Feasibility of a Portable Weigh-in-Motion System for Axle Load Data Collection on Secondary Highways

By

Michael Olfert

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

In Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

Department of Civil Engineering

University of Manitoba

Winnipeg, Manitoba

ABSTRACT

This research investigates the feasibility of using a commercially available portable weigh-in-motion (WIM) system to improve the spatial representation of axle load data from trucks operating on secondary highways. The research evaluated the validity of axle load data collected from three on-road installations next to comparison load data sources, as well as considering the validity of raw and post-processed data.

Analysis revealed that portable WIM data cannot achieve the accuracy standard required of ASTM E1318 type II WIM data, but proper installation, calibration, and post-processing can allow portable WIM to have 95% of loads within ±36% of comparison loads for GVW, within ±57% for tandem axles, within ±87% for tridem axles, and within ±63% for individual axles.

Additionally, aggregation and post-processing of the data can allow selected vehicle statistics, such as fully loaded vehicles or GVW distributions, to achieve higher accuracies, potentially allowing for various indirect applications of the data.

ACKNOWLEDGEMENTS

First, I'd like to thank my advisor, Dr. Jonathan Regehr, P.Eng, for inviting me to join the team at UMTIG, fuelling my interest in transportation engineering, and giving me opportunities to conduct and present interesting work. Your guidance has been invaluable; I couldn't have asked for a better advisor.

I'd also like to extend my thanks to Giuseppe, Karalee, Abby, Karen, and the rest of the UMTIG family. I've enjoyed our roundabouts, conference trips, ITE events, and long days assembling MHTIS reports together; thanks for your support and friendship.

I will further thank all those whose support made this project, and my degree, possible. This includes financial support from the Natural Sciences Engineering Research Council of Canada and the International Society for Weigh-In-Motion and assistance with data collection from Manitoba Infrastructure. Special thanks to Kevin Gavin and Bofei Sun at Traffic Engineering, as well as Tim Brown and the personnel at Motor Carrier Enforcement. Additional thanks to Roy Czinku at International Road Dynamics for sharing from your extensive knowledge of WIM.

Last, I'd like to thank my family. Thanks to my parents, who have supported my engineering education, in addition to everything else, since day one. Finally, thank you to my wife, Rebecca, for your love, patience, feigned interest in my work, and encouragement. You're the best.

Contents

List of Figures

List of Tables

1 INTRODUCTION

1.1 PURPOSE

This research investigates the feasibility of using a portable weigh-in-motion (WIM) system to achieve greater spatial coverage of axle load data from trucks on secondary highways. The portable WIM system tested consists of commercially available components. In this investigation, the system's data is evaluated according to its accuracy, precision, and potential for increased data quality through various post-processing methods and analyses. By determining the system's capabilities in these areas, this research evaluates the feasibility of collecting load data that is useful for design and other applications and gives recommendations for a potential portable WIM data collection program.

1.2 BACKGROUND AND NEED

Professionals require vehicle axle load data for a variety of engineering and planning applications, including design and analysis of highway structures and freight planning studies. To be useful for these applications, this load data must be sufficient in four principal data quality dimensions:

1. Data Validity: The data must be both precise and accurate.
2. Temporal Coverage: The data must be representative of the patterns of vehicle loads in time.
3. Spatial Coverage: The data must be representative of the geographic areas under analysis.

4. Data Availability: It must be possible to collect the data for a reasonable level of effort and financial investment (Wood, 2017).

The only technology that sufficiently meets these criteria, particularly the requirement of data availability, is weigh-in-motion (WIM). High-speed WIM systems, which are capable of obtaining axle loads as vehicles pass over them at highway speeds, are able to achieve excellent temporal coverage and acceptable data validity. However, given their high costs and requirements for regular calibration, WIM stations can only be installed at a limited number of locations in any jurisdiction, limiting the spatial coverage of the data (Zhang, Haas and Tighe, 2007; Refai, Othman and Tafish, 2014).

While vehicle load data is useful for a variety of applications, the current state of the practice in the application of axle load data is defined with specific regard to the use of axle load data for highway design in the Mechanistic Empirical Pavement Design Guide (MEPDG). The MEPDG, developed by the American Association of State Highway Transportation Officials (AASHTO), is a tool for improving the design and analysis of pavement structures. The MEPDG requires load data to be input to the analysis in the form of axle load spectra (ALS) for each axle type, which can only be obtained from WIM sensors (AASHTO, 2015).

Due to the difficulties with achieving sufficient spatial coverage of axle load data using WIM systems, the MEPDG specifies three possible levels of data quality for any given highway segment:

- Level 1: Data is available from a nearby WIM site,
- Level 2: Data is available from a WIM site on a similar segment in the same region, and
- Level 3: No site specific data is available; default loading values included in the MEPDG software are used (AASHTO, 2015).

Level 1 data is not available in most locations, and level 2 and 3 data cannot always be accepted as valid, particularly in regions outside of the United States, where the MEPDG's default values, which are based on typical traffic patterns in the United States, may not apply (Swan *et al.*, 2008). Because of this, various technologies for improving the spatial coverage of axle load data have been and are being developed. These technologies include portable WIM systems (Refai, Othman and Tafish, 2014; Petersen, 2015), bridge WIM (BWIM) systems (Lydon *et al.*, 2016), on-board weighing (OBW) systems (Labarrere, 2018b), and vehicle signature matching (Jeng and Chu, 2015). Additionally, due to patterns in the trucking industry, truck loading characteristics can be correlated spatially and temporally to factors such as road regulations, patterns of industry, and climatic variables (Haider *et al.*, 2011; Reimer and Regehr, 2013), and efforts have been made to model truck axle loads based on these variables. This approach would allow more accurate application of level 2 and 3 data over areas for which the variables are known.

In order to generate more level 1 data, some agencies have used portable WIM systems as a way of collecting data over a wider spatial coverage. However, due to a number of issues that stem from the portability of the systems, portable WIM systems have not been able to generate data of similar validity to that of permanent WIM systems (Selezneva and Von Quintus, 2014). Several state DOTs have performed research into how to improve portable WIM performance throughout the 2010s (Refai, Othman and Tafish, 2014; Petersen, 2015; Faruk *et al.*, 2016). Five state DOTs have taken an additional approach to the data validity issues of portable WIM by applying portable WIM data indirectly to design. This approach typically involves using portable WIM data from a site to select a representative axle load spectra from a predetermined list (Selezneva and Von Quintus, 2014). To do this, the individual vehicle records are not considered to be correct; rather, only certain parameters of the data are taken to be accurate and used to match the portable WIM load spectra to load spectra from the predetermined list.

Though some efforts have been made to improve the spatial coverage of axle load data, many jurisdictions still have inadequate coverage. Furthermore, the MEPDG's default axle load values are based on the context of the United States and are less applicable in other countries due to differences in road regulations, distribution of truck classifications, and climate. These areas stand to benefit most from the validation and application of a portable WIM system.

1.3 OBJECTIVES AND SCOPE

The objectives of this research are as follows:

1. Determine if a commercially available portable WIM system is capable of obtaining axle load data of sufficient validity to be used directly for engineering and planning applications in Canada.
2. Determine if a commercially available portable WIM system is capable of obtaining axle load data that can be applied indirectly to engineering and planning applications in Canada.
3. Generate recommendations for installations of a portable WIM system.

The scope of this research was constrained to the following:

- This research examines the performance of only a single portable WIM system consisting of components selected by Manitoba Infrastructure.
- This research only analyzes the data collected from 3 portable WIM installations. All three of these installations were performed in southern Manitoba.
- This research does not consider the data requirements of inputs to pavement design other than the accuracies outlined in ASTM E1318: Standard Specification for Highway Weigh-In-Motion (WIM) Systems with User Requirements and Test Methods.
- This research is not intended to provide all the necessary information for the implementation of a portable WIM data collection program. It is intended only to discuss best installation practices, examine the data collection capabilities of a portable WIM system, and give those

recommendations about the formation of a data collection program that can be inferred directly from these capabilities.

1.4 APPROACH

This research examines the capabilities of a portable WIM system consisting of commercially available components. To do so, the accuracy of the portable WIM system is calculated in reference to two other sources of axle load data: 1) a permanent piezo-quartz WIM system, and 2) a static truck weigh scale. The portable WIM system is installed alongside each of these sources so that individual truck records may be paired between the data sources in order to compare the axle loads.

The analysis of the axle loads is done by calculating the accuracies of the portable WIM system's axle load data as compared to axle load data from both a static weigh scale and a permanent WIM system. The static weigh scale's axle load data are considered 'ground truth'—that is, data that is as accurate to the true axle loads as is possible to obtain, while the permanent WIM system's data is included due to its high-quality axle load data at an greater sample size. The accuracy standards as stated in ASTM E1318 are considered as a benchmark. The analysis considers the validity of the portable WIM data on both a per-vehicle and aggregated basis. The per-vehicle analysis examines the accuracies of individual vehicles axles, axle groups, and gross vehicle weights (GVWs), while the aggregated data analysis examines the relative parameters of axle load spectra between the portable WIM and comparison datasets.

To determine the maximum validity that can be achieved by the portable WIM system, the datasets are examined in the state they are collected in, as well as after two post-processing methods. One of these corrects for any relevant effects from temperature and vehicle speed, while the other is an autocalibration procedure that continuously corrects the data based on pre-set values.

1.5 THESIS ORGANIZATION

This thesis is organized into five chapters:

Chapter 2 – Environmental Scan: This chapter presents background information on the need for obtaining axle load data over a wider spatial representation than is currently available in Canada. It then evaluates several methods of obtaining axle load data over that wider spatial representation and includes a detailed discussion of past studies and implementation of portable WIM.

Chapter 3 – Research Methodology: This chapter provides a detailed description of the portable WIM system used to conduct the research, the methods and locations of its installation, and methods of data analysis that were applied.

Chapter 4 – Analysis: This chapter presents the results of the data analyses and evaluates the validity of the collected portable WIM data on both a per-vehicle and aggregated basis.

Chapter 5 – Conclusions and Recommendations: This chapter presents the conclusions drawn from this research, including recommendations for a potential portable WIM data collection.

1.6 TERMINOLOGY

Accuracy – The degree of agreement between a measured value and an accepted reference value.

Autocalibration – A method of calibrating a WIM system wherein the applied calibration factor is continuously updated based on the factor required to match a selection of the measured axle loads to an accepted reference value. Autocalibration is applied in this research as a post-processing method.

Automatic Vehicle Classifier (AVC) – A type of permanent count station that provides vehicle classification on the basis of axle spacing, but does not provide axle load data.

Axle Group – A defined group of adjacent axles. Includes steering axles, single axles, tandem axles, and tridem axles.

Axle Load – The sum of all tire loads of the wheels on an axle.

Axle Load Spectra (ALS) – A method of representing axle load data using histograms of loads of single, tandem, and tridem axles and GVWs, often separated by vehicle configuration. A primary input for mechanistic-empirical design.

Bridge Weigh-in-Motion (BWIM) – A WIM system that calculates vehicle axle loads by instrumenting a bridge with sensors to measure deflection of the structure.

Calibration – The adjustment of a WIM system's settings to produce valid loads. Typically done using a test truck of known axle loads to generate a calibration factor and dynamic compensation factor.

Calibration factor – The factor generated by a calibration process that is applied to all axles to correct the measured values to the actual values.

Data Quality – The measure of how well data represents actual conditions and can be reasonably obtained. Includes validity, temporal coverage, spatial coverage, and data availability.

Dynamic Compensation Factor – The factor generated by a calibration process that is applied only to all front axles to correct for dynamic effects on the front axle.

Front Axle – A vehicle's articulated steering axle; almost always a single axle group.

Gaussian Mixture Model (GMM) – An analysis method wherein an uneven distribution is separated into a specified number of 'mixing components', with each component having the form of a gaussian (normal) distribution.

Gross Vehicle Weight (GVW) – The combined loads of all tires on all axles of a vehicle.

Individual Axle – One axle, either in an axle group by itself or as part of a tandem or tridem axle group.

Post-Processing Calibration – A calibration procedure developed for this research wherein a physical test truck is replaced with paired vehicle records from the nearby static weigh scale.

Mechanistic-Empirical (ME) – A pavement design approach involving calculation of pavement responses (mechanistic) and predicting incremental damage over time (empirical). Uses axle load spectra as inputs.

Precision – The consistency, or stability, of a set of results.

Selected Configurations – For the purposes of this research, 3-S2, 3-S3, 3-S3-S2, and 3-S2-4 vehicles, which are isolated for specific analysis.

Tandem Axle Group – An axle group consisting of two axles. Also referred to as a ‘tandem axle’.

Tridem Axle Group – An axle group consisting of three axles. Also referred to as a ‘tridem axle’

Validity – A measure of data quality that indicates a value was measured as it was intended to be. Validity includes elements of both accuracy and precision.

Weigh-in-Motion (WIM) – The process of estimating a vehicle’s axle loads, axle group loads, and gross vehicle weight by measuring and analyzing dynamic tire forces.

2 ENVIRONMENTAL SCAN

This chapter discusses weigh-in-motion (WIM) technology, including its applications and common issues with the quality of WIM data. The chapter includes a summary of the current state of axle load data collection in Canada, including both WIM and other sources. The chapter closes with a review of research into methods of obtaining axle load data for broader spatial representation.

2.1 COLLECTION AND USE OF AXLE LOAD DATA

WIM systems have been used around the world for several decades to collect detailed axle load data for a variety of applications. Most WIM systems are permanent, meaning that they are installed in the road at a particular location, allowing high quality data to be collected, but only at a single point in a highway network (van Loo and Znidaric, 2019). All permanent, in-road WIM systems all share several common major components:

- Scales or sensors installed in the roadway which measure the loads of passing vehicles,
- An electronic control system equipped with an algorithm that converts the signal from the sensors to axle load measurements, and
- Support equipment including power sources and communication devices to transmit the load measurements off-site.

A variety of sensor technologies can be employed in these WIM systems. The FHWA identifies five primary WIM sensor technologies (U.S. Department of Transportation, 2018a):

- Piezoelectric sensors consist of a copper strand encased in a piezoelectric polymer material, all enclosed in a flexible brass sheath. Piezoelectric sensors measure axle loads by generating a voltage charge proportional to the pressure applied on the sensor, which is then translated into a load measurement. Piezoelectric sensors are sensitive to temperature changes, which can change the loads measured by the sensors.
- Piezo-quartz sensors are strip sensors that function much the same as piezoelectric sensors but use quartz crystal technology to translate vertical force into an electric charge. Piezo-quartz sensors are less sensitive to temperature changes than piezoelectric sensors and typically produce more accurate results.
- Bending plates are wide sensors (typically 50 cm wide) that have strain gauges mounted to a semi-rigid plate's underside to measure the bend in the plate as a vehicle passes over the sensor and translate that bend to an axle load. They are one of the most accurate WIM sensor technologies available.
- Load cells are wide plate sensors (typically 90 cm wide) that have two transducer sensors underneath the right and left sides of the rigid plate. Loads on the sensor surface are transferred to the transducers which

generate an electric signal that can be translated into a load measurement. Load cells are the most accurate WIM sensor technology, but also the most expensive and so are used only in selective applications.

- Strain gauge strip sensors are strip sensors that use strain gauges directly to generate electric charges to translate to axle loads, rather than mounting them on a plate as in the bending plate sensor. Strain gauge strip sensors are roughly comparable with piezo-quartz sensors in the accuracy of load measurements that they offer.

Axle load data collected by WIM systems can be used for a variety of applications. The following list of applications is not comprehensive, but provides an overview of some of the primary areas where WIM data is used.

- Pavement design is a common application for axle load data. By providing axle load data at known points, axle loads on new roads or at locations of road improvements can be predicted, and thus the road can be designed to avoid premature failure and increased maintenance costs while avoiding overdesign (Turochy, Timm and Tisdale, 2005). The two primary accepted methods of doing this are the ESAL method, wherein all vehicle weights are converted to an equivalent number of standardized single axles (U.S. Department of Transportation, 2018a), and the mechanistic-empirical approach described in the Mechanistic-Empirical Pavement Design Guide (MEPDG) (Haider *et al.*, 2011; Nassiri, Bayat and Kilburn, 2014), which is a software-assisted iterative approach that combines a variety of factors to predict a pavement’s response to a particular loading. The M-E approach requires axle load spectra separated by axle group type (single, tandem, tridem) as inputs (AASHTO, 2015).

- Bridge design also requires axle load data. Data collected from WIM systems can be used to calibrate design codes that govern bridge design to ensure both the cost-effectiveness and safety of the final designs (Van de Lindt *et al.*, 2002; Chotickai and Bowman, 2006; Helmi, Bakht and Mufti, 2014). Long-span bridges in particular require precise inputs to ensure safe designs and so benefit from high-quality axle load data (van Loo and Znidaric, 2019). Furthermore, bridge weigh-in-motion (BWIM), which measures axle loads by instrumenting a bridge with various types of sensors, has more recently contributed to bridge design and evaluation efforts (Cantero and González, 2015; Lydon *et al.*, 2016).
- A third use for axle load data is for freight planning studies. These studies examine the movement of goods with the goal of developing policy and regulations relating to freight and generating recommendations for investment in road and rail infrastructure (Krisztin, 2017). These studies typically examine the movement of freight on key highway corridors or between regions; in them, vehicle load data can be used to estimate freight volumes (Koniditsiotis, Buckmaster and Fraser, 1995). Therefore, policy decisions based on these studies are better informed when axle

load data is available in more locations and at higher quality (van Loo and Znidaric, 2019).

- The final use of axle load data and WIM systems discussed here is in the enforcement of vehicle weight regulations. Enforcement is important to maintaining and managing a highway network, and WIM systems are often used to assist with these tasks. WIM systems can be used in two ways here: direct and indirect enforcement. Direct enforcement allows a citation issued to a driver or company based only on the load measurements from a WIM system (Richardson *et al.*, 2014). This allows complete temporal coverage of an enforcement location. However, direct enforcement from WIM has only been used in a limited capacity due to the requirement for the system to be certified to meet stringent accuracy standards (Jacob and Cottineau, 2016). More common is indirect enforcement from WIM, wherein vehicles pass a WIM system before coming to a static weigh scale, which is the typical technology for enforcement. The WIM system is able to identify only those vehicles that are near or over gross vehicle and/or axle load limits, and can be connected to a system to then instruct these drivers to report to the scale for an accurate and precise weight measurement and possible citation (Žnidarič, Kalin and Kreslin, 2018).

2.2 ISSUES WITH WEIGH-IN-MOTION SYSTEMS

When using data collected from WIM systems, there are a number of different issues that may cause the data to be inaccurate or otherwise unreliable. As WIM sensors measure weight dynamically, axle load data collected from WIM systems are never completely accurate as compared to static scale load measurements, and their error can only be reduced through careful calibration. The primary standard for WIM accuracy in North America is ASTM E1318 (ASTM International, 2009); the primary European standard is COST 323 (Jacob, O'Brien and Jehaes, 2002). The standards' required accuracies for axle loads, axle group loads, and GVW are summarized in Table 1. Both standards also state required accuracies for axle groups and individual axles, which require higher accuracies than GVW for the same application and class.

Table 1 – WIM Accuracy Standards

Application	ASTM E1318				COST 323				
	Class	Allowed Error			Class	Allowed Error			
		GVW	Axle Group	Axle Load		GVW	Axle Group	Single Axle Load	Axle within Group
Direct enforcement	Type IV*	±1100 kg	±500 kg	±200 kg	A	±5%	±7%	±8%	±10%
					B+	±7%	±10%	±11%	±14%
Indirect enforcement	Type III	±6%	±10%	±15%	B	±10%	±13%	±15%	±20%
					C	±15%	±15%	±20%	±25%
Infrastructure design	Type I	±10%	±15%	±20%	B	±10%	±13%	±15%	±20%
	Type II	±15%	±20%	±30%	C	±15%	±15%	±20%	±25%
Statistics	Type I	±10%	±15%	±20%	D+	±20%	±23%	±25%	±30%
	Type II	±15%	±20%	±30%	D	±25%	±28%	±30%	±35%
Rough Statistics	N/A	N/A	N/A	N/A	E	>25%	>28%	>30%	>35%

* ASTM WIM type IV has not yet been approved for use in the United States, and is a specification for low speed (3-16 km/h) WIM for direct enforcement. The allowed error is specified in load, with the tolerances applying if the axle is greater than 5400 kg, the axle group is greater than 11 300 kg, or the GVW is greater than 27 200 kg.

To fall within a type defined by ASTM E1318, 95% of results must be within the stated tolerance. COST 323 does not define a strict percentage of results that must be within the stated tolerances, but rather allows a user to set this percentage based on the specific requirements of the site and data. Both standards prescribe a method for calibrating the systems. In this method, a test truck of known weight is to be driven over the sensors a minimum number of times to ensure that the weight measured by the WIM system matches the truck's known weight. An uncalibrated site will not be able to achieve acceptable accuracy and will not deliver meaningful data.

Even with proper calibration, a WIM system may be unable to achieve acceptable accuracy due to several factors in the site. A WIM site's accuracy can be affected by pavement type, smoothness, and condition, longitudinal and lateral alignment of the road, and traffic composition and characteristics. If one or more of these factors creates issues for the WIM system, a given site may never be able to reach acceptable accuracy. Additionally, an improperly installed WIM system will also experience issues, necessitating proper installation by qualified personnel (U.S. Department of Transportation, 2018b). Furthermore, to ensure ongoing accuracy of a WIM system, regular calibration must take place using the test truck method, as a WIM site will drift in its calibration due to environmental factors. An annual calibration is recommended by the United States' Federal Highway Administration (U.S. Department of Transportation, 2018c); independent research in Canada has shown this to be an appropriate calibration schedule (Wood, 2017).

An additional issue with WIM systems is that of temperature sensitivity. Piezoelectric WIM sensors, which are used by many jurisdictions, will generally give axle load measurements that get larger as pavement temperature increases (Gajda *et al.*, 2013). Some WIM equipment manufacturers correct for this effect by applying automatic temperature-based correction factors to WIM data when temperature is recorded at a WIM site; while this can improve the accuracy of the system's output, better accuracies can be achieved by developing site-specific temperature correction factors (Burnos and Gajda, 2016; Wood, 2017).

A final issue with WIM systems is that of cost and logistics. WIM units are expensive to install and maintain, and so jurisdictions are typically limited in terms of the number of systems they can deploy (Hallenbeck and Weinblatt, 2004). This sometimes leads to the absence of representative load data for certain geographic areas or highways. To deal with this, the MEPDG specified a method of assigning data to any given highway segment, resulting in three possible levels of data quality for a given highway location:

- Level 1: Data is available from nearby WIM sites,
- Level 2: Data can be applied from a WIM site on a similar segment in the same region, and
- Level 3: No site specific data is available so default loading values included in the MEPDG software must be used (AASHTO, 2015).

This system is an imperfect solution to the problem of spatial representation, but is recognized as an industry standard approach.

2.3 EXISTING SOURCES OF AXLE LOAD DATA IN CANADA

Canada consists of ten provinces and three territories, all but one of which collect axle load data in some capacity. These sources of load data, as are described in publicly available sources, are summarized in Table 2. In addition, the number of weigh scale locations listed by Allstays in 2020 are included (Allstays LLC, 2020). Allstays is a crowdsourced online resource that includes data for those in the trucking industry. This data is generally close to the officially reported numbers of weigh scales but does differ in some provinces; this can be attributed to the fact that some of the official sources are dated, and that Allstays sometimes lists multiple weigh scales at a single location due to directionality. Sources of official information are described below Table 2. A value of N/A indicates that no publicly available information was available.

Table 2 – Sources of axle load data in Canada

Jurisdiction	WIM Sites	Static Weigh Scale Stations		Static Weigh Scale Mobile Units
		Official Sources	Allstays	
British Columbia	8	22	14	≥1
Alberta	6	17	18	24
Saskatchewan	15	10	12	≥1
Manitoba	7	14	8	≥1
Ontario	4	32	46	N/A
Quebec	19	≥1	19	N/A
New Brunswick	4	4	7	N/A
Prince Edward Island	5	2	3	1
Nova Scotia	≥3	5	2	≥1
Newfoundland and Labrador	4	4	8	N/A
Yukon	N/A	4	N/A	N/A
Northwest Territories	1	2	N/A	N/A
Nunavut	N/A	0	N/A	0

As of 2019, the province of Alberta collected load data from 7 permanent WIM sites covering a total of 24 lanes of traffic (Poapst, 2020). The sites each have a calibration verification performed every two months (Transportation Modelling and Analysis Section, 2014). In addition, as of 2019 the province operates 17 manned static weigh scale stations and 24 unmanned static weigh scales mobile units that are used for vehicle weight regulation enforcement, rather than for data collection (Government of Alberta, 2020).

The province of British Columbia, as part of its traffic data program, has 5 WIM sites listed on its GIS application out of a total of 120 permanent traffic count stations. The GIS application provides traffic data up until 2019, however the WIM sites provide data only up to 2013, or earlier for some sites (Government of British Columbia, 2019). The province operates a further 3 WIM sites as part of its commercial vehicle inspection program; these sites weigh vehicles as part of a pre-screening for more detailed vehicle inspections (Government of British Columbia, no date). A further 22 sites are equipped with static weigh scales, and an unspecified number of mobile units assist with vehicle weigh regulation enforcement (B.C. Ministry of Transportation and Infrastructure, no date).

The province of Manitoba operates 6 WIM sites as part of its traffic monitoring program, however not all of these sites monitor all lanes of the highway. Over all 6 sites, a total of 11 lanes have WIM sensors installed. Three of these WIM sites are equipped with piezoelectric sensors and three are equipped with piezoquartz sensors (Sun, 2020); all are calibrated approximately annually (Wood, 2017). In addition to traffic monitoring with WIM, the province operates 14 permanent static weigh scale sites and 5 temporary static weigh scale sites for vehicle weight regulation enforcement (Manitoba Infrastructure, 2017).

The province of New Brunswick operates 4 WIM sites; these sites are primarily used for pre-screening commercial vehicles at inspection stations, which are located near each of the WIM sites (New Brunswick Department of Transportation and Infrastructure, no date). The data from these sites is also stored for later use in engineering applications (Hanson, Klashinsky and McGibney, 2010).

The province of Newfoundland and Labrador had 4 WIM sites installed to be used for pre-screening of commercial vehicles at inspection stations in 2007 (Hanson, Klashinsky and McGibney, 2010) and used them for this purpose until 2013 when the use for pre-screening stopped and the systems were used only to collect vehicle load and speed data (Fitzpatrick, 2016). This resulted in the province’s Department of Transportation and Works ending the contract for the use of the WIM systems in 2015 (Auditor General of Newfoundland and Labrador, 2012).

The Northwest Territories have one WIM site that is used both for enforcement of heavy vehicle tolls and for planning purposes (Thom, 2020), and had 2 static weigh scales for vehicle weight enforcement as of 2015 (Northwest Territories Department of Infrastructure, 2017).

The province of Nova Scotia uses WIM systems as part of its traffic count program. It has 36 permanent count stations, some of which are WIM systems using piezo-electric sensors, but does not provide a precise number of WIM stations in use (Nova Scotia Department of Transportation and Infrastructure Renewal, 2017). The province does have three WIM stations used as a pre-screening tool for vehicle weight enforcement (Hanson, Klashinsky and McGibney, 2010); these stations are in addition to 5 static weigh scale stations and an unspecified number of mobile scale units (Province of Nova Scotia, 2013).

The territory of Nunavut does not currently employ any static weigh scales for vehicle monitoring or enforcement (Saskatchewan Highway Traffic Board, 2019); no sources report any use of WIM in the territory.

The province of Ontario uses WIM at 4 truck inspection stations, where sensors are used to pre-screen trucks as they enter the station for inspection. WIM data is not used for vehicle data collection on provincial highways (Sureshan, 2020) as WIM data from the province has historically not been accurate (Swan *et al.*, 2008). The primary source of commercial vehicle load data in the province is from an annual Commercial Vehicle Survey (CVS), which is taken at over 200 locations across the province approximately every 5 years since 1967 (Ontario Ministry of Transportation, 2012). Vehicle weight regulation enforcement is performed at 32 static weigh scale locations (Government of Ontario, 2020); there is no reported use of mobile scale units.

The province of Prince Edward Island operated 5 WIM stations for vehicle load data collection as of 2010. At that point, vehicle weight enforcement was performed at two static weigh scale stations and one mobile scale unit (Prince Edward Island Department for Transportation and Infrastructure Renewal, 2009); in 2011, one of the static scales had a WIM station installed for pre-screening of commercial vehicles at a static weigh scale station (Evans and Klashinsky, 2012).

The province of Quebec operates 19 WIM stations as of 2019. These stations use piezoelectric sensors and collect vehicle load data for use in road design (Laplante, 2020). Static weigh scale and mobile scale numbers for the province are not reported.

The province of Saskatchewan operated a total of 15 WIM stations in 2016; these are used for vehicle load data collection (Traffic Services Saskatchewan Ministry of Highways and Infrastructure, 2016). Additionally, the province has performed several trials with a portable WIM system from 2012-2015 in an effort to substitute its existing program of low-speed WIM surveys with a program of high-speed portable WIM surveys. This system used piezo-quartz sensors installed with a pocket tape enclosure on the road surface. The system has not been verified for accuracy and as such there is no formal plan to substitute the portable WIM system for the low-speed surveys (Jaworski, 2018). It currently has 10 static weigh scale locations and an unspecified number of mobile scale units that it uses to enforce vehicle weight regulations (Government of Saskatchewan, 2020). The province is targeting 900 mobile weight checks in the 2019-20 fiscal year (Saskatchewan Ministry of Highways and Infrastructure, 2019).

The territory of Yukon has no reported use of WIM systems, though the government is scheduled to acquire and install one between 2026 and 2028 (Task Force on Vehicle Weights and Dimensions Policy, 2019). The territory currently operates four static weigh scale stations for vehicle weight restriction enforcement, with no reported mobile scale units (Government of Yukon, 2020).

Overall, the thirteen provinces and territories in Canada have disparate approaches to the collection of axle load data, which has led to widely varying levels of axle load data spatial coverage. While some provinces have higher levels of WIM and static scale coverage, many provinces have extensive highway network with relatively low coverage from WIM systems, resulting in axle load data that is frequently inadequate for planning and design purposes.

2.4 METHODS FOR GATHERING SPATIALLY REPRESENTATIVE AXLE LOAD DATA

Due to their cost and the logistical issues that come with maintaining a network of WIM systems, many jurisdictions do not have extensive networks of WIM systems. This creates a need to find a method of gathering axle load data over a broader geographic area. This section presents a literature review of research into methods of doing this; five methods are described and compared here: portable WIM, bridge WIM, on-board weighing, vehicle reidentification, and axle load modelling.

2.4.1 Portable Weigh-In-Motion

Portable WIM is a technology that operates in much the same manner as permanently installed WIM. A portable WIM installation consists of sensors installed on top of the roadway that record the axle loads of trucks that pass over them at highway speed, allowing axle load data for every vehicle that passes the sensor to be collected (Kwon, 2012). A portable WIM system, like a permanent WIM system, consists of three main components—a sensor, a controller, and support equipment for supplying power and transmitting data (Refai, Othman and Tafish, 2014). Often, the equipment used for portable WIM installations is the same as that used in permanent WIM applications, though the full range of options that is afforded to permanent WIM is not available for portable use due to the size and/or installation requirements for the equipment (Hallenbeck and Weinblatt, 2004; Stephens *et al.*, 2017). The sensors are attached to the surface of the pavement, creating visible, raised sensors which are typically installed for between one day and one month to gather axle load data (Kwon, 2012; Refai, Othman and Tafish, 2014; Petersen, 2015). A high-speed portable WIM system can cost from $12,000 - $25,000 USD, making its low cost relative to permanent WIM a point of appeal (Texas A&M Transportation Institute, 2015).

Portable WIM has been used for monitoring vehicle axle loads as early as 2004, when the National Cooperative Highway Research Program reported it as a desired method of collecting axle load data on highways in the United States, stating a preference for a combination of permanent and portable installations. However, the portable WIM technology required a lengthy calibration procedure for each installation, which increased its cost to the point of making its use impractical in many cases. Furthermore, the sensors were frequently less accurate than permanent WIM sensors due to the inability to calibrate them as often as they required. These issues led to many states choosing to substitute low-speed WIM or semi-permanent installations for portable WIM traffic monitoring (Hallenbeck and Weinblatt, 2004). More recently, the state of Montana discussed various portable WIM technologies as a potential part of its traffic monitoring strategy, as it currently makes extensive use of portable automatic traffic recorder (ATR) and automatic vehicle classifier (AVC) technology. However, portable WIM is not currently used in the state due to its low accuracy and high labor costs (Stephens *et al.*, 2017). The state of Georgia used portable WIM as an additional source of axle load data in the 2000s, collecting 107 portable WIM samples at 56 unique locations between 2002 and 2012. An analysis of these samples concluded that the data collected from portable WIM provided quality insight to traffic loading patterns, but should not be considered to be accurate enough to be a MEPDG level 1 data input for pavement design (Selezneva and Von Quintus, 2014).

Since 2012, various jurisdictions in North America have expressed a renewed interest in portable WIM and performed studies to determine the accuracy and practicality of portable WIM technology. These studies are summarized in Table 3. The two studies from Minnesota (Kwon, 2012 and Peterson, 2015) are connected; the 2015 study is a field evaluation of the system that was developed in the 2012 study.

Table 3 – Recent studies of portable WIM

	Kwon	Refai et al.	Jaworski	Peterson	Faruk et al.
Year	2012	2014	2015	2015	2016
Location	Minnesota	Oklahoma	Saskatchewan	Minnesota	Texas
WIM Controller	Custom-built unit	IRD iSINC Lite	IRD TCC 540	Custom-built unit	TRS portable WIM
Sensor type	Roadtrax BL	Roadtrax BL	Roadtrax BL	Roadtrax BL	Roadtrax BL
Sensor Setup	Conveyor belt weigh pad	Metal sheet loading pads	Pocket tape enclosure	Conveyor belt weigh pad	Pocket tape enclosure
Test Truck Used	Yes	Yes	Yes	Yes	Yes
Comparison with Permanent WIM	Yes	Yes	No	Yes	No
GVW Accuracy	3.88% (NRMSE)	29.05% (NRMSE)	-	3.94-8.14%	-
Temperature Consideration	Yes	Yes	No	No	Yes
Preferred Deployment Duration	Unstated	4 days	Unstated	2 days	7 days
Reference	(Kwon, 2012)	(Refai, Othman and Tafish, 2014)	(Jaworski, 2018)	(Petersen, 2015)	(Faruk *et al.*, 2016)

Each of these five studies used Roadtrax BL sensors, a piezoelectric technology often installed in the road as part of a permanent WIM system. These sensors were installed either using ‘pocket tape’, which is a product designed to allow the sensor to be attached to the road without leaving any adhesive on the sensor itself, or with custom solutions using high-strength materials like rubber conveyor belts or sheet metal to protect the sensor from the impact of traffic and thus decrease sensor wear. All five studies calibrated the sensors before installation with a test truck of known weight and axle spacings, and three of them installed the sensors alongside permanent WIM sensors in order to compare the accuracy of the portable unit to the baseline values of the permanent WIM. The error in these three studies was not reported uniformly; two studies reported vehicle accuracy using normalized root mean square error (NRMSE), while Petersen (2015) reported average percent error by vehicle class, providing a range of GVW accuracies.

The five studies considered different aspects of portable WIM technology. Three of the studies considered the effect of pavement temperature on the sensor's axle load measurements, though Kwon (2012) concluded that temperature has no significant effect on the axle load measurements and Refai et al. (2014) concluded that temperature has a strong correlation to the portable WIM's axle load measurements. Additionally, Kwon (2012) examined the effect of vehicle speed on the load measurements, concluding that higher vehicle speed resulted in higher load values due to the dynamic effects of the bump created by the sensor. Peterson (2015) extensively examined the difficulties of installing the sensors in a real-world environment, such as problems with installation and removal of the sensors and wear on the sensors by vehicles. Refai et al. (2014) noted that due to the installation method, which is relatively exposed to traffic, vehicle pressure can cause the sensor to vibrate within its enclosure, creating signal noise that must be filtered out. The studies also had various recommended durations of sensor deployment based on the observed wear on the sensors and degradation of data quality over time; these durations ranged from two to seven days.

Once developed, a portable WIM system will require an appropriate sampling program in order to most effectively gather spatially representative axle load data. This review found no studies examining a portable WIM sampling program.

2.4.2 Bridge Weigh-In-Motion

Bridge weigh-in-motion (Bridge WIM, or BWIM), is a WIM technology that calculates vehicle axle loads by instrumenting a bridge with sensors to measure deflection of the structure. A typical BWIM system consists of a number of strain gauges or fiber optic sensors attached to the underside and supports of a bridge and an axle detection system installed on the pavement approaching the bridge. A processing algorithm is required to determine when a vehicle is present from the axle detections and use that information to translate the strains experienced by the bridge into axle loads. The axle detection system has typically used pneumatic tubes or tape switches to detect axles. The benefits of BWIM sensors include ease of installation, as they can be installed without interrupting traffic flow, portability, cost-effectiveness, and accuracy (Yu, Cai and Deng, 2016)).

BWIM sensors are currently used in many countries around the world, including Australia, Japan, various European countries, the United States (Richardson *et al.*, 2014; Labarrere, 2018a), and Canada (Lydon *et al.*, 2016). Development of the necessary technology for a functioning BWIM system has been ongoing since the 1970s (Koniditsiotis, Buckmaster and Fraser, 1995; Richardson *et al.*, 2014; Yu, Cai and Deng, 2016), with the earliest development occurring in Australia, where BWIM development was driven by three desired applications: 1) infrastructure design and management, 2) trade and regulation, and 3) vehicle weight restriction enforcement (Koniditsiotis, Buckmaster and Fraser, 1995). This contrasts experience in other jurisdictions, which focused their development of the same technologies on vehicle weight restriction enforcement and bridge monitoring (Cantero and González, 2015; Lydon *et al.*, 2016; Yu, Cai and Deng, 2016). Despite the efforts invested into development for enforcement, modern BWIM systems are only accurate enough for indirect enforcement of weight regulations, as they are typically unable to achieve a GVW accuracy within +-5% (Lydon *et al.*, 2016), which is required for direct enforcement. When properly calibrated, BWIM systems can be accurate to within 7% of static scale GVW, but typically achieve a lower accuracy, with some installations unable to achieve accuracies within 25% of static scale GVW (Richardson *et al.*, 2014).

