A Statistical Approach to Statewide Traffic Counting

STEPHEN G. RITCHIE

A statistical framework that can be used for analysis of statewide traffic count data is described. A basis for designing a streamlined and cost-effective statewide traffic data collection program is also provided. The procedures described were developed as part of an in-depth evaluation study for the Washington State Department of Transportation. They were used to develop recommendations for an improved, statistically based, statewide highway data collection program. The program is intended to be implemented readily and is consistent with the FHWA Highway Performance Monitoring System and the recent FHWA draft Traffic Monitoring Guide. Several modifications (improvements) to the statistical framework of the latter for volume counting and vehicle classification were investigated, particularly methods of deriving estimates of annual average daily traffic (AADT) from short-duration axle counts at any location on the state highway system. AADT estimates can be derived for each vehicle type, if desired. The estimation of associated seasonal, axle correction, and growth factors is also described. The methodology enables the statistical precision of all estimates to be determined. The results obtained from applying these procedures to Washington State traffic data are presented.

For many years state departments of transportation (DOTs) have had responsibility for collecting a large amount of highway data. This has been undertaken to assist planning, design, and operations functions, as well as to comply with requirements and needs of other agencies including those at the federal level. However, collection of large amounts of data is costly. In a climate of increasing fiscal austerity at all levels of government and in all program areas, it is important not only that the right type of data is collected but that data are collected efficiently. Moreover, the data should meet the needs of the users with respect to type, amount, form, accuracy, and availability. A statewide highway data collection program should satisfy these criteria in an up-to-date and cost-effective manner.

In this paper a statistical framework that can be used for analysis of statewide traffic count data is described, and a basis for designing a streamlined and cost-effective statewide traffic data collection program is provided. The procedures described were developed as part of an in-depth evaluation study for the Washington State Department of Transportation (WSDOT) and were used to develop recommendations for an improved, statistically based, statewide highway data collection program (see paper by Ritchie and Hallenbeck in this Record).

Several studies have been reported in recent years that relate to general efforts to develop more cost-effective approaches to statewide highway data collection. These include the work of Hallenbeck and Bowman *(1)*, who proposed a general statewide traffic-counting program based on the Highway Performance Monitoring System (HPMS) *(2)*; the study by Wright Forssen Associates *(3)*, which evaluated, and developed improvement recommendations for, the highway data program of the Alaska Department of Transportation and Public Facilities; and work by the New York State Department of Transportation to streamline and reduce the cost of its traffic-counting program *(4)*. Although each of these studies provides useful background and guidance, the conceptual basis of Hallenbeck and Bowman *(1)*—utilizing the HPMS framework for purposes of statewide highway data collection—was explored in this study. There are a number of other relevant and useful works in the general area *(5–13)*. A comprehensive account of sampling theory as it has been developed for use in sample surveys is given by Cochran *(14)*.

In this paper, a statistical framework is presented for volume counting and vehicle classification, particularly for deriving estimates of annual average daily traffic (AADT) from short-duration axle counts at any location on a state highway system, using Washington State and WSDOT as a case study.

Department of Civil Engineering and Institute of Transportation Studies, University of California, Irvine, Calif. 92717.

ANNUAL AVERAGE DAILY TRAFFIC

Basic Model

A basic model for estimating AADT for a particular highway segment based on a single, short-duration count is

$$AADT = VOL(F_S)(F_A)(F_G) \tag{1}$$

where

- VOL = average 24-hr volume from a standard WSDOT 72-hr Tuesday-Thursday short count;
- F_S = seasonal factor for the count month;
- F_A = weekday axle correction factor if VOL is in axles; equal to 1 if VOL is in vehicles; and
- F_G = growth factor if VOL is not a current year count; equal to 1 otherwise.

To determine the relative precision of an estimated AADT from Equation 1, the coefficient of variation (ratio of standard deviation to mean) must be found. This can be obtained from the following approximate expression:

$$cv2(AADT) = cv2(F_S) + cv2(F_A) + cv2(F_G) \tag{2}$$

where each $cv2$ is the squared coefficient of variation of each variable. Thus the coefficient of variation of the AADT estimate is

$$cv(AADT) = [cv2(F_S) + cv2(F_A) + cv2(F_G)]^{0.5} \tag{3}$$

The relative precision (percentage) at a 100 (1 − α) percent confidence level is then given approximately by

$$\text{Precision } (AADT) = \pm\ 100Z_{\alpha/2}cv(AADT)\ \% \tag{4}$$

where $Z_{\alpha/2}$ is a standard normal statistic corresponding to the 100 $(1 - \alpha)$ percent confidence level (found in tables of any statistics book).

