Weigh-in-Motion Auto-Calibration Using Automatic Vehicle Identification

Fu Ren Zhang Zhang Durandal
University of Arkansas, Fayetteville

Follow this and additional works at: https://scholarworks.uark.edu/etd

Part of the Civil Engineering Commons, and the Transportation Engineering Commons

Citation

This Thesis is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact ccmiddle@uark.edu.
Weigh-in-Motion Auto-Calibration Using Automatic Vehicle Identification

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering

by

Fu Zhang Durandal
University of Arkansas
Bachelor of Science in Civil Engineering, 2017

December 2019
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

Sarah Hernandez, Ph.D.
The thesis Director

Clinton Wood, Ph.D.
Committee Member

Gary Prinz, Ph.D.
Committee Member
Abstract

Weigh-in-Motion (WIM) sensors are installed on mainline lanes at highway locations to record vehicle weights, axle spacing, vehicle class, travel speed, vehicle length, and traffic volume. These data elements support effective transportation planning, infrastructure design, and policy development. Therefore, it is important that WIM sensors supply accurate data. After initial installation and calibration, WIM systems may experience measurement drifts in weight and axle detection. Recalibration takes two general forms: (a) On-site calibration involving running trucks of known weight over WIM scales and (b) Auto-calibration methods involving comparisons to assumed reference weights. Auto-calibration can be more cost and time effective than on-site calibration. This paper leverages the increasing prevalence of truck tracking technologies like Global Positioning Systems (GPS) to improve auto-calibration methods and was divided into three aims: (i) data collection, (ii) data processing and (iii) model development. Truck GPS data from a national provider, WIM recorded truck weights, and static weights collected at weight enforcement station were gathered at several highway locations in Arkansas. A “matching” algorithm was developed to automatically match each GPS record to a WIM record based on timestamp and vehicle configuration. Algorithm performance was assessed via manual video verification of matches. Approximately, 75% of WIM and truck GPS records were correctly paired. Lastly, an auto-calibration model was developed to estimate lane and site specific calibration factors. The algorithm estimates hourly calibration factors by comparing the front axle weight of the same truck as it passes multiple WIM sites. Algorithm performance was measured by comparing estimated front axle and gross vehicle weights to known weights of the same truck measured at a static enforcement scale. The algorithm achieved Median Absolute Percent Error (MdAPE) of 11-23% for front axle weight and 15-45% for gross vehicle weight. These results can be improved by
increasing the number of trucks that are able to be tracked across WIM sites with Automatic Vehicle Identification.
Acknowledgements

This research and thesis was made possible with the help and guidance of Dr. Sarah Hernandez, Ph.D., Johnson Baker, Taslima Akter, ARDOT, Drivewyze, and Dr. Hernandez lab research team.
Contents
Chapter 1. Introduction .............................................................................................................1
Chapter 2. Background .............................................................................................................5
  2.1 WIM Sensor Overview .................................................................................................5
    2.1.1 Sensor Types ........................................................................................................5
    2.1.1 Sensor Accuracy Tolerances ...............................................................................11
  2.2. Calibration Methods for WIM Sensors .....................................................................11
    2.2.1. On-site calibration methods .............................................................................12
    2.2.2 Off-site calibration methods ..............................................................................12
    2.2.3. Auto-calibration methods ...............................................................................13
  2.3. Automatic Vehicle Identification for Auto-Calibration ..........................................14
Chapter 3. Data .........................................................................................................................17
  3.1 Data Descriptions ........................................................................................................17
    3.1.1 WIM PVR .........................................................................................................17
    3.1.2 AVI Records ......................................................................................................17
  3.2 Data Collection ............................................................................................................18
    3.2.1 Site Selection .....................................................................................................18
    3.2.2 Field Data Collection ......................................................................................19
  3.3 Data Pre-Processing ....................................................................................................22
    3.3.1 Matching Trucks Crossing Static Scales and WIM Sites ..................................22
    3.3.2 Matching WIM and GPS Records ....................................................................23
Chapter 4. Methods .................................................................................................................26
  4.1 Truck Matching Algorithm .........................................................................................27
    4.1.1 Offset Calculation .............................................................................................27
    4.1.2 Match Filtering ..................................................................................................28
    4.1.3 Data Pairing ......................................................................................................29
  4.2 Auto-Calibration Algorithm .......................................................................................30
  4.3 Auto-calibration Algorithm Implementation Example .............................................34
    4.3.1 Truck Matching Example ................................................................................34
    4.3.2 Auto-Calibration Example ..............................................................................37
Chapter 5. Results ...................................................................................................................41
  5.1 Summary of Data Collection Efforts ..........................................................................41
5.2 Truck Matching Algorithm Performance ........................................................42
5.2 Auto-Calibration Algorithm Performance .......................................................45
Chapter 6. Discussion and Future Work ..................................................................51
Chapter 7. Conclusion .............................................................................................54
References ................................................................................................................56
Appendix ..................................................................................................................58
List of Tables

Table 1. Data produced by WIM Systems (Ref: ASTM E 1318 E – 09, 2017). ..........................5
Table 2. Data collection summary for different WIM types. .........................................................6
Table 3. Range of costs for commonly used scale types on WIM systems. .........................6
Table 4. Deviation tolerances by WIM sensor type (FHWA, 2018 Part 3 table 1) ..................11
Table 5. Video to WIM time offset calculation example. ...............................................................23
Table 6. MAPE Comparison for FAW and GVW by site and Auto-Calibration Method ........50
Table 8. Arkansas weight limits for a given axle type by USDOT. ........................................59
Table 9. Highway Pavement design table to calculate ESALs ....................................................60
List of Figures

Figure 1. Normalized volume distribution by vehicle class from 143 sites (Tam & Quintus, 2003) ................................................................. 2
Figure 2. Normalized volume distribution for truck classes from 43 sites (Tam & Quintus, 2003) ................................................................. 3
Figure 3. Gross vehicle weight distribution for five axle tractor trailers (Abbas et al., 2014) ........ 3
Figure 4. WIM sensor type II embedded in the pavement at I-30 Arkansas (from Google Maps). 7
Figure 5. Piezoelectric WIM sensor configuration (from McCall & Vodrazka, 1997) ............... 8
Figure 6. Bending Plate WIM sensor (from McCall & Vodrazka, 1997) ................................. 9
Figure 7. Eastbound Alma static weight station in Arkansas (from google maps) ................. 10
Figure 8. Drivewayze Truck Traffic Patterns ........................................................................ 19
Figure 9. Example of a weight receipt (A) and weight recording configuration (B) ............... 21
Figure 10. Trucks being weighed at a weigh station ............................................................... 21
Figure 11. Truck images from traffic recordings at WIM sites ................................................ 21
Figure 12. Example truck images from video at Lamar WIM site .......................................... 24
Figure 13. Example of truck re-identified at Lonoke WIM site ............................................ 24
Figure 14. Example of notes used to match trucks across WIM sites .................................. 25
Figure 15. Example of finalized matches between video, GPS, and WIM records .................. 25
Figure 16. Example of Cluster Analysis for AVI Auto-Calibration ...................................... 31
Figure 17. Error Rate by time of day at the Lonoke WIM site during the March 2018 data collection .......................................................................... 45
Figure 18. Shared AVI truck from Lamar to Lonoke during data collection period ............... 45
Figure 19. Comparison of static and WIM weights for Lamar and Lonoke. ............................ 47
Figure 20. Comparison of MAPE by Auto-Calibration Method ............................................. 49
Figure 21. Lamar and Lonoke per vehicle calibration factors by time of day ....................... 53
Figure 22. FHWA vehicle classes 1 thru 13 (FHWA, 2019). .................................................... 58
Figure 23. Lamar PVR RAW FAWs where 1 denotes weights recorded at lane 1, the inner lane and 2 denotes weights recorded at lane 2, the outer lane. ................................. 63
Figure 24. Lonoke PVR RAW FAWs where 1 denotes weights recorded at lane 1, the inner lane and 2 denotes weights recorded at lane 2, the outer lane. ................................. 63
Figure 25. Glen Rose PVR ARDOT Calibrated FAWs where 4 denotes weights recorded at lane 4, the inner lane, and 3 denotes weights recorded at lane 3, the outer lane. ................. 64
Figure 26. Arkadelphia PVR ARDOT Calibrated FAWs where 4 denotes weights recorded at lane 4, the inner lane, and 3 denotes weights recorded at lane 3, the outer lane. ................. 64
Figure 27. Texarkana PVR ARDOT Calibrated FAWs where 4 denotes weights recorded at lane 4, the inner lane, and 3 denotes weights recorded at lane 3, the outer lane.
Chapter 1. Introduction

Effective transportation decisions, such as congestion mitigation and infrastructure planning, can be made based on available data collected by Weigh-in-Motion (WIM) sensors. In the 1980’s, WIM sensors became widely used in the U.S. to gather traffic data (Bunnell et. Al, 2017). WIM sensors record traffic volume, vehicle weights, and axle configurations from which classification into vehicle classes (VCs) are made. The commonly used FHWA scheme F categorizes vehicles into 10 classes based on axle configuration. Of these, class 2 (passenger vehicles), class 5 (single unit trucks), and class 9 (five axle tractor trailers) are the most common. These data are needed to track freight movements, regulate truck weights and sizes, and plan for transportation infrastructure. These data are used by metropolitan planning organizations (MPOs), state departments of transportation (DOTs), and federal DOTs to make effective transportation investment and maintenance decisions.

An important application of WIM data is for pavement design. WIM systems convert dynamic weights of moving vehicles to equivalent single axle loads (ESALs). An ESAL represents the equivalent of any axle configuration and weight to an 18,000-pound single axle. An ESAL is calculated by the following equation:

\[
ESAL_i = f_d \times G_m \times AADT_i \times 365 \times N_i \times F_{et}
\]

Eq. 1 (AASHTO, 1993)

where:

- \(ESAL_i\) = equivalent accumulated 18,000 lb single axle load for each category \(i\)
- \(i\) = vehicle class axle category
- \(f_d\) = design lane factor
- \(G_m\) = Growth factor
- \(AADT\) = Annual average Daily Traffic for each category \(i\)
- \(N_i\) = number of axles for each category \(i\)
- \(F_{et}\) = load equivalency factor for category \(i\)
More recently however DOTs such as TXDOT are transitioning from the use of ESAL measurements to the use of axle load spectra. Axle load spectra represent axle weight distributions by vehicle class (VC) and axle configuration (e.g., single, tandem, tridem) are preferred to be collected specifically for each site. Heavier VCs cause a larger ESAL and more damage to pavement structures. Passenger vehicles (Class 2), two axle trucks (Class 5), and five axle tractor trailer (Class 9) are the most common vehicle configurations (Figures 1 and 2). For five axle tractor trailers, the gross vehicle weight (GVW) distribution tends to have a bi-modal shape (Figure 3) and similar characteristics have been shown for single and tandem axle load spectra. With the transition to site-specific axle load spectra which have more detail than traditional ESAL estimations, there is a need to ensure WIM data are accurate and reliable. Thus, calibration of WIM systems is increasingly important.

![Normalized volume distribution by vehicle class from 143 sites](image)

**Figure 1.** Normalized volume distribution by vehicle class from 143 sites (Tam & Quintus, 2003)
WIM sites provide rich data for design and planning, however, without proper maintenance and calibration programs, WIM sensors can drift producing inaccurate data especially for weight. Intermittent calibration is performed on WIM systems to ensure the collection of accurate data through on-site calibration and auto-calibration. Unfortunately, on-site calibration is performed infrequently due to higher cost.