Many studies have been conducted in the 2010s investigating improvements to various aspects of BWIM technology. These studies primarily focus on issues with the specific technologies that are used as part of the system with a view to developing hardware or software with new capabilities and/or accuracy. The focuses of these studies are summarized in Table 4.

Table 4 – Focus of studies of BWIM technology

Technological Improvement	References
Improved accuracy of BWIM systems	(Lydon *et al.*, 2016; Kalyankar and Uddin, 2017; Obrien *et al.*, 2018; Žnidarič, Kalin and Kreslin, 2018)
Discovery of better locations on the bridge structure to install sensors	(Zolghadri *et al.*, 2016; Kalyankar and Uddin, 2017)
New bridge sensor technologies	(Lydon *et al.*, 2016; Sekiya, Kubota and Miki, 2018)
New and/or improved processing algorithms	(Zhao *et al.*, 2014; Ieng, 2015; Lydon *et al.*, 2016; Zolghadri *et al.*, 2016; Žnidarič, Kalin and Kreslin, 2018; Algohi, Mufti and Bakht, 2020)
New strategies for axle detection	(Lydon *et al.*, 2016; Žnidarič, Kalin and Kreslin, 2018)

These studies generally presented BWIM as a technology that justifies use for a variety of applications in its current form, but has enough accuracy issues to warrant further research into a variety of methods of improving the technology. It is also notable that no BWIM system has yet been able to solve the problem of how to accurately detect and weigh individual vehicles when multiple vehicles are on the bridge at once, which is frequently required (Yu, Cai and Deng, 2016; Sekiya, Kubota and Miki, 2018). Furthermore, due to the many possible combinations of sensor types, sensor installation methods, axle detection technologies, and processing algorithms, no one BWIM system configuration is presented as superior to the others, and none appears likely to become an industry standard in the near future.

In addition to the studies concerning improvements to the BWIM system, there are some studies that do not seek to develop the technology further, but instead examine what can be achieved with the existing technology. Cantero and

Gonzalez (2015) sought to characterize the structural integrity of a given bridge by comparing the load measurements of a nearby pavement-based WIM system to a bridge WIM system, and concluded that this method is a valid indicator of the structural integrity of the bridge. Lydon et al. (2016) discussed a study which used BWIM data to develop live load factors for bridge design in Alabama, determining that the BWIM loads provided more accurate factors than the previous values. Faraz et al. (2017) were able to identify errors in bridge fatigue evaluation using a BWIM system and therefore were able to improve the fatigue estimate for the bridge that was instrumented.

One paper in particular warrants closer examination. While BWIM is a useful technology for gaining vehicle load data due in part to its portability, that portability raises the question of where to install it and for how long. The only report examining this issue is from Ireland (Enright, Leahy and O’Brien, 2012), which examined possible data collection strategies for a fully functioning BWIM system. Three strategies were examined: two focusing on randomized application of the system for 1 week periods across different networks of defined road segments, and one focusing on targeted application of the system at known truck origins and destinations. The study emphasized BWIM’s potential use for direct enforcement of overweight vehicles, and so reported its results as percent likelihood of recording any one overweight truck. The study determined that the likelihood was between 12.1% and 23.7%, with the randomized approaches performing better than the targeted approach. The report notes that a system of reinstallation every week would create high operating costs for the system but would provide data for a wide range of locations thus, in theory, making the benefits worth the costs. The methodologies examined here could also be applied to portable WIM technology, which has no studies of this kind dedicated to its implementation.

2.4.3 On-board Weighing

On-board weighing is a method of obtaining truck weights wherein the sensors to measure the vehicle's axle loads are installed on individual trucks, rather than in or on the roadway. This allows the axle load data to be continuously recorded and transmitted to a database that stores the load measurements and corresponding locations (Labarrere, 2018b). If used across a sufficient number of trucks that cover a desired geographic area, these individual point measurements could be collected in a database of load measurements with corresponding locations, grouped by location, and used to assemble sets of axle load data for any given location on the road network. No system such as this has been implemented yet, but on-board weighing technology has been developed and used to enforce weight regulations (Labarrere, 2018b).

On-board weighing systems are currently used for vehicle weight restriction enforcement in Australia, but are not used elsewhere in any notable capacity. The systems used are 'static' on-board weighing systems, which only measure axle loads when the vehicle is stationary (Todts *et al.*, 2013). Using a dynamic system, which can take continuous load measurements while the vehicle is in motion, is considered incapable of measuring axle loads at the accuracy required without increasing the cost by 5-10 times (Labarrere, 2018b). Recently, costs for an on-board weighing unit have come down due to advances in the technology and ranged from $800 to $15 000 per installation in 2015 (Texas A&M Transportation Institute, 2015). Furthermore, wider implementation of on-board weighing for enforcement is restricted by legal issues, as they can only be used to enforce weight regulations if legislation allows (Todts *et al.*, 2013).

In addition to the problem of accuracy, significant issues must be overcome with respect to the use of on-board weighing. Because the equipment is installed on privately own trucks instead of publicly owned roads as in a traditional WIM system, work must be done to garner cooperation from carriers before the systems can be deployed (Texas A&M Transportation Institute, 2015). This proves to be a greater issue when the goal of the program is to provide accurate axle load data for specific locations on the road network, necessitating significant penetration of the technology into an area's truck fleet.

One model for the potential implementation of a wide scale program of on-board load data collection is the United States' National Corridors Analysis and Speed Tool (N-CAST). Developed since 2002, N-CAST characterizes the performance of the national truck fleet of the United States by combining travel data from many individual trucks to produce a GIS map of truck performance characteristics across 75% of the nation's heavy truck network. This data is a useful tool for carriers and transportation planners; however, due to the manner in which the data is collected, there are limitations to its use. N-CAST reports average spot speeds and percentage of trucks by time of day for each segment on its network, but cannot report traffic volumes as the fleet used to generate it was only a sampling of the total truck fleet (American Transportation Research Institute, 2012). Furthermore, it does not report any vehicle load data. Truck volume data will be needed to accurately predict the structural impacts of trucks on infrastructure, and so a database of axle load data generated by on-board weighing must have data from the entire truck fleet if it is to be truly accurate.

To achieve complete penetration of a given trucking fleet, it would be useful to capitalize on synergies with the digital tachograph, or electronic logging device (ELD) (Todts *et al.*, 2013), which measures and records location and hours spent driving and is currently mandatory in Europe (European Commision, 2018) and is in the process of becoming mandatory in Canada (Sanderson, 2018) and the United States (Federal Motor Carrier Safety Administration, 2018). Additionally, further synergies might be generated with Intelligent Transportation Systems (ITS) technologies that assist with commercial vehicle inspection and enforcement such as Drivewyze (Drivewyze, 2020) or British Columbia’s Weigh2GoBC program (Government of British Columbia, no date). These similar programs allow vehicles to bypass weigh scales and inspection stations if they have been previously confirmed to be in compliance by communicating between a smartphone in the truck and a receiver in the weigh scale or inspection station. The inclusion of axle load data in these systems is a logical next step, though there has been no plan to implement them with any kind of on-board weighing system yet.

2.4.4 Vehicle Reidentification

Vehicle reidentification is a method of traffic monitoring that has been developed to measure a variety of traffic parameters but has only been recently integrated with the collection of axle load data. The principle of the technology is that an individual vehicle can be identified by a sensor and assigned a unique signature that can be reidentified at a later point and time, with the vehicle then being known to have traveled the path between the two points in the intervening time (Sun *et al.*, 1999). The advantage of this technology is that it can provide individual vehicle-level travel information while maintaining anonymity, and thus does not require cooperation from carrier companies (Cetin *et al.*, 2011). In theory, when combined with a known weight of a vehicle obtained from a WIM system, a reidentification system would be able to match a vehicle's known weight to its route between two identification points, and with enough measurements would be able to generate an axle load spectra at a number of points along a route (Hyun, Tok and Ritchie, 2017).

This technology was first contemplated in 1999, when researchers sought to reidentify a vehicle signature using inductive loops in order to calculate vehicle travel times on highways (Sun *et al.*, 1999). In 2005, this was developed further by applying an algorithm to the reidentification procedure in order to identify traffic level of service (LOS) from the vehicle travel times (Oh, Tok and Ritchie, 2005).

Further research into vehicle reidentification technology using inductive loop signatures investigated the possibility of identifying a vehicle's FHWA vehicle class from its inductive signature to improve the reidentification algorithm (Jeng and Ritchie, 2008; Jeng, Chu and Hernandez, 2013), and, more recently, to improve the algorithm by correcting for different signal conditioning within the inductive loop (Marszalek and Duda, 2018).

The first integration of WIM technology to vehicle reidentification technology came in 2011, when Cetin et. al. examined factors affecting the reidentification of trucks using only WIM data, noting that sensor accuracy and truck volumes had a significant effect on the system's accuracy (Cetin *et al.*, 2011). These researchers additionally examined the reidentification model and determined that trucks can be reidentified with 91% accuracy when using only WIM systems (Cetin, Monsere and Nichols, 2011). This research then led to the development of a new vehicle reidentification model and a method of analyzing WIM accuracy using vehicle reidentification algorithms (Cetin, Nichols and Chou, 2014).

The next improvements in vehicle reidentification technology came when Jeng and Chu (2015) developed a vehicle reidentification method that uses both WIM sensors and inductive loops to obtain a more complete vehicle signature to allow for greater accuracy in reidentification. Their method used the loops to identify a vehicle signature, and the axle loads and configuration from the WIM sensor to provide a check on this signature. An improvement to this method came when Hyun et. al. (2017) tested several algorithms that better integrated the WIM and inductive loop data to provide a more accurate reidentification method. Overall, vehicle reidentification as a method of vehicle load data collection is still in development, but has made significant advances in recent years.

2.4.5 Axle Load Modelling

Axle load modelling is a method of characterizing vehicle loads that seeks to achieve greater spatial representation in axle load data by applying previously gathered data on vehicle load patterns in a more comprehensive way, rather than gathering more data. This method relies on the fact that the volume and loads of trucks that travel a given section of road are influenced by that road's surrounding land use and industry, weight regulations, and climate, as well as how those variables vary in time and space. This method seeks to use data about how each of these variables influences the axle load spectra at a given point to estimate axle load spectra at that point.

There are two basic approaches to axle load modelling:

1. Generate axle load spectra at a location based purely on a model of factors that correlate to patterns in axle load spectra.
2. Apply axle load spectra from road segments that are most similar across all correlating factors.

The first of these approaches, axle load spectra generatation, has not been performed using vehicle loads, but insight to this approach can be gained by examining the 4-stage freight model. The 4-stage freight model is a tool for estimating transportation demand from freight, wherein a region's demographic, economic, and spatial characteristics are linearly regressed on freight volumes to provide a framework for estimating changes to transportation activity. The first stage of the process, trip generation, is described by Kulpa in a review of existing methods of regional freight trip generation. The author notes that a multiple regression model is capable of correlating variables such as employment in various industries and number of companies in a given industry to number of truck trips, which have individual trips sorted into the categories of 'light' and 'heavy' (Kulpa, 2014). There has been some research into methods of improving this step of the process. A study from Austria (Krisztin, 2017) examines a few shortcomings of the existing model, notably that they do not take spatial dependencies between nearby regions into effect, and that the relationships of some variables are nonlinear, which does not fit the standard linear model. Alternatively, research by Dybing and Dakota (2017) focuses on trips related to the agriculture industry, and notes that to accurately forecast the number of trips generated by this industry, data on acreage and production rates must be obtained. However, these data are typically generated from estimates rather than actual data, and so must be first verified in a detailed analysis. Overall, the 4-stage freight model provides a theoretical basis for axle load modelling, but the focus on the overall activity of a region, rather than trips on individual roads, decreases the benefit to a future load-based model.

The second approach, where axle load spectra from known locations are applied to similar roads, has been applied in some instances to vehicle loads, and in others to only truck volumes or classifications. This most often occurs in the application of traffic inputs for the MEPDG. The existence of three levels of data as specified by the MEPDG creates a need to use higher-quality data to inform truck traffic characteristics where only lower-quality data is available. There are several similar, but distinct, approaches to this.

Regehr and Reimer (2013) used a methodological framework for developing system-wide estimates of truck travel on rural highways in Manitoba wherein pattern groups of truck class distribution are created from MEPDG data level 1 sites, and have sites of lower data quality assigned to these groups. The assigning process takes into account whether a route is affected by the forestry or agricultural industry, but provides no detail as to the degree of the influence and considers no other factors or industries. A strategy similar to this was implemented in Montana (Stephens *et al.*, 2017), where it was proposed to create pattern groups of axle load spectra based on areas that have similar land use and industries or based on functional classification. The study found that grouping based on industry is impractical due to its subjectivity, and that the grouping scheme using functional classification is best due to its objectivity and ease of implementation. It proposed a final grouping scheme that uses distinctions between commercial and non-commercial vehicles, interstate and non-interstate roads, and urban and rural land uses.

Haider et al. (2011) and Nassiri et al. (2014) also conducted studies that used this general approach in determining traffic inputs for Michigan and Alberta, respectively. Haider's method was to determine the most appropriate vehicle load input when MEPDG level 1 data is unavailable for a site. Nassiri's method examined the difference between the MEPDG level 3 default axle load spectra and axle load spectra reported by six permanent WIM sites across the province. The studies found that axle load spectra generated as state/province-wide values were more accurate than the default values provided by the MEPDG due to the unique mix of local industry, regulations, and climate.

Turochy et al. (2005) reached the same conclusion, which sought to determine axle load modelling's value to pavement design in Alabama. Focusing on temporal variation and variation by direction, they found that there were statistically significant differences in truck axle load spectra between different regions of the state, but that the implications of these differences on pavement design were minimal, resulting in a recommendation to use statewide default values for axle load spectra where MEPDG class 1 data was unavailable.

Overall, no efforts have been made to generate axle load spectra at a location based purely on any combination of location-based correlating factors as is done with truck traffic volumes in freight traffic demand modelling, while some efforts have been made to apply axle load data to sites that have none from sites with similar characteristics. Significant further development of axle load modelling methods will be required before it can be implemented on a wide scale.

2.5 SUMMARY

Axle load data is important for a variety of applications, including design of pavements and bridges, freight planning studies, and vehicle weight restriction enforcement. WIM systems are the best method of collecting axle load data for these applications, and a variety of WIM sensor technologies are currently used, with the various sensor types each having advantages and disadvantages regarding accuracy, sensitivity to temperature, and durability.

While excellent at collecting axle load data, WIM systems all have issues that must be considered when using them. As they measure vehicles dynamically, WIM sensors are prone to a degree of inaccuracy which can be partially mitigated through regular calibration of the systems and consideration of the impact of temperature on WIM measurements, which is often significant. Standards organizations have set minimum accuracy requirements for axle load data which vary by application. Furthermore, as WIM sensors are installed at single locations, WIM data is frequently unavailable at locations where it would be desired, and other data must be applied, though it is not always spatially representative of the desired location. The MEPDG’s three level system is recognized as an industry standard approach to this process.

Eleven of 13 of Canada’s provinces and territories currently report use of WIM systems. WIM sites in Canada are used for several applications, including general data collection for engineering and planning applications, pre-screening for vehicle weight restriction enforcement, and more specific applications such as the administration of bridge tolls. WIM coverage varies widely between jurisdictions, from 19 WIM stations in Quebec to just 1 in the Northwest Territories.

As many jurisdictions suffer from inadequate spatial representation in their axle load data, various methods of improving the spatial representation of axle load data have been developed. Five of these methods were explored; the methods were in different stages of development, with each having unique research needs.

Axle load modelling and on-board weighing are both in the early stages of development and cannot be implemented for load data collection until significant research needs are met. Axle load modelling requires further studies into appropriate methodologies, while on-board weighing requires more development of the weighing technology as well as legal work done to achieve sufficient penetration of the technology into the truck fleet.

Vehicle reidentification also has significant needs. While recent progress has demonstrated that the technology can be used to gather useful data, the lack of implementation outside of research contexts and absence of an accepted deployment methodology indicates that this technology is not yet ready for implementation in a load data collection program.

Bridge WIM and portable WIM both have been deployed in limited capacity in jurisdictions' vehicle load data collection programs. An abundance of research into bridge WIM shows that it is capable of gathering spatially representative axle load data, but still suffers from issues with multiple vehicle detection and overall accuracy. Furthermore, the wide variety of BWIM technologies leaves a greater number of research needs in the technology's development. Alternatively, portable WIM, though it has been deployed less frequently than BWIM, has had consistency in the technology used that creates a clear path forward for the technology's development. Furthermore, the similarity of the technology to existing permanent WIM systems will aid in near future implementation. Further research into methods of gathering and applying vehicle load data from portable WIM systems would benefit jurisdictions seeking to gather more spatially representative axle load data in the near future.

3 RESEARCH METHODOLOGY

This chapter presents the methodology developed and applied in this research, including (1) the sources of both portable WIM data and comparison data used, (2) the data collection processes, (3) the method by which portable WIM records were paired with comparison data records and initially assessed for data quality, and (4) the analyses in which the portable WIM system's accuracy, precision, and potential for increased data quality through various post-processing methods and analyses will be evaluated.

3.1 DATA SOURCES

To evaluate the capabilities of the portable WIM system, this research compares it to two existing sources of vehicle load data: (1) permanent WIM stations and (2) static weigh scales.

3.1.1 Portable WIM System Design

In order to provide the most accurate assessment of the feasibility of using portable WIM to collect useful vehicle load data, the portable WIM system used in this research was designed to be as similar as possible to one that would be used in an actual data collection program. For this reason, the selection of components was done with the two goals of (1) optimum system performance and (2) cost minimization. The system that resulted from this selection process used components that were recommended for use in portable WIM applications both by past studies of portable WIM and by the equipment supplier, and whenever possible used components that were already in Manitoba

Infrastructure's (MI's) stock due to their use in MI's current permanent WIM data collection program. Table 5 lists the components selected for use in the study.

Table 5 - Portable WIM System Components

Component	Model	Number Required	Expected Lifetime
WIM controller	IRD iSINC Lite	1	>20 years
WIM sensors	Piezoelectric Roadtrax BL (12' length)	2	Unknown for portable application
Power supply	80W Solar panels	2	>20 years
Installation materials	Pocket tape	1 roll (60')	Single use
	Reinforcing tape	1 roll (60')	Single use
	Metal strapping	48'	>6 weeks of installation
Temperature probe	IRD temperature sensor	1	Unknown

The rationale for selecting each of these particular components follows:

- WIM controller: The iSINC Lite WIM controller from International Road Dynamics (IRD) is the current WIM controller being installed at new permanent WIM sites in Manitoba and was used in one previous portable WIM study.
- WIM sensors: The piezoelectric Roadtrax BL sensors were used in all of the previous portable WIM studies discussed in chapter 2.
- Power supply: A solar power supply allowed greater flexibility in installation location as compared to a grid-based power supply, which will be necessary for any future implementation of a portable WIM sampling program.

- Installation materials: Pocket tape is a low-cost attachment method, and was previously used in the portable WIM study in Texas in 2016. Reinforcement from additional tape and metal strapping was added due to previous studies' issues with longevity of the sensors, as well as precedent for a similar heavily reinforced system used in the two studies in Minnesota in 2012 and 2015.
- Temperature sensor: The IRD temperature sensor was already in MI's stock.

The equipment was installed on a trailer as seen in Figure 1. The WIM controller was housed in a cabinet attached to the trailer for security. The temperature sensor, which is typically installed in the pavement at a permanent WIM site, was left open to the air on the trailer. The solar panels were also mounted to the trailer.

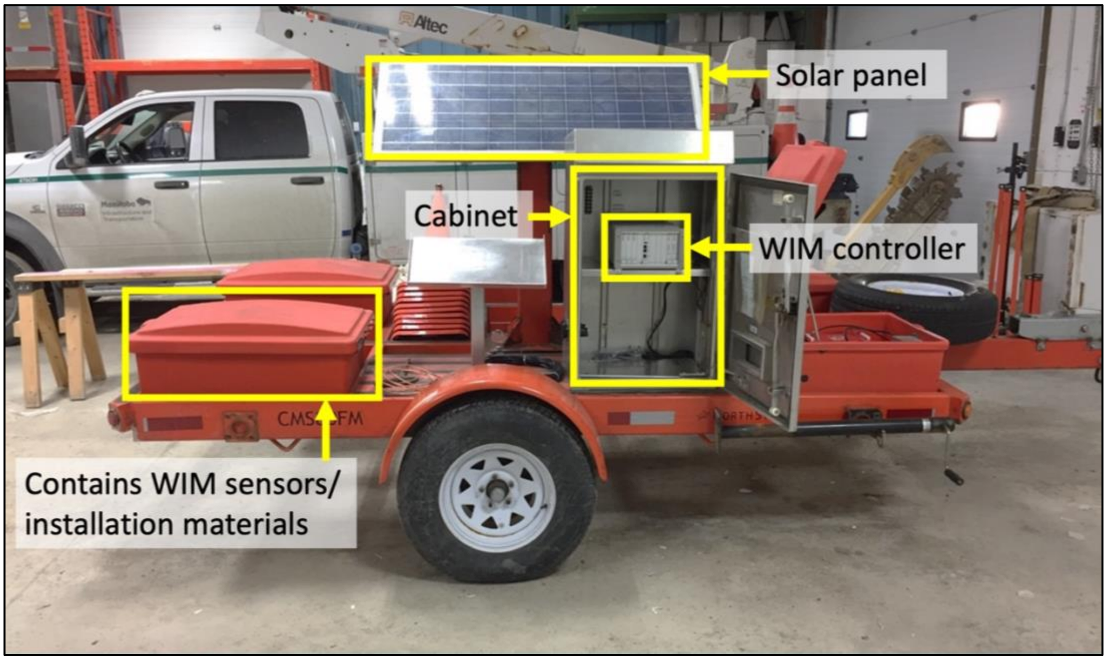


Figure 1 – Portable WIM trailer

When the system was installed, the trailer was parked on the side of the highway as far from the traveled lane as possible. Connections between the sensors and the trailer were configured as seen in Figure 2. This configuration follows the installation guidelines outlined in the Federal Highway Administration's Weigh-In-Motion Pocket Guide Part 2 (Federal Highway Administration, 2018).

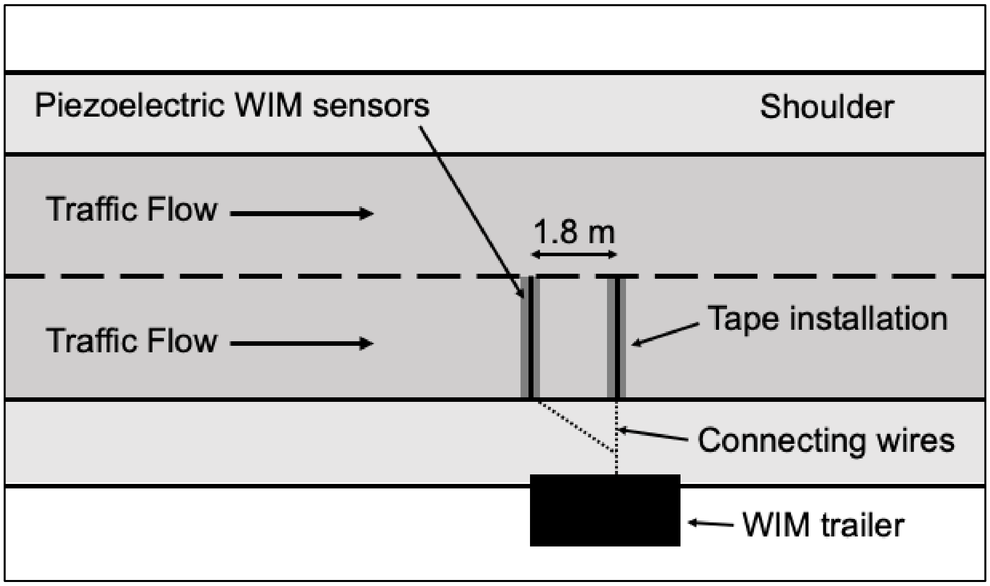


Figure 2 - Portable WIM system configuration

The tape installation was designed to be both durable and easy to install. Pocket tape, which encloses the sensor between two layers of tape approximately 15 centimeters wide and has an adhesive underside, was used to hold the sensors in place. In addition to this, a second, thicker layer of protective tape approximately 20 centimeters wide was installed over the pocket tape to provide additional protection for the sensors. Finally, to prevent lifting of the tape, metal strapping was installed over the leading and trailing edges of the tape using concrete screws placed approximately 15 centimeters apart. Figure 3 shows the pocket tape installation.

Figure 3 - Portable WIM sensor pocket tape installation

3.1.2 WIM Data in Manitoba

Manitoba currently operates six permanent WIM sites; three of these are equipped with piezoelectric sensors and three sites are equipped with higher-quality piezo-quartz sensors. Figure 4 shows the locations of these sites, along with Manitoba's static weigh scale stations.

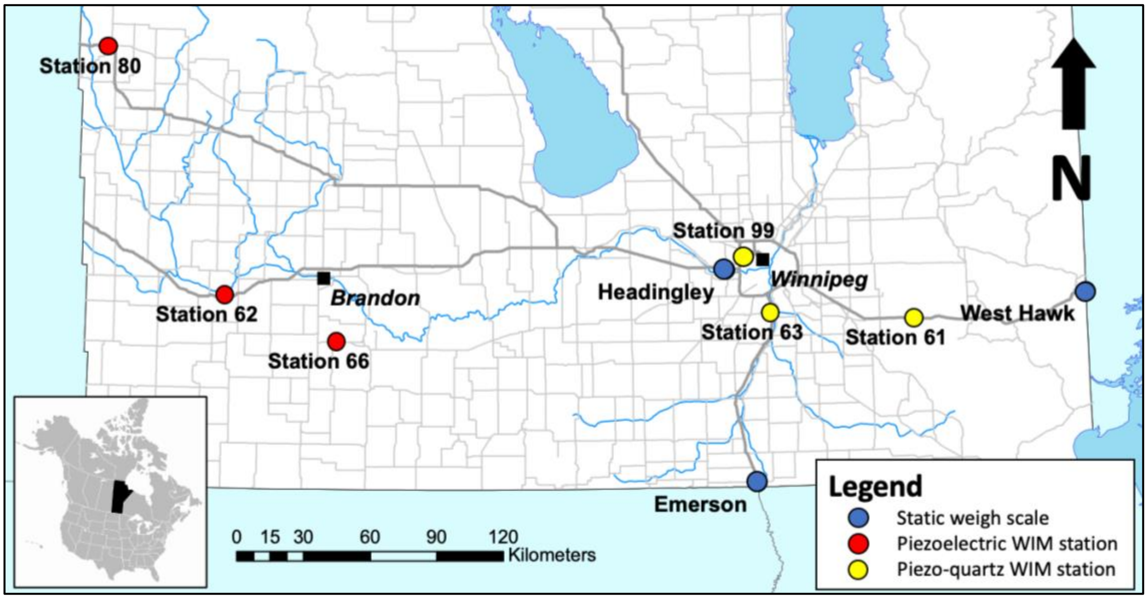


Figure 4 - Map of Manitoba WIM stations and static weigh scales

The six permanent WIM sites have two WIM sensors installed per lane in one to four lanes of traffic. For those sites where WIM sensors are not present in all lanes, automatic vehicle classifiers (AVCs) are installed in the remaining lanes. The data collected at these sites is used to support pavement design projects and research in Manitoba and is also submitted to the FHWA in support of the Long-Term Pavement Performance (LTPP) program. To correct for the effects of long-term calibration drift, the sites are calibrated approximately once per year. However, despite this calibration, the piezoelectric sensors are known to provide axle load measurements that are heavily influenced by temperature, and additionally are occasionally out of service due to equipment malfunctions (Wood, 2017). Due to these shortcomings, Manitoba Infrastructure has implemented a program of gradually phasing out piezoelectric WIM sensors and replacing them with more reliable piezo-quartz sensors. In 2019, this program resulted in the replacement of the sensors at two sites with piezo-quartz, resulting in a total of three sites with piezo-quartz sensors.

3.1.3 Static Weigh Scale Data in Manitoba

Manitoba Infrastructure operates three primary static weigh scales for vehicle weight enforcement; Figure 4 shows the locations of these scales. As static scale axle load measurements can be used to issue citations to drivers and carriers, they must be accurate to within 200 kg for a single axle load. To this end, load cell weighing devices are used, and the scales are calibrated once every two years. Scale measurements are transmitted from the scale to a display inside the scale building that is visible to the enforcement officer on duty. Trucks are typically instructed to slowly roll over the scales, with the enforcement officer checking each axle group against the legal limit for that group type. While this is happening, the axle group load measurements can fluctuate due to the truck's movement; only when an infraction is deemed probable is the driver instructed to stop with each axle group on the scale until the fluctuations stop. This occurs for less than 1% of the trucks surveyed. Furthermore, it is only in these cases that a load measurement is recorded by the officer for issuing a citation; if no citation is required, the loads are not recorded.

The University of Manitoba Transport Information Group (UMTIG) has conducted regular 'weigh scale surveys' at each of the three scales approximately every 2-3 years. During these surveys, a researcher stationed inside the scale building manually recorded the loads of each vehicle axle group as it passed over the scale in addition to the configuration, body type, lift axles, and if necessary, identifying visual characteristics of each vehicle. As most trucks do not come to a complete stop for each axle group, the researcher must choose a 'best' value for each axle group based on the fluctuating load measurement display. Despite this, the data collected from UMTIG's weigh scale surveys has been shown to be sufficiently accurate to be considered 'ground truth' data (Wood, 2017). This research utilized data from these static weigh scale surveys.

3.2 DATA COLLECTION

For this research, the portable WIM system was installed three times, each time near to a comparison data source. These installations all took place in 2019. The portable WIM unit was installed twice at permanent WIM station 99 (piezo-quartz) and once at the Headingley static weigh scale. Table 6 provides basic information about all three installations.

Table 6 - Portable WIM installations summary

	Installation 1	Installation 2	Installation 3
Location	WIM Station 99	WIM Station 99	Headingley Static Weigh Scale
Start Date	June 27, 2019	July 17, 2019	October 16, 2019
End Date	July 15, 2019	July 31, 2019	October 28, 2019
Duration (Days)	19	15	13
Sensor lane coverage	Full	Half	Half
Initial calibration performed	Yes	No	No

3.2.1 Installations at Permanent WIM Station 99

The first and second installations of the portable WIM system occurred near Manitoba Infrastructure's permanent WIM station 99 located on CentrePort Canada Way (PTH 190) in Winnipeg, Manitoba. During the installation, WIM station 99 was used as the comparison dataset with which to evaluate the portable WIM system.

CentrePort Canada way is a 4-lane divided highway with a concrete surface oriented in the east-west direction with a speed limit of 90 km/h. The speed limit decreases to 70 km/h in the westbound direction near the Perimeter Highway (PTH 101) approximately 900 meters west of WIM station 99. WIM station 99 is equipped with piezo-quartz WIM sensors in the westbound drive lane and AVCs in the other 3 lanes. This location was selected for the first two installations due to the relative accuracy of the piezo-quartz WIM sensors and the station's convenient location inside Winnipeg.

Figure 5 shows the installation location of the portable sensors in relation to CentrePort Canada Way and WIM station 99. The portable WIM system was installed in the westbound drive lane approximately 50 meters west of WIM station 99 in order to capture the same vehicles as the permanent WIM system. The portable WIM system was located downstream from the permanent WIM due to the possibility of the raised portable sensors creating dynamic effects in passing vehicles that could influence the permanent WIM load measurements if they were taken downstream from the portable WIM. A 50 meter separation between the WIM systems ensured that the portable WIM did not interfere with the operation of the permanent WIM, while allowing vehicle timestamps to be virtually identical for vehicle record pairing.

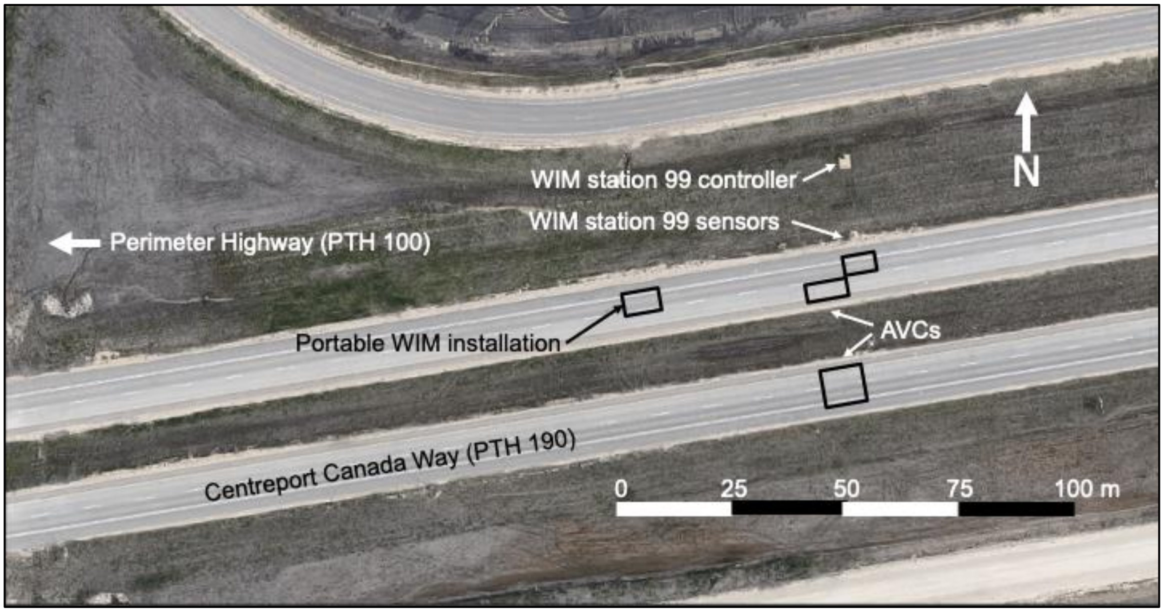


Figure 5 - Map of portable WIM installation at permanent WIM station 99

Calibration of both the portable WIM system and WIM station 99 was performed at the beginning of Installation 1 at WIM station 99. Installation 2 at WIM station 99 was performed shortly after the first to provide two key points of comparison: (1) whether reducing the lane coverage of the sensor to half would provide more accurate axle load measurements by eliminating signal noise, and (2) whether the elimination of a calibration at the beginning of Installation 2 could be compensated for by post-processing of the data. While Installation 2 covered only half the lane with the sensor, the sensor was the same 12-ft sensor that was used for Installation 1.

During both installations, the system was visually inspected every 2 to 4 days to monitor the physical wear on the sensors. No significant visible damage to the sensors was observed during these inspections.

3.2.2 Installation at Headingley Static Weigh Scale

The third installation of the portable WIM system occurred at the static weigh scale located in Headingley, Manitoba. This scale was chosen due to its convenient location near Winnipeg, Manitoba; the scale is located less than 1 kilometer to the west of the town of Headingley on the Trans-Canada Highway (PTH 1), which at this location is a 4-lane divided highway. The scale is located in the highway's median, with lanes located on either side to allow both eastbound and westbound trucks to exit from the highway's passing lane to the lane leading to the weigh scale. In both directions, there is a sign instructing 'All Trucks Report When Lights Flashing' with lights connected to the scale mounted to the sign. In the westbound direction, this sign is located approximately 350 meters upstream from the scale. In the eastbound direction, this sign is located approximately 1.4 kilometers upstream from the scale. At the location of the scale, the speed limit is 70 km/h; however, immediately to the west of the scale the speeds are 100 km/h eastbound and 110 km/h westbound.

Figure 6 shows the installation location of the portable sensors in relation to the Headingley static weigh scale. The portable WIM system was installed in the eastbound drive lane of the Trans-Canada Highway approximately 2.0 kilometers west of the static weigh scale. This location was chosen as it allows trucks to pass the portable WIM system at full highway speed, as they would in an actual deployment of the portable WIM system. Locating the portable WIM system 2.0 kilometers west of the scale ensures that most vehicles are still in the drive lane when they pass the portable WIM system, allowing axle load measurements to be taken. Furthermore, the eastbound direction was necessary as construction work closed the westbound scale lane during the time of the installation.