Also, a 100 $(1 - \alpha)$ percent confidence interval is defined approximately as

$$AADT \pm Z_{\alpha/2} AADT\ cv(AADT) \quad (5)$$

The Z-statistics corresponding to 95, 90, and 80 percent confidence levels are 1.96, 1.645, and 1.282, respectively.

Seasonal Factor Analysis

Factor Grouping

The data for analyzing seasonal factors were basically obtained from WSDOT *Annual Traffic Reports (15)*, which list the monthly permanent traffic recorder (PTR) traffic volumes throughout each year.

Several alternative methods for performing seasonal factoring were evaluated. The primary ones considered were

- Continued use of existing WSDOT Data Office procedures (see paper by Ritchie and Hallenbeck in this Record),
- Cluster analysis of PTRs,
- Procedures suggested in the FHWA draft counting guide *(13)*, and
- A revised FHWA procedure using linear regression.

The chosen strategy was the fourth of these options. The approach uses the basic method recommended by FHWA. The state highway system is stratified by geographic region and functional classification. The strata are then examined to determine which have similar seasonal patterns and might therefore be combined. PTR data from 1980 through 1984 were used to calculate the appropriate factor groups. The chosen groups were

- Rural Interstates,
- Urban roads,
- Other rural roads in the northeastern part of the state,
- Other rural roads in the southeastern part of the state,
- Other rural roads in the northwestern part of the state,
- Other rural roads in the southwestern part of the state, and
- Central mountain passes.

With the exception of the central mountain group, each factor group is defined by functional class of road and county boundaries. (Note that the urban group contains all state highways classified as urban regardless of county location.)

The advantages of the adopted approach are that

- The seasonal factors are statistically valid, meaning that the precision associated with any AADT estimate based on these factors can be calculated;
- The overall errors associated with this approach are equal to or smaller than the errors associated with any other seasonal factoring approach considered; and
- The factoring procedure is transparent to any user of volume information and thus allows the recalculation of the raw traffic count at some later time if desired.

Each of the other seasonal factor procedures had drawbacks that were judged unacceptable. For example, in the case of cluster analysis,

- The clusters computed were not consistent across years (i.e., PTRs changed groups from year to year), which means that roads should change groups as well, but no method was available to make that adjustment each year (see paper by Ritchie and Hallenbeck in this Record);
- Individual road sections are not easily or accurately assigned to cluster groups, irrespective of the difficulties mentioned previously; and
- The total error in the AADT estimate (including seasonal variation, daily variation, and variation in the axle correction factor) was only marginally better than that obtained by the recommended approach before inclusion of the indeterminate error that is present as a result of the first two points.

Regression Models

Seasonal factors for each month of the year were therefore derived for each of the seven factor groups described earlier. The modified FHWA approach adopted basically involved a regression analysis for each factor group for each month of AADT versus the average 24-hr short-count volumes that could be formed for each PTR from 72-hr Tuesday-Thursday counts in that month. The resulting regression coefficient of the short-count volume is then the derived seasonal factor for that factor group and month. This approach corresponds to the manner in which short counts are actually taken and converted to AADT estimates by WSDOT.

The first seasonal factor regression model estimated was as follows (note that the constant term is suppressed):

$$AADT = \beta\ VOL + u \quad (6)$$

where *AADT* and *VOL* are as defined previously, β is the regression coefficient (seasonal factor) to be estimated, and u is the error term. Such an equation would typically be estimated by ordinary least squares *(16)*. However, one of the required assumptions of that method is homoscedasticity, which means that the variance of the error term (u) is constant regardless of the magnitude of *VOL*. It often happens that this assumption is not valid (the case of heteroscedasticity) and the model must be reduced (by a transformation) to a form in which the error term does have a constant variance.

Estimation of Equation 6 revealed the presence of heteroscedasticity for some factor group and monthly traffic count data sets. Further, a consequence of this problem was that estimated variances would be biased and would underestimate the true variance. To address this issue, a commonly used transformation was employed to reduce Equation 6 to a homoscedastic form. It was assumed that the variance of the error term was known up to a multiplicative constant:

$$var\ (u) = \sigma^2\ VOL^2 \quad (7)$$

Dividing through Equation 6 by *VOL* yields

$$AADT/VOL = \beta + (u/VOL) \quad (8)$$

Substituting $e = u/VOL$ gives

$$AADT/VOL = \beta + e \tag{9}$$

where

$$\begin{aligned} var(e) &= (1/VOL^2)\, var(u) \\ &= (1/VOL^2)\, \sigma^2\, VOL^2 \\ &= \sigma^2 \end{aligned}$$

