Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic Estimation

Assessing AADT Accuracy Issues Related to Short-Term Count Durations

Publication No. FHWA-PL-16-008

October 2015

U.S. Department of Transportation
Federal Highway Administration

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-PL-16-008		

4. Title and Subtitle Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic Estimation: Assessing Accuracy Issues Related to Short-Term Count Durations	5. Report Date October 2015 6. Performing Organization Code
7. Author(s) Robert Krile, Fred Todt, Jeremy Schroeder (Battelle)	8. Performing Organization Report No.
9. Performing Organization Name and Address Battelle 505 King Ave. Columbus, Ohio 43201	10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTFH61-13-D-00012
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Highway Policy Information 1200 New Jersey Avenue SE Washington, DC 20590	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplementary Notes

Project performed in cooperation with the U.S. Department of Transportation, Federal Highway Administration and a Technical Advisory Committee (TAC) consisting of Alaska DOT & Public Facilities, Georgia DOT, Illinois DOT, Minnesota DOT, Pennsylvania DOT, Texas DOT, and Wisconsin DOT. Steven Jessberger (Task Manager)

16. Abstract

AADT for many roadways is estimated through a temporary count obtained over anywhere from a few hours to one week, and subsequently expanded to a full year using factors derived from permanent count stations with similar characteristics. Historic research has focused on determining the optimum duration for the short-term count to balance resource restrictions, which favor shorter count durations, against statistical accuracy, which favors longer count durations. This task quantifies the relative accuracy and precision associated with different durations. It also evaluates the impact of day of week and month of year for AADT estimation. Results are evaluated for different factoring methods, and different factor grouping methods.

FHWA Travel Monitoring Analysis System (TMAS) volume data from 14 years consisting of hourly counts by day from nearly 43,000 continuous permanent volume traffic data sites/years in the United States comprised the reference dataset for this research. A subset of 320 of these were utilized which include complete data for all 24 hours of every day of the year. These sites collectively represented a wide range of AADT volumes, nine functional classes, 32 states, and years 2000 through 2012. This report is a final task report that summarizes accuracy and precision of AADT estimation for various short-term count durations, by day of week and month of year in which count was obtained, and includes the analysis methodology and summary statistics findings.

17. Key Words	18. Distribution Statement
Annual average daily traffic, AADT, continuous traffic monitoring	No restrictions. This document is available to the public.

19. Security Classif.(of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	45	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Preface

We would like to acknowledge the following Technical Advisory Committee members for their contributions, support and technical guidance during this project.

Jennifer Anderson	Alaska DOT & Public Facilities
Scott Susten	Georgia DOT
William Morgan	Illinois DOT
Gene Hicks, Thomas Nelson, Mark Flinner	Minnesota DOT
Jeremy Freeland, Joseph Pipe	Pennsylvania DOT
William Knowles	Texas DOT
Rhonda McDonald	Wisconsin DOT

Table of Contents

List of Figures

List of Tables

Executive Summary

Annual average daily traffic volume (AADT) for many roadways is estimated through a temporary count obtained over anywhere from a few hours to one week, and subsequently expanded to a full year using factors derived from permanent count stations with similar characteristics. This research, Task 3 of a larger evaluation project entitled “Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic Estimation”, examined the inherent accuracy and precision of expanding short-term counts, Average Daily Traffic (ADT) to AADT. Federal Highway Administration (FHWA) Travel Monitoring Analysis System (TMAS) data from 14 years consisting of 24 hours of the day and seven days of the week volume data from nearly 6000 continuous permanent volume traffic data sites in the United States comprised the reference dataset for this research.

This task was completed using a methodology that leveraged the known traffic volumes of a national cross-section of permanent traffic count stations, treating these permanent stations as surrogates for temporary count stations, and then comparing the estimated AADT from short-term counts extracted from the permanent count stations to the true known AADT for each permanent count station. This method permitted a comparison of AADT estimation accuracy (i.e., on average, how close the estimated AADT is to the true AADT) and precision (i.e., how variable the estimated AADT results are from one short-term count instance to another) as a function of short-term count durations ranging from six hours to a week. Different factoring methods and factor groupings were also studied, as well as how accuracy and precision relate to the day of week and month of year in which the short-term counts are obtained. This is a final task report that includes the analysis methodology, summary statistics findings, and separate supplemental documentation.

Introduction

Annual Average Daily Traffic (AADT) volume is an important measure of road use. For a set of locations nationally known as permanent traffic count stations, AADT can be measured quite accurately and precisely, even if some periods of time are missing or unavailable. For most roads, it is not economically or logistically practical to maintain a permanent traffic count station, but an accurate AADT is still needed. To determine AADT for such roads, the typical practice is to deploy a temporary counter for anywhere from a few hours to as many as seven days. After the counter is retrieved, the hourly traffic volume is extracted and then converted to an expected annual average daily value based on the characteristics of the roadway and of the period sampled. This conversion, or factoring, is based on the same period (e.g., Tuesday, or Tuesday in November) and for a set of sites with permanent count stations that are considered similar to the location of the temporary count.

The expected error in the process of using a temporary count to develop an AADT has been an issue of significant study over many years. While it is possible to use statistical methods to simulate the impact of varying practices in temporary counts (ADT) and then to assess how they project to an AADT, these methods have difficulty adequately capturing the complex dynamics of variability that exist in true count data over time and across sites.

An overall research effort was begun in 2014 in which Federal Highway Administration (FHWA) Travel Monitoring Analysis System (TMAS) data from 14 years consisting of 24 hours of the day volume data from nearly 43,000 continuous permanent volume traffic data sites (25 million records) in the United States comprised the reference dataset. In the report, "Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic (AADT) Estimation: Assessing Accuracy Issues with Current Known Methods in AADT Estimation from Continuous Traffic Monitoring Data,"[1] the expected bias and precision associated with AADT estimation of permanent traffic count stations were established under different analytical methods to account for daily weighting and missing data. This task uses much of the same data, but identifies what duration of a short-term count provides optimal accuracy and precision for estimating AADT. It also examines the accuracy and precision as a function of the factoring methods used, the factor grouping methodology, and the time period (e.g., day of week, month of year) in which the temporary count is obtained.

FHWA has undertaken this evaluation that samples from existing permanent count station data where AADT is fully measured as if these stations were temporary count locations. In this method, a slice of traffic volume data within the reference dataset is removed from a permanent count site (e.g., 24 hours) to simulate a short duration count. The simulated short-term count then has day-of-week and monthly correction factors applied to arrive at an estimated AADT. The adjustment factors would be those typical for the site evaluated had it truly been a temporary count location. This estimated AADT from a simulated short-term count is compared to the true AADT for the site since it is truly a permanent count station with a definitively known AADT. The difference between the estimated and true AADT value is the bias in using that particular short-term count to estimate AADT.

[1] Publication No. FHWA-PL-015-008, November 2014.

The basic methodology described was ultimately applied to a set of nearly 320 permanent count stations. By applying this across such a broad number of sites, summary statistics could be generated that both characterize the overall bias inherent in short-term count use to estimate AADT, and could be utilized to compare the error under different short-term count durations. Of particular research interest was to compare the error distribution under a range of different short-term durations to include:

- Partial day (6 hours and 12 hours);
- Single day (24 hours and a calendar day);
- Two days (48 hours and two calendar days);
- Three days (72 hours and three calendar days); and
- One week.

Additionally, with the robust nature of this evaluation, it is possible to examine whether factors such as functional classification (FC), volume range, day of week, month, or year influence the comparisons.

This following sections of this report begin by discussing Methods employed to select the sites for evaluation, to simulate short-term counts from those sites and appropriately factor them up to AADT, to determine the AADT estimation error, and then to calculate and characterize the subsequent bias (accuracy) and precision of the results for a number of different reporting levels. The Methods section is followed by Results, which presents summary results and discusses how they address the key research objectives of the task. This is followed by Conclusions of the key findings of the research. An Appendix and a Supplemental Set of Tables complete the report.

Methods

Traffic Monitoring Site Data

For this evaluation, FHWA TMAS traffic volume data were used as the input data. Following a series of data processing steps, which are discussed in more detail in the report, “Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic (AADT) Estimation: Assessing Accuracy Issues with Current Known Methods in AADT Estimation from Continuous Traffic Monitoring Data,”[1] a complete dataset of available sites with corresponding hourly traffic volumes was generated. This dataset consisted of 42,876 site and year combinations for which adequate hourly data existed to determine an AADT using the American Association of State Highway and Transportation Officials (AASHTO) method. The AASHTO method was used in this evaluation even though improvements to accuracy and precision of AADT estimation were identified in [1], notably the Highway Policy Steven Jessberger-Battelle (HPSJB) Method. The advantage of the HPSJB method is most apparent for data with some hourly volumes but not all hours on a day. This scenario is not common in current TMAS data, since the AASHTO method excludes all such partial days and therefore there is no benefit to submitting partial daily data to TMAS. Additionally, the currently utilized AASHTO method is more familiar to the audience, and since this evaluation looks at comparative accuracy and precision for different subsets of time, rather than absolute accuracy and precision, the estimation of AADT with the AASHTO method was judged to be acceptable. For 5,681 of the site and year combinations, data were available for every hour of the year. It was this latter set of data which constitutes the core of this task, since an exact AADT is known for each of these sites (i.e., no missing data). By sub-setting portions of the complete data for these sites to simulate a temporary count, and then applying appropriate weighting factors, simulated short-term count duration estimates of AADT are produced, which can be compared against the true value for the site.

In the portion of the methodology that features the application of weighting factors, there are a number of candidate methods that could have been employed. Each method has some basis in use either through the Highway Performance Monitoring System (HPMS) or in practice. Development of weights requires a set of site data separate from the sites to be evaluated directly as temporary count stations. For these sites, having complete hourly data for the entire year is not strictly necessary, but they must support calculation of weighting factors at the most detailed level desired. As such, the sites that established the weighting factors could be part of the 5681 complete data sites, or they could simply be within the 42,876 sites where an AASHTO AADT calculation was possible. Additionally, the minimum number of sites that formed a weighting group is an important consideration. Following the Traffic Monitoring Guide (TMG), a minimum of six sites would be required for developing factor groupings, which would then produce factors that could be applied to a seventh site with complete data that was being used as the proxy for the temporary count station. For computer programming efficiency, the maximum number of sites for a particular factor grouping was 15.

A number of other criteria for selection of sites were provided by FHWA to assure that the aggregate results would be representative of traffic count patterns as a whole across the United States:

- A minimum of 400 sites;
- To the extent possible, all states, ranges of estimated AADT volumes, and roadway functional classifications; and
- All years from 2000 through 2012 (data from 2013 were available, but due to a later task on this work assignment to associate multi-year factoring results with the single year results of this task, it was necessary to restrict sites to only the first 13 years of the period).

Ultimately three distinct, but largely overlapping, sets of sites were selected for this evaluation. In each of the three evaluations, the sites that made up factor groupings were organized based on a different characteristic of the hourly data volumes for those sites. The groupings included:

1. Subset for functional classification;
2. Subset for volume range; and
3. Subset for clustering algorithm.

Site Selection

Choosing which sites to include in each of the three types of grouping analyses started with the list produced in the Task 2 analysis of this evaluation.[1] The site selection for each of the grouping methods is discussed in greater detail below.

Functional Classification Groups

Based on input from the Technical Advisory Committee (TAC) states on this project regarding sites that might not be most representative of typical traffic volumes and subsequent discoveries of potential data reporting issues in some sites, the functional classification sites selected for the evaluation were a very similar, though slightly expanded, set of sites as were used in the Task 2 evaluation under this contract. They are summarized in Table 1 below. In aggregate, there are 552 total sites used in the analysis, organized into 44 unique functional classification, state, and year groups. Each group contains between seven and 15 sites, at least one of which must have complete hourly data for the year. A total of 206 sites of the 552 met this criteria. In aggregate, the groups represent nine of the 14 functional classifications. There were not adequate numbers of sites with complete data to include groups for the other five functional classifications. In total, 32 different states, and all years from 2000 through 2012 are represented by at least one group.

Table 1. Summary of Sites Selected for Functional Classification Analysis

FC	State	Year	Number of Sites per Group Complete	AADT	Total
1R	Indiana	2007	1	14	15
	Pennsylvania	2004	3	5	8
1U	Hawaii	2010	3	9	12
	Massachusetts	2000	9	1	10
	Michigan	2006	3	11	14
	Minnesota	2008	10	4	14
	Wisconsin	2009	1	13	14
2U	California	2003	3	10	13
	New York	2006	3	12	15
	Oklahoma	2008	1	7	8
3R	Idaho	2004	7	8	15
	Kansas	2004	2	13	15
	Kentucky	2003	4	11	15
	Maine	2008	15	0	15
	Mississippi	2008	1	10	11
	Nebraska	2008	7	8	15
	New Mexico	2005	2	7	9
	Washington	2005	8	6	14
	Wyoming	2000	3	12	15
3U	Alabama	2000	2	7	9
	California	2003	4	11	15
	Florida	2005	1	14	15
	Illinois	2011	2	9	11
	Nevada	2008	1	9	10
	Utah	2005	9	2	11
	Wisconsin	2010	2	13	15
	Wyoming	2000	4	11	15

FC	State	Year	Number of Sites per Group Complete	AADT	Total
4R	Colorado	2011	2	7	9
	Florida	2006	1	14	15
	Michigan	2003	2	8	10
	Montana	2006	1	12	13
	Oregon	2005	9	6	15
	Pennsylvania	2011	6	5	11
	Texas	2002	15	0	15
	Vermont	2011	3	9	12
4U	Alaska	2010	5	10	15
	Iowa	2004	7	0	7
	Nevada	2008	3	6	9
5R	Kansas	2009	2	8	10
	Maine	2011	15	0	15
	Ohio	2005	2	7	9
	Texas	2001	11	4	15
	Vermont	2012	9	6	15
5U	Kentucky	2006	2	7	9
Grand Total			206	346	552

Source: Battelle

Volume Range Groups

The selection of sites for the volume range groups were derived from the sites in the functional classification analysis. Where possible, groups from the functional classification analysis were left the same (e.g., if all sites in a functional classification group fell within the same volume range). Some additional supplementation was necessary to achieve groups in all four volume ranges. Ultimately, the volume range site selection generated 52 total groups, comprising 493 total permanent count stations, 192 of which had complete hourly data for their respective state and year. The breakdown of clusters by volume range groups was as follows:

- Volume Range 1 (AADT<1000) – 4 groups;
- Volume Range 2 (1000 ≤ AADT < 10,000) – 24 groups;
- Volume Range 3 (10,000 ≤ AADT < 100,000) – 20 groups; and
- Volume Range 4 (AADT ≥ 100,000) – 4 groups.

Similarly to the functional classification analysis, the aggregate set of sites for the volume range groups represented 32 total states and included at least one group for each of years 2000 through 2012.

