Final Report

Viability of Vehicle Length in Estimating Vehicle Classification and Axle Factors

Transportation Pooled Fund Study TPF-5(340), led by Wisconsin Department of Transportation

Pooled Fund Partner State DOTs:
Georgia, Idaho, Illinois, Iowa, Kansas, Minnesota, New York, North Dakota, Ohio, Pennsylvania, Texas and Utah

Prepared by:

Scott Petersen and Erik Minge
SRF Consulting Group, Inc.

Raul Avelar and Sruthi Ashraf
Texas A&M Transportation Institute

October 30, 2017

Contents

Table of Figures

List of Tables

Executive Summary

Single-tube traffic counts are widely used because they are inexpensive and accurate enough to meet the needs of many travel monitoring applications. However, to convert single tube counts from an axle count to annual average daily traffic (AADT) volume, the axle count must be factored. The most common way to factor is to analyze axle-based classification data and determine the ratio of total vehicles to the axle count. Many agencies call this an "axle factor." Such axle-based vehicle classification sites have historically been significantly costlier and require more maintenance than length-based traffic detection methods. Non-intrusive sensors, such as sidefire radar sensors, can cost-effectively collect vehicle length data, which can then be used to determine axle factors.

The primary objectives of this project are to understand the accuracy of axle factors determined from vehicle length and to develop and evaluate methods for converting length data to axle-based classifications.

Methodology

Two sets of data analysis tests were conducted. The first test compared the accuracy of eight proposed methods to estimate axle factors. From this data set, two were selected for a second round of analysis and a robust evaluation of accuracy that considered season, facility type, and whether the site was in an urban or rural location. The research team assembled a multistate dataset using Wisconsin and Long Term Pavement Performance (LTPP) data for comprehensive performance tests of the methods.

Results

Based on the results of the evaluation, the performance of the two recommended methods (Methods 1 and 5) were found to be within the expected limits of performance – within 2 percent error. Axle factor accuracy was found to not vary significantly depending on the road character, facility type, and number of lanes.

Method 1 described in the report is recommended if the intent is to develop axle factors. The axle factor estimates developed using this method ranged from 0 percent to 1 percent error.

The other method that was tested in the second round of testing (Method 5) used algorithms to classify the length data into FHWA vehicle classifications as a precursor to developing the axle factor. This method estimates vehicle classes of about 90 percent of the vehicles correctly. The classes with poor performance are those that generally have very little traffic and thus their error has a small impact on the total result. This method can also be used to generate axle factors and the results were only slightly less accurate than Method 1, from 0 percent to 2 percent error.

While both methods were found to successfully use length data to estimate axle factors, the ultimate goal is to have the methods actively used by the transportation community. A next phase of the project will provide implementation guidance, training and case studies to document the process and encourage agencies to adopt these methods.

Introduction

This project assessed alternative methods to estimate axle factors and vehicle class from length-based data. Accordingly, an initial analysis investigated and compared the performance of various methods to obtain axle factors from vehicle length data, including a subset of methods that estimate axle-based vehicle classifications. A set of eight methods was proposed and evaluated in this initial set of tests. This initial set of tests was carried out by comparing absolute errors in axle factors, and errors in the estimated proportion of vehicles per FHWA class.

The two best performing methods were selected based on the results of the initial analyses. An additional dataset from the nationwide Long Term Pavement Performance Program (LTPP) was prepared for further analyses using the two selected methods. This report first describes the initial analysis which used only Wisconsin data and then describes the further analyses performed using the new set of nationwide LTPP data.

Initial Analysis to Select Test Methods

These tests were conducted over a dataset from 61 Wisconsin (WI) sites with complete information on both vehicle length and axle-based vehicle class. Among these sites, seven typical sites were used to calibrate all methods of analysis. The research team performed two batteries of tests over the rest of the data. The first battery of tests used data from ten other sites with characteristics similar to the selected seven calibration sites (i.e., homogeneous subset). A second set of tests were performed using the data from the remaining 44 sites to evaluate performance on a more heterogeneous dataset. The dataset, provided by the Wisconsin DOT, consisted of 491,156,794 individual records with both vehicle length and vehicle class variables available. Figure 1 shows a histogram of number of records by vehicle length.

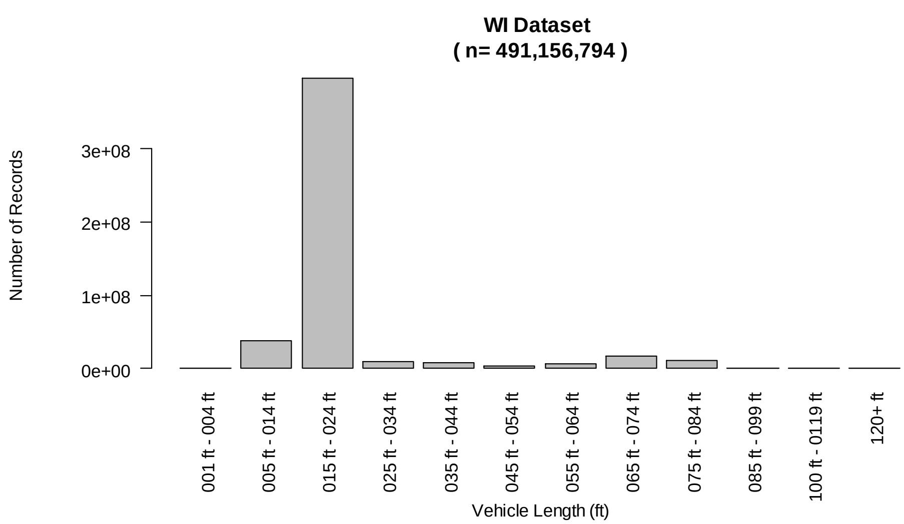


Figure 1. Histogram of Vehicle Lengths in Wisconsin Dataset

The calibration subset consisted of 106,803,612 records from seven sites with typical proportions of vehicles, per a preliminary assessment of the whole dataset. The dataset for the first phase testing on ten typical sites contains 94,472,209 individual records. 289,880,973 records from 44 sites are available for the second phase tests.

Axle Factoring Methods Summary

Eight methods for estimating axle factors from length data were studied. The following diagram illustrates the inputs and outputs of the eight methods.

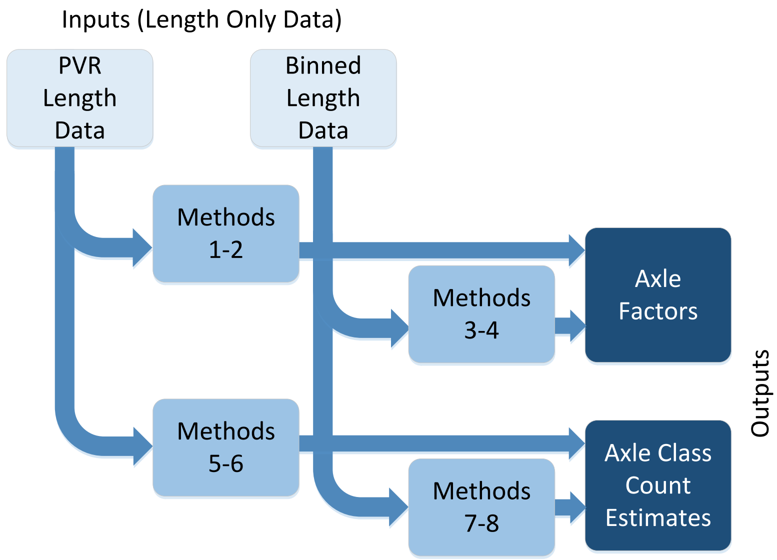


Figure 2. Methods Overview

These eight methods were initially evaluated. Other methods that were evaluated are discussed in Appendix A - Data Collection and Analysis Methods. Methods 1 and 5 were selected for further analysis and are described here.

Method 1 Summary

Use per vehicle axle class data to determine typical numbers of axles per length grouping ("band") and generate an AF.

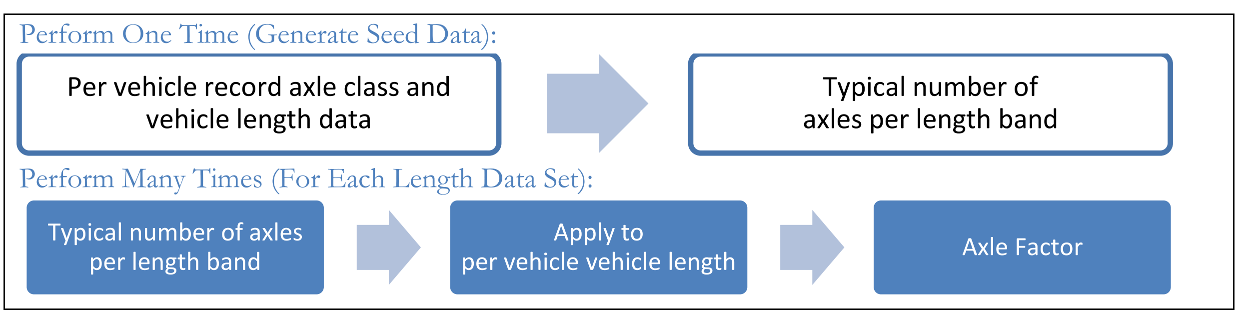


Figure 3. Method 1 Data Flow Chart

- Generate “seed” data from axle class data. Use per vehicle axle class data (with both axle count and length) to determine the average number of axles per length band.
    - Gather per vehicle axle class data with vehicle lengths and axle counts and generate a list of vehicles (sort the list by vehicle length).
    - Determine length “band break points” based on visual inspection of histogram peaks or via an algorithmic method.
    - Determine the average number of axles per length band.
- Gather PVR VL data and determine the vehicle count per length band.
- Apply the average number of axles to vehicle counts per length band and calculate the axle factor.

Method 1 Example

The following example summarizes the computation using Method 1. North Dakota provided both data from an ATR site and a collocated Wavetronix radar sensor. The ATR site reports length, class, and number of axles. The Wavetronix site only reports lengths (and a length class 1 - 4).

North Dakota DOT uses the following length classes:

- Length Class 1 - 1 foot to 6 feet
- Length Class 2 - 7 feet to 29 feet
- Length Class 3 - 30 feet to 44 feet
- Length Class 4 - 45+ feet

The average axles per length class were calculated based on one sample 24-hour period of ATR data (October 11, 2016, a Tuesday outside of the test data range).

Table 1. North Dakota DOT Length Classes

Length Class	Count	Axles	Average Axles
Length Class 1	557	1,119	2.008976661
Length Class 2	70,515	141,841	2.011501099
Length Class 3	1,716	5,559	3.23951049
Length Class 4	5,488	26,550	4.837827988

Two days of Wavetronix data were compiled (October 4 - 5, 2016):

- Length Class 1 Count: 1,192
- Length Class 2 Count: 85,520
- Length Class 3 Count: 3,545
- Length Class 4 Count: 8,454

Then, by multiplying the counts by the average number of axles, an estimate of the total axles is computed to be 226,801.

The estimated axle factor from Wavetronix data is 0.435.

The axle factor determined directly from the ATR data is 0.451.

Method 5 Summary

Use per vehicle axle class data to classify the per vehicle length data into axle classes with a parameterized function.

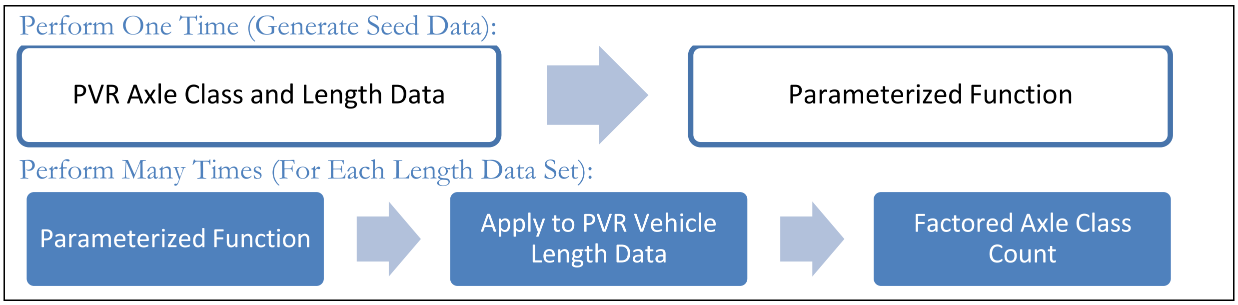


Figure 4. Method 5 Data Flow Chart

- Use per vehicle axle class data (axle class and length) to generate seed data (parameterized function relationship between length and axle class).
    - Input data into algorithm to generate the parameterized function.
- Gather and input per vehicle length into the algorithm to apply the parameterized function.
    - The algorithm determines the vehicle count per axle class.

Method 5 Example: Modeling the Observed Distribution of Lengths as a Mixed Distribution of Vehicle Classes

In simple terms, the method of analysis takes the observed distribution of lengths and produces a combination of known vehicle classes that resembles it. For each individual vehicle class, the length distributions are known in detail (average, min, max, standard deviation, distribution shape, etc.) and are condensed in a set of probability density functions. The algorithm then constructs the corresponding mixed distribution of lengths varying the proportions of each vehicle classes. The algorithm takes as inputs the observed distribution from a site, the calibrated set of functions for known vehicle-classes, and a length-based dataset. The output of the algorithm is the set of most-likely proportions of each class, given observed distribution of lengths at that site (this is shown in the two schematics below). However, some recalibration of the class distribution functions will be necessary from time to time, so that the methodology remains representative of the most-current fleet of cars.