On-site calibration is a common practice for WIM stations in many state DOTs. This method requires running test trucks of known weights and vehicle parameters over WIM stations to calibrate axle weight, and gross vehicle weights (GVWs) and inter-axle length. This practice
requires running at least two types of trucks over a WIM scale multiple times per the American Society of Testing Materials (ASTM Standard E1318-09, 2017). Many states vary the test truck configurations and the number of test runs per station due to state needs and convenience as traffic volumes have different characteristics in different regions. These On-site calibration operations are time consuming and costly. As a result, many states do not perform on-site calibration often nor by itself. Instead, they rely on auto-calibration to correct for systematic measurement errors, and sporadic on-site calibration for major adjustments.

Auto-calibration adjusts vehicle parameters in real-time. For instance, it may calibrate inter axle spacing and axle weights using reference values that may include variations based on temperature, seasonality and frequency of a type of truck. A five-axle tractor-trailer steering axle (VC 9) weight of 10kips is commonly used as a reference measurement as this is one of the most common truck configurations. The frequency of auto-calibration is determined by the sample size of trucks crossing a WIM site or over a pre-specified time span. These sequences are selected by the volumes experienced at each site. For instance, auto-calibration can be set to be performed every 48 hours or once a sample size of 250 vehicles have crossed the WIM sensor.

This project developed a WIM auto-calibration method based on Automatic Vehicle Identification (AVI). AVI systems record unique and identifiable characteristics of a truck and can be used to track trucks across multiple locations. In the context of auto-calibration, AVI provides a way to compare the weight of the same truck across multiple WIM sites and then use the multiple weights to determine site-specific calibration factors. This thesis presents a methodology to use AVI data for auto-calibration and evaluates the algorithm against static weights collected at enforcement weigh stations.
Chapter 2. Background

2.1 WIM Sensor Overview

WIM sensors are continuous collection devices embedded in pavement structures. Unlike other common traffic data collection devices, such as loop detectors or cameras that record traffic volume, WIM have the ability to capture vehicle weights. The data items captured by WIM are listed in Table 1. Being able to capture these data items allow the WIM sensor to perform automatic vehicle classification (AVC). In addition, WIM systems can estimate ESALs (FHWA, 2017).

Table 1. Data produced by WIM Systems (Ref: ASTM E 1318 E – 09, 2017).

<table>
<thead>
<tr>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Wheel Load</td>
</tr>
<tr>
<td>2. Axle Load</td>
</tr>
<tr>
<td>3. Axle-Group-Load</td>
</tr>
<tr>
<td>4. Gross-Vehicle Weigh</td>
</tr>
<tr>
<td>5. Speed</td>
</tr>
<tr>
<td>6. Center-to-Center Spacing Between Axles</td>
</tr>
<tr>
<td>7. Vehicle Class</td>
</tr>
<tr>
<td>8. Site Identification Code</td>
</tr>
<tr>
<td>9. Lane and Direction of Travel</td>
</tr>
<tr>
<td>10. Time Stamp</td>
</tr>
<tr>
<td>11. Sequential Vehicle Record Number</td>
</tr>
<tr>
<td>12. Wheelbase (front to back axle)</td>
</tr>
<tr>
<td>13. Equivalent Single-Axle Loads (ESALs)</td>
</tr>
<tr>
<td>14. Violation code</td>
</tr>
</tbody>
</table>

2.1.1 Sensor Types

There are four types of WIM systems: Type I, II, III, and IV. The WIM sensor type has an effect on data recording capabilities (ASTM E 1318 E – 09, 2017). WIM sensor types I, II and III can be situated on highways as they can record data for vehicles moving at speeds between 10 and 80 mph (15 to 80 mph for type II systems). Type IV WIM systems are strictly for weight-enforcement stations to detect any weight violations and wheel loads using tire force sensors supporting the
contact area of all tires. The data items recorded by specific WIM type are listed in Table 2 (ASTM E 1318 E – 09, 2017).

Each WIM system type contains a different type of sensor to capture and weigh vehicles including: piezoelectric (e.g., polymer and quartz), bending plate (e.g., strain gauge and bending plate), and load cell sensors. Scales with longer life and higher reliability are more costly on average (Table 3). Arkansas’ WIM sites use polymer piezoelectric sensors, e.g., Type II (Figure 4).

Table 2. Data collection summary for different WIM types.

<table>
<thead>
<tr>
<th>Data</th>
<th>WIM Type I</th>
<th>WIM Type II</th>
<th>WIM Type III</th>
<th>WIM Type IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>15. Wheel Load</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>16. Axle Load</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>17. Axle-Group-Load</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>18. Gross-Vehicle Weigh</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>19. Speed</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>20. Center-to-Center Spacing Between Axles</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>21. Vehicle Class</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22. Site Identification Code</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>23. Lane and Direction of Travel</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>24. Time Stamp</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>25. Sequential Vehicle Record Number</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>26. Wheelbase (front to back axle)</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27. Equivalent Single-Axle Loads (ESALs)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28. Violation code</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 3. Range of costs for commonly used scale types on WIM systems.

<table>
<thead>
<tr>
<th>Sensor Type</th>
<th>Life (years)</th>
<th>Sensor Installation (costs for one lane installation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Polymer Piezo</td>
<td>2 – 3</td>
<td>4,000</td>
</tr>
<tr>
<td>Quartz Piezo</td>
<td>3 – 5</td>
<td>16,000</td>
</tr>
<tr>
<td>Strain Gauge</td>
<td>3 – 5</td>
<td>16,000</td>
</tr>
<tr>
<td>Bending Plate</td>
<td>6 – 8</td>
<td>18,000</td>
</tr>
<tr>
<td>Load Cell</td>
<td>10 -12</td>
<td>44,000</td>
</tr>
</tbody>
</table>
To collect weight and vehicle movement, piezoelectric sensors detect the change in voltage exerted on the pavement surface by vehicle tires. These electrical charges are recorded and transformed into dynamic loads which are then used to estimate a static weight. The physical configuration of a piezoelectric sensor consists of at least one piezo sensor and one inductive loop. A typical configuration though has two inductive loops and two piezo sensors as shown in Figure 5. The upstream inductive loop identifies incoming traffic and the downstream loop serves to measure time-based travel speed and axle spacing, and the piezo sensors in between record the weight. Axle spacing and speed can be measured using the sensors by measuring the time the front axle of the vehicle crosses the loop.
Bending plate sensors use strain gauges to record strain caused by vehicles traveling over a bending plate. Dynamic loads are calculated from the strain recordings. Dynamic loads are then converted to static loads by the WIM host computer taking into account vehicle speed and pavement/suspension dynamics. Bending plate sensors can be permanent or portable. Portable bending plates cannot record vehicles traveling at high speeds. The physical configuration of a permanent bending plate sensor consists of at least one scale and two inductive loops, where the scale or scales are placed perpendicular to the traffic flow as shown below in Figure 2.
If two scales are used, then each scale is placed at the wheel path to weigh vehicle wheels individually. These scales can be placed side by side or staggered at a maximum of 16 feet apart. Inductive loops are situated both upstream and downstream from the scale. The inductive loop configuration above is able to detect traveling vehicles and measure axle spacing. An axle sensor may also be included for this type of configuration. Axle spacing is measured using one of the following three methods: (1) measuring axle distance and time from weigh pad to an inductive loop (2) gauging the axle distance using the weigh pad to the axle sensor, (3) measure axle distance from weigh pad to weigh pad if they are staggered.
Load cell sensors employ a single load cell and two scales to record axle weight and the weight of the left and right wheels individually. Its physical configuration is similar to a bending plate sensor. A WIM system using a load cell consists of at least one load cell, one inductive loop and one axle sensor. The inductive loops are place in similar fashion to the previously shown configurations.

Static weight scales, in contrast to WIM sensors, do not record continuous nor dynamic weight data from a traffic stream. Static weight scales require trucks to drive into weigh stations in order to record their weights. The weights recorded are typically not stored for data collection purposes but for weight enforcement. Static axle weights and gross vehicle weights (GVW) of trucks recorded at these stations are “true weights”, which are the calibration target for WIM sensors. Unfortunately, static weight scales measure only a small percentage of the total trucks operating along a roadway. Figure 7 portrays how weight enforcement static scales are configured alongside highways. The weigh station in image is the Alma eastbound weigh station alongside Arkansas I-40.

Figure 7. Eastbound Alma static weight station in Arkansas (from google maps)
2.1.1 Sensor Accuracy Tolerances

The level of accuracy for different weight recording functions for each WIM type is different. The most accurate WIM types for weight recordings are type III and I and are also capable of recording wheel loads, whereas Type II sensors cannot. This is due to the fact that Type III and I systems are typically equipped with bending plate and load cell scales which have more accuracy and are able to capture wheel loads. Type II WIM are typically equipped with piezoelectric sensors which are not capable of recording wheel loads and have higher deviation. Table 4 presents the typical deviation tolerances for each WIM type for 95% compliance, deviations that are representative for each WIM type.

<table>
<thead>
<tr>
<th>Function</th>
<th>Tolerance for 95% Compliance</th>
<th>Type I</th>
<th>Type II</th>
<th>Type III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheel Load</td>
<td></td>
<td>±25%</td>
<td>NA</td>
<td>±20%</td>
</tr>
<tr>
<td>Axle Load</td>
<td></td>
<td>±20%</td>
<td>±30%</td>
<td>±15%</td>
</tr>
<tr>
<td>Axle-Group Load</td>
<td></td>
<td>±15%</td>
<td>±20%</td>
<td>±10%</td>
</tr>
<tr>
<td>Gross Vehicle Weight</td>
<td></td>
<td>±10%</td>
<td>±15%</td>
<td>±6%</td>
</tr>
<tr>
<td>Speed</td>
<td></td>
<td>±1 mph</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axle-Spacing</td>
<td></td>
<td>±0.5 ft</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.2. Calibration Methods for WIM Sensors

WIM sensors measure dynamic weights and output static vehicle weights or ESALs that are an estimate of the true weight. Inaccuracies are caused by the conversion of dynamic to static weights, vehicle characteristics (e.g. vehicle classification, length, weight, and speed), and site and environmental conditions (e.g. pavement condition, site levelness, and weather). Therefore, WIM scales always tend to have some degree of inaccuracy.

The goal of calibration is to minimize the discrepancy between static weights or reference parameters and WIM vehicle parameter measurements. Calibration factors are applied to WIM sensors to reduce discrepancies between the measurements and true vehicle parameters.
Calibration factors are typically calculated by dividing the WIM recorded vehicle parameter (like GVW) by a reference value such as a known static weight, predefined axle spacing, or other known vehicle parameter. There are three types of calibration methods commonly used: (1) on-site calibration (2) off-site calibration, and (3) auto-calibration.

2.2.1. On-site calibration methods

On-site calibration requires using (1) test trucks or (2) trucks and vehicles from the traffic stream. For test truck runs, per ASTM E 1318 at least two different test trucks of known weight are used to travel over a site several times to compare the known weights of these trucks versus the recorded weights at a WIM scale. Many DOTs perform these tests varying the number of runs at a site and the type and quantity of test trucks used (Papagiannakis et al., 2008). For calibration using vehicles from the traffic stream, true weights or vehicle parameters are obtained at weight enforcement stations for vehicles that later will travel over WIM sensors in order to calibrate the measured WIM parameters against the true weights measured at the enforcement station. This method is convenient when WIM sensors and weight enforcement scales are close together and lie along the same route so that trucks can be tracked more easily.