Clustering Algorithm Groups

The objective of aggregating sites for factoring by functional classification or volume range is to establish a group of sites that has a similar trend in hourly volumes as the site to be used for the temporary count. A third group of sites was generated in this evaluation that directly analyzed hourly time trends to establish the groups. The methodology consisted of first sub-setting the 42,890 separate sites to just those within the states and years previously included in the other two grouping methods (i.e., 32 states and 13 years of data from 2000 through 2012). This resulted in 4,405 individual site and year combinations.

The volume data, normalized by hour of day, day of week, and day or year was then fit to a k-means clustering algorithm. From the clustering algorithm, groups were generated for each of these separate time trends (i.e., daily, weekly, hourly). Graphical depictions of this clustering process are shown in Figure 1, with Alabama sites in 2000 as the example. The number of groups within each category was determined by where an inflection point was reached in the percentage of variability explained by the addition of more clusters. Adding more clusters provides better differentiation between patterns but at the cost of not having an adequately robust number of sites to develop factors. From Figure 1, the first row of plots illustrates this process by which the daily and weekly patterns (left and center plots in top row) were separated optimally into two different unique patterns. There was not adequate differentiation of the hour of day pattern (right plot in top row) to develop separate clusters. The subsequent three rows of Figure 1 show the mean trends that resulted from the clustering process, as well as those sites (denoted by blue lines) that were associated with each trend. The Combination: 1,1,1 row (n=10) shows sites with a strong seasonal trend of higher summer volumes and a day of week pattern with relatively highest volume on Friday (Day of Week=6). The Combination: 2,1,1 (n=20) and 2,2,1 (n=26) rows shows a day of year trend that is more evenly distributed, but the two clusters are differentiated by the day of week trends of either more even traffic distribution (2,1,1) or a very pronounced drop in weekend, especially Sunday (Day of Week=1) volume (2,2,1). As can be seen in all three clusters, the hour of day pattern on the far right is consistent across sites with a modest AM peak, building volume through the later morning and early afternoon, and then a strong PM peak.

Generalizing to all state and year clusters, within each time trend group, every site was assigned to one group, and only one group, based on the k-means calculation. If k_d clusters were generated by the daily trend, k_w clusters by the weekly trend, and k_h clusters by the hourly trend, a total of $k_d \times k_w \times k_h$ separate profiles could exist for the state and year. Every site would be assigned to only one of these aggregate clusters (though it is possible a cluster might have no site assigned to it).

To select individual sites for the clustering algorithm procedure, the mathematical distances between each of the sites and the three centroids of the clusters were reduced to a normalized distance metric within each cluster. A spreadsheet was generated that included all 4,405 available sites, the grouping they had been assigned by the clustering algorithm, their distance metric relative to the centroid of their cluster, and an indicator for whether the site was contained in the functional classification analysis. A manual process was conducted in which groups of sites were selected for the clustering algorithm analysis based on the following guidance:

- Include wherever possible the sites with complete data that were part of the functional classification analysis; and
- Around the complete data sites, selecting sites immediately adjacent with respect to the cluster distance metric.

The final dataset for the clustering algorithm analysis contained 44 clusters, ranging in size from seven to 17 sites. The clusters were comprised of 567 total site and year combinations, of which 201 had complete hourly data for the year. Overall, the 567 clustering algorithm sites encompass 32 states and all years from 2000 through 2012.

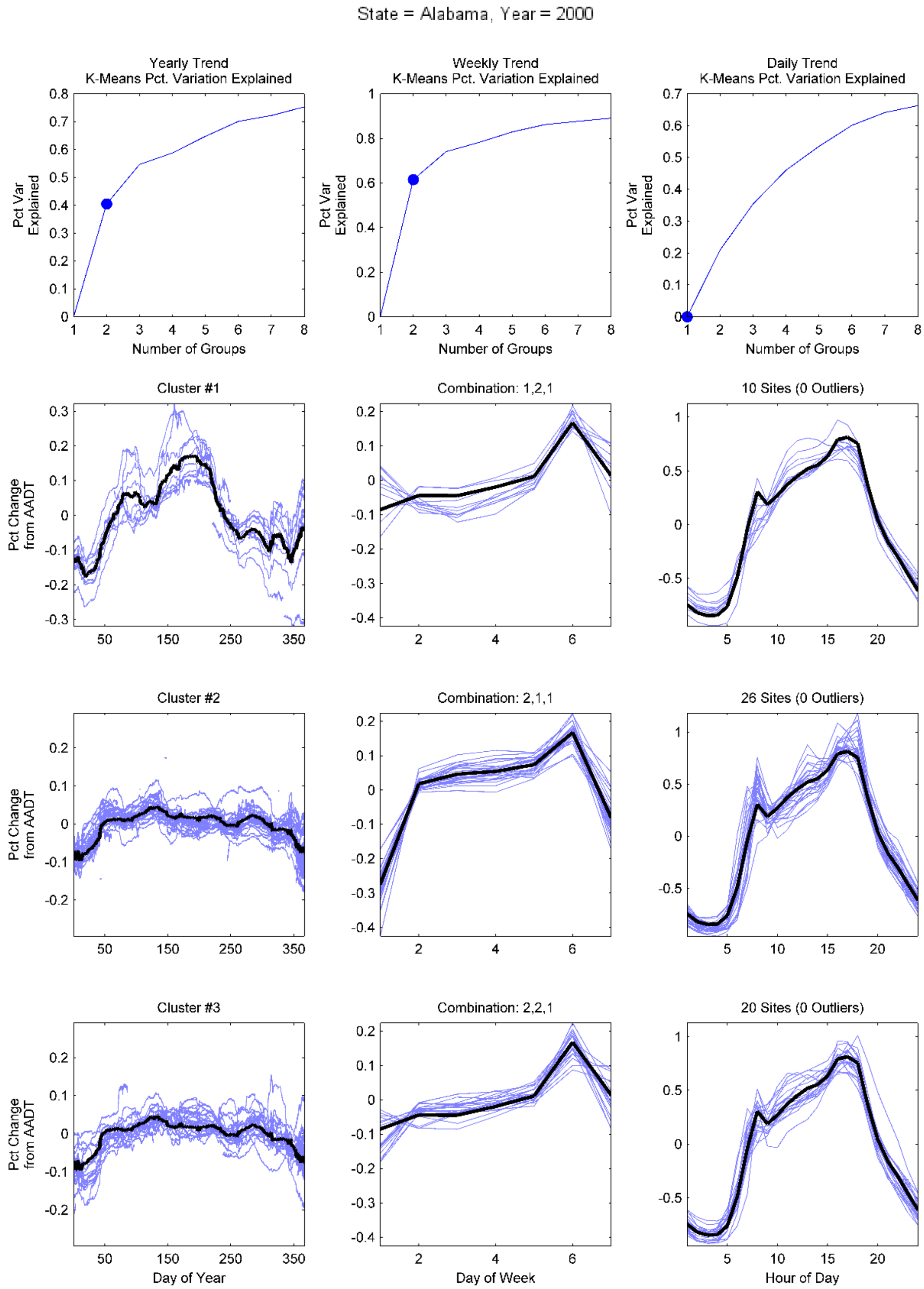

Figure 1. Clustering Algorithm Demonstrated for One State and Year (Source: Battelle)

Site Selection – Overlap

For the three factor grouping methods: functional classification, volume range, and clustering algorithm, site selection was carried out to achieve as much overlap as possible. This was done to allow for comparison of the factor grouping method results to avoid possibly being confounded with the volume characteristics of the sites themselves. The achievement of this overlap is documented in Figure 2 and Figure 3. In Figure 2, a Venn diagram is provided for the overlap of all the sites with 100 percent complete data, which shows that a grand total of 320 sites were used for the evaluation, with 82 of them part of all three factor grouping methods. The overlap of the functional classification and volume range factor groups was very high, while the cluster algorithm required inclusion of some different sites. Figure 3 is the Venn diagram that shows overlap of all sites used. This figure shows that 956 total sites were used either as a 100 percent complete site or as a part of a factoring group, with 155 of these common to all three factoring methods.

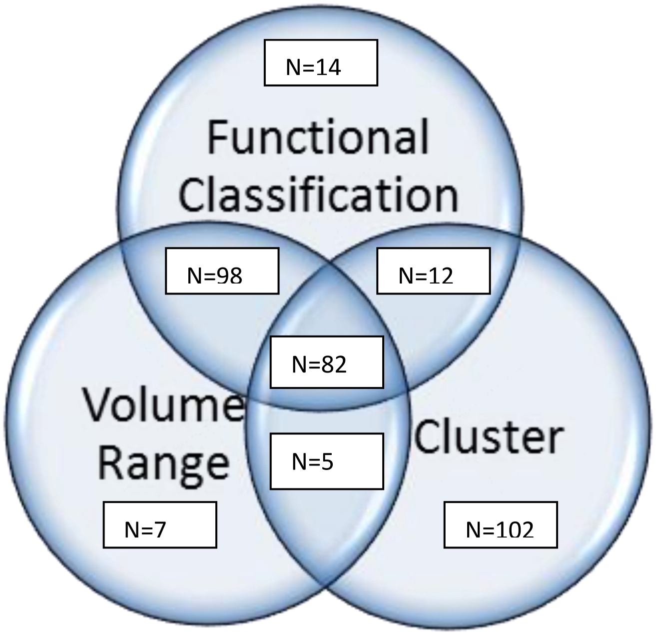

Figure 2. Overlap of Sites with 100 Percent Reporting for Different Factor Grouping Methods (Source: Battelle)

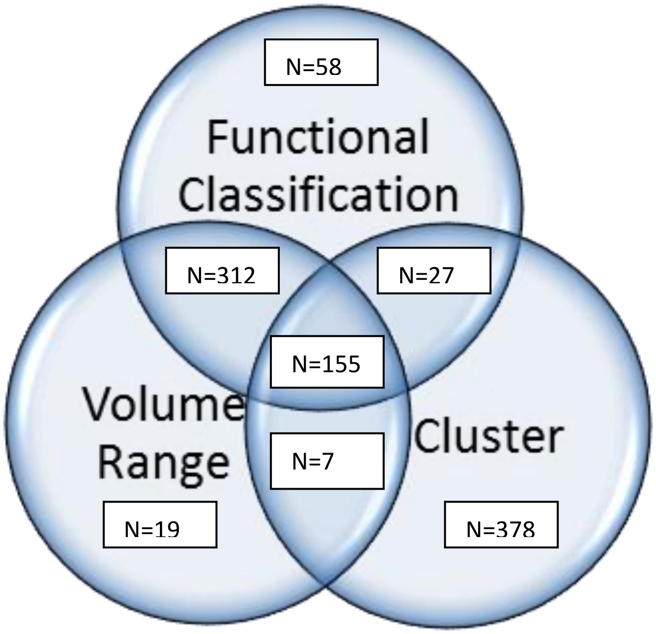

Figure 3. Overlap of All Sites Used by Different Factor Grouping Methods (Source: Battelle)

Methodology for Estimating AADT from Short-term Counts

The basic approach to the evaluation is to select a site with complete hourly data for a year, and hence a known AADT (assumed to be calculated using the AASHTO method). From this site, the generation of a temporary traffic count is simulated using a subset of time ranging from a few hours to as many as seven days, An appropriate weighting factor is then applied to this temporary count to produce an estimated AADT. We then calculate the difference between this AADT based on simulated short-term count with the true AADT for the site to determine the bias, or accuracy. These bias statistics can then be summarized over time periods or site geographies to determine the potential error that could be observed in practice with short-term counts expanded to AADT.

To execute the basic methodology, there are some important considerations that include:

- What are the weighting factors applied to the temporary counts to determine an estimated AADT?
- What short-term count durations should be evaluated?

Each of these topics is covered in greater detail below.

Weighting Factors

For a particular factor group, assume that there are k total sites. Of this total, k_c sites have 100 percent complete data for the entire year (i.e., 24 hours of reported volumes every day for all 365 (or 366) days in the year). The remaining (k-k_c) sites have adequate data to determine the AADT using the AASHTO method (i.e., at least one complete day of each day of week in each of the twelve months of the year).

The analysis for this task was completed using three different weighting factor methods:

- Separate monthly and day of week factors;
- Monthly day of week factors; and
- Day of year factors.

The reason for evaluating the different factoring method is that each is or has been in use within the travel monitoring community. Each of these methods is detailed below.

Separate monthly and day of week factors

The following weighting factor method is considered the primary method within this evaluation. It is comparable to what is presented in Section 3.4.2 and 3.4.3 of the TMG (September 2013, p. 3-78 through 3-80).

Develop monthly factors:

1. Use the AASHTO method to calculate AADT for each of k site and year combination.
2. For a particular one, i, of the k_c site and year combinations, calculate monthly average daily traffic (MADT) for the k-1 site and year combinations excluding the i^{th} site. Specifically, first get average traffic volume for each day of week within each month. Then average day of week traffic volume within every month.
3. Calculate 12 monthly factors for each of k-1 site and year combinations as AADT divided by corresponding MADT.
4. Average monthly factors across the k-1 site and year combinations to get monthly factors (one for each month) for the overall group.
5. Repeat steps 2 through 4 for $i=2,\ldots,k_c$. In this manner, a separate set of monthly factors are generated for each of the sites that has 100 percent complete data.

Develop day of week factors:

1. Treat Monday, Tuesday through Thursday, Friday, Saturday, and Sunday as five separate groups.
2. For a particular one, i, of the k_c site and year combinations, average all the day of week averages from the AASHTO AADT calculation to generate five day of week averages (i.e., one for Fridays, one for Saturdays, one for Sundays, one for Mondays, and one that includes all Tuesdays through Thursdays.) Exclude the i^{th} site and year in this calculation.
3. For each of the k-1 site and year combinations, divide its AADT by the average traffic volume for each day of week group to get day of week factors.
4. Average day of week factor across the k-1 site and year combinations to get day of week factors for the group. There will be five day of week factors (one for Tuesday through Thursday, and one each for Monday, Friday, Saturday, and Sunday).

From the factors generated above, a short-term count is expanded to an AADT by multiplying both the monthly and day of week factors that apply to that short-term count.

Monthly day of week factors

An alternative to the separate monthly and day of week factors is to have day of week factors that are month-specific. These factors are derived as follows:

1. Use the AASHTO method to calculate AADT for each of k site and year combination.
2. For a particular one, i, of the k_c site and year combinations, calculate a day of week by month average daily traffic volume for the k-1 site and year combinations excluding the i^{th} site. Specifically, average all January Mondays, January Tuesdays through Thursdays, January Fridays, January Saturdays, January Sundays, February Mondays,…, December Sundays.
3. Calculate 60 factors for each of k-1 site and year combinations as AADT divided by corresponding day of week by month traffic volume.
4. Average day of week by month factors across the k-1 site and year combinations to get day of week by month factors (60 total) for the overall group.
5. Repeat steps 2 through 4 for $i=2,\ldots,k_c$. In this manner, a separate set of day of week by month factors are generated for each sites that has 100 percent complete data.