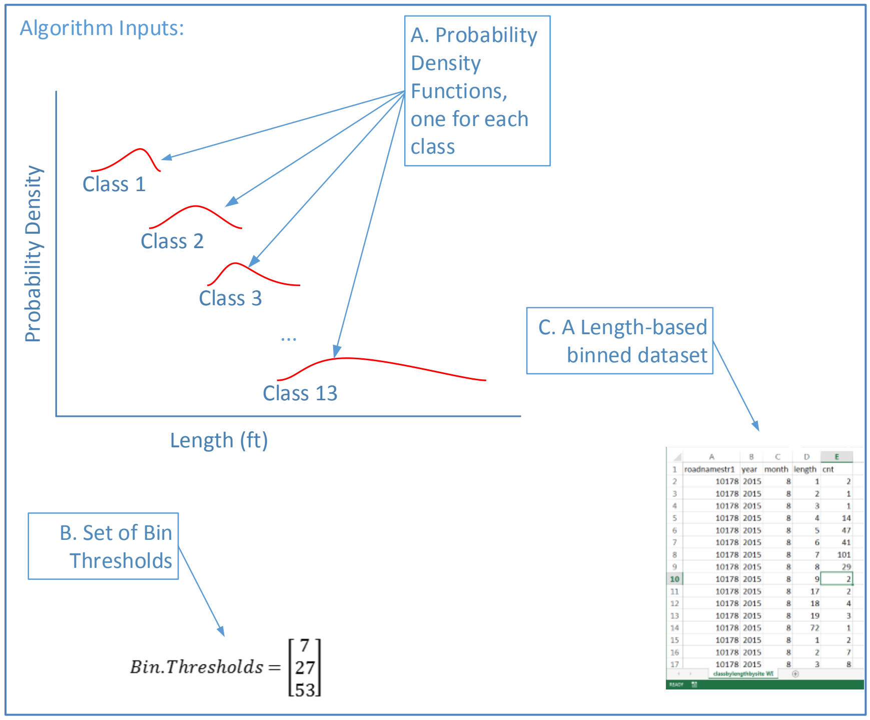


Figure 5. Method 5 Process Diagram

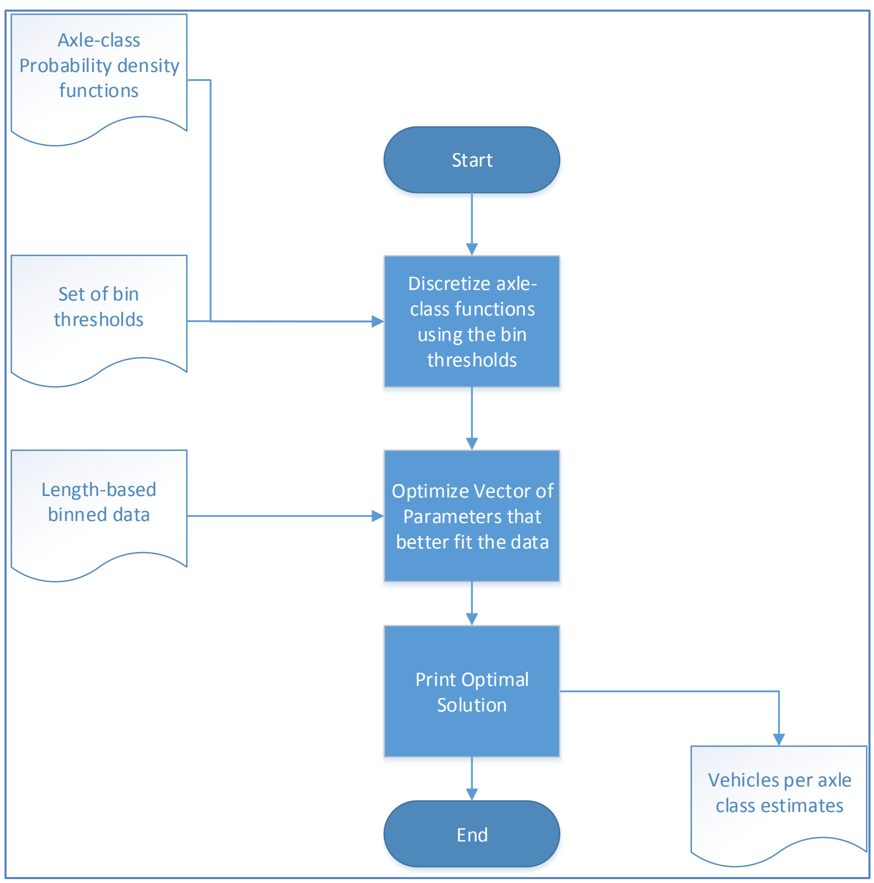


Figure 6. Method 5 Process Flow Chart

Example:

The following table shows the results from this algorithm for data from site ID 116051 from the Wisconsin dataset. This shows that the method estimated many classes within a low percentage error including the most common class 2 and class 3. However, there was an approximately 6 percent error on class 2 which scaled up to a larger error on class 9.

Table 2. Real and Estimated Vehicle Counts by Class

Real Vehicle Category	Real Count	Estimated
1	12823	16070
2	1276511	1390023
3	463205	268677
4	10407	47449
5	54853	42841
6	13163	28642
7	5431	15036
8	29159	52059
9	90377	90602
10	2946	6355
11	189	7377
12	191	863
13	370	260
14	6926	506
Total	1966551	1966760

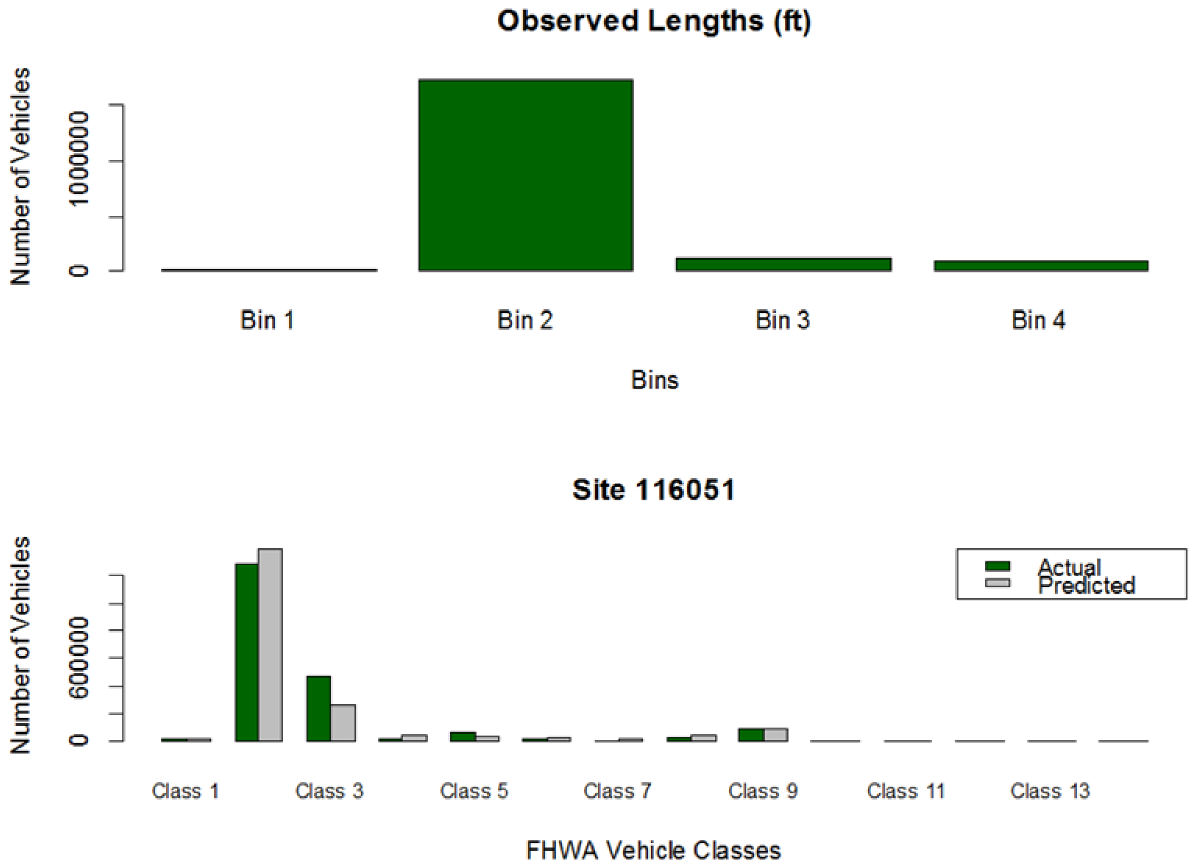


Figure 7. Number of Vehicles by Length and Class

Comparison by Axle Factors on Homogeneous Data

Figure 8 below shows the performance of each method when estimating axle factors from length-based data.

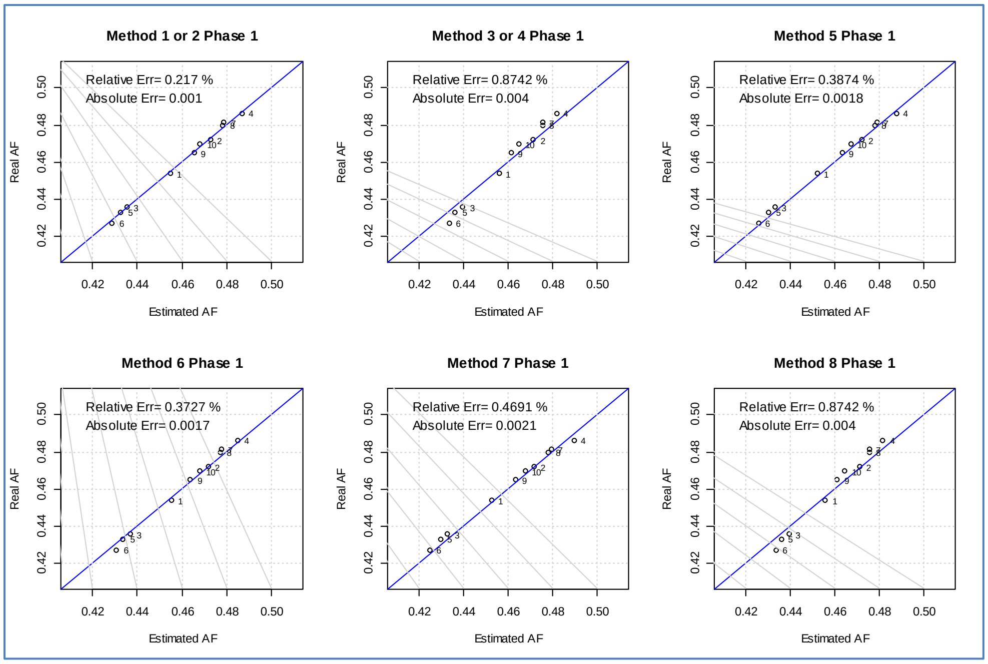


Figure 8. Axle Factor Estimation Performance

All methods performed with a relative error smaller than 1 percent, or equivalently, all performed with absolute errors at the third decimal place. All methods were found to be accurate relative to practical considerations for data collection and analysis.

Comparison by Axle Factors on Heterogeneous Data

Figure 9 shows the performance of each method when estimating axle factors from length-based data (either binned or banded, depending on the method) using heterogeneous data.

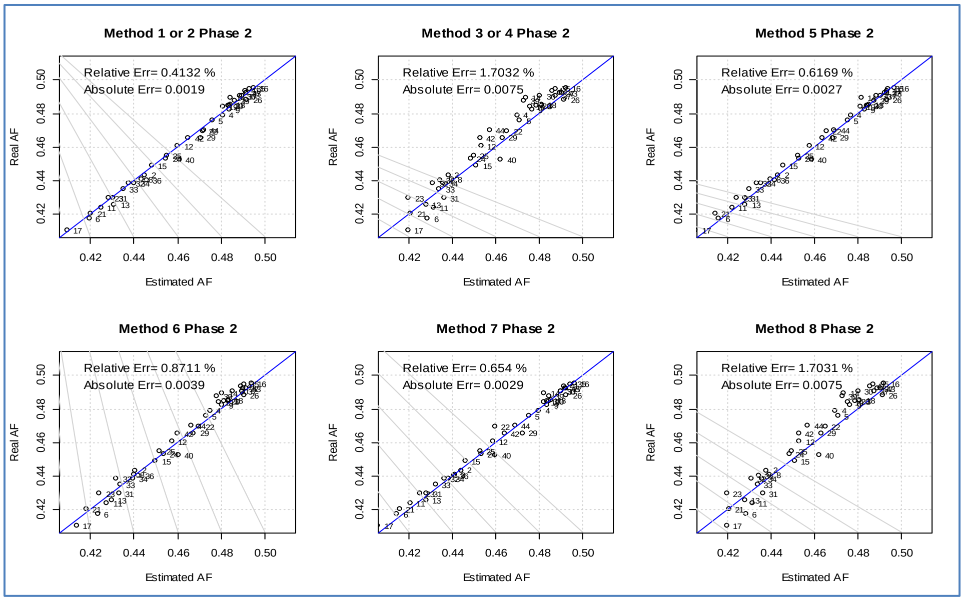


Figure 9. Axle Factor Estimation Performance

Figure 9 indicates that, except for Methods 3, 4 and 8, all methods performed with relative errors smaller than 1 percent. In contrast, Methods 3, 4 and 8 had relative errors of about 1.7 percent.

Comparison by Proportions of Vehicle Classes on Homogeneous Data

The research team next compared the performance of the set of methods that produce estimates of the number of vehicles per class. Results of these comparisons are shown in Figure 10.