2.2.2 Off-site calibration methods

Off-site calibration is performed by comparing WIM measure vehicle parameters to reference parameters. For example, comparisons may be made for five-axle tractor trailer (FHWA class 9) using front axle weights (FAW) where the reference value is 10 kips. Reference values may be site specific using vehicle parameter values that are most common for freight in the area for better results.
2.2.3. Auto-calibration methods

Auto-calibration algorithms are an alternative that decreases on-site labor making it less time consuming and less costly compared to running test trucks, for example. Auto-calibration compares vehicle parameter data collected at a WIM site to reference parameters and calculates corresponding calibration factors. Auto-calibration methods vary in the frequency at which calibration is performed and the ways in which reference parameters are used. For instance, auto-calibration can be performed each time a sample of a pre-specified size is collected (e.g., when a predetermined number of trucks have crossed a WIM site) or periodically (e.g., after a predetermined period of time like each 48 hours).

Ideally, the auto-calibration procedure should contain class-based, speed-based, and weather or seasonal-based calibration factors (Susor, 2010). Different vehicle classes pertain to different dimensions and weight ranges that influence measurement accuracy at a site. Vehicles traveling at different speeds exert different dynamic loads on the pavement. Reference parameters for auto-calibration may vary by state and even by site based on common vehicle parameters that are observed frequently in the state or site pertaining.

Pavement conditions and temperature influence WIM sensors accuracy. For example, FAW data from a WIM sensor at temperatures lower than 40 ° F when the pavement is harder experience less accuracy and more dispersion in FAW recordings. On the other hand, at higher temperature ranges as in the months of March-August, the pavement is more flexible and thus FAW measurements vary less. (Nassif et al., 2017, and Bunnell et al., 2017) (Nassif et al., 2017). Pavement smoothness, temperature, vehicle composition are examples of environmental conditions that are typically accounted for by calibration factors and can act independently and simultaneously. For example, in the case of a FHWA Class 9 truck that crosses a WIM sensor at
65 mph during the summer months the corresponding class, speed or seasonal calibration factors may be applied one at the time. Alternatively, a single calibration factor that takes into account all characteristics can be applied to calibrate WIM vehicle parameters (Susor, 2010).

The current ARDOT auto-calibration method performs weight adjustments every 50 vehicles. The weights of these trucks recorded at the WIM site are evaluated to a global reference weight in order to produce calibration factors. For ARDOT the global reference weight is 10 kips for the front axle weight. Then the recorded weights of the 50 vehicles are multiplied by the calibration factor to adjust the weights and the process repeats. Two primary limitations arise from the ARDOT method: 1) for low volume WIM sites, the accrual of 50 vehicles is often slower than the change in pavement temperature which leads to poor calibration results, and 2) data collected through this project shows that the reference front axle weight of 10 kips does not hold true at all sites. In fact, the reference front axle weight varies by site and by truck GVW.

An alternative to the ARDOT algorithm that corrects some of its limitations was developed by the Minnesota Department of Transportation (MNDOT). The MNDOT auto-calibration method estimates calibration factors every 250 vehicles or every 48 hours allowing for temporal flexibility. Also, instead of a single reference weight the MNDOT algorithm considers three different front axle weight reference values, pertaining to three GVW bins. This is an important difference that allows for the reference value to vary for unloaded and loaded vehicles.

### 2.3. Automatic Vehicle Identification for Auto-Calibration

Instead of using reference values to calculate calibration factors, it is possible to track trucks across WIM sites and compare their weights to generate calibration factors. Automatic Vehicle Identification (AVI) is a method of using computers to identify and track vehicles. Common
examples of AVI technologies include vision based systems like cameras, license plate readers, and Radio Frequency Identification (RFID). AVI is used for traffic enforcement in border and customs check points, electronic toll collection, intersection violations and for transportation analysis (Ozbay et al., 2007). AVI has the capability of recording path flow information and tracking a vehicle’s trip from origin to destination. In addition, AVI methods may have the capability to capture a larger sample than traditional surveys and traffic counts if desired.

The two most common methods using visual recordings in AVI are tag-based, and license plate-based recognition. Other forms of AVI are cellular phone based (Dixon and Rilett, 2002), GPS-based (Hyun K., Tok A., Ritchie S. G., 2017), transponder number (Nichols & Cetin, 2015), and inductive signature (Jeng & Chu, 2015 & Hyun K., Tok A., Ritchie S. G. 2017). Inductive loops and transponders create unique records using inductive signatures and transponder numbers, uniquely identifying trucks with their respective time stamps when they travel over these sensors. Then these trucks may be identified in other sensor sites of the network with their unique IDs or signatures. Hyun K., Tok A. and Ritchie S. G. (2017) and Jeng and Chu (2015) used ILDs along with WIM sensors to collect vehicle attributes of shared trucks for their studies.

The type of AVI technology applied depends on traffic sensor, infrastructure and equipment available and budget. For example, California uses advanced ILDs to track trucks across multiple scales while Oregon uses transponder identification numbers (Cetin and Nichols, 2015). The state of Arkansas does not have ILD available and automatic license plate recognition is not permitted. Thus, a suitable form of AVI for Arkansas is to use passive GPS tracking. There are many private data providers that collect and share GPS tracking data with public agencies for various applications including INRIX, HERE, and Drivewyze. Each of these companies collect data from
only a small percentage of the total traffic. However, this type of probe vehicle data is often enough to estimate performance measures like travel time and speed.

For this study, the GPS AVI data used was provided by a private company that operates as an app providing pre-clearance for trucks through weigh and enforcement stations. This AVI data was used to track trucks traveling Arkansas WIM sites. WIM, GPS AVI data, and camera recordings at designated sites were in this study were utilized to perform a truck matching task in order to create the inputs for the auto-calibration algorithm to calibrate multiple WIM sites with the data recorded at each station. Flow Chart 1 depicts the AVI WIM auto-calibration process. Input data include WIM recorded weights and GPS tracking data. Calibration factors (CFs) are calculated by comparing the weight of each truck as it crosses multiple WIM sites. The output is produced by applying the site and time dependent CFs to the WIM recorded measurements of trucks not seen in the GPS data set.

**Flow Chart 1. Auto-calibration process**
Chapter 3. Data

Data collection consisted of two field data collection efforts in which WIM, AVI, video, and static enforcement station data were gathered. Data collection and pre-processing are described in this section.

3.1 Data Descriptions

3.1.1 WIM PVR

WIM Per Vehicle Record (PVR) files include data of each vehicle detected by the WIM sensors. PVR data contains a record number, traveled lane, direction of travel, speed, VC, weights, and axle spacing. Note that the WIM sensor may not detect all vehicles and may also produce duplicate records (e.g., when vehicles change lanes over the sensors). WIM records for each vehicle may be represented in an array for vehicle parameters of interest having WIM site $D_s$ with vehicle parameter $W_i$:

$$D_s = [W_{s,1}, W_{s,2}, W_{s,3} ... W_{s,n}]$$

Where:

$D_s$ is WIM site s.

$W_{s,i}$ is a vehicle parameter data record at s for vehicle parameter I, e.g. axle weight, axle spacing, vehicle length, etc..

3.1.2 AVI Records

Each AVI truck record contained a unique ID that remains constant across traveled sites, time of day, and day of week. AVI records can be represented as an array for each unique truck $d_k$, with
the origin WIM site and time stamp at which the truck crossed the site, \( S_w \), and \( t_k,s \), and the next term \( S_{w'}, t_{k,s'} \) corresponding to site and timestamp of the next WIM site crossed:

\[
d_k = [ (S_w, t_{k,s}), (S_{w'}, t_{k,s'}) \ldots ]
\]

Where:
- \( d_k \) is the AVI record, where \( k \) denotes the unique id.
- \( S_w \) is the ID of the WIM site that the truck traveled over.
- \( t_{k,s} \) is the timestamp of truck \( k \) at site \( s \).

### 3.2 Data Collection

Data include WIM PVR, GPS truck tracking data, and video recordings at select sites. Still images taken at selected static enforcement sites and video footage of trucks crossing selected WIM sites were recorded and used for model development and validation.

#### 3.2.1 Site Selection

Static scale and WIM sites used for data collection were selected based on an analysis of the common truck paths, i.e. ‘shared traffic’ observed in Arkansas using historical GPS data. Traffic flows for the month of March 2018, for example (Figure 8), among WIM sites show a large portion of truck volumes from Texarkana to Malvern along I-30 and from Little Rock to West Memphis along I-40.

For the first round of data collection on March 2018 the selected static scale was the Alma Eastbound weigh station along I-40 and the selected WIM sites were Lamar and Lonoke along I-40 and Bald Knob along Highway 67. In this case, Lamar (WIM 360009) and Lonoke (WIM 430037) were higher volume sites with a significant proportion of shared truck volume, and Bald Knob (WIM 730068) was selected as a low volume site. It is to note that due to WIM relocation
the data collection stations in Bald Knob and Lonoke were set up on previous coordinates one mile upstream from the relocated WIM sites. For the second round of data collection on March 2019 the WIM sites selected were Glen Rose (WIM 301769) and Arkadelphia (WIM 100019) on I-30 and Texarkana (WIM 460286) on I-49. The selected static scale for this instance was Hope on I-30 located between the Glen Rose and Arkadelphia WIM sites.

Figure 8. Drivewayze Truck Traffic Patterns (map prepared by Fu Ren Zhang Durandal in QGIS)

3.2.2 Field Data Collection

Data collection was performed on March 15th, 2018 and March 19th, 2019. In the 2018 data collection raw or “un-calibrated” weights at WIM sites were recorded because auto-calibration was turned off at the selected sites. In the 2019 data collection auto-calibration was running for all
the selected sites thus the PVR records for the 2019 sites have calibrated axle and gross vehicle weights per the ARDOT calibration method. The static weight data collection and visual data recordings procedure in 2018 was repeated in 2019 were cameras were set up at each static and WIM site to record passing trucks. Static weights were collected from trucks that stopped at the static enforcement scales using weight receipts that recorded axle and gross vehicle weights (Figure 9). The FAW was categorized as ‘Steer’, the second and third axles where weighed as ‘Drive’, and third and fourth axle weighed were as ‘Trailer’ further distinguished as ‘TrailerA’ and ‘TrailerB’ if multiple trailers were present. Still images and video recordings were collected at the static sites for traffic that got off the interstate into the weight station. Video was recorded for all the selected WIM sites for the study. GPS records were gathered for all WIM sites in the Arkansas network. Figures 8 and 9 show the truck visual data recorded by the cameras, all images were logged into a spreadsheet.

It should be noted that during the March 2018 data collection the WIM sites at Lonoke and Bald Knob were not located at the latitude/longitude positions indicated in the WIM site specifications. Instead, they had been moved about 1 mile upstream of their current locations. Since the incorrect positions were shared with the GPS data provider, the AVI screenline point and the WIM station did not correspond to the same location. The camera was set up at the correct WIM station. Further complicating the data collection, a traffic incident occurred upstream of the WIM site. This reduced our ability to match WIM and AVI data at the Lonoke site.
Figure 9. Example of a weight receipt (A) and weight recording configuration (B)

Figure 10. Trucks being weighed at a weigh station (Photos taken by Research Hernandez lab team)

Figure 11. Truck images from traffic recordings at WIM sites (Photos taken by Research Hernandez lab team)
3.3 Data Pre-Processing

Data pre-processing consisted of (1) matching trucks crossing the static scale to trucks crossing the WIM sites using static images and video recordings and (2) matching truck GPS records to WIM records using video recordings. This data was used to validate the automatic proposed truck matching and calibration methods.

3.3.1 Matching Trucks Crossing Static Scales and WIM Sites

Prior to re-identifying shared trucks crossing the selected stations, the time offset between the cameras and the WIM sensors had to be determined. This was performed by looking at truck sequencing patterns from the WIM vehicle records and then finding this truck sequence in the traffic videos using time stamps as reference (e.g. one-minute buffer of video watching around the WIM record timestamps). First, we compared vehicle headways by vehicle class (Table 4). This was repeated for the morning, noon, and afternoon at each study site to find an average time offset between the WIM and video. The example below pertains to the time offset for the Texarkana video and WIM for the morning which was a difference of 17 seconds.

