	20 s	± 1 vehicle
Mainline Flow Rate	veh/h/lane	0-2500	20 s	± 2.5% *
Mainline Occupancy	%	0-100	20 s	± 2%
Mainline Speed	mph	0-80	20 s	± 5 mph

Strategic Parameters (Central Control)

Parameter	Units	Range	Collection Interval	Allowable Error
Mainline Flow Rate	veh/h/lane	0-2500	5 min	± 2.5% *
Mainline Occupancy	%	0-100	5 min	± 2%
Mainline Speed	mph	0-80	5 min	± 5 mph

Historic Parameters (Pretimed Operation)

Parameter	Units	Range	Collection Interval	Allowable Error
Mainline Flow Rate	veh/h/lane	0-2500	15 min or 1 hour	± 2.5% *
Mainline Occupancy	%	0-100	15 min or 1 hour	± 2%
Mainline Speed	mph	0-80	15 min or 1 hour	± 5 mph
On-Ramp Flow Rate	veh/h/lane	0-1800	15 min or 1 hour	± 2.5% *
Off-Ramp Flow Rate	veh/h/lane	0-1800	15 min or 1 hour	± 2.5% *

* @ 500 veh/h/lane

The aforementioned study selected 19 detectors (three ultrasonic, one active IR, two passive IR, five microwave radar, five VIP, one acoustic, one inductive loop, and one magnetometer) for potential evaluation with laboratory and field testing. The laboratory testing focused on operating parameters such as power consumption, operating frequency, minimum detectable signal, and detection zone size. While valuable in their own right, these laboratory test results are not directly relevant to the comparison of accuracies of various detector technologies in the field.

The field test quantified performance of detectors as it related to their measured values of flow rate, speed, and density (or occupancy, as is commonly detected as a proxy for density). Intersection and freeway field testing sites were selected in Minnesota,

Florida, and Arizona in order to include a wide variety of environmental conditions. The evaluated detectors included three ultrasonic detectors, five microwave detectors, four infrared detectors (including active and passive infrared detectors), five video image processing detectors, one magnetometer, one microloop, and one pneumatic tube detector. Manual observation of video recordings of the traffic scene was used to establish the ground truth against which the detector technologies were compared. Speed ground truth was determined through the use of a probe vehicle, with the driver recording his speedometer reading at the detector location. These field test results were evaluated to determine the best technologies for the following applications, with the following results: the best- performing non-intrusive technologies for collecting both low and high volume count data were microwave radar and video image processors; the best-performing non-intrusive technologies for low and high volume speed data were microwave radar detectors for per-vehicle results. Other technologies, such as video image processors, enter the scene when average speed data over some aggregation interval is needed. Microwave detectors were also found to be the most unaffected by inclement weather. The technologies with the most noticeable inclement weather limitations were ultrasonic, infrared, acoustic, and VIP.

Based solely on count accuracy, it was found that the inductive loop detectors provided the most accurate data, with an error rate below 1% (19). These were followed by the overhead RTMS-X1 microwave radar and one lane of the Autoscope 2003 VIP outputs, with 1-2% error rates, which were in turn proceeded by the following detectors, having 3-7% error rates: Whelen TDN-30 microwave radar; the other lane of Autoscope 2003 VIP; Microwave Sensors TC-30C ultrasonic; Sumitomo SDU-300 ultrasonic;

Midian Electronics SPVD magnetometer; side-fire EIS RTMS-X1 microwave radar; and Eltec 833 passive IR. The detectors with the least accurate counts in this study were the Eltec 842 passive IR, AT&T SmartSonic passive acoustic, and Microwave Sensors TC-26 microwave radar.

The primary author of these studies, Lawrence Klein, went on to publish a book entitled *Sensor Technologies and Data Requirements for ITS* (6). In it, he draws on his experience from the above studies, as well as the findings of previous studies, in order to provide an overview of various detector technologies available for ITS. The book also addresses the application of sensor data to various ITS strategies and the data processing necessary for these applications. It provides a broad overview of traffic data in ITS, ranging from data acquisition by sensors and communications protocols to data processing, fusion, and archival at a traffic management center (TMC).

Klein has been involved in two other seminal studies relating to traffic detection. The first of these was the *Traffic Detector Handbook*, published in its third edition in 2006 (5). This document was intended as a primer on intersection and freeway traffic detection for the practicing traffic engineer. It addresses the operational mechanics of the various detector technologies, detector applications, in-roadway detector design, detector installation, and detector maintenance. The second (2007) study compiled manufacturer and model information for over 50 commercially available traffic detector models (20). This study also provided brief descriptions of the functionality of each type of traffic detection technology.

2.4.3 Minnesota Guidestar Studies

Since 1997, a series of studies has been conducted under the Minnesota Guidestar program to assess state-of-the-art non-intrusive traffic detectors. In the first phase of this study, 17 different traffic detectors were analyzed at both freeway and signalized intersection locations (21). The primary sources of ground truth data were loop detectors embedded in the roadway with select 15-minute periods, rather than manual observation from recorded video. While confidence in the results may be limited due to the loop detector ground truth data method, this form of ground truth is less labor intensive than manual observation, and allows for larger data sets to be efficiently processed. A subsample with 15-minute manual observation ground truth reveals similar error rates to the error rates with loop detectors as the ground truth thereby increasing confidence in the results from the larger data sets where loop detectors served as the ground truth.

The (1997) study also included a section on the influence of weather on the various detectors, though the results presented were qualitative in nature. Though the results involved the impact upon a given detector technology by a given weather condition, the study lacked a statistical analysis of the significance of these effects. Graphs showed apparent correlations between error rates and precipitation rates or other environmental phenomenon, but were utilized only for a qualitative visual assessment. The value of weather-based assessment is to offer potential explanations for errors based on environmental conditions. One example is an assessment of an active infrared device which states, “Overcounting was also observed during periods of heavy snowfall when snow in the air may have been detected by the device” (21). Table 2.6 shows the 17 devices evaluated in the initial study and their reactivity to environmental factors. Of

particular interest in this table is the fact that the video and radar technologies appeared to perform well in all weather conditions tested, with the exception of leakage in the housing of the radar unit, which caused electrical problems following the weather event. This can be viewed as a minor problem which should not be counted against the potential utility of this technology. Finally, the magnetic detectors appeared to demonstrate poorer performance in rain and low temperature conditions.

Table 2.6 Environmental Factors Affecting Device Performance (22)

	Freeway						Intersection				Both Test Sites				
Technology	High Speeds	Low Speeds	High Volumes	Low Volumes	Geometrics	Lighting Effects	High Volumes	Low Volumes	Geometrics	Lighting Effects	Rain	Freezing Rain	Snow (1)	High Temperature	Low Temperature
Device															
Inductive Loop	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Passive Infrared															
Eltec Model 833	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-	+/-
ASIM IR 224 (2)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Active Infrared															
Autosense I	+	+	+	+	+	+	?	?	?	?	-	-	-	+	+
Magnetic															
IVHS 232E (2)	+	+	+	+	+	+	?	?	?	?	-	+	+	+	-
Radar															
RTMS X1	+	+	+	+	+	+	?	?	?	?	-*	-*	+	+	+
Doppler Microwave															
PODD	+	+	+/-	+	+	+	-	-	-	-	+	+	+	+	+
TDN-30	+	+	+	+	+	+	-	-	-	-	+	+	+	+	+
Pulse Ultrasonic															
Lane King	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
TC-30	+	+	+	+	+	+	+/-	+/-	+/-	+/-	+	+	+	+	+
Passive Acoustic															
SmartSonic (2)	+/-	+	+/-	+	+	+	+/-	+/-	+/-	+/-	+	+	+	+	-
Video															
EVA 2000s	+	+	+	+	+	+	?	?	?	?	+	+	+	+	+
Autoscope 2004	+	+	+	+	+	-	+/-	+	+	-	+	+	+	+	+
TraffiCam - S	+	+	+	+	+	?	?	?	?	?	+	+	+	+	+
Video Trak-900	+	+	+	+	+	-	?	?	?	?	+	+	+	+	+

(1) Snow is evaluated here as a direct factor in affecting device performance, secondary factors such as vehicle tracking patterns are not included.

(2) Two detectors of this model were analyzed.

* The RTMS unit was observed to miscount following periods of rain and freezing rain due to water entering the housing.

\+ Denotes a device which performs satisfactorily in the stated condition.

+/- Denotes a device which meets some but not all of the criteria for satisfactory performance.

\- Denotes a device which does not perform satisfactorily in the stated condition.

? Denotes a situation that could not be confirmed.

Phase 2 of the Minnesota Guidestar non-intrusive detector evaluation study was published five years later, in 2002 (23). The methodology of this study was modeled after that of the first phase, but placed greater emphasis upon assessment in freeway traffic detection. The nine detector models evaluated in this phase differed from those of the previous phase, though some were simply newer-generation models of the same technology, from the same manufacturer. A summary of detector performance, similar to that given for phase 1 of the same study, is provided in table 2.7. Due to the study schedule coinciding with a mild winter, weather impacts were not assessed in this phase.

Table 2.7 Summary of Sensor Performance (23)

Sensor Model	Technology	Freeway Test Site			Ease of Installation	Ease of Calibration	Reliability
		Speed Performance	Volume Performance				
			Peak	Off Peak			
Autosense II	Active Infrared	+	+	+	+/-	+	+
3M Canoga	Magnetic	+	+	+	-	+/-	+
ECM Loren (1)	Microwave				+/-	-	-
SmarTek	Passive Acoustic	+	+/-	+	+	+	+
ASIM IR 254 (2)	Passive IR (PIR)	+/-	+/-	+	+	+/-	+
ASIM DT 272 (3)	PIR/Ultrasonic	N/A	+/-	+	+	+	+/-
ASIM TT 262	PIR/Ultrasonic/ Radar	+	+	+	+	+	+/-
Autoscope Solo	Video	+	+	+	+	+/-	+
Traficon VIP D	Video	+	+	+	+	+/-	+

(1) The EMC Loren did not function in the test. No data available.
(2) ASIM IR 254 was difficult to calibrate for side-fire installation because of alignment complications.
(3) Data collection problem presented difficulty in fully evaluating the ASIM DT 272.
+ Denotes a device which performs satisfactorily in the stated condition.
+/- Denotes a device which meets some but not all of the criteria for satisfactory performance in the stated condition.
- Denotes a device which does not perform satisfactorily in the stated condition.

The next phase of the study concentrated on the design and assessment of a portable, non-intrusive traffic detection system (PNITDS) (24). A successful PNITDS should be able to be installed and calibrated quickly, easily, and safely without disrupting traffic flow, in order to facilitate short-term traffic studies. There were three different system concepts presented in the paper under review. A pole-mounted system was tested, which allowed a non-intrusive detector to be mounted to any roadside signpost or lamppost. This system was tested with three different detectors: a Wavetronix SmartSensor, a RTMS X3, and a SmarTek SAS-1. The second system was trailer-mounted PNITDS which consisted of a Wavetronix SmartSensor mounted on a retractable mast arm on a heavy-duty trailer designed as a platform for a mobile dynamic message sign. The third system was relatively new to the market (i.e., The Infra-Red Traffic Logger [TIRTL], an axle-based vehicle classifier, developed in Australia).

In the analysis of the various detectors installed with the pole-mounted system at an eight-lane freeway test site, the following results were found over 24-hour test periods (24): The Wavetronix SmartSensor had a per-lane volume detection error ranging from 1.4%-4.9% and speed detection error between 3.0% and 9.7%. It also provided reasonable length-based classification when properly calibrated. The RTMS X3 had volume detection errors ranging between 2.4% and 8.6% and speed detection errors ranging between 4.4% and 9.0%. This detector also provided reasonable length-based classification when properly calibrated. Finally, the SmartTek SAS-1, which was mounted in a non-optimal location, had volume errors ranging between 9.9% and 11.8% (performing particularly poorly in congested traffic conditions) and speed detection errors ranging between 5.6% and 6.8%. When properly calibrated, this detector provided

accurate percent-passenger-vehicle estimates, but poor accuracy in estimates of percent-medium and percent-large vehicles.

The most recent phase of the Minnesota Guidestar study returned to the detector test bed used in the first two phases in order to assess newer detector technologies in a long-term installation scenario (8). In this phase of the study, the following five technologies were tested: Wavetronix SmartSensor HD, GTT Canoga Microloops, PEEK AxleLight, TIRTL, and Miovision.

The analysis of the SmartSensor HD found that the volume absolute percent error was 1.6% and the absolute percent error for speed was 1.0% at an average speed of 60.9 mph. The classification percent error was 3.0% incorrectly classified vehicles, based on a length-based, three-class system. The test period for the SmartSensor HD included extreme cold, rain, snow, and fog conditions, with fog being the only condition to noticeably affect performance. The volume error remained below 5%, even in foggy conditions.

The analysis of the Canoga Microloops found that the volume absolute percent error was 2.5%, and the absolute percent error for speed was 0.6% at an average speed of 60.9 mph. The classification percent error was 2.9% incorrectly classified vehicles, based on a length-based, three -class system. The only potential weather effect noted in the study was snow on the roadway, which might have caused drivers to maintain poor lane position, potentially affecting the accuracy of volume data.

The analysis of the AxleLight found that vehicles were initially undercounted by 9.1%. As the AxleLight is an axle-based detector, it was found that this error was due to two cars with a small spacing (tailgating) being classified as a multiple unit truck. After

further calibration, the undercounting was 5.4%. The study found that speed was consistently underreported by the AxleLight, but claimed that this could be addressed by recalibration, as a simple speed trap configuration is used by this detector. While not analyzed during the study, the manufacturer recommended that the AxleLight not be used in heavy rain conditions, as significant amounts of water kicked up by wheels could decrease accuracy.

The analysis of TIRTL found that it generally reported volume with a 2% overcount, but a few outliers with greater error could not be explained. The absolute average percent error in reported speed was found to be 2%, or 1.2 mph, at an average speed of 58 mph. Testing in rainy conditions revealed that at the test site, rain did not affect the performance of TIRTL. However, the study reported that locations with poor drainage, wheel path rutting, ponding, or extremely heavy rain could produce wheel spray capable of degrading performance.

This phase of the research concluded with an analysis of the Miovision system, a non-traditional approach to video image processing. At the freeway test site, the Miovision collected volume data within the accuracy of the baseline (2%). Speed data was not analyzed. However, turning movement counts were conducted at two different intersections. These movement counts were very accurate, each movement volume having an error of less than 0.5% for the two-hour test period.

All four of the detector studies conducted under the Minnesota Guidestar program were well-executed, and prove to be invaluable reference works. In addition to scientific analyses of detector performance, the experiences of the research team with installation, calibration, maintenance, and cost were well-documented.

2.4.4 Texas Transportation Institute Studies

In recent years, the Texas Transportation Institute (TTI) has also conducted research related to non-intrusive traffic detectors and their data. In 2000, a TTI report focused specifically on freeway application of the following three detectors: PEEK Videotrak 900 VIP, 3M Microloop magnetic, and SmarTek SAS-1 acoustic (25). In this study, count and speed detection accuracy were only part of the evaluation criteria. The other factors assessed were the ease with which the different systems were set up and configured, and installation cost. While the study did not set out to evaluate the effects of environmental conditions on performance, a rainstorm on one of the eight days of data collection introduced a discussion of the impact this rain had on detection accuracy. It appeared that the rain negatively affected the performance of both the video and acoustic detectors, but there was no statistical analysis of the significance of these effects beyond demonstration that the error rates were greater during wet weather. The error rates of the detectors under evaluation were not presented as straightforward mean percent errors or mean absolute percent errors. The study reported the percent of intervals in which the error was 0-5%, 5-10%, or greater than 10%. For results of the study, refer to the source (25).

A subsequent report, published in 2002, highlighted the experiences of Texas and various other states with loop detectors and non-intrusive detectors (26). This study also analyzed the performance of five detector models for freeway data collection. First, the Peek ADR-6000 was assessed for its classification, count, and speed accuracy, in order to determine its viability as a baseline against which non-intrusive detectors could be tested. This system was found to have a classification accuracy of 98.9%, count accuracy greater than 99.9%, and speed accuracy within +/- 1 mph of a speed gun for 95.0% of vehicles.

The Peek ADR-6000 was determined to be an adequate baseline for the testing of the four non-intrusive detectors.

The non-intrusive detectors were assessed based on per-lane five-minute counts and average speed, and 15-minute occupancy (26). The Autoscope Solo Pro was found to undercount by up to 5% in free flow conditions, by 10-25% in congested conditions in lane one, and by 0-10% in all other lanes in free flow and congested conditions. The Solo Pro speed was found to be within 3 mph of the baseline for lane one, 2 mph for lanes two and three and 5 mph for lane four. Of the three detectors tested for occupancy, the Solo Pro was found to have the greatest agreement with loop occupancy, within 1% of loop occupancy for most intervals. The Iteris Vantage was found to have less count bias than the Solo Pro, but had the greatest standard deviation of count accuracy, undercounting by as much as 22% in lane one and overcounting by as much as 10% in lanes one and two. The speeds reported by the Vantage were found to generally be within 5 mph for all lanes, with the exception of lane two, which occasionally reported speeds 15 mph greater than the baseline. The Vantage was found to report occupancy within 6% of loop occupancy for most intervals. The EIS RTMS was found to provide counts generally within 10% of loop counts for lane one and within 5% of loop counts for lanes two, three, and four. The RTMS speeds in lane three were found to be within 5 mph of baseline speeds, except for intervals where the average speed dropped below 50 mph, in which case speeds were up to 10 mph above the baseline. Lane four consistently overestimated speeds by 2-5%. Lane one speeds differed from baseline speeds by up to 15% in congested conditions. Occupancy tests were not performed on the RTMS. The SmarTek SAS-1 was the final detector analyzed. Lane one counts were found to be up to 32%

below baseline during congested conditions. Other lanes were found to overcount by as much as 6% and undercount by as much as 18%. The SAS-1 was found to overestimate speeds in lane one during congested conditions by as much as 25 mph, but was within 5 mph during free flow speeds. Lanes two, three, and four were generally within 5 mph of the baseline. The occupancy reported by the SAS-1 was generally found to be within 4% of the baseline.

In 2007, TTI selected an urban freeway site and developed a detector test bed for the Arizona Department of Transportation (ADOT), recommending four state-of-the-art detectors to be analyzed in the first round of tests at the new test bed (27). While the report did not present the results of detector analyses, it addressed many key considerations in the design process of a detector evaluation program. The report recommended that the detectors be analyzed in the conditions under which they are expected to perform, which may include some or all of the following: "a.m. peak" period, "p.m. peak" period, off-peak, dry weather, wet weather, congested conditions with slow speeds, free-flow conditions, intense fog, blowing dust, full sunlight, full dark, light transitions (dawn and dusk), or snow/ice conditions. The report recommended the following as potential statistical measures of data accuracy: mean absolute error, mean absolute percent error (MAPE), mean percent error, and root mean squared error (RMSE). It recommended the use of a Peek ADR 6000 system for a baseline against which other detectors would be tested, based on the confidence TTI had gained in that particular product during a previous study (26). A search for a subsequent report from ADOT that included information on the implementation of the TTI test bed design or results of detector testing at such a site did not return any results.

2.4.5 Purdue University Studies

In recent years, researchers at Purdue University have conducted a number of studies for the Indiana Department of Transportation (INDOT) relating to traffic detection, most being focused on video detectors. The first of these studies evaluated the performance of two VIP systems at signalized intersection, in comparison to loop detectors (28). The two systems evaluated were the Econolite Autoscope and Peek VideoTrak-905. As stated earlier, performance metrics at an intersection do not necessarily imply similar performance for freeway installations, but data trends are worth acknowledging. For example, this study noted that at night, vehicle headlights extended far enough ahead of vehicles to prevent gap out, whereas it would have occurred during daylight conditions. It was also determined that at night it was possible for a vehicle to pull too far forward at the stop bar so that headlights were out of the detection area and the dark vehicle was not detected in the detection area. It is possible that additional illumination at the intersection could reduce the effect of both issues. Based on the findings of this report, INDOT suspended the deployment of VIP detectors at signalized intersections. As this relates to freeway installations of video detectors, it could imply a potential for errant vehicle length and classification information at night if headlights are detected instead of vehicles.

Another report by Purdue researchers examined methods of identifying errors in ITS data from freeway detectors when the data are recorded and archived (29). While most detectors are evaluated immediately after installation, there is generally a lack of data quality control performed throughout the life of the detector, during which time data quality could deteriorate. The authors proposed a set of automatic tests that could be run

periodically to ensure data quality. The first test addressed flow continuity, comparing five-minute, all-lane vehicle counts for two closely spaced freeway detectors with no ingress or egress between the two detectors. Significant departures indicated erroneous data from at least one of the detectors. The second test addressed speed continuity, comparing five-minute per-lane average speeds as reported by two closely-spaced detectors with no ingress or egress between them. Any significant departure or consistent offset in values indicated erroneous data from at least one of the detectors. The third test addressed data availability, using statistical modeling based on the expected traffic volume to estimate the number of set-duration time periods (i.e., 30-sec, 1-min., 5-min., etc.) in a day, during which it could be expected that there would be zero volume. If the actual number of zero volume intervals was significantly different, it was possible that the detector was malfunctioning. Finally, the fourth test addressed average effective vehicle length (AEVL), assessing the relationships between reported volume, speed, and occupancy to determine whether these relationships were practically feasible. Values outside of a preset range of expectations indicated erroneous data. The tests were demonstrated on data from RTMS radar and Canoga microloop detectors along the Borman expressway (I-80/94). It was proposed that the tests be automated on INDOT traffic data archives to help maintain freeway sensor data quality.

The next three relevant reports by Purdue researchers all focused on the assessment of VIP detector accuracy at signalized intersections. The first of these studies assessed the stop bar detection performance of Autoscope Solo Pro VIP detectors at different mounting locations, as compared to loop detectors at a high speed intersection (30). The mounting locations were 40 feet above the pavement, 165 feet downstream of

the stop bar, and 60, 48, or 36 feet from the mast arm standard, with 60 feet being the optimal location, aligned with the lane marking between the left turn lane and leftmost through lane. It was concluded that, even with optimal camera location, the VIP still had statistically significantly more missed and false calls than the stop bar loop detectors. The difference in performance at the three mounting locations was minimal.

The second of these three signalized intersection VIP studies was published in 2006. The study compared the performance of the following three detector models: Autoscope Solo Pro, Peek UniTrak, and Iteris Vantage (31). All three VIP systems were found to have many more false calls and missed calls than the traditional loop detectors. Depending on when in a signal cycle a false or missed call occurs, it can have either safety or efficiency implications. As a result, it was determined that the INDOT moratorium on VIP detectors at signalized intersections, in place since 2001, was still justified.

The next VIP study focused specifically on the question of detection zone activation and deactivation during daytime and nighttime conditions (32). This study addressed a specific issue with video detection at night, that is, when the reflection of headlights on the pavement ahead of the vehicle is detected instead of (or in addition to), the vehicle itself. The analysis found that 15 of the 16 camera mounting locations at the intersection had a statistically significant difference in activation residual between daytime and nighttime conditions. This is to say that, at night, presence detection was activated significantly earlier than during the day. The deactivation times were found to differ significantly between daytime and nighttime for 9 of the 16 cameras, but the average difference in deactivation time was much smaller than the average difference in

activation time. These findings supported the hypothesis that headlight reflection on pavement causes early detector activation. While this paper focused on activation and deactivation of presence sensors at a signalized intersection, this type of error could have potential implications for occupancy and length-based classification in freeway detection scenarios.

In 2008, another report was published on freeway detector monitoring for data verification (33). This report further developed the concept of Average Effective Vehicle Length (AEVL), detailed in an earlier report (29), and presented a user interface through which detector reliability could be monitored. The AEVL is used as a monitoring metric because it combines the effects of volume, occupancy, and speed into a single variable. Once a range of reasonable values is determined, it is possible to automate analysis of detector data for intervals during which the AEVL lies outside of the acceptable range. The remainder of the report focused on the design of a user interface which would allow traffic management center (TMC) personnel to easily monitor the health of numerous detectors in the TMC coverage region. The essence of this user interface was a geographic information system (GIS) map, which classified the AEVL from each detector in the database as acceptable or unacceptable and created either a green or red marker at the physical location of each detector, based on that detector's AEVL. By clicking a marker, the user was directed to that detector's data in the database. This allowed the user to determine whether the detector required maintenance.

2.4.6 University of Nebraska Studies

A previous study conducted by researchers at the University of Nebraska-Lincoln evaluated the performance of three non-intrusive detectors for freeway installation (34).

The three detector models evaluated were the EIS RTMS microwave radar detector, Wavetronix SmartSensor microwave radar detector, and Autoscope RackVision VIP detector. The analysis considered various data aggregation levels by addressing per-vehicle data, 1-minute interval data, and 15-minute interval data. The primary focus was on volume, but speed and classification were also addressed. The study found that the 15-minute interval mean percent volume errors for the RTMS, SmartSensor, and RackVision were -1.4%, 1.4%, and 0.7%. The 15-minute mean absolute percent volume errors for the RTMS, SmartSensor, and RackVision were 3.6%, 3.2%, and 1.8%. These results indicate that each of the above detectors was capable of providing reasonably accurate historical volume data. Analysis of rainy and clear weather data indicated that there was no significant difference in the performance of any of these detectors based on weather. Analysis of light and heavy traffic indicated that the SmartSensor was most affected by traffic, having a 15-minute mean percent volume error of 1.5% in normal traffic and -0.5% in heavy traffic. Analysis of lighting conditions indicated that the RackVision was minimally impacted by lighting, with a mean percent volume error of 0.8% in daylight and -0.8% in dark conditions. 15-minute average speed analysis was included, but appears to be primarily an indication of calibration accuracy, rather than detector capability, since no ground truth data was provided. Analysis of length-based classification was performed on the SmartSensor and RackVision. The results indicated that the RackVision classified more vehicles in the small class (0-23 feet long) while the SmartSensor classified more vehicles in the medium class (24-45 feet long).

Manual counts were not conducted at the 1-minute interval; therefore, error rates were not reported for this less-aggregated level. Instead the detectors were compared to

one another to reveal relative differences. For 1-minute mean volume, it was determined that there was not a significant statistical difference between values reported by different detectors. A speed analysis was performed on a small sample of 20 minutes, using data from a Lidar gun to serve as ground truth. The results of this analysis showed that, as configured, the RTMS provided the most accurate speed data across all lanes. The difference between RackVision speeds and Lidar speeds was consistent across lanes. This indicates that a single calibration factor for the RackVision could have significantly improved speed performance. The differences between SmartSensor speeds and Lidar speeds were more erratic across lanes, indicating that each lane would require a unique calibration factor to improve performance. Per-vehicle, length-based classification results were given for the SmartSensor and RackVision, but not for the RTMS. The SmartSensor classified 79%, 16%, and 5% of the traffic as small, medium, and large vehicles, respectively, while the RackVision classified 91%, 6%, and 3% in the same categories. While no ground truth data was given, these results indicate that the large vehicles were approximately consistent, while the SmartSensor classified some of the vehicles as medium that the RackVision classified as small. These results were consistent with the 15-minute results presented above.

Another paper from the University of Nebraska was recently presented which outlined the research plan and some preliminary results of the study completed for this thesis (35). This paper expressed the need for a side-by-side comparison of detector technologies in order to eliminate any bias due to each detection technology being subjected to a unique set of environmental and traffic conditions. In a side-by-side comparison, all detectors are analyzed under the same set of operating conditions. The

statistics of mean absolute percent difference (MAPD) and mean percent difference (MPD) were proposed to compare the results of pairs of detectors, as a ground truth source had not yet been established. The detectors compared in the study were the Wavetronix SmartSensor, ISS RTMS G4, and Autoscope Solo Pro II. Based on 119 one-minute samples, it was determined that the Autoscope reported volumes 9% and 7% greater than the SmartSensor and RTMS G4, respectively. As a proxy for length-based classification, percent passenger vehicles (vehicles less than 21 feet long) was reported for each detector. This comparison found that the Autoscope reported percent passenger vehicles 37% and 26% higher than the SmartSensor and RTMS G4, respectively. This preliminary study also analyzed six probe vehicle speed runs (with GPS ground truth speeds) finding that the mean percent errors (MPE) in speed were 4%, -3%, and 14% for the SmartSensor, RTMS G4, and Autoscope.

2.4.7 Illinois Center for Transportation Studies

The Illinois Center for Transportation recently completed a study further examining sources of error in VIP detection at intersections. For this study, the following three VIP detectors were mounted side-by-side: Autoscope Solo Pro, Peek Unitrak, and Iteris Edge 2. The first volume of this study addressed the impacts of configuration changes on VIP performance (36). The stop bar and advance detection zones were analyzed for false, missed, stuck-on, and dropped calls in day and night conditions after preliminary configuration. The results were presented to the VIP manufacturer representatives, who made configuration changes before a second round of analysis was performed. The report presented extensive quantified changes in each type of detection error. The general trend was that after recalibration, the missed and dropped calls were decreased, but at the cost

of increased false and stuck-on calls. Thus, it was concluded that when recalibrating a VIP detector to diminish a specific type of error, it is important to be cognizant of the effect that the recalibration has on overall VIP performance.

The next volume of this study analyzed lighting effects on VIP performance (37). The various lighting conditions for which data were collected were dawn, sunny morning, cloudy noon, dusk, and night. In cloudy noon (ideal) conditions, false calls were the only concern, with tall vehicles triggering calls in the lane adjacent to their travelled lane in addition to a call in their travelled lane. At the stop bar, the false calls in lanes one and two were less than 3% for each VIP, but were up to 20% for lane three. False calls in lane three were also problematic for advance detection zones. Missed, dropped, and stuck-on calls were nearly non-existent in cloudy noon conditions. Dawn conditions increased false calls for the Autoscope and Peek detectors (due to headlight spillover), while increasing missed calls for the Iteris detector. Sunny morning conditions increased false calls for all detectors (due to shadow spillover), and stuck-on calls were increased for Autoscope and Peek detectors. Dusk conditions increased false calls for all detectors and increased missed calls in lane one for the Peek detector. Night conditions increased false calls (due to headlight spillover) for Autoscope in lanes one and two and Peek in lane two, while decreasing false calls for Peek in lane three. Missed calls increased for Peek in lane one at night. This portion of the study was valuable, primarily for its qualitative explanations for detection errors such as headlight and shadow spillover and tall vehicle occlusion.

The third volume of this study addressed the effects of windy conditions on VIP detector performance (38). While windy condition performance is determined primarily

by the rigidity of the structure on which the camera is mounted, this portion of the study provided information on the relative sensitivity of the different VIP detectors to camera movement. It is important to note that all three cameras were mounted side-by-side on a luminaire arm at an approximate height of 40 feet above the roadway. The researchers observed that VIP reaction to wind was greatly dependent on lighting conditions. They found that under cloudy noon lighting, wind effects were minimal. Under sunny morning lighting (when long shadows were present), there was a significant increase in false calls for all detector models, while advance zone missed calls increased for the Peek detector, and decreased for the Iteris and Autoscope detectors. Under nighttime lighting, false calls significantly increased for all three detector models at both stop bar and advance zones.

The final volume of this study analyzed the effects of adverse weather conditions on VIP detector performance (39). The conditions for which data were collected were rain and snow under both day and night lighting, and light and dense fog under daytime lighting. Results indicated that daytime light fog conditions moderately increased false calls for Autoscope and Iteris detectors. During daytime dense fog, Iteris and Autoscope registered image contrast loss and went into permanent call modes, while missed calls were registered for the Peek detector. Both daytime and nighttime snow greatly increased false calls for all three systems. False calls also increased in daytime rain and to a greater extent nighttime rain (purportedly due to headlight spillover from adjacent lanes). Detailed performance analysis for each detector zone is provided in the report.

Another detector evaluation study, performed at the Illinois Center for Transportation, looked at the performance of wireless magnetometers under various weather conditions at intersection and railroad crossing installations (40). The

magnetometers under investigation were manufactured by Sensys Networks. It was found that at the stop bar, false calls made up 5.6% to 7.2% of total calls per lane in favorable weather and 7.7% to 15.4% in winter weather. These were primarily due to a vehicle placing a call in its lane as well as the adjacent lane. At the advance detection zone (approximately 250 feet upstream of the stop bar), missed calls were the most prevalent type of error, ranging from 0.7%-9.7% depending on lane and weather. While these missed calls varied with weather conditions, they were not found to correlate with the weather conditions. The missed calls were primarily attributed to lane change maneuvers. The results at the railroad grade crossing indicated that the detectors were configured in such a way so as to reduce missed and dropped calls at the expense of more frequent false and stuck-on calls.

2.4.8 Other Studies

While most of the relevant traffic detection technology assessment studies have been conducted in series, or by authors who established themselves by conducting ongoing research in the field, there are a few studies worth noting that were conducted as standalone works relating to traffic detection technology. The first of these is *A Comparative Study of Non-Intrusive Traffic Monitoring Sensors* by Gregory Duckworth et al. (41). This study emphasized recognition of the intrinsic limitations of various technologies for traffic detection. While commercially available detectors employed various technologies at the time the study was conducted, the authors developed their own low-cost detectors and signal processing algorithms based on video, Doppler radar, Doppler ultrasound, pulsed ultrasound, passive acoustic, and passive infrared technologies. The basic analysis of each of their detectors is given in table 2.8. The final

conclusion was that the most promising low-cost replacement for an inductive loop detector was a combination detector with pulsed ultrasonic and either pulsed-Doppler ultrasound or Doppler radar.

Table 2.8 Duckworth Tested Sensors and Characteristics (41)

Sensor Type	Sensor Cost	Commun-ications Bandwidth	Processing Load	Detection Performance	Speed Estimation Performance	Vehicle Classification Performance
Video Camera	High ($150-500)	Med-High (10-4500 kbs)	Med-High (10 MOPS)	Good	Very Good	Very Good
Doppler Radar	Medium (<$100)	Medium (2-10 kbs)	Low (0.2 MOPS)	Fair/Good	Excellent	Poor
Doppler Ultrasound	Low (<$75)	Medium (8 kbs)	Medium (0.12 MOPS)	Good	Fair	N/A
Pulsed Ultrasound	Low (<$75)	Very Low (0.32 kbs)	Very Low (0.01 MOPS)	Very Good	N/A	Good
Passive Acoustic	Low (<$25)	Medium (10 kbs)	High (2.2 MOPS)	Poor	Fair	Poor
Passive Infrared	Low (<$30)	Very Low (0.32 kbs)	Very Low (0.001 MOPS)	Very Good	N/A	N/A

Two more relevant studies have been published in recent years. The first of these papers attempted to determine the feasibility of mounting an ultrasonic detector in a side-fire configuration instead of the overhead configuration in which ultrasonic detectors have traditionally been mounted (42). The designed system was implemented at two test sites (a highway with light traffic and an arterial with heavy traffic). While mounted in the side-fire configuration, it was arranged with such a detection zone as to only detect vehicles in one lane. At the highway test bed, the five-minute count mean absolute percent error was 0.7% across daylight, night, and heavy rain conditions. Five-minute average speed was also calculated at this test bed, based on an assumed average vehicle length, with the following root mean squared errors for each condition: 7.4% (daylight), 6.9% (night), and 7.6% (heavy rain). At the arterial test bed with heavier traffic, the five-

minute count MAPE was 3.4% during testing, which included dusk, night, and heavy snow conditions.

The second study developed a neural-edge-based tracking video detection algorithm (43). Most other video detectors with tracking algorithms employed either background subtraction or edge detection, but the neural-edge-based detection method was shown to outperform other methods. The count accuracy of the new algorithm was over 98% in overcast daylight conditions. Dawn, dusk, and night conditions caused the greatest difficulty for the new algorithm, but count accuracy was still above 96% in these conditions. Vehicle classification ranged from 80% correct classification in dawn, dusk, and night conditions, to over 95% correct detection in daylight conditions.

A report published in 2003 by the University of Utah did not provide any unique detector accuracy assessment, but provided a well-organized review that was state-of-the-art for that time period (2). The report begins with a presentation of the various traffic data needs and explanations of the functionality of various types of traffic detector technologies. It then assessed detector technologies based on various selection criteria such as data type, data accuracy, ease of installation, and cost. Finally the report provided a procedure for the selection of an appropriate traffic detector for a specific installation.

In 2004, a study was conducted to assess the accuracy of VIP detectors installed at intersections in Utah (44). While the results of detector functionality at an intersection installation cannot directly be related to results at a freeway installation, there is value in recognizing trends that emerge when different environmental and lighting conditions are considered. Eight detectors from four manufacturers were analyzed in the study. It was found that the detection performance was good under day and dusk conditions, with

87.2% correct detection across all manufacturers in both conditions with reduced performance: in inclement weather (81.3% correct detection) and at night (73.4% correct detection). This indicates that weather and lighting had an impact on VIP performance at street intersections, and, potentially, at freeway installations.

A 2009 report was written in an attempt to apply the ISO *Guide to the Expression of Uncertainty in Measurement (GUM)* to the quantification of traffic detector performance (45, 46). In an effort to conform to this broad standard, the method of evaluation was rigid, and produced results which were conveyed in very general statistical measures, such as the mean and standard deviation of count error. The study recommended the use of traffic stream videotape for baseline volume and classification data and probe vehicles (with onboard GPS) for baseline speed data. The test methodology was demonstrated on the DataCollect SDR radar detector and Traficon VIP/D video detector. Count error histograms showed that both detectors tended to undercount, provided the tested calibration. Length histograms showing reported lengths of multiple passes with a probe vehicle of known length showed that the VIP/D underestimated length more severely than the SDR, and also had a greater standard deviation of length measurements for the probe vehicle. This analysis was used as a proxy for length-based classification. Finally, speed error histograms showed that both detectors tended to overestimate speed, but the VIP/D overestimated more severely, and also had a greater standard deviation of speed error. The authors determined that the VIP/D did not necessarily have worse detection capabilities than the SDR, but required more precise configuration and calibration in order to facilitate comparable detection. It was also concluded that there were both advantages and disadvantages to attempting to

apply a broad standard such as the ISO GUM to a specific task such as analyzing traffic detectors.

Finally, a recent study conducted in Hawaii evaluated three different detector technologies based on their vehicle classification capabilities (47). The three systems chosen for this study were the Autoscope Rack Vision Terra VIP, Custom Electronic and Optical Solutions TIRTL active infrared detector, and Wavetronix SmartSensor HD radar detector. TIRTL, an axle-based detector, was analyzed based on a 15-class scheme composed of the standard FHWA 13 category scheme plus a 14^{th} class for "unclassified" vehicles and 15^{th} class for 8-15 axle road trains (48). The Rack Vision Terra and SmartSensor HD provide vehicle lengths, and were analyzed based on a five-class, length-based scheme, with classes designed to represent motorcycles, light-duty vehicles, single-unit heavy vehicles, articulated heavy vehicles, and multiple-trailer trucks. The detectors were tested at four sites with varying levels of truck traffic. The study concluded that the Rack Vision Terra was adequate for daytime classification on low volume arterials, but was affected by poor lighting and weather conditions, and had difficulty discerning motorcycles from other light-duty vehicles. It was concluded that neither the Rack Vision Terra nor SmartSensor HD provided desirable accuracy at freeways, primarily due to congestion. TIRTL was found to have good classification performance, but was inhibited by the need for a flat cross-section in order to achieve optimal performance. This study did not address the relative cost of the detectors or the environmental impacts on classification.

2.5 Chapter Summary

The literature review presented in this chapter provided background information on traffic detector technologies, the standards that direct their implementation, and previous research on the assessment of competing traffic detectors. Brief explanations were given of the functional characteristics of different types of intrusive detectors (including inductive loop, pneumatic road tube, magnetometer, and weigh-in-motion systems) and non-intrusive detectors (including video image processors, microwave radar, magnetic, ultrasonic, passive acoustic, infrared, and combined systems). This introduction to the functional characteristics of the various detectors was followed by a review of existing standards governing the selection and performance evaluation of traffic detectors. Of particular relevance were two ASTM standards that provided a basis for a traffic detector performance-based specification, as well as an accompanying standard method for evaluating the performance of an installed detector. While the latter of these two standards can be used to assess detectors, it is based on a duplex *accept* or *reject* decision that relates to the performance specification it was intended to complement, but is less relevant to research on the common sources of error for various detector technologies. The remainder of the literature review focused on the methodologies and findings of traffic detector assessments performed over the past two decades. Tables 2.9 and 2.10 provide an overview of the findings of these previous studies as they relate to the Wavetronix SmartSensor 105 and 3M (or subsequent GTT) Canoga Microloop 702, which are evaluated in the current study. No previous studies have specifically assessed the Autoscope Solo Pro II or ISS RTMS G4 at freeway installations.

Table 2.9 Previous Field Test Results for the Wavetronix SmartSensor 105

Organization	Mounting Location	Volume	Speed	Classification	Response to Environment
Minnesota Guidestar – PNITDS (24)	8 lane freeway – various sidefire locations	24-hr APD per lane 1.4% -4.9%	24-hr APD per lane 3.0% - 9.7%	3 lanes, 3 length bins APD per lane 0.4% -5.6%	no impact
	4 lane freeway – sidefire 17’ height, 17’ offset	24-hr APD per lane 1.6% -3.9%			
	4 lane freeway – sidefire 19’ height, 15’ offset	1-hr APD per lane 0.0% - 0.7% for 2 far lanes, 9.7% - 20% for 2 near lanes			
	4 lane arterial – sidefire 17’ height, 15’ offset	24-hr APD per lane 0.6% - 2.7% in 3 farthest lanes, over-counting in near lane			
UNL (34)	5 lane freeway – sidefire 18’ height, 19’ offset	15-min APD per lane 1.4% - 5.8%			affected by traffic volume

Table 2.10 Previous Field Test Results for the 3M Canoga Microloop 702

Organization	Mounting Location	Volume	Speed	Classification	Response to Environment
TTI (25)	3 lane arterial – under bridge	3.25-hr APD per lane 0.1% - 1.5%			
	2 lane highway – conduit under pavement	15-min APD per lane within 5% for 93.5% - 99.5% of intervals	1-min average speed error μ = -0.25 mph σ = 3.6 mph		
Minnesota Guidestar – Phase 2 (23)	3 lane freeway – conduit under pavement	24-hr APD < 2.5% (within accuracy of baseline loop)	24-hr APD per lane 1.4% - 4.8%		
	1 lane freeway – under bridge	24-hr APD 1.8%			
Minnesota Guidestar – Phase 3 (8)	3 lane freeway – conduit under pavement	24-hr APD per lane 1.1% - 3.7%	24-hr APD per lane 0.0% - 1.3%	3 lanes, 3 length bins 30-min APD 2.9%	no performance degradation

CHAPTER 3 NTC/NDOR NON-INTRUSIVE DETECTOR TEST BED SETUP

The NTC/NDOR non-intrusive detector test bed used in this study was developed with limitations identified by previous research conducted at the University of Nebraska in mind. The previous evaluation of three non-intrusive detector technologies found that conclusions were limited by the fact that the detectors under evaluation were installed at locations separated by approximately 900 feet, and that the installations were temporary (34). As a result of these limitations, as well as the recognition of a need for future research at a permanent test facility with collocated detectors, the Nebraska Department of Roads (NDOR) and Nebraska Transportation Center (NTC) planned a permanent, non-intrusive traffic detector test bed along an urban section of I-80 in Omaha, near the Giles Road interchange. Figure 3.1 shows the location of the NTC/NDOR non-intrusive detector test bed. This chapter outlines the configuration of that test bed and calibration of the installed detectors.

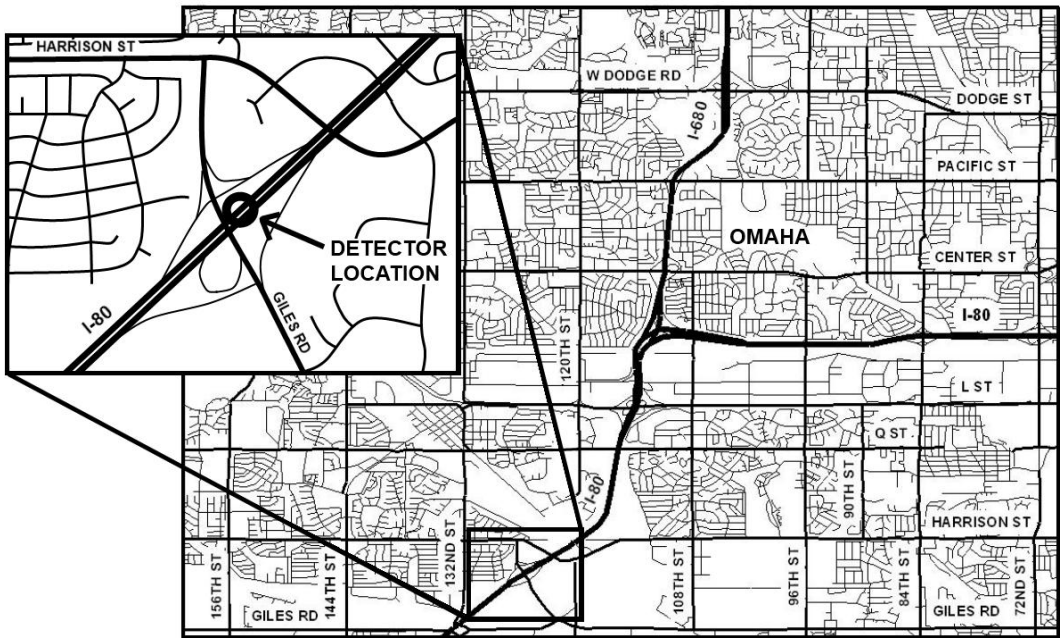

Figure 3.1 Test Bed Location

3.1 Test Bed Organization

In 2007, NDOR installed fixtures at the NTC/NDOR non-intrusive detector test bed in Omaha. The original installation included three above-ground detection systems and one buried detection system, along with appropriate support infrastructure. The buried detector was a 3M Canoga Microloop 702 magnetic induction system, and the three above-ground systems were the Autoscope Solo Pro II VIP system, EIS RTMS microwave radar system, and Wavetronix SmartSensor 105 microwave radar system. Three of these systems remained in place throughout the duration of this study. However, the EIS RTMS was replaced in October 2009 with the newer generation ISS RTMS G4 by NDOR personnel. This technology is examined in the current study.

The test bed layout is shown in figure 3.2. This figure shows the locations of the detectors, as well as additional support infrastructure at the site. This support infrastructure includes two 41-foot-tall support towers for the sidefire radar detectors, two NEMA 332 cabinet enclosures with necessary electronic fixtures, a third 41-foot-tall support structure with PTZ surveillance camera and wireless communication link to the NDOR network, and conduit, along with appropriate electrical and communications wiring, to link site fixtures.

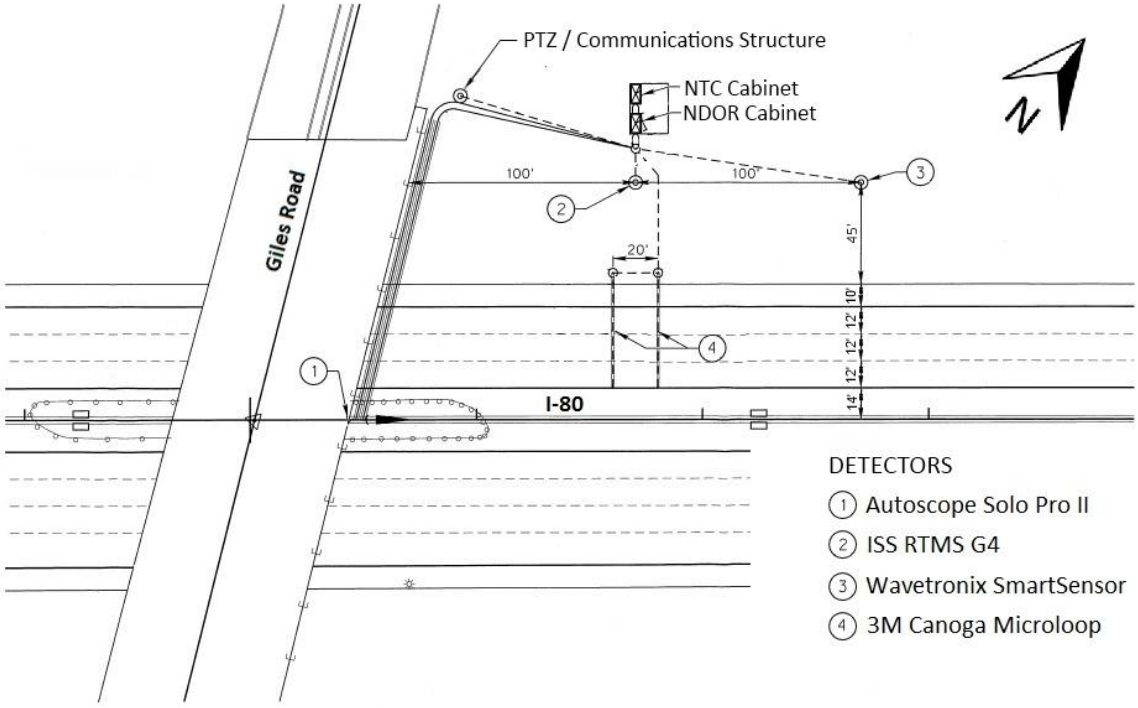

Figure 3.2 Test Bed Layout

The physical locations of the detectors are shown in figure 3.2. The detection zones (areas in which they detect the presence of vehicles) are shown in figure 3.3. The Canoga Microloop 702 and Solo Pro II were arranged to have overlapping detection zones, which is to say that they detect vehicles at the same location along the roadway. These two systems were configured for detection in the three westbound lanes. The RTMS G4 has a detection zone which overlaps those of the Canoga Microloop 702 and Solo Pro II detectors; it was configured to detect traffic in both the eastbound and westbound lanes. The SmartSensor 105 also detects traffic in both directions, but has a detection zone 100-feet east of the other three systems. While this offset was not ideal for data comparison, it diminished the likelihood of crosstalk between the RTMS G4 and SmartSensor 105. Crosstalk is a phenomenon in which the electromagnetic signals from these two detectors could interact in a manner that would degrade performance if they

were closer together. Because two of the detector technologies were installed for the three westbound lanes only, this thesis focused on analysis of data reported for those lanes.

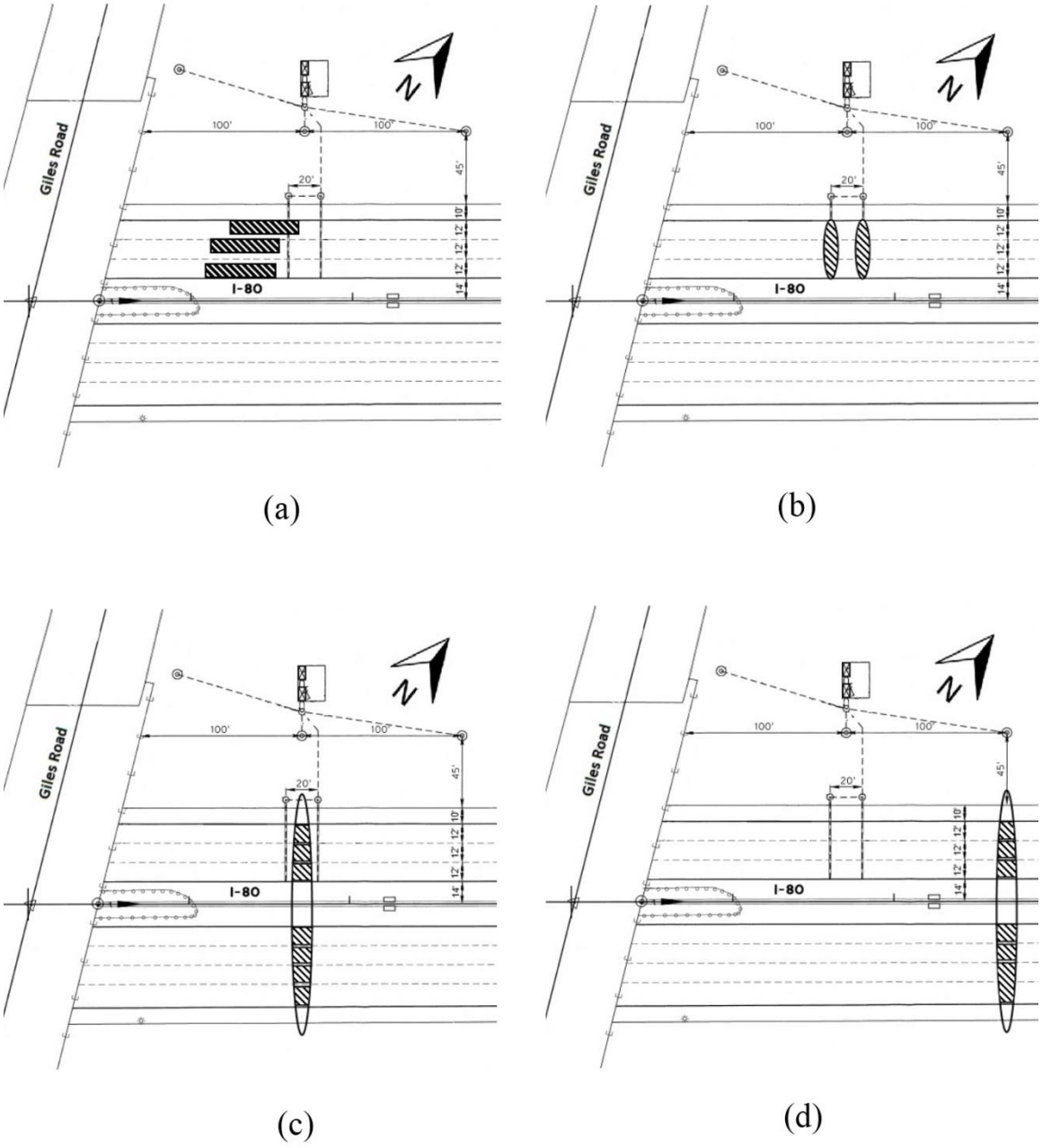

Figure 3.3 Detection Zones of the Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

The NDOR cabinet was outfitted with additional electronic equipment at the time the detectors were installed in order to support the various detectors. This equipment

included a Transition Networks SISTM10XX-180 Ethernet switch. This switch connects all of the detectors to the wireless bridge, which facilitates communication with the NDOR Ethernet backbone. This allows the detectors to be accessed by NDOR personnel from remote locations, including the NDOR main office. It should be noted that the reason for data eventually being collected on-site was due to bandwidth limitations of this wireless bridge. In addition to the Ethernet switch, support electronics for the various detectors are housed in the NDOR cabinet. The Autoscope Solo Pro II VIP requires an ACIP4E communications panel, which sends power to the VIP and converts the data and video signal from the 11 conductor cable into an Ethernet output, as well as an NTSC coaxial video output.

The ISS RTMS G4 has a native Ethernet output and thus does not require additional hardware in the cabinet. The Wavetronix SmartSensor 105 is connected via RS-485 serial communication to a Wavetronix Click!200 lightning surge protector in the cabinet. This is connected to a Wavetronix Click!301 serial to Ethernet converter, which sends an Ethernet output to the switch.

The three westbound lanes of I-80 were outfitted with two 3M Canoga Microloop 702 detectors per lane. These were connected via RS-485 serial communication to three rack mounted 3M 942 Traffic Monitoring Cards (one per lane). These each output RS-232 serial communication, which was connected to three Wavetronix Click!301 serial to Ethernet converters, which send three Ethernet connections to the switch.

At the outset of this study, a second NEMA 332 cabinet enclosure was installed for NTC next to the existing NDOR cabinet, as shown in figure 3.2. The two cabinets were connected via conduit. Power and an Ethernet connection from the NDOR cabinet

were supplied to the new NTC cabinet. A field-hardened VIA AMOS-3001 embedded computer was installed in the NTC cabinet. This computer accesses the detector output from the NDOR cabinet through an Ethernet connection. The computer stores detector data from all four non-intrusive detectors being evaluated. In addition, the computer stores video from the Autoscope Solo Pro II camera. This computer was selected based on its operating specifications, which allow it to operate reliably under the harsh environmental conditions encountered in Nebraska, such as extreme heat and humidity. An AXIS 241Q video server was also installed in the NTC cabinet. This equipment is used to digitize the NTSC video from the Autoscope camera so that it can be recorded by the NTC data collection computer. Finally, a D-Link DGS-2205 Ethernet switch was installed in the NTC cabinet. This switch allowed the data collection computer to communicate with the AXIS video server as well as the detector fixtures in the NDOR cabinet. Figure 3.4 shows the electronic components installed at the test bed, and communications protocols linking the various components. Additionally, figures 3.5-3.8 show the components as they are laid out in the NDOR and NTC cabinets. The resulting test bed is maintained jointly by NDOR and NTC.

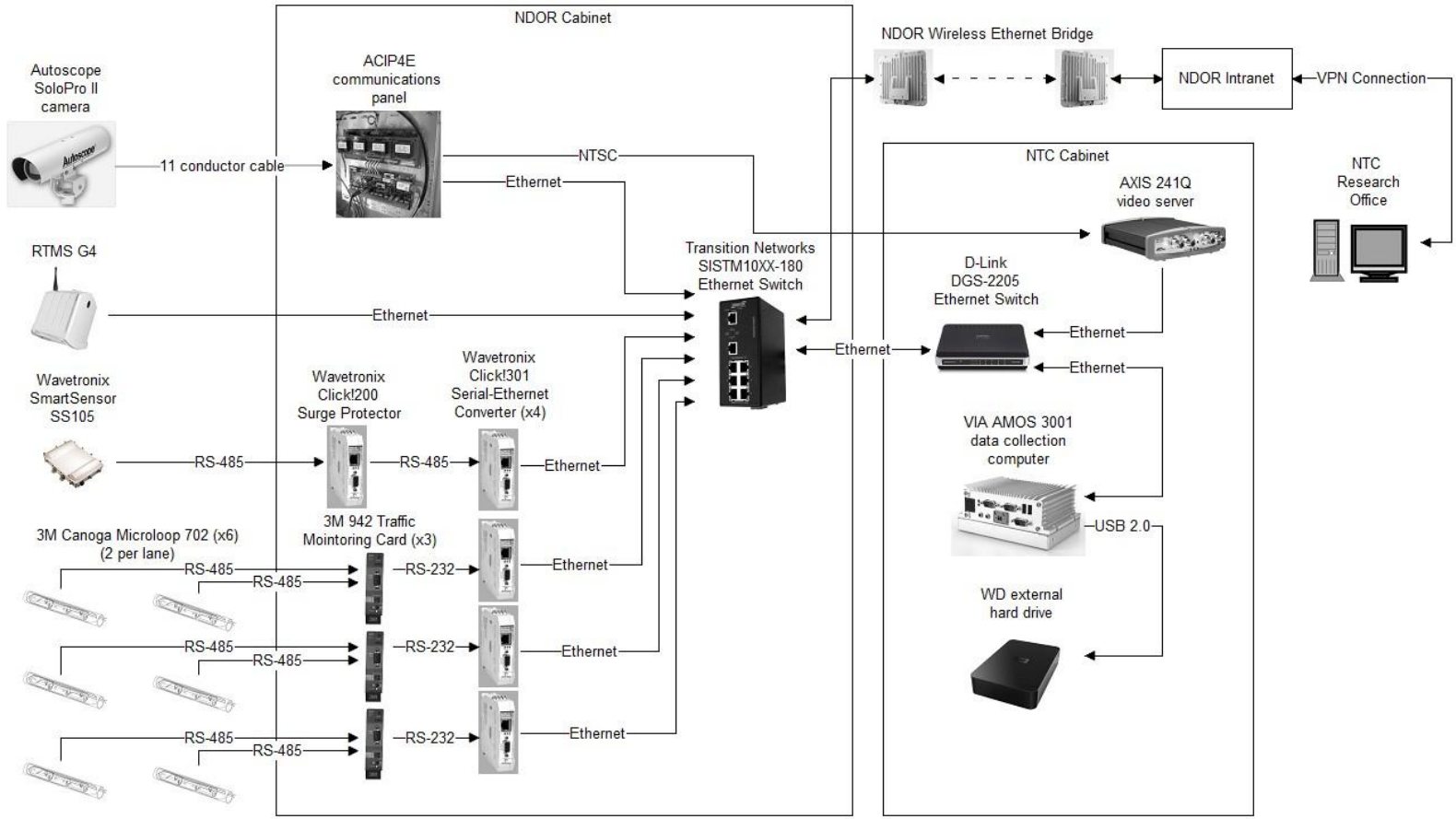

Figure 3.4 Test Bed Fixture Block Diagram

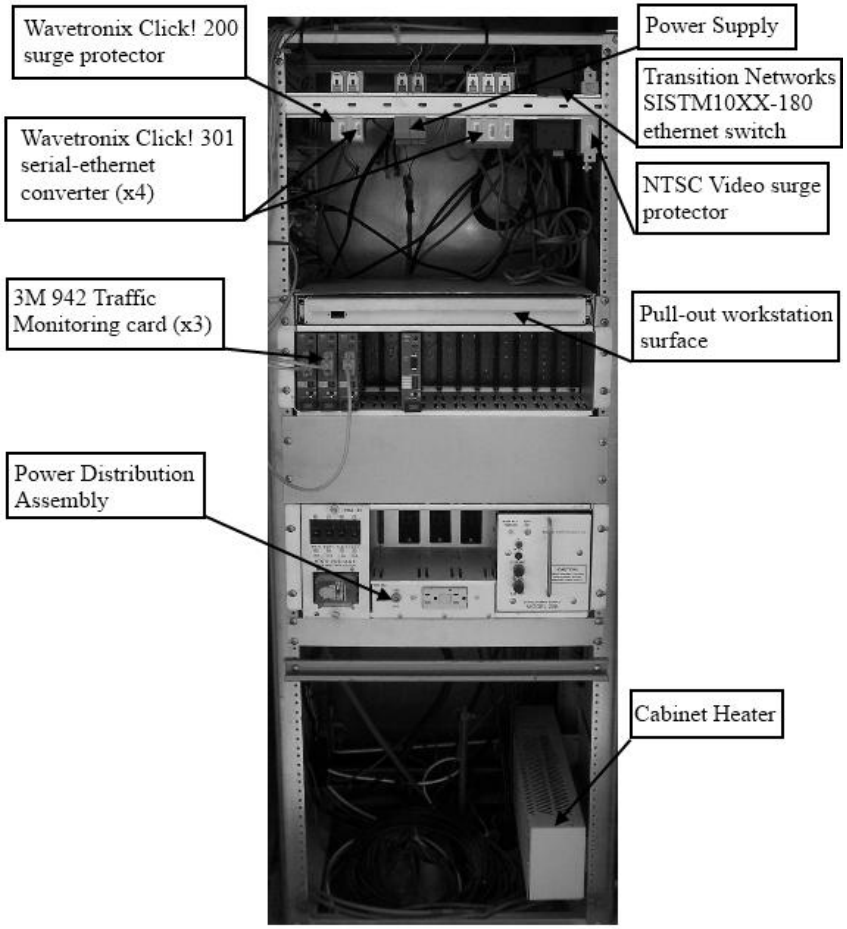

Figure 3.5 Front of NDOR Cabinet

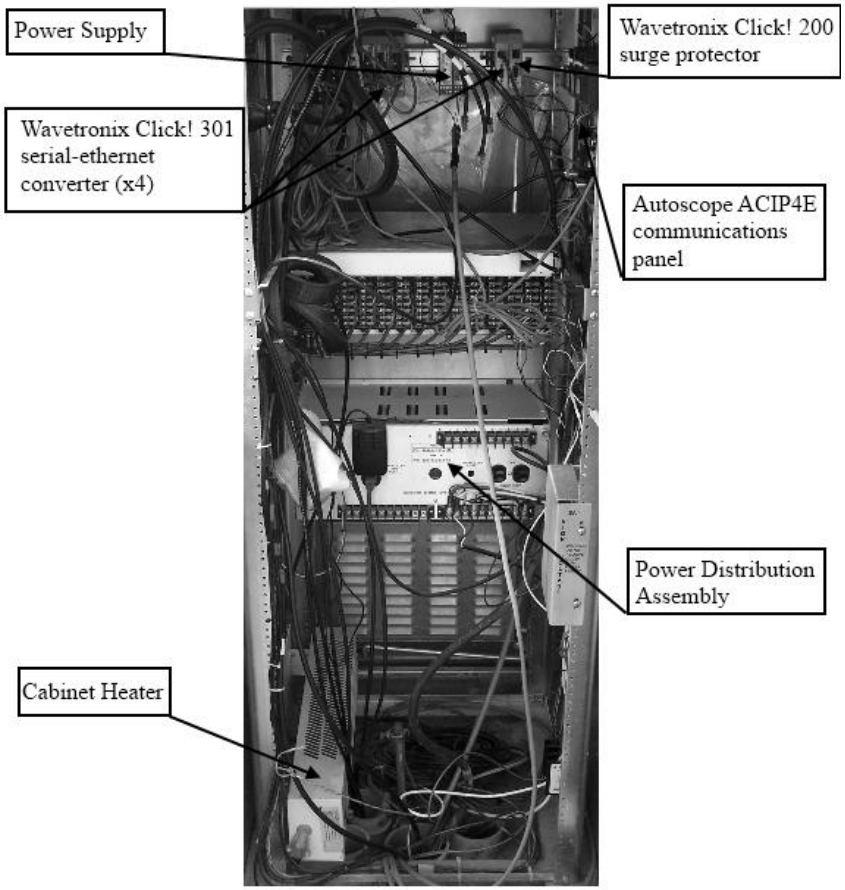

Figure 3.6 Back of NDOR Cabinet

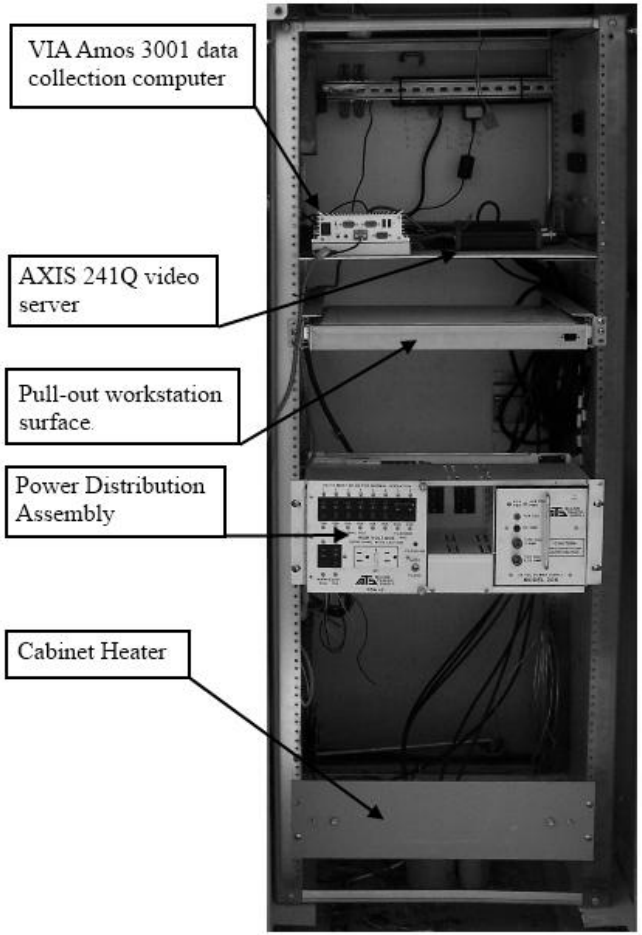

Figure 3.7 Front of NTC Cabinet

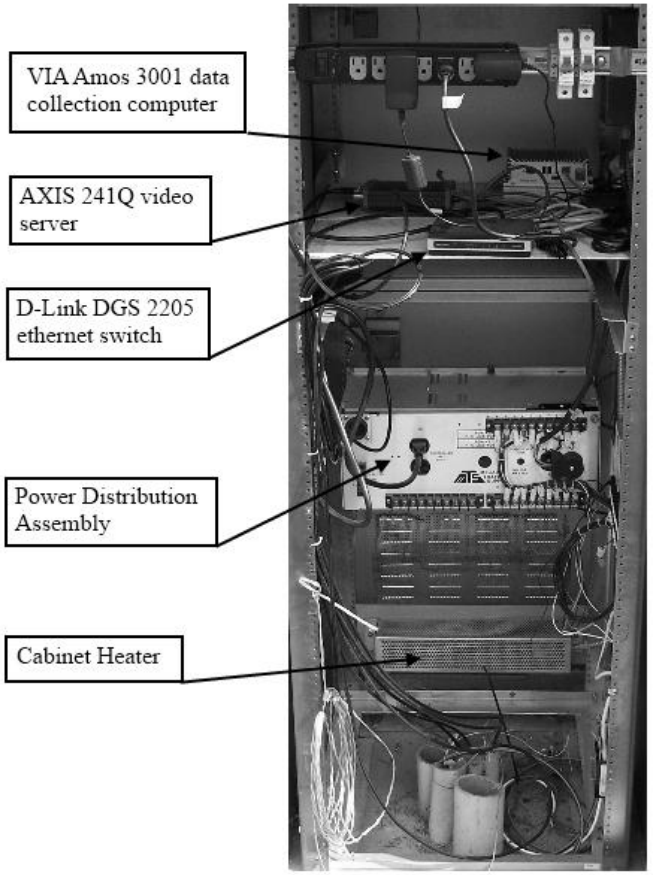

Figure 3.8 Back of NTC Cabinet

At the NTC/NDOR non-intrusive detector test bed, there are two methods of obtaining detector data. The first alternative is to physically visit the test bed and 1) download the information from the NTC data collection computer, or 2) remove the hard drive from the NTC data collection computer and retrieve the information later. The second alternative is to connect to the NDOR intranet through a VPN login, which was provided to the researcher. Once connected through the VPN, Microsoft's Remote Desktop Connection software can be used to control the NTC data collection computer from the NTC ITS Laboratory. For detector calibration, it was advantageous to physically be at the test bed so that the vehicle detection could be manually verified. However, the ability to remotely access the data collection computer allowed the researcher to commence and terminate data collection intervals, while mitigating risk by limiting time spent at a potentially dangerous roadside location.

Because a goal of this thesis was to examine the impact of environmental factors on the performance of the various detection technologies, it was necessary to also collect weather data. While the test bed is not instrumented with a weather station, weather data was available at the Millard Airport (KMLE) weather station, which is located 0.5 miles north of the test bed. A full METAR weather report is logged online every 20 minutes (49). This information was automatically recorded at the NTC ITS Laboratory. It was determined that this weather information, along with confirmation of conditions through manual review of video from the test bed, would provide the necessary weather data for the proposed analysis.

3.2 Detector Locations and Configuration Process

Each of the detectors evaluated in this study required a specific mounting location and configuration process. This section of the report outlines the location and configuration tasks for each detector.

3.2.1 Autoscope Solo Pro II

The Autoscope Solo Pro II camera was mounted 47 feet above the roadway on a street light pole on the Giles Road overpass bridge, as seen in figure 3.9. At this location, it is offset 14 feet from the nearest detected lane, as seen in figure 3.2. The detection zones, which are shown in figure 3.3(a), begin 130 feet upstream and end 65 feet upstream of the camera's location.

Figure 3.9 Solo Pro II Camera Mounting Location

The hardware components of the Solo Pro II detection system as it is installed at the NTC/NDOR non-intrusive detector test bed include the Solo Pro II camera, a pan/tilt

head on which the camera is mounted, an Autoscope ACIP4E communications panel, and an 11-conductor cable connecting the pole-mounted camera to the communications panel in the NDOR traffic cabinet. The software components of the system include the image processing software, which is run on hardware in the camera housing, and the Autoscope Software Suite, which is run on the NTC data collection computer and is used to configure the image processing software and collect detection data. The version of the Autoscope Software Suite used for this data collection effort was the Autoscope Network Browser Version 8.3.2.

At the outset of this study, it was necessary to calibrate the Solo Pro II to conditions at the site. This initial calibration was conducted using the Autoscope Network Browser Version 8.3.2. First, the pan, tilt, and zoom were adjusted so that the desired section of roadway was in the frame. Next, the geometry of the image was calibrated by placing a series of lines in the image that were longitudinal and transverse with respect to the roadway, and dimensioning the offsets between these lines. The camera height above the roadway was also required for this calibration. This geometric information allowed the image processing software to calculate parameters such as vehicle speed and length. The next step was to place virtual detectors on the image. These virtual detectors, which can be seen overlaid on the image in figure 3.10, defined which pixels were monitored for changes by the image processing software to be registered as detections. The boxes in figure 3.10 are speed detectors. They calculate vehicle speed and length based on a speed trap algorithm, which analyzes the time of pixel color/hue change at the upstream end of the box and, subsequently, at the downstream end of the box. The final step in the calibration of these detectors was to adjust their placement for optimal detection

accuracy. Due to the oblique angle of the image, it was necessary to offset the virtual detectors so that they were not directly over the center of each lane. In figure 3.10, it can be seen that the speed detector for lane 1 (the lane farthest from the camera) was actually located mostly over the shoulder of the roadway. This was to avoid large vehicles in lane 2 (the center lane in the image) from being detected in both lanes 1 and 2. Once the virtual detectors were configured, the detection file was uploaded to the image processing software in the camera and detection could commence.

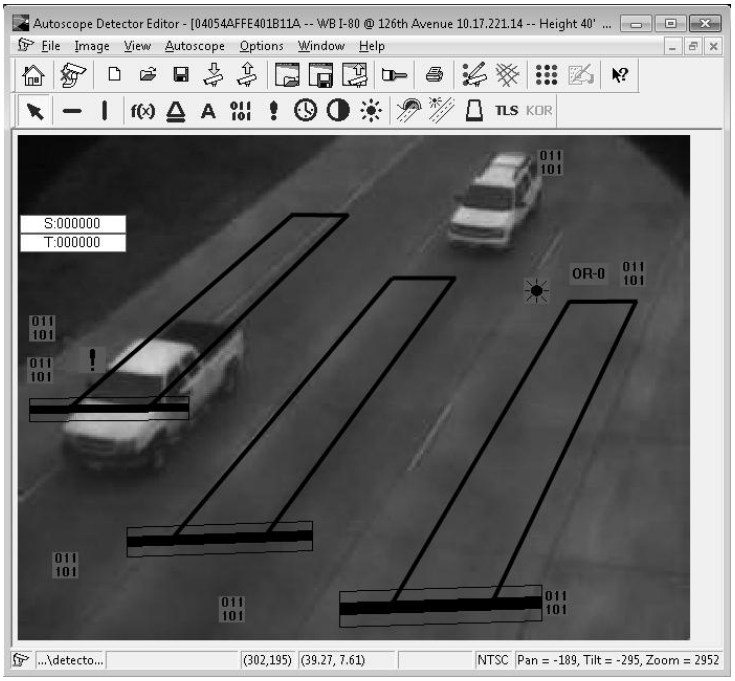

Figure 3.10 Autoscope Virtual Detector Layout

During a site visit on May 11, 2011, Mr. Jordan Schwening, a representative of Mid American Signal, reviewed the detector layout and confirmed that the detector placement was appropriate for the camera location. He made two qualifying comments, first, noting that the oblique angle of the camera view made the detector susceptible to errors related to occlusion (though the camera height reduced the severity of this issue). Occlusion refers to a scenario in which a large vehicle in a lane closer to the camera hides

(occludes) a smaller vehicle in a lane farther from the camera. In an ideal situation the camera would be mounted directly over the middle lane to minimize occlusion. For this site (and presumably most other traffic detection sites) it would be cost-prohibitive to install the necessary support infrastructure to provide such a mounting location. Therefore, the planned data collection commenced at the current mounting location.

The second concern noted by Mr. Schwening was regarding noise in the video signal, which he thought could cause false detections. This noise was found to be due to a loose connection between the camera and the cable leading back to the cabinet, and was addressed by tightening the loose connection. As the detection algorithm is implemented in the camera itself, and because this noise was introduced to the video signal after the signal had left the camera, it was determined that this issue had not affected previous detection results.

Another concern pertained to the presence of the pan/tilt mounting for the Solo Pro II camera; it could have been easy for someone to inadvertently adjust the video alignment, which would have moved the virtual detectors to less ideal locations. As a quality control measure, a reference screenshot was created when the virtual detector configuration was finalized. This reference screenshot was used throughout the study to confirm that the camera angle was not altered. For each day on which data was collected, a video frame was visually compared to the reference screenshot. No camera realignment was noted throughout the data collection phase of this study.

3.2.2 3M Canoga Microloop 702

The 3M Canoga Microloop 702 detectors were installed in two parallel three-inch diameter PVC conduits, which were bored 21 inches below the road surface. The boring

process was such that the lower and upper boundary on the depth were 18 and 24 inches. These conduits were offset by 20 feet, as seen in figure 3.2. Microloop 702 probes were installed in each conduit under each of the three westbound lanes of I-80. By offsetting the conduits, and therefore the microloops under each lane, in this manner, the detectors could function as a standard speed trap. Pull box covers at the end of both conduits can be seen on the shoulder of the road in figure 3.11.

Figure 3.11 Microloop 702 Pull Box Locations

The hardware components in the microloop detection system included the three-inch PVC conduits, the Microloop 702 probes installed in the conduits, a pull box at the end of each conduit on the shoulder of the roadway, cabling from the pull boxes to the NDOR traffic cabinet, 3M 942 Traffic Monitoring Cards in the NDOR traffic cabinet for each lane, and Wavetronix Click!301 serial to Ethernet converters for each lane to allow the serial output from the Traffic Monitoring Cards to be transmitted to the NTC data collection computer via Ethernet communications. The software components of the

system included the detection software on the 3M 942 Traffic Monitoring Cards, the Lantronix CPR Version 4.3 virtual serial port package, and the Global Traffic Technology ITS Link Version 3.4.0.8 software package, which is run on the NTC data collection computer. A screenshot in figure 3.12 shows the user interface of the ITS Link software, which has tools for calibration of the detectors and the collection of traffic data.

Figure 3.12 ITS Link Software Screenshot

The Microloop 702 detection system was installed during the initial construction of the NTC/NDOR non-intrusive detector test bed in 2007. Due to personnel turnover since that time and an inability to find previous documentation of communications protocols, there was difficulty establishing communications between the ITS Link software and 3M 942 Traffic Monitoring Cards at the outset of this study. While the ITS Link software can communicate over RS-232 serial communications, the RS-232 output of the 942 Traffic Monitoring Cards was converted to Ethernet by the Wavetronix Click!301 serial to Ethernet converters for networking with the NDOR intranet and NTC data collection computer. Ultimately, the Lantronix CPR Version 4.3 software package was installed on the NTC data collection computer. This software package created virtual serial ports which directed the IP addresses of the three Wavetronix Click!301 to "virtual"

RS-232 serial ports on the NTC data collection computer. With this software in place, the ITS Link software was able to communicate with the 3M 942 Traffic Monitoring Cards.

Once communication was established with the Microloop 702 detection system, it was found that the detectors had been calibrated at the initial installation. A preliminary comparison of ten minutes of detector data with ground truth from 4:10 PM to 4:20 PM on March 3, 2011 indicated that the volume error was below 3%. Additionally, the speed and length output for this period provided reasonable values, although ground truth values were not available for comparison at that time. It was therefore determined that there was no need for recalibration of the Microloop 702 detection system, specifically, for this study. When Jordan Schwening, a product representative with Mid American Signal, visited the site on May 11, 2011, he agreed that these detectors were calibrated correctly and were functioning as intended.

3.2.3 Image Sensing Systems RTMS G4

The RTMS G4 was mounted on a support structure (see figure 3.13) at the location shown in figure 3.2. Its mounting height was 30 feet above the roadway, and it is offset 54 feet from the nearest lane. This mounting location is consistent with manufacturer recommendations for optimal performance (50).

(a) (b)

Figure 3.13 G4 Mounting Support Structure (a) and Unit (b)

The hardware components of the G4 detection system include the RTMS G4 radar unit, a power supply in the NDOR traffic cabinet, and an Ethernet cable over which data can be transmitted to the NTC data collection computer. The software components of the system include the internal signal processing software within the radar unit and the WinRTMS4 Version 4.5.0.0 software utility, which is run on the NTC data collection computer. A screenshot of the user interface for the WinRTMS4 utility can be seen in figure 3.14. This utility is used to calibrate the G4 detector as well as collect G4 data.

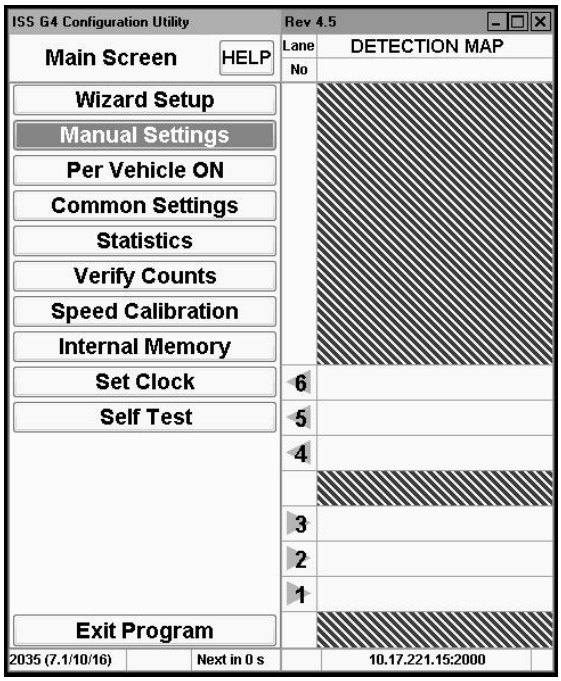

Figure 3.14 WinRTMS4 Screenshot

Communications were established with the G4 relatively easily due to its native Ethernet output. A preliminary comparison of ten minutes of data with ground truth from 4:10 PM to 4:20 PM on March 3, 2011 indicated that the volume error was below 1%. Speed and length output also appeared to be reasonable, although ground truth for these parameters was not available at the time. Therefore, it was determined that there was no need to adjust the configuration of the G4. Additionally, it was deemed unnecessary to request a product representative visit the site based on a report from NDOR personnel that a representative was present for the initial installation of the device, and the positive findings of the preliminary ten-minute data collection period. While the RTMS G4 was configured to detect traffic in both directions, this study focused its analysis on the three westbound lanes (those nearest to the detector) due to limitations of the ground truth data source.

3.2.4 Wavetronix SmartSensor 105

The Wavetronix SmartSensor 105 was mounted on a support structure (figure 3.15), located 100 feet from the G4, as seen in figure 3.2. Because the two devices use similar microwave radar technology, they were separated to prevent "crosstalk," or signal interference. The detector was offset from the nearest lane by 54 feet and mounted at a height of 30 feet above the roadway. While the mounting offset from the nearest lane was greater than the minimum specified by the manufacturer, and all lanes were less than the maximum distance away from the detector, it was not within the "recommended offset" range of 25 to 35 feet (51). It was understood that the support structure was installed at its current location to maintain a specified clear zone next to the roadway. Due to this clear zone consideration, the offset of 54 feet was considered a typical mounting location, though it fell outside the manufacturers recommended offsets but within its acceptable offsets. The mounting height of 30 feet above the roadway matches the manufacturers recommendation for the given offset of 54 feet.

(a) (b)

Figure 3.15 SmartSensor 105 Mounting Support Structure (a) and Unit (b)

The hardware components of the SmartSensor 105 detection system include the SmartSensor 105 microwave radar unit, SmartSensor cable from the radar unit to the NDOR traffic cabinet, Wavetronix Click!200 Surge Protector, and Wavetronix Click!301 serial to Ethernet converter in the NDOR traffic cabinet. The software components of the system include the internal signal processing software within the radar unit and the SmartSensor Manager Version 3.0.0 software utility, which is run on the NTC data collection computer. This software package, which can be seen in figure 3.16, includes tools for calibration as well as data collection.

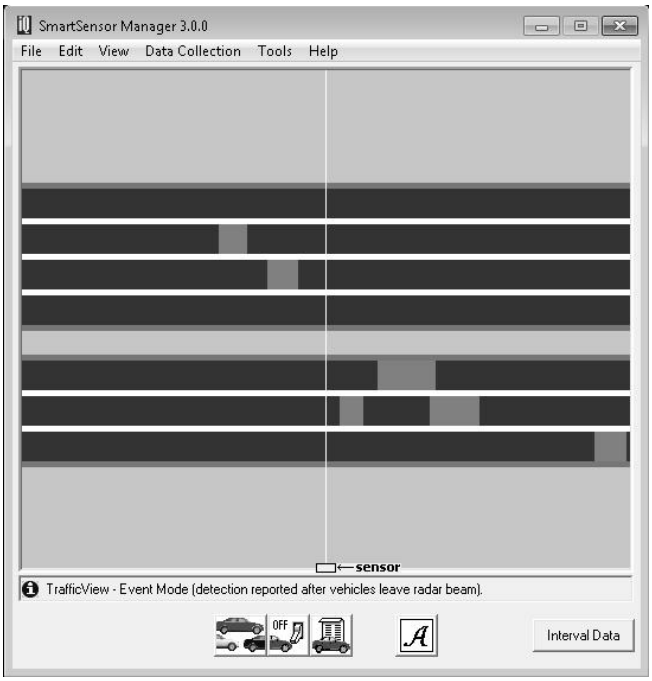

Figure 3.16 SmartSensor Manager Screenshot

When communication with the SmartSensor 105 was established, the existing detection zones appeared to align with the existing traffic lanes. Therefore, a preliminary data set was collected. This comparison of ten minutes of data with ground truth from 4:10 PM to 4:20 PM on March 3, 2011 showed that the volume error was greater than 30%, indicating a problem with the detector configuration. This information was provided to Mr. Jordan Schwening, a product representative with Mid American Signal. During a site visit on May 11, 2011, Mr. Schwening adjusted the per-lane sensitivity settings, which appeared to correct most of the detection issues based on observed performance during the remainder of the site visit. The SmartSensor 105 configuration resulting from these adjustments was used for the duration of the data collection for this thesis.

3.3 Chapter Summary

The test bed setup outlined in this chapter provides information on the NTC/NDOR non-intrusive detector test bed from which the data utilized in conducting this thesis were obtained. This test bed is located in Omaha, along I-80, at the Giles Road interchange. While the site was intended to be representative of typical urban freeway traffic data collection sites in Nebraska, the exact mounting configuration and calibration, as well as site geometrics, will undoubtedly vary slightly between installation locations. The information in this chapter outlines the characteristics of this site and the detector calibration, in order to demonstrate its representative nature, while also examining its unique characteristics. The first portion of this chapter also details the communications infrastructure and describes the NTC data collection computer, both of which were installed at the test bed.

In addition to defining the characteristics of the NTC/NDOR non-intrusive detector test bed site, information is provided in this chapter on the four detectors under evaluation, along with their supporting infrastructure. These four non-intrusive traffic detectors are the Autoscope Solo Pro II, 3M Canoga Microloop 702, Image Sensing Systems RTMS G4, and Wavetronix SmartSensor 105. The mounting locations of these detectors are described, as well as the system components for each detector. Finally, the calibration of each detector for this evaluation is presented. The installation and calibration of these detectors is summarized in table 3.1. Once the detectors were calibrated, the NTC/NDOR non-intrusive detector test bed was ready for data collection.

Table 3.1 Detector Calibration Summary

Detector	Installed / Configured	Initial Calibration	Further Calibration
Autoscope Solo Pro II	Yes (Spring 2007)	Yes (12-15-2010)	Yes (06-07-2011)
3M Canoga Microloop 702	Yes (Spring 2007)	No	No
Image Sensing Systems RTMS G4	Yes (10-15-2009)	Yes (12-14-2010)	No
Wavetronix SmartSensor 105	Yes (Spring 2007)	Yes (12-14-2010)	Yes (05-11-2011)

CHAPTER 4 DATA COLLECTION AND REDUCTION

4.1 Data Collection

The data collection effort for this study took place during the six months between March and August of 2011 at the NTC/NDOR non-intrusive detector test bed described in chapter 3. While data was not collected continuously during the entire study period, the duration of the selected collection period allowed for representative data to be collected under various environmental conditions. Preliminary analysis of the initial data led to a recalibration of some of the detectors during March, April, and May of 2011. It was determined that only data collected after the final calibration of all detectors would be analyzed. Therefore, the data analyzed in this thesis was collected in June, July, and August of 2011.

Output from the four detector systems at the site was available for collection through connections to an Ethernet switch in the NDOR cabinet. Through this connection, time-stamped vehicle observations (with speed information) were archived to the NTC data collection computer. Additionally, video from the Autoscope Solo Pro II camera was routed to the data collection computer by the AXIS video server, and was recorded. The recorded data and video were transferred to a WD external hard drive located in the NTC research cabinet. At intervals of approximately 14-days, the external hard drive in the field was manually retrieved and brought back to the NTC ITS laboratory. When one external hard drive was retrieved, a comparable unit was left in its place for the next 14-day interval. The data and video on the retrieved external hard drive were then transferred to a server at the NTC ITS laboratory. In addition to the data retrieved from the test bed, weather data were obtained in real-time from the Millard

Airport (KMLE) weather station via an internet connection (49). This information was also archived at the NTC ITS laboratory. The reported weather conditions were stored with each dataset in order to expand the scope of the analysis, and video recordings were referenced to confirm the reported weather condition.

An issue arose during data collection involving the Microloop 702 detectors. It was noted that the ITS Link software for the microloops was utilizing up to 90% of the NTC data collection computer's 1.2GHz processor. Initially this was noted as peculiar but unimportant. Later, it became apparent that, while the microloops were reporting accurate vehicle volumes at the beginning of each data collection interval, a large number of vehicles were not recorded after approximately two hours of data collection. During a site visit, it was noted that indicators on the detector card were signaling detections that were not being recorded on the data collection computer. It was concluded that the large percentage of "missed vehicles" was the result of a communications issue between the detector card and data collection computer, and not a result of poor detection. The detector manufacturer was contacted, but was unable to provide an explanation for the communications issue. It was determined that only data collected during the initial period (i.e., the first two hours) of each data collection interval would be used in the analysis, as the factor under investigation was detection capability, not the testing of a specific communications medium. This limitation reduced the amount of collected Microloop 702 data that was available for analysis. One potential solution for future studies would be to collect contact closure data through a traffic counter/classifier such as a PEEK ADR instead of a data collection computer.