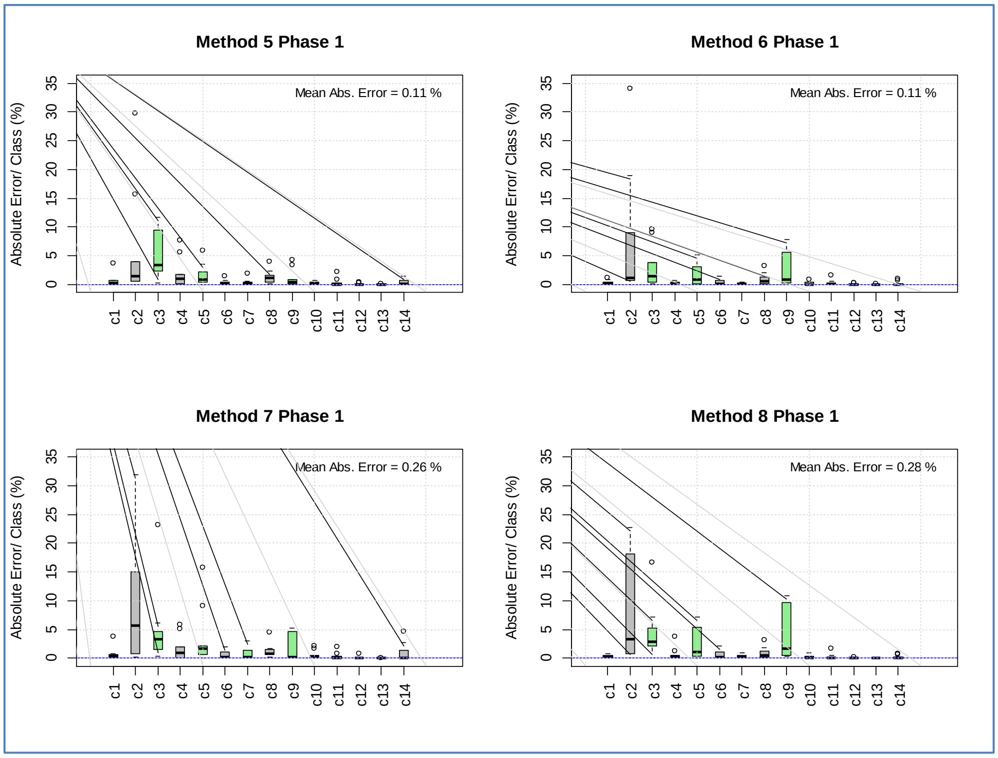


Figure 10. Absolute Error in Vehicle Classification (Methods 5 through 8)

Figure 10 presents the comparison of method performance by their absolute error (i.e., error expressed as percent of the total number of vehicles per site). This figure shows that, on average, all methods performed within 0.3 percent absolute error. The two best performing methods are Methods 5 and 6, with an average absolute error of 0.11 percent. These two methods rely on bins selected by considering seed data. Such bins that were defined based on seed data and not predefined and input in the classifier are referred to as "bands" later in this report. In contrast, the errors from the two methods that use bins was about twice as large, i.e., 0.26 percent mean absolute error.

Comparison by Proportions of Vehicle Classes on Heterogeneous Data

Compared to Figure 10, Figure 11 shows a better performance using heterogeneous data in terms of absolute error: all methods performed within 0.16 percent absolute error.

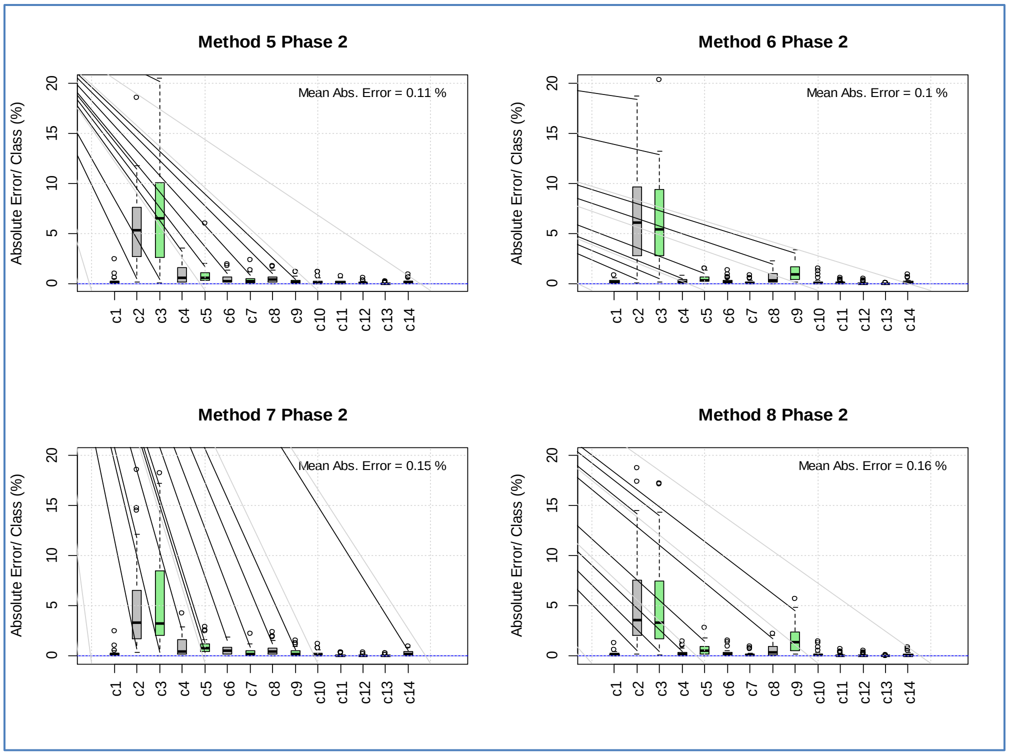


Figure 11. Absolute Error in Vehicle Classification for Heterogeneous Data (Methods 5 through 8)

The two best performing methods are again Methods 5 and 6, with an average absolute error of 0.11 percent. However, the sizes of the boxes in Figure 11 indicate that Methods 5 and 7 had smaller absolute errors in recognizing the proportions of Classes 2 and 9.

The Effect of Heterogeneous Data

A final set of assessments was performed by comparing the changes in errors between the sets of analysis when the only difference is the heterogeneity of the data. These results are of interest because they provide an estimate of the effect of using a more heterogeneous dataset with each method so that it may help anticipate the performance on a larger data set (as presented later, in the third part of this report). For this purpose, the research team computed the percent changes in various error estimates for each method. This assessment in shown in Table 3.

Table 3. Ratios in Absolute Error from Homogeneous to Heterogeneous Data

Method	Homogeneous / Heterogeneous Absolute Error Ratio (for Axle Factors)	Homogeneous / Heterogeneous Absolute Error Ratio (for Class Proportions)
1	1.90	NA
3	1.95	NA
5	**1.66**	**4.32**
6	2.34	8.06
7	**1.49**	**4.77**
8	1.95	6.78

This last set of comparisons indicates that Methods 5 and 7 (i.e., parametric methods) are most robust to data heterogeneity when estimating both axle factors and proportions of vehicle class. In any case, it should be noted that all the magnitudes of the errors in these comparisons are rather small.

Summary of Findings in the Initial Analysis

The following points summarize the findings about performance in estimating axle factors:

- Considering only the magnitude of the absolute errors, any of the methods seems appropriate. Among all methods, Method 1 is recommended for implementation if the desired outcome is to have axle-factor-estimate accuracy within the second decimal or within 1 percent of the real axle factor per site.

- Regarding axle factor estimation, all six methods evaluated on homogeneous data performed within 1 percent error when estimating axle factors. In fact, all methods performed below 0.5 percent error except for Methods 3, 4, and 8, whose performance averaged about 0.87 percent error.

- When using heterogeneous data, the performance in axle factor estimation was generally worse. All methods experienced an increase in errors with respect to their errors when using homogeneous data (per Table 9), but remained within acceptable boundaries (per Figure 10 and Figure 11). However, this deterioration in performance was expected as heterogeneous data will generally pose a steeper challenge to the competing methods. Methods 5 and 7 proved the most resilient to the increase in heterogeneity of the data.

- Comparing the performance of Methods 1 and 3 on one hand, and Methods 5 and 7 on the other, non-parametric estimation of axle classes might be especially sensitive to the choice of bins over bands (or alternatively, to varying the band specification); while parametric estimation seems robust against that particular choice. This finding is supported at both Phase 1 and Phase 2 evaluations.

The following points summarize the findings of this investigation regarding the methods that estimate the proportions of vehicle classes:

- The methods under evaluation performed comparably when using homogeneous data but with significant differences when using heterogeneous data.
- Comparing the change in relative error in the proportion of vehicles per class, when using heterogeneous data to the corresponding estimates from homogeneous data, Methods 5 and 7 again showed significantly more resilience to data heterogeneity (per Table 3).
- Method 5 is recommended for implementation or further analysis, because of its resilience to data heterogeneity and performance. Method 1 is recommended for its simplicity while maintaining generally accurate results.

Data Analysis Including LTPP sites

Based on the results of the tests with the Wisconsin dataset, Methods 1 (or 2) and 5 were recommended, given their ease of implementation and high accuracy. On the one hand, Methods 1 or 2 performed the best in estimating axle factors and were recommended if there is no need to estimate axle classes.

Method 5 was recommended in case axle classes are required, as it showed robustness against both the choice of bands or bins as well as the heterogeneity in the data. This method produced: (1) smaller relative errors overall; (2) smaller absolute errors for two of the key classes of vehicles: classes 2 and 9; and (3) robustness against heterogeneity in the data compared to other competing methods that also estimate vehicle classes.

This part of the report documents further analyses using Methods 1 and 5 to determine the impact of key variables in axle factor and vehicle class estimation. An additional dataset of 19 sites from LTPP was identified for further analysis. However, when PVR data were obtained for these sites, one site from Delaware and one site from Maine were found to not have data in the additional subset and thus were excluded from analysis. The final analyses were performed using the aggregated dataset comprised of the 61 sites from Wisconsin and 17 sites from the LTPP. The geographic distribution of the final dataset is shown in Figure 12.

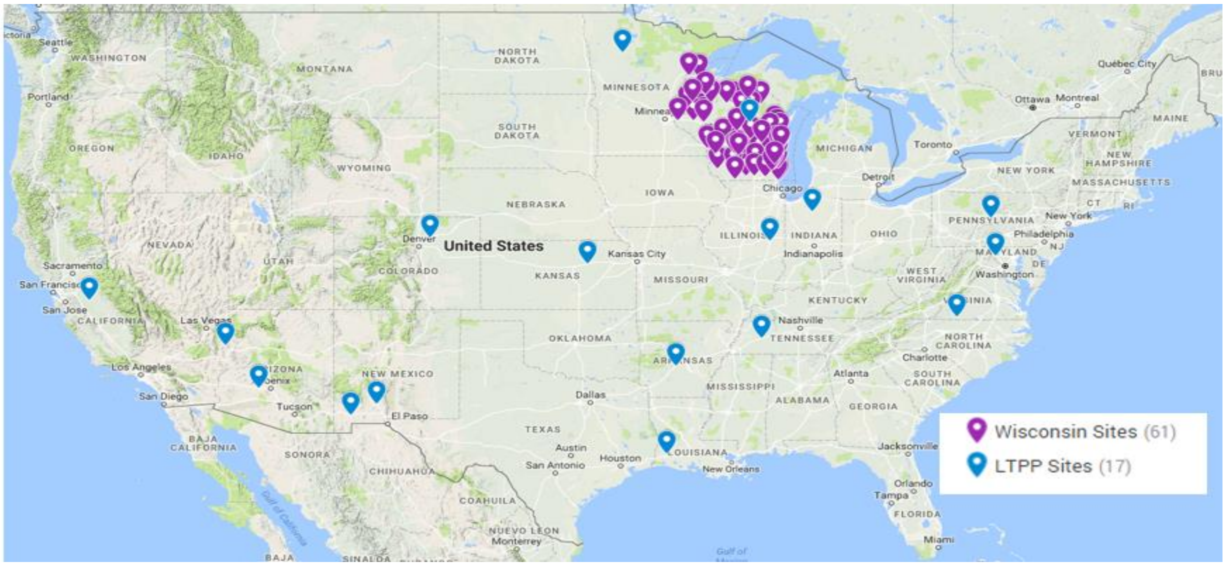


Figure 12. Geographic Distribution of Locations in the Final Dataset

As shown in this figure, the additional LTPP sites are located in multiple states across the country. These sites provided a more heterogeneous mix of traffic. In addition to the length based and axle-based vehicle counts, the research team manually collected data on vehicle length, road character (rural or urban), type of facility (freeway or highway), and number of lanes (2, 4, or 6). The term freeway is defined as a limited access highway and the term highway is defined as other highways. The basic contingency table showing the distribution of the final dataset is presented in Table 4.

Table 4. Contingency Table with Sites Characteristics

Source	Road Character		Type of Facility		Cross Section		
	Urban	Rural	Highway	Freeway	Two Lanes	Four Lanes	Six Lanes
Wisconsin	16	45	43	18	31	26	4
LTPP	0	17	9	8	1	16	0

It can be seen from this table that the data consists mostly of rural Wisconsin sites (45 sites) when classified by source and road character; Wisconsin highways are most prominent when source and type of facilities are considered; and Wisconsin two-lane sites when dividing the data by source and cross section. Regarding other categories, they are represented by at least eight sites, except for the following categories:

- Urban LTPP sites and LTPP sites with six lanes (no sites)
- LTPP sites with two lanes (only one site)
- Wisconsin sites with six lanes (only four sites)

Seasonal Evaluation

One evaluation of interest of this research regards the differences in performance by season of the year. In discussions with other team members, it was determined that the following breakdown of the data would serve the purposes of the seasonal evaluation:

- Data from July and August was analyzed to evaluate summer performance.
- Data from October and November was analyzed to evaluate fall performance.
- Data from January and February was analyzed to evaluate winter performance.

Performance Metrics

The research team selected the metrics in Table 5 to compare performance of the final analytical methods. In the interest of succinctness, these metrics defer from the ones in the rounds of analysis using Wisconsin data only.

Table 5. Metrics of Performance

Performance Measure	Metric	Description	Units of Analysis
Axle Factor	Relative Error	The average of the deviation of the estimated Axle Factor per unit of analysis (as a proportion) relative to the true Axle Factor	Per site, per test
	Absolute Error	The average of the absolute deviation of the estimated Axle Factor per unit of analysis (in axle factor units)	Per site, per test
Proportion of Vehicles per Class	Proportion of Vehicles Misclassified	The ratio of the cumulative deviation of the estimated Vehicle Class proportion per unit of analysis (in number of vehicles) to the total number of vehicles	Per site, per test

All error metrics and misclassified vehicles are given as proportions (i.e., between 0 and 1, corresponding to 0 percent and 100 percent) in the following evaluations.