Thus, the variance of the error term (e) in Equation 9 is constant (σ^2) and ordinary least squares estimation methods can be applied. The form of Equation 9 is now so simple that computerized regression packages are not really required. The estimation results can be obtained as follows:

$$\hat{\beta} = \sum_{i=1}^{n} (AADT/VOL_i)/n \tag{10}$$

$$\hat{\sigma}^2 = \left\{ \sum_{i=1}^{n} [(AADT_i/VOL_i) - \hat{\beta}]^2 \right\} /(n-1) \tag{11}$$

$$var(\hat{\beta}) = \hat{\sigma}^2/n \tag{12}$$

and the t-statistic on $\hat{\beta}$ is

$$t_{\hat{\beta}} = \hat{\beta}(n^{1/2})/\hat{\sigma} \tag{13}$$

In Equations 10 and 11 the subscript i refers to each short count in the month for the factor group, and n represents the number of counts.

Finally, the relative precision of the AADT estimates must be derived. When the seasonal factors from Equation 9 are applied to counts in the following year, the value of the ratio $AADT/VOL$ in the equation is forecast. Therefore the appropriate variance measure is the variance of the prediction error for the forecast ratio of $AADT$ to VOL. It can be shown that this variance is given by

$$\sigma^2 (1 + 1/n) \tag{14}$$

for each factor group and month. The required coefficient of variation for Equation 3 is then

$$cv(F_S) = \hat{\sigma}(1 + 1/n)^{0.5}/\hat{\beta} \tag{15}$$

It is interesting to note that this theoretically derived result is equivalent to that obtained by more qualitative reasoning *(1, 13)*.

Results

The seasonal factors for 1984, derived using the procedures described, are given in Table 1 for April through September (the period when WSDOT performs the vast majority of its traffic counting) and in Table 2 for October through March. Because of the high variability of factors for the central mountain group, this group was treated separately.

The coefficients of variation, based on Equation 15, are given in Table 3. These have been used to calculate relative precision levels of April through September AADT estimates, as given in Table 4, without incorporating axle correction or growth factors.

TABLE 1 1984 SEASONAL FACTORS FOR APRIL THROUGH SEPTEMBER

	Month					
Group	April	May	June	July	August	September
Rural Interstate	1.132	1.126	0.960	0.907	0.849	0.990
Urban	0.966	0.952	0.903	0.894	0.878	0.907
Northwestern	1.023	0.995	0.921	0.848	0.812	0.957
Southwestern	1.087	1.055	0.935	0.823	0.769	0.925
Southeastern	1.137	1.077	0.956	0.896	0.855	0.979
Northeastern	1.025	0.927	0.895	0.754	0.779	0.882

It is also interesting to note how the AADT precision levels vary as a function of the number of PTRs in each factor group. Little improvement in relative precision was obtained beyond about six to eight PTRs per group. Thus, in terms of statistical precision of AADT estimates only, little is gained by having additional PTRs. However, as discussed by Ritchie and Hallenbeck in this Record, there may be other reasons for maintaining large numbers of PTRs in any group, such as the automatic collection of vehicle classification data.

Axle Correction Factor Analysis

Axle correction factors are required to convert short-count volumes to AADT estimates when those short counts are obtained using equipment that records axles rather than vehicles. Calculation of the factors requires vehicle classification information (percentage of vehicles in each class) as well as knowledge of the number of axles per vehicle in each vehicle class.

The average number of axles per vehicle (A_V) in a given factor group (typically highway functional class) is given by

$$A_V + \sum_C (Axles_C)(P_C) \tag{16}$$

where $Axles_C$ is the number of axles per vehicle in Class C and P_C is the proportion of vehicles in Class C (system-level estimate). The variance of A_V is then given by

$$var(A_V) = \sum_C (Axles_C)^2\, var(P_C) \tag{17}$$

where $var(P_C)$ is the variance of Vehicle Class C proportion, from a vehicle classification study.