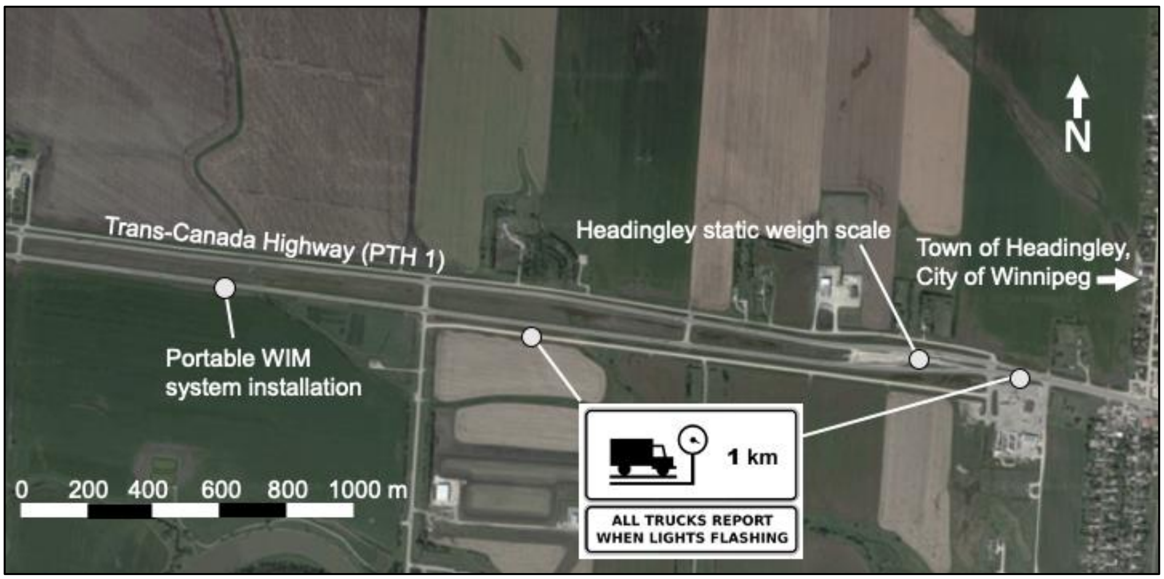


Figure 6 - Map of portable WIM installation at Headingley static weigh scale

As described in section 3.1.3, axle load measurements are not automatically recorded for most vehicles passing through the static weigh scale. This necessitates a researcher to manually record axle load measurements as vehicles pass through the scale. To allow these measurements to be matched to vehicles recorded by the portable WIM system, time stamps are automatically generated in the vehicle recording spreadsheet. Additionally, identifying information including axle configuration, body type and a visual characteristic of the truck were included with each vehicle record. A Miovision® traffic recording camera was installed to record traffic passing the portable WIM system with timestamps to allow for matching visually and by relative timestamp.

Installation 3 at the Headingley weigh scale remained in place from October 16 to 28, 2019, though equipment malfunctions resulted in all data from October 16 and all data from the afternoon of October 24 onward being unusable. During the installation, 3 to 6 hours of data from the Headingley weigh scale was manually recorded every weekday, as the scale was closed on weekends. This resulted in useful data being recorded on 6 days during the 13 day installation.

3.2.3 WIM System Calibration

To ensure the accuracy of its axle load measurements, a WIM system requires regular calibration. The calibration procedure used by Manitoba Infrastructure for its permanent WIM stations, which was used once for this research for both permanent and portable WIM systems, follows the method described in ASTM E1318 (ASTM International, 2009). This consisted of driving a fully loaded five-axle tractor semitrailer (3-S2) truck of known weight (the 'test truck') over the sensors at typical highway speed 10 to 15 times, after which the WIM controller was programmed with two factors – the calibration factor ($F_{calibration}$) and the dynamic compensation factor ($F_{dynamic}$). These factors, calculated and applied individually for each WIM sensor, adjust the raw load measurements to match the actual weight of the test truck as closely as possible. Equation 1 and

Equation 2 show how these calibration factors are applied to the front axle load and to all other axle loads, respectively.

Equation 1

$$CalibratedFrontAxleWeight = RawMeasurement * F_{calibration} * F_{dynamic}$$

Equation 2

$$CalibratedNonFrontAxleWeight = RawMeasurement * F_{calibration}$$

$F_{calibration}$ and $F_{dynamic}$ were calculated using the sums of the relevant axle group loads from the WIM system and the static scale. Equation 3 and Equation 4 show this calculation for *n* passes of the WIM system. As a FHWA class 9 vehicle has 3 axle groups, they will here be described as *Front* (single axle), *Drive* (tandem axle group), and *Rear tandem* (tandem axle group).

Equation 3

$$F_{calibration} = \frac{n * (Load_{DriveStaticScale} + Load_{RearTandemStaticScale})}{\left(\sum_{pass=1}^{n} Load_{DriveWIM} + \sum_{pass=1}^{n} Load_{RearTandemWIM}\right)}$$

Equation 4

$$F_{dynamic} = \frac{n * Load_{FrontStaticScale}}{\left(\sum_{pass=1}^{n} Load_{FrontWIM}\right) * F_{calibration}}$$

Both the portable WIM unit and permanent WIM station 99 required this calibration at the beginning of Installation 1; for cost efficiency, a single test truck performed passes of both sets of sensors at the same time, allowing simultaneous calibration of the two systems.

3.3 DATA PREPARATION

This section details how the portable WIM data was processed and checked before analyses of its quality were performed. The two primary steps in this process were the vehicle pairing procedure and the three preliminary data assessments.

3.3.1 Vehicle Pairing Procedure

To assess the per-vehicle validity of the portable WIM axle load measurements, the individual vehicle records generated by the portable WIM system and the comparison datasets needed to be paired. The two sources of comparison data – the Headingley weigh scale and WIM station 99 – have several key differences, and two different vehicle pairing methods were required.

WIM Station 99 Vehicle Pairing Procedure

The two portable WIM installations at WIM station 99 each had two datasets that required pairing: (1) the per-vehicle records from the portable WIM system, and (2) the per-vehicle records from WIM station 99. The pairing process was done by matching the vehicle record timestamps between these two datasets, and was done entirely automatically. A detailed description of this vehicle pairing procedure is provided in Appendix A.

This method successfully paired 60380 of 69580 (87%) portable WIM records for the Installation 1, and 23304 of 59442 (39%) portable WIM records for Installation 2. The difference in the percentage of successfully paired records can be attributed to erroneous vehicle records. During Installation 1, 4% of records had an error code, while during Installation 2, 28% of records had an error code. These error codes indicate that a vehicle did not pass over the two sensors as planned, and can therefore be likely attributed to the half-lane coverage of the sensors during Installation 2. This is further discussed in section 4.1.

Headingley Weigh Scale Vehicle Pairing Procedure

Portable WIM Installation 3 at the Headingley weigh scale yielded 3 sources of data: (1) the per-vehicle records from the portable WIM system, (2) the manually entered per-vehicle records from the Headingley weigh scale, and (3) the video taken of vehicles passing the portable WIM system. This pairing process was done by first manually recording vehicle timestamps from the video of vehicle passing the portable WIM system and visually matching the vehicles on video to the weigh scale records, then automatically pairing these timestamps to the portable WIM records. A detailed description of this vehicle pairing procedure is provided in Appendix A.

This method successfully paired 877 of 1852 Installation 3 static scale vehicle records to the portable WIM system (47%). The primary reason for not matching records was that they were traveling at least partially in the passing lane as they passed the portable WIM system, resulting in no record or an error record being recorded.

3.3.2 Preliminary Data Quality Assessments

Before the portable WIM system's feasibility and data quality could be evaluated, two preliminary data quality assessments needed to be performed to assess the validity of the datasets that would be compared. These assessments served to validate datasets and procedures which are used in the assessment of the portable WIM's data quality, but are not a direct assessment of the portable WIM's data.

Station 99 Data Quality Assessment

The first data quality assessment was of the permanent WIM station 99 data. As WIM data is always specified to have a certain degree of inaccuracy as compared to static scale data, this assessment evaluated whether the WIM data from station 99 was:

1) Self-consistent, i.e. not subject to excessive calibration drift, and
2) Sufficiently accurate to static scale data.

The assessment reveals that the WIM data from station 99:

1) Did not experience excessive amounts of calibration drift, and
2) Was typically accurate to within ±7% of static scale data.

This assessment is presented in detail in Appendix B.

Post-Processing Calibration Feasibility Assessment

The second data quality assessment was of the post-processing calibration procedure. This procedure was required as there was no data captured comparing calibrated portable WIM system data to static weigh scale data. The post-processing calibration used a selection of records from the static weigh scale data to calibrate the portable WIM system, allowing it then to drift from that calibration as it would with a real calibration. This assessment determined that the post-processing calibration procedure was valid so long as the vehicle records used in it were taken only from the beginning of the portable WIM system installation.

This assessment is presented in detail in Appendix C.

3.4 PORTABLE WIM DATA QUALITY ASSESSMENT METHODS

An evaluation of the portable WIM system's capabilities must consider its accuracy, precision, and potential for improved accuracy and precision through post-processing methods and analyses. In order to adequately evaluate the system in each of these areas, a series of five analyses are conducted on the available datasets. Figure 7 provides an overview of the order of these analyses. In this figure, arrows indicate that some analyses must be done prior to others as their findings impact the processes used in the subsequent analyses.

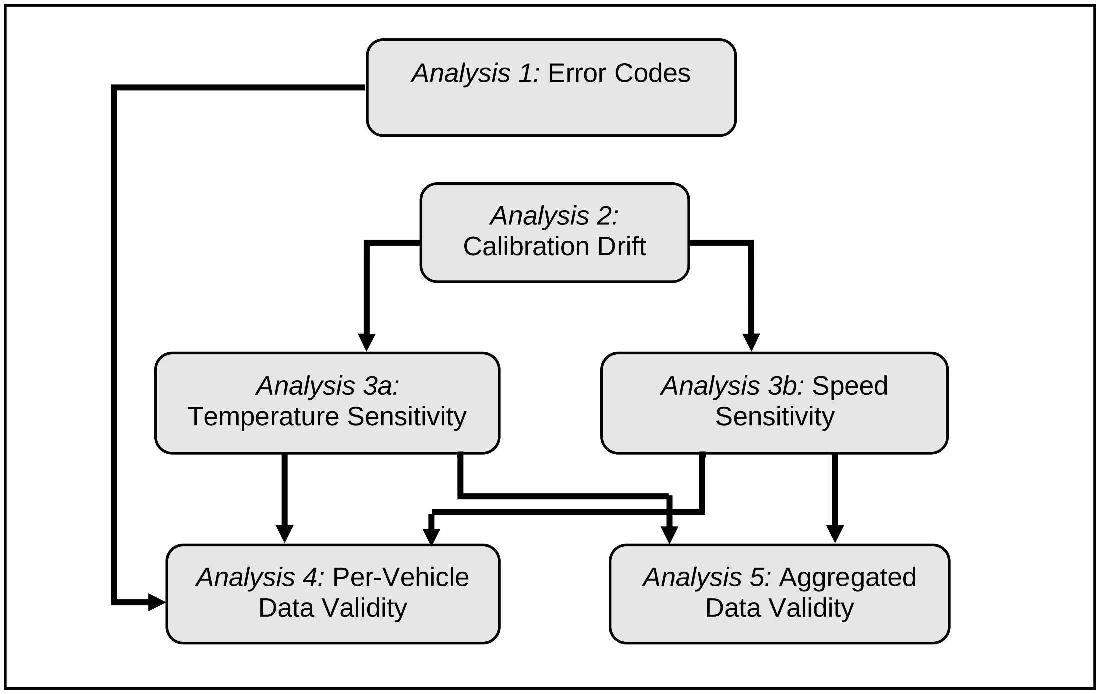


Figure 7 – Overview of portable WIM system analyses

The five analyses are described in more detail as follows:

Analysis 1 investigates the rate of erroneous portable WIM records during each installation, as well as the rate of successful vehicle pairing.

Analysis 2 investigates the tendency of the portable WIM axle loads to change over time as environmental factors cause the system to drift from an initially calibrated state. The analysis seeks to determine the best explanatory variable for the system’s calibration drift, which will be used to separate the data into temporal bins for future analyses.

Analysis 3 determines the impact of two measurable variables on the portable WIM data: temperature and speed. The analysis investigates whether these variables have a significant impact on the portable WIM axle load measurements, the nature of those impacts, and how that impact changes as the calibration of a portable WIM system drifts during an installation.

Analysis 4 determines the per-vehicle validity of the portable WIM data. It considers both calibrated and uncalibrated data, data adjusted for impacts from temperature and speed, and data that is autocalibrated throughout the installation. Accuracies measured here are compared to the standards set in ASTM E1318 to determine their sufficiency for various applications. Sensor coverage recommendations also inform the findings from this analysis.

Analysis 5 determines the aggregated validity of the portable WIM data. In this analysis, the axle load spectra in the data bins of the portable WIM system are reduced to several parameters, which are then tested in comparison to data from WIM station 99. This analysis considers both calibrated and uncalibrated data, data adjusted for impacts from temperature and speed, and data that is autocalibrated throughout the installation. Sensor coverage recommendations also inform the findings from this analysis.

Each of these analyses uses data from one or more of the 3 portable WIM installations outlined in Table 6 and described in sections 3.2.1 and 3.2.2.

3.4.1 Erroneous Vehicle Records

This analysis examines the differences in the rate of erroneous vehicle records between the three installations. During the three portable WIM installations, when a vehicle is not properly registered by the sensors, an error record is generated, indicating that a vehicle passed the sensors but not recording its axle loads. The WIM software used generates various error codes for various ways that a vehicle may not register properly in the sensors; the relative rates of these error codes may have implications for recommendations for best installation practices. Additionally, the relative rates of successful vehicle pairing are also analyzed for differences between the installations, as the rate of successful pairing by vehicle axle count can explain how erroneous records affect different types of vehicles.

3.4.2 Calibration Drift

WIM systems are known to drift in calibration over time, and so analyses of a portable WIM system's validity must consider the time elapsed since the calibration. This analysis investigates the portable WIM system's calibration drift, using a combination of correlation and regression analysis to determine which variable best explains the observed calibration drift over the course of the portable WIM's installation. Three variables are considered:

1) Time elapsed since calibration,
2) Number of vehicles passing over sensors since calibration, and
3) Number of trucks passing over sensors since calibration.

Regression analysis was used to determine the best explanatory variable; the strength of the correlation between each of the variables and the front axle loads of 3-S2, 3-S3, 3-S3-S2, and 3-S2-4 vehicles is used as the basis for selecting one variable as the best option; Appendix D displays the algorithm used to identify these vehicle configurations. Table 7 displays these configurations with their FHWA classifications and schematics.

Table 7 - Truck configurations used in specific analyses

Configuration	FHWA Classification	Schematic
3-S2 (5-axle tractor semitrailer)	Class 9	
3-S3 (6-axle tractor semitrailer)	Class 10	
3-S3-S2 (8-Axle B-Train)	Class 13	
3-S2-4 (Turnpike Double)	Class 13	

The correlation analysis considers only the front axles of these configurations—referred to as the 'selected' configurations—as previous research has shown the first three of these configurations' front axle loads to be very consistent in Manitoba, and therefore appropriate for isolating the effects of individual factors (Tan, 2002). 3-S2-4 vehicles have been included as they are particularly common at the two portable WIM installation sites used in this research, and typically have front axle loading similar to 3-S2s.

The outcomes of this analysis are twofold: first, further analyses consider the impact of time and calibration drift in their results, and so must separate the data into bins to consider this impact. The selected explanatory variable is used as the basis for the separation of these bins. Second, the explanatory variable selected as the best is considered when making a recommendation about the optimal duration of a portable WIM installation.

As WIM data changes most significantly immediately following a calibration, only data from Installation 1 is used in this analysis, as that installation had a calibration performed at the beginning and had a sufficient sample sizes for its data bins.

3.4.3 Temperature and Speed Sensitivity

All WIM systems have issues with error in their data; this error is introduced to the data in several ways. Errors can be considered to be either ‘systematic’, where a trend to values that are consistently high or low is present, ‘random’, where no trend is present, or a combination of the two. Figure 8 provides a conceptual visualization of how the sources of error in WIM measurements are added to the data.

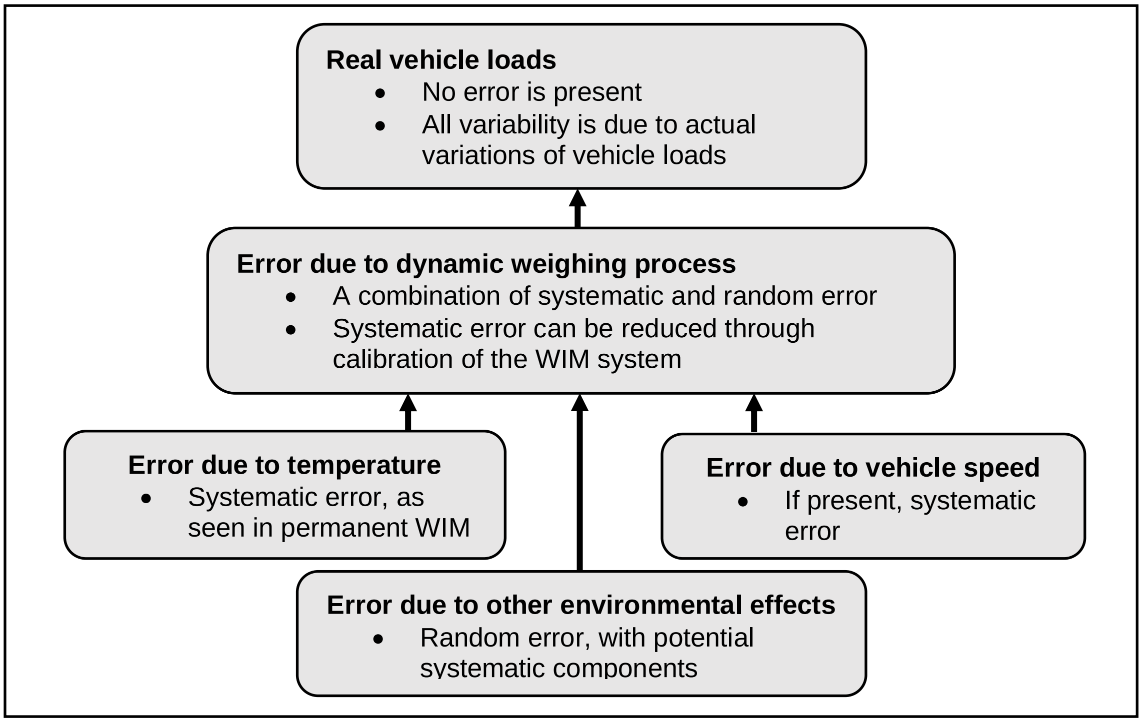


Figure 8 - Conceptual visualization of sources of WIM error

It is not possible to remove all error from the WIM data so as to reveal the real vehicle loads, as there is random error inherent in the dynamic weighing process and from environmental effects that are not measured. However, corrections to the data based on the measured temperature and vehicle speed could eliminate these sources of systematic error, and therefore improve the quality of the WIM data. This analysis investigates the effects of temperature and vehicle speed on the portable WIM axle load data. The analysis seeks first to determine if adjustments to the axle loads based on either or both of these variables will make a significant difference in the validity of the portable WIM system's data, and second, if so, what the nature of those adjustments should be.

Only front axle weights of the four selected truck configurations (see section 3.4.2) were considered in order to use their consistent load values to isolate the effects of the two variables. Regression analysis of these front axle weights and the variables are used to reveal if a statistically significant relationship exists between the axle loads and other variables. If a relationship does exist between the axle loads and one or both of the variables, the relevant regression equation is used to generate correction factors that can be applied to reduce or eliminate the impact that the relevant variable(s) has on the load measurements. In a potential case where speed is a predictor of axle loads, separate regression equations are constructed for each configuration's axle groups to examine how speed affects the different axle groups separately, as it is possible that dynamic effects are systematically experienced differently for the different axle groups of individual vehicles.

If a relationship between either of these variables and the portable WIM axle loads exists for an entire installation, it is likely that this relationship changes as the portable WIM system's calibration drifts. To examine whether this is the case, the root mean square error (RMSE) of each day of data is calculated using the regression equation from the relevant variable and the actual data, and the resulting values are examined for trends. The change of these relationships as the system's calibration drifted is further examined by binning the data and calculating separate regression equations for each bin, then comparing the resulting equation parameters.

This analysis uses data from all three portable WIM installations, but the examination of the relationships between temperature and/or speed and the calibration drift only uses data from Installation 1.

3.4.4 Per-Vehicle Data Validity

This analysis investigates the validity of the portable WIM data on a per-vehicle basis; the validities of individual vehicle axles, axle groups, and gross vehicle weights (GVWs) are examined. Historically, portable WIM data has not achieved per-vehicle accuracies sufficient for direct application to design and planning applications according to the standards set in ASTM E1318 (see Table 1). This analysis considered four portable WIM datasets from three portable WIM installations:

1. *Calibrated, WIM-Compared Data*

   This dataset, collected from portable WIM Installation 1 (at WIM station 99), has a large sample size and had a calibration performed on the portable WIM system at the beginning of the installation. A simultaneous calibration was performed on WIM station 99 at the beginning of this dataset's collection, ensuring the validity of the comparison WIM data. This is described further in Appendix B.

2. *Uncalibrated, WIM-Compared Data*

   This dataset, collected from portable WIM Installation 2 (at WIM station 99), has a large sample size and had no calibration performed on the portable WIM unit, though the calibration performed on WIM station 99 was determined to have resulted in valid comparison data in Appendix B.

3. *Calibrated, Static-Scale Compared Data*

   This dataset, collected from portable WIM Installation 3 (at the Headingley weigh scale), has a small sample size and has the post-processed calibration procedure performed on it (see Appendix C).

4. *Uncalibrated, Static-Scale Compared Data*

   This dataset, collected from portable WIM Installation 3 (at the Headingley weigh scale), has a small sample size and has no calibration performed on it.

In order to achieve the highest quality data and thereby demonstrate the maximum capability of the portable WIM system, statistics were calculated for the portable WIM data for three data processing methods:

1. *Uncorrected Data*

   Uncorrected data has not been processed beyond any initial calibration that has been performed; therefore, this analysis considers data as it is typically considered coming from a permanent WIM station. All 4 datasets are considered in their uncorrected form.

2. *Corrected Data*

   In this analysis, 'corrected data' refers to data that has been adjusted to compensate for the effects of temperature and/or speed using the regression models developed in section 4.3 . All 4 datasets are considered in their corrected form.

3. *Autocalibrated Data*

   Autocalibration is a procedure wherein WIM data records the loads of a sample of specific vehicles (for example, FHWA class 9 vehicles), then checks those loads against reference values, generating an adjustment factor to make the average sample value equal to the reference value. The adjustment factor is then applied to vehicles until the next check, at which point the factor is re-calculated (Burnos and Gajda, 2020). Though there are issues regarding validation of autocalibration systems, it can be a useful tool for improving WIM load measurement accuracy at high truck volume sites (Selezneva and Wolf, 2017; U.S. Department of Transportation, 2018a). As autocalibration removes the need for physical calibration, datasets 1 and 2 are analyzed as a single WIM-compared autocalibrated dataset and dataset 4 is analyzed as the only static scale-compared autocalibrated dataset.

This analysis compares the quality of the data between the four installation datasets and the three processing methods. These comparisons are made through histograms of percent error and several key statistics calculated on the accuracies of each vehicle:

- The mean error serves as an overall indicator of the data's accuracy. A mean of 0% is the best possible indicator of accurate data.

- The median error is more resistant to outliers than the mean. A median of 0% is the best possible indicator of accurate data.
- The standard deviation of errors indicates how precise the data is; in theory, it is unaffected by initial calibration. A standard deviation of 0% indicates perfectly precise data.
- The percent of records in compliance with the ASTM type II accuracy standard provides a more applicable measure of data accuracy. Type II is used as it has the largest accuracy tolerance of the ASTM specified WIM data types. Compliance of 100% indicates perfectly accurate data.
- A statistic labelled ‘95% of Results Within’ provides a general look at the data quality, considering both accuracy and precision. As the ASTM standard specifies that 95% of results are within the specified percent error range, this statistic states the percent error tolerance that 95% of portable WIM records are within. For example, ‘95% of Results Within’ of 50% indicates that 95% of records have percent error within ±50%.

The percent error histograms and key statistics are calculated for GVW, axle group loads, and individual axles separately. The GVW and individual axle analyses are conducted on all vehicles with three or more axles, which focuses the analysis on trucks, though this method does include some non-truck vehicles. The axle group analyses are conducted on the four selected classes described in Table 7 on page 68 to ensure that axle groups are grouped correctly.

Additionally, this analysis considers how the per-vehicle accuracy changes over the course of installations 1 and 2 by calculating some of the above statistics daily, rather than for the installation as a whole.

3.4.5 Aggregated Validity

This analysis investigates the validity of the portable WIM data in aggregate; this analysis considers whether aggregating data into several vehicle loading scenarios may result in random errors cancelling each other out and some systematic errors being isolated to individual loading scenarios, thus producing a more valid representation of axle loads than the per-vehicle data. This aggregated data may then be appropriate for indirect application to engineering and planning applications. This indirect application would examine specific parameters of loading scenario axle load spectra, rather than individual vehicle records. This is to overcome the historical inability of portable WIM systems to achieve individual vehicle accuracies at, or even near, the standards set in ASTM E1318 (see Table 1).

The specific parameters of each axle load spectra were calculated using Gaussian Mixture Models (GMMs), which separate uneven distributions into a specified number of 'mixing components', with each component having the form of a Gaussian (normal) distribution. This approach to axle load spectra analysis was developed by Regehr et. al. (2020). Figure 9 presents a graphical representation of this process as developed in research by Regehr et. al. (2020) (used with permission).

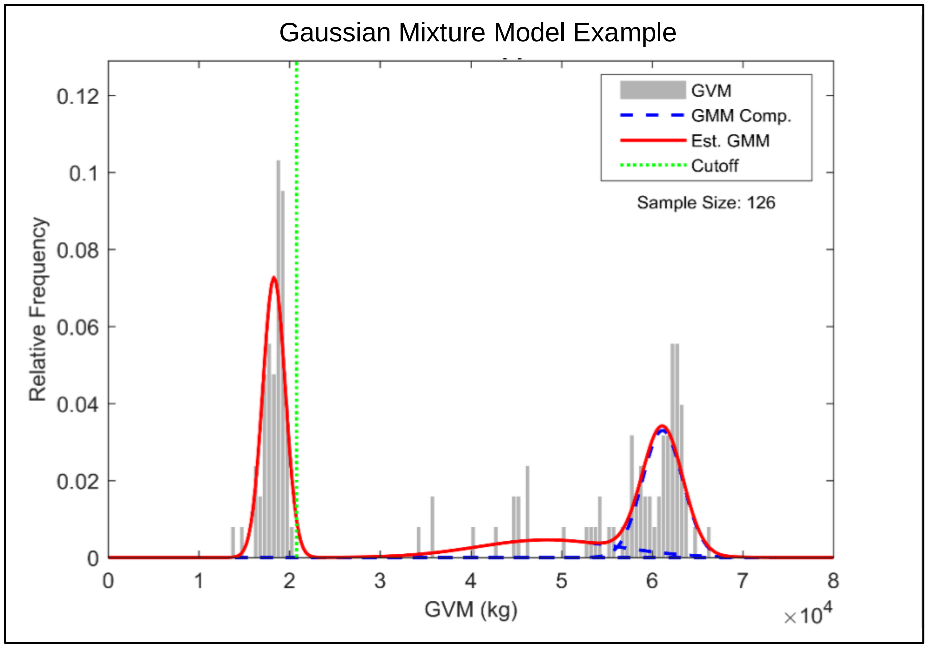


Figure 9 - Example of a Gaussian mixture model applied over a gross vehicle mass (GVM) distribution

GMMs were applied to axle load distributions that were derived from the front, tandem, and tridem axle groups as well as GVW of each of the four selected vehicle configurations. Models of the front axle load distributions used a single mixing component, while models of all other axle group loads and GVW used three mixing components to represent unloaded, partially loaded, and fully loaded vehicles. These models were done for the same four portable WIM datasets for each of the same three data processing methods as the analysis of per-vehicle accuracy (see section 3.4.4).

After calculating the GMMs for the portable WIM and comparison datasets, each individual normal distribution are compared in three ways:

1. The percent difference of the distributions' means are compared.

2. The percent overlap of the distributions' 95% confidence intervals are compared.
3. t-Tests of the normal distributions determine whether the portable and comparison datasets are statistically significantly different. It is expected that most or all of the distributions will be significantly different, as even a small systematic error would indicate significant difference for a sample of sufficient size.

By comparing the distributions in these three ways, this analysis seeks both to determine the general level of data quality in the same way as the analysis of per-vehicle accuracy and to discover any ways that using axle load data in this form may be particularly suited to any potential application. Additionally, this analysis has the potential to reveal aspects of the data validity that the per-vehicle analysis does not; any aspects of the accuracy that can be further explored using insights from the GMM process and analysis will be done so in this analysis.

4 ANALYSIS

This chapter presents the results of the analyses performed on the portable WIM data and evaluates the quality of the data against the relevant standards.

4.1 ERRONEOUS VEHICLE RECORDS RESULTS (ANALYSIS 1)

This section presents the results of the analysis of the rates of erroneous vehicle records and successful vehicle pairing. In this analysis, the relative rates of erroneous vehicle records are compared between all the three installations; due to Installation 3 undergoing a separate vehicle pairing procedure, the rates of successful vehicle pairing are only be compared between installations 1 and 2. This analysis examines whether these statistics highlight any differences between the full lane sensor installation (Installation 1) and the half lane sensor installation (Installation 2), as well as whether they have any implications for recommended installation practices.

4.1.1 Percentage of Erroneous Records

During the portable WIM installations, not all WIM records were successfully recorded. The WIM software assigned erroneous records a number of error codes. Table 8 summarizes the rate and type of these error codes for the three portable WIM installations.

Table 8 - Error code summary for portable WIM Installations 1 and 2

Error Code	Percentage of Records			Meaning
	Installation 1 (Full Lane Sensor)	Installation 2 (Half Lane Sensor)	Installation 3 (Half Lane Sensor)	
0	95.91%	71.70%	96.40%	No error
5	0.00%	0.01%	0.00%	Vehicle too fast to record
6	0.39%	0.00%	2.44%	Unequal axle count on all sensors
9	0.03%	0.26%	0.01%	Maximum axle count exceeded
10	0.00%	0.00%	0.00%	Zero axles detected
11	0.31%	6.59%	1.14%	Only one axle detected
12	3.36%	21.44%	3.16%	Vehicle too slow to record
Total Errors	**4.09%**	**28.30%**	**3.60%**	

Table 8 shows that Installation 2 had a higher rate of erroneous records than either Installation 1 or Installation 3. For all three installations, most erroneous records were a result of the vehicle being too slow to record, though a significant number of records during Installation 2 were erroneous due to only one axle being detected. While a number of variables separate Installation 3 from Installations 1 and 2, such as location, vehicle pairing method, and presence of the static weigh scale, which caused many trucks to change lanes at or near the portable WIM system, the error rates during Installation 3 are similar enough to those of Installation 1 to show that the sensor lane coverage was not likely the cause of the increased error rates.

To investigate the cause of the increased error rates during Installation 2, the rate of each error code for each day of both Installation 1 and Installation 2 are calculated. Installation 3 is omitted due to its separation in time from the other two installations and short duration. Figure 10 shows a graph of daily error code rates for both installations.

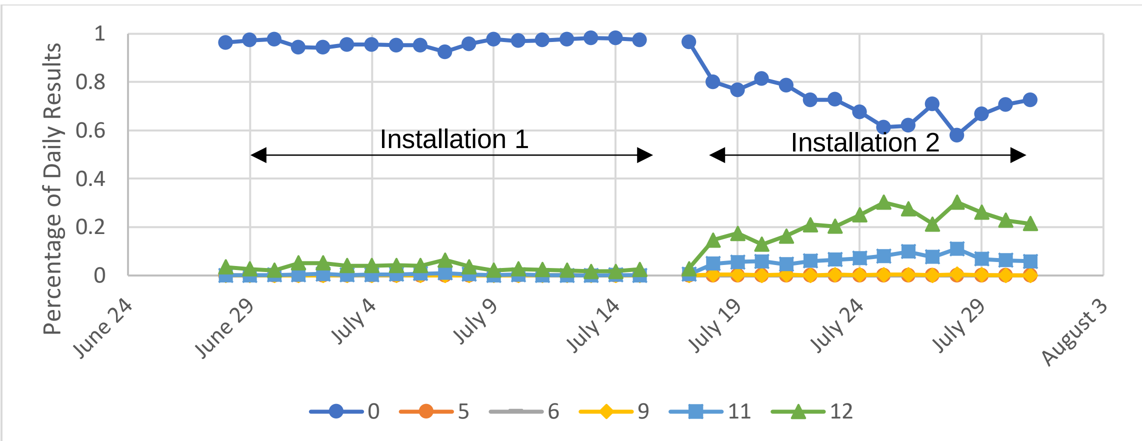


Figure 10 - Scatterplot of daily error code rates during portable WIM installations 1 and 2

Figure 10 shows that the increased rate of error codes, especially codes 11 and 12, persists in approximately equal magnitude through all of Installation 2 with the exception of the first day of the installation, which shows an approximately equal rate of errors to Installation 1, which is consistent for its entire duration. This indicates that something may have happened to the installation after the first day of Installation 2 to cause the increased rate of error codes.

4.1.2 Percentage of Successfully Paired Records

Erroneous vehicle records have their data rejected and thus have 0 axle loads recorded. This prevents the erroneous records from being assigned to vehicles of a particular class or axle count. However, as this means that erroneous vehicle records could not be paired during the vehicle pairing process, the proportion of vehicles recorded in the comparison data that were able to be successfully paired to the portable WIM data can show how the erroneous records occurred for vehicles of each axle count. Figure 11 displays a graph of the percentage of vehicles recorded at WIM station 99 that were successfully paired to the portable WIM data for installations 1 and 2 for each axle count. Due to the different vehicle pairing procedure used, Installation 3 is not included.

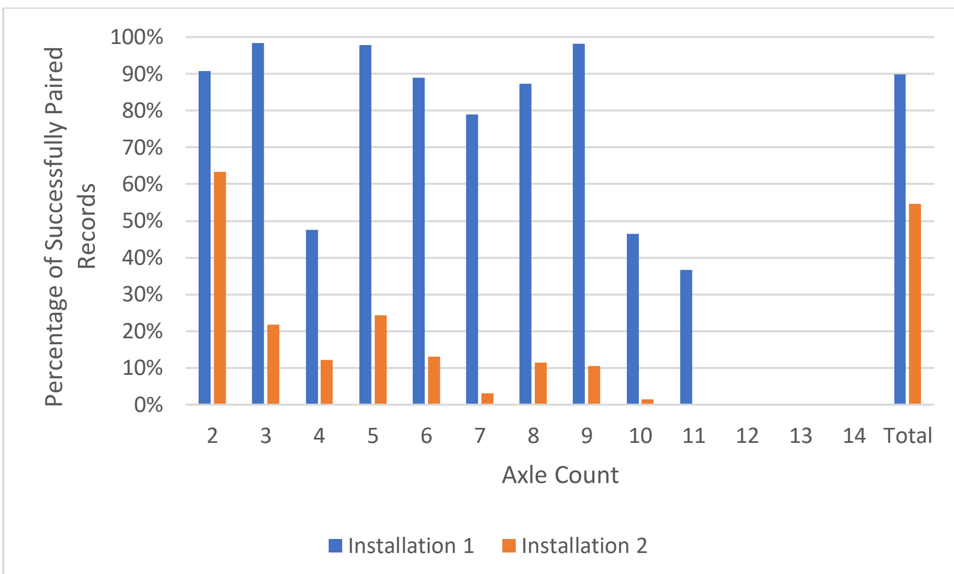


Figure 11 - Percentage of successfully paired vehicle records by axle count for portable WIM installations 1 and 2

This graph shows that while both installations had a relatively high rate of successful pairings for 2 axle vehicles, a lower percentage of vehicles with 3 or more axles were paired during Installation 2. This indicates that the increased rate of error codes had a strong impact on the truck data that a portable WIM installation would consider most valuable.

4.2 CALIBRATION DRIFT RESULTS (ANALYSIS 2)

This section presents the results of the analysis of the portable WIM's calibration drift during the installation at permanent WIM station 99 from June 27 to July 15, 2019 (Installation 1).

To provide a baseline context, the scatterplot in Figure 12 illustrates the front axle loads of 3-S2, 3-S3, 3-S3-S2, and 3-S2-4 vehicles (referred to as the 'selected' configurations) over the course of the installation. The calibration drift is seen in a non-horizontal trend in the portable WIM front axle loads over the course of the installation. In comparison, the front axle loads of the same vehicles as recorded by WIM station 99 exhibit no such trend.

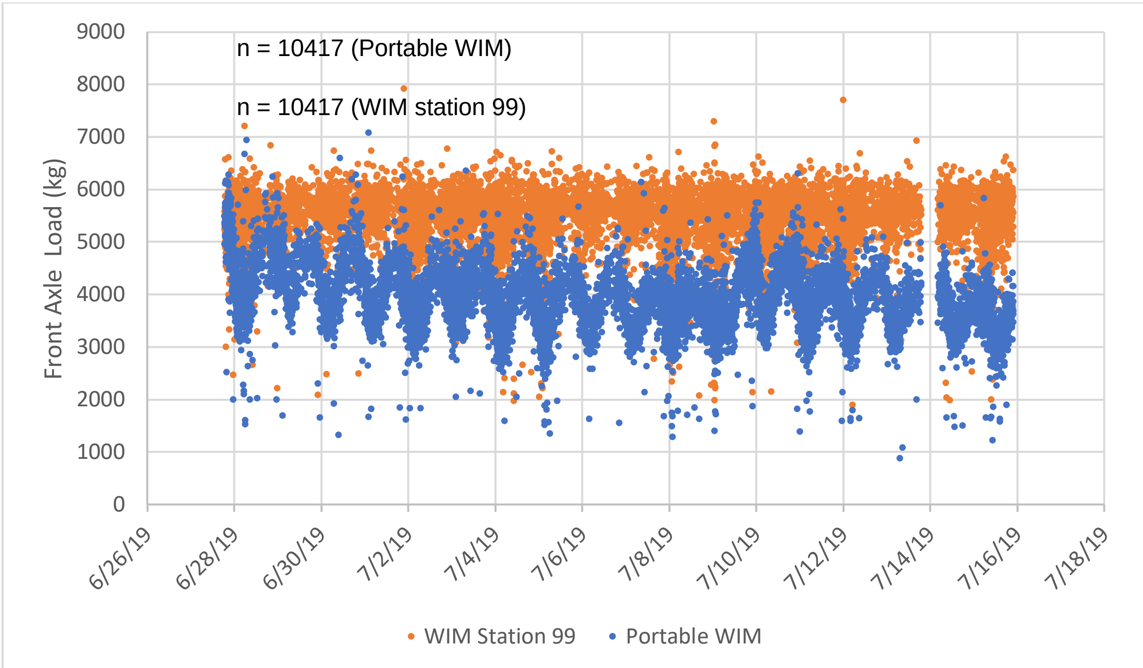


Figure 12 - Calibration drift of front axle loads of selected configurations during first portable WIM installation at WIM station 99

Several important observations about the front axle loads can be made from Figure 12:

1. The portable WIM front axle loads fluctuate regularly over the course of a day, likely due to changes in temperature, while the WIM station 99 loads do not.
2. The portable WIM front axle loads exhibit a downward trend over the first half of the installation, while the WIM station 99 loads do not.
3. The downward trend in portable WIM front axle loads does not appear to be consistent over the course of the installation.

