Statistical Evaluation of the Effects of Highway Geometric Design on Truck Accident Involvements

SHAW-PIN MIAOU AND HARRY LUM

Illustrated in this paper are ways in which the Poisson regression model can be used to evaluate the effects of highway geometric design on truck accident involvement rates and estimate and quantify the uncertainties of the expected reductions in truck accident involvements from various improvements in highway geometric design. The data source used in this study was the Highway Safety Information System, a highway safety data base administered by the Federal Highway Administration. Among the five Highway Safety Information states currently available for analysis, Utah was considered to be the state that had the most complete information on highway geometric design and was selected for illustration. Five years of highway geometric, traffic, and truck accident data for rural Interstate highways from 1985 to 1989 were used.

The effects of roadway characteristics on traffic safety are substantial, according to the nation's highway safety performance records (*1*). For example, in 1988 the fatality rates on rural Interstate, other rural Federal-aid primary arterial, and rural non-Federal-aid arterial are, respectively, 9.7, 21.7, and 50.9 fatalities/billion vehicle km (1.56, 3.48, and 8.20 fatalities/100 million vehicle mi). The records also suggest that if all urban and rural travel were at the same fatality rate as the corresponding Interstate rate, then fatalities would be 23,491 instead of 47,093 in 1988, a reduction of over 50 percent (*2*). Potential factors that make vehicle accident rate different from one roadway class to another include the physical nature of the roadway, such as geometric design, roadway markings, and traffic signs, and the type of incurred travel, traffic control, and traffic conditions.

Highway geometric design elements, such as horizontal curvature, vertical grade, lane width, shoulder width, and median, are logical engineering factors that contribute to the differences in vehicle accident rate among roadway classes (*3*). Their effects on vehicle accidents are, however, difficult to quantify because of large confounding influences from the human factor, the environment (including lighting and weather conditions), traffic, and vehicles. Previous studies suggested that roads were rarely the sole factors associated with a traffic accident—only about 2 percent according to Rumar (*4*). It was mainly through the interactions with other factors, especially human and environmental, that roads were associated with traffic accidents.

Ideally, to investigate the effects of highway geometric design on vehicle accidents, roadway, traffic, accident, environmental, road user, vehicle, and exposure data for individual road sections are needed. In practice, many of these factors are qualitative in nature, especially human factors, and are not likely to be available for individual road sections in any accident data base. In addition, detailed vehicle exposure data (e.g., by vehicle type, time of day, weather, and vehicle speed) may not be available for individual road sections. This means that many factors that may have influence on the occurrence of vehicle accidents would not be available for study of the geometric design effects.

In view of this inevitable omitted variable problem, when any geometric design effect is discussed, we have in mind the average observed effect, which includes the collective influence of all the interacting effects. This includes the influence of interacting factors such as the driver's physical condition, driving skill, mood, and knowledge; vehicle speed; weather; and so on. Thus, the geometric design effects are estimated to be conditional on the omitted variables. That is, the effect of the same highway geometric design on vehicle accidents would be different if some of the omitted variables change over time. For example, changes in socioeconomic, legislative, and law-enforcement conditions over the years would change the driver's behavior and, therefore, would change the geometric design effects on vehicle accidents even if nothing is done to the road. For this reason, the analysts should always be careful in interpreting the estimated effects, be conscious of any potential bias, and be cautious in using the effects derived from one area for other areas.

To give another example, consider two hypothetical road sections of the same roadway class, say $C10$ and $C1$, the geometry of which are different only in horizontal curvature: $C10$ is a 10-degree curve (per 30.48-m or 100-ft arc) and $C1$ is a 1-degree curve. The distribution of vehicle speed on $C10$ is expected to be different from that on $C1$, and the average vehicle speed on $C10$ is expected to be less than that on $C1$. Given that the vehicle speed distribution on these two curves is not known, estimated curvature effects on vehicle accidents for $C10$ and $C1$ will be the effects averaged over their respective vehicle speed distribution and, therefore, conditional on their vehicle speed distribution. If the underlying vehicle speed distribution of any curve changes because of speed limit change, for example, then its average curvature effect is likely to change too.

S.-P. Miaou, Center for Transportation Analysis, Energy Division, Oak Ridge National Laboratory, P.O. Box 2008, MS 6366, Building 5500A, Oak Ridge, Tenn. 37831. H. Lum, Office of Safety and Traffic Operations Research and Development, Federal Highway Administration, U.S. Department of Transportation, Turner-Fairbank Highway Research Center, 6300 Georgetown Pike, McLean, Va. 22101-2296.