The video processing for identifying shared trucks between the static scales and WIM was performed by manually examining pictures of trucks weighed at the static scale and re-identifying them in the video recordings from the WIM sites. For the 2018 data collection the high-volume sites Lamar and Lonoke had a total of 106 shared trucks (e.g., trucks that crossed both WIM sites) of which 69 were also seen at the static scale at Alma. Bald Knob, the low volume site, had only two trucks that crossed the static scale. In the 2019 data collection for the higher volume WIM sites of Glen Rose and Arkadelphia had 97 and 157 shared trucks with the Hope static scale respectively. For the site with lower volume, Texarkana, 22 shared trucks were identified.
Table 5. Video to WIM time offset calculation example.

<table>
<thead>
<tr>
<th>WIM Records</th>
<th>Video</th>
<th>Estimated Time Offset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record</td>
<td>Class</td>
<td>Timestamp</td>
</tr>
<tr>
<td>604</td>
<td>5</td>
<td>10:02:40</td>
</tr>
<tr>
<td>605</td>
<td>7</td>
<td>10:02:52</td>
</tr>
<tr>
<td>606</td>
<td>9</td>
<td>10:03:05</td>
</tr>
<tr>
<td>607</td>
<td>9</td>
<td>10:03:07</td>
</tr>
<tr>
<td>608</td>
<td>4</td>
<td>10:03:12</td>
</tr>
</tbody>
</table>

**Average time offset** -0:00:17

### 3.3.2 Matching WIM and GPS Records

The time offset between the WIM and AVI GPS data had to be established to match WIM records to GPS records in order to compare recorded weights of each truck. The processing began by examining 20 minutes of video for each GPS truck e.g. a 10-minute buffer around the GPS timestamp. Images and descriptions of each truck were recorded. Then the same process was followed when observing the video at the second WIM site looking for trucks that had previously crossed first WIM site within the 10-minute buffer of the GPS time stamp. After trucks started being successfully matched or re-identified crossing both WIM sites, the video to GPS time offset was determined. At Lamar the offset was 12 to 17 seconds and at Lonoke it was 1 min 45 seconds to 3 minutes. The greater time variability at Lonoke was attributed to traffic congestion caused by a traffic accident during the data collection.

The following steps explain how trucks were identified and matched once the time offsets were found using the Lamar and Lonoke sites as an example:

1. Watch the video at Lamar of trucks near the Drivewyze timestamp considering the 17s offset between the video and DW records. Take screen shots (Figure 13) and notes of the trucks.
2. Observe the video footage from Lonoke near the GPS timestamp considering the offset between the video and GPS records at Lonoke (1min 45s - 3min) to find if any of the trucks from the captured images at Lamar cross the Lonoke WIM site. Figure 14 shows the truck found at Lonoke, which corresponds to a truck that previously crossed Lamar site in Figure 13.

3. Record the “matched” truck (Figure 14).
This process resulted in a list of trucks and their WIM measurements for trucks that crossed Lamar and Lonoke (Figure 15). A challenge was that the AVI to video time offset had some variability due to traffic flow at each site. GPS records were matched more precisely to WIM records by examining headways of WIM and video records. This was performed in order to obtain one to one matches at Lamar and Lonoke between the GPS and WIM vehicle records also looking at the truck lane and class sequence to find the exact match in order to develop the data to validate truck matching and auto-calibration algorithms. The time offset at Lonoke was much greater as the traffic video collection station was 1 mile away from the actual site and an accident near the data collection site that caused traffic discontinuity.

Figure 15. Example of finalized matches between video, GPS, and WIM records
Chapter 4. Methods

The WIM auto-calibration model consisted of two parts as highlighted in Flow Chart 2: (1) A truck matching algorithm and (2) An auto-calibration algorithm. Each are described in this section.

Flow Chart 2. Project task overview
4.1 Truck Matching Algorithm

Tracking trucks traveling over more than one WIM sensor allows comparison of vehicle characteristics recorded at the different sites thus producing calibration factors. As part of data pre-processing, we manually identified offsets between the AVI and WIM records and manually matched AVI truck records to WIM records using video data. In this section, we describe the automatic algorithm, called “Truck Matching”, used to perform the same task.

The Truck Matching algorithm followed three steps as shown in Flow Chart 3: (1) Time Offset Calculation, (2) Match Filtering, and (3) Data Pairing. Each is described below.

**Flow Chart 3. Truck matching inputs, process and outputs.**

4.1.1 Offset Calculation

Since the GPS data provides unique identifiers (IDs) for each truck record, trucks can be tracked across WIM sites solely based on their ID (e.g., advanced truck re-identification was not necessary). However, it was necessary to find the time offset between the GPS records and the WIM PVRs via an automated process. The algorithm steps were as follows:
1. Query the GPS truck records for a given day and year to produce a list of stations crossed by GPS trucks.

2. Loop through each station, then loop through each GPS truck at each station.

3. For each GPS truck, query WIM PVRs within a specified time window around the GPS truck timestamp. The initial time window was set to three minutes (180 seconds).

4. For each WIM PVR returned by the query, calculate the offsets between the PVRs and the GPS truck timestamps.

5. Find the ‘mode’ (e.g., the most frequently occurring value) among all GPS trucks and PVR offsets.

6. If no mode exists or there are multiple modes within the initial time window, then widen the window and repeat Steps 3-6. The time window was widened by 10 seconds each iteration and allowed to increase to five minutes. Note that no mode exists if all offset values occur only once.

7. When a mode is found, assign it as the offset for the station.

4.1.2 Match Filtering

AVI records of trucks tracked across multiple WIM sites were subjected to a match filter. A travel time filter was applied first to reduce the possibility of truck weight differences due to pick-up and deliveries between sites. A maximum travel time threshold ensured the same truck with the same cargo and trailer was found. The maximum threshold was based on observed travel time distributions among WIM sites. A minimum travel time threshold controlled for recording errors inherent in the GPS data. A temporal window of one day was applied to the AVI matched trucks ($d_k$) to filter out potential variation in weights due to drop-off, pick-up, cargo and trailer changes.
4.1.3 Data Pairing

Data pairing between the AVI and WIM records was needed since when matching AVI to WIM records there are typically several candidate matches within the time window of each AVI record even after filtering out records outside the match filter. There are many more WIM records than AVI records, e.g. the AVI data represents less than 10% of the total truck population. Therefore, the set of candidate matches was reduced by examining the timestamps and axle spacing recorded by the WIM. The objective of data pairing is to assign each AVI record uniquely to a WIM record. The algorithm was carried out in two steps: (1) identify candidate WIM records for each AVI record, and (2) assign a unique WIM record to each AVI record.

**Step 1. Identify Candidate WIM records.** A time buffer, $\Delta$, around the AVI timestamp $(t_k,s + \Delta)$ for each site of interest, for each truck $d_k$ was established based on time offsets and was used to obtain candidate WIM records. The set of candidate matches for truck $d_k$ was:

$$C(t_k,s) = [ W((t+x)-\Delta)_{s,i}, ..., W(t+x)_{s,i,+n}, ..., W((t+x) +\Delta)_{s,m}]$$

Where:

- $C(t_k,s)$ is the set of WIM records corresponding to the time stamp of an AVI truck at site $s$ at time $t_k,s$.
- $W(t)_{s,i}$ is the WIM record of the vehicle at site $s$, timestamp $t$ such that the set of candidates is within a buffer, $\Delta$, around $t_k,s$ ($t-\Delta$, $t$, $t+\Delta$), $i = 1... m$.

**Step 2. Assign WIM record to AVI record.** Finally, for an AVI truck $d_k$ crossing stations $s$ and $s'$ the set of candidate WIM records were filtered to find a unique match such that the time stamp and vehicle parameters from each corresponding WIM record were minimized. A matrix $D_{s,s'}$ representing all pairwise combinations of WIM to WIM pairs ($W_{s,i}$ to $W_{s',j}$) contained the candidate sets $C(t_k,s)$ and $C(t_k,s')$ for sites $s$ and $s'$. The sum of absolute differences of vehicle parameters was used as a metric to find the unique match. The WIM records that produce the
minimum difference \( \argmin(D_{s,s'}, \{W_{s,i}, W_{s',j}\}) \) and were within all temporal constrains are selected as a unique match. The parameters compared in the study were inter axle spacings.

\[
D_{s,s'} = \begin{bmatrix}
|W_{s,1} - W_{s',1}| & \ldots & |W_{s,1} - W_{s',m}| \\
\vdots & \ddots & \vdots \\
|W_{s,n} - W_{s',1}| & \ldots & |W_{s,n} - W_{s',m}|
\end{bmatrix}
\]

Where

- \( D_{s,s'} \) is the matrix of differences between all WIM records at sites \( s \) and \( s' \) pertaining to \( t_{k,s} \) and \( t_{k,s'} \)
- \( W_{s,i} \) is WIM record \( i \) at site \( s \) contained in \( C(t_{k,s}) = [W(t-\Delta)_{s,i}, \ldots, W(t)_{s,i}, \ldots, W(t+\Delta)_{s,i}] \), \( i = 1 \ldots n \)
- \( W_{s',j} \) is WIM record \( j \) at site \( s' \) contained in \( C(t_{k,s'}) = [W(t-\Delta)_{s',j}, \ldots, W(t)_{s',j}, \ldots, W(t+\Delta)_{s',j}] \), \( j = 1 \ldots m \)
- \( |W_{s,n} - W_{s',m}| = \sum_{p=1}^{p} |y_{p,s} - y_{p,s'}| \), the sum of the absolute differences between vehicle parameters, \( y \), for sites \( s \) and \( s' \).

### 4.2 Auto-Calibration Algorithm

The output of the Truck Matching Algorithm was WIM records paired to each AVI truck record. This data was then used to compare weights of the same vehicle at different WIM sites as a form of auto-calibration (Flow Chart 4) in which calibration factors were determined based on the weight of the same truck measured at different WIM sites.

The proposed auto-calibration method computed hourly; site specific calibration factors derived from the \textit{steering axle weights} of AVI VC 9 ‘3-S2’ configured trucks that crossed more than one WIM site at all visited WIM sites, the algorithm first calculated the deviation among \textit{steering axle weights} (also referred to as FAW) for the same AVI truck. Pairwise differences between \textit{steering axle weights} were calculated and the differences were measured against a predefined threshold, \( \delta_{S} \).

If percent of sites with \textit{steering axle weights} above \( \delta_{S} \) was greater than a predefined threshold on the number of sites in agreement, \( P_{S} \), the calibration factor for each site for the specified hour was
set to 1.0, e.g., all sites were in agreement on the *steering axle weight* and thus are in calibration. Otherwise, for sites with *steering axle weights* in “disagreement”, a calibration factor was calculated as follows.

First, high volume sites were used to determine the *likely weight* \( \omega_a \) of the truck’s steering axle. The *likely weight* was found through a cluster analysis in which *steering axle weights* corresponding to the AVI truck across multiple WIM sites were compared to find a common measurement. The inputs to clustering were the *steering axle weight* and the GVW.