4.2.1 Duration of Calibration Drift Trend

The third observation from Figure 12 is significant to this analysis as the goal is to determine what variable best explains the change in portable WIM calibration. To do this, the analysis must determine over what period the portable WIM's calibration drift trend was consistent, so that this consistent relationship can be compared to the candidate variables. To examine this, the selected front axle load data is divided into 24-hour periods, and correlation analyses are conducted on datasets that each include all records from the beginning of the installation to the end of one 12-hour period. Sample sizes for these correlation analyses ranged from 447 to 10331.

Figure 13 shows the scatterplots that were the basis of the correlation analysis of the front axle load samples that were taken for the first 1, 5, and 10 days. These scatterplots illustrate that as long as the calibration drift trend is consistent, a greater number of vehicle records will result in the daily cyclical pattern having less influence on the overall trend, the overall calibration drift trend having greater influence on the overall trend, and the R^2 value being higher due to a better fit for a linear trendline.

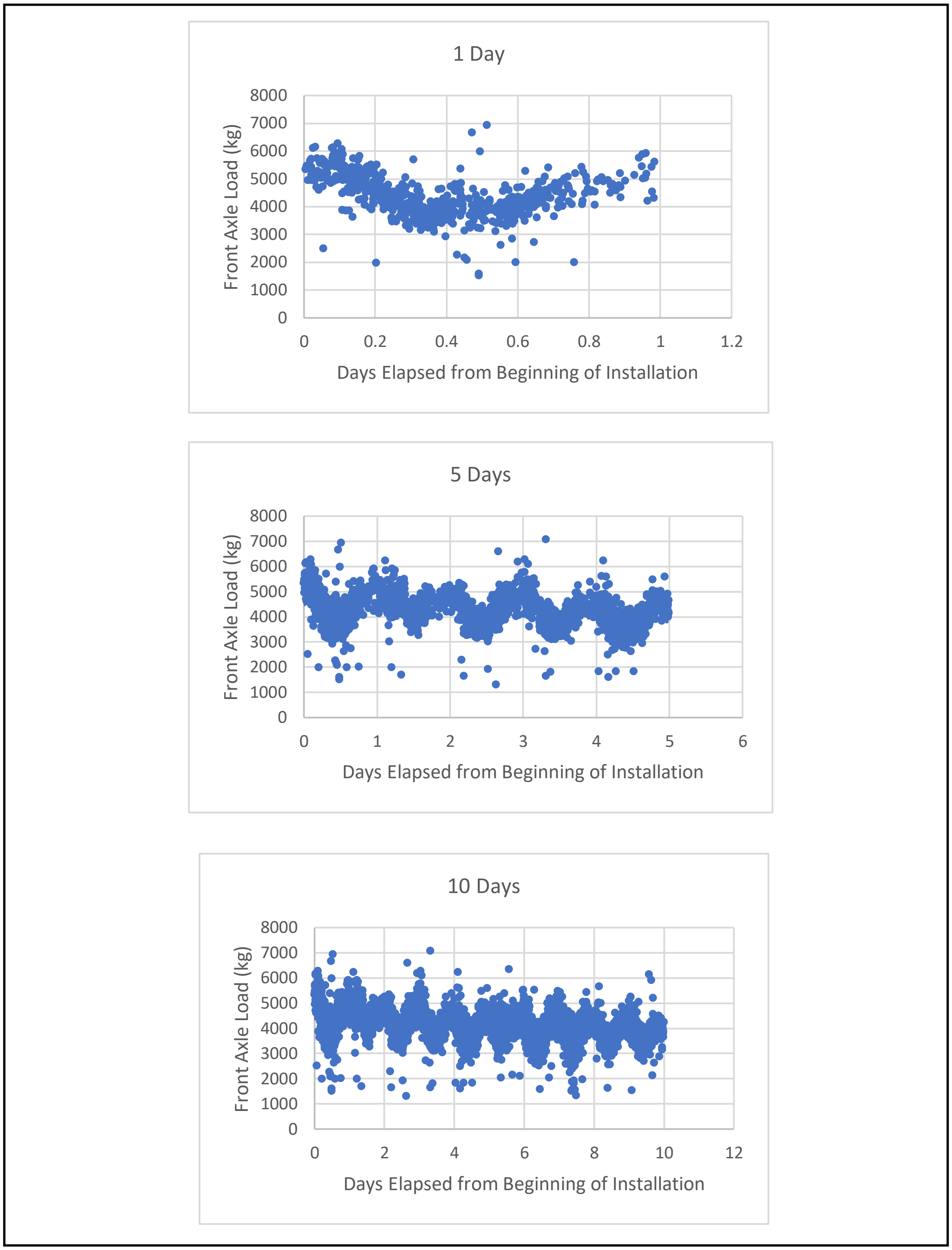


Figure 13 - Example scatterplots for analysis of calibration drift trend

The scatterplots in Figure 13 display the data behind the correlation analysis of the front axle loads. However, two correlations are calculated for each dataset:

1. Time elapsed vs. portable WIM front axle load, and
2. Time elapsed vs. portable WIM front axle percent error as compared to WIM station 99.

These two analyses are expected to be similar, as WIM station 99 showed no signs of drift from the calibration that was performed at the beginning of the installation. Therefore, while the front axle load correlation accounts for all variability in the loads, including variability of the actual loads, the percent error calculations correct for the variability in the actual loads, and the correlation of the front axle percent error accounts for only variability due to one or more sources of error.

The R^2 values from these correlation analyses are plotted with the expectation being that after some initial instability, similar correlations will be observed for the duration of the initial calibration drift trend, after which the R^2 values will drop and a second trend would take its place at the lower R^2 values. R^2 values are expected to be low, as the variability in the front axle loads and front axle load errors is primarily due to a combination of WIM errors and variability in the actual loads, with only a small portion of the variability being due to the calibration drift. However, any R^2 values of non-negligible magnitude indicate that the correlation does explain some portion of the variability in the front axle loads or errors. Therefore, this analysis is not primarily concerned with the absolute magnitude of the R^2 values but with their relative magnitude, specifically how the value changes as the time bins get longer. Figure 14 shows the plot of R^2 values.

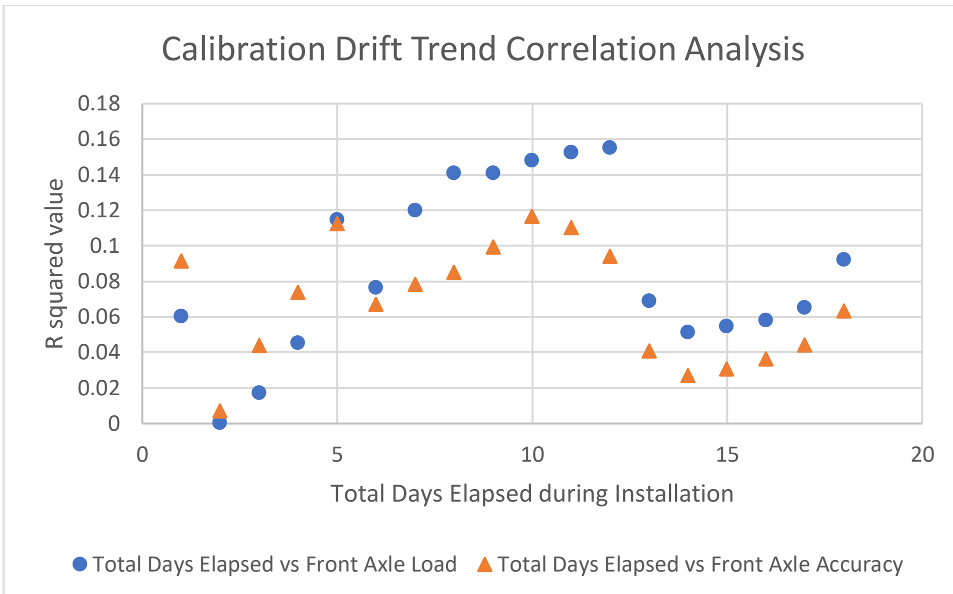


Figure 14 – Scatterplot of R squared values of time elapsed vs portable WIM front axle load and front axle load percent error during portable WIM installation 1

These results show that, after some initial instability, R^2 values for both front axle loads and percent errors stabilize for the samples taken of 7 days to 12 days, after which there is a decrease in correlation of 55% for the load measurements and 56% for the percent error measurements from day 12 to 13. The correlation values for samples of 13 days long or longer are consistent for both load and percent error measurements. This indicates that the calibration drift relationship is consistent for the first 12 days of the installation, after which the daily change in calibration becomes less consistent and the calibration drift no longer explains as large a portion of the variability in the front axle loads and errors.

Based on these findings, the rest of the analyses considering the effects of the calibration drift will only reference data from the first 12 days of the installation.

4.2.2 Selection of Explanatory Variable

This analysis considered three variables as possible explanations for the calibration drift:

1) Time elapsed since calibration,
2) Number of vehicles passing over sensors since calibration, and
3) Number of trucks passing over sensors since calibration.

To explain the calibration drift, regression models of these variables are compared by their strength. However, first, to determine whether models with more than one of these variables should be considered, a multicollinearity analysis is conducted on the three. Sample sizes for this analysis were all 6657. Table 9 displays the results of this analysis.

Table 9 - Results of multicollinearity analysis of candidate explanatory variables for calibration drift

	Time Elapsed	Vehicles Passed	Trucks Passed
Time Elapsed	1		
Vehicles Passed	0.9922346	1	
Trucks Passed	0.9882845	0.9990901	1

These results show extremely high multicollinearity (>0.98) for all 3 variable pairs; therefore the analysis will only consider single-variable models. It is expected that the three single-variable models would have similar R^2 values; that is, they will all predict the same proportion of the variability in the selected front axle loads. To confirm this, the three regression equations are calculated and their R^2 values compared. As in the investigation into the duration of a consistent calibration drift trend, both the front axle loads and the front axle percent errors of the selected classes are used to construct the equations. Linear least-squares regression models are used as a linear trend can be visually observed in the scatterplot in Figure 12. Sample sizes are 6657 for all analyses, as all three use the first 12 days of data from the installation. Table 10 presents the results of the regression analyses.

Table 10 - Results of regression analyses of front axle loads and front axle load percent errors on candidate explanatory variables

Variable	Y Variable	Regression Equation	R^2	P-value of F Statistic
Time Elapsed	Load	$Y = -67.372*X + 4317.2$	0.155	<0.001
	Percent Error	$Y = -0.011506*X - 0.20660$	0.0937	<0.001
Vehicles Passed	Load	$Y = -0.0173106*X + 4296.5$	0.150	<0.001
	Percent Error	$Y = -0.00000297*X - 0.20978$	0.0917	<0.001
Trucks Passed	Load	$Y = -0.071425*X + 4296.7$	0.152	<0.001
	Percent Error	$Y = -0.0000122*X - 0.21022$	0.0915	<0.001

These results show that:

1. The regression models are able to explain a greater percentage of the variability in the front axle loads (~15%) than of the front axle percent errors (~9%).
2. All parameters of all equations have a statistically significant F statistic, which measures the overall significance of the regression model (p-value <0.05), indicating that despite the low predictive power of the models, the predictions the models can make are valid.
3. As expected, there are no significant differences between the R^2 values of the three candidate explanatory variables for both the front axle load and front axle percent error regression models.

As none of the variables are demonstrably superior to the others, the time elapsed since the calibration is used as the explanatory variable for the calibration drift in all further analyses.

4.3 TEMPERATURE AND SPEED SENSITIVITY RESULTS (ANALYSIS 3)

This section presents the results of the analysis of the relationships between the temperature and speed and the WIM axle load measurements. All 3 portable WIM installations were considered.

4.3.1 Temperature Sensitivity Analysis

Selection of Temperature Dataset

During both portable WIM installations at WIM station 99, two sources of temperature data were collected: temperature from an in-pavement sensor at

WIM station 99, and temperature from a sensor left open to the air on the portable WIM unit. These two sensor installation methods were expected to record different temperature values, as they are recording different parameters: pavement temperature and air temperature, respectively. While the open-air sensor was easier to install, if the in-pavement sensor provides sufficiently higher quality temperature data, future portable WIM installations could install the temperature sensor in-pavement. Therefore, this section assesses the relative quality of the in-pavement and open-air temperature measurements.

The scatterplot in Figure 15 displays the temporal temperature patterns as measured by both temperature data sources during both portable WIM installations at WIM station 99.

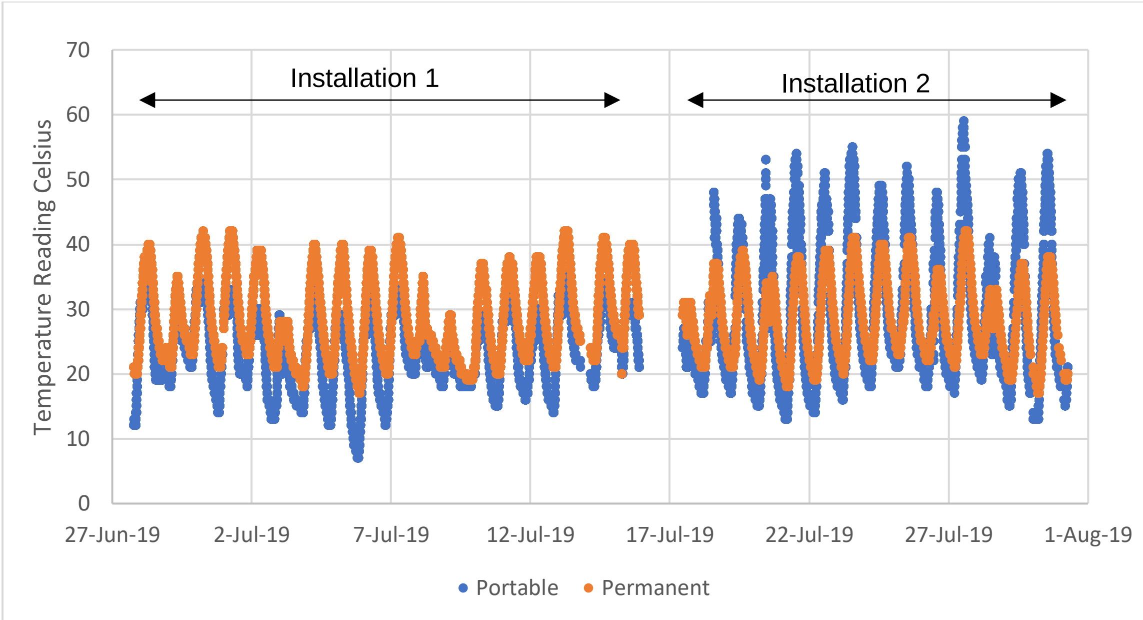


Figure 15 - Temperature measurements from WIM systems during portable WIM installations at WIM station 99

The scatterplot in Figure 15 reveals two characteristics of the WIM temperature data:

1) The in-pavement and open-air temperature sensors record different measurements, and
2) The open-air sensor temperature data exhibits a significantly different temperature pattern during the second portable WIM installation, despite the in-pavement sensor temperature data exhibiting consistent patterns.

As the daily high and low ambient and in-pavement temperatures were approximately the same, the most probable explanation for the second characteristic is that the portable WIM sensor was placed differently during the second installation, making it more susceptible to heating and cooling throughout the day. However, as the exact placement of the sensor was not documented, this cannot be confirmed. Regardless, the two open-air sensor installations are individually assessed.

The temperature data that is considered ‘best’ in this assessment is the data that predicted the greatest proportion of the changes in front axle load measurements, as the greater this predictive power was, the more effective the temperature data would be for correcting errors in the axle load measurements. Quadratic least-squares regression analysis is used to determine the relative predictive power of each temperature data source. A quadratic model is used to ensure a better fit to potential non-linear relationships and therefore have greater predictive power. To most effectively isolate the effects of temperature, only the front axles from the four selected vehicle configurations are used in the regression analysis, as past research has shown them to be relatively consistent. The regressions in section 4.2 examine how well temperature and speed influence the observed variability in front axle loads. Since these loads still vary, despite have general consistency, the R^2 values of the regression models are expected to be low, as temperature and speed are expected to be poor predictors of the actual loads when other sources of variability are not considered (see Figure 8 on page 70 for a conceptual illustration of the sources of variability in WIM front axle loads). However, any non-zero R^2 values in a statistically significant regression model show that these parameters impose some error in the axle load measurements beyond the other sources of variability. Consequently, corrections for parameters that meet these criteria are expected to improve the accuracy of the axle load measurements. Table 11 shows the results of the regression analysis.

Table 11 - Results of regression analysis of 3 temperature data sources against select front axle loads

Temperature Data Source	Regression Equation	R^2	n	P-value of F statistic
Open Air (Installation 1)	$Y = 1.1437*T^2 – 102.35*T + 5649.6$	0.2066	10417	0
Open Air (Installation 2)	$Y = -0.39745*T^2 + 8.4843*T + 2330.7$	0.1945	848	<0.001
In-Pavement (Both Installations)	$Y = 0.65236*T^2 – 73.056*T + 5273.7$	0.0972	11265	0

Table 11 shows that the temperature data from the two open air installations have approximately R^2 values and therefore have equal predictive power. Additionally, both open air temperature datasets have approximately double the predictive power of the in-pavement temperature data. This indicates that while the specific placement of the temperature sensor during a portable WIM installation affects the temperature measurements, it does not affect the temperature dataset’s ability to predict, and therefore correct for, variations in the portable WIM’s axle load measurements. Due to these findings, the temperature measurements from the portable WIM unit are used in all further analyses.

Comparison Dataset Temperature Influence

To determine the nature of the relationship between axle load and temperature, regression analyses are performed, revealing what effect the temperature has on both the front axle loads and front axle percent error. However, as the front axle percent error is calculated with reference to the axle loads as measured by the comparison datasets (WIM station 99 and the Headingley static weigh scale), this regression analysis first determines to what degree the temperature influences the axle loads in the comparison datasets. If the axle load variability in the comparison datasets can be predicted by temperature to the same extent as the portable WIM system axle load variability, the percent error calculations will effectively cancel out the temperature effect between the two datasets, and will not support meaningful conclusions about the temperature relationship.

Only the front axle loads of the four selected vehicle configurations were considered. Table 12 displays the results of the regression analysis.

Table 12 - Results of regression analysis of temperature against select front axle loads from WIM station 99

Installation	Regression Equation	R^2	n	P-value of F statistic
1	$Y = 0.66785*T^2 – 32.434*T + 5807.8$	0.00229	10417	<0.001
2	$Y = 0.20970*T^2 – 16.124*T + 5632.9$	0.00460	848	0.14
3	$Y = 1.0081T^2 – 16.640*T + 5209.5$	0.000361	551	0.91

The low R^2 values (<0.01) for all 3 installations and lack of model significance (p-value>0.05) for installations 2 and 3 indicate that the temperature cannot be used to predict the variability of the front axle loads in the comparison datasets in any consequential way. Therefore, predictions of the portable WIM load percent error with reference to the comparison datasets effectively isolate the temperature effect on only the loads from the portable WIM.

Nature of Axle Load-Temperature Relationship

To determine the nature of the relationship between axle load and temperature, quadratic least-squares regression analyses are used to determine what portion of the variability in the selected front axle loads can be accounted for by temperature. To ensure that the variability that is being accounted for is variability due to changes in temperature, rather than the natural variability in these front axle loads, regression analyses of the portable WIM's front axle load percent error as compared to the comparison datasets are also performed. Higher R^2 values in these models will lead to more effective front axle load corrections. Additionally, percent error-based models with R^2 values close to the R^2 values of the corresponding load-based models indicates that the variability being explained is, only that caused by temperature, rather than other sources of error. Table 13 shows the results of the regression analyses.

Table 13 - Results of regression analysis of temperature against select portable WIM front axle loads and front axle load percent errors

Installation	Y Variable	Regression Equation	R^2	n	Significance of F Statistic
1	Load	$Y = 1.1437*T^2 – 102.35*T + 5649.6$	0.207	10417	<0.001
	Percent Error	$Y = 0.0001149*T^2 – 0.014213*T + 0.00001149$	0.146	10417	<0.001
2	Load	$Y = -0.39745*T^2 – 8.4847*T + 2331.0$	0.194	848	<0.001
	Percent Error	$Y = -0.0000967*T^2 – 0.0032690*T -0.58919$	0.136	848	<0.001
3	Load	$Y = -2.0175*T^2 – 4.1178*T + -3734.3$	0.045	551	<0.001
	Percent Error	$Y = -0.0005724*T^2 – 0.0021581*T - 0.28381$	0.048	551	<0.001

These results show that temperature accounts for a non-negligible portion of the variability in the selected portable WIM front axle loads, though less for the portable WIM during Installation 3. All the predictive models are significant. Furthermore, as the regression models using the front axle percent error have similar predictive power to the front axle load regression models, corrections based on these models are expected to improve the accuracy of the portable WIM axle load measurements.

Effect of Calibration Drift on Temperature Relationship

Now that it has been established that there is a statistically significant relationship between temperature and the selected front axle loads, it must be determined whether the nature of the relationship changes as the calibration drifts. To analyze this, the first installation at WIM station 99 is examined, as it was initially calibrated and has a measurable calibration drift.

First, a regression model is calculated using only the first 12 days of data from installation 1, as this is the period over which the calibration drift trend is determined to be consistent in analysis 1. Figure 16 displays a scatterplot of the residuals of this regression model throughout the first 12 days of installation 1; this plot allows the relationship between temperature and front axle load to be visually assessed for changes during this period.

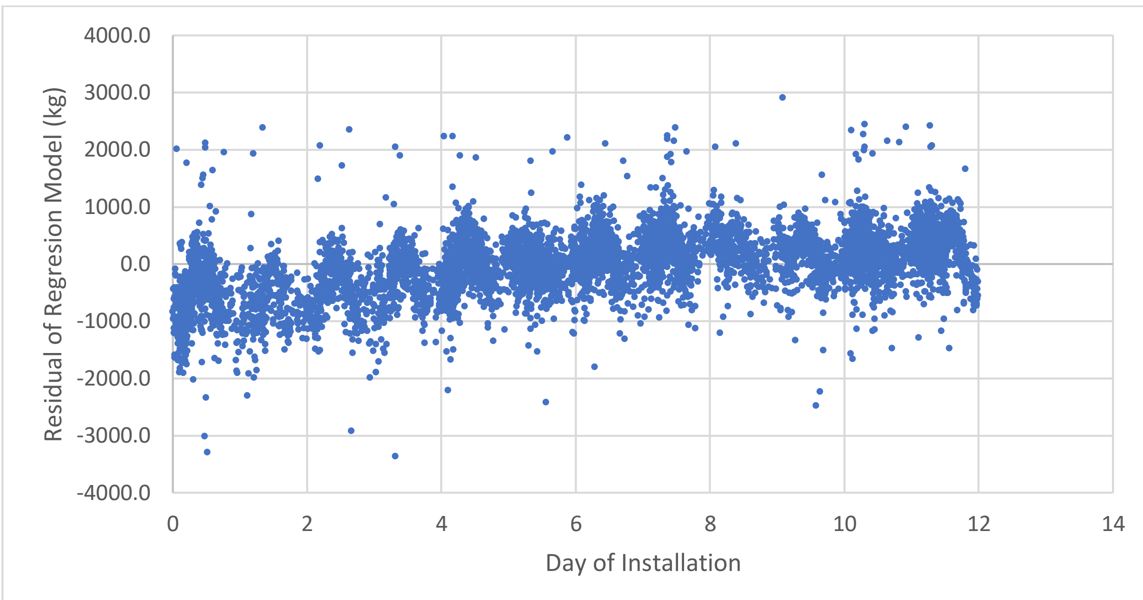


Figure 16 - Scatterplot of residuals of temperature regression model for first 12 days of portable WIM installation 1

The scatterplot in Figure 16 shows a cyclic pattern in the residuals corresponding to the daily fluctuations in load value caused by variations in temperature. This pattern can be ignored as the daily pattern is known to not fit the linear regression, and this analysis is concerned only with the approximate average residual for each day.

From Figure 16, it can be seen that the first 4 days of the installation have a different temperature-front axle load relationship than days 5 through 11. To further investigate the first 4 days of the installation, quadratic regression equations for the relationship between temperature and front axle load are calculated for each of the 12 days individually. The graphs in Figure 17 display how the coefficients and intercepts of the individual day’s regression equations change over time.

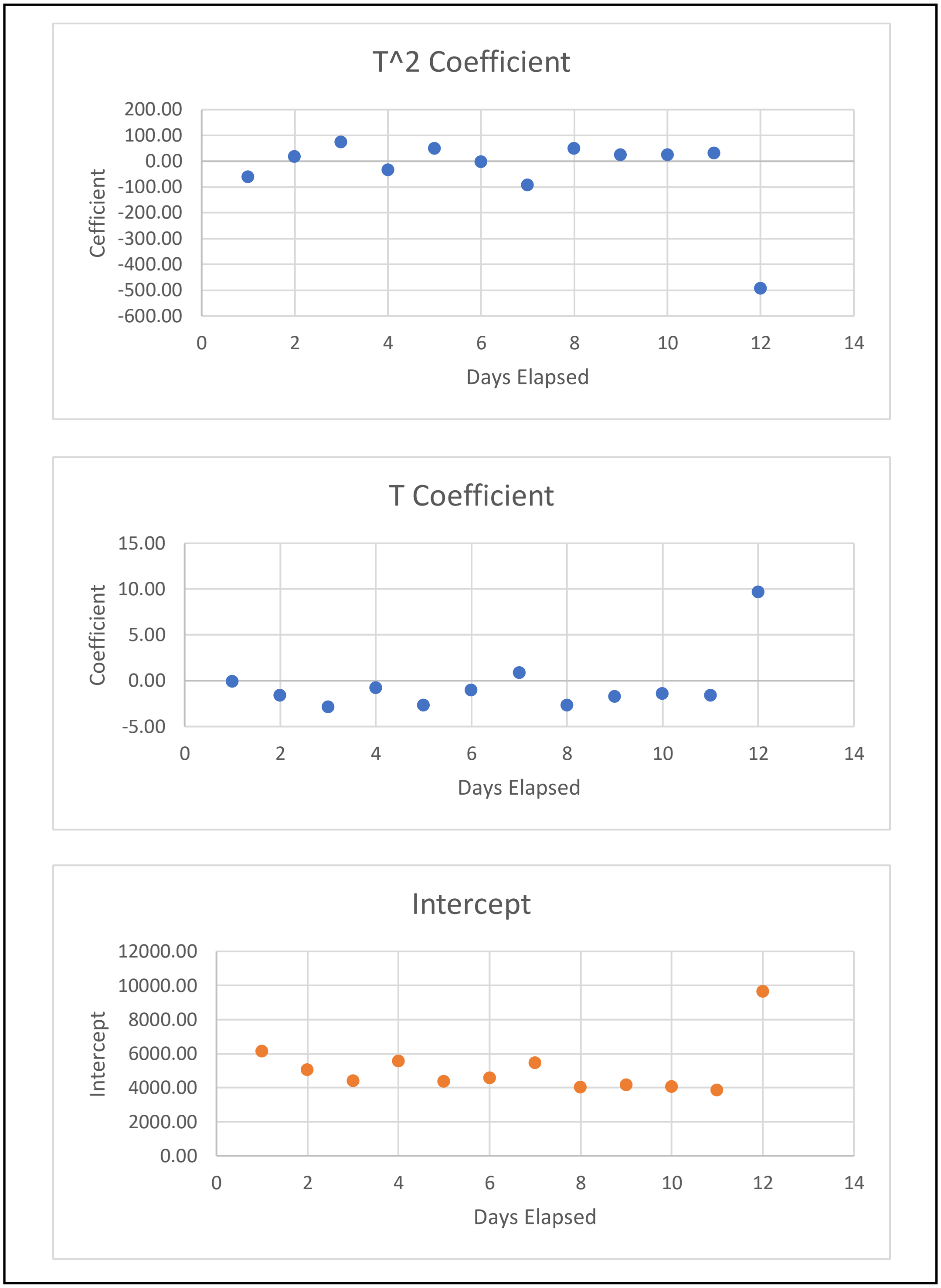


Figure 17 - Values from individual day temperature regression equations during first installation at WIM station 99

The regression equation values in Figure 17 show relatively consistent values with the exception of the 12th day, which is a clear outlier for all three equation parameters. Notably, the first 4 days of installation have similar equation parameter values to those from days 5 to 11, which indicates that while days 1 through 4 were a poorer fit for the regression model using the entire installation's data, the best possible regression equations for those days have similar coefficients and intercept to those generated by days 5 through 11. While it is unclear why the model for day 12 is such a clear outlier, the scatterplot in Figure 16 shows that this day had similar residuals to days 5 to 11 for the regression model containing all 12 days of data, indicating that it is fit approximately equally well to the all-days regression model as these days. Therefore, the outlier on day 12 will be ignored.

These results indicate that the single regression model for the entire installation cannot be substantially improved upon by time-separated regression models and that therefore the calibration drift does not have a significant impact on the temperature-front axle load relationship. Due to these findings, a single regression equation for temperature correction will be applied for the entire installation.

4.3.2 Vehicle Speed Sensitivity Results

Comparison Dataset Speed Influence

To determine the nature of the relationship between axle load and speed, regression analyses are performed using vehicle speed to predict both the front axle loads and front axle percent error. However, as the front axle accuracy is calculated with reference to the axle loads as measured by the comparison datasets (WIM station 99 and the Headingley weigh scale), this regression analysis first determines to what degree vehicle speed influences the comparison data from WIM station 99. If the axle load variability in the comparison datasets can be predicted by speed to the same extent as the portable WIM system axle load variability, the percent error calculations would effectively cancel out the effect of speed, and would not support meaningful conclusions about the speed relationship.

Only the front axle loads of the four selected vehicle configurations are considered in the analysis, and the installation at the Headingley weigh scale (Installation 3) is disregarded, as speed is not measured at a static scale. Table 14 displays the results of the regression analysis.

Table 14 - Results of regression analysis of speed against select front axle loads from WIM station 99

Installation	Regression Equation	R^2	n	P-value of F statistic
1	$Y = -0.25969*T^2 - 43.565*T + 3627.1$	0.00324	10417	<0.001
2	$Y = 0.036978*T^2 - 7.4957*T + 5738.2$	0.000510	848	0.81

The low R^2 values here (<0.01) and insignificance of the second installation's model indicate that the speed cannot be used to predict the variability of the front axle loads as measured at WIM station 99 in any consequential way. Therefore, the predictions from this analysis of the portable WIM load accuracy compared to WIM station 99 effectively isolate the effect of speed only on the loads from the portable WIM.

Nature of Axle Load-Speed Relationship

To determine the nature of the relationship between axle load and speed, quadratic least-squares regression analyses are used to determine what portion of the variability in the selected front axle loads can be accounted for by speed. To ensure that the variability that was being accounted for was variability due to changes in speed, rather than the natural variability in these front axle loads, regression analyses of the portable WIM's front axle load percent error as compared to the comparison datasets was also performed. Higher R^2 values in these models will lead to more effective front axle load corrections. Additionally, percent error-based models with R^2 values close to the R^2 values of the corresponding load-based models indicates that the variability being explained is, in fact, only that caused by speed. Table 15 shows the results of the regression analyses.

Table 15 - Results of regression analysis of speed against select portable WIM front axle loads and front axle accuracies

Installation	Y Variable	Regression Equation	R^2	n	P-value of F statistic
1	Load	$Y = -0.093167*S^2 - 24.063*S + 2463.93$	0.0114	10417	0
	Percent Error	$Y = 0.000016537*S^2 - 0.0011533*S - 0.31376$	0.0109	10417	0
2	Load	$Y = -0.11182*S^2 + 25.571*S + 796.7$	0.0146	848	0.002
	Percent Error	$Y = -0.00002641*S^2 - 0.0058888*S - 0.90330$	0.0146	848	0.002
3	Load	$Y = -0.31985*S^2 + 58.515*S - 938.01$	0.0046	543	0.29
	Percent Error	$Y = -0.0000404*S^2 + 0.0074516*S - 0.64231$	0.002	543	0.61

These results show that speed does not have a substantial impact on the front axle load measurements taken during any of the three installations. This indicates that corrections based on only the speed-front axle load relationship will not substantially improve the accuracy of the portable WIM load measurements.

However, any effect of speed on the axle loads would be due to dynamic effects that may affect certain parts of a vehicle differently, so examination of the effect of speed on other axle groups is warranted.

Axle Load-Speed Relationship for Non-Front Axles

If vehicle speed has any meaningful effect on the portable WIM axle loads, it is due to dynamic effects introduced by the bump created by the sensors. These dynamic effects may affect certain parts of a vehicle differently. To investigate this, regression models are created that predicted the variability in the axle group load error from the vehicle's speed for each axle group of the four selected vehicle configurations. These models consider axle group load error rather than the axle group loads themselves as these axle groups are known to be highly variable, and the axle load error values have the effect of reducing this variability. These models are created using only data from the first installation, which ensures that each model had as large a sample size as possible, thus giving each model the highest possible chance of achieving statistical significance. Both quadratic and linear models are created, as the nature of the speed-axle load is not known. Table 16 presents the results of these models. The models themselves are presented in Appendix E; this table presents the R^2 values of the F statistic and the coefficients (C). In the quadratic models, C1 is the coefficient of $speed^2$ and C2 is the coefficient of speed.

Table 16 - Results of regression analysis of vehicle speed on non-front axle load errors of selected vehicle configurations

Configuration	Axle Group	n	Quadratic			Linear		
			R^2	P-value of F	P-value of C1, C2	R^2	P-value of F	P-value of C
3-S2	Single (Front)	6611	0.011	<0.001	0.151, 0.440	0.011	<0.001	<0.001
	Tandem (Drive)		0.057	0	0.004, 0.192	0.056	<0.001	<0.001
	Tandem (Trailer)		0.056	0	<0.001, 0.059	0.054	<0.001	<0.001
3-S3	Single (Front)	1726	0.015	<0.001	0.731, 0.928	0.014	<0.001	<0.001
	Tandem (Drive)		0.051	0	0.100, 0.411	0.050	<0.001	<0.001
	Tridem (Trailer)		0.085	0	0.002, 0.049	0.080	<0.001	<0.001
3-S3-S2	Single (Front)	427	0.053	<0.001	0.138, 0.297	0.049	<0.001	<0.001
	Tandem (Drive)		0.15	<0.001	0.011, 0.077	0.14	<0.001	<0.001
	Tridem (Trailer 1)		0.16	<0.001	0.014, 0.101	0.15	<0.001	<0.001
	Tandem (Trailer 2)		0.15	<0.001	0.057, 0.270	0.14	<0.001	<0.001
3-S2-4	Single (Front)	1653	0.005	0.013	0.351, 0.275	0.004	0.005	0.005
	Tandem (Drive)		0.058	0	<0.001, <0.001	0.055	<0.001	<0.001
	Tandem (Trailer 1)		0.034	<0.001	<0.001, <0.001	0.030	<0.001	<0.001
	Tandem (Front of Trailer 2)		0.035	<0.001	<0.017, 0.050	0.014	<0.001	<0.001
	Tandem (Rear of Trailer 2)		0.123	<0.001	0.012, 0.052	0.113	<0.001	<0.001

These results reveal several facts about the speed-axle load relationship during the first installation.

1. All models are significant according to the F statistic.
2. The linear regression models are able to predict a similar proportion of the variability in the axle group load errors as the quadratic models, but have all significant regression coefficients, while not all the quadratic models do. This indicates that the linear models are a better fit for the speed-axle group load error relationships.
3. None of the individual configuration's front axle load models, apart from the 3-S3-S2 model, are able to predict more than 2% of the variability in the front axle load error. This confirms the results of the regression models calculated for all selected configurations combined as displayed in Table 15.
4. All but two of the models for non-front axle group loads are able to predict more than 5% of the variability in the axle group load errors (for both quadratic and linear models).

As these results show a non-trivial relationship between vehicle speed and axle load error, appropriate corrections for speed are expected to improve the accuracy of the portable WIM axle load measurements. However, in a deployment outside of the testing phase, a comparison dataset will be unavailable, so corrections to the axle load measurements must be made on the basis of regression of the axle loads themselves. Furthermore, many regression equations have been calculated for the various configurations and axle groups, and they must be compared to determine if the axle load-speed relationship is consistent enough between the various axle groups to use as the basis for corrections. Figure 18 illustrates the regression coefficients and intercepts for regression of vehicle speed against (1) axle group load error and (2) axle group load for each of the four selected configurations.