Thus the coefficient of variation of A_V is

$$cv(A_V) = \left[\sum_C (Axles_C)^2\, var(P_C)\right]^{0.5} \Big/ \left[\sum_C (Axles_C)(P_C)\right] \tag{18}$$

However, the desired axle correction factor (F_A) is actually the inverse of A_V:

$$F_A = A_V^{-1} \tag{19}$$

It can be shown by a first-order Taylor series approximation that

$$cv(F_A) = cv(A_V) \tag{20}$$

TABLE 2 1984 SEASONAL FACTORS FOR OCTOBER THROUGH MARCH

Group	Month					
	October	November	December	January	February	March
Rural Interstate	1.274	1.220	1.116	1.554	1.425	1.238
Urban	1.045	1.006	0.935	1.088	1.033	0.988
Northwestern	1.236	1.124	1.067	1.296	1.558	1.075
Southwestern	1.467	1.283	1.067	1.408	1.259	1.145
Southeastern	1.500	1.318	1.043	1.595	1.472	1.259
Northeastern	1.339	1.176	0.981	1.200	1.184	1.163

TABLE 3 COEFFICIENTS OF VARIATION OF 1984 SEASONAL FACTORS, $cv(F_S)$

Month	Factor Group					
	Rural Interstate	Urban	Northwestern	Southwestern	Southeastern	Northeastern
January	0.172	0.090	0.149	0.216	0.196	0.074
February	0.150	0.073	0.105	0.154	0.190	0.100
March	0.113	0.057	0.102	0.147	0.180	0.146
April	0.109	0.062	0.095	0.132	0.144	0.123
May	0.089	0.070	0.078	0.108	0.138	0.080
June	0.064	0.057	0.095	0.082	0.118	0.077
July	0.057	0.063	0.092	0.077	0.115	0.104
August	0.064	0.042	0.090	0.143	0.090	0.097
September	0.090	0.059	0.069	0.129	0.112	0.086
October	0.167	0.112	0.150	0.217	0.239	0.176
November	0.255	0.090	0.130	0.186	0.250	0.115
December	0.078	0.073	0.084	0.114	0.088	0.083

This result permits the coefficient of variation of the axle correction factor to be derived readily from Equation 18 for insertion into Equation 3.

Table 5 gives the estimated axle correction factors for eight functional classes of highway, together with relative precisions and coefficients of variation.

Growth Factors

Growth factors often represent a relatively minor part of the factoring process to obtain AADT estimates from short counts. However, at times an old count must be converted to a more recent AADT by means of a growth factor. Several methods exist for estimating growth factors. In general, the approaches are fairly crude ways of attempting to account for traffic growth or decline over time. The analysis discussed in this section was exploratory only, although the results appear reasonable.

Simple growth factors were estimated for each of the previously identified seasonal factor groups for 1982–1983 and 1983–1984. The factors were obtained by forming the ratio of AADT in the more recent year to that in the earlier year for each PTR in a group and applying the regression analysis procedure discussed previously. In one group there was one PTR, and in a second group no PTR, for both years, so that coefficients of variation of the factors (F_G) could not be formed. Table 6 gives the estimated growth factors for each period together with their coefficients of variation.

TABLE 4 RELATIVE PRECISION (%) OF SEASONALLY ADJUSTED AADT ESTIMATES FROM SHORT COUNTS IN EACH MONTH (without incorporating axle correction or growth factors)

Month	Factor Group					
	Rural Interstate	Urban	Northwestern	Southwestern	Southeastern	Northeastern
April	18	10	16	22	24	20
May	15	12	13	18	23	13
June	11	9	16	13	19	13
July	9	10	15	13	19	17
August	11	7	15	24	15	16
September	15	10	11	21	18	14

Note: 90 percent confidence level.

TABLE 5 AXLE CORRECTION FACTORS

Functional Class	F_A[a]	Percentage Precision[b]	$cv(F_A)$
Rural Interstate	0.423	10.2	0.062
Rural principal arterial	0.461	8.8	0.053
Rural minor arterial	0.471	4.8	0.029
Rural collector	0.459	10.7	0.066
Urban Interstate	0.454	3.9	0.023
Urban principal arterial	0.463	6.8	0.041
Urban minor arterial	0.482	2.1	0.013
Urban collector	0.495	1.6	0.010

[a]Weekday factors.
[b]90 percent confidence level.

TABLE 6 GROWTH FACTORS

Group	1982–1983		1983–1984	
	F_G	$cv(F_G)$	F_G	$cv(F_g)$
Rural Interstate	1.065	0.020	1.024	0.037
Urban	1.175	0.306	1.046	0.066
Northwestern	1.052	0.110	1.016	0.055
Southwestern	1.059	—	1.094	—
Southeastern	1.041	0.060	1.041	0.042
Northeastern	—	—	—	—

VEHICLE CLASSIFICATION

Data Analysis

Because of the limited nature of vehicle classification counts taken by WSDOT in recent years, the best available data set for statistical analysis was from a 1980–1981 study that was done for FHWA. Unlike volume counts, which utilize a system of PTR stations for continuous monitoring, it is not presently possible to derive vehicle classification seasonal factors for conversion of a single (say 24-hr) classification count to an annual average estimate for a given highway segment. Rather, the data available permit only an approximate systemwide plan to be developed for an annual counting program on different functional classes, in order to derive annual average vehicle classification results. Improvements to the department's current vehicle classification activities are discussed further by Ritchie and Hallenbeck in this Record.