Many statistical models have been developed to establish the empirical relationships between vehicle accidents and highway geometric design for different roadway classes, vehicle configurations, and accident severity types (*3*). However, most of these models were developed on the basis of the conventional multiple linear regression approach, and have been shown to be lacking the distributional property to adequately describe discrete, nonnegative, and typically sporadic vehicle accident events on the road (*5–7*). These unsatisfactory properties of the linear regression models have led to the investigation of the Poisson regression and negative binomial regression models in recent studies (*6–9*). In general, most of these studies found the Poisson regression model to be appropriate for studying the relationships among vehicle accidents and the contributing factors under their study. In addition, despite the limitations in existing highway geometric data, some encouraging relationships have been developed for horizontal curvature, vertical grade, and shoulder width using the Poisson regression model.

The objective of this paper is to illustrate how the Poisson regression model can be used to evaluate the effects of highway geometric design elements on truck accident involvement rates. Also, described in this paper is the way in which the model can be used to estimate and quantify the uncertainties of the expected reductions in truck accident involvements from various improvements in highway geometric design. The data source used in this study was the Highway Safety Information System (HSIS), a highway safety data base administered by the Federal Highway Administration (FHWA) (*10*).

POISSON REGRESSION MODEL

The Poisson regression model employed in this paper was proposed by Miaou et al. (*8*) to develop the relationship between vehicle accidents and highway geometric design. In theory, the model can be applied to any roadway class, vehicle configuration, and accident severity type of interest. The following presentation focuses on accidents of all severity types involving large trucks (more than 4,545 kg or 10,000 lb) on a particular roadway class.

The Model

Consider a set of n road sections of a particular roadway type; for example, a rural Interstate. Let Y_i be a random variable representing the number of trucks involved in accidents on road section i during a period of 1 year, where $i = 1, 2, \ldots, n$. Here the same road section in different sample periods can be considered as separate road sections. This allows the year-to-year changes in geometric design and traffic conditions to be considered in the model. Further, the actual observation of Y_i during the period is denoted as y_i, where $y_i = 0, 1, 2, 3, \ldots$ and $i = 1, 2, \ldots, n$. The amount of truck travel (or truck exposure) during the sample year on road section i, denoted by v_i, is computed as

$$365 \times \text{AADT}_i \times (T\%/100) \times l_i$$

where

AADT_i = average annual daily traffic (in number of vehicles),
$T\%_i$ = average percentage of trucks (or percent trucks) in the traffic stream (e.g., 15), and
l_i = length (in km or mi) of road section i.

Note that $\text{AADT}_i \times (T\%_i/100)$ represents the "truck AADT" of road section i during the year. Associated with each road section i, there is a $k \times 1$ covariate vector, x_i, describing its geometric design characteristics, traffic conditions, and other relevant attributes. The transpose of the covariate vector is denoted by $x_i' = (x_{i1}, x_{i2}, \ldots, x_{ik})$. Without loss of generality, let the first covariate x_{i1} be a dummy variable equal to one for all i (i.e., $x_{i1} = 1$).

To the extent possible, these n road sections should be selected to cover as much variation in geometric design, traffic conditions, and other relevant attributes as possible. In addition, to avoid the bias in estimating the truck accident-geometric design relationship, the selection of road sections should not be based on the outcomes of the dependent variable (i.e., y_is).

Under the assumption that (a) truck exposure data and other covariates are free from errors, (b) the occurrences of truck accidents on different road sections are independent, and (c) the number of trucks involved in accidents on a particular road section i, Y_i, follows a Poisson distribution, Miaou et al. (*8*) proposed the following model to establish the relationship between truck accidents and highway geometric design:

$$p(Y_i = y_i) = p(y_i) = \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!},$$
$$y_i = 0, 1, 2, 3, \ldots, i = 1, 2, 3, \ldots, n, \tag{1}$$

where

$$\mu_i = \text{E}(Y_i) = v_i[e^{\underline{x}_i'\underline{\beta}}]$$
$$= v_i[e^{\sum_{j=1}^{k} x_{ij}\beta_j}] = v_i\lambda_i \quad i = 1, 2, 3, \ldots, n. \tag{2}$$