![Cluster Analysis Diagram](image)

**Figure 16. Example of Cluster Analysis for AVI Auto-Calibration**

Then, the *likely weight* was compared to a reference weight \( W_R \) (e.g. the same used in the traditional ARDOT auto-calibration method, 10 kips) to assess its reasonableness. The *likely weight* was used to compute the calibration factor if it fell within a certain deviation, \( \delta_W \), of the reference weight, otherwise the reference weight was used. The *likely weight* found through the cluster analysis of the high volume sites were used to compute the calibration factors for the high and low volume sites. Thus, each truck \( k \) produced a calibration factor corresponding to each site \( i \) during the corresponding hour \( h \) it was detected at the site:
\[ CF_{k,i,h} = \frac{W_{k,i,h}}{\omega_a} \quad \text{Eq. 2} \]

where
\[ W_{k,i} \] is the WIM recorded weight of the truck \( k \) at site \( i \)

Next, calibration factors for each truck \( (CF_{k,i,h}) \) were averaged for each site to determine the average calibration factor, \( CF_{i,h} \), for site for each hour. Finally, the adjusted weights for every truck at each WIM site were computed as:

\[ \hat{W}_{k,i,h} = CF_{i,h} \times W_{k,i,h} \quad \text{Eq. 3} \]

All terms previously defined.

The key distinction between traditional and the proposed auto-calibration was the method to incorporate reference axle weights to compute calibration factors. In traditional auto-calibration algorithms, WIM measured weights are compared to a predetermined, non-changing reference weight, such as a reference FAW, to compute a calibration factor. In the proposed auto-calibration method the use of a reference weight was replaced by a likely weight \( (\omega_a) \) defined from the AVI-WIM pairs.

Rather than averaging, choosing the mode, or using a median steering axle weight from the set of WIM steering axle weights for a truck, the clustering approach was adopted to ensure that the likely weight reflected the majority among all measurements while also allowing, by varying the number of clusters, the ability to detect outliers or steering axle weight discrepancies resulting from different GVWs (e.g. if the truck made a pick up or delivery between WIM sites). Moreover, the proposed auto-calibration algorithm distinguished between high and low volume sites when estimating the likely weight. This was an important distinction because FAW measurements taken at high volume WIM sites (e.g., sites with more than 50 FHWA Class 9 trucks per hour) tended to be more accurate than those collected at low volume sites as more values could be used for
clustering. Differences in accuracy can be attributed to the increased auto-calibration frequency at high volume sites that, in turn, tracks with temperature changes. Ambient and pavement temperatures significantly affect WIM piezo sensor accuracy.
Flow Chart 4. AVI Auto-Calibration Procedure

4.3 Auto-calibration Algorithm Implementation Example

An illustrative example of the Truck Matching and Auto-calibration methods are provided here using WIM stations Lamar, Lonoke and Bald Knob.

4.3.1 Truck Matching Example

Vehicle inter axle spacing in feet were utilized in the example as the vehicle parameter to compare in order to find the WIM records for AVI trucks.
**Step 1: Determine time offsets.** The AVI to WIM time offsets at Lamar, Lonoke and Bald Knob are found to be 15s, 20s, and 12s, respectively.

**Step 2: Identify Candidate WIM records.** An AVI truck with unique ID ‘123’ is identified crossing these sites within one day at the following times.

\[
[t\text{uck}_{123}] : \\
\text{Lamar @ } t_1 = 9:00:00 \text{ AM} \\
\text{Lonoke @ } t_2 = 10:00:00 \text{ AM} \\
\text{Bald Knob @ } t_3 = 10:30:00 \text{ AM}
\]

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Time 1</th>
<th>Site 2</th>
<th>Time 2</th>
<th>Site 3</th>
<th>Time 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lamar\text{360009}</td>
<td>9:00:00 AM</td>
<td>Lonoke\text{430037}</td>
<td>10:00:00 AM</td>
<td>Bald knob\text{730068}</td>
<td>10:30:00 AM</td>
</tr>
</tbody>
</table>

**Step 2: Assign WIM records to AVI records.** A buffer of $\Delta = 5$ minutes was used in this example to find sets of candidate WIM records at each site.

**Candidate WIM records for $t_{\text{uck}_{123}}$ 5 min around $t_1 = 9:00:00 \text{ AM at Lamar}$ →**

<table>
<thead>
<tr>
<th>Record</th>
<th>Time</th>
<th>AVI Time</th>
<th>WB 1</th>
<th>WB 2</th>
<th>WB 3</th>
<th>WB 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>8:55:15AM</td>
<td>8:55:00AM</td>
<td>17.50</td>
<td>4.02</td>
<td>28.78</td>
<td>4.64</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>95</td>
<td>9:00:15AM</td>
<td>9:00:00AM</td>
<td>17.00</td>
<td>4.20</td>
<td>32.00</td>
<td>4.15</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>100</td>
<td>9:05:15AM</td>
<td>9:05:00AM</td>
<td>15.88</td>
<td>4.23</td>
<td>32.98</td>
<td>4.71</td>
</tr>
</tbody>
</table>

Where AW 1, 2… is the weight of the 1\textsuperscript{st}, 2\textsuperscript{nd}, etc. axle. Note that inter-axle spacing could also be included.

**Candidate WIM records for $t_{\text{uck}_{123}}$ 5 min around $t_2 = 10:00:00 \text{ AM at Lonoke}$ →**

<table>
<thead>
<tr>
<th>Record</th>
<th>Time</th>
<th>AVI Time</th>
<th>WB 1</th>
<th>WB 2</th>
<th>WB 3</th>
<th>WB 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>9:55:20AM</td>
<td>9:55:00AM</td>
<td>16.72</td>
<td>4.21</td>
<td>32.93</td>
<td>4.00</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>220</td>
<td>10:00:20AM</td>
<td>10:00AM</td>
<td>17.00</td>
<td>4.20</td>
<td>32.00</td>
<td>4.15</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>240</td>
<td>10:05:20AM</td>
<td>10:05:00AM</td>
<td>15.99</td>
<td>4.13</td>
<td>30.84</td>
<td>4.02</td>
</tr>
</tbody>
</table>
Candidate WIM records for truck \textsubscript{123} at Bald Knob →

<table>
<thead>
<tr>
<th>Record</th>
<th>Time</th>
<th>AVI Time</th>
<th>WB 1</th>
<th>WB 2</th>
<th>WB 3</th>
<th>WB 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>172</td>
<td>10:25:12AM</td>
<td>10:25:00AM</td>
<td>19.50</td>
<td>4.28</td>
<td>31.50</td>
<td>3.96</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>175</td>
<td>10:30:12AM</td>
<td>10:30:00AM</td>
<td>17.00</td>
<td>4.20</td>
<td>32.00</td>
<td>4.15</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>180</td>
<td>10:35:12AM</td>
<td>10:35:00AM</td>
<td>17.69</td>
<td>5.72</td>
<td>29.56</td>
<td>4.32</td>
</tr>
</tbody>
</table>

The vehicle parameters differences were used to determine potential matches. The vehicle parameter used in this example was axle weights. For instance, WIM record 95 at Lamar would be compared to records 200 thru 240 at the Lonoke WIM site and to records 172 thru 180 at the Bald Knob site. Notice that records 95, 220 and 175 would have the least overall difference. The following are GVW differences:

A. Data vector for Lamar: \( W_{95, \text{Lamar}} = [ ] \)

B. Differences between truck 95 at Lamar and candidates at Lonoke:

\[
W_{95, \text{Lamar}} - W_{220, \text{Lonoke}} = 0 \\
W_{95, \text{Lamar}} - W_{200, \text{Lonoke}} = 1.47 \\
W_{95, \text{Lamar}} - W_{240, \text{Lonoke}} = 2.37
\]

Resulting match: \( \text{argmin}(\{0, 1.47, 2.37\}, \{W_{s,i}, W_{s’,j}\}) = ( W_{95, \text{Lamar}}, W_{220, \text{Lonoke}} ) \)

C. Differences between truck 95 at Lamar and candidates at Bald Knob:

\[
W_{95, \text{Lamar}} - W_{175, \text{Bald Knob}} = 0 \\
W_{95, \text{Lamar}} - W_{172, \text{Bald Knob}} = 3.27 \\
W_{95, \text{Lamar}} - W_{345, \text{Bald Knob}} = 4.82
\]

Resulting match: \( \text{argmin}(\{0, 3.27, 4.82\}, \{W_{s,i}, W_{s’,j}\}) = ( W_{95, \text{Lamar}}, W_{228, \text{Bald Knob}} ) \) as \( W_{175, \text{Bald Knob}} \) has a closer time stamp to the AVI given the time offset, and least overall difference in axle weights as well.

Records 95, 220 and 175 have the lowest overall difference therefore these records are assigned to truck 123 for auto-calibration. Thus, the final unique WIM pairings for AVI truck ‘123’ are as follows:
The example above is suited for a WIM system network that does not have a lot of variability in weight recordings preferably using WIM systems with bending plates or load cells which are more accurate than piezoelectric sensors and WIM systems that are not type II as they have a higher variability of 30% combined with the sensitivity to temperatures and pavement conditions of piezoelectric sensors (FHWA, 2018). For the case study inter axle spacing was selected as the comparable vehicle parameter as these are more consistent measurements across FHWA VCs and due to the wide range in weight variability across WIM sites that was experienced in the recorded data.

### 4.3.2 Auto-Calibration Example

The following is an idealized example of the AVI auto-calibration method presented in Flow Chart 4. In this example a 1-hour sample of trucks taken from 9 AM to 10 AM at WIM Station A includes three trucks: truck IDs 101, 105, and 203. WIM records for the same AVI trucks found at WIM Station A within time window $T$ of 3.5 hours included three additional stations: B, C, and F. The reference steering axle weight, $W_R$, was set to 10 kips. The deviation among steering axle weights among the sites ($\delta_s$) was 10%. The deviation ($\delta_w$) between the reference weight and likely weight was 10%. The volume of FHWA Class 9 trucks at each WIM site on the given day were: 300, 250, 190, and 50, for sites A, B, C, and F, respectively. The algorithm is as follows:

1. **Select samples:** Obtain a set of FHWA Class 9 trucks crossing WIM site A from the AVI data.

   AVI truck sample at WIM A from 9 AM to 10 AM.
2. **Find AVI trucks:** Find each of the AVI trucks from site A that traversed other WIM sites in the 3.5 travel time window and count the number of sites crossed by each truck: truck 105 crossed 4 sites, truck 101 crossed 3 sites, and truck 203 crossed three sites.