Data were collected over 48 days during the months of June, July, and August of 2011. These 48 data collection days are shown in table 4.1. However, due to the extensive manual labor requirements of data reduction, only a subsample of this data set was ultimately included in the analysis. The analyzed data set included 1,467 minutes (slightly more than 24 hours) of data. Intervals were chosen for this analysis data set so that various combinations of environmental factors such as lighting, precipitation, and traffic volume would be represented. Table 4.2 outlines the data intervals that were ultimately included in the analyzed data set.

Table 4.1 Data Collection Dates

6/7/2011	6/19/2011	7/5/2011	7/30/2011
6/8/2011	6/20/2011	7/6/2011	8/2/2011
6/9/2011	6/21/2011	7/7/2011	8/3/2011
6/10/2011	6/22/2011	7/8/2011	8/4/2011
6/11/2011	6/25/2011	7/11/2011	8/5/2011
6/12/2011	6/26/2011	7/12/2011	8/6/2011
6/13/2011	6/27/2011	7/13/2011	8/12/2011
6/14/2011	6/30/2011	7/14/2011	8/15/2011
6/15/2011	7/1/2011	7/21/2011	8/16/2011
6/16/2011	7/2/2011	7/22/2011	8/18/2011
6/17/2011	7/3/2011	7/28/2011	8/19/2011
6/18/2011	7/4/2011	7/29/2011	8/30/2011

Table 4.2 Data Intervals Included in Analysis

Date	Time	Date	Time
6/9/2011	8:04 - 9:59	6/20/2011	21:22 - 21:50
6/9/2011	10:01 - 11:19	6/20/2011	22:14 - 22:19
6/9/2011	11:36 - 12:39	6/20/2011	22:41 - 23:58
6/9/2011	22:50 - 23:58	6/25/2011	6:23 - 11:08
6/20/2011	16:01 - 16:42	6/25/2011	11:10 - 11:25
6/20/2011	16:44 - 17:11	7/6/2011	5:28 - 6:08
6/20/2011	17:13 - 17:15	7/6/2011	15:36 - 15:40
6/20/2011	17:17 - 17:22	7/6/2011	15:42 - 16:59
6/20/2011	17:24 - 17:39	7/6/2011	17:27 - 21:30
6/20/2011	17:41 - 18:27	7/28/2011	5:30 - 6:29
6/20/2011	18:31 - 21:04		

4.2 Data Reduction

The data reduction procedure for this study involved two distinct steps. The first step was the derivation of ground truth vehicle time stamps and length-based classifications from video of the traffic stream. The second step was the compilation of ground truth data and data from the various detectors at the test bed into a consolidated data set.

4.2.1 Step 1: Ground Truth

The derivation of the ground truth data from video of the traffic stream was a laborious task because it had to be done manually (i.e., based on video observations). To facilitate the task, a series of macros (i.e., customized programs) were written for implementation in Microsoft Excel. With these macros and the input of video start time and playback speed, it was possible to correlate various keystrokes to a vehicle passage timestamp. Nine different keys were assigned to represent each combination of three vehicle classes and three traffic lanes. The user would watch the video at a particular location, and every time the front bumper of a vehicle reached this location, the user entered the appropriate key stroke. For example, *Ctrl+r* indicated a long vehicle in lane 1 (the westbound lane nearest to the shoulder). The final result was an output file that contained vehicle timestamps, traveled lanes, and classifications (see table 4.3).

Table 4.3 Ground Truth Output Sample

Timestamp	Lane	Class
5:29:31	2	1
5:29:34	2	1
5:29:37	1	1
5:30:00	2	1
5:30:18	3	1
5:30:19	2	2
5:30:29	2	1
5:30:44	2	1
5:30:52	1	2
5:31:13	2	1

As a quality control measure, ground truth data were reduced by two users independently for 30 minutes (2% of the final data set). Comparison of the observations recorded by the two users revealed a volume agreement of 99.9%. Additionally, there was agreement on the lane assignment for 99.1% of the vehicles and agreement on the classification of 96.5% of the vehicles. Length-based classification ground truth was more susceptible to human error than volume or lane assignment due to the subjectivity of interpreting a vehicle's length from the video.

4.2.2 Step 2: Data Compilation

The compilation of the data into a consolidated data set was also accomplished through macros implemented in Microsoft Excel. A separate macro was required for each detector technology because each had a unique data file. These files were retrieved from the test bed, as outlined in section 4.1. The output files from the Microloop 702 detectors were XML-formatted while the other detectors provided various types of delimited text files. A unique macro was written to parse the output files from each detector into similar Excel worksheets. While the data files from each technology included various parameters, each included per-vehicle timestamps, speeds, and either lengths or length-based classifications. Once clock synchronization was performed (as discussed in section 4.2.3), the data in these worksheets was formatted for per-vehicle analysis. At this point, information regarding the environmental factors under consideration was incorporated into the data worksheets using another macro. This consolidated data file was saved and the data in this file was also aggregated for one, five, and fifteen minute aggregation files. These per-vehicle and aggregate files were converted to comma delimited tables, which were imported into the R software environment for statistical analysis.

4.2.3 Clock Synchronization

A major issue that must be considered when collecting time-stamped data from multiple sources with independent internal clocks is clock drift. Clock drift occurs when the internal clocks of two or more different devices deviate relative to one another over the passage of time. If clock drift does occur, clock synchronization is required so that the error associated with this drift is reduced or eliminated. The clock synchronization process utilized for this thesis is shown in figure 4.1, and described in the following paragraphs.

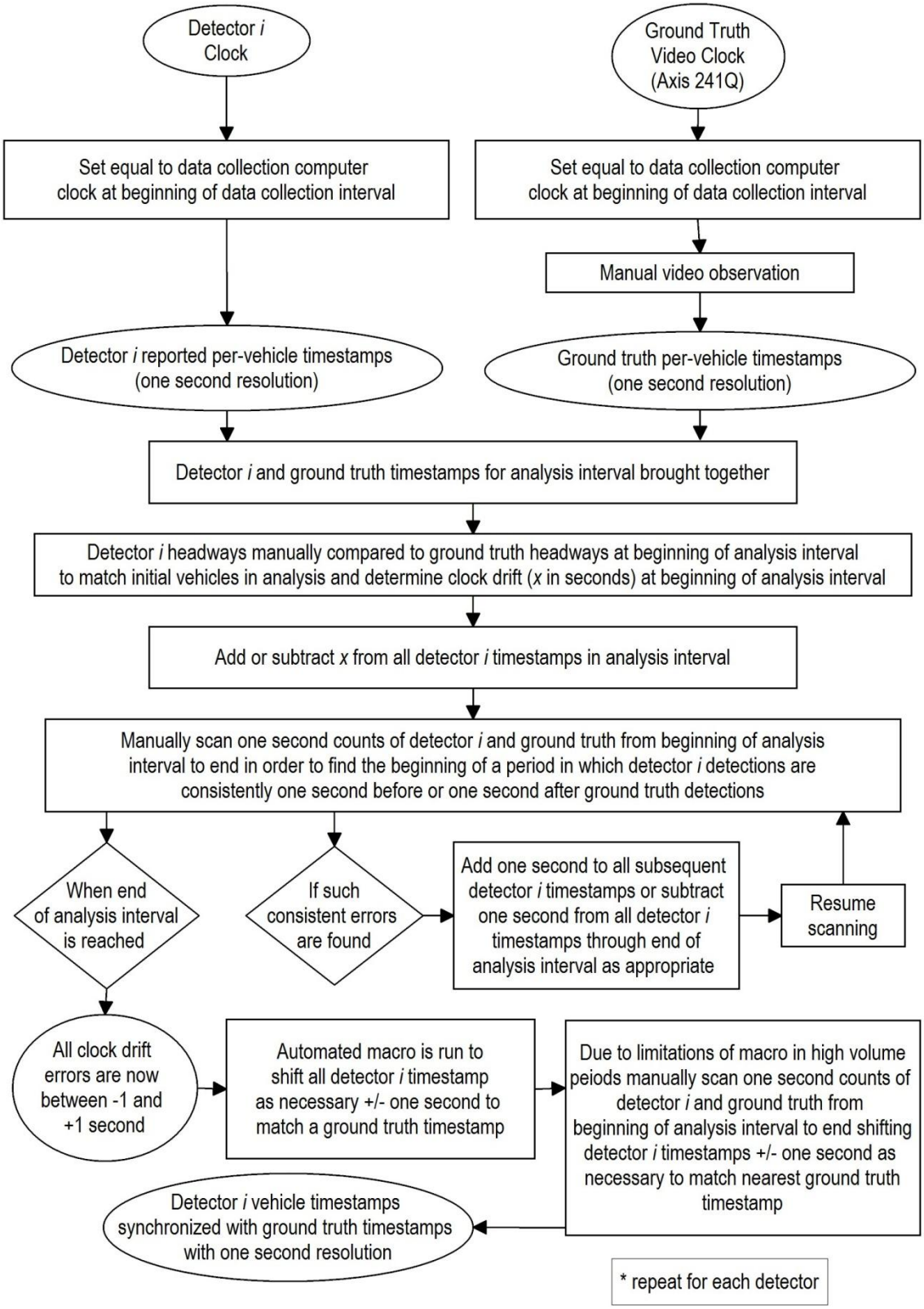

Figure 4.1 Clock Synchronization Flow Chart

The standard methodology in these situations is to establish a reference clock. For this study, the reference clock was the time stamp on the ground truth video, which was taken from the internal clock of the Axis 241Q video server. Before each data collection period commenced, this clock was set equal to the time on the NTC data collection computer using the Axis Camera Station Client software. Using the Autoscope Network Browser, ITS Link, WinRTMS4, and SmartSensor Manager software tools, it was then possible to set each of the detectors' internal times equal to the NTC data collection computer's internal time. This approach gave all data sources an equal starting point at the beginning of each data collection interval (most of which lasted less than 24 hours).

From this common starting point, clock drift throughout the data collection interval was relatively small. Based on analysis of clock drift in the data set under analysis, it was found that clock drift with respect to the Axis 241Q video server clock never exceeded 10 seconds per 24 hours for any of the detectors under evaluation. While the SmartSensor Manager software had a tool to automatically synchronize the SmartSensor 105 time with the NTC data collection computer time at regular intervals, the software tools for the other detectors and the Axis 241Q video server did not have this capability. Therefore, compensation for this clock drift within a data collection interval was made during the data reduction stage.

This compensation during data reduction involved both manual and automated procedures. The first manual procedure involved in this process was to observe one-second per-lane counts from each source after the source data files were aggregated into a common Microsoft Excel workbook. These counts were observed in the format shown in table 4.4. When any detector's data were observed to consistently deviate from the ground

truth by more than one second in either direction, an adjustment factor of one second was either added to or subtracted from all subsequent detections from that detector, and the manual analysis continued until the end of the data set. This approach is best demonstrated by example:

Table 4.4 Sample Count Aggregation Before (a) and After (b) Manual Time Shift

	Video	SP II	ML 702	G4	SS 105
8:05:00					
8:05:01					
8:05:02	1	1	1	1	1
8:05:03					
8:05:04	1	1		1	1
8:05:05			1		
8:05:06					
8:05:07					
8:05:08	1	1			1
8:05:09			1		
8:05:10					

(a)

	Video	SP II	ML 702	G4	SS 105
8:05:00					
8:05:01					
8:05:02	1	1	1	1	1
8:05:03					
8:05:04	1		1	1	1
8:05:05		1			
8:05:06					
8:05:07					
8:05:08	1	1	1		1
8:05:09					
8:05:10					

(b)

The hypothetical example in table 4.4 shows one second being subtracted from all Microloop 702 timestamps after 8:05:03. At this point all clock drift had been reduced to ±1 second from the ground truth (video) timestamp. An Excel macro was written that was able to shift times by ±1 second to match the timestamp of the nearest ground truth detection not already correlated to a matched detection from the given detector. A flow chart (figure 4.2) demonstrates the logical process used by this macro to automate this portion of the clock synchronization. As a reference for the algorithm applied by this macro, the code is included in Appendix B. A sample data interval is shown in table 4.5 before and after running this macro.

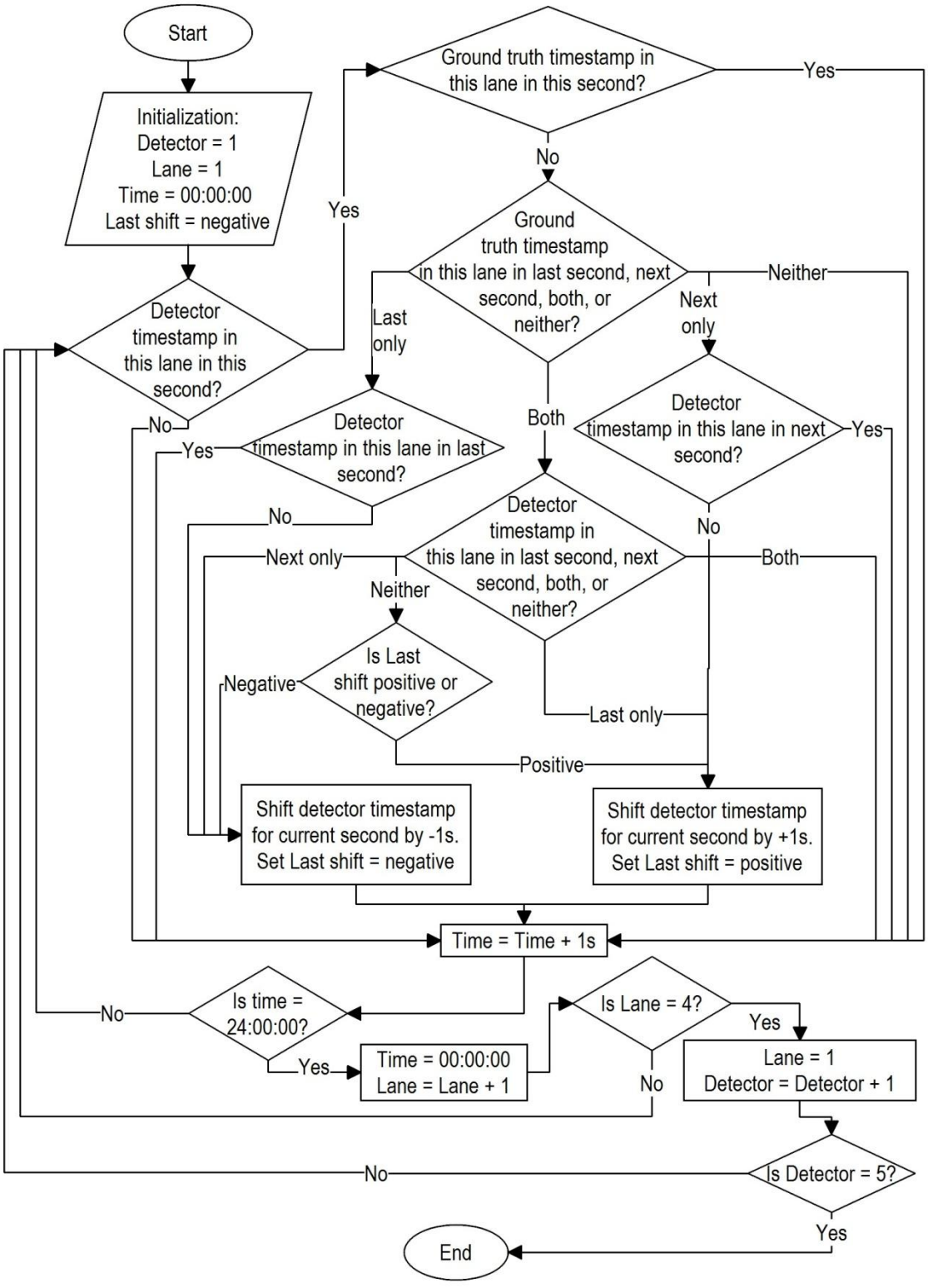

Figure 4.2 Clock Synchronization Macro Flow Chart

Table 4.5 Sample Count Aggregation Before (a) and After (b) Automated Macro Time Shift

	Video	SP II	ML 702	G4	SS 105
8:05:00					
8:05:01					
8:05:02	1	1	1	1	1
8:05:03					
8:05:04	1		1	1	1
8:05:05		1			
8:05:06					
8:05:07					
8:05:08	1	1	1		1
8:05:09					
8:05:10					

(a)

	Video	SP II	ML 702	G4	SS 105
8:05:00					
8:05:01					
8:05:02	1	1	1	1	1
8:05:03					
8:05:04	1	1	1	1	1
8:05:05					
8:05:06					
8:05:07					
8:05:08	1	1	1		1
8:05:09					
8:05:10					

(b)

While this macro functioned appropriately in low-volume times, it was difficult to develop a function that could assess matched detections during high volume periods, such as the high volume period shown in table 4.6(a). Therefore, the final procedure in clock synchronization was to manually shift detection times in these high volume periods (as seen in table 4.6(b) with the SP II or Solo Pro II detector).

Table 4.6 Sample High Volume Count Aggregation Before (a) and After (b) Second Manual Time Shift

	Video	SP II	ML 702	G4	SS 105
17:15:00					
17:15:01	1		1		
17:15:02		1			
17:15:03	1	1	1	1	1
17:15:04					
17:15:05	1	1	1	1	1
17:15:06	1		1		
17:15:07	1		1	1	
17:15:08		1			
17:15:09	1	2	1	1	1
17:15:10	1		1	1	1
17:15:11					

(a)

	Video	SP II	ML 702	G4	SS 105
17:15:00					
17:15:01	1	1	1		
17:15:02					
17:15:03	1	1	1	1	1
17:15:04					
17:15:05	1	1	1	1	1
17:15:06	1		1		
17:15:07	1	1	1	1	
17:15:08					
17:15:09	1	1	1	1	1
17:15:10	1	1	1	1	1
17:15:11					

(b)

The result of this clock synchronization process was that the detector-reported timestamps were shifted as necessary to minimize clock drift. The resulting timestamps had a resolution of one second, which was adequate for the analysis to be performed in this thesis. Once clock synchronization was completed, data compilation was resumed as outlined in section 4.2.2 which ultimately resulted in a data set for statistical analysis.

4.3 Chapter Summary

The data collection and reduction process outlined in this chapter have provided an overview of the processes employed in gathering the appropriate data from the evaluated detectors and formatting it for analysis. Once detectors had been appropriately calibrated, as was outlined in the previous chapter, data collection was relatively simple. The only significant issues involving data collection were related to communications between the detectors and the data collection computer. These issues were overcome through frequent monitoring during data collection. Data reduction was facilitated through the development of Excel macros, but remained a labor-intensive task. The manual

derivation of ground truth data from recorded video was a limiting factor in the size of the analyzed data set. The primary difficulty encountered during data reduction was the need to account for clock drift in the various detectors. This was accomplished through manual and automated procedures. The data collection and reduction endeavors resulted in tabulated detection data for statistical analysis.

CHAPTER 5 STATISTICAL METHODS

A number of statistical methods were used in the data analysis for this thesis. Explanations of the various methods are presented in this chapter, so as to avoid muddling the presentation of the results in the chapters with interspersed theory. These statistics are applied in the following analysis chapters.

5.1 Simple Statistics

Throughout the analyses of the four traffic detectors under examination in this study, a number of simple statistical methods were applied in order to define their accuracies and the distributions of values they reported.

5.1.1 Mean Percent Error

The mean percent error (MPE) is a simple statistic that provides the arithmetic mean of the deviations of detected values from ground truth values, scaled as a percentage of the ground truth value. When no ground truth was available, a baseline was selected, and the statistic was referred to as the mean percent difference (MPD), instead of MPE. The MPE was defined according to the following equation:

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - y_i}{y_i} * 100\right) \tag{5.1}$$

where: n is the number of observations,

x_i is the detector reported value for observation i, and

y_i is the ground truth value for observation i.

The MPE is negative when the mean detector-reported value is less than the mean ground truth value, and positive when the mean detector reported value is greater than the mean ground truth value. While the MPE is useful for determining the direction and magnitude

of a detector's bias for a specified parameter, its weakness is that percent errors with opposite signs have a balancing effect.

5.1.2 Mean Absolute Percent Error

The mean absolute percent error (MAPE) is a statistic that accounts for the balancing effect of positive and negative percent errors, which is problematic in MPE. It represents the arithmetic mean of the absolute values of deviations of detected values from ground truth values, scaled as a percentage of the ground truth value. When no ground truth is available, a baseline is selected and the statistic is referred to as the mean absolute percent difference (MAPD), instead of MAPE. The MAPE is defined according to the following equation:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|x_i - y_i|}{y_i} * 100\right) \tag{5.2}$$

where: n is the number of observations,

x_i is the detector reported value for observation i, and

y_i is the ground truth value for observation i.

5.1.3 Correlation Coefficient

Another simple statistic is the correlation coefficient. The correlation coefficient (r) indicates the strength of a linear relationship between two variables, and is calculated according to the following equation:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{5.3}$$

where: n is the number of observations,

x_i is the detector reported value for observation i,

$\bar{x}$ is the mean of detector reported values for all observations,

y_i is the true value for observation i, and

$\bar{y}$ is the mean of true values for all observations.

The correlation coefficient is on the range $-1 \leq r \leq 1$. A value near 1 indicates a strong positive linear relationship between *x* and *y* while a value near -1 indicates a strong inverse linear relationship, and a value near 0 indicates a weak linear relationship.

5.2 Skewness and Kurtosis

Skewness and kurtosis are two statistics that describe the distribution of a set of values. For example, the distributions from various detectors of observed speeds over a given time period will each have a skewness and kurtosis. Specifically, skewness and kurtosis are the third and fourth standardized moments of the distribution. Skewness is a measure of the asymmetry of a distribution. A negative skew indicates that the left tail is longer and that the bulk of the values are greater than the mean. A positive skew indicates that the right tail is longer and that the bulk of the values are less than the mean. A symmetrical distribution will have zero skewness. The magnitude of the skewness can be interpreted as a measure of asymmetry. The skewness of a sample is calculated according to the following equation:

$$g_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\right)^{3/2}} \tag{5.4}$$

where: n is the number of observations,

x_i is the value for observation i, and

$\bar{x}$ is the mean of the values for all observations.

Kurtosis is a measure of the peakedness of a distribution. A platykurtic distribution has a kurtosis of less than three and is characterized by broad peaks and thin tails. A leptokurtic distribution has a kurtosis of greater than three and is characterized by a slender peak and fatter tails. Lastly, a mesokurtic distribution has a kurtosis of exactly three. All normal distributions are mesokurtic regardless of their parameters. As it applied to analysis of pre-vehicle speed detection in this thesis, kurtosis provided a measure of sensitivity to differences in speed. The kurtosis of a sample as defined here is calculated according to the following equation:

$$b_2 = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\right)^2} \tag{5.5}$$

where: n is the number of observations,

x_i is the value for observation i, and

$\bar{x}$ is the mean of the values for all observations.

This definition of kurtosis is not to be confused with the kurtosis excess ($g_2 = b_2 - 3$), such that a normal distribution has a kurtosis excess of zero.

An example demonstrating the interpretation of skewness and kurtosis is given based on pre-vehicle speed data taken from the NTC/NDOR non-intrusive detector test bed between 4:35 PM and 5:35 PM on June 20th, 2011. Histograms of the distributions of

speeds from the four detectors under evaluation in this thesis for this time period are given in figure 5.1, along with the skewness and kurtosis of each distribution.

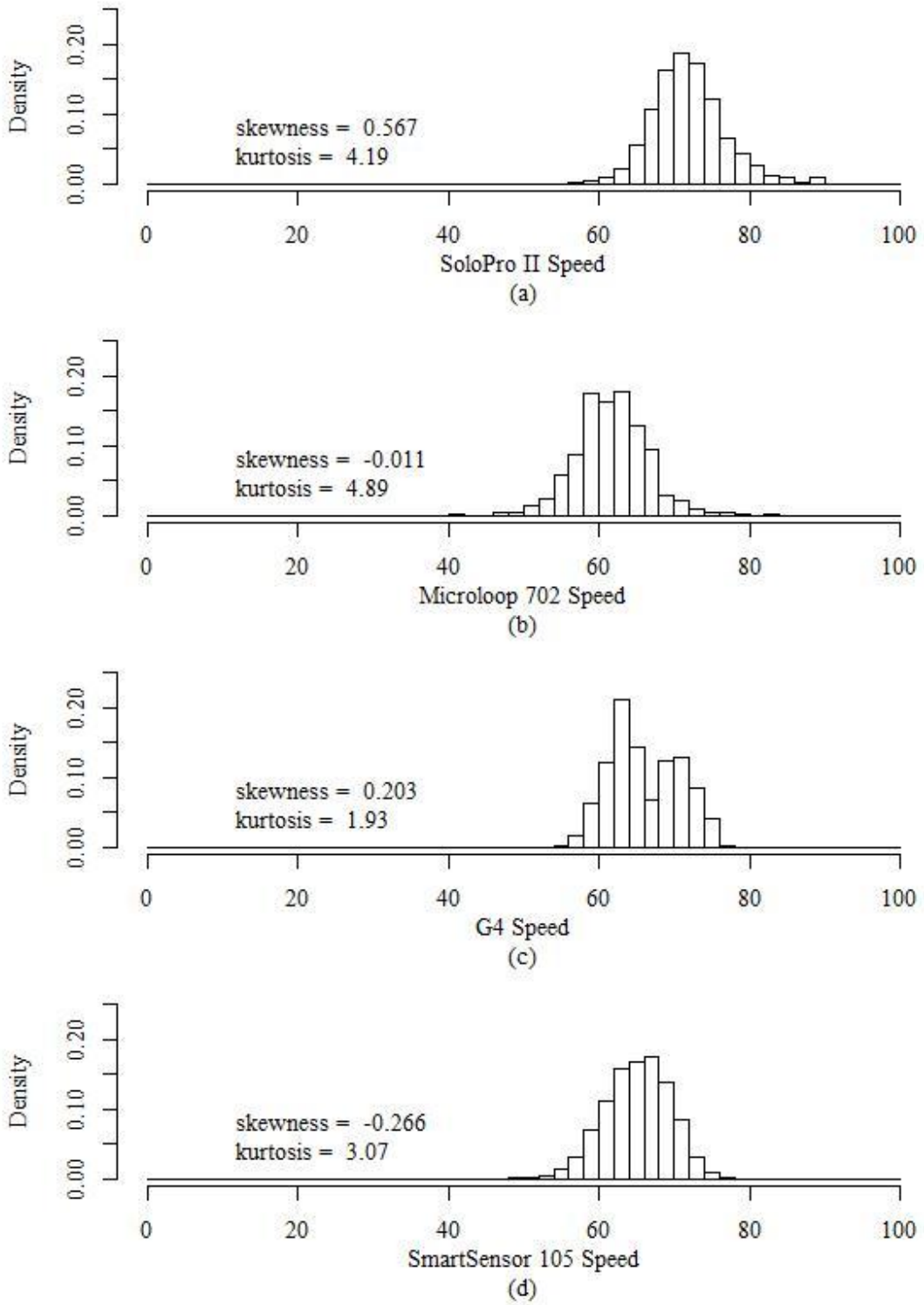

Figure 5.1: Small Sample Histograms of Per-Vehicle Speed Distributions for the Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

All of the distributions in figure 5.1 have skewness values between -1 and 1, indicating relatively balanced distributions. However, the distribution of speeds from the Solo Pro II, which had the largest positive skewness value, can be seen to have a longer right tail than left tail. Also, the distribution of speeds from the SmartSensor 105, which had the most negative skew, appears to have a slightly longer left tail than right tail. Regarding kurtosis, the SmartSensor 105 distribution, which had a kurtosis of 3.07 (nearly mesokurtic), has a distribution with a peakedness similar to a normal distribution. It can also be seen that the two significantly leptokurtic distributions (those with kurtoses significantly greater than three), compared to the other distributions, are characterized by having long tails and slender peaks. Lastly, the G4 with a platykurtic distribution (kurtosis of 1.93) has a broad peak and nearly non-existent tails.

5.3 GEH Statistic

The GEH statistic is a self-weighting test statistic used in assessment of traffic volume estimates, which has most frequently been applied to validate traffic microsimulation models (52). The self-weighting characteristic, which makes it appealing for microsimulation model validation, also made it appropriate for analysis of traffic volume detection accuracy. The GEH statistic for time period i is calculated according to the following equation:

$$GEH_i = \sqrt{\frac{2(v_{di} - v_{ri})^2}{v_{di} + v_{ri}}} \tag{5.6}$$

where: v_{di} is the detector reported volume for time period i and

v_{ri} is the reference volume for time period i.

A GEH statistic of 0 indicates perfect detection for the given time period, while higher values indicate more severe errors.

In detection accuracy, a large percent error at a low volume should not necessarily receive the same weight as a large percent error at a high volume. By self-weighting, the GEH statistic assigns greater weights to errors at high volumes than errors at low volumes. The following example demonstrates the applicability of the GEH statistic. Consider a hypothetical example with two time intervals: (a) 7 of 8 vehicles detected and (b) 70 of 80 vehicles detected. The percent errors for the two intervals (12.5% for each) suggest equal performance in both intervals. The absolute errors for the two intervals (1 and 10 missed vehicles, respectively) suggest that the detector performance was far worse for the high volume interval (b). The GEH statistics for the two intervals (0.37 and 1.15) suggest a more significant error during the high volume interval, without suggesting that the error was 10 times as bad as the low volume interval, as was suggested by the absolute error.

5.4 Theil's Inequality Coefficient

In measuring the difference between detected and true values, or detected values from two detection sources, it is useful to have a numerical representation of the degree of agreement, or inversely the degree of inequality, between the two sets of values. While statistics such as a correlation coefficient, mean percent error, and mean absolute percent error are useful for this purpose, they do not in themselves convey information about the nature of the differences between two sets of data. Theil's inequality coefficient provides a similar metric that can be deconstructed in such a way as to indicate the nature of the differences between two sets of data (53). Originally, this inequality coefficient was

developed to measure the goodness-of-fit of economic forecasts, but was recently introduced into the traffic engineering field for validation of microscopic simulation models (54). Because this validation of simulated data with respect to observed values is similar to the current application of validating observed data with respect to ground truth, Theil's inequality coefficient was included in this thesis. Theil's inequality coefficient (*U*) is defined by the following equation:

$$U = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}} \tag{5.7}$$

where: n is the number of observations,

y_i is the ground truth value for observation i, and

x_i is the detector reported value for observation i.

The numerator of this equation is the root mean square error, and the denominator scales *U* such that it will always lie on the range $0 \leq U \leq 1$. If $U = 0$, the detector reported values are equal to the true values for all observations. If $U = 1$, the detection performance is as bad as possible. The mean square error, as seen in the numerator of the above equation can be deconstructed into three components, as shown in the following equation:

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2 = \left(\bar{y} - \bar{x}\right)^2 + \left(\sigma_x - \sigma_y\right)^2 + 2\left(1 - r\right)\sigma_y\sigma_x \tag{5.8}$$

where: $\bar{y}$ is the mean true value of all observations,

$\bar{x}$ is the mean detector reported value of all observations,

σ_x is the standard deviation of detector reported values for all observations,

σ_y is the standard deviation of true values for all observations, and

r is the correlation coefficient of detector reported and true values.

When these three components are each divided by their sum, as shown in the following equations, they become proportions such that $U_m + U_s + U_c = 1$.

$$U_m = \frac{\left(\bar{y} - \bar{x}\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2} \tag{5.9}$$

$$U_s = \frac{\left(\sigma_x - \sigma_y\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2} \tag{5.10}$$

$$U_c = \frac{2\left(1 - r\right)\sigma_y \sigma_x}{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2} \tag{5.11}$$

where: U_m is the bias proportion, which indicates the proportion of the inequality that can be contributed to a systematic tendency toward over- or under-estimation of the true value (a small value of U_m indicates good detector calibration);

U_s is the variance proportion, which indicates the proportion of the inequality that can be attributed to unequal variances between the detector's reported values and true values (a large U_s indicates

that the variance of the detected values is significantly different from the variance of the true values); and

U_c is the covariance proportion, which indicates the proportion of the inequality that is unsystematic (ideally U_c should represent the largest proportion of the inequality).

5.5 Analysis of Variance

Analysis of variance (ANOVA) is utilized numerous times throughout the following chapters in order to determine which factors significantly affect a given detector's ability to correctly detect a specific parameter, such as volume, speed, or vehicle classification. In these analyses, there were two factors (lighting and rain) with four and two levels, respectively (day, night, dawn, dusk; and clear, rain). Therefore, the model chosen was the following four-by-two factorial ANOVA:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk} \tag{5.12}$$

where:

$i = 1, \ldots, a$

$j = 1, \ldots, b$

$k = 1, \ldots, n$

μ is the overall mean for all y,

α_i is the effect of the ith level of factor A,

β_j is the effect of the jth level of factor B,

$\alpha\beta_{ij}$ is the interaction effect between the ith level of factor A and the jth level of factor B, and

ε_{ijk} is the random effect or error term.

Each use of ANOVA in the following chapters began with this model, having different y variables, such as volume percent error or classification error percentage. However, in models that found the interaction term ($\alpha\beta_{ij}$) to not be statistically significant, this term was eliminated from the model in order to increase the power of the analysis for significance of the independent factors of lighting and rain.

These analyses were conducted on an unbalanced sample, meaning that there were different numbers of sample points in different combinations of the levels of factors for lighting and rain. Due to this fact, type III (marginal) sums of squares were used. Type III sums of squares are calculated in such a way that the sum of squares for each factor is calculated given the effects of all other factors. It is the only type of sums of squares that does not convolute the hypotheses being tested to be about the order in which factors are added to the model or number of sample points in each cell. Rather, the hypothesis tested by this ANOVA with type III sums of squares is whether the effect of a factor, given all other factors, is statistically significant (55).

There are three basic assumptions for the ANOVA model. The first of these assumptions is normality of random effect (ε_{ijk}). While this assumption was not strictly met by the majority of the models in the following chapters, ANOVA can be appropriate in some instances where this assumption is not met. One statistical text states that "for large samples, more radical departures are acceptable since the central limit theorem comes into play" (56). The sample sizes for the analyses in this report were of a magnitude which made this qualification applicable. The next assumption is independence of the random effect (ε_{ijk}). In time series data such as that used in this analysis, autocorrelation (a lack of independence) can be an issue. For that reason, each

of the ANOVA models in the following analyses were tested for autocorrelation and thinned appropriately to eliminate autocorrelation and meet the assumption of independence. An example of this thinning procedure is given in Appendix B. The final assumption for ANOVA is homoscedasticity (homogeneity of variances). The tests for this assumption are sensitive to non-normality (Bartlett's test) or unequal sample sizes (Hartley's & Cochran's tests), which made them inappropriate for these data. Also, the F tests (which underlie ANOVA) are robust with respect to departures from homogeneity (56). Therefore, while this third assumption was not checked, there was a great deal of confidence in the ANOVA models employed in this study.

5.6 Multiple Regression Model

A series of multiple regression models were used throughout the data analysis in this study. The general form of these models is given by the following equation:

$$p_i = \alpha + \beta_1 X_{i1} + \gamma_{11} D_{i11} + \gamma_{12} D_{i12} + \gamma_{13} D_{i13} + \gamma_{21} D_{i21} + \gamma_{31}\left(D_{i11} * D_{i21}\right) + \gamma_{32}\left(D_{i12} * D_{i21}\right) + \gamma_{33}\left(D_{i13} * D_{i21}\right) + \varepsilon_i \quad (5.13)$$

where:

p_i is the predicted percent error or deviation for a single given time period i,

α is the theoretical mean percent error or deviation for the specified detector given daylight non-rainy conditions with true volume of 0 vehicles,

β_1 is the coefficient for the average effect of one more vehicle in the true volume,

X_{i1} is the true volume for time period i,

$\gamma_{11}$is the coefficient for the average effect of night lighting conditions,

D_{i11} is a dummy variable taking the value of *1* during night lighting conditions,

γ_{12} is the coefficient for the average effect of dawn lighting conditions,

D_{i12} is a dummy variable taking the value of *1* during dawn lighting conditions,

γ_{13} is the coefficient for the average effect of dusk lighting conditions,

D_{i13} is a dummy variable taking the value of *1* during dusk lighting conditions,

γ_{21} is the coefficient for the average effect of rainy conditions,

D_{i21} is a dummy variable taking the value of *1* during rainy conditions,

γ_{31} is the coefficient for the average interaction effect of night and rainy conditions,

γ_{32} is the coefficient for the average interaction effect of dawn and rainy conditions,

γ_{33} is the coefficient for the average interaction effect of dusk and rainy conditions, and

ε_i is the residual error for time period i.

The definitions of the dependent variable were specific to the various applications of the model, and were therefore given with each application of the model in the following chapters.

Regression analysis also posits a number of assumptions that must hold in order for the model to be valid. The first of these assumptions is independence of the residual error. As with ANOVA, this assumption was met through appropriate thinning of the data

in the manner demonstrated in Appendix B. Another assumption is that there is no multicollinearity in the predictors. This condition was the reason that certain variables were not explicitly included in the model. For example, the day lighting condition was not explicitly included in the model. Instead, it is implied when the dummy variables for night, dawn, and dusk were all 0. In the same way, clear weather was not explicitly stated, but rather, was implied when the dummy variable for rain was 0. Homoscedasticity is also assumed for linear regression, but was not confirmed for this analysis. Lastly, it is assumed that the independent variables are measured without error. This assumption was met through the experimental design.

5.7 Chapter Summary

The preceding chapter defined the statistics used in the following chapters to analyze the data and draw appropriate conclusions. The analyses in this thesis begin with elementary statistics, such as mean percent error, mean absolute percent error, and correlation coefficients, which have been used in many of the previous detector evaluation studies documented in the literature review. There are also a number of graphical representations of the data, which are enhanced by descriptive statistics such as skewness and kurtosis. This analysis attempted to go one step further by introducing statistics borrowed from other specializations within transportation systems engineering, such as the GEH statistic and Theil's inequality coefficient. Lastly, established statistical models such as ANOVA and regression were applied to test hypotheses regarding the statistical significance of environmental factors on the accuracies of various traffic detectors.

CHAPTER 6 AGGREGATE ANALYSIS AND RESULTS

This thesis compared the relative accuracy of reported traffic parameters from particular detector technologies under various conditions. The following analysis employed a variety of graphical representations and statistical tests in order to convey the strengths and weaknesses of the various detection technologies. The analysis was divided principally between aggregate analysis of one-minute, five-minute, and fifteen-minute interval data, and disaggregate, per-vehicle analysis. This chapter focuses on the aggregate analysis while chapter 7 covers disaggregate per-vehicle analysis.

This aggregate analysis was based on vehicle detections in the 1467 minute (24-hour) analysis data set defined in section 4.1. In this data set there were a total of 36,124 time-stamped ground truth vehicle presence detections with associated vehicle classification. The data set also included time-stamped detector-reported vehicle detections with individual speeds and vehicle classifications from each of the four analyzed detection systems. These detections were aggregated over one-minute, five-minute, and fifteen-minute intervals to obtain interval volumes, interval average speeds, and interval classification proportions. Additionally, lighting, precipitation conditions, and traffic volume were noted for each minute so that potential effects of these factors on the performance of the various detector technologies could be determined.

When traffic volume was considered as a factor in this analysis, each one-minute period was classified as either a low-volume or high volume period. Low volume periods were defined as periods when the traffic stream had a level of service of A or B (i.e., one-minute periods during which the three-lane passenger car equivalency did not exceed 54). High-volume periods were characterized by a level of service of C or D (i.e., one-minute

periods during which the three-lane passenger car equivalency exceeded 54). Because the worst level of service observed in the data set was LOS D, it was inappropriate to extrapolate this analysis to conditions representing level of service E or F.

6.1 One-Minute Aggregation Interval Analysis

The focus of this section is the one-minute interval data analysis performed on volume, speed, and vehicle classification.

6.1.1 One-Minute Volume Analysis

The analysis of volume begins with simple graphics comparing the reported one-minute volumes from each detector with the ground truth one-minute volumes obtained by manual observation of video. Figure 6.1 shows detector-reported one-minute volume versus ground truth one-minute volume for each detector. While the Solo Pro II, Microloop 702, and G4 one-minute volumes all appeared to have strong linear relationships with the ground truth volume, figure 6.1(d) shows that the SmartSensor 105 tended to under-report volume when the ground truth volumes were high (e.g. greater than 40 veh/min). This led to a correlation coefficient (r) of 0.91 for the SmartSensor 105, lower than the correlation coefficients of the other detectors, which were all greater than 0.99.

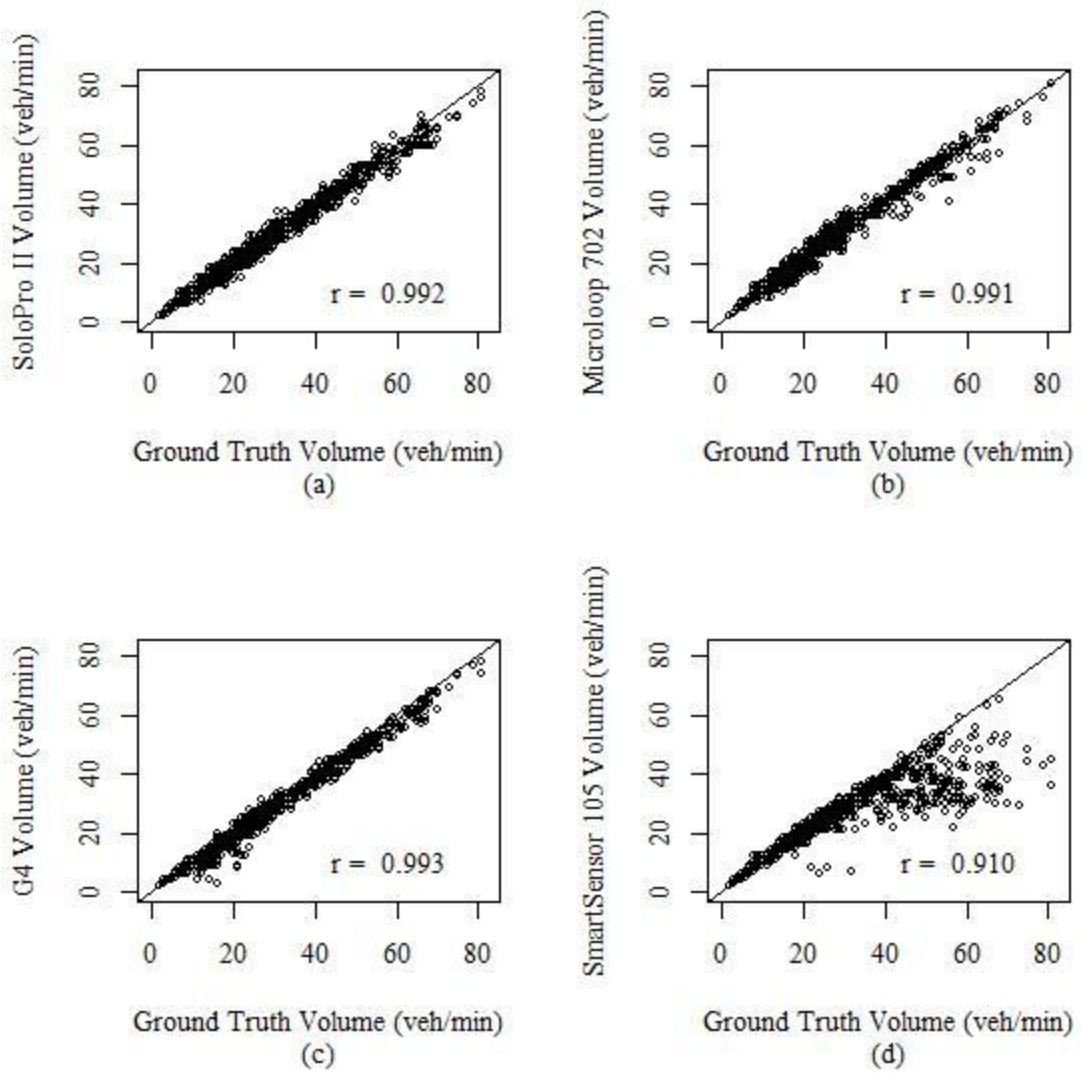

Figure 6.1: One-Minute Volume Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

Figure 6.2 shows box plots for each detector's reported volume. Again, this figure shows that the SmartSensor 105 did not report as many high volumes (60+ vehicles per minute) as the other detectors and the ground truth. When comparing the 75th percentile one-minute volumes (i.e., the upper boundaries of the inter-quartile ranges in figure 6.2) of the detectors with that of the ground truth volumes, the relatively lower 75th percentile values from the Solo Pro II, G4, and SmartSensor 105 may indicate a tendency to under-

report one-minute volume, while the relatively higher 75th percentile value from the Microloop 702 may indicate a tendency to over-report one-minute volume.

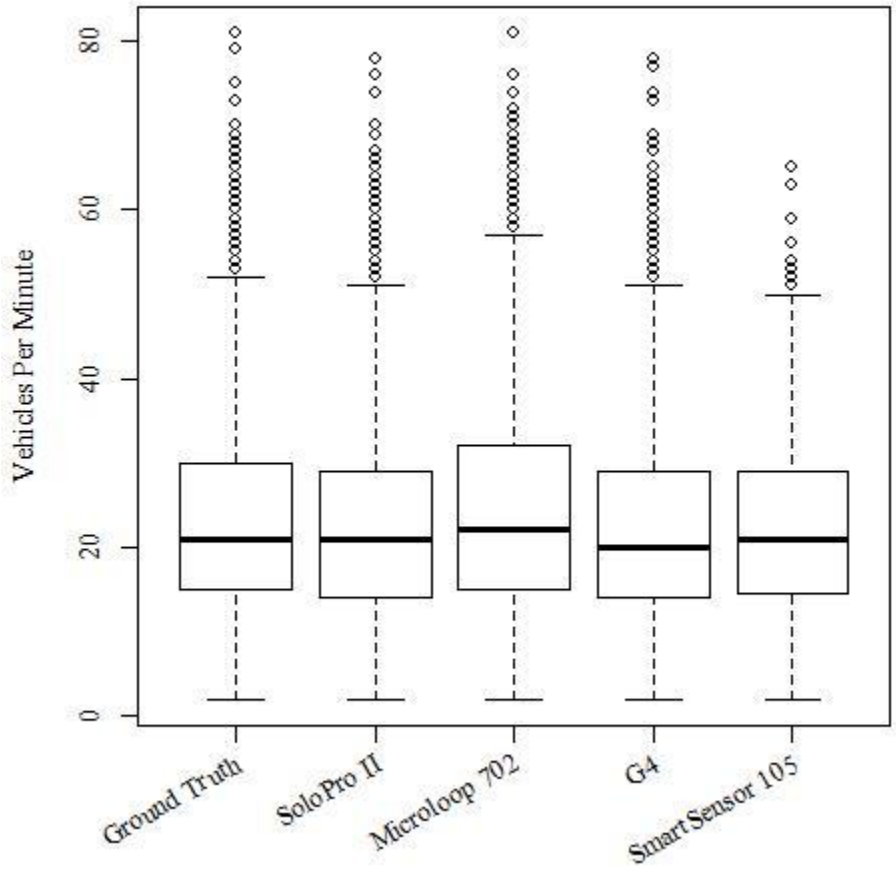

Figure 6.2: Box Plot of Reported One-Minute Volumes

The histograms in figure 6.3 again show that the SmartSensor 105 was missing the extreme upper tail of the ground truth and other detectors. This is quantified in the values of skewness and kurtosis given along with the histograms. The skewness of the ground truth distribution of one-minute volumes (1.190) was relatively high because of the impact of the long right tail of the distribution. While the Solo Pro II, Microloop 702, and G4 one-minute volume distribution skewnesses was similar to the ground truth, the SmartSensor 105 had a lower value of skewness (0.660).

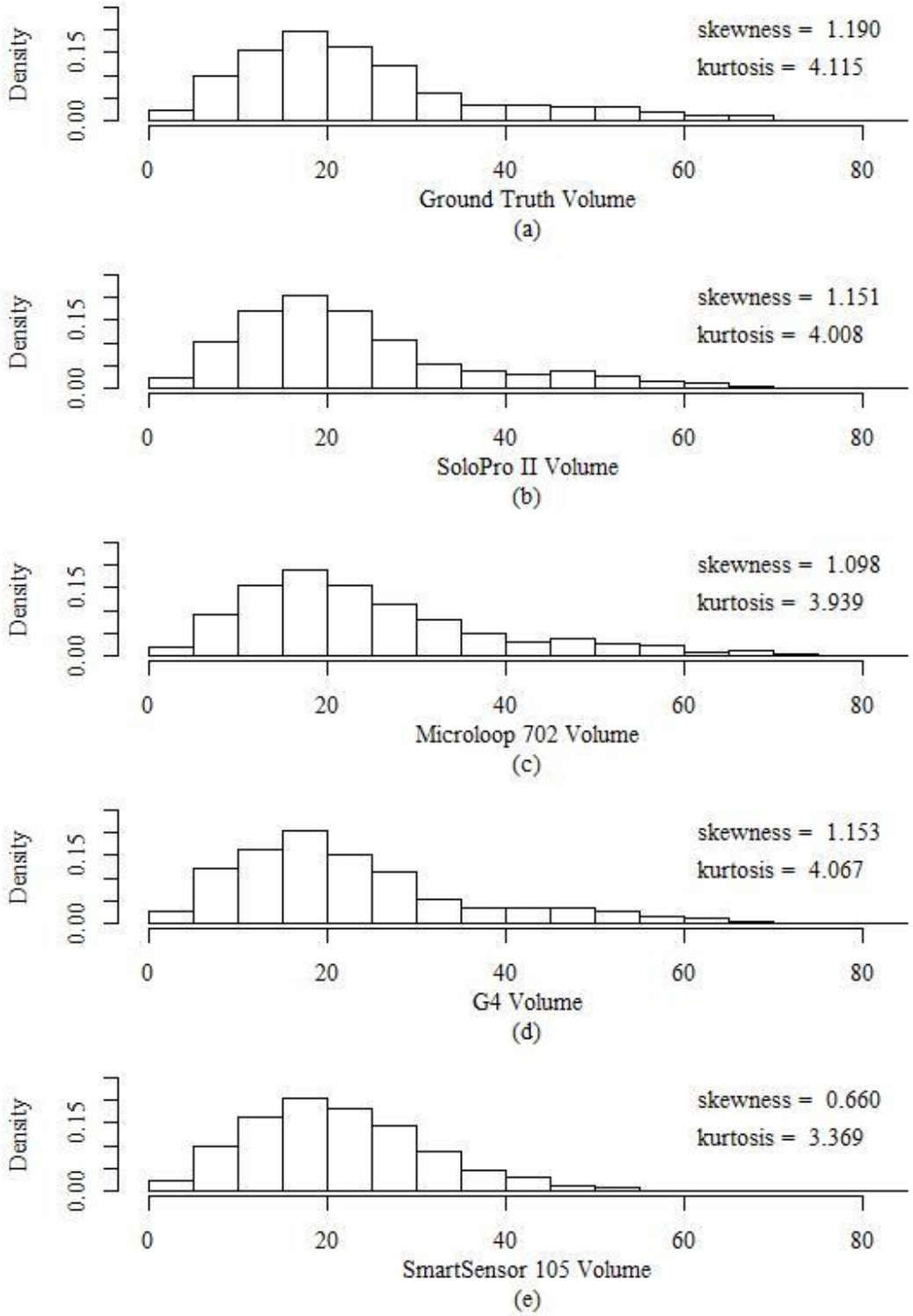

Figure 6.3: Histograms of One-Minute Volume Distributions for Ground Truth (a), Solo Pro II (b), Microloop 702 (c), G4 (d), and SmartSensor 105 (e)

Figure 6.4 gives a cumulative distribution plot of the one-minute volumes from the ground truth and four detectors under evaluation. This plot also shows that the ground truth, Solo Pro II, Microloop 702, and G4 had similar distributions, while the upper end of the SmartSensor 105 distribution had a distinctly different shape.

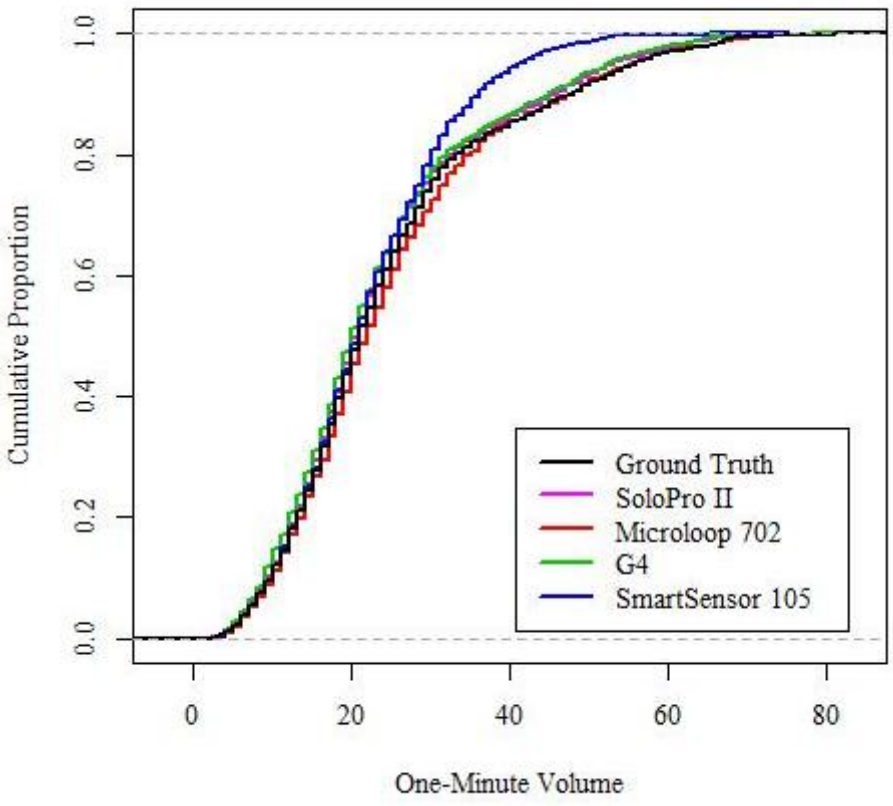

Figure 6.4: Cumulative Distribution Plot of One-Minute Volume Distributions for Ground Truth and All Detectors

Summary one-minute volume statistics were calculated for the ground truth data, as well as each detector, and are given in table 6.1. The values for mean one-minute volume indicate that the Solo Pro II, G4, and SmartSensor 105 tended to under-report one-minute volume compared to the ground truth, while the Microloop 702 mean one-minute volume indicates that it tended to over-report volume. The standard deviation (i.e., 10.3) of the SmartSensor 105 one-minute volume distribution provides further

indication of the lack of high one-minute volumes present in the ground truth one-minute volume distribution, which had a standard deviation of 14.5.

Table 6.1 One-Minute Volume Summary Statistics

	Mean	Median	Standard Deviation
Ground Truth	24.6	21	14.5
Solo Pro II	23.9	21	13.8
Microloop 702	25.2	22	14.4
G4	23.4	20	13.7
SmartSensor 105	22.0	21	10.3

These summary depictions of the one-minute volume data were followed by calculation of the percent error, absolute percent error, and GEH statistic for each detector and each one-minute interval. The distributions one-minute volume percent error are shown in the box plots in figure 6.5. Volume percent error was calculated such that a negative value indicated undercounting and a positive value indicated overcounting. Based on the placement of the inter-quartile ranges with respect to zero percent error, it can be seen that the Solo Pro II, G4, and SmartSensor 105 each tended to undercount more frequently than they overcounted. In contrast, the Microloop 702 can be seen to overcount more frequently than it undercounted. It is also worth noting that while the inter-quartile ranges of the four detectors were all approximately equal in height, the total range of one-minute volume percent errors was much greater for the G4 and SmartSensor 105 than for the Solo Pro II and Microloop 702.

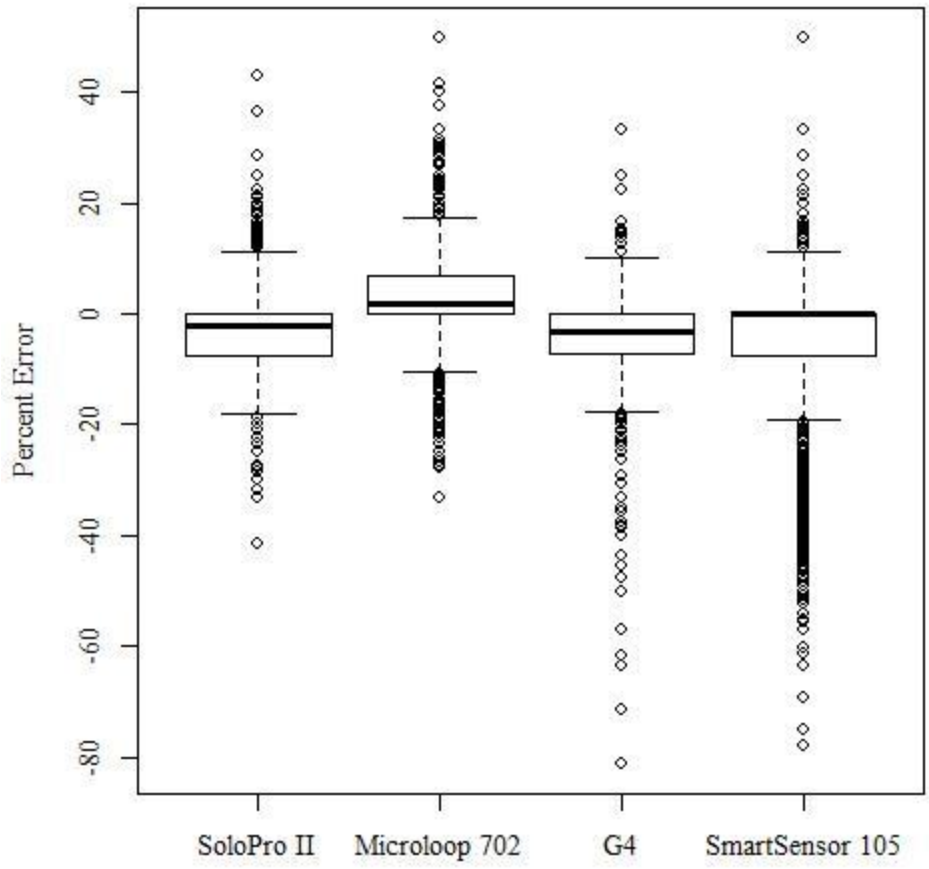

Figure 6.5: One-Minute Volume Percent Error Box Plot

Figure 6.6 shows histograms of the volume percent error distributions for the four detectors. The tendency of each detector to either overcount or undercount is readily observed in these histograms. While the negative values of skewness indicated longer left tails than right tails for the G4 and SmartSensor 105, the Solo Pro II and Microloop 702 had positive values of skewness with relatively equal left and right tails. While difficult to see in the histogram in figure 6.6(a), the density of the upper "outliers" compared to lower "outliers" in the Solo Pro II and Microloop 702 box plots in figure 6.5 provides evidence of the more prominent upper tail of these distributions, leading to the positive value of skewness.

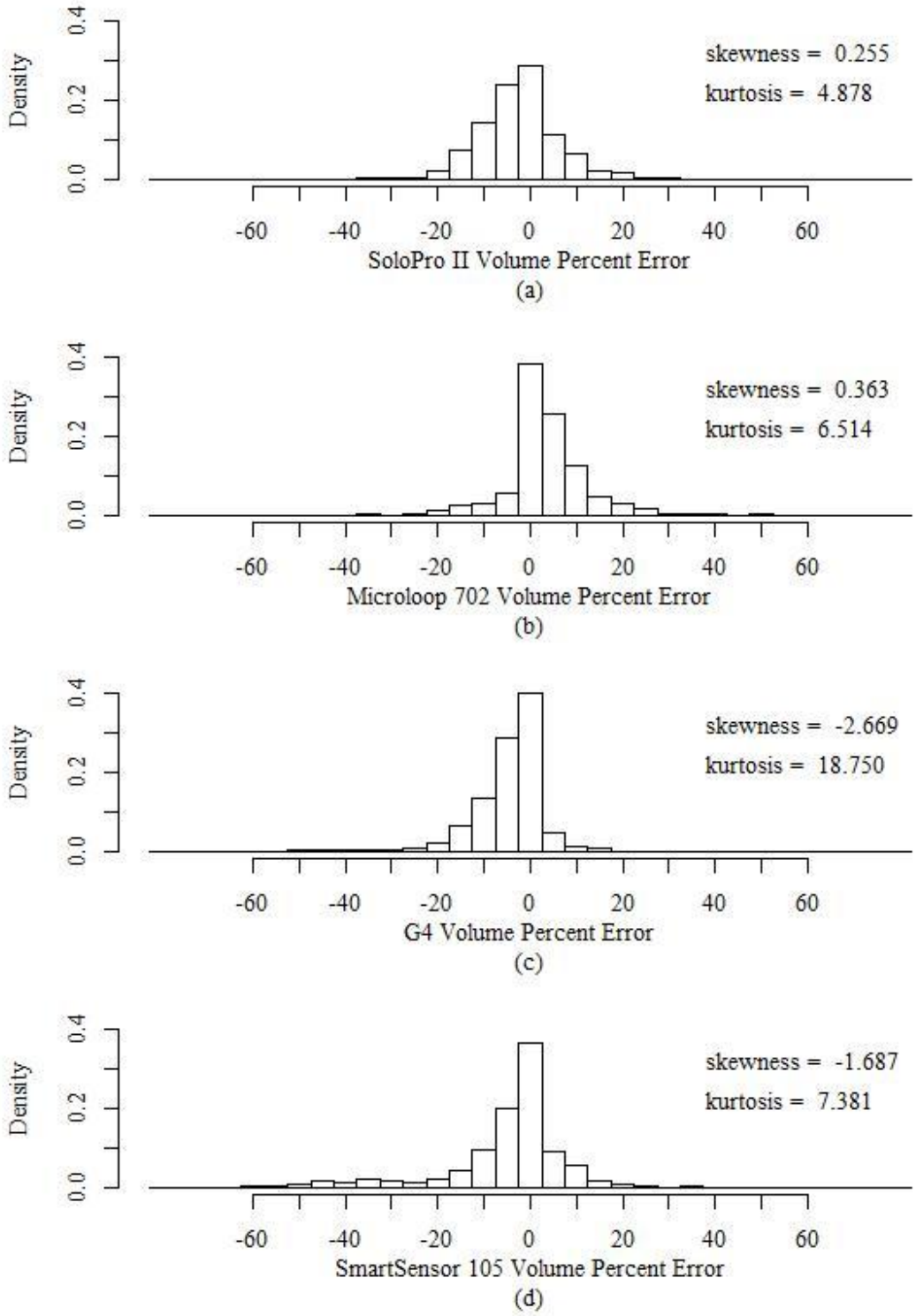

Figure 6.6: Histograms of One-Minute Volume Percent Error Distributions for Solo Pro II (a), Microloop (b), G4 (c), and SmartSensor 105 (d) Detectors

Figure 6.7 shows cumulative distribution plots of the one-minute volume percent error distributions for the four detectors. The vertical portion of each curve at 0% error represents the proportion of one-minute intervals for which the respective detector correctly reported the volume. This graph shows very clearly that the Microloop 702 had the lowest proportion of intervals in which volume was under-reported, while having the largest proportion of intervals in which volume was over-reported. The long left tail of the SmartSensor 105 in figure 6.7 was a result of its under-reporting during high volume intervals.

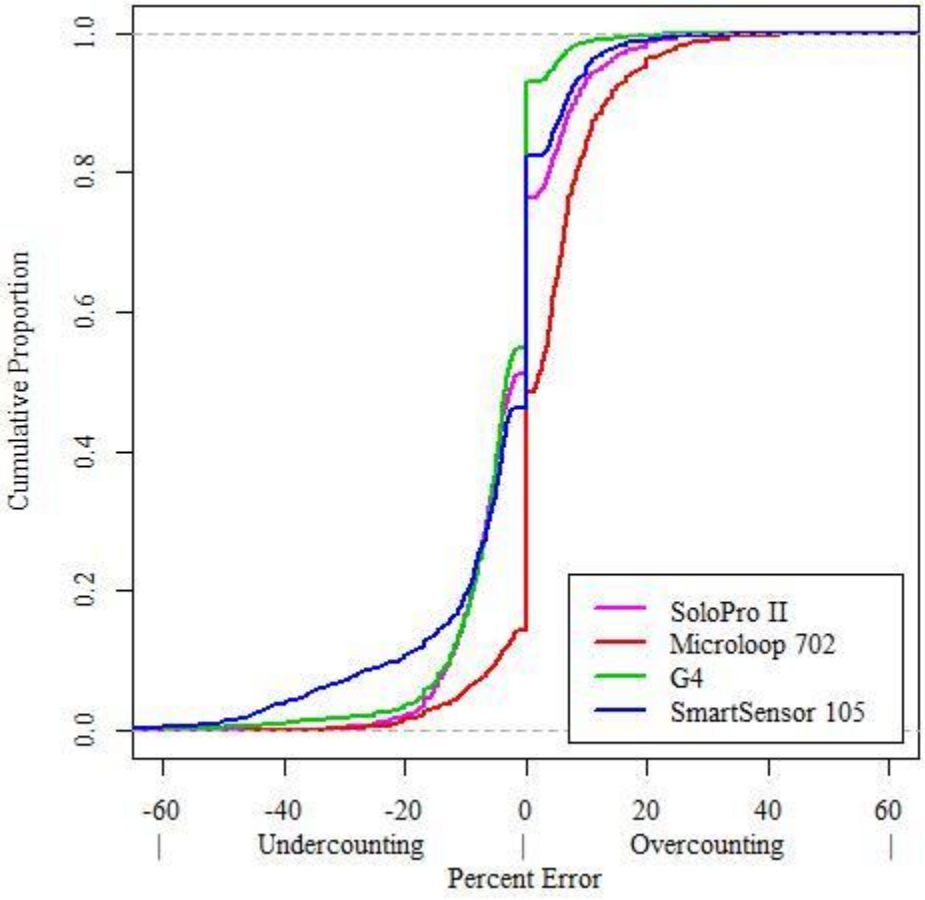

Figure 6.7: One-Minute Volume Percent Error Cumulative Distribution Plot

Appropriate statistics, such as correlation coefficient, mean percent error (MPE), mean absolute percent error (MAPE), percent error variance, mean GEH statistic, 85th

percentile GEH statistic, and GEH variance, are given in table 6.2. A review of the correlation coefficients shows a very strong linear relationship between each of the Solo Pro II, Microloop 702, and G4 one-minute volumes and the ground truth one-minute volumes, and a slightly weaker correlation between the SmartSensor 105 one-minute volumes and ground truth one-minute volumes. Mean percent error values indicate a tendency for under-reporting one-minute volumes by the Solo Pro II, G4, and SmartSensor 105, while indicating a tendency for over-reporting one-minute volumes by the Microloop 702. Mean absolute percent error values indicate that the G4 had, on average, the one-minute volume closest to the ground truth one-minute volume of the four detectors. While MAPE indicates that the G4 reported the most accurate one-minute volumes of the four detectors, the GEH statistic indicated that the Microloop 702 was more accurate than the G4 when absolute error was considered in conjunction with percent error.

Table 6.2: Detector One-Minute Volume Error Statistics

	Correlation Coefficient	MPE	MAPE	Percent Error Variance	Mean GEH	85th Percentile GEH	GEH Variance
Solo Pro II	0.992	-2.34%	6.53%	0.00749	0.304	0.577	0.0712
Microloop 702	0.991	3.30%	6.07%	0.00764	0.270	0.552	0.0852
G4	0.993	-4.52%	5.54%	0.00700	0.276	0.555	0.137
SmartSensor 105	0.910	-5.07%	8.18%	0.0178	0.516	0.707	0.920

Additionally, Theil's inequality coefficient (U) was calculated and presented in table 6.3, along with its proportional components for each detector. This goodness-of-fit measure is explained in section 5.4. It is useful here because of the additional components, indicating the nature of the errors. The first additional component is the bias proportion (U_m) which is a measure of systematic error indicative of consistent overestimation or underestimation of volume. The second additional component is the

variance proportion (U_s) which is a measure of the degree of equality between variance in the reported volumes and variance in actual volumes. The third additional component is the covariance proportion (U_c) which is a measure of the unsystematic error. As mutually exclusive proportions, U_m, U_s, and U_c sum to one.

Table 6.3: One-Minute Volume Theil's Inequality Coefficients

	U	U_m	U_s	U_c
Solo Pro II	0.037	0.136	0.113	0.752
Microloop 702	0.035	0.086	0.002	0.913
G4	0.040	0.292	0.112	0.596
SmartSensor 105	0.135	0.131	0.342	0.527

The values of *U* in table 6.3 indicate that the Microloop 702, Solo Pro II, and G4 one-minute volumes had similar degrees of inequality when each was compared to the ground truth one-minute volumes. The SmartSensor 105 was found to have an inequality coefficient higher than the other three detectors, indicating a comparatively greater inequality when its one-minute volumes are compared to the ground truth one-minute volumes. The fact that the G4 had the highest value of U_m indicates that it had the greatest bias proportion of the three detectors, and could benefit most from further fine tuning of its calibration. The fact that the SmartSensor 105 had the highest U_s indicates that it had the greatest variance proportion of the three detectors, and that the variance in one-minute SmartSensor 105 volumes differed most from the variance in one-minute ground truth volumes. Lastly, the high value of U_c for the Microloop 702 indicates that it had the greatest covariance proportion or unsystematic error. That is to say that a large proportion of the Microloop 702's one-minute volume error could not be explained by consistent bias or a different variance than the ground truth one-minute volumes.

Next, the data set was broken down by environmental conditions. Percent error distributions were determined for data subsets with similar conditions for factors such as

lighting (day, night, dawn, dusk), precipitation (clear, rain), and traffic volume. It is important to note that the effects of dawn and dusk could be expected to differ by installation location. A VIP detector specifically would be affected by these lighting transitions differently if the camera is pointed north, east, south, or west.

Effects of lighting, precipitation, and volume on the Solo Pro II one-minute volume percent error are shown in the distributions in figures 6.8-6.10. Figure 6.8 shows that the largest undercounting and overcounting errors occurred during dawn lighting conditions. It was hypothesized that this was due to long shadows causing problems for the video image processing algorithm employed by this detector. Figure 6.9 shows that rain tended to decrease undercounting by the Solo Pro, while increasing overcounting. This could be attributed to headlight spillover due to a more reflective pavement surface in rainy conditions. The potential causes of this phenomenon are further explored in section 7.1.3 of this thesis. Next, figure 6.10 shows an intuitive effect of volume on Solo Pro II one-minute volume percent error. The frequency and magnitude of overcounting were lower for high volume periods than low volume periods, while undercounting was more frequent, but with a smaller magnitude, for high volume periods.

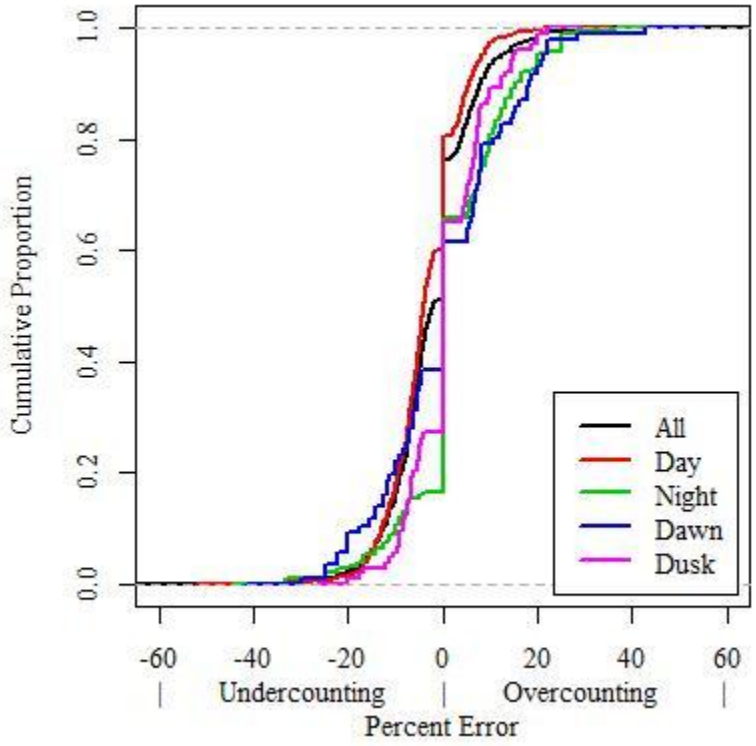

Figure 6.8: Solo Pro II One-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

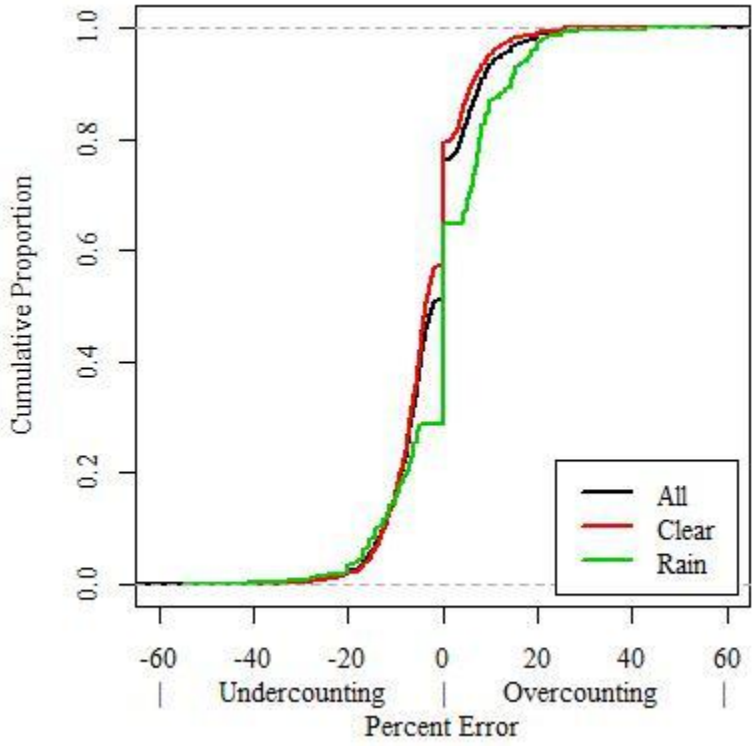

Figure 6.9: Solo Pro II One-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

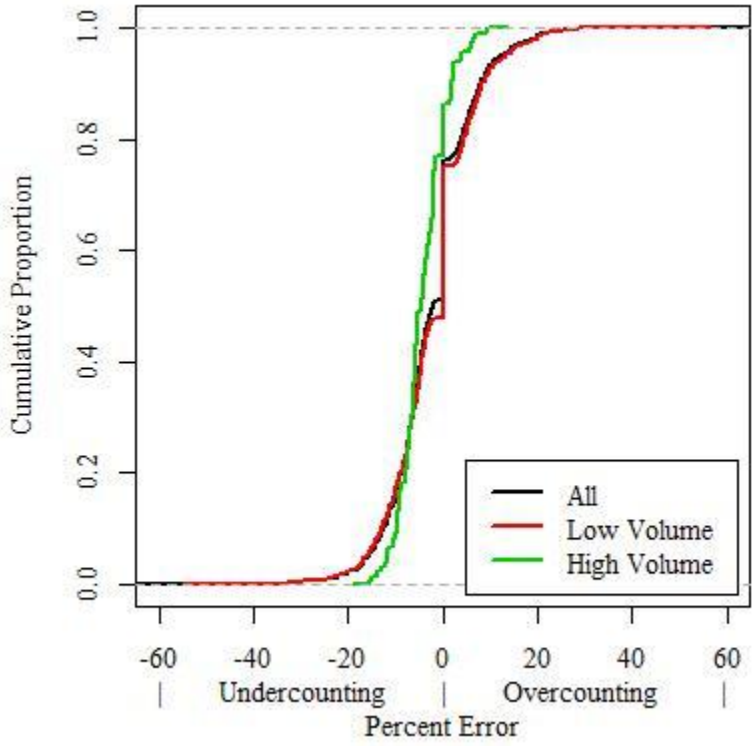

Figure 6.10: Solo Pro II One-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

Figures 6.11-6.13 depict similar plots of the effects of lighting, rain, and volume on the Microloop 702 one-minute volume percent error distributions. Figure 6.11 shows greater undercounting by the Microloop 702 under dusk lighting conditions and greater overcounting under night and dawn lighting conditions. One possible explanation of these trends involves inconsistent vehicle lane position, which could result in either undercounting or overcounting. Figure 6.12 shows similar distributions of one-minute volume percent error under clear and rainy conditions for the Microloop 702. Lastly, the effects of volume seen in figure 6.13 indicate that at higher volumes, overcounting by the Microloop 702 decreased and undercounting increased in both frequency and magnitude.

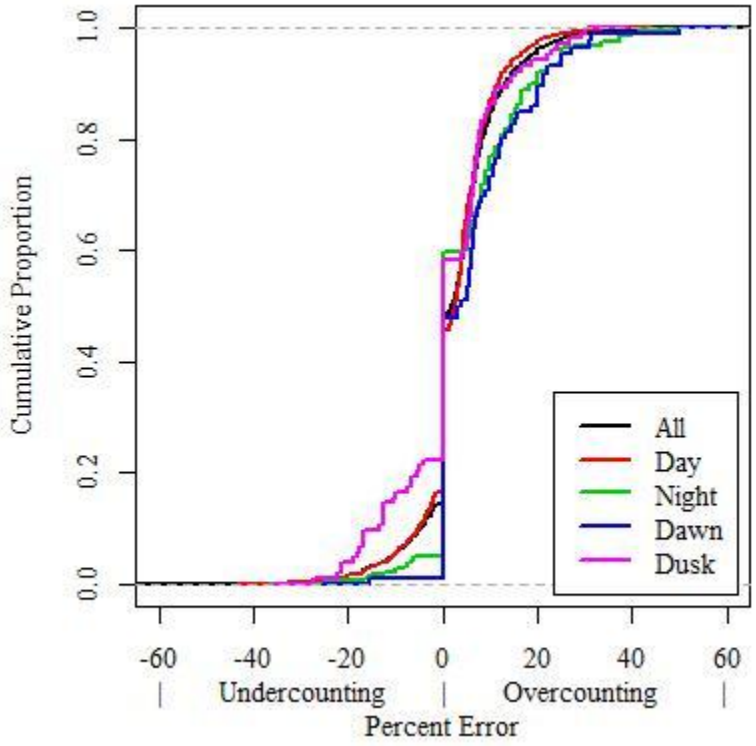

Figure 6.11: Microloop 702 One-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

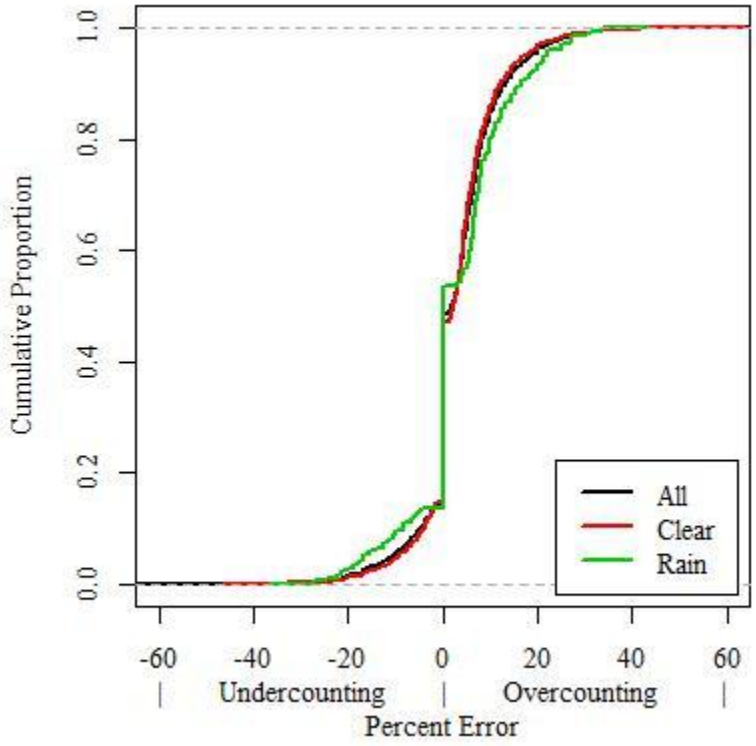

Figure 6.12: Microloop 702 One-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

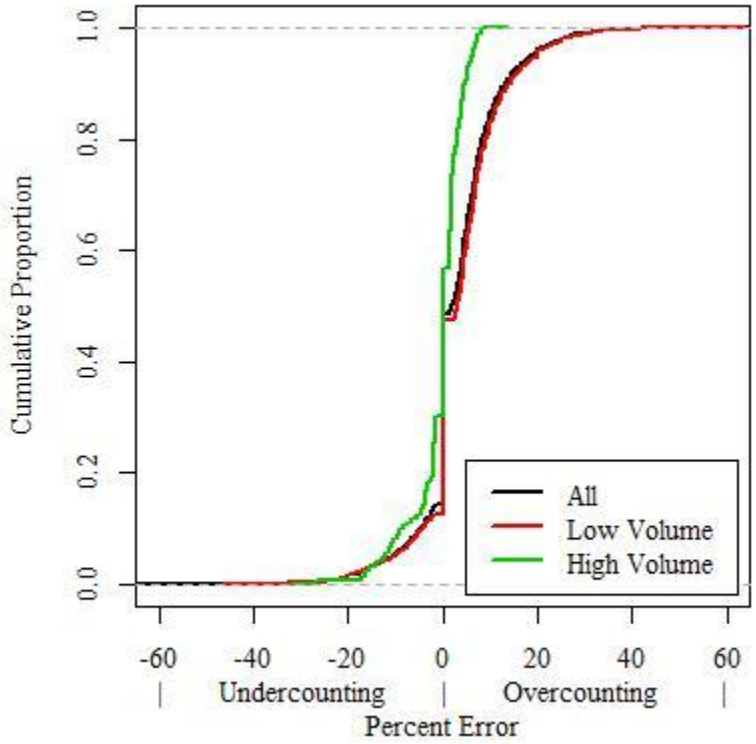

Figure 6.13: Microloop 702 One-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

Figures 6.14-6.16 depict the effects of lighting, rain, and volume on the G4 one-minute volume percent error distributions. In figure 6.14, the greater proportion and magnitude of undercounting during dusk conditions, compared to other lighting conditions, stands out. Further review of the ground truth video revealed that the heaviest period of rain in the dataset took place during dusk conditions on June 20th, 2011. It was hypothesized that the severe undercounting during dusk conditions was due to the heavy rain. This hypothesis was supported by the severe undercounting during rainy conditions, shown in figure 6.15. Figure 6.16 shows that high volume tended to reduce overcounting by the G4, while generally increasing the frequency and decreasing the magnitude of undercounting.

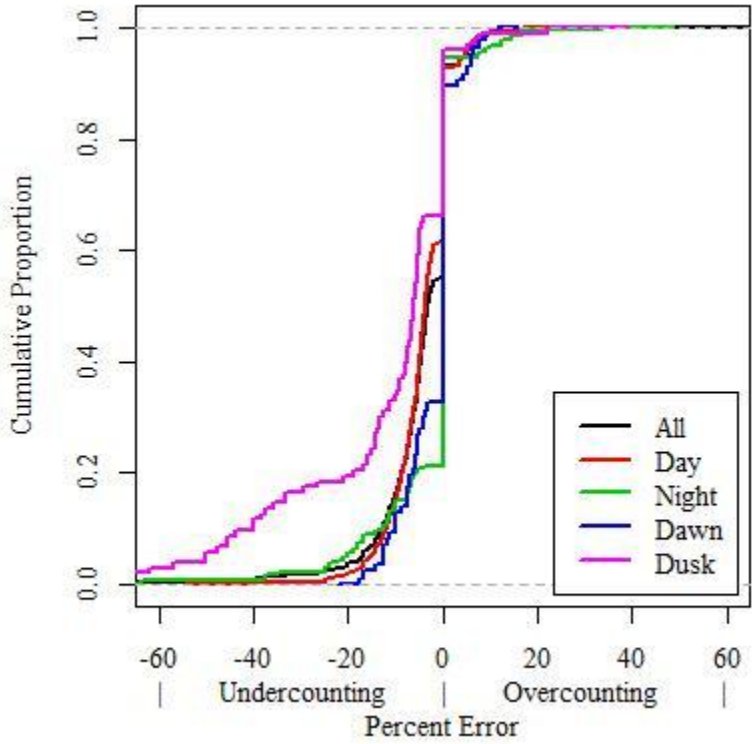

Figure 6.14: G4 One-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

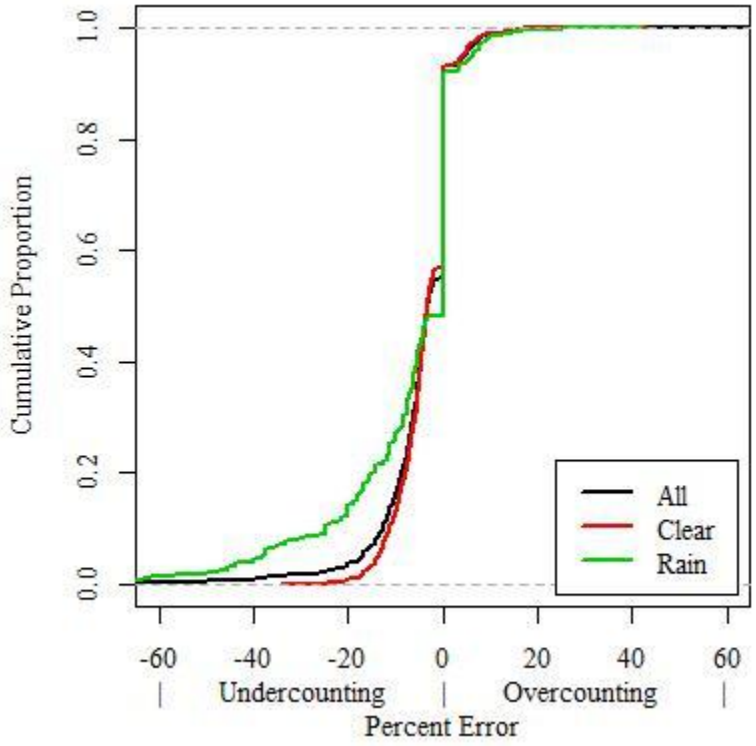

Figure 6.15: G4 One-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

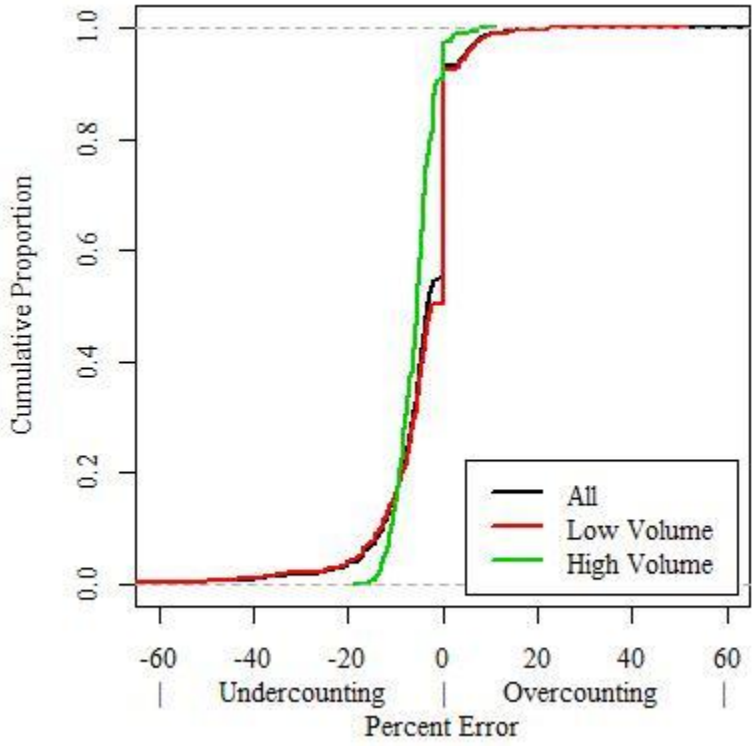

Figure 6.16: G4 One-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

Figures 6.17-6.19 show the effects of lighting, rain, and volume on the SmartSensor 105 one-minute volume percent error distributions. In order to interpret figures 6.17 and 6.18, it is important to first recognize the strong impact of high volume traffic on the undercounting of the SmartSensor 105, as shown in figure 6.19. Under high volume conditions (LOS C or D), the SmartSensor 105 undercounted 96.9% of the one-minute intervals. 50% of those high volume intervals were undercounted by 30.6% or more. This severe impact of high traffic volume provides an explanation of the severe undercounting in day lighting conditions as well as clear (i.e. rain-free) conditions, as all high traffic volume intervals occurred during clear, day lighting periods.