and $\underline{\beta}$ is a $k \times 1$ vector of unknown regression coefficients to be estimated from the data, the transpose of which is denoted by $\underline{\beta}' = (\beta_1, \beta_2, \ldots, \beta_k)$. This model assumes that the number of trucks involved in accidents $Y, i = 1, 2, \ldots, n$, are independently and Poisson distributed with mean μ_i, and the mean μ_i (i.e., the expected number of trucks involved in accidents) is proportional to truck travel v_i. The model also assumes an exponential rate function, $\lambda_i = \text{E}(Y_i)/v_i = \exp(\underline{x}_i'\underline{\beta})$, which ensures that accident involvement rate is always nonnegative. This type of rate function has been widely employed in statistical literature and found to be very flexible in fitting different types of count data (*11,12*). Note that whenever appropriate, higher order and interaction terms of covariates can be included in Equation 2 without difficulty.

On the basis of the model, the variance, $Var(Y_i)$, and coefficient of skewness, $skew(Y_i)$, of the underlying distribution of Y_i are μ_i and $\mu_i^{-1/2}$, respectively. The variance $Var(Y_i)$, which is equal to the mean μ_i, depends on its rate function and thus involves unknown regression coefficients. In addition, $Var(Y_i)$ grows linearly with truck exposure v_i. The model supposes a positive skewness coefficient that varies from road section to road section, depending on their means [$skew(Y_i) = \mu_i^{-1/2}$]. As mean μ_i increases, either as a result of an increase in vehicle exposure v_i or an increase in the rate function λ_i, the skewness coefficient $skew(Y_i)$ decreases, and as μ_i decreases, $skew(Y_i)$ increases.

Estimation and Statistical Inference

In this paper, the regression coefficients of the Poisson regression model are estimated using the maximum likelihood method. The maximum likelihood estimates (MLE) of the regression coefficients are obtained by maximizing the following loglikelihood function:

$$L(\underline{\beta}) = \log\left(\prod_{i=1}^{n} p(y_i)\right) = \log\left(\prod_{i=1}^{n} \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!}\right)$$

$$= \sum_{i=1}^{n} [y_i \log(\mu_i) - \mu_i - \log(y_i!)]$$

$$= \sum_{i=1}^{n} [y_i(\underline{x}_i'\underline{\beta}) + y_i \log(v_i) - v_i e^{\underline{x}_i'\underline{\beta}} - \log(y_i!)] \quad (3)$$

The first derivative of the loglikelihood function with respect to the jth regression coefficient can be shown to be

$$\frac{\partial L(\underline{\beta})}{\partial \beta_j} = \sum_{i=1}^{n} \{y_i - v_i e^{\underline{x}_i'\underline{\beta}}\} x_{ij} \quad (4)$$

where $j = 1, 2, \ldots, k$ and must all vanish at the MLE $\hat{\underline{\beta}}$. Because the first covariate x_{i1} is a dummy variable equal to 1 for all i, the MLE requires that $\Sigma_i y_i = \Sigma_i v_i \exp(\underline{x}_i'\hat{\underline{\beta}})$. That is, the (estimated) expected total number of accident involvements, $\Sigma_i \hat{\mu}_i$, has to be equal to the observed total $\Sigma_i y_i$, where $\hat{\mu}_i = v_i \exp(\underline{x}_i'\hat{\underline{\beta}})$. This is a desirable statistical property in modeling vehicle accidents (*6*). Note that most of the suggested conventional multiple linear regression models for establishing geometric design-vehicle accident relationships do not have such a property (*6*).

The asymptotic covariance and t-statistics of the estimated coefficients, as the sample size n becomes infinite, can be determined using the second derivative of the loglikelihood function (i.e., Fisher's information matrix), as follows. The second derivative, or the Hessian matrix, of the loglikelihood function can be derived as

$$h_{jq} = \frac{\partial^2 L(\underline{\beta})}{\partial \beta_j \partial \beta_q} = -\sum_i (v_i e^{\underline{x}_i'\underline{\beta}}) x_{ij} x_{iq}$$

$$j = 1, 2, \ldots, k, \; q = 1, 2, \ldots, k \quad (5)$$

which is a function of unknown regression coefficient $\underline{\beta}$, and does not involve dependent variable y_i. Provided the Poisson assumption is adequate and the sample size is reasonably large, the asymptotic covariance matrix of the MLE can be obtained as