The 1980–1981 data consist of 248 manual 24-hr vehicle classification counts. The data were collected at 31 locations across the state with 4 weekday counts (one per season) and 4 weekend counts (one per season) at each location. For purposes of analysis, the data were reduced to six vehicle types:

1. Cars,
2. Two-axle trucks,
3. Three-axle trucks,
4. Four-axle trucks,
5. Five-axle trucks, and
6. Trucks with six or more axles.

In addition, a slightly more detailed set of functional classifications than was used in the seasonal factor development was retained for initial analysis. These functional classes consisted of eight groups: Interstates, principal arterials, minor arterials, and collectors for both rural and urban locations.

The principal analysis method used was a two-stage cluster sampling approach with multiple strata. The first set of strata corresponded to functional classes. Within strata, the primary sampling units or clusters were possible count locations, and the secondary or elementary sampling units were days at each location (required to be the same at each location in a stratum). The second stratification was introduced with respect to weekdays and weekend days because vehicle classifications were noticeably different across these strata; truck percentages were often considerably lower on weekend count days. The population sizes for each stage were taken to be the number of HPMS population sections in each functional class in the case of locations and, at the second stage, simply the number of weekdays or weekend days, or both, in a year. Allowance was also made in the analysis for the unequal size of the second-stage units (as is often assumed in cluster analysis) due to the daily variations in traffic volume throughout the year.

Within each functional class, and for each Vehicle Class C, the average (weighted) vehicle proportion (P_C) was estimated as

$$P_C = \left(\sum_{i=1}^{n} p_i\right)/n \qquad (21)$$

where

$p_i = w_1 p_{i1} + w_2 p_{i2}$;
p_i = proportion at location i;
p_{i1} = weekend proportion at location i

$$= \left(\sum_{k=1}^{m_1} C_{ik1}\right)\Big/\left(\sum_{k=1}^{m_1} X_{ik1}\right);$$

p_{i2} = weekday proportion at location i

$$= \left(\sum_{j=1}^{m_2} C_{ij2}\right)\Big/\left(\sum_{j=1}^{m_2} X_{ij2}\right);$$

C_{ik1} = total number of vehicles of type C at station i on weekend day k;
C_{ij2} = total number of vehicles of type C at station i on weekday j;
X_{ik1} = total number of vehicles at station i on weekend day k;
X_{ij2} = total number of vehicles at station i on weekday j;
p_{i1k} = proportion observed on weekend day k;
p_{i2j} = proportion observed on weekday j;
m_1 = number of weekend days at each location;
m_2 = number of weekdays at each location;
w_1 = 2/7;
w_2 = 5/7; and
n = number of count locations.

The variance was obtained from

$$var\,(P_C) = (1 - f_1)(s_1^2/n) + [w_1^2(1 - f_{21})s_{21}^2/(nm_1) + w_2^2(1 - f_{22})s_{22}^2/(nm_2)] \qquad (22)$$

where

f_1 = n/N,
N = population size of HPMS segments for functional class,
f_{21} = $m_1/104$,
f_{22} = $m_2/261$,

$$s_{21}^2 = \left(\sum_{i=1}^{n} s_{21i}^2\right)/n,$$

$$s_{21i}^2 = \sum_{k=1}^{m_1} (p_{i1k} - p_{i1})^2/(m_1 - 1),$$

$$s_{22}^2 = \left(\sum_{i=1}^{n} s_{22i}^2\right)/n,$$

$$s_{22i}^2 = \sum_{j=1}^{m_2} (p_{i2j} - p_{i2})^2/(m_2 - 1), \text{ and}$$

$$s_1^2 = \sum_{i=1}^{n} (p_i - p_c)^2/(n - 1).$$

Thus the coefficient of variation of the estimate is

$$cv(P_C) = [var(P_C)]^{0.5}/P_C \tag{23}$$

The relative precision (percentage) at a 100 (1 − α) percent confidence level is then given approximately by

$$\text{Precision } (P_C) = \pm\ 100\ Z_{\alpha/2}cv(P_C) \tag{24}$$

In addition to this analysis approach, which distinguishes between counts on weekdays and weekends by introducing sample stratification, estimates for P_C were also calculated without this stratification by pooling weekday and weekend counts at each location. For this simpler formulation, P_C is calculated from

$$P_C = \left(\sum_{i=1}^{n}\sum_{j=1}^{m} C_{ij}\right)\Big/\left(\sum_{i=1}^{n}\sum_{j=1}^{m} X_{ij}\right) \tag{25}$$

where

C_{ij} = total number of vehicles of type C at station i on day j,
X_{ij} = total number of vehicles at station i on day j,
f_1 = n/N,
f_2 = $m/365$,
m = number of days sampled at each station, and
n = number of count locations.

The variance of P_C is then calculated from

$$var(P_C) = (1 - f_1)(S_1^2/n) + f_1(1 - f_2)(s_2^2/mn) \tag{26}$$

where s_1^2 is as previously defined, and

$$s_2^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} (p_{ij} - \bar{p}_i)^2/[n(m - 1)],$$

$$\bar{p}_i = \sum_{j=1}^{m} C_{ij}/\sum_{i=1}^{m} X_{ij}, \text{ and}$$

$$p_{ij} = C_{ij}/X_{ij}.$$

The coefficient of variation and precision of P_C are then calculated as before by Equations 23 and 24, respectively.

Results

Table 7 gives the classification count results for each functional class. These averages are based on the weighted weekday and weekend counts. Table 8 gives the relative precision of these results at a 90 percent confidence level. Clearly, the precision of the estimates for large trucks (five or more axles) is relatively poor, although this was not unexpected given the limited nature of the counts and the inherent variability of truck travel as a percentage of total daily volume. Table 9 gives the coefficients of variation for each vehicle class proportion.

TABLE 7 PERCENTAGE OF VEHICLES BY TYPE IN EACH FUNCTIONAL CLASS

	Vehicle Class					
Functional Class	1	2	3	4	5	6
Rural Interstate	87.0	3.1	0.6	0.3	8.3	0.8
Rural principal arterial	90.3	3.2	1.0	0.1	5.0	0.3
Rural minor arterial	92.2	2.9	0.9	0.1	3.5	0.5
Rural collector	89.3	3.5	3.0	0.3	3.6	0.3
Urban Interstate	91.1	2.8	0.7	0.4	4.5	0.4
Urban principal arterial	90.8	3.1	0.6	0.2	4.9	0.4
Urban minor arterial	94.4	2.8	0.8	0.2	1.7	0.2
Urban collector	95.1	3.4	0.4	0.1	0.9	0.1

The estimation of annual average daily truck traffic (AADTT) volume can be accomplished readily by applying the analysis results and extending the AADT estimation equations:

$$AADTT = VOL\ (F_S)(F_A)(F_G)(P_C) \tag{27}$$

where P_C is the appropriate vehicle proportion estimate from Table 7 and all other notations are as defined previously. It must be remembered that this AADTT estimate is based on system-level vehicle classification data not a specific truck count for the section where the volume count *(VOL)* was taken.

The coefficient of variation can be obtained from

$$cv(AADTT) = [cv2(F_S) + cv2(F_A) + cv2(F_G) + cv2(P_C)]^{0.5} \tag{28}$$

where $cv(P_C)$ is as given in Table 9. The relative precision at a 100 (1 − α) percent confidence level is then given approximately by

$$\text{Precision } (AADTT) = \pm\ 100\ Z_{\alpha/2}\ cv(AADTT)\ \% \tag{29}$$

TABLE 8 RELATIVE PRECISION (%) OF VEHICLE CLASSIFICATION RESULTS

	Vehicle Class					
Functional Class	1	2	3	4	5	6
Rural Interstate	4	11	13	35	35	33
Rural principal arterial	3	7	50	43	43	48
Rural minor arterial	2	9	22	45	33	68
Rural collector	7	29	82	62	91	69
Urban Interstate	1	8	13	22	20	14
Urban principal arterial	3	17	22	39	41	40
Urban minor arterial	1	26	31	67	19	44
Urban collector	1	25	35	43	34	86

Note: 90 percent confidence level.

TABLE 9 COEFFICIENTS OF VARIATION FOR VEHICLE PROPORTIONS FROM TABLE 7

	Vehicle Class					
Functional Class	1	2	3	4	5	6
Rural Interstate	0.024	0.068	0.079	0.213	0.215	0.201
Rural principal arterial	0.018	0.044	0.303	0.263	0.259	0.294
Rural minor arterial	0.010	0.057	0.134	0.271	0.201	0.416
Urban Interstate	0.007	0.050	0.077	0.131	0.119	0.088
Urban principal arterial	0.018	0.103	0.134	0.237	0.247	0.241
Urban minor arterial	0.008	0.157	0.187	0.405	0.114	0.266
Urban collector	0.007	0.150	0.216	0.260	0.207	0.522

As an example, consider the calculation of an annual average daily five-axle truck volume on a rural Interstate segment, based on a short duration axle count in June:

Average 24-hr volume *(VOL)* = 50,000 axles,

F_S = 0.960 (Table 1),
F_A = 0.423 (Table 5),
F_G = 1.0 (because this is a current-year count),
P_C = 0.083 (Table 7),
$cv(F_S)$ = 0.064 (Table 3),
$cv(F_A)$ = 0.062 (Table 5),
$cv(F_G)$ = 0.0 (because an estimated factor is not used), and
$cv(P_C)$ = 0.215 (Table 9).

Thus, from Equation 21, the estimate of daily five-axle trucks is

$$AADTT = 50{,}000\,(0.960)(0.423)(1.0)(0.083) = 1{,}685 \text{ five-axle trucks.}$$

From Equation 22, the coefficient of variation of this estimate is

$$CV(AADTT) = [(0.064)^2 + (0.062)^2 + (0.0)^2 + (0.215)^2]^{0.5} = 0.233.$$

Finally, from Equation 23, the relative precision of this estimate at a 90 percent confidence level is

$$\text{Precision } (AADTT) = \pm\, 100\,(1.645)(0.233)\ \% = \pm\, 38.3\ \%,$$

which means there is 90 percent confidence that the true value of AADTT is within about 40 percent of the estimate of 1,685 five-axle trucks per day.

Sample Design

The results obtained from these analyses of vehicle classification data provided some basis for developing the study recommendations for this data item (see paper by Ritchie and Hallenbeck in this Record). Some of the findings related to design of a sample for collecting vehicle classification data are presented in this subsection.

Of interest is how the statistical precision of classification estimates is affected by sample size and choice of confidence level. To gain further insight into these relationships, a number of tabular and graphic reports were generated.

For example, Table 10 gives the variation in precision achieved with a number of different sample designs in the case of rural Interstates. These results are based on a cluster analysis, as before, but with pooled weekend and weekday counts without stratification. It can be seen that the precision levels are more sensitive to the number of locations chosen than the number of days surveyed per location. For a given number of classification counts, the results indicate that it is better to take all of those counts at different locations, with only one count per location, on randomly chosen days during the year.

TABLE 10 RELATIVE PRECISION (%) OF RURAL INTERSTATE VEHICLE CLASSIFICATIONS FOR DIFFERENT SAMPLE DESIGNS

No. of Locations	No. of Days	No. of Counts	Vehicle Class 1	2	3	4	5	6
2	1	2	9	37	81	95	105	105
2	5	10	6	23	39	68	71	69
4	1	4	7	26	57	67	74	74
4	5	20	4	16	27	48	50	49
8	1	8	5	18	40	47	52	52
8	5	40	3	11	19	34	35	34
20	1	20	3	12	25	29	33	33
20	5	100	2	7	12	20	21	21
40	1	40	2	9	18	20	23	23
40	5	200	1	5	8	13	14	14

Note: 90 percent confidence level.

To avoid the added complexity and cost of having to take at least two counts per location (one weekday, one weekend) at every sampled location, as required by the stratified cluster analysis procedure, it was decided that, for purposes of sample design and implementation, a pooled cluster analysis approach should be used without stratification by day of week. All that this would mean in practice is that the count day or days at a location would be chosen randomly from all days in the year. Given the nature of the data on which the analyses were based and the interim nature of any recommended manual count program [due to introduction of automatic vehicle classifiers by the department (see paper by Ritchie and Hallenbeck in this Record)], this approach was judged appropriate.

Also investigated was the effect of both confidence level and number of counts (or locations counted) on the precision of vehicle proportions. Achieving both smaller precision levels and higher confidence levels requires that more counts be taken. In the case of five-axle trucks on rural Interstates it was noted for example that the major improvement in precision came from taking approximately 20 counts and that the improvement in precision for successive counts was relatively small. However, the magnitude of the precision was still undesirably high. The implication is that, to achieve precise results, a much larger number of vehicle classification counts than the department currently collects are required. The detailed recommendations that were developed on the basis of these results are reported by Ritchie and Hallenbeck in this Record.

