ANALYSIS OF VEHICLE CLASSIFICATION AND TRUCK WEIGHT DATA OF THE NEW ENGLAND STATES: IS DATA SHARING A GOOD IDEA?

Rick Schmoyer
and
Patricia S. Hu
Center for Transportation Analysis
Oak Ridge National Laboratory

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6073
managed by
LOCKHEED MARTIN ENERGY RESEARCH CORP.
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-96OR22464

ABSTRACT

This paper is about a statistical research analysis of 1995-96 classification and weigh in motion (WIM) data from seventeen continuous traffic-monitoring sites in New England. Data screening is discussed briefly, and a cusum data quality control procedure is proposed. The main purpose of the analysis, however, is to infer statistical methods for using data from multiple states in a common resource data pool. Because data sharing means cross-state extrapolation, the combined data should not be used without a proper statistical accounting for extrapolation error. Another major concern in implementing a data-sharing procedure is operational simplicity. Of particular interest are the possible analytical simplifications of combining vehicle classes (i.e., reducing the number of vehicle classes used in practice) or combining HPMS roadway functional classes. Adjusting for seasonal and day-of-week effects is also a concern.

Conclusions based on the analysis are still preliminary. Analysis of the ultimate use of the data suggests that from the perspective of vehicle load estimation, there is little advantage to combining vehicle classes. Analysis of both the WIM and classification data suggests that differences among HPMS functional classes are sufficient to warrant against combining functional classes. But even without these simplifications, data sharing among states is a good idea. The analysis method used here, one-way analysis of variance, is reasonably simple (can be done with an ordinary spreadsheet program), provides an accounting for statistical error, and is thus an appropriate analysis tool for data sharing.

1. INTRODUCTION

For many years the six New England States (U.S. DOT standard Region 1) have been collecting vehicle classification and truck weight data to meet programmatic needs of the state and Federal governments. Each state has a well-developed traffic monitoring system. In addition, a good working relationship exists among the states. This is evident from technology sharing meetings held several times a year, from regular exchanges of data, and from the states' desire and commitment to improve existing traffic monitoring programs, particularly for trucks. Currently, the Region 1 states are reviewing the cost-effectiveness of their data collection and analysis activities, and exploring the possibility of combining traffic data programs.

Although never formally demonstrated, it is reasonable to think that truck travel in each of these states is similar, because of geographic location, the small size of each state, continuity of major truck routes across the states, and similarity in economic activities. It is also reasonable to think that the six states may have other similarities and that a combined data collection effort may significantly reduce the resource demand on each state. Unfortunately, available resources have limited detailed analyses of each state's data. These analyses are crucial to determine similarities in data and to establish an effective way of combining their traffic data.

The work described here is an analysis of classification and weigh-in-motion data from several of the Region 1 states. The classification data were vehicle counts for FHWA classes 1-13. The WIM data were axle weights and spacings for trucks (classes 4-13). Details about data availability and decisions about was kept for further analysis are documented in [1]. The decisions were based on an analysis of missing data, and several preliminary data-quality checks. For the classification data, the checks were based on class frequency ratios, frequency changes, and three-standard-deviation control limits. For the WIM data, the checks were based on a graphical analysis of front-axle and gross-vehicle weights of five-axle single-trailer trucks (vehicle class 9). Table 1 gives basic descriptive information about the sixteen classification sites and eleven continuous-monitoring WIM sites kept for further analysis. The total number of different sites is seventeen—ten sites were kept for both their class and WIM data. Figure 1 shows the locations of the sites.

The data analysis was a research task. The objective was to analyze the traffic volume and classification data, and at the same time to explore and develop statistical methods for (1) combining data across states, (2) combining vehicle classes, (3) combining HPMS roadway functional classes, and (4) making seasonal and day-of-week adjustments to short-term class or WIM data. Both the need and the methodology for day-of-week and seasonal adjustments (i.e. adjustment factors) in short-term traffic volume data are well understood [2]. Here we use the same approaches for classification and WIM data. Therefore issue (4) will only be considered briefly.

Combining data across states means cross-site extrapolation beyond state borders. Cross-state extrapolations are subject to site-differences attributable not simply to differences in location but also to differences in weight-limit regulations. Therefore, it is especially important that any methodology for data-sharing across state boundaries should include measures of the extrapolation error, that is, standard errors of estimates based on extrapolating. Reasonable approximate standard errors allow