Day of year factors

A most detailed level of factoring would develop a separate factor for each day of the year. In this manner, the factors applied to a short-term count are perfectly paired to the average factor group volumes of the same day. These factors are derived as follows:

1. Use the AASHTO method to calculate AADT for each k site and year combinations.
2. For a particular one, i, of the k_c site and year combinations, calculate daily traffic volume factors for the k-1 site and year combinations excluding the i^{th} site as AADT divided by the daily traffic volume.
3. Average daily factors across the k-1 site and year combinations to get day of year factors (365 or 366 total) for the overall group.
4. Repeat steps 2 through 3 for $i=2,\ldots,k_c$. In this manner, a separate set of day of year factors are generated for each sites that has 100 percent complete data.

Hour of day factors

In addition to the factoring above, some of the short-term count methods will result in data with less than a full day of hourly counts or with counts split across more than one calendar day. For this reason, a single set of hour of day factors are generated that apply to all the higher level factoring methods.

1. Treat Tuesday, Wednesday and Thursday as a group. Monday, Friday, Saturday, and Sunday are each treated individually.
2. For each of the k-1 site and year combinations, calculate the percentage of traffic volume for each hour within every day where all hourly volumes are available for that day.
3. For each of the k-1 site and year combinations, average the percentages by each hour and day of week group.
4. Average the percentages across the k-1 sites in the group. There will be 120 hour of day factors (i.e., 24 hours of the day x 5 day of week groups).

Short-term Count Duration

To examine the impact of different short-term count scenarios, a total of ten different durations were selected, ranging from just a few hours to as many as seven days. This range effectively bounds the shortest to longest count durations employed in practice for temporary traffic counts. The ten durations included the following:

1. 6 hours to include the 6 AM – 9 AM interval and the 3 PM – 6 PM interval;
2. 12 hours to include the 6 AM – 6 PM interval;
3. 12 hours to include the 9 AM – 9 PM interval;
4. One day (i.e., midnight to midnight) to include the 24 hours of one calendar day;
5. 24 hours to include a random start time on one day and through 24 subsequent hours;
6. Two days (i.e., midnight to midnight) to include the 48 hours of two calendar days;
7. 48 hours to include a random start time on one day and through 48 subsequent hours;
8. Three days (i.e., midnight to midnight) to include the 72 hours of three calendar days;
9. 72 hours to include a random start time on one day and through 72 subsequent hours; and
10. Seven days (i.e., midnight to midnight) to include the 168 hours of seven calendar days.

For a particular site and year, each duration can be evaluated by its starting date. For instance, in a year with 365 days, there are 365 unique scenarios for evaluating the short-term count of one calendar day. For each of the durations less than one day, the number of scenarios is the same. For durations that span two, three, and seven days, the number of unique scenarios was limited because of the restriction of keeping the evaluation within a calendar year. For instance, the three-day evaluation could be evaluated from January 1 through December 29, but not for December 30 and 31, as this would require extending into another calendar year. Therefore, the three-day evaluation in a year with 365 days is based on 363 days. A similar restriction applies to the 24-hour, 48-hour, and 72-hour durations. Additionally, as these scenarios involved a randomly assigned starting time to each day, up to 24 different values could have applied to any particular starting day. For these evaluations, only one randomly determined starting time was applied to each day.

Short-term Count Duration and Factor Groupings

Bringing together the weighting factors and the short-term durations, the following methodology applies:

Estimate AADT from six-hour counts:

1. For one of the k_c sites within a factor group, select six hourly counts (6 AM – 9 AM and 3 PM – 6 PM).
2. Use the hour of day factors corresponding to the factor group excluding the k_c site and the six hourly counts to estimate traffic volume for the day. Traffic volume for the day=(total traffic volume of the six hours)/(total percentage of six hours traffic volume for the day of week group). For example, if the total of the six hourly counts is 676, and the total of the corresponding six hourly percentages from the hour of day factors for that day of week group is 30.78 percent, then the estimated daily volume is 676/0.3078 = 2196.
3. Apply the appropriate monthly and day of week group factors specific to the factor group excluding the k_c site to get an estimated AADT. For example, if the previously identified count of 676 was obtained on Saturday, January 1, 2000, the estimated daily volume was 2196, the monthly factor for January is 1.1389 and the day of week factor for Saturday is 1.039, then the estimated AADT = 2196*1.1389*1.039=2599.
4. Since the k_c sites have a known AADT, the error in estimation can be calculated as the percent difference between the calculated AADT and the true AADT for the site (based on the simple average of all hourly counts in the year).
5. Repeat steps 1 – 3 sequentially for each day of the year.

Estimate AADT from 12-hour counts:

1. Separately calculate for 6 AM – 6 PM or 9 AM – 9 PM.
2. The steps are similar to estimating AADT from six-hour counts, except that the specific 12 hours are used.

Estimate AADT from 24-hour (single-day) counts:

1. For the full 24-hour period selected, apply just the monthly and day of week group factors to get estimated AADT.
2. Since the k_c sites have a known AADT, the error in estimation can be calculated as the percent difference between the calculated AADT and the true AADT for the site (based on the simple average of all hourly counts in the year).
3. Repeat steps 1 – 2 sequentially for each day of the year.

Estimate AADT from full day two-, three-, and seven-day counts:

1. Average the requisite number of separate 24-hour AADT estimates as calculated from the single day procedure above to get short-term AADT.
2. Repeat steps sequentially for each day of the year until there are not sufficient additional days in the year (December 26 or later depending on the length of the counts).

Estimate AADT from 24-hour (two-day) counts:

1. Randomly select a starting hour of the day, which then establishes the last hour as one hour earlier on the next day (e.g., 6:00 AM start on Monday means that the 24 hourly counts from 6:00 AM on Monday through 5:00 AM on Tuesday are used).
2. Use the hours from the first day to estimate traffic volume for the first day, similar to using six-hour counts to estimate traffic volume for the day. Then complete the same process for the second day.
3. Applying monthly factor and day of week group factors separately to each day's AADT. One is based on the first day and the other one is based on the second day.
4. Calculate a weighted average of AADT based on the two estimated AADT. For example, if the first day has 16 hourly counts (i.e., hour nine (starts at 8:00 AM) through hour 24 (starts at 11:00 PM)) and the second day has eight hourly counts (i.e., hour one (starts at Midnight) through hour

eight (starts at 7:00 AM)), the weighted average will be 16/24*AADT based on the first day + 8/24*AADT based on the second day.

5. Repeat steps 1 through 4 for each day of the year, except for the last day of the year.

Estimate AADT from 48-hour (three-day), 72-hour (four-day), and 168-hour (seven-day) counts:

1. Use the same process as the 24-hour (two-day) counts except that the weighting is calculated for all applicable days.
2. Repeat steps sequentially for each day of the year until there are not sufficient additional days in the year (December 25 or later depending on the length of the counts).

Calculation and Characterization of Bias (Accuracy) and Precision

The bias incurred by any particular short-term count duration, factor group, and factoring method was calculated as a simple percent change:

$Bias_i = 100*(AADT_i – AADT_{true})/AADT_{true}$

Where i is a short-term count for a site within a factor group calculated under a given factoring method

The distribution of the $Bias_i$ terms was then characterized with a set of descriptive statistics. Both the mean and standard deviation of the bias terms were calculated and reported. However, due to the potential for the bias distribution to show some skewness, the median and the 2.5th and 97.5th percentiles of the distribution were the statistics selected for graphical presentation. The median is a measure of the central tendency of the bias results and is not sensitive to a small number of possibly extreme values as the mean. A median bias less than zero indicates a scenario that prevalently underestimates AADT, while a median greater than zero indicates a scenario that prevalently overestimates AADT. The bias estimates will vary from day to day throughout a year. To characterize the magnitude of this variability, the difference between the 97.5th and 2.5th percentile estimates is a useful statistic. This range provides the observer with 95 percent probabilistic confidence of bracketing the bias that exists for a particular short-term count duration scenario. A narrow range for this statistic provides better assurance that the short duration count expanded to an AADT is likely to estimate that AADT closely for the condition.

Sub-setting the Data

Weighting factors were calculated for all days in a year, and subsequent short-term duration counts expanded to estimated AADT were produced for every day of the year possible. However, many of these simulated AADT values would never be produced in practice. For instance, it is highly unlikely that a state DOT would take a temporary count at a location on Christmas Day. To prevent the introduction of bias into the results for the analysis, FHWA elected to have certain days removed from the tabulation of summary results. The list for the removals is provided as Appendix A, and is comprised of the Federal holidays from the years 2000 through 2013. It is noted that certain Federal holidays are not recognized in some states (e.g., Columbus Day). Additionally, local holidays (e.g., Mardi Gras in New Orleans) or non-Federal holidays (e.g., Easter) may impact decisions for whether to perform short-term counts. However, adjustment of results at this level was determined to be beyond the scope of the analysis, so the Federal holiday adjustment is the only removal from the results.

Beyond the holiday removal, there is a basic research question of how bias differs by days of the week. As such, the analysis is further summarized into the following five groups:

- All temporary counts are included, except those that would overlap with a holiday. These results are labelled as “All”.
- Temporary counts are included so long as the days are within Monday-Thursday, and assuming no holidays within the time span. These results are labelled as “Monday-Thursday”.
- Temporary counts are included that contain Friday, and assuming no holidays within the time span. These results are labelled as “Friday”.

- Temporary counts are included that contain Saturday, and assuming no holidays within the time span. These results are labelled as “Saturday”.
- Temporary counts are included that contain Sunday, and assuming no holidays within the time span. These results are labelled as “Sunday”.

Understanding these summaries is critical to understanding the results. For the Monday-Thursday results, and assuming holidays are already removed, these results will include every six-hour, 12-hour, and one-day estimate where the day of the estimate is either Monday, Tuesday, Wednesday, or Thursday. For the 24-hour estimate, since the 24 hours will usually span two calendar days, the results are only provided where the two days spanned are Monday-Tuesday, Tuesday-Wednesday, and Wednesday-Thursday. Hence, 24-hour results for Sunday-Monday and Thursday-Friday, while partially falling within the period or interest, are not included. This particular reporting division was selected based on a group consensus within the TAC that it was most representative of actual DOT field practices (though not necessarily universally representative). The inclusion of the Friday, Saturday, and Sunday divisions of data were more exploratory, as it is suspected that they are rarely done in practice. However, there was an interest in seeing whether there were statistical reasons for this. All three days have the same methodology. For instance, a “Friday” count for the six-hour, 12-hour, and one-day estimate have to be obtained only on a Friday. A two-day count could either be Thursday-Friday or Friday-Saturday, and so forth.

In the case of the weekly count duration, the results for the different days all collapse to a single outcome. Every seven-day period is included which does not include a holiday. As such, there is no difference between the subset of data used for the “All”, “Monday-Thursday”, “Friday”, “Saturday”, and “Sunday” results.

Additional subsets of data were examined by individual day of week where the day was simply the starting day of the week. For Monday, this would include all count durations of one day and less that occurred on Monday. For 24 hours are longer durations, these results would only include estimates where the first portion of the data for the short-term count occurred on Monday. This could be a seven-day count that was Monday through Sunday, or a 24-hour count that started on Monday at 6:00 PM. Conversely, a 24-hour count that started at 6:00 PM on Sunday would not count as a Monday result, despite most of the hours of that count being on Monday itself.

When looking at sub-setting for month of year, it was decided to assign each result to one, and only one, month. If a result spanned more than one month (e.g., two-day count started on the last day of one month and completed on the first day of next month), some determination was required for how to assign it. The following algorithm was used to determine how cross-month summaries would be assigned:

- 24 hours, two-day – assign to starting month;
- 48 hours, three-day, and 72-hour – assign to month of second day; and
- One week – assign to month of fourth day.

Statistical Methods for Comparing Bias (Accuracy) and Precision

An additional research objective was to compare the level of bias and precision between the primary short term sampling durations of interest; 24 hour and 48 hour. From all the bias and precision estimates discussed above, this evaluation was completed with those that were based on Monday through Thursday counts, did not include federal holidays, and were derived from separate day of week and month of year group factors. Both the clustering algorithm and functional classification groupings were evaluated. Two important statistics generated from these data were the difference between the 48 and 24 hour duration median percent bias values across days and sites, where the median bias is a measure of the accuracy of the estimates, and the percentage reduction in the width of the 95 percent interval for percent bias between the 48 and 24 hour duration results, where the width of the 95 percent interval for percent bias is a measure of precision.

To perform a statistical evaluation of the two key statistics, a bootstrap approach was utilized. This was necessary because of the high degree of correlation between sample results since a particular site, year, and day result for a 24 hour count duration was nearly always matched with a similar 48 hour count. To perform the bootstrap, the aggregate set of 24 and 48 hour duration percent error from true AADT estimate results were randomly resampled with replacement. For each realization of sample, the statistics of differences in median percent bias and percent reduction in 95 percent confidence interval widths, were calculated for each of the durations. This process was repeated 1,000 times, and the final set of 1,000 outcomes were ranked and used to generate 95 percent confidence intervals of each statistic as the 2.5th and 97.5th percentile values from the bootstrapped resamples. If these intervals overlap “0”, they do not provide adequate evidence (Type 1 error rate of 5 percent) of a difference between the 24 and 48 hour results. If the intervals do not overlap “0”, they provide evidence (with no larger than a five percent risk of error) that there is a statistically significant difference between the accuracy and precision of 24 and 48 hour counts to estimate AADT. The results were further evaluated to determine a p-value if the results were significant.

The statistical analysis was conducted in SAS ® v 9.3. Bootstrap samples were generate with the SurveySelect procedure.

Results

The complete results for this evaluation are made up of a percentage bias in AADT between the short-term count and the true AADT for the following sets of conditions:

- 206 sites with complete hourly data for the functional classification analysis;
- 192 sites with complete hourly data for the volume range analysis; and
- 201 sites with complete hourly data for the clustering algorithm analysis.

Each of the above is evaluated for ten separate short-term count durations, extending from the first day of the calendar year to the last one for which the duration fits into the year, and under three separate weighting factor methods.

The aggregate number of records estimated above is in the millions. These data have been provided to FHWA as a separate SAS dataset. Several important summaries of these records will be discussed here and provide the basis for examining the research questions of interest in this task.

National Summary: Accuracy and Precision of AADT Estimation by Count Duration

Under the site groups defined by state, year, and roadway functional classification, the bias results (see All Days section of Table 2) show that extending the duration of the short-term count has little impact on the median accuracy of estimating AADT. After removing AADT errors related to counts taken over a Federal holiday, results from the analysis show a median over-count from the short-term counts of 0.57 to 1.20 percent, depending on count duration. While median error does not exhibit a strong relationship with count duration, the variability of the errors (i.e., precision) is reduced as count duration increases. A 95 percent confidence interval on error in AADT estimation based on a six-hour count is (-32.41, +42.97) percent, while the same interval on a seven-day count is (-17.85, +24.12). The results of this analysis are depicted graphically in Figure 4. While there is a large variability in precision of AADT estimation between the shortest (i.e., six-hour) and longest (i.e., seven day) durations, the two durations most frequently encountered in practice (one-day, and two-day – or 24 hours and 48 hours) show very little difference. At (-24.13, +32.49), the 95 percent confidence interval on error is only slightly smaller for the two-day count duration than the one-day (-28.03, +36.06). The difference in total uncertainty range between these two durations is only about 12 percent.