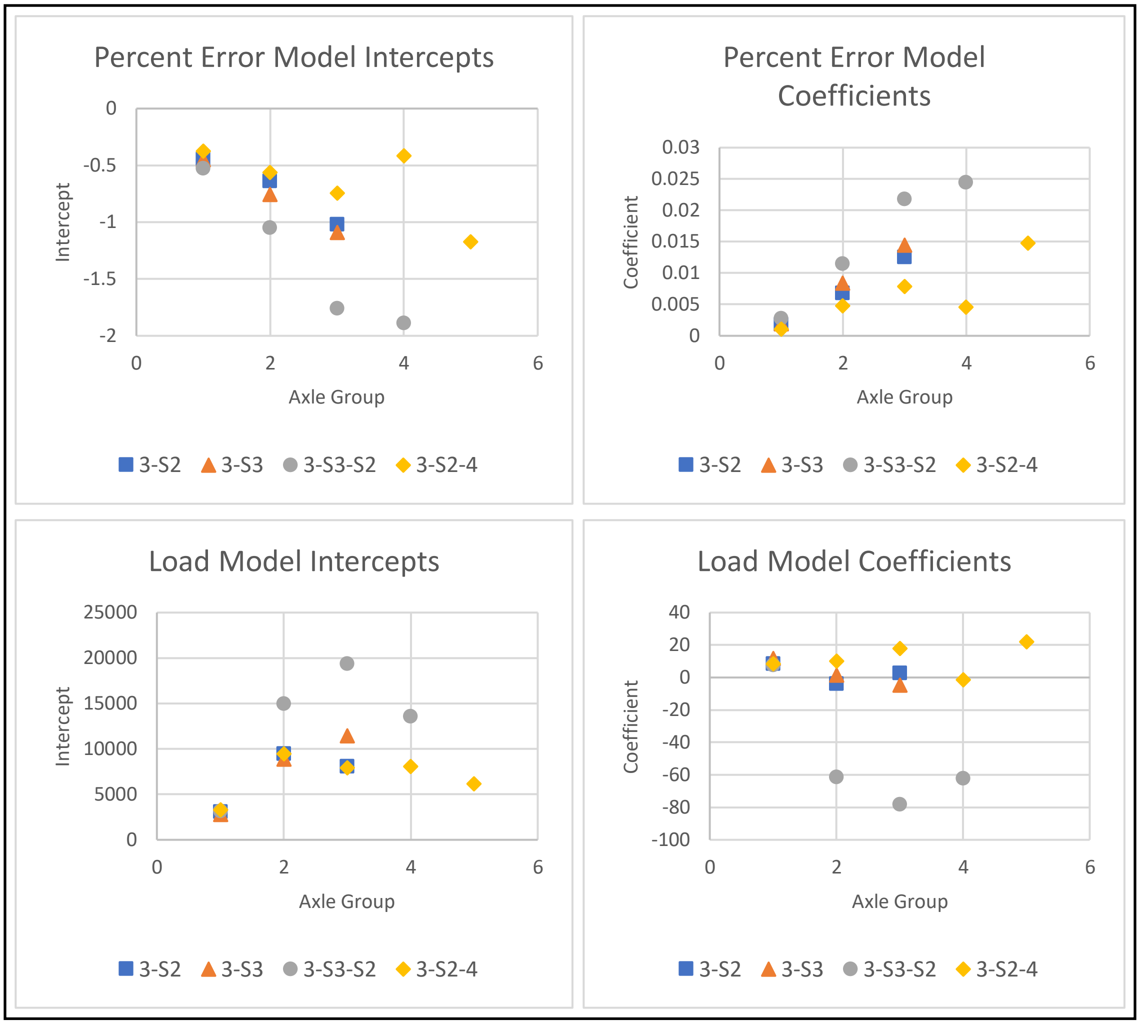


Figure 18 - Scatterplots of parameters of regression equations of speed versus axle group load errors and speed versus axle group loads

These graphs show that the regression equations of the axle group load percent errors and the axle group loads are fundamentally different. For a given configuration and axle group, the equation parameters for the axle group load percent error regression and the axle group load regression will not necessarily be proportional; the equations will not even necessarily slope in the same direction. This can be seen most clearly in the coefficients of the equations for 3-S3-S2 vehicles, which has all positive coefficients for non-front axle group percent errors, while having negative coefficients for non-front axle groups. Furthermore, the relationship between the axle group load equation parameters of the 4 selected classes is not always constant; for example, the coefficients of 3-S3 axle groups trend downwards from front to back, while 3-S2-4 axle groups trend upwards and 3-S2 axle groups exhibit a different downward trend. Additionally, 5 out of 11 regression models for non-front axle group loads do not even have a statistically significant F statistic (that is, $P<0.05$), and while the front axle group load models for both axle group load error and axle group load are all significant, none of them can describe more than 2% of the variability in the front axle loads. This means that while corrections would likely improve the accuracy of the data if they were based on the error of the axle group loads, the variability of the axle-load speed relationship between the various configurations and the differences between the regression models of axle group load error and axle group load make this infeasible. While correction factors based on the front axle loads would account for 1 to 2% of the variability in those loads, this is not guaranteed to improve the accuracy of subsequent axle group load measurements and at best would provide marginal benefits to the accuracy of the load measurements. Therefore, no corrections based on speed will be applied.

4.4 PER-VEHICLE DATA VALIDITY RESULTS (ANALYSIS 4)

This section presents the results of the analysis of the portable WIM data's validity on a per-vehicle basis. In this analysis, all statistics were calculated on a per-vehicle basis for single axle loads, axle group loads, and GVWs. This analysis assesses the data's accuracy as outlined in ASTM E1318: Standard Specification for Highway Weigh-In-Motion (WIM) Systems with User Requirements and Test Methods. ASTM E1318 outlines data accuracy requirements for various data applications, which are summarized in Table 1 on page 17 and are referenced as target values for the WIM accuracies calculated here.

The analysis is performed on all four installation datasets, which are installations 1 through 3 and the data from Installation 3 that has been calibrated with the post-processing calibration method. Each dataset had its GVWs, axle group loads, and tandem and tridem axle groups analyzed in three forms:

1. Uncorrected.
2. Temperature corrected based on regression models of the effect of temperature on front axle loads of the 4 selected vehicle configurations. The details of this correction process are described in Appendix G.
3. Autocalibrated with a post-processing autocalibration procedure. The autocalibration procedure was only applied to the Installation 3 dataset with no post-processing calibration applied, as the effects of the post-processing calibration procedure would be eliminated by the autocalibration. The details of this calibration process are described in Appendix G.

The autocalibration and temperature correction procedures have not been combined at any point, as the effects of the temperature correction would be largely eliminated by subsequent application of the autocalibration.

In all tables in this analysis, the labels C and UC will refer to datasets that were initially calibrated and uncalibrated, respectively.

4.4.1 Overall Validity of Axle Loads

This analysis applies the ASTM E1318 standards for the GVW, axle groups, and individual axles. All percent error histograms are presented with data binned by groups of 5%.

GVW

Figure 19, Figure 20, and Figure 21 display histograms of the percent errors of the GVWs of the four portable WIM datasets in their uncorrected, corrected, and autocalibrated forms, respectively.

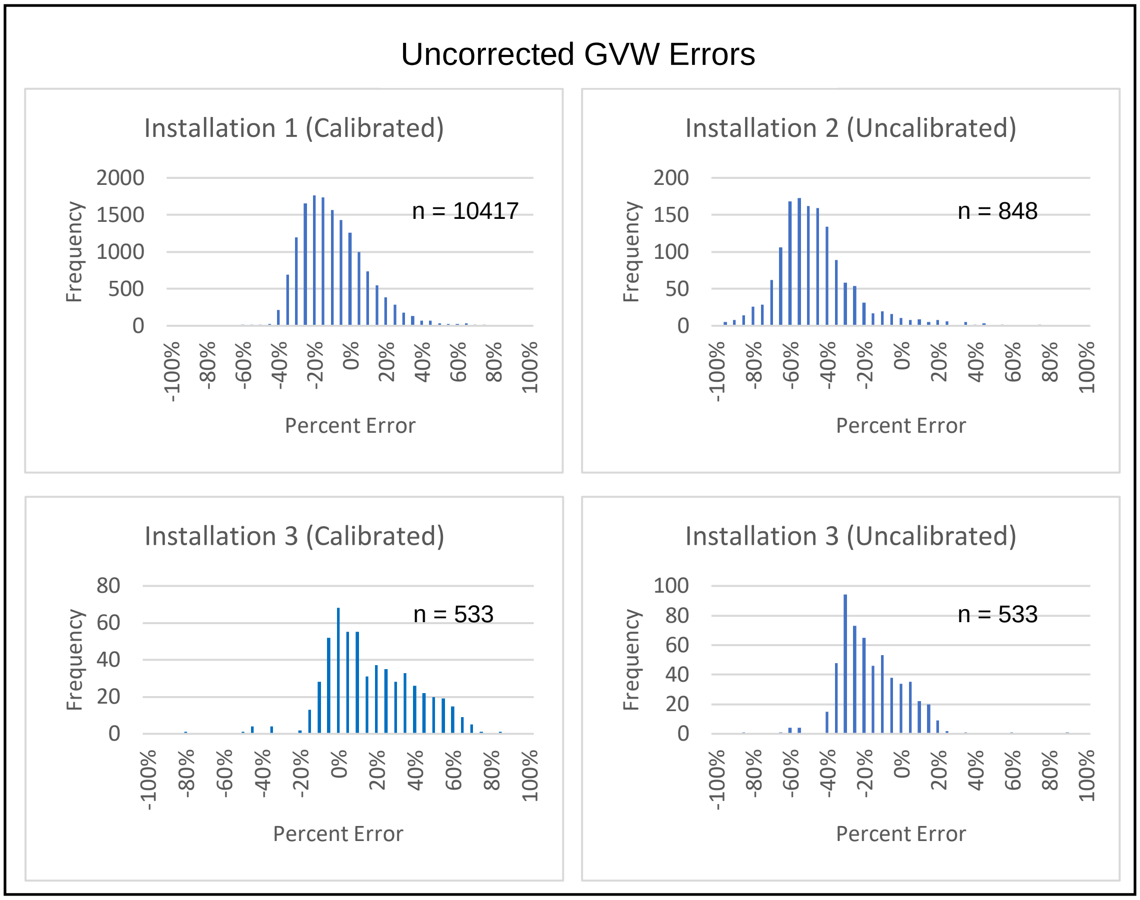


Figure 19 - Histograms of errors of uncorrected GVWs of trucks from all portable WIM datasets

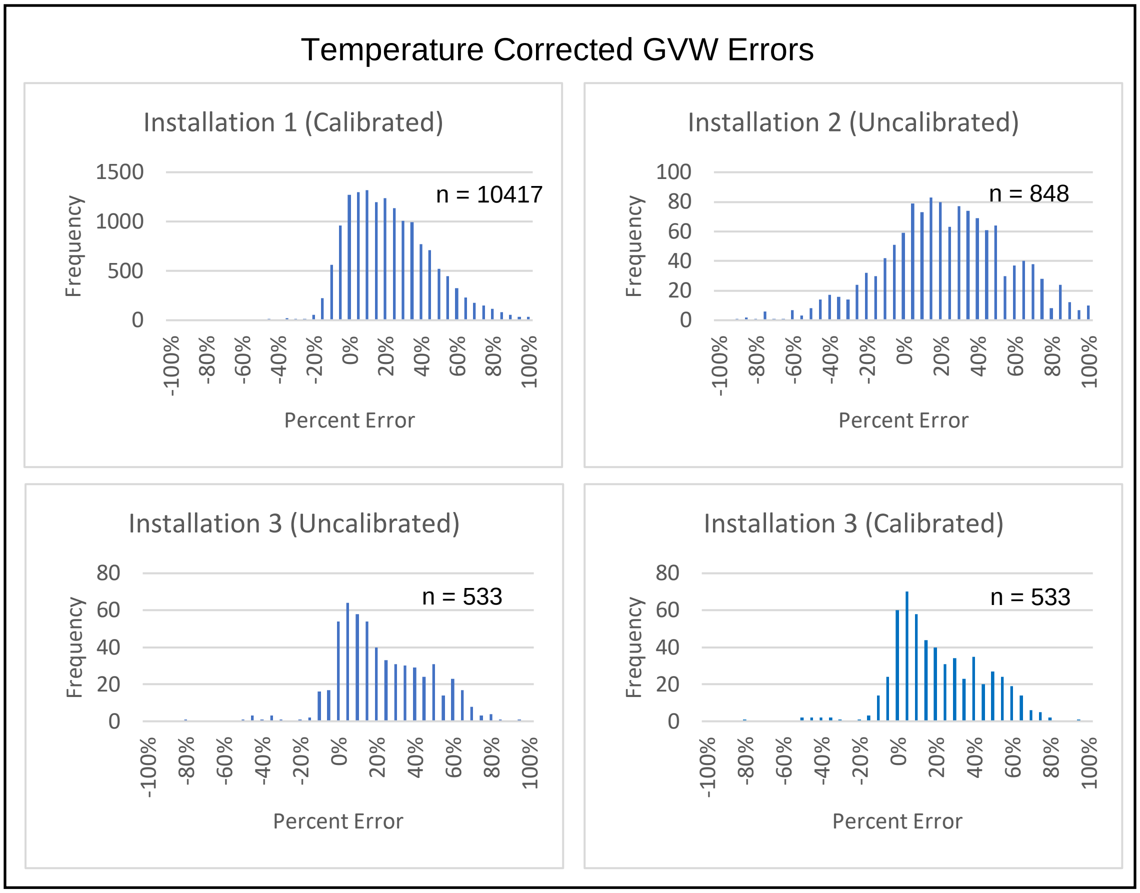


Figure 20 - Histograms of errors of temperature corrected GVWs of trucks from all portable WIM datasets

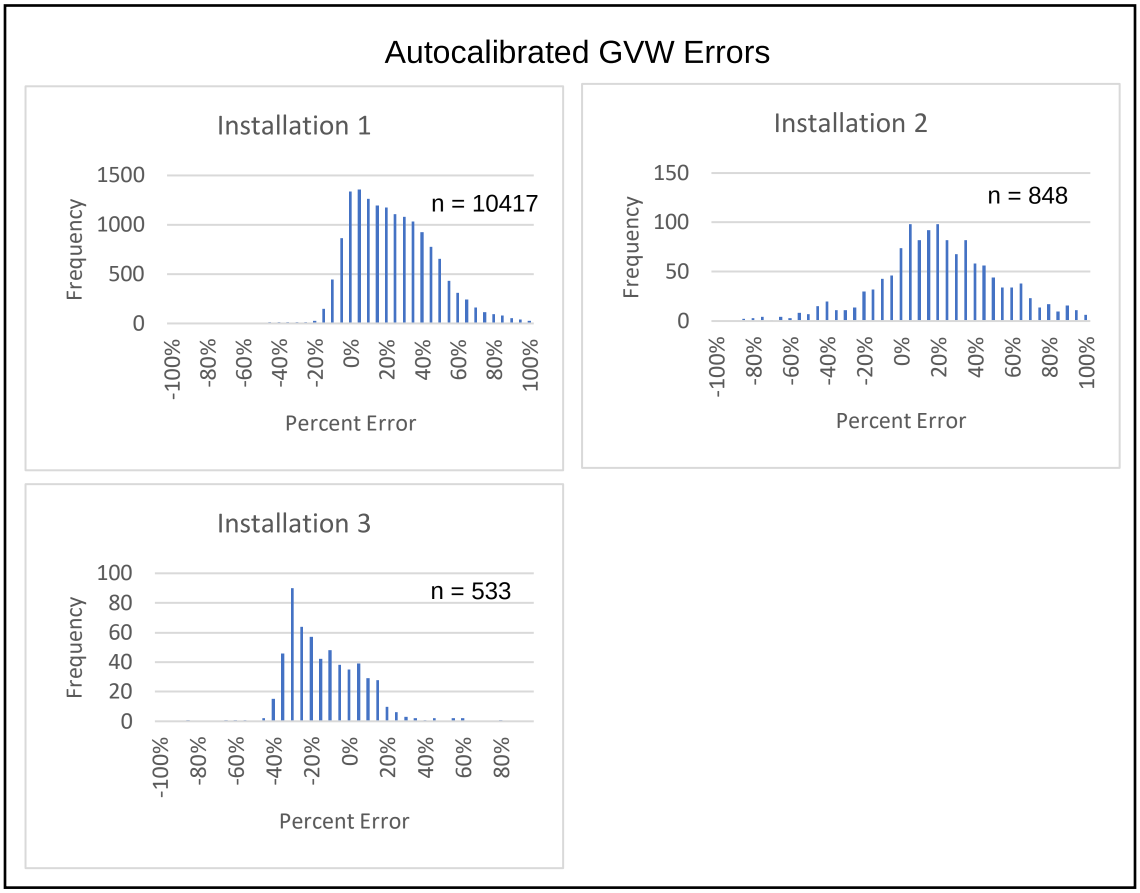


Figure 21 - Histograms of errors of autocalibrated GVWs of trucks from 3 portable WIM datasets

These figures show that all of the GVW percent error histograms have single-peak, positively skewed distributions, though the amount of the skewness varies. Histograms from installations 1, 2, and 3 (uncalibrated) all have peaks between -20% and -50% for uncorrected data; the calibrated measurements from Installation 3 have peaks approximately at 0%. For the temperature corrected and autocalibrated datasets, the peaks are at and above 0% for installations 1 and 2, respectively. Installation 3 datasets have peaks at 0% for temperature corrected datasets and above 0% for the autocalibrated dataset. These characteristics are better observed in the datasets from Installation 1, which have a substantially higher sample size than the other installations and, as a result, a more defined pattern in their distribution.

To better examine the portable WIM datasets in relation to one another, several key statistics were calculated for each GVW distribution. These are mean, median, standard deviation, ASTM type II compliance and ‘95% of results within’. These are described in greater detail in section 3.4.4.Table 17 presents these statistics for each dataset’s GVW measurements for each of the 3 processing methods.

Table 17 - Summary statistics of portable WIM GVW percent errors

Statistic	Installation	Uncorrected Value	Temperature Corrected Value	Autocalibrated Value
Mean	Installation 1 (C)	-6.0%	26.8%	27.7%
	Installation 2 (UC)	-41.6%	41.1%	39.1%
	Installation 3 (UC)	-13.6%	44.5%	-10.2%
	Installation 3 (C)	20.1%	42.0%	-
Median	Installation 1 (C)	-9.1%	22.7%	23.8%
	Installation 2 (UC)	-46.9%	29.4%	25.1%
	Installation 3 (UC)	-17.0%	32.4%	-14.4%
	Installation 3 (C)	15.1%	29.2%	-
Standard Deviation	Installation 1 (C)	21.7%	29.0%	28.6%
	Installation 2 (UC)	38.2%	96.1%	98.7%
	Installation 3 (UC)	18.2%	26.2%	25.3%
	Installation 3 (C)	25.3%	26.1%	-
ASTM Type II Compliance (±15%)	Installation 1 (C)	50.7%	37.0%	35.5%
	Installation 2 (UC)	5.8%	23.8%	26.7%
	Installation 3 (UC)	40.2%	37.2%	40.7%
	Installation 3 (C)	47.8%	40.4%	-
95% of Results Within	Installation 1 (C)	±36%	±73%	±71%
	Installation 2 (UC)	±74%	±139%	±140%
	Installation 3 (UC)	±35%	±67%	±36%
	Installation 3 (C)	±62%	±66%	-

These statistics reveal several key features of the GVW data:

1. For the uncorrected data, the calibrated Installation 1 dataset had the smallest mean percent error (-6.0%) and median percent error (-9.1%). In contrast, the uncalibrated Installation 2 dataset had the largest mean percent error (-41.6%) and median percent error (-46.9%).
2. Installation 2 had a low percentage of GVW measurements in ASTM compliance for all processing methods (5.8% to 26.7%), while all other installations had between 35.5% and 50.9% of GVW measurements in ASTM compliance for all processing methods.
3. Uncalibrated datasets required tolerances of ±35% to ±140% to include 95% of GVW measurements for all processing methods, while initially calibrated datasets required tolerances of ±36% to ±71% to capture 95% of GVW measurements for all processing methods. In particular, Installation 2 required higher tolerances to include 95% of measurements than the other installations.
4. For the ASTM compliance and the ‘95% of results within’ statistics, the temperature correction process, relative to the uncorrected data, improved 1 of 8 statistics, had a negligible effect on 2 of 8 statistics, and made 7 of 8 statistics worse.
5. For the ASTM compliance and the ‘95% of results within’ statistics, the autocalibration process, relative to the uncorrected data, improved 1 of 6

statistics, had a negligible effect on 2 of 6 statistics, and made the data worse for 3 of 6 statistics.

6. Installation 2 had the highest standard deviation for uncorrected data at approximately 39%, while the other 3 datasets had standard deviations between 18% and 26%.
7. Both the temperature correction procedure and the autocalibration procedure increased the standard deviation of the errors for all installations.

Axle Groups

Figure 22, Figure 23, and Figure 24 display histograms of the percent errors of the axle groups of the four portable WIM datasets in their uncorrected, corrected, and autocalibrated forms, respectively. To ensure that axle groups are calculated correctly, only axle groups of the four selected configurations are displayed. Tridem and tandem axle groups are displayed separately. Single (front) axle loads are not displayed as they are included in the analysis of individual axle loads.

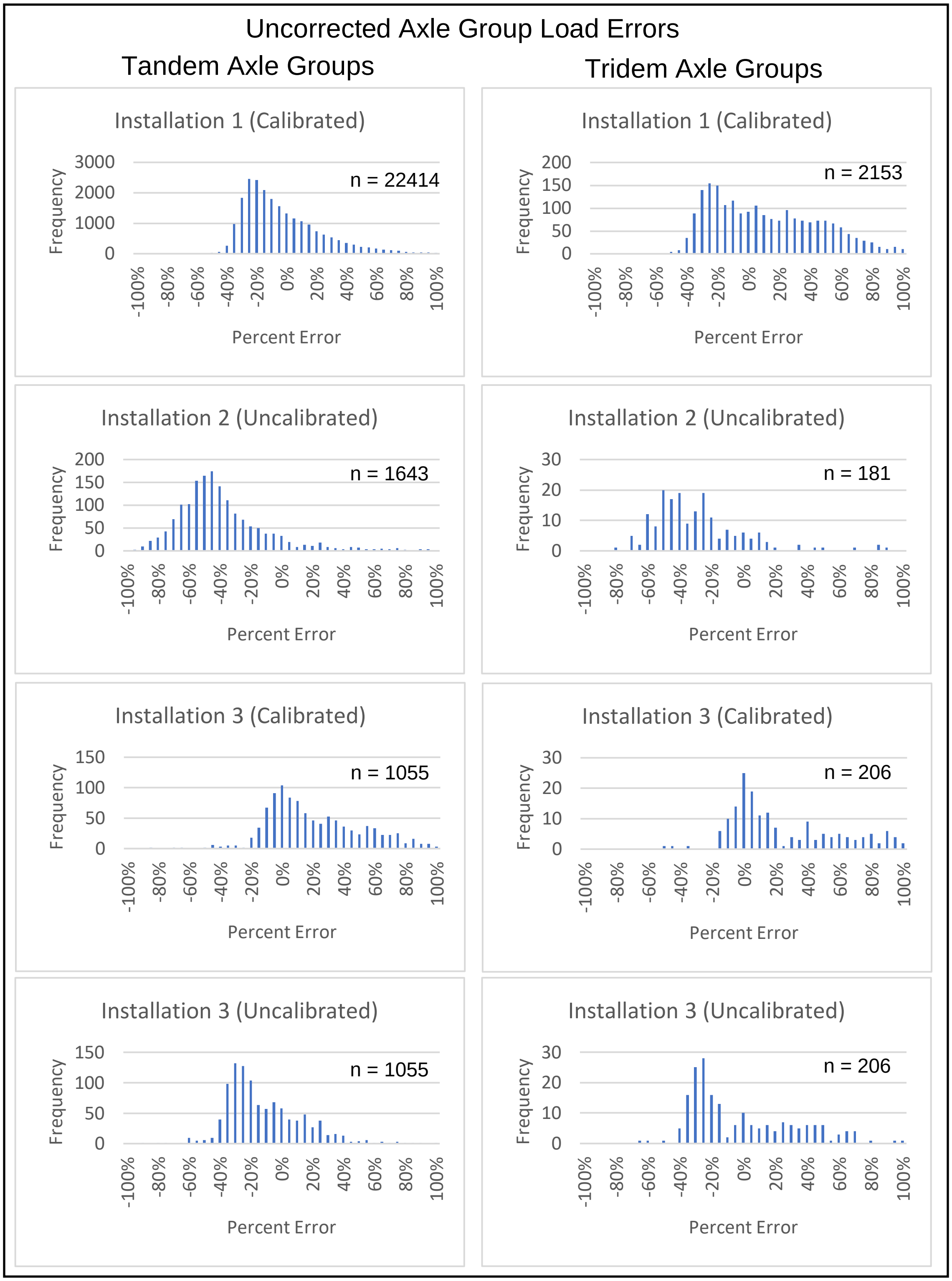


Figure 22 - Histograms of errors of uncorrected axle group loads of trucks from all portable WIM datasets

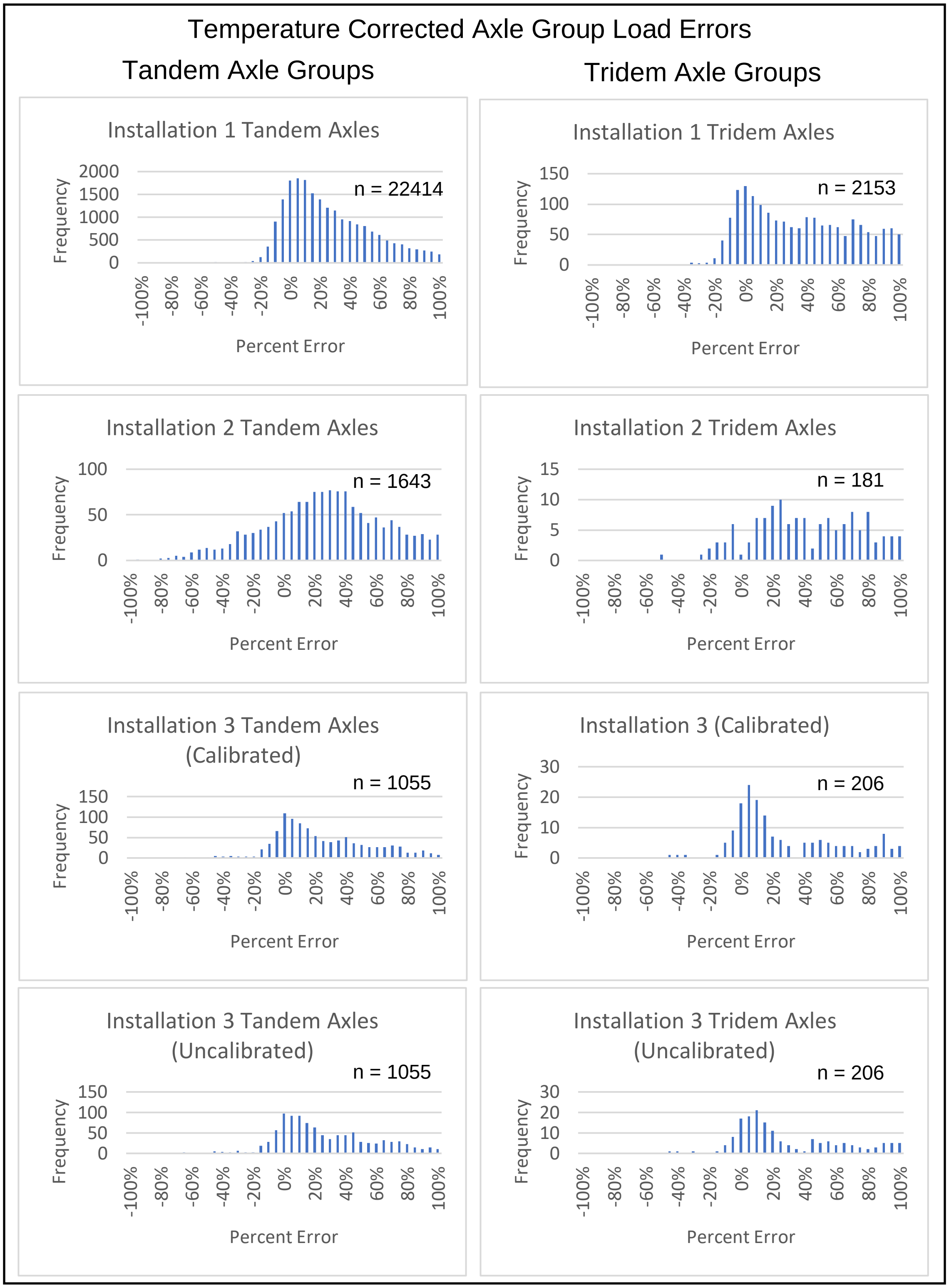


Figure 23 - Histograms of errors of temperature corrected axle group loads of trucks from all portable WIM datasets

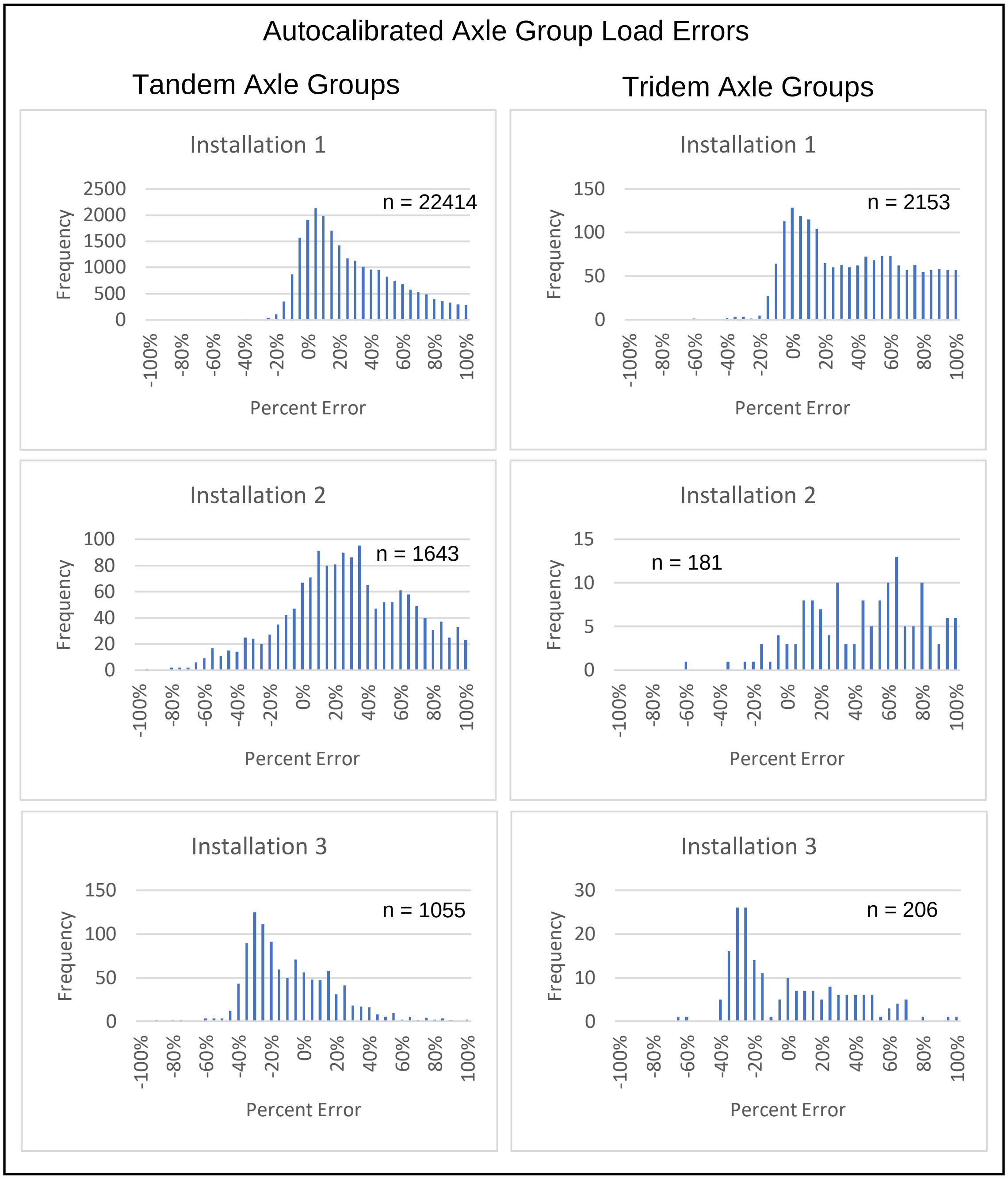


Figure 24 - Histograms of errors of autocalibrated axle group loads of trucks from 3 portable WIM datasets

These figures show that, similar to the GVW percent error histograms, all of the axle group load percent error histograms have single-peak, positively skewed distributions, though the amount of the skewness varies. Moreover, many of the histograms have tails that are long and flat, appearing in some of the tridem axle distributions to be a two-regime distribution, where to the right of the positive skewed curve there is a flat section. This is likely due to the smaller sample sizes for tridem axle groups and the tendency for these distributions to be positively skewed. Distribution peaks are approximately equal for tandem and tridem axles histograms for each dataset. Datasets from installations 1, 2, and 3 (uncalibrated) all have peaks between -20% and -50% for uncorrected data; the uncorrected, calibrated installation 3 dataset had a peak around 0%. For the temperature corrected and autocalibrated datasets, the peaks are at and above 0% for installations 1 and 2, respectively; the temperature corrected Installation 3 datasets have peaks at 0% and the autocalibrated Installation 3 dataset has a peak below 0%. The locations of the peaks of the graphs are all approximately equal to the peaks observed in the corresponding GVW distributions.

Table 18 presents the same summary statistics as Table 17 – median, ASTM compliance, and ‘95% of results within’. ASTM compliance is ±20% when considering axle group loads. These statistics were calculated for each dataset’s selected axle group load measurements for each of the 3 processing methods, with tandem and tridem axle groups calculated separately.

Table 18 - Summary statistics of portable WIM axle group load percent errors

Statistic	Installa-tion	Uncorrected Value		Temperature Corrected Value		Autocalibrated Value	
		Tandem	Tridem	Tandem	Tridem	Tandem	Tridem
Mean	1 (C)	0.7%	17.3%	35.5%	57.2%	35.9%	59.4%
	2 (UC)	-32.8%	-25.4%	60.0%	75.4%	56.4%	76.0%
	3 (UC)	-8.5%	3.6%	32.5%	44.5%	-4.8%	5.6%
	3 (C)	26.4%	43.0%	34.1%	42.0%	-	-
Median	1 (C)	-7.2%	9.3%	24.6%	47.3%	24.3%	50.7%
	2 (UC)	-41.6%	-31.1%	-39.3%	60.4%	35.7%	63.9%
	3 (UC)	-15.3%	-13.0%	31.8%	26.7%	-29.0%	-42.5%
	3 (C)	16.9%	20.0%	28.7%	24.2%	-	-
Standard Deviation	1 (C)	32.9%	42.2%	44.4%	56.4%	44.5%	56.6%
	2 (UC)	39.6%	30.9%	94.5%	71.7%	92.0%	75.0%
	3 (UC)	25.9%	37.2%	37.7%	54.5%	29.3%	39.3%
	3 (C)	35.8%	51.3%	37.2%	53.7%	-	-
ASTM Type II Comp-liance (±20%)	1 (C)	55.2%	38.0%	43.7%	31.6%	44.3%	31.4%
	2 (UC)	15.5%	23.4%	23.0%	17.7%	26.1%	17.1%
	3 (UC)	45.1%	31.1%	43.7%	40.8%	19.0%	13.6%
	3 (C)	50.6%	47.1%	46.2%	43.7%	-	-
95% of Results Within	1 (C)	±57%	±87%	±110%	±150%	±111%	±155%
	2 (UC)	±78%	±66%	±192%	±206%	±218%	±194%
	3 (UC)	±51%	±103%	±105%	±206%	±154%	±202%
	3 (C)	±95%	±180%	±100%	±196%	-	-

These summary statistics reveal several key features of the axle group load data:

1. For the uncorrected data, initially calibrated datasets (Installation 1 and Installation 3 calibrated) had mean percent errors from 0.7% to 43.0% and median percent errors from -7.2% to 20.0%, while uncalibrated datasets (Installation 2 and Installation 3 uncalibrated) had mean percent errors from -32.8% to 3.6% and median percent errors from -41.6% to -13.0%.
2. Installation 2 had a low percentage of axle group load measurements in ASTM compliance for all processing methods (15.5% to 26.1%), while Installation 1 had the highest average ASTM compliance across all processing methods, with compliance between 31.4% and 55.2%.
3. For the ASTM compliance and the ‘95% of results within’ statistics, the temperature correction process, relative to the uncorrected data, improved 2 of 16 statistics, had negligible effects on 4 of 16 statistics, and made 10 of 16 statistics worse.
4. For the ASTM compliance and the ‘95% of results within’ statistics, the autocalibration process, relative to the uncorrected data, improved 1 of 12 statistics and made 11 of 12 statistics worse.
5. Similar to the GVW data, Installation 3 (UC) had the lowest standard deviations, and Installation 2 had the highest standard deviations for all but one axle group-processing method combination.
6. As in the GVW data, both the temperature correction process and the autocalibration process increased the standard deviation for all installations for both tandem and tridem axles.

Individual Axle Loads

Figure 25, Figure 26, and Figure 27 display histograms of the percent errors of the individual axle loads of the two portable WIM datasets collected at WIM station 99 (installations 1 and 2) in their uncorrected, temperature corrected, and autocalibrated forms, respectively. As the comparison data collected at the Headingley static weigh scale was only collected by axle group, individual axle accuracies cannot be calculated and only the data from WIM station 99 is analyzed.