Table 1. The Seventeen Classification/WIM Sites Kept for Classification Analysis

Site	HPMS Class	Location (also see map)	Yrs.	Dir.	Avg. Ann. Daily Traffic	Pct. in Class 4-13	Avg. Daily Trks.	Avg. GVW (kips)
CT974	**Rural—Major Collector (7)**	**Rt. 117—.9 m N of Rt. 184**	**95**	**N**	4,802	4.61	190	9.8
CT978	**Urban—Principal Arterial Other Free/Expwy (12)**	**Rt. 2—2.5 m W of Rt. 83**	**95**	**W**	16,094	4.15	600	30.0
CT990	**Urban—Principal Arterial Interstate (11)**	**I-84—2 m W of Rt. 30**	**95**	**W**	42,681	8.59	3,430	40.1
CT991	**Urban—Principal Arterial Interstate (11)**	**I-84—.75 m W of Rt. 31**	**95**	**W**	33,009	10.09	Class Only	Class Only
MA001	**Urban—Principal Arterial Interstate (11)**	**I-93—N of Rt. 28**	**96**	**N, S**	93,070	5.64	9,420	32.7
MA002	**Urban—Principal Arterial Interstate (11)**	**I-391—N of I-90**	**96**	**N, S**	13,659	3.47	210	16.3
MA003	**Urban—Principal Arterial Other (14)**	**Rt. 27—S of Hospital Rd.**	**95, 96**	**N, S**	3,363	5.41	Class Only	Class Only
MA004	**Urban—Principal Arterial Interstate (11)**	**I-95—E of Acushnet River**	**96**	**E, W**	16,156	4.52	Class Only	Class Only
MA005	**Urban—Principal Arterial Interstate (11)**	**I-95—S of Rt. 38**	**95, 96**	**N, S**	85,172	5.69	5,450	30.7
RI350	**Urban—Principal Arterial Other Free/Expwy (12)**	**Rt. 146 at Mass. State Line**	**95, 96**	**N, S**	7,817	11.62	1,770	41.7
VT132	**Rural—Principal Arterial Other (2)**	**U.S. 7—Charlotte**	**95, 96**	**N, S**	5,131	8.09	Class Only	Class Only
VT249	**Rural—Principal Arterial Other (2)**	**VT 103, Rockingham**	**95**	**E, W**	2,512	11.35	Class Only	Class Only
VTa41	**Rural—Principal Arterial Other (2)**	**U.S. 7, New Haven**	**95, 96**	**N, S**	3,135	8.51	Class Only	Class Only
VTd92	**Rural—Principal Arterial Interstate (1)**	**I-91—Fairlee**	**95, 96**	**N, S**	WIM Only	WIM Only	1,420	38.6
VTn01	**Rural—Principal Arterial Interstate (1)**	**I-91—Fairlee**	**95**	**N, S**	3,893	11.55	860	41.6
VTr01	**Rural—Principal Arterial Other (2)**	**U.S. 4—New Haven**	**95, 96**	**E, W**	3,194	14.39	860	45.1
VTx73	**Rural—Principal Arterial Interstate (1)**	**I-91—Putney**	**96**	**N, S**	6,385	12.54	1,530	43.0
All	**1, 2, 7, 11, 12, 14**	**CT, MA, RI, VT**	**95, 96**	**N, S E,W**	18,865	6.66	2,320	35.4

Figure 1. The Seventeen Classifications/WIM Sites Kept for Analysis. *Classification analysis only; **WIM analysis only; other sites were used in both analyses.

for decisions about whether cross-site extrapolations are adequate. In addition, error analysis can identify where resources might best be spent in improving cross-site estimates (e.g., longer monitoring at short-term sites vs. more continuous sites).

The process of converting short- or long-term WIM data or axle or classification counts into estimates of loads and other useful statistics is deceptively complex. Thus, in addition to technical defensibility, a major concern in data-sharing methodology is simplicity of operation. Concerns about operational simplicity (and cost) have lead to interest in combining vehicle classes or roadway functional classes, for the purpose of data analysis, and these possible simplifications should be considered in decisions about methods for data-sharing.

Another reason for investigating the possibility of combining vehicle classes is that because the traffic for some of the classes is low-frequency, statistical properties of estimates (particularly the relative error) for those classes tend to be poor. (Combining the vehicle classes might improve the relative error.) In addition, validation "ground-truthing" experiments [3] have indicated that FHWA vehicle classes 2 and 3 might well be combined because of the incapability of classification equipment to differentiate those two classes. The same rationale about statistical properties applies to HPMS functional classes, and, similarly, there is doubt that some of the HPMS classes are sufficiently different to warrant separate consideration.

Analysis of variance (ANOVA) provides a mechanism for cross-site extrapolation with a formal accounting for extrapolation error. It provides tests for differences between classes of sites, such as HPMS functional classes. In its simplest form, one-way ANOVA, it is simple enough to implement with an ordinary spreadsheet program such as Excel.

In Section 2 of this paper we discuss data quality checks. Although the basic selection of the data for analysis is discussed elsewhere, several additional quality control checks were performed as part of the analysis discussed here. In Section 3 we briefly consider seasonal and day-of-week effects. In Section 4, propagation of errors is considered as a basis for deciding about whether vehicle classes should be combined. The basic conclusion is that for the purpose of load estimation, there is little to gain by combining vehicle classes. In Section 5 we consider limitations on the data structure, which affect how the data should be analyzed. ANOVAs are discussed in Section 6. The ANOVAs suggest that many HPMS functional classes differ substantially enough that they should not be combined, and for the others there is not sufficient basis to combine them either.

The ANOVA and propagation of error methods together form a methodology that can be used for cross-state data sharing and extrapolation, a methodology that is reasonably simple and provides an accounting for statistical error incurred in cross-site extrapolations. Conclusions and several important areas for a more detailed analysis are mentioned in Section 7.

2. DATA SCREENING

The data selection criteria (discussed in [1]) we used for the classification data were more extensive than for the WIM data. The checks for the classification data had explicitly defined rejection criteria, whereas the checks for the WIM data were graphical and more subjective. Therefore, although the preliminary checks for both the class and WIM data were used to decide whether to keep or exclude the class or WIM data for entire site-years, the checks for the class data were also used as a basis for excluding smaller sections from the “kept” data. For the WIM data, a number of additional checks were made. The additional WIM data checking was done in two steps: (1) comparing data values to internal and external references checks, and (2) serial checks and graphical inspection.

The internal and external reference checks include, for example, comparisons: of axle weights to minimum and maximum limits; of the number-of-axles data entry to the number of axles having positive weight and to the number of axles implied by the six-digit code; of the total wheelbase data-entry to the sum of individual axle-spacings. VTRIS [4] default limits were used for minimum or maximum limits. For more details about these checks see [5].