$$cov(\hat{\underline{\beta}}) = [I(\hat{\underline{\beta}})]^{-1} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ s_{21} & s_{22} & \cdots & s_{2k} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ s_{k1} & s_{k2} & \cdots & s_{kk} \end{bmatrix} \quad (6)$$

where

$$I(\hat{\underline{\beta}}) = -\left.\frac{\partial^2 L(\underline{\beta})}{\partial \underline{\beta} \partial \underline{\beta}'}\right|_{\underline{\beta}=\hat{\underline{\beta}}} = -\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ h_{k1} & h_{k2} & \cdots & h_{kk} \end{bmatrix}_{\underline{\beta}=\hat{\underline{\beta}}} \quad (7)$$

is the Fisher information matrix evaluated at the MLE $\hat{\underline{\beta}}$ (*13*). The asymptotic t-statistic for each estimated regression coefficient $\hat{\beta}_j$ is computed as $\hat{\beta}_j/(s_{jj})^{1/2}$, where $(s_{jj})^{1/2}$ is the asymptotic standard deviation of $\hat{\beta}_j$ and $j = 1, 2, \ldots, k$, and its significance level can be assessed using a t distribution table with n-k degrees of freedom or simply using a normal probability table because of large n. The asymptotic correlation matrix of the estimated regression coefficients can be constructed as $\hat{\rho}_{ij} = s_{ij}/(s_{ii}s_{jj})^{1/2}$, for $i = 1, 2, \ldots, k$, and $j = 1, 2, \ldots, k$. (Note that $\hat{\rho}_{ii} = 1$ for $i = 1, 2, \ldots, k$.)

A limitation of using the Poisson regression model, which is well known in the statistical literature (*14,15*), is that the variance of the data is restrained to be equal to the mean. In many applications, count data were found to display extra variation or overdispersion relative to a Poisson model (*15*). That is, the variance of the data was greater than the Poisson model indicated.

If the overdispersion exists in the data, the MLE of the regression coefficients, $\hat{\underline{\beta}}$, under the Poisson regression model, will still be close to the true coefficients, $\underline{\beta}$, when the sample size n is large. (This is assuming that the rate function in Equation 2 has the correct form.) However, under the Poisson regression model, the variances of the estimated coefficients (i.e., s_{jj}, $j = 1, 2, \ldots, k$) would tend to be underestimated and, therefore, the associated t-statistics $\hat{\beta}_j/(s_{jj})^{1/2}$, $j = 1, 2, \ldots, k$, would tend to be overestimated (*16*). Following Wedderburn (*17*), to correct for the overdispersion problem for the Poisson regression model, it can be assumed that the variance of Y_i is $\tau\mu_i$ instead of μ_i, as originally assumed in the Poisson model, where τ is called the overdispersion parameter (typically, $\tau \geq 1$). Furthermore, a moment estimator of the overdispersion parameter τ is $\hat{\tau} = X^2/(n - k)$, where X^2 is the Pearson's chi-square statistic, n is the number of observations (i.e., the number of road sections in this case), and k is the number of unknown regression coefficients in the Poisson regression model. The Pearson's X^2 statistic is computed as $\Sigma_i(y_i - \hat{\mu}_i)^2/\hat{\mu}_i$. A better estimate of the asymptotic covariance matrix of the estimated coefficients is $\hat{\tau} \times cov(\hat{\underline{\beta}})$ and, therefore, a better estimate of the t-statistic for regression coefficient $\hat{\beta}_j$ is $\hat{\beta}_j/(\hat{\tau}\, s_{jj})^{1/2}$, $j = 1, 2, \ldots, k$ [see, e.g., Agresti (*18*)].

Model Applications

To illustrate how the Poisson regression model can be used to estimate the expected reduction in truck accident involvements caused by improvements in some geometric design elements, consider a particular road section i, and let the value of its covariates before and after the improvement be x_{ij}^b and x_{ij}^a, for $j = 1, 2, \ldots, k$. Also, let v_i^b and v_i^a be the amount of truck travel in one year on road section i before and after the improvement.