CONCLUSIONS

A rigorous statistical approach to statewide data collection and program design permits the estimation of data precision and can provide a rational basis to assist in allocating limited resources among the various possible data collection activities. A statistical approach is also important because the desired precision and confidence level have a major impact on sample design and cost. There is little point in collecting more precise sample data at a higher level of confidence than is required by the data users, particularly when considerable cost savings can be realized by using smaller sample sizes. Conversely, when resources are limited and insufficient for the desired sample size, trade-offs between precision and level of confidence can be made explicit. Further discussion of this issue is presented in a companion paper by Ritchie and Hallenbeck in this Record.

A statistical framework for volume counting and vehicle classification, and particularly for deriving estimates of AADT from short-duration axle counts at any location on a state highway system, has been presented. AADT estimates can be derived for each vehicle type, if desired. The estimation of associated seasonal, axle correction, and growth factors was also described. The methodology enables the statistical precision of all of these estimates to be determined.

ACKNOWLEDGMENT

Many individuals from WSDOT, the Washington State Transportation Center (TRAC), and FHWA contributed significantly to this study, and their input is gratefully acknowledged. This paper is based on a research project conducted by the University of Washington and TRAC for the Washington State Transportation Commission and Department of Transportation, in cooperation with the FHWA, U.S. Department of Transportation.

REFERENCES

1. M. E. Hallenbeck and L. A. Bowman. *Development of a Statewide Traffic Counting Program Based on the Highway Performance Monitoring System.* FHWA, U.S. Department of Transportation, 1984.
2. *Highway Performance Monitoring System Field Manual for the Continuing Analytical and Statistical Data Base.* Office of Highway Planning, FHWA, U.S. Department of Transportation, 1984.
3. Wright Forssen Associates, Comsis Corporation, and TRANSPO Group. *Highway Traffic Data Study,* Vol. I: *Final Report.* Alaska Department of Transportation and Public Facilities, Planning and Programming Division, Southeast Region, 1983.
4. D. T. Hartgen and J. H. Lemmerman. "Streamlining Collection and Processing of Traffic Count Statistics." In *Transportation Research Record 928,* TRB, National Research Council, Washington, D.C., 1983, pp. 11–20.
5. Peat, Marwick and Mitchell. *Guide to Urban Traffic Volume Counting.* Final Report. FHWA, U.S. Department of Transportation, 1980.
6. J. DiRenzo, L. Bowman, and M. Hallenbeck. *Review of the Department's Traffic Count Program.* Pennsylvania Department of Transportation, Harrisburg, 1983.
7. John Hamburg & Associates. *Improved Methods for Vehicle Counting and Determining Vehicle-Miles of Travel—Final Report on State of the Art.* NCHRP, TRB, National Research Council, Washington, D.C., 1982.
8. John Hamburg & Associates. *Improved Methods for Vehicle Counting and Determining Vehicle-Miles of Travel—Traffic Counting Program Design.* NCHRP, TRB, National Research Council, Washington, D.C., 1982.
9. L. T. Hoang and V. P. Poteat. "Estimating Vehicle Miles of Travel by Using Random Sampling Techniques." In *Transportation Research Record 779,* TRB, National Research Council, Washington, D.C., 1980, pp. 6–10.
10. L. M. Rudman. "Vehicle Kilometers Traveled: Evaluation of Existing Data Sources." In *Transportation Research Record 726,* TRB, National Research Council, Washington, D.C., 1979, pp. 19–28.
11. D. L. Green and A. S. Loebl. *Vehicle-Miles of Travel Statistics, Lifetime Vehicle-Miles of Travel and Current Methods of Estimating Vehicle-Miles of Travel.* Oak Ridge National Laboratory, Oak Ridge, Tenn., 1979.
12. J. P. Mahoney. *Traffic Data Needs for Pavement Rehabilitation.* Washington State Transportation Commission, 1983.
13. *Draft Traffic Monitoring Guide.* Office of Highway Planning, FHWA, U.S. Department of Transportation, July 1984.
14. W. G. Cochran. *Sampling Techniques.* John Wiley & Sons, Inc., New York, 1977.
15. *Annual Traffic Reports 1980–1984.* Planning, Research and Public Transportation Division, Washington State Department of Transportation, Olympia, 1983, 1983, 1983, 1984, and 1985.
16. G. S. Maddala. *Econometrics.* McGraw-Hill Book Company, New York, 1977.

The contents of this paper reflect the views of the author who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Washington State Transportation Commission, Department of Transportation, or the FHWA. This paper does not constitute a standard, specification, or regulation.