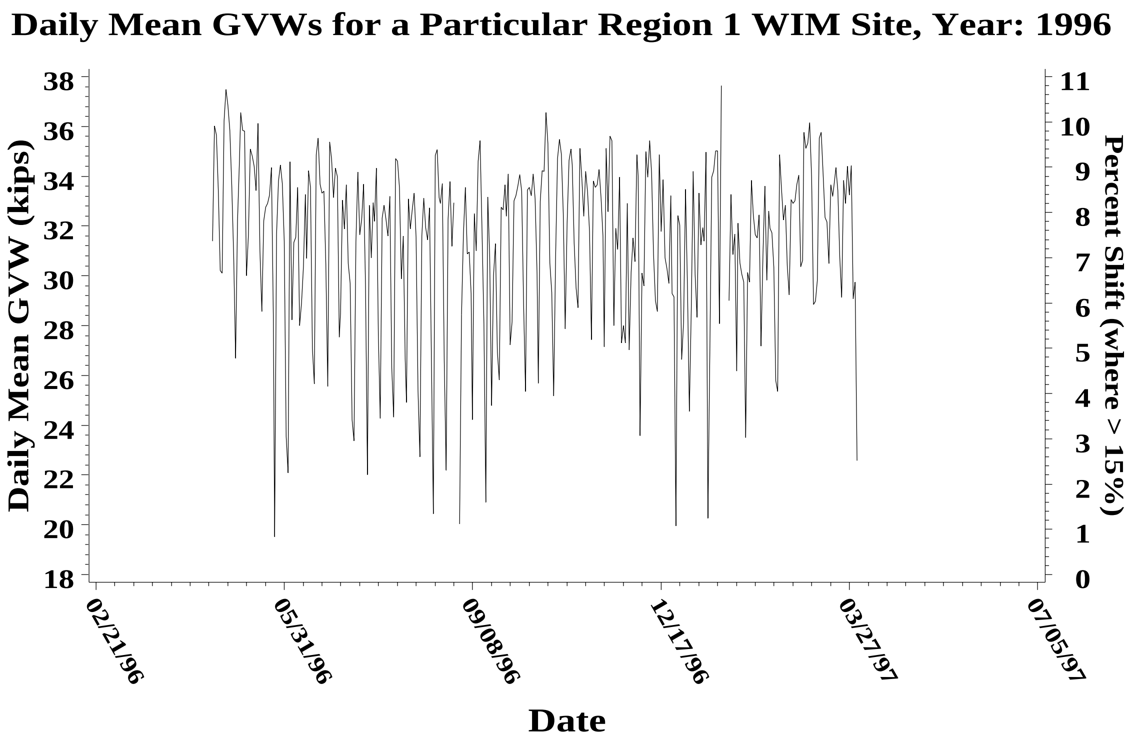
The serial WIM data checks were performed as follows. For each site, direction, and year, daily average gross vehicle weights (GVWs) were plotted over time, and marked, using a change-point algorithm, wherever appreciable jumps or change-points—possibly bad data—seemed to occur. The change-point algorithm is based on the statistic:

$$T = 200 \;x \left| \frac{\text{mean for two weeks post - mean for two weeks prior}}{\text{mean for two weeks post + mean for two weeks prior}} \right|,$$

evaluated at each point in the data series. The statistic T is actually a *cusum* (cumulative sum) statistic from statistical quality control theory [6]. A change in the series is suggested at any point for which the mean for the last two weeks is appreciably different from the mean for the next two weeks. “Appreciably different” must be defined, of course, and should achieve a reasonable balance of false positives and false negatives. Here, after several iterations, “appreciably different” was defined as “greater than 15 percent.”

Two of these plots, for two Region 1 sites in 1996, are in Figure 2. Appreciable changes are marked as “Percent Shift (where > 15%).” There are no appreciable change points in the series for the first site, but there is a change in the series for the second site, near the beginning of November 1996.

Upon inspection, the change in the second series is obvious, and the change may represent a change in real traffic conditions, rather than an instrumentation problem. Nevertheless, the convenient feature of the cusum approach is that an “alarm” can be sounded (e.g., email sent) when there appears to be a substantial change. The series can then be examined graphically, with special attention paid to those cases more likely to need attention, and without the need for routine (e.g., daily) inspection of all the data plots. Cusum and other data quality checking can be incorporated into automatic data downloading procedures.