Based on the Poisson regression model (Equations 1 and 2), the expected number of truck accident involvements on road section i before and after the improvements of geometric design elements are, respectively, $v_i^b \exp(\Sigma_j x_{ij}^b \beta_j)$ and $v_i^a \exp(\Sigma_j x_{ij}^a \beta_j)$. The percentage reduction in the expected truck accident involvements can be computed as

$$R_i = \left[\frac{v_i^b \exp\left(\sum_{j=1}^{k} x_{ij}^b \beta_j\right) - v_i^a \exp\left(\sum_{j=1}^{k} x_{ij}^a \beta_j\right)}{v_i^b \exp\left(\sum_{j=1}^{k} x_{ij}^b \beta_j\right)} \right] \times 100$$

$$= \left\{ 1 - \left(\frac{v_i^a}{v_i^b}\right) \exp\left[\sum_{j=1}^{k} (x_{ij}^a - x_{ij}^b)\beta_j \right] \right\} \times 100 \tag{8}$$

The percentage reduction R_i is sometimes referred to as the truck accident involvement reduction factor. If v_i is the same before and after the improvement (i.e., $v_i^b = v_i^a$) then R_i also represents the percentage reduction in truck accident involvement rate. By substituting β_j with the MLE $\hat{\beta}_j$ in Equation 8 for $j = 1, 2, \ldots, k$, a MLE of the reduction in the expected number of truck accident involvements can be obtained, denoted by $\hat{R}_i$. Because, for a large sample, $\underline{\hat{\beta}}$ is approximately normally distributed with mean $\underline{\beta}$ and with covariance matrix $\hat{\tau} \times cov(\underline{\hat{\beta}})$ [see, for example, Agresti (*18*)], it can be shown that the standard deviation ($s.d.$) of $\hat{R}_i$ is approximately as follows:

$$s.d.(\hat{R}_i) \approx \left(\frac{v_i^a}{v_i^b}\right) \times \left\{ \exp\left[\sum_{j=1}^{k} (x_{ij}^a - x_{ij}^b)\hat{\beta}_j + \frac{\hat{\tau}}{2} \sum_{m=1}^{k} \sum_{q=1}^{k} \right.\right.$$
$$\left.\left. \times (x_{im}^a - x_{im}^b)(x_{iq}^a - x_{iq}^b)\hat{\rho}_{mq}(s_{mm}s_{qq})^{1/2} \right] \right\}$$
$$\times \left\{ \exp\left[\hat{\tau} \sum_{m=1}^{k} \sum_{q=1}^{k} (x_{im}^a - x_{im}^b) \right.\right.$$
$$\left.\left. \times (x_{iq}^a - x_{iq}^b)\hat{\rho}_{mq}(s_{mm}s_{qq})^{1/2} \right] - 1 \right\}^{1/2} \times 100 \tag{9}$$

The derivation uses the property that if z is normally distributed with mean μ and variance σ^2, then the variance of $\exp(z)$ is $\{\exp[\mu + (1/2)\sigma^2]\}^2 [\exp(\sigma^2) - 1]$ [see Lindgren, (*19*), page 191]. Equation 9 allows the uncertainty of the estimated reduction to be assessed by quoting plus or minus one standard deviation.

ILLUSTRATION AND DISCUSSION

Data Source

To illustrate the use of the Poisson regression model, data from the HSIS were employed to develop relationships between truck accidents and key highway geometric design variables. The HSIS currently has data from five states. A general description of the HSIS data base is given in Council and Paniati (*10*). Specifically, accidents involving trucks (of more than 4,545 kg or 10,000 lb) on rural Interstate highways from Utah were used. Among the five HSIS States, Utah was considered to be the state that had the most complete information on highway geometric design. In addition, Utah was the only HSIS state with a historical road inventory file in which year-to-year changes on highway geometric design element and traffic conditions were recorded. Thus, accidents in a given year could be matched to the road inventory information of the same period. Data from 1985 to 1989 were used for the illustration.

Utah data in HSIS were stored in six files: roadlog, horizontal curvature, vertical grade, accident, vehicle, and occupant files. Thus, these files had to be linked before any analysis could be performed. Each record in the roadlog file represented a homogeneous section in terms of its cross-sectional characteristics, such as number of lanes, lane width, shoulder width, median type and width, AADT, and percent trucks. However, these road sections were not necessarily homogeneous in terms of their horizontal curvatures and vertical grades. Road sections in the horizontal curvature and vertical grade files, on the other hand, were delineated in such a way that they were homogeneous in terms of their horizontal curvatures and vertical grades, respectively, but not necessarily in terms of other road characteristics.