Roadway Characteristics and Axle Factor

In order to assess the variability present in the dataset, the research team computed the real axle factors from the axle based data and prepared comparisons by road character and type of facility.

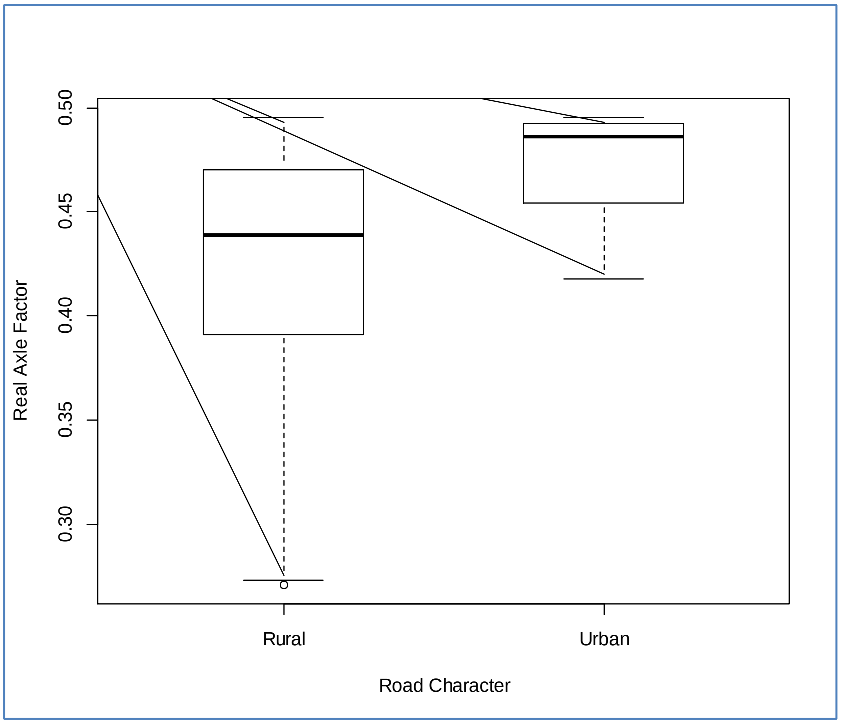


Figure 13. Axle Factor Based on Road Character

Figure 13 shows that the real axle factor distribution along urban roads is skewed toward small values and has a slightly higher median value than the distribution of real axle factor at rural roads (0.49 and 0.44).

Highways tend to have a compact distribution of axle factors with larger values, compared to freeways as indicated in Figure 14. Both distributions tend to be skewed toward lower values, but the range of values at freeway facilities is clearly wider.

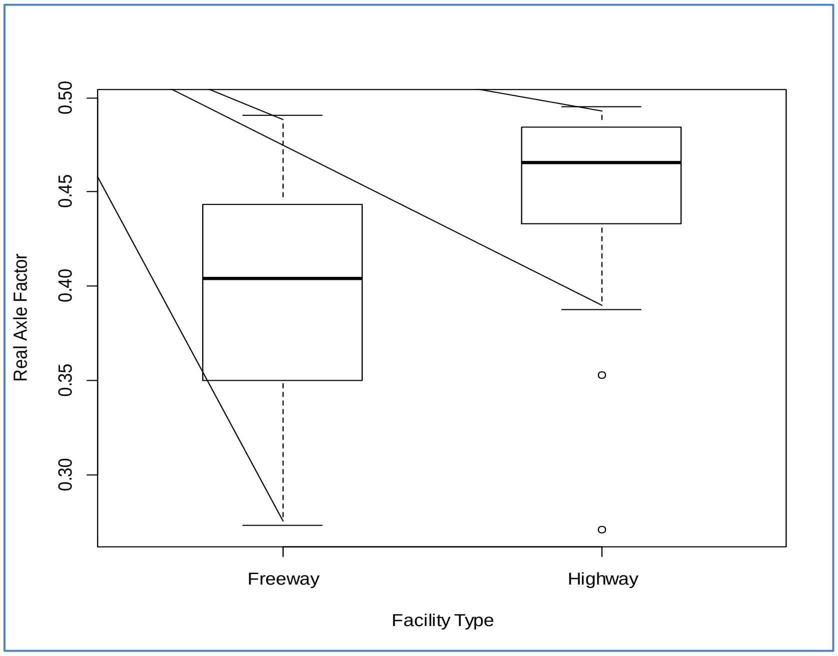


Figure 14. Axle Factor Based on Type of Facility

Two-lane roads show axle factors with higher values, compared to four-lane roads, where the values tend to be lower and spanning a wider range. However, caution is advised when interpreting the trend change from four- to six-lane sites in Figure 15. This trend is most likely meaningless, given the limited number of six-lane sites in the dataset (four sites, per Table 4).

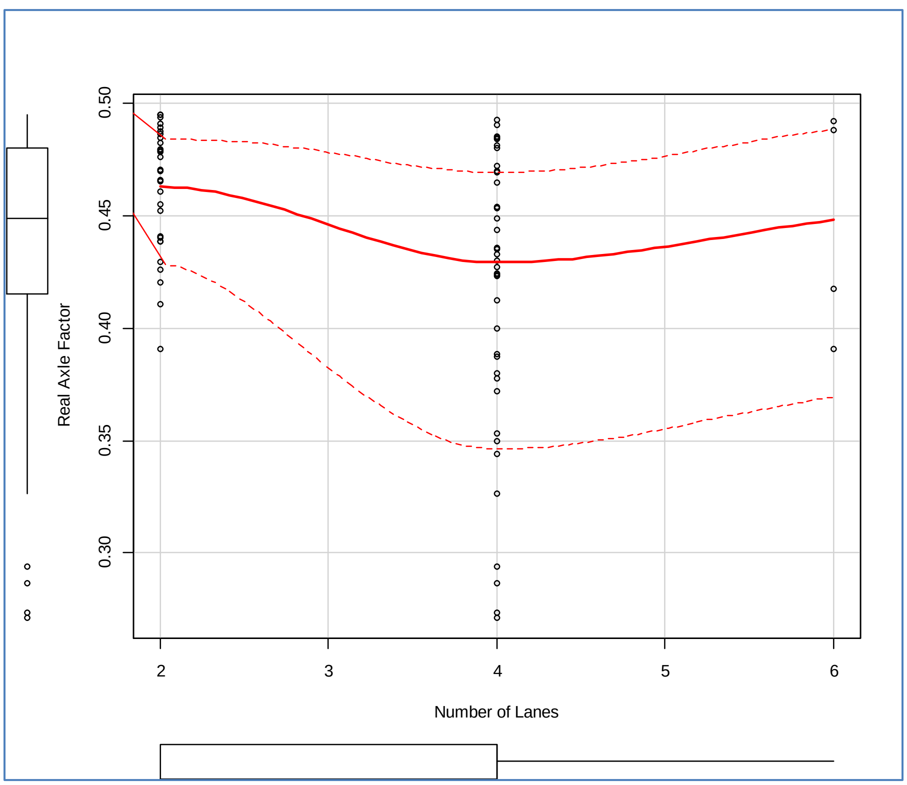


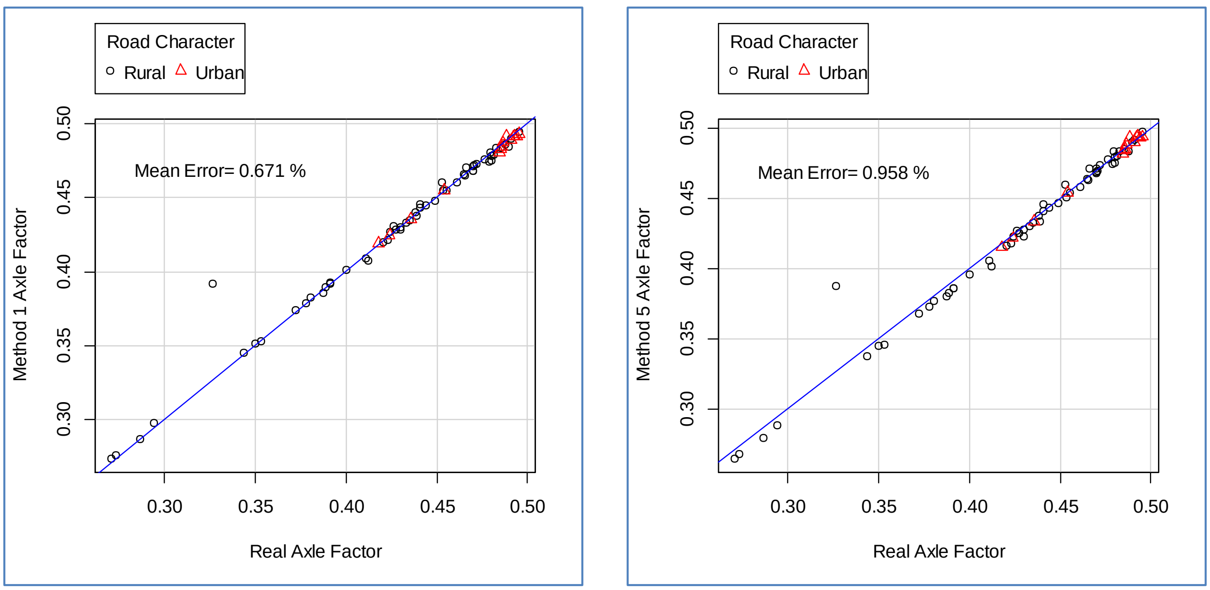
Figure 15. Axle Factor Based on Number of Lanes

Summary of Initial Data Assessment

The examination of the combined dataset showed a wide range of conditions to test the performance of the methods selected for further analysis. The bulk of the data tend to represent rural, two-lane and four-lane highways in the state of Wisconsin. However, Table 5 shows that a variety of other conditions are also represented with few exceptions.

Performance in Axle Factor Estimation

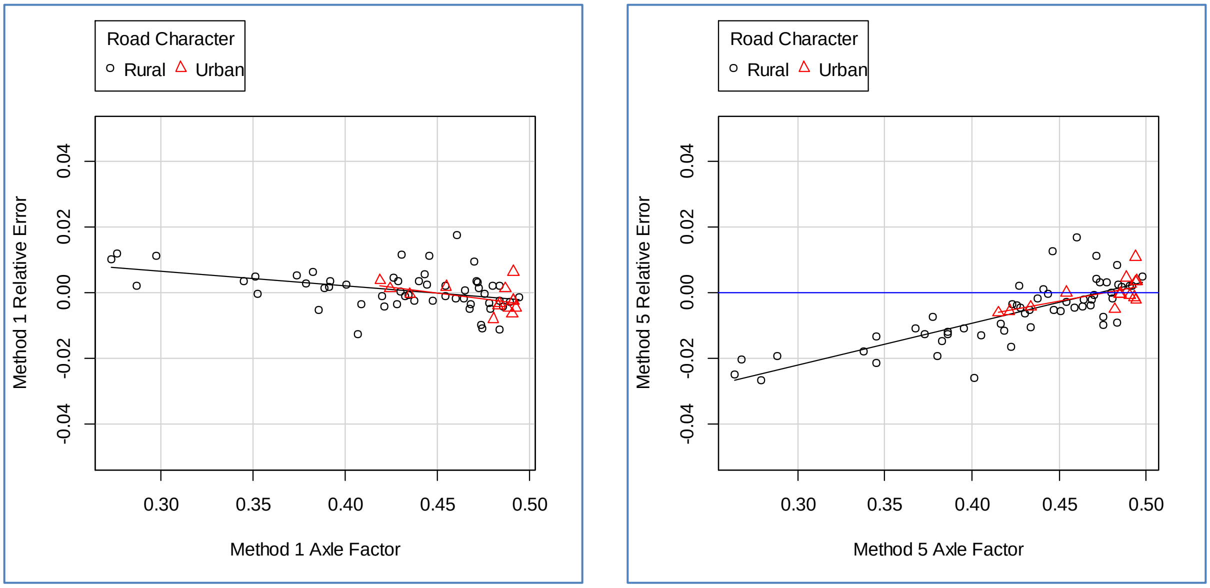
The research team computed axle factors using both Methods 1 and 5 to be compared to actual axle factor values. A graphic representation of such comparison is shown in Figure 16. Consistent with the first round of analysis, Method 5 produced better results with a mean error of 0.67 percent compared to 0.958 percent for Method 1. Both methods were found to produce axle factors within one percent of the baseline and this is within practical considerations for data collection.



(a) Method 1 (b) Method 5

Figure 16. Comparison of Predicted and Actual Axle Factors by Method 1 and Method 5

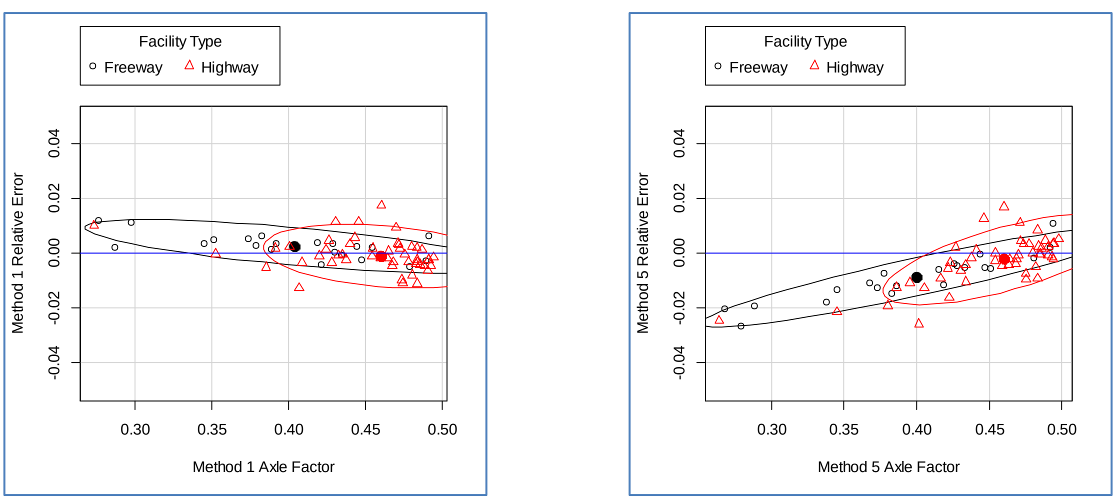
A closer look at the error produced by both methods across the range of axle factor values is shown in Figure 17. As part (a) of this figure shows, Method 1 tends to slightly overestimate axle factor values when the values are smaller. Such an overestimation is reduced as the axle factor approaches 0.5. This is true for both rural and urban sites. In contrast, part (b) of this figure shows that Method 5 tends to underestimate axle factors in the low end of the values, with such underestimation being reduced as the axle factor approaches 0.5. In any case, errors for either of the estimation methods are within 2 percent of the actual value, with only four sites where the error was between 2 and 3 percent when Method 5 is used.