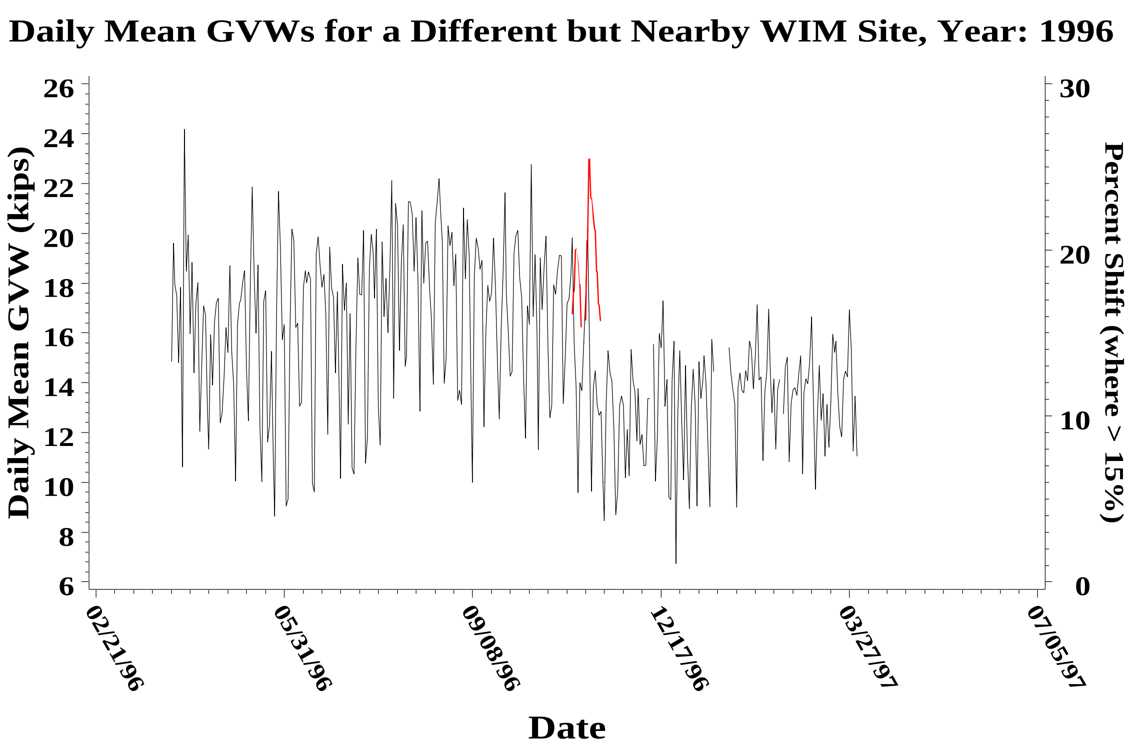


Figure 2. WIM data cusum control charts for two Region 1 sites. The percent shift indicates a possible change-point, that is, either an actual change in traffic or an instrumentation change.

Analogous cusum plots for classification counts are also useful. Several other plots were also made as data quality checks. They are discussed in [5]. For the most part, however, suspicious data points were not thrown out, because, for this project, we wanted to assess the data in the context of the full data variability we expect to see in practice. Control limit values were, however, substituted for data values found to be outside the limits.

3. SEASONAL AND DAY-OF-WEEK EFFECTS

Figure 3 illustrates the effect of season (months) on Region 1 passenger car volumes. Figure 4 shows the effect of day-of-week on total load in kips for two truck classes: 5 (two-axle, six-tire, single-unit trucks) and 9 (five-axle single-trailer trucks). There is also a seasonal effect on truck loads in Region 1, though it is much smaller than the day-of-week effect. Both the need and methods for day-of-week and seasonal adjustments in short-term traffic volume data are discussed in [2]. The same methods were applied to class counts and WIM loads to compute the adjustment factors used for the analyses discussed in the following sections.

4. PROPAGATION OF ERRORS

The statistical distribution most commonly used to model counting processes is the Poisson distribution (see, for example [7], p 223). For the Poisson distribution, the relative precision, expressed as the coefficient of variation (standard deviation divided by mean) is given by

$$CV = \frac{1}{\sqrt{\text{mean}}} \ .$$

From this equation an approximate standard deviation for daily counts can be obtained from mean daily counts, as in Table 2, which was computed from daily count means for the sixteen classification sites. As illustrated in the table, CV's for combined vehicle classes are smaller than any of the CV's from the individual classes in the combination.

Table 2. CV's (Percent) by Vehicle Class From Average Annual Daily Traffic Estimates

Vehicle Class	1	2	3	4	5	6	7	8	9	10	11	12	13	All
Mean Veh./Day	64	15115	2430	39	464	86	7	132	491	16	19	2	1	18865
CV	12.5	.81	2.0	16.1	4.6	10.8	36.7	8.7	4.5	25.4	22.8	74.5	141.4	.72
Comb. Classes	Passenger Vehicles			Single-Unit Trucks				Trailer Trucks						All
Mean Veh./Day	17609			596				660						18865
CV	.75			4.1				3.9						.72