Therefore, after matching road sections in the horizontal curvature and vertical grade files with the road sections in the roadlog file, each road section in the road inventory file may have contained more than one horizontal curvature or vertical grade. In this illustration, those road sections with multiple curvatures and grades were further disaggregated into smaller subsections so that each subsection contained a unique set of horizontal curvature and vertical grade. Each subsection, which was totally homogeneous in cross-sectional characteristics, horizontal curvature, and vertical grade, was then considered as an independent road section in the model. In order to test the effects of the length of curve and grade, information on the length of the original curve and grade, from which the subsection was delineated, was maintained for each subsection.

Accidents, Characteristics of Road Sections, and Covariates

The time period considered was 1 year, which means that the same road section, even if nothing had changed, was considered as five independent sections—one for each year from 1985 to 1989. As indicated earlier, this allowed the year-to-year changes on highway geometric design and traffic conditions to be considered in the model. A total of 8,263 homogeneous road sections during the 5-year period were considered to have reliable data. These road sections covered about 99 percent of the entire rural Interstate highway mileage in Utah and constituted 23,570 lane-km or 14,731 lane-mi of roadway. Data for each year contained roughly ⅕ of the total sections and lane-km. The section lengths varied from 0.016 to 12.43 km (0.01 to 7.77 mi)—with an average of 0.72 km (0.45 mi). Descriptive statistics of these 8,263 road sections on truck accident involvements and truck miles (km) traveled are given in Table 1.

TABLE 1 Variable Definitions and Summary Statistics of the 8,263 Rural Interstate Road Sections

Variable	Notation & Definition (for section *i*)	Min	Max	Mean	% Zero
Number of Trucks Involved in Accidents	y_i	0	8	0.20	86
Section Length (in mi)	ℓ_i	0.01	7.77	0.45	0
Truck Miles or Truck Exposure (in 10^6 truck-miles)	v_i = $[365 \times AADT_i \times (T\%_i/100) \times \ell_i]/10^6$, where $T\%_i$ is percent trucks (366 for leap years).	8×10^{-4}	5.03	0.25	0
Dummy Intercept	$x_{i1} = 1$				
Dummy Variable for Year 1986, representing year-to-year changes due to random fluctuations, annual trend, and omitted variables such as weather.	$x_{i2} = 1$, if the road section is in year 1986 $= 0$, otherwise				
Dummy Variable for Year 1987 (See above explanation)	$x_{i3} = 1$, if the section is in 1987 $= 0$, otherwise				
Dummy Variable for Year 1988 (See above explanation)	$x_{i4} = 1$, if the section is in 1988 $= 0$, otherwise				
Dummy Variable for Year 1989 (See above explanation)	$x_{i5} = 1$, if the section is in 1989 $= 0$, otherwise				
AADT per Lane (in 1000's of vehicles), a surrogate variable to indicate traffic conditions or traffic density.	$x_{i6} = (AADT_i/\text{number of lanes}_i)/1000$	0.35	12.04	1.80	0
Horizontal Curvature, HC, (in degrees per 100-ft arc)	x_{i7}	0	12.00	1.00	67
Length of Original Horizontal Curve, LHC, (in mi) from which this curve was subdivided for creating homogeneous sections; only for HC > 1 and LHC ≤ 1.	x_{i8} = LHC, if $x_{i7}>1$ and LHC ≤ 1 mi. = 1.0, if $x_{i7}>1$ and LHC > 1 mi. = 0, if $x_{i7} \leq 1$	0	0.96	0.05	81
Vertical Grade, VG, (in percent)	x_{i9}	0	8.00	2.14	20
Length of Original Vertical Grade, LVG, (in mi) from which this section was subdivided for creating homogeneous sections; only for sections with VG > 2 and LVG ≤ 2.	$x_{i,10}$ = LVG, if $x_{i9}>2$ and LVG ≤ 2 mi. = 2.0, if $x_{i9}>2$ and LVG > 2 mi. = 0, if $x_{i9} \leq 2$	0	2.00	0.21	74
Deviation of Paved Inside Shoulder Width (per direction) from an "ideal" width of 12 ft (3.66 m).	$x_{i,11}$ = max{0, 12 - paved inside shoulder width}	4.00	12.00	8.16	0
Percent Trucks in the traffic stream (e.g., 15)	$x_{i,12}$	7.00	57.00	24.13	0
HC × LHC	$x_{i,13} = x_{i7} \times x_{i8}$	0	2.88	0.18	81
VG × LVG	$x_{i,14} = x_{i9} \times x_{i,10}$	0	13.37	0.97	74

(1 mi = 1.61 km; 1 ft = 0.3048 m)

During the 5-year period, 1,643 large trucks were reported to be involved in accidents on these highway sections, regardless of truck configuration and accident severity type. With the total truck travel estimated to be 3,248 million truck km (MTK) or 2,030 million truck mi (MTM), the overall truck accident involvement rate was therefore 0.51 truck accident involvements/MTK or 0.81 truck accident involvements/MTM. These accidents occurred on only 14 percent of the 8,263 road sections. The maximum number of trucks involved in accidents on an individual road section in one year was 8. On average, each section had 0.20 trucks involved in accidents in 1 year.