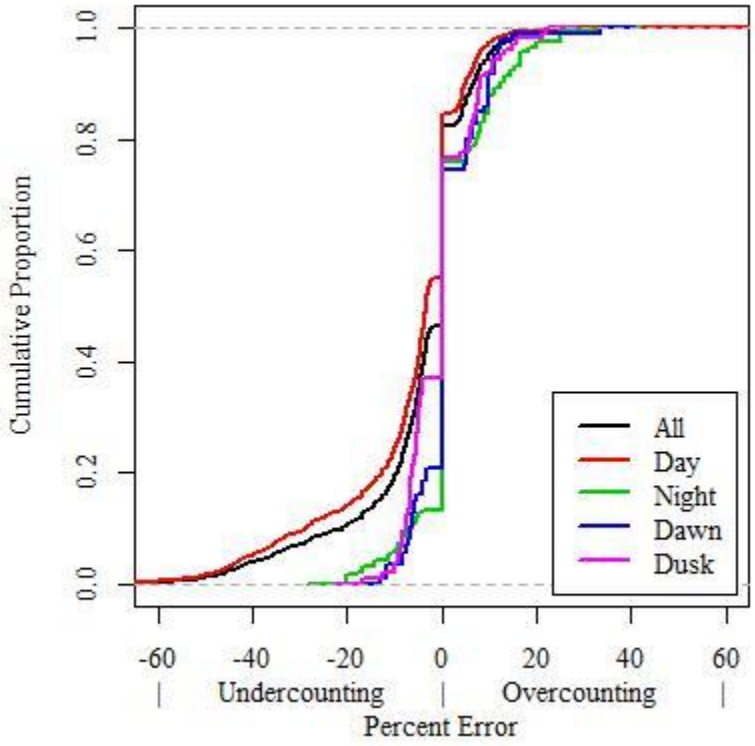

Figure 6.17: SmartSensor 105 One-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

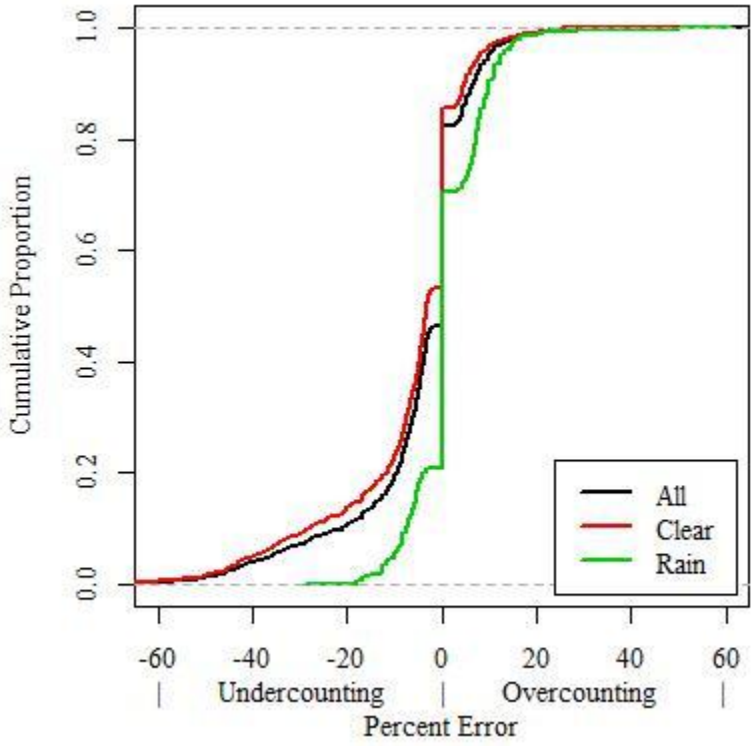

Figure 6.18: SmartSensor 105 One-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

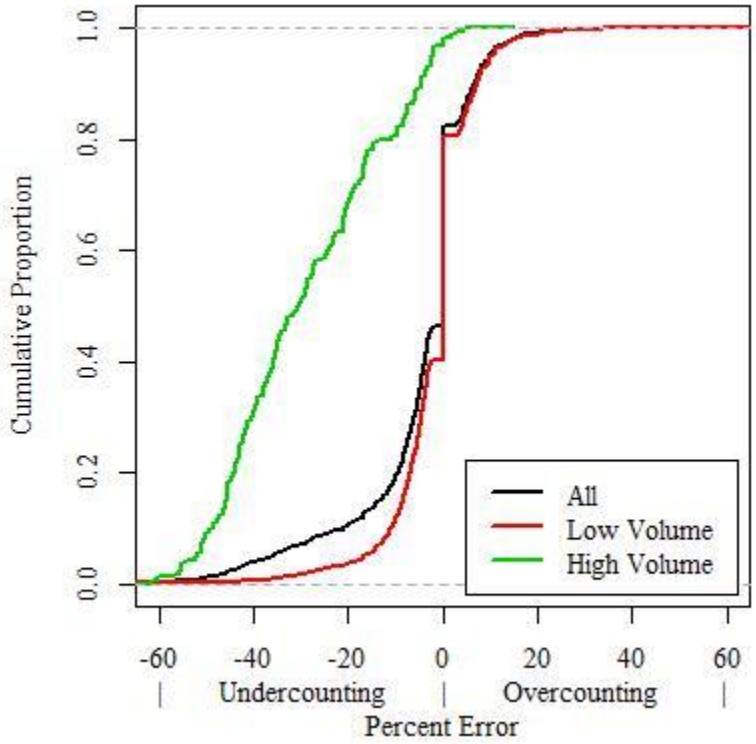

Figure 6.19: SmartSensor 105 One-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

Next, the statistical significance of the effects of various environmental conditions on one-minute volume percent error was assessed through analysis of variance (ANOVA). Specifically, ANOVA based on the model defined in section 5.5 was performed on each detector's one-minute volume percent error with factors for lighting (levels=Day, Night, Dawn, and Dusk) and precipitation (levels = None and Rain). In order to minimize the effects of serial correlation, the data sets for the Solo Pro II, Microloop 702, and G4 were thinned by a factor of 10 (that is, the ANOVA models were developed using every 10th minute of data in the initial data set) to include 147 data points, while the Smartsensor 105 data were thinned by a factor of 20 to include 74 data points. The decision to thin the data sets with the stated factors is documented in Appendix B. Statistical significance is reported at an $\alpha = 0.05$ level.

The output of the Solo Pro II ANOVA found in table 6.4 indicates that lighting and rain each had statistically significant impacts on the Solo Pro II's one-minute volume percent error. These effects could be attributed to vehicle shadows in specific lighting conditions and headlight glare in rainy conditions. The results of the Microloop 702 ANOVA, found in table 6.5, indicate that the interaction between lighting and rain had a statistically significant impact on Microloop 702 one-minute volume percent error. This effect could be attributed to vehicle lane position under different precipitation and lighting conditions. The results of the G4 ANOVA, found in table 6.6, indicate that lighting, rain, the interaction between lighting and rain, and the intercept all had statistically significant impacts on G4 volume percent error. These results defied expectations, as there exists no intuitive, practical explanation for this technology to be affected by both lighting and rain. Further review of the data found that this detector performed the most poorly during a nearly three-hour rainy period that spanned day, dusk, and night lighting. One potential explanation is that water or water vapor entered the detector housing and caused malfunction during this period. This hypothesis was based on a similar issue documented with an earlier model in this detector family in a previous study (22). An independent study of this issue was beyond the scope of this thesis. Lastly, the results of the SmartSensor 105 ANOVA, found in table 6.7, indicate that rain had a statistically significant impact on the SmartSensor 105 volume percent error. One possible explanation of this effect could be that the radar signal reflected off of large raindrops and created false detections.

Table 6.4: Solo Pro II One-Minute Volume Percent Error ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	0.001	1	0.203	0.653	
Lighting	0.089	3	5.422	0.001	*
Rain	0.041	1	7.473	0.007	*
Residuals	0.777	142			

Table 6.5: Microloop 702 One-Minute Volume Percent Error ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	0.008	1	1.593	0.209	
Lighting	0.005	3	0.326	0.806	
Rain	0.013	1	2.705	0.102	
Lighting:Rain	0.071	3	4.814	0.003	*
Residuals	0.684	139			

Table 6.6: G4 One-Minute Volume Percent Error ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	0.268	1	34.2355	0.000	*
Lighting	0.141	3	6.0312	0.001	*
Rain	0.033	1	4.1616	0.043	*
Lighting:Rain	0.129	3	5.4895	0.001	*
Residuals	1.086	139			

Table 6.7: SmartSensor 105 One-Minute Volume Percent Error ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	0.017	1	1.271	0.264	
Lighting	0.014	3	0.353	0.787	
Rain	0.139	1	10.177	0.002	*
Residuals	0.941	69			

Type III sums of squares were selected based on the fact that the analysis was unbalanced, meaning that there were unequal numbers of observations at each level of the given factors. This type of sum of squares tests each factor with the effect of all other factors including the interaction as givens. In cases where the interaction effect was found to not be statistically significant, it was eliminated from the model and a subsequent model was analyzed. It was concluded that the lighting-precipitation effect was not significant for the Solo Pro II (table 6.4) or SmartSensor 105 (table 6.7).

Next, an attempt was made to fit a multiple regression model for the one-minute volume percent error for each detector to support trends noticed in the graphical representation of the data. The model for this regression takes the form presented in section 5.6, with the dependent variable (p_i) being the volume percent error of the given detector for minute i, and the first dependent variable (α) being the theoretical mean volume percent error for the specified detector given daylight, non-rainy conditions with a true volume of 0 vehicles. The same thinning methodology presented in Appendix B for ANOVA analyses was used in this regression analysis, however, different required thinning factors were dictated by these regression models. In this case, the data for all detectors was thinned by a factor of 10.

The Solo Pro II one-minute volume percent error model has coefficients given in table 6.8. The statistically significant factors in this model were night lighting and the combined effect of dawn lighting and rain. It was hypothesized that night and the interaction effect of dawn and rain were significant due to headlight spillover. The adjusted R-squared for this model was 0.1476, indicating a low correlation between the predicted and observed values for Solo Pro II one-minute volume percent error.

Table 6.8: Solo Pro II One-Minute Volume Percent Error Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	-2.50	1.879	-1.331	0.185	
V.Truth (β1)	-0.03	0.053	-0.612	0.542	
Night (γ11)	7.70	2.328	3.309	0.001	*
Dawn (γ12)	-7.18	3.878	-1.852	0.066	
Dusk (γ13)	-0.43	3.152	-0.135	0.893	
Rain (γ21)	2.69	2.114	1.27	0.206	
Night:Rain (γ31)	-4.46	4.651	-0.959	0.339	
Dawn:Rain (γ32)	12.71	5.606	2.267	0.025	*
Dusk:Rain (γ33)	4.98	4.846	1.029	0.305	

A similar model was created next, but with independent variables not found to be significant in the first model excluded. The coefficients in this model are shown in table 6.9. While this model had an even lower adjusted R-squared value of 0.1085, the average effect of the significant factors from the first model on the Solo Pro II one-minute volume percent error are shown more clearly in the "Estimate" column of this model. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more accurately depict the effects of the significant independent variables on Solo Pro II one-minute volume percent error.

Table 6.9: Solo Pro II One-Minute Volume Percent Error Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	-2.94	0.670	-4.384	0.000	*
Night (γ11)	7.39	1.838	4.021	0.000	*
Dawn:Rain (γ32)	8.16	3.790	2.152	0.033	*

The Microloop 702 one-minute volume percent error model coefficients are shown in table 6.10. The only statistically significant factor in this model was the combined effect of dusk lighting and rain. It was hypothesized that this effect was found to be significant due to erratic vehicle lane position caused by either driver fatigue or heavy rain occurring during one of the dawn periods in the data set. The adjusted R-squared for this model was 0.0832, indicating a low correlation between the predicted and observed values for Microloop 702 one-minute volume percent error.

Table 6.10: Microloop 702 One-Minute Volume Percent Error Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	2.99	1.807	1.657	0.100	
V.Truth (β1)	-0.05	0.051	-1.035	0.303	
Night (γ11)	4.22	2.238	1.884	0.062	
Dawn (γ12)	-6.20	3.728	-1.662	0.099	
Dusk (γ13)	5.34	3.030	1.763	0.080	
Rain (γ21)	-0.14	2.033	-0.069	0.945	
Night:Rain (γ31)	-7.81	4.472	-1.746	0.083	
Dawn:Rain (γ32)	7.28	5.390	1.351	0.179	
Dusk:Rain (γ33)	-12.81	4.659	-2.749	0.007	*

Another similar model was created that excluded independent variables which were not found to be significant in the first model. The coefficients in this model are shown in table 6.11. While this model had an even lower adjusted R-squared value of 0.0272, the average effect of the significant factors from the first model on the Microloop 702 one-minute volume percent error are shown more clearly in the "Estimate" column of this model. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more accurately depict the effects of the significant independent variables on Microloop 702 one-minute volume percent error.

Table 6.11: Microloop 702 One-Minute Volume Percent Error Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	2.06	0.606	3.392	0.001	*
Dusk:Rain (γ33)	-7.41	3.287	-2.255	0.026	*

The G4 one-minute volume percent error model coefficients are shown in table 6.12. The statistically significant factors in this model were the intercept and the combined effect of dusk lighting and rain. It was hypothesized that the intercept was significant because of the low variance in G4 one-minute volume percent-error. It was

also hypothesized that the combined effect of dusk and rain was significant due to heavy rain occurring during one of the dawn periods in the data set. The adjusted R-squared for this model was 0.1380, indicating a low correlation between the predicted and observed values for G4 one-minute volume percent error.

Table 6.12: G4 One-Minute Volume Percent Error Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	-5.99	2.284	-2.622	0.010	*
V.Truth (β1)	0.03	0.064	0.500	0.618	
Night (γ11)	4.02	2.829	1.422	0.157	
Dawn (γ12)	-0.49	4.713	-0.105	0.917	
Dusk (γ13)	-0.80	3.830	-0.210	0.834	
Rain (γ21)	4.17	2.569	1.622	0.107	
Night:Rain (γ31)	-10.92	5.652	-1.932	0.055	
Dawn:Rain (γ32)	-2.58	6.813	-0.379	0.705	
Dusk:Rain (γ33)	-22.68	5.888	-3.852	0.000	*

Another similar model was created, but with the removal of independent variables not found to be significant in the first model. The coefficients in this model are shown in table 6.13. This model had a slightly higher adjusted R-squared value of 0.1477. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more accurately depict the effects of the significant independent variable on G4 one-minute volume percent error.

Table 6.13: G4 One-Minute Volume Percent Error Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	-4.28	0.740	-5.79	0.000	*
Dusk:Rain (γ33)	-20.57	4.011	-5.129	0.000	*

The SmartSensor 105 one-minute volume percent error model coefficients are shown in table 6.14. The statistically significant factors in this model were the intercept and true volume. It was hypothesized that the intercept was found to be significant due to

the SmartSensor 105's high average volume percent error, and that the true volume was significant due to increased volume percent error under high volume conditions. The adjusted R-squared for this model was 0.3687, which, while higher than the adjusted R-squared values from the models for the other detectors, also indicated a low correlation between the predicted and observed values for SmartSensor 105 one-minute volume percent error. The reason this adjusted R-squared is so high compared to those of the other detectors was because of the strong effect of true volume on the SmartSenser 105 volume percent error, as can be seen in figure 6.1(d).

Table 6.14: SmartSensor 105 One-Minute Volume Percent Error Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	9.34	2.742	3.406	0.001	*
V.Truth (β1)	-0.60	0.077	-7.788	0.000	*
Night (γ11)	-4.31	3.397	-1.270	0.206	
Dawn (γ12)	-6.02	5.659	-1.063	0.289	
Dusk (γ13)	1.36	4.599	0.296	0.767	
Rain (γ21)	-0.49	3.085	-0.159	0.874	
Night:Rain (γ31)	2.13	6.787	0.314	0.754	
Dawn:Rain (γ32)	1.68	8.180	0.206	0.837	
Dusk:Rain (γ33)	3.64	7.070	0.515	0.608	

Another similar model was created with independent variables not found to be significant in the first model excluded. The coefficients in this model are shown in table 6.15. This model had a slightly higher adjusted R-squared value of 0.3784. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more accurately depict the effects of the significant independent variable on SmartSensor 105 one-minute volume percent error.

Table 6.15: SmartSensor 105 One-Minute Volume Percent Error Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	7.63	1.713	4.452	0.000	*
V.Truth (β1)	-0.56	0.059	-9.48	0.000	*

While the low adjusted R-squared values for these models suggest a weak linear relationship between the independent factors and the one-minute volume percent error, this is to be expected in this application, due to variability in detection based on factors other than the environmental factors considered herein. If it were possible to consistently predict the volume percent error of a specific detector for any given minute based on a model of this character, it would be possible to eliminate these errors. While these models are not as accurate as one might hope, as evidenced by their low adjusted R-squared values, they remain useful in their ability to demonstrate the average effect of potential environmental factors (see "Estimate" column in the previous tables) and to show which of these effects are consistent enough to be deemed statistically significant.

6.1.2 One-Minute Speed Analysis

The analysis of one-minute mean speed is the focus of this section. As a particular ground truth speed measurement was not available at the test site, the Microloop 702 was selected as a baseline against which the other detectors were compared. The results of this analysis are tempered by the acknowledgement that there were potential errors in the baseline speed from the Microloop 702. The reason that this system was selected as the baseline was that its practical implementation most closely resembled the legacy system of loop detector "speed traps."

The one-minute mean speed analysis began with graphical representations of the reported one-minute mean speeds for each detector. The box plot in figure 6.20 indicates

that the Solo Pro II tended to report a higher speed than the other detectors. However, this bias could potentially be reduced with further calibration. For further information on potential calibration tools available to remove this bias, refer to section 7.2. A more important concern was the variability in the reported one-minute mean speeds. The histograms in figure 6.21, as well as the cumulative distribution curves in figure 6.22, depict similar shapes for the distributions of the Solo Pro II, Microloop 702, and SmartSensor 105, with a distinct shape for the G4's distribution, which has a shorter left tail.

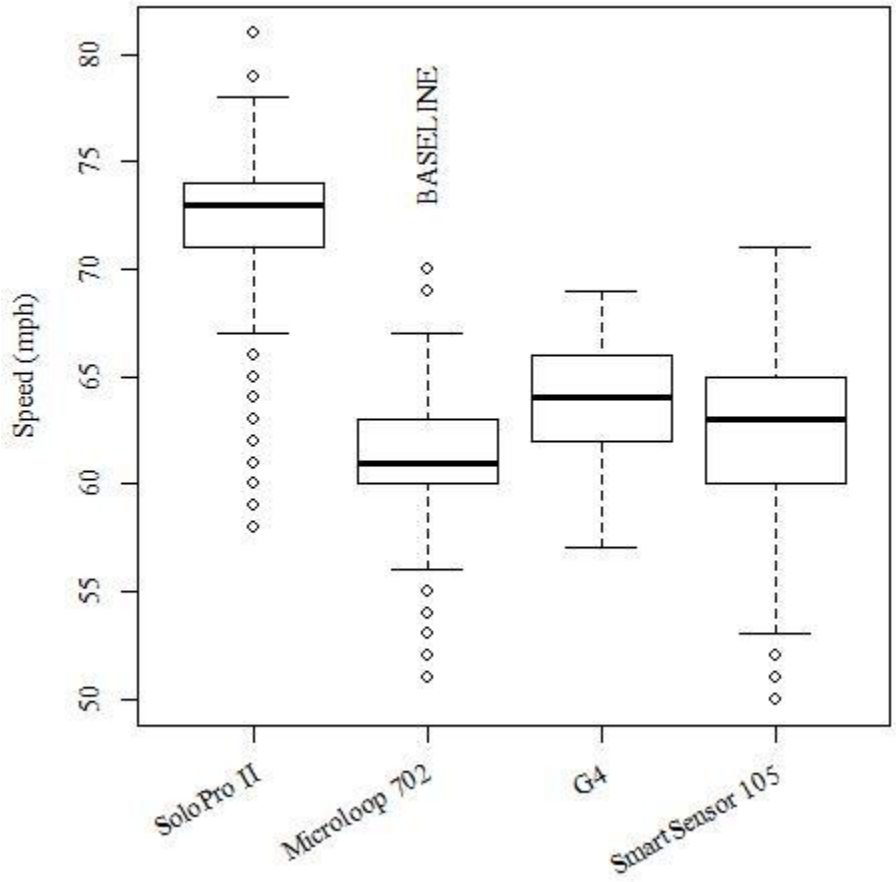

Figure 6.20: Box Plot of Reported One-Minute Mean Speeds

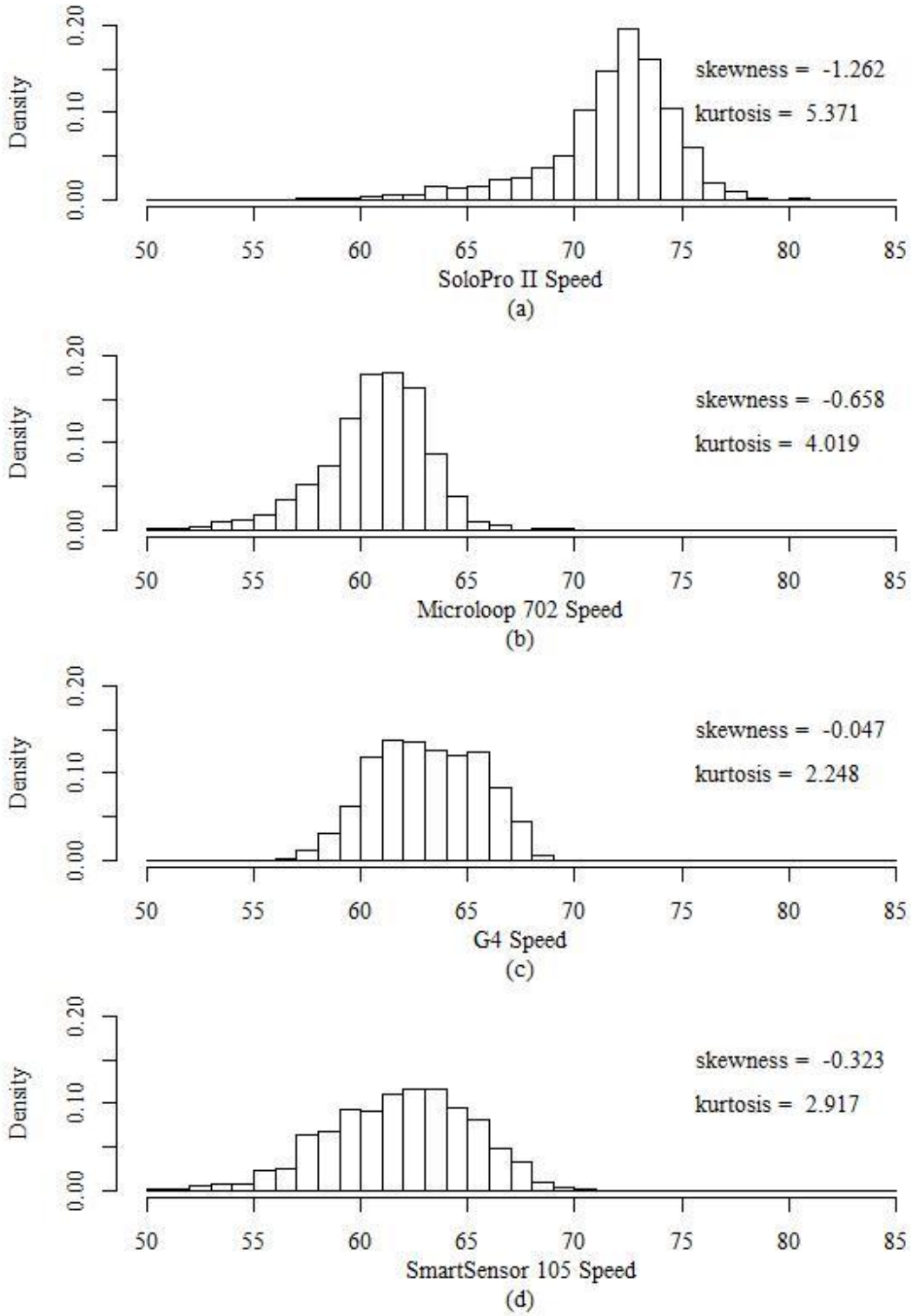

Figure 6.21: Histograms of One-Minute Mean Speed Distributions for the Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

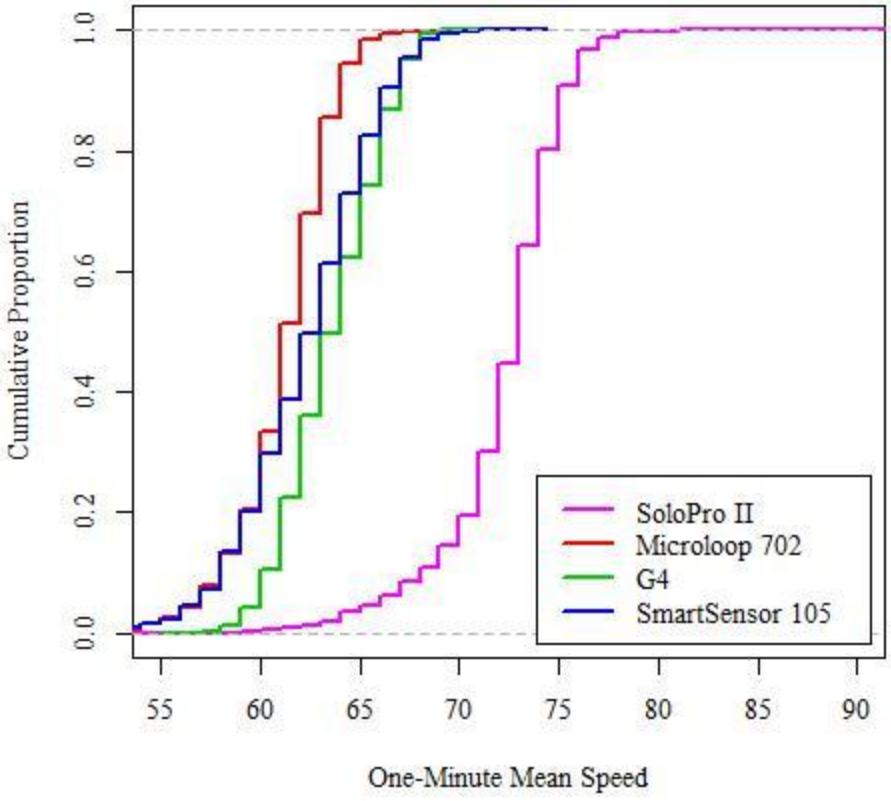

Figure 6.22: Cumulative Distribution Plot of One-Minute Mean Speed Distributions for All Detectors

Summary statistics for the one-minute mean speed distributions are given in table 6.16. In this table, the speed bias of the Solo Pro II is again evident, with the mean Solo Pro II speed being approximately 11 miles per hour higher than the mean baseline speed from the Microloop 702. It is also interesting to note that while the G4 speed distribution appeared to be different from the baseline Microloop 702 distribution, it had a standard deviation very similar to the baseline distribution. The kurtosis (as shown in figure 6.21) is a good measure of the difference between the G4 and baseline one-minute speed distributions. The Microloop 702 distribution, which was much more peaked than the G4 distribution, had a kurtosis of 4.019, in comparison to 2.248.

Table 6.16 One-Minute Mean Speed Summary Statistics

	Mean	Median	Standard Deviation
Solo Pro II	72	73	3.09
Microloop 702	61	61	2.43
G4	64	64	2.45
SmartSensor 105	62	63	3.32

*all units are (mph)

Next the detected speeds from the Solo Pro II, G4, and SmartSensor 105 were compared to the one-minute mean speed of the Microloop 702 baseline detector. The scatter plots are shown in figure 6.23. The accompanying correlation coefficients (*r*) indicate that the Solo Pro II had the strongest linear relationship to the baseline one-minute mean speeds, with a correlation coefficient of 0.736, compared to 0.327 for the G4 and 0.433 for the SmartSensor 105.

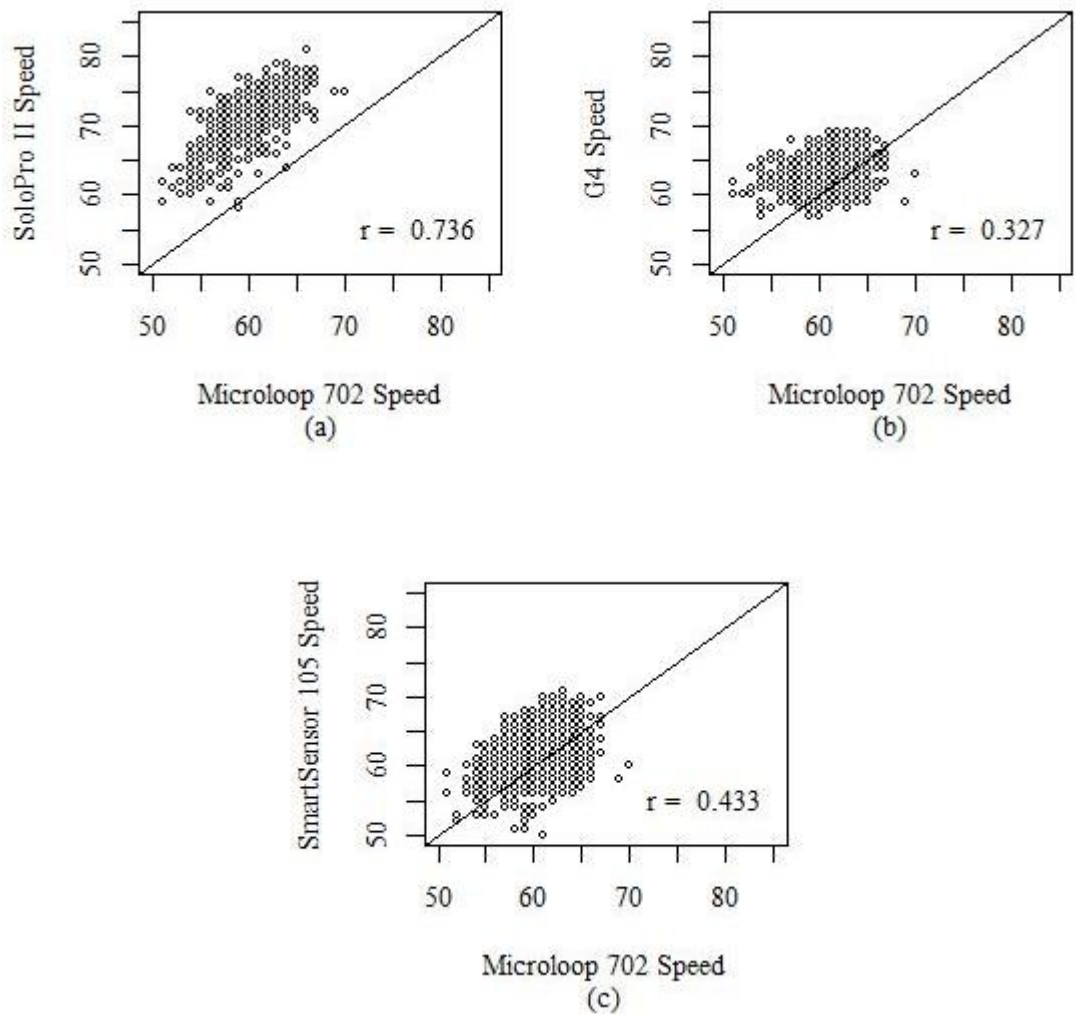

Figure 6.23: One-Minute Mean Speed Scatter Plots Against Baseline for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

This step was followed by the calculation of the percent deviation and absolute percent deviation from the baseline for each detector and each one-minute interval. The distributions of the percent deviation values for each detector are displayed graphically in figures 6.24-6.26. In figure 6.24, the inter-quartile range of the Solo Pro II is shorter than the inter-quartile ranges of the other detectors, indicating less variance in the percent deviation between the Solo Pro II and the baseline than between either the G4 and the baseline of the SmartSensor 105 and the baseline. The histograms in figure 6.25 further

indicate that the percent deviation from the baseline one-minute speeds was more consistent for the Solo Pro II than for the other detectors. This is quantified by the kurtosis values given with the histograms. The kurtosis of the Solo Pro II one-minute mean speed percent deviation distribution was 6.317, indicating a peaked distribution, while the G4 and SmartSensor 105 distributions had kurtoses of 3.279 and 3.202, respectively, indicating distributions with a peakedness similar to a normal distribution. The relative steepness of the middle portion of the Solo Pro II cumulative distribution curve in figure 6.26 provides another depiction of the consistency of its one-minute speed deviation from the baseline.

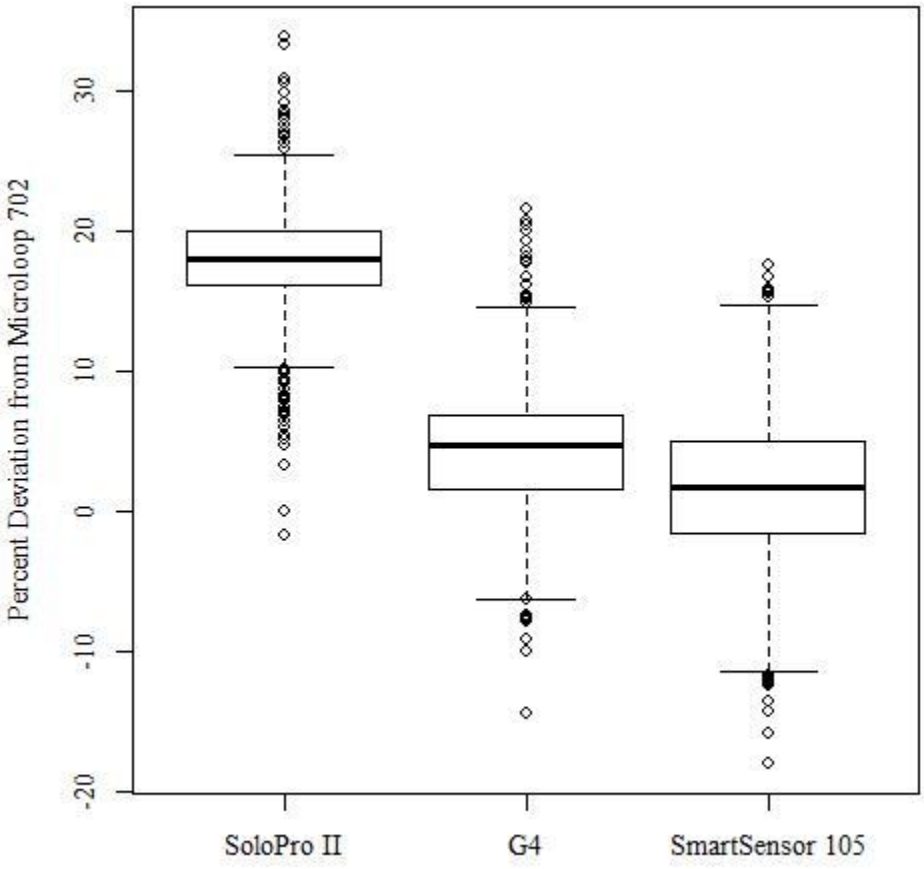

Figure 6.24: One-Minute Mean Speed Percent Deviation Box Plot

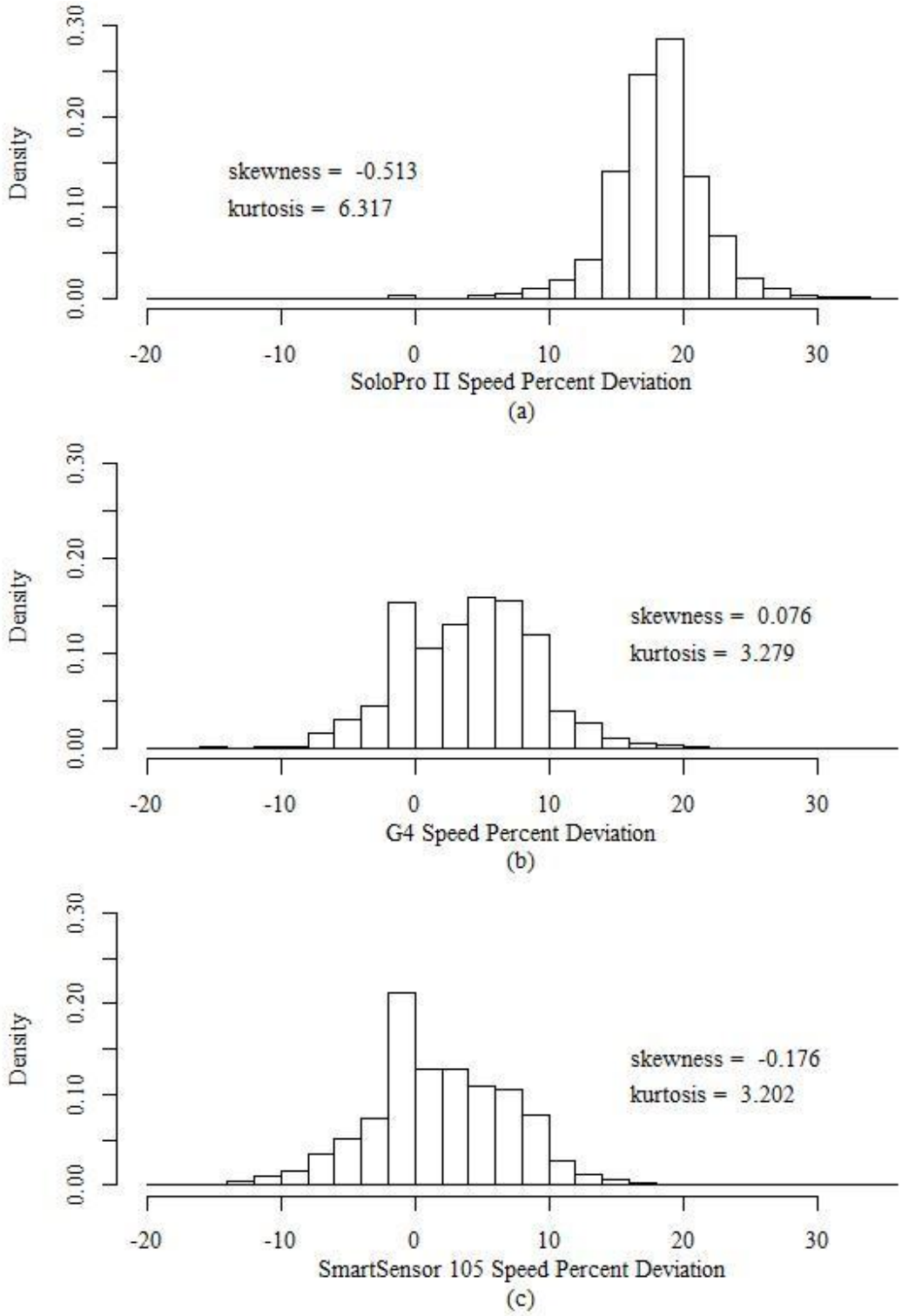

Figure 6.25: Histograms of One-Minute Mean Speed Percent Deviation Distributions for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

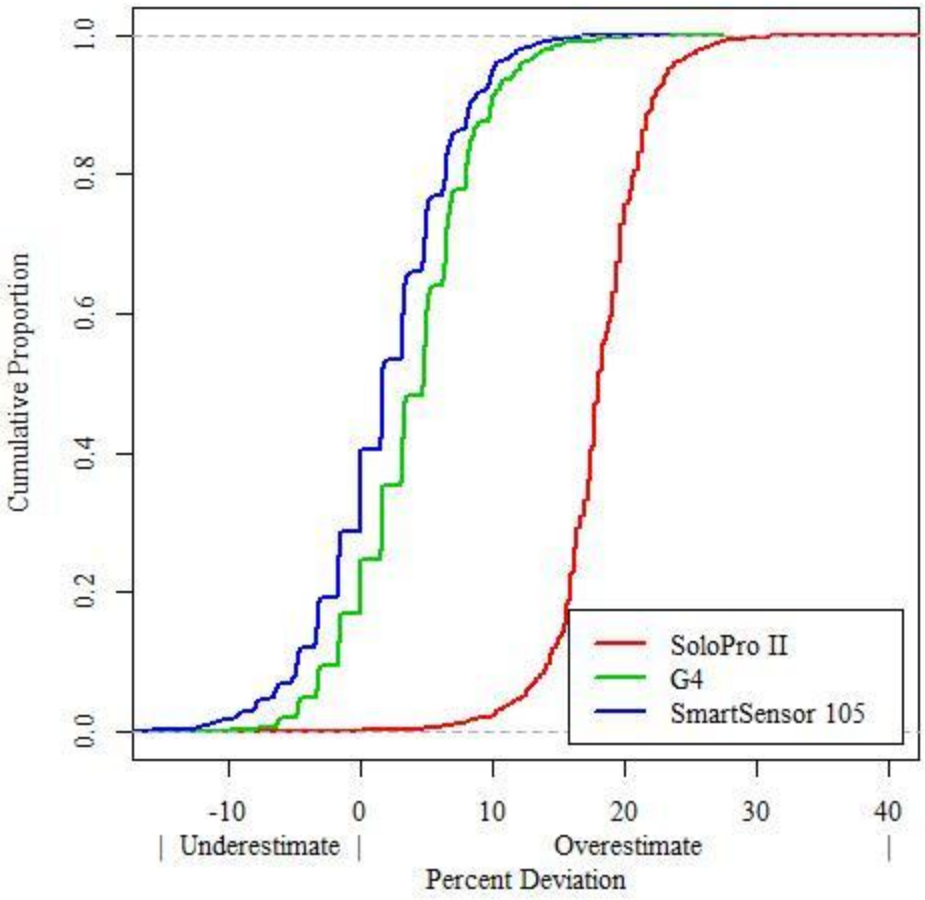

Figure 6.26: One-Minute Mean Speed Percent Deviation Cumulative Distribution Plot

Appropriate one-minute mean speed deviation statistics, such as mean percent deviation (MPD), mean absolute percent deviation (MAPD), and percent deviation variance are given in table 6.17. Comparison of the MPD values in this table indicates that the SmartSensor 105 was calibrated so that its mean speed most closely reflected the mean speed of the baseline detector. The percent deviation variances quantify the observations regarding the preceding figures. The Solo Pro II had a percent deviation variance much lower than the other two detectors, indicating that its deviation from the baseline was more consistent. It is again worth noting that this consistent bias could be removed with further appropriate calibration.

Table 6.17: Detector One-Minute Mean Speed Deviation Statistics

	MPD	MAPD	Percent Deviation Variance
Solo Pro II	18.10%	18.10%	0.00130
G4	4.03%	5.11%	0.00228
SmartSensor 105	1.92%	4.38%	0.00269

Theil's inequality coefficient was also calculated for one-minute mean speeds, and is presented, along with its proportion components, in table 6.18. This goodness-of-fit measure is explained in section 5.4. The proportion components provided further understanding of the characteristics of the differences in each detector's reported speed from the baseline. The bias proportion (U_m) is a measure of proportion of the deviation due to consistent bias in the detection of speed. The variance proportion (U_s) is a measure of the proportion of the deviation due to inequality baseline and detector variances in one-minute mean speeds. The covariance proportion (U_c) is a measure of the proportion of the deviation that is unsystematic or random. As mutually exclusive proportions, U_m, U_s, and U_c sum to one.

Table 6.18: One-Minute Mean Speed Theil's Inequality Coefficients

	U	U_m	U_s	U_c
Solo Pro II	0.084	0.965	0.003	0.031
G4	0.030	0.417	0.000	0.583
SmartSensor 105	0.027	0.114	0.070	0.817

The values of *U* in table 6.18 indicate that the G4 and SmartSensor 105 one-minute mean speeds had similar degrees of inequality when each was compared to the baseline one-minute mean speeds. The Solo Pro II was found to have an inequality coefficient higher than the other detectors, indicating a comparatively greater inequality when its one-minute mean speeds were compared to the baseline one-minute mean speeds. The fact that the Solo Pro II had the highest U_m indicates that it had the greatest

bias proportion of the three detectors, and could benefit most from further calibration. The fact that the SmartSensor 105 had the highest value of U_s indicates that it had the greatest variance proportion of the three detectors, and that the variance in one-minute SmartSensor 105 mean speeds was the most significantly different from the variance in one-minute baseline mean speeds. Lastly, the high value of U_c for the SmartSensor 105 indicates that it has the greatest covariance proportion or unsystematic error. That is to say that a large proportion of the SmartSensor 105's one-minute speed percent deviation cannot be explained by consistent bias or a different variance than the baseline one-minute speeds.

Next, the data set was broken down by environmental conditions; percent deviation distributions were determined for data subsets with similar conditions for factors such as lighting (day, night, dawn, dusk), precipitation (clear, rain), and traffic volume.

Effects of lighting, precipitation, and volume on the Solo Pro II one-minute mean speed percent deviation are shown in the distributions in figures 6.27-6.29. Figure 6.27 shows that there was more variation in the one-minute speed percent deviations under night, dawn, and dusk lighting conditions than under day lighting conditions. It was hypothesized that headlight use during night, dawn, and dusk periods created a gradient of hues on the image, which the VIP software cannot interpret as precisely as it interprets the stark contrast of vehicle on pavement during day lighting periods. Similarly, the effect of rain, as shown in figure 6.28, was to increase variation in speed deviations. This could again be attributed to greater headlight use in rainy conditions, or to image quality reduction with rain and mist in the air. Lastly, figure 6.29 shows that under higher traffic

volumes, the Solo Pro II one-minute speed percent deviation was more consistent. This could be attributed to an aggregation effect. When volume was higher, the one-minute mean speed was based on more vehicle speeds. If one of those vehicle speeds was misreported by the detector, it had less impact on the one-minute mean speed than a similarly misreported single speed during a low volume minute.

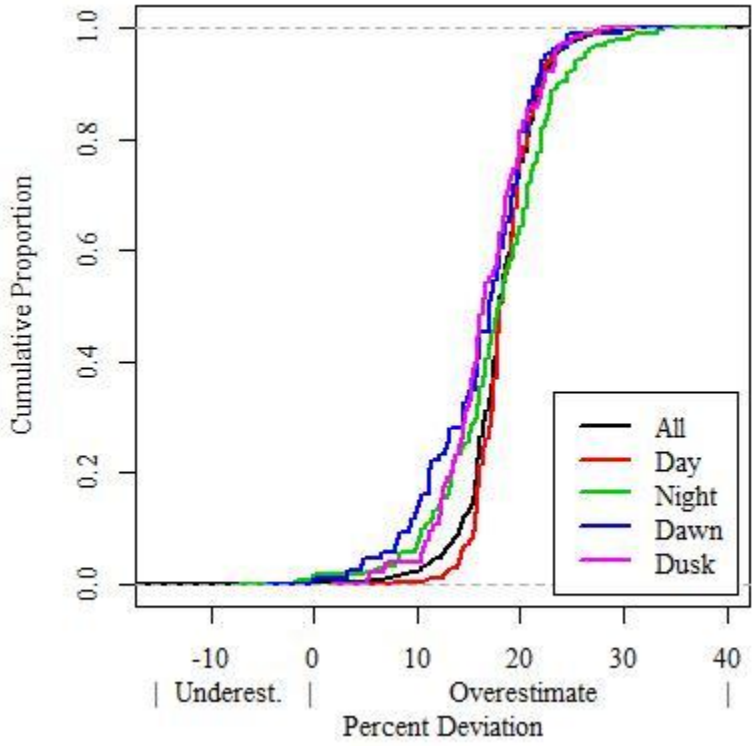

Figure 6.27: Solo Pro II One-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

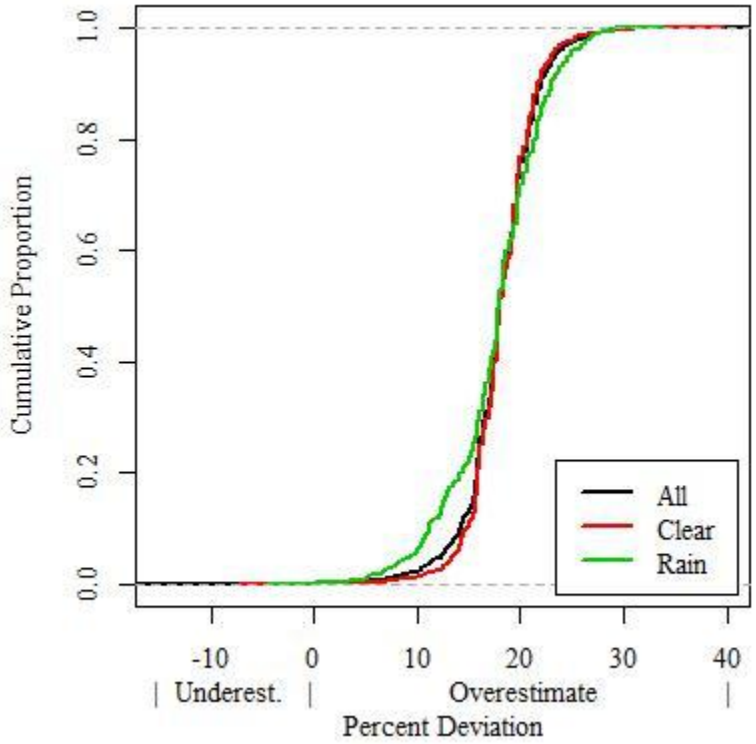

Figure 6.28: Solo Pro II One-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

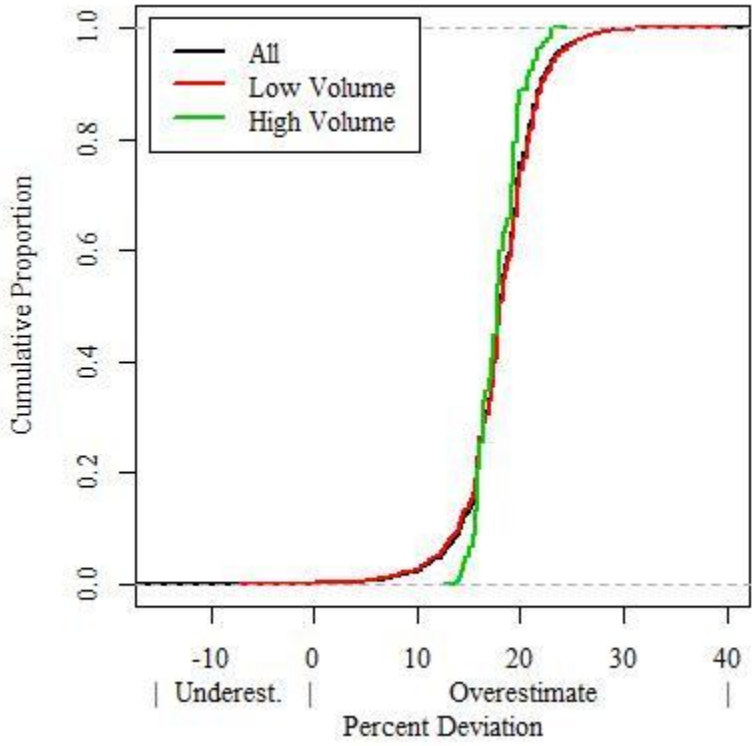

Figure 6.29: Solo Pro II One-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

Figures 6.30-6.32 depict similar plots of the effects of lighting, rain, and volume on the G4 one-minute mean speed percent deviation distributions. In figure 6.30, it appears that dawn lighting conditions shifted G4 speeds so that more one-minute mean speeds were underestimated and fewer were overestimated. No practical explanation for this trend was found. Figure 6.31 shows that the variability of G4 one-minute speed percent deviation increased in rainy weather. This could be due to disruption of the radar signal by rain droplets in the air, which in turn decreased detection precision. Figure 6.32 shows reduced variability of G4 speed percent deviation under high volume conditions. This could be attributed to an aggregation effect, as was previously explained for the Solo Pro II.

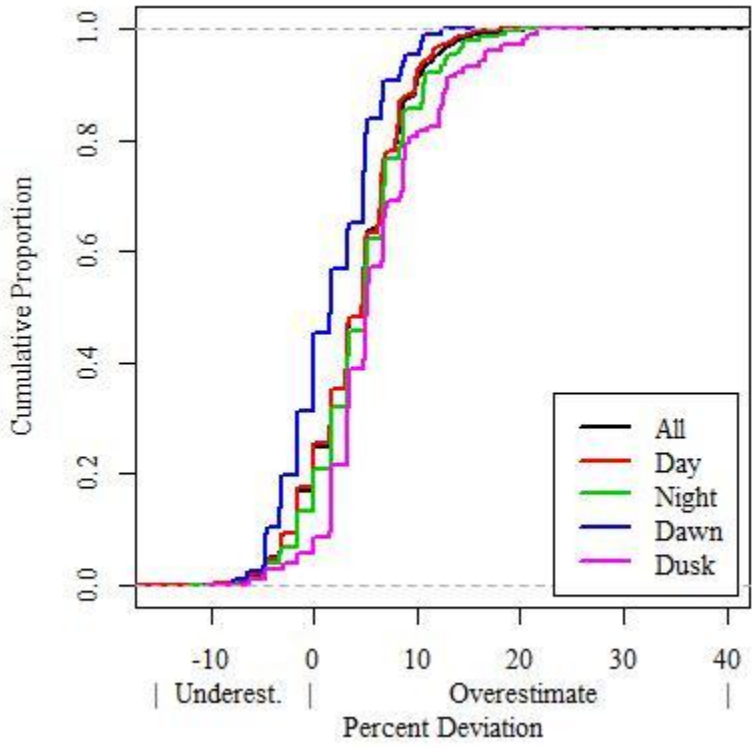

Figure 6.30: G4 One-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

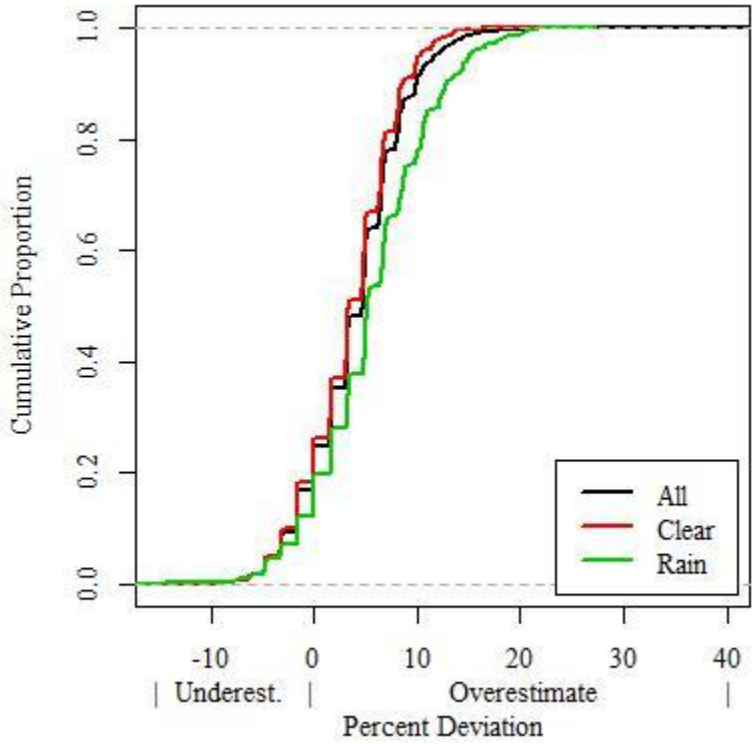

Figure 6.31: G4 One-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

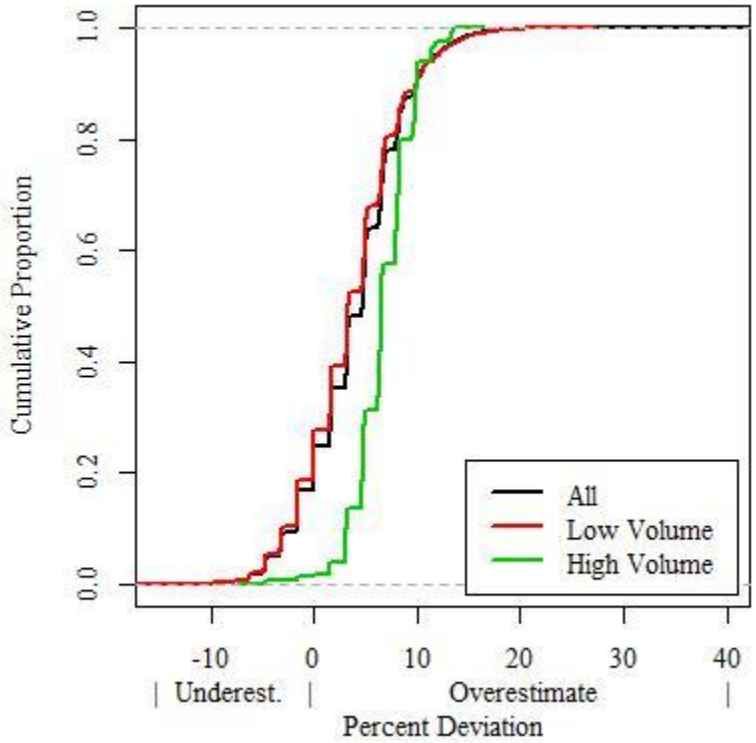

Figure 6.32: G4 One-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

Figures 6.33-6.35 show the effects of lighting, rain, and volume on the SmartSensor 105 one-minute mean speed percent deviation distributions. Figures 6.33 and 6.34 show that the SmartSensor 105 one-minute speed detection appeared to be relatively consistent under various lighting conditions and the absence or presence of rain. Figure 6.35 shows reduced variability of SmartSensor 105 speed percent deviation under high volume conditions. It was again hypothesized that this was due to an aggregation effect, as was previously posited for the Solo Pro II.

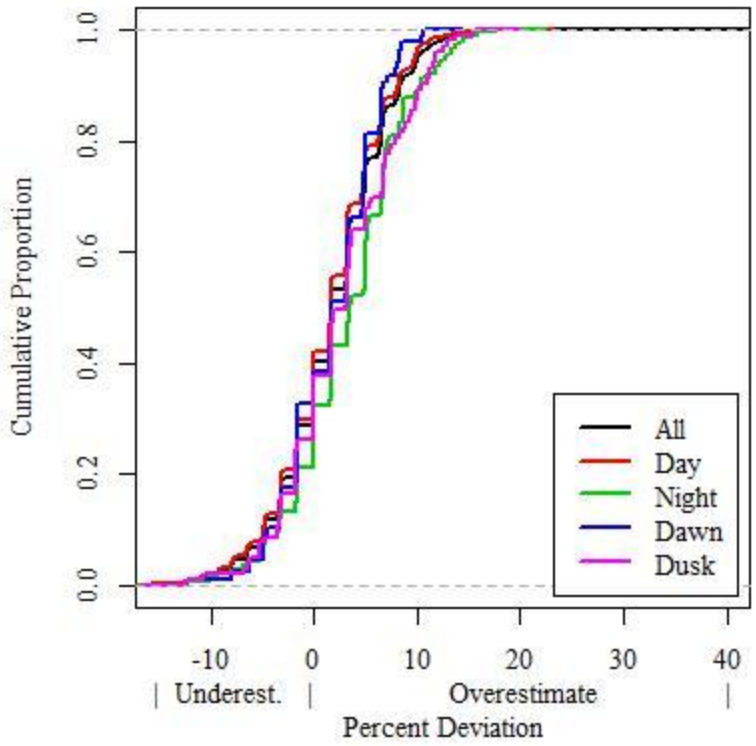

Figure 6.33: SmartSensor 105 One-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

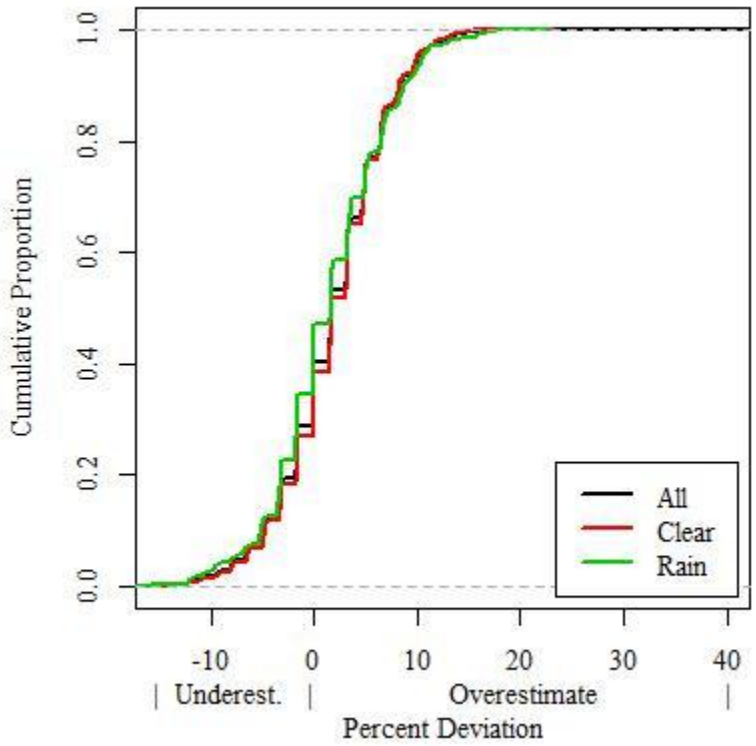

Figure 6.34: SmartSensor 105 One-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

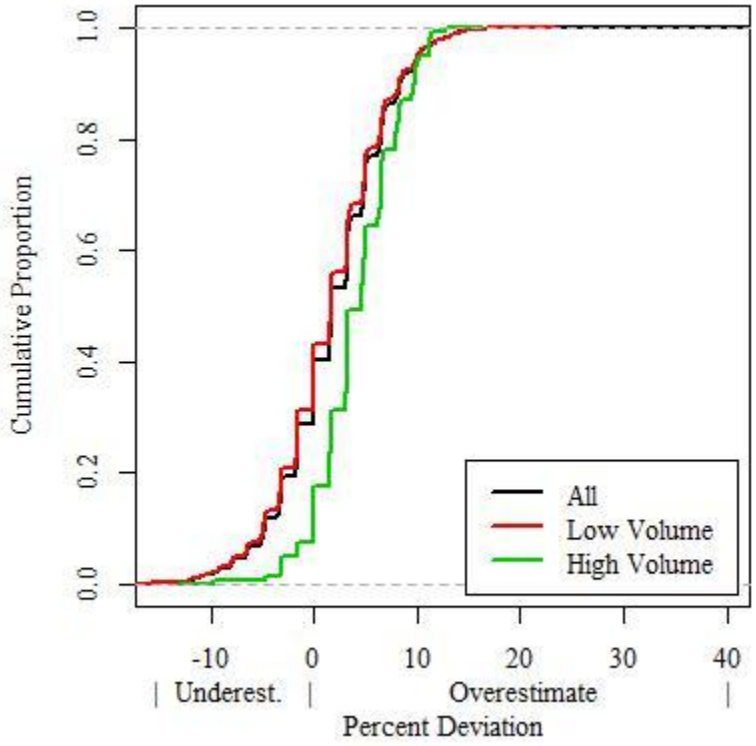

Figure 6.35: SmartSensor 105 One-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

The statistical significance of these environmental effects on speed detection was determined through ANOVA. As with the volume percent error ANOVA, this will be an unbalanced four-by-two factorial ANOVA based on the model presented in section 5.5. This analysis was performed on each detector's one-minute mean speed percent deviation, with factors for lighting (levels=Day, Night, Dawn, and Dusk) and precipitation (levels = None and Rain). In order to minimize the effects of serial correlation, thinning was performed in a manner similar to that outlined in Appendix B for one-minute volume ANOVA. The models for one-minute mean speed ANOVA dictated that a thinning factor of 10 would eliminate autocorrelation for all detectors. Statistical significance was reported at an $\alpha = 0.05$ level.

The output of the Solo Pro II speed ANOVA found in table 6.19 indicates that the intercept, as well as the effects of rain and an interaction effect between lighting and rain, were statistically significant. The results of the G4 ANOVA, found in table 6.20, indicate the mean one-minute mean speed percent deviation was significant, as was the effect of lighting and an interaction effect between lighting and rain. Lastly, the results of the SmartSensor 105 ANOVA, found in table 6.21, indicate that the mean one-minute mean speed percent deviation was statistically significant, while the effects of lighting and rain were not found to be statistically significant. As the interaction effect between lighting and rain was found not to be statistically significant for the SmartSensor 105, it was eliminated from the underlying model to provide greater power to the test of significance for the independent effects of lighting and rain, respectively.

Table 6.19: Solo Pro II One-Minute Mean Speed Percent Deviation ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	1.510	1	1687.807	0.000	*
Lighting	0.007	3	2.551	0.058	
Rain	0.014	1	15.945	0.000	*
Lighting:Rain	0.018	3	6.619	0.000	*
Residuals	0.124	139			

Table 6.20: G4 One-Minute Mean Speed Percent Deviation ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	0.165	1	104.524	0.000	*
Lighting	0.025	3	5.179	0.002	*
Rain	0.001	1	0.581	0.447	
Lighting:Rain	0.019	3	4.007	0.009	*
Residuals	0.220	139			

Table 6.21: SmartSensor 105 One-Minute Mean Speed Percent Deviation ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	0.053	1	17.851	0.000	*
Lighting	0.007	3	0.788	0.502	
Rain	0.001	1	0.214	0.645	
Residuals	0.421	142			

Lastly, multiple regression models for the one-minute mean speed percent deviation for each detector were developed to support trends observed in the graphical representation of the data. This regression was based on the equation given in section 5.6, with the dependent variable (p_i) being the mean speed percent deviation for minute i, and the first dependent variable (α) being the theoretical mean speed percent deviation for the specified detector given daylight non-rainy conditions, with a true volume of 0 vehicles. As with the other analyses in this chapter, the effects of serial correlation were minimized through data thinning performed in a manner similar to that outlined in Appendix B for one-minute volume ANOVA. The models for one-minute mean speed regression dictated that a thinning factor of 10 would eliminate autocorrelation for all detectors. Statistical significance of model factors was reported at a level of $\alpha = 0.05$.

The coefficients of the Solo Pro II one-minute mean speed percent deviation model are shown in table 6.22. The statistically significant factors in this model were the intercept, the combined effect of dawn lighting and rain, and the combined effect of dusk lighting and rain. It was hypothesized that headlight reflection off of pavement, which was made more reflective by rain, caused issues for Solo Pro II speed detection. Based on this hypothesis, it was expected that the interaction effect of night lighting and rain would also be significant. While that was not the case at an $\alpha = 0.05$ level, the p-value of 0.084 indicates that this interaction effect would have been significant under a slightly less stringent analysis. The adjusted R-squared for this model was 0.1202, indicating a low correlation between the predicted and observed values for Solo Pro II one-minute mean speed percent deviation.

Table 6.22: Solo Pro II One-Minute Mean Speed Percent Deviation Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	18.36	0.773	23.771	0.000	*
V.Truth (β1)	-0.01	0.022	-0.570	0.570	
Night (γ11)	0.69	0.957	0.718	0.474	
Dawn (γ12)	1.20	1.594	0.750	0.454	
Dusk (γ13)	0.88	1.296	0.680	0.498	
Rain (γ21)	0.71	0.869	0.821	0.413	
Night:Rain (γ31)	-3.33	1.912	-1.740	0.084	
Dawn:Rain (γ32)	-7.93	2.304	-3.440	0.001	*
Dusk:Rain (γ33)	-5.18	1.992	-2.599	0.010	*

A similar model was created next, through removal of the non-significant independent variables from the first model. The coefficients in this model are shown in table 6.23. This model had a slightly higher adjusted R-squared value of 0.1229. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more

accurately depict the effects of the significant independent variables on the Solo Pro II one-minute mean speed percent deviation.

Table 6.23: Solo Pro II One-Minute Mean Speed Percent Deviation Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	18.23	0.255	71.549	0.000	*
Dawn:Rain (γ32)	-6.08	1.519	-4.003	0.000	*
Dusk:Rain (γ33)	-3.63	1.363	-2.663	0.009	*

The coefficients of the G4 one-minute mean speed percent deviation model are shown in table 6.24. The statistically significant factors in this model were the true volume, night lighting, rain, the combined effect of dawn lighting and rain, and the combined effect of dusk lighting and rain. The adjusted R-squared for this model was 0.1845, indicating a low correlation between the predicted and observed values for G4 one-minute mean speed percent deviation.

Table 6.24: G4 One-Minute Mean Speed Percent Deviation Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	1.21	0.998	1.218	0.225	
V.Truth (β1)	0.08	0.028	2.955	0.004	*
Night (γ11)	3.76	1.236	3.041	0.003	*
Dawn (γ12)	3.62	2.059	1.760	0.081	
Dusk (γ13)	2.55	1.673	1.524	0.130	
Rain (γ21)	2.48	1.122	2.211	0.029	*
Night:Rain (γ31)	-2.63	2.469	-1.065	0.289	
Dawn:Rain (γ32)	-7.75	2.976	-2.605	0.010	*
Dusk:Rain (γ33)	5.44	2.572	2.115	0.036	*

A similar model was created by removing the independent variables not found to be significant in the first model. The resulting model showed both rain and the interaction effect of dawn and rain to be non-significant. Therefore, another model was created with these factors removed as well. The coefficients in the resulting model are shown in table 6.25. While this model had an even lower adjusted R-squared value of 0.1577, the

average effect of the significant factors from the first model on the G4 one-minute mean speed percent deviation are shown more clearly in the "Estimate" column of this model. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more accurately depict the effects of the significant independent variables on G4 one-minute mean speed percent deviation.

Table 6.25: G4 One-Minute Mean Speed Percent Deviation Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	2.76	0.746	3.697	0.000	*
V.Truth (β1)	0.05	0.024	2.002	0.047	*
Night (γ11)	2.59	1.042	2.489	0.014	*
Dusk:Rain (γ33)	9.43	1.824	5.170	0.000	*

The coefficients of the SmartSensor 105 one-minute mean speed percent deviation model are shown in table 6.26. The statistically significant factors in this model were true volume, night lighting, and the combined effect of night lighting and rain. A hypothesis could not be formulated to explain why these factors were found to be significant. The adjusted R-squared for this model was 0.0231, indicating a low correlation between the predicted and observed values for SmartSensor 105 one-minute mean speed percent deviation.

Table 6.26: SmartSensor 105 One-Minute Mean Speed Percent Deviation Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	-0.31	1.380	-0.224	0.823	
V.Truth (β1)	0.08	0.039	2.129	0.035	*
Night (γ11)	4.71	1.709	2.755	0.007	*
Dawn (γ12)	1.73	2.847	0.607	0.545	
Dusk (γ13)	0.26	2.314	0.111	0.911	
Rain (γ21)	1.21	1.552	0.780	0.437	
Night:Rain (γ31)	-7.03	3.415	-2.058	0.042	*
Dawn:Rain (γ32)	1.36	4.116	0.331	0.741	
Dusk:Rain (γ33)	1.52	3.558	0.428	0.669	

An attempt was made to create a similar model by removing the independent variables not found to be significant in the first model. This resulting model found both true volume and the interaction effect of night and rain to be non-significant. When another model was created having the intercept and night as the only factors, night was found to be non-significant. Therefore, it was concluded that none of the tested factors were significant by themselves in a model for the SmartSensor 105 one-minute mean speed percent deviation.

While the low adjusted R-squared values for these models suggests a weak linear relationship between the independent factors and the one-minute mean speed percent deviation, this is to be expected in this application due to variability in speed detection based on factors other than the environmental conditions considered herein. If it were possible, based on a model similar to one of these, to accurately predict the percent deviation in speed of a specific detector for any given minute, it would be possible to eliminate these errors. As this is not the case, we present these models in spite of their low adjusted R-squared values, in order to demonstrate the average effect of potential environmental factors (see "Estimate" column in the above tables) and to demonstrate which of these effects were consistent enough to be deemed statistically significant.

6.1.3 One-Minute Classification Analysis

The final detection parameter to be analyzed at the one-minute aggregation interval was vehicle classification. This analysis assessed the ability of each detector to correctly identify in which of three length-based bins a vehicle belonged. The three length bins were (in length): under 25 feet, 25 to 40 feet, and over 40 feet. They were intended to represent passenger vehicles, single unit heavy vehicles, and multiple unit heavy vehicles. Throughout the remainder of this section, these three classes will be referred to as *short*, *medium*, and *long* vehicles. The mean one-minute proportions of short, medium, and long vehicles, as reported in the ground truth and by each detector, are depicted in figure 6.36. These mean one-minute classification proportions are also given in table 6.27. This figure and table indicate that the Solo Pro II had a tendency to classify more vehicles as short and medium, and fewer as long, than did the ground truth. The other detectors appeared to average approximately the same proportions as the ground truth.

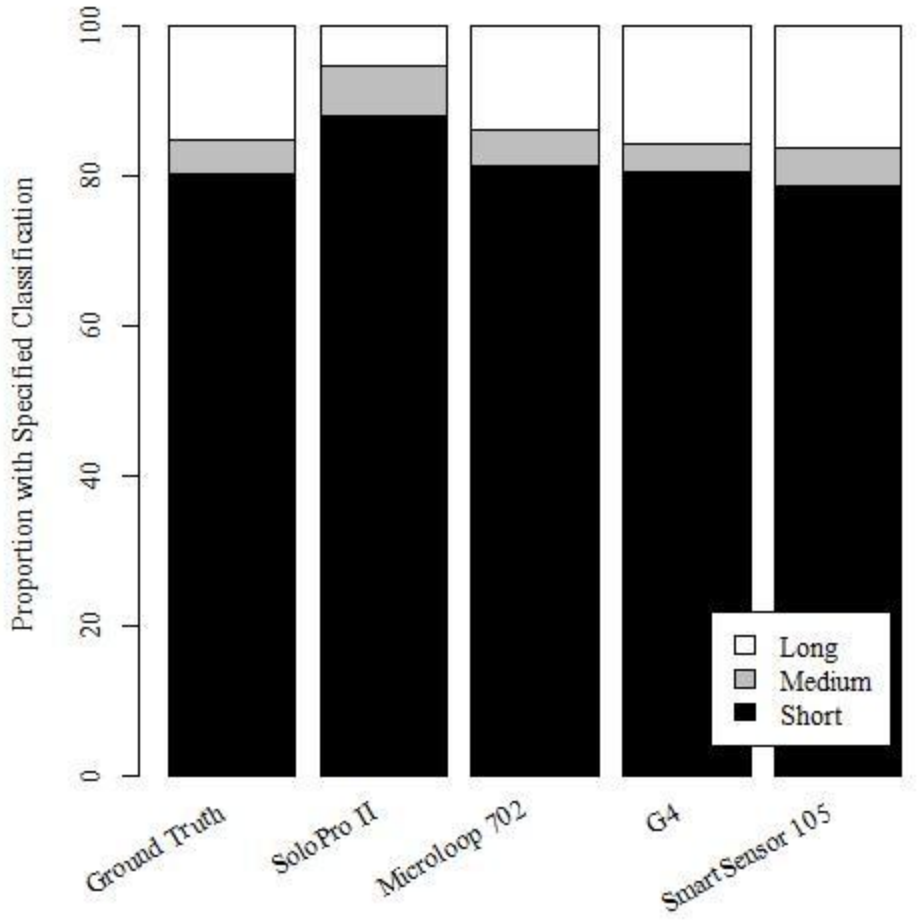

Figure 6.36: Mean One-Minute Proportion Short, Medium, and Long Vehicles Bar Chart

Table 6.27: Mean One-Minute Classification Proportions

	Ground Truth	SoloPro II	Microloop 702	G4	Smartsensor 105
Short	80.2%	88.0%	81.3%	80.4%	78.5%
Medium	4.4%	6.7%	4.7%	3.8%	5.0%
Long	15.4%	5.4%	13.9%	15.8%	16.5%

These tendencies, indicated by the mean proportions, can be further investigated by examining the distributions of one-minute percent short, medium, and long vehicles, as reported by the ground truth and each detector. Box plots of the distributions for percent short, medium, and long vehicles are given in figures 6.37-6.39. It is shown in these figures that distributions of Microloop 702, G4, and SmartSensor 105 one-minute

percent short, medium, and long vehicles closely resembled the ground truth distributions. It is worth noting that while the Solo Pro II long and short vehicle proportion distributions appeared to differ greatly from the ground truth distributions, the Solo Pro II proportion medium vehicle distribution bore a greater resemblance to the ground truth proportion medium vehicle distribution.

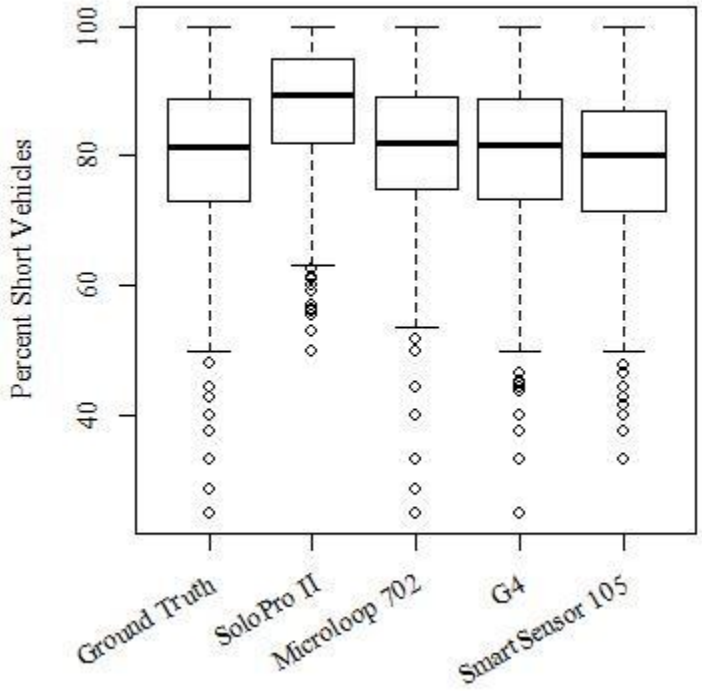

Figure 6.37: Box Plot of One-Minute Percent Short Vehicle Distributions

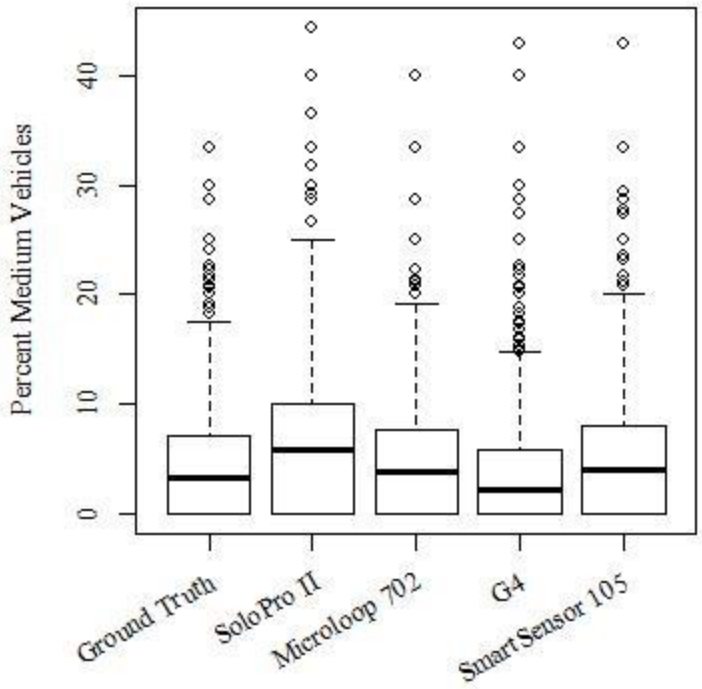

Figure 6.38: Box Plot of One-Minute Percent Medium Vehicle Distributions

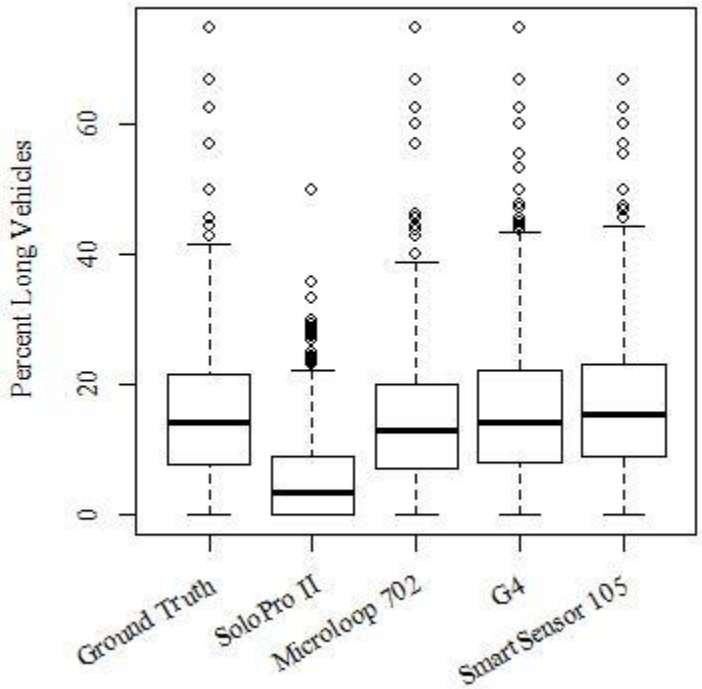

Figure 6.39: Box Plot of One-Minute Percent Long Vehicle Distributions

Scatter plots in figures 6.40-6.42 illustrate the correlations between one-minute true and detected percent short, medium, and long vehicles. The correlation coefficients included in the figures indicate that the G4 exhibited the strongest correlations between reported and ground truth classification proportions, while the Microloop 702 and Smartsensor 105 also exhibited good correlation with the ground truth. The Solo Pro II had lower correlation coefficients, and appeared to over-report short vehicle proportions and under-report long vehicle proportions.

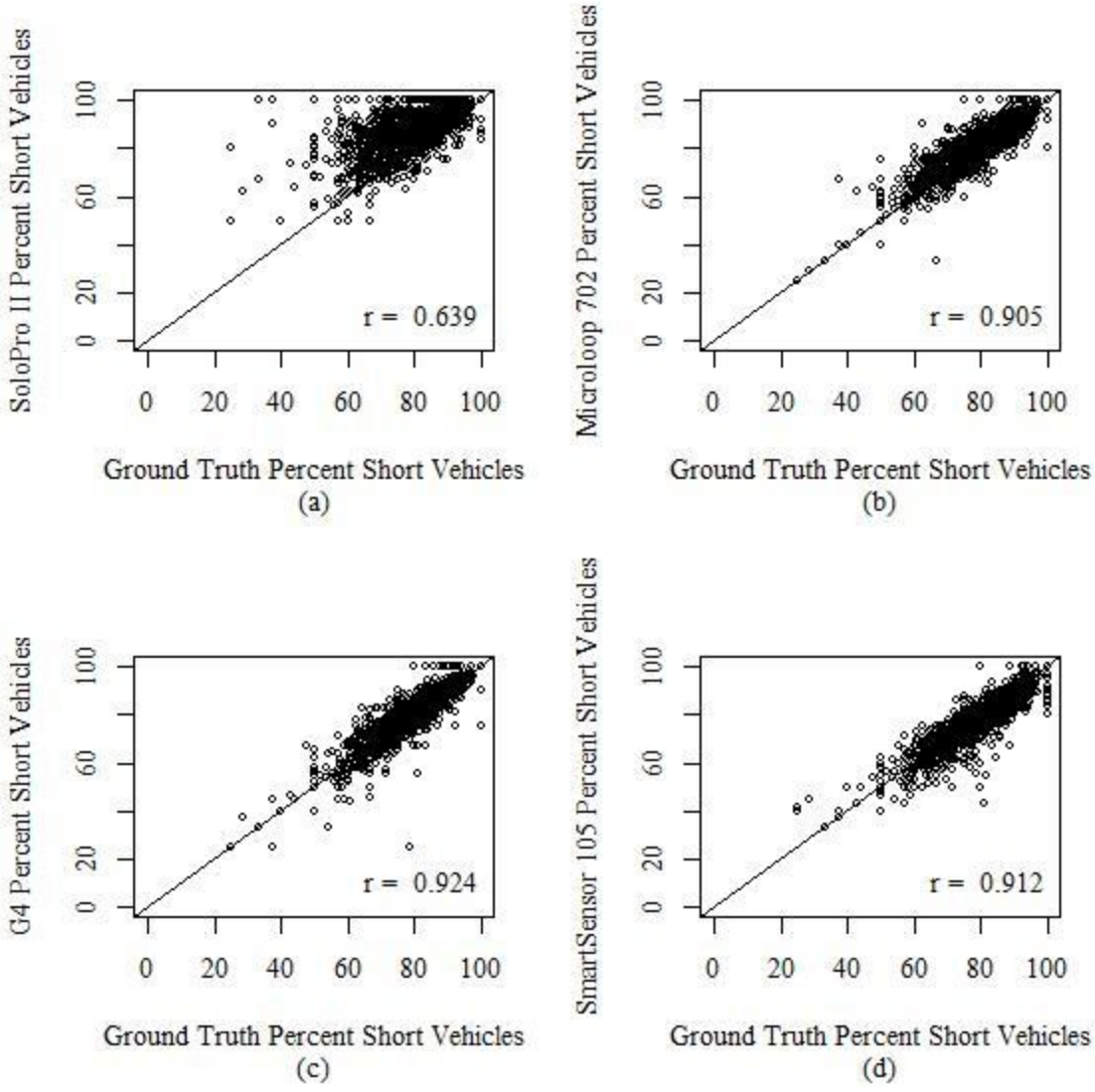

Figure 6.40: One-Minute Percent Short Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

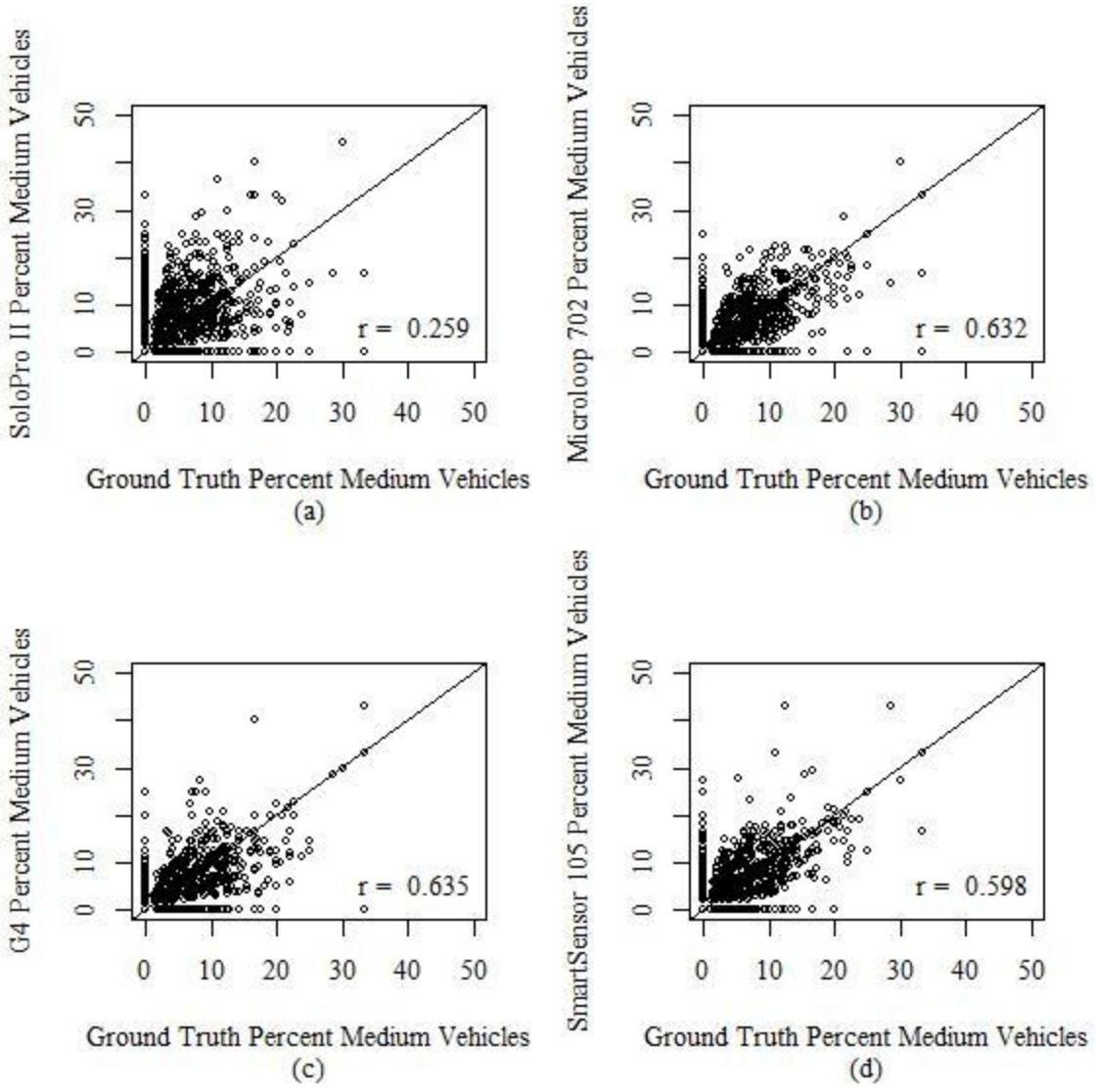

Figure 6.41: One-Minute Percent Medium Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

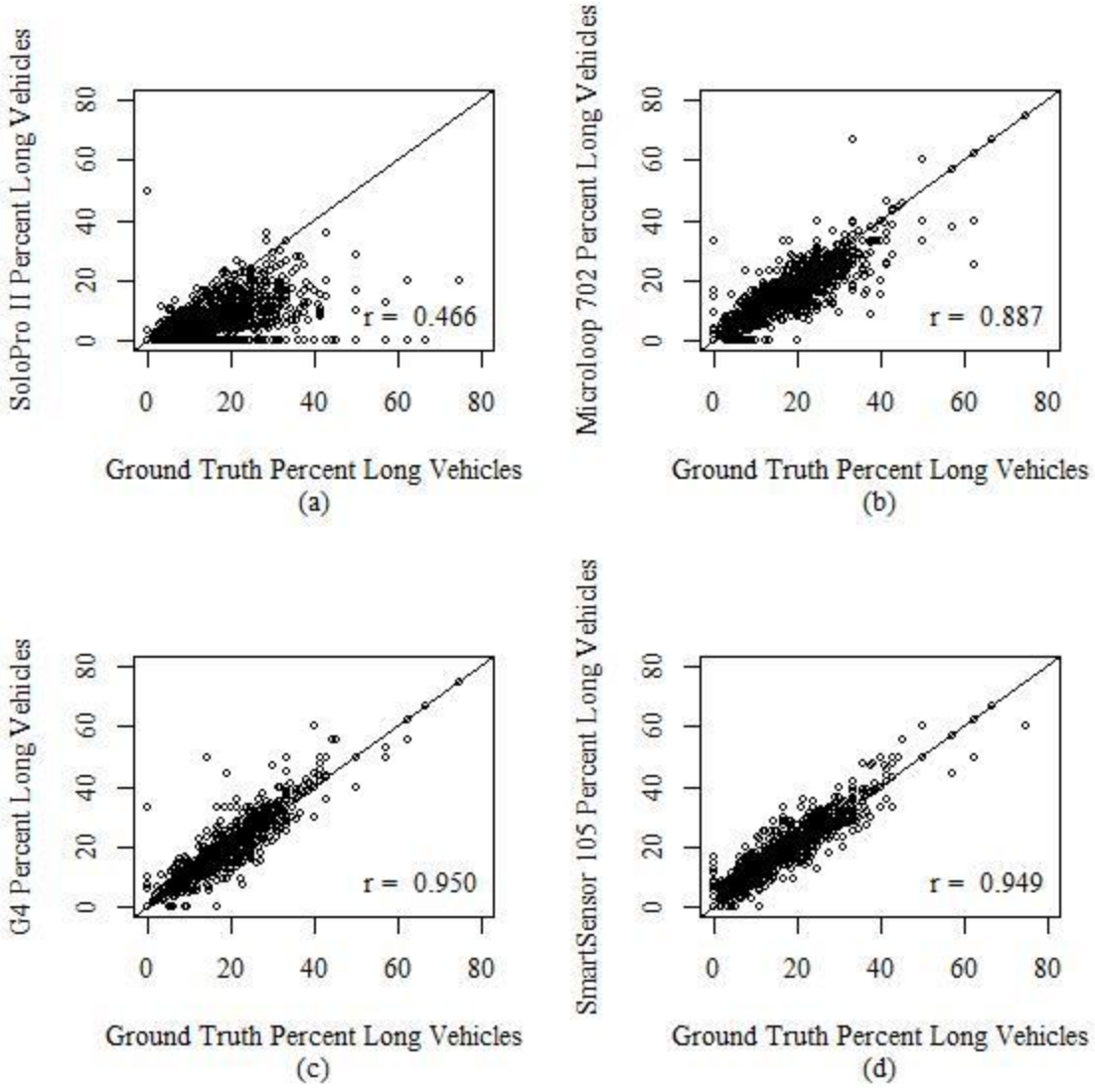

Figure 6.42: One-Minute Percent Long Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

The next step in the analysis was to determine the one-minute proportion errors for each minute in the dataset for each detector. This was accomplished by subtracting the ground truth short vehicle proportion from the detector-reported short vehicle proportion for each minute, and likewise for the medium and long vehicle proportions. A positive error value indicates that the detector reported a higher percentage of the specified class in a given minute than the ground truth percentage. A negative error value

indicates that in a given minute the detector reported a lower percentage of vehicles of the specified class than were reported by the ground truth percentage. An error value of zero indicates that the detector reported a proportion of the specified class equal to the ground truth proportion belonging to that class for the given minute.