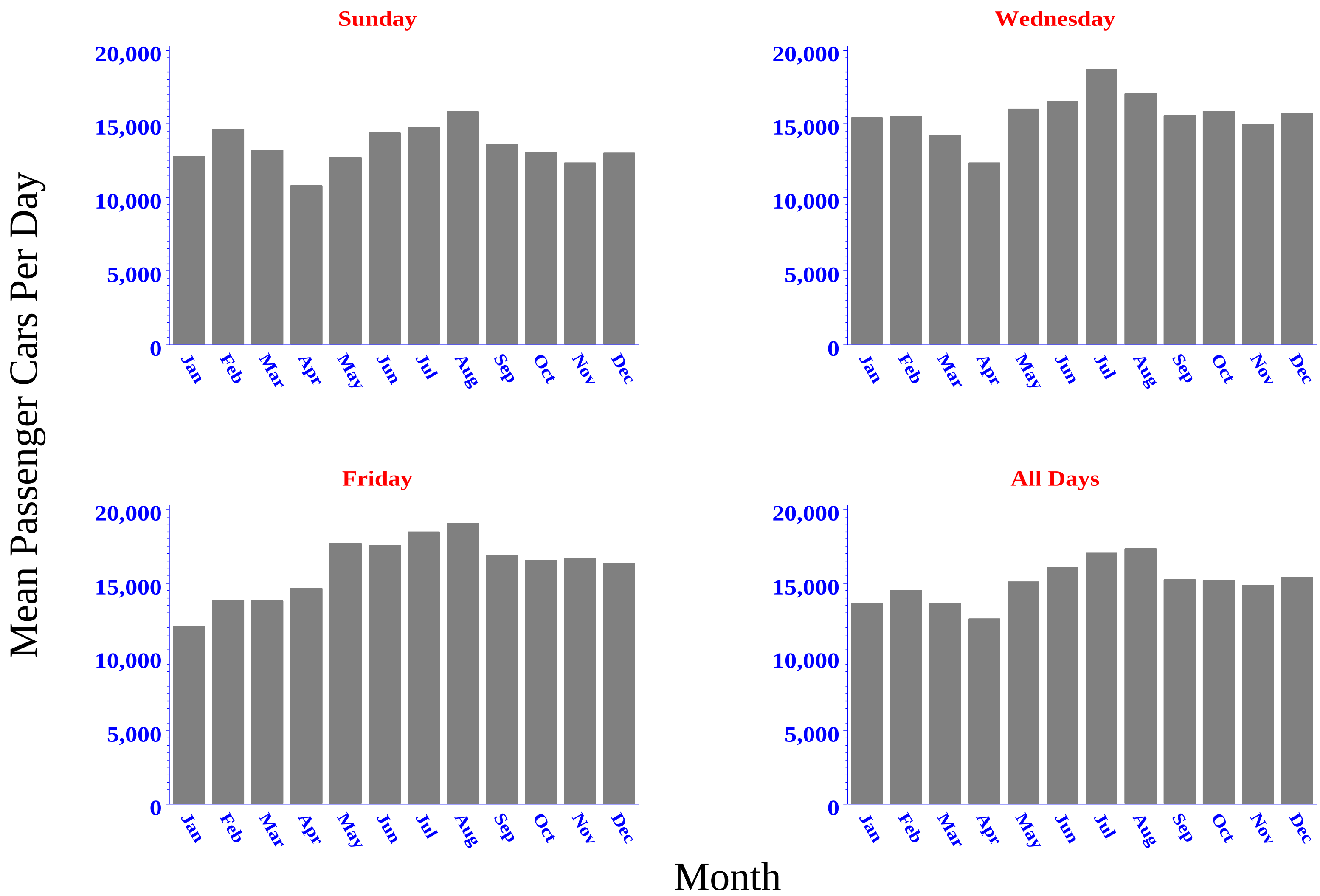


Figure 3. Effect of seasonality on passenger car volume for selected days-of-the-week, average for sixteen class sites.

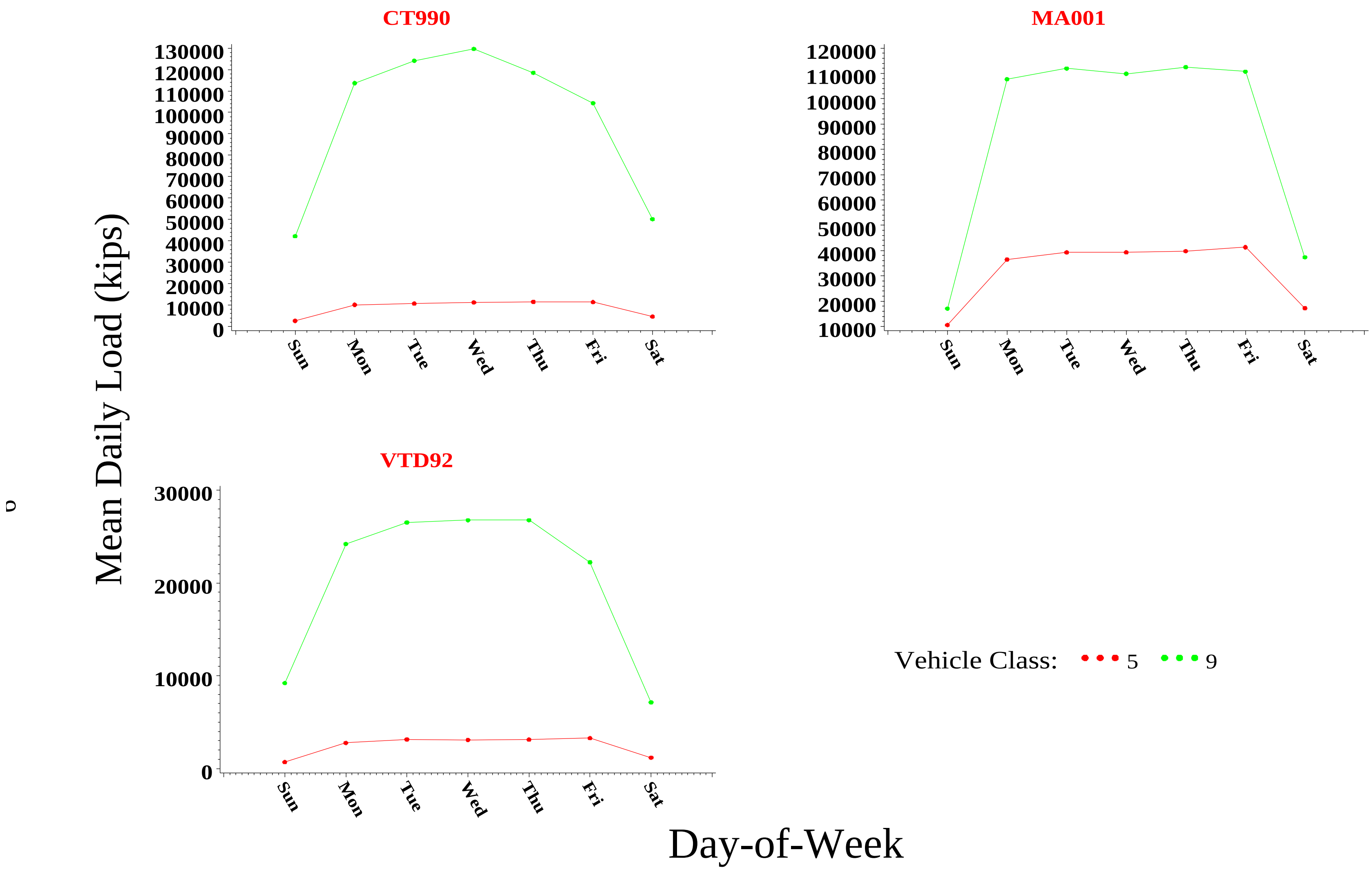


Figure 4. Effect of day-of-week, for selected WIM sites, on Class 5 and 9 truck loads.

Of course, combining class counts results in a loss of information (How many light trucks in Passenger Vehicles?). But the table does suggest that combining vehicle classes would lead to class totals with better statistical properties. Tables of counts, however, are not the only product of classification count statistics. The counts are also used to estimate loads.