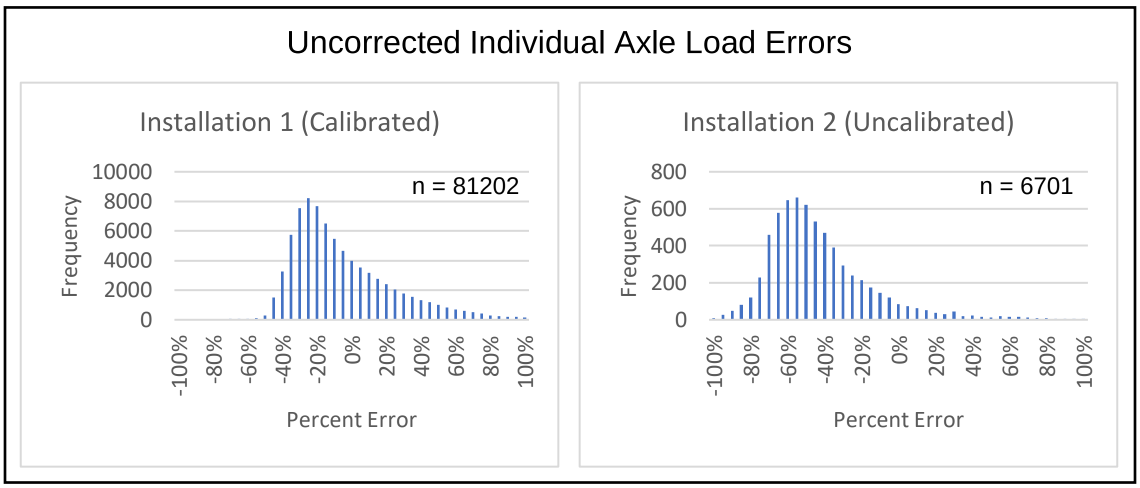


Figure 25 - Histograms of errors of uncorrected axle loads of trucks from 2 portable WIM datasets

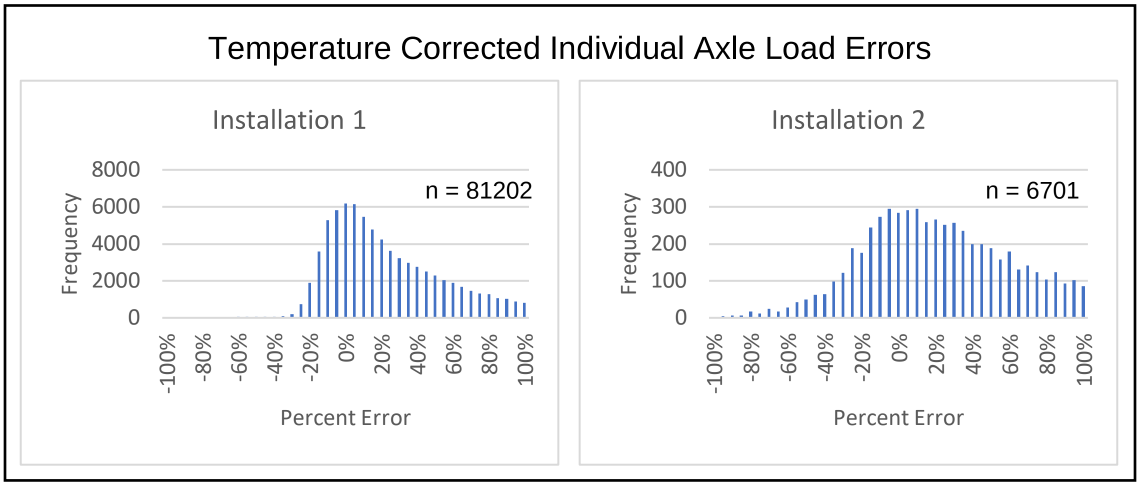


Figure 26 - Histograms of errors of temperature corrected axle loads of trucks from 2 portable WIM datasets

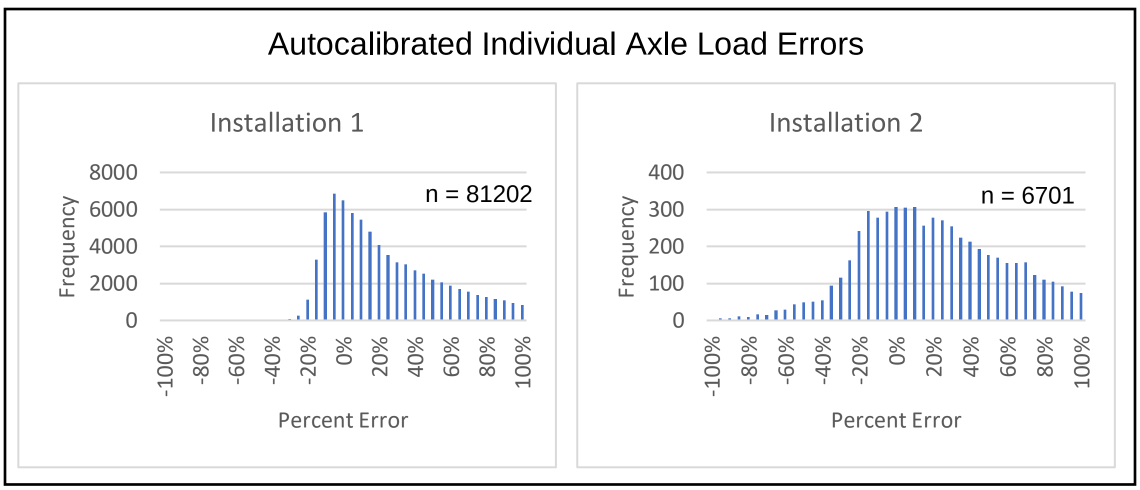


Figure 27 - Histograms of errors of autocalibrated axle loads of trucks from 2 portable WIM datasets

These figures show that, similar to the GVW and axle group load percent error histograms, all of the axle group load percent error histograms have single-peak, positively skewed distributions, with relatively consistent amounts of skewness. Datasets from installations 1 and 2 have peaks approximately at -30% and -60% for the uncorrected datasets. For the temperature corrected and autocalibrated datasets, the peaks are approximately at 0% for both installations. The locations of the peaks of the graphs are all similar to the peaks observed in the corresponding GVW distributions.

Table 19 presents the same summary statistics as Table 17 – median, ASTM compliance, and '95% of results within'. ASTM compliance is ±30% for single axles. These statistics were calculated for the installation 1 and 2 datasets' selected axle load measurements for each of the three processing methods.

Table 19 - Summary statistics of portable WIM axle load percent errors

Statistic	Installation	Uncorrected Value	Temperature Corrected Value	Autocalibrated Value
Mean	Installation 1 (C)	-0.7%	33.8%	34.4%
	Installation 2 (UC)	-35.6%	55.0%	52.7%
Median	Installation 1 (C)	-10.4%	20.3%	20.4%
	Installation 2 (UC)	-46.0%	29.4%	26.6%
Standard Deviation	Installation 1 (C)	54.7%	71.8%	70.9%
	Installation 2 (UC)	106.9%	280.3%	278.2%
ASTM Type II Compliance (±30%)	Installation 1 (C)	71.4%	59.0%	58.6%
	Installation 2 (UC)	22.8%	43.9%	46.4%
95% of Results Within	Installation 1 (C)	±64%	±118%	±119%
	Installation 2 (UC)	±78%	±194%	±193%

These summary statistics reveal several key features of the individual axle load data:

1. The Installation 1 datasets performed better in all three processing methods for all five statistics than the corresponding Installation 2 datasets. The Installation 1 datasets had means and medians closer to 0%, lower standard deviations, higher percentage of results within ±30% error, and lower tolerance range for 95% of results.
2. Both temperature correction and autocalibration had the effect of increasing the portable WIM loads, and resulted in median percent errors of 20% to 30% compared to -10% and -46% in uncorrected data.
3. Both temperature correction and autocalibration performance were erratic, but they typically resulted in similar values; both processing methods improved Installation 2's ASTM compliance, but made Installation 1's compliance worse, and made '95% of results within' worse for both installations.
4. Standard deviations were higher than in the GVW and axle group load data.
5. As in the previous data, both the temperature correction process and the autocalibration process had the effect of increasing the standard deviations.

4.4.2 Accuracy of Axle Loads During Calibration Drift

In addition to analyzing the accuracy of the portable WIM measurements during an entire installation, an analysis examining accuracy trends as the calibration of the system drifts was conducted to inform recommendations about duration of installation and calibration practices. To examine these trends with reference to the ASTM standards, the percentage of GVW, tandem axle loads measurements that fall within the ASTM type II standards for each day of each installation were calculated. These calculations also sought to clarify if the increased rate of erroneous records during Installation 2 had any effect on the validity of the load data. Figure 28 displays graphs of the daily measurements of ASTM type II compliance for GVW, tandem axle group loads, and individual axle loads during installations 1 and 2. For each load category, both installations are displayed on a single graph chronologically.

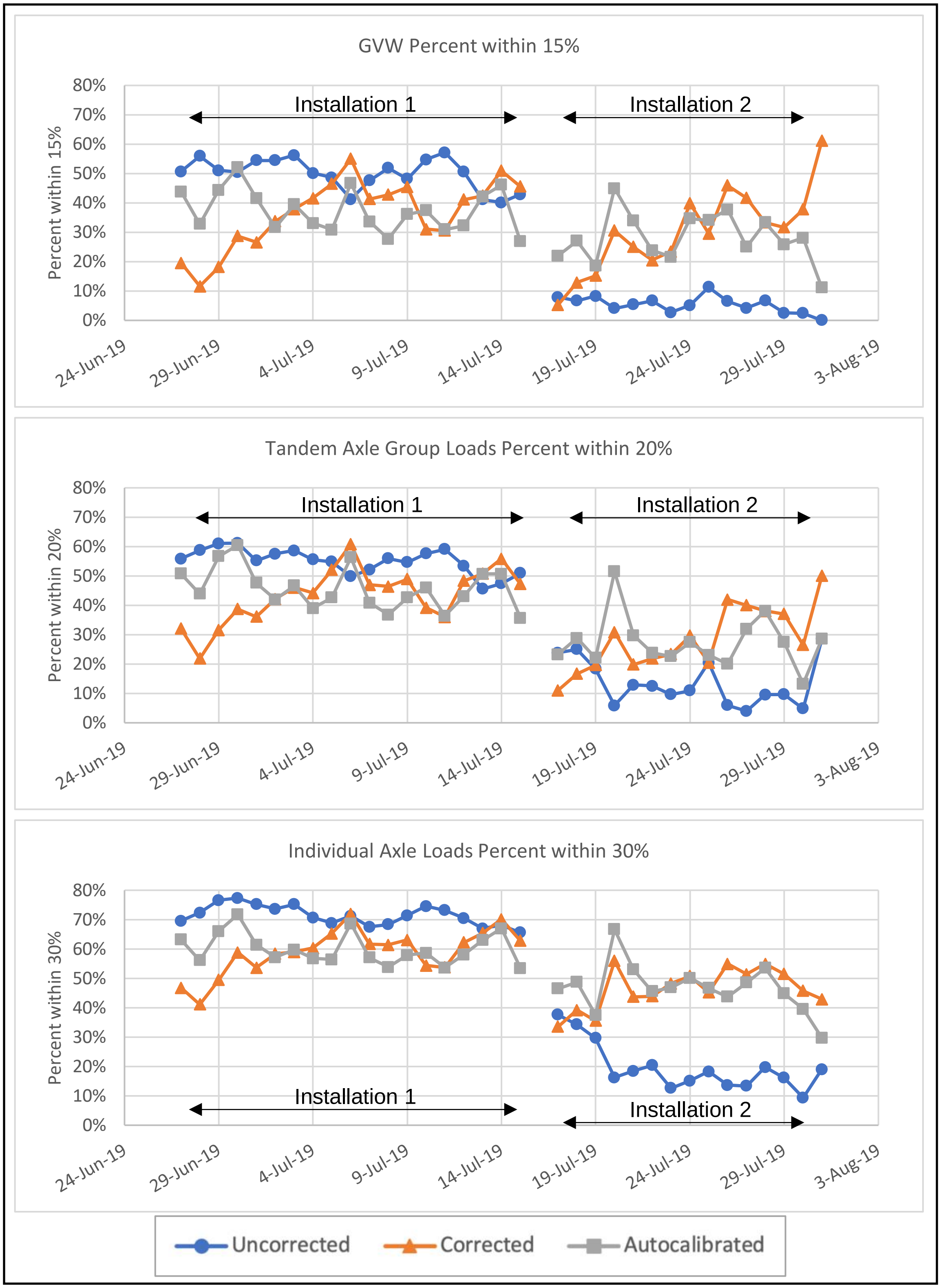


Figure 28 - Daily measurements of ASTM type II compliance during portable WIM installations 1 and 2

These graphs reveal that:

1. In addition to not meeting the ASTM type II accuracy standards for the installations as a whole, the first two portable WIM installations did not meet the ASTM type II accuracy standards on any day of the installation for any processing method.
2. There was a downward trend in the compliance during both of the installations for the uncorrected data; the temperature corrected data trended upwards, and the autocalibrated data, as expected, did not trend significantly throughout either installation.
3. For uncorrected GVW and tandem axle group loads, compliance on the first day of Installation 2 was approximately representative of Installation 2 as a whole. As this was the day of the installation that had fewer erroneous records than the rest of the installation, this indicated that the error rate had no consequential effect on the load accuracy.
4. The second portable WIM installation yielded data that was less accurate than the first. The decrease in accuracy was seen not just through the steady decrease in accuracy over time, but immediately at the beginning of the second installation.
5. Both the temperature correction process and the autocalibration process decreased the accuracy of the portable WIM load data during the first installation but improved it during the second, making the difference in

accuracy between Installation 1 and Installation 2 less than in the uncorrected data.

To provide a more detailed look at the accuracy of the portable WIM system over the course of installations 1 and 2, key error percentiles were calculated and graphed. Appendix F displays and provides commentary on these graphs.

4.5 AGGREGATED DATA VALIDITY RESULTS (ANALYSIS 5)

This section presents the results of the analysis of the portable WIM data's validity on an aggregated basis. The analysis of aggregated data looks at axle load spectra for front, tandem, and tridem axles as well as GVW spectra. These spectra are considered separately for each of the four selected vehicle configurations during each of the four portable WIM datasets (derived from the three portable WIM installations) as noted in the analysis of per-vehicle accuracy (see section 4.4). All three of the data processing methods used in the per-vehicle accuracy analysis are considered: uncorrected, temperature corrected, and autocalibrated. As in the analysis of per-vehicle accuracy, the autocalibration and temperature correction procedures were not be combined at any point, as the effects of the temperature correction would be largely eliminated by subsequent application of the autocalibration.

Each of the load spectra noted above was subjected to a Gaussian Mixture Model (GMM), which separated the spectra into three normal distributions, each representing one of three possible loading scenarios: unloaded, partially loaded, and fully loaded. Due to the low variability regardless of loading scenario, front axles were modeled with a single normal distribution, and because the models perform best when using large sample sizes, only the first installation's data was considered in the analysis. Additionally, only vehicles that were paired and considered in the analysis of per-vehicle accuracy were considered in order to best demonstrate how the aggregated accuracy method is able to perform compared to the per-vehicle accuracy method.

This analysis first considered the distribution means, to gain a general picture of the models' accuracy, then considered the models' confidence intervals, and finally used statistical t-tests to determine if the models were statistically similar to the comparison data.

4.5.1 Comparison of Distribution Means

Figure 29 presents graphs of the percent error of the means of the normal distributions of the portable WIM GVW, front axle, tandem axle, and tridem axle measurements, respectively, as compared to data from WIM station 99 during the first installation. These comparisons are made for uncorrected, temperature corrected, and autocalibrated data.

Figure 29 - Percent error of means of portable WIM loading scenarios during portable WIM Installation 1

The graphs of the percent error of the means reveal several notable observations about the aggregated portable WIM data:

1. Front axle percent errors were lower for the temperature corrected and autocalibrated data (~5%) than for the uncorrected data (~30%)
2. For GVWs, and tandem/tridem axle groups, temperature corrected and autocalibrated data generally had higher errors for unloaded and partially loaded distributions, and lower errors for fully loaded distributions.
3. Unloaded vehicles more often had positive percent errors (portable WIM values higher), and partially loaded and fully loaded more often had negative percent errors (portable WIM values lower).
4. No GVW or axle group loading scenario distributions had error of their mean lower than the lowest mean per-vehicle error for that axle group or GVW as reported in Section 4.4.

4.5.2 Comparison of Distribution Confidence Intervals

Figure 30 presents graphs of the percent of the portable WIM normal distribution 95% confidence intervals that overlaps with the WIM station 99 normal distribution 95% confidence intervals for each loading scenario for GVWs, front axles, tandem axles, and tridem axles, respectively. These comparisons are made for uncorrected, temperature corrected, and autocalibrated data. The percent overlap is calculated such that a portable WIM confidence interval fully contained by the comparison confidence interval has a 100% overlap, while one wholly outside the comparison confidence interval will have a 0% overlap.

Figure 30 - Percent of 95% CI overlap portable WIM loading scenarios during portable WIM Installation 1

The graphs of the percent overlap of the 95% confidence intervals reveal several notable observations about the aggregated portable WIM data:

1. For GVW, tandem axle groups, and tridem axle groups, uncorrected data had higher percent overlap for unloaded distributions than the other processing methods, but for GVW and tridem axles, temperature corrected and autocalibrated data had higher percent overlap for fully loaded distributions than the uncorrected data. Partially loaded distributions had no processing method that revealed consistently higher percent overlap.
2. For front axles, temperature corrected and autocalibrated data had higher percent overlap than uncorrected data; autocalibrated data had slightly higher percent overlap than temperature corrected data.
3. There were no systematic trends differentiating the percent overlap of the different vehicle configurations.

4.5.3 t-Test Results

Table 20 presents the p-values resulting from statistical t-tests comparing the loading scenario normal distributions of GVWs for uncorrected, temperature corrected, and autocalibrated portable WIM records, respectively, to records from WIM station 99 during the first installation. P-values greater than 0.05, which indicate that there is insufficient evidence to state that the portable and comparison distributions are statistically different, are highlighted.

Table 20 - Results of t-tests comparing portable WIM GVW loading scenarios to comparison data

Processing Method	Configuration	P-value		
		Unloaded	Partially Loaded	Fully Loaded
Uncorrected	3-S2	0.00	0.00	0.00
	3-S2-4	0.00	0.00	0.00
	3-S3	0.00	0.00	0.00
	3-S3-S2	0.00	0.00	0.00
Temperature Corrected	3-S2	0.00	0.00	0.49
	3-S2-4	0.00	0.00	0.59
	3-S3	0.00	0.00	0.69
	3-S3-S2	0.00	0.00	0.94
Autocalibrated	3-S2	0.00	0.00	0.45
	3-S2-4	0.00	0.00	0.00
	3-S3	0.00	0.00	0.73
	3-S3-S2	0.00	0.00	0.93

These results reveal that:

1. No distributions were statistically similar for uncorrected data.
2. For temperature corrected and autocalibrated data, the fully loaded distributions were more likely to be statistically similar than the unloaded or partially loaded distributions.
3. Both the temperature corrected and autocalibrated data had a majority of t-test results indicate statistically similar distributions.

Table 21 presents the results of statistical t-tests comparing the loading scenario normal distributions of front axles for uncorrected, temperature corrected, and autocalibrated portable WIM records, respectively, to records from WIM station 99 during the first installation.

Table 21 - Results of t-tests comparing portable WIM front axle load spectra to comparison data

Processing Method	Configuration	P-value
Uncorrected	3-S2	0.00
	3-S2-4	0.00
	3-S3	0.00
	3-S3-S2	0.00
Temperature corrected	3-S2	0.00
	3-S2-4	0.00
	3-S3	0.00
	3-S3-S2	0.00
Autocalibrated	3-S2	0.00
	3-S2-4	0.00
	3-S3	0.00
	3-S3-S2	0.00

These results reveal that all portable WIM front axle load distributions were statistically different than those generated from WIM station 99, regardless of the processing method or configuration.

Table 22 presents the results of statistical t-tests comparing the loading scenario normal distributions of tandem and tridem axles for uncorrected, temperature corrected, and autocalibrated portable WIM records, respectively, to records from WIM station 99 during the first installation.

Table 22 - Results of t-tests comparing portable WIM axle group loading scenarios to comparison data

Processing Method	Config-uration	Unloaded		Partially Loaded		Fully Loaded	
		Tandem	Tridem	Tandem	Tridem	Tandem	Tridem
Uncorrected	3-S2	0.00	-	0.90	-	0.47	-
	3-S2-4	0.00	-	0.00	-	0.11	-
	3-S3	0.00	0.00	0.00	0.00	0.77	0.06
	3-S3-S2	0.00	0.00	0.00	0.00	0.00	0.00
Temperature Corrected	3-S2	0.00	-	0.00	-	0.30	-
	3-S2-4	0.00	-	0.00	-	0.09	-
	3-S3	0.00	0.00	0.00	0.00	0.00	0.34
	3-S3-S2	0.00	0.00	0.00	0.00	0.88	0.86
Autocalibrated	3-S2	0.00	-	0.00	-	0.34	-
	3-S2-4	0.00	-	0.00	-	0.10	-
	3-S3	0.00	0.00	0.00	0.00	0.96	0.31
	3-S3-S2	0.00	0.00	0.00	0.00	0.97	0.97

These results do not contradict any of the findings from the analysis of the difference in the means or the overlap of the confidence intervals, and additionally reveal the following:

1. For all processing methods, as in the GVW results, the fully loaded distributions were more likely to be statistically the same than the unloaded or partially loaded distributions. Most test results of fully loaded distributions were found to be statistically similar.
2. All three processing methods performed similarly, with uncorrected, temperature-corrected, and autocalibrated data returning 5, 5, and 6

statistically similar test results, respectively. Only uncorrected data returned statistically similar results for partially loaded vehicles.

4.5.4 Loading Separated Temperature Corrected Results

The analysis of the GMM-aggregated data reveals that when applying either the temperature correction process or the autocalibration process, fully loaded vehicle GVW and axle group distributions often have statistically similar results, but have much higher error than front axle load distributions, which have means of under 5% error. This suggests that per-vehicle accuracy is different under different loading scenarios, and that correction made on the basis of front axle loads may not result in appropriate corrections to other axles. To investigate these findings, an updated temperature correction procedure was been developed that applies separate temperature correction factors generated from GVWs, rather than front axle loads, to unloaded and fully loaded vehicles. The details of this correction procedure, along with the other correction procedures, are found in Appendix G.

Figure 31 displays histograms of the per-vehicle errors from Installation 1 with both the initial and loading separated temperature correction procedures applied. As the loading separated temperature correction process requires separation of unloaded and fully loaded vehicles, only the selected configurations were processed and are included in the results in this section, including Figure 31.

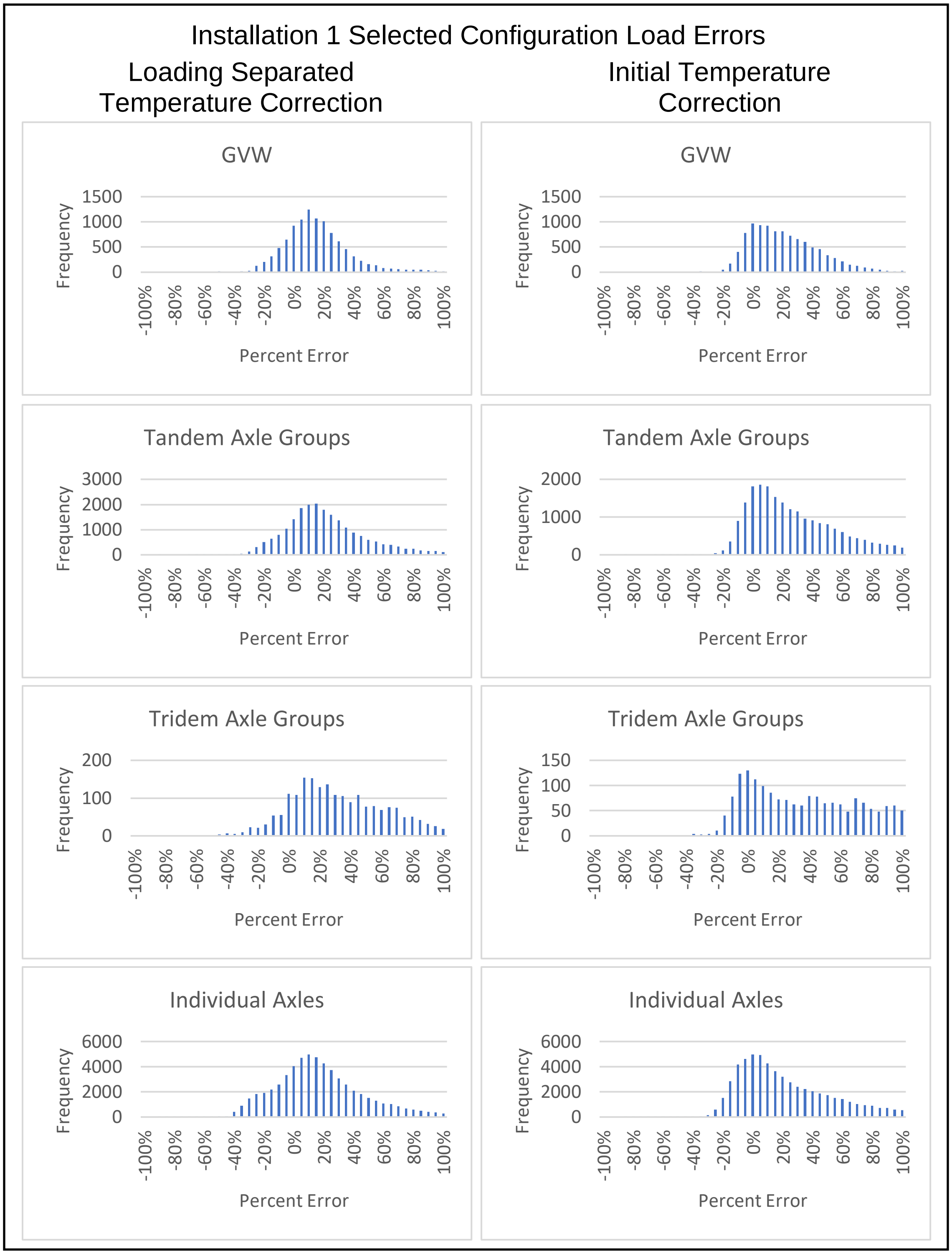


Figure 31 - Histograms of errors of temperature corrected GVW, axle groups, and axles of selected configurations during Installation 1

Figure 31 shows that the selected axle percent error histograms for the initial temperature corrected data have the same single-peak, positively skewed distributions that their corresponding histograms for all trucks did, but that the load separated temperature corrected data has virtually no skewness. The initial temperature corrected data have peaks at approximately 0% error, and the load separated temperature corrected data have peaks at approximately 10% error. Table 23 further examines these differences, and additionally compares these datasets to uncorrected data for selected configurations, through the display of key statistics.

Table 23 - Key statistics of selected configuration axle load errors for temperature correction procedures during Installation 1

Statistic	Axle Group	Uncorrected Value	Load Separated Temperature Corrected Value	Initial Temperature Corrected Value
Mean	GVWs	-7.0%	19.0%	25.3%
	Tandem Axles	0.7%	28.4%	35.5%
	Tridem Axles	17.3%	42.0%	57.2%
	Individual Axles	-2.3%	23.9%	31.3%
Median	GVWs	-10.5%	15.5%	20.6%
	Tandem Axles	-7.2%	21.0%	24.6%
	Tridem Axles	9.3%	33.4%	47.3%
	Individual Axles	11.6%	17.2%	18.6%
Standard Deviation	GVWs	20.8%	26.6%	27.7%
	Tandem Axles	32.9%	41.6%	44.4%
	Tridem Axles	42.2%	44.7%	56.4%
	Individual Axles	36.9%	46.2%	49.5%
ASTM Type II Compliance	GVWs (±15%)	48.3%	44.7%	40.3%
	Tandem Axles (±20%)	55.2%	45.7%	43.7%
	Tridem Axles (±20%)	38.0%	32.0%	31.6%
	Individual Axles (±30%)	72.3%	64.5%	60.9%
95% of Results Within	GVWs	±36%	±61%	±71%
	Tandem Axles	±57%	±92%	±110%
	Tridem Axles	±87%	±116%	±150%
	Individual Axles	±61%	±91%	±114%

These statistics reveal several key observations about the temperature correction procedures:

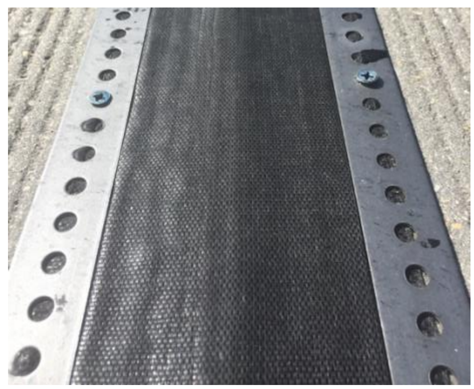
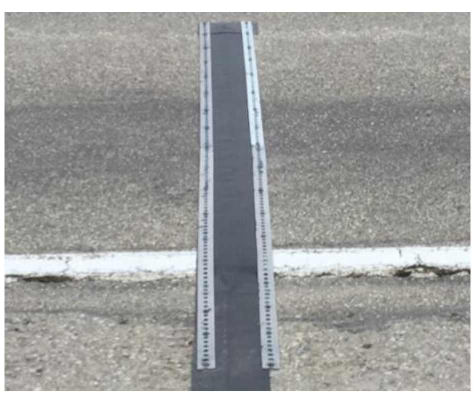
1. Both temperature correction procedures shifted the dataset means and medians positively; the uncorrected data had negative or near-zero means and medians, while the temperature corrected data had positive means and medians.
2. Both temperature correction procedures had poorer results than the uncorrected data for standard deviation, ASTM type II compliance, and ‘95% of results within’, but by small margins (e.g., ASTM compliance is 3.6% to 11.4% lower).
3. The load separated temperature corrected data had better results than the initial temperature corrected data for standard deviation, ASTM type II compliance, and ‘95% of results within’, but by small to insignificant margins (e.g., ASTM compliance is 0.4%-4.4% higher).

4.6 DISCUSSION

4.6.1 Qualitative Observations

During all three installations, the portable WIM unit was observed periodically to inspect physical wear on the sensors. Table 24 shows pictures from near the beginning and the end of each installation, noting the total traffic volume that each installation experienced, including during calibration. The ‘Late Photograph’ column does not show photos taken after the volume stated has passed, only one taken near the end of the installation.

Table 24 - Photographs of portable WIM sensors at beginning and end of each installation

Installation	Early Photograph	Late Photograph
1	 June 27, 2019 (Installation Day 1)	 July 15, 2019 (Installation Day 19) Total Traffic Volume: 70030
2	 July 17, 2019 (Installation Day 1)	 July 27, 2019 (Installation Day 11) Total Traffic Volume: 59457
3	 October 18, 2019 (Installation Day 3)	 October 23, 2019 (Installation Day 8) Total Traffic Volume: 42504

These photographs show that over the course of all three installations, very little visible physical damage occurred to the sensors; the effects were primarily seen in a small amount of bunching of the pocket tape and protective tape layers at the leading edge of the sensors (left side in all pictures), and a small amount of pulling up in some of the installation screws. This confirms that the installation method used was effective in securing the sensors in the long term, though the comparative accuracy of a different installation method is unknown.

4.6.2 Implications for a Portable WIM Data Collection Program

The objectives of this research were to evaluate the feasibility of using a portable WIM data collection program to collect useful axle load data and to inform how to most effectively conduct such a program. To do this, recommendations have been developed surrounding installation practices, data processing procedures, and how the data can be applied most effectively.

The analysis of erroneous vehicle records (section 4.1) revealed that for the one installation (Installation 2) that had an increased rate of erroneous vehicle records, the most common error codes were those for ‘vehicle too slow’ and ‘only one axle detected’. These indicate that one of the two sensors may have been intermittently malfunctioning during the collection of these records, resulting in vehicles that missed recording one or more axles. This may indicate equipment damage, but for the fact that the subsequent Installation 3 recorded a low rate of error codes, as in Installation 1. This suggests that the intermittent malfunction was possibly the result of improper installation (e.g., improperly connected wires). Regardless of the cause, the observation serves to highlight the importance of taking proper care during installation. This is especially important when noting that the analysis also found that erroneous vehicle records were most likely to affect vehicles with three or more axles, which are typically the vehicles of most interest to WIM data analysts.

The error codes did not indicate any difference between a full and half lane installation, so a recommendation for sensor lane coverage cannot be made on that basis. However, the analysis of the daily error percentiles and ASTM compliance (section 4.4.2) provided some insight into the differences based on accuracy. When comparing Installation 1 and Installation 2, which were identical except for the sensor lane coverage and conduct of an initial calibration for Installation 1, the daily accuracies of the half-lane sensor coverage Installation 2 (0 to 7% ASTM compliance for uncorrected GVWs) were consistently lower for uncorrected data than those of the full lane sensor coverage Installation 1 (40% to 58% ASTM compliance for uncorrected GVWs). This could be attributable either to the re-installation without re-calibration, which may have changed the way the axle loads were felt by the sensors in small, but consequential, ways, or the sensor lane coverage. The sensor lane coverage might have had decreased accuracy because half lane coverage does not account for any left to right imbalances in the measured vehicles, and the cross-slopes of the highways would likely lead to systematic imbalances in the vehicles.

The standard deviations of error calculated for each installation showed no clear bias as to whether full or half lane sensor coverage provided more precise results. Installation 2 (half lane) consistently had the highest standard deviation, while Installation 1 and Installation 3 uncalibrated (full and half lane, respectively), had approximately equal standard deviations. Based on these findings, full lane sensor coverage is recommended for future portable WIM installations that use 12-foot sensors; further research is needed to determine if using 6-foot sensors would provide increased accuracy or precision.

The analysis of the calibration drift revealed that the calibration drift relationship during Installation 1 remained consistent for 12 days. Additionally, analysis of the daily ASTM compliance during Installations 1 and 2 showed a pattern of decreasing compliance over the duration of both installations. Therefore, regardless of whether an initial calibration is performed, shorter installations will generally yield higher accuracy data. However, in order to capture any weekly loading patterns and avoid bias to higher- or lower-load days, installations should have durations covering intervals of 7 days. A 7-day installation will allow a full week of data to be collected while staying within the period of consistent calibration drift and is recommended for these reasons.

The analysis of the per-vehicle data validity revealed that the portable WIM system was unable to achieve the accuracies required for ASTM type II WIM data for any installation or processing method. This suggests that portable WIM data is not accurate enough to be applied directly to design applications.

However, the differences between installations and processing methods revealed how a reasonable level of data quality could be achieved.

The per-vehicle data validity analysis revealed that an initial calibration generally improves data validity. When considering uncorrected data, Installation 2, which had no calibration, performed the worst in almost all statistics for GVW, axle group loads, and individual axle loads, while Installation 1, which was initially calibrated, performed the best. The difference in accuracy was further illustrated in the daily measure of ASTM compliance, which showed a gradual decline in accuracy during both installations, with a sharp drop in accuracy from the end of Installation 1 to the beginning of Installation 2. Though by smaller margins, the calibrated dataset from Installation 3 also had higher accuracy than the uncalibrated dataset from Installation 3, which confirms that an initial calibration improves the accuracy of the load data. It is, however, notable that for the two datasets from Installation 3, the calibrated dataset had higher standard deviations, indicating a loss of precision that comes with the increase in accuracy. This loss of precision was also reflected in greater values for ‘95% of results within’ for the calibrated Installation 3 dataset. An additional important observation is that both temperature correction and autocalibration reduced the improvement in accuracy gained from the initial calibration to the point of eliminating it in many datasets; however, percent error histograms showed that initially calibrated datasets still often had peaks closer to 0% than uncalibrated datasets, so calibration for these datasets can still be said to have benefits. Due to these findings, an initial calibration is recommended for all portable WIM installations to ensure load data of maximum accuracy.

The analysis of the benefits of post-processing the data revealed that the two processes examined in the per-vehicle analysis, temperature correction and autocalibration, did not increase the validity of the data as measured by the ASTM compliance and ‘95% of results within’, except in the case of Installation 2, which had low validity both before and after processing. In every dataset tested, both processing methods increased the standard deviation. However, observation of the percent error histograms revealed that the processed data, particularly the temperature corrected data, typically had peaks closer to 0% error. Therefore, it can be seen that the post processing was able to correct a portion of the data to within a small margin of error, but pushed an equally large portion of the data further out of compliance.

This effect was further explored in the analysis of the loading separated temperature correction process, which, though it was unable to achieve ASTM type II accuracy standards for the selected loads on which it was performed, did improve on the initial temperature correction process in all statistics, and showed that the positive skewness seen in all other error histograms can be corrected. While all the processing methods tested in this research failed to improve upon the uncorrected data set’s accuracy and precision, the load-separated temperature correction process showed that more sophisticated data correction can lead to improved accuracy. This suggests that further research may be able to develop correction processes that can improve significantly on the validity of the uncorrected data.

Overall, Installation 1, which was the only initially calibrated, full-lane installation performed, had ASTM type II compliance of between 38.0% and 71.4% for uncorrected data for the GVW, axle groups, and individual axles. As these compliance rates indicate that the portable WIM is unable to achieve ASTM type II accuracy under any conditions, they beg the question of what standard they are able to achieve. The calculation of the ‘95% of results within’ statistic indicates that the Installation 1 uncorrected data falls within error tolerances 2 to 3 times the magnitude of the ASTM type II standard for GVW (±30 to 45%), tandem axles (±40 to 60%), and individual axles (±60 to 90%), though tridem axles fall within an error tolerance 4.5 times the ASTM type II standard (90%). As tolerances 2-3 times the ASTM type II standard represent a significant lower accuracy than the standard demands, research into the accuracy standard required for various indirect applications of load data would further indicate how useful the per-vehicle data truly is.

The analysis of aggregated data validity revealed that during Installation 1, the percent errors of GVW distribution means for uncorrected data were between -31% and 27%; these values were generally larger than the difference in the per-vehicle mean GVW error, indicating no overall accuracy benefit due to the aggregation.