Throughout the results section, the only factoring method that will be detailed is the one which featured separate day of week and month of year factors. This is due to the very minor differences observed with results that include either of the other two factoring methods (i.e., day of week by month factoring and day of year factoring). However, complete results that include these methods are provided in the Appendix for reference.

National Summary: Accuracy and Precision of AADT Estimation by Day of Week

The national results in Figure 4 are based on converting short-term count duration for every day of the year that is not a Federal holiday, to an AADT. In practice, though, short-term counts are most often performed on weekdays. To examine the merits of this practice from a statistical accuracy and precision perspective, the previous functional classification grouping results were subset to the day of the week and these are shown in Figure 4. To simplify the interpretation, only the one day duration is shown, though results are provided for all durations in the appendix. From Figure 4, the important observations are the following:

- Use of Thursday counts to estimate AADT produces the most biased median error, at 3.4 percent. The second largest error is attributed to Monday, at 1.83 percent. In both cases, the AADT estimates tend to over-count.
- All weekdays show fairly similar total uncertainty in AADT estimation as provided by the 95 percent confidence interval. The 95 percent confidence interval range (i.e., 97.5th percentile – 2.5th percentile) does not exceed 60 percent for any of the five weekdays. The range is slightly smaller on the two days with larger median bias (Monday, Thursday), but only slightly.
- Both Saturday and Sunday show larger range of errors, over 70 percent for Saturday, and over 90 percent for Sunday.

These results support the statistical rationale for obtaining counts only on weekdays. This practice will lead to less error in the AADT values to be estimated.

National Summary: Accuracy and Precision of AADT Estimation by Count Duration – Restricted Days

In discussions with the TAC for this task, it was found that common practice was not to obtain short-term counts on Friday, Saturday, or Sunday. This was not universal, but common enough to warrant a decision that results for this analysis be subset to just short-term counts that were contained within the Monday through Thursday time period. Accordingly, Figure 6 shows the same results as Figure 4, with the restriction that estimated AADT errors are removed where the sampling period was not fully contained within a Monday through Thursday time period. The Figure 4 and Figure 6 data are also contained within Table 2. The seven-day sampling period is not shown in Figure 6 since it cannot fit into the Monday-Thursday window. The seven-day results can still be referenced for comparison in Figure 4, if desired. For the remainder of the sampling durations, the 95 percent confidence interval range of errors is universally smaller for the Monday-Thursday sampling as compared to sampling that includes all days of the week. This reflects the fact that the higher variability results for Saturday and Sunday have been excluded from the estimates.

For count durations of one day and less, the median error is actually higher for the Monday-Thursday counts. While this is not desirable, the over-count values for the Monday-Thursday sampling are within 0.5 percent of sampling on all days in all cases except the six-hour duration, and even here the loss of accuracy is less than a percent. This effect occurs because the short-duration periods for Monday-Thursday include the two days of the week that show the greatest median bias in errors, Monday and Thursday. For durations that span more than one day, the results for Monday and Thursday are less prominently weighted since durations longer than one day contained within Monday-Thursday must all contain a Tuesday or Wednesday, while some of these durations may exclude Monday (e.g., three days from Tuesday-Thursday) or Thursday (e.g., three days from Monday-Wednesday).

For the restricted days scenario, a 95 percent confidence interval on error in AADT estimation based on a six-hour count is (-30.36, +39.94) percent, while the same interval on a 72-hour count is (-21.98, +26.24). The results of this analysis are depicted graphically in Figure 6. While there is a large variability in precision of AADT estimation between the shortest (i.e., six-hour) and longest (i.e., 72-hour) durations, the two durations most frequently encountered in practice (one-day, and two-day – or 24 hours and 48 hours) show very little difference. At (-22.53, +28.28), the 95 percent confidence interval on error is only slightly smaller for the two-day count duration than the one-day (-24.31, +30.51). The difference in total uncertainty range between these two durations is only about seven percent.

Due to the TAC agreement that Monday-Thursday sampling was most frequent, subsequent summary results in this report are restricted to these days. However, the more detailed estimates that include all days have been provided to FHWA.

Table 2. Comparison of Median and 95 Percent Confidence Interval (CI) Error on AADT Estimation for Functional Classification Factoring Between All Days and Monday-Thursday Counts, Excluding Holidays – Reference Figures 5 and 7

Duration	All Days		Monday-Thursday	
	Median	95% CI	Median	95% CI
Six Hours	1.16	(-32.41, 42.97)	1.89	(-30.36, 39.94)
Twelve Hours 6AM-6PM	1.20	(-29.72, 38.92)	1.64	(-24.72, 31.87)
Twelve Hours 9AM-9PM	0.97	(-30.29, 38.90)	1.45	(-26.79, 32.45)
One Day	0.91	(-28.03, 36.06)	1.35	(-24.31, 30.51)
Twenty Four Hours	0.57	(-27.06, 35.27)	0.40	(-24.91, 30.46)
Two Days	0.81	(-24.13, 32.49)	0.79	(-22.53, 28.28)
Forty-Eight Hours	0.70	(-23.53, 31.73)	0.13	(-23.19, 27.36)
Three Days	0.79	(-21.43, 29.61)	0.67	(-21.48, 26.95)
Seventy-Two Hours	0.73	(-21.09, 28.97)	0.24	(-21.98, 26.24)
Seven Days	0.89	(-17.85, 24.12)	N/A	N/A

Source: Battelle

National Summary: Accuracy and Precision of AADT Estimation by Functional Classification

The results provided in Figure 6 are averaged across a large number of individual sites in different states and years and of different functional classifications. Figure 7 separates these national level results for the one day duration to each of the nine functional classifications available for use in the study. This figure shows that accuracy of AADT estimation, on median, ranges from as much as 3.17 percent underestimation (FC 5U) to 4.28 percent overestimation (FC 2U). Functional Classification 5U is the only one that showed median underestimation of the AADT. With respect to precision, the variability of AADT errors is larger for the rural designation of each classification number than for the urban designation.

Functional Classification Group Summary: Precision of AADT Estimation

To understand the extent of variability observed, Figure 8 and Figure 9 show the functional classification grouping results at the group level. Each group is a state, year, and functional classification. Figure 8 shows sampling durations of one day and lower, while Figure 9 shows sampling durations of one through three days. The results are the detailed grouping data that underlie Figure 6. The plotted response variable in Figure 8 and Figure 9 is half the difference between the upper 97.5th percentile and the 2.5th percentile of the error distribution. While the half width of the 95 percent confidence interval on the national level summary in Figure 6 for one day sampling duration was 27.41 percent (i.e., [30.51-(-24.31)]/2), the half width for an individual group could be as low as about seven percent (1U, Hawaii, 2010) or as high as about 71 percent (5R, Vermont, 2012), with most one day group results in the 10 to 30 percent range. The very large error for the Vermont 5R sites in 2012 underscores the importance of associating a count location to a properly matched set of permanent count stations. Further review of this particular factor group revealed that the functional classification group included a mix of sites with and without a highly seasonal (possibly recreational) pattern of volumes. This resulted in some days used as temporary counts having a highly overestimated AADT upon application of the monthly factors.

Another important observation from Figure 8 and Figure 9 is that the pattern of reduced error associated with longer duration, but limited reduction for the most typical durations (i.e., one and two day), holds at the group level as well as it did at the national level. Across the plotted symbols within each figure for a particular group, the half-widths are generally reduced for longer durations, but only by a small amount for similar durations.

National Summary: Accuracy and Precision of AADT Estimation by Volume Range

Many states develop factor groups for their temporary counts using permanent count stations of the same functional classification. This fact led to all the results already provided. However, there are other methods that can be used to develop factor groups. A more simplistic method is to group just by the AADT volume itself. The analysis already described was completed in this manner with factor groups in each state and year based on the following ranges of AADT (0<AADT<1000, 1000≤AADT<10,000, 10,000≤AADT<100,000, and AADT≥ 100,000). The national results for this analysis are comparable to what was shown in Figure 6 is provided in Figure 10. The results are summarized in Table 3 below. They show very little difference in the overall accuracy and precision of AADT estimation.

Table 3. Comparison of Median and 95% Confidence Interval (CI) Error on AADT Estimation for Functional Classification Factoring and Volume Range Factoring, Monday-Thursday Counts, Excluding Holidays – Reference Figures 6 and 10

Duration	Functional Classification		Volume Range	
	Median	95% CI	Median	95% CI
Six Hours	1.89	(-30.36, 39.94)	1.86	(-31.52, 39.54)
Twelve Hours 6AM-6PM	1.64	(-24.72, 31.87)	1.63	(-26.20, 32.68)
Twelve Hours 9AM-9PM	1.45	(-26.79, 32.45)	1.63	(-28.04, 33.38)
One Day	1.35	(-24.31, 30.51)	1.43	(-25.43, 31.39)
Twenty Four Hours	0.40	(-24.91, 30.46)	0.52	(-25.60, 31.51)
Two Days	0.79	(-22.53, 28.28)	0.89	(-23.32, 29.00)
Forty-Eight Hours	0.13	(-23.19, 27.36)	0.29	(-23.83, 27.96)
Three Days	0.67	(-21.48, 26.95)	0.75	(-22.01, 28.12)
Seventy-Two Hours	0.24	(-21.98, 26.24)	0.32	(-22.46, 25.86)

Source: Battelle

The accuracy and precision of AADT estimation with volume range factoring was reviewed at the volume range level, with results shown in Figure 11 for one day duration sampling. These results are comparable to the functional classification results of Figure 7. The results here very clearly show a relationship between the AADT itself and the precision of estimation. Sites with lower volumes are subject to much larger errors in estimation, as a percentage, than those with larger volumes.

National Summary: Accuracy and Precision of AADT Estimation by Clustering Algorithm

The functional classification and volume range groupings are methods that use higher-level summary information about a permanent count site to associate it with a temporary count location. Another recommended FHWA method is to group sites by the pattern of volume measurement over time, potentially as fine as at an hourly level. This approach is detailed in the TMG (p. 3-10, Determining Monthly Pattern Groups by Cluster Analysis). A clustering algorithm approach was applied to this task in which sites were associated to factor groups by the aggregate of their day of year, day of week, and hour of day volume patterns. The national level results of AADT estimation with this form of factoring are provided in Figure 12. These results can be compared to the same presentations shown in Figure 6 (Functional Classification Groups) and Figure 10 (Volume Range Factor Groups). The very apparent result of this comparison is that the more mathematically rigorous method of associating short-duration count locations to factor groups by the clustering algorithm substantially improves the precision of AADT estimation, and results in most durations on a less biased median estimation. In can be calculated from

Table 4 that the 95 percent confidence interval widths are 26 to 38 percent smaller with the clustering algorithm factoring method than for the functional classification factoring method.

Table 4. Comparison of Median and 95 Percent Confidence Interval (CI) Error on AADT Estimation for Clustering Algorithm Factoring in Comparison to Functional Classification Factoring and Volume Range Factoring, Monday-Thursday Counts, Excluding Holidays – Reference Figures 6, 10, and 12

Duration	Functional Classification		Volume Range		Clustering Algorithm	
	Median	95% CI	Median	95% CI	Median	95% CI
Six Hours	1.89	(-30.36, 39.94)	1.86	(-31.52, 39.54)	2.36	(-21.89, 29.05)
Twelve Hours 6AM-6PM	1.64	(-24.72, 31.87)	1.63	(-26.20, 32.68)	1.92	(-17.98, 23.46)
Twelve Hours 9AM-9PM	1.45	(-26.79, 32.45)	1.63	(-28.04, 33.38)	1.22	(-17.95, 23.22)
One Day	1.35	(-24.31, 30.51)	1.43	(-25.43, 31.39)	0.96	(-16.55, 19.48)
Twenty Four Hours	0.40	(-24.91, 30.46)	0.52	(-25.60, 31.51)	-0.06	(-17.41, 19.01)
Two Days	0.79	(-22.53, 28.28)	0.89	(-23.32, 29.00)	0.48	(-15.33, 17.18)
Forty-Eight Hours	0.13	(-23.19, 27.36)	0.29	(-23.83, 27.96)	-0.22	(-15.78, 16.24)
Three Days	0.67	(-21.48, 26.95)	0.75	(-22.01, 28.12)	0.35	(-14.18, 16.37)
Seventy-Two Hours	0.24	(-21.98, 26.24)	0.32	(-22.46, 25.86)	-0.14	(-14.57, 15.16)

Source: Battelle

While the clustering method provided more precise AADT estimation for all durations, and more accurate for most, it should be noted that the full benefit of this type of factoring method likely cannot be realized outside of this evaluation. The reason is that temporary count locations in practice will not have the day of week and day of year data that were available in this evaluation, and which were used to assign each permanent count station acting as a temporary count station to an optimally matched set of permanent count stations. However, further research may be able to leverage the hour of day data for a temporary count in conjunction with other characteristics of the temporary count station to dynamically assign it a factor group with greater expected match in expected volume to ultimately provide a more accurate and precise AADT estimate.

National Summary: Accuracy and Precision of AADT Estimation by Month

Monthly factors are developed and applied to temporary count data, which should assure that the accuracy of estimating AADT is independent of month of year. Using the clustering algorithm results of one-day duration on Monday-Thursday counts (one day results in Figure 12 as the basis), this was evaluated and the results shown in Figure 13. There is an obvious pattern in these results of median AADT estimates overstating the true values in the months of November, December, January, and February. This may be a result of removing all estimates for Federal holidays. The volumes for Federal holidays are included in the development of the monthly factors. If actual traffic volumes on these days are lower than other days, failure to include them in the AADT estimates will have the effect of biasing the AADT errors on the high side.

Even if the factoring produced AADT estimates that were unbiased for each month of the year, there could still be different precision in AADT estimation between months of the year. This could be due to greater relative variability in traffic volumes across the days of one month relative to another. Indeed, this appears to be the case, as the precision of estimating AADT is best in the April-September time period, slightly more variable in March and October, and markedly more variable in November, December, January, and February. To visualize the potential impact of this effect as a function of the number of days sampled, Figure 14 repeats the data points of Figure 13 (symbols now reduced to outlines) and superimposes the two day duration data (shown as blue diamonds). Consistent with previously discussed results at the national level (see Figure 12), the short-term count duration of two days provides very little improvement in accuracy or precision, here at the monthly level.

National Summary: Accuracy and Precision of AADT Estimation by Year

To perform this evaluation, data were obtained for states and sites across 13 years from 2000 through 2012. Using the clustering algorithm results of one day duration on Monday-Thursday counts (one day results in Figure 11 as the basis), results were separated by year and are shown in Figure 15. Since the numbers of sites by state and functional classification are not consistent from year to year, it is expected that the accuracy and precision by year would vary. However, this presentation looks for any systematic time trend across the duration of the evaluation period that could suggest issues with data quality. No such pattern is apparent as the accuracy and precision vary a little from year to year, but not to a repeatable pattern.