<table>
<thead>
<tr>
<th>Truck ID</th>
<th>Timestamp</th>
<th>FAW (kips)</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>9:03:00</td>
<td>10</td>
</tr>
<tr>
<td>101</td>
<td>9:05:00</td>
<td>12</td>
</tr>
<tr>
<td>203</td>
<td>9:30:00</td>
<td>9</td>
</tr>
</tbody>
</table>

**Trucks crossing WIM B:**

<table>
<thead>
<tr>
<th>Truck ID</th>
<th>Timestamp</th>
<th>FAW (kip)</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>10:03:00</td>
<td>9</td>
</tr>
<tr>
<td>101</td>
<td>10:05:00</td>
<td>11</td>
</tr>
<tr>
<td>203</td>
<td>10:30:00</td>
<td>10</td>
</tr>
</tbody>
</table>

**Trucks crossing WIM C:**

<table>
<thead>
<tr>
<th>Truck ID</th>
<th>Timestamp</th>
<th>FAW (kip)</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>11:33:00</td>
<td>10</td>
</tr>
<tr>
<td>101</td>
<td>11:35:00</td>
<td>10</td>
</tr>
<tr>
<td>203</td>
<td>12:00:00</td>
<td>12</td>
</tr>
</tbody>
</table>

**Trucks crossing WIM F:**

<table>
<thead>
<tr>
<th>Truck ID</th>
<th>Timestamp</th>
<th>FAW (kip)</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>13:03:00</td>
<td>8</td>
</tr>
</tbody>
</table>

2. **Check Deviation:** Check the deviation, $\delta_s$, in *steering axle weight* of each AVI truck recorded at each WIM site to see if the sites require calibration. For this example, $\delta_s$ was 10% so that if the difference in *steering axle weights* (or FAW) recorded at two WIM sites for the same truck differed by more than 10%, we considered them to need calibration. The following table shows necessary calculations:

| Truck ID | Timestamp A | FAW A | Timestamp B | FAW B | Difference $|A-B|/A$ |
|----------|-------------|-------|-------------|-------|----------|
| 105      | 9:03:00     | 10    | 10:03:00    | 9     | 10%      |
| 101      | 9:05:00     | 12    | 10:05:00    | 11    | 8%       |
| 203      | 9:30:00     | 9     | 10:30:00    | 10    | 11%      |
Deviation for sites B and C:

| Truck ID | Timestamp B | FAW B | Timestamp C | FAW C | Difference | |B-C|/B |
|----------|-------------|-------|-------------|-------|------------|--|----|
| 105      | 10:03:00    | 9     | 11:33:00    | 10    | 11%        |
| 101      | 10:05:00    | 11    | 11:35:00    | 10    | 9%         |
| 203      | 10:30:00    | 10    | 12:00:00    | 12    | 20%        |

Deviation for C and F:

| Truck ID | Timestamp C | FAW C | Timestamp F | FAW F | Difference | |C-F|/C |
|----------|-------------|-------|-------------|-------|------------|--|----|
| 105      | 11:33:00    | 10    | 13:03:00    | 8     | 20%        |

By computing the deviations above it may be observed that in most cases the weights for a an AVI truck are significantly different in all cases except for truck 101 where the weights recorded in sites A and B are below the deviation therefore they are similar. In this case, the calibration factor resulting from truck 101 recorded at sites A and B is 1.0.

3. **Find Likely Weight**: To compute the likely weights, \( \omega_a \), we first differentiate between high and low volume sites based on historical AVI data such that a site with over 50 AVI trucks per day was considered to be high volume. This is referenced via a look up table. Clustering is used to find \( \omega_a \) when there is more than one high volume site. Then the \( \omega_a \) was compared to the reference FAW, \( W_R \), of 10 kips to see if it is within a weight deviation, \( \delta_W \), of 10%. If deviation between \( \omega_a \) and \( W_R \) exceeds \( \delta_W \) then \( W_R \) was used to compute the calibration factor, otherwise \( \omega_a \) was used. A calculation for Truck 105 was as follows:

**Truck 105:**
- FAWs were 10, 9, 10, and 8 for sites A, B, C, and F
- High volume sites = A, B, C
- \( \omega_a = 9.75 \) kips from clustering analysis (e.g., cluster with 10, 9, and 10 kip steering axle weights and GVWs)
- Deviation to reference weight: \( (9.75-10.00)/10.00 \times 100\% = 2.5\% \)
- Comparison to threshold: \( \delta_W = 10\% > 2.5\% \), therefore use \( \omega_a = 9.75 \) kips

4. **Calculate Calibration Factors**: The calibration factors were calculated as the ratio of the likely weights, \( \omega_a \), to the recorded FAW. An example for site A is:
Site A:

<table>
<thead>
<tr>
<th>Truck ID</th>
<th>Timestamp A</th>
<th>FAW A</th>
<th>Likely Weight A</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>9:03:00</td>
<td>10</td>
<td>9.75</td>
<td>9.75/10 = 0.975</td>
</tr>
<tr>
<td>101</td>
<td>9:05:00</td>
<td>12</td>
<td>11.20</td>
<td>11.20/12 = 0.933</td>
</tr>
<tr>
<td>203</td>
<td>9:30:00</td>
<td>9</td>
<td>10.40</td>
<td>10.40/9.0 = 1.15</td>
</tr>
<tr>
<td>Average</td>
<td>9:00 to 10:00</td>
<td>-</td>
<td>-</td>
<td>1.02</td>
</tr>
</tbody>
</table>

Since site F was a low volume site the likely weight of truck 105 determined from clustering FAWs from sites A, B, and C was used to calculate the calibration factor for Site F as follows.

Site F:

<table>
<thead>
<tr>
<th>Truck ID</th>
<th>Timestamp F</th>
<th>FAW F</th>
<th>Likely Weight A</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>13:03:00</td>
<td>8</td>
<td>9.75</td>
<td>9.75/8 = 1.22</td>
</tr>
<tr>
<td>Average</td>
<td>13:00 to 14:00</td>
<td>-</td>
<td>-</td>
<td>1.22</td>
</tr>
</tbody>
</table>

5. Calibrate Site: The resulting calibration factors generated from the AVI trucks were used to adjust the weights recorded by the WIM for all trucks by dividing each of the WIM measured weights by the calibration factor.
Chapter 5. Results

This section presents the results in terms of measurement accuracy according to three metrics. The Truck Matching and AVI Auto-Calibration methods were evaluated using data collected in the field in March 2018 as described in Section 3.2. The Truck Matching algorithm was evaluated against manually matched AVI to WIM records based on video analysis. The Auto-Calibration algorithm was evaluated by comparing algorithm-adjusted weights to the static weights recorded at the weight enforcement stations.

5.1 Summary of Data Collection Efforts

Data was collected on March 15, 2018 and March 20, 2019 at different sites (Table 6). The WIM sites selected for the 2018 data collection were eastbound along I-40 at Lamar and Lonoke and east/northbound along Highway 167 at Bald Knob. The eastbound Alma weigh station along I-40 was used as the static enforcement site. The WIM PVR records for 2018 show that at around 6 AM the auto-calibration algorithm was turned off at Lamar, Lonoke, and Bald Knob and from that time raw weight records were reported. For the 2018 collection, the number of PVR records at Lamar, Lonoke and Bald Knob were 5,346, 10,801 and 1,963 respectively. At the Alma static scale, 263 trucks were recorded, from these 106 were re-identified at Lamar, 69 at Lonoke and 2 at Bald Knob. A total of 121 AVI trucks crossed Lamar and Lonoke during video recording hours (8AM – 6 PM), 44 of these trucks also crossed the Alma weigh station. After removing trucks with WIM error flags or mismatched vehicle classes between the WIM sites, we had 33 samples to use for algorithm validation.

The WIM sites for the data collection in 2019 were south/westbound on I-30 at Glen Rose and Arkadelphia, and southbound on I-49 at Texarkana. The south/westbound weigh station on I-30
at Hope was used as the static enforcement site. The WIM PVR for this instance were recorded with the ARDOT calibration method on, therefore the PVR are all adjusted. In 2019, the number of PVR records for Glen Rose, Arkadelphia and Texarkana were 9,905, 11,007, and 2,159 respectively. The number of trucks weighed at Hope weigh station was of 261, from these 88 were re-identified at Glen Rose, 157 at Arkadelphia, and 17 at Texarkana. Unfortunately, a data logging error occurred during the data collection and the AVI data for the southbound WIM sites at Glen Rose and Arkadelphia were not available. Therefore, we were not able to evaluate the WIM to AVI truck matching algorithm. But we were able to replicate the AVI data with video records to evaluate the auto-calibration method at these sites. There were 71 trucks re-identified trucks crossing Glen Rose, Arkadelphia and Hope stations and 8 trucks were re-identified crossing Arkadelphia, Hope, And Texarkana. Only two trucks crossed all the selected sites, Glen rose, Arkadelphia, Hope and Texarkana.

Table 6. Data Collection Summary

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>Site</th>
<th>WIM PVRs</th>
<th>AVI Trucks Matched to Static Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>March 15, 2018</strong>&lt;br&gt;1-40 EB/Hwy 167 NB</td>
<td>Lamar</td>
<td>5,346</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Lonoke</td>
<td>10,801</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Bald Knob</td>
<td>1,963</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Alma (static scale)</td>
<td>263</td>
<td></td>
</tr>
<tr>
<td><strong>March 20, 2019</strong>&lt;br&gt;1-30 SB/I-49 SB</td>
<td>Glen Rose</td>
<td>9,905</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Arkadelphia</td>
<td>11,007</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>Texarkana</td>
<td>2,159</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Hope (static scale)</td>
<td>261</td>
<td></td>
</tr>
</tbody>
</table>

5.2 Truck Matching Algorithm Performance

The Truck Matching Algorithm was evaluated using three performance indexes: true match rates (TMR), correct match rates (CMRs) and error rates (ER) as follows (Equations 2-4):
The True Match Rate, $TMR = \frac{M_{true}}{T_{veh}}$ (Eq. 2)

Correct Match Rate, $CMR = \frac{M_{TM}}{M_{true}}$ (Eq. 3)

Error Rate, $ER = \frac{M_{miss}}{M_{true}}$ (Eq. 4)

Where:
- $T_{veh}$ = total number of vehicles observed at the site
- $M_{true}$ = total number of actual true matches from groundtruth
- $M_{TM}$ = number of successful matches obtained using algorithm
- $M_{miss}$ = number of mismatched trucks selected by algorithm

The TMR reflects the groundtruth process in terms of our ability to match all vehicles seen in the video. There were three reasons for not being able to manually match all AVI trucks to their corresponding WIM record to achieve 100% TMR. First, trucks had to be visually confirmed to have passed both stations but with the camera recording from a side-fire position it was not possible to view trucks in the inner lane due to occlusion. Second, trucks were not able to be visually confirmed if a site had a high variability in the time offset between the records and the video. This was found to occur at Lonoke due to traffic congestion upstream of the data collection site. Third, some trucks seen in the video and recorded in the AVI data were not recorded by the WIM sensor. This is likely due to sensor error or the truck travelling off center to the sensors causing measurement error.

During the March 2018 data collection, a total of 121 AVI trucks traveled from Lamar to Lonoke. Out of these 121 trucks, 93 (e.g., TMR of 77%) were successfully matched with their respective WIM PVR at Lamar. At Lonoke, matches between PVR and AVI were only sought for the trucks also found at Lamar. Thus, all 93 AVI trucks were successfully matched to their WIM PVR record, e.g., 100% TMR.

The CMR and ER reflect the ability of the matching algorithm to correctly match WIM PVR and AVI truck records. CMR assesses the truck matching algorithm success rate such that a value
closer to 100% is better. It used the 93 successfully matched WIM records that were able to be correctly matched to AVI records. The CMR at Lamar was 75% and at Lonoke 52%. An initial time window of 180 seconds (3 minutes) was found to produce the highest CMR across all sites. The selected time window was based on trial and error, running the algorithm under different time window settings which yielded the best CMR. ER captures the same concept as CMR but is represented as error, e.g., the goal is to achieve a low ER. Thus, the ER for Lamar was 25% and Lonoke was 48%.

Lower CMR can be attributed to WIM sensor errors like missed detections, ghost detections (detections of vehicles that were not actually there), counting vehicles with two trailers as two separate vehicles, and counting vehicles straddling two lanes as two separate vehicles. Traffic flow was also a contributing factor, as demonstrated in Lonoke. A lower CMR (higher ER) at Lonoke was also attributed to an upstream accident that occurred around noon during data collection which caused larger variability in the time offset between the WIM and AVI records (Figure 17). Recall the location of the WIM site and the AVI screenline and camera were about 1 mile apart. Most of the shared AVI trucks between Lamar and Lonoke crossed Lonoke around noon (Figure 18) also contributing to the lower CMR. The temporal inputs and sequencing of the records were two central inputs components of the Truck Matching algorithm therefore having uninterrupted traffic flow is critical if the WIM and AVI locations differ.
Figure 17. Error Rate by time of day at the Lonoke WIM site during the March 2018 data collection

Figure 18. Shared AVI truck from Lamar to Lonoke during data collection period

5.2 Auto-Calibration Algorithm Performance

Absolute Percent Error (APE), Mean Absolute Percent Error (MAPE), and Median Absolute Percent Error (MdAPE) were used to measure the discrepancy between the auto-calibration algorithm outputs and static (or true) weights. The MdAPE was used as it is less sensitive to outliers than MAPE. The performance measures were obtained as follows:

\[
APE = \left( \frac{|\text{Calibrated WIM Weight} - \text{Static Weight}|}{\text{Static Weight}} \right) \times 100 \quad \text{(Eq. 5)}
\]

\[
MAPE = \frac{\sum APE}{n} \quad \text{(Eq. 6)}
\]

\[
MdAPE = \text{median}(APE) \quad \text{(Eq. 7)}
\]

Where:
*Calibrated WIM Weight* is the truck axle or GVW weight adjusted using calibration factors produced by the AVI Auto-calibration method.