The aggregated data analysis was not as conclusive as the per-vehicle analysis with regard to which data processing method was the best option. While the distribution comparisons for tandem and tridem axles were erratic even within a single processing method, those for GVW and front axles showed greater consistency and revealed some key trends in the data. The analysis of the difference in GMM distribution means and 95% confidence interval overlaps showed that for the three loading scenarios (empty, partially loaded, fully loaded), uncorrected data had the lowest mean difference and highest confidence interval overlap across the three scenarios. Temperature corrected and autocalibrated data often had very low error of distribution means and high confidence interval overlaps for fully loaded vehicles, performing better than the uncorrected data. This resulted in 9 of 10 t-tests of temperature corrected, fully loaded vehicle distributions and 9 of 10 t-tests for autocalibrated, fully loaded distributions noting statistically similar distributions. Additionally, the percent difference of distribution means was lower for front axles of temperature corrected and autocalibrated data than uncorrected data. This can be attributed to both procedures' use of only front axle load data to calibrate the system. As the temperature correction has a stronger theoretical basis, temperature correction is recommended over autocalibration for examining fully loaded vehicle distributions, though development and application of more sophisticated correction procedures, such as was done with the loading separated temperature correction procedure, may prove beneficial in the future.

The specific areas of increased accuracy in the aggregated data could support specific indirect applications of the axle load data. For example, the high accuracy noted for processed, fully loaded vehicle distributions could allow for a selection of data to be highly useful for freight planning purposes. Alternatively, as the GVW and front axle loads for uncorrected data showed significant, but consistent errors, these could form the basis for a procedure to select a representative axle load spectrum from a predefined list, as has been done in other jurisdictions. These efforts would require more research and development, but the aggregated data validity analysis showed that this approach can lead to more insights than simple per-vehicle analysis.

5 CONCLUSIONS AND RECOMMENDATIONS

This chapter presents the key findings and recommendations from the research and presents opportunities for future research in portable WIM systems.

5.1 SUMMARY OF KEY FINDINGS

This research investigated many aspects of the data collected from the portable WIM system, developed and applied several data processing methods, and used those investigations and processing methods to contribute new knowledge about the validity of portable WIM data. The key findings from this research follow:

- Error codes were recorded for 4% of Installation 1 portable WIM records, 28% of Installation 2 records, and 4% of Installation 3 records. The codes recorded indicated that the sensor was functioning only intermittently, and that the errors may have resulted from installation issues. Analysis of successfully paired vehicles by axle count revealed that the erroneous records were more likely to occur for vehicles with three or more axles.
- A measurable calibration drift was found to occur for the duration of Installation 1, which was initially calibrated. The measured trend was found to be consistent for the first 12 days of the installation. Investigation into the best explanation for the drift found no significant differences between time, number of vehicles passed, and number of trucks passed as the explanatory variable.

- Analysis of the effect of temperature on the measured axle loads found that an open-air temperature sensor was better able to predict the front axle loads of the portable WIM system than an in-pavement sensor. The relationship between temperature and selected front axle loads was found to be statistically significant and remained consistent as the calibration of the portable WIM system drifted. Corrections to the load values based on the measured temperature were applied to the loads, but were found to not change the portable WIM values by significant amounts.
- Analysis of the effect of vehicle speed on the measured front axle loads found that speed had a statistically significant effect on non-front axle load accuracies for selected vehicle. However, these effects were not significant in their impact on the loads and were inconsistent across different configurations and axle groups, resulting in an inability to construct correction factors on the basis of vehicle speed.
- Analysis of the per-vehicle data validity found that initial calibration was a strong factor in increasing the accuracy of the portable WIM loads. Temperature correction and autocalibration were found to either decrease both accuracy and precision or provide no improvement.
- Analysis of the first two portable WIM installations found a significant drop in ASTM compliance between the end of Installation 1 and the beginning of Installation 2, indicating that half lane sensor coverage provides less accuracy than full lane sensor coverage when 12-foot sensors are used for both installation methods.

- Analysis of the aggregated data validity separated the portable WIM data into normal distributions for each loading scenario (empty, partially-loaded, fully-loaded) and found that front axle and GVW generally had consistent percent differences in distribution means and percent overlap of the distributions' 95% confidence intervals. This indicated potential usefulness for several indirect applications, but no accuracy gained beyond that seen in the per-vehicle validity analysis. Tandem and tridem axle groups were inconsistent and not recommended for use. Uncorrected data was the most consistent across all three GVW loading scenarios, but post-processed data showed consistently high accuracy for fully loaded vehicles, with most statistical t-tests indicating that the portable WIM's fully loaded, temperature corrected or autocalibrated distributions were not statistically different than those from the comparison dataset.
- Application of the findings of the aggregated data validity analysis in the loading separated temperature correction procedure resulted in per-vehicle data that was not as accurate or precise as the uncorrected data, but was measurably improved from the initial temperature correction procedure, and provided the only non-skewed error histograms.
- Analysis of the per-vehicle data revealed that even the dataset with the most accurate data is unable to meet the accuracy required of ASTM type II WIM data. With the processing and calibration used in this research, error tolerances 2 to 3 times higher than those specified by ASTM type II would be required to allow most load categories to meet this standard.

5.2 SUMMARY OF KEY RECOMMENDATIONS

The key recommendations arising from this research follow:

- Portable WIM installations should be installed carefully and checked to ensure that all parts of the system are operating as expected.
- If 12-foot sensors are the only sensors available, portable WIM installations should install sensors to cover a full lane of traffic.
- Portable WIM installations should be initially calibrated using a test truck.
- When seeking to use portable WIM data on a per-vehicle basis, the data should not be post-processed with any of the methods tested in this research unless no calibration is conducted at the beginning of the installation.
- Any temperature correction should be based on temperature measurements taken by a probe that is left open to the air, rather than installed in the pavement.
- The GMM-based aggregated data approach can be used to gain insights about specific loading scenarios for GVWs, but should not be considered to be more accurate overall than per-vehicle data. When seeking to use GMM analysis to examine fully loaded vehicle GVWs, temperature correction should be applied.
- Portable WIM installations (initially calibrated) should be installed for 7 days (1 week) to ensure sufficient data of the highest possible accuracy.

- Practitioners seeking to utilize portable WIM data should be mindful of the accuracy and precision of the data, and the follow-on implications for various applications.

5.3 OPPORTUNITIES FOR FUTURE PORTABLE WIM RESEARCH

This research has identified the following opportunities for future research into portable WIM systems:

- This research analyzed three installations of the portable WIM system. Performing this research again with a greater number of installations, with attention paid to various installation practices and recorded variables, would serve to clarify and expand upon the findings presented in this research in three ways. First, by installing the portable WIM at sites with varying traffic volumes and classification distributions, the multicollinearity seen between the calibration drift explanatory variables could be eliminated to better explore the difference in explanatory power between the variables. Second, by gathering more data, the relationship between speed and axle loads could be better explained. If a consistent relationship was found, correction factors based on speed could be calculated, applied, and analyzed. Third, a greater number of installations would serve to reinforce and/or clarify the findings of the efficacy of the GMM-based aggregated data approach.
- This research found that both temperature correction and autocalibration decreased both accuracy and precision, but adjusted the peaks of the

error histograms to be closer to 0% error. Moreover, the loading separated temperature correction procedure was able to improve on the initial temperature correction procedure's results. Future research could explore additional alternative versions of these correction procedures to fully realize their benefits and minimize the issues they create.

- This research found that portable WIM data is unable to achieve ASTM type II accuracy standards. Research into the level of accuracy required for freight planning analyses and other indirect applications could reveal to what degree portable WIM data are able to fulfill the data requirements of these applications.
- The data validity analyses suggest a possible use case of the portable WIM data for indirect applications or to support high-level freight transportation planning efforts. Future research could determine what level of data quality is necessary to apply portable WIM data in this way, as well as how to best apply portable WIM data in these applications.

6 REFERENCES

AASHTO (2015) *Mechanistic-Empirical Pavement Design Guide - A Manual of Practice*. 2nd edn. American Association of State Highway and Transportation Officials.

Algohi, B., Mufti, A. and Bakht, B. (2020) 'BWIM with constant and variable velocity: theoretical derivation', *Journal of Civil Structural Health Monitoring*. Springer Berlin Heidelberg, 10(1), pp. 153–164. doi: 10.1007/s13349-019-00377-0.

Allstays LLC (2020) *Weigh Scales Locations*. Available at: https://www.allstays.com/c/weigh-scales-locations.htm (Accessed: 20 February 2020).

American Transportation Research Institute (2012) *National Corridors Analysis and Speed Tool (N-CAST)*, *American Transportation Research Institute*. Available at: http://truckingresearch.org/2012/10/23/n-cast/ (Accessed: 5 March 2018).

ASTM International (2009) *Standard specification for highway weigh-in-motion (WIM) systems with user requirements and test methods*. doi: 10.1520/E1318-09R17.1.4.

Auditor General of Newfoundland and Labrador (2012) 'Report of the Auditor General to the House of Assembly on Reviews of Departments and Crown Agencies', (June).

B.C. Ministry of Transportation and Infrastructure (no date) *List of Inspection Stations*. Available at: https://www.th.gov.bc.ca/cvse/inspection_stations.htm (Accessed: 18 February 2020).

Burnos, P. and Gajda, J. (2016) 'Thermal property analysis of axle load sensors for weighing vehicles in weigh-in-motion system', *Sensors (Switzerland)*, 16(12). doi: 10.3390/s16122143.

Burnos, P. and Gajda, J. (2020) 'Optimised autocalibration algorithm of weigh-in-motion systems for direct mass enforcement', *Sensors (Switzerland)*, 20(11). doi: 10.3390/s20113049.

Cantero, D. and González, A. (2015) 'Bridge damage detection using weigh-in-motion technology', *Journal of Bridge Engineering*, 20(5), pp. 1–10. doi: 10.1061/(ASCE)BE.1943-5592.0000674.

Cetin, M. *et al.* (2011) 'Key factors affecting the accuracy of reidentification of trucks over long distances based on axle measurement data', *Transportation Research Record*, (2243), pp. 1–8. doi: 10.3141/2243-01.

Cetin, M., Monsere, C. M. and Nichols, A. P. (2011) 'Bayesian models for reidentification of trucks over long distances on the basis of axle measurement data', *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 15(1), pp. 1–12. doi: 10.1080/15472450.2011.544571.

Cetin, M., Nichols, A. R. and Chou, C. S. (2014) 'Reidentification of trucks on basis of axle-spacing measurements to facilitate analysis of weigh-in-motion accuracy', *Transportation Research Record*. National Research Council, pp. 96–105. doi: 10.3141/2443-11.

Chotickai, P. and Bowman, M. D. (2006) 'Truck Models for Improved Fatigue Life Predictions of Steel Bridges', *Journal of Bridge Engineering*, 11(1), pp. 71–80. doi: 10.1061/共ASCE兲1084-0702共2006兲11:1共71兲.

Drivewyze (2020) *For Drivers*. Available at: https://drivewyze.com/for-drivers/ (Accessed: 20 February 2020).

Dybing, A. and Dakota, N. (2017) 'Agricultural Trip Generation - Linking Spatial Data and Travel Demand Modeling using SAS ®', pp. 1–9.

Enright, B., Leahy, C. and O'Brien, E. (2012) 'Portable Bridge WIM Data Collection Strategy for Secondary Roads', pp. 402–409. doi: 10.21427/D7G51N.

European Commision (2018) *Tachograph*. Available at: https://ec.europa.eu/transport/modes/road/social-provisions/tachograph_en (Accessed: 10 April 2018).

Evans, D. and Klashinsky, R. (2012) 'Weigh-In-Motion System for the Reduction of Commercial Vehicle Traffic Delays at the Borden-Carleton, PEI Weight Inspection Station', in *Transportation Association of Canada Annual Conference Proceedings*. Frederickton, NB, p. 14.

Faraz, S. *et al.* (2017) ‘Sources of errors in fatigue assessment of steel bridges using BWIM’, *Journal of Civil Structural Health Monitoring*. Springer Berlin Heidelberg, 7(3), pp. 291–302. doi: 10.1007/s13349-017-0225-8.

Faruk, A. N. M. *et al.* (2016) ‘Traffic volume and load data measurement using a portable weigh in motion system: A case study’, *International Journal of Pavement Research and Technology*. Chinese Society of Pavement Engineering, 9(3), pp. 202–213. doi: 10.1016/j.ijprt.2016.05.004.

Federal Motor Carrier Safety Administration (2018) *Electronic Logging Devices*. Available at: https://www.fmcsa.dot.gov/hours-service/elds/electronic-logging-devices (Accessed: 10 April 2018).

Fitzpatrick, A. (2016) ‘Weight cameras not being used for enforcement’, *The Telegram*, 14 November. Available at: https://www.thetelegram.com/news/local/weight-cameras-not-being-used-for-enforcement-128037/.

Gajda, J. *et al.* (2013) ‘The Influence of Temperature on Errors of WIM Systems Employing Piezoelectric Sensors’, *Metrology and Measurement Systems*, 1(2), pp. 171–182. doi: 10.2478/mms-2013-0015.J.Gajda.

Government of Alberta (2020) *Commercial vehicle inspection stations and branch offices*. Available at: https://www.alberta.ca/commercial-vehicle-inspection-stations-offices.aspx (Accessed: 18 February 2020).

Government of British Columbia (2019) *Traffic Data Program GIS Application*. Available at: https://prdoas3.pub-apps.th.gov.bc.ca/tsg/ (Accessed: 13 February 2020).

Government of British Columbia (no date) *Weigh2GoBC Inspection Stations*. Available at: https://www2.gov.bc.ca/gov/content/transportation/vehicle-safety-enforcement/services/weigh2gobc-join/weigh2gobc-inspection-stations (Accessed: 18 February 2020).

Government of Ontario (2020) *Inspection stations*. Available at: https://511on.ca/list/inspectionstations?start=0&length=25&order%5Bi%5D=0&order%5Bdir%5D=asc (Accessed: 20 February 2020).

Government of Saskatchewan (2020) *Weigh scale locations*. Available at: https://www.saskatchewan.ca/business/transportation-and-road-construction/information-for-truckers-and-commercial-trucking-companies/regulations-and-road-restrictions/weigh-scale-locations (Accessed: 19 February 2020).

Government of Yukon (2020) *Find a weigh station*. Available at: https://yukon.ca/en/driving-and-transportation/weigh-stations/find-weigh-station (Accessed: 19 February 2020).

Haider, S. *et al.* (2011) 'Development of Traffic Inputs for *Mechanistic-Empirical Pavement Design Guide* in Michigan', *Transportation Research Record: Journal of the Transportation Research Board*, 2256, pp. 179–190. doi: 10.3141/2256-21.

Hallenbeck, M. and Weinblatt, H. (2004) *Equipment for collecting traffic load data*. Washington, D.C.

Hanson, R., Klashinsky, R. and McGibney, S. (2010) ‘ITS Technologies for Commercial Vehicle Compliance in the Maritimes’, in *Transportation Association of Canada Conference Proceedings*. Halifax, NS, p. 15.

Helmi, K., Bakht, B. and Mufti, A. (2014) ‘Accurate measurements of gross vehicle weight through bridge weigh-in-motion: A case study’, *Journal of Civil Structural Health Monitoring*. Springer Verlag, 4(3), pp. 195–208. doi: 10.1007/s13349-014-0076-5.

Hyun, K. (Kate), Tok, A. and Ritchie, S. G. (2017) ‘Long distance truck tracking from advanced point detectors using a selective weighted Bayesian model’, *Transportation Research Part C: Emerging Technologies*. Elsevier Ltd, 82, pp. 24–42. doi: 10.1016/j.trc.2017.06.004.

Ieng, S. S. (2015) ‘Bridge influence line estimation for bridge weigh-in-motion system’, *Journal of Computing in Civil Engineering*, 29(1), pp. 1–4. doi: 10.1061/(ASCE)CP.1943-5487.0000384.

Jacob, B. and Cottineau, L. M. (2016) ‘Weigh-in-motion for Direct Enforcement of Overloaded Commercial Vehicles’, *Transportation Research Procedia*. Elsevier B.V., 14(0), pp. 1413–1422. doi: 10.1016/j.trpro.2016.05.214.

Jacob, B., O’Brien, E. and Jehaes, S. (2002) *COST 323 Weigh-in-Motion of*

Road Vehicles: Final Report. Paris.

Jaworski, D. (2018) ‘Personal Communication’. Winnipeg.

Jeng, S.-T. and Chu, L. (2015) ‘Tracking Heavy Vehicles Based on Weigh-In-Motion and Inductive Loop Signature Technologies’, *IEEE Transactions on Intelligent Transportation Systems*, 16(2), pp. 632–641. doi: 10.1109/tits.2014.2333003.

Jeng, S. T., Chu, L. and Hernandez, S. (2013) ‘Wavelet-k nearest neighbor vehicle classification approach with inductive loop signatures’, *Transportation Research Record*. National Research Council, (2380), pp. 72–80. doi: 10.3141/2380-08.

Jeng, S. T. and Ritchie, S. G. (2008) ‘Real-time vehicle classification using inductive loop signature data’, *Transportation Research Record*, (2086), pp. 8–22. doi: 10.3141/2086-02.

Kalyankar, R. and Uddin, N. (2017) ‘Axle detection on prestressed concrete bridge using bridge weigh-in-motion system’, *Journal of Civil Structural Health Monitoring*. Springer Berlin Heidelberg, 7(2), pp. 191–205. doi: 10.1007/s13349-017-0210-2.

Koniditsiotis, C., Buckmaster, R. and Fraser, P. (1995) ‘Australian high speed weigh-in-motion - an overview’, *Proceedings of the International Symposium on Heavy Vehicle Weights and Dimensions*, pp. 143–151.

Krisztin, T. (2017) ‘The Determinants of Regional Freight Transport: A Spatial, Semiparametric Approach’, *Geographical Analysis*, 49(3), pp. 268–308. doi: 10.1111/gean.12125.

Kulpa, T. (2014) ‘Freight Truck Trip Generation Modelling at Regional Level’, *Procedia - Social and Behavioral Sciences*. Elsevier B.V., 111, pp. 197–202. doi: 10.1016/j.sbspro.2014.01.052.

Kwon, T. M. (2012) *Development of a Weigh-Pad-Based Portable Weigh-In-Motion System*. Minneapolis.

Labarrere, J. (2018a) *Bridge WIM*, *International Society for Weigh In Motion*. Available at: http://www.is-wim.org/index.php?nm=2&nsm=6&lg=en (Accessed: 3 April 2018).

Labarrere, J. (2018b) *On-board WIM*, *International Society for Weigh In Motion*. Available at: http://www.is-wim.org/index.php?nm=2&nsm=7&lg=en (Accessed: 23 April 2019).

Laplante, J. (2020) ‘Personal Communication’. Winnipeg, MB.

Van de Lindt, J. W. *et al.* (2002) ‘Locality of Truck Loads and Adequacy of Bridge Design Load’, *Journal of Bridge Engineering*, 10(5), pp. 622–629. doi: 10.1061/(ASCE)1084-0702(2005)10.

van Loo, H. and Znidaric, A. (2019) *Guide for Users of Weigh-In-Motion*. 1st edn. International Society for Weigh-In-Motion.

Lydon, M. *et al.* (2016) ‘Recent developments in bridge weigh in motion (B-WIM)’, *Journal of Civil Structural Health Monitoring*. Springer Berlin Heidelberg, 6(1), pp. 69–81. doi: 10.1007/s13349-015-0119-6.

Manitoba Infrastructure (2017) ‘Manitoba Weigh Stations’. Winnipeg, MB. Available at: https://www.gov.mb.ca/mit/mcd/resources/pdf/twlm/weigh_stations.pdf.

Manitoba Infrastructure Motor Carrier Division (2017) *Vehicle Weights and Dimensions Guide*. Winnipeg, MB.

Marszalek, Z. and Duda, K. (2018) ‘Vehicle Magnetic Signature Compatibility for Inductive Loop Sensor with Various Signal Conditioning Systems’, in *International Conference on Signals and Electronic Systems (ICSES)*. Krakow, Poland, pp. 33–36.

Nassiri, S., Bayat, A. and Kilburn, P. (2014) ‘Traffic inputs for mechanistic-empirical pavement design guide using weigh-in-motion systems in Alberta’, *International Journal of Pavement Engineering*. Taylor & Francis, pp. 483–494. doi: 10.1080/10298436.2013.797978.

New Brunswick Department of Transportation and Infrastructure (no date) *Weigh-in-Motion*. Available at: https://www2.gnb.ca/content/gnb/en/departments/dti/trucking/content/weigh-in-motion.html (Accessed: 13 February 2020).

Northwest Territories Department of Infrastructure (2017) *Northwest Territories 2016 Highway Traffic Report*. Yellowknife.

Nova Scotia Department of Transportation and Infrastructure Renewal (2017) *Traffic Counts and Analysis 2008-2017*. Halifax, NS.

Obrien, E. J. *et al.* (2018) 'Probabilistic bridge weigh-in-motion', *Canadian Journal of Civil Engineering*, 45(8), pp. 667–675. doi: 10.1139/cjce-2017-0508.

Oh, C., Tok, A. and Ritchie, S. G. (2005) 'Real-time freeway level of service using inductive-signature-based vehicle reidentification system', *IEEE Transactions on Intelligent Transportation Systems*, 6(2), pp. 138–146. doi: 10.1109/TITS.2005.848360.

Ontario Ministry of Transportation (2012) *Commerical vehicle activities*. Available at: https://mto-on-ca.maps.arcgis.com/apps/MapSeries/index.html?appid=a0cc97d7ca8643998033b5ba7183e838 (Accessed: 20 February 2020).

Petersen, S. (2015) *Portable Weigh-in-Motion System Evaluation*. Minneapolis.

Poapst, R. (2020) 'Personal Communication'. Winnipeg, MB.

Prince Edward Island Department for Transportation and Infrastructure Renewal (2009) *Transportation and Infrastructure Renewal Annual Report 2009-2010*. Charlottetown, PEI.

Province of Nova Scotia (2013) *Vehicle Compliance*. Available at: https://novascotia.ca/tran/trucking/compliance.asp (Accessed: 14 February 2020).

Refai, H., Othman, A. and Tafish, H. (2014) *Portable Weigh-In-Motion for Pavement Design - Phases 1 and 2*. Oklahoma City.

Regehr, J. D. *et al.* (2020) 'Gaussian mixture model to characterize payload distributions for predominant truck configurations and body types', *Journal of Traffic and Transportation Engineering, Part B: Pavements*, 2(146). doi: 10.1061/JPEODX.0000174.

Reimer, M. and Regehr, J. (2013) 'Hybrid Approach for Clustering Vehicle Classification Data to Support Regional Implementation of the Mechanistic-Empirical Pavement Design Guide', *Transportation Research Record: Journal of the Transportation Research Board*, 2339(204), pp. 112–119. doi: 10.3141/2339-13.

Richardson, J. *et al.* (2014) 'On the use of bridge weigh-in-motion for overweight truck enforcement', *International Journal of Heavy Vehicle Systems*, 21(2), pp. 83–104. doi: 10.1504/IJHVS.2014.061632.

Sanderson, B. (2018) *No more 'cheat sheets': Electronic logs for truckers mandatory by 2020*, *CBC News*. Available at: https://www.cbc.ca/news/canada/nova-scotia/canada-logging-trucks-electronic-devices-1.4540607 (Accessed: 20 February 2020).

Saskatchewan Highway Traffic Board (2019) ‘Nunavut Professional Driver’s Manual’. Nunavut Economic Development and Transportation. Available at: https://www.gov.nu.ca/sites/default/files/professional_drivers_handbook.pdf.

Saskatchewan Ministry of Highways and Infrastructure (2019) *Ministry of Highways and Infrastructure Plan for 2019-20*. Regina, SK.

Sekiya, H., Kubota, K. and Miki, C. (2018) ‘Simplified Portable Bridge Weigh-in-Motion System Using Accelerometers’, *Journal of Bridge Engineering*, 23(1), pp. 1–15. doi: 10.1061/(ASCE)BE.1943-5592.0001174.

Selezneva, O. and Von Quintus, H. (2014) *Traffic load spectra for implementing and using the mechanistic-empirical pavement design guide in georgia*. Forest Park.

Selezneva, O. and Wolf, D. (2017) *Successful Practices in Weigh-in-Motion Data Quality with WIM Guidebook Volume 1*. Phoenix.

Stephens, J. *et al.* (2017) ‘Montana Weigh-In-Motion and Automatic Traffic Recorder Strategy’, p. 276.

Sun, B. (2020) ‘Personal Communication’. Winnipeg, MB.

Sun, C. *et al.* (1999) *Use of vehicle signature analysis and lexicographic optimization for vehicle reidentification on freeways*. Available at: www.elsevier.com/locate/trc.

Sureshan, S. (2020) ‘Personal Communication’. Winnipeg, MB.

Swan, D. *et al.* (2008) ‘Development of Regional Traffic Data for the Mechanistic-Empirical Pavement Design Guide’, *Transportation Research Record: Journal of the Transportation Research Board*, 2049(2049), pp. 54–62. doi: 10.3141/2049-07.

Tan, E. (2002) ‘Weight characteristics of predominant truck configurations in Manitoba’. Winnipeg, MB: University of Manitoba Department of Civil Engineering.

Task Force on Vehicle Weights and Dimensions Policy (2019) ‘Territorial Developments Round Table Report’. Government of Yukon. Available at: http://www.comt.ca/english/programs/trucking/2019/Yukon Update.pdf.

Texas A&M Transportation Institute (2015) ‘Intelligent Freight Monitoring : A Review of Potential Technologies’.

Thom, R. (2020) ‘Personal Communication’. Winnipeg, MB.

Todts, W. *et al.* (2013) ‘Study on heavy vehicle on-board weighing, Final report’, *Rapp Trans*, p. 49.

Traffic Services Saskatchewan Ministry of Highways and Infrastructure (2016) *Travel on Saskatchewan Highways 2016*. Regina, SK.

Transportation Modelling and Analysis Section (2014) *Weigh in Motion Report*

2014. Edmonton, AB.

Turochy, R. E., Timm, D. H. and Tisdale, S. M. (2005) *Truck Equivalency Factors , Load Spectra Modeling and Effects on Pavement Design*.

U.S. Department of Transportation (2018a) *Weigh-In-Motion Pocket Guide Part 1 WIM Technology, Data Acquisition, and Procurement Guide*. Washington.

U.S. Department of Transportation (2018b) *Weigh-In-Motion Pocket Guide Part 2 WIM Site Selection, Design, and Installation Guide*. Washington.

U.S. Department of Transportation (2018c) *Weigh-In-Motion Pocket Guide Part 3 WIM Calibration and Maintenance Guide*. Washington.

Wood, S. (2017) *Options for Providing Quality Axle Load Data for Pavement Design*. University of Manitoba.

Yu, Y., Cai, C. S. and Deng, L. (2016) 'State-of-the-art review on bridge weigh-in-motion technology', *Advances in Structural Engineering*, 19(9), pp. 1514–1530. doi: 10.1177/1369433216655922.

Zhang, L., Haas, C. and Tighe, S. L. (2007) 'Evaluating Weigh-In-Motion Sensing Technology for Traffic Data Collection', *Transportation Association of Canada*, pp. 1–17.

Zhao, H. *et al.* (2014) 'Identification of vehicular axle weights with a bridge weigh-in-motion system considering transverse distribution of wheel loads', *Journal of*

Bridge Engineering, 19(3), pp. 1–16. doi: 10.1061/(ASCE)BE.1943-5592.0000533.

Žnidarič, A., Kalin, J. and Kreslin, M. (2018) 'Improved accuracy and robustness of bridge weigh-in-motion systems', *Structure and Infrastructure Engineering*. Taylor & Francis, 14(4), pp. 412–424. doi: 10.1080/15732479.2017.1406958.

Zolghadri, N. *et al.* (2016) 'Field Verification of Simplified Bridge Weigh-in-Motion Techniques', *Journal of Bridge Engineering*, 21(10), pp. 1–17. doi: 10.1061/(ASCE)BE.1943-5592.0000930.

APPENDIX A: VEHICLE PAIRING PROCEDURES

To assess the per-vehicle validity of the portable WIM axle load measurements, the individual vehicle records generated by the portable WIM system and the comparison datasets needed to be paired. The two sources of comparison data – the Headingley weigh scale and WIM station 99 – have several key differences, and two different vehicle pairing methods were required.

WIM Station 99 Vehicle Pairing Procedure

The 2 portable WIM installations at WIM station 99 each had 2 datasets that required pairing: (1) the per-vehicle records from the portable WIM system, and (2) the per-vehicle records from WIM station 99. The pairing process was done primarily by matching the vehicle record timestamps; Figure 32 provides a flowchart of the steps in the vehicle pairing process. The steps referenced in Figure 32 are described in detail after the figure.

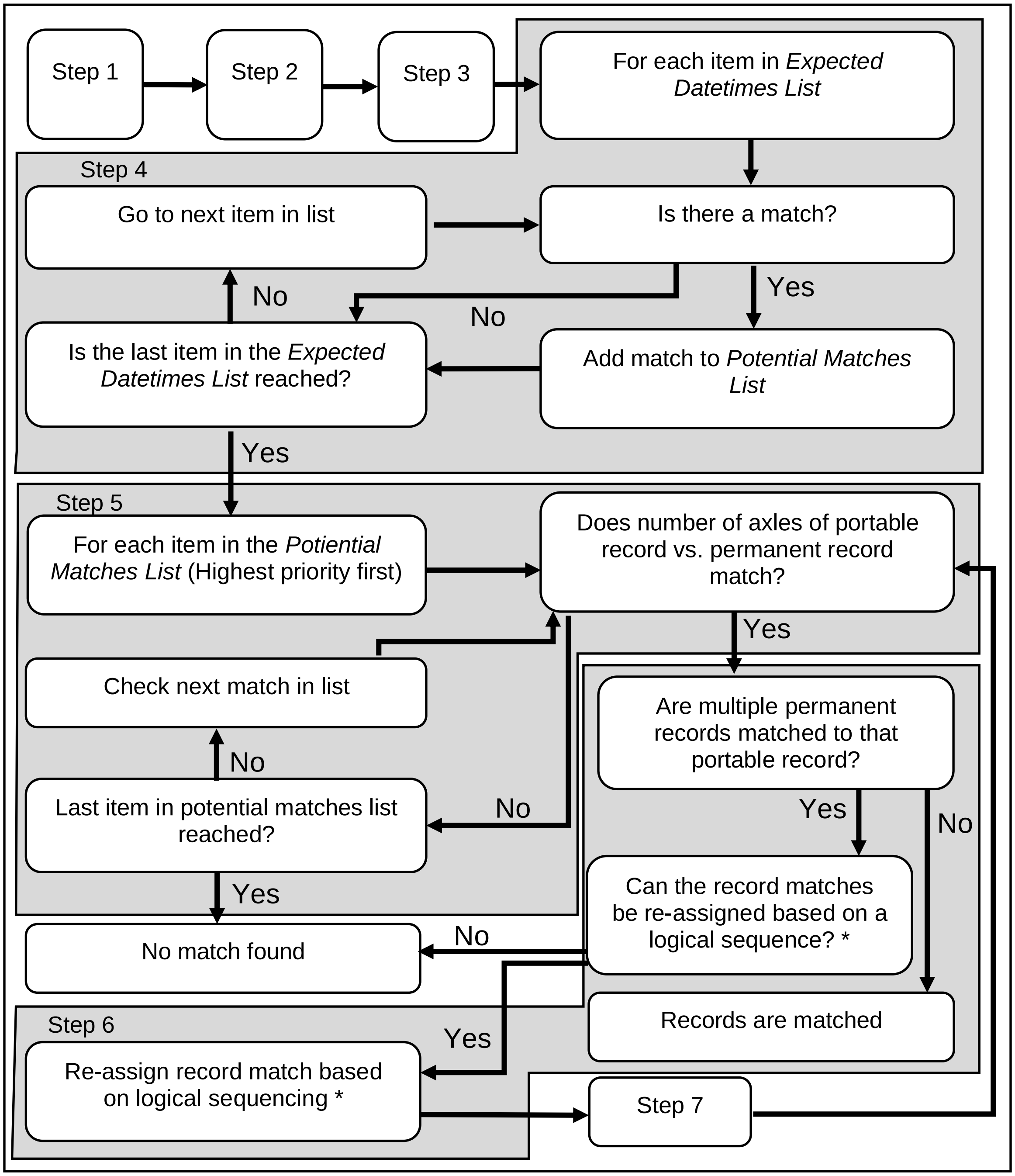


Figure 32 - WIM station 99 vehicle pairing procedure algorithm

The steps listed in Figure 32 are described below.

1. The fields *Year, Month, Day, Hour, Minute,* and *Second* were combined into a single *Datetime* field for both the portable and permanent WIM datasets. Due to different date settings, *Datetime* was not an exact match between the two datasets. To calculate the difference in *Datetime*, two uncommon truck classes were isolated using the automatically calculated *Class* field. These two classes (class 15 and class 4) had low sample sizes in Installation 1 so they could be easily manually paired, ensuring that the difference in *Datetime* between the portable and permanent datasets was approximately equal in all pairs. 11 of the 12 class 15 records were paired, giving a consistent difference in *Datetime* (*DatetimeDifference*). *DatetimeDifference* was found to change part of the way through Installation 1 as the clock settings were re-set to be more accurate; the two values of *DatetimeDifference* were used to pair 103 of the 110 class 4 records. The paired class 4 record set was of sufficient size to change in the datetime difference for each day of installation within a pair of records, or 'drift' ($\Delta_{DatetimeDifference}$). These variables were used to create Equation 5, which predicts the permanent WIM *Datetime* value for each portable WIM vehicle record.

Equation 5

$$Datetime_{Permanent} = \Delta_{DatetimeDifference} * DaysElapsed + Datetime_{Portable} + DatetimeDifference$$

2. Create a tentative match for each portable WIM vehicle record based on its *Datetime* value and number of axles. Each record had a value of

Datetime$_{permanent}$ calculated using Equation 5 to serve as a starting point when looking for a matching permanent WIM *Datetime*.

3. Create a list of additional potential permanent WIM *Datetime* values (the *Expected Datetimes List*) to consider to account for small variations from the expected value. This list of potential permanent WIM *Datetime* values was given an order of priority; Table 25 shows these assigned priorities.

Table 25 – Priority values of the Expected *Datetimes* List

Priority	Expected Permanent WIM *Datetime* Value
1	*Datetime*$_{permanent}$
2	*Datetime*$_{permanent}$ + 1 second
3	*Datetime*$_{permanent}$ - 1 second
4	*Datetime*$_{permanent}$ + 2 seconds
5	*Datetime*$_{permanent}$ - 2 seconds
6	*Datetime*$_{permanent}$ + 3 seconds
7	*Datetime*$_{permanent}$ - 3 seconds
8	*Datetime*$_{permanent}$ + 4 seconds
9	*Datetime*$_{permanent}$ - 4 seconds
10	*Datetime*$_{permanent}$ + 5 seconds
11	*Datetime*$_{permanent}$ - 5 seconds

4. Consider each item in the *Expected Datetimes List* in order of priority, as a potential match value for a permanent WIM *Datetime*. Any permanent WIM *Datetime* that matched one of the predicted values from the list was considered a potential match and was added to the *Potential Matches List*.
5. Starting with the highest priority value item in the *Potential Matches List*, compare the number of axles between the potentially matched records.

The first match that had an equal number of axles for the portable WIM and permanent WIM records was the tentative vehicle record match.

6. Fix those situations where multiple portable WIM vehicle records were initially paired to a single permanent WIM vehicle record by re-assigning them to permanent records based on 'logical sequencing'. The logical sequencing method ignores the vehicle *Datetime* values and uses the order in which the portable and permanent records are listed in to make a match. If the number of consecutive portable vehicle records that matched with a single permanent WIM vehicle record is equal to the number of permanent WIM vehicle records between successful matches (matches with only one portable WIM vehicle record matched with the permanent WIM vehicle record), the portable WIM vehicle records that matched to the single permanent WIM vehicle record are re-assigned to the permanent WIM vehicle records between successful matches, maintaining the order of the permanent WIM records. Figure 33 displays an example of this process.

Successful match

Multiple match

Portable WIM Vehicle Record Number	Permanent WIM Vehicle Record Number
1554	1679
1555	1680
1556	1682
1557	1682
1558	1682
1559	1684

Initial Match: 3 portable WIM records are matched to a single permanent WIM vehicle record

Portable WIM Vehicle Record Number	Permanent WIM Vehicle Record Number
1554	1679
1555	1680
1556	1681
1557	1682
1558	1683
1559	1684

Re-assigned Match: 3 portable WIM records are now assigned to unique permanent WIM records

Figure 33 - Example of logical sequencing vehicle record pairing method

7. Re-check all records that were re-matched with the logical sequencing method to ensure that the new permanent WIM record fell within +/- 5 seconds of $Datetime_{permanent}$ and that the number of axles of the two records were equal. If either of these conditions is not met, the match is discarded and no match is used.

<u>Headingley Weigh Scale Vehicle Pairing Procedure</u>

Portable WIM Installation 3 at the Headingley weigh scale yielded 3 sources of data: (1) the per-vehicle records from the portable WIM system, (2) the manually entered per-vehicle records from the Headingley weigh scale, and (3) the video taken of vehicles passing the portable WIM system. Figure 34 displays a flowchart of the steps in the vehicle pairing process. The steps referenced in the figure are described in detail after the figure.