National Summary: Statistical Comparisons of Accuracy and Precision of AADT Estimation between 24- and 48-Hour Count Durations

Particular interest exists for the comparison of 24- and 48-hour duration short-term counts. In making this comparison, both the accuracy and the precision are of interest. With respect to accuracy, a statistical evaluation was done for the 24-hour and 48-hour duration count data at the national level under the clustering algorithm approach and restricted to Monday-Thursday counts. Using a statistical bootstrap procedure, the median bias across all sites and days of year was very slightly negative (-0.06 percent) for the 24-hour count, though not statistically distinguishable from zero since the 95 percent confidence interval was (-0.17, 0.02). For the 48-hour counts, the median bias was -0.22 percent with 95 percent confidence interval (-0.33, -0.10). For the difference in the day to day 24- and 48-hour bias results, the 48-hour estimate was more biased (-0.15 percent, with 95 percent confidence interval (-0.24, -0.06)). This particular difference was statistically significant ($p<0.001$). Hence, 48-hour counts for Monday through Thursday and excluding holidays, on median, underestimated true AADT slightly more than 24-hour counts.

In regards to precision, as was shown in Table 4 and Figure 12, increasing the short duration count period resulted in less variability in the AADT estimation error. Due to the asymmetric nature of the error variation, and consequent lack of normality, the variability comparison was conducted using a statistic of the width of the 95 percent confidence interval on the bias in the estimated AADT compared to the true AADT. For the 48-hour count duration, the 95 percent confidence interval width of 32.02 percent is 12.1% smaller than the 36.42 percent width of the 24-hour count duration. This 12.1 percent improvement (i.e., reduction) in error is statistically significant ($p<0.001$), with 95 percent confidence interval of (10.4%, 13.8%).

Table 5. Statistical Comparison of Median and 95% CI Error on AADT Estimation at 24 Hour and 48 Hour Durations for Clustering Algorithm Factoring, Monday-Thursday Counts, Excluding Holidays – Reference Figure 12.

		Bias (% Difference from True AADT)		
	Mean	Median	95% CI	Width
Twenty-Four Hour	0.08 (-0.03,0.20)	-0.06 (-0.17, 0.02)	(-17.41, 19.01)	36.42
Forty-Eight Hour	-0.10 (-0.21,0.02)	-0.22 (-0.33, -0.10)	(-15.78, 16.24)	32.02
Difference	-	-0.15	Reduction	12.1%
95% Bootstrap Confidence Interval	-	(-0.24,-0.06)	95% Bootstrap Confidence Interval	(10.4%,13.8%)
	-	p<0.001	-	p<0.001

Source: Battelle

With respect to accuracy, a statistical evaluation was done for the 24-hour and 48-hour duration count data at the national level under the functional classification factoring approach and restricted to Monday-Thursday counts. Using a statistical bootstrap procedure, the median bias across all sites and days of year was positive (0.40 percent) for the 24-hour count, and statistically distinguishable from zero since the 95 percent confidence interval was (0.26, 0.53). For the 48-hour counts, the median bias was 0.13 percent, and not distinguishable from zero with 95 percent confidence interval (-0.01, 0.32). For the difference in the day to day 24- and 48-hour bias results, the 48-hour estimate was less biased (-0.26 percent, with 95 percent confidence interval (-0.40, -0.13)). This particular difference was statistically significant ($p<0.001$). Hence, 48-hour counts for Monday through Thursday and excluding holidays, on median, better estimated true AADT than twenty-four hour counts.

In regards to precision, as was shown in Table 4 and Figure 6, increasing the short duration count period resulted in less variability in the AADT estimation error. Due to the asymmetric nature of the error variation, and consequent lack of normality, the variability comparison was conducted using a statistic of the width of the 95 percent confidence interval on the bias in the estimated AADT compared to the true AADT. For the 48-hour count duration, the 95 percent confidence interval width of 50.55 percent is 8.7 percent smaller than the 55.36 percent width of the 24-hour count duration. This 8.7 percent improvement (i.e., reduction) in error is statistically significant ($p<0.001$), with 95 percent confidence interval of (6.5%, 10.3%).

Table 6. Statistical Comparison of Median and 95% CI Error on AADT Estimation at 24 Hour and 48 Hour Durations for Functional Classification Factoring, Monday-Thursday Counts, Excluding Holidays – Reference Figure 6.

		Bias (% Difference from True AADT)		
	Mean	Median	95% CI	Width
Twenty-Four Hour	0.92 (0.76,1.07)	0.40 (0.26,0.53)	(-24.91, 30.46)	55.36
Forty-Eight Hour	0.58 (0.40,0.77)	0.13 (-0.01,0.32)	(-23.19, 27.36)	50.55
Difference	-	-0.26	Reduction	8.7%
95% Bootstrap Confidence Interval	-	(-0.40,-0.13)	95% Bootstrap Confidence Interval	(6.5%,10.3%)
	-	p<0.001	-	p<0.001

Source: Battelle

Ultimately, both the clustering algorithm and functional classification factoring showed similar results. When restricting the percentage bias differences between calculated and true AADT to Monday through Thursday excluding holidays, the 48-hour counts showed significantly reduced variability (on the order of 10 percent) compared to the 24-hour counts. The center of the distribution of errors was very slightly (on the order of magnitude of 0.2 percent) but still statistically significantly shifted left (toward lower estimated AADT) for the forty-eight hour counts compared to the twenty-four hour counts.

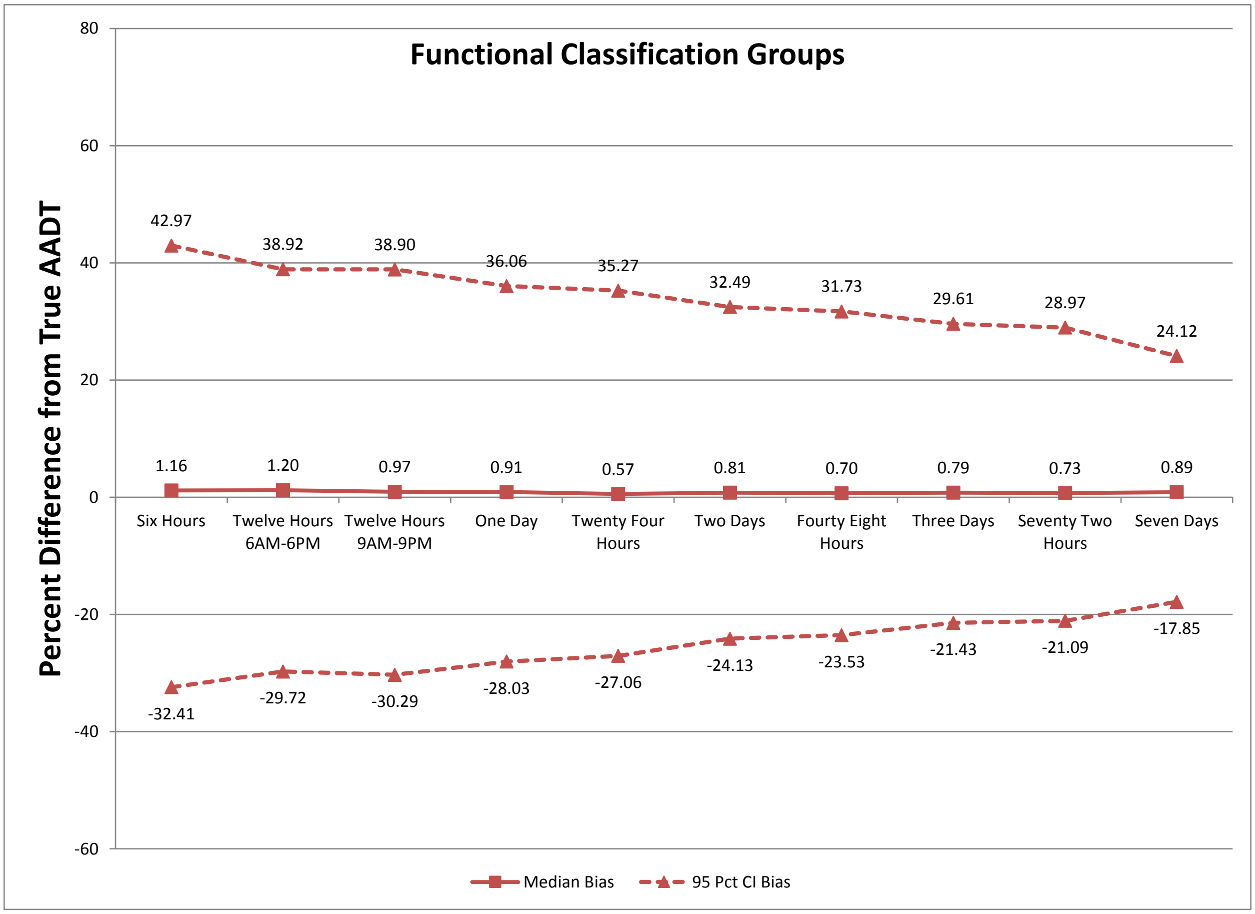

Figure 4. National Level Bias by Short-Term Count Duration for Functional Classification Groups, Separate Day of Week and Month of Year Factors, All Days Excluding Federal Holidays (Source: Battelle, based on data from TMAS)

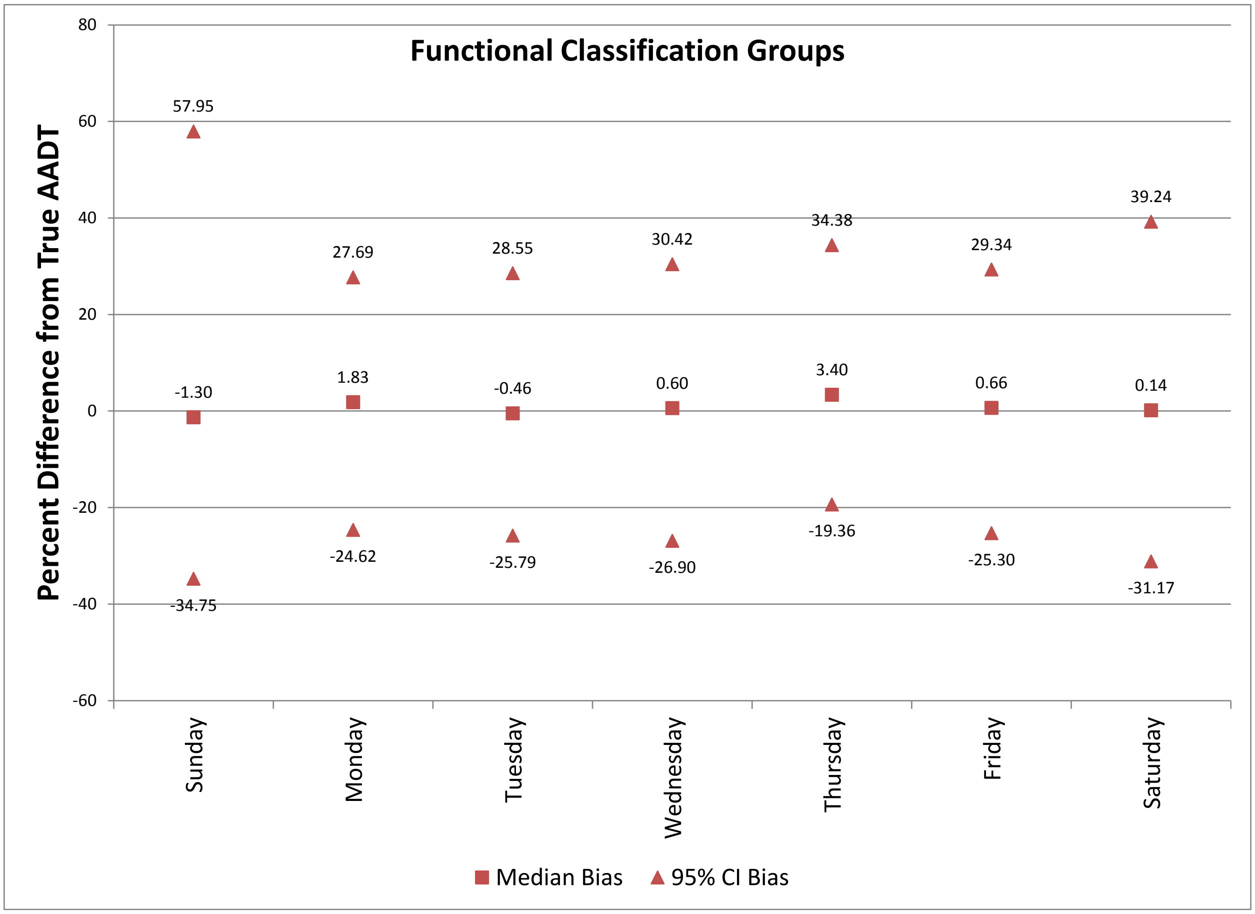

Figure 5. National Level Bias by Day of Week for One Day Count Duration for Functional Classification Groups, Separate Day of Week and Month of Year Factors, All Days Excluding Federal Holidays (Source: Battelle, based on data from TMAS)

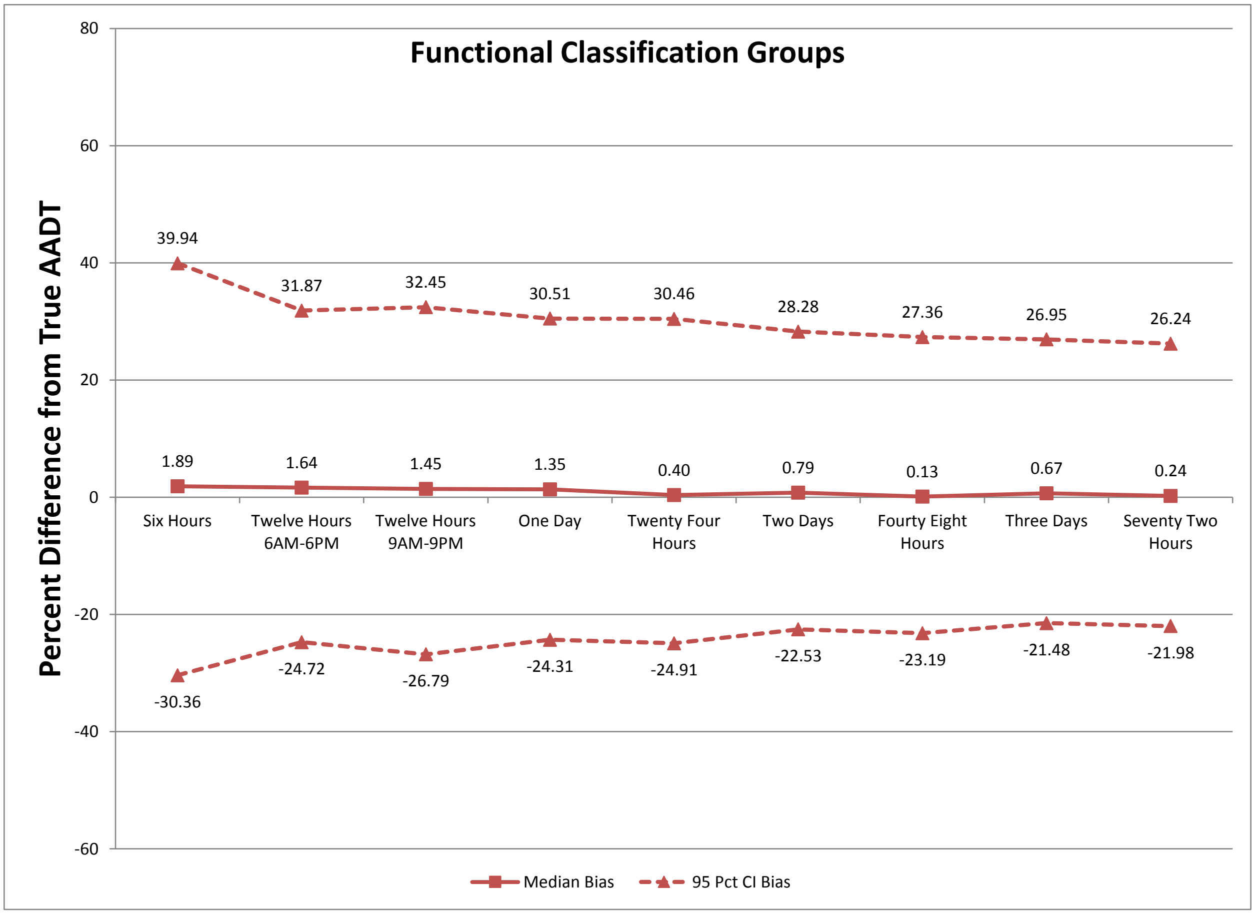

Figure 6. National Level Bias by Short-Term Count Duration for Functional Classification Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