Load estimates are computed under a variety of data-collection schemes. For example, average annual daily loads (AADLs) can be estimated from short-term class counts, seasonal/day-of-week adjustment factors (AFs), and average loads per vehicle (estimated from WIM data), as in the following equation:

$$\text{AADL estimate} = \sum_{\text{vehicle classes}} (\text{short class count}) \times (\text{class AF}) \times (\text{load per vehicle}).$$

In [5] we considered propagation of errors under several estimation schemes, including equation (1). The details of these analyses are fairly technical, but the basic idea is straightforward: although vehicle counts and load estimates for low-frequency vehicle classes do have high relative variability, their contributions to overall loads (i.e., combined over all vehicle classes) are small, and because errors in the various individual estimates tend to cancel, the high variability of the low-frequency classes does not matter much in overall load estimates. After the statistical error in counts or weights propagates to overall load estimates, there is little advantage to combining vehicle classes.

Thus, combining vehicle classes leads to better relative precision in class counts, but also to a loss of information. Combining vehicle classes is of little value in load estimation. In view of this and because the vehicle classes do clearly differ, we do not think vehicle classes should be combined. Therefore, we performed analyses—for investigating HPMS class differences and for cross-site extrapolations—on a vehicle class-by-class basis.

5. DATA LIMITATIONS

Before proceeding to analyze the Region 1 data, one should understand certain limitations on the data structure itself. Table 1 shows that in terms of the HPMS functional classes, states, and years, the Region 1 data is convolved: comparisons of any one of these are, for the most part, not easily separated from the others. For example, out of the eleven WIM sites, the only comparisons of states that can be made that are free from differences due to HPMS class or year are (1) CT990 with MA005 and (2) CT978 with RI350. All other comparisons of states also involve year-to-year or HPMS class differences. This does not leave much statistical room for directly deciding about combining data across states. For WIM comparisons of HPMS classes, there are three different HPMS classes in CT (7,11, and 12; all 1995), and two different HPMS classes in VT (1 and 2 for 1995 and 1996). All MA sites are HPMS class 11, and Rhode Island has only one site (class 12). Again, statistically, there is little basis for direct conclusions about combining HPMS classes.

Hallenbeck [8] confronted a similar situation in working with data from 99 sites from 19 states and HPMS classes 1, 6, 7, 11, 12, 14. Siting a continuum (rather than clustering) of day-of-week patterns, and differences between automobile and truck day-of-week patterns as primary reasons for the difficulty in developing roadway factor groups, Hallenbeck concluded (p 11) "there is insufficient data in the LTPP database at this time to support the creation of these [factor] groups." The difficulty with sparsity is similar in the Region 1 data discussed here, though here the focus is only on one region.

Therefore, several simplifying assumptions about the joint effects of state, year, and HPMS class differences were made for the data analyses here. The first assumption is that the selection of Region 1 sites emulates a simple random sample. That is, the selection of one site is assumed to be statistically independent of the selection of other sites. This is clearly an approximation. Because permanent classification or WIM sites are expensive, selection of their locations is usually purposive rather than random. Nevertheless, the sites are approximately randomly scattered over a subset of the total New England area (Fig. 1). For more on the importance of random sampling in traffic monitoring, see [9].

The second simplifying assumption is that results (counts, loads, means, totals) for separate directions and years are also statistically independent. The rationale for this assumption is that because the Region 1 data is sparse and uneven, it would be too complicated to account for year-to-year and direction-within-site differences while simultaneously measuring the effect of HPMS class differences. The assumption is also clearly an approximation. Traffic at the same site but different directions tends to be similar (though it can be quite different—as it is for example at site VTr01). Traffic at the same site in different years is also generally similar. Note, however, that to some extent, the consequence of departures from independence is limited in that there are at most two years and two directions for any given site.

The two independence assumptions imply that results for different site-direction-years are statistically independent. Thus departures of results for different site-direction-years from their HPMS class means are statistically independent. This is a requirement for a valid ANOVA.

6. ANOVAs in HPMS Class

Sharing traffic monitoring data means extrapolating across sites. Because statistics such as total traffic or load vary substantially from site to site, extrapolation is ordinarily in terms of some normalized rate or adjustment factor (e.g., ratio of average-annual-daily-traffic to average-daily-traffic) for a particular type of day (e.g., August Wednesdays) or perhaps a weight per vehicle (GVW) if that is reasonably assumed constant across sites. But even though they are normalized, AFs, GVWs, and similar statistics can still differ from site to site.

Figure 5 illustrates “raw” AFs and their site-to-site variability for August Wednesdays (arbitrary choice). The raw AFs are simply the AADTs divided by the average daily traffic for August Wednesdays, for each classification site, direction, and year. The AFs were entered into one-way