*Static Weight* is the truck axle or GVW weight measured by static scales.

The performance of the AVI, ARDOT, and MNDOT auto-calibration methods were evaluated using the above performance metrics which compare dynamically measured to static weights. It is important to note that the adjusted weights estimated by applying calibration factors will always differ from the true static weights due to differences in how the data are collected.

The proposed AVI Auto-calibration algorithm which aggregates CFs hourly produced MAPEs for Lamar of 26% and 39% for FAW and GVW, respectively. For Lonoke the MAPEs were 16% and 37%, for FAW and GVW, respectively. The MdAPEs for Lamar were 23% and 45% for FAW and GVW, respectively. Lonoke MdAPEs were 11% and 15% for FAW and GVW, respectively. The improvements in FAW and GVW accuracy can be seen in Figure 19. The adjustment at Lonoke was more pronounced than at Lamar. Disaggregate CFs which were produced for each vehicle in the method produced more accurate results with MAPEs of 20% and 35% were found for Lamar for FAW and GVW, respectively. The MAPEs for Lonoke were 16% and 35% for FAW and GVW respectively.
The results of the proposed AVI algorithm were compared to the current method used by ARDOT and a MNDOT method for Lamar and Lonoke (Table 6 and Figure 20). The 2018 sites were used to evaluate the method using raw measurements. The AVI auto-calibration method produced at times close or better results for FAW and GVW in comparison to the ARDOT method. Similar MdAPE results were observed at Lonoke and an underperformance at Lamar. It may also be observed that using raw weights at least more consistent GVW MAPEs between Lamar and Lonoke exist of 39% and 37%.

Since AVI data was not available for the 2019 WIM sites, we instead used the video data to replicate the AVI data. To do this, we used video matched trucks as the AVI records and computed calibration factors from each matched truck. We then averaged the calibration factors for each hour and applied the calibration factors back to each truck. Note that this process differs in two keyways. First, the video matched data from all three WIM sites and the static scale is a much smaller sample size than the AVI matches. Second, since the video matched data is only from the three WIM sites, it does not allow weight comparisons to WIM sites around the state, which is different from the proposed AVI-based auto-calibration algorithm. For ease of comparison, we applied the video matched truck replication process to the March 2018 data. This produced lower
MAPES and especially lower MdAPEs around 10% and lower error for the AVI method (Table 7).

Unlike the ARDOT and MnDOT methods, the proposed algorithm calibrates the FAW to measured FAWs rather than a static reference value due to the position of the king-pin, loading configuration, type of cargo, etc., affect the exact weight carried by the FAW. Therefore, we see some variation in its measurement reference FAW. This allows for minor variations in the FAW to be incorporated into the CF. Therefore, although the FAW is used for calibration factors it allows for more accurate adjustment of the tandem axles that sum to estimate the GVW of the vehicle. This principle simplifies having to separate weights into different categories having to measure several static reference FAWs instead of a fluctuating one. However, two main challenges were faced in the study to achieve more accuracy which was the number of trucks used to compute CFs and truck matching. The number of AVI shared trucks able to compute calibration factors at each site were too low ranging from under 1% to 2% of the traffic at each site. The truck matching also faced some issues were traffic irregularity was experienced and the Lonoke site was 1 mile apart from the data collection site lowering the number of correctly matched AVI records to WIM as the only link between them were their time stamps.
Figure 20. Comparison of MAPE by Auto-Calibration Method

(A) FAW

(B) GVW
Table 7. MAPE Comparison for FAW and GVW by site and Auto-Calibration Method

<table>
<thead>
<tr>
<th>Site</th>
<th>No. Records</th>
<th>Method</th>
<th>MAPE (%)</th>
<th>MdAPE (%)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>FAW</td>
<td>GVW</td>
<td>FAW</td>
<td>GVW</td>
</tr>
<tr>
<td>Lamar</td>
<td>33</td>
<td>ARDOT (current)</td>
<td>12.3</td>
<td>17.3</td>
<td>6.6</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MNDOT</td>
<td>10.9</td>
<td>11.4</td>
<td>2.2</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AVI-based</td>
<td>26</td>
<td>39</td>
<td>23</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance Change</td>
<td>-13.7</td>
<td>-21.7</td>
<td>-16.4</td>
<td>-37.4</td>
</tr>
<tr>
<td>Lonoke</td>
<td>33</td>
<td>ARDOT (current)</td>
<td>14.1</td>
<td>29.1</td>
<td>9.9</td>
<td>12.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MNDOT</td>
<td>21</td>
<td>28.9</td>
<td>26</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AVI-based</td>
<td>16</td>
<td>37</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance Change</td>
<td>-1.9</td>
<td>-7.9</td>
<td>-1.1</td>
<td>-2.1</td>
</tr>
<tr>
<td>Glen Rose</td>
<td>71</td>
<td>ARDOT (current)</td>
<td>15.5</td>
<td>19.5</td>
<td>12.7</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MNDOT</td>
<td>23.5</td>
<td>21.4</td>
<td>18.1</td>
<td>17.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AVI-based</td>
<td>41</td>
<td>39</td>
<td>24</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance Change</td>
<td>-25.5</td>
<td>-19.5</td>
<td>-11.3</td>
<td>-11</td>
</tr>
<tr>
<td>Arkadelphia</td>
<td>77</td>
<td>ARDOT (current)</td>
<td>15.3</td>
<td>14.1</td>
<td>12.9</td>
<td>12.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MNDOT</td>
<td>17.6</td>
<td>15.7</td>
<td>12.9</td>
<td>12.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AVI-based</td>
<td>19</td>
<td>23</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance Change</td>
<td>-3.7</td>
<td>-8.9</td>
<td>-2.1</td>
<td>+1.8</td>
</tr>
<tr>
<td>Texarkana</td>
<td>8</td>
<td>ARDOT (current)</td>
<td>13.1</td>
<td>13.5</td>
<td>13.3</td>
<td>22.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MNDOT</td>
<td>12.2</td>
<td>13.8</td>
<td>34.1</td>
<td>26.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AVI-based</td>
<td>8</td>
<td>23</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance Change</td>
<td>+22</td>
<td>-0.3</td>
<td>+7.3</td>
<td>+12.3</td>
</tr>
</tbody>
</table>

We also assessed a variation of the AVI based method in which we estimate a *likely* GVWs instead of a *likely* FAWs. This approach resulted in less accurate results, e.g., higher MAPE and MdAPE. This can be attributed to the high variability in GVWs which makes it impossible to assume a reference GVW to compare the likely GVW. To adapt the algorithm, we used the FAW of each truck to determine which GVW was “correct”, e.g. if the FAW of a truck was outside the tolerance of the reference FAW, then we would not use that truck’s GVW as the *likely* GVW.
Chapter 6. Discussion and Future Work

An AVI based auto-calibration method was developed consisting of a truck matching method and a calibration procedure. WIM PVR records from ARDOT, AVI data from a national truck GPS data provider, and static weight recordings were collected and used to develop and evaluate the proposed method. The AVI auto-calibration procedure produced different FAW and GVW error compared to ARDOT's current method and a more robust but similar method developed by MNDOT. The AVI-based method is adaptable to slight changes in FAW that result from different GVW and loading patterns and thus is able to calibrate GVWs without separating weights into loaded and unloaded bins.

Using a single data provider for AVI data (in this case GPS data) could be considered a limitation of the current methodology since the data may not be representative of all truck industries and cargo types. Although we did not note the cargo configurations of all trucks in the AVI and static weight sample, most trucks were van trailers. This means that calibration factors do not incorporate different trailer types that might have different loading patterns. For example, liquid bulk tanks, livestock, and logging trailers may have very different loading patterns that effect the FAW variation and resulting calibration factors calculated via our proposed auto-calibration algorithm.

In future work, we would like to consider a broader spectrum of AVI data sources such as various GPS and Electronic Logging Device (ELD) providers or license plate matching technology. Another related issue was the size of the AVI data sample. Our sample represented only a very small proportion of the total truck volumes which initially was 121 AVI trucks. With a larger sample, we could compute more accurate likely weights which would potentially increase the accuracy of the calibration factors.
A limitation of this study was the need to generate “groundtruthed” matches to estimate model performance. This was time consuming process due to the low number of AVI trucks that entered the weight enforcement station relative to the total number of trucks that crossed each WIM site during the data collection and the need to manually verify matches using video recordings. In future work, it would be highly beneficial to use license plate readers to automatically match trucks across sites during data collection.

A major factor contributing to the inability to produce accurate FAW and GVW measurement was the quality of the WIM sensors. Although they are maintained adequately, the piezoelectric sensor in the WIM system produces errors as large as 30% (FHWA, 2018 Part 3). With no temperature sensors at the sites to adjust weight measurements in accordance with pavement and ambient temperature changes, it is difficult to produce accurate weight measures, even with the proposed AVI auto-calibration algorithm. Our results show variation in CFs by time of day (Figure 21) indicating the effect of temperature on sensor performance.

Further, piezoelectric sensors have short life spans (2-3 years) but it is likely infeasible due to budget restrictions to replace sensors this frequently. As the sensors degrade, they become more sensitive to weather and pavement conditions. We found that WIM sites along high traffic areas were less accurate than sites with lower traffic for all methods. A solution for this might be to transition the higher volume sites into higher quality scales such as strain gage or bending plate scales and possibly even relocate and drop some sites that experience low volumes in order to adjust the budget for improved WIM systems. Bending plates and load cells have a 6 to 10% error in GVW and 15 to 20% error in axle loads while piezo electric sensors have 15% error associated with GVW and up to 30% error in axle loads.
Figure 21. Lamar and Lonoke per vehicle calibration factors by time of day
Chapter 7. Conclusion

WIM systems are important to track loads and monitor traffic behavior on a transportation network. WIM may provide detailed data being able to record many vehicle characteristics of each vehicle traveling over the sensor. Recording dynamic weights which then must be calibrated to be closer to static weights typically through an auto-calibration method. However, over time WIM sensors tend to develop systematic errors where weight recordings are consistently higher or lower. The proposed AVI method sought to mitigate systematic error using a different approach in using inter site WIM data in computing calibration factors with shared AVI trucks.

WIM systems may come in a variety of different configurations and may implement different types of scales and sensors for data recording. Therefore, the error in data recording capabilities in each different configuration must be noted. There are three types of WIM system widely used, types I, II and III. Type I and III typically employ bending plate or load cell scales which have less error associated in measuring weights while piezo electric sensors usually used in type II systems have a higher error. The traffic detection sensors may also vary using different types of inductive loops or even adding transponders to the WIM system to be able detect a vehicle approaching or leaving the site and track it across the network. Another way to track vehicles across the network is GPS pinpoint data which is utilized in this study as AVI data in order to find trucks at WIM sites.

The study consisted of four types of data. The first data type gathered were WIM PVR data provided by ARDOT containing detailed vehicle characteristics of vehicles traveling over Arkansas WIM sites. The second data type gathered were GPS AVI records provided by Drivewyze which provided the location of sets of trucks that utilized the Drivewyze app at WIM sites. The third data type gathered were traffic video recordings at selected WIM sites. The final
data type gathered were static weights of trucks that stopped at the weigh stations. Together with these data the groundtruth, models, and model testing were carried out.

The AVI auto-calibration method performs truck matching and calibration. The truck matching process relates WIM records to a set of AVI trucks making possible to track these tucks’ weights across WIM sites. Then the these AVI trucks with their corresponding WIM records are used in the calibration part of the model. Calibration factors were computed after a process of discerning how many sites AVI trucks cross, the weight deviation between WIM sites, and weight clustering based on higher volume sites to select the reference steering axle weight for calibration.

The proposed auto-calibration method is an alternative calibration method able to calculate per vehicle calibration factors in order to adjust the overall accuracy at a WIM site. It developed a truck matching algorithm able to replace manual truck matching using video traffic recordings matching trucks at about the same success rate and in a much lower time. It is a resourceful alternative to track trucks without the continuous reliability on visual or sensor aid using only time stamps and the vehicle parameters recorded at different WIM sites. The calibration method can create continuous calibration factors using data from the WIM sites themselves able to change and use the WIM data from other sites to correct themselves creating more uniform and accurate recordings for WIM sites on a network. The proposed method with a larger number of AVI trucks could improve the calibration results of the AVI method and consequently may be used to calibrate raw WIM sites or as a method complementing an existing calibration method for WIM sites.
References


12. FHA. “Verification, Refinement, and Applicability of Long-Term Pavement Performance Vehicle Classification Rules.” U.S. Department of Transportation/Federal Highway


Appendix

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Class 7</th>
<th>Class 8</th>
<th>Class 9</th>
<th>Class 10</th>
<th>Class 11</th>
<th>Class 12</th>
<th>Class 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motorcycles</td>
<td>Four or more axle, single unit</td>
<td>Four or less axle, single trailer</td>
<td>5-Axle tractor semitrailer</td>
<td>Six or more axle, single trailer</td>
<td>Five or less axle, multi trailer</td>
<td>Six axle, multi-trailer</td>
<td>Seven or more axle, multi-trailer</td>
</tr>
<tr>
<td>Class 2</td>
<td>Class 3</td>
<td>Class 4</td>
<td>Class 5</td>
<td>Class 6</td>
<td>Class 7</td>
<td>Class 8</td>
<td>Class 9</td>
</tr>
<tr>
<td>Passenger cars</td>
<td>Four tire, single unit</td>
<td>Buses</td>
<td>Two axle, six tire, single unit</td>
<td>Three axle, single unit</td>
<td>Four or more axle, single unit</td>
<td>Four or less axle, single trailer</td>
<td>5-Axle tractor semitrailer</td>
</tr>
</tbody>
</table>

Figure 22. FHWA vehicle classes 1 thru 13 (FHWA, 2019).
Table 8. Arkansas weight limits for a given axle type by USDOT.

<table>
<thead>
<tr>
<th>Axle Type</th>
<th>Weight Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Axle</td>
<td>20,000 lbs.</td>
</tr>
<tr>
<td>Tandem Axle</td>
<td>34,000 lbs.</td>
</tr>
<tr>
<td>Tridem Axle</td>
<td>50,000 lbs.*</td>
</tr>
<tr>
<td>Gross Weight</td>
<td>Per State weight table</td>
</tr>
<tr>
<td></td>
<td>80,000 lbs.</td>
</tr>
<tr>
<td>Other</td>
<td>Steering axle 20,000 lbs.**</td>
</tr>
<tr>
<td></td>
<td>Tandem-steer axle 24,000 lbs.</td>
</tr>
</tbody>
</table>

* Provided that, within a tri-axle group, no single axle exceeds 18,000 lbs. and no tandem axle group exceeds 32,000 lbs. This number is derived by adding the weight limit for a tandem to the weight limit for a single axle, as specified in Ark. Stat. Ann. §27-35-203(e)(1), consistent with the definition of a tridem given in Ark. Admin. Code §001.01.3-IV.

**12,000–20,000 lbs., depending on the manufacturer’s steering axle weight rating §27-35-203(c)(1).
Table 9. Highway Pavement design table to calculate ESALs

<table>
<thead>
<tr>
<th>Gross Axle Load</th>
<th>Load Equivalency Factors</th>
<th>Gross Axle Load</th>
<th>Load Equivalency Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kN</td>
<td>lb</td>
<td>Single</td>
<td>Tandem</td>
</tr>
<tr>
<td>4.45</td>
<td>1,000</td>
<td>0.00002</td>
<td></td>
</tr>
<tr>
<td>8.9</td>
<td>2,000</td>
<td>0.00018</td>
<td></td>
</tr>
<tr>
<td>17.8</td>
<td>4,000</td>
<td>0.00209</td>
<td></td>
</tr>
<tr>
<td>22.25</td>
<td>5,000</td>
<td>0.00500</td>
<td></td>
</tr>
<tr>
<td>26.7</td>
<td>6,000</td>
<td><strong>0.01043</strong></td>
<td></td>
</tr>
<tr>
<td>35.6</td>
<td>8,000</td>
<td>0.0343</td>
<td></td>
</tr>
<tr>
<td>44.5</td>
<td>10,000</td>
<td>0.0877</td>
<td>0.00688</td>
</tr>
<tr>
<td>53.4</td>
<td>12,000</td>
<td>0.189</td>
<td>0.0144</td>
</tr>
<tr>
<td>62.3</td>
<td>14,000</td>
<td>0.360</td>
<td>0.0270</td>
</tr>
<tr>
<td>66.7</td>
<td>15,000</td>
<td>0.478</td>
<td>0.0360</td>
</tr>
<tr>
<td>71.2</td>
<td>16,000</td>
<td>0.623</td>
<td>0.0472</td>
</tr>
<tr>
<td>80.0</td>
<td>18,000</td>
<td>1.000</td>
<td>0.0773</td>
</tr>
<tr>
<td>89.0</td>
<td>20,000</td>
<td>1.51</td>
<td>0.1206</td>
</tr>
<tr>
<td>97.8</td>
<td>22,000</td>
<td>2.18</td>
<td>0.180</td>
</tr>
<tr>
<td>106.8</td>
<td>24,000</td>
<td>3.03</td>
<td>0.260</td>
</tr>
<tr>
<td>111.2</td>
<td>25,000</td>
<td>3.53</td>
<td>0.308</td>
</tr>
<tr>
<td>115.6</td>
<td>26,000</td>
<td>4.09</td>
<td>0.364</td>
</tr>
<tr>
<td>124.5</td>
<td>28,000</td>
<td>5.39</td>
<td>0.495</td>
</tr>
<tr>
<td>133.5</td>
<td>30,000</td>
<td>6.97</td>
<td>0.658</td>
</tr>
<tr>
<td>142.3</td>
<td>32,000</td>
<td><strong>8.88</strong></td>
<td><strong>0.857</strong></td>
</tr>
<tr>
<td>151.2</td>
<td>34,000</td>
<td>11.18</td>
<td>1.095</td>
</tr>
<tr>
<td>155.7</td>
<td>35,000</td>
<td>12.50</td>
<td>1.23</td>
</tr>
<tr>
<td>160.0</td>
<td>36,000</td>
<td>13.93</td>
<td>1.38</td>
</tr>
<tr>
<td>169.0</td>
<td>38,000</td>
<td>17.20</td>
<td>1.70</td>
</tr>
<tr>
<td>178.0</td>
<td>40,000</td>
<td><strong>21.08</strong></td>
<td><strong>2.08</strong></td>
</tr>
</tbody>
</table>

Note: kN converted to lb are within 0.1 percent of lb shown

Calculation of ESALs in the Introduction Section:

1st calculate the growth factor with:

\[ G_m = \frac{(1+r)^n-1}{r} \quad \text{Eq. 8 (NCEES, 2013)} \]

Where:

\[ n = \text{design life} \]
\[ r = \text{growth rate} \]
Solving for growth factor:

\[ G_m = \frac{(1 + 0.03)^{25} - 1}{0.03} = 35.46 \]

Calculating ESALs using Equation 1:

Load equivalency factors where retrieved from Table 7.

\[ ESAL_{Passenger\ Cars} = 0.9 \times 35.46 \times (0.5 \times 19,000) \times 365 \times 2 \times 0.00018 \]
\[ = 0.0398 \times 10^6 ESALs \]

\[ ESAL_{VC_9\ front\ axle} = 0.9 \times 35.46 \times (0.2 \times 19,000) \times 365 \times 1 \times 0.189 = 8.366 \times 10^6 ESALs \]
\[ ESAL_{VC_9\ tandem\ axle} = 0.9 \times 35.46 \times (0.2 \times 19,000) \times 365 \times 1 \times 1.095 \]
\[ = 193.879 \times 10^6 ESALs \]

\[ ESAL_{VC_9} = ESAL_{VC_9\ front\ axle} + ESAL_{VC_9\ tandem\ axle} = 202.245 \times 10^6 ESALs \]
Flow Chart 4. AVI Calibration method Using GVW.

For each hour of the day, \( t = 1 \ldots 24 \)

1. Get AVI-WIM truck records for specified hour, \( t \)
2. Determine number of WIM sites, \( n \), crossed by each AVI truck
3. Record not used for auto-calibration
4. Set Calibration Factor for WIM sites equal to 1, \( \overline{CF}_i = 1 \)

For each AVI truck corresponding to a ‘3-5’ axle configuration, \( k = 1 \ldots K \)

1. Distinguish between high and low volume sites using lookup table
   - Low Volume Sites, \( l \)
   - High Volume Sites, \( h \)

2. Cluster analysis to determine likely GVW, \( \omega_a \), of truck, \( k \)

3. Calculate the percent of WIM sites for which the difference in GVW weights is \( \leq \pm \delta_S \)
   - \( n = 2^+ \)
   - \( 100\% \)
4. Use weight record at high volume site(s) as likely GVW, \( \omega_a \)
5. Compare likely steering weight \( \omega_a \) to reference steering weight, \( W_R \), if \( \omega_a \) is within the weight deviation, \( \pm \delta_W \), or closer to \( W_R \) use corresponding \( \omega_a \), else use \( W_R \) as \( \omega_a \)

For each WIM site, \( i = 1 \ldots N \)

1. Compute calibration factor for each AVI truck at each site, \( CF_{k,h} = \frac{W_{k,h}}{\Omega_h} \)
2. Apply average calibration factors to all WIM records to estimate calibrated weight, \( \hat{W}_{k,i} = \overline{CF}_i \times W_{k,i} \)
Figure 23. Lamar PVR RAW FAWs where 1 denotes weights recorded at lane 1, the inner lane and 2 denotes weights recorded at lane 2, the outer lane.

Figure 24. Lonoke PVR RAW FAWs where 1 denotes weights recorded at lane 1, the inner lane and 2 denotes weights recorded at lane 2, the outer lane.
Figure 25. Glen Rose PVR ARDOT Calibrated FAWs where 4 denotes weights recorded at lane 4, the inner lane, and 3 denotes weights recorded at lane 3, the outer lane.

Figure 26. Arkadelphia PVR ARDOT Calibrated FAWs where 4 denotes weights recorded at lane 4, the inner lane, and 3 denotes weights recorded at lane 3, the outer lane.
Figure 27. Texarkana PVR ARDOT Calibrated FAWs where 4 denotes weights recorded at lane 4, the inner lane, and 3 denotes weights recorded at lane 3, the outer lane.