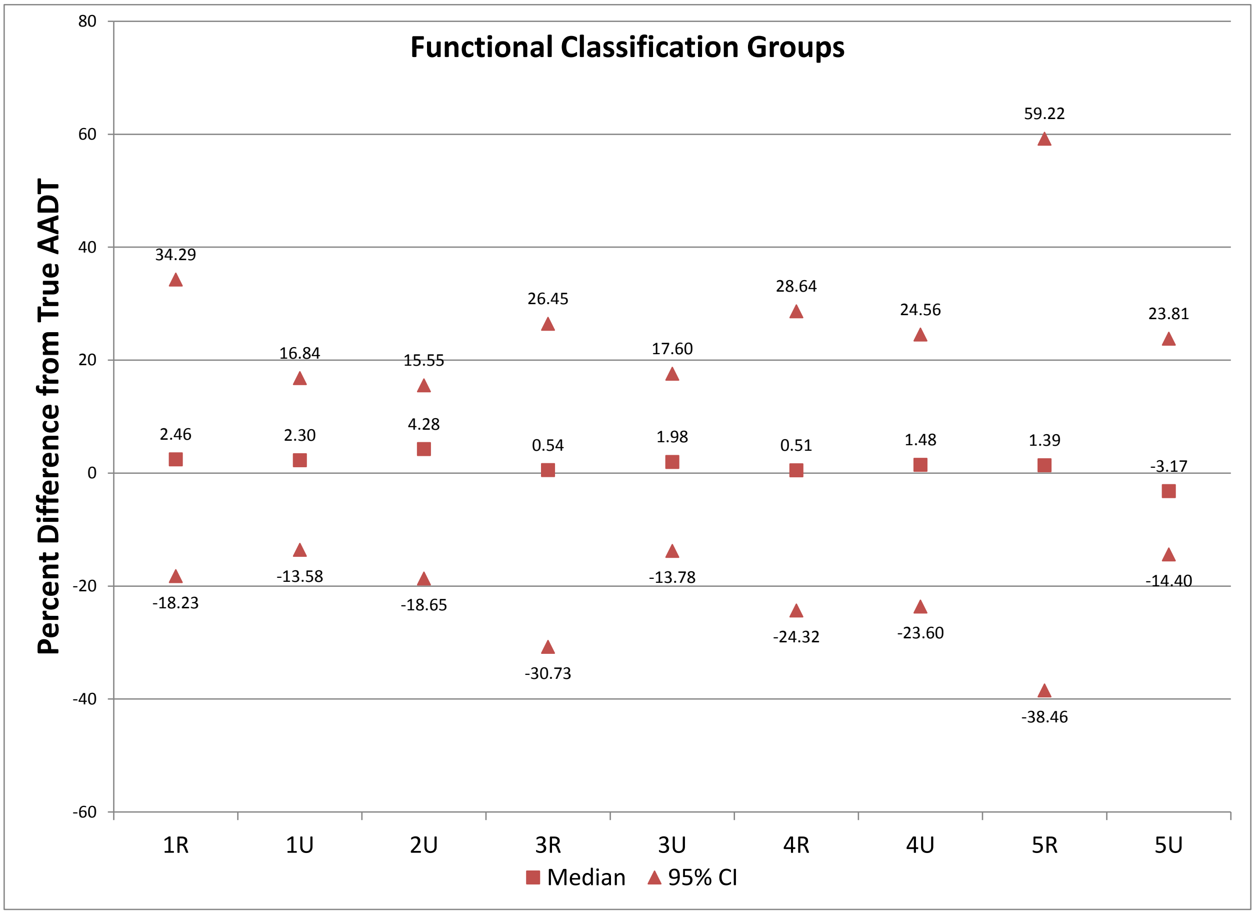

Figure 7. National Level Bias by Functional Classification Groups and Factoring for Short-Term Count Duration of One Day, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts on Federal Holidays (Source: Battelle, based on data from TMAS)

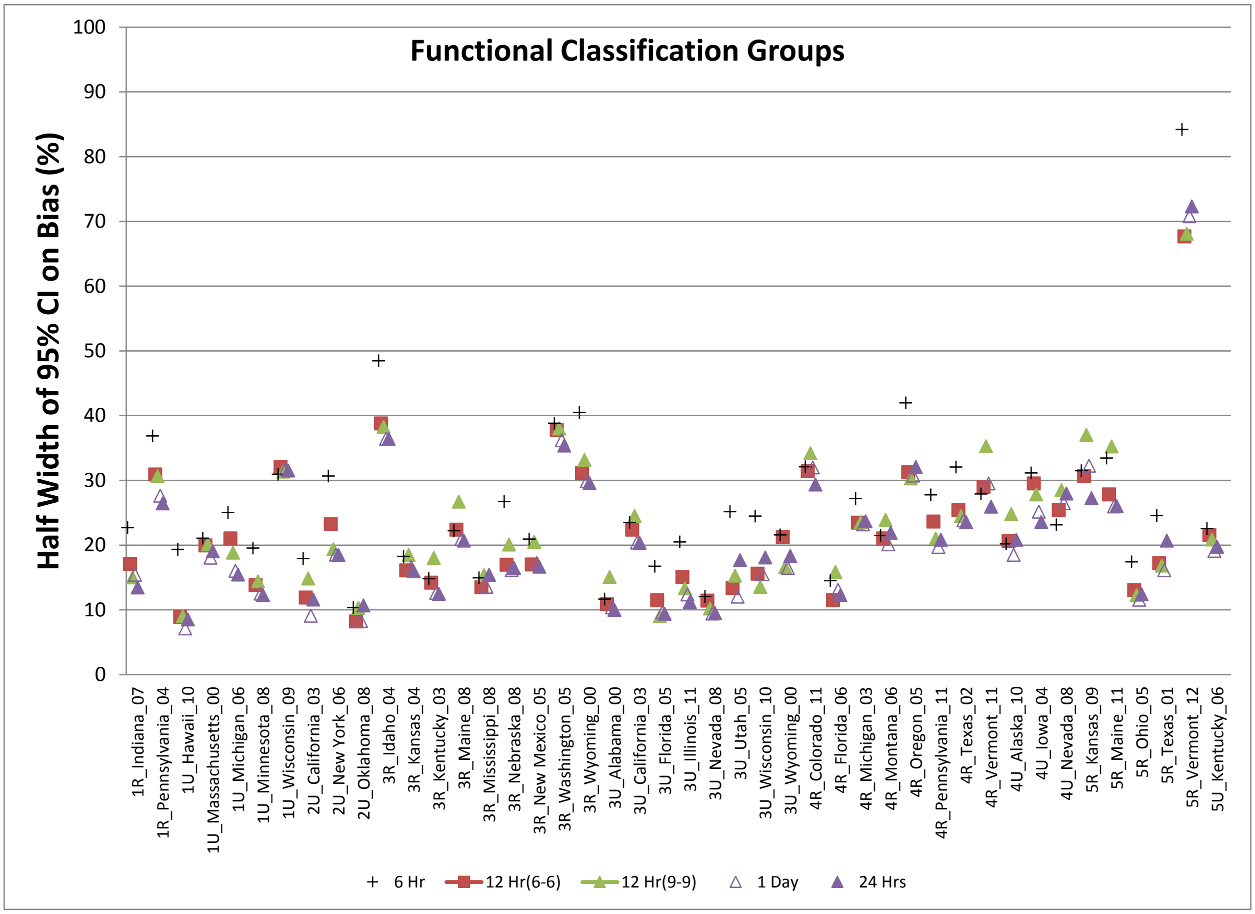

Figure 8. National Level Bias by Short-Term Count Durations ≤1 Day for Functional Classification Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

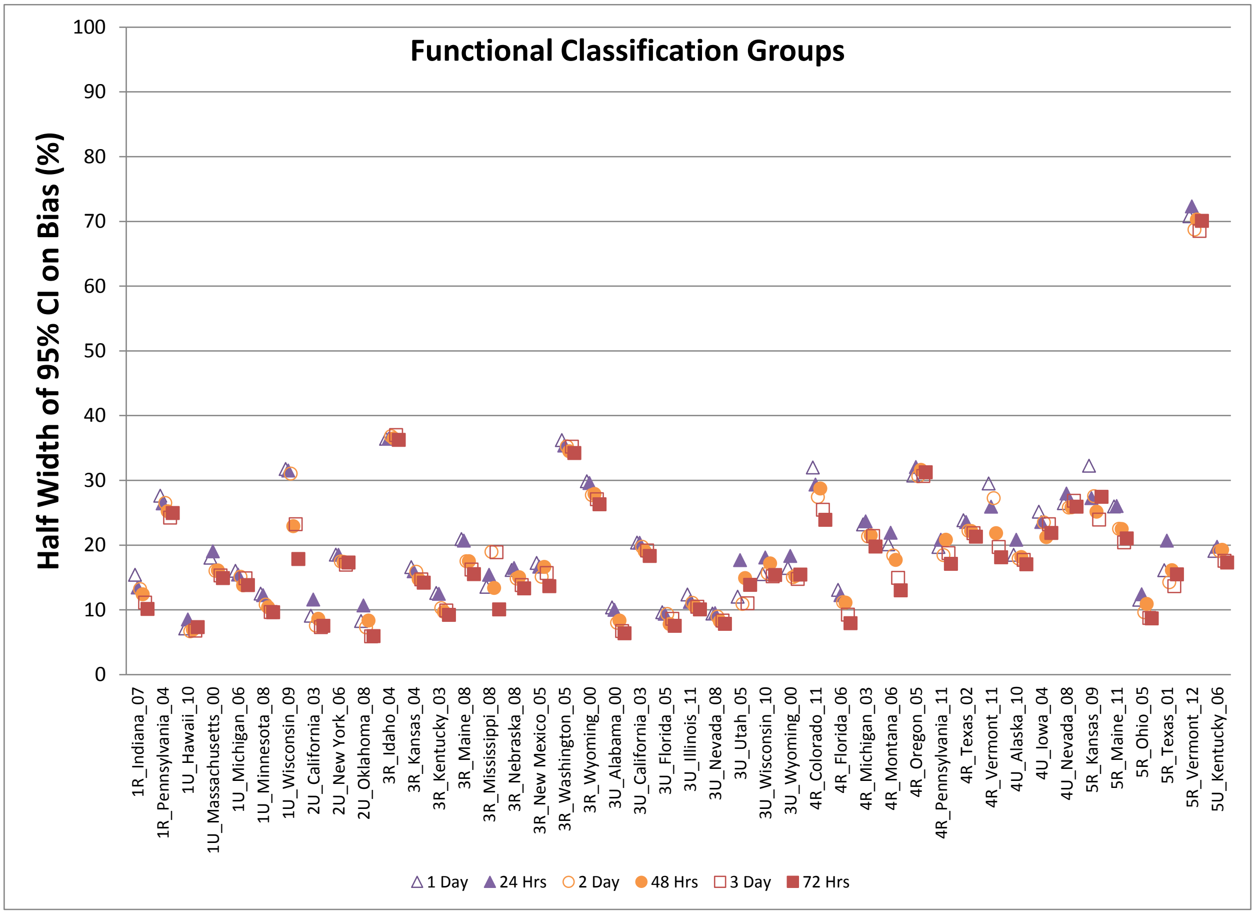

Figure 9. National Level Bias by Short-Term Count Durations ≥1 Day for Functional Classification Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

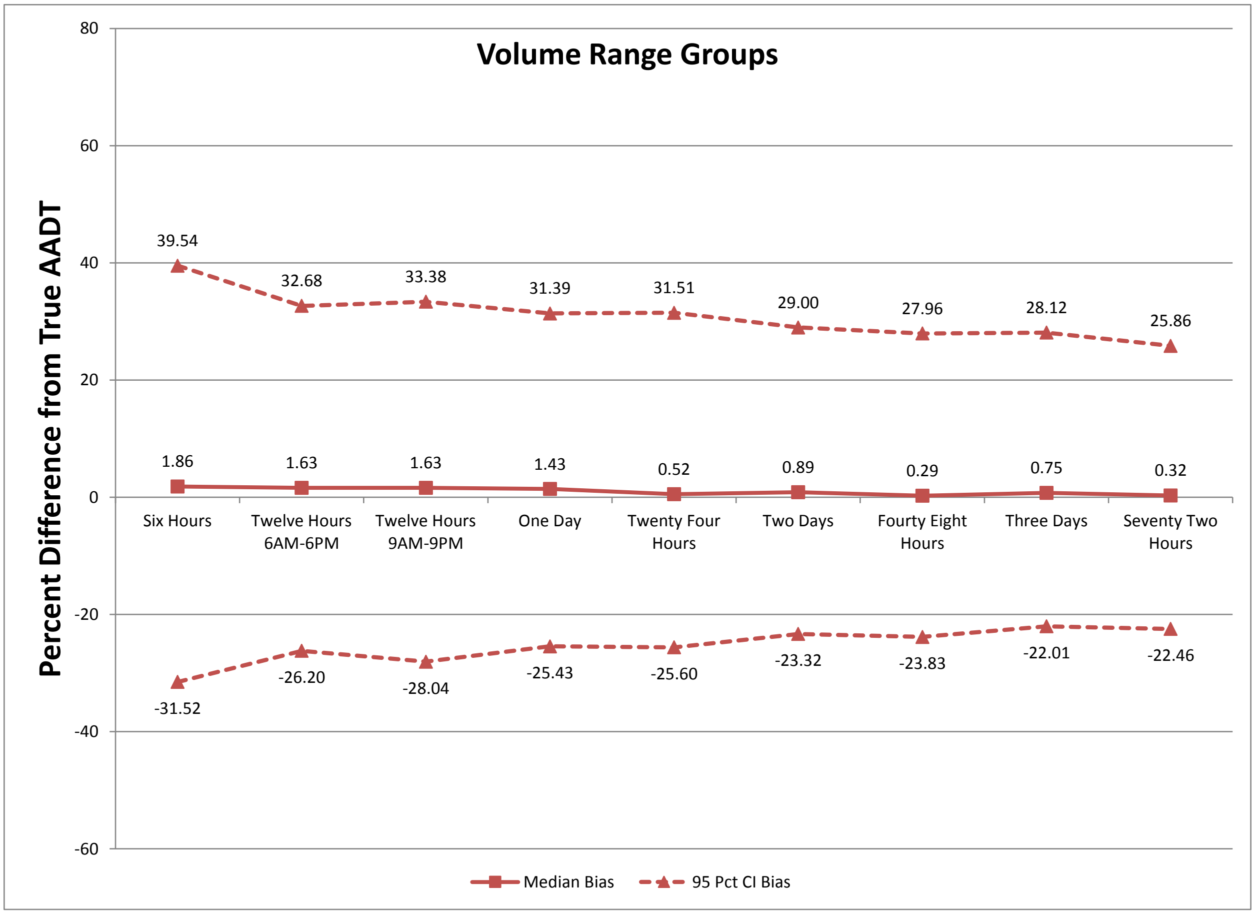

Figure 10. National Level Bias by Short-Term Count Duration for Volume Range Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