(a) Method 1 (b) Method 5

Figure 17. Error in Axle Factor Prediction by Method 1 and Method 5 Based on Road Character

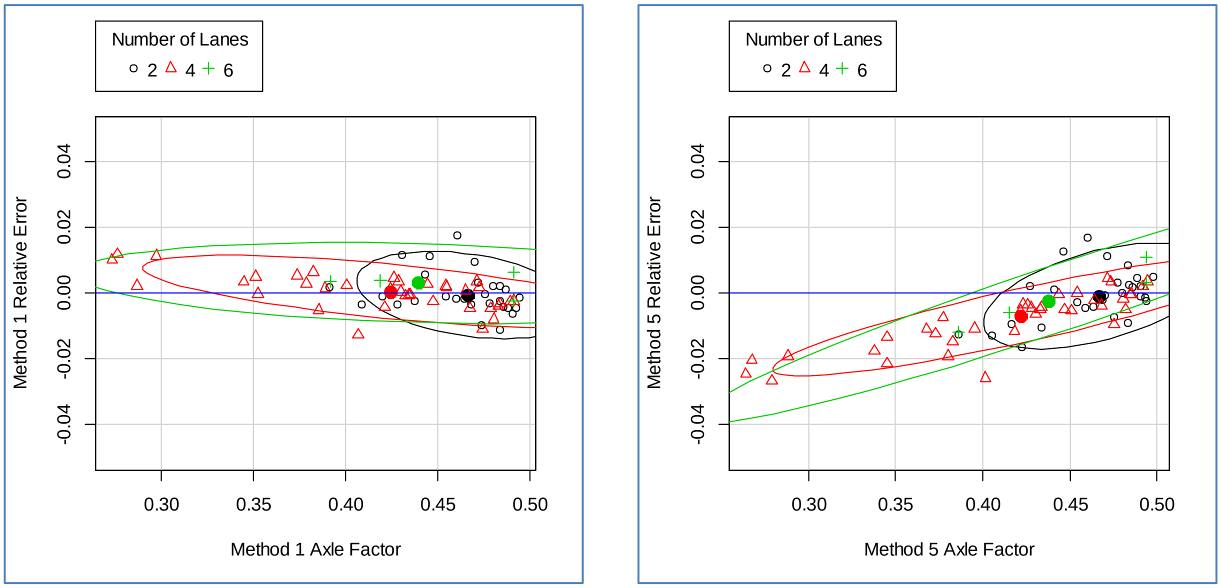
The research team setup another set of comparisons for axle factors at freeways and at highways as shown in Figure 18. This figure shows a 95 percent confidence ellipse for each trend in order to help visualize spread. The trends for the errors when breaking down the data this way remain unchanged: part (a) of this figure shows that axle factors predicted by Method 1 are within 2 percent error and a trend to slightly overestimate appears at low axle factor values. This observation is applicable for both freeway and highway axle factors. Similar to Figure 17, part (b) of this figure shows that Method 5 tends to underestimate axle factors in the low end of the values, an observation that is applicable for both freeways and highways. Again, errors for either estimation method are within 2 percent of the actual value, with only four sites where the error was between 2 and 3 percent using Method 5.



(a) Method 1 (b) Method 5

Figure 18. Error in Axle Factor Prediction by Method 1 and Method 5 Based on Type of Facility

The research team developed an additional set of comparisons of the errors produced by each method, this time comparing performance by the number of lanes in each site. The results of these comparisons are shown in Figure 19.



(a) Method 1 (b) Method 5

Figure 19. Error in Axle Factor Prediction by Method 1 and Method 5 Based on Number of Lanes

Figure 19 depicts the error in axle factor prediction based on number of lanes. The trends are unchanged from the previous two figures. Regardless of a site having two, four, or six lanes the following observations remain:

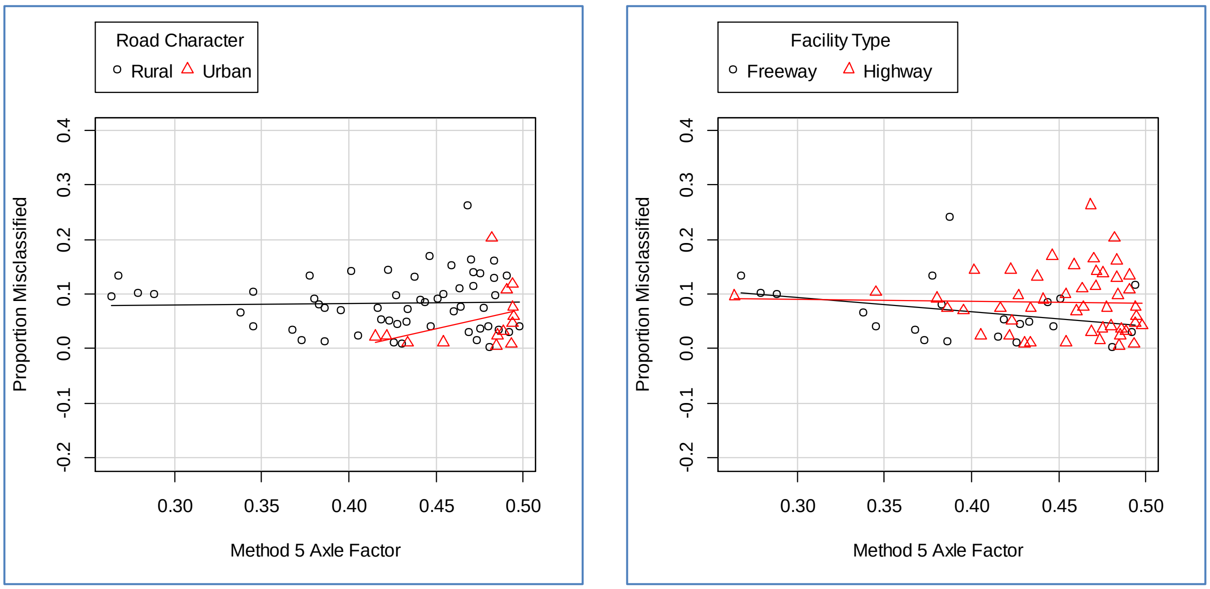
- Method 1 tends to slightly overestimate the true axle factor.
- Method 5 tends to slightly underestimate the true axle factor.

Even with this trend of slight bias, the predicted axle factor remains within 2 percent of the true axle factor value. The next section shows a set of tests on the performance of Method 5 concerning the prediction of vehicle classes.

Prediction of Vehicle Classes

The research team performed further analyses on Method 5 regarding its ability to recognize vehicle classes correctly. To perform this assessment, the research team determined the proportion of vehicles misclassified. It should be noted that classifying vehicles is, in essence, a zero-sum procedure (i.e., the total of all classified vehicles should add up to the total number of vehicles at a site). That being so, the net number of misclassifications is simply half the sum of absolute deviations from all classes. The assumption of this estimation procedure is that a misclassified vehicle should produce a negative deviation in the count of the correct class and a positive deviation in the count of the incorrect class.

It should be noted that it is still possible that misclassifications among similar classes cancel each other in such a way that they do not produce any deviations between estimated and real counts, though such scenario is very unlikely and, in any case, it should have no impact on the performance of the classification. Even if equal-size sets of individual vehicles of different class happen to be counted under one another's class (such that no deviations occur between the estimates and real proportions), the goal of correctly estimating the counts (or proportions) of vehicles of each class for the site would be achieved.



(a) Based on Road Character (b) Based on Type of Facility

Figure 20. Percent of Vehicles Misclassified by Method 5

Using the proportion of misclassifications as a performance metric, Figure 20 shows how such proportion compares to the axle factor by road character and facility type. In general, it can be seen that the proportion of misclassifications tend to be flat at around 0.1 (or about 10 percent), regardless of the axle factor value. It is interesting to notice smaller proportion of misclassified vehicles at urban sites and at freeway sites to a lesser extent.

The research team performed another set of tests to assess the performance of Method 5. In this battery of tests, the real and the estimated proportions of vehicles of each class were plotted against each other. Figure 21 shows these comparisons.

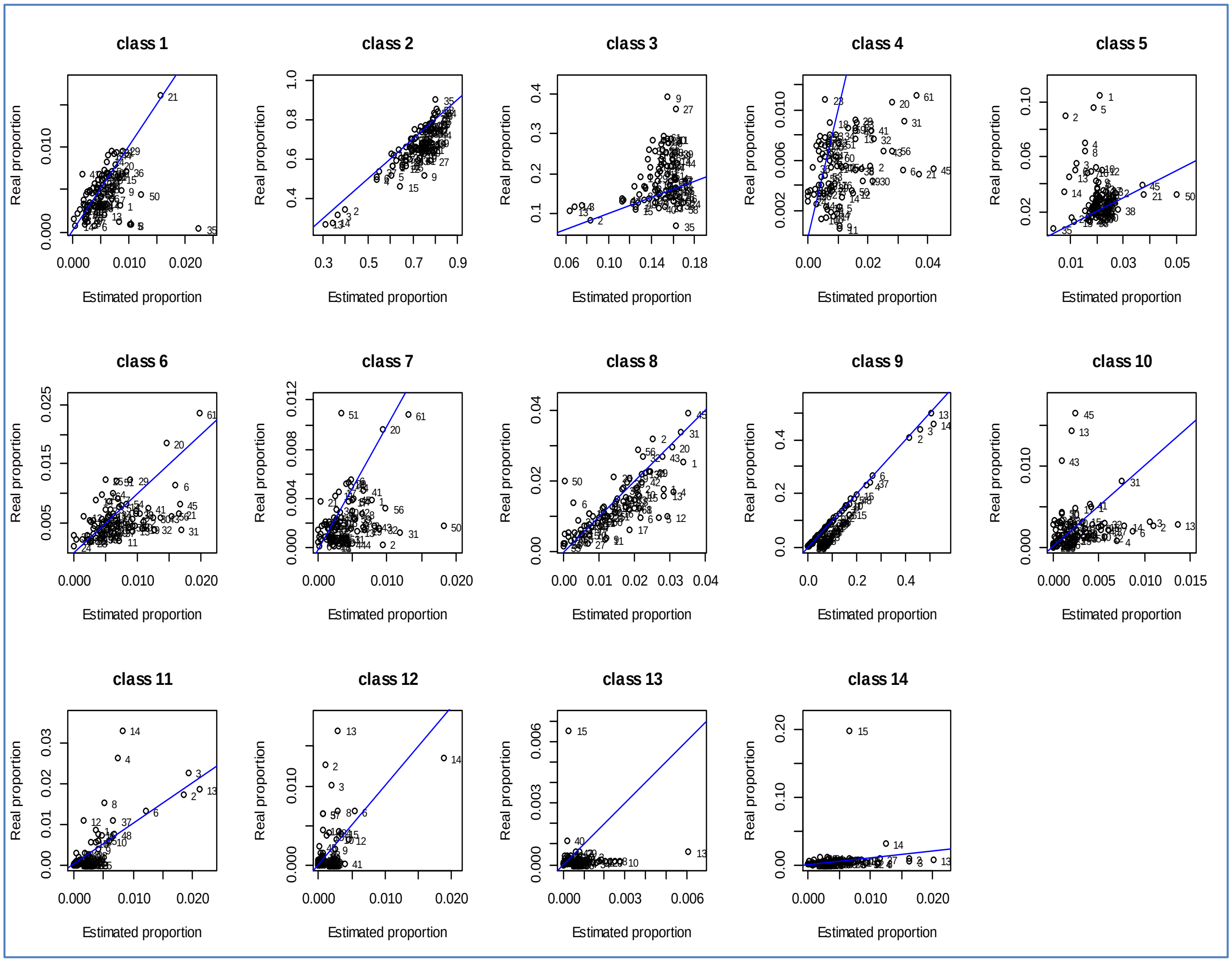


Figure 21. Real and Estimated Proportion of Vehicle Classes for Wisconsin and LTPP Dataset

As this figure shows, accuracy of classification is highest for classes 1, 2, 8, 9 and for the vehicles of unknown class. The most precise classifications occurred for class 9 and vehicles of unknown class. The precision in proportion of vehicles of unknown class is not surprising, given that these vehicles are most likely class 9 vehicles with a long gap between axles.

Figure 22 shows the performance of Method 5 using only sites from Wisconsin. The performance of recognizing classes 1, 8, and 9 are the highest. The classes that seem to have less precision are classes 4, 13, and vehicles of unknown class.