The covariates considered for individual road sections and their definitions are also presented in Table 1. They include

1. Yearly dummy variables to capture year-to-year changes in the overall truck accident involvement rate caused by, for example, long-term trend, annual random fluctuations, changes in posted speed limit, and changes in omitted variables such as weather;
2. AADT/lane, used as a surrogate measure for traffic flow density;
3. Horizontal curvature (HC);
4. Vertical grade (VG); and
5. Deviation of paved inside (or left) shoulder width/direction from an "ideal" width of 3.66 m (12 ft).

Note that paved inside shoulder width (ISH) of 3.66 m (12 ft)/direction is recommended by the "Greenbook" for roads with heavy truck traffic (*20*). Because all of the road sections were 3.66 m (12 ft) in lane width, more than 89 percent of them had 4 lanes, and all road sections had paved outside (or right) shoulder widths of 3.05 m (10 ft), the effects of these variables could not be determined in this study.

It has been suggested that as the length of grade increases to a point that can slow a truck to a speed significantly slower than the speed of the traffic stream (e.g., 16 km/hr or 10 mi/hr), the accident rate increases (*3*). Also, for a fixed curvature degree, as the length of curve increases, the accident rate increases (*21*). To test the effects of length of curve and length of grade on truck accident involvement rate, two covariates—length of original curve (LHC), $x_{i,8}$, and length of original grade (LVG), $x_{i,10}$,—were considered. As indicated earlier, each curve or grade considered in the model may have been subdivided from a longer curve or grade for achieving total homogeneity. Thus, for each road section in the model, these two covariates were defined as the length of the original undivided curve or undivided grade to which this section belonged. In addition, these two covariates were defined only for curves with horizontal curvatures greater than 1 degree per 30.48-m (100-ft) arc and sections with grade greater than 2 percent. (Note that these two covariates were set equal to 0 if horizontal curvature is less than or equal to 1 degree or if vertical grade is less than or equal to 2 percent.) This definition was based on an assumption that the length of a mild curve or grade has no aggravated effect on truck accident involvement rate. On the basis of engineering judgments, it was further assumed that there were no additional effects after LHC reached 1.6 km (1.0 mi) or after LVG reached 3.2 km (2.0 mi). This assumption makes the effects of LHC and LVG on truck accident involvement rate more robust to unusually long curves and grades. The interactions of HC and LHC ($x_{i,13} = x_{i,7} \times x_{i,8}$), VG and LVG ($x_{i,14} = x_{i,9} \times x_{i,10}$), and HC and VG ($x_{i,7} \times x_{i,9}$) were also considered.

Percent trucks in the traffic stream was included in the model to evaluate the effects of automobile-truck mix. Previous studies suggested that as percent trucks increases, truck accident involvement rate decreases. One possible reason is that, for a constant vehicle density, as percent trucks increases, the frequency of lane changing and overtaking movements by automobiles decreases. Also, previous records showed that more trucks were involved in truck-automobile multivehicle accidents than in truck-truck accidents [e.g., see Jovanis and Chang (*5*)].

Model Results

The estimated regression coefficients of some of the tested models using the 8,263 homogeneous road sections and the associated *t*-statistics are presented as Models 0–7 in Table 2. The estimated overdispersion parameter ($\hat{\tau}$), loglikelihood function evaluated at the estimated coefficients, $L(\hat{\beta})$, and the Akaike Information Criterion (AIC) value (*22*) for each model are also given in the table. Note that AIC $= -2L(\hat{\beta}) + 2k$, where k is the total number of regression coefficients in the model, and the estimated models with high loglikelihood function and low AIC values are preferred. Furthermore, the expected total number of trucks involved in accidents across road sections ($\Sigma_i \hat{\mu}_i$) was compared with the observed total ($\Sigma_i y_i$).

These eight models in Table 2 are arranged as follows.

Model 0: This is the simