Volume-Based Model for Forecasting Truck Lane Use on the Rural Interstate

D. ALBRIGHT AND C. BLEWETT

Truck lane use is important when assessing the impact of trucks on multilane roads. A study was conducted to estimate truck lane use on tangent sections of the New Mexico rural interstate. Following review of related literature, a field study methodology was developed, data were collected, and a model was derived for estimating a truck lane distribution factor. The study determined that total vehicle volume is statistically significant in explaining truck lane use. The percent trucks of total volume also adds to the explanatory ability of the model. Three models were developed from the field data, based on three traffic volume groups. To diminish the possibility of underbuilding a facility, precision ranges were established for the predictive ability of the models. The models were modified to account for variability in the data. An analysis was conducted of the impact of study models on pavement structural design. The study models, when compared with previous lane use estimates, typically reduced pavement thickness by ¾ in. For all study sites the reduction in pavement thickness was between ¾ and 1 in.

The lane distribution factor (LDF) is an integral part of load calculations on multilane facilities. The LDF describes the percent of heavy commercial vehicles in the right lane. This factor is used in design of pavement thickness, affecting pavement design and construction costs. For the New Mexico rural interstate system, the truck lane distribution factor for pavement design assumed that 94% of the heavy commercial traffic travels in the right-hand lane and the remaining 6% occupies the left lane. These values were estimates and were not specifically based on New Mexico traffic data. To determine if these factors accurately described New Mexico rural interstate truck lane use, a research program was begun to review the literature on truck lane distribution.

The review of existing literature yielded two conclusions. First, little research has been published on the subject. Second, the published studies available were completed before enactment of the Surface Transportation Assistance Act of 1982. Because heavy commercial vehicle characteristics have changed significantly since that legislation, the New Mexico State Highway Department felt that the issue was important enough to warrant a study. The objectives of the study were to evaluate the existing procedure and, if appropriate, to develop a new truck lane distribution formula.

STUDY DESIGN

The truck lane distribution study involved three elements: (1) the methodology to facilitate accurate data collection and ensure the validity of the data for statistical inquiry; (2) the process of data collection, assimilation and summarization, including the verification of base data; and (3) the statistical analysis of summary data and generation of a valid set of lane distribution postulates.

New Mexico State Highway Department, Santa Fe, N. Mex. 87504-1149.

Methodology

Based in part on the literature, the research procedure was modified to address the problem of personnel constraints and accurate data collection. The first step was to determine which variables were most important for the study. The percent of heavy commercial traffic in the right lane was defined by the nature of the study as the dependent variable. The remaining issues to be answered by the design were the selection of measurable independent variables, determination of sample size, and site selection.

The selection of independent variables is typically limited by resources. Variables are selected which previous experience or inquiry indicate have promise. These variables are then reduced by some scale of priority, according to the resources available.

Based on previous research, notably that conducted in 1971 by M. M. Alexander and R. A. Graves for the Georgia Highway Department (*1*), the independent variables for the study were limited to total traffic volume, percent heavy commercial with five and more axles, and percent heavy commercial with less than five axles. The last two independent variables were totaled separately by lane during data collection and analysis.

Vehicle speed was considered as a variable but was not selected. Driver detection of radar might result in atypical data, particularly among heavy commercial vehicles. Additionally, if speed were determined to improve the lane use model, it would have to be collected for each site where lane use was to be calculated, increasing the cost of model implementation. For the remaining variables used in the study, system-wide data were available so that lane use calculations could be readily made. This is a critical element of a useful model for determining truck lane use distribution.

Sample Size and Site Selection

After selection of independent variables, the next step was to define the universe of the study and identify an adequate sample size. A previous study determined that New Mexico's 900-mile rural interstate system could be subdivided into fourteen sections with unique volume characteristics. Because these fourteen sections represented the entire rural interstate system, and because the determinant criteria for a unique section were volume and percent heavy commercial of total traffic, the sections were appropriate units for determining sample size. A sample of five sites was randomly selected from these sections. After the five sites were selected, they were reviewed to determine if they were representative of the New Mexico rural interstate. The review resulted in the addition of another site representative of lower traffic volumes and lower percent heavy commercial of total traffic. All subsequent analysis in this study is based on data from these sites.

Once the six sections were identified, the road segment for the count was identified. Each section has a corresponding permanent counter located in a unique road segment. In the interest of accurate data collection, the field data collector would count only heavy commercial traffic. The total volume would be determined by the permanent counter. Therefore, the road segment for the count was located between the same two interchanges as the permanent counter. Because access to the rural interstate system is controlled, the heavy commercial counts could be compared directly with total volume counts of the permanent counter for the same hours to complete the data base.

The remaining concern for site selection was the minimum distance the site had to be located from the nearest interchange. Based on standard acceleration rates, it was determined that the site for the count would preferably be 2 miles from the nearest interchange. This constraint ensures that the site will not be affected by atypical lane distributions that result from merging. With typically 5 miles between interchanges, there was a minimum of a 1-mile road segment within which the specific count site was selected.

Count Days and Hours

In order to minimize seasonal factors affecting the study, the month of September was chosen as the time to complete the field counts. Historical volume data for the New Mexico rural interstate indicate that the traffic ratio of September mean daily traffic to annual mean daily traffic is close to unity.

At the time of site selection, it was decided that each count would be made on three consecutive days. To avoid weekend variation in traffic volume and to provide adequate time for personnel travel and set-up, the study count days were Tuesday, Wednesday, and Thursday.

An analysis of the daily hourly volumes at the six sites during previous Septembers revealed no discernible morning or afternoon peak in all cases. The typical daily pattern reflected a fairly continuous flow of traffic from 9:00 a.m. until 5:00 p.m. The data did suggest that in most cases the volume decreased from noon to 2:00 p.m. The count hours for the study were selected as 9:00 a.m. to noon, and from 2:00 p.m. to 5:00 p.m. for all six locations. Based on previous years' data, these hours would include any peaking characteristics for all locations and represent characteristic traffic volume at the sites.

DATA COLLECTION, ASSIMILATION, AND SUMMARIZATION

After all field data were collected, the raw data were assimilated and summarized for analysis. This process involved surveying the raw data, eliminating unusable records, aggregating the data to different levels, and determining the most useful summary data set for the development of lane distribution factors.

For each of the six sites, data were collected for both directions. This, in effect, doubled the site size for data summarization purposes. Under optimal conditions, 36 hours of data would be available, by direction, for each study site. Of the total of 216 possible hours of observation, it was determined that 44 could not be used in the analysis. Of these, 2 hours were discarded because of lane closures, 6 were lost due to field transportation problems, and the remaining 36 were dropped because volumes from the permanent counter were suspect. Fortunately, the 172 good records were distributed evenly among the six sites and included data for traffic moving in both directions at each site. No sites were dropped from the study because of insufficient data, and no recounting was required.

To analyze the variance of the data, the mean square error of each variable was examined at each of the six sites, by direction. A high within-group variance would suggest that sampled traffic and lane patterns were not uniform. Alternatively, with review of the standard error of the mean to estimate sampling error, high within-group variance might suggest that the three-day sample was inadequate and the study would have to be redesigned. On the other hand, given low within-group variance as indicated by mean square error and standard error, consistent site-specific traffic and lane use patterns could be concluded, and further analysis of the data would be indicated. The within-group mean square error was found to be low. The mean statistics for each site, and each site by direction, adequately represent traffic in the lane distribution factor model.

The relationship among mean traffic data for the sites, by direction, was addressed through regression analysis. The best fit equation was determined for mean values by site and direction. The statistical analysis system (SAS) procedure used was PROC STEPWISE MAX R.

In the stepwise procedure the use of the two independent variables provided a better fit than single variable. The *R*-squared value for the best one-variable model, with total volume the independent variable, was .5845. The *R*-squared value for the best fit two-variable model, total volume and percent trucks of total volume, was .6087.

The two-variable model was selected despite relatively low improvement in the fit of the model. On some segments of the rural interstate, percent trucks can affect the resulting LDF equation. When the percent trucks of total volume exceeds 16%, this variable affects the LDF. At 32% this variable reduces the LDF by 1%. Some New Mexico rural interstate segments currently exceed 38% trucks of total volume. This variable is also intuitively sensible. As the percent trucks of the volume increases, the trucks with higher weight-to-horsepower ratio will on occasion be forced into the left lane.

The two-variable model, or base equation, for the study, is shown below. Figure 1 is a plot of equation 1 with the assumption of 26% truck traffic.

$$\text{LDF} = .98144 - .0004(\text{Volume}) - .000293(\text{Percent Trucks}) \quad (1)$$

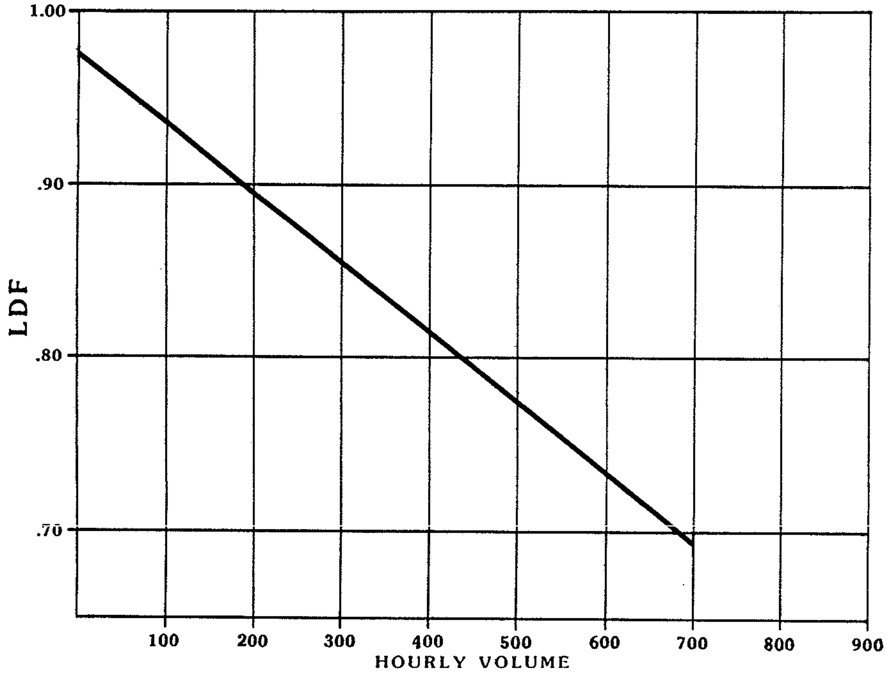

FIGURE 1 Equation 1 (assume 26% trucks).

This equation is only for tangent segments of the rural interstate. Because the data used for model development cover a limited range of volumes, it is important to note the traffic volumes at which the model does not provide useful information. The equation fails at volumes greater than 1,230 per hour, per direction. At this point the equation provides less than half the vehicles in the right lane. No more than a 50/50 split, without other access/egress factors, would occur, even at level of service E. The volume or heavy truck percent at which the rural interstate LDF reduces to equilibrium, a 50/50 split, is not known and cannot be accurately extrapolated from the data collected for this study.

What is the volume at which the LDF equation should not be applied? The equation should not be used with volumes greater than 700 vehicles per hour, per direction. At this level there is an LDF of 70%—the lowest observed right lane percent of heavy commercial trucks on the rural interstate. Extrapolation beyond the observed percent is not recommended because it could result in underbuilding a section of rural interstate.

While the primary purpose of the study is to check the assumption of 94/6 distribution between the right and left lanes and perhaps to prevent overbuilding of a facility, underbuilding is also of concern. This concern not to underbuild a section of interstate directed the study to modifications of equation 1.

The results of the study demonstrate that current distribution assigns too much of the heavy commercial traffic to the right lane. At no location surveyed was the LDF as high as the standard LDF value previously employed by the New Mexico State Highway Department. This suggests that the interstate construction cost could be reduced through a corrected load analysis resulting from more accurate LDF.

Equation 1, based around a line fit to the mean data by site and direction, reduces the possibility of overbuilding through error in the LDF. To prevent underbuilding through LDF error, the LDF must not be underestimated. This concern is the basis for modification to the equation.

RANGE OF ACCEPTABLE ERROR

To use the equation would essentially "miss" the mean as high as to miss it low because of the sum of squares step in fitting the regression line to the data. The equation should be modified so it is fit not only to the mean, but also to the range of desirable values. This range will be characterized by a smaller margin of acceptable error below the mean, and a larger margin of acceptable error above the mean.

Two ranges of error were identified in discussions with personnel from the Materials Laboratory and Technical Services Engineering. These can be used to compare and evaluate the results of equation 1 with field observation data. A standard for the result of the equation is an LDF equal to or greater than −1 percentage point of the observed value, and equal to or less than +3 percentage points of the observed value.

A less precise, but acceptable, standard would be the equation LDF providing dependent variable values equal to or greater than −2 percentage points of the observed value, and equal to or less than +5 percentage points of the observed value. The ranges of values identified reflect the concern to overestimate rather than underestimate the LDF.

At the lower precision level, equation 1 results in eight of twelve sites being defined within the range. However, at the higher precision level, equation 1 drops to four of twelve sites within the range. Equation 1 results in low mean and cumulative errors, of −.00907 and −.10885 respectively. However, both errors are minus, more typically below the mean field observations.