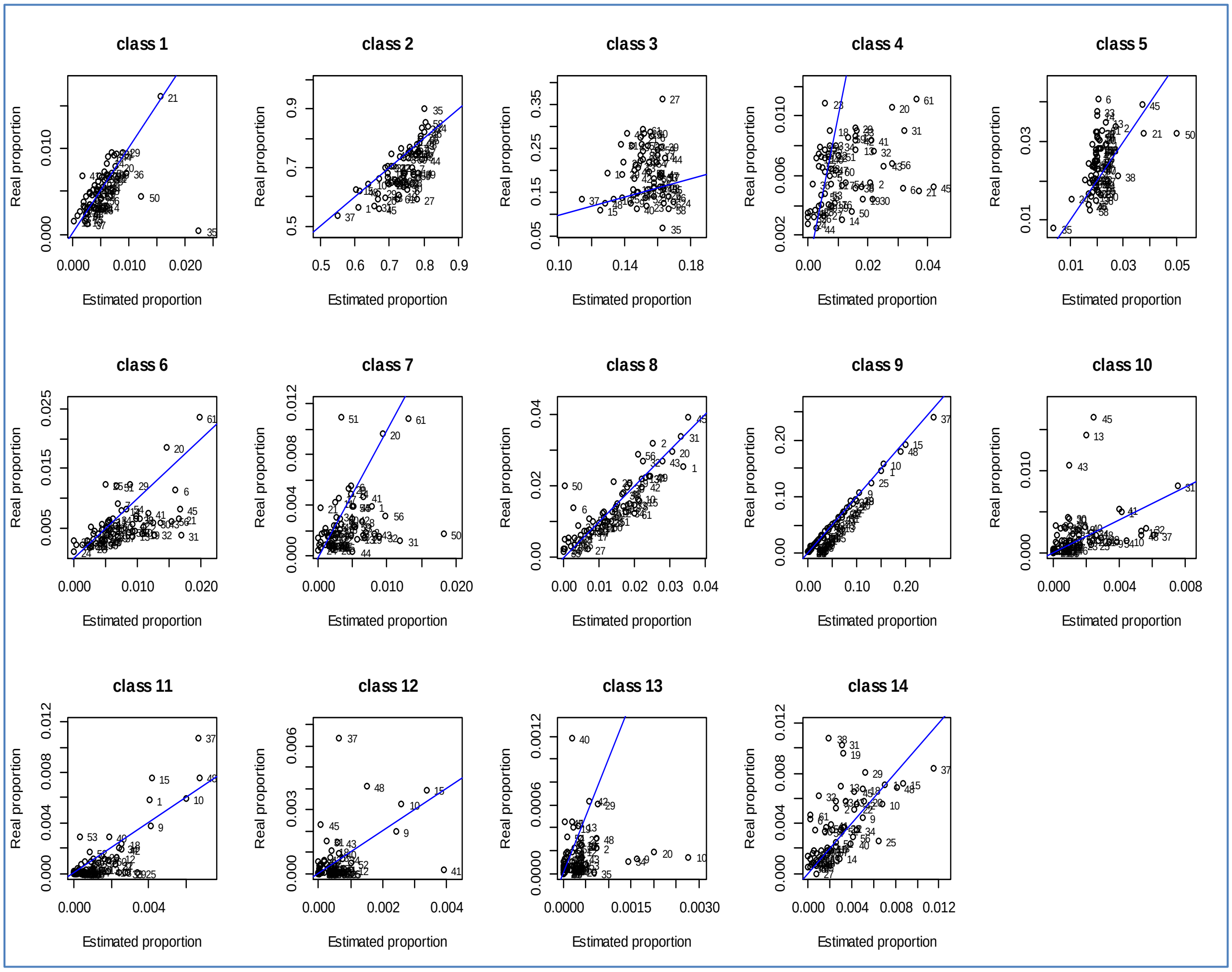


Figure 22. Real and Estimated Proportion of Vehicle Classes for Wisconsin Dataset

Finally, Figure 23 shows the performance of vehicle classification on the LTPP data only. Again, the performance of estimating the proportion of classes 2 and 9 vehicles are the best, but performance was marginal for almost every other class. The research team speculates that this decreased performance is the result of differences between the Wisconsin and LTPP data sets, and that performance would improve significantly if the algorithm was recalibrated using LTPP data as well as Wisconsin data.

Seasonal Variability in Performance

As mentioned earlier in this report, data were further subdivided by fall (months of October and November), winter (January and February) and summer (July and August) to examine seasonal differences in performance. The research team applied both Methods 1 and 5 to each subset of data and developed the corresponding statistics for each season by data source and facility type.

Seasonal Performance of Method 1 for Axle Factor Estimation by Data Source

The axle factor values were found generally higher for Wisconsin sites (compared to LTPP) for all seasons with distributions closer to 0.50. This difference probably results from the fact that LTPP sites are only rural and about half of them are freeways, whereas only 74 percent of Wisconsin sites are rural roadways with a majority of sites (73 percent) being highways (see Table 4).

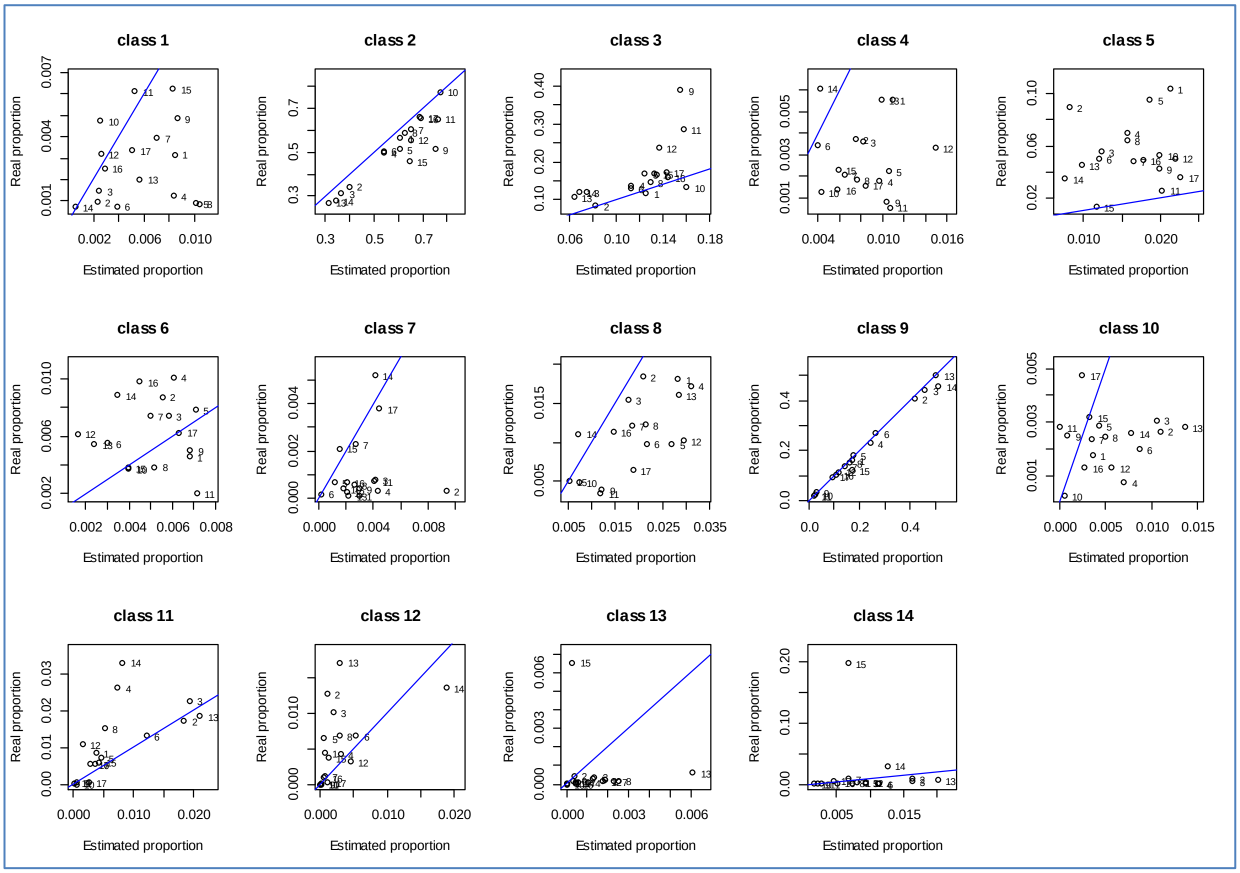


Figure 23. Real and Estimated Proportion of Vehicle Classes for LTPP Dataset

As shown in Figure 24, axle factor values during summer and fall are very similar, but during winter, axle factors tend to be more dispersed for both datasets, especially for LTPP sites.

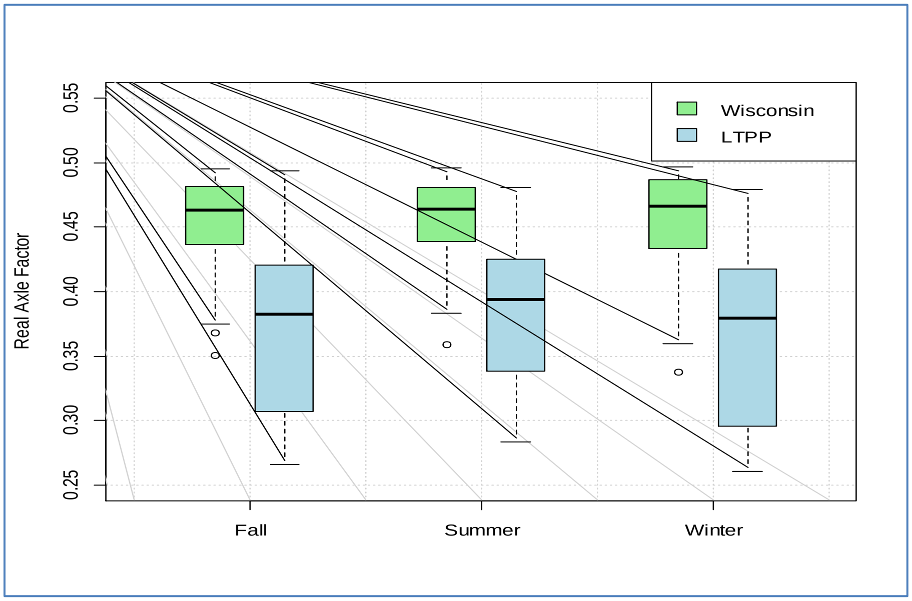


Figure 24. Axle Factor by Source and Season

Figure 25 shows that proportional error in axle factor estimation from Method 1 by season and data source. Interestingly, this plot shows that Method 1 is essentially unbiased in estimating axle factors for both data sources during fall and summer, except perhaps for LTPP sites during winter, where the error seem to be most likely positive. In other words, there is a trend to overestimate axle factors in winter for LTPP sites, but that expected overestimation is, in average, less than 1 percent. In general, Method 1 results in smaller and more consistent errors at Wisconsin sites. Also for Wisconsin, the results seem to remain as accurate during winter, as opposed to sites from the LTPP dataset.

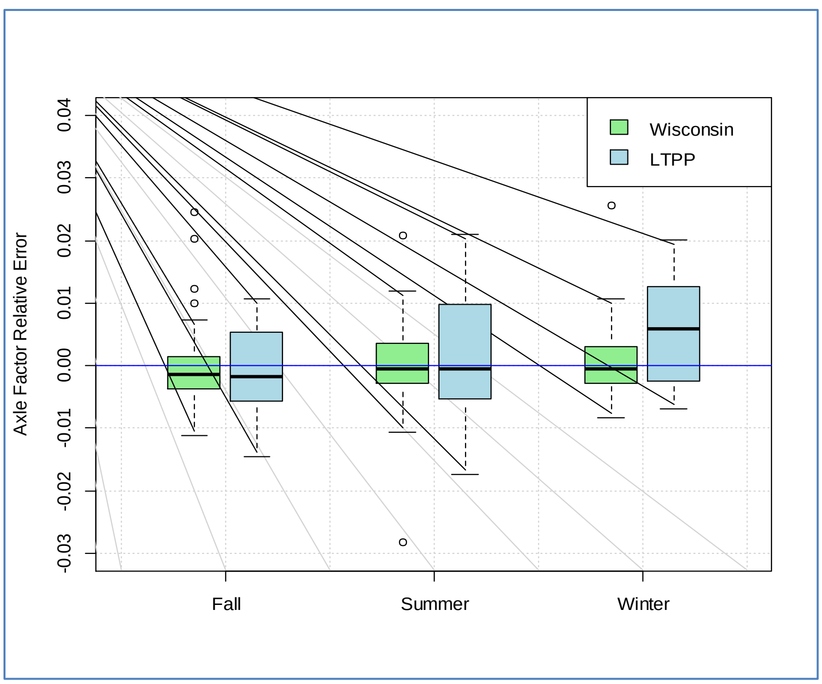


Figure 25. Error in Axle Factor Estimation using Method 1 by Source and Season

Seasonal Performance of Method 1 for Axle Factor Estimation by Facility Type

Figure 26 shows that deviation of axle factor values from the median value is lower during summer for both highways and freeways. This figure shows that highways have higher axle factors compared to freeways. This difference simply indicates that, on average, freeway facilities tend to carry vehicles with more axles than highway facilities. Again, winter months show larger dispersion compared to summer and fall for both types of facilities.

Figure 27 shows that proportional error in axle factor estimation from Method 1 by season and facility type. Axle factors do not appear to be estimated with bias (at least within 1 percent on average) for all the three seasons and at both highway and freeway facilities.