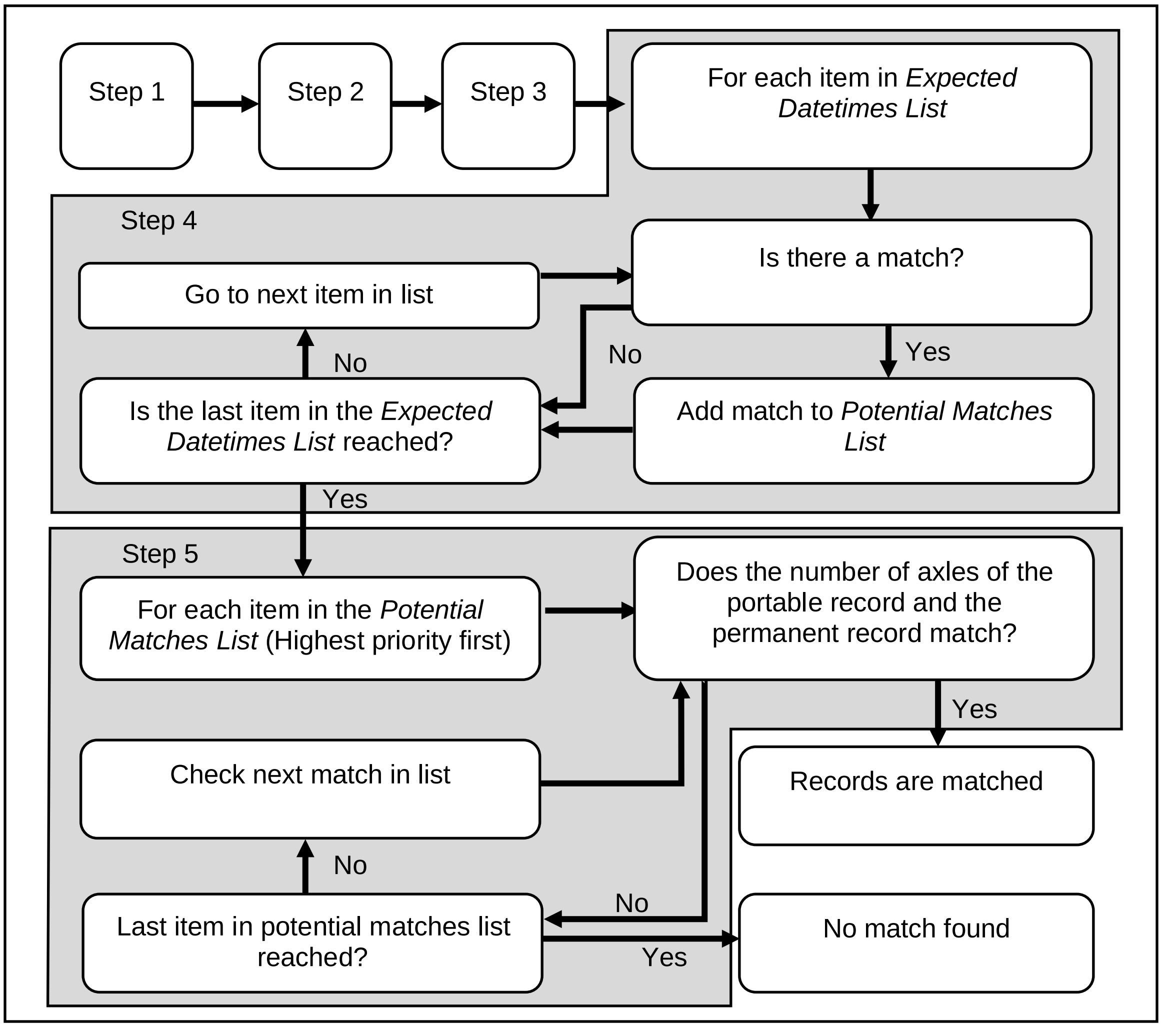


Figure 34 - Static weigh scale vehicle pairing algorithm

The steps listed in Figure 34 are described below.

1. Identify each vehicle that passed over the static weigh scale in the Miovision video taken at the portable WIM system installation. This was

done manually; the vehicles in the video were identified by axle configuration, body type, and visual characteristics, combined with the knowledge that the timestamp on the static scale data would be between 4 and 8 minutes behind the timestamp on the video. Figure 35 gives an example video frame with the descriptive information that was used to identify the pictured vehicle.

Description	Blue cab
Configuration	3-S3

Body Type	Flat deck

Figure 35 - Screenshot from camera installed at portable WIM system with accompanying information

2. Calculate the expected $Datetime_{portable}$ for each static scale record. To do this, the video timestamps for each vehicle record from the static scale were combined with the dates they were recorded on to create *Datetime* values for each static scale record. Then, the first several of these *Datetimes* were matched with portable WIM records that had the same

number of axles, ensuring that the difference in timestamp was consistent between these pairs. This established the *DatetimeDifference* for the installation. The short duration of the test meant that the drift in *Datetime* difference was not significant enough to require calculation.

3. Create a list of additional potential portable WIM *Datetime* values (the *Expected Datetimes List*) whose consideration would account for small variations from the expected value. This list of potential portable WIM *Datetime* values was given an order of priority as in step 3 of the WIM station 99 pairing procedure; Table 25 shows these assigned priorities.
4. Consider each item in the *Expected Datetimes List*, in order of priority, as a potential match value for a permanent WIM *Datetime*. Any permanent WIM *Datetime* that matched one of the predicted values from the list was considered a potential match, and was added to the *Potential Matches List*.
5. Starting with the highest priority value item in the *Potential Matches List*, compare the number of axles between the potentially matched records. The first match that had an equal number of axles for the portable WIM and permanent WIM records was the tentative vehicle record match.

There were no instances of multiple records from either dataset pairing to a single record from the other, so the logical sequencing method was not required for this process.

APPENDIX B: PERMANENT WIM STATION 99 DATA EVALUATION

The first preliminary data quality assessment was of the permanent WIM station 99 data. WIM data is always specified to have accuracy only within a certain range (ASTM International, 2009), so a degree of inaccuracy and imprecision was expected. Therefore, while the in-road piezo-quartz sensors of WIM station 99 provide higher quality data than the portable WIM system, the permanent WIM system cannot be considered to be a source of ground truth data in the same way the static weigh scale can. Here, ground truth data refers to data collected from the most accurate source available, which in the case of axle load data refers to data collected from a static weigh scale. This assessment determined the validity of the permanent WIM station data by comparing the five weeks of permanent WIM data during the portable WIM installations to each other, then comparing the permanent WIM data to the static scale data.

To calculate the precision of the WIM station 99 data over the relevant period (the five weeks of the portable WIM system installations), the four selected vehicle configurations were considered: 3-S2, 3-S3, 3-S3-S2, and 3-S2-4 vehicles. These four configurations are FHWA class 9, 10, 13, and 15 respectively, but as each of these classes includes configurations other than the four listed, the selection algorithm described in Appendix D was used to select only the correct configurations. These configurations were selected because they are among the most common configurations found in southern Manitoba, which means both that it is important to weigh these configurations in particular accurately, and that sufficient sample sizes of all four configurations can be found for statistical purposes.

To evaluate the precision over time of the selected configurations, statistical t-tests were used to compare the load distribution of each configuration's axle groups and GVW in the first week of portable WIM installation, which was also the first week after station 99's calibration, to each of the other 4 weeks of the portable WIM system installations. Table 26 shows the p-value of each of these t-tests, with those t-tests that failed to reject the null hypothesis at a p-value of 0.05 highlighted. These highlighted p-values are the tests that show statistically similar axle group load or GVW distributions, while unhighlighted cells show statistical differences.

Table 26 – P-values from t-tests comparing weeks of WIM station 99 axle load data over the portable WIM installation

Configuration	Axle Group	2 Weeks Post-Calibration	3 Weeks Post-Calibration	4 Weeks Post-Calibration	5 Weeks Post-Calibration
3-S2	Single	0.0297	0.0700	0.331	0.836
	Tandems	0.00535	0.0215	0.196	0.427
	GVW	0.0310	0.0757	0.384	0.576
3-S3	Single	0.0429	0.119	0.491	0.133
	Tandem	0.632	0.00345	0.444	0.879
	Tridem	0.867	0.0363	0.735	0.758
	GVW	0.869	0.0109	0.880	0.734
3-S3-S2	Single	0.755	0.998	0.769	0.836
	Tandems	0.0828	0.000768	0.00214	0.000736
	Tridem	0.134	0.0124	0.0295	0.00211
	GVW	0.165	0.0117	0.0259	0.00289
3-S2-4	Single	0.148	0.719	0.116	0.0218
	Tandems	0.467	0.426	0.0263	0.00345
	GVW	0.645	0.580	0.0882	0.0225

These results show the station 99 axle load measurements to be largely self-consistent in time. Of the 56 t-tests conducted, 35 (63%) failed to find statistically significant differences between the datasets, and in any one week, at least 7 out of 14 tests failed to find differences. Those tests that did find statistically significant differences in the data that could be attributed to one of three explanations:

1) The vehicles that passed WIM station 99 in that week had significantly different axle loads than the first week of the portable WIM installation,

2) The WIM sensors shifted in their calibration and recorded different axle loads despite the axle loads truly being similar,
3) The t-test results are displaying Type I error, where a null hypothesis that is correct is rejected due to the specific samples that were used.

The only one of these explanations that would indicate that the WIM station 99 data is not an appropriate comparison dataset is explanation 2. However, the lack of a clear trend toward greater differences in the datasets indicates that this is not likely the explanation. Therefore, the WIM station 99 data can be said to be sufficiently self-consistent to use as a comparison dataset for the portable WIM system.

To determine the accuracy of the WIM station 99 data, percentiles from 0.05 to 0.95 in increments of 0.05 were calculated for each axle group's weight for each configuration for both WIM station 99 and the Headingley weigh scale. With the Headingley weigh scale percentiles taken as the real values, the percent error of each the WIM station 99 percentiles was calculated. Calculating the error of each percentile between the two data sources allows the analysis to identify the magnitude of any systematic differences between the two data source's weight measurements to be identified. Figure 36 displays graphs of these percent error values; these graphs show how the error of the permanent WIM load values changes for different load magnitudes

Figure 36 – Percent error of WIM station 99 weight percentiles as compared to static weigh scale data

These graphs show that the single (front) axles of trucks are measured approximately 7% higher at WIM station 99 that at the static weigh scale, but that other axle groups and GVWs have percentile errors of up to 26% in 3-S3 vehicles, 70% in 3-S3-S2 vehicles, and 16% in 3-S2-4 vehicles. The fact that this is seen most pronounced in 3-S3-S2 vehicles, specifically in the tridem axle group, and at higher loads of the 3-S3 vehicles, indicates that these differences are not solely attributable to issues with the WIM's accuracy, but also indicate that the vehicle samples captured at the two locations have different loading scenarios (i.e. a higher percentage of loaded vs. unloaded vehicles), causing inaccuracy in those configurations that typically carry heavier loads. Due to the consistently low values of the single (front) axle load error, as well as the relatively low error values of the other axle groups and GVW of 3-S2 vehicles, which have the largest sample size of the four configurations, the weight data from WIM station 99 will be said to be valid as a comparison dataset for the portable WIM system.

APPENDIX C: POST-PROCESSING CALIBRATION PROCEDURE AND EVALUATION

One of the primary data limitations of the research was the inability to capture data comparing the calibrated portable WIM system measurements to static weigh scale data. To compensate for this, a post-processing calibration procedure was devised which would allow the uncalibrated portable WIM data collected at the static weigh scale to be calibrated using a limited number of static weigh scale records, after which the calibration would be complete and the portable WIM system allowed to drift in its calibration as it would in a real calibration scenario. This process is an entirely separate procedure from autocalibration, which is described in 3.4.4.

The post-processing calibration procedure is similar to the calibration procedure described in section 3.2.3. Equation 6 and Equation 7 generated the values for $F_{calibration}$ and $F_{dynamic}$ in the post-processing calibration. These equations were derived from Equation 3 and Equation 4 in section 3.2.3, but instead of a single static scale load, they use the paired static scale loads from the manual data collection at the Headingley weigh scale.

Equation 6

$$F_{calibration} = \frac{\left(\sum_{record=1}^{n} Load_{Drive,\ Static\ Scale} + \sum_{record=1}^{n} Load_{Rear\ tandem,\ Static\ Scale}\right)}{\left(\sum_{record=1}^{n} Load_{Drive,\ WIM} + \sum_{record=1}^{n} Load_{Rear\ tandem,\ WIM}\right)}$$

Equation 7

$$F_{dynamic} = \frac{\left(\sum_{record=1}^{10} Load_{Steer,\ WIM}\right)}{\left(\sum_{record=1}^{10} Load_{Steer,\ WIM}\right) * F_{calibration}}$$

These factors were calculated and applied once each for the entire system, instead of calculating them separately for each WIM sensor. To imitate the actual calibration procedure as closely as possible, only successfully paired 3-S2 vehicles that were loaded to >=35000 kg as determined by the static scale and were traveling >=80 km/h as measured by the portable WIM system were used for calibration; 59 records met this criteria.

As the post-processing calibration uses multiple unique 'test truck' records from the static scale data instead of a single test truck, it was possible that the post-processing calibration would not accurately represent a real calibration. The two issues here were:

1) Vehicles of multiple weights were used in the post-processing calibration, and
2) The vehicles that were used were recorded over an entire week instead of on a single day.

To ensure that no inaccuracy in the calibration was introduced by the multiple vehicle weights, the accuracy of the uncalibrated WIM measurement of each vehicle's steering and combined tandem axles was calculated; Figure 37 displays the scatterplots of these accuracies.

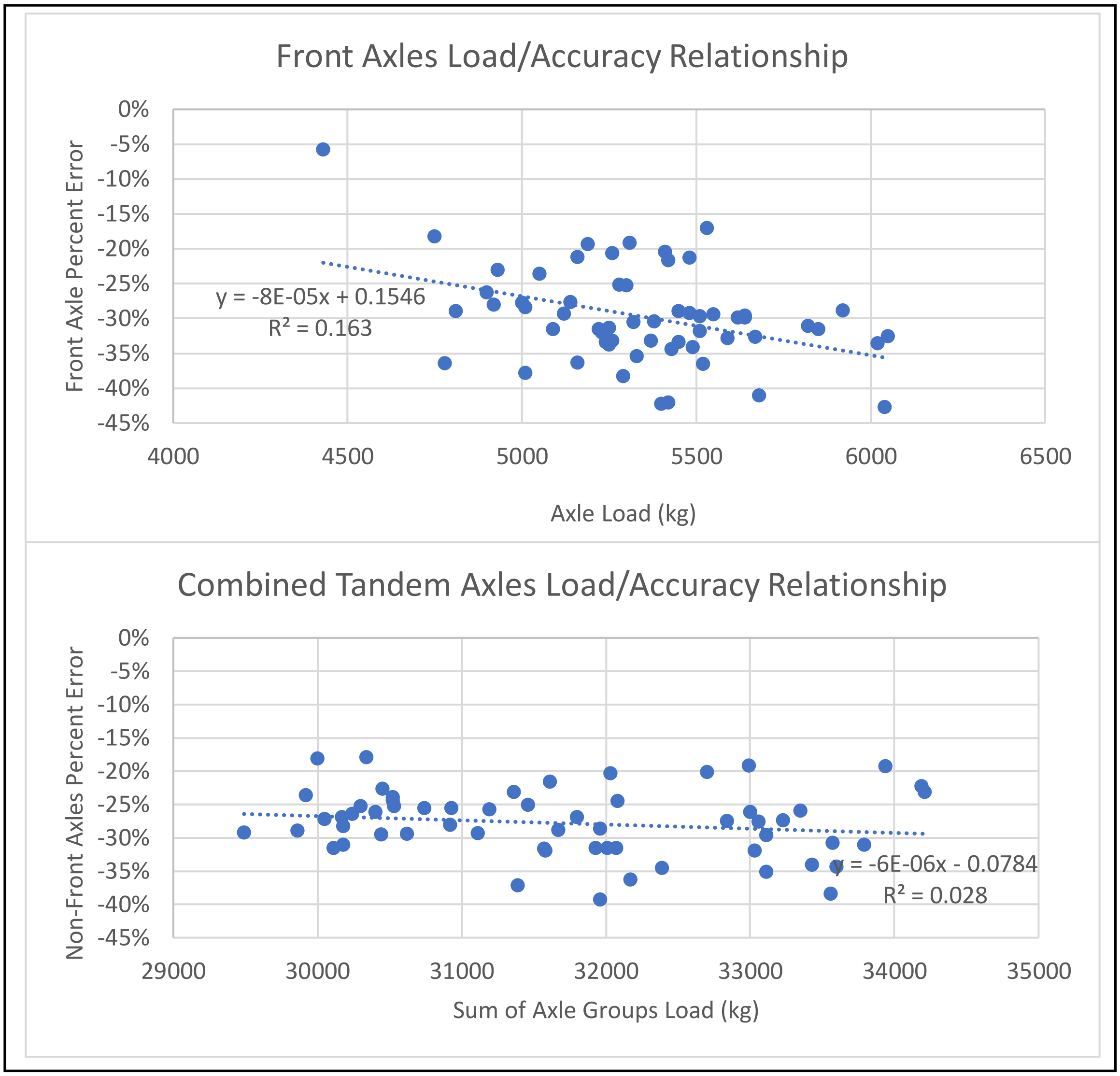


Figure 37 - Relationship of axle load measurements to percent error in portable WIM post-processing calibration

The scatterplots in Figure 37 show that that there is an approximately equal range of percent error values at all axle load measurement values for both the front axles and the combined tandem axle groups. This indicates that the generated calibration factors will be approximately the same for any vehicle weight. The low R^2 values of the linear trendlines (<0.2 for both graphs) support this finding.

To ensure that no inaccuracy in the calibration was introduced by the passage of time between the axle load measurements, the measurements were grouped by day of installation, and each group was used to generate the calibration factors. Table 27 compares these calibration factors.

Table 27 - Post-Processing calibration factors for each day of portable WIM installation

Date	Number of Qualifying Measurements	$F_{calibration}$	$F_{dynamic}$
Oct. 17	8	1.398	1.043
Oct. 18	13	1.402	1.048
Oct. 21	10	1.396	1.060
Oct. 22	10	1.382	1.008
Oct. 23	18	1.367	1.006

To determine if the generated calibration factors from this process were affected by the passing of time, the R^2 values for the correlation of both $F_{calibration}$ and $F_{dynamic}$ to the date were generated. $F_{calibration}$ had an R^2 value of 0.71, indicating a strong relationship, while $F_{dynamic}$ had an R^2 value of 0.41, indicating a moderate relationship. As these relationships are non-trivial, the passage of time can be said to have a statistically significant impact on the calibration factors, despite the fact that the values of $F_{calibration}$ and $F_{dynamic}$ do not change by a significant amount (<0.05 for both). Due to this, the calibration factors generated from vehicles recorded on October 17 at the beginning of the installation were used in the post-processing calibration procedure as performed on the third installation portable WIM dataset.

APPENDIX D: VEHICLE CONFIGURATION SELECTION ALGORITHMS

Four vehicle configurations of interest were selected using the following algorithms. The algorithms selected vehicles on the basis of axle spacings. Each axle spacing has required minimum and maximum values, with several exceptions that had only a minimum value or were unrestricted. A vehicle was only considered to be of the listed configuration if it had the correct number of axles and all of its axle spacings fell within the ranges specified by the algorithms. The axle spacings used were based on the requirements of the Manitoba Vehicle Weights and Dimensions Guide (Manitoba Infrastructure Motor Carrier Division, 2017). Table 28 displays the spacing ranges required by the algorithms.

Table 28 - Vehicle configuration selection algorithm axle spacing ranges

Configuration		Axle Spacing (cm)							
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
3-S2	Max	615	185	1240	185				
	Min	300	100	500	100				
3-S3	Max	615	185	1240	185	185			
	Min	300	100	550	120	120			
3-S3-S2	Max	615	185	-	165	165	-	185	
	Min	300	100	550	120	120	550	100	
3-S2-4	Max	615	185	-	185	-	185	-	185
	Min	300	100	500	100	-	100	500	100

APPENDIX E: REGRESSION EQUATIONS

This appendix displays the regression equations calculated to analyze the relationship between vehicle speed and the various non-front axle loads and axle load errors for several selected vehicle configurations in section 4.3.2. Table 29 displays the regression equations with the R^2 values, sample sizes, and significances.

Table 29 - Regression equations of speed and axle group loads for selected vehicle configurations

Configuration	Axle Group	Regression Equation	R^2	n	P-value of F statistic
3-S2	Tandem (Drive)	$Y = -0.98849*S^2 + 164.31*S + 22357.8$	0.00043	6561	<0.001
	Tandem (Trailer)	$Y = -0.85585*S^2 + 148.38*S + 1937.1$	0.0018	6561	0.0026
3-S3	Tandem (Drive)	$Y = -0.74844*S^2 + 124.15*S + 3854.4$	0.015	1592	0.33
	Tridem (Trailer)	$Y = -0.52084*S^2 + 79.083*S + 8042.1$	0.00098	1592	0.46
3-S3-S2	Tandem (Drive)	$Y = -2.7676*S^2 + 417.89*S – 5583.6$	0.064	416	<0.001
	Tridem (Trailer 1)	$Y = -4.2925*S^2 + 665.53*S – 12573.4$	0.047	416	<0.001
	Tandem (Trailer 2)	$Y = -3.3477*S^2 + 517.29*S – 11319.8$	0.051	416	<0.001
3-S2-4	Tandem (Drive)	$Y = -1.3903*S^2 + 248.45*S – 688.27$	0.012	1631	<0.001
	Tandem (Trailer 1)	$Y = -1.0335*S^2 + 197.65*S + 161.96$	0.0091	1631	<0.001
	Tandem (Front of Trailer 2)	$Y = 0.40864*S^2 + 72.824*S + 4673.3$	0.00079	1631	0.53
	Tandem (Rear of Trailer 2)	$Y = -0.58839*S^2 + 126.04*S + 1597.4$	0.017	1631	<0.001

APPENDIX F: LOAD ERROR PERCENTILES IN TIME

During the per-vehicle analysis (section 4.4), the daily ASTM compliance over the course of Installations 1 and 2 is calculated and graphed. To provide a more detailed look at the trends of the data during these installations, key percentiles (5%, 25%, 50%, 75%, and 95%) of the load measurement percent errors were calculated for each day of installations 1 and 2 for each of the three processing methods. Time was used as the measure of the calibration drift due to the findings of the calibration drift analysis (section 4.2). These percentiles were calculated for GVW, tandem axle loads, and single axle loads. Tridem axle loads were not included due to their much smaller sample size. Figure 38, Figure 39, and Figure 40 display scatterplots of the GVW, tandem axle group loads, and individual axle loads percentile values, respectively. For each axle group type/GVW and processing method combination, both installations are displayed on a single graph by date. Additionally, on some of the autocalibrated graphs, the 95% percentile load error is not displayed due to the scale of the graph. This was left in place to allow a consistent y-axis that effectively displayed most of the data points. A graph of this type of an ideal portable WIM installation would have a median at 0% error, with the other percentiles tightly clustered about 0%. Figure 38, Figure 39, and Figure 40 display the scatterplots of key eror percentiles.

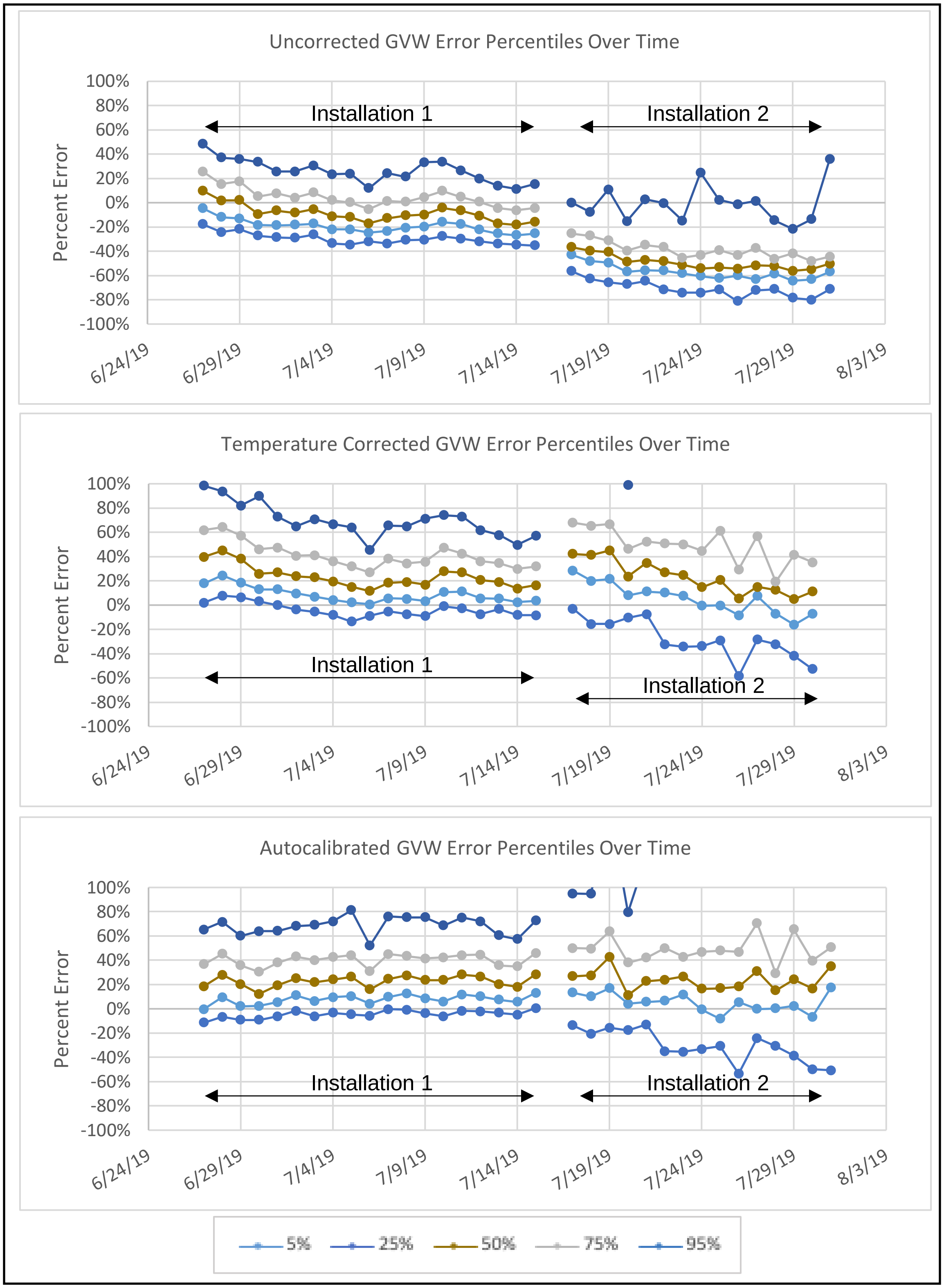


Figure 38 - Key percentiles of portable WIM GVW percent errors over installations 1 and 2

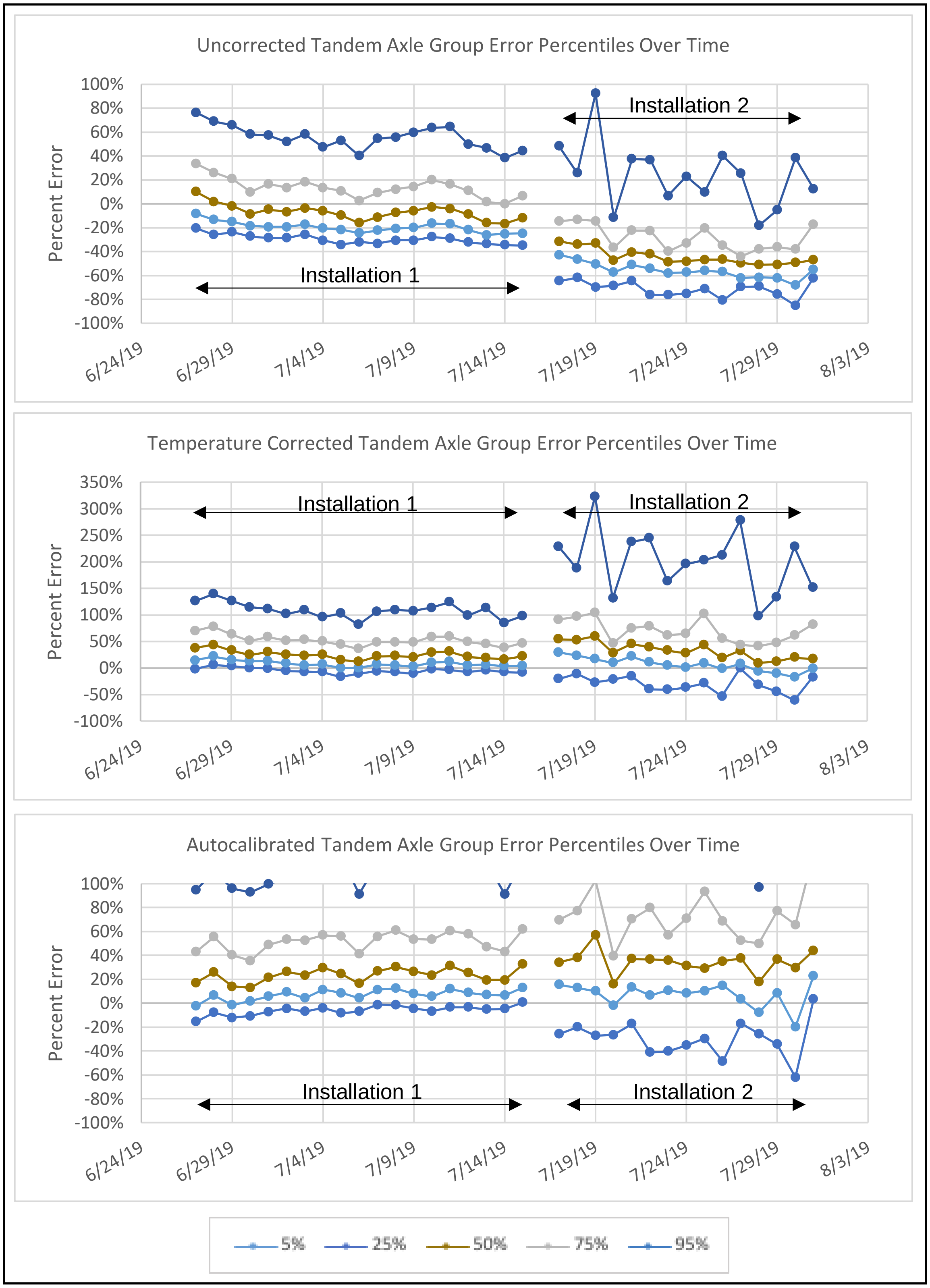


Figure 39 - Key percentiles of portable WIM tandem axle load percent errors over installations 1 and 2

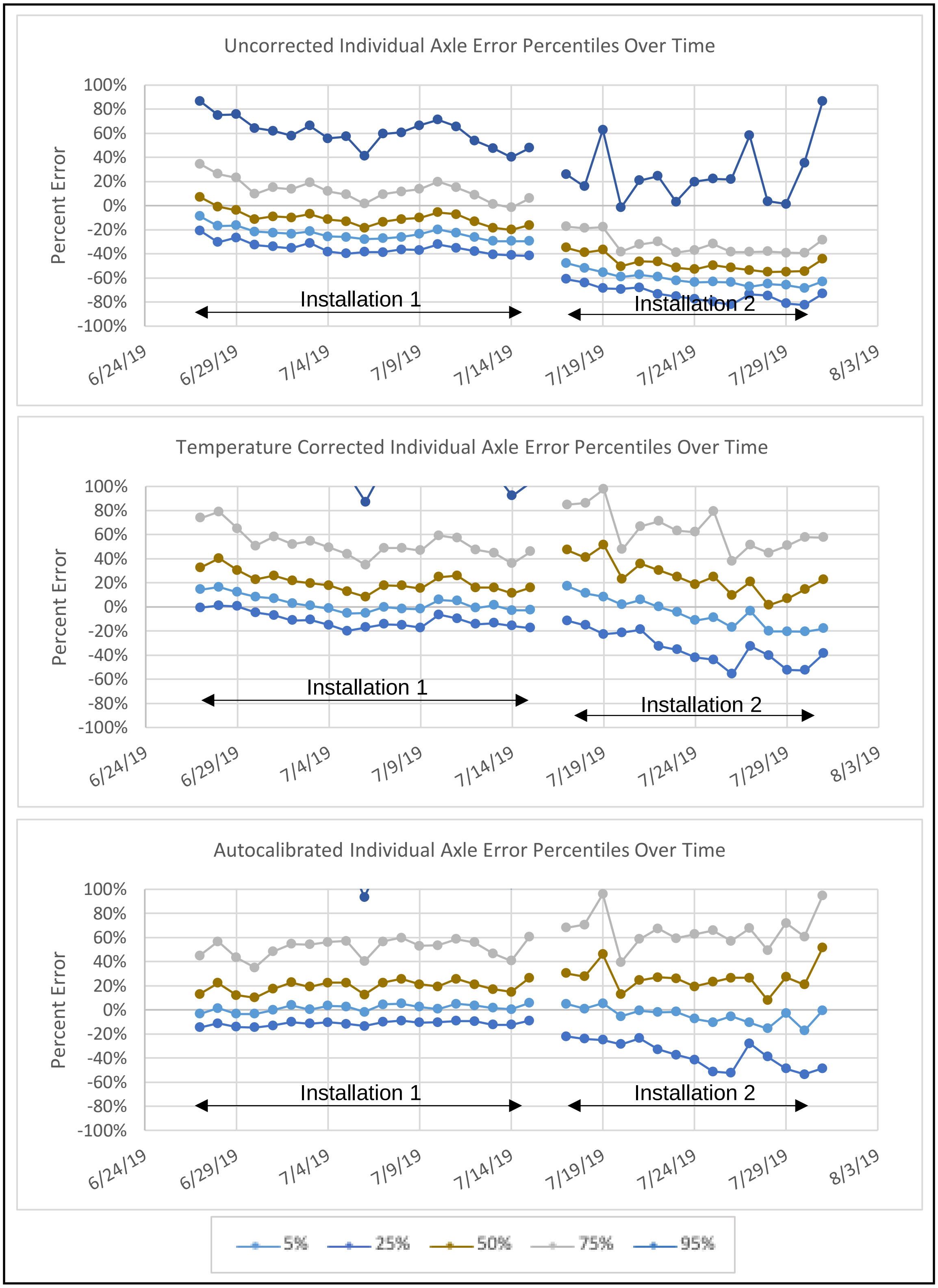


Figure 40 - Key percentiles of portable WIM individual axle load percent errors over installations 1 and 2

These figures reveal several key facts about the way the portable WIM load percent errors changed over the course of installations 1 and 2, many of which corroborate the same findings from the ASTM compliance data:

1. For uncorrected data, the data is best calibrated at the beginning of installation 1 (median nearest to 0% error with other percentiles tightly clustered about 0%).
2. For uncorrected data, the dataset as a whole (all 5 measured percentiles) trends steadily downwards during the first half of each installation and does not trend significantly during the second half of each installation. This trend was explored in greater detail for installation 1 in section 4.2.1.
3. For all 3 data processing methods, the initial error percentiles of installation 2 are substantially different than the final error percentiles of installation 1.
4. The day-to-day consistency is much stronger during installation 1 than installation 2; this is true mainly for the 95th percentile of all datasets.
5. The autocalibrated datasets do not show any sign of trending either up or down during either installation, as would be expected due to the constant re-calibration.
6. GVW measurements generally have a lower percent error range than tandem axle groups or individual axles, i.e. the 5th and 95th percentiles are closer together for GVW data.

APPENDIX G: DATA CORRECTION PROCEDURES

This appendix provides step-by step descriptions of the three data correction procedures that were performed on the portable WIM data: the initial temperature correction procedure, the autocalibration procedure, and the loading separated temperature correction procedure.

Initial Temperature Correction Procedure

1. Generate a quadratic regression equation by regressing the front axle loads of the four selected vehicle configurations (3-S2, 3-S3, 3-S3-S2, 3-S2-4) on temperature.
2. Use the regression equation to calculate the expected selected configuration front axle loads at each temperature.
3. Calculate the average selected configuration front axle load from the most recent dataset collected at the nearest source of static weigh scale data. For this research, the static weigh scale data collected during Installation 3 is the dataset used, and the average front axle load is 5170 kg. This is the target front axle load.
4. Calculate temperature correction factors by dividing the target front axle load by the expected selected configuration at each temperature value.
5. Apply the temperature correction factors by multiplying all axle loads in a record by that record’s corresponding temperature correction factor.

Autocalibration Procedure

1. Calculate the average 3-S2 front axle load from the most recent dataset collected at the nearest source of static weigh scale data. For this research, the static weigh scale data collected during Installation 3 is the dataset used, and the average front axle load is 5180 kg. This is the target front axle load.
2. Separate out the dataset to be autocalibrated into bins each consisting of groups of 50 3-S2 records and the records between them.
3. Calculate the average 3-S2 front axle load for each bin.
4. Calculate autocalibration factors for each bin by dividing the target front axle load by the average 3-S2 front axle load for each bin.
5. Apply the autocalibration factors by multiplying all records in each bin by the autocalibration factor.

Loading Separated Temperature Correction Procedure

1. Apply a two-component gaussian mixture model to 3-S2 GVWs from the most recent dataset collected at the nearest source of static weigh scale data; this model will separate the data into two distributions representing unloaded and fully loaded vehicles. For this research, the static weigh scale data collected during Installation 3 is the dataset used, and the mean GVWs for unloaded and fully loaded vehicles are 17700 kg and 32860 kg, respectively. These means are the target GVWs.
2. Separate out all 3-S2 vehicles from the dataset to be temperature corrected and define a cutoff point between vehicles considered to be

unloaded and fully loaded. The cutoff point can be defined by visually inspecting a histogram of the 3-S2 GVW values and placing the cutoff point at the lowest point between the two peaks of a bimodal distribution.

3. Generate two quadratic regression equations by regressing the unloaded and loaded 3-S2 GVWs from the dataset to be temperature corrected on temperature.
4. Use the regression equations to calculate the expected 3-S2 GVWs at each temperature.
5. Calculate temperature correction factors by dividing the target unloaded and fully loaded GVWs by the expected unloaded and fully loaded 3-S2 GVWs at each temperature value.
6. Separate out the four configurations of interest from the dataset to be temperature corrected and define cutoff points between vehicles considered to be unloaded and fully loaded for each configuration separately.
6. Apply the temperature correction factors for unloaded and loaded vehicles to the unloaded and loaded vehicles of each configuration separately by multiplying all axle loads in a record by that record’s corresponding temperature correction factor.
7. Combine the temperature corrected records of the four configurations back into a single dataset.