The distributions of these errors for the short, medium, and long vehicles are shown in the histograms in figures 6.43-6.45. The peakedness of the distributions for the Microloop 702, G4, and SmartSensor 105 in these figures indicates that for many of the data intervals these detectors exhibited small or non-existent departures from the ground truth proportions. The Solo Pro II histograms for the short and long proportions in figures 6.43 and 6.45 indicate that this detector had a bias for over-reporting the proportion of short vehicles and under-reporting the proportion of long vehicles.

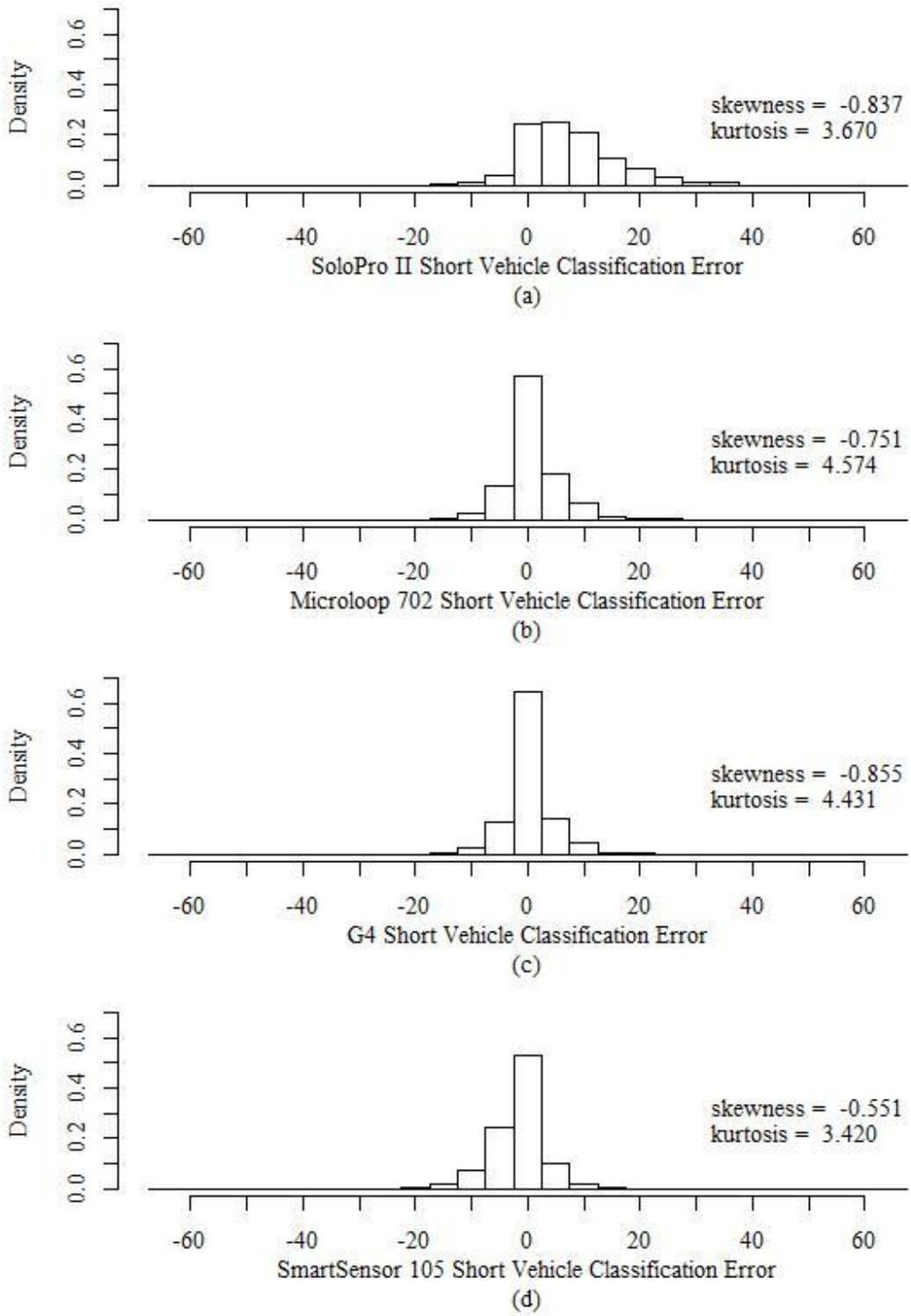

Figure 6.43: Histograms of One-Minute Percent Short Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

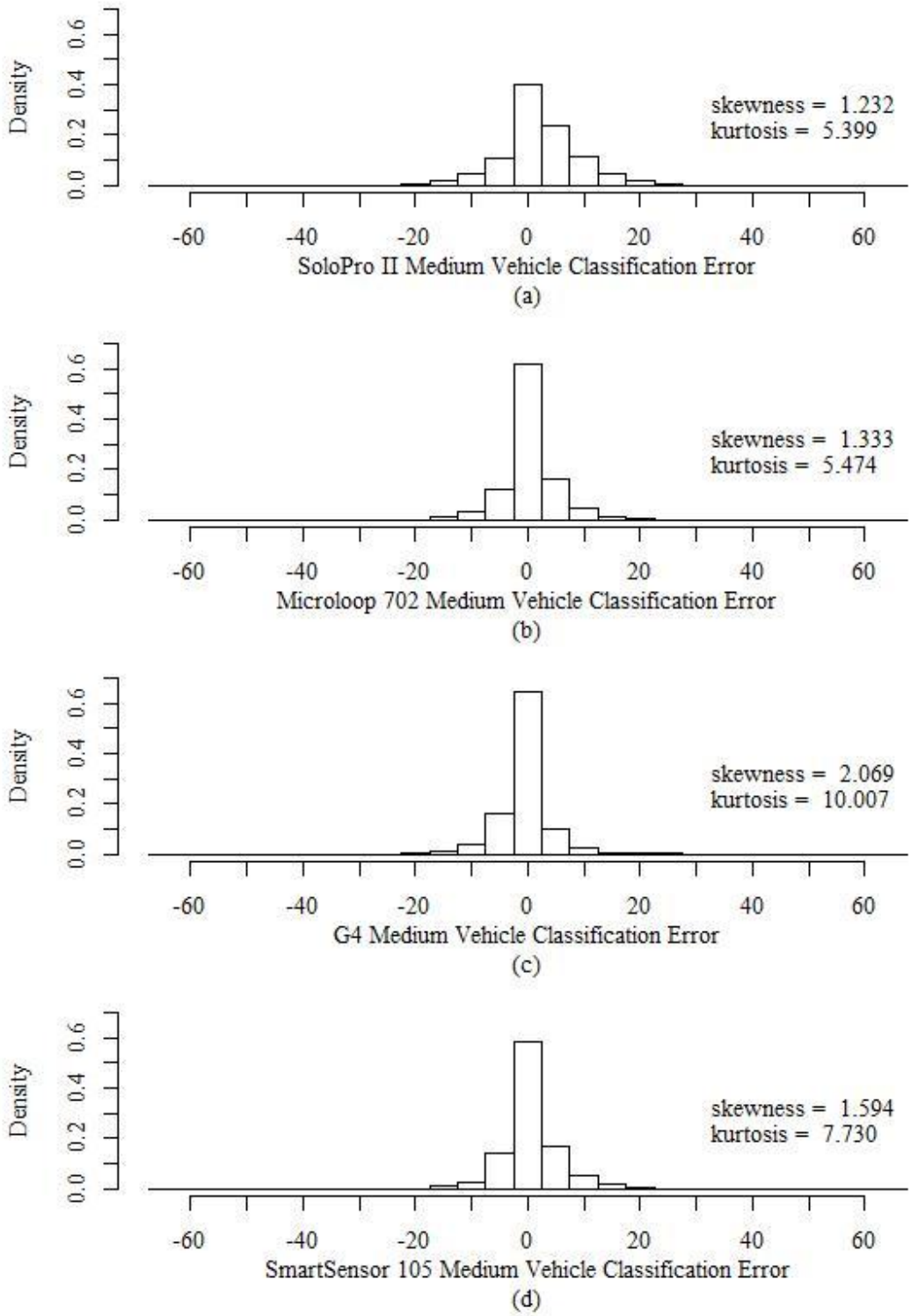

Figure 6.44: Histograms of One-Minute Percent Medium Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

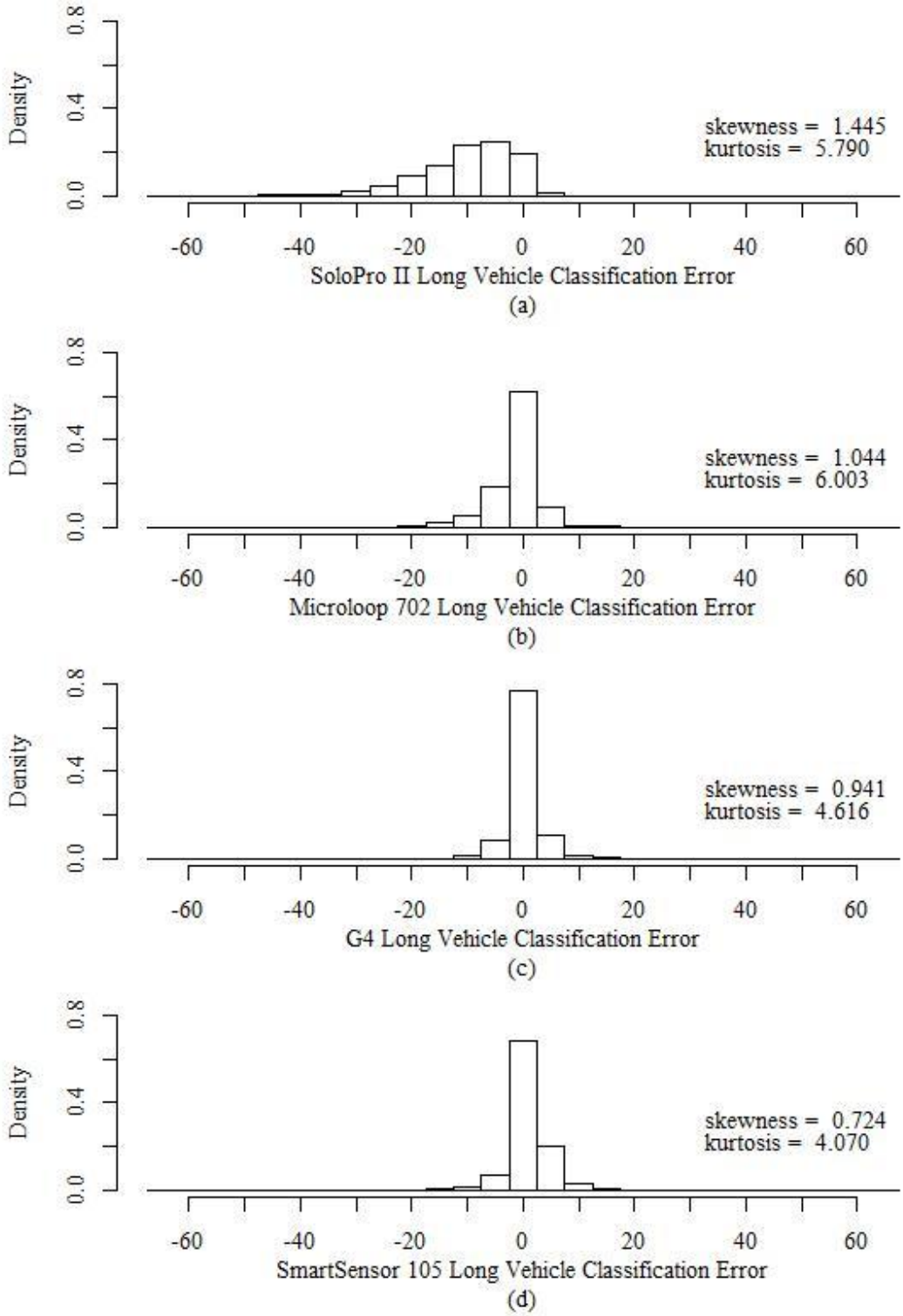

Figure 6.45: Histograms of One-Minute Percent Long Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

Another visual representation that draws attention to the distributions of these one-minute proportion errors is a cumulative distribution function. Figures 6.46-6.48 depict cumulative distribution functions for short, medium, and long vehicle proportions for each detector, which illustrate the nature of the undercounting and overcounting of the respective classes. These figures again show that the Solo Pro II had the largest errors in classification of the analyzed detectors. The distributions of the other three detectors appeared to be very similar.

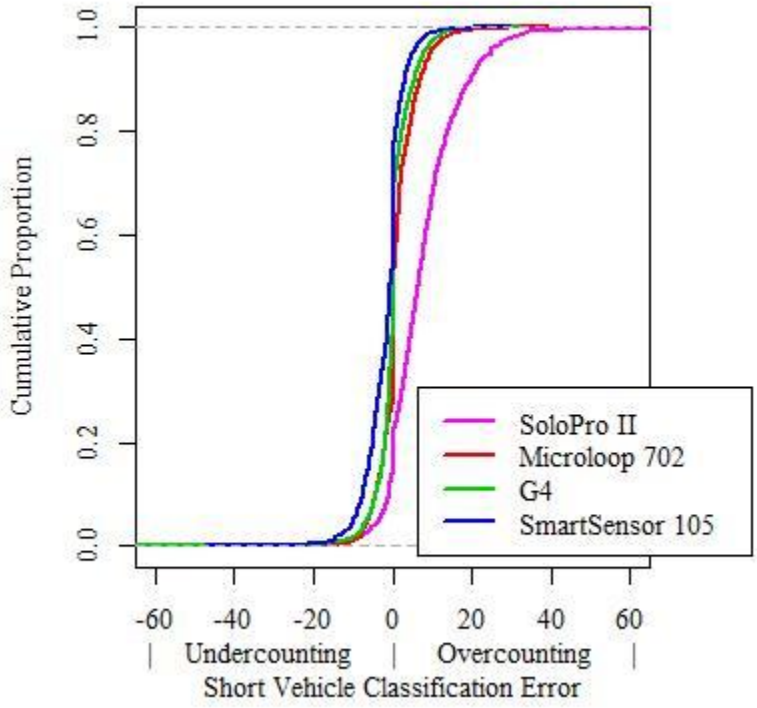

Figure 6.46: One-Minute Percent Short Vehicles Error Cumulative Distribution Plot

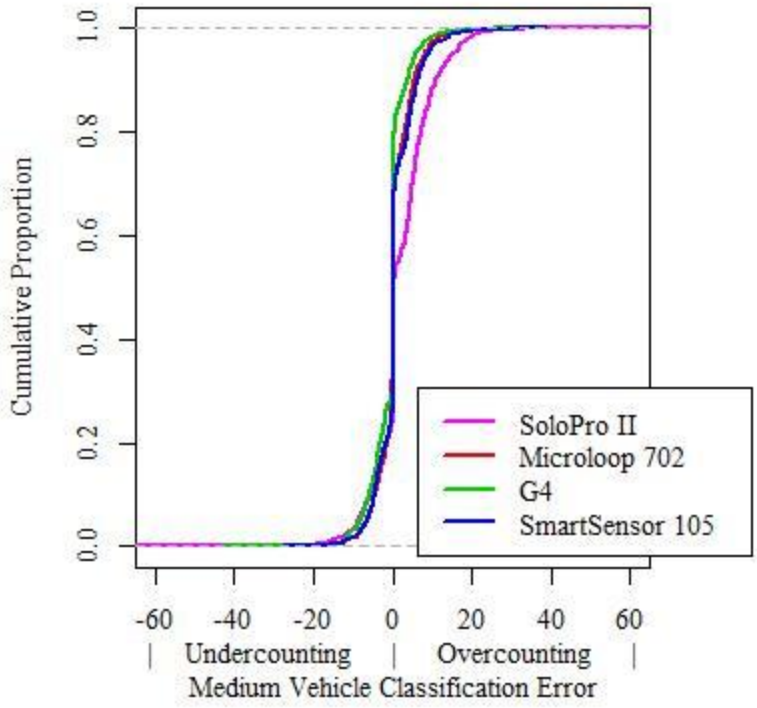

Figure 6.47: One-Minute Percent Medium Vehicles Error Cumulative Distribution Plot

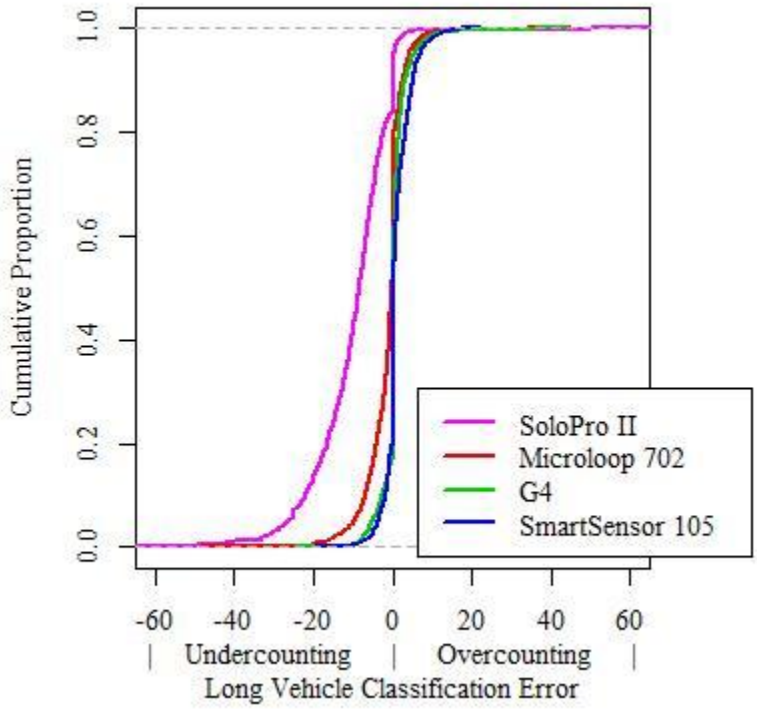

Figure 6.48: One-Minute Percent Long Vehicles Error Cumulative Distribution Plot

An additional statistic was used to define the classification error without replicating the analyses in triplicate for short, medium, and long vehicle classes. This statistic will be referred to as the one-minute classification error percentage, and is defined by the following equation:

$$CE_{ij} = \frac{\left|pst_i - psd_{ij}\right| + \left|pmt_i - pmd_{ij}\right| + \left|plt_i - pld_{ij}\right|}{2} \quad (6.1)$$

where: pst_i is the true percent short vehicles for minute i,

psd_{ij} is the percent short vehicles for minute i reported by detector j,

pmt_i is the true percent medium vehicles for minute i,

pmd_{ij} is the percent medium vehicles for minute i reported by detector j,

plt_i is the true percent long vehicles for minute i, and

pld_{ij} is the percent long vehicles for minute i reported by detector j.

The factor of two in the denominator is necessary to eliminate overestimation of misclassification errors. The need for this factor is demonstrated by the following hypothetical example: During a minute with 10 short, 0 medium, and 0 long vehicles, a detector reports 9 short, 1 medium, and 0 long vehicles. The intuitive classification error percentage is 10%, as 1 of 10 vehicles was incorrectly classified. The numerator of the above equation would equal 20% as $\left|pst_i - psd_{ij}\right| = 10\%$ and $\left|pmt_i - pmd_{ij}\right| = 10\%$. The denominator eliminates the double-counting of vehicles that are missed in one class and counted in another class. Summary statistics for the classification error percentage are given in table 6.28.

Table 6.28 One-Minute Classification Error Percentage Summary Statistics

	Mean	Median	Standard Deviation
Solo Pro II	12.0%	10.5%	8.84
Microloop 702	4.4%	3.4%	4.58
G4	3.4%	2.2%	4.17
SmartSensor 105	4.2%	3.5%	4.17

The statistical significance of the effect of environmental factors on the various detectors' ability to classify vehicles was determined through ANOVA. As with the volume percent error ANOVA, this was an unbalanced four-by-two factorial ANOVA based on the model presented in section 5.5. This analysis was performed on each detector's one-minute classification error percentage, with factors for lighting (levels=Day, Night, Dawn, and Dusk) and precipitation (levels = None and Rain). In order to minimize the effects of serial correlation, thinning was performed in a manner similar to that outlined in Appendix B for one-minute volume ANOVA. The models for one-minute classification error percentage ANOVA dictated that a thinning factor of 5 would eliminate autocorrelation for all detectors. Statistical significance was reported at a level of $\alpha = 0.05$. The initial models for each detector were tested with consideration of a potential interaction effect between lighting and rain. As this interaction effect was found to not be statistically significant for any of the detector's models, it was removed from the models to increase the statistical power of the analysis on the independent effects of lighting and rain factors.

The output of the Solo Pro II classification ANOVA, found in table 6.29, indicates that the intercept, as well as the effects of lighting and the effects of rain, were statistically significant. The results of the Microloop 702 ANOVA found in table 6.30 indicate that the intercept was the only statistically significant parameter in the model.

The results of the G4 ANOVA found in table 6.31 indicate that the intercept, as well as the effects of lighting and the effects of rain, were statistically significant. Lastly, the results of the SmartSensor 105 ANOVA, found in table 6.32, indicate that the intercept was the only statistically significant parameter in the model.

Table 6.29: Solo Pro II One-Minute Classification Error Percentage ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	22595.7	1	348.597	0.000	*
Lighting	2759.9	3	14.193	0.000	*
Rain	394.8	1	6.091	0.014	*
Residuals	18732.7	289			

Table 6.30: Microloop 702 One-Minute Classification Error Percentage ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	1742.5	1	104.022	0.000	*
Lighting	91.7	3	1.825	0.143	
Rain	28.9	1	1.726	0.190	
Residuals	4841.1	289			

Table 6.31: G4 One-Minute Classification Error Percentage ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	2020.0	1	89.333	0.000	*
Lighting	271.4	3	4.001	0.008	*
Rain	100.1	1	4.425	0.036	*
Residuals	6534.8	289			

Table 6.32: SmartSensor 105 One-Minute Classification Error Percentage ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	1976.8	1	96.604	0.000	*
Lighting	45.7	3	0.744	0.526	
Rain	1.2	1	0.059	0.808	
Residuals	5913.9	289			

Next, multiple regression models for the one-minute classification error percentage for each detector were developed to support trends noticed in the graphical representation of the data. This regression was based on the equation given in section 5.6,

with the dependent variable (p_i) being the classification error percentage for minute i, and the first dependent variable (α) being the theoretical classification error percentage for the specified detector given daylight, non-rainy conditions with true volume of 0 vehicles. As with other analyses in this chapter, the effects of serial correlation were minimized through data thinning, performed in a manner similar to that outlined in Appendix B for one-minute volume ANOVA. The models for one-minute classification error percentage regression dictated that a thinning factor of 5 would eliminate autocorrelation for all detectors. Statistical significance of model factors was reported at a level of $\alpha = 0.05$.

The Solo Pro II one-minute classification error percentage model coefficients are shown in table 6.33. The statistically significant factors in this model were the intercept, true volume, and night lighting. It was hypothesized that the true volume was significant because higher volume periods generally had higher short vehicle proportions, which diminished the Solo Pro II's tendency to overestimate short vehicle proportion. The increase in classification error under night lighting conditions was attributed to the impact of vehicle headlights. The adjusted R-squared for this model was 0.1616, indicating a low correlation between the predicted and observed values for Solo Pro II one-minute classification error percentage.

Table 6.33: Solo Pro II One-Minute Classification Error Percentage Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	13.64	1.422	9.594	0.000	*
V.Truth (β1)	-0.11	0.040	-2.856	0.005	*
Night (γ11)	7.79	1.791	4.349	0.000	*
Dawn (γ12)	-2.33	2.965	-0.785	0.433	
Dusk (γ13)	2.78	2.429	1.144	0.254	
Rain (γ21)	2.44	1.611	1.513	0.131	
Night:Rain (γ31)	-3.05	3.587	-0.849	0.396	
Dawn:Rain (γ32)	1.30	4.211	0.309	0.758	
Dusk:Rain (γ33)	-5.07	3.818	-1.328	0.185	

A similar model was created, removing independent variables not found to be significant in the first model. The coefficients in this model are shown in table 6.34. This model had a slightly higher adjusted R-squared value of 0.1658. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more accurately depict the effects of the significant independent variable on Solo Pro II one-minute classification error percentage.

Table 6.34: Solo Pro II One-Minute Classification Error Percentage Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	14.79	1.036	14.280	0.000	*
V.Truth (β1)	-0.14	0.034	-4.170	0.000	*
Night (γ11)	6.80	1.472	4.620	0.000	*

The coefficients of the Microloop 702 one-minute classification error percentage model are shown in table 6.35. The only statistically significant factor in this model was the intercept. The adjusted R-squared for this model was 0.0190, indicating a low correlation between the predicted and observed values for Microloop 702 one-minute classification error percentage.

Table 6.35: Microloop 702 One-Minute Classification Error Percentage Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	5.14	0.731	7.027	0.000	*
V.Truth (β1)	-0.04	0.021	-1.796	0.074	
Night (γ11)	1.37	0.921	1.485	0.139	
Dawn (γ12)	-1.31	1.525	-0.861	0.390	
Dusk (γ13)	-0.17	1.250	-0.138	0.891	
Rain (γ21)	-1.06	0.829	-1.283	0.201	
Night:Rain (γ31)	-2.15	1.846	-1.163	0.246	
Dawn:Rain (γ32)	0.61	2.166	0.282	0.778	
Dusk:Rain (γ33)	0.70	1.964	0.358	0.721	

The G4 one-minute mean speed percent deviation model coefficients are shown in table 6.36. The statistically significant factors in this model were the intercept, true volume, and the combined effect of dusk lighting and rain. The impact of ground truth volume on this model could be attributed to increased short vehicle proportions under higher volume conditions. It was noted earlier (in the analysis of one-minute volume) that the G4 was adversely affected by heavy rain occurring during one of the dusk data intervals. It was hypothesized that this heavy rain was the reason that the combined effect of dusk and rain was found to be significant in this model. The adjusted R-squared for this model was 0.0627, indicating a low correlation between the predicted and observed values for G4 one-minute classification error percentage.

Table 6.36: G4 One-Minute Classification Error Percentage Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	5.03	0.845	5.954	0.000	*
V.Truth (β1)	-0.05	0.024	-1.979	0.049	*
Night (γ11)	-1.89	1.064	-1.775	0.077	
Dawn (γ12)	-2.07	1.761	-1.175	0.241	
Dusk (γ13)	0.32	1.443	0.223	0.824	
Rain (γ21)	-0.03	0.957	-0.029	0.977	
Night:Rain (γ31)	1.72	2.131	0.805	0.422	
Dawn:Rain (γ32)	0.54	2.501	0.218	0.828	
Dusk:Rain (γ33)	5.98	2.268	2.639	0.009	*

A similar model was created excluding independent variables not found to be significant in the first model. This model showed the ground truth volume to be non-significant. Therefore, it was removed and another model created. The coefficients in this resulting model are shown in table 6.37. While this model had an even lower adjusted R-squared value of 0.0609, the average effect of the significant factors from the first model on the G4 one-minute classification error percentage are shown more accurately in the "Estimate" column of this model. While the estimates of the significant factors in the first model were affected by the inclusion of additional non-significant independent variables, the estimates in this model more accurately depict the effects of the significant independent variables on G4 one-minute classification error percentage.

Table 6.37: G4 One-Minute Classification Error Percentage Significant Factors Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	3.53	0.279	12.630	0.000	*
Dusk:Rain (γ33)	7.14	1.595	4.473	0.000	*

The SmartSensor 105 one-minute mean speed percent deviation model coefficients are given in table 6.38. The only statistically significant factor in this model was the intercept. The adjusted R-squared for this model was -0.0137, indicating a low correlation between the predicted and observed values for SmartSensor 105 one-minute classification error percentage.

Table 6.38: SmartSensor 105 One-Minute Classification Error Percentage Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	4.58	0.815	5.620	0.000	*
V.Truth (β1)	-0.01	0.023	-0.279	0.780	
Night (γ11)	0.57	1.026	0.554	0.580	
Dawn (γ12)	-0.04	1.699	-0.023	0.982	
Dusk (γ13)	-0.75	1.392	-0.541	0.589	
Rain (γ21)	0.61	0.923	0.656	0.512	
Night:Rain (γ31)	-1.83	2.055	-0.890	0.374	
Dawn:Rain (γ32)	-0.44	2.413	-0.182	0.856	
Dusk:Rain (γ33)	-2.08	2.188	-0.952	0.342	

The extremely low adjusted R-squared values for these models suggest that volume, lighting, and rain factors were not appropriate variables for predicting the classification error percentage. The models were presented in spite of their low adjusted R-squared values in order to demonstrate the average effect of potential environmental factors (see "Estimate" column in the above tables) and demonstrate which of these effects were consistent enough to be deemed statistically significant.

Throughout this analysis of one-minute classification, one observation has recurred. The Solo Pro II appears to have a propensity for misclassifying long vehicles as short. Figures 6.49-6.51 graphically represent the extent of this issue and show that the problem was exacerbated during night lighting conditions. One potential practical explanation for this is that the headlights of the vehicle were detected while the body of the vehicle was not. This would potentially cause the headlights of a long vehicle to register a vehicle length of less than 25 feet.

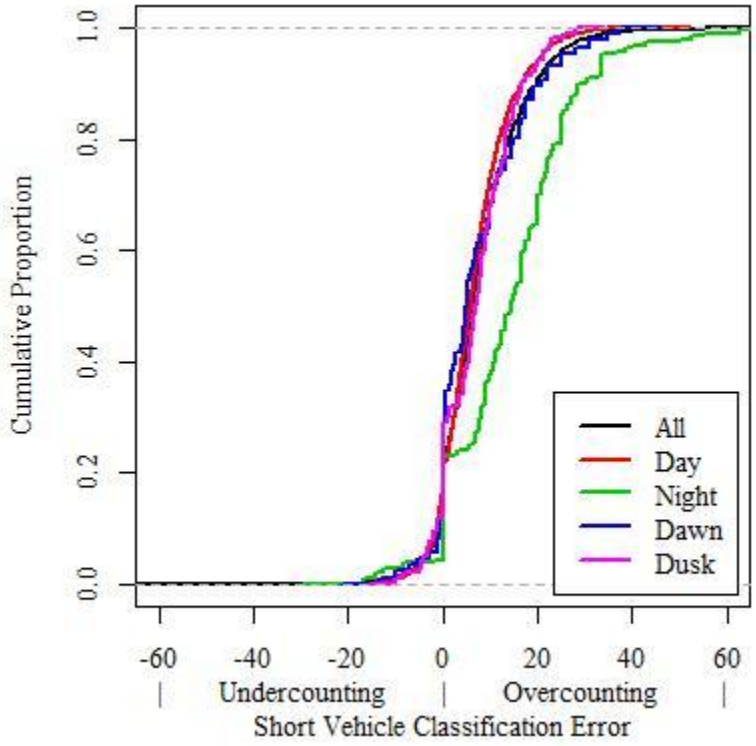

Figure 6.49: Solo Pro II One-Minute Percent Short Vehicles Error Lighting Factor Cumulative Distribution Plot

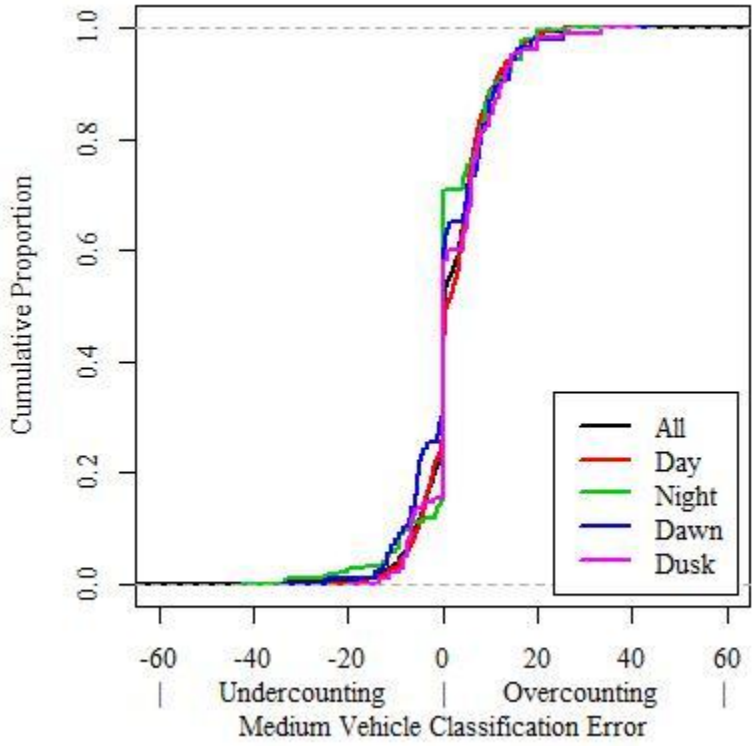

Figure 6.50: Solo Pro II One-Minute Percent Medium Vehicles Error Lighting Factor Cumulative Distribution Plot

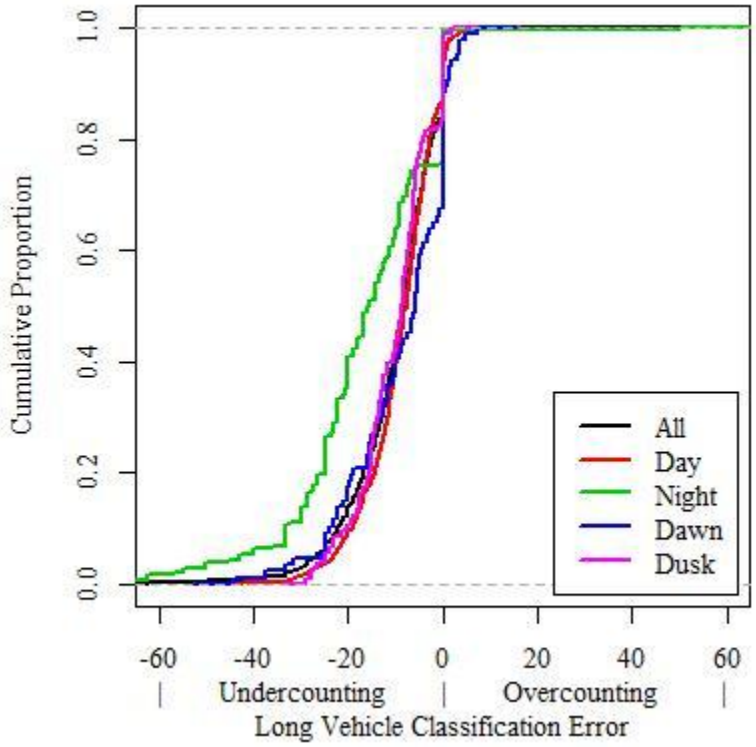

Figure 6.51: Solo Pro II One-Minute Percent Long Vehicles Error Lighting Factor Cumulative Distribution Plot

6.2 Five-Minute and Fifteen-Minute Aggregation Interval Analysis

In addition to the aggregate analysis performed at the one-minute interval, similar analyses were replicated at five-minute and fifteen-minute intervals. Due to the repetitive nature of these analyses and the degree to which the results were similar to the one-minute analysis results, a full description of these analyses is not given in this thesis. However, the differences introduced by various aggregation intervals are highlighted in this section. Additionally, many of the five-minute and fifteen-minute counterparts to the figures and tables in the one-minute analysis are given in appendices D and E.

6.2.1 Five-Minute and Fifteen-Minute Volume Analysis

The first and most noteworthy difference that occurred with more extensive aggregation was the loss of information. Most of the more specific observations that follow stem from this initial finding. For example, as scatter plots were developed and correlation

coefficients calculated for detector versus ground truth volumes, correlation coefficients increased with the aggregation interval, as shown in table 6.39. Due to this loss of variability, the volume MAPE and variance of volume percent error for each detector decreased from one-minute to five-minute and from five-minute to fifteen-minute aggregation intervals.

Table 6.39: Interval Volume Correlation Coefficients At Various Aggregation Levels

	1-minute	5-minute	15-minute
Solo Pro II	0.992	0.996	0.997
Microloop 702	0.991	0.994	0.995
G4	0.993	0.997	0.998
SmartSensor 105	0.910	0.925	0.938

Regarding the analysis of volume inequality using Theil's inequality coefficient and its proportional components, the actual inequality coefficient decreased with greater aggregation, similar to MAPE. It was also noted that the bias proportion and variance proportion both increased with greater aggregation, while the covariance (unexplained) proportion decreased with greater aggregation. Based on equations 5.9-5.11, and the understanding that mean volumes are larger over longer aggregation intervals, and that the variance of observations decreasing with greater aggregation, these trends follow logically.

When the effects of various lighting, rain, and traffic volume conditions on volume detection were considered at different aggregation intervals, the same trends were recognizable at each level of aggregation. The cumulative distribution plots of five-minute and fifteen-minute volume percent error in appendices D and E have the same basic shapes as the one-minute cumulative distribution plots presented earlier in this chapter, but generally have curves that are less smooth, since, when the same data are

aggregated over longer intervals, the result is fewer observation points from which to create the cumulative distribution curves.

6.2.2 Five-Minute and Fifteen-Minute Speed Analysis

The analyses of five-minute and fifteen-minute mean speeds gave results very similar to the one-minute mean speed analysis, the primary difference being reduced variability of interval mean speeds at greater aggregation intervals. This can be seen in table 6.40, where the standard deviation of fifteen-minute mean speeds was lower than those of the five-minute mean speeds for each detector. As with the aggregation of volume data, the interval mean speed correlation coefficients with respect to the baseline Microloop 702 increased for each detector as aggregation interval length increased.

Table 6.40: Five-Minute and Fifteen-Minute Mean Speed Summary Statistics

	Five-Minute			Fifteen-Minute		
	Mean	Median	Standard Deviation	Mean	Median	Standard Deviation
Solo Pro II	72	73	2.54	72	73	2.37
Microloop 702	61	62	1.88	61	62	1.78
G4	64	63	2.21	64	64	2.09
SmartSensor 105	62	63	2.60	62	63	2.14

* all units are (mph)

Regarding the analysis of speed inequality using Theil's inequality coefficient and its proportional components, the actual inequality coefficient decreased with greater aggregation, as it did for volume analysis. Also, the bias proportions and variance proportions increased with greater aggregation, while the covariance (unexplained) proportion decreased, for the same reasons provided for the volume application of Theil's inequality coefficient. Lastly, the shapes of speed percent deviation cumulative distribution plots were similar at various aggregation intervals, with a slight increase in

the steepness of the middle of some of these plots with greater aggregation due to reduced variability. These plots can be found in appendices D and E.

6.2.3 Five-Minute and Fifteen-Minute Classification Analysis

The reduced variability with greater aggregation becomes most obvious upon analysis of classification at five-minute and fifteen-minute intervals. One-minute intervals can produce extremely diverse proportions of short, medium, and long vehicles (especially during very low-volume periods throughout the night when three long vehicles out of five total vehicles in a minute can cause a long vehicle proportion of 60%). When aggregation over a longer temporal interval is considered, chance distributions of vehicle classes such as this balance out and variability in the data is decreased. This can be readily seen by comparing the five-minute and fifteen-minute percent long vehicle distributions in figures 6.52 and 6.53 with each other and the one-minute distributions in figure 6.39.

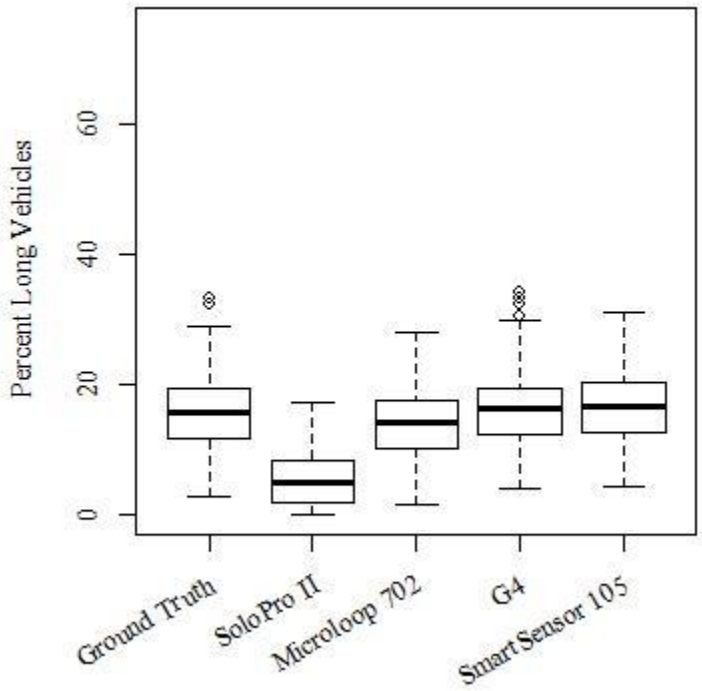

Figure 6.52: Box Plot of Five-Minute Percent Long Vehicle Distributions

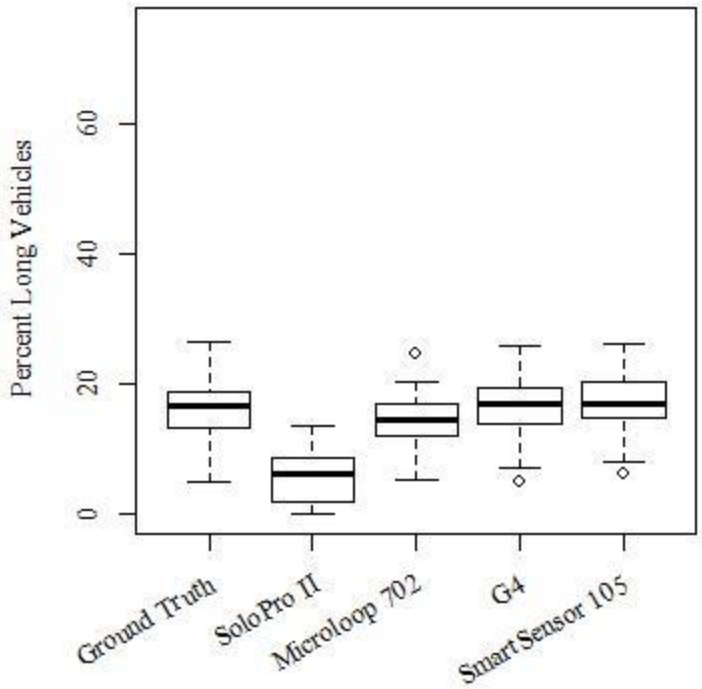

Figure 6.53: Box Plot of Fifteen-Minute Percent Long Vehicle Distributions

When classification error percentage (as defined by equation 6.1) was analyzed at five-minute and fifteen-minute aggregation intervals, the classification error decreased with further aggregation. This was again due to the loss of information which takes place with further aggregation. This loss of information can be understood by imagining a short vehicle in one minute being misclassified as a long vehicle and a long vehicle in the next minute being misclassified as a short vehicle. Assuming no other vehicles were detected in this two-minute period, aggregation at the one-minute interval would report a mean 100% classification error, while aggregation at the two-minute interval would report a mean 0% classification error. While this example is unrealistic, it serves to demonstrate how the mean G4 classification error percentages were 3.4%, 2.1%, and 1.6% at the one-minute, five-minute, and fifteen-minute aggregation intervals, respectively. Refer to

appendices D and E for further information on five-minute and fifteen-minute classification error.

6.3 Chapter Summary

This chapter has provided analysis of the time aggregate detection abilities of the four detectors under evaluation. The relative strengths and weaknesses of the different detectors were demonstrated in the results of this analysis. One-minute, five-minute, and fifteen-minute aggregation intervals were selected to represent the effect of various levels of aggregation on detector accuracy. Specific ITS applications require data at various intervals, and one detector may be well-suited for an application that uses fifteen-minute aggregate data while not providing appropriately accurate data for an application requiring one-minute aggregate data. The aggregate data analysis presented in this chapter focused on interval traffic volume, mean speed over the interval, and traffic composition over the interval (proportion short, medium, and long vehicles).

The analysis of interval traffic volume detection in this chapter indicated that the G4 had the strongest correlation with ground truth volumes, with correlation coefficients of 0.993, 0.997, and 0.998 for one, five, and fifteen minute intervals, respectively. The Solo Pro II and Microloop 702 had correlation coefficients nearly as strong as the G4, and had mean percent errors closer to zero than the G4. The SmartSensor 105 was found to underreport volume when higher traffic volumes were present. It was found that while mean percent error was relatively unchanged by longer aggregation intervals, mean absolute percent error decreased for every detector with longer aggregation intervals. Regression analysis found that the environmental conditions that significantly affected Solo Pro II volume detection were night lighting and the combined effect of dawn

lighting and rain. The Microloop 702 and G4 were found to be significantly affected by the combined effect of dusk lighting and rain, while the SmartSensor 105 was not found to be significantly affected by lighting or rain conditions.

The analysis of interval mean speed was conducted with the Microloop 702 data serving as a baseline due to the lack of ground truth speeds. The distributions of one, five, and 15-minute mean speeds indicated that the Solo Pro II was reporting interval mean speeds much higher than the other three systems, including the baseline Microloop 702. However, it was concluded that this could be corrected with further calibration. The more intriguing finding was that while the Microloop 702, Solo Pro II, and SmartSensor 105 mean speed distributions all had similar shapes, the G4's mean speed distribution had a more symmetrical shape which lacked the significant left tail that was present in the other detectors' distributions. This was interpreted as the G4 being relatively insensitive to reductions in speed. Interval mean speed analysis provided very similar results at the one, five, and 15-minute aggregation levels, with the primary difference being a reduction in the variance of reported values from each detector as aggregation increased. This was consistent with expectations for data aggregation. The interval speed detection analysis also considered the influence of environmental factors with mixed results.

Lastly, the interval classification analysis indicated strong length-based classification from the Microloop 702, G4, and SmartSensor 105, with mean classification error percentages below 5% for all three at one-minute intervals. The Solo Pro II struggled with classification, with the most frequent problem being the misclassification of long vehicles as short. The Solo Pro II's mean classification error was 12% at the one-minute aggregation interval. It was found that greater aggregation levels

decreased mean classification error percentages for all four detection systems, and also decreased the variance of these classification error percentages. Analysis involving the influence of environmental factors indicated that night lighting conditions exacerbated the Solo Pro II's classification problem. The G4's classification ability was found to be affected by the combination of dusk lighting and rain. This effect was hypothesized to be a result of heavy rain which took place during one of the dusk lighting intervals. The classification abilities of the other detectors appeared to be relatively uninfluenced by the documented environmental factors.

CHAPTER 7 DISAGGREGATE ANALYSIS AND RESULTS

While aggregate interval analysis provided information on detector performance over temporal intervals, representing what may be used in practical planning and ITS implementations, disaggregate per-vehicle analysis provides a powerful tool for the determination of factors which affect detector performance. The following analysis focused on disaggregate analysis of per-vehicle detection.

This disaggregate analysis was based on vehicle detections in the 1467 minute analysis data set defined in section 4.1. In this data set there were a total of 36,124 time-stamped ground truth vehicle presence detections with vehicle classification. The data set also included time-stamped detector reported vehicle detections with individual speeds and vehicle classifications from each of the four analyzed detection systems. Additionally, lighting and precipitation conditions and traffic volume were noted at the time of each detection, so that potential effects of these factors on the performance of the various detector technologies could be determined.

7.1 Presence Detection Analysis

The first detection parameter analyzed at the per-vehicle disaggregate level was presence detection. Each detection reported by one of the traffic detectors could be classified as either a correct detection or a false detection. If the detection could be correlated to a ground truth detection during the same second and in the same lane, it was classified as a correct detection. If there was no corresponding ground truth detection in the same lane at the same second, it was classified as a false detection. Additionally, if there was a ground truth detection without a corresponding reported detection from the given detector, this was classified as a missed detection for that detector. Table 7.1 gives the number of

correct, missed, and false detections for each analyzed detector during the entire data set, as well as percent correct, missed, and false detections.

Table 7.1 Presence Detection Summary Statistics

	Correct Detections	Missed Detections	False Detections	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	33785	2339	1204	90.5%	6.3%	3.2%
Microloop 702	35177	947	1816	92.7%	2.5%	4.8%
G4	33934	2190	431	92.8%	6.0%	1.2%
SmartSensor 105	31189	4935	1137	83.7%	13.2%	3.1%

The values in this table indicate that the Microloop 702 and G4 had the best overall presence detection rates, while the SmartSensor 105 had a comparatively high number of missed detections. Figure 7.1 provides a graphical depiction of the information presented in the table above. It is interesting to note that while the Microloop 702 and G4 had similar percent correct detections, the Microloop 702's errors were primarily false detections, while the G4's errors were primarily missed detections.

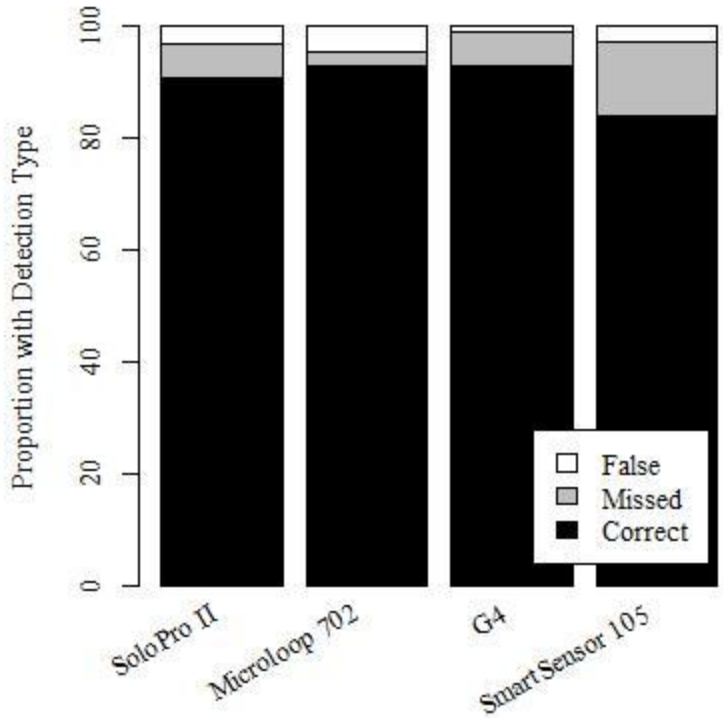

Figure 7.1: Presence Detection Stacked Bar Chart

The next step in the analysis was to separate the data into subsets representing the various factors being considered as potentially affecting detection performance, and to determine the percent correct, missed, and false detections for these subsets.

7.1.1 Volume Effect

The first division was by traffic volume at the time of the detection into low volume and high volume subsets. Low volume periods were defined as periods when the traffic stream had a level of service of A or B (i.e., one-minute periods during which the three-lane passenger car equivalency did not exceed 54). High volume periods were characterized by a level of service of C or worse (i.e., one-minute periods during which the three-lane passenger car equivalency exceeded 54). Table 7.2 gives the presence detection performance for low volume periods, while table 7.3 gives the presence detection performance for high volume periods.

Table 7.2 Low Volume Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	90.0%	6.3%	3.7%
Microloop 702	92.3%	2.2%	5.4%
G4	93.0%	5.8%	1.3%
SmartSensor 105	89.0%	7.5%	3.6%

Table 7.3 High Volume Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	92.0%	6.2%	1.8%
Microloop 702	93.9%	3.4%	2.7%
G4	92.4%	6.7%	1.0%
SmartSensor 105	67.3%	31.2%	1.5%

As would be expected, there was a tradeoff between missed detections and false detections at different volumes of traffic. At a higher traffic volume, there were generally more missed detections and fewer false detections. It is noteworthy, however, that the

percent correct detections remained fairly similar at different volumes. The one major exception is the SmartSensor 105 which appears to have performed much better at low volumes than at high volumes. This supports the finding in section 6.1.1 that the SmartSensor 105 tended to under-report volumes when the ground truth volume was high. Figure 7.2 depicts visually the effects of volume on presence detection for the various detectors analyzed. This figure again shows that the SmartSensor 105 performed much better under low volume conditions than high volume conditions.

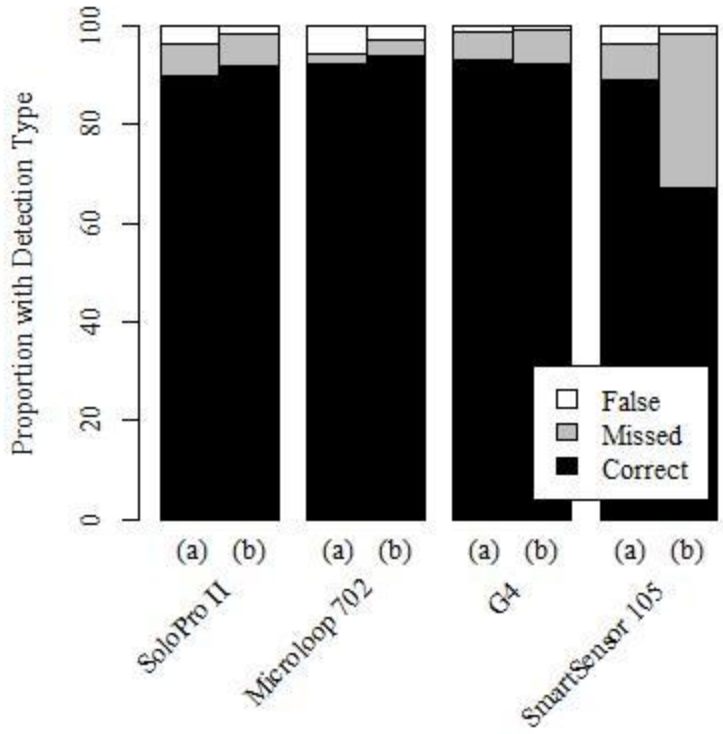

Figure 7.2: Presence Detection Volume Factor Stacked Bar Chart
*where (a) represents low volume periods and (b) represents high volume periods

7.1.2 Precipitation Effect

The next factor to be considered was precipitation. A division was made between clear and rainy subsets of the data. Table 7.4 gives the presence detection performance for clear weather periods, while table 7.5 gives the presence detection performance for rainy periods. Rainy periods were defined as any minute in the data set during which liquid

precipitation was noted. This absence or presence of rain was determined based on weather reports from the nearby Millard Airport in conjunction with manual observation of the ground truth video from the NTC/NDOR non-intrusive detector test bed.

Table 7.4 Clear Weather Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	90.7%	6.5%	2.9%
Microloop 702	93.0%	2.4%	4.5%
G4	93.4%	5.5%	1.1%
SmartSensor 105	82.6%	14.6%	2.7%

Table 7.5 Rainy Weather Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	89.5%	5.1%	5.5%
Microloop 702	90.6%	2.9%	6.5%
G4	88.8%	9.4%	1.7%
SmartSensor 105	90.7%	4.3%	5.0%

The correct detection rates of the Solo Pro II, Microloop 702, and G4 all decreased with rain by varying magnitudes. One contrast that emerged in these two tables was the improvement of the SmartSensor 105's percent correct detections by 8.1 percentage points between clear and rainy conditions. In the search for a logical explanation for this result, it was noted that all high volume periods (i.e., LOS C or D) were also clear periods. This unintentional correlation could have been reintroducing the strong negative effect of high volume on SmartSensor 105's presence detection as a pseudo-positive effect of rain. Therefore, it should not be concluded that the SmartSensor 105 performed better in rainy conditions based on these data. Figure 7.3 visually depicts the contrasts between the values in tables 7.4 and 7.5.

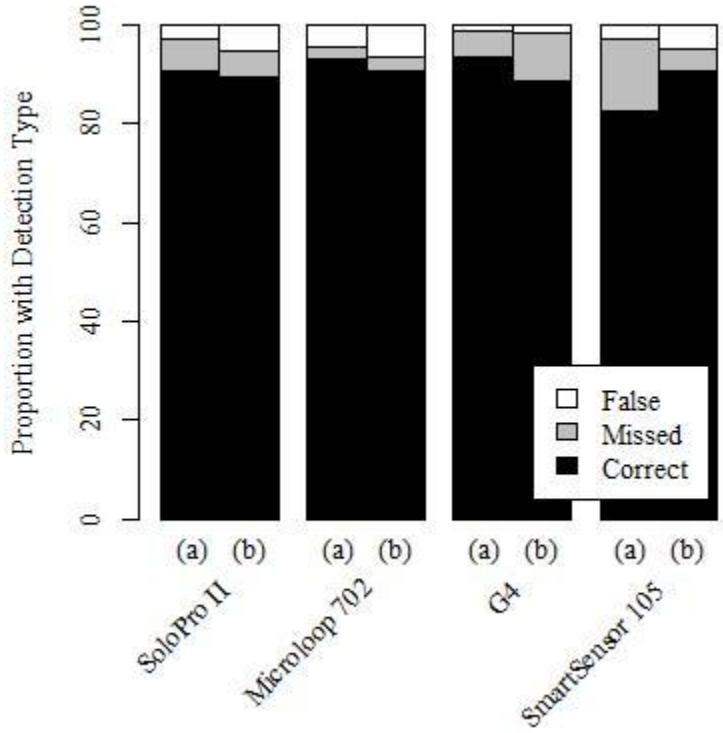

Figure 7.3: Presence Detection Rain Factor Stacked Bar Chart
*where (a) represents clear weather periods and (b) represents rainy weather periods

7.1.3 Lighting Effect

The final factor to be considered was lighting. For lighting, a division was made between day, night, dawn, and dusk subsets of the data. The definitions of these lighting conditions were related to time of day. For the purpose of this study, dawn was defined as the one hour period centered around sunrise. Dusk was defined as the one hour period centered around sunset. Review of video of the traffic stream confirmed that the lighting transition from day to night took place during this one hour period, as shown in figure 7.4. Day was defined as the period from the end of the dawn period to the beginning of the dusk period. Night was defined as the period from the end of the dusk period to the beginning of the dawn period.

Figure 7.4: Dusk Lighting Transition on 06/20/2011

*where (a) is sunset - 30 min, (b) is sunset -15 min, (c) is sunset, (d) is sunset +15 min, and (e) is sunset + 30 min

Table 7.6 gives the presence detection performance for day lighting periods, while table 7.7 gives the presence detection performance for night lighting periods. Table 7.8 gives the presence detection performance for dawn lighting periods, and table 7.9 gives the presence detection performance for dusk lighting periods.

Table 7.6 Day Lighting Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	90.6%	6.6%	2.8%
Microloop 702	92.9%	2.5%	4.5%
G4	93.1%	5.7%	1.1%
SmartSensor 105	82.4%	14.8%	2.9%

Table 7.7 Night Lighting Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	89.7%	3.8%	6.5%
Microloop 702	92.1%	1.5%	6.5%
G4	94.2%	4.9%	0.9%
SmartSensor 105	93.1%	2.8%	4.1%

Table 7.8 Dawn Lighting Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	87.1%	5.7%	7.1%
Microloop 702	92.4%	0.9%	6.7%
G4	95.2%	3.4%	1.3%
SmartSensor 105	90.9%	4.2%	4.9%

Table 7.9 Dusk Lighting Presence Detection Statistics

	% Correct Detections	% Missed Detections	% False Detections
Solo Pro II	91.7%	3.8%	4.5%
Microloop 702	90.1%	4.2%	5.7%
G4	83.8%	14.2%	2.0%
SmartSensor 105	91.6%	4.4%	4.0%

There are a few noteworthy values in these tables. First, the 14.8% missed detections for the SmartSensor 105 under day lighting conditions were 10.4% to 12.0% higher than the missed detections for this unit under the three other conditions. The most rational explanation for this is that the volume effect was, again, showing up unintentionally due to the fact that all high volume periods were during day lighting conditions. Another error rate that stood out was the 14.2% missed detections for the G4 under dusk lighting conditions. Further analysis of the data set indicated that this severe error rate may have been due to the effect of heavy rain during portions of the dusk subset. There were much higher missed detection rates during this heavy rain period than during the remainder of the dusk period. Another noteworthy trend was the increase in Solo Pro II false detections under night and dawn lighting. This could potentially be attributed to headlight spillover at night and long shadow spillover at dawn. Spillover is a phenomenon where a vehicle artifact, such as shadow or headlight reflection on pavement, is detected in a lane adjacent to the lane in which the vehicle is actually travelling. A potential instance of headlight spillover in lane two from the vehicle travelling in lane one can be seen in figure 7.5(a), while a potential instance of shadow spillover in lane two from the truck in lane one can be seen in figure 7.5(b). Next, figure 7.6 visually depicts the contrasts between the presence detection rates under various lighting conditions.

Figure 7.5: Potential Spillover Situations

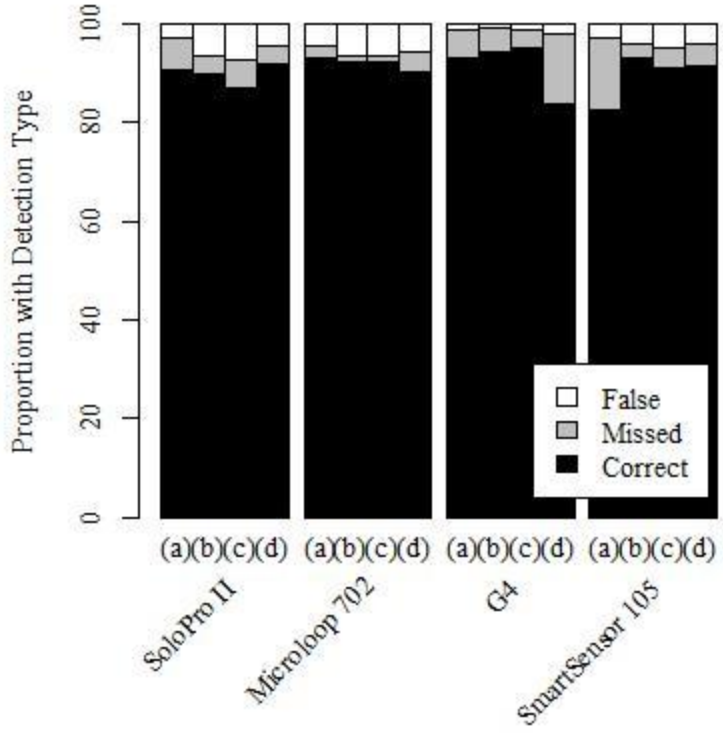

Figure 7.6: Presence Detection Lighting Factor Stacked Bar Chart

*where (a) represents day periods, (b) represents night periods, (c) represents dawn periods, and (d) represents dusk periods

While disaggregate presence detection may be considered the most basic metric of traffic detector accuracy, it should not be overemphasized in the assessment of traffic detectors. Most ITS applications for which a traffic detector would be required utilize data aggregate on some time interval. As presence detection is aggregate, it is represented

by volume over the set interval. This aggregation allows for a balancing effect of missed and false detections, which is not represented in the disaggregate analysis. For that reason, the metric of disaggregate presence detection was presented in conjunction with a number of other metrics.

7.2 Per-Vehicle Speed Analysis

As a ground truth speed was not available throughout the duration of the data collection period, the Microloop 702 was selected as a baseline against which the other detectors were compared. This system was chosen as the baseline because its magnetic induction technology and functional procedure for collecting speed data through a "speed trap" configuration most closely represented the legacy system of inductive loop detectors.

This speed trap configuration introduced a potential type of error that is not present in the other detectors. While other detectors use one detection zone to calculate speed, the speed trap correlates detections from two discrete sources to calculate speed. If only one of the sources registers a detection, no correlation occurs and the vehicle is assigned a speed of zero. Additionally, if the two sources falsely correlate detections of two different vehicles as one, extreme high or low speeds can be calculated as a result. These specific errors must be removed from the data set before analysis commences. This was done by defining an interval of reasonable speeds and removing detections having speeds outside this reasonable interval.

Based on the fact that "operating speeds have been found to be normally distributed," the speeds of vehicles at the detector test bed were assumed to be normally distributed (57). Under this assumption, the 40,395 vehicle sample should only have included approximately three vehicles (0.0063%) outside the range of 36 - 87 mph

($\mu \pm 4\sigma$). The range defined by four standard deviations from the mean was selected based on sample size and the number of expected values outside the ranges for that sample size. In reality, there were 185 values outside of this range (still less than 0.5% of the sample), rather than three. Many of these values were zero speeds. Other values near 160 mph resulted when vehicles in adjacent lanes occasionally confounded the speed trap calculation for speed. These 185 values were labeled "outliers," and were removed from the data set for the per-vehicle speed analysis. The remaining data set included speed data for 40,210 vehicles.

This analysis began with graphical representation of the distributions of detected per-vehicle speeds from each detector. The box plot in figure 7.7 indicates that the G4 reported the smallest distribution of speeds, while the Solo Pro II reported the largest distribution of speeds. The inter-quartile ranges speeds in this box plot also shows that the Solo Pro II frequently reported speeds much higher than the other three detectors.

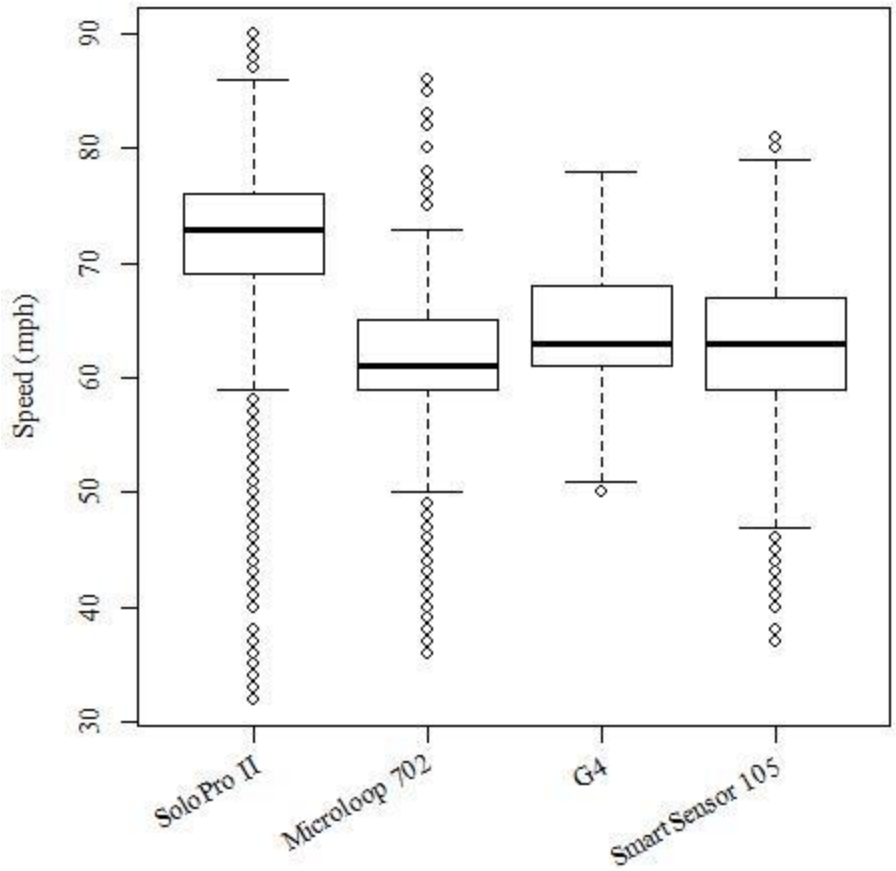

Figure 7.7: Box Plot of Reported Per-Vehicle Speeds

The histogram that follows (figure 7.8) depicts even more clearly the distributions of reported speeds from the various detectors. Additionally, the values for the first four central moments were given to further characterize each distribution. The mean speed values again showed that the Solo Pro II mean speed was 8.4 to 11.2 mph higher than the other detectors. It is also worth noting that the variance of the G4 speeds was lower than that of the other three detectors. This supports the hypothesis from chapter 6 that the G4 was less sensitive to differences in speed than were the other three detectors.

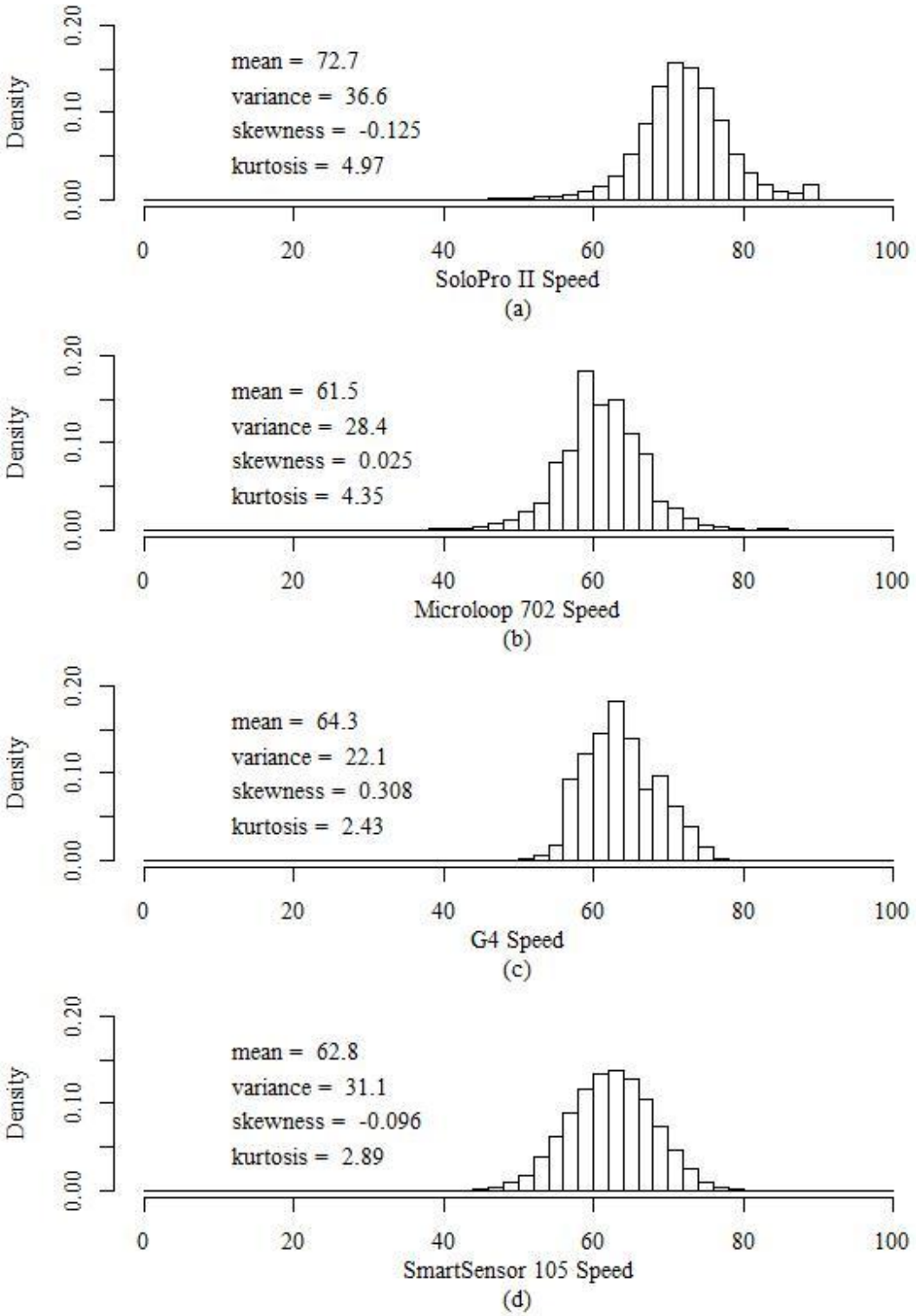

Figure 7.8: Histograms of Per-Vehicle Speed Distributions for the Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

The cumulative distribution plot in figure 7.9 provides one more graphical representation of the speed distributions for the four detectors under consideration. In this plot, the higher Solo Pro II speeds were again obvious. Closer examination revealed that, while the G4 detected higher speeds similarly to the Microloop 702 and SmartSensor 105, it did not detect the same lower speeds as the Microloop 702 and SmartSensor 105 (i.e., speeds below approximately 55 mph).

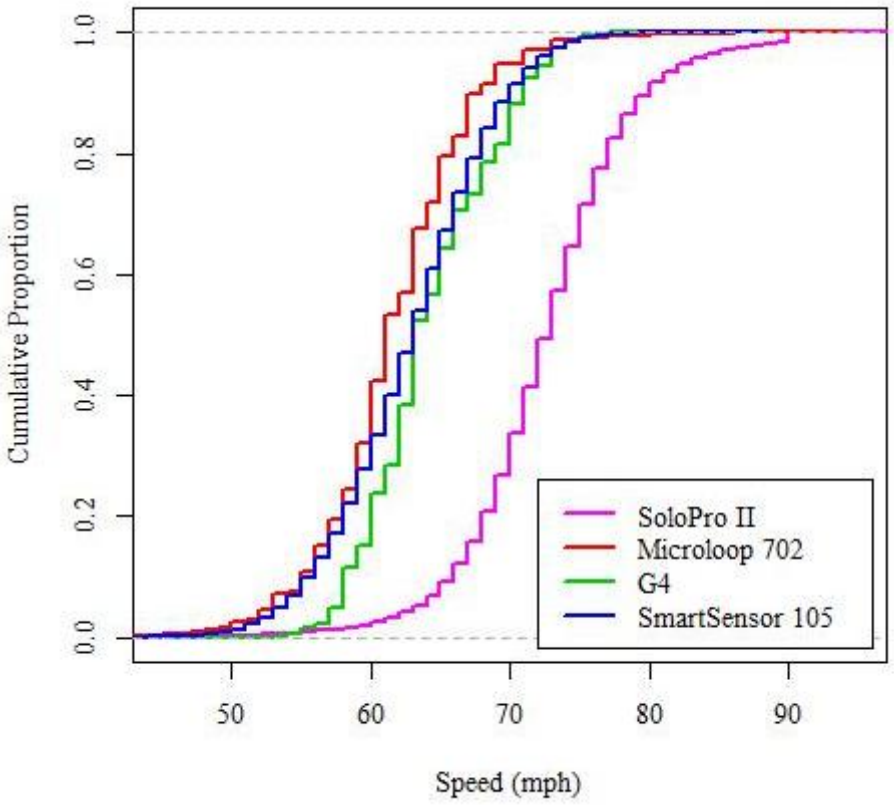

Figure 7.9: Cumulative Distribution Plot of Per-Vehicle Speed Distributions for All Detectors

The most obvious information available in the above figures is that the mean of the Solo Pro II reported speeds (72.7 mph) was much higher than the other three detectors, which all had similar mean speeds (61.5 mph - 64.3 mph). While the Solo Pro II software contained a speed calibration adjustment factor (a multiplicative factor which

can be applied to every vehicle speed), this factor was not adjusted since the initial installation of the detector, based on the fact that its inclusion would be purely empirical and not based on the theory behind how speed is calculated by this detector. It is noted here that configuration of the detectors and recalibration for this thesis was primarily focused on optimizing presence detection. Recalibration after a preliminary data collection interval did not address speed detection. As such, the mean speed bias alone should not be considered as a detriment for any of the detectors. Figure 7.10 shows how closely the distributions of per-vehicle speeds from each detector represented one-another when appropriate multiplicative factors were applied to each speed so that all detectors had the same mean speed as the baseline Microloop 702.

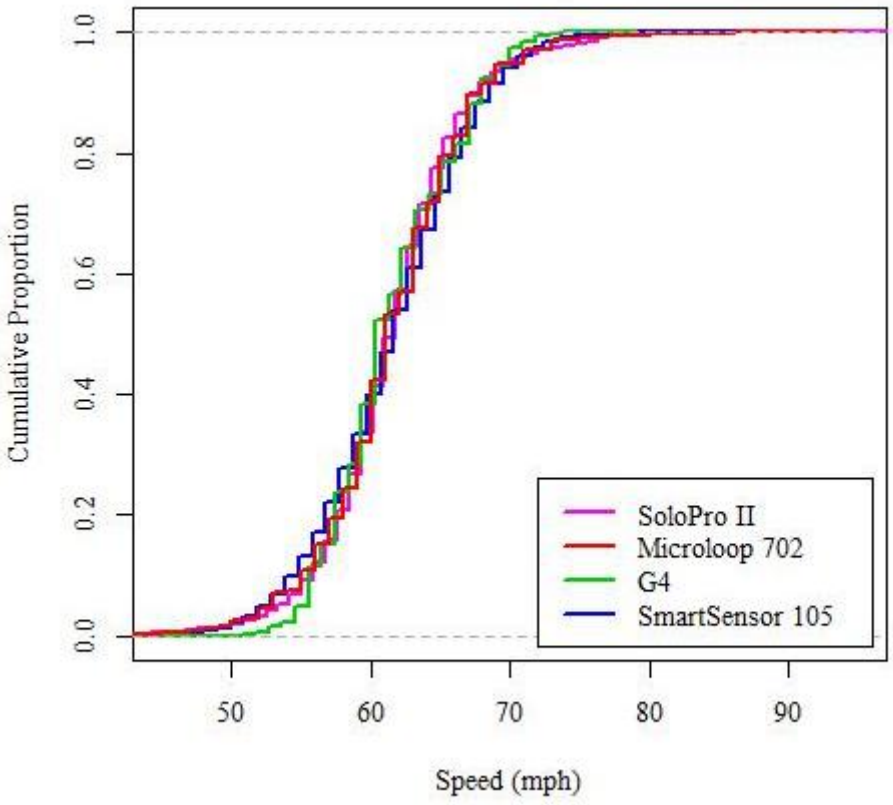

Figure 7.10: Cumulative Distribution Plot of Per-Vehicle Speed Distributions for All Detectors with Respective Multiplicative Factors Applied

After noting the speed distributions reported by each detector, the detected speeds from the Solo Pro II, G4, and SmartSensor 105 were compared to the speeds reported by the Microloop 702 baseline detector. The scatter plots in figure 7.11 and the accompanying correlation coefficients (*r*) indicated that the Solo Pro II speeds had the strongest linear relationship to the baseline speeds. Figure 7.11 also shows that the range of G4 speeds was narrower than the range of speeds from the other detectors, suggesting that it may be relatively insensitive to changes in speed when compared to the other detectors.

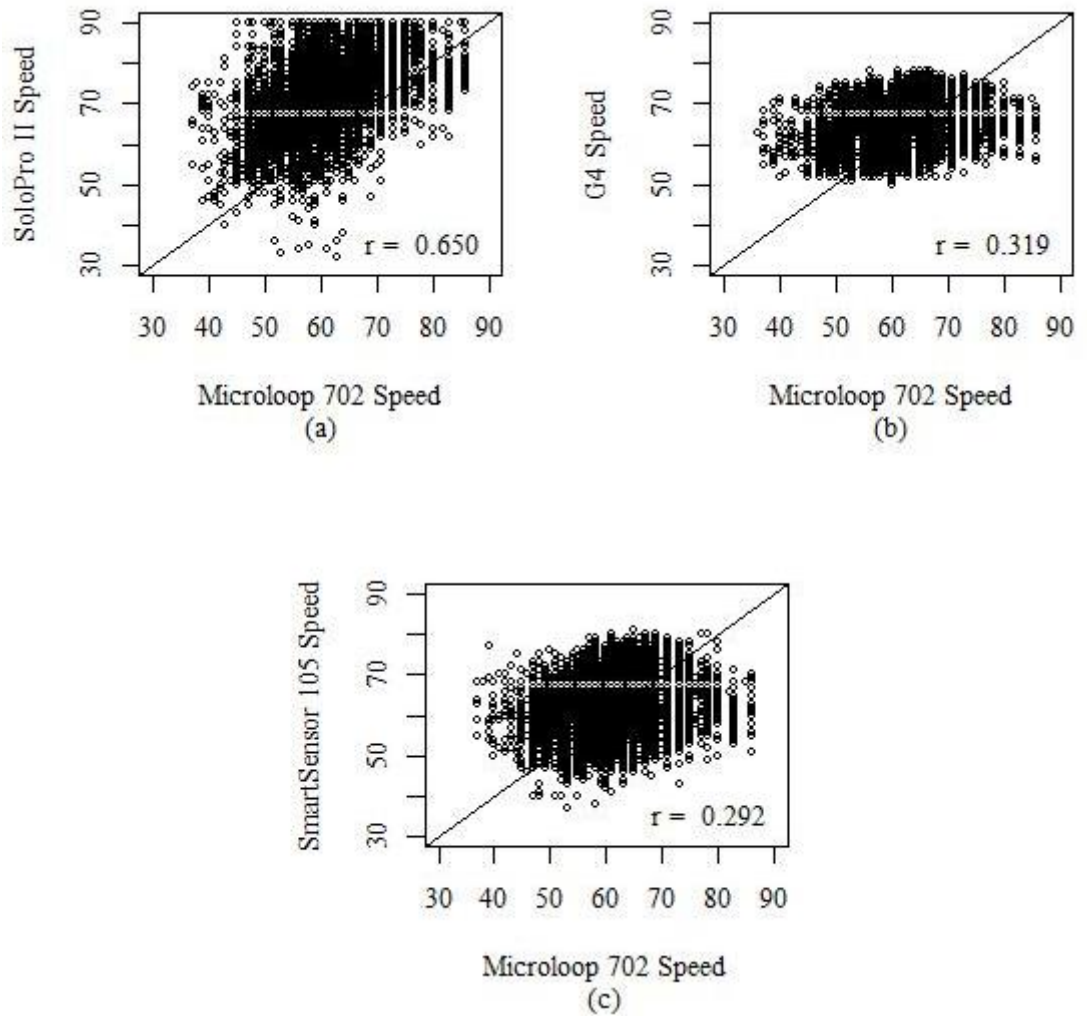

Figure 7.11: Per-Vehicle Speed Scatter Plots Against Baseline for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

This was followed by the calculation of the percent deviations and absolute percent deviations from the baseline for each detection. The distributions of the percent deviation values for each detector are displayed graphically in figures 7.12-7.14. Appropriate per-vehicle speed deviation statistics such as MPD, MAPD, and variance of percent deviation are given in table 7.10. There are a few observations worth noting in these figures and the table. Figure 7.12 shows that the inter-quartile range of the Solo Pro II was narrower than those of the G4 and SmartSensor 105, indicating that it had a

relatively consistent deviation from the baseline. The relatively high kurtosis of the Solo Pro II speed percent deviation in figure 7.13 provides further evidence of this fact, as does the steep central portion of its cumulative distribution curve (figure 7.14) and the relatively small percent deviation variance of the Solo Pro II (table 7.10). Also worth noting are the similarities between the G4 and the SmartSensor 105. It was hypothesized that the similar distributions of these two detectors' speed percent deviations, shown in figures 7.12 and 7.13, indicated that the common technology of microwave radar employed by these detectors led to a specific bias in speed detection. Additionally, the differences between these two detectors, indicated by the values in table 7.10, indicate that other attributes of reported speeds were unique to each detector model with the same technology.

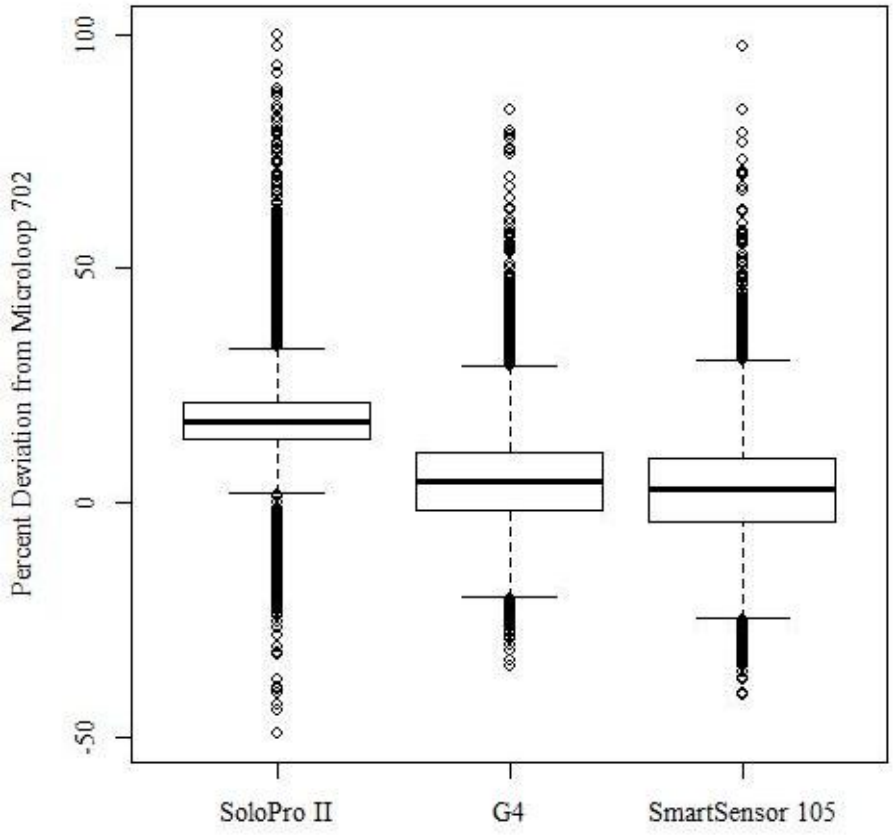

Figure 7.12: Per-Vehicle Speed Percent Deviation Box Plot

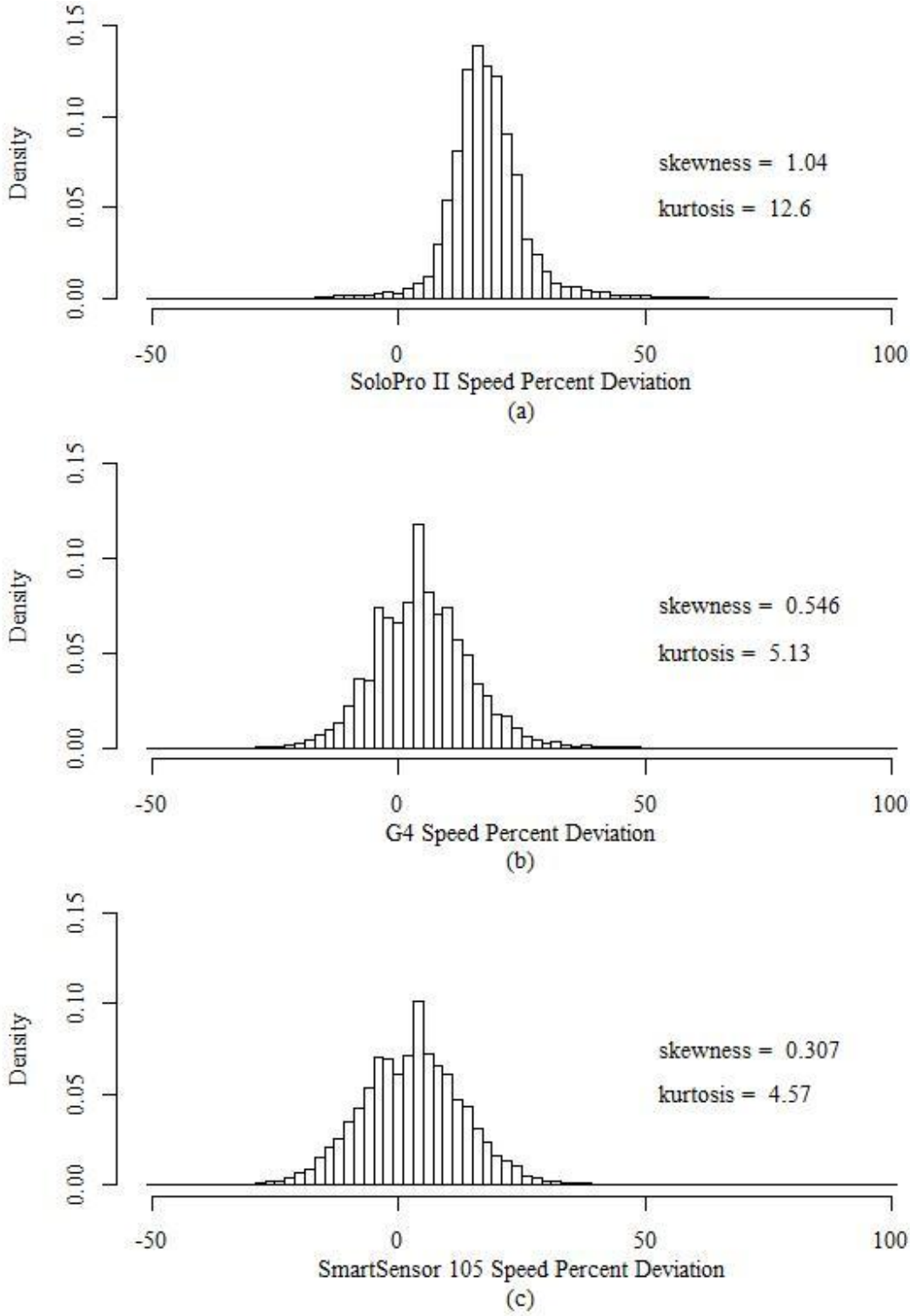

Figure 7.13: Histograms of Per-Vehicle Speed Percent Deviation Distributions for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

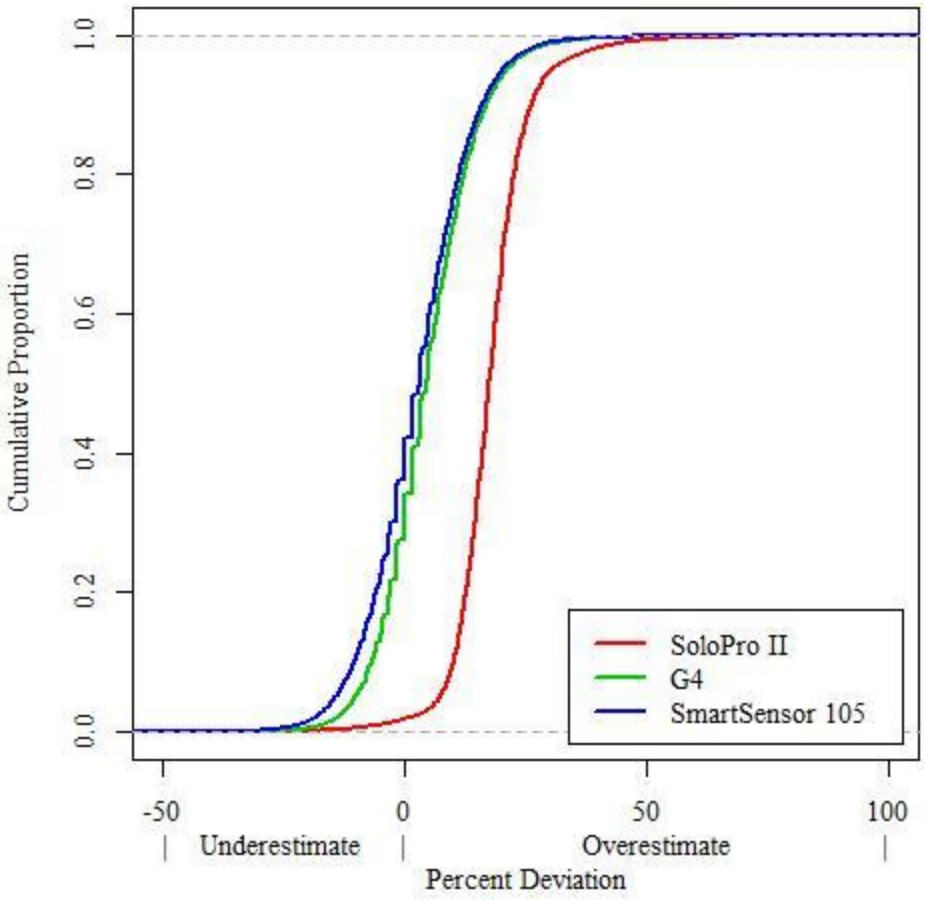

Figure 7.14: Per-Vehicle Speed Percent Deviation Cumulative Distribution Plot

Table 7.10: Detector Per-Vehicle Speed Deviation Statistics

	MPD	MAPD	Percent Deviation Variance
Solo Pro II	17.9%	18.2%	0.00694
G4	4.85%	8.33%	0.00959
SmartSensor 105	2.88%	8.59%	0.0115

Theil's inequality coefficient (*U*) was calculated for per-vehicle speeds for each detector, and is presented along with its proportion components in table 7.11. This goodness-of-fit measure was explained in section 5.4. *U* can take values from zero to one, with higher values indicating greater inequality between the detector-observed speeds and baseline speeds. The proportion components provide further understanding of the

character of differences of each detector's reported speed from the baseline. The bias proportion (U_m) is a measure of proportion of the deviation due to consistent bias in the detection of speed. The variance proportion (U_s) is a measure of the proportion of the deviation due to inequality between the baseline and detector variances in per-vehicle speed distributions. The covariance proportion (U_c) is a measure of the proportion of the deviation that is unsystematic, or random. As mutually exclusive proportions, U_m, U_s, and U_c sum to one.

Table 7.11: Per-Vehicle Speed Theil's Inequality Coefficients

	U	U_m	U_s	U_c
Solo Pro II	0.088	0.848	0.002	0.150
G4	0.050	0.174	0.006	0.820
SmartSensor 105	0.053	0.049	0.003	0.949

The values for *U* in table 7.11 indicate that the Solo Pro II had the greatest per-vehicle speed inequality with respect to the baseline speeds. This was to be expected based on the previous data presented on per-vehicle speed. However, the value of U_m indicated that 84.8% of the Solo Pro II's inequality with respect to baseline speeds was attributable to bias (a consistent error that can be addressed with further calibration). The remainder of table 7.11 indicates that the G4 could also benefit from additional calibration with a bias proportion (U_m) of 17.4%, and that the SmartSensor 105 had the highest proportion of unsystematic inequality (U_c = 94.9%).

Next, the data set was broken down by environmental conditions, and percent deviation distributions were determined for data subsets with similar conditions for factors such as lighting (day, night, dawn, dusk), precipitation (clear, rain), and traffic volume (low volume, high volume).

Effects of lighting, precipitation, and volume on the Solo Pro II per-vehicle speed percent deviation are shown in the distributions in figures 7.15-7.17. Figure 7.15 indicates that the Solo Pro II was prone to greater speed errors under night lighting in comparison to the other lighting conditions, as evidenced by relatively fat tails at both ends of the cumulative distribution line for night lighting. Figure 7.16 indicates that under rainy conditions, the severity of Solo Pro II speed overestimation may be slightly reduced relative to clear conditions. It was hypothesized that both of these environmental impacts could be attributed to headlight reflection off of the pavement in night or wet conditions. However, testing this hypothesis was beyond the scope of this thesis. Traffic volume did not appear to greatly impact Solo Pro II reported speeds (figure 7.17).

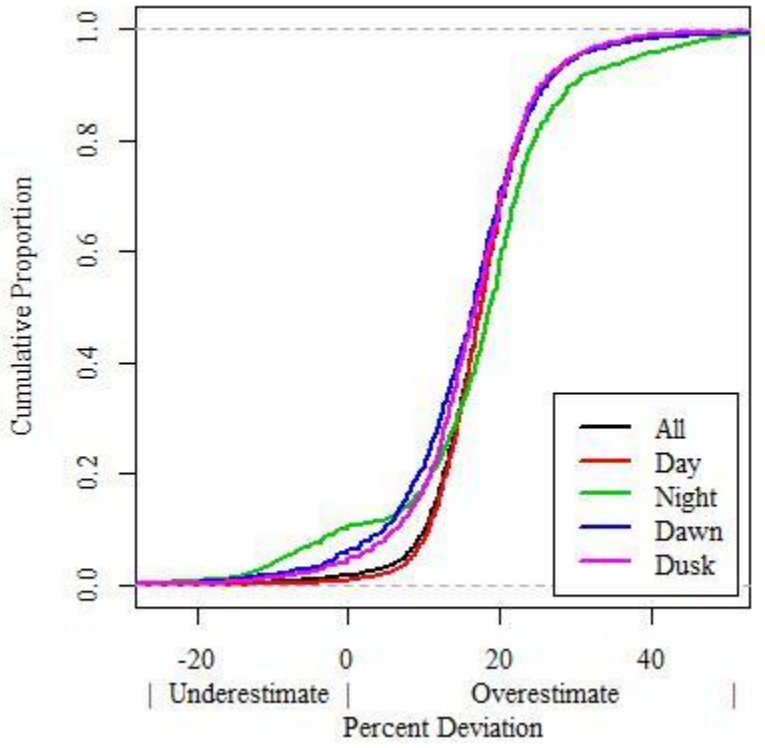

Figure 7.15: Solo Pro II Per-Vehicle Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

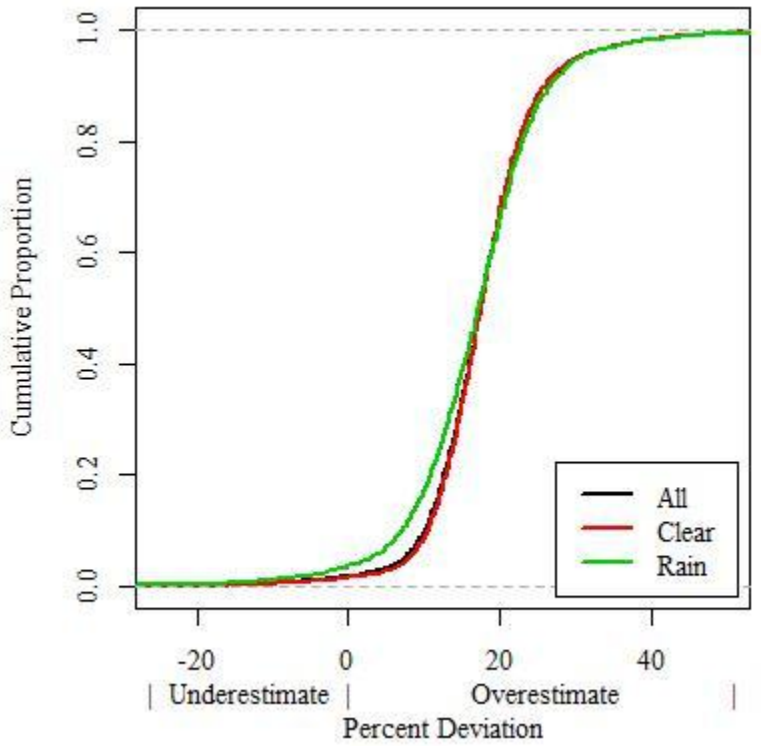

Figure 7.16: Solo Pro II Per-Vehicle Speed Percent Deviation Rain Factor Cumulative Distribution Plot

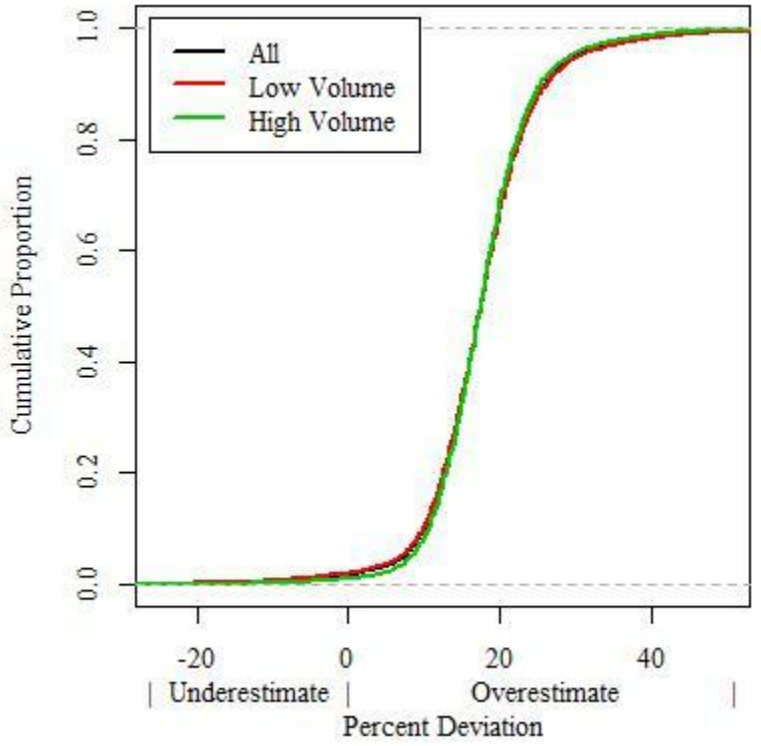

Figure 7.17: Solo Pro II Per-Vehicle Speed Percent Deviation Volume Factor Cumulative Distribution Plot

Figures 7.18-7.20 represent the effects of lighting, rain, and traffic volume on G4 speed detection. The cumulative distribution lines in figure 7.18 indicate that while the G4 generally overestimated speed, the severity of this overestimation was diminished in dawn lighting conditions. As the microwave radar technology employed by the G4 should not have been affected by light, an alternative explanation was required. The most practical explanation implied that the G4 was insensitive to changes in speed in comparison to the other detector systems evaluated. The three other systems each had similar mean speeds for dusk and night conditions and a mean speed approximately 2 mph higher during dawn and day conditions, indicating more aggressive driver behavior at those times. In contrast, the G4 had similar mean speeds for dusk, night, and dawn conditions, and a mean speed approximately 2 mph higher during day lighting conditions. Figure 7.19 indicates that the G4 was relatively unaffected by rain conditions. Lastly, figure 7.20 indicates that the G4 overestimated speed by 7.5% during high volume conditions, as compared to 4.0% during low volume conditions. Based on the fundamental speed-density relationship, it was anticipated that actual speeds would be lower at high densities (and thus also high volume). Therefore, the greater overestimation of speed under high volume conditions again indicates that the G4 was relatively insensitive to changes in speed.

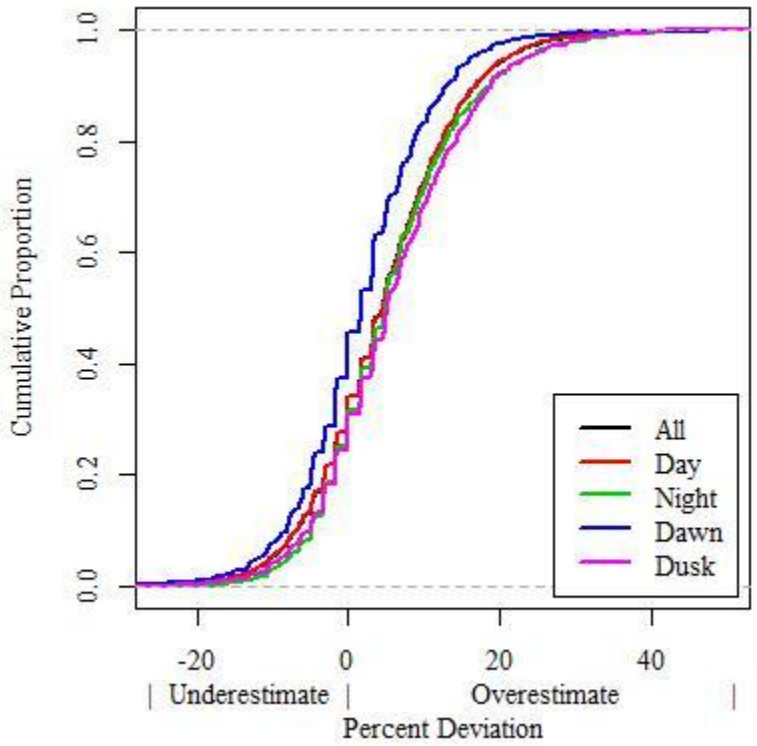

Figure 7.18: G4 Per-Vehicle Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

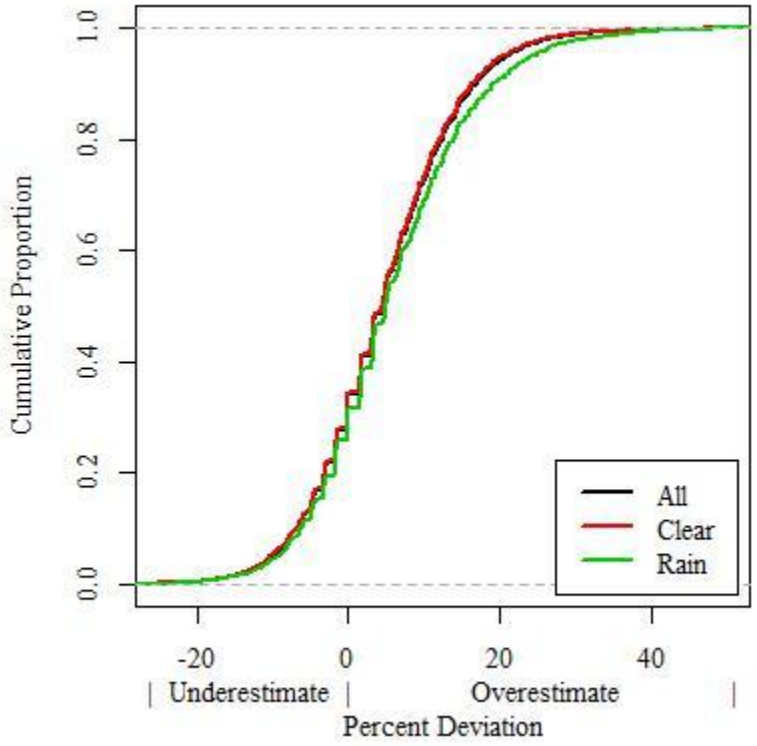

Figure 7.19: G4 Per-Vehicle Speed Percent Deviation Rain Factor Cumulative Distribution Plot

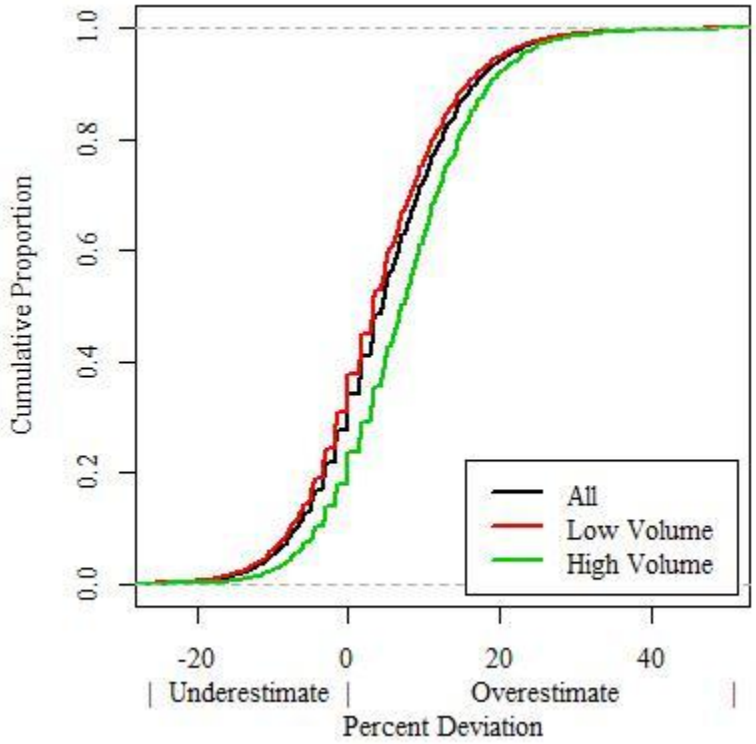

Figure 7.20: G4 Per-Vehicle Speed Percent Deviation Volume Factor Cumulative Distribution Plot

The observed speeds from the SmartSensor 105 for different lighting, rain, and traffic volumes are shown in figures 7.21-7.23. The similar cumulative distribution lines in figure 7.21 indicate that the SmartSensor 105 speed detection was unaffected by various lighting conditions. Similarly, figure 7.22 indicates that the SmartSensor 105 speed detection was relatively unaffected by rain. Lastly, figure 7.23 indicates that traffic volume did have some impact on the reported speeds of the SmartSensor 105. It appears that higher traffic volume increased the percent deviation of the SmartSensor 105 speed relative to the baseline speed by an average of 2.5 percentage points (4.9% mean deviation in high volume compared to 2.4% mean deviation in low volume).

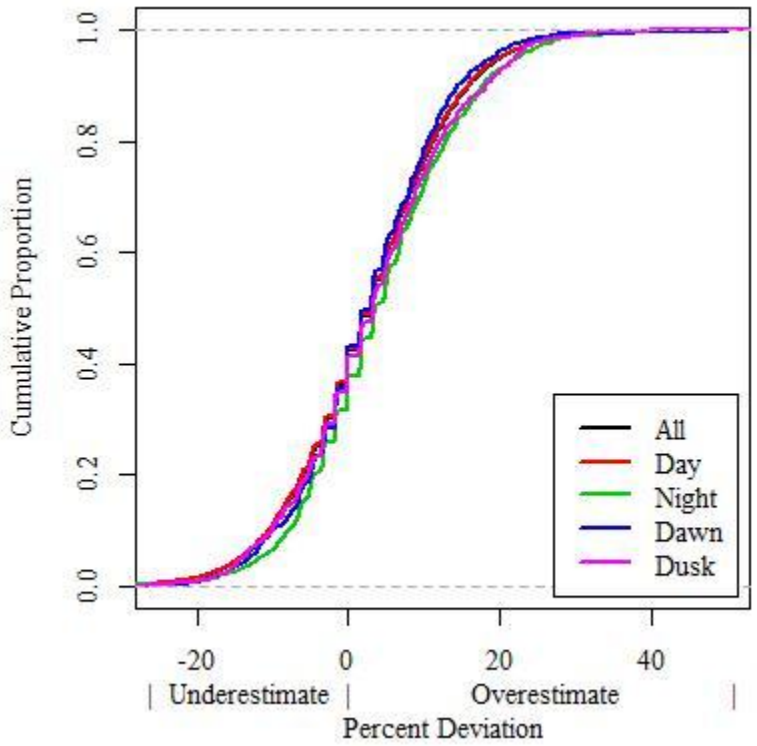

Figure 7.21: SmartSensor 105 Per-Vehicle Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

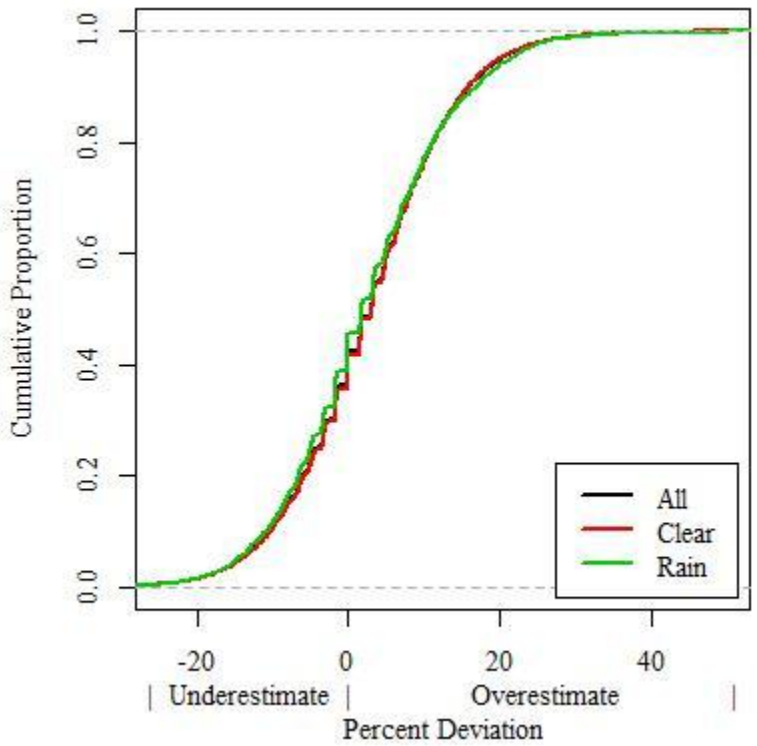

Figure 7.22: SmartSensor 105 Per-Vehicle Speed Percent Deviation Rain Factor Cumulative Distribution Plot

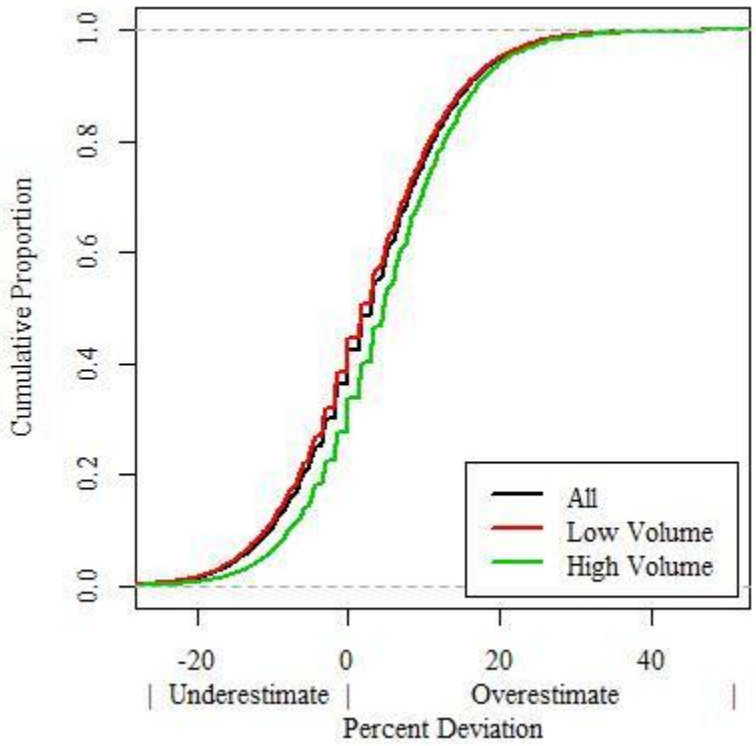

Figure 7.23: SmartSensor 105 Per-Vehicle Speed Percent Deviation Volume Factor Cumulative Distribution Plot

The effects of environment on speed detection were studied using ANOVA. An unbalanced four-by-two factorial ANOVA, based on the model presented in section 5.5, was used due to the unequal numbers of vehicles observed in each category, defined by the four lighting levels and two precipitation levels. This analysis was performed on each detector's per-vehicle speed percent deviation, with factors for lighting (levels=Day, Night, Dawn, and Dusk) and precipitation (levels = None and Rain). In order to minimize the effects of serial correlation, thinning was performed in a manner similar to that outlined in Appendix B for the one-minute volume ANOVA. The models for per-vehicle speed ANOVA dictated that a thinning factor of 10 would eliminate autocorrelation for all detectors. Statistical significance was reported a level of $\alpha = 0.05$. It is important to note that statistical significance reported here does not imply practical significance. This is to say that, due to the large sample size, a factor could be found to have a statistically

significant effect on the speed percent deviation, but the magnitude of that effect could be so small as to be meaningless from an engineering perspective.

The output of the Solo Pro II speed ANOVA, found in table 7.12, indicates that the intercept, as well as the effects of lighting, rain, and an interaction effect between lighting and rain, were statistically significant. The results of the G4 ANOVA, found in table 7.13, indicate the intercept was significant, as were the effects of lighting, rain, and an interaction effect between lighting and rain. Lastly, the results of the SmartSensor 105 ANOVA, found in table 7.14, indicate that intercept was statistically significant, while the effects of lighting and rain were not found to be statistically significant. As the interaction effect between lighting and rain was found not to be statistically significant for the SmartSensor 105, it was eliminated from the underlying model to provide greater power to the test of significance for the independent effects of lighting and rain.

Table 7.12: Solo Pro II Per-Vehicle Speed Percent Deviation ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	20.496	1	2913.207	0.000	*
Lighting	0.169	3	7.987	0.000	*
Rain	0.066	1	9.321	0.002	*
Lighting:Rain	0.141	3	6.691	0.000	*
Residuals	23.527	3344			

Table 7.13: G4 Per-Vehicle Speed Percent Deviation ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	1.944	1	204.036	0.000	*
Lighting	0.320	3	11.193	0.000	*
Rain	0.057	1	5.974	0.015	*
Lighting:Rain	0.167	3	5.855	0.001	*
Residuals	32.051	3364			

Table 7.14: SmartSensor 105 Per-Vehicle Speed Percent Deviation ANOVA

	Sum Sq	Df	F value	Pr(>F)	Sig.
(Intercept)	0.725	1	60.694	0.000	*
Lighting	0.056	3	1.562	0.197	
Rain	0.014	1	1.212	0.271	
Residuals	37.119	3106			

Next, multiple regression models for the per-vehicle speed percent deviation for each detector were developed to test whether the relationships found in the graphical representation of the data were statistically significant. This regression was based on the equation given in section 5.6, with the dependent variable (p_i) being the speed percent deviation for vehicle i, and the first dependent variable (α) being the theoretical mean speed percent deviation for the specified detector given daylight, non-rainy conditions. As with other analyses in this chapter, the effect of serial correlation was minimized through data thinning performed in a manner similar to that outlined in Appendix B for one-minute volume ANOVA. The models for per-vehicle speed regression dictated that a thinning factor of 10 would eliminate autocorrelation for all detectors. Statistical significance of model factors was reported at $\alpha = 0.05$.

Table 7.15 lists the Solo Pro II's one-minute mean speed percent deviation model coefficients. The statistically significant factors in this model were the intercept, rain, the combined effect of dawn lighting and rain, and the combined effect of dusk lighting and rain. The adjusted R-squared for this model was 0.0101 signifying a low correlation between the predicted and observed values for speed percent deviation.

Table 7.15: Solo Pro II Per-Vehicle Speed Percent Deviation Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	18.27	0.164	111.518	0.000	*
Night (γ11)	-0.19	0.728	-0.257	0.797	
Dawn (γ12)	-0.25	1.186	-0.207	0.836	
Dusk (γ13)	-0.20	0.876	-0.225	0.822	
Rain (γ21)	1.25	0.545	2.291	0.022	*
Night:Rain (γ31)	-2.87	1.495	-1.918	0.055	
Dawn:Rain (γ32)	-6.10	1.625	-3.755	0.000	*
Dusk:Rain (γ33)	-3.78	1.474	-2.567	0.010	*

The coefficients of the G4 per-vehicle speed percent deviation model are shown in table 7.16. The statistically significant factors in this model were the intercept, rain, and the combined effect of dawn lighting and rain. The adjusted R-squared for this model was 0.0150, signifying a low correlation between the predicted and observed values for speed percent deviation.

Table 7.16: G4 Per-Vehicle Speed Percent Deviation Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	4.72	0.190	24.889	0.000	*
Night (γ11)	1.15	0.841	1.369	0.171	
Dawn (γ12)	-1.21	1.354	-0.892	0.373	
Dusk (γ13)	-0.62	1.019	-0.613	0.540	
Rain (γ21)	1.81	0.628	2.888	0.004	*
Night:Rain (γ31)	3.08	1.797	1.712	0.087	
Dawn:Rain (γ32)	-5.89	1.865	-3.155	0.002	*
Dusk:Rain (γ33)	3.06	1.875	1.633	0.102	

The coefficients of the SmartSensor 105 per-vehicle speed percent deviation model are shown in table 7.17. The only statistically significant factor in this model was the intercept. The adjusted R-squared for this model was 0.0010, signifying a very low correlation between the predicted and observed values for speed percent deviation.

Table 7.17: SmartSensor 105 Per-Vehicle Speed Percent Deviation Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)	Sig.
(Intercept) (α)	2.93	0.224	13.074	0.000	*
Night (γ11)	0.83	0.944	0.881	0.379	
Dawn (γ12)	0.92	1.504	0.612	0.540	
Dusk (γ13)	0.61	1.144	0.531	0.596	
Rain (γ21)	-0.88	0.706	-1.249	0.212	
Night:Rain (γ31)	3.63	1.970	1.842	0.066	
Dawn:Rain (γ32)	-1.81	2.079	-0.870	0.384	
Dusk:Rain (γ33)	0.21	1.935	0.111	0.912	

While the low adjusted R-squared values for these models suggest a weak fit, that was to be expected in this application. If it were possible to accurately predict the speed percent error of a specific detector for any given vehicle based on one of the models listed above, it would be possible to eliminate these errors. As this is not the case, these models were presented in spite of their low adjusted R-squared values to demonstrate the average effect of potential environmental factors (see "Estimate" column in the above tables), and to surmise which of these effects were consistent enough to be deemed statistically significant.

7.3 Per-Vehicle Classification Analysis

The final detection parameter to be analyzed was vehicle classification. This analysis assessed the ability of each detector to correctly identify in which of three length-based bins a vehicle belonged. The three length bins were: under 25 feet, 25 to 40 feet, and over 40 feet in length, and were intended to represent passenger vehicles, single unit heavy vehicles, and multiple unit heavy vehicles, respectively. These length bin divisions were chosen based on the stated practice of NDOR officials responsible for the collection of planning data. Throughout the remainder of this section, these three classes will be referred to as *short*, *medium*, and *long* vehicles. The proportions of vehicles classified as

short, medium, and long by ground truth observation and each detector are depicted in figure 7.24. These classification proportions are also given in table 7.18. This figure and table indicate that the Solo Pro II had a tendency to classify more vehicles as short and medium, and fewer as long, than the actual ground truth. The other detectors appeared to provide classification proportions similar to the ground truth.

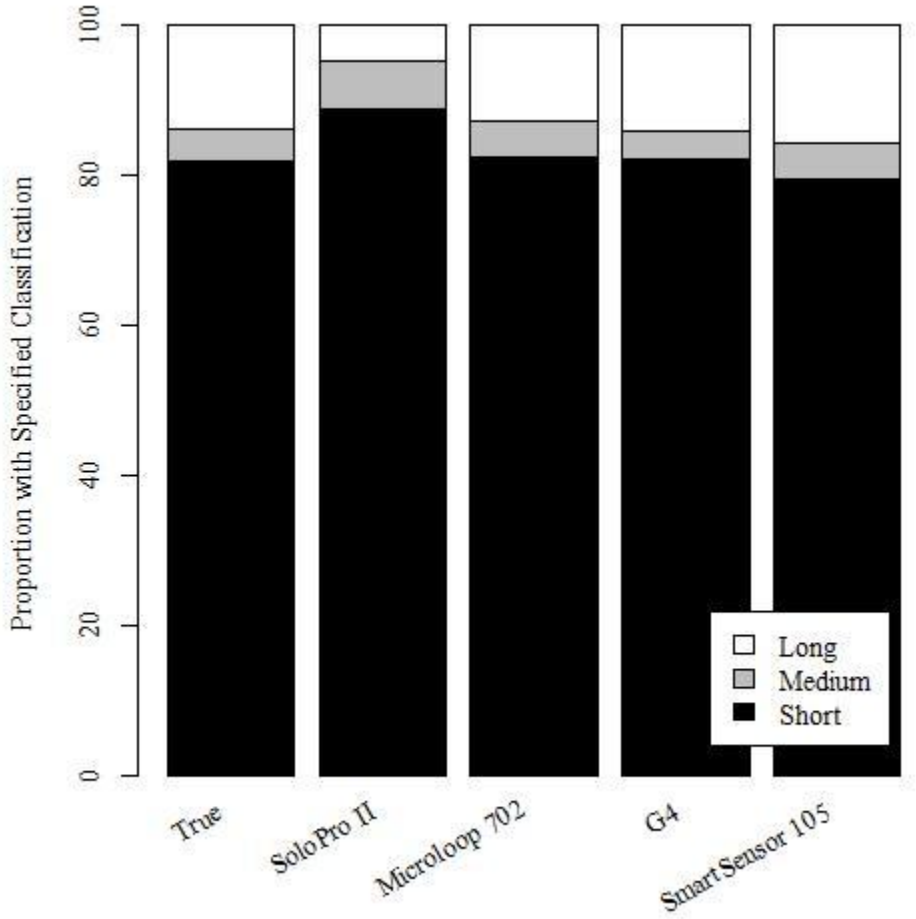

Figure 7.24: Per-Vehicle Classification Proportion Bar Chart

Table 7.18: Per-Vehicle Classification Proportions

	Ground Truth	Solo Pro II	Microloop 702	G4	Smartsensor 105
Short	81.7%	88.8%	82.3%	82.0%	79.4%
Medium	4.4%	6.4%	4.8%	3.8%	5.0%
Long	13.9%	4.8%	13.0%	14.2%	15.7%

In the analysis of a classification problem such as this one, confusion matrices provide a useful tool. A confusion matrix is an *n*-by-*n* matrix where *n* is the number of classes. For this vehicle classification problem, the confusion matrix was 3-by-3, with the rows representing ground truth classifications and the columns representing detector-reported classifications. The values in each cell represented the number of vehicles that had the specific combination of ground truth and detector-reported classification, based on the row and column, respectively. As can be seen in the following tables, the diagonal of the matrix represents correctly classified vehicles, while the non-diagonal cells represents misclassified vehicles. Also, row sums gave the total number of vehicles in the given class, while column sums gave the number of detector-reported vehicles in the given class.

The confusion matrix for the Solo Pro II classification is given in table 7.19. The sum of the diagonal cells indicates that 85.4% of the vehicles were correctly classified. Examination of the cells off the diagonal indicates that the most common classification error made by the Solo Pro II was to misclassify long vehicles as short, which it did with 2410 vehicles (7% of the total traffic stream). Other frequent errors included misclassifying long vehicles as medium vehicles (3.2% of the total traffic stream) and medium vehicles as short vehicles (3.1% of the total traffic stream).

Table 7.19: Solo Pro II Classification Confusion Matrix

		Solo Pro II Class							
		Short		Medium		Long		Row Total	
Ground Truth Class	Short	27274	(79.4%)	380	(1.1%)	47	(0.1%)	27701	(80.6%)
	Medium	1078	(3.1%)	468	(1.4%)	38	(0.1%)	1584	(4.6%)
	Long	2410	(7%)	1093	(3.2%)	1582	(4.6%)	5085	(14.8%)
	Column Total	30762	(89.5%)	1941	(5.6%)	1667	(4.9%)		

The confusion matrix for the Microloop 702 classification is given in table 7.20. The sum of the diagonal cells indicates that 94.9% of the vehicles were correctly classified. Examination of the cells off the diagonal indicates that all potential misclassifications had similar occurrence rates, ranging from 0.5% to 1.1% of the total traffic stream.

Table 7.20: Microloop 702 Classification Confusion Matrix

		Microloop 702 Class							
		Short		Medium		Long		Row Total	
Ground Truth Class	Short	28593	(80%)	365	(1%)	255	(0.7%)	29213	(81.8%)
	Medium	404	(1.1%)	1000	(2.8%)	180	(0.5%)	1584	(4.4%)
	Long	364	(1%)	246	(0.7%)	4312	(12.1%)	4922	(13.8%)
	Column Total	29361	(82.2%)	1611	(4.5%)	4747	(13.3%)		

The confusion matrix for the G4 classification is given in table 7.21. The sum of the diagonal cells indicates that 96.2% of the vehicles were correctly classified. Examination of the cells off the diagonal indicates that the most common classification error made by the G4 was to misclassify medium vehicles as short, which it did to 556 vehicles (1.6% of the total traffic stream). Other types of potential misclassifications all had infrequent occurrence rates, ranging from 0.3% to 0.6% of the total traffic stream.

Table 7.21: G4 Classification Confusion Matrix

		G4 Class							
		Short		Medium		Long		Row Total	
Ground Truth Class	Short	27617	(80%)	203	(0.6%)	97	(0.3%)	27917	(80.8%)
	Medium	556	(1.6%)	908	(2.6%)	113	(0.3%)	1577	(4.6%)
	Long	161	(0.5%)	185	(0.5%)	4698	(13.6%)	5044	(14.6%)
	Column Total	28334	(82%)	1296	(3.8%)	4908	(14.2%)		

The confusion matrix for the SmartSensor 105 classification is given in table 7.22. The sum of the diagonal cells indicates that 95.4% of the vehicles were correctly classified. Examination of the cells off the diagonal indicates that the most common classification error made by the SmartSensor 105 was to misclassify short vehicles as

medium, which it did to 575 vehicles (1.8% of the total traffic stream). Other types of potential misclassifications all had infrequent occurrence rates, ranging from 0.2% to 1.0% of the total traffic stream.

Table 7.22: SmartSensor 105 Classification Confusion Matrix

		SmartSensor 105 Class							
		Short		Medium		Long		Row Total	
Ground Truth Class	Short	24850	(78%)	575	(1.8%)	109	(0.3%)	25534	(80.2%)
	Medium	257	(0.8%)	903	(2.8%)	307	(1%)	1467	(4.6%)
	Long	147	(0.5%)	63	(0.2%)	4644	(14.6%)	4854	(15.2%)
Column Total		25254	(79.3%)	1541	(4.8%)	5060	(15.9%)		

The next step in the analysis was to break the data into subsets representing the various factors that may affect detector classification performance, and to determine the percent correctly classified at each level of a given factor. The first factor to be considered was lighting, and the four levels were day, night, dawn, and dusk, as defined in section 7.1. Figure 7.25 depicts the classification proportions for the ground truth and various detectors under each of the four lighting conditions. Additionally, confusion matrices such as those already presented were analyzed for the various lighting levels, with the percent correctly classified by each detector under each lighting level presented in table 7.23. The Solo Pro II had difficulty classifying long vehicles appropriately under all lighting conditions, as evidenced by figure 7.25, but this problem was most severe at night. This observation is supported by table 7.23, which shows that the percent of vehicles correctly classified by the Solo Pro II dropped 6% during night lighting compared to other lighting conditions. The other detectors under evaluation appeared to function consistently across lighting conditions.

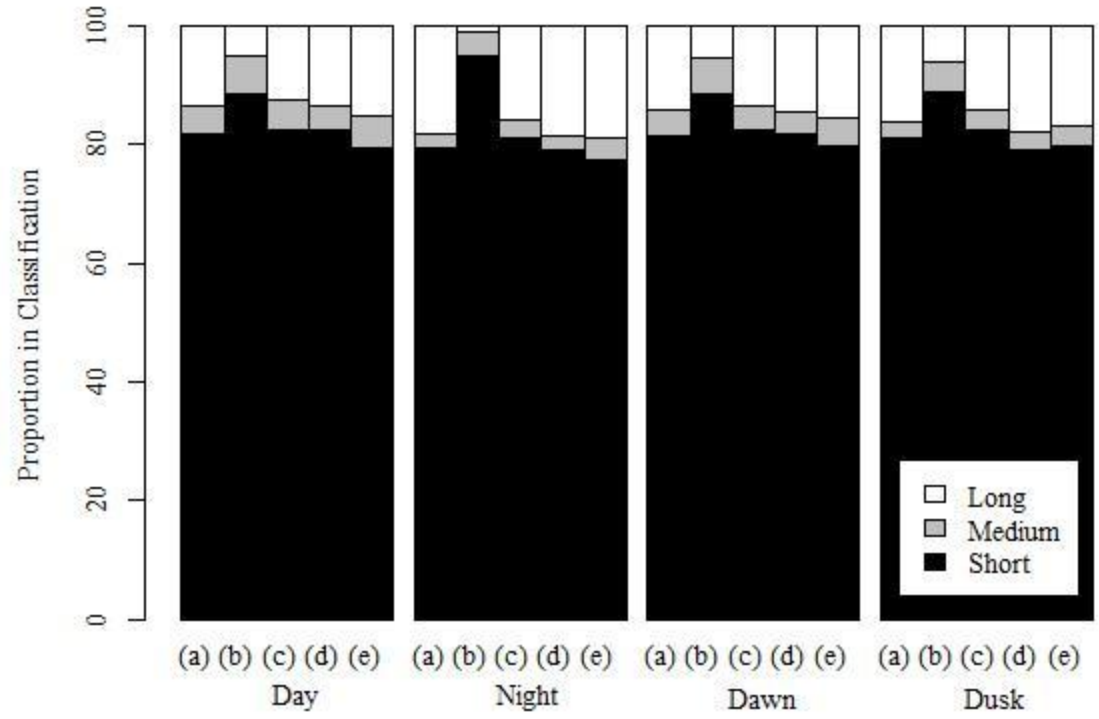

Figure 7.25: Classification Proportions Lighting Factor Stacked Bar Chart
*where (a) represents ground truth, (b) represents Solo Pro II, (c) represents Microloop 702, (d) represents G4, and (e) represents SmartSensor 105

Table 7.23: Percent Correctly Classified by Lighting Levels

	Day	Night	Dawn	Dusk
Solo Pro II	85.6%	79.8%	86.5%	86.0%
Microloop 702	94.8%	96.1%	95.9%	95.9%
G4	96.0%	97.8%	97.4%	97.1%
SmartSensor 105	95.3%	96.2%	95.3%	96.8%

The next factor to be considered was precipitation. Figure 7.26 depicts the classification proportions for the ground truth and various detectors under clear and rainy conditions. Additionally, confusion matrices such as those already presented were analyzed for data subsets of clear and rainy weather, with the percent correctly classified by each detector shown in table 7.24. Based on table 7.24, it appears that the Solo Pro II was more affected by the presence of rain than were any of the other detectors. However, close examination of the ground truth bars in figure 7.26 reveals that there was a higher proportion of long vehicles in the rain subset than the clear subset. Because it was found

that the Solo Pro II had difficulty correctly classifying long vehicles, the decreased correct classification in table 7.24 was probably more closely linked to the proportion of long vehicles in the traffic stream than to the precipitation.

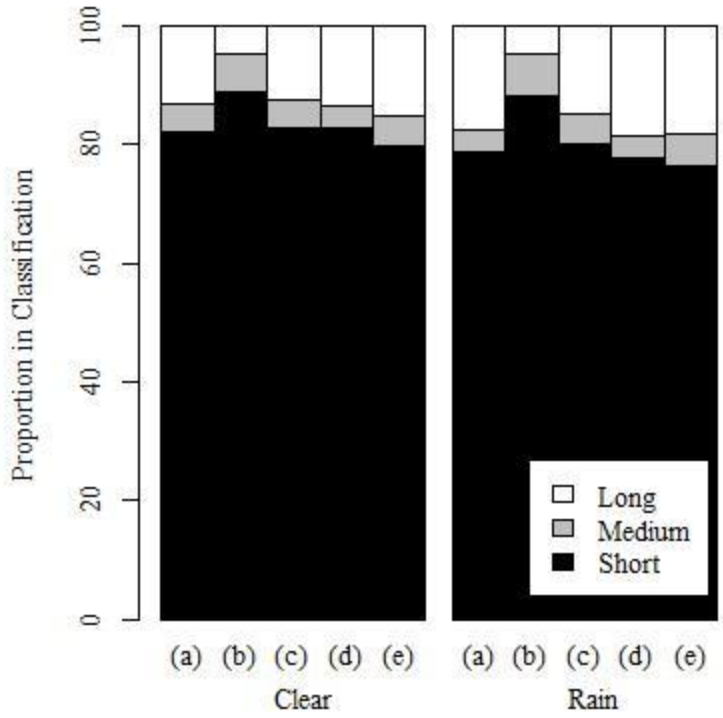

Figure 7.26: Classification Proportions Rain Factor Stacked Bar Chart
*where (a) represents ground truth, (b) represents Solo Pro II, (c) represents Microloop 702, (d) represents G4, and (e) represents SmartSensor 105

Table 7.24: Percent Correctly Classified by Rain Factor

	Clear	Rain
Solo Pro II	85.7%	82.8%
Microloop 702	95.0%	94.6%
G4	96.2%	96.3%
SmartSensor 105	95.4%	95.4%

The final factor to be considered was traffic volume. Figure 7.27 depicts the classification proportions for the ground truth and various detectors under low volume (LOS A or B) and high volume (LOS C or worse) conditions. Additionally, confusion matrices were analyzed for data subsets of low and high volume periods, with the percent correctly classified by each detector presented in table 7.25. While table 7.25 indicates

that all detectors evaluated had either relatively unchanged or improved classification ability in high volume traffic, figure 7.27 reveals that this was most likely due to the higher proportion of short vehicles during high volume periods. For example, note that the percent correctly classified by a null model detector, which classified every vehicle as short, would increase from 79.6% in low volume to 87.9% in high volume based on the ground truth in this data set.

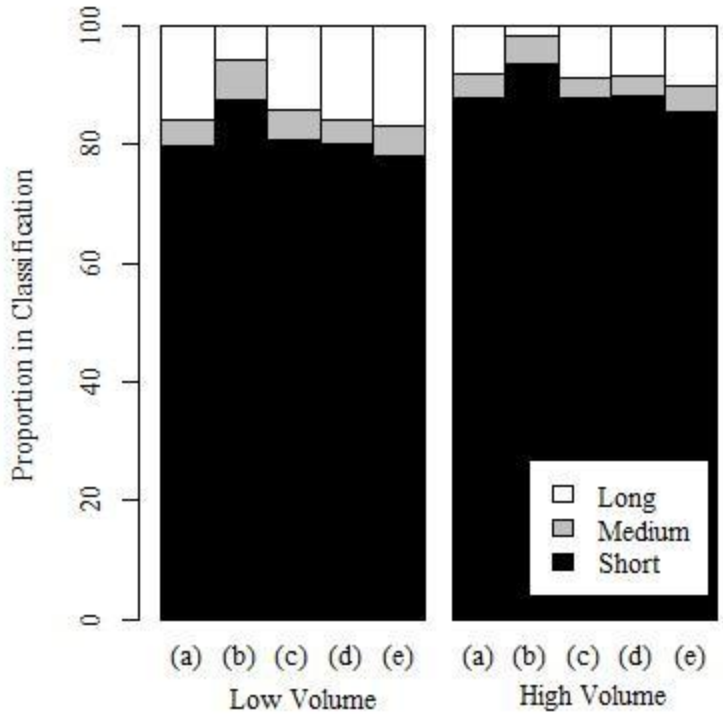

Figure 7.27: Classification Proportions Volume Factor Stacked Bar Chart
*where (a) represents ground truth, (b) represents Solo Pro II, (c) represents Microloop 702, (d) represents G4, and (e) represents SmartSensor 105

Table 7.25: Percent Correctly Classified by Traffic Volume Factor

	Low Volume	High Volume
Solo Pro II	84.3%	88.4%
Microloop 702	94.7%	95.5%
G4	96.2%	96.2%
SmartSensor 105	95.5%	95.2%

The per-vehicle classification analysis performed here indicates that the Microloop 702, G4, and SmartSensor 105 each correctly classified approximately 95% of all vehicles they detected. It is also demonstrated that the correct classification rates of these three detectors were relatively unaffected by lighting, rain, or traffic volume. In contrast, the Solo Pro II correctly classified only 85% of the vehicles it detected. The most frequent classification error committed by the Solo Pro II was to misclassify a long vehicle as a short vehicle. It was found that this type of misclassification by the Solo Pro II was most prevalent under night lighting conditions.

7.4 Chapter Summary

This chapter has provided analyses of the individual vehicle-level detection abilities of the four detectors under evaluation. The relative strengths and weaknesses of the different detectors were demonstrated in the results of this analysis. The disaggregate analysis presented in this chapter indicates the nature of error committed by the different technologies, while aggregate analysis (as presented in chapter 6) indicates the magnitude of these errors in intervals consistent with practical ITS applications.

The analysis of presence detection in this chapter indicated that the G4 and Microloop 702 had the strongest presence detection abilities, with 92.8% and 92.7% correct detection rates, respectively, while the Solo Pro II had a 90.5% correct detection rate, and the SmartSensor 105 lagged with an 83.7% correct detection rate. Further, the SmartSensor 105 correct presence detection rate was found to drop to 67.3% in periods of high traffic volume, compared to 89.0% in low volume periods.

The analysis of per-vehicle speed was conducted with the Microloop 702 data serving as a baseline due to the lack of ground truth speeds. While the SmartSensor 105

had the lowest mean percent deviation from the baseline speed at 2.88%, the variance in percent deviation indicated that the Solo Pro II could most closely resemble the baseline speeds if further calibration was conducted to remove the extreme speed detection bias. As calibrated, the Solo Pro II had a mean percent deviation from the baseline of 17.9%. The speed detection analysis also considered the influence of environmental factors, with mixed results.

Lastly, the per-vehicle classification analysis indicated strong length-based classification from the Microloop 702, G4, and SmartSensor 105, with correct classification rates of 94.9%, 96.2%, and 95.4%. The Solo Pro II struggled with classification, the most frequent problem being the misclassification of long vehicles as short. The Solo Pro II's correct classification rate was 85.4%. Analysis involving the influence of environmental factors indicated that night lighting conditions exacerbated the Solo Pro II's classification problem, correct classification rate dropping to 79.8% in this condition. The classification abilities of the other detectors appeared to be relatively uninfluenced by the documented environmental factors.

CHAPTER 8 CONCLUSIONS

8.1 Summary

In this thesis, four non-intrusive detection systems were evaluated for their ability to detect traffic parameters on a typical urban freeway segment in Nebraska. The four detectors evaluated were the Autoscope Solo Pro II video image processing system, 3M Canoga Microloop 702 magnetic induction system, RTMS G4 microwave radar system, and Wavetronix SmartSensor 105 system. These systems were installed at the NTC/NDOR Non-Intrusive Detector Test Bed along I-80 near the Giles Road interchange in Omaha, Nebraska. The detectors were each calibrated using recommended procedures, and preliminary data were collected so that further calibration could fine-tune detection. After the fine-tuning, all detectors were functioning as expected, and ready for data collection. Vehicle presence/volume, speed, and length-based classification data were collected between March and August of 2011. Additionally, ground truth data was collected through manual observation of video from the test bed. Statistical analysis of the data was performed at both the disaggregate per-vehicle level and various temporal aggregation intervals. Comparisons of the performance of the various detectors were made on a variety of statistical measures relating to accuracy. The analysis also investigated the impact of environmental factors such as lighting and rain on the performance of the various detectors. Lastly, generalized conclusions about the detection performance of the evaluated systems were drawn from the numerous investigated analytical metrics.

8.2 Conclusions

The analysis of vehicle presence detection at the per-vehicle level generally revealed a tradeoff between missed detections and false detections. The G4 and Microloop 702 detectors had the strongest presence detection abilities, with 92.8% and 92.7% correct detection rates, while the Solo Pro II had a 90.5% correct detection rate, and the SmartSensor 105 lagged with an 83.7% correct detection rate. Similar results were found at the one-minute aggregation interval. The G4 had a mean absolute percent error (MAPE) of 5.5%, while the Microloop 702, Solo Pro II, and SmartSensor 105 followed with MAPEs of 6.1%, 6.5%, and 8.2%. The MAPEs of all detectors decreased at the greater aggregation levels of five and fifteen minutes, but at these levels, the Solo Pro II MAPEs were the lowest, followed by the G4, Microloop 702, and SmartSensor 105. This indicates that detector selection could be influenced by aggregation level of required data. Analysis of the effects of various lighting and rain conditions found that the Solo Pro II volume detection accuracy was affected by night lighting conditions and the combined effect of dawn lighting and rain. Microloop 702 and G4 volume detection were found to be affected by the combined effect of dusk lighting and rain, while SmartSensor 105 volume detection was found to not be significantly affected by lighting or rain conditions.

The analysis of speed detection was conducted with the Microloop 702 data serving as a baseline due to the lack of ground truth speeds. The distributions of per-vehicle as well as one, five, and fifteen minute mean speeds indicated that the Solo Pro II was reporting speeds much higher than the other three systems, including the baseline Microloop 702. However, it was concluded that this could be corrected with further calibration. The more intriguing finding was that, while the Microloop 702, Solo Pro II,

and SmartSensor 105 speed distributions all had similar shapes, the G4's mean speed distribution lacked the significant left tail that was present in the other detector's distributions. This was interpreted as the G4 being relatively insensitive to reductions in speed. The primary effect of longer aggregation intervals on speed detection was a reduction in the variance of reported values from each detector as aggregation increased. This was consistent with expectations for data aggregation. The consideration of the impact of environmental factors on speed detection for the various detectors provided mixed results.

Lastly, the detectors were assessed for their ability to classify vehicles into one of three length-based classifications (0-24 ft, 25-40 ft, or 41+ ft). This analysis indicated strong length-based classification from the Microloop 702, G4, and SmartSensor 105, with 94.9%, 96.2%, and 95.4% of vehicles being correctly classified by these three systems, respectively. As the data were temporally aggregated, the accuracies improved (due to an aggregation effect) to the extent that the mean fifteen-minute classification error percentages for the Microloop 702, G4, and SmartSensor 105 were 2.1%, 1.6%, and 2.1%. In contrast, the Solo Pro II struggled with classification, having a per-vehicle correct classification rate of 85.4% and a mean fifteen-minute classification error of 10.4%. The most frequent type of error made by the Solo Pro II classification was misclassifying long vehicles as short. Analysis involving the influence of environmental factors indicated that night lighting conditions exacerbated the Solo Pro II's classification problem. The G4 classification ability was found to be affected by the combination of dusk lighting and rain, which ultimately led to the hypothesis that this detector's classification ability was affected by heavy rainfall. The classification abilities of the

other detectors appeared to be relatively uninfluenced by the documented environmental factors.

When the results of this thesis were compared to results of previous studies which evaluated similar parameters, they were found to generally be comparable but with slightly higher error rates. The fact that the errors rates were on similar orders of magnitude indicated that the results of this thesis were consistent with the body of knowledge on these detectors. The slightly higher error rates were attributed to the fact that this data set included a greater proportion of data from inclement conditions than most of the comparable studies. Also influential in the higher error rates in this study was the fact that most of the analysis herein was performed at a more disaggregate level than many of the previous studies. As discussed in chapter 6, the effect of greater aggregation is generally to decrease error rates.

8.3 Future Research

While this thesis answered a number of questions that aid in the comparison of alternative traffic detection technologies currently available on the market, it also left a number of questions unanswered. As was stated throughout, the evaluation criteria for traffic detectors is application specific. The accuracy assessment provided here represents only one such criterion. Other comparative criteria are system cost, number of traffic parameters estimated, ease of installation, maintenance concerns, power consumption, communications, onboard data storage availability, and reliability. Some of these represent simple questions that can be addressed when a detector is selected for a specific application. Other analytical criteria relating to the life of a detector, such as reliability and maintenance concerns, could warrant future research. Analysis over a longer data

collection period could also provide useful information on the drift or potential deterioration of performance over time. It would be valuable to understand at what intervals a permanent detector should be recalibrated over its life to maintain a desired degree of accuracy.

Additionally, a number of new questions relating to detector accuracy are raised by the results found in this thesis. For example, this thesis found various environmental factors to significantly affect accuracy of some of the detectors evaluated herein. Further analysis is necessary to determine if these affects apply to whole classes of detectors (such as video image processors, microwave radar, magnetic induction, etc.), or specifically to the models tested in this thesis. Analysis of accuracy under snowy conditions could add to the knowledge of precipitation effects on various detection technologies. There is also a continual need to analyze the newest detectors on the market representing each technology.

REFERENCES

1. Middleton, D., D. Gopalakrishna, and M. Raman. Advances in Traffic Data Collection and Management. *Intelligent Transportation Systems: U.S. Department of Transportation.* January 31, 2003. http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS_TE/13766.html. Accessed February 25, 2010.
2. Martin, P. T., Y. Feng, and X. Wang. *Detector Technology Evaluation.* Publication UT-03.30. Utah Department of Transportation, Salt Lake City, UT, 2003.
3. Coifman, B. Freeway Detector Assessment: Aggregate Data from Remote Traffic Microwave Sensor. In *Transportation Research Record: Journal of the Transportation Research Board,* No. 1917, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 149-163.
4. Coifman, B. Estimating Density and Lane Inflow on a Freeway Segment. *Transportation Research Part A,* Vol. 37, No. 8, 2003, pp. 689-701.
5. Klein, L. A., M. K. Mills, and D. R. P. Gibson, *Traffic Detector Handbook: Third Edition - Volume I.* Turner-Fairbank Highway Research Center, McLean, VA, 2006.
6. Klein, L. A. *Sensor Technology and Data Requirements for ITS.* Artech House, Boston, 2001.
7. Zhang, L. *An Evaluation of the Technical and Economic Performance of Weigh-In-Motion Sensing Technology.* Master's thesis, University of Waterloo, Waterloo, Ontario, Canada, 2007.

8. Minge, E., J. Kotzenmacher, and S. Petersen. *Evaluation of Non-Intrusive Technologies for Traffic Detection.* Publication MN/RC 2010-36. Minnesota. Department of Transportation, St. Paul, MN, 2010.
9. ASTM Subcommittee E17.52: Published standards under E17.52 jurisdiction. *ASTM International - Standards Worldwide.* www.astm.org/COMMIT/SUBCOMMIT/E1752.htm. Accessed Sept. 8, 2011.
10. ASTM Standard E2532, 2009, "Standard Test Methods for Evaluating Performance of Highway Traffic Monitoring Device," ASTM International, West Conshohocken, PA, 2009, DOI: 10.1520/E2532-09, www.astm.org.
11. ASTM Standard E2300, 2009, "Standard Specification for Highway Traffic Monitoring Device," ASTM International, West Conshohocken, PA, 2009, DOI: 10.1520/E2300-09, www.astm.org.
12. MacCarley, C. A., S. L. Hockaday, D. Need, and S. Taff. Evaluation of Video Image Processing Systems for Traffic Detection. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1410, Transortation Research Board of the National Academies, Washington, D.C., 1992, pp. 46-49.
13. Malik, J. and S. Russell. *Traffic Surveillance and Detection Technology Development: New Traffic Sensor Technology Final Report*. Publication UCB-ITS-PRR-97-6. UC Berkeley: California Partners for Advanced Transit and Highways (PATH), 1997.

14. MacCarley, A. *City of Anaheim/Caltrans/FHWA Advanced Traffic Control System Field Operational Test Evaluation: Task C Video Traffic Detection System*. Publication UCB-ITS-PRR-98-32. UC Berkley: California Partners for Advanced Transit and Highways (PATH), 1998.
15. Coifman, B. Vehicle Level Evaluation of Loop Detector and the Remote Traffic Microwave Sensor. *Journal of Transportation Engineering,* Vol. 132, 2006, pp. 213-226.
16. MacCarley, C. A. and J. Slonaker. Automated Consensus-Based Data Verification in Caltrans Detector Testbed. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1993, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 124-130.
17. MacCarley, A. and J. Slonaker. *Video Vehicle Detector Verification System (V2DVS).* Publication UCB-ITS-PRR-2008-21. UC Berkley: California Partners for Advanced Transit and Highways (PATH), 2008.
18. Klein, L. A. and M. R. Kelley. *Detection Technology for IVHS Volume I: Final Report.* Publication FHWA-RD-95-100. FHWA, U.S. Department of Transportation, 1996.
19. Klein, L. A., M. R. Kelley, and M. K. Mills. Evaluation of Overhead and In-Ground Vehicle Detector Technologies for Traffic Flow Measurement. *Journal of Testing and Evaluation*, Vol. 25, No. 2, 1997, pp. 205-214.

20. Mimbela, L. E. Y. and L. A. Klein. *Summary of Vehicle Detection and Surveillance Technologies Used In Intelligent Transportation Systems.* The Vehicle Detector Clearinghouse, Southwest Technology Development Institute, New Mexico State University. Las Cruces, NM, 2007.
21. Minnesota Department of Transportation; SRF Consulting Group, Inc. *Field Test of Monitoring of Urban Vehicle Operations Using Non-Intrusive Technologies.*Publication FHWA-PL-97-018. FHWA, U.S. Department of Transportation, 1997.
22. Bahler, S. J., J. M. Kranig, and E. D. Minge. Field Test of Nonintrusive Traffic Detection Technologies. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1643, Transportation Research Board of the National Academies, Washington, D.C., 1998, pp. 161-170.
23. Minnesota Department of Transportation; SRF Consulting Group, Inc. *NIT Phase II: Evaluation of Non-Intrusive Technologies for Traffic Detection.* 2002.
24. Kotzenmacher, J., E. Minge, and B. Hao. *Evaluation of Prtable Non-Intrusive Traffic Detection System*. Publication MN-RC-2005-37. Minnesota Department of Transportation, 2005.
25. Middleton, D. and R. Parker. *Initial Evaluation of Selected Detectors to Replace Inductive Loops on Freeways*. Publication FHWA/TX-00/1439-7. Texas Department of Transportation, 2000.
26. Middleton, D. and R. Parker. *Vehicle Detector Evaluation*. Publication FHWA/TX-03/2119-1. Texas Department of Transportation, 2002.

27. Middleton, D., R. Longmire, and S. Turner. *State of the Art Evaluation of Traffic Detection and Monitoring Systems: Volume I - Phase A & B: Design*. Publication FHWA-AZ-07-627(1). Arizona Department of Transportation, 2007.
28. Grenard, J., D. Bullock, and A. Tarko. *Evaluation of Selected Video Detection Systems at Signalized Intersections*. Publication FHWA/IN/JTRP-2001/22. Indiana Department of Transportation, 2001.
29. Achillides, C. D. and D. M. Bullock. *Performance Metrics for Freeway Sensors*. Publication FHWA/IN/JTRP-2004/37. Indiana Department of Transportation, 2004.
30. Rhodes, A., D. M. Bullock, J. Sturdevant, Z. Clark, and D. G. Candey Jr. Evaluation of the Accuracy of Stop Bar Video Vehicle Detection at Signalized Intersections. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1925, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 134-145.
31. Rhodes, A., E. J. Smaglik, and D. Bullock. *Vendor Comparison of Video Detection Systems*. Publication FHWA/IN/JTRP-2005-30. Indiana Department of Transportation, 2006.
32. Rhodes, A., K. Jennings, and D. M. Bullock. Consistency of Video Detection Activation and Deactivation Times Between Day and Night Periods. *Journal of Transportation Engineering*, Vol. 133, No. 9, 2007, pp. 505-512.
33. Wells, T. J., E. J. Smaglik, and D. M. Bullock. *Health Monitoring Procedures for Freeway Traffic Sensors*. Publication FHWA/IN/JTRP-2006/40. Indiana Department of Transportation, 2008.

34. Zhang, L. *Comparison of Non-Intrusive Traffic Detection Devices for Freeway Application.* Master's Thesis, University of Nebraska-Lincoln, 2006.

35. Grone, B., J. Appiah, and L. Rilett. A Methodology for Comparing Non-Intrusive Traffic Detectors Under Different Operating Conditions. Presented at the 1st International Conference on Transportation Information and Safety (ICTIS), Wuhan, China, 2011.

36. Medina, J. C., R. F. Benekohal, and M. Chitturi. *Evaluation of Video Detection Systems, Volume 1 - Effects of Configuration Changes in the Performance of Video Detection Systems*. Publication FHWA-ICT-08-024. Illinois Department of Transportation, 2008.

37. Medina, J. C., R. F. Benekohal, and M. Chitturi. *Evaluation of Video Detection Systems, Volume 2 - Effects of Illumination Conditions in the Performance of Video Detection Systems*. Publication FHWA-ICT-09-046. Illinois Department of Transportation, 2009.

38. Medina, J. C., R. F. Benekohal, and M. Chitturi. *Evaluation of Video Detection Systems, Volume 3 - Effects of Windy Conditions in the Performance of Video Detection Systems*. Publication FHWA-ICT-09-047. Illinois Department of Transportation, 2009.

39. Medina, J. C., R. F. Benekohal, and M. V. Chitturi. *Evaluation of Video Detection Systems, Volume 4 - Effects of Adverse Weather Conditions in the Performance of Video Detection Systems*. Publication FHWA-ICT-09-039. Illinois Department of Transportation, 2009.

40. Medina, J. C., A. Hajbabaie, and R. F. Benekohal. *Detection Performance of Wireless Magnetometers at a Signalized Intersection and a Railroad Grade Crossing Under Different Weather Conditions.* Presented at 90th Annual Meeting of the Transportation Research Board, Washington, D.C., 2011.

41. Duckworth, G. L., M. L. Frey, C. E. Remer, S. Ritter, and G. Vidaver. *A Comparative Study of Non-Intrusive Traffic Monitoring Sensors*. National Traffic Data Acquisition Conference Proceedings. Vol. 2, pp. 485-514. Publication FHWA-CT-RD 2304-F2-94-1. Connecticut Department of Transportation, 1994.

42. Kim, H., J. H. Lee, S. W. Kim, J. I. Ko, and D. Cho. Ultrasonic Vehicle Detector for Side-Fire Implementation and Extensive Results Including Harsh Conditions. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 2, No. 3, 2001, pp. 127-134.

43. Ha, D. M., J. M. Lee, and Y. D. Kim. Neural-Edge-Based Vehicle Detection and Traffic Parameter Extraction. *Image and Vision Computing*, Vol. 22, No. 11, 2004, pp. 899-907.

44. Martin, P. T., G. Dharmavaram, and A. Stevanovic. *Evaluation of UDOT's Video Detection Systems: System's Performance in Various Test Conditions*. Publication UT-04.14. Utah Department of Transportation, 2004.

45. International Organization for Standardization *Guide to the Expression of Uncertainty in Measurement (GUM)*. ISO, Geneva, Switzerland, 1995.

46. Di Leo, G., A. Pietrosanto, and P. Sommella. Metrological Performance of Traffic Detection Systems. *IEEE Transactions on Instrumentation and Measurement*, Vol. 58, No. 9, 2009, pp. 3199-3206.

47. Yu, X., P. D. Prevedouros, and G. Sulijoadikusumo. Evaluation of Autoscope, SmartSensor HD, and Infra-Red Traffic Logger for Vehicle Classification. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2160, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 77-86.
48. *Traffic Monitoring Guide*. Publication FHWA-PL-01-021. FHWA, U.S. Department of Transportation, 2001.
49. Weather Underground. *Weather History & Data Archive*. www.wunderground.com/history. Accessed October 4, 2011.
50. *RTMS G4 User Manual*. Image Sensing Systems, Inc., 2010.
51. *SmartSensor 105 User Guide v2.2*. Wavetronix LLC, Lindon, Utah, 2009.
52. Balakrishna, R., C. Antoniou, M. Ben-Akiva, H. N. Koutsopoulos, and Y. Wen. Calibration of Microscopic Traffic Simulation Models: Methods and Application. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1999, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 198-207.
53. Theil, H. *Economic Forecasts and Policy*. North-Holland Publishing Company, Amsterdam, 1970.
54. Hourdakis, J., P. G. Michalopoulos, and J. Kottommannil. Practical Procedure for Calibrating Microscopic Traffic Simulation Models. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1852, Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 130-139.

55. Chen, G. AFNI and NIfTI Server for NIMH/NIH/PHS/DHHS/USA/Earth. *Types of Sums of Squares*. National Institutes of Health, May 31, 2011. http://afni.nimh.nih.gov/sscc/gangc/SS.html. Accessed December 22, 2011.

56. Dowdy, S., S. Wearden, and D. Chilko. *Statistics for Research.* Third Edition. John Wiley & Sons, Inc., Hoboken, NJ, 2004.

57. Donnell, E. T., S. C. Hines, K. M. Mahoney, R. J. Porter, and H. McGee. *Speed Concepts: Informational Guide*. Publication FHWA-SA-10-001. FHWA, U.S. Department of Transportation, 2009.

APPENDICES

Appendix A Glossary

Key Terms

Active Detector	A traffic detector which transmits electromagnetic energy to be reflected back toward the detector by a passing vehicle.
Active Infrared Detector	An infrared detector which transmits energy in the infrared portion of the electromagnetic spectrum and detects the portion of this energy reflected off a vehicle in the detection zone.
Advance Detection Zone	A detection zone generally 250 feet or more upstream of an intersection stop bar, where traffic detection can be used to augment signal timing to provide dilemma zone protection.
Baseline	Detector-provided data against which other detectors are analyzed. While the presence of errors in the baseline data is acknowledged, it is assumed to represent a fair standard against which the other detectors can be analyzed.
Call	When a traffic detector installed at an intersection registers vehicle presence in a detection zone and requests right-of-way for that vehicle at the intersection.
Clock Drift	A phenomenon whereby the reported times from two clocks which were once set to the same time tend to diverge as time passes.
Coil	A loop of wire which uses the principle of electromagnet induction to cause a change in current.
Conduit	A tube in which wire or other electrical components can be installed to protect them from environmental conditions.
Correct Detection	A presence detection from a detector that can be correlated to a ground truth detection in the same lane during the same second.
Crosstalk	Unintended interaction between two distinct electromagnetic signals. Can be caused by interaction of two proximate inductive coils or other proximate detectors functioning at similar frequencies.
Density	A measure of the concentration of vehicles on a segment of roadway generally expressed in vehicles per mile or vehicles per mile per lane.

Detection Zone — The physical location on a roadway where a vehicle must be located in order for a traffic detector to register its presence or passage.

Detector — See *Traffic Detector*.

Doppler Radar Detector — A type of microwave radar detector which is capable of registering the passage of moving vehicles in the detection zone, but not presence of stopped vehicles. Also known as a continuous wave radar detector.

Dropped Call — A detector activation which ends before the detected vehicle has vacated the detection zone.

False Call — An improper detector activation when no vehicle was present in the detection zone.

False Detection — A presence detection from a detector that cannot be correlated to a ground truth detection because no ground truth detection was registered in the same lane during the same second.

Frequency — The number of times that an electromagnetic waveform repeats its cycle in 1 second.

Frequency Modulated Continuous Wave Radar Detector — A type of microwave radar detector capable of registering both passage of moving vehicles and presence of stopped vehicles in the detection zone. This is achieved by constantly changing the waveform of the transmitted electromagnetic energy.

Ground Truth — The manually-collected vehicle time stamps and classification assignments obtained by observation of recorded video of the traffic stream. Numerous precedents for manual ground truth are documented in the literature review of this thesis.

Inductive Loop Detector — An active traffic detector composed of one or more coils of wire embedded in or under the roadway, as well as an associated electronics unit. The presence of a vehicle in the detection zone causes the inductance of the wire coils to decrease. This change is registered by the electronics unit as a vehicle passage.

Infrared Detector — A traffic detector which senses electromagnetic waves in the portion of the electromagnetic spectrum between wavelengths of 0.74 μm and 300 μm and frequencies of 400 THz and 1 THz. There are infrared detectors with either passive or active wave sources.

Intrusive Detector	A traffic detector which, by nature of its installation procedure, requires part of the roadway to be blocked during its installation or maintenance. Generally these detectors are installed in the subgrade of the roadway, in the pavement, or directly on the surface of the pavement.
Long Vehicle	A class of vehicle that is defined as having a total length of greater than 40 feet. This length-based class is intended to represent multiple unit heavy vehicles.
Loop Detector	See *Inductive Loop Detector*.
Macro	A procedure which can be defined by a block of code to perform a set of tasks. Macros are frequently used within Microsoft Excel to automate repetitive tasks.
Magnetic Detector	A traffic detector which functions by passively sensing the vertical component of the earth's magnetic field. A perturbation of the earth's magnetic field due to the passage of a large ferrous object through the detection zone is registered as a vehicle detection. Magnetic detectors are generally installed under the roadway and can be either intrusive or non-intrusive depending on the installation procedure.
Magnetometer Detector	More specifically known as a two-axis fluxgate magnetometer, this traffic detector senses both the vertical and horizontal components of the earth's magnetic field. A change in the magnetic field due to a large ferrous object in the detection zone is registered as either a vehicle presence or passage.
Medium Vehicle	A class of vehicle that is defined as having a total length between 25 and 40 feet. This length-based class is intended to represent single unit heavy vehicles.
Microwave Radar Detector	An active, non-intrusive traffic detector installed above or beside the roadway which functions by transmitting and receiving electromagnetic energy in the microwave range of the electromagnetic spectrum (wavelengths from 1 mm to 1 m and frequencies from 300 GHz to 300 MHz).
Missed Call	The lack of a detector activation when a vehicle was present in the detection zone.
Missed Detection	A ground truth detection that cannot be correlated to a detector-reported detection because no detector-reported detection was registered in the same lane during the same second for the specified detector.

Non-Intrusive Detector — A traffic detector which, by nature of its installation procedure, allows the roadway to remain fully operational during its installation or maintenance. Generally these detectors are installed above the roadway surface either offset from the nearest lane in a side-fire configuration or directly over the roadway in an overhead configuration.

Occlusion — A phenomenon whereby a tall vehicle in a lane nearer to an overhead or side-fire detector either causes false activation of a detection zone in a lane further from the detector, or "hides" a vehicle in a lane further from the detector, causing a missed detection.

Occupancy — A measure of the percentage of time in which a detection zone is occupied by a vehicle. Occupancy is frequently used as a proxy for density.

Overhead Configuration — An installation in which a non-intrusive detector is mounted on a support structure directly over the roadway in order to detect vehicles passing beneath it.

Passive Acoustic Detector — A non-intrusive traffic detector which functions by passively sensing audible noise created by a vehicle's engine, exhaust, and tires.

Passive Detector — A traffic detector which does not transmit electromagnetic energy of its own but rather detects energy emitted by objects in its detection zone or emitted by an external source and reflected off objects in the detection zone.

Passive Infrared Detector — An infrared detector which does not transmit energy of its own, but detects energy emitted by the vehicle and energy emitted by the sun and atmosphere reflected off the vehicle.

Pull Box — An underground container into which electrical conduit runs so that appropriate wire or cable splices can be created or serviced through a removable cover flush with the ground level.

Short Vehicle — A class of vehicle that is defined as having a total length of less than 25 feet. This length-based class is intended to represent passenger vehicles.

Side-Fire Configuration — An installation in which a non-intrusive detector is mounted on a support structure on the side of the road and offset a given distance from the nearest lane of traffic.

Speed Trap	A configuration of detectors in which two detectors are placed in the same lane at a known distance apart. Speed and vehicle length are able to be determined based on rising and falling edge time stamps for the two detectors. This configuration is typical for loop detectors.
Spillover	A phenomenon whereby a vehicle's headlights, shadow, or large magnetic footprint cause a detection to be registered in the detection zone of an adjacent lane.
Stuck-On Call	A detector activation which persists after the detected vehicle has vacated the detection zone. This type of error can result in messed calls for subsequent vehicles entering the same detection zone.
Test Bed	An intersection or segment of roadway outfitted with appropriate infrastructure for comparative analysis of traffic detectors.
Tracking	A class of video image processing algorithm which functions by following or "tracking" a moving object from the time it enters the image until the time it leaves the image.
Traffic Detector	A device which is capable of registering the presence or passage of automotive vehicles at a given point on the roadway. In addition to presence and passage, traffic detectors can also potentially provide data on other physical characteristics of the detected vehicles.
Trip-Line	A class of video image processing algorithm which functions by determining when a moving object moves through a specific area of the video image, thereby "tripping" the detector.
Ultrasonic Detector	An active traffic detector which functions by transmitting high frequency sound waves (above the human audible range) and registering the reflection of the wave from a vehicle in the detection zone.
Video Image Processor	A passive traffic sensor which functions by processing a video signal through a series of algorithms which separate moving objects from the background image and interpret the moving objects as vehicles in a detection zone.
Virtual Detector	An image overlay which is used in video image processing traffic detectors to define which pixels are to be monitored for changes by the image processing software and how those changes are to be interpreted as detections.

Weigh-in-Motion Detector A class of traffic detector employed for the specific purpose of determining wheel, axle, or axle group weight and aggregating this into vehicle weight for vehicles moving at high speeds. Weigh-in-Motion detectors are generally based on piezoelectric, bending plate, or load cell technologies.