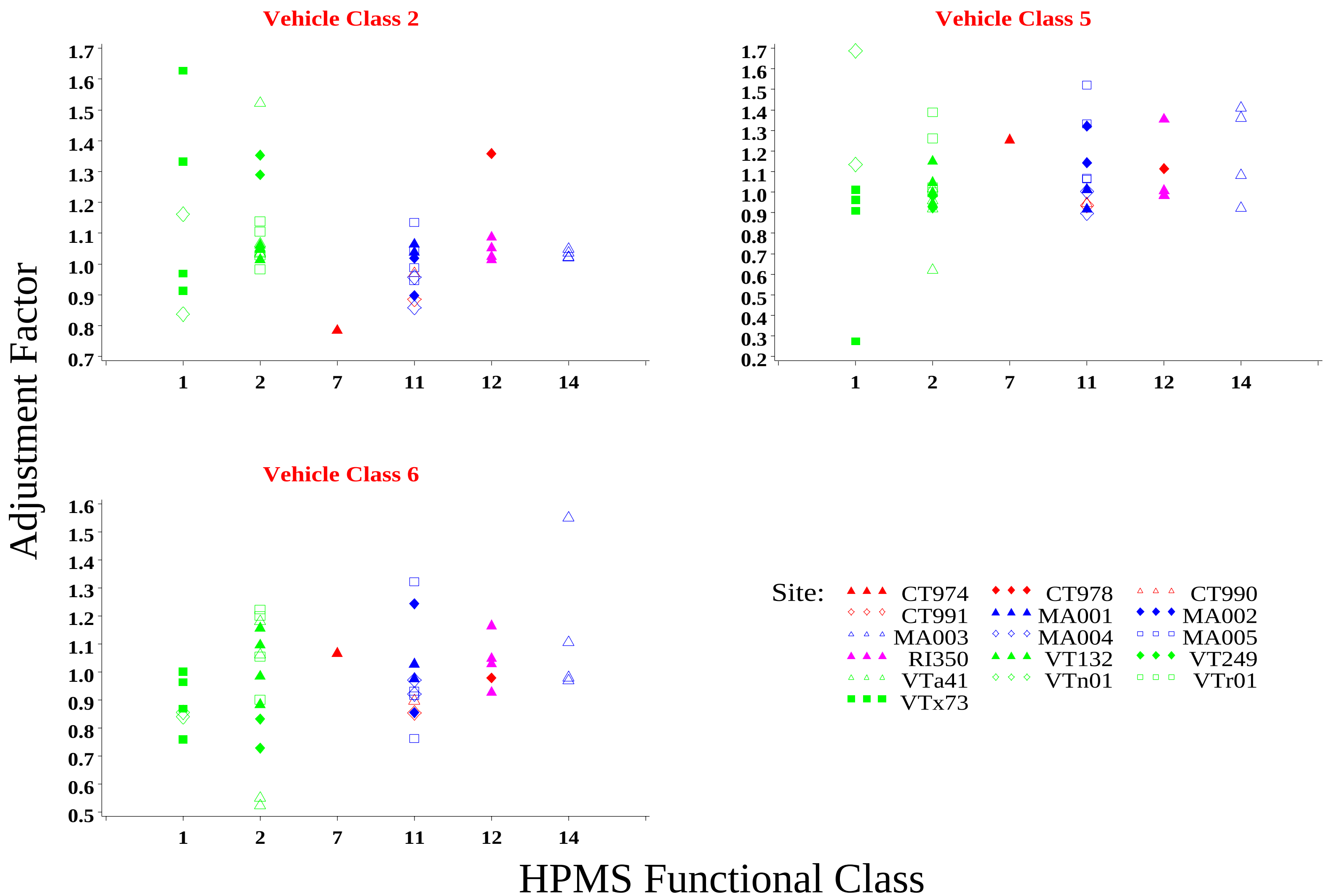


Figure 5. Example of adjustment factors—for August Wednesdays, selected vehicle classes—for input into a one-way ANOVA.

ANOVAs in HPMS functional class. As in Table 3, the ANOVAs produce the means for each HPMS class, which are the ANOVA AF estimates, standard errors for the means, and prediction standard errors. Standard errors for the means are useful for comparing the AF estimates, for example to investigate HPMS class differences. Prediction standard errors are larger than the corresponding standard errors for the means, because they reflect the error in the mean estimates plus the error in the data itself. (The standard error for the mean depends on not just the data error, but also the number of observations that go into each mean estimate.) The prediction standard errors can be entered into propagation of error formulas to yield overall standard errors for AADT or AADL estimates computed from short-term monitoring data.

Table 3. Class-Count AF Estimates and Standard Errors for August Wednesdays, Vehicle Class 5 (2-axle, 6-tire, single-unit trucks)

HPMS Class	AF Estimate	Prediction Std. Err.	Std. Err. Mean
1	0.830	0.198	0.075
2	0.781	0.189	0.049
7	1.467	0.259	0.183
11	0.804	0.191	0.053
12	0.882	0.201	0.082
14	1.338	0.205	0.092

To illustrate, suppose at a new site, single-day, August Wednesday classification counts are taken, and the count for Vehicle Class 5 is 500. Also suppose the new site is classed in HPMS functional class 11. To put the count of 500 on an annual basis, we need to multiply it by an adjustment factor. From Table 3, the AF for HPMS Class 11 is .804. Thus the Class 5 AADT estimate for the new site is AF $\times$ Count = .804 $\times$ 500 = 402. As in all good science we would also like to know how accurate the AADT estimate actually is. From propagation of error theory (variance of products), the standard

$$\text{Std. Err.(AFxCount)} = \text{Mean(AF) xMean(Count)} \left[[\text{CV(AF) xCV(Count)}]^2 + [\text{CV(AF)}]^2 + [\text{CV(Count)}]^2 \right]^{1/2}.$$

error of the AADT (AF $\times$ Count) is

From Section 3, an approximate CV for the count is $1/500^{1/2} = .045$. From Table 3, the CV for the AF (prediction std. Err.) is .191/.804=.238. Entering these into the above equation, and the AF and count for their means, gives 97.5, the standard error of the AADT estimate for the new site. Thus the AADT estimate, 402, should be qualified with plus-or-minus 97.5 or perhaps plus or minus 1.64 $\times$ 97.5 = 160 (90% confidence interval) to indicate that the “402” is far from exact. It is also interesting to note that because the AF CV (.238) is much bigger than the Count CV (.045), most of the variability in the AADT is coming from the AF, not the single-day count. Examples of propagation-of-error calculations for load estimates, via equation (1) for example, are in [5].

Table 3 represents only one of 84 possible day-of-week and month combination and only one vehicle class, but the table is not atypical in the sense of exhibiting substantial differences between HPMS classes. In Table 3, HPMS classes 1, 2, 11, and 12 appear similar, but different from classes 7 and 14. A more thorough examination of data for the other months, days-of-the-week, and vehicle classes, demonstrates many other AF differences that are big enough to be of practical importance (e.g., greater than 10%), and are also statistically significant (as indicated by a t-test based on the ANOVA standard errors of means). For example, in the following table for passenger cars (Vehicle Class 2), HPMS class 11 is substantially different from HPMS classes 1 and 2.