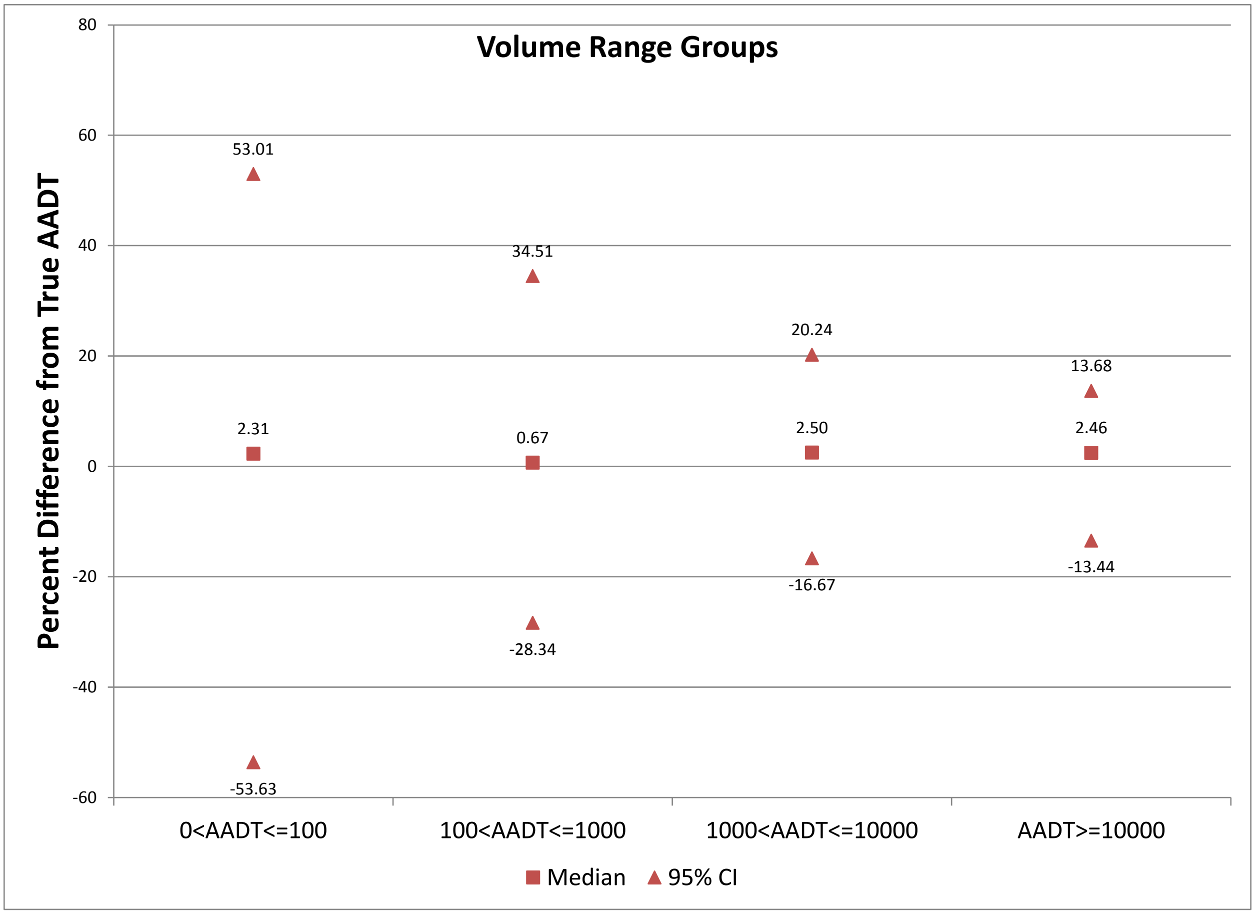

Figure 11. National Level Bias for Short-Term Count Duration of One Day for Volume Range Groups and Factoring, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

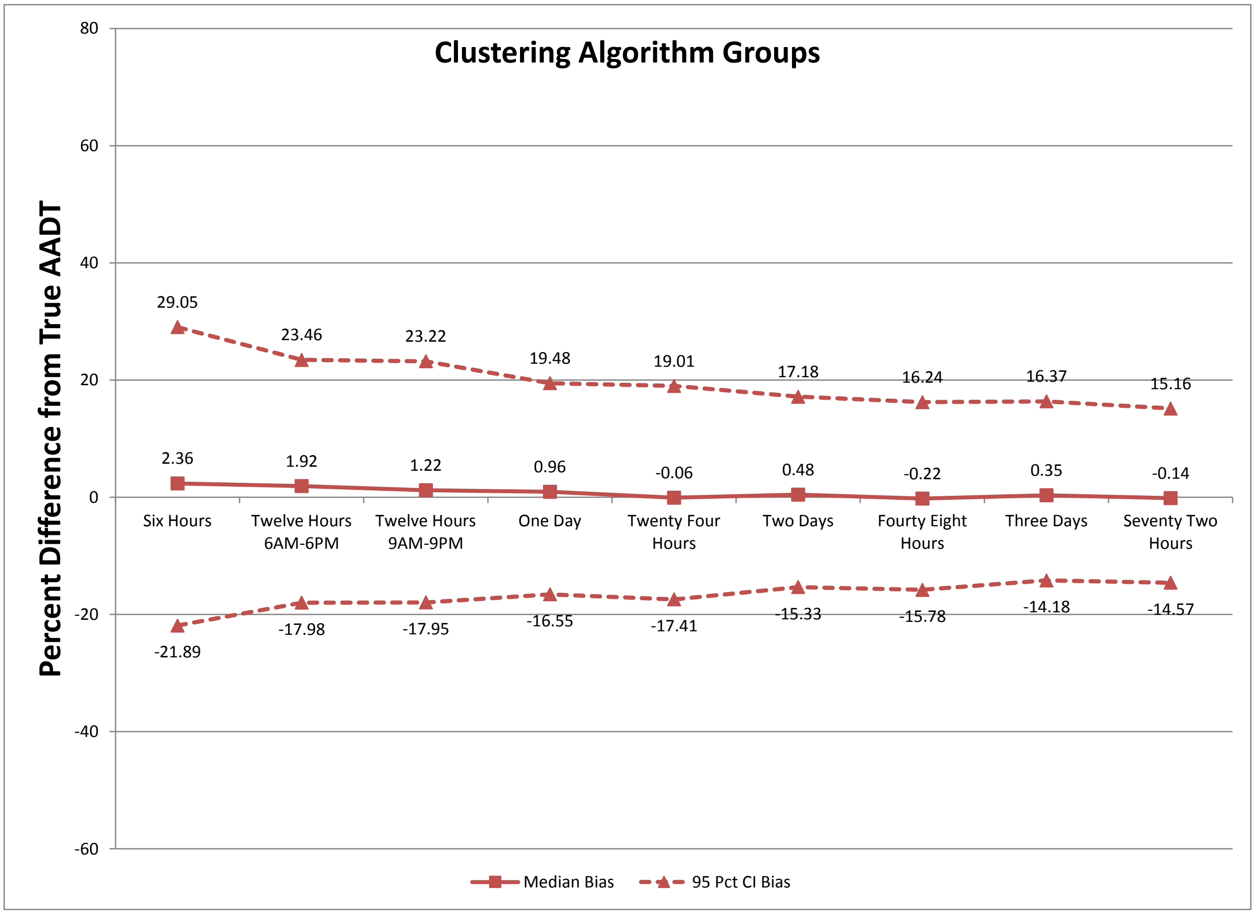

Figure 12. National Level Bias by Short-Term Count Duration for Clustering Algorithm Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

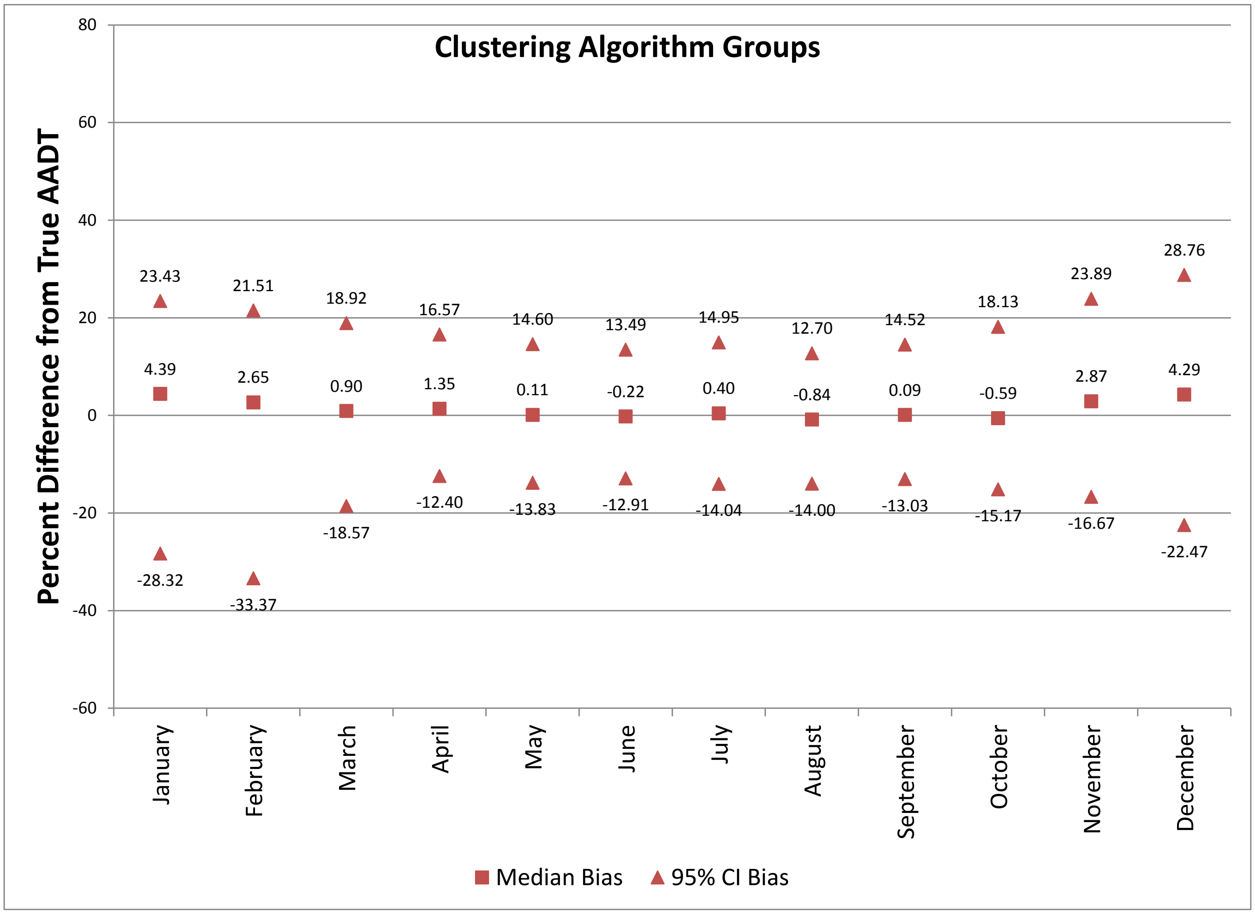

Figure 13. National Level Bias by Month for One Day Count Duration for Clustering Algorithm Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays
(Source: Battelle, based on data from TMAS)

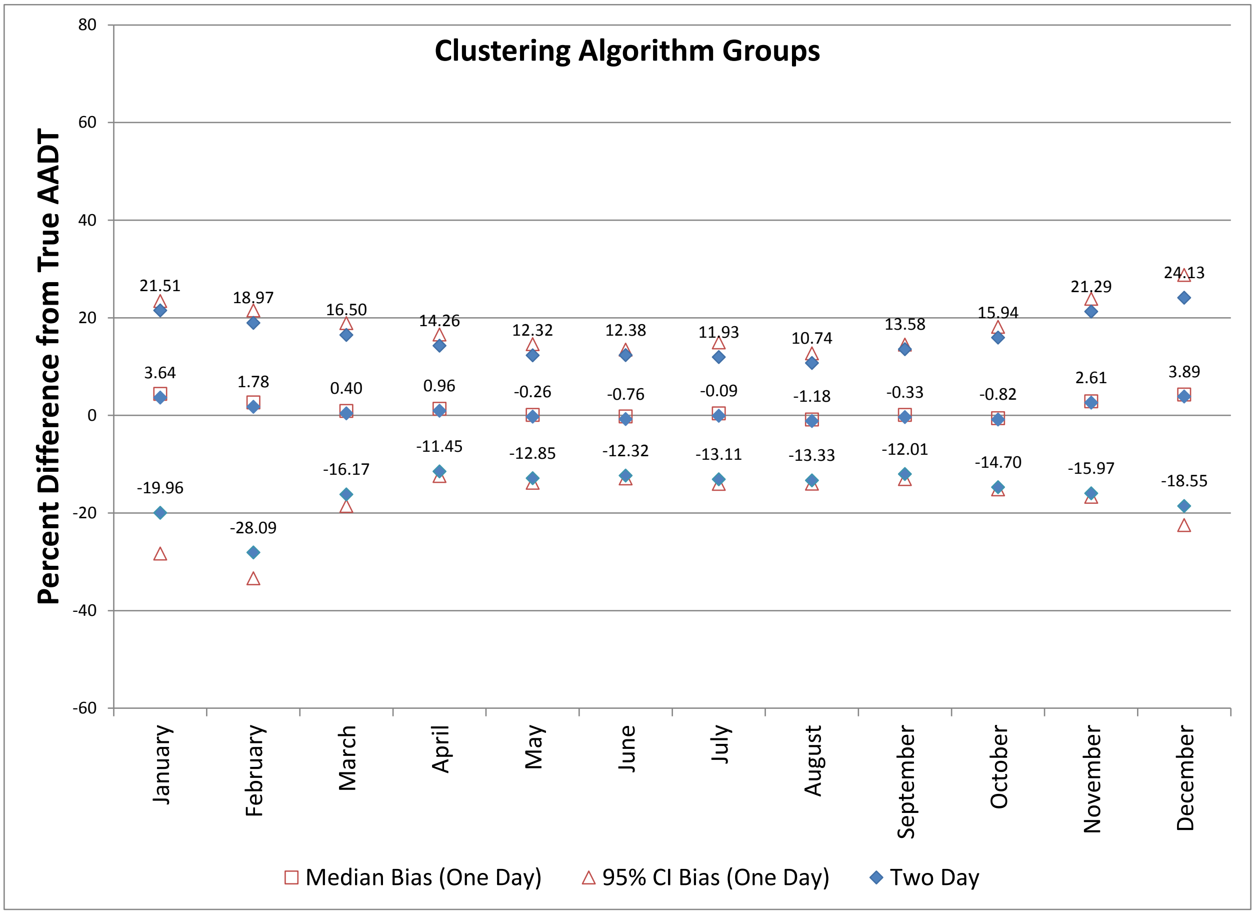

Figure 14. National Level Bias by Month for Two Day Count Duration for Clustering Algorithm Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