Acronyms

ADOT	Arizona Department of Transportation
AEVL	Average Effective Vehicle Length
ANOVA	Analysis of Variance
APD	Absolute Percent Deviation / Absolute Percent Deviation
AVI	Automatic Vehicle Identification
CW	Continuous Wave
FHWA	Federal Highway Administration
FMCW	Frequency Modulated Continuous Wave
GIS	Geographic Information System
GPS	Global Positioning System
GUM	Guide to the Expression of Uncertainty in Measurement
INDOT	Indiana Department of Transportation
IR	Infrared
ISO	International Organization for Standardization
ITS	Intelligent Transportation Systems
IVHS	Intelligent Vehicle-Highway System
LOS	Level of Service
MAPD	Mean Absolute Percent Difference
MAPE	Mean Absolute Percent Error
MPD	Mean Percent Difference
MPE	Mean Percent Error
NDOR	Nebraska Department of Roads
NEMA	National Electrical Manufacturers Association
NTC	Nebraska Transportation Center

NTSC	National Television System Committee
PATH	Partners for Advanced Transportation technology
PNITDS	Portable Non-Intrusive Traffic Detection System
PTZ	Pan Tilt Zoom
PVC	Polyvinyl Chloride
RMSE	Root Mean-Square Error
RTMS	Remote Traffic Microwave Sensor
SCOOT	Split Cycle Offset Optimization Technique
TIRTL	The Infra-Red Traffic Logger
TMC	Traffic Management Center
TMD	Traffic Monitoring Device
TTI	Texas Transportation Institute
V2DVS	Video Vehicle Detector Verification System
VIP	Video Image Processor
VPN	Virtual Private Network
VTDS	Video Traffic Detection System
WIM	Weigh-in-Motion
XML	Extensible Markup Language

Appendix B Macros for Automated Step in Clock Synchronization

There were five powerful macros employed in the clock synchronization process that significantly reduced the amount of manual work required to synchronize clocks for the analyzed detectors. The following is the macro code written for this purpose.

```
Sub ClockSynchAllDetectors()

    ' this macro runs the four macros that adjust timestamps of the four detectors
    Debug.Print "Beginning " & Now
    Call clockSynchAutoscope    ' this line runs Sub clockSynchAutoscope()
    Debug.Print "Autoscope " & Now
    Call clockSynchMicroloop    ' this line runs Sub clockSynchMicroloop()
    Debug.Print "Microloop " & Now
    Call clockSynchG4           ' this line runs Sub clockSynchG4()
    Debug.Print "G4 " & Now
    Call clockSynchSmartSensor  ' this line runs Sub clockSynchSmartSensor()
    Debug.Print "SmartSensor " & Now

End Sub


Sub clockSynchAutoscope()

    ' this macro adjusts Autoscope timestamps +/- 1 second to match the nearest ground truth
    ' timestamp in the same lane
Debug.Print "Beginning " & Now
    ' the next lines define variables
Dim A As Worksheet
Dim S(1 To 3) As Worksheet
Dim i As Integer
Dim t1 As Date
Dim t2 As Date
Dim rFound As Range
Dim last As Boolean
    ' the next lines define which worksheets are referred to as S(1), S(2), and S(3)
Set S(1) = Sheets("Lane1")
Set S(2) = Sheets("Lane2")
Set S(3) = Sheets("Lane3")
    ' the next lines format the timestamps in the Autoscope worksheet so that the .Find method
    ' works correctly later on
Worksheets("Autoscope").Columns("K:M").NumberFormat = "[$-F400]h:mm:ss AM/PM"
For i = 1 To 3  ' this for loop loops through the worksheets for the three lanes
    S(i).Activate   ' this activates one of the lane worksheets
    Range("C2").Select  ' column C is the column with autoscope one second counts in it; row 2
                        ' represents 00:00:00 (midnight) for the given day
    last = False    ' initializes last (false for last value moved up, true for last value moved down)
    Do Until ActiveCell.Row = 86402 ' row 86401 represents 11:59:59 therefore this Do Until loop
                                    ' does every second for the day
```

```
If ActiveCell.Value <> "" Then ' if the autoscope one second count for the current second is
                                ' not "" (null) then
  If ActiveCell.Offset(0, -1).Value = "" Then ' if the ground truth one second count for the
                                               ' current second is null then
    If ActiveCell.Offset(-1, 0).Value = "" And _
    ActiveCell.Offset(-1, -1).Value <> "" And _
    ActiveCell.Offset(1, 0).Value = "" And _
    ActiveCell.Offset(1, -1).Value <> "" Then
        ' if the autoscope count for the previous second is null and the ground truth count
        ' for the previous second is not null and the autoscope count for the next second
        ' is null and the ground truth count for the next second is not null then
      If last = False Then ' if the last autoscope timestamp adjustment was to subtract one
                           ' second then
        t1 = ActiveCell.Offset(0, -2).Value ' t1 is the current second
        t2 = ActiveCell.Offset(-1, -2).Value ' t2 is the previous second
        With Worksheets("Autoscope").Columns(i + 10)
                ' go to the autoscope worksheet and the column corresponding to lane i
          Set rFound = .Find(What:=t1, LookIn:=xlValues)
                ' find the autoscope timestamp matching the current second
          rFound.Value = t2   ' replace that autoscope timestamp with a timestamp of the
                              ' previous second (i.e. subtract 1 second from that
                              ' autoscope timestamp)
        End With
        last = False    ' set last equal to false
      Else    ' if last is true
        t1 = ActiveCell.Offset(0, -2).Value ' t1 is the current second
        t2 = ActiveCell.Offset(1, -2).Value ' t2 is the next second
        With Worksheets("Autoscope").Columns(i + 10)
                'go to the autoscope worksheet and the column corresponding to lane i
          Set rFound = .Find(What:=t1, LookIn:=xlValues)
                'find the autoscope timestamp matching the current second
          rFound.Value = t2   ' replace that autoscope timestamp with a timestamp of the
                              ' next second (i.e. add 1 second from that autoscope timestamp)
        End With
        last = True ' set last equal to true
      End If
    ElseIf ActiveCell.Offset(-1, 0).Value = "" And _
    ActiveCell.Offset(-1, -1).Value <> "" Then
                ' if ther is no ground truth timestamp for the same second as the current
                ' autoscope timestamp, and there is a ground truth timestamps 1 second
                ' before but not 1 second after the current autoscope timestamp then
        t1 = ActiveCell.Offset(0, -2).Value ' t1 is the current second
        t2 = ActiveCell.Offset(-1, -2).Value    ' t2 is the previous second
        With Worksheets("Autoscope").Columns(i + 10)
                ' go to the autoscope worksheet and the column corresponding to lane i
          Set rFound = .Find(What:=t1, LookIn:=xlValues)
                ' find the autoscope timestamp matching the current second
          rFound.Value = t2   ' replace that autoscope timestamp with a timestamp of the
                ' next second (i.e. add 1 second from that autoscope timestamp)
        End With
        last = False ' set last equal to false
    ElseIf ActiveCell.Offset(1, 0).Value = "" And _
    ActiveCell.Offset(1, -1).Value <> "" Then ' if there is no ground truth timestamp for the
                ' same second as the current autoscope timestamp, and there is a
                ' ground truth timestamps 1 second after but not 1 second before the
                ' current autoscope timestamp then
```

```
                t1 = ActiveCell.Offset(0, -2).Value ' t1 is the current second
                t2 = ActiveCell.Offset(1, -2).Value ' t2 is the next second
                With Worksheets("Autoscope").Columns(i + 10)
                        ' go to the autoscope worksheet and the column corresponding to lane i
                   Set rFound = .Find(What:=t1, LookIn:=xlValues)
                        ' find the autoscope timestamp matching the current second
                   rFound.Value = t2   ' replace that autoscope timestamp with a timestamp of the
                        ' next second (i.e. add 1 second from that autoscope timestamp)
                End With
                last = True ' set last equal to true
             End If
          End If
       End If
       ActiveCell.Offset(1, 0).Select  ' select autoscope one second count for the next second
   Loop    ' go back to beginning of Do Until loop
Next i  ' go back to beginning of For loop with i incremented

Worksheets("Autoscope").Columns("K:M").NumberFormat = _
"h:mm:ss;@"     ' revert autoscope timestamps to original time format
          ' next lines erase autoscope 1 second counts from worksheets Lane1, Lane2, and Lane3
For i = 1 To 3
   S(i).Activate
   Range("C2:C86500").Select
   Selection.ClearContents
   Range("C1").Select
Next i
Debug.Print "calcAutoscope " & Now
          ' next line calls the Sub calcAutoscope() macro which calculates 1 second autoscope
          ' counts in worksheets Lane1, Lane2, and Lane3 based on the newly synchronized
          ' autoscope timestamps
Call calcAutoscope
Debug.Print "Ending " & Now

End Sub


Sub clockSynchMicroloop()

   ' this subroutine employs similar logic to Sub clockSynchAutoscope() with the major exception
   ' that while the three lanes of autoscope timestamps are in three columns of the same
   ' worksheet, the three lanes of microloop timestamps are in similar columns of three distinct
   ' worksheets called Microloop1, Microloop2, and Microloop3
Debug.Print "Beginning " & Now
Dim S(1 To 3) As Worksheet
Dim i As Integer
Dim t1 As Date
Dim t2 As Date
Dim rFound As Range
Dim last As Boolean
Set S(1) = Sheets("Lane1")
Set S(2) = Sheets("Lane2")
Set S(3) = Sheets("Lane3")
For i = 1 To 3
   Worksheets("Microloop" & i).Columns("G:G").NumberFormat = _
   "[$-F400]h:mm:ss AM/PM"
Next i
```

```
For i = 1 To 3
   S(i).Activate
   Range("D2").Select
   last = False    ' false for last value moved up, true for last value moved down
   Do Until ActiveCell.Row = 86402
      If ActiveCell.Value <> "" Then
         If ActiveCell.Offset(0, -2).Value = "" Then
            If ActiveCell.Offset(-1, 0).Value = "" And _
            ActiveCell.Offset(-1, -2).Value <> "" And _
            ActiveCell.Offset(1, 0).Value = "" And _
            ActiveCell.Offset(1, -2).Value <> "" Then
               If last = False Then
                  t1 = ActiveCell.Offset(0, -3).Value
                  t2 = ActiveCell.Offset(-1, -3).Value
                  With Worksheets("Microloop" & i).Columns(7)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = False
               Else
                  t1 = ActiveCell.Offset(0, -3).Value
                  t2 = ActiveCell.Offset(1, -3).Value
                  With Worksheets("Microloop" & i).Columns(7)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = True
               End If
            ElseIf ActiveCell.Offset(-1, 0).Value = "" And _
            ActiveCell.Offset(-1, -2).Value <> "" Then
                  t1 = ActiveCell.Offset(0, -3).Value
                  t2 = ActiveCell.Offset(-1, -3).Value
                  With Worksheets("Microloop" & i).Columns(7)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = False
            ElseIf ActiveCell.Offset(1, 0).Value = "" And _
            ActiveCell.Offset(1, -2).Value <> "" Then
                  t1 = ActiveCell.Offset(0, -3).Value
                  t2 = ActiveCell.Offset(1, -3).Value
                  With Worksheets("Microloop" & i).Columns(7)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = True
            End If
         End If
      End If
      ActiveCell.Offset(1, 0).Select
   Loop
Next i
For i = 1 To 3
   Worksheets("Microloop" & i).Columns("G:G").NumberFormat = _
   "h:mm:ss;@"
```

```
Next i
For i = 1 To 3
   S(i).Activate
   Range("D2:D86500").Select
   Selection.ClearContents
   Range("D1").Select
Next i
Debug.Print "calcMicroloop " & Now
For i = 1 To 3
   Call calcMicroloop(i)
Next i
Debug.Print "Ending " & Now

End Sub


Sub clockSynchG4()

   ' this subroutine employs similar logic to Sub clockSynchAutoscope()
Debug.Print "Beginning " & Now
Dim A As Worksheet
Dim S(1 To 3) As Worksheet
Dim i As Integer
Dim t1 As Date
Dim t2 As Date
Dim rFound As Range
Dim last As Boolean
Set S(1) = Sheets("Lane1")
Set S(2) = Sheets("Lane2")
Set S(3) = Sheets("Lane3")
Worksheets("G4").Columns("I:K").NumberFormat = _
"[$-F400]h:mm:ss AM/PM"
For i = 1 To 3
   S(i).Activate
   Range("E2").Select
   last = False    ' false for last value moved up, true for last value moved down
   Do Until ActiveCell.Row = 86402
      If ActiveCell.Value <> "" Then
         If ActiveCell.Offset(0, -3).Value = "" Then
            If ActiveCell.Offset(-1, 0).Value = "" And _
            ActiveCell.Offset(-1, -3).Value <> "" And _
            ActiveCell.Offset(1, 0).Value = "" And _
            ActiveCell.Offset(1, -3).Value <> "" Then
               If last = False Then
                  t1 = ActiveCell.Offset(0, -4).Value
                  t2 = ActiveCell.Offset(-1, -4).Value
                  With Worksheets("G4").Columns(i + 8)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = False
               Else
                  t1 = ActiveCell.Offset(0, -4).Value
                  t2 = ActiveCell.Offset(1, -4).Value
                  With Worksheets("G4").Columns(i + 8)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
```

```
                    rFound.Value = t2
                End With
                last = True
            End If
        ElseIf ActiveCell.Offset(-1, 0).Value = "" And _
        ActiveCell.Offset(-1, -3).Value <> "" Then
            t1 = ActiveCell.Offset(0, -4).Value
            t2 = ActiveCell.Offset(-1, -4).Value
            With Worksheets("G4").Columns(i + 8)
                Set rFound = .Find(What:=t1, LookIn:=xlValues)
                rFound.Value = t2
            End With
            last = False
        ElseIf ActiveCell.Offset(1, 0).Value = "" And _
        ActiveCell.Offset(1, -3).Value <> "" Then
            t1 = ActiveCell.Offset(0, -4).Value
            t2 = ActiveCell.Offset(1, -4).Value
            With Worksheets("G4").Columns(i + 8)
                Set rFound = .Find(What:=t1, LookIn:=xlValues)
                rFound.Value = t2
            End With
            last = True
        End If
      End If
    End If
    ActiveCell.Offset(1, 0).Select
  Loop
Next i
Worksheets("G4").Columns("I:K").NumberFormat = _
"h:mm:ss;@"
For i = 1 To 3
  S(i).Activate
  Range("E2:E86500").Select
  Selection.ClearContents
  Range("E1").Select
Next i
Debug.Print "calcG4 " & Now
Call calcG4
Debug.Print "Ending " & Now

End Sub


Sub clockSynchSmartSensor()

  ' this subroutine employs similar logic to Sub clockSynchAutoscope()
Debug.Print "Beginning " & Now
Dim A As Worksheet
Dim S(1 To 3) As Worksheet
Dim i As Integer
Dim t1 As Date
Dim t2 As Date
Dim rFound As Range
Dim last As Boolean
Set S(1) = Sheets("Lane1")
Set S(2) = Sheets("Lane2")
```

```
Set S(3) = Sheets("Lane3")
Worksheets("SmartSensor").Columns("J:L").NumberFormat = _
"[$-F400]h:mm:ss AM/PM"
Worksheets("SmartSensor").Columns("J:L").ColumnWidth = 11
For i = 1 To 3
   S(i).Activate
   Range("F2").Select
   last = False    ' false for last value moved up, true for last value moved down
   Do Until ActiveCell.Row = 86402
      If ActiveCell.Value <> "" Then
         If ActiveCell.Offset(0, -4).Value = "" Then
            If ActiveCell.Offset(-1, 0).Value = "" And _
            ActiveCell.Offset(-1, -4).Value <> "" And _
            ActiveCell.Offset(1, 0).Value = "" And _
            ActiveCell.Offset(1, -4).Value <> "" Then
               If last = False Then
                  t1 = ActiveCell.Offset(0, -5).Value
                  t2 = ActiveCell.Offset(-1, -5).Value
                  With Worksheets("SmartSensor").Columns(i + 9)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = False
               Else
                  t1 = ActiveCell.Offset(0, -5).Value
                  t2 = ActiveCell.Offset(1, -5).Value
                  With Worksheets("SmartSensor").Columns(i + 9)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = True
               End If
            ElseIf ActiveCell.Offset(-1, 0).Value = "" And _
            ActiveCell.Offset(-1, -4).Value <> "" Then
                  t1 = ActiveCell.Offset(0, -5).Value
                  t2 = ActiveCell.Offset(-1, -5).Value
                  With Worksheets("SmartSensor").Columns(i + 9)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = False
            ElseIf ActiveCell.Offset(1, 0).Value = "" And _
            ActiveCell.Offset(1, -4).Value <> "" Then
                  t1 = ActiveCell.Offset(0, -5).Value
                  t2 = ActiveCell.Offset(1, -5).Value
                  With Worksheets("SmartSensor").Columns(i + 9)
                     Set rFound = .Find(What:=t1, LookIn:=xlValues)
                     rFound.Value = t2
                  End With
                  last = True
            End If
         End If
      End If
      ActiveCell.Offset(1, 0).Select
   Loop
Next i
```

```
Worksheets("SmartSensor").Columns("J:L").NumberFormat = _
"h:mm:ss;@"
For i = 1 To 3
    S(i).Activate
    Range("F2:F86500").Select
    Selection.ClearContents
    Range("F1").Select
Next i
Debug.Print "calcSmartSensor " & Now
Call calcSmartSensor
Debug.Print "Ending " & Now

End Sub
```

Appendix C One-Minute Volume ANOVA Thinning

One of the assumptions for an analysis of variance is independence of data or a lack of autocorrelation. The autocorrelation of a data set can be seen in index plots and correlograms. Figure C.1 displays the one-minute volume percent error ANOVA residuals for each detector, while Figure C.2 shows the correlograms associated with this data. The dashed lines in correlograms indicate the 95% confidence interval for no statistically significant correlation. Autocorrelation factors (ACFs) outside this interval indicate potentially significant correlations. In Figure C.2 it can be seen that all four detectors appear to have significant autocorrelation. An attempt was made to remove this correlation through thinning the full data set by a factor of 10, which left 147 data points of an original 1,467. The index plots for this thinned data set are given in Figure C.3, and the correlograms are given in Figure C.4. The autocorrelation factors for the Solo Pro II, Microloop 702, and G4 were mostly non-significant at this level of thinning, with potentially significant factors having no recognizable patterns, indicating that the potentially significant factors can be attributed to white noise. Therefore, the data thinned at this level was selected to be analyzed with ANOVA for these three detectors. The autocorrelation for the SmartSensor 105 appears to remain significant at this level of thinning based on Figure C.4(d). Therefore, the data set for this detector was thinned by a factor of 20, leaving 74 data points. The index plot and correlogram for this thinned data set are given in figures C.5 and C.6. As there is only one potentially significant autocorrelation factor at this level of thinning, it was determined to conduct the ANOVA for this detector on the factor 20 thinned data.

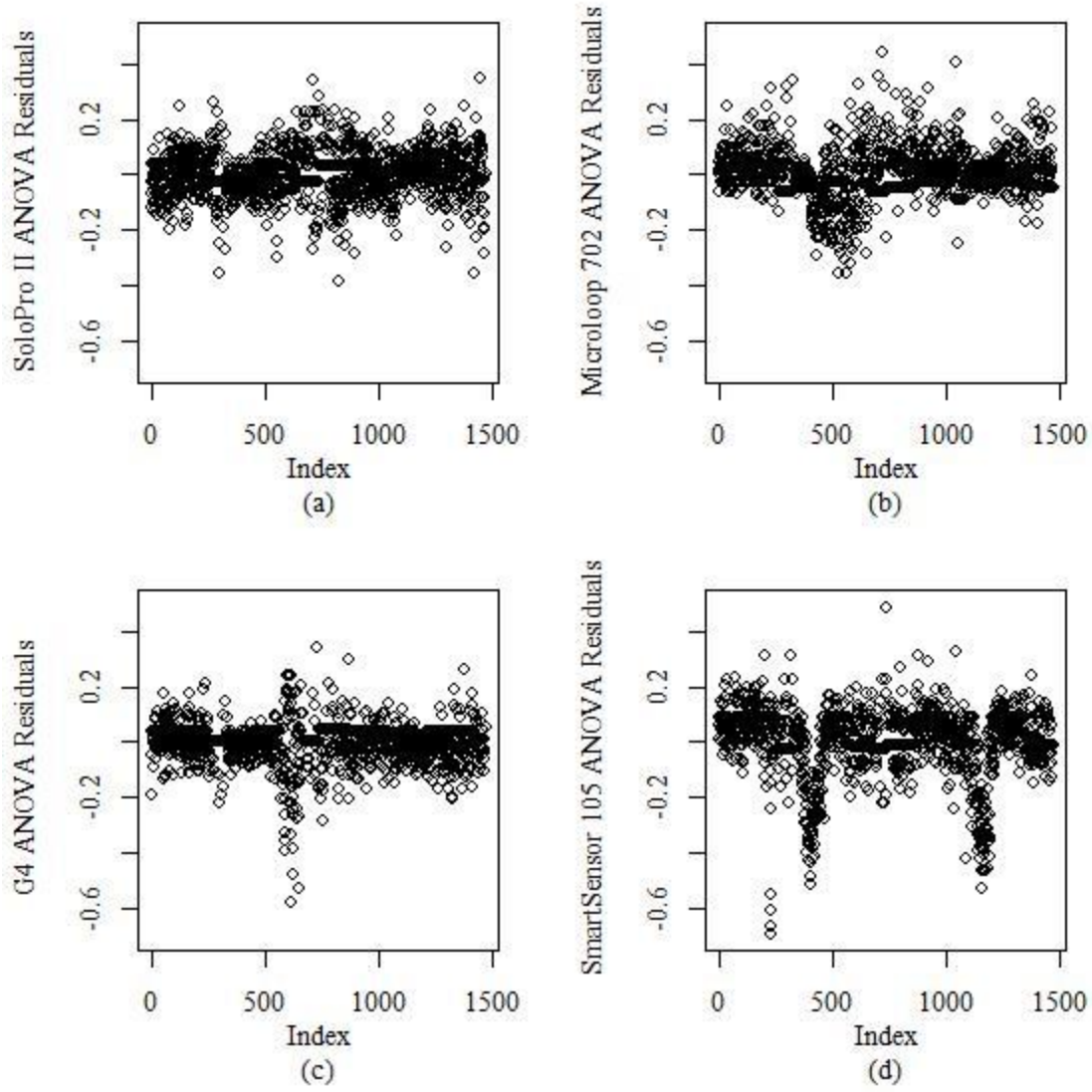

Figure C.1: Full Data One-Minute Volume Percent Error ANOVA Residual Index Plots for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

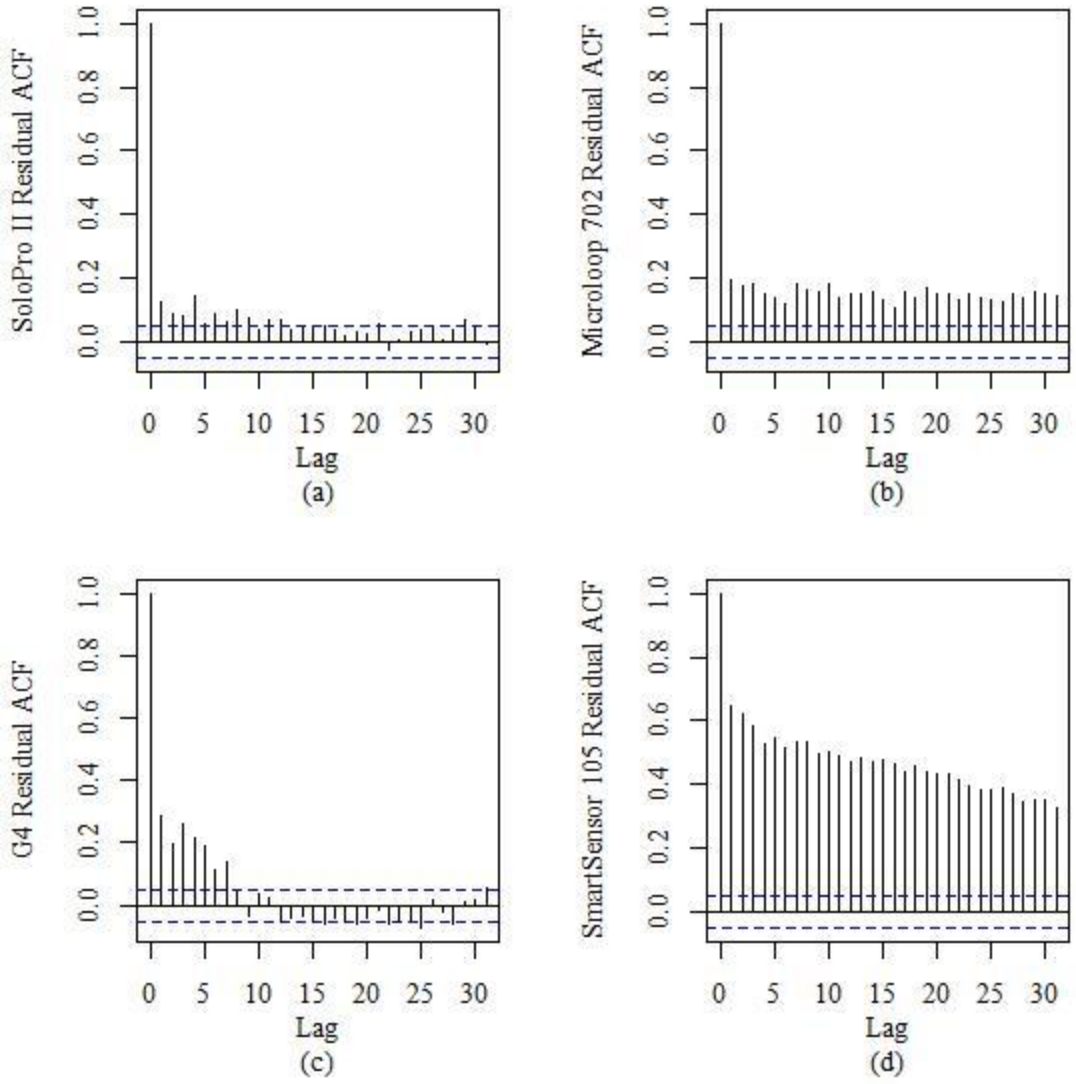

Figure C.2: Full Data One-Minute Volume Percent Error ANOVA Residual Correlograms for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

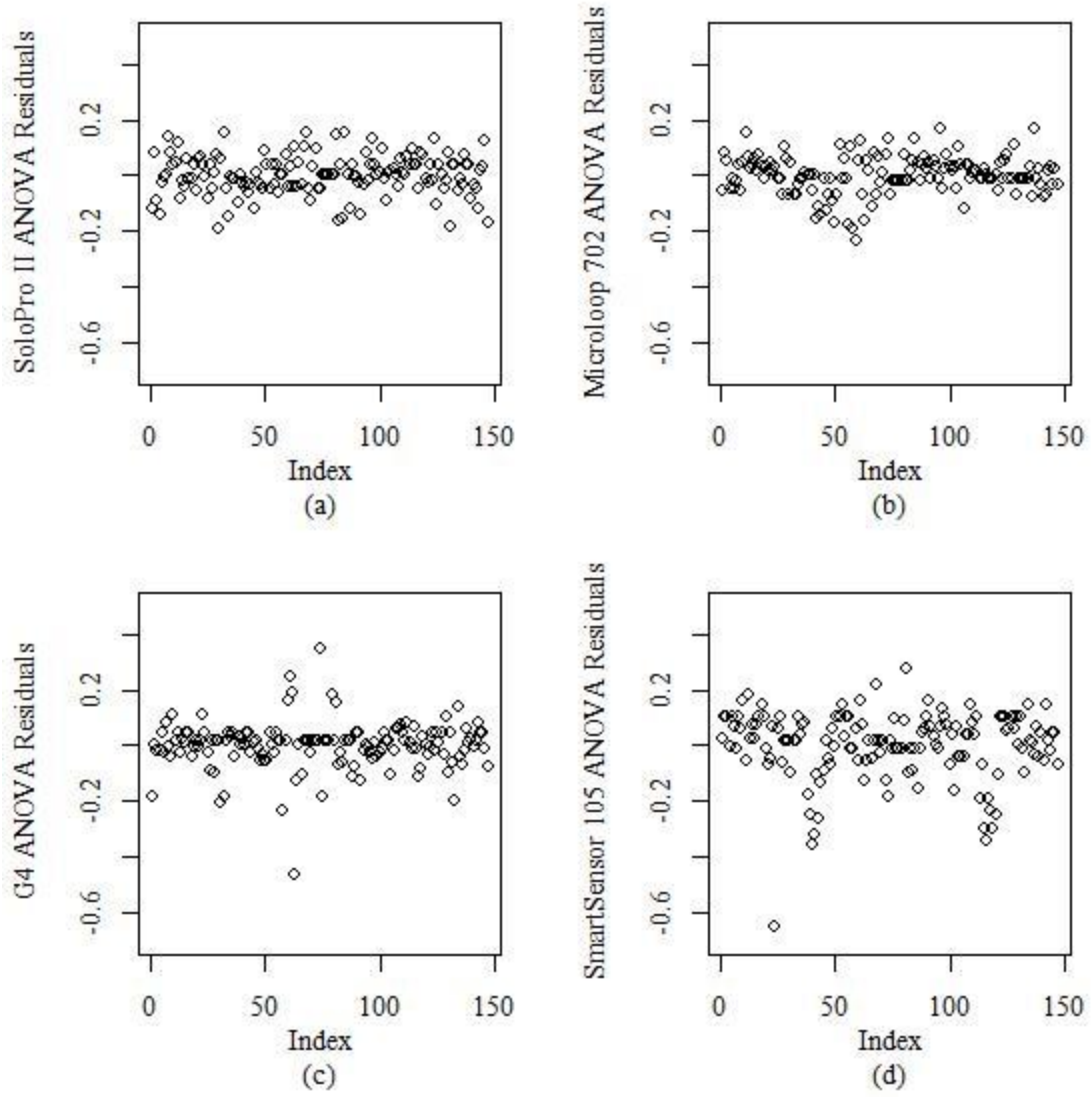

Figure C.3: Factor 10 Thinned One-Minute Volume Percent Error ANOVA Residual Index Plots for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

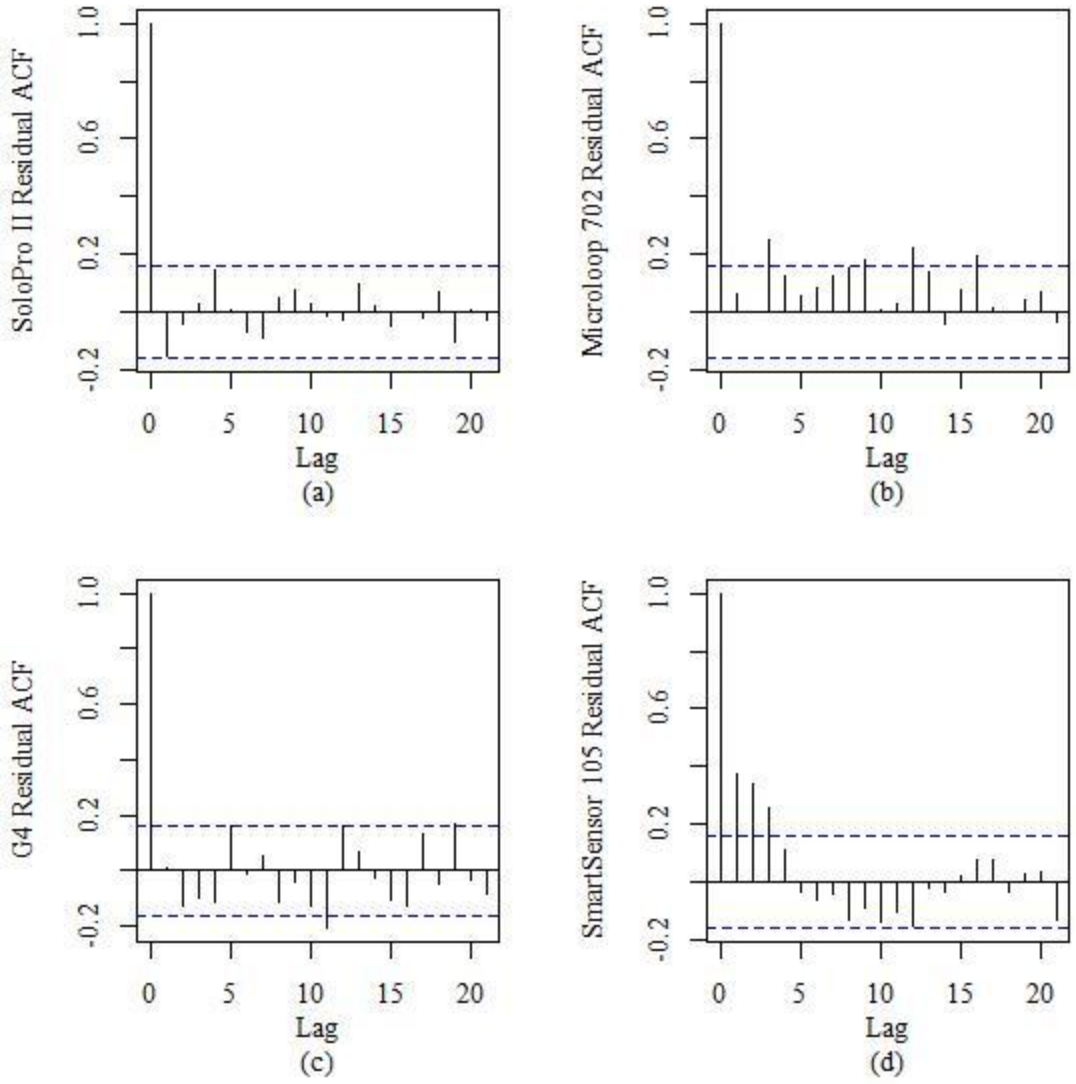

Figure C.4: Factor 10 Thinned One-Minute Volume Percent Error ANOVA Residual Correlograms for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

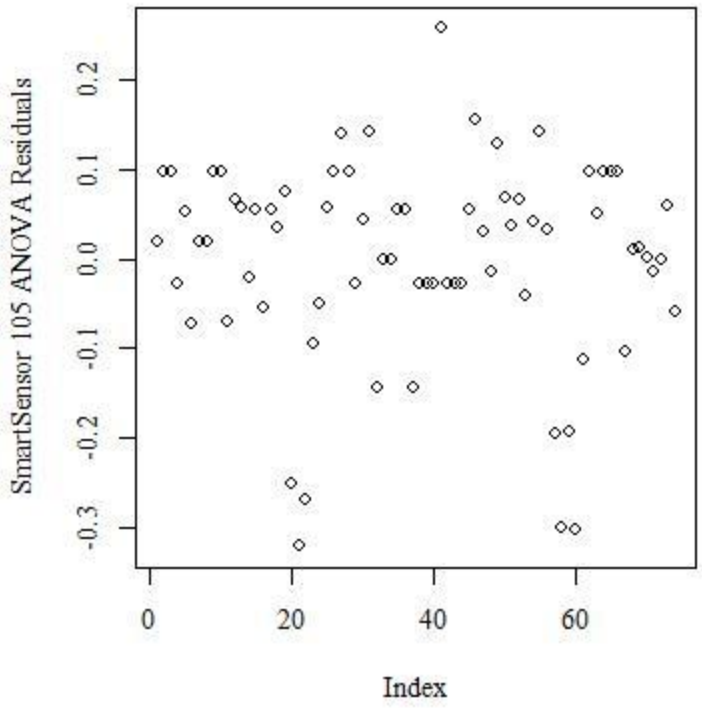

Figure C.5: Factor 20 Thinned One-Minute Volume Percent Error ANOVA Residual Index Plot for SmartSensor 105

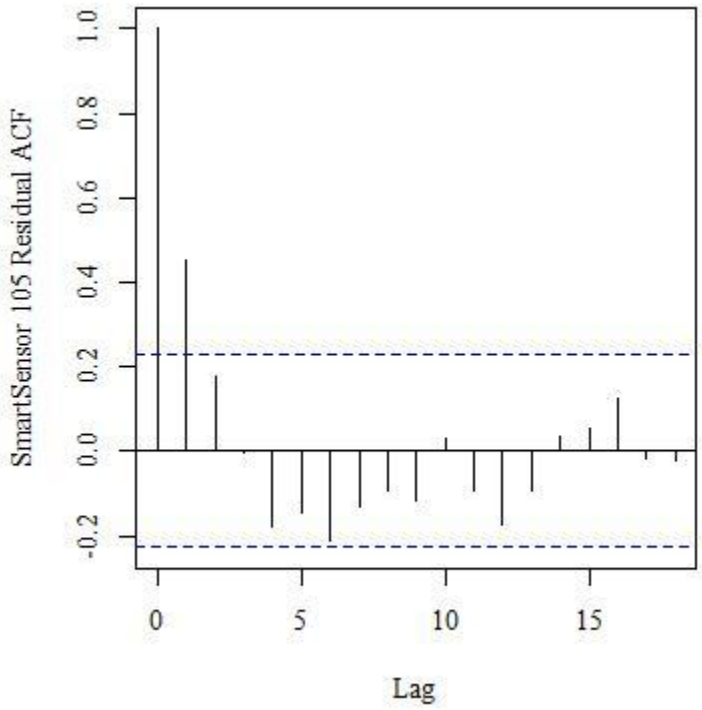

Figure C.6: Factor 20 Thinned One-Minute Volume Percent Error ANOVA Residual Correlogram for SmartSensor 105

Appendix D Five-Minute Analysis Additional Figures and Tables

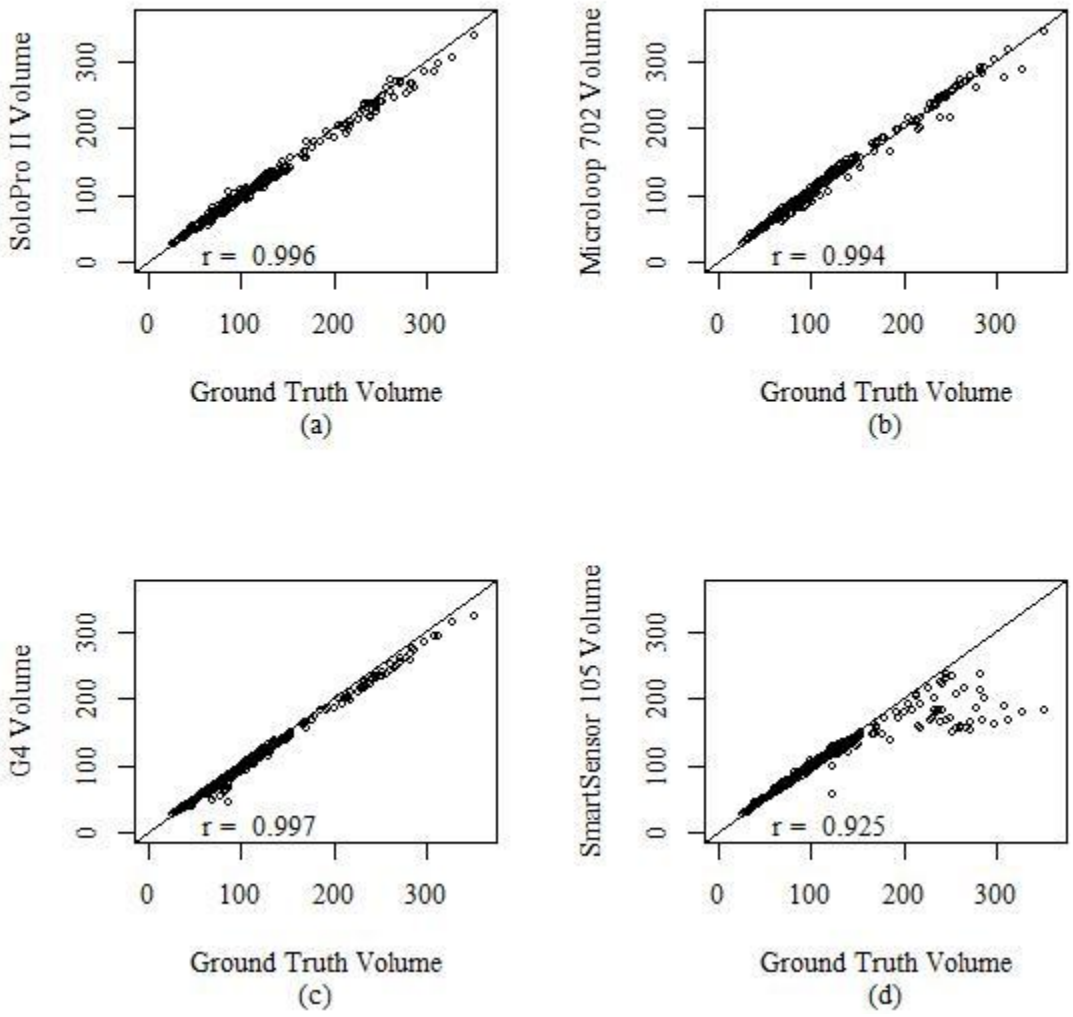

Figure D.1: Five-Minute Volume Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

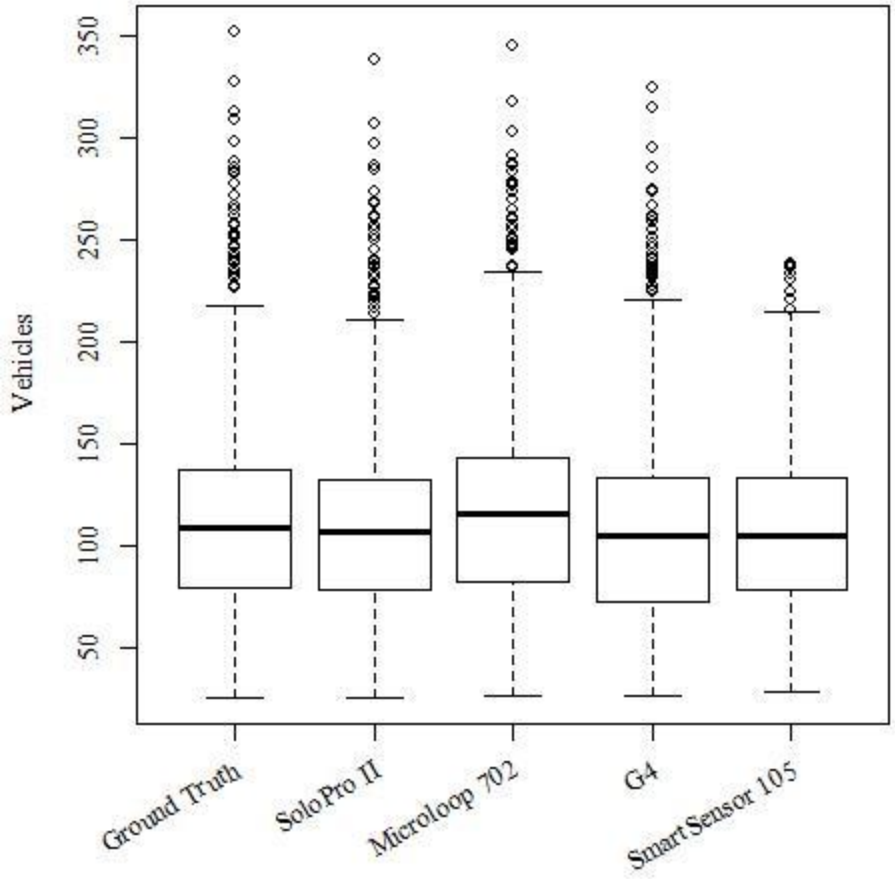

Figure D.2: Box Plot of Reported Five-Minute Volumes

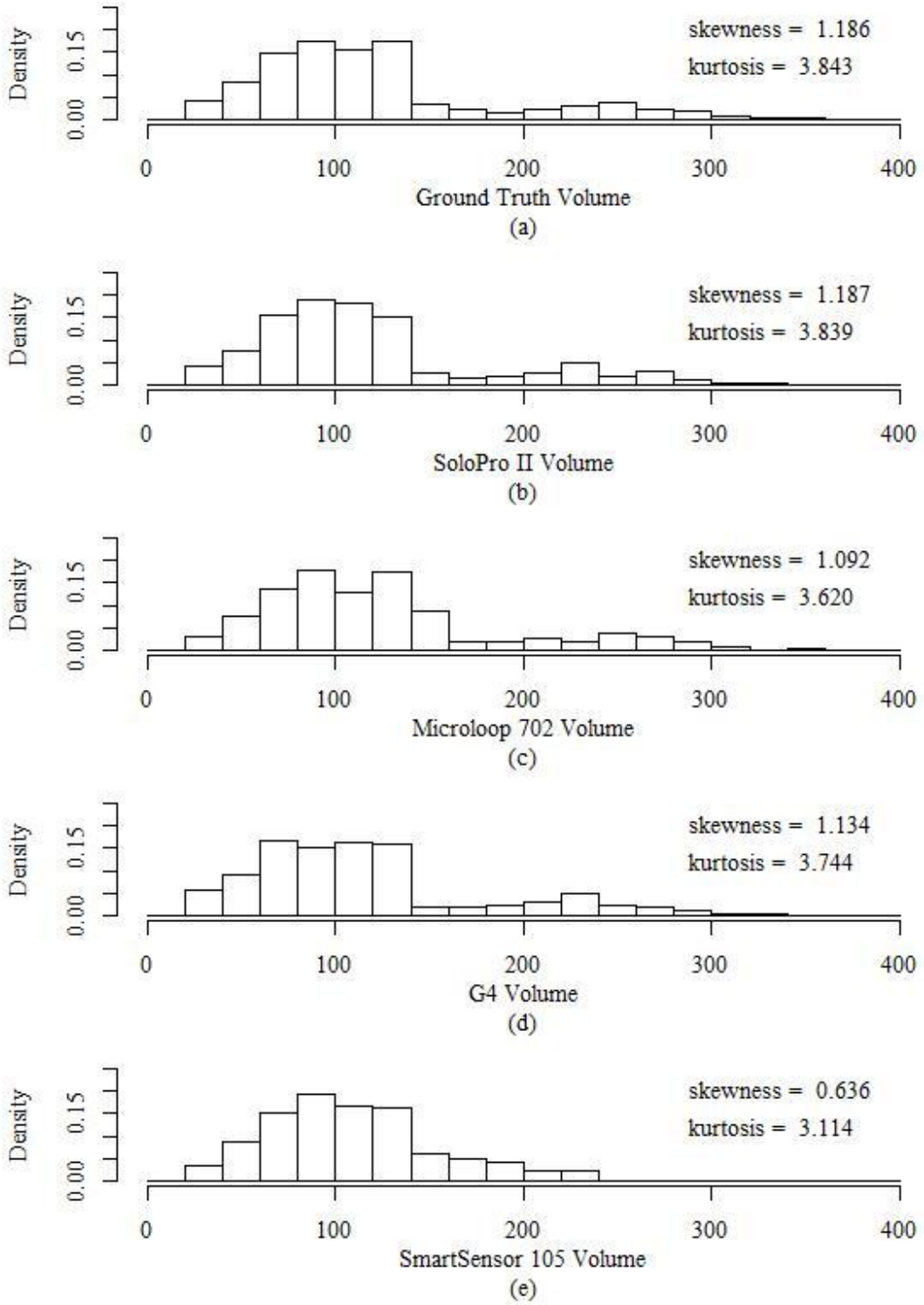

Figure D.3: Histograms of Five-Minute Volume Distributions for Ground Truth (a), Solo Pro II (b), Microloop 702 (c), G4 (d), and SmartSensor 105 (e)

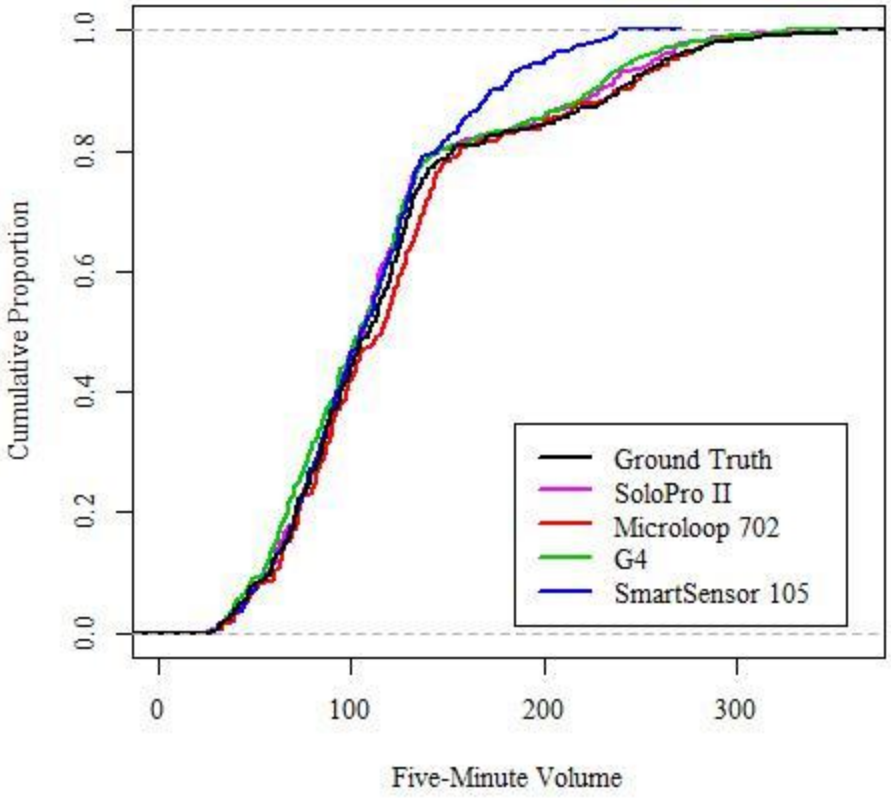

Figure D.4: Cumulative Distribution Plot of Five-Minute Volume Distributions for Ground Truth and All Detectors

Table D.1 Five-Minute Volume Summary Statistics

	Mean	Median	Standard Deviation
Ground Truth	123	109	66.1
Solo Pro II	119	107	62.6
Microloop 702	126	116	65.1
G4	117	105	62.4
SmartSensor 105	110	105	45.3

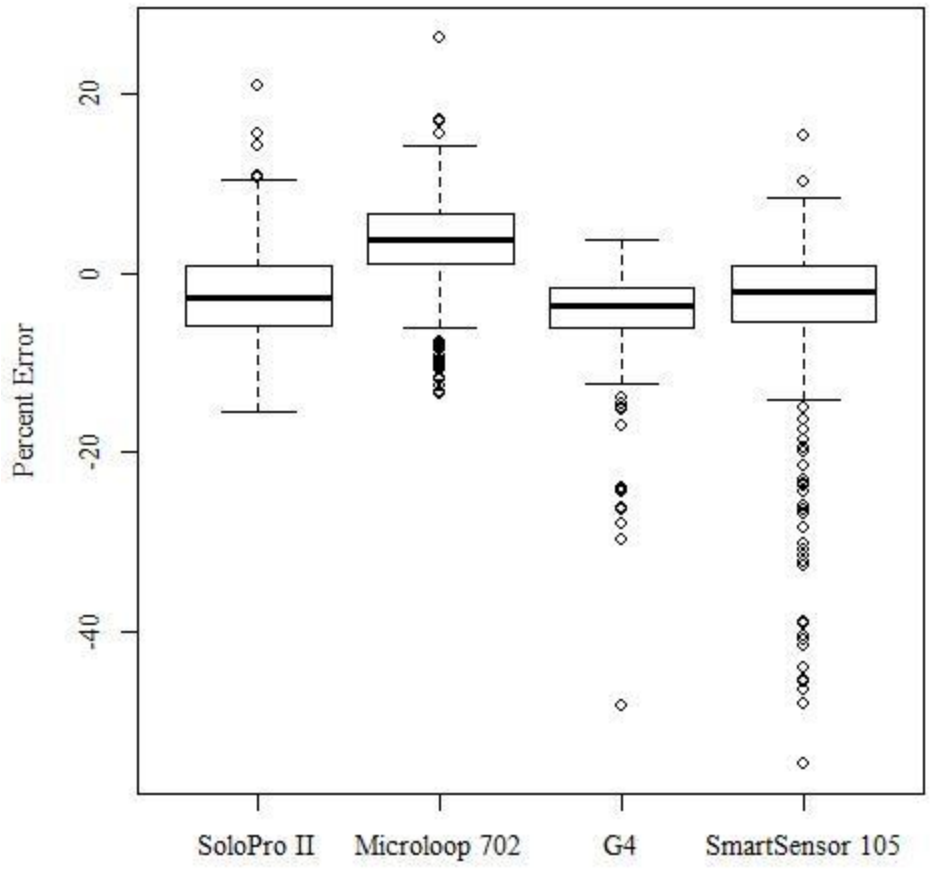

Figure D.5: Five-Minute Volume Percent Error Box Plot

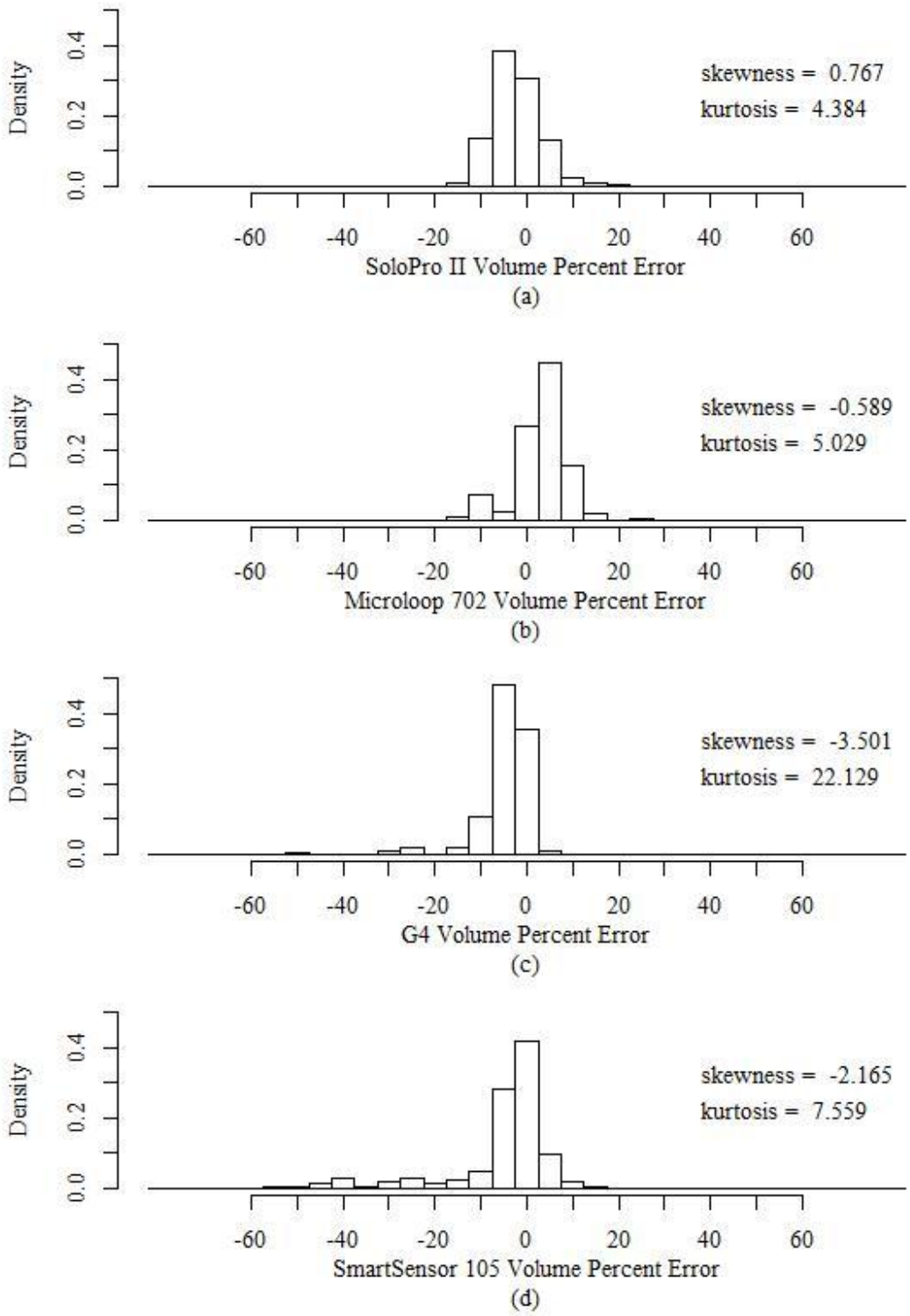

Figure D.6: Histograms of Five-Minute Volume Percent Error Distributions for Solo Pro II (a), Microloop (b), G4 (c), and SmartSensor 105 (d) Detectors

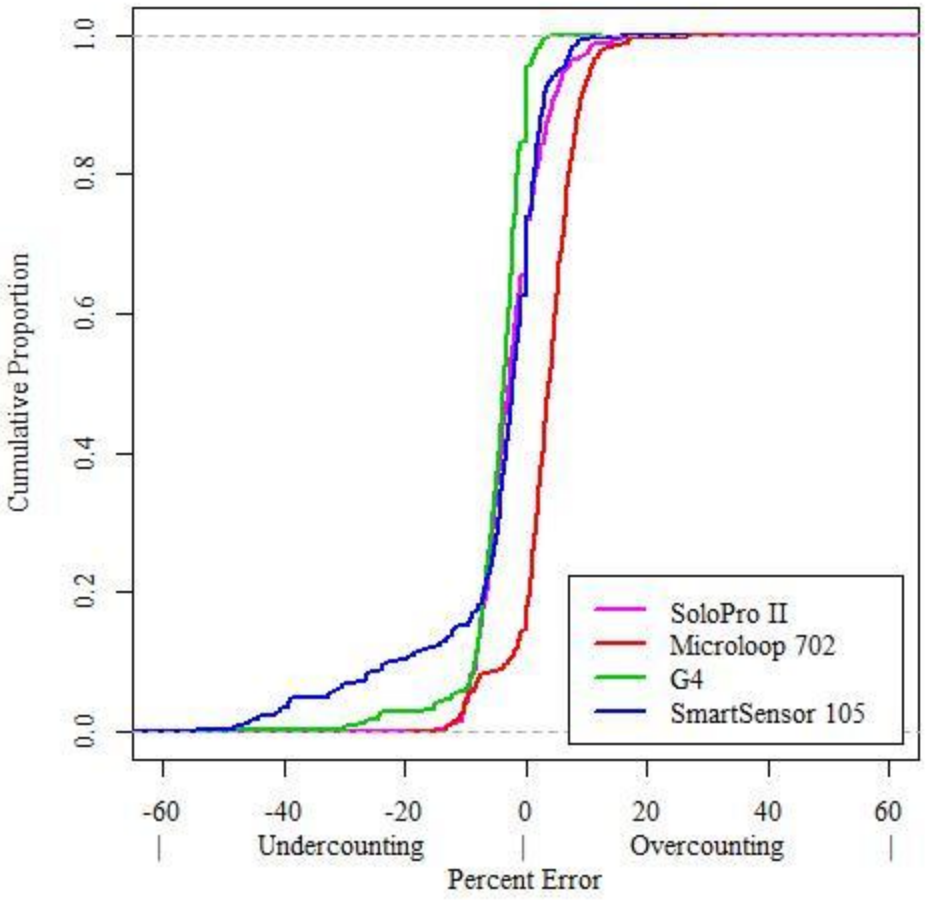

Figure D.7: Five-Minute Volume Percent Error Cumulative Distribution Plot

Table D.2: Detector Five-Minute Volume Error Statistics

	Correlation Coefficient	MPE	MAPE	Percent Error Variance	Mean GEH	85th Percentile GEH	GEH Variance
SoloPro II	0.996	-2.24%	4.58%	0.00270	0.495	0.885	0.139
Microloop 702	0.994	3.35%	5.28%	0.00306	0.532	0.897	0.139
G4	0.997	-4.58%	4.75%	0.00295	0.531	0.921	0.311
SmartSensor 105	0.925	-5.24%	6.96%	0.0132	1.02	1.60	3.77

Table D.3: Five-Minute Volume Theil's Inequality Coefficients

	U	U_m	U_s	U_c
SoloPro II	0.028	0.234	0.210	0.559
Microloop 702	0.027	0.153	0.019	0.831
G4	0.032	0.469	0.187	0.346
SmartSensor 105	0.124	0.152	0.419	0.431

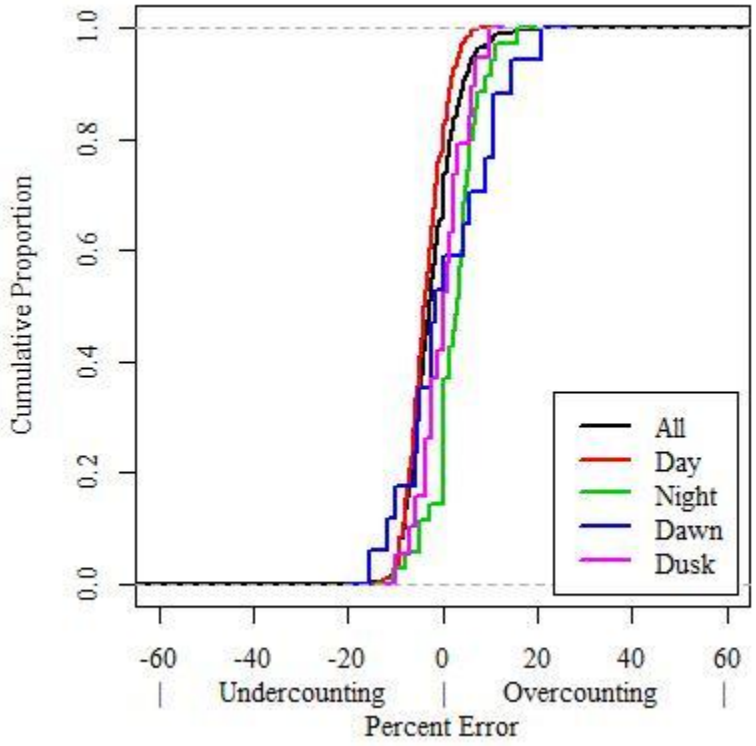

Figure D.8: Solo Pro II Five-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

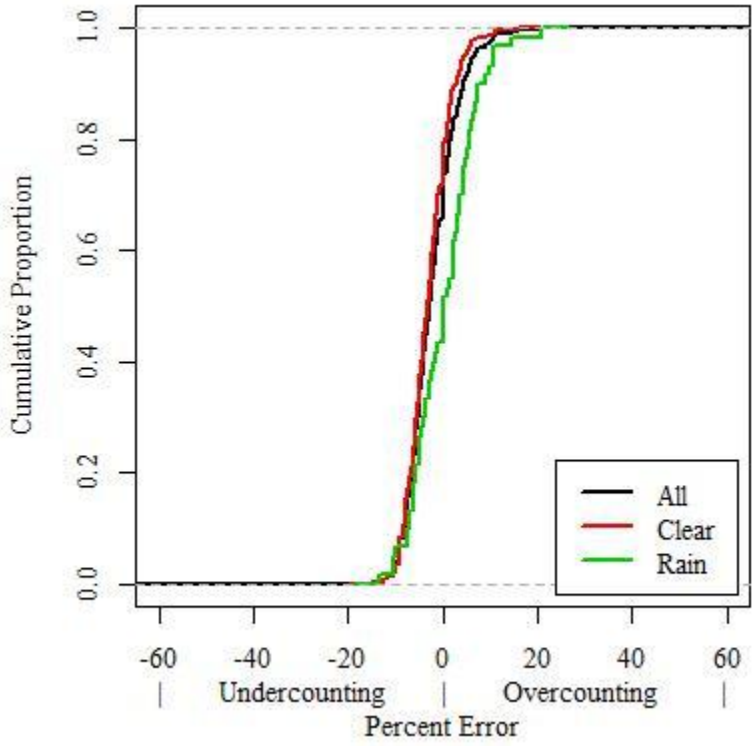

Figure D.9: Solo Pro II Five-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

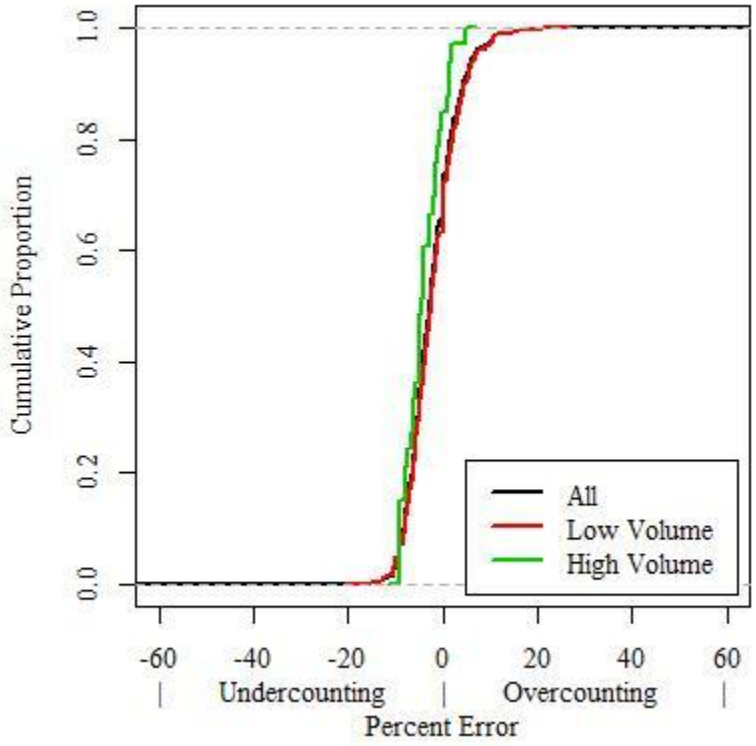

Figure D.10: Solo Pro II Five-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

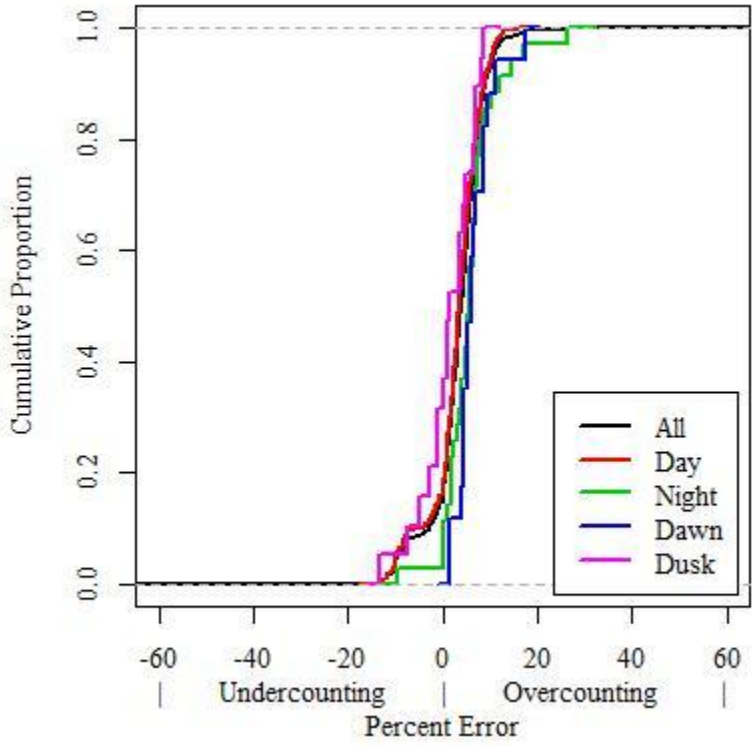

Figure D.11: Microloop 702 Five-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

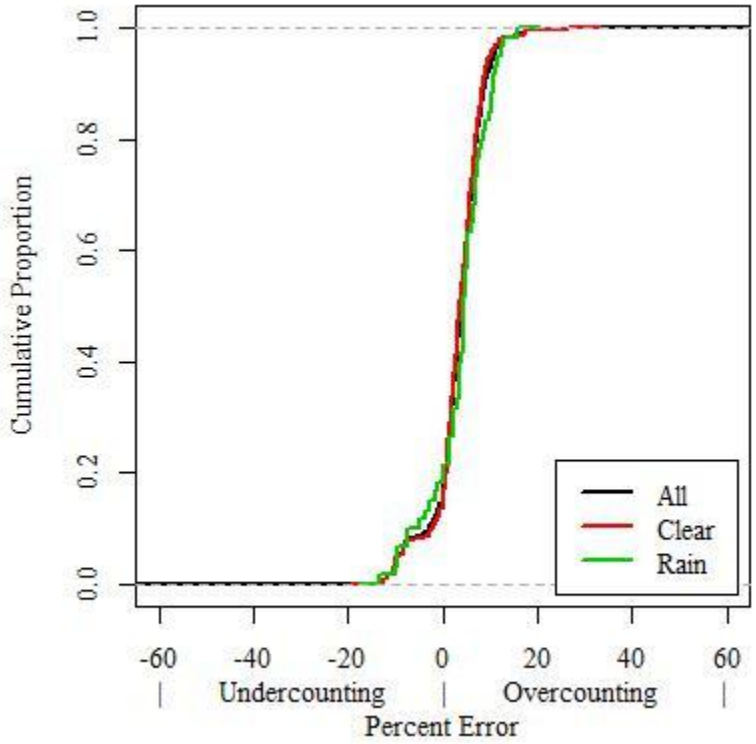

Figure D.12: Microloop 702 Five-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

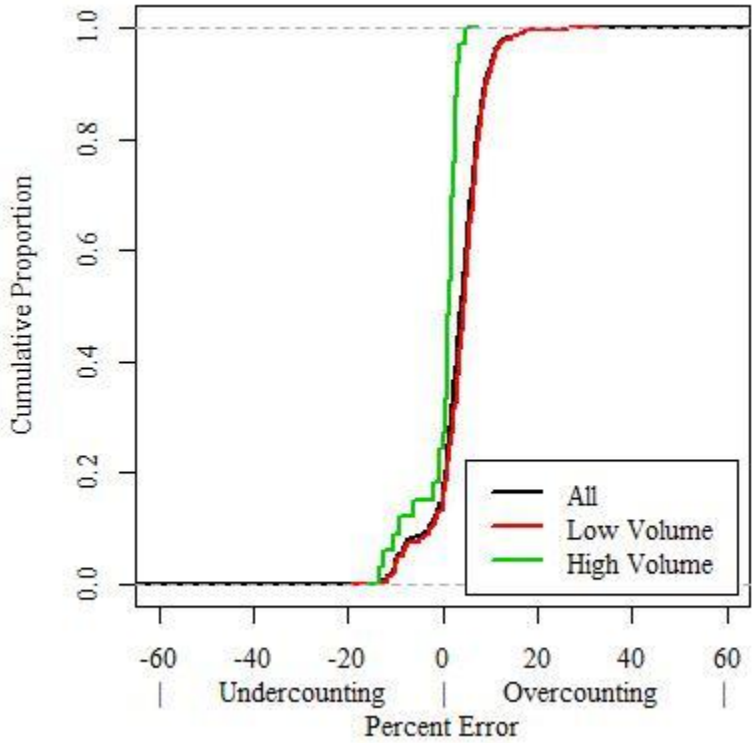

Figure D.13: Microloop 702 Five-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

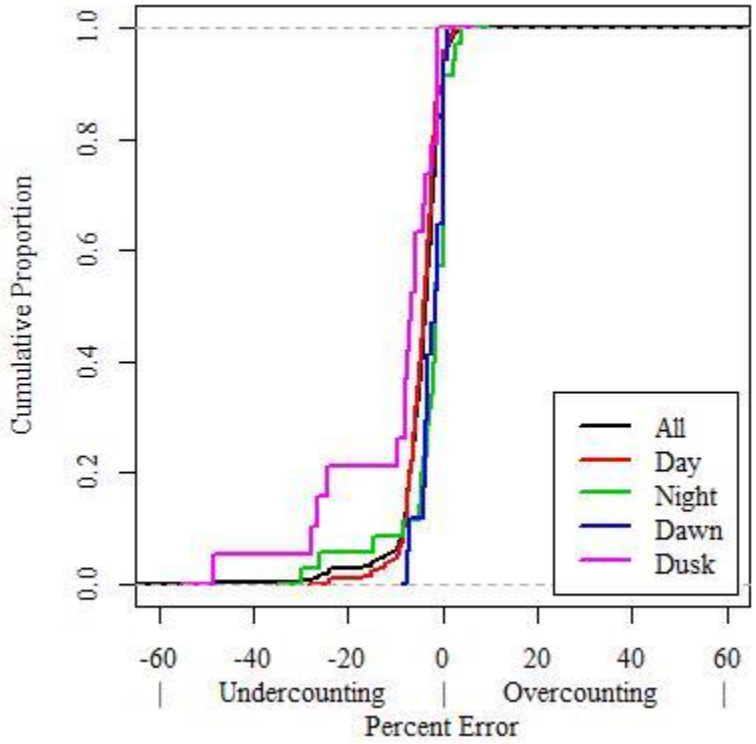

Figure D.14: G4 Five-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

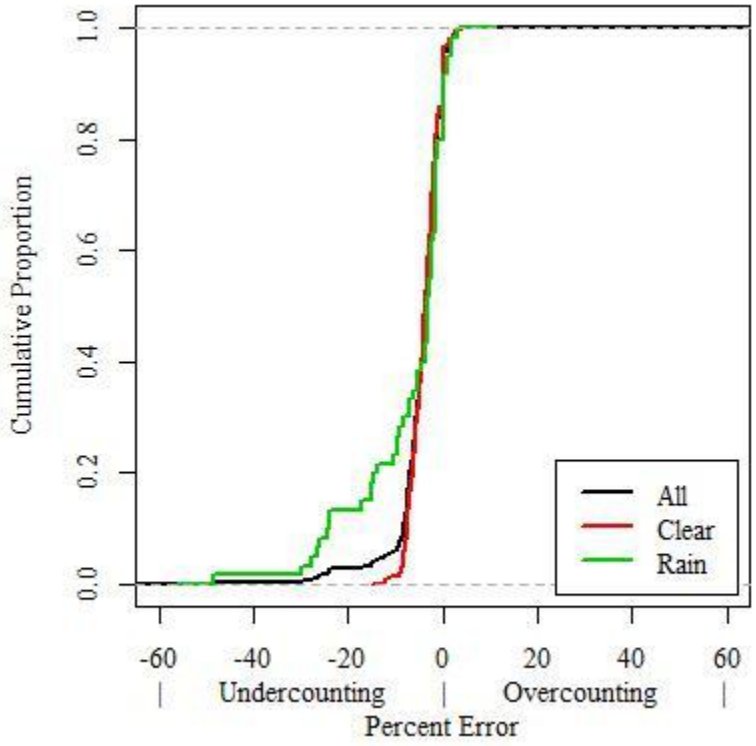

Figure D.15: G4 Five-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

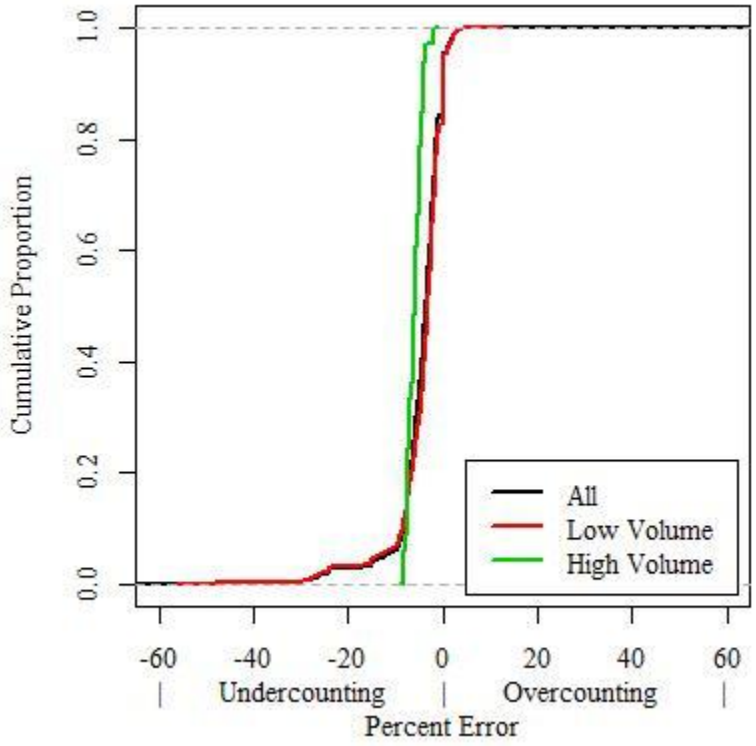

Figure D.16: G4 Five-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

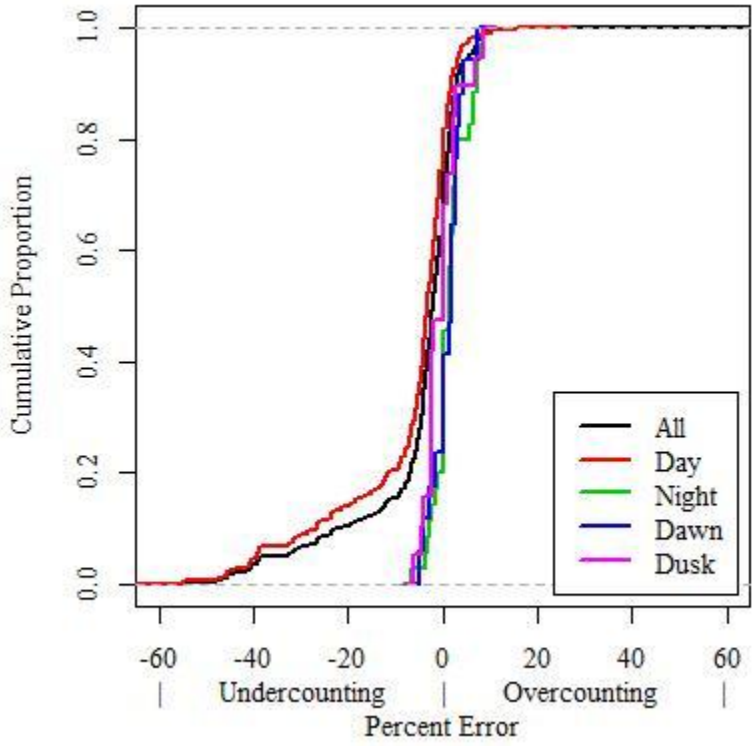

Figure D.17: SmartSensor 105 Five-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

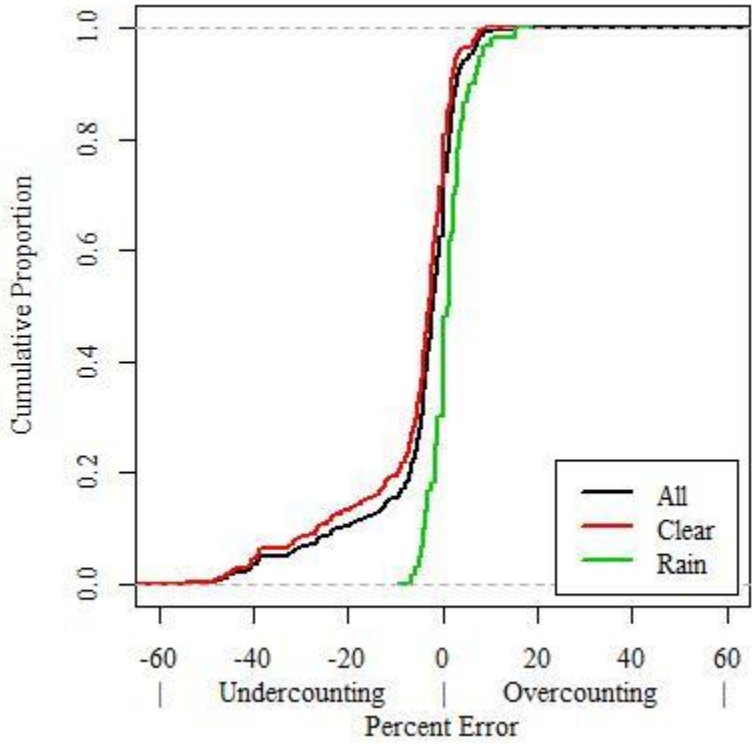

Figure D.18: SmartSensor 105 Five-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

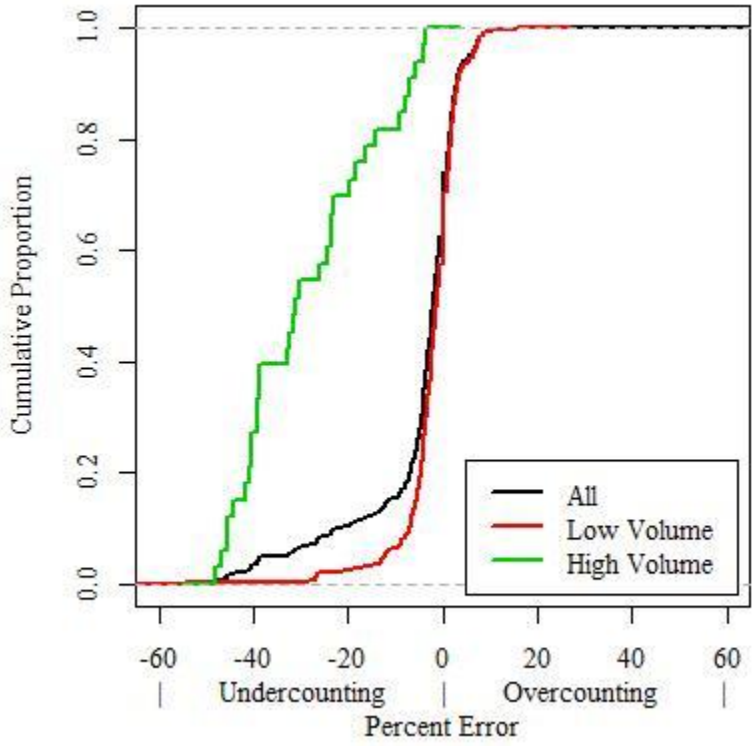

Figure D.19: SmartSensor 105 Five-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

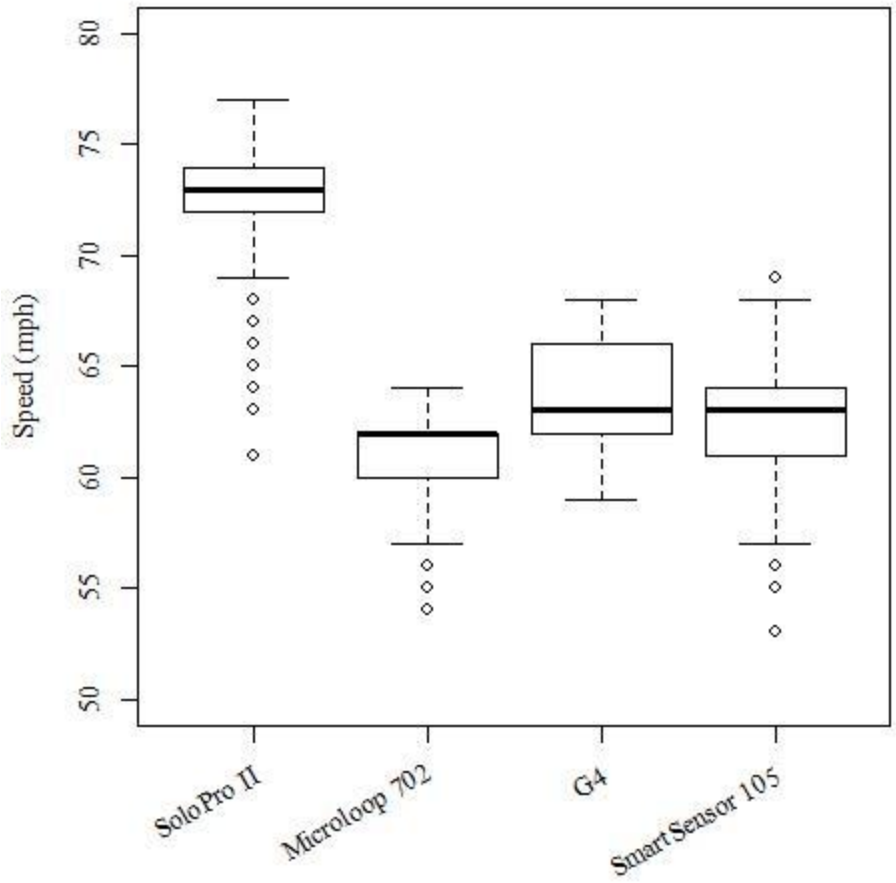

Figure D.20: Box Plot of Reported Five-Minute Mean Speeds

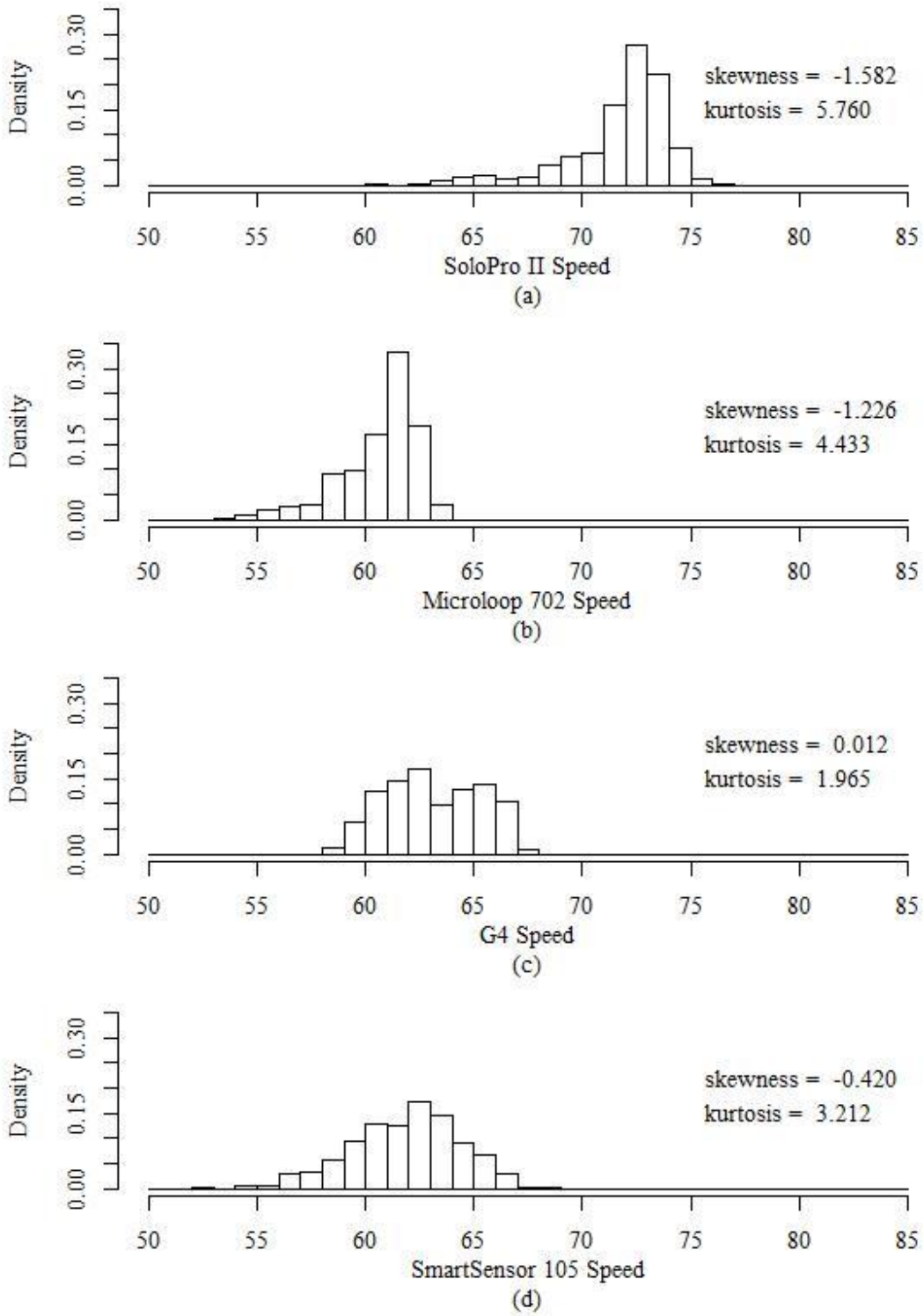

Figure D.21: Histograms of Five-Minute Mean Speed Distributions for the Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

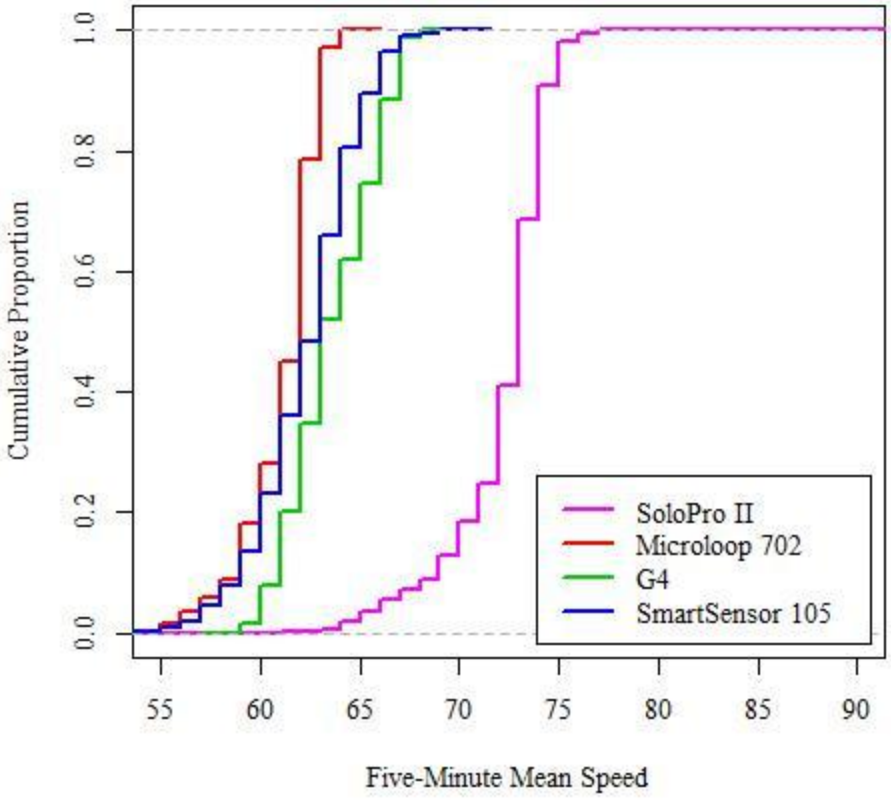

Figure D.22: Cumulative Distribution Plot of Five-Minute Mean Speed Distributions for All Detectors

Table D.4 Five-Minute Mean Speed Summary Statistics

	Mean	Median	Standard Deviation
Solo Pro II	72	73	2.54
Microloop 702	61	62	1.88
G4	64	63	2.21
SmartSensor 105	62	63	2.60

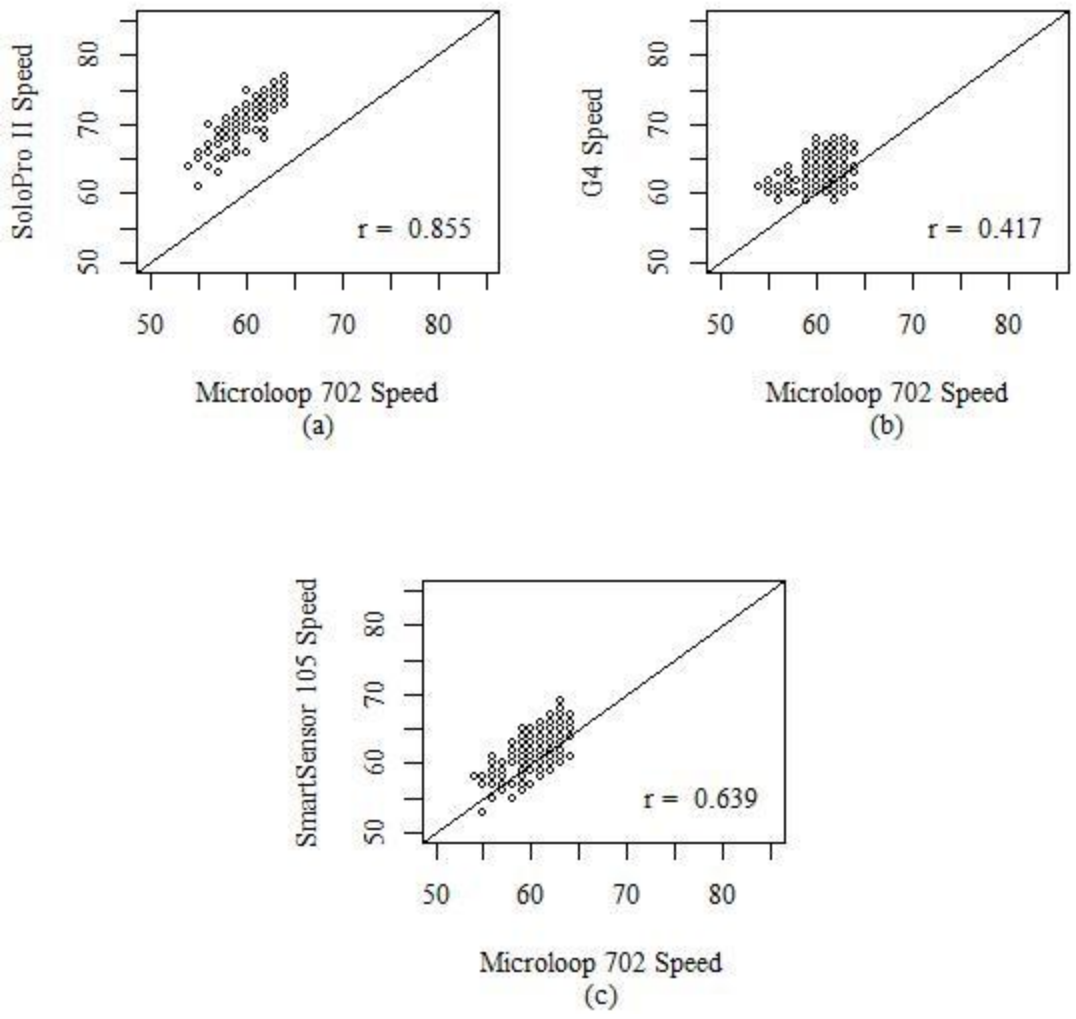

Figure D.23: Five-Minute Mean Speed Scatter Plots Against Baseline for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

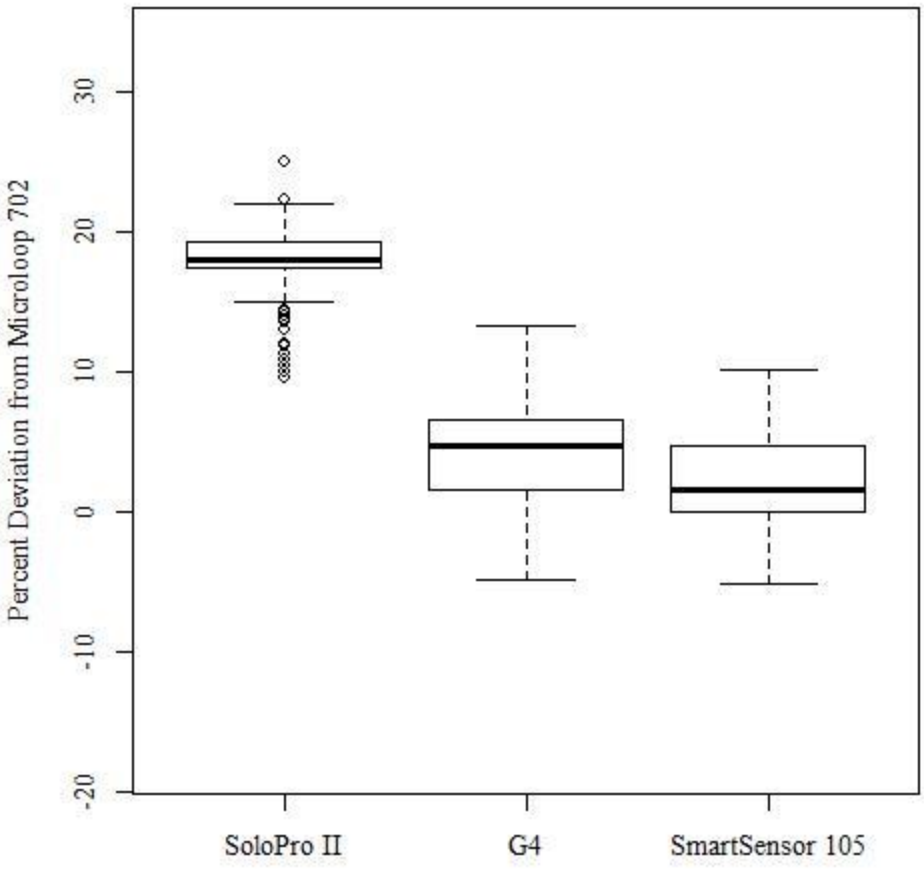

Figure D.24: Five-Minute Mean Speed Percent Deviation Box Plot

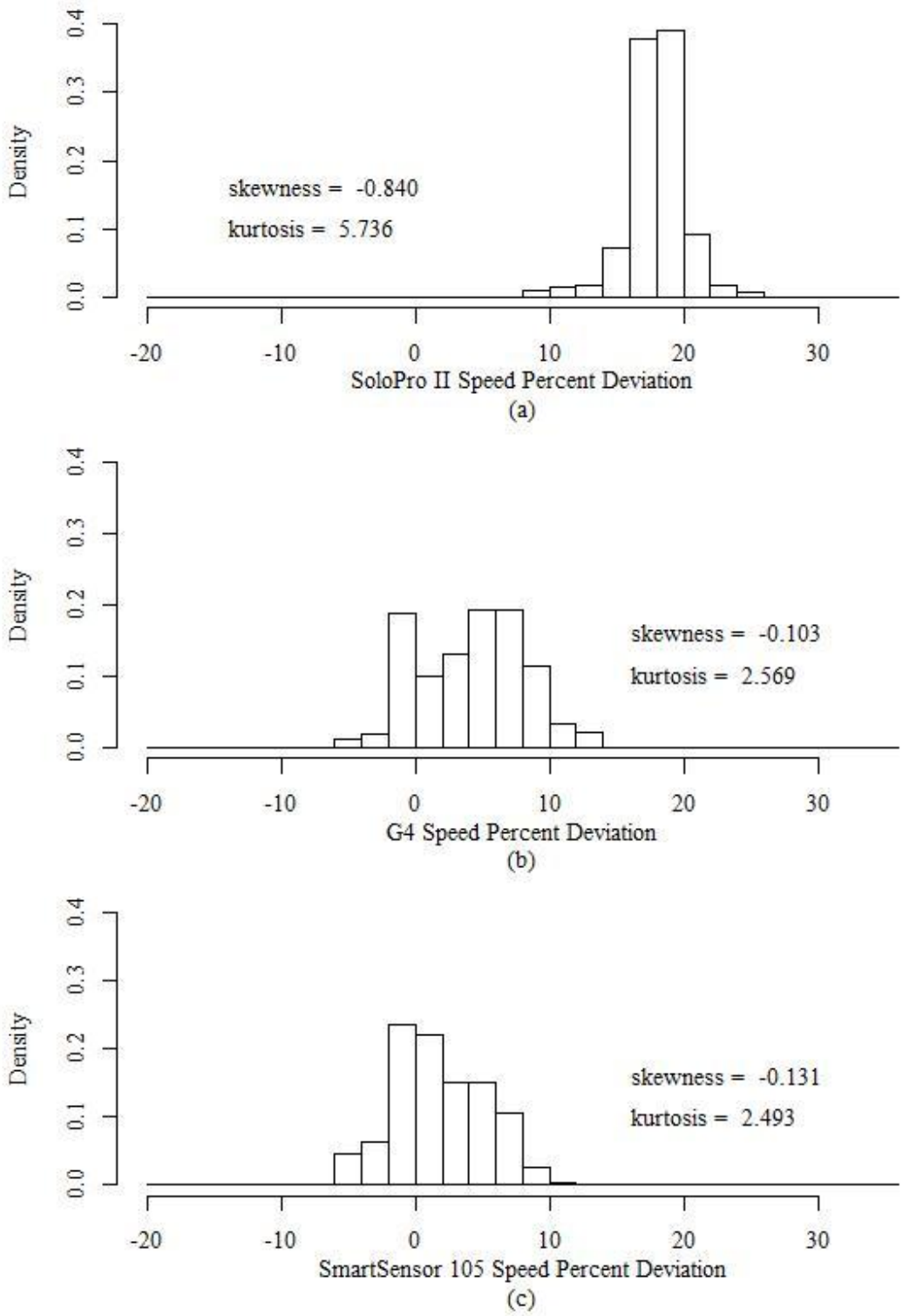

Figure D.25: Histograms of Five-Minute Mean Speed Percent Deviation Distributions for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

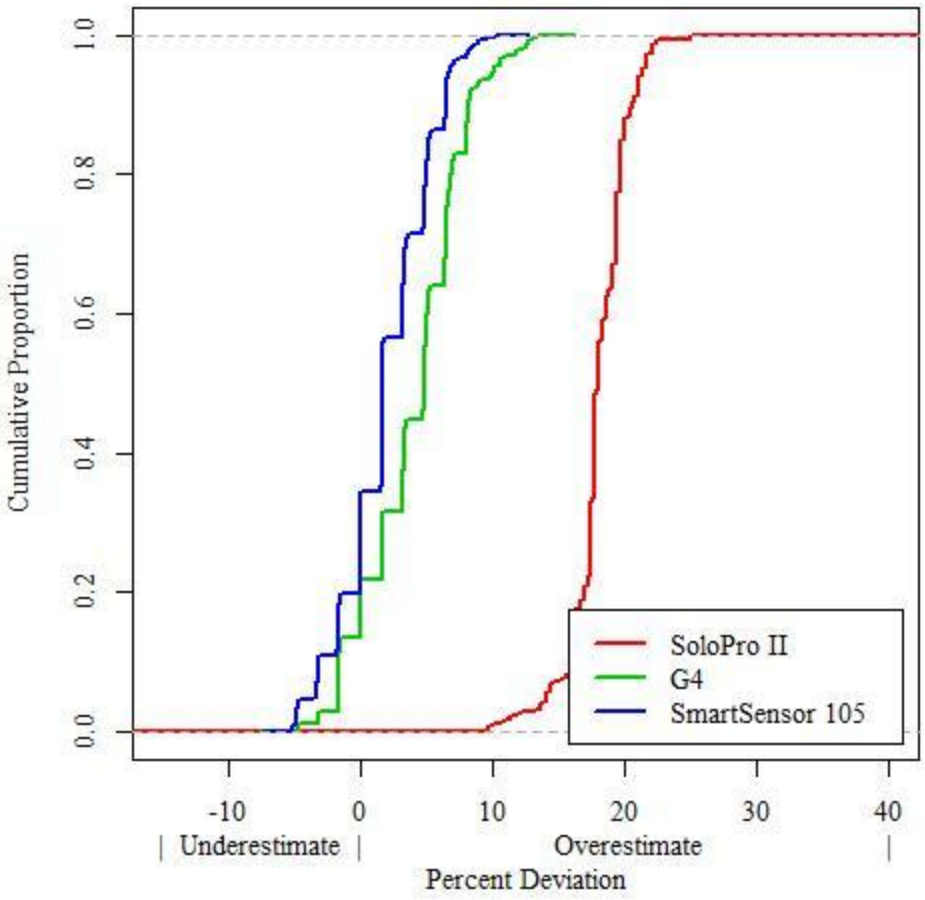

Figure D.26: Five-Minute Mean Speed Percent Deviation Cumulative Distribution Plot

Table D.5: Detector Five-Minute Mean Speed Deviation Statistics

	MPD	MAPD	Percent Deviation Variance
SoloPro II	18.07%	18.07%	0.00049
G4	4.10%	4.66%	0.00139
SmartSensor 105	1.96%	3.13%	0.00110

Table D.6: Five-Minute Mean Speed Theil's Inequality Coefficients

	U	U_m	U_s	U_c
SoloPro II	0.083	0.985	0.004	0.011
G4	0.027	0.552	0.010	0.440
SmartSensor 105	0.019	0.261	0.094	0.648

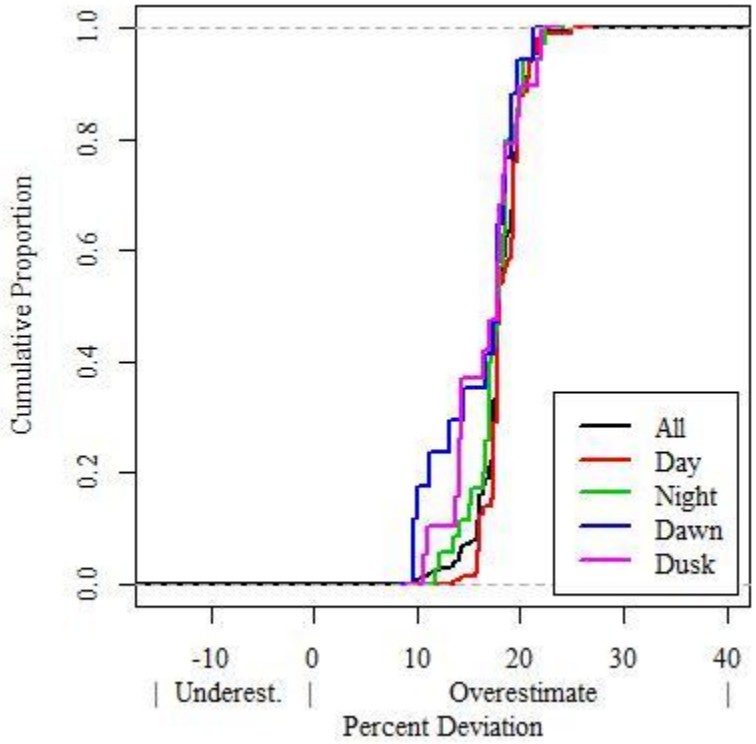

Figure D.27: Solo Pro II Five-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

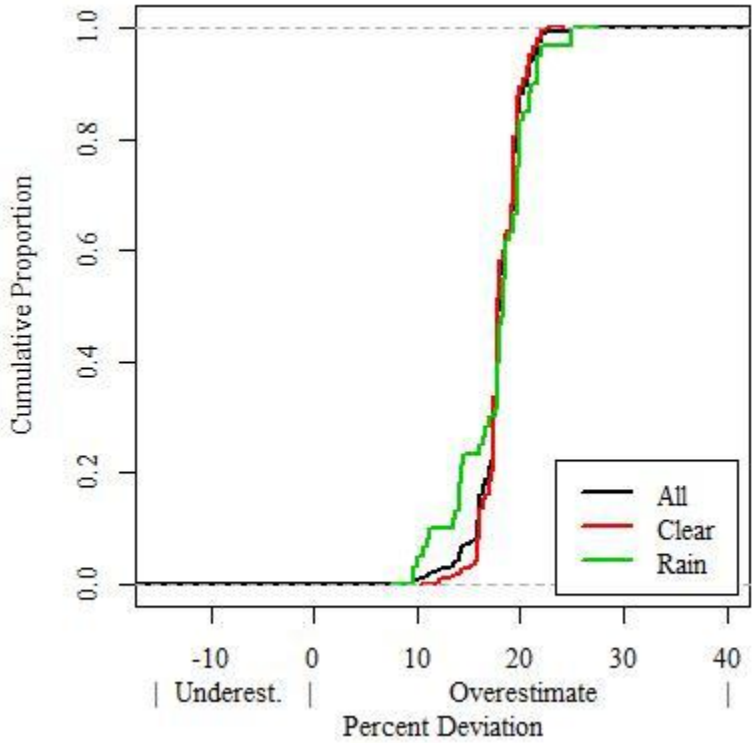

Figure D.28: Solo Pro II Five-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

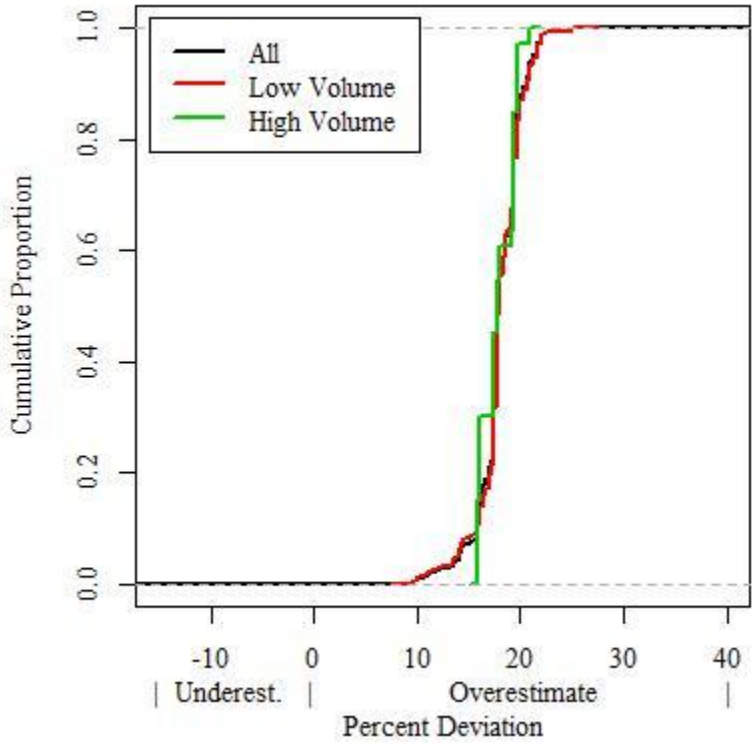

Figure D.29: Solo Pro II Five-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

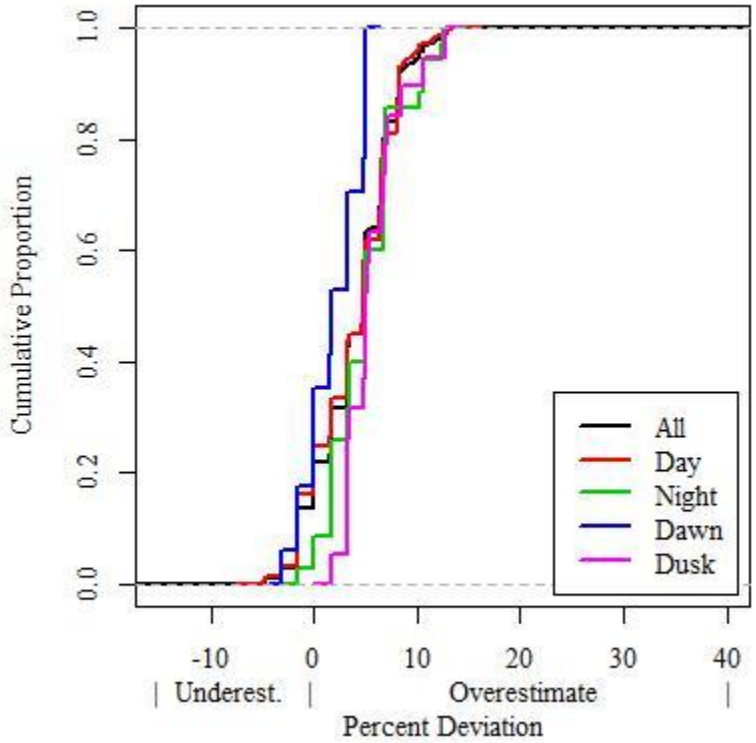

Figure D.30: G4 Five-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

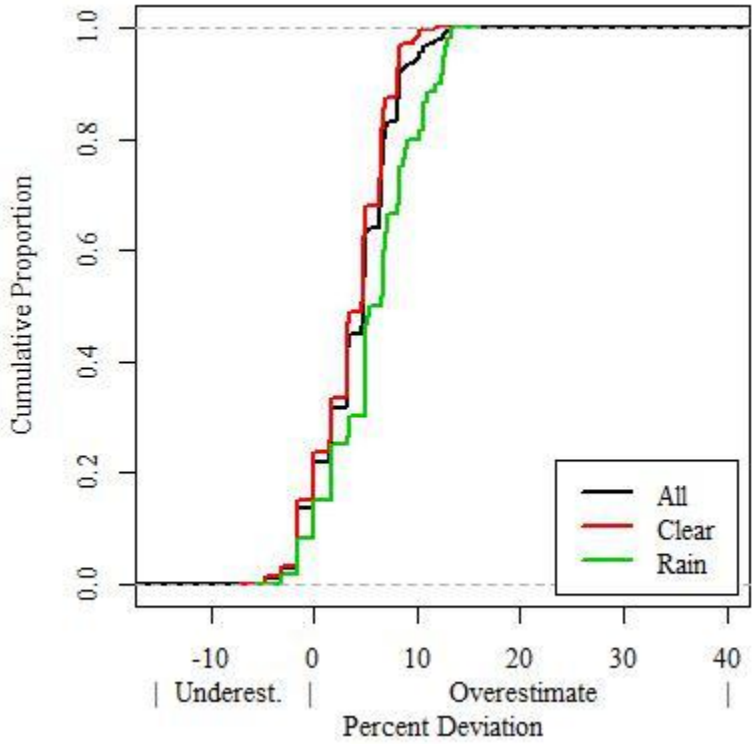

Figure D.31: G4 Five-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

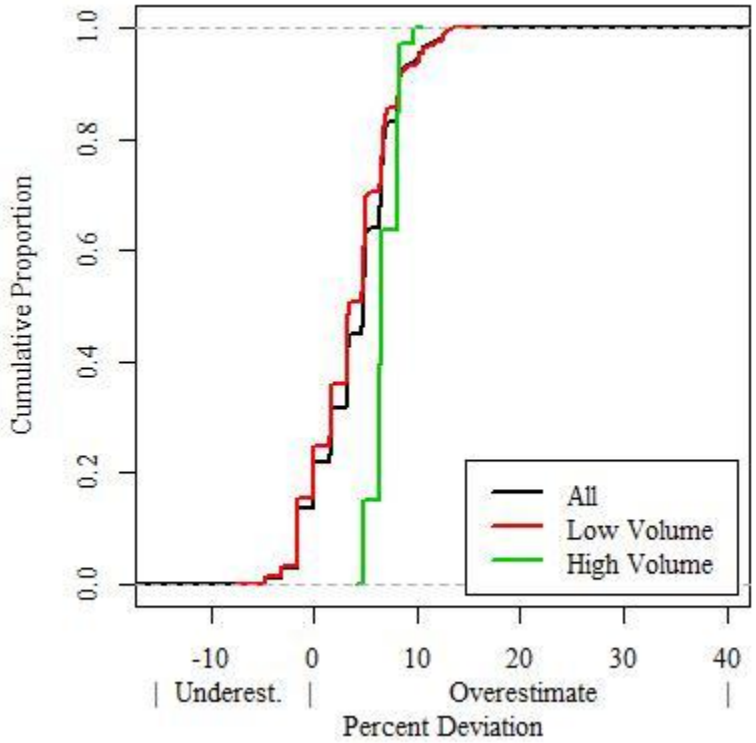

Figure D.32: G4 Five-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

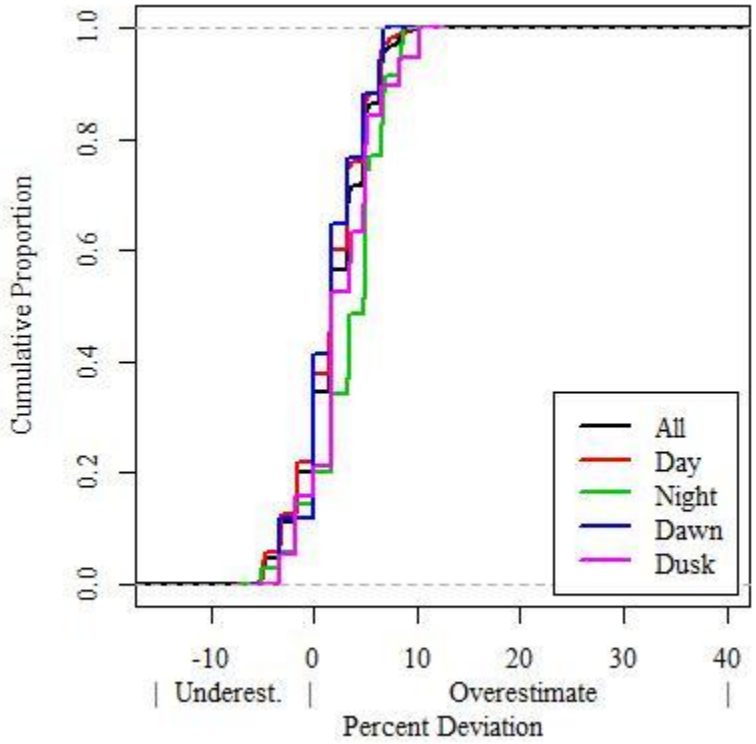

Figure D.33: SmartSensor 105 Five-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

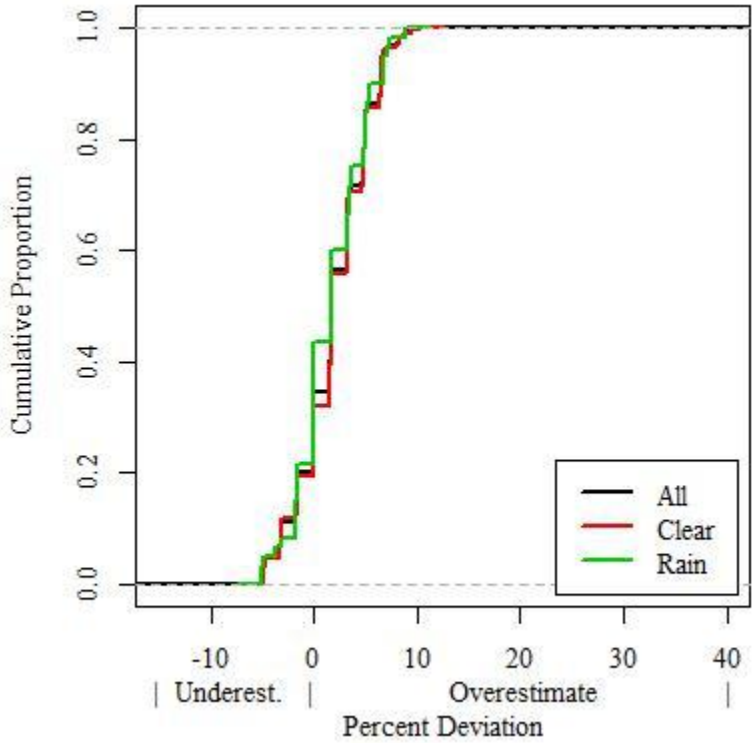

Figure D.34: SmartSensor 105 Five-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

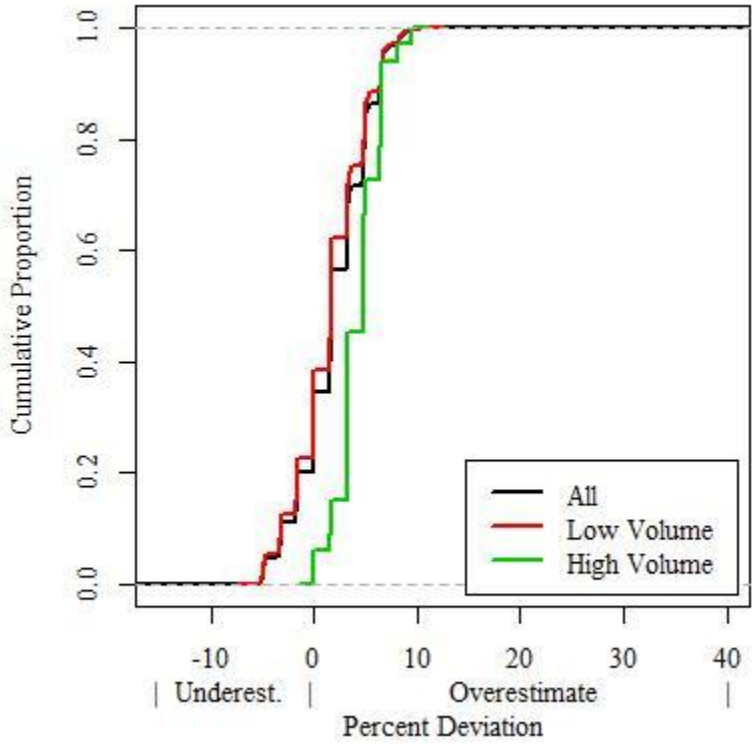

Figure D.35: SmartSensor 105 Five-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

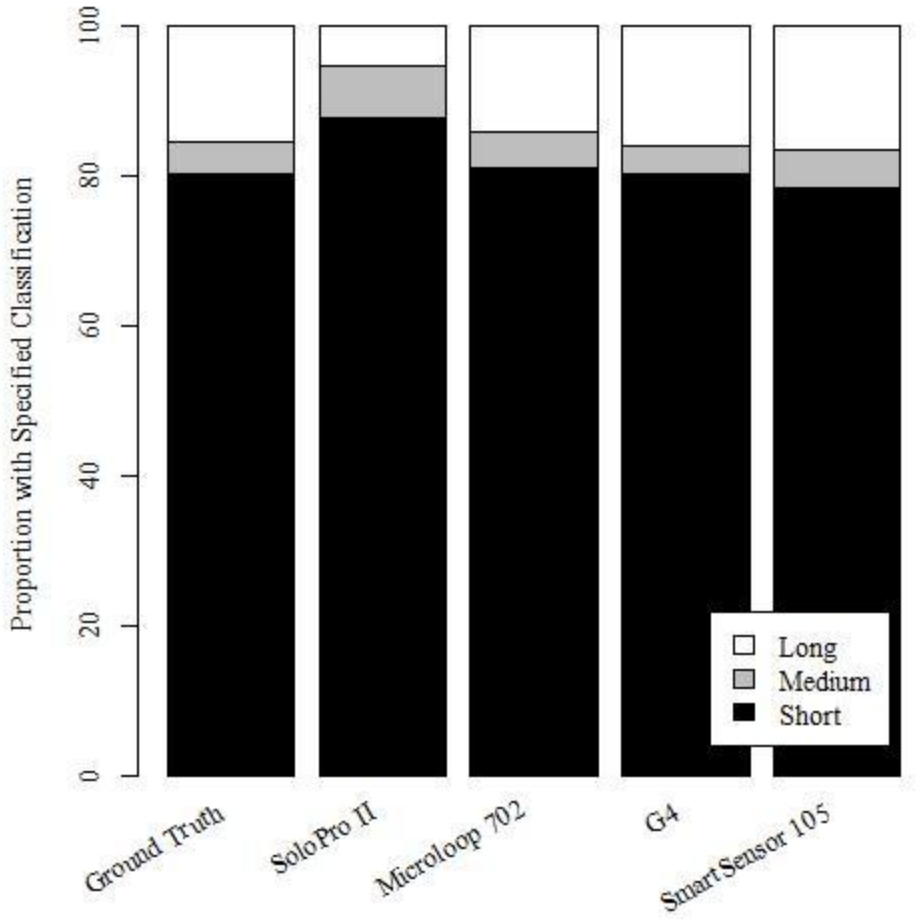

Figure D.36: Mean Five-Minute Proportion Short, Medium, and Long Vehicles Bar Chart

Table D.7: Mean Five-Minute Classification Proportions

	Ground Truth	SoloPro II	Microloop 702	G4	Smartsensor 105
Short	80.1%	87.8%	81.1%	80.3%	78.3%
Medium	4.3%	6.8%	4.8%	3.8%	5.0%
Long	15.6%	5.4%	14.1%	16.0%	16.7%

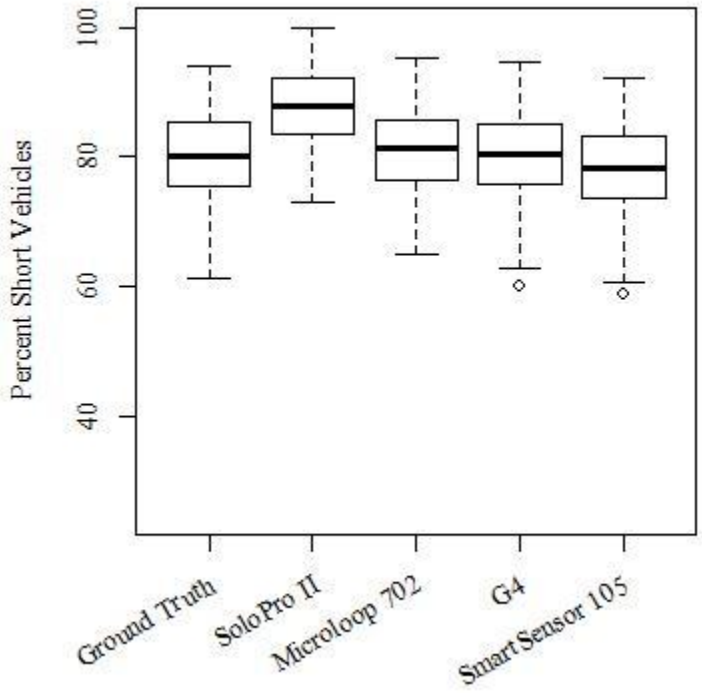

Figure D.37: Box Plot of Five-Minute Percent Short Vehicle Distributions

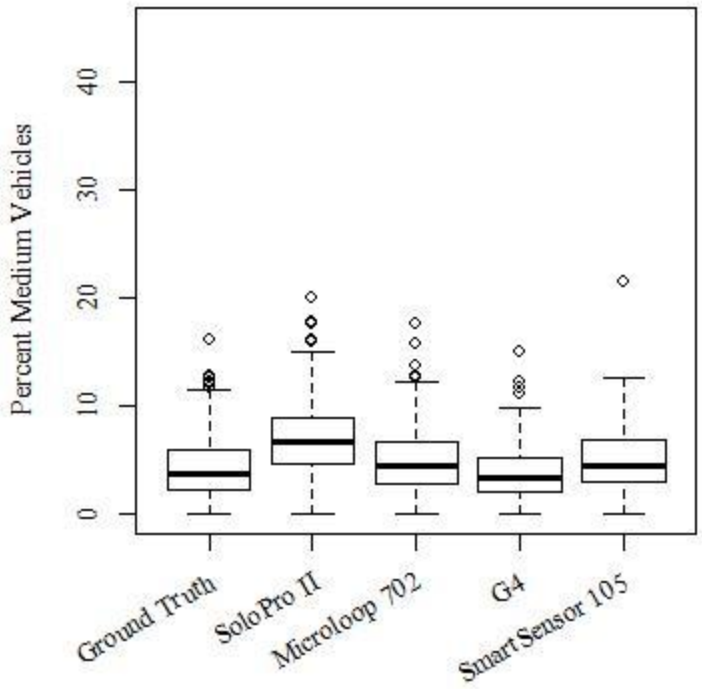

Figure D.38: Box Plot of Five-Minute Percent Medium Vehicle Distributions

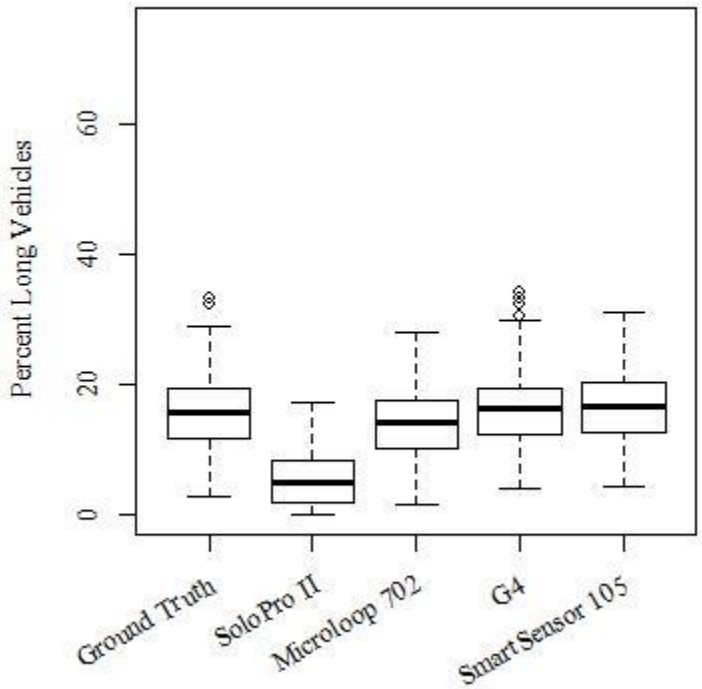

Figure D.39: Box Plot of Five-Minute Percent Long Vehicle Distributions

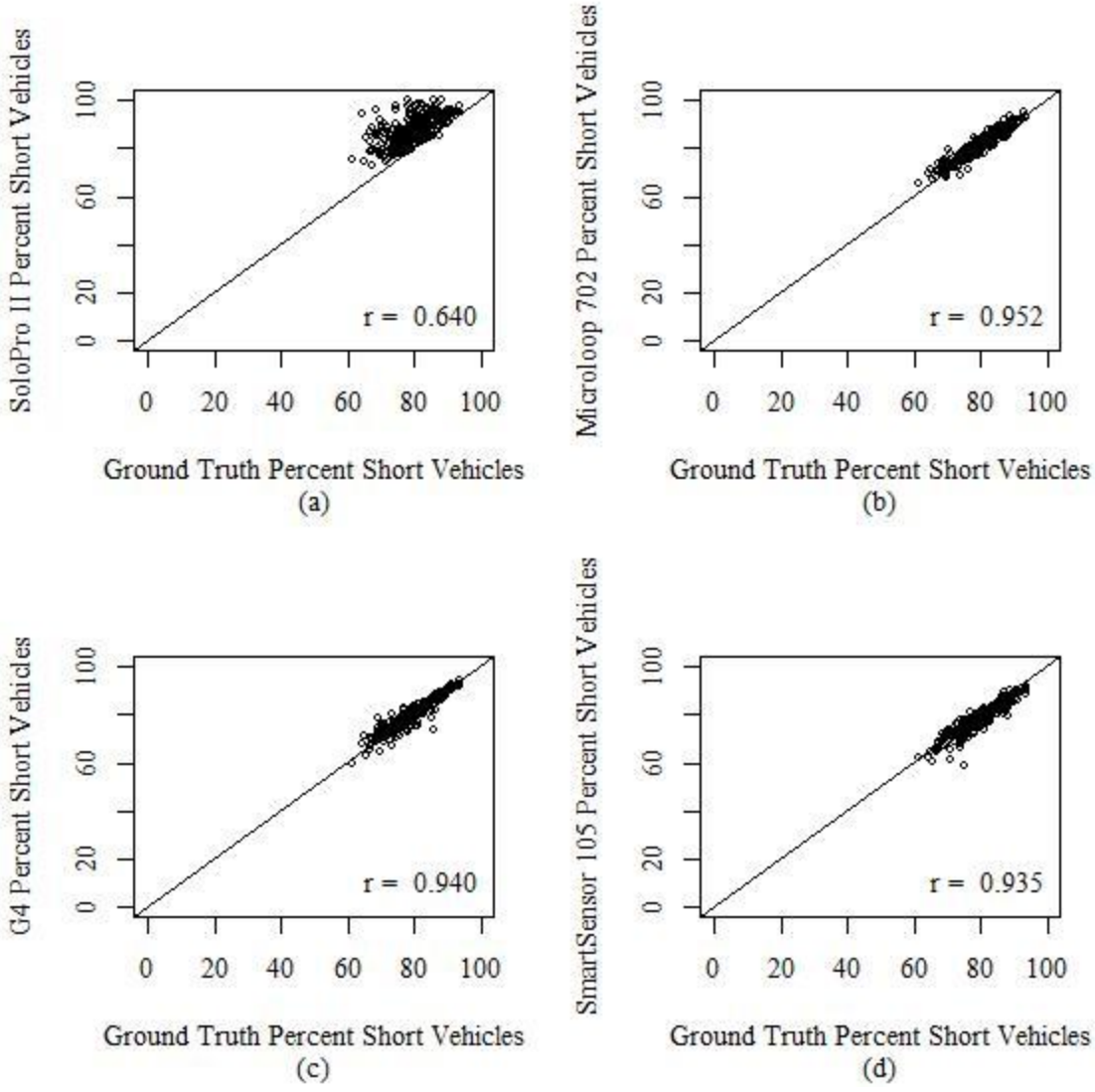

Figure D.40: Five-Minute Percent Short Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

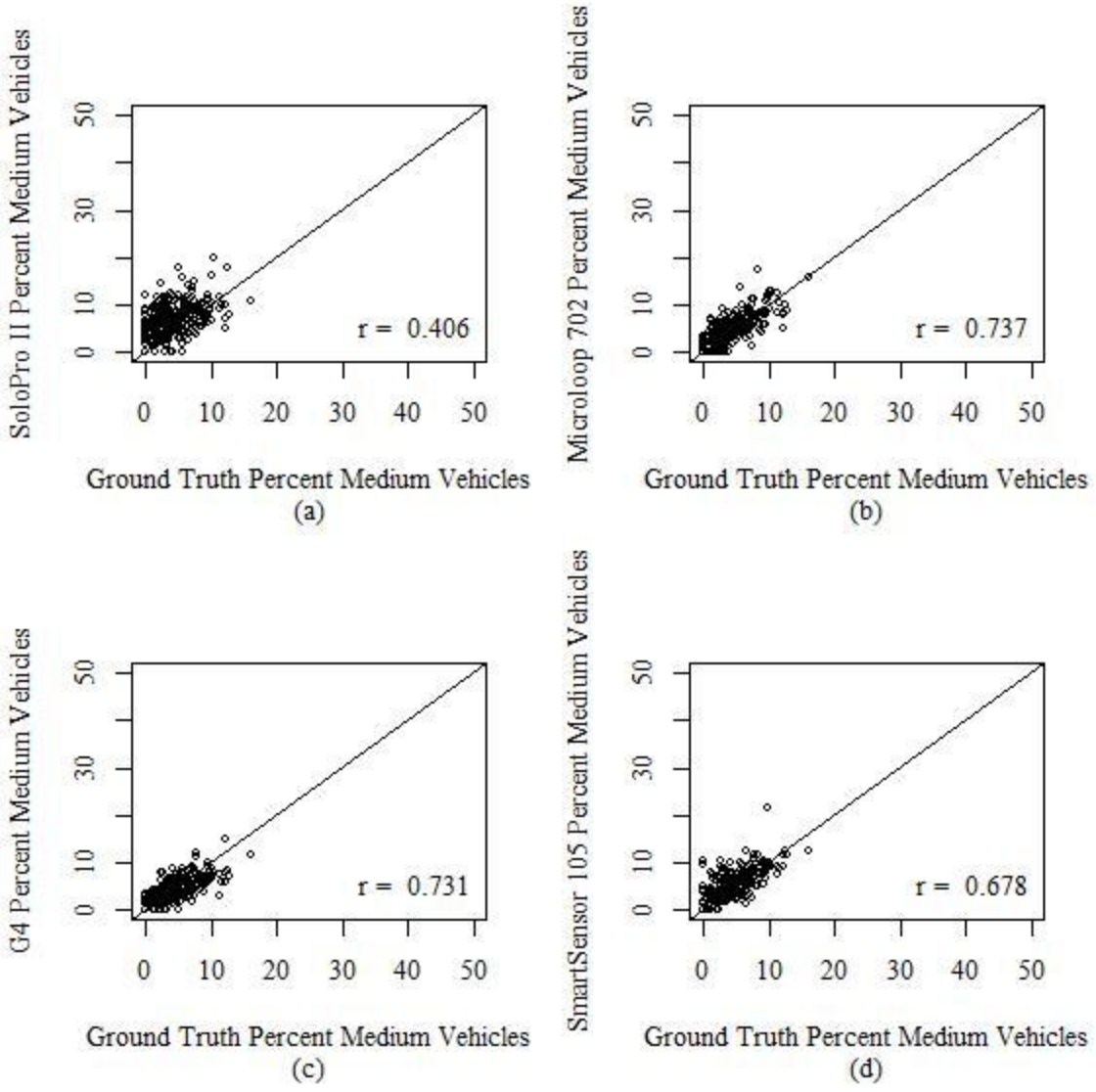

Figure D.41: Five-Minute Percent Medium Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

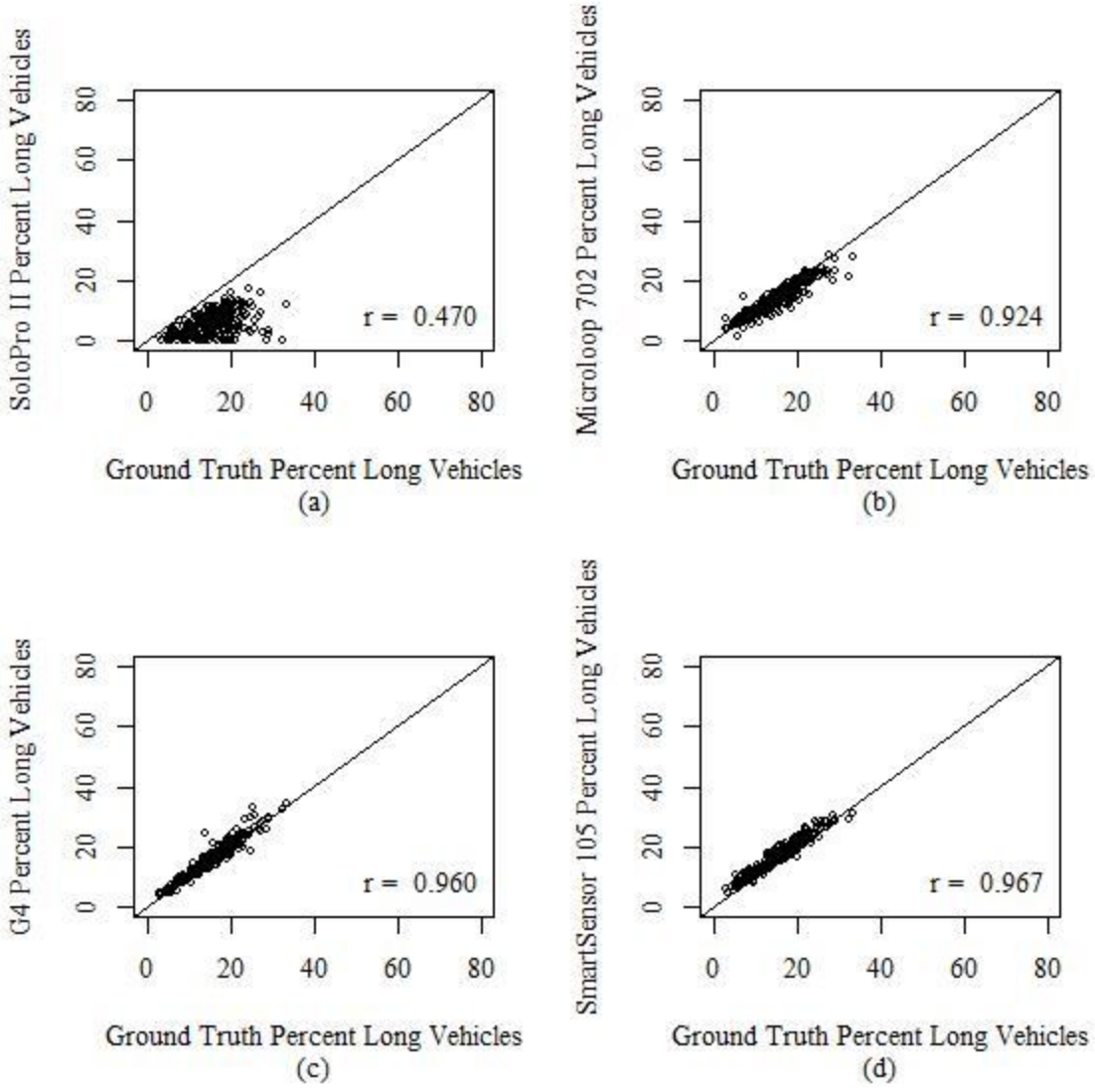

Figure D.42: Five-Minute Percent Long Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

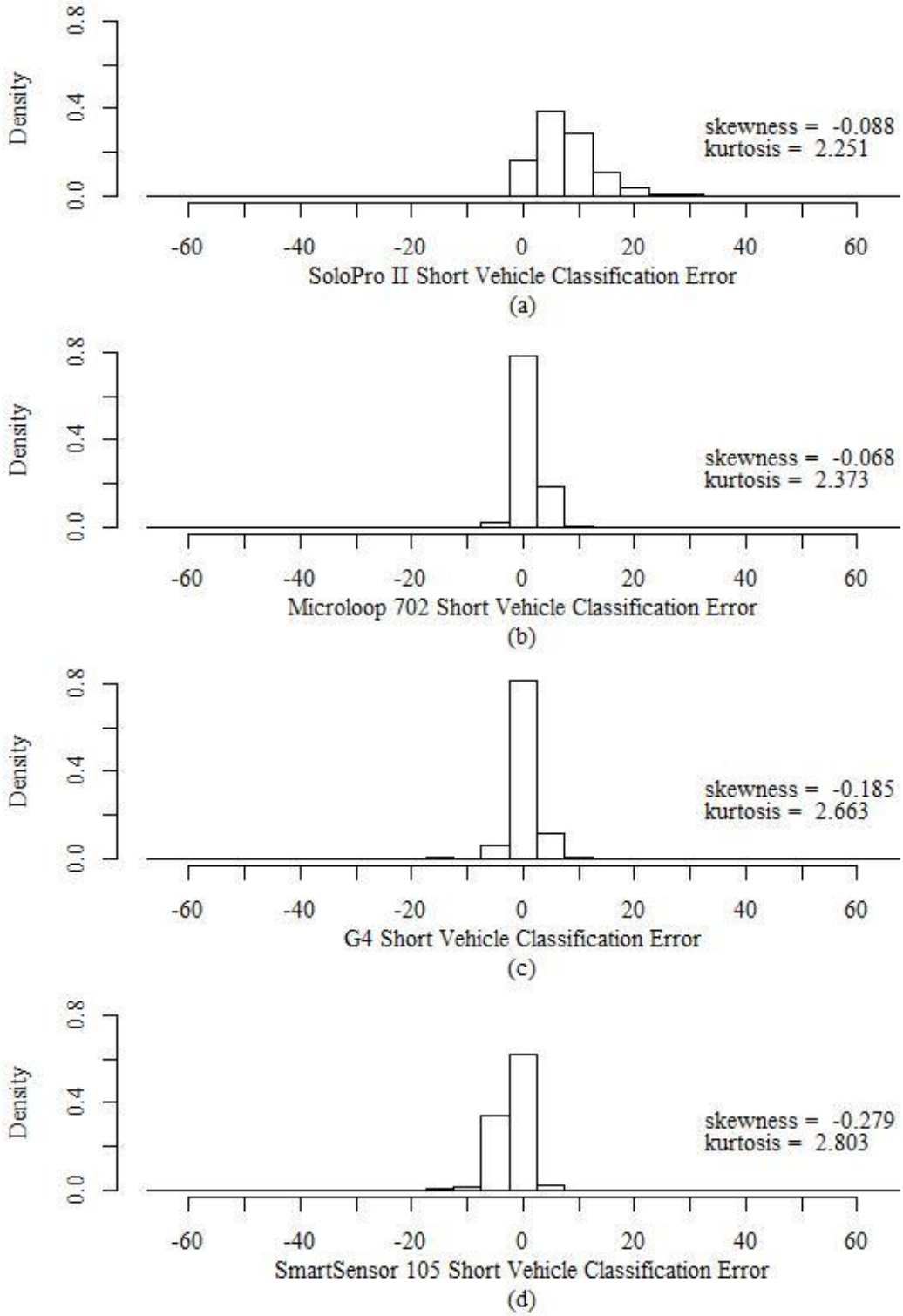

Figure D.43: Histograms of Five-Minute Percent Short Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

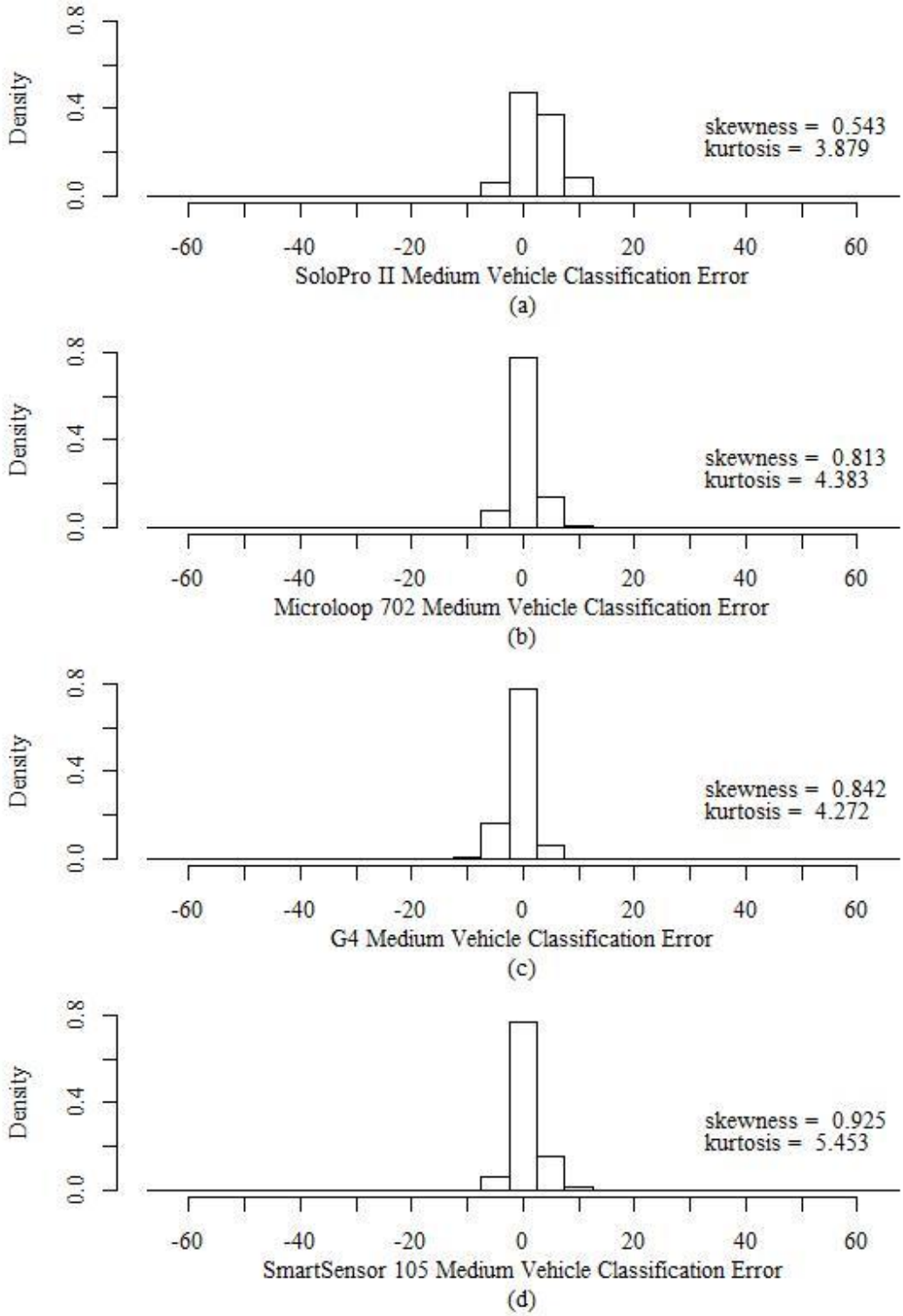

Figure D.44: Histograms of Five-Minute Percent Medium Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

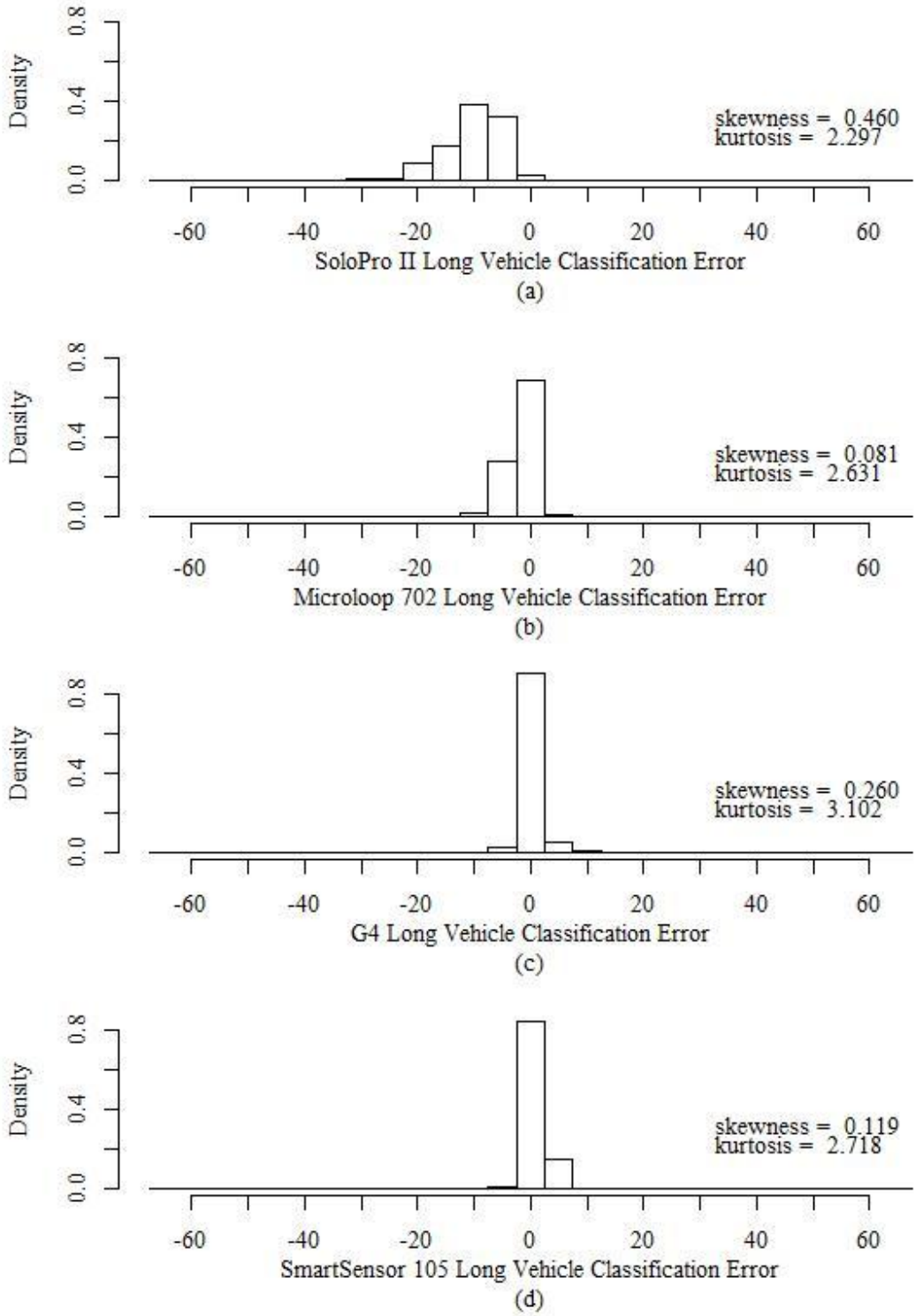

Figure D.45: Histograms of Five-Minute Percent Long Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

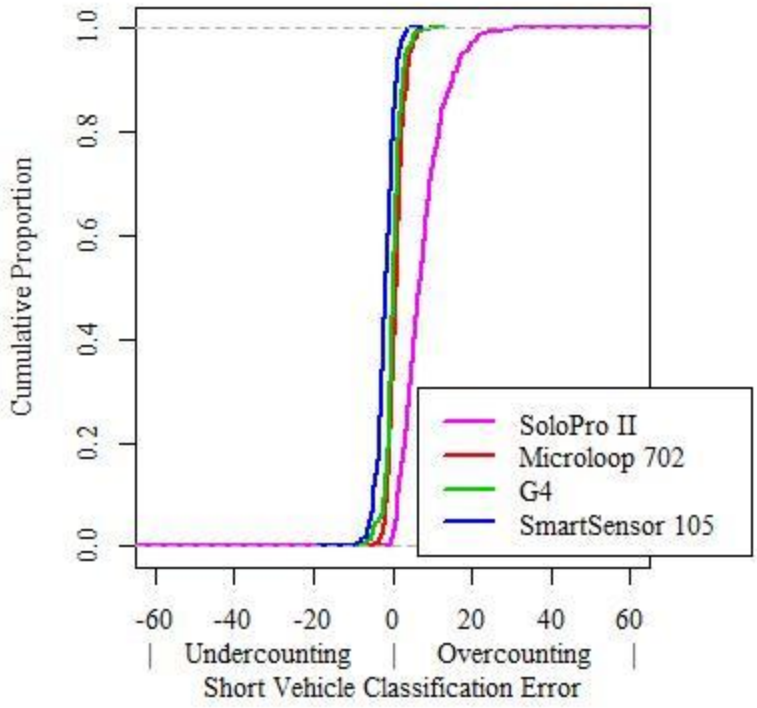

Figure D.46: Five-Minute Percent Short Vehicles Error Cumulative Distribution Plot

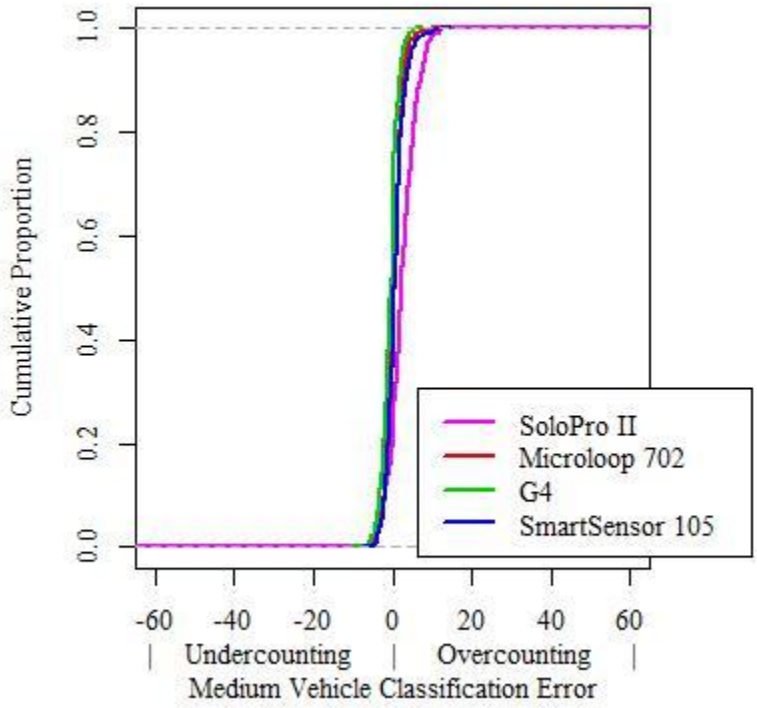

Figure D.47: Five-Minute Percent Medium Vehicles Error Cumulative Distribution Plot

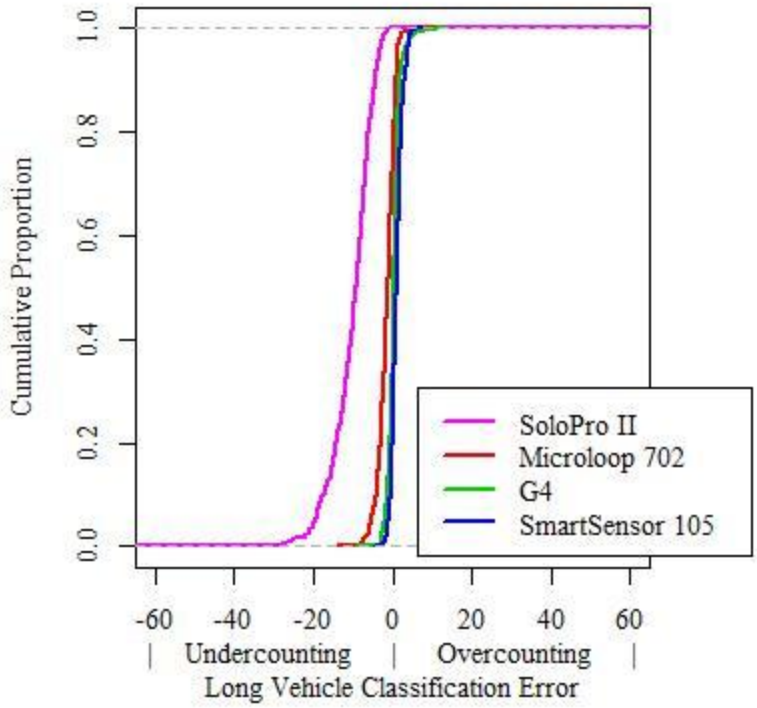

Figure D.48: Five-Minute Percent Long Vehicles Error Cumulative Distribution Plot

Table D.8 Five-Minute Classification Error Percentage Summary Statistics

	Mean	Median	Standard Deviation
Solo Pro II	10.6%	9.8%	5.22
Microloop 702	2.6%	2.2%	1.77
G4	2.1%	1.7%	1.70
SmartSensor 105	2.7%	2.4%	1.82

Appendix E Fifteen-Minute Analysis Additional Figures and Tables

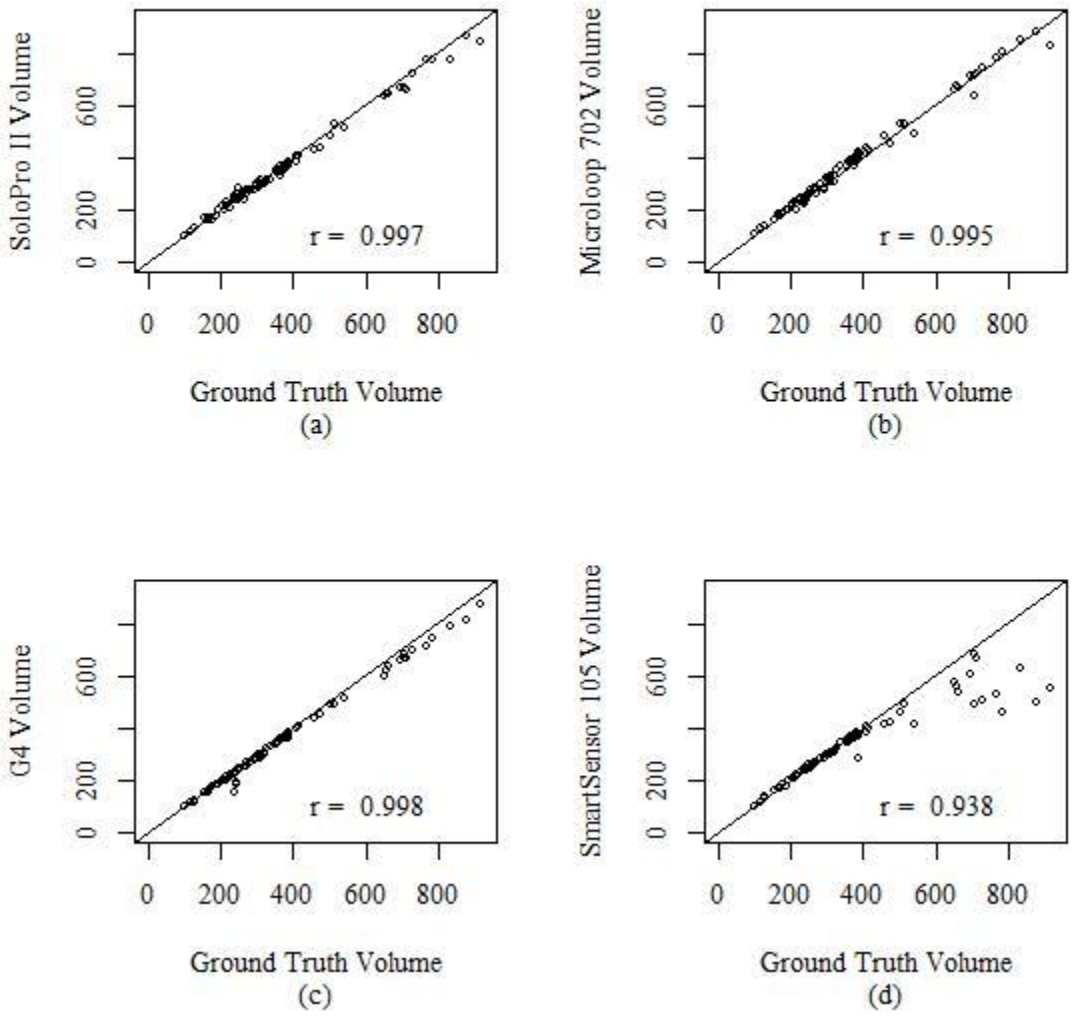

Figure E.1: Fifteen-Minute Volume Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

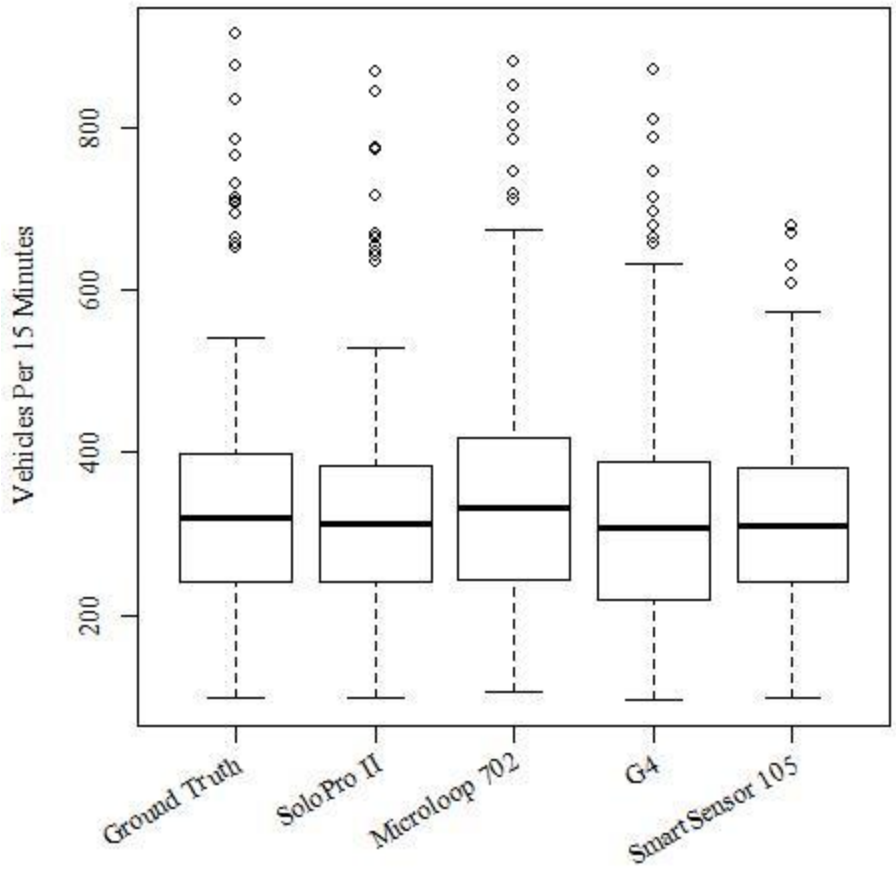

Figure E.2: Box Plot of Reported Fifteen-Minute Volumes

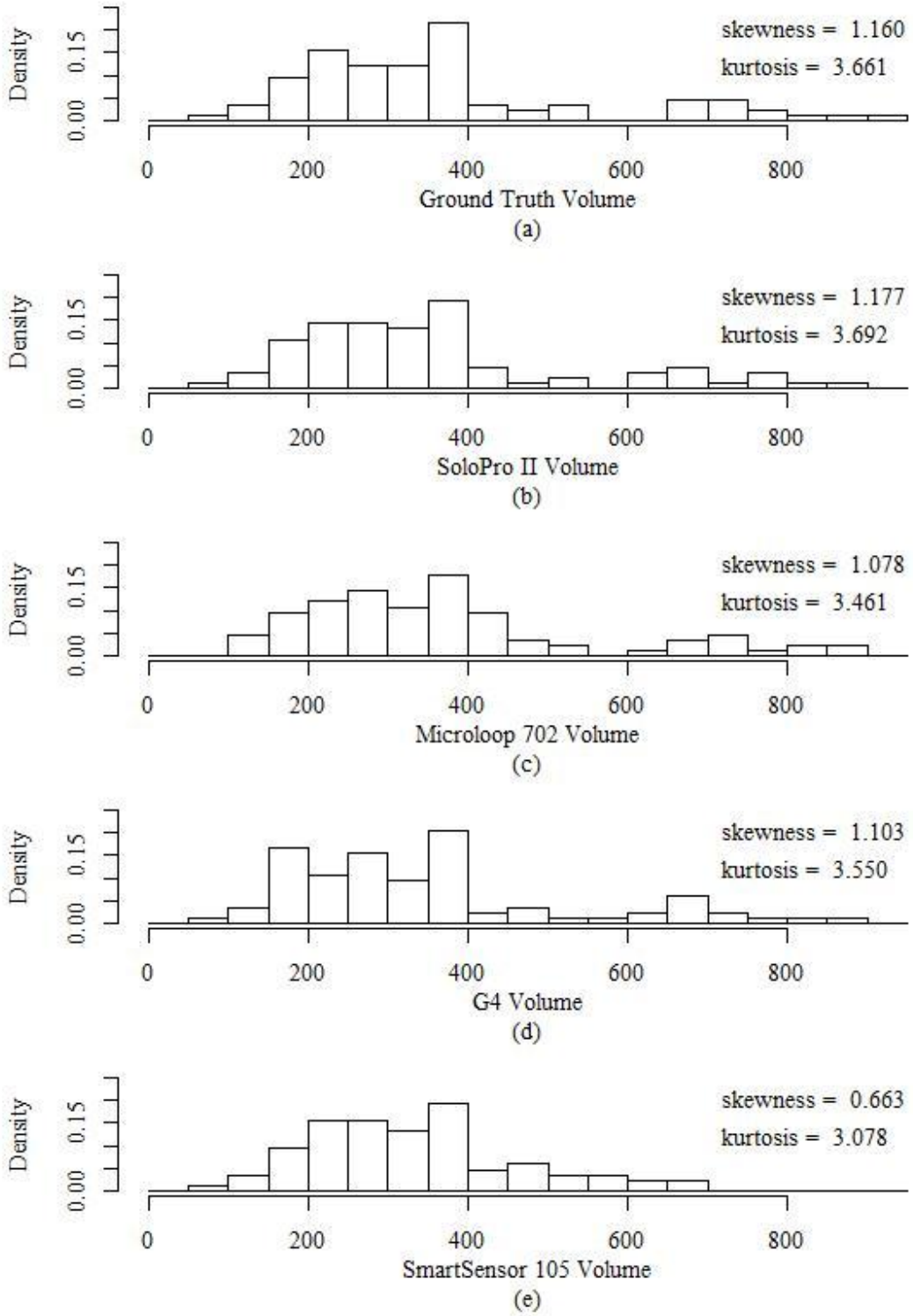

Figure E.3: Histograms of Fifteen-Minute Volume Distributions for Ground Truth (a), Solo Pro II (b), Microloop 702 (c), G4 (d), and SmartSensor 105 (e)

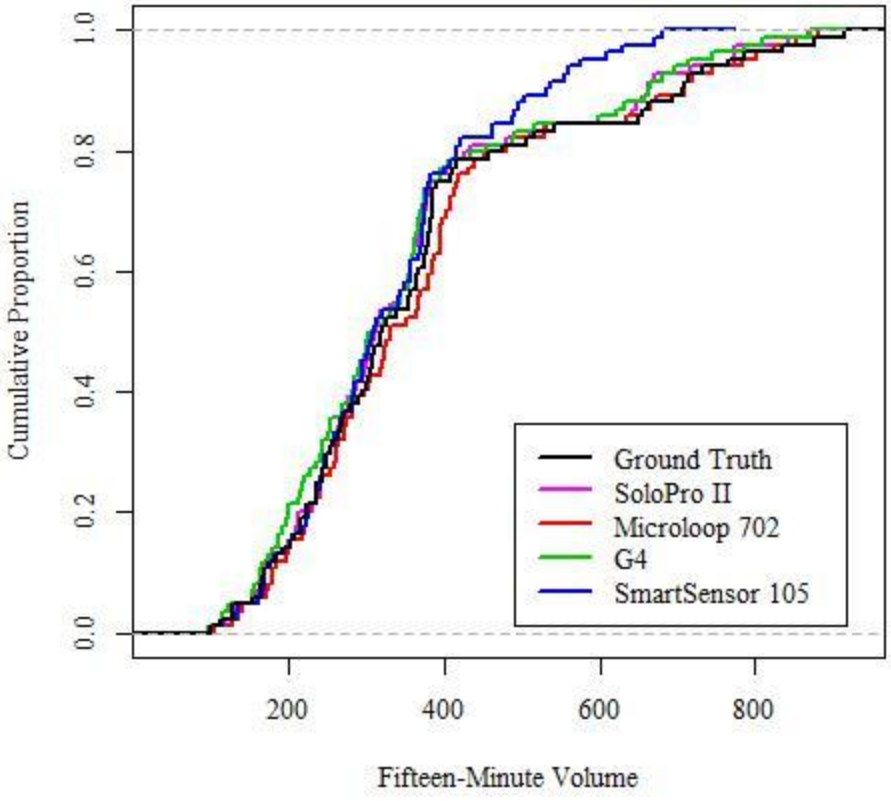

Figure E.4: Cumulative Distribution Plot of Fifteen-Minute Volume Distributions for Ground Truth and All Detectors

Table E.1 Fifteen-Minute Volume Summary Statistics

	Mean	Median	Standard Deviation
Ground Truth	368	320	189
Solo Pro II	357	312	180
Microloop 702	376	332	185
G4	350	307	179
SmartSensor 105	332	310	130

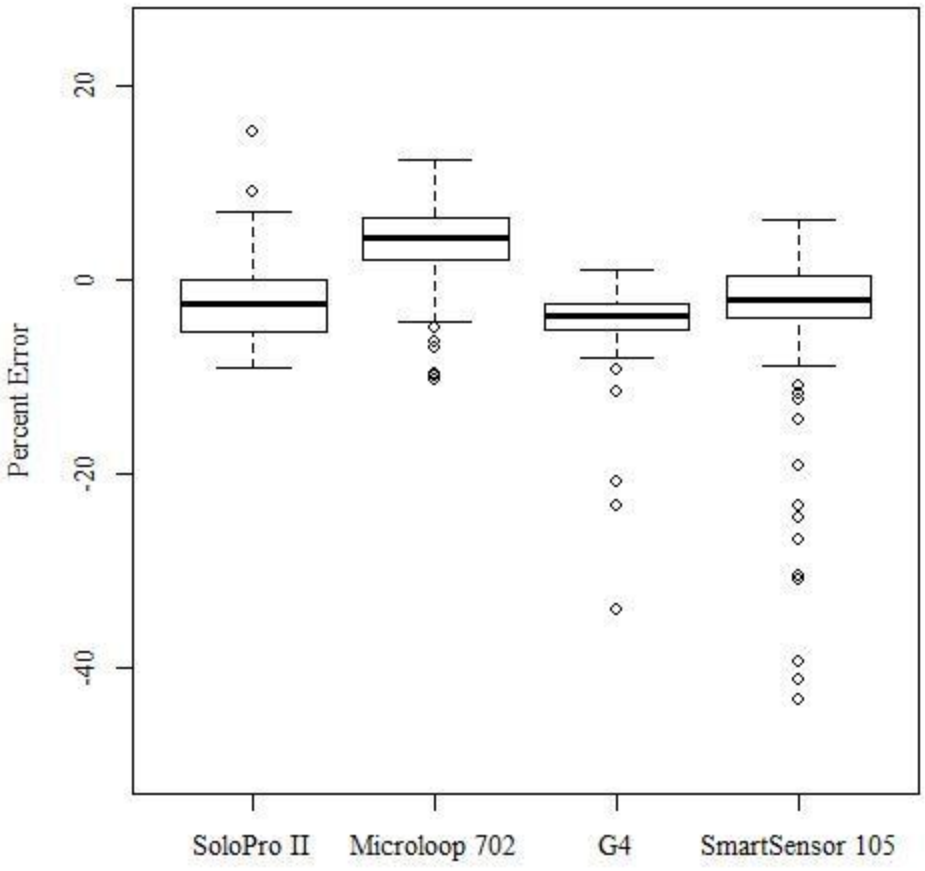

Figure E.5: Fifteen-Minute Volume Percent Error Box Plot

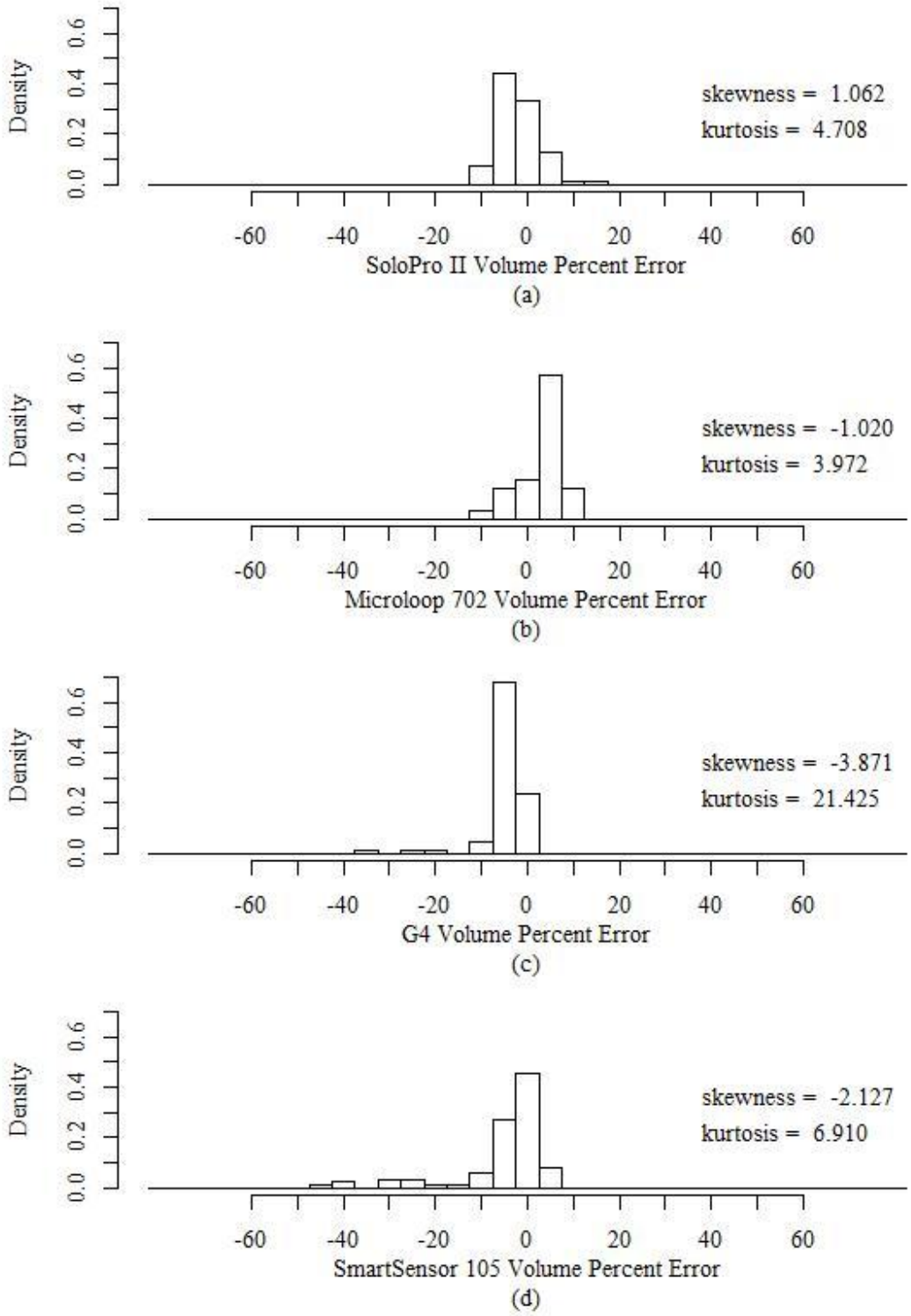

Figure E.6: Histograms of Fifteen-Minute Volume Percent Error Distributions for Solo Pro II (a), Microloop (b), G4 (c), and SmartSensor 105 (d) Detectors

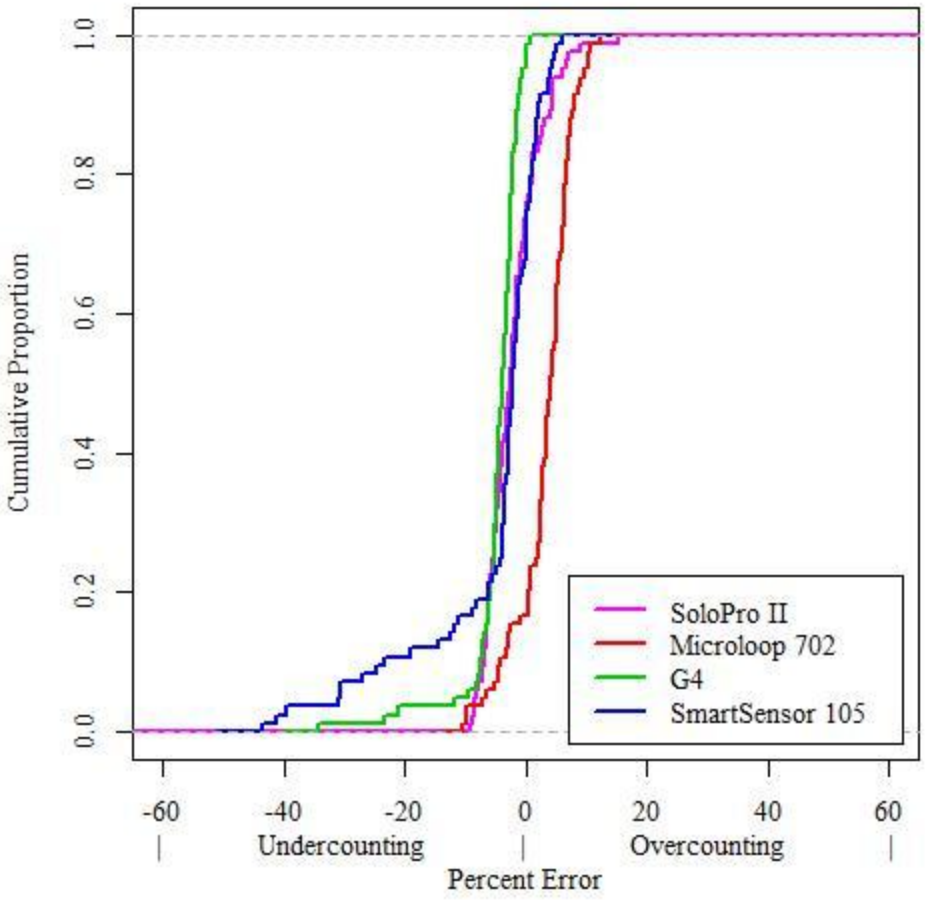

Figure E.7: Fifteen-Minute Volume Percent Error Cumulative Distribution Plot

Table E.2: Detector Fifteen-Minute Volume Error Statistics

	Correlation Coefficient	MPE	MAPE	Percent Error Variance	Mean GEH	85th Percentile GEH	GEH Variance
SoloPro II	0.997	-2.14%	4.08%	0.00199	0.766	1.25	0.313
Microloop 702	0.995	3.26%	5.03%	0.00221	0.880	1.27	0.265
G4	0.998	-4.71%	4.73%	0.00233	0.913	1.41	0.744
SmartSensor 105	0.938	-5.22%	6.47%	0.0112	1.64	2.82	9.59

Table E.3: Fifteen-Minute Volume Theil's Inequality Coefficients

	U	U_m	U_s	U_c
SoloPro II	0.025	0.275	0.239	0.495
Microloop 702	0.025	0.156	0.039	0.816
G4	0.030	0.539	0.191	0.276
SmartSensor 105	0.115	0.166	0.453	0.391

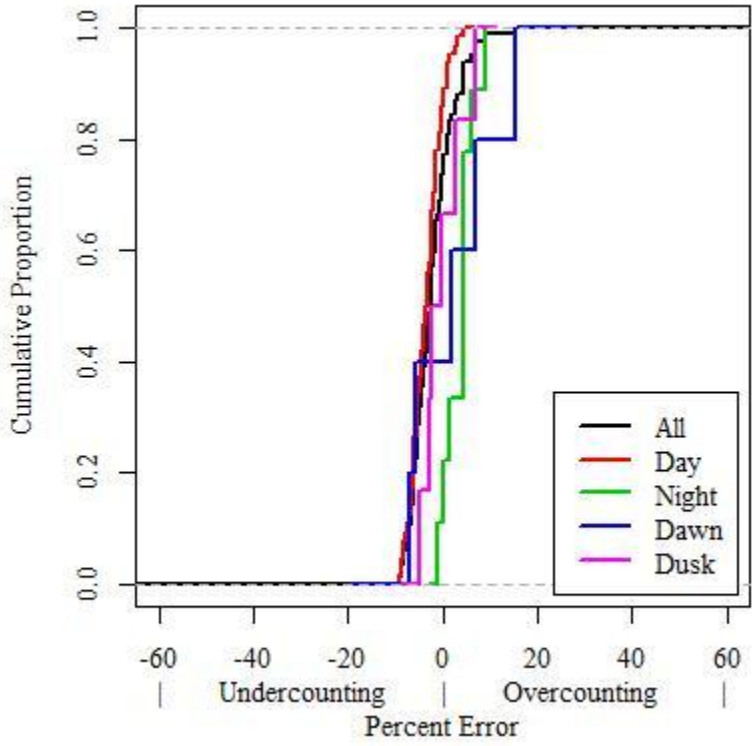

Figure E.8: Solo Pro II Fifteen-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

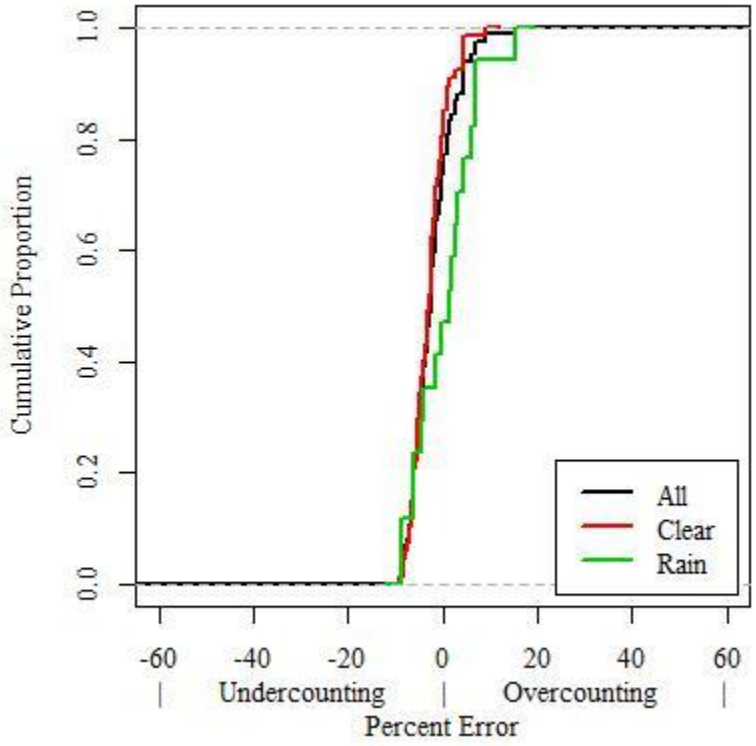

Figure E.9: Solo Pro II Fifteen-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

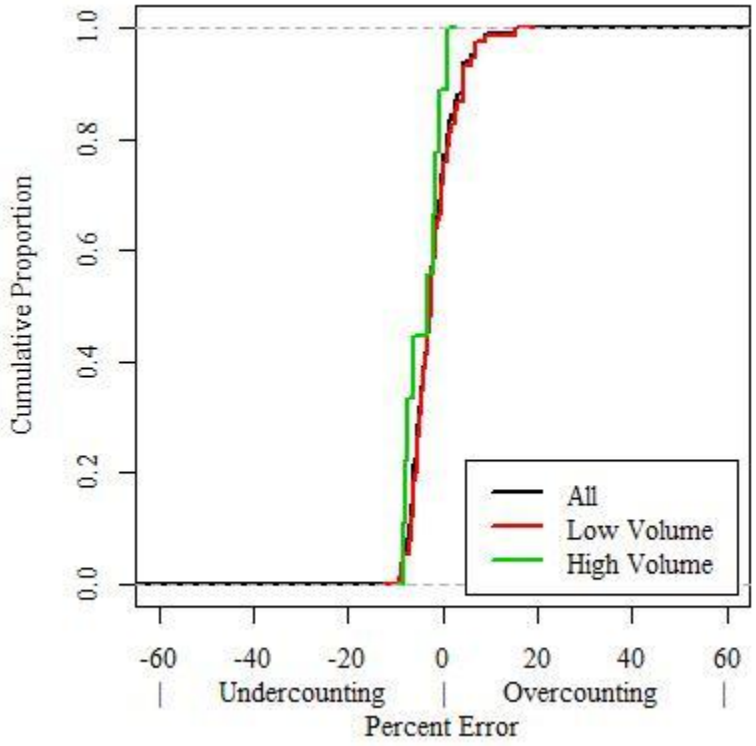

Figure E.10: Solo Pro II Fifteen-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

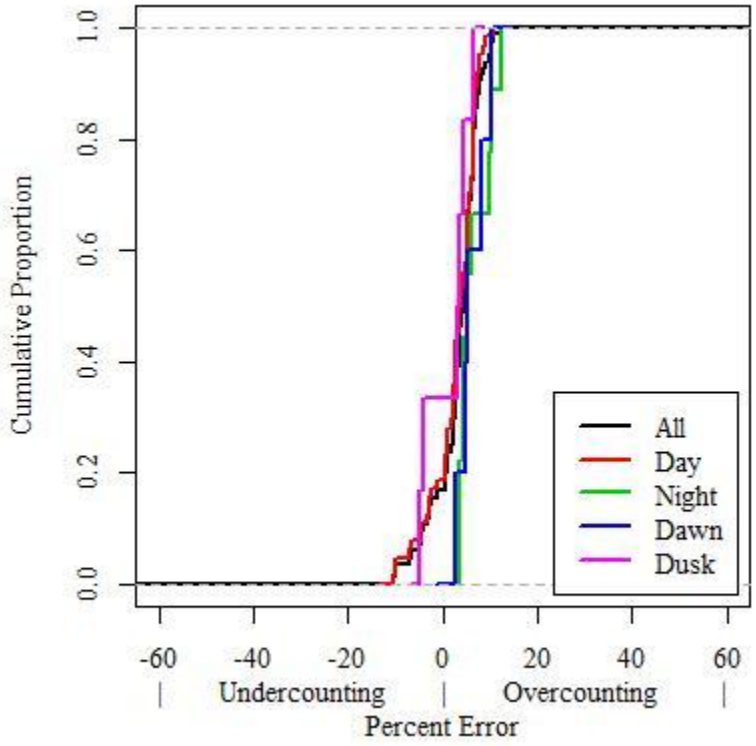

Figure E.11: Microloop 702 Fifteen-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

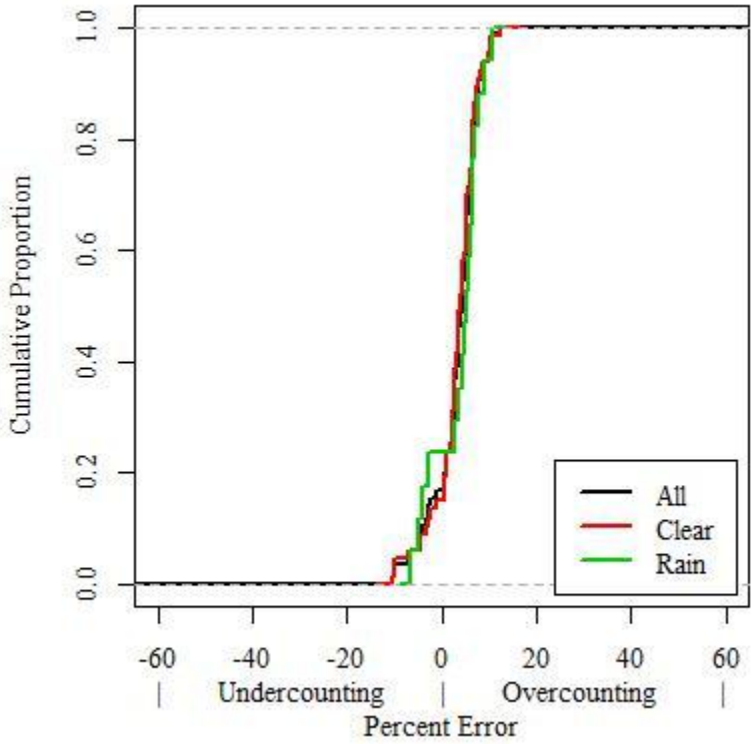

Figure E.12: Microloop 702 Fifteen-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

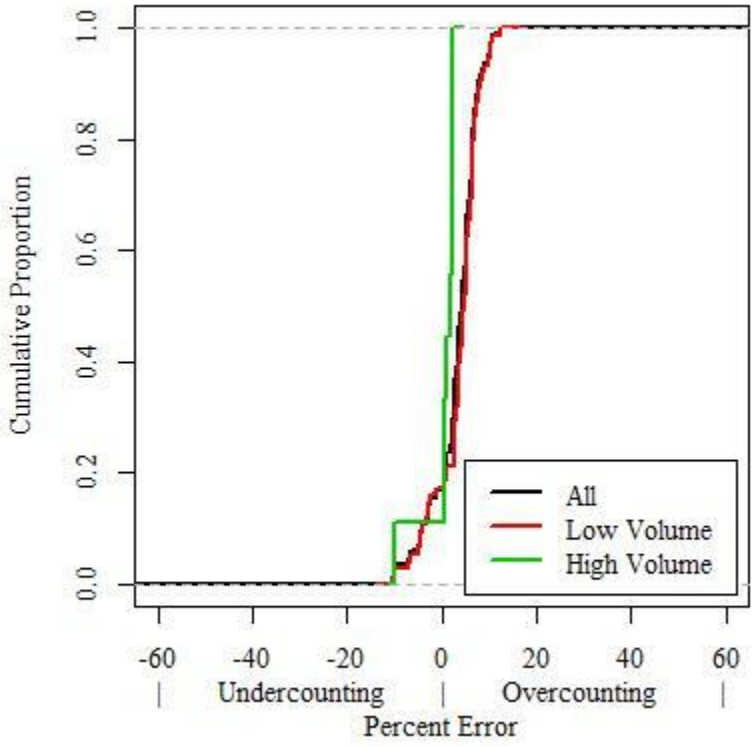

Figure E.13: Microloop 702 Fifteen-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

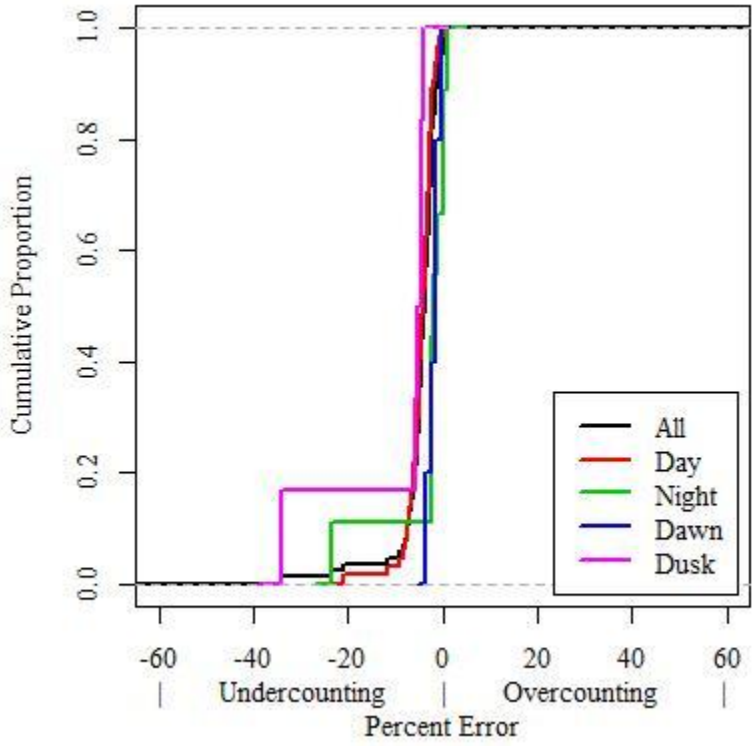

Figure E.14: G4 Fifteen-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

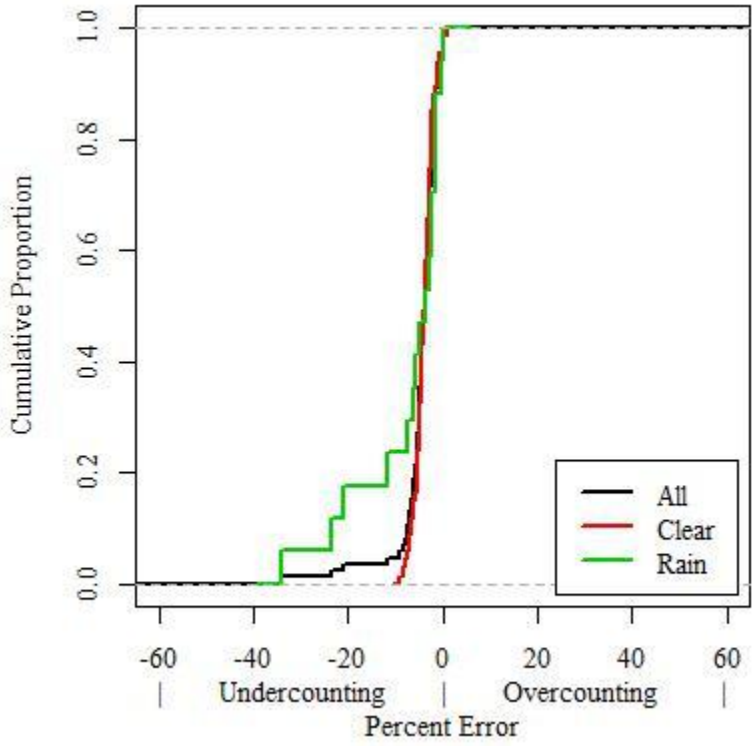

Figure E.15: G4 Fifteen-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

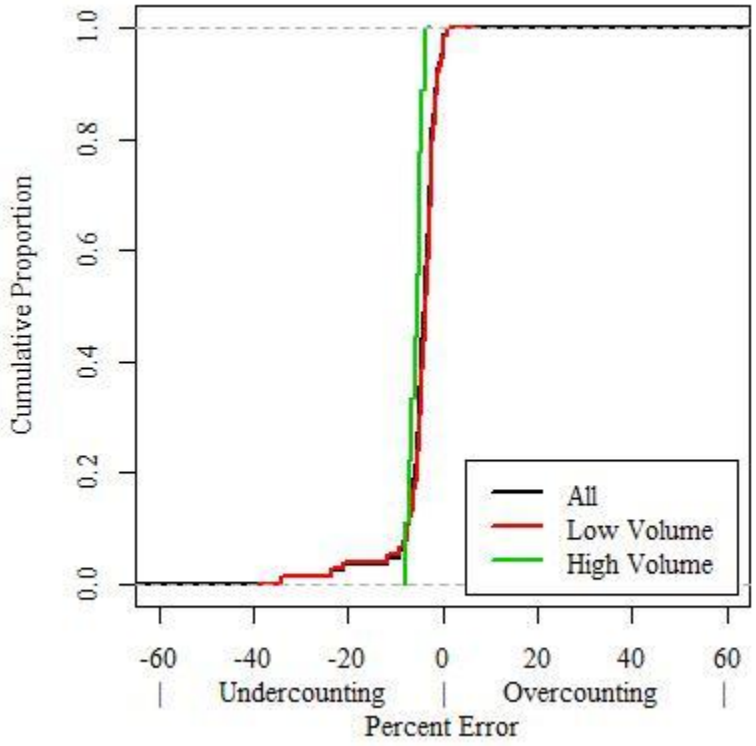

Figure E.16: G4 Fifteen-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

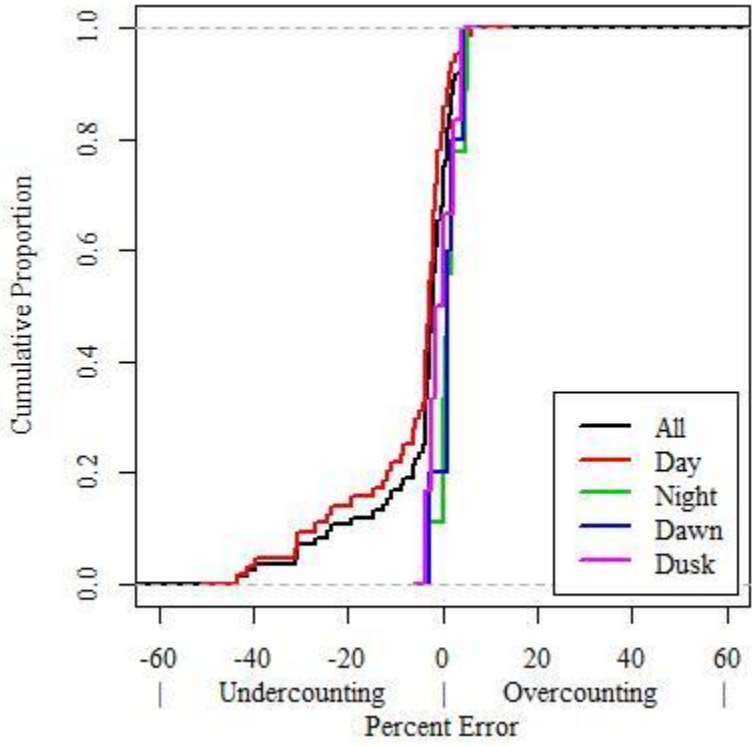

Figure E.17: SmartSensor 105 Fifteen-Minute Volume Percent Error Lighting Factor Cumulative Distribution Plot

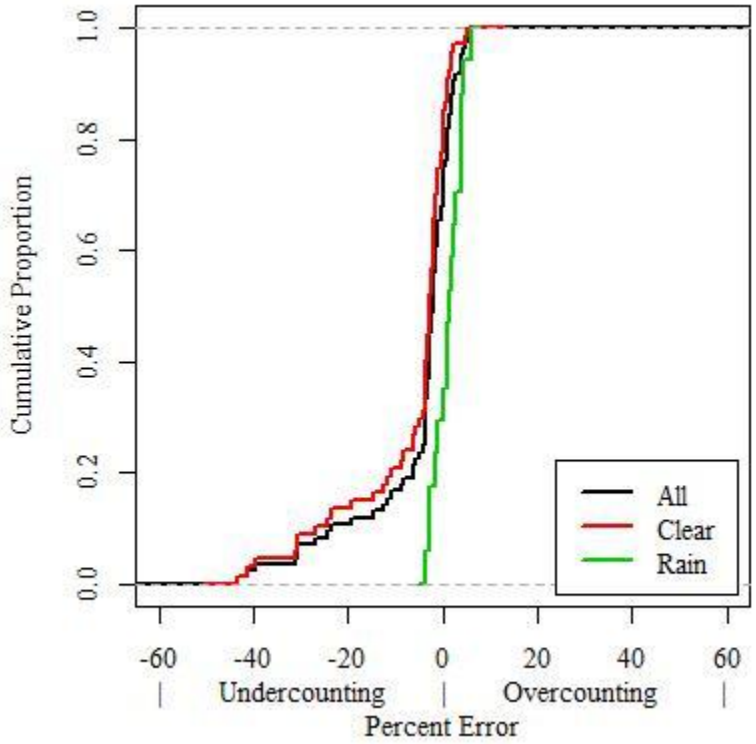

Figure E.18: SmartSensor 105 Fifteen-Minute Volume Percent Error Rain Factor Cumulative Distribution Plot

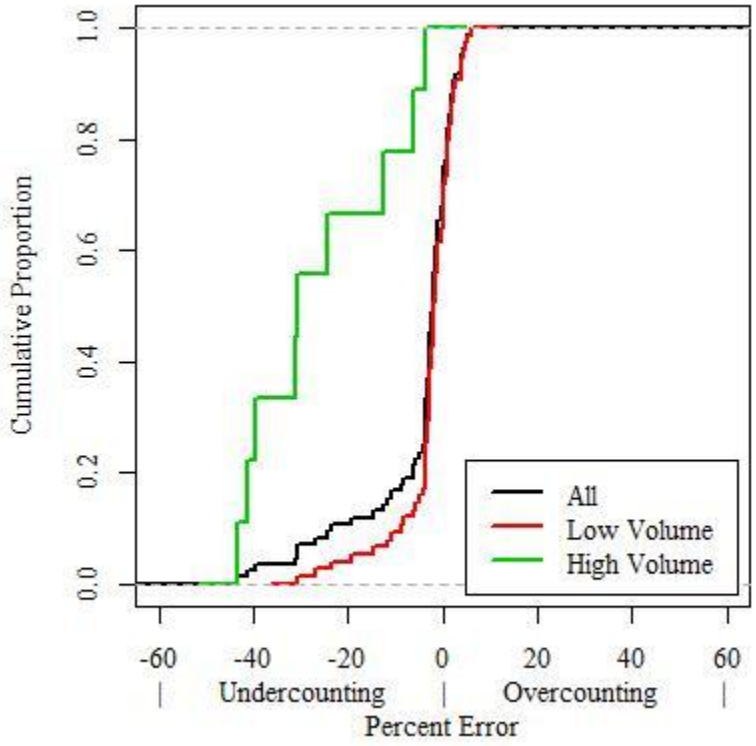

Figure E.19: SmartSensor 105 Fifteen-Minute Volume Percent Error Volume Factor Cumulative Distribution Plot

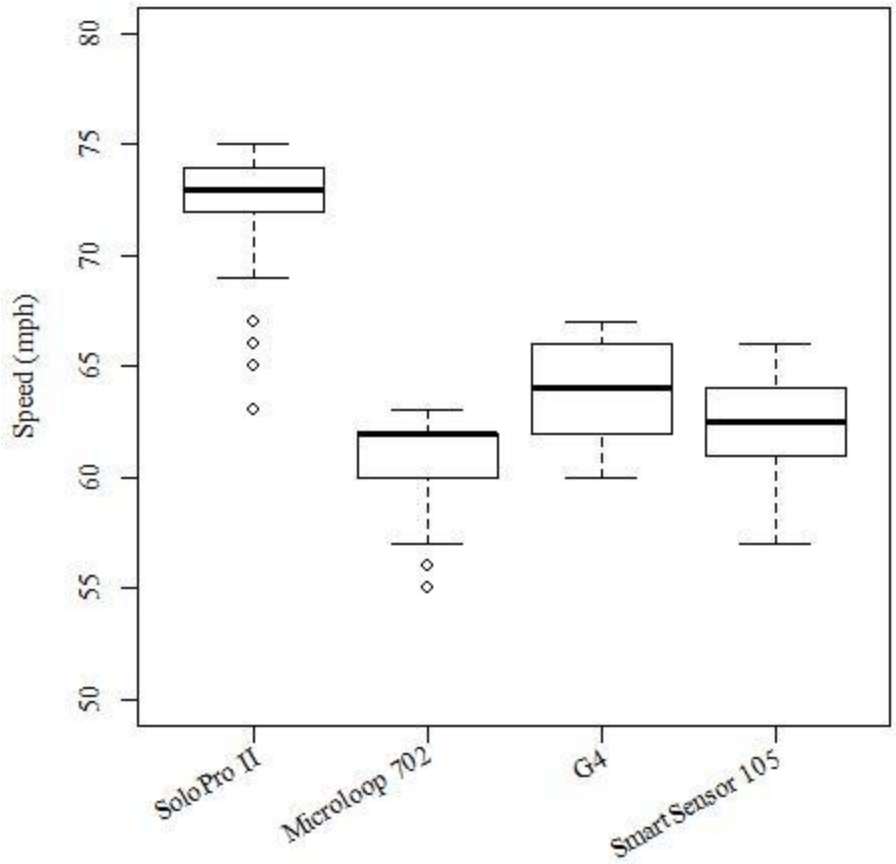

Figure E.20: Box Plot of Reported Fifteen-Minute Mean Speeds

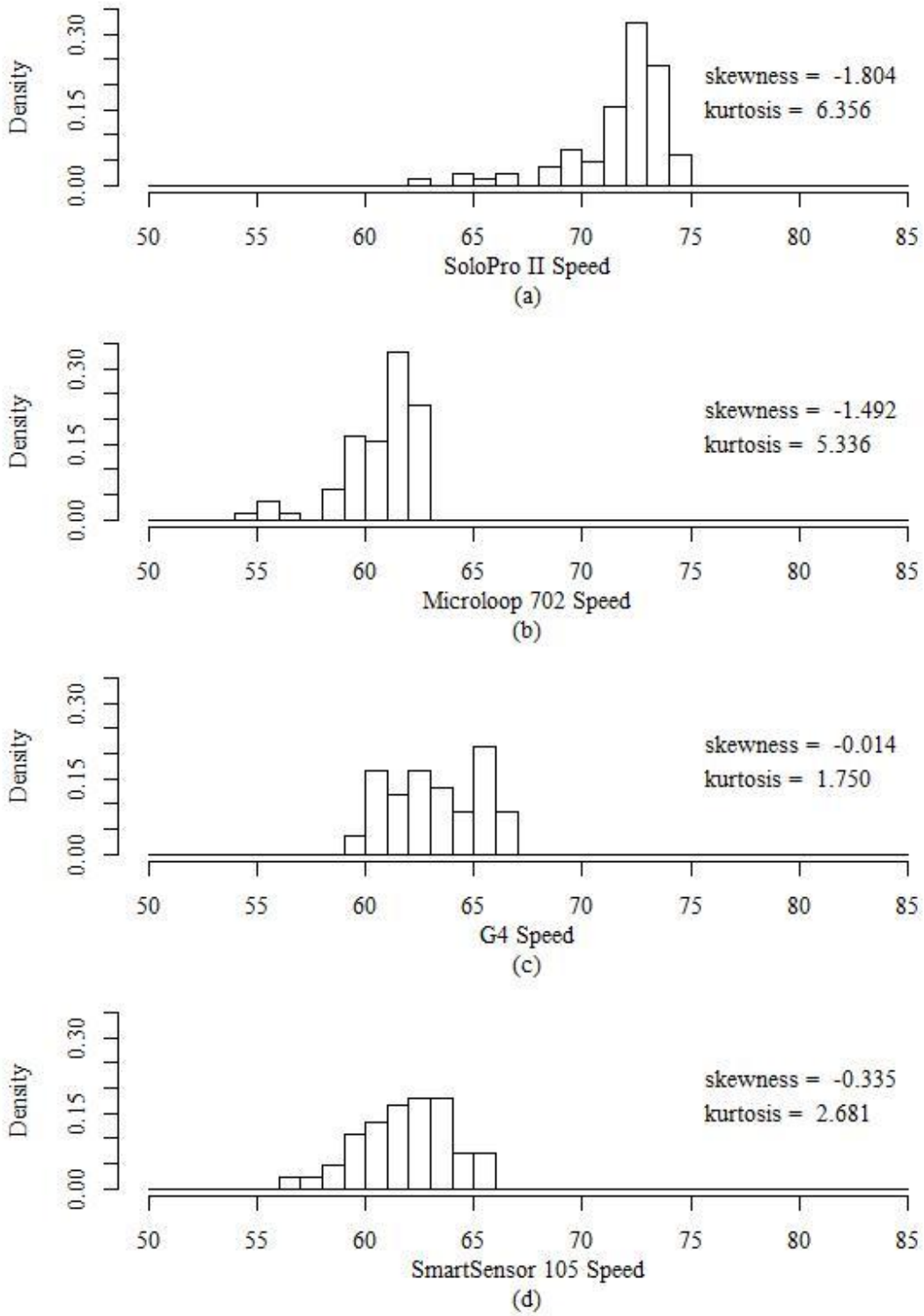

Figure E.21: Histograms of Fifteen-Minute Mean Speed Distributions for the Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

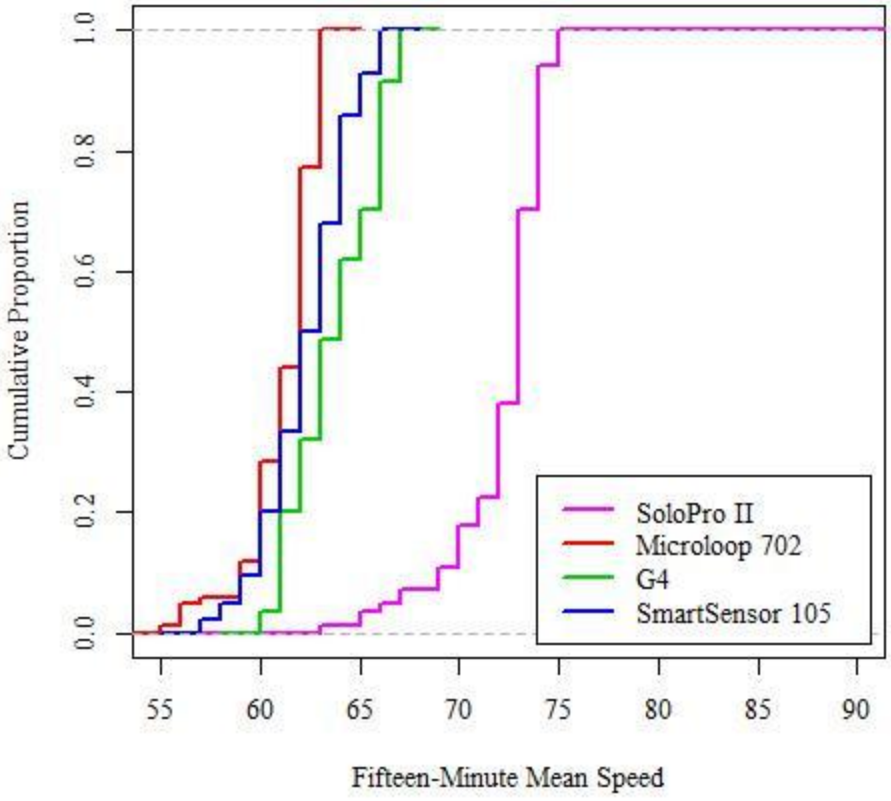

Figure E.22: Cumulative Distribution Plot of Fifteen-Minute Mean Speed Distributions for All Detectors

Table E.4 Fifteen-Minute Mean Speed Summary Statistics

	Mean	Median	Standard Deviation
Solo Pro II	72	73	2.37
Microloop 702	61	62	1.78
G4	64	64	2.09
SmartSensor 105	62	63	2.14

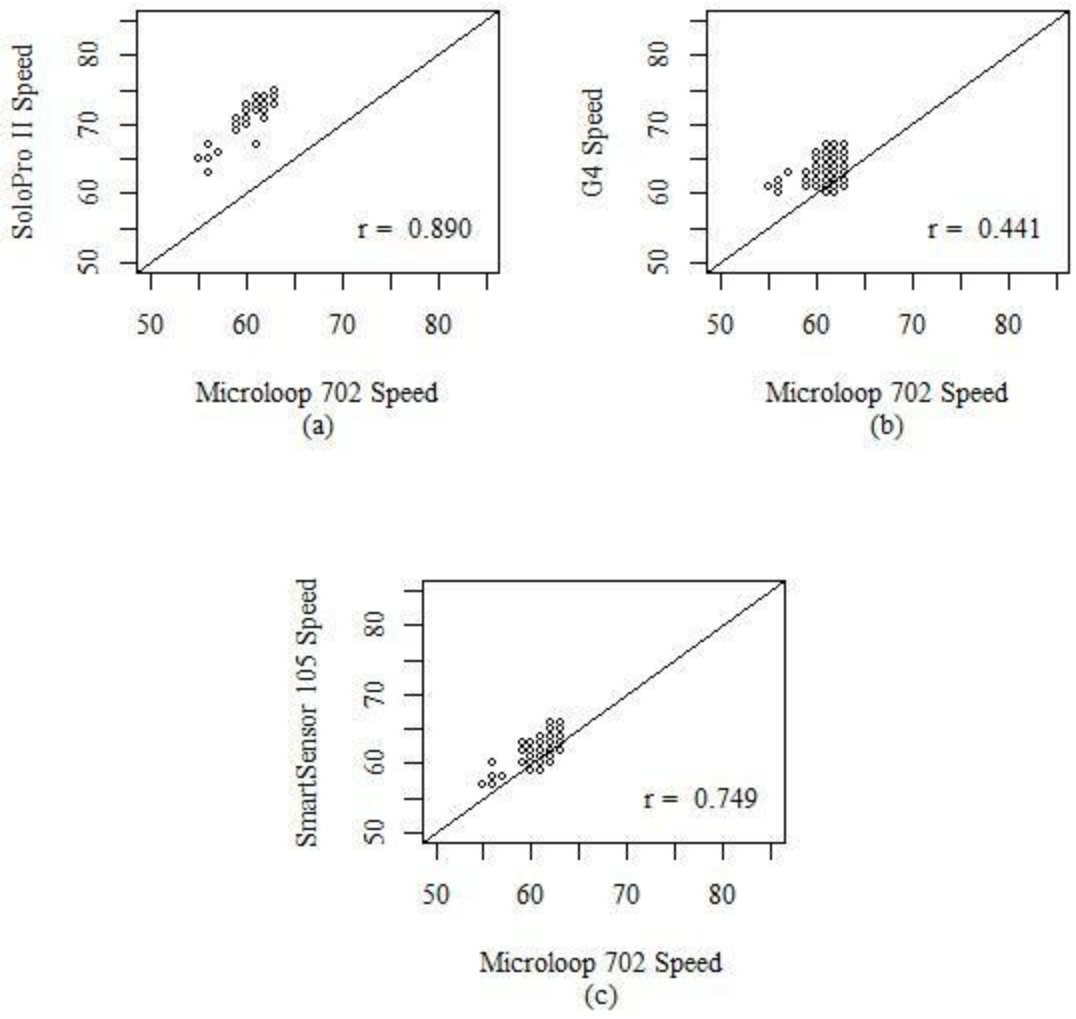

Figure E.23: Fifteen-Minute Mean Speed Scatter Plots Against Baseline for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

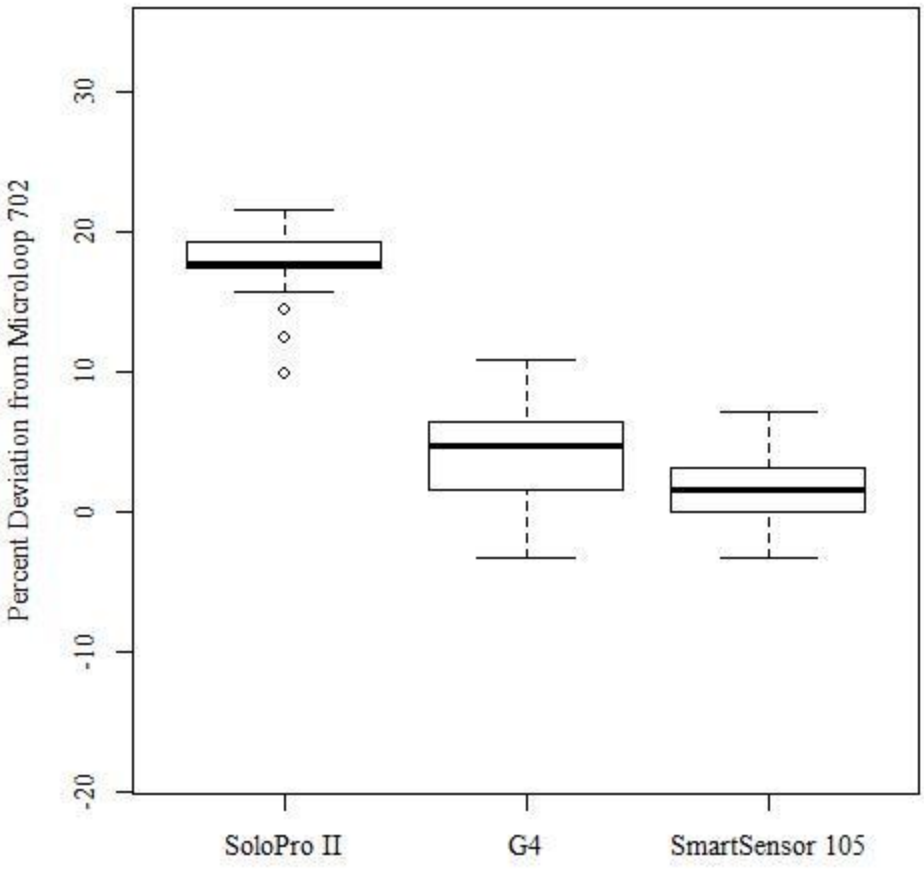

Figure E.24: Fifteen-Minute Mean Speed Percent Deviation Box Plot

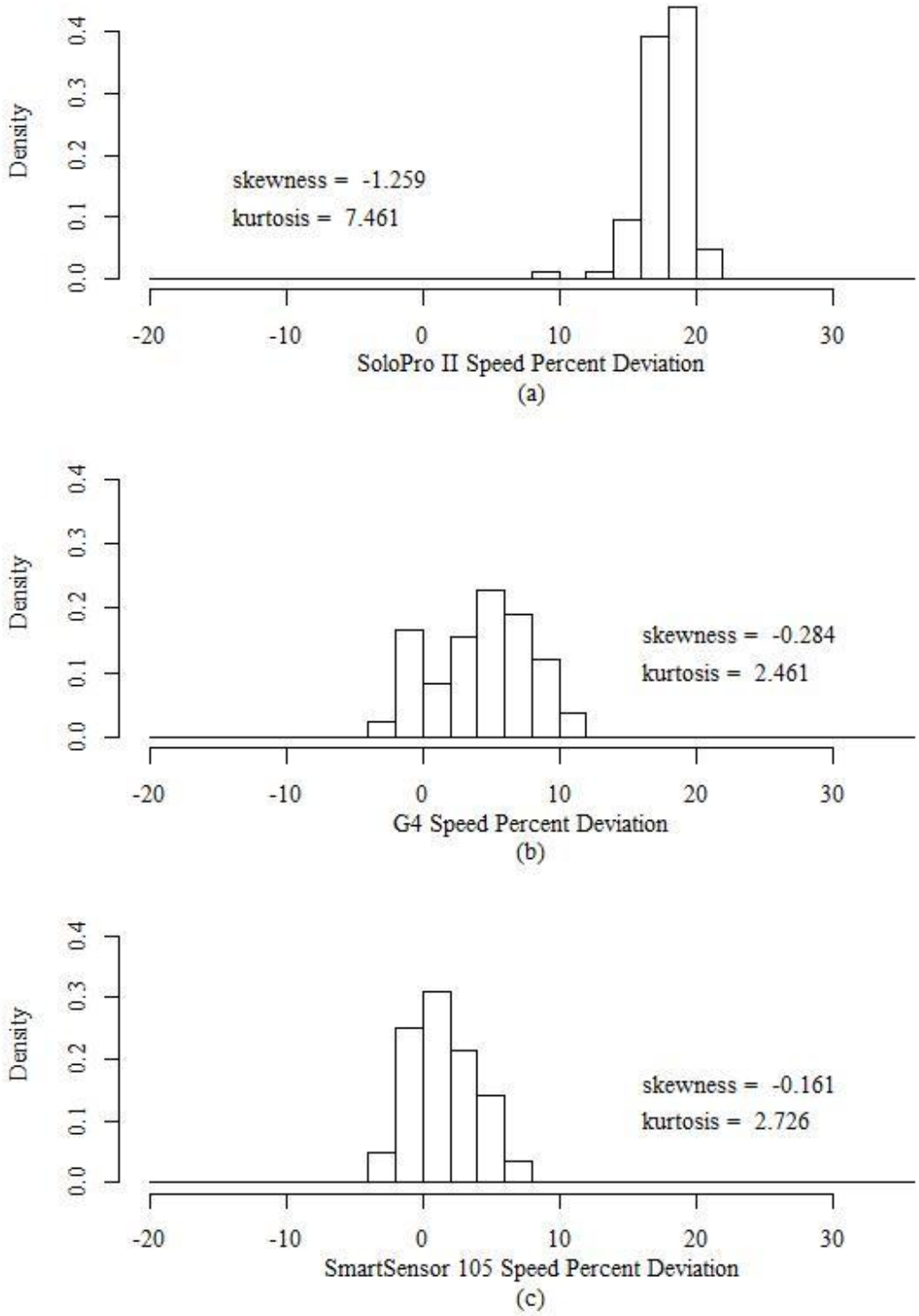

Figure E.25: Histograms of Fifteen-Minute Mean Speed Percent Deviation Distributions for Solo Pro II (a), G4 (b), and SmartSensor 105 (c) Detectors

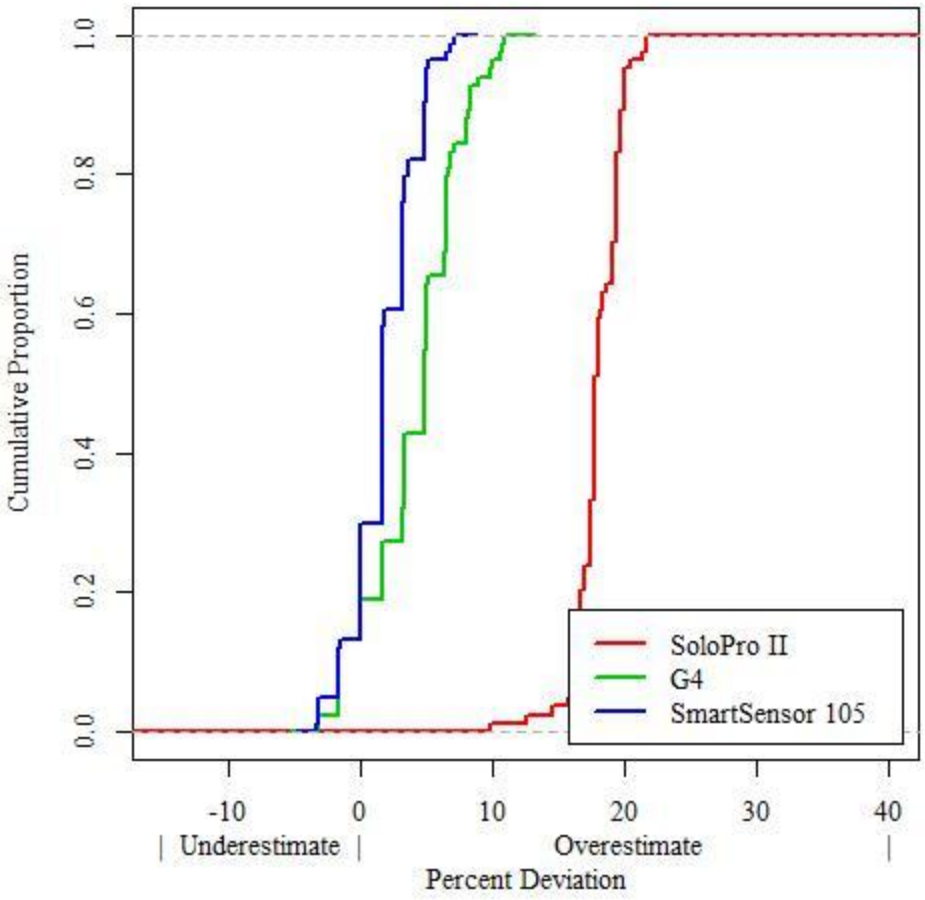

Figure E.26: Fifteen-Minute Mean Speed Percent Deviation Cumulative Distribution Plot

Table E.5: Detector Fifteen-Minute Mean Speed Deviation Statistics

	MPD	MAPD	Percent Deviation Variance
SoloPro II	17.99%	17.99%	0.00032
G4	4.15%	4.65%	0.00118
SmartSensor 105	1.86%	2.44%	0.00055

Table E.6: Fifteen-Minute Mean Speed Theil's Inequality Coefficients

	U	U_m	U_s	U_c
SoloPro II	0.083	0.990	0.003	0.008
G4	0.026	0.600	0.010	0.395
SmartSensor 105	0.015	0.388	0.041	0.579

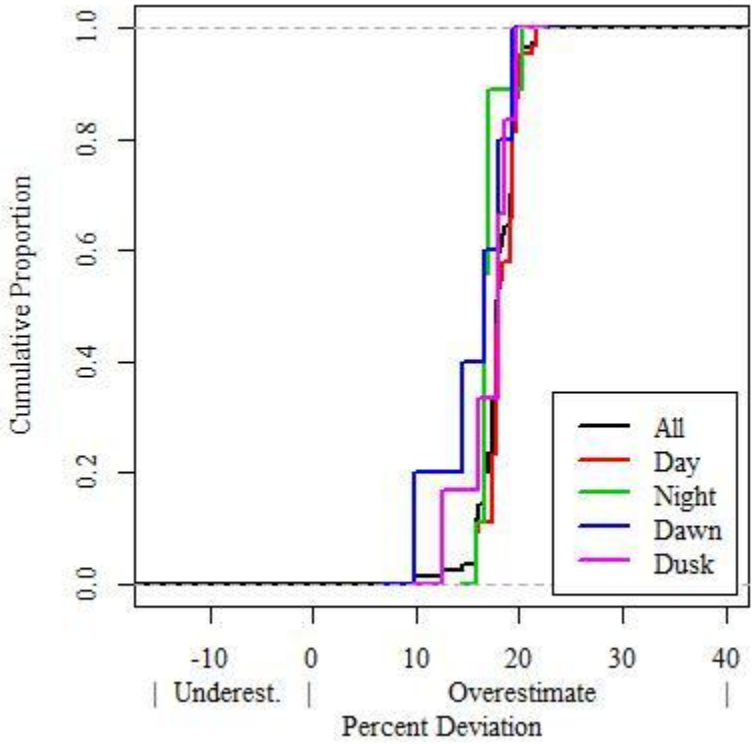

Figure E.27: Solo Pro II Fifteen-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

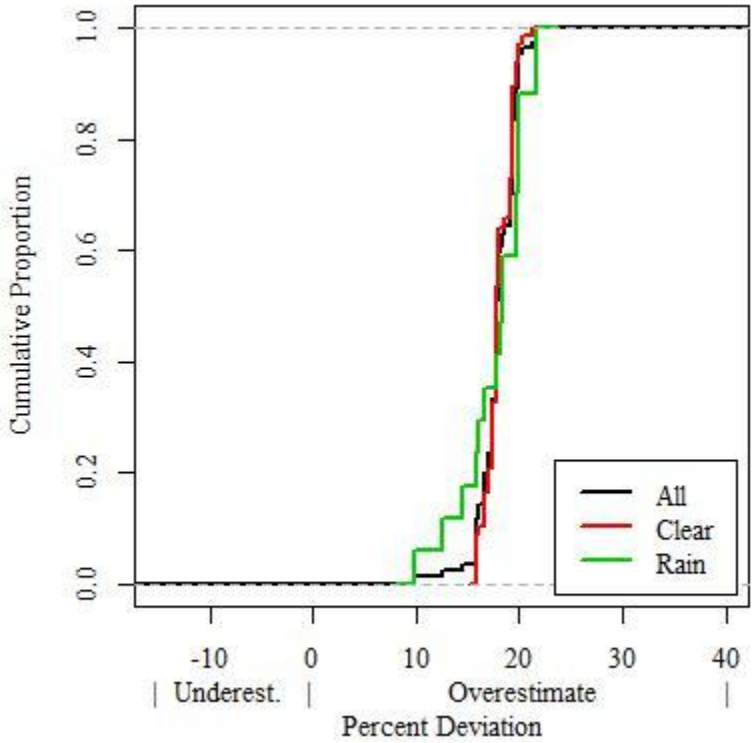

Figure E.28: Solo Pro II Fifteen-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

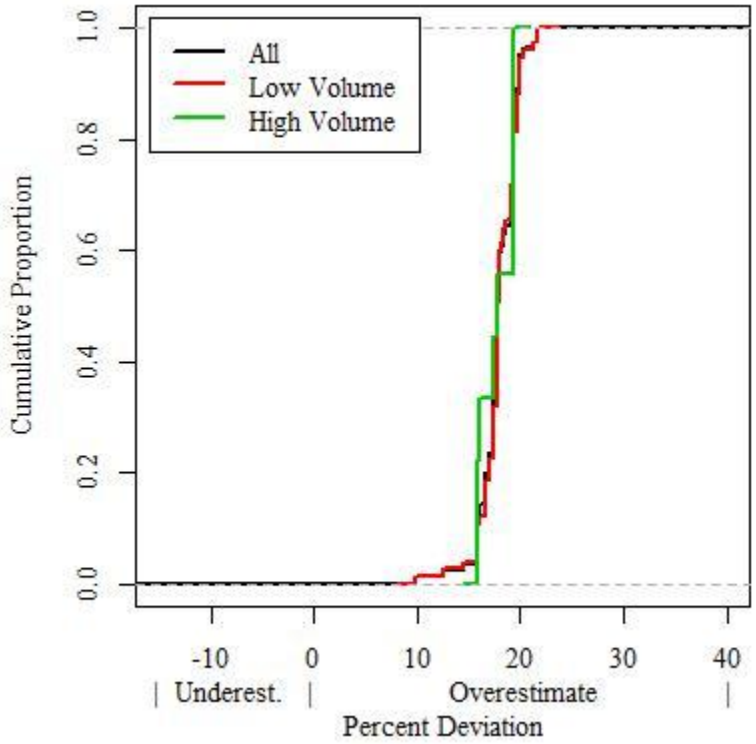

Figure E.29: Solo Pro II Fifteen-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

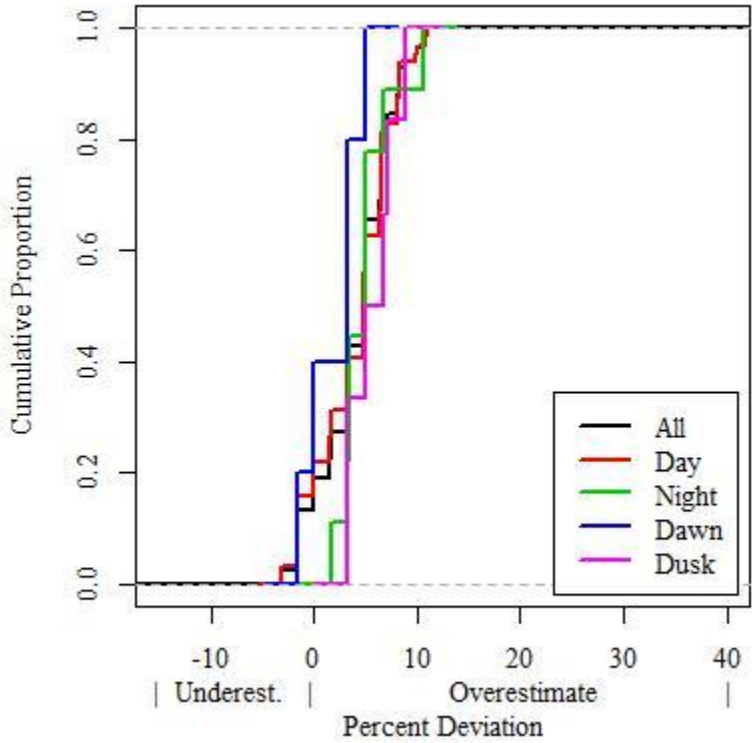

Figure E.30: G4 Fifteen-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

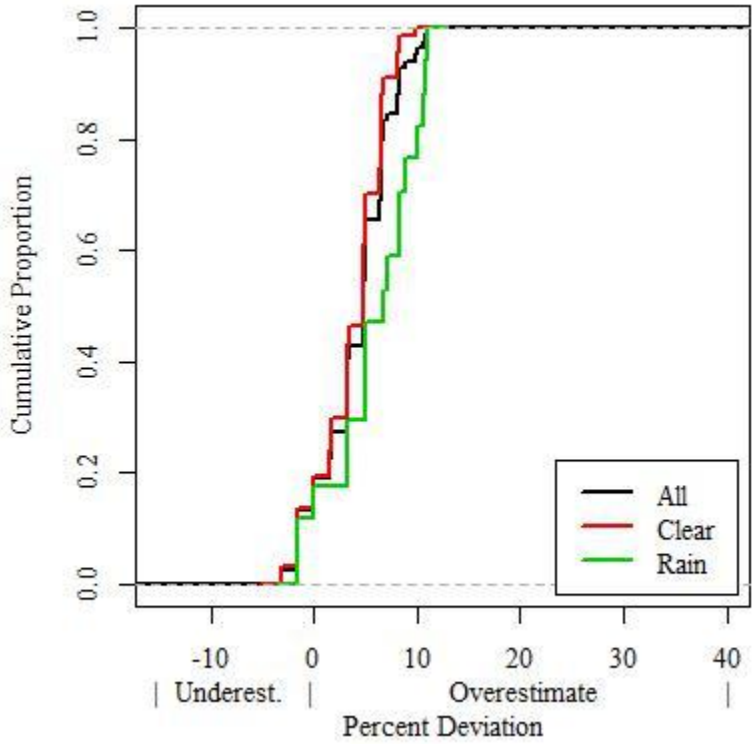

Figure E.31: G4 Fifteen-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

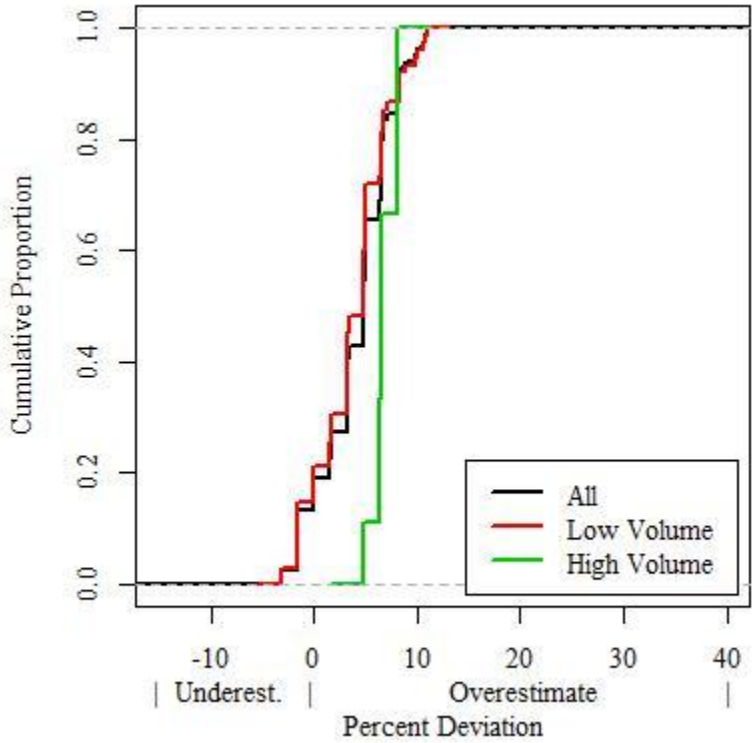

Figure E.32: G4 Fifteen-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

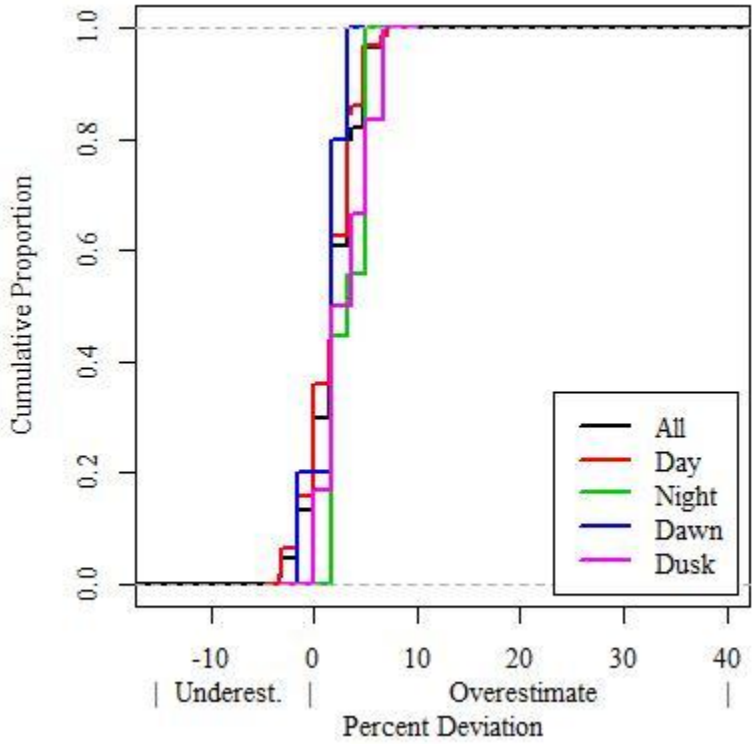

Figure E.33: SmartSensor 105 Fifteen-Minute Mean Speed Percent Deviation Lighting Factor Cumulative Distribution Plot

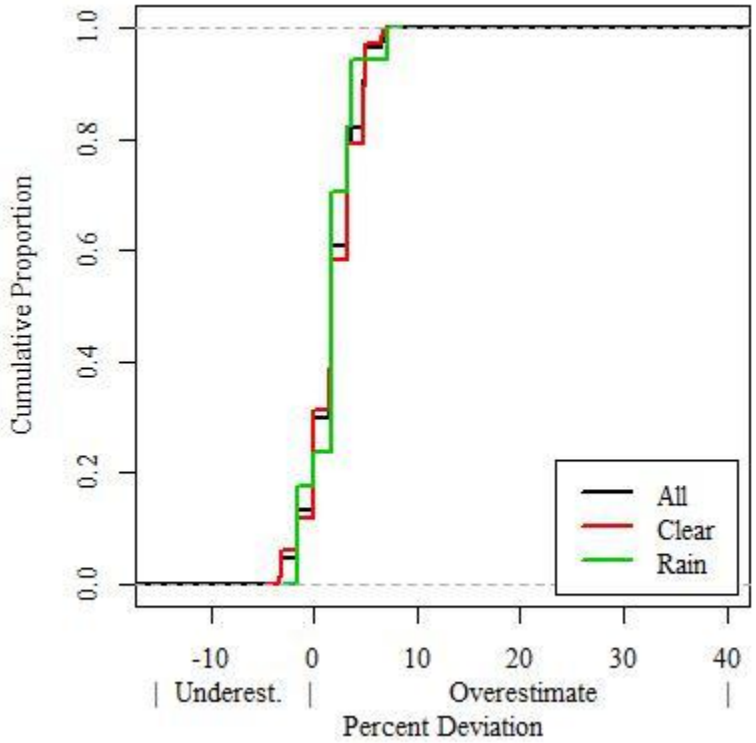

Figure E.34: SmartSensor 105 Fifteen-Minute Mean Speed Percent Deviation Rain Factor Cumulative Distribution Plot

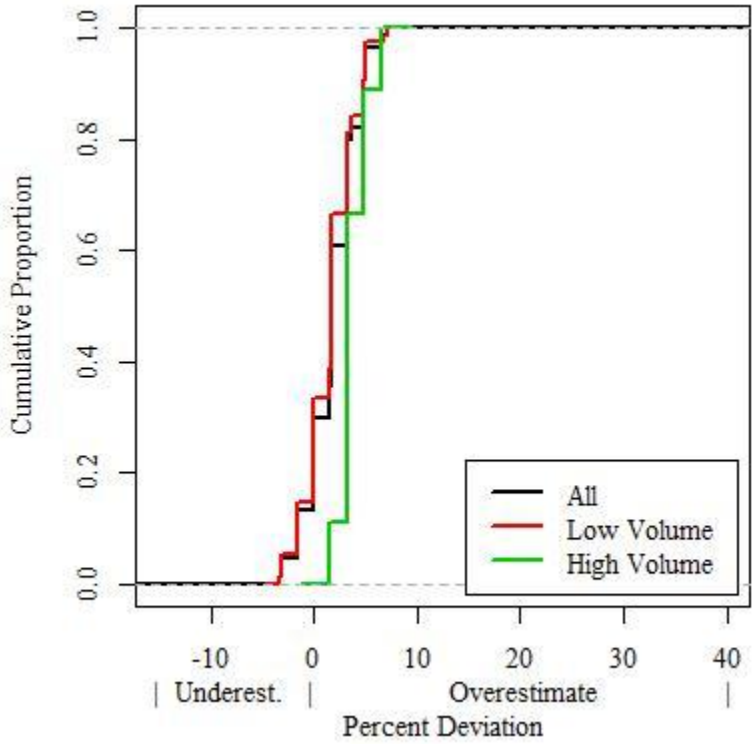

Figure E.35: SmartSensor 105 Fifteen-Minute Mean Speed Percent Deviation Volume Factor Cumulative Distribution Plot

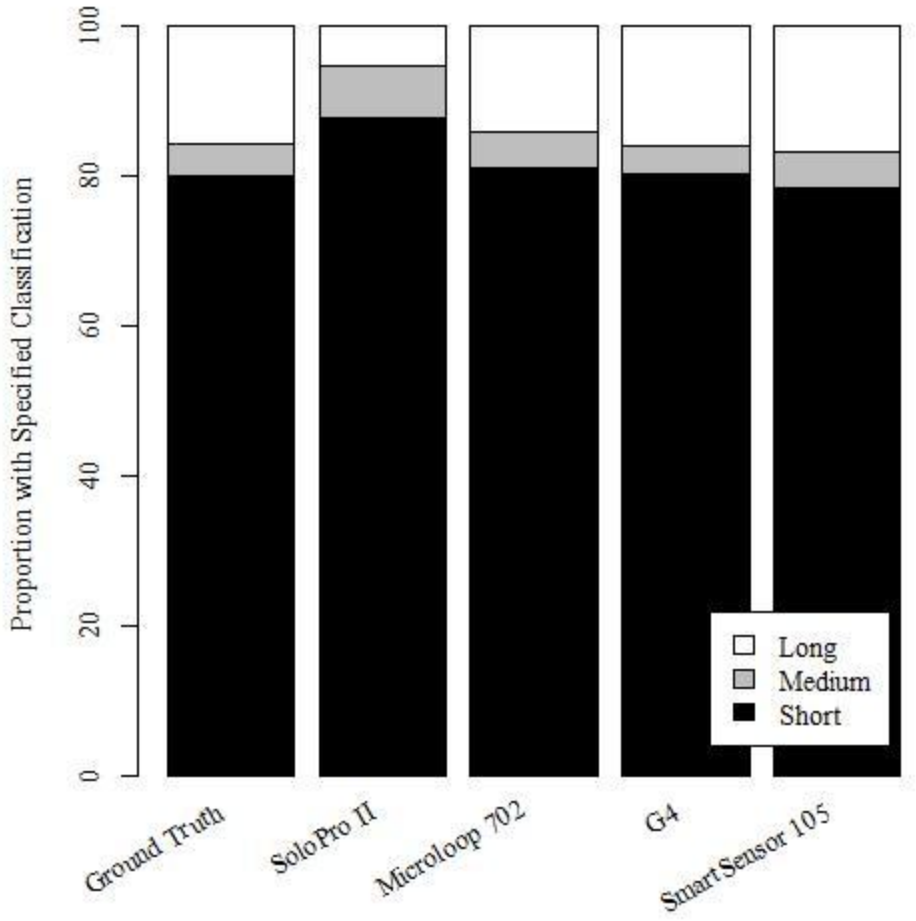

Figure E.36: Mean Fifteen-Minute Proportion Short, Medium, and Long Vehicles Bar Chart

Table E.7: Mean Fifteen-Minute Classification Proportions

	Ground Truth	SoloPro II	Microloop 702	G4	Smartsensor 105
Short	80.0%	87.6%	80.9%	80.2%	78.3%
Medium	4.3%	6.8%	4.8%	3.7%	4.9%
Long	15.8%	5.5%	14.3%	16.1%	16.8%

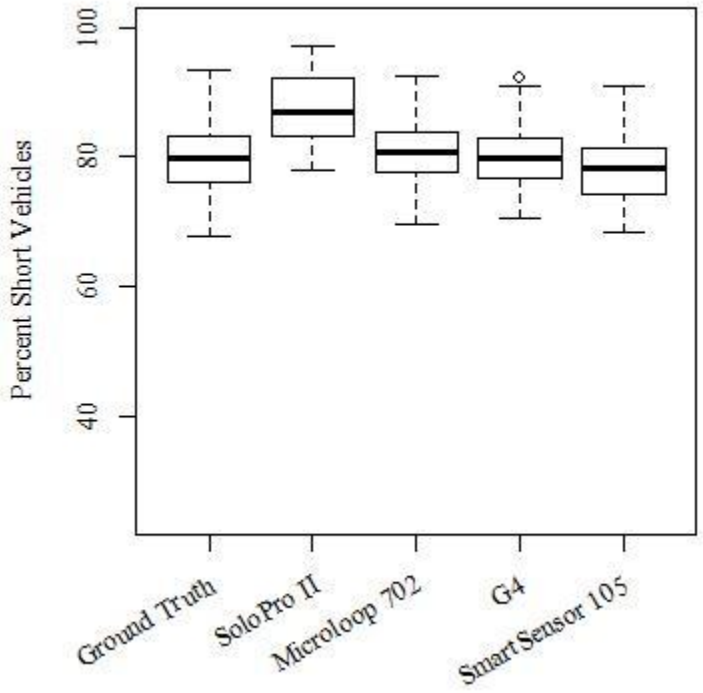

Figure E.37: Box Plot of Fifteen-Minute Percent Short Vehicle Distributions

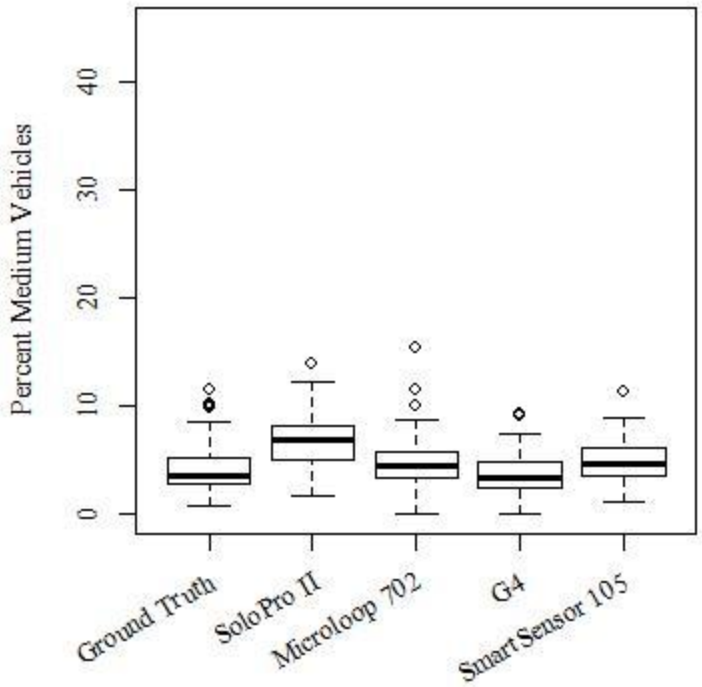

Figure E.38: Box Plot of Fifteen-Minute Percent Medium Vehicle Distributions

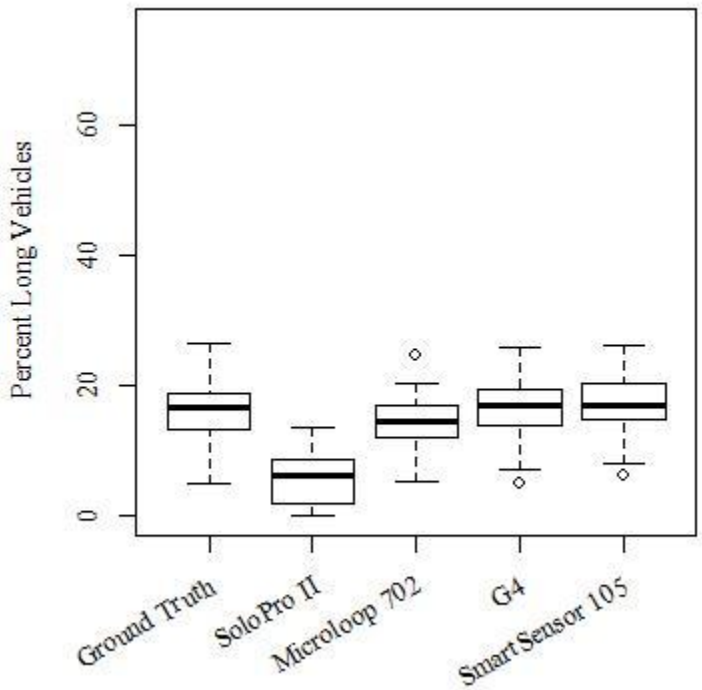

Figure E.39: Box Plot of Fifteen-Minute Percent Long Vehicle Distributions

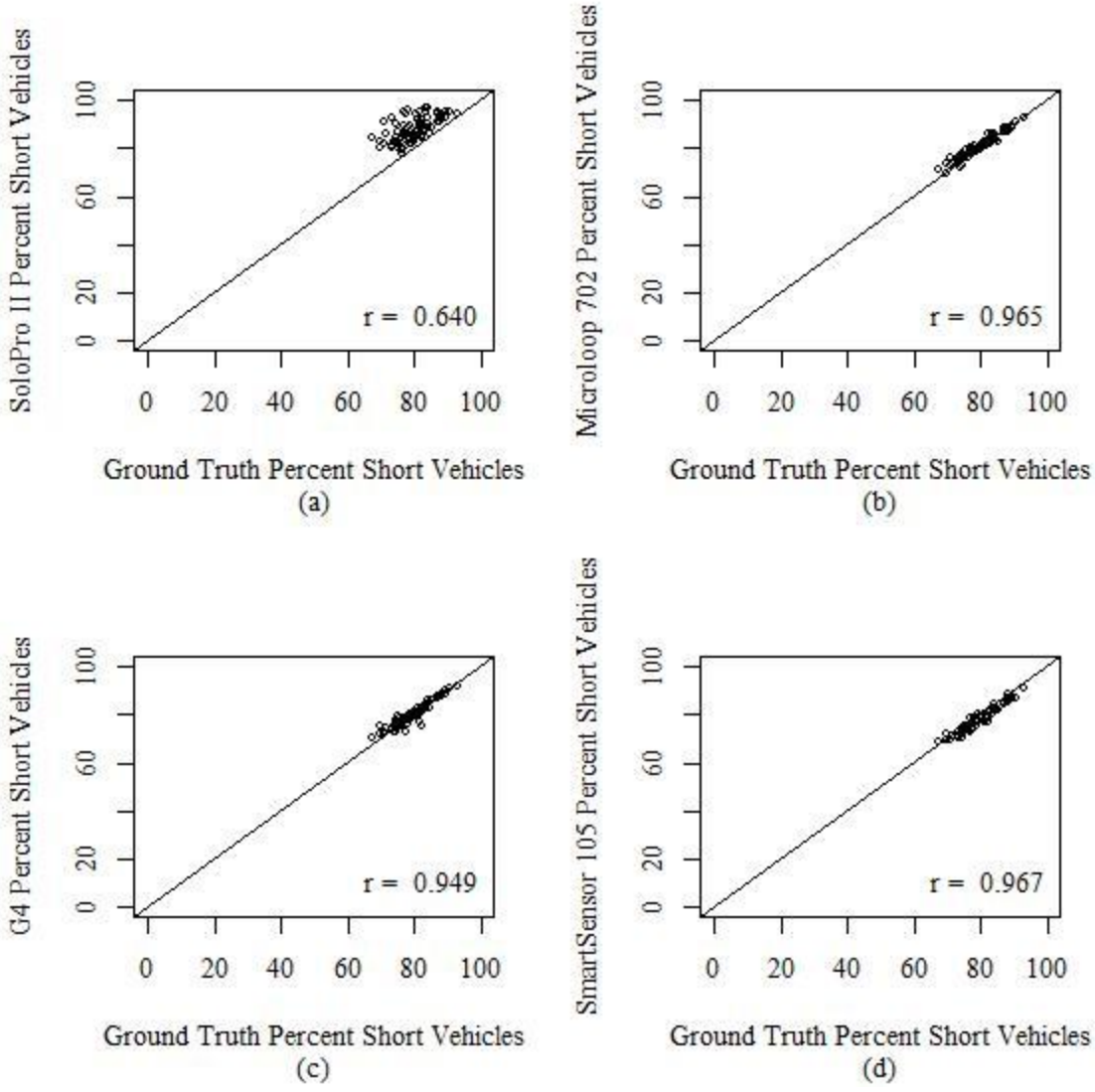

Figure E.40: Fifteen-Minute Percent Short Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

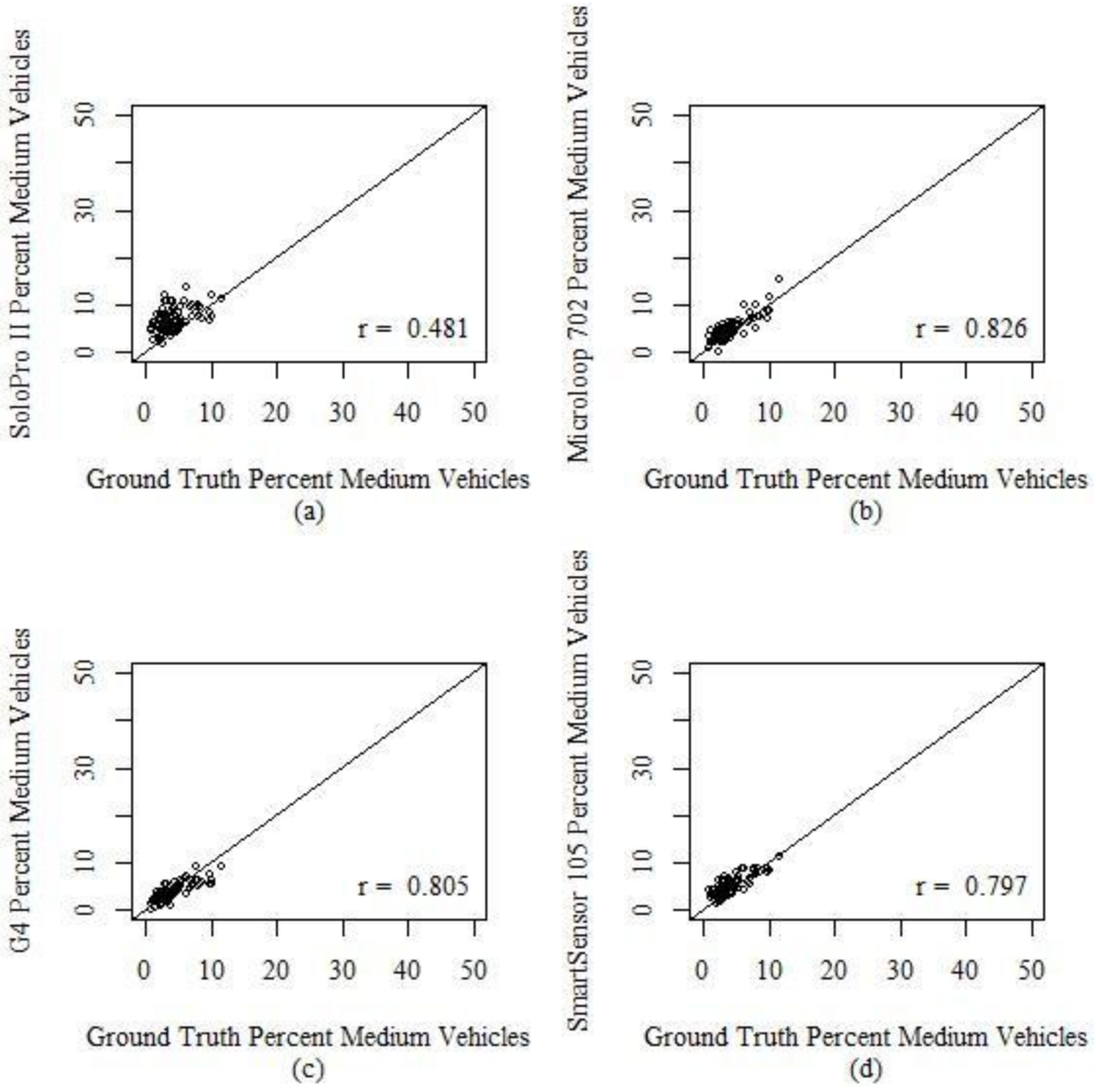

Figure E.41: Fifteen-Minute Percent Medium Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

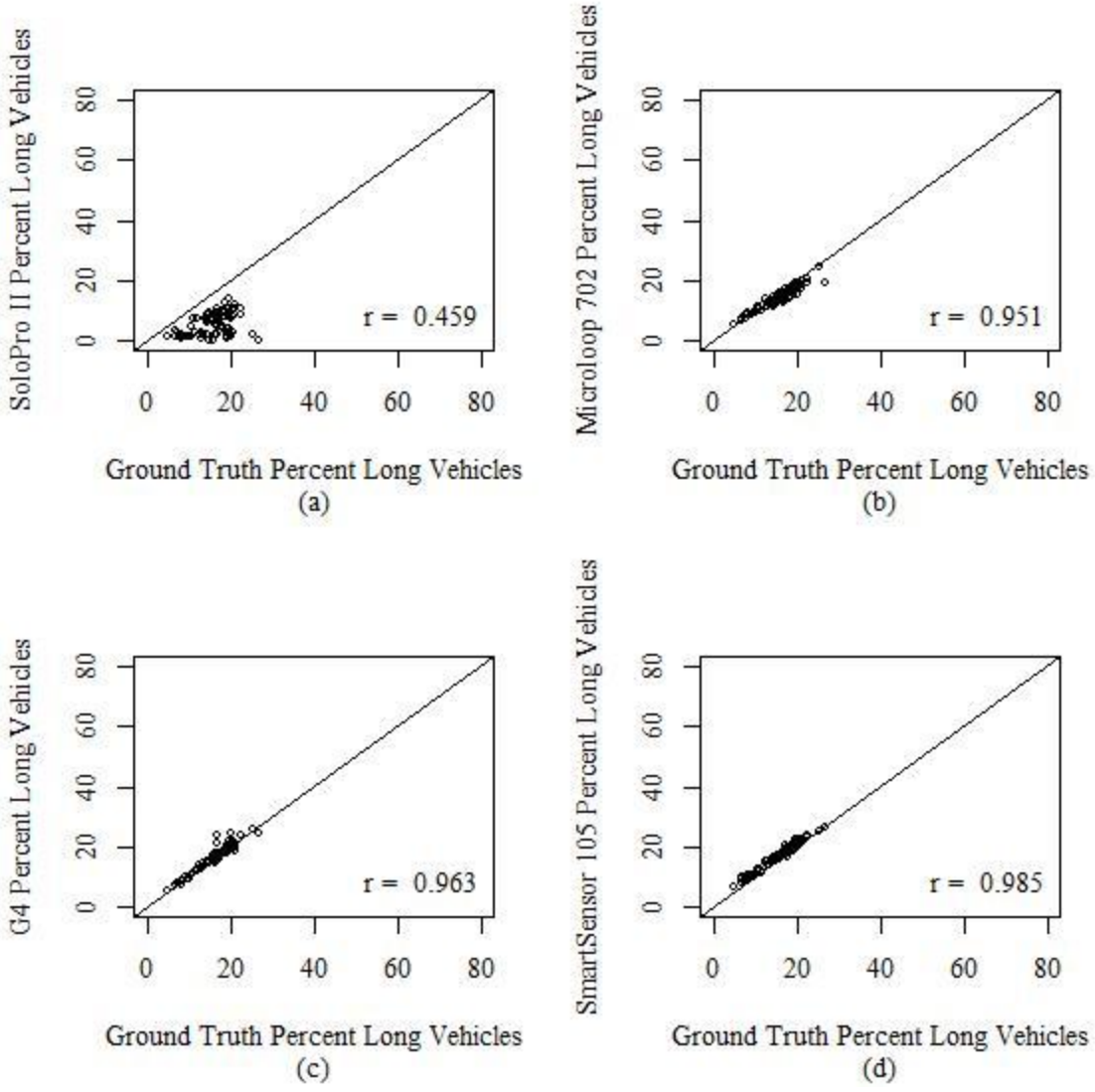

Figure E.42: Fifteen-Minute Percent Long Vehicles Scatter Plots Against Ground Truth for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d) Detectors

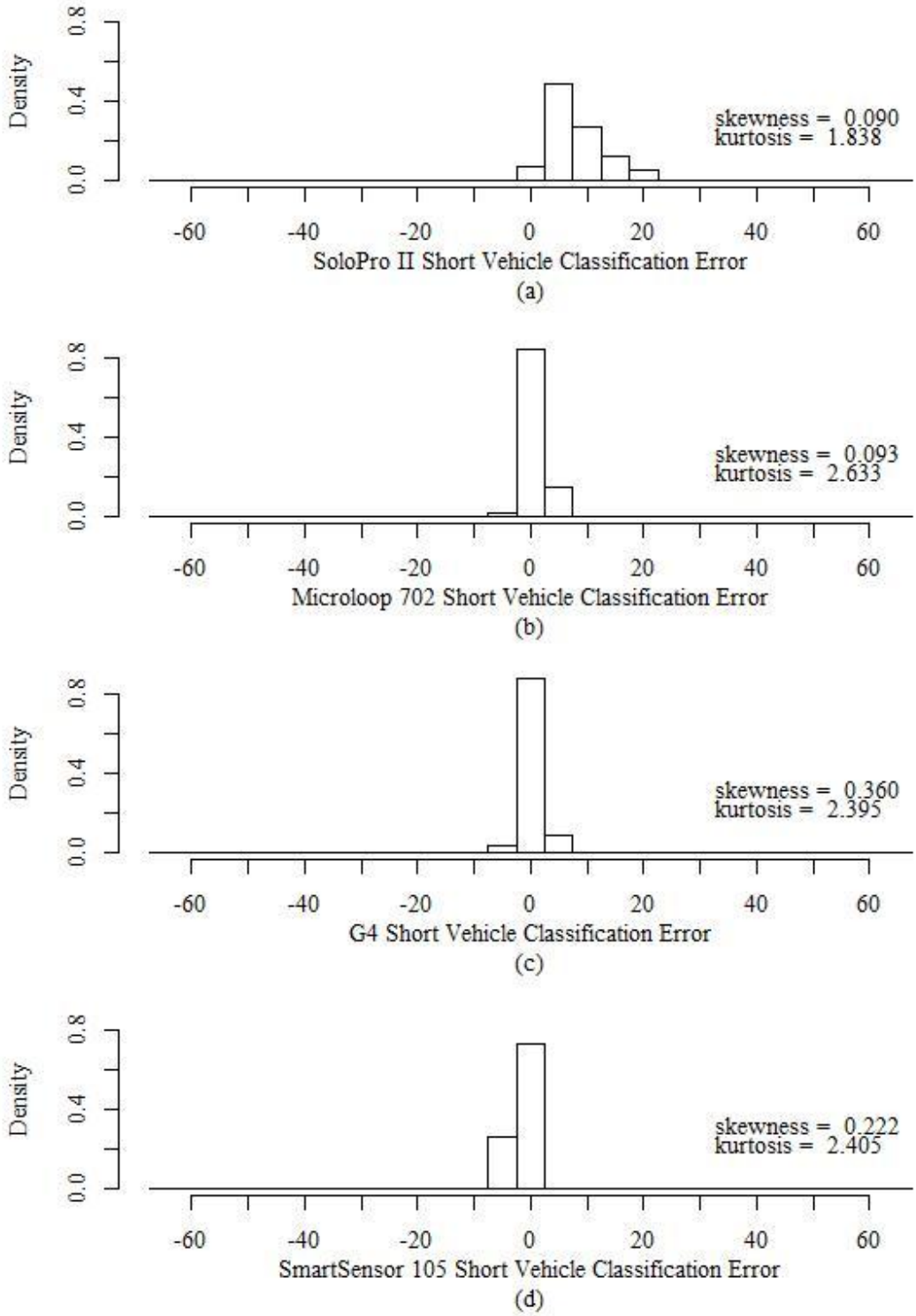

Figure E.43: Histograms of Fifteen-Minute Percent Short Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

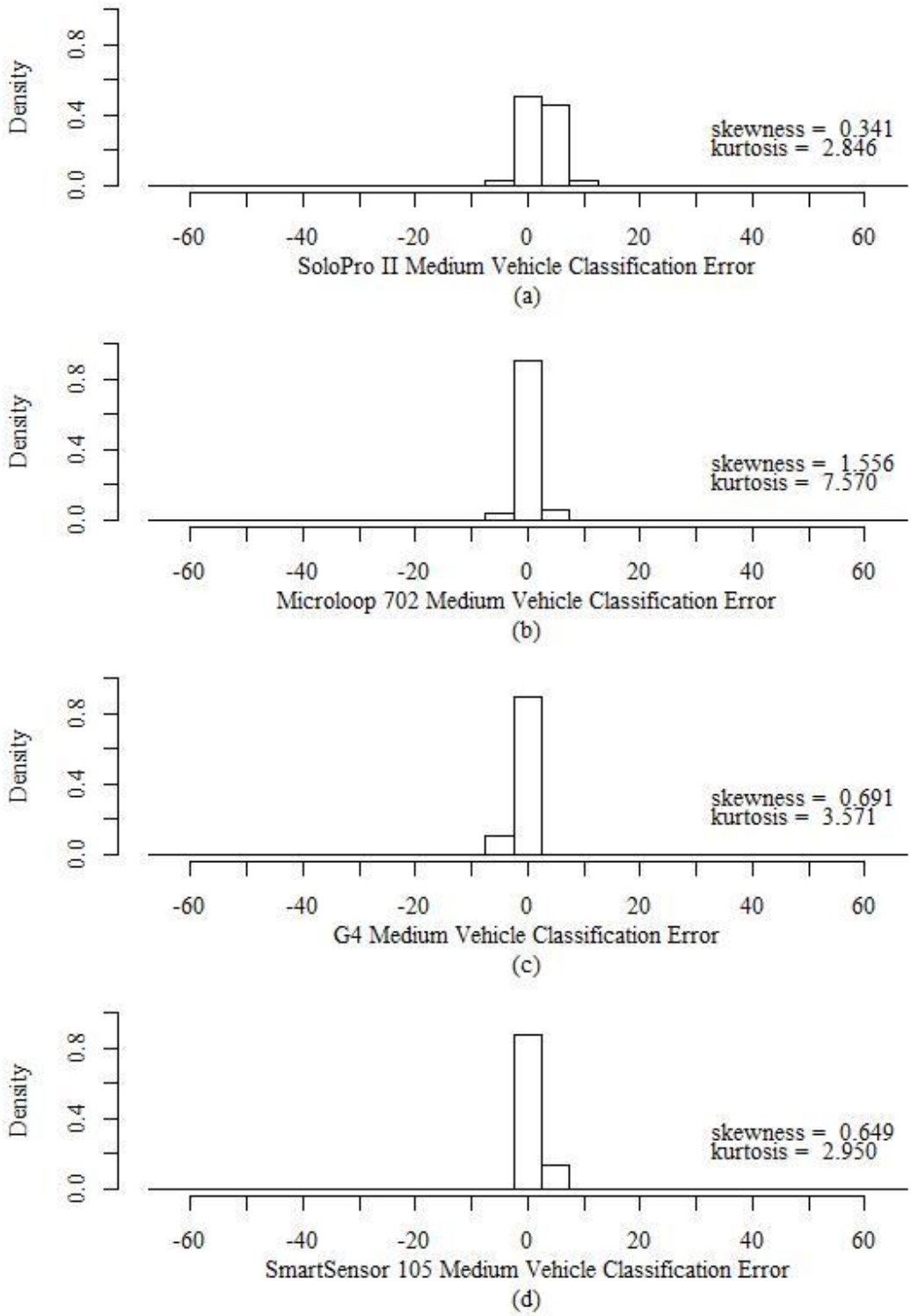

Figure E.44: Histograms of Fifteen-Minute Percent Medium Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

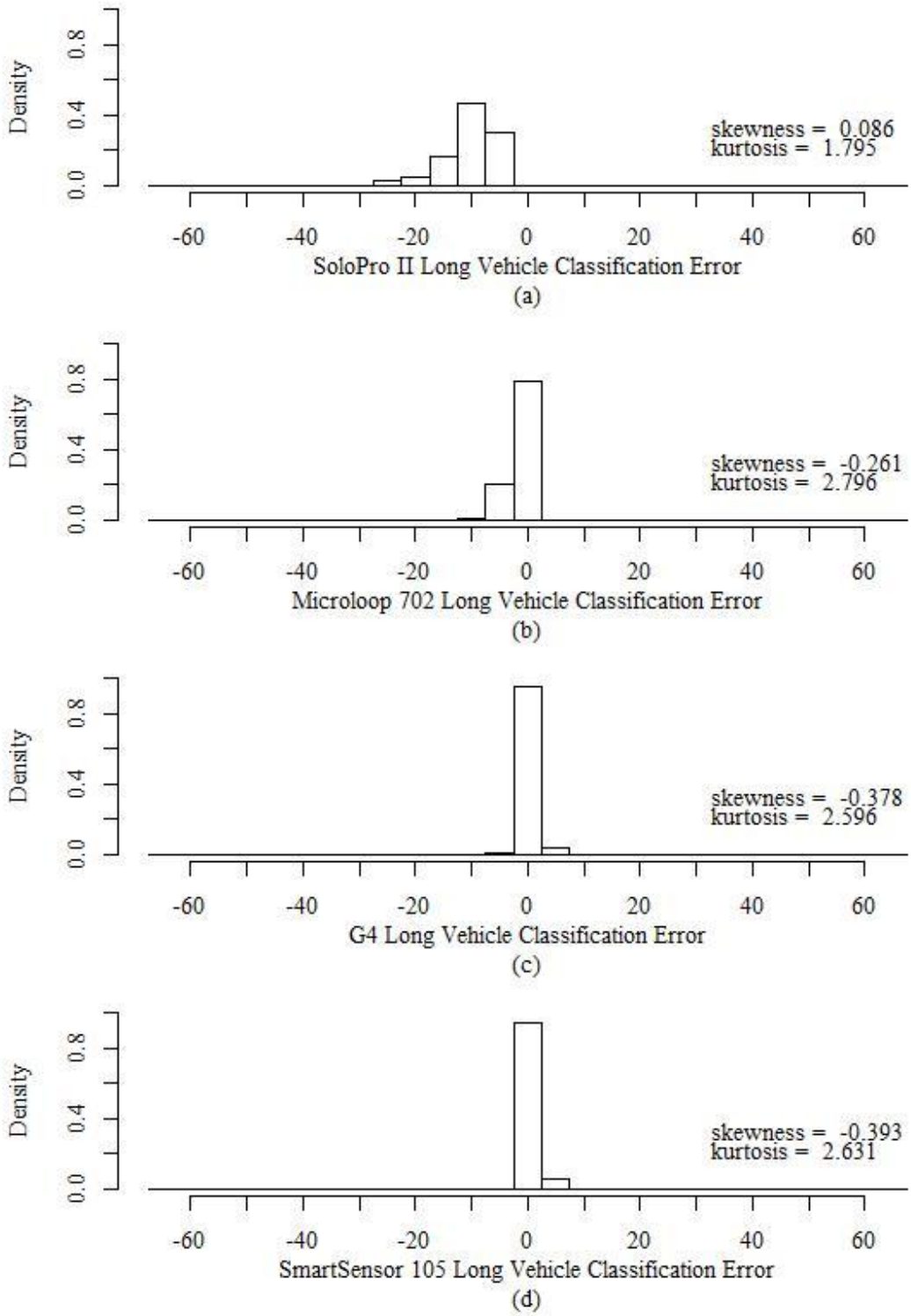

Figure E.45: Histograms of Fifteen-Minute Percent Long Vehicles Error Distributions for Solo Pro II (a), Microloop 702 (b), G4 (c), and SmartSensor 105 (d)

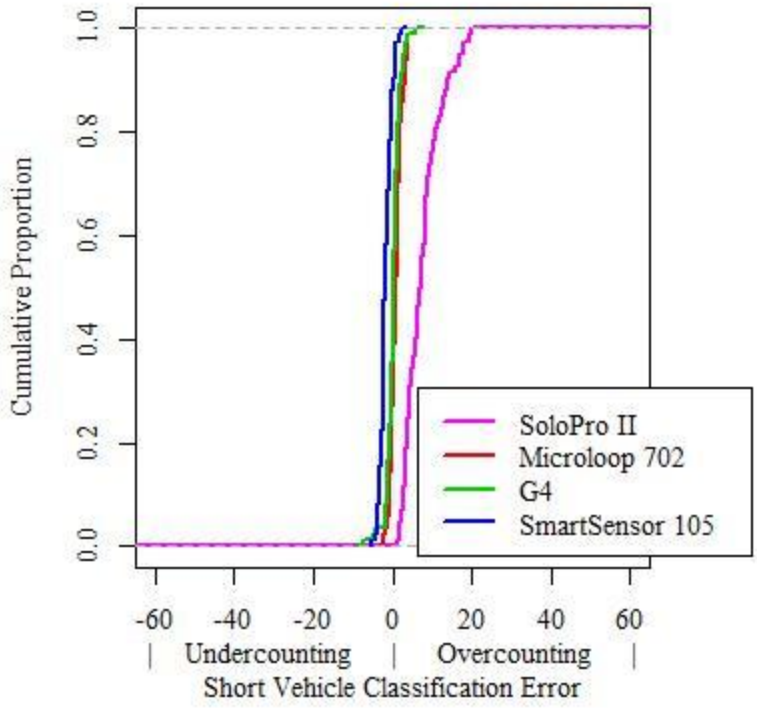

Figure E.46: Fifteen-Minute Percent Short Vehicles Error Cumulative Distribution Plot

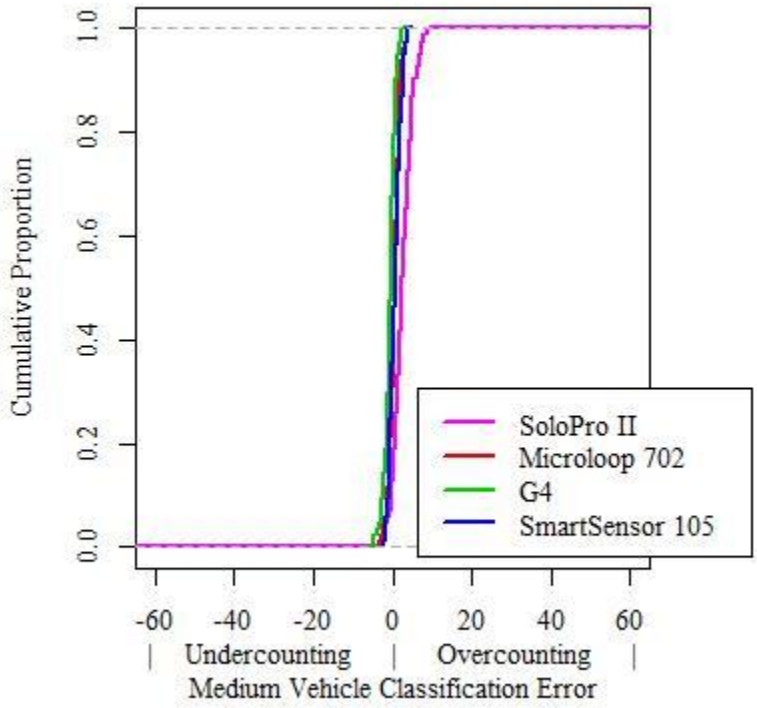

Figure E.47: Fifteen-Minute Percent Medium Vehicles Error Cumulative Distribution Plot

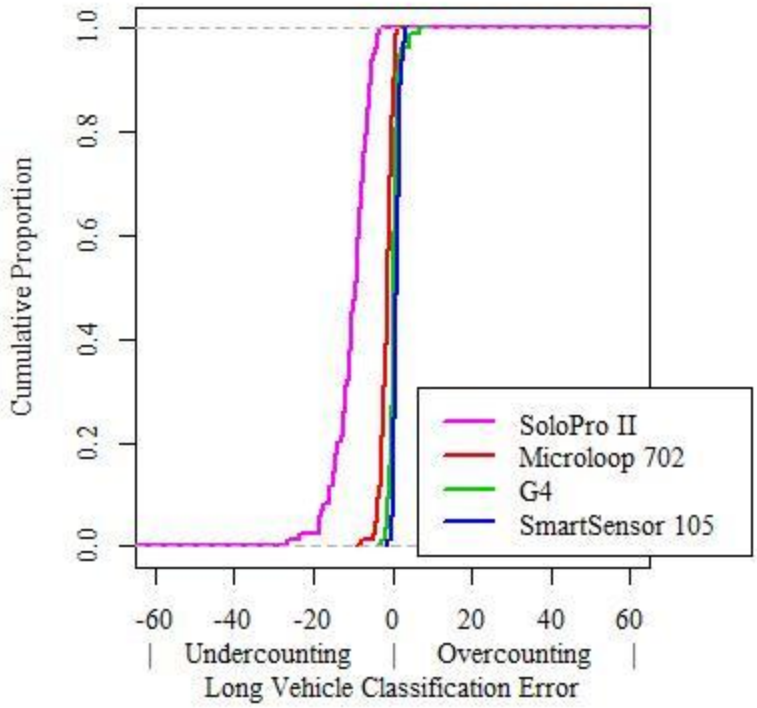

Figure E.48: Fifteen-Minute Percent Long Vehicles Error Cumulative Distribution Plot

Table E.8 Fifteen-Minute Classification Error Percentage Summary Statistics

	Mean	Median	Standard Deviation
Solo Pro II	10.4%	9.5%	4.41
Microloop 702	2.1%	1.9%	1.29
G4	1.6%	1.2%	1.28
SmartSensor 105	2.1%	2.1%	0.97