The model developed to this point does not satisfactorily interpret all of the stations. Equation 1 produces acceptable results for study sites with traffic volumes between 400 and 450 vph.

VOLUME GROUPS

There are three groups of total volumes. Given the desirable range around the mean, the equation 1 intercept parameter operates in a linear manner with the coefficients of the independent variables only within one range of total traffic volume. Volumes less than 400 and greater than 450 vph are modeled to the desirable range. This was accomplished by adding half a standard deviation to the intercept for the low volume group, and one standard deviation for the high volume group. Adding the standard deviation to the equation increases the LDF. The equation is expressed with the standard deviation added to the intercept parameter. The intercept parameter is, therefore, modified to model the characteristics of the three volume groups. The volume groups, and the modified parameters, are noted below by range.

Equation	*Volume Range (vph)*	*Equation Number*
LDF = 1.02144 − .0004(Volume) − .000293(T)	405–700	2
LDF = .98144 − .0004(Volume) − .000293(T)	400–449	1
LDF = 1.00144 − .0004(Volume) − .000293(T)	10–399	3

With these equations, the projected LDF would be within the more precise range five of twenty sites, and twelve of twelve within range at the less precise range. For the less precise range, six sites were overestimated but in range, four sites were underestimated but in range, and there was no error at two sites.

Equation 1 when applied to all traffic volumes performed satisfactorily to 700 vph. The use of three equations for the LDF by volume range applies to 800 vph.

FURTHER ADJUSTMENT FOR THE HIGHER VOLUME RANGE

A further adjustment may be made to ensure that there is no underestimation of the LDF. No adjustment is needed for volume ranges 10–399 and 400–449. Underestimation at higher volumes suggests that the intercept adjustment for this volume is not sufficient. An additional .03 added to this intercept would provide, at higher volume sites, LDFs that are overestimated and in range, rather than underestimated and in range. This modification will ensure that the LDF is not understated, and that the design and cost gains through the study are only slightly diminished. The effect of this additional modification is to increase the equation by slightly under two standard deviations. The precision range result is similar to that of a confidence interval modification of the base equation.

LDF EQUATIONS

The final equations by volume on the rural interstate are noted below.

Equation	*Volume Range (vph)*	*Equation Number*
LDF = 1.05144 − .0004(Volume) − .000293(T)	405–700	4
LDF = .98144 − .0004(Volume) − .000293(T)	400–449	1
LDF = 1.00144 − .0004(Volume) − .000293(T)	10–399	3

The above equations were evaluated by testing for homogeneity of variance. PROC TABLES from the SAS Supple-

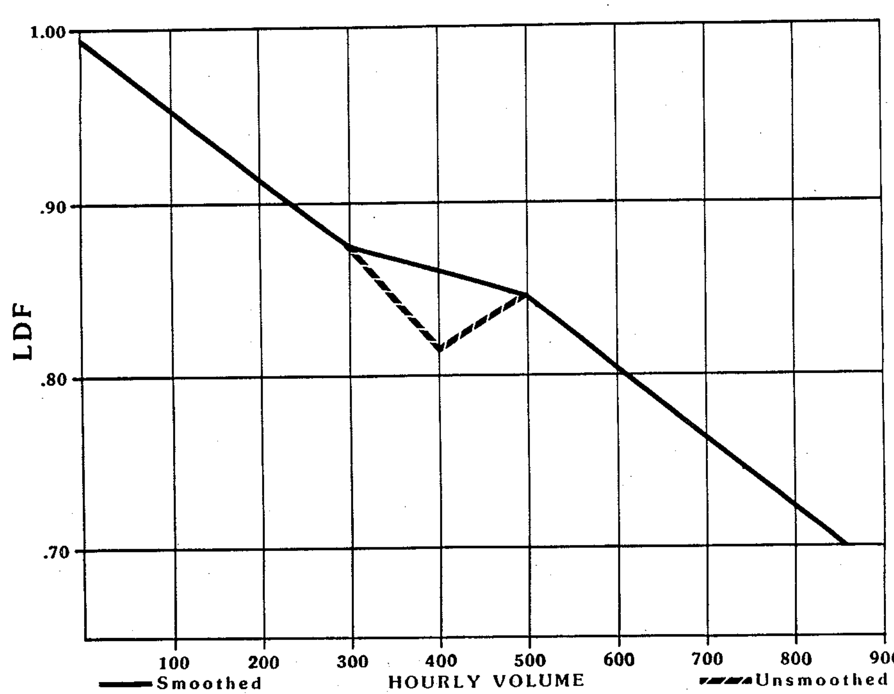

FIGURE 2 Adjusted equation (assume 26% trucks).

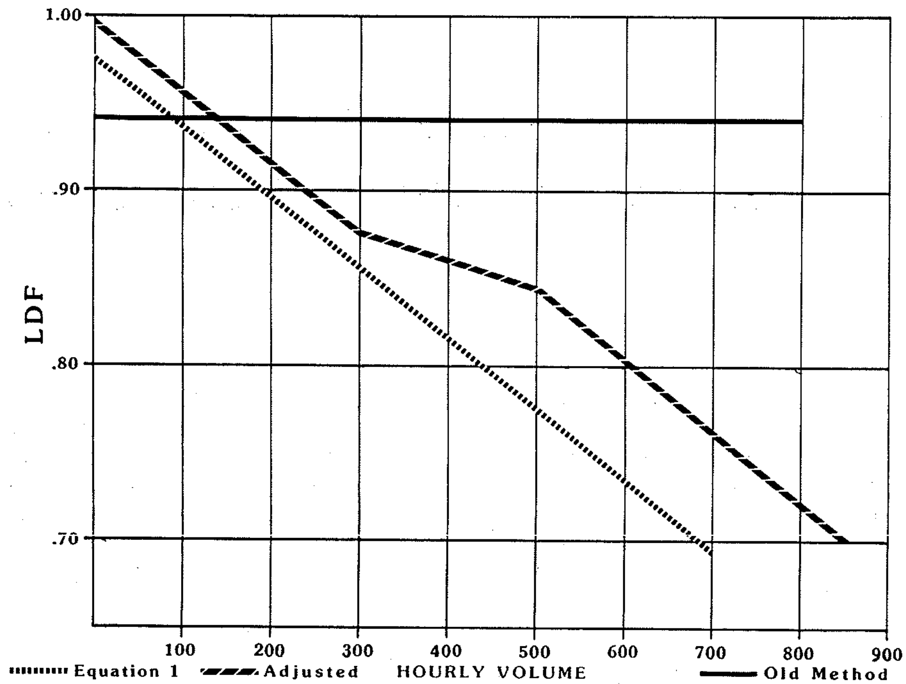

FIGURE 3 Comparison of methods (assume 26% trucks).

mental Library was used to test the homogeneity of variance in volume and percent trucks to volume. Using Bartlett's procedure at $p \leq 0.001$, the variance was found to be homogeneous. For purposes of comparison with the above equations, despite acceptance of homogeneity, weighted regression analysis was conducted using SAS PROC REG. The resulting high and low volume group equations derived lower LDFs than calculated by the above equations. If the concern is not to underbuild a facility, the equations would not be acceptable.

To present the resulting equations for daily departmental use, the LDF relationship to volume and trucks is presented in graphic form. The equations line is smoothed at the middle volume range. In the process of smoothing the regression equation line, the resulting LDF for the middle volume range is also higher than that calculated from the weighted regression equation. Figure 2 illustrates the use of the three equations with volume group 2 smoothed.

CONCLUSION

The lane distribution factors put forth in the study represent a significant improvement in accuracy over earlier methods. Figure 3 compares the earlier method with the results of the present study, for the case of 26% truck traffic. The models can be applied easily because the data required are available on a system-wide basis and because the LDFs can be represented in graphic form. The models reflect the differences in volume and heavy commercial percent of total volume that are characteristic of New Mexico's rural interstate system.

Using the lane distribution factors will result in better road design and may provide considerable savings. A sensitivity analysis of the impact of this procedure on pavement structural characteristics was conducted by the Materials Testing Laboratory and Planning Bureau of the New Mexico State Highway Department. The analysis determined that for longer-term pavement design the lane distribution factors reduce pavement thickness by an average of ¾ in. For all of the sites in the present study, the pavement design was reduced between ¾ and 1 in. pavement thickness. It may be concluded that the lane distribution factors derived from this study improve the accuracy of determining truck lane usage, and that this improvement significantly affects pavement design.

Additional research will be undertaken, including analyzing the effect of positive grades on truck lane use, sampling for monthly variation, and examining the time stability of LDF equations.

REFERENCE

1. M. M. Alexander, Jr., and R. A. Graves II. *Determination of the Lanal Distribution of Truck Traffic on Freeway Facilities.* Georgia Highway Department, Atlanta, Nov. 1971.

Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.