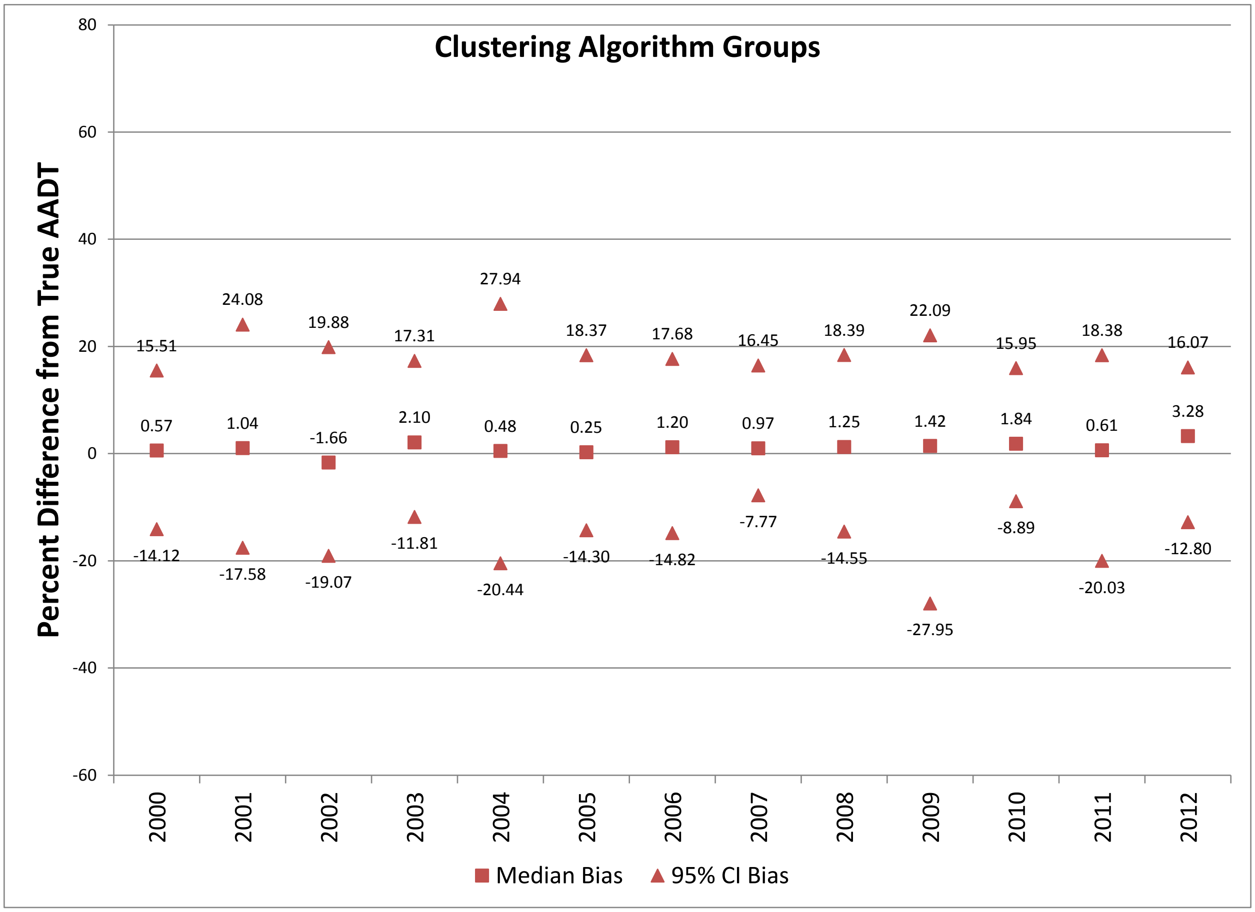

Figure 15. National Level Bias by Year for One Day Count Duration for Clustering Algorithm Groups, Separate Day of Week and Month of Year Factors, Only Counts Contained within Monday-Thursday, Excluding Counts Encompassing Federal Holidays (Source: Battelle, based on data from TMAS)

Supplemental Results

The full results of the analysis are provided in a separate document entitled “Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic Estimation: Assessing AADT Accuracy Issues Related to Short-Term Count Durations – Supplemental Tables”. The tables in that document are organized in groupings of the three different factoring methods used: Functional Classification, Volume Range, and Clustering Algorithm. Each table is presented in three sections, with the factoring method by day of week and month as the top section, day of week by month as the middle section, and day of year factoring as the bottom section. The columns in each table are the short-term count duration (ranging from six hours to seven days), the number of count durations, n, in a summary, and the mean, standard deviation, 2.5th percentile, median, and 97.5th percentile of the percent error results in estimated versus true AADT across all the sites for that particular duration. Note in each table that the results reflect the removal of any scenario with a national holiday, with these holidays separately shown in Table A-1.

A number of different levels of summarization are employed, with the tables organized as follows:

Functional Classification Groups Section

- National Summary – All Days of Week
- National Summary – Counts Beginning and Ending Monday-Thursday
- National Summary – Counts Containing Friday, Saturday, and Sunday, respectively
- National Summary – Counts Beginning Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday, respectively
- Functional Classification Summary – Counts Beginning and Ending Monday-Thursday for FC 1R, 1U, 2U, 3R, 3U, 4R, 4U, 5R, and 5U, respectively
- National Summary – Counts Beginning and Ending Monday-Thursday for January, February, March, April, May, June, July, August, September, October, November, and December, respectively
- National Summary – Counts Beginning and Ending Monday-Thursday for 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, and 2012, respectively
- Functional Classification Group Summary - Counts Beginning and Ending Monday-Thursday, with one table for each Functional Classification Group from Alabama/2000/3U through Wyoming/2000/3U

Volume Range Groups Section

- National Summary – All Days of Week
- National Summary – Counts Beginning and Ending Monday-Thursday
- National Summary – Counts Containing Friday, Saturday, and Sunday, respectively
- National Summary – Counts Beginning Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday, respectively
- Functional Classification Summary – Counts Beginning and Ending Monday-Thursday for Volume Range 1, 2, 3, and 4, respectively
- National Summary – Counts Beginning and Ending Monday-Thursday for January, February, March, April, May, June, July, August, September, October, November, and December, respectively
- National Summary – Counts Beginning and Ending Monday-Thursday for 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, and 2012, respectively
- Volume Range Group Summary - Counts Beginning and Ending Monday-Thursday, with one table for each Volume Range Group from Alabama/2000/Volume Range 3 through Wyoming/2000/Volume Range 3

Clustering Algorithm Groups Section

- National Summary – All Days of Week
- National Summary – Counts Beginning and Ending Monday-Thursday
- National Summary – Counts Containing Friday, Saturday, and Sunday, respectively
- National Summary – Counts Beginning Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday, respectively
- National Summary – Counts Beginning and Ending Monday-Thursday for January, February, March, April, May, June, July, August, September, October, November, and December, respectively
- National Summary – Counts Beginning and Ending Monday-Thursday for 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, and 2012, respectively
- Clustering Algorithm Groups Summary - Counts Beginning and Ending Monday-Thursday, with one table for each Clustering Algorithm Group from Alabama/2000 through Wyoming/2000/Group B

Conclusions

Several important conclusions can be distilled from this evaluation:

1. Short-term counts of longer duration provide a more precise estimate of AADT. For the most common durations of 24 hours and 48 hours at the national level, the width of the 95 percent confidence interval for error in estimating AADT in the longer duration was 12.1 percent smaller than for the shorter duration. This difference was found to be statistically significant with 95 percent confidence interval of (10.4, 13.8) percent reduction. This result applied under typical Monday-Thursday sampling, excluding holidays, and with separate day of week and month factors developed from factor groups created by a clustering algorithm. The similar functional classification results were an 8.7 percent reduction (95 percent confidence interval of 6.5 percent, 10.3 percent) in the width of the 95 percent interval on percent bias in the 48-hour results as compared to the 24-hour results. Both results were statistically significant at $p<0.001$.

2. Accuracy of short term counts, as measured by the median error relative to true AADT, was very good (i.e., close to zero) for both factoring methods. The 48-hour counts did produce lower median estimated AADT by 0.15 percent (95 percent confidence interval of (0.06, 0.24)) for the clustering algorithm, and 0.26 percent (95 percent confidence interval of (0.13, 0.40)) for the functional classification factoring, both cases statistically significant at $p<0.001$.

3. The simple methods of associating short-term count locations to factor groups using functional classification or volume ranges result in less precise AADT estimates than a more rigorous clustering algorithm that better associates the temporal volume pattern for the temporary count station to comparable permanent count stations.

4. Restricting short-term counts to weekdays results in more precise AADT estimation. Precision of estimates obtained on or including Friday is comparable to other weekdays, and it does not appear to be necessary to exclude Friday from volume counting from either an accuracy or a precision perspective.

5. The precision of AADT estimates from short-term counts taken March through October, and especially April through September, is substantially better than for the months November, December, January, and February.

APPENDIX:

Federal Holidays

Table A-1. Federal Holidays Excluded from Summarized Results

Holiday	Year													
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
New Year's Day	Saturday 01 Jan	Monday 01 Jan	Tuesday 01 Jan	Wednesday 01 Jan	Thursday 01 Jan	Saturday 01 Jan	Sunday 01 Jan	Monday 01 Jan	Tuesday 01 Jan	Thursday 01 Jan	Friday 01 Jan	Saturday 01 Jan	Sunday 01 Jan	Tuesday 01 Jan
New Year's Day observed							Monday 02 Jan						Monday 02 Jan	
Martin Luther King Day	Monday 17 Jan	Monday 15 Jan	Monday 21 Jan	Monday 20 Jan	Monday 19 Jan	Monday 17 Jan	Monday 16 Jan	Monday 15 Jan	Monday 21 Jan	Monday 19 Jan	Monday 18 Jan	Monday 17 Jan	Monday 16 Jan	Monday 21 Jan
Presidents' Day (Washington's Birthday)	Monday 21 Feb	Monday 19 Feb	Monday 18 Feb	Monday 17 Feb	Monday 16 Feb	Monday 21 Feb	Monday 20 Feb	Monday 19 Feb	Monday 18 Feb	Monday 16 Feb	Monday 15 Feb	Monday 21 Feb	Monday 20 Feb	Monday 18 Feb
Memorial Day	Monday 29 May	Monday 28 May	Monday 27 May	Monday 26 May	Monday 31 May	Monday 30 May	Monday 29 May	Monday 28 May	Monday 26 May	Monday 25 May	Monday 31 May	Monday 30 May	Monday 28 May	Monday 27 May
Independence Day observed										Friday 03 Jul				
Independence Day	Tuesday 04 Jul	Wednesday 04 Jul	Thursday 04 Jul	Friday 04 Jul	Sunday 04 Jul	Monday 04 Jul	Tuesday 04 Jul	Wednesday 04 Jul	Friday 04 Jul	Saturday 04 Jul	Sunday 04 Jul	Monday 04 Jul	Wednesday 04 Jul	Thursday 04 Jul
Independence Day observed					Monday 05 Jul						Monday 05 Jul			
Labor Day	Monday 04 Sep	Monday 03 Sep	Monday 02 Sep	Monday 01 Sep	Monday 06 Sep	Monday 05 Sep	Monday 04 Sep	Monday 03 Sep	Monday 01 Sep	Monday 07 Sep	Monday 06 Sep	Monday 05 Sep	Monday 03 Sep	Monday 02 Sep
Columbus Day	Monday 09 Oct	Monday 08 Oct	Monday 14 Oct	Monday 13 Oct	Monday 11 Oct	Monday 10 Oct	Monday 09 Oct	Monday 08 Oct	Monday 13 Oct	Monday 12 Oct	Monday 11 Oct	Monday 10 Oct	Monday 08 Oct	Monday 14 Oct
Veterans Day observed	Friday 10 Nov						Friday 10 Nov							
Veterans Day	Saturday 11 Nov	Sunday 11 Nov	Monday 11 Nov	Tuesday 11 Nov	Thursday 11 Nov	Friday 11 Nov	Saturday 11 Nov	Sunday 11 Nov	Tuesday 11 Nov	Wednesday 11 Nov	Thursday 11 Nov	Friday 11 Nov	Sunday 11 Nov	Monday 11 Nov
Veterans Day observed		Monday 12 Nov						Monday 12 Nov					Monday 12 Nov	
Thanksgiving Day	Thursday 23 Nov	Thursday 22 Nov	Thursday 28 Nov	Thursday 27 Nov	Thursday 25 Nov	Thursday 24 Nov	Thursday 23 Nov	Thursday 22 Nov	Thursday 27 Nov	Thursday 26 Nov	Thursday 25 Nov	Thursday 24 Nov	Thursday 22 Nov	Thursday 28 Nov
Christmas Day observed					Friday 24 Dec						Friday 24 Dec			
Christmas Day	Monday 25 Dec	Tuesday 25 Dec	Wednesday 25 Dec	Thursday 25 Dec	Saturday 25 Dec	Sunday 25 Dec	Monday 25 Dec	Tuesday 25 Dec	Thursday 25 Dec	Friday 25 Dec	Saturday 25 Dec	Sunday 25 Dec	Tuesday 25 Dec	Wednesday 25 Dec
Christmas Day observed						Monday 26 Dec						Monday 26 Dec		
New Year's Day observed					Friday 31 Dec						Friday 31 Dec			

Source: Battelle

U.S. Department of Transportation
Federal Highway Administration

U.S. Department of Transportation
Federal Highway Administration
Office of Highway Policy Information
1200 New Jersey Ave., SE
Washington, D.C. 20590
https://www.fhwa.dot.gov/policyinformation
October 2015
FHWA-PL-16-008