AF Estimates and Standard Errors for April Sundays, Vehicle Class 2 (Passenger Cars)

HPMS Class	AF Estimate	Prediction Std. Err.	Std. Err. Mean
1	0.988	0.193	0.073
2	1.011	0.185	0.048
7	1.148	0.253	0.179
11	1.240	0.186	0.052
12	1.111	0.196	0.080
14	1.384	0.200	0.089

The same kind of differences can also be seen in WIM estimates. For example, consider the following table, similar to the above two (computed by ANOVA) but with average annual daily loads (kips per vehicle) rather than AFs.

Average Annual Load per Vehicle Estimates and Standard Errors for
Vehicle Class 5 (2-axle, 6-tire, single-unit trucks)

HPMS Class	Mean kips per Vehicle	Prediction Std. Err.	Std. Err. Mean (kips per Vehicle)
1	11.91	1.34	0.40
2	11.76	1.43	0.64
7	8.04	1.81	1.28
11	13.43	1.36	0.45
12	13.06	1.40	0.57

Here HPMS Class 11 also differs appreciably from Class 1 and 2, though it is close to Class 12. In general we found that HPMS Classes 7, 11, 12, and 14 are all different and different from Classes 1 and 2. Classes 7, 11, 12, and 14 should be kept separate. Classes 1 and 2 tend to be similar (though there are exceptions). However, all of the HPMS Class 1 and 2 data considered here is from Vermont. Therefore, we feel the data is inadequate to support a recommendation to combine classes 1 and 2.

7. CONCLUSION

A primary consideration in this data-analysis research has been to arrive at reasonably simple, workable, approaches. On the other hand, many areas could be further explored: Directions (within sites) and different years were regarded as independent. In a more detailed analysis, their correlations could be modeled. Transformations such as the log transform should be considered as a means of making the data more normal and the ANOVAs more valid. Analyses of ANOVA residuals (observed minus predicted values) can also yield useful information about both transformations and data outliers. Although HPMS-classes are roughly similar in data scatter, heterogeneity of variance should also be investigated.

Seasonal and day-of-week effects were not discussed in detail here, but the need to adjust for them is clear. Adjustment factor calculations could be explored. Adjustment factors are inherently biased high, because their short-term (e.g., average for particular day-of-week and month) components enter as denominators. The bias follows from a statistical property of reciprocals (that the expectation of a reciprocal equals or exceeds the reciprocal of the expectation). This bias was also observed empirically in a study of traffic monitoring data from Florida and Washington [10]. The impact of the bias in cross-site extrapolations should be evaluated.

The data analysis did involve a lot of simplification, and our conclusions are still preliminary, but the main (tentative) conclusions are:

- Although combining vehicle classes does reduce the relative error in traffic counts, from the perspective of the statistical precision of load estimates, there is little advantage to combining vehicle classes.
- There is not sufficient evidence in the Region 1 data to support combining any of the HPMS functional classes.
- Data-sharing among the New England States is reasonable, as long as there is a proper accounting for the statistical error of estimates based on the common data. ANOVA provides a reasonably straightforward method for that accounting.

ACKNOWLEDGMENT

The authors would like to thank Tony Esteve and Fred Orlowski for help with many requests for general guidance, and, for help with particular questions, Mike Sprague (Rhode Island DOT), Anne-Marie McDonnel (Connecticut DOT), Ralph Gillman (FHWA), Phil Hughes and Steve Greene (Massachusetts DOT), Mike Morgan (Maine DOT), David Scott (Vermont AOT), and Subramanian

Sharma (New Hampshire DOT).

REFERENCES

[1] Hu, P. S., and Swank, P. (1997). *Analysis of Vehicle Classification and Truck Weight Data of the New England States, Task 1: Edit Vehicle Classification and Truck Weight Data, Task 2: Develop Descriptive Profiles of Data Collected at Each Site*, [one report each for:] *Connecticut, Massachusetts, Maine, Rhode Island, Vermont*, Drafts.

[2] Federal Highway Administration (1995). *Traffic Monitoring Guide, Third Edition, February, 1995*, U.S. DOT, Office of. Highway Information Management.

[3] Harvey, B. A., Champion, G. H., Ritchie, S. M., and Ruby, C. D. (1995) *Accuracy of Traffic Monitoring Equipment*, Georgia Technical Research Institute, A-9291.

[4] Federal Highway Administration, Office of Highway Information Management, *Vehicle Travel Information System, VTRIS, User's Guide, Version 2.6*.

[5] Schmoyer, R. L., and Hu, P. S. (1998). *Analysis of Vehicle Classification and Truck Weight Data of the New England States, Task 5: Analyze Truck Weight Data,* Draft.

[6] Page, E. S. (1961) "Cumulative sum charts," *Technometrics*, 3, 1-9.

[7] Snedecor, G. W., and Cochran, W. G. (1967), *Statistical Methods*, Iowa University Press, Ames, Iowa.

[8] Hallenbeck, M. E., Smith, B., Cornell-Martinez, Wilkinson, J. (1997), "Vehicle Volume Distributions by Classification," FHWA-PL-97-025.

[9] Kinateder, J. G., McMillan, N. J., Orban, J. E., Skarpness, B. O., and Wells, D. (1997), "Sampling Designs and Estimators for Monitoring Vehicle Characteristics Under Limited Inspection Capacity," *Transportation Research Record*, 1581, 72-81.

[10] Wright, T., Hu, P. S., Young, J., and Lu, A. (1997), *Variability in Traffic Monitoring Data,* Oak Ridge National Laboratory Technical Report.