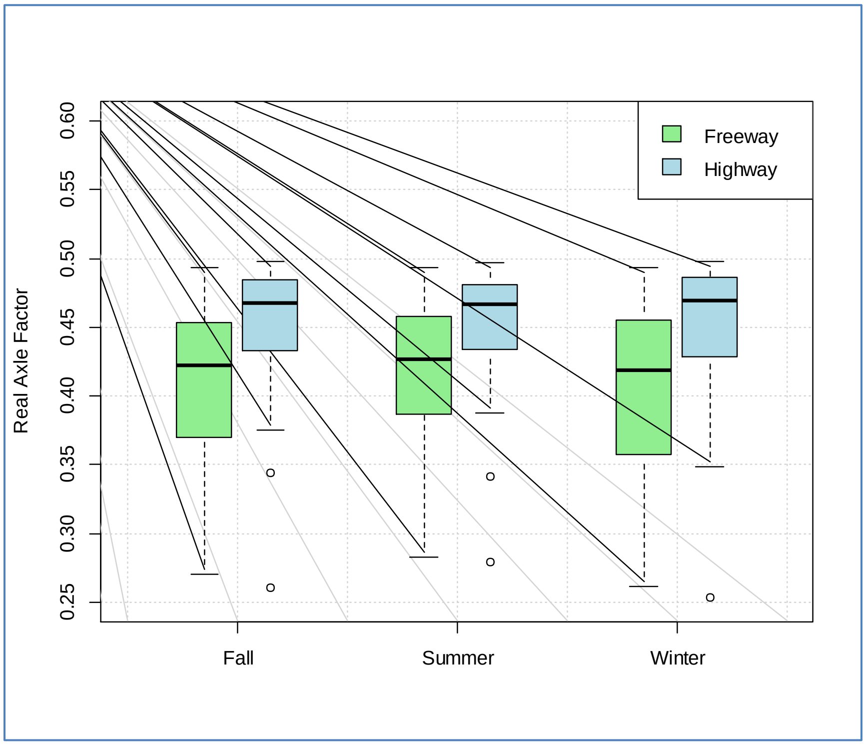


Figure 26. Axle Factor by Facility and Season

Per Figure 27, the maximum errors in estimation of axle factors are observed during winter at freeway facilities, yet more than 75 percent of the freeway sites (about 20 out of 26 sites) have errors within 1 percent of the real value.

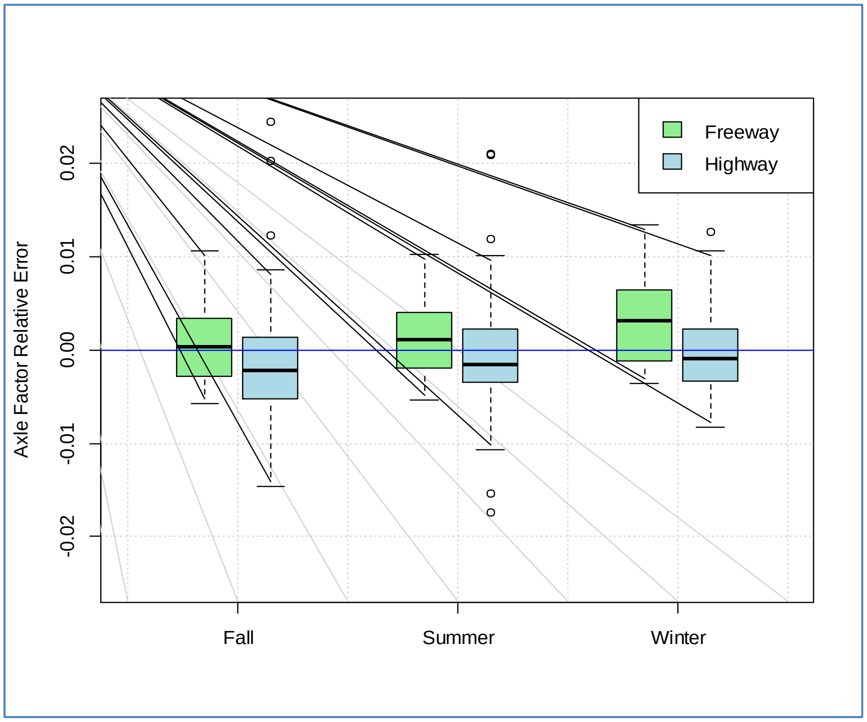


Figure 27. Error in Axle Factor Estimation using Method 1 by Facility and Season

Seasonal Performance of Method 5 for Vehicle Class Estimation by Data Source

Figure 28 shows the proportion of misclassified vehicles by season and data source. It can be seen that the median classification error was within 10 percent for both datasets during the three seasons considered. However, it is also clear that the LTPP subset had higher classification errors in general. The distribution of errors remained more or less unchanged during all seasons for both Wisconsin and LTPP datasets. In the case of the LTPP subset, however, more variability was observed in the performance of the classification algorithm during the summer months.

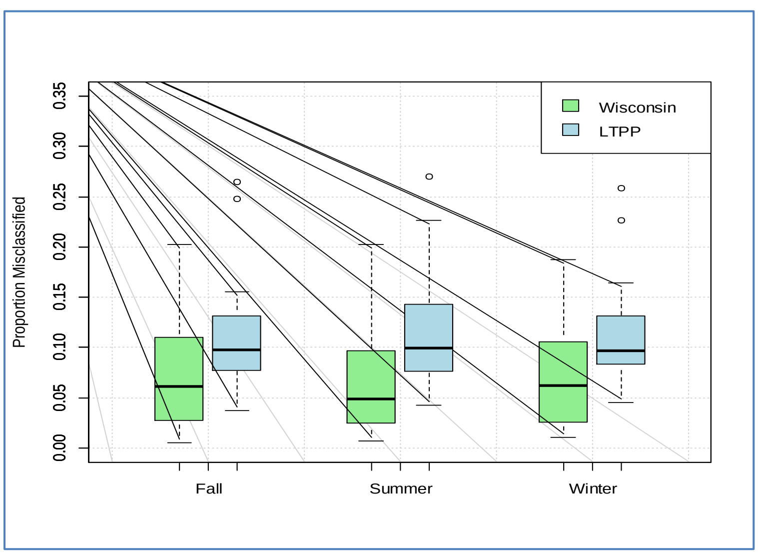


Figure 28. Classification Error by Source and Season using Method 5

Seasonal Performance of Method 5 for Vehicle Class Estimation by Facility Type

Figure 29 shows the performance of vehicle classification by season and facility type. The majority of classification errors shown in this figure are 13 percent or smaller for both highways and freeways. Freeway sites tend to have lower classification errors in general. Classification error seems to be disperse during fall for highway sites, as shown in Figure 29.

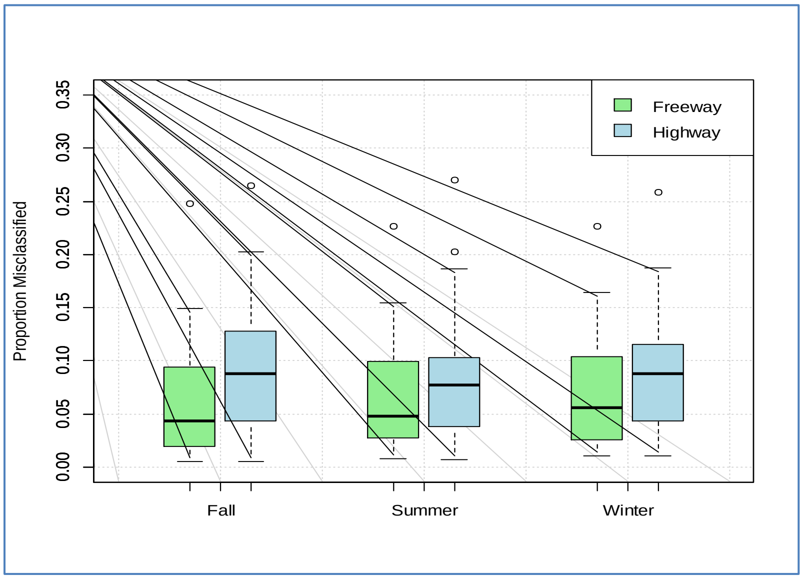


Figure 29. Classification Error by Facility and Season using Method 5

Conclusions and Recommendations

The research team assembled a multistate dataset using Wisconsin and LTPP data for comprehensive performance tests on Methods 1 and 5. Based on the results of the final round of evaluations, the performance of the two recommended methods was found within the expected limits of performance, as suggested by the first round of analysis. An implementation phase for both estimation methods is recommended.

The evaluations based on axle factor prediction confirmed the performance of these methods to be within 2 percent of the axle factor value being predicted. A trend to over predict axle factor at low axle factor values was identified for Method 1 and the opposite trend was identified for Method 5, but none of these trends exceeded acceptable amounts of error. The trends were invariant when looking at subsets of data by road character, facility type, and number of lanes. This invariability strongly suggests that a simple adjustment of the predicted values produced by either methods may help control for the bias and reduce the overall error. In any case, Method 1 is recommended if the intent is to develop axle factors only.

Based on the systematic trend of the bias, the research team recommends developing a post-hoc adjustment in the prediction of axle factors and incorporating such adjustment in the implementation of this method. However, given the small magnitude of the bias observed, acceptable performance is expected even if no systematic adjustment is developed at all.

Regarding the prediction of vehicle classes by Method 5, the evaluations showed a median misclassification proportion of about 10 percent, regardless of differences between data sources, and facility types. Interestingly, only meaningful trend in this error was found for urban sites (Figure 20). The proportion of misclassification was minimal at urban sites with lower axle factors and at freeways with higher axle factors. Similar to the error trends in axle factor estimation, these trends have the potential to inform a post-hoc adjustment of the results of estimates from Method 5.

Although a post-hoc adjustment is possible for Method 5, the research team recognized patterns from the performance evaluations by vehicle classes that could help improve the estimation procedure. For example, relationships of joint probabilities between vehicle classes can be incorporated into the estimation, potentially resulting in improved results. A revision and incorporation of such joint probabilities is recommended for the implementation phase.

Appendix A – Data Collection and Analysis Methods

SRF Consulting Group, Inc. | ENGINEERS PLANNERS DESIGNERS

Memorandum

SRF No. 9295

To: TPF5-340 Technical Advisory Committee

From: Scott Petersen, Associate

Date: November 16, 2016

Subject: TPF5-340 Axle and Length Classification Factor Analysis and Effects on AADT Outline of Data Collection Methods and Considerations

As the project team prepares to move into the data analysis portion of the project, this memorandum summarizes the data needs and proposes analysis methods to perform. A webinar will be held to prioritize analysis methods and further discuss the data sets to be analyzed.

Data Fields

For the analysis, all data must be provided as per vehicle record (PVR), even when the analysis method calls for binned data—data will be binned as appropriate for the analysis. PVR data can be binned to generate such data sets. This will make each methods' results directly comparable.

All data must be time and date stamped in the respective time zone or in UTC. Adjustments for Daylight Savings Time will be made if relevant. Agency-defined length bins must be provided for length data.

All axle-based data must contain the following vehicle attributes:

- Number of axles
- Axle-based vehicle class (classified per the respective jurisdiction)
- Speed (only to be used for a quality check if needed)

All length-based data must contain the following vehicle attributes:

- Length
- Speed (only to be used for a quality check if needed)

Data Quality Standards

Representatives from the agency that collected the data must vouch for the quality of the data. The data must be "production-level" data, meaning that this data meets the data quality standards of the respective agency.

The analysis will be conducted for a variety of durations to determine the seasonal and year-to-year change in results. Thus, data must be provided for as long of a duration as practical such that the agency remains confident about the quality and consistency of the data (avoid submitting data of unknown quality).

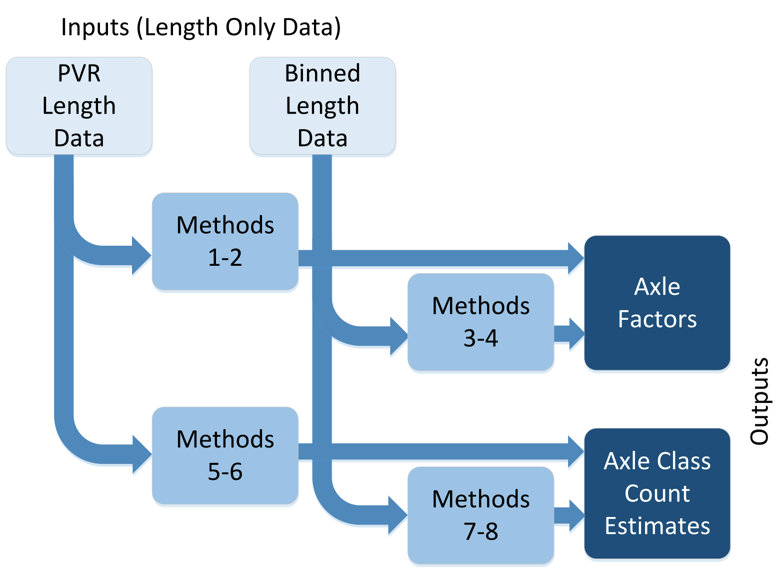
Calibration Method

The project scope calls for examining calibration methods with the expectation that calibrated sites have higher accuracy data. However, during the September 2016 project meeting, it was found that agencies have varying calibration methods ranging from testing the speed with a radar sensor to evaluating the average vehicle length per lane. Additionally, it was noted that some sensors, such as the Wavetronix sensors, report different lengths based on the range (distance from the sensor) of the subject vehicle. Since calibration methods vary significantly, it was determined that the analysis should not have a minimum standard for calibration, but would instead rely on the agencies' methods.

Analysis Methods

Axle Factoring Methods Summary

The following diagram illustrates eight methods that are suggested to meet the various needs of the participating agencies. These methods use a variety of inputs and outputs to meet the needs suggested at the September 2016 workshop.



Abbreviations and Terminology

Throughout the following sections, the following terms and abbreviations are used to minimize repetition.

- **AF**: axle factor, often used for factoring single tube counts
- **AC**: axle class (for example, FHWA 13 class scheme)
- **LC**: length class (for example, motorcycle/small/medium/large/very large)
- **PVR**: per vehicle record
- **VL**: vehicle length
- **Length bins**: mutually exclusive ranges of vehicle length that are generally configured in the classifier to produce a length classification
- **Length band**: similar to a length bin, but here is used to mean that the break points were selected during analysis, not predetermined
- **Parameterized function**: algorithm that uses probability distributions of vehicle lengths to categorize them into axle classes
- **Fractional assignment**: simple algorithm for factoring vehicle lengths into axle classes

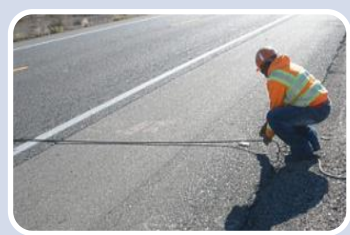
Site Type Purposes for Proposed Methods

The following three types of data collection sites are considered in this project. The use of each type is listed.

Use PVR data form axle class sites to generate seed data (use axle count and lengths).

Data from a single AC site may be used per cluster of many LC sites

Use length class sites (PVR or binned) to generate axle factors or factored axle class data.



Convert single tube counts to AADT using axle factor for the respective cluster.

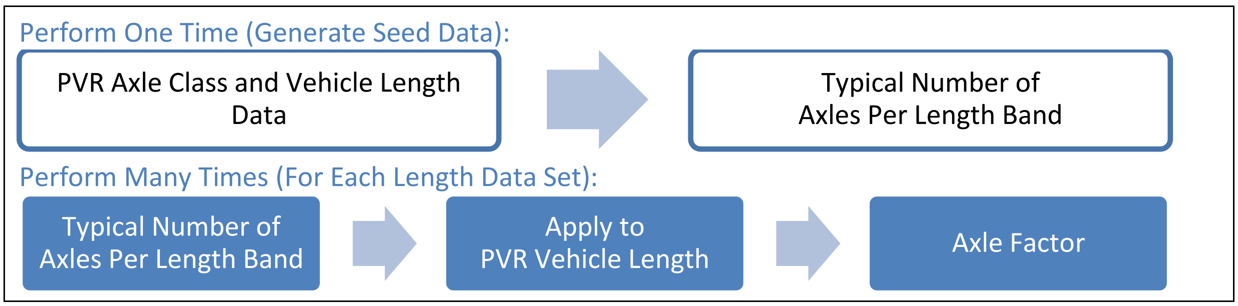
Method Detailed Description

Axle Factor Methods

The following four methods produce axle factors from length data. This axle factor can be applied to single tube counts to determine AADT.

1. **Method 1 (Uses PVR AC data and PVR VL data)**

   Summary: Use PVR AC data to determine typical numbers of axles per length grouping ("band") and generate an AF.



1.1. <u>Generate "seed" data from AC data.</u> Use PVR AC data (with both axle count and length) to determine the average number of axles per length band

1.1.1. Gather PVR AC data with vehicle lengths and axle counts and generate a list of vehicles (sort the list by vehicle length)

1.1.2. Determine length "band break points" based on visual inspection of histogram peaks or via an algorithmic method

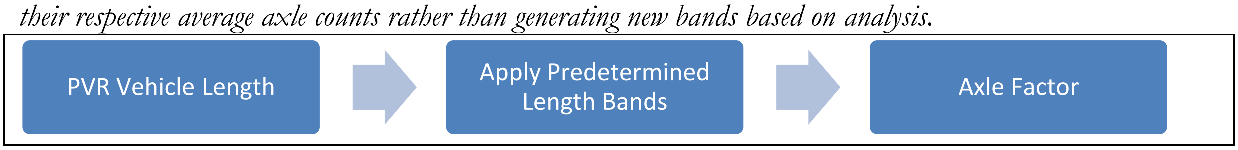
1.1.3. Determine the average number of axles per length band

1.2. Gather PVR VL data and determine the vehicle count per length band (bands determined in 1.1.2)

1.3. Apply the average number of axles (from 1.1.3) to vehicle counts per length band and calculate the AF

2. **Method 2 (Uses PVR VL data only)**

   Summary: This method is similar to Method 1 and produces an AF, but uses predetermined length bands and their respective average axle counts rather than generating new bands based on analysis.



2.1. Select a set of predetermined length bands and their average axle counts per band

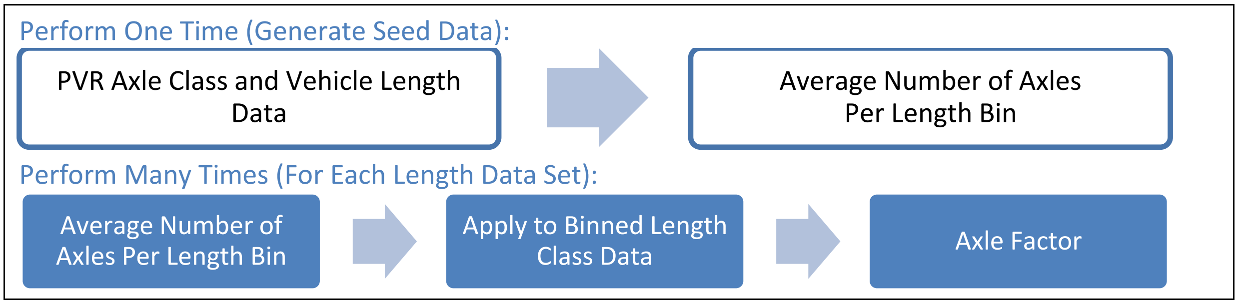
2.2. Gather PVR VL data and classify them using predetermined length bands

2.3. Determine the vehicle count per length band

2.4. Use a predefined conversion algorithm to generate an AF

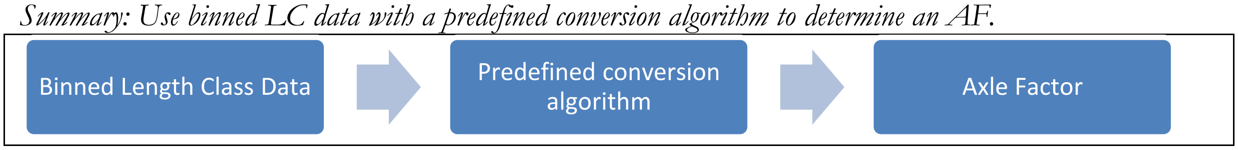
2.4.1. Example: AF = vehicle count / (2.0*motorcycle+2.0*small+3.6*medium+5.3*large+7.3*very_large)

3. **Method 3 (Uses PVR AC and binned LC data)**
Summary: Use PVR AC data to determine the average number of axles per length bin and apply it to the binned data to generate an AF.



3.1. Generate seed data. Use PVR AC data (axle count and vehicle length) to determine the average number of axles per length bin
  3.1.1. Gather PVR AC data (with vehicle lengths and axle counts) and generate list of vehicles with axle count and vehicle length (sort by length)
  3.1.2. Determine average number of axles per band (using length bins)
3.2. Gather binned LC data
3.3. Apply average number of axles (from 3.1.2) to the vehicle counts per length bin to determine the total axle count. Divide the number of vehicles by the axle count to calculate AF

4. **Method 4 (Uses binned LC data only)**
Summary: Use binned LC data with a predefined conversion algorithm to determine an AF.

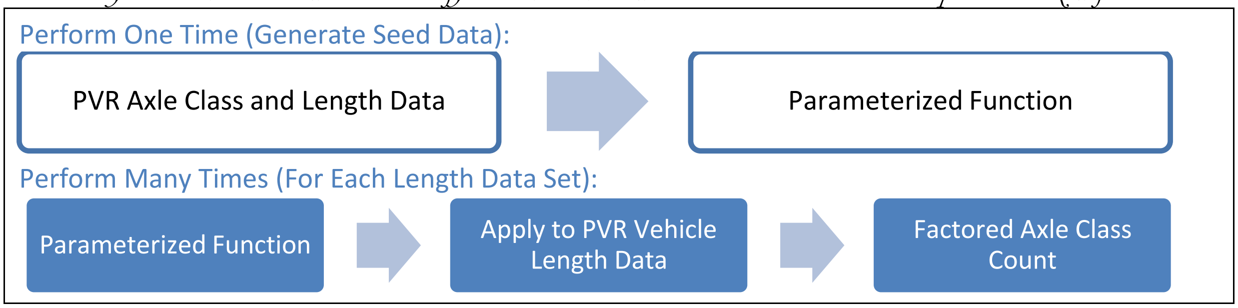


4.1. Gather binned LC length data
4.2. Use predefined conversion algorithm
  4.2.1. Example: AF = vehicle count / (2.0*bin1+2.0*bin2+3.6*bin3+5.3*bin4+7.3*bin5)

Axle Class Methods

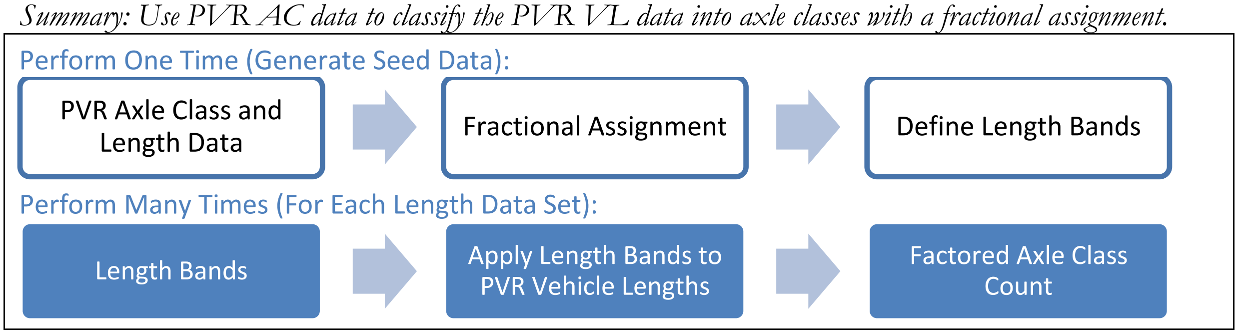
The following four methods are used to convert length data to AC data.

5. **Method 5 (Uses PVR AC and PVR VL data)**
Summary: Use PVR AC data to classify the PVR VL data into axle classes with a parameterized function.



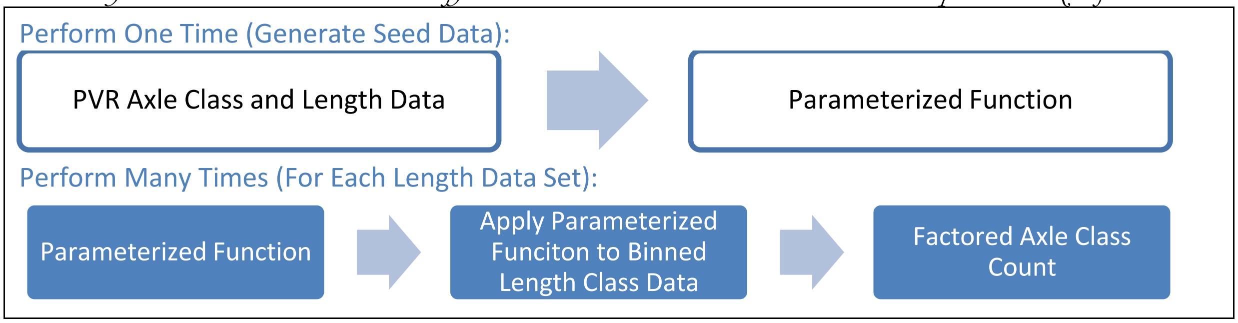
5.1. Use PVR AC data (axle class and length) to generate seed data (parameterized function relationship between length and axle class)
    5.1.1. Complete spreadsheet to generate the parameterized function
5.2. Gather and input PVR VL into the spreadsheet to apply the parameterized function
    5.2.1. The spreadsheet determines the vehicle count per axle class

6. **Method 6 (Uses PVR AC data and PVR VL data)**
Summary: Use PVR AC data to classify the PVR VL data into axle classes with a fractional assignment.



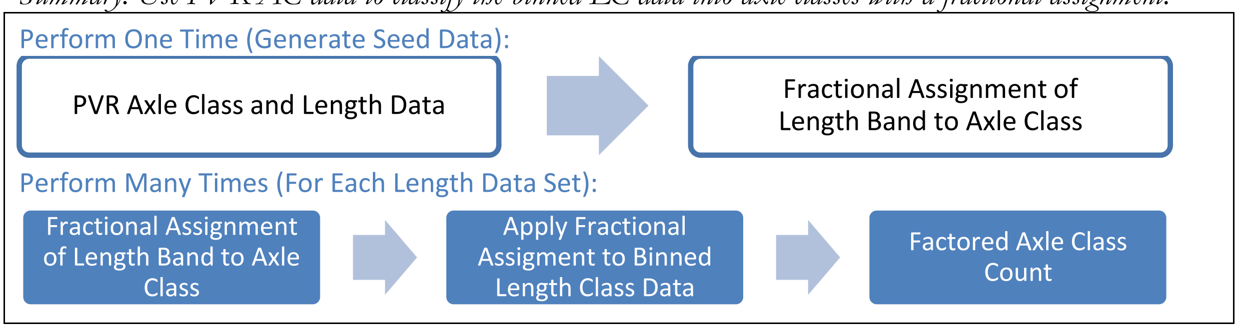
6.1. Generate seed data for fractional assignment
    6.1.1. Gather list of vehicle lengths and axle classes and sort each of these parameters independently (both ascending)
    6.1.2. Define length bands that correlate to axle class
6.2. Perform fractional assignment of PVR vehicle length data to axle class

7. **Method 7 (Uses PVR AC data and binned LC data)**
Summary: Use PVR AC data to classify the binned LC data into axle classes with a parameterized function.



7.1. Use PVR AC data (axle class and length) to generate seed data (parameterized function relationship between length and axle class)
 7.1.1. Complete spreadsheet to generate the parameterized function
7.2. Gather and input binned LC data into the spreadsheet to apply the parameterized function and determine the vehicle count per axle class

8. **Method 8 (Uses PVR AC data and binned LC data)**
Summary: Use PVR AC data to classify the binned LC data into axle classes with a fractional assignment.



8.1. Generate seed data for fractional assignment
 8.1.1. Gather list of vehicle lengths and axle classes and sort each of these parameters independently (both ascending)
 8.1.2. Define length bands that correlate to axle class
8.2. Perform fractional assignment of binned LC data to axle class

Next Steps

A project conference call will be held in November or December to determine which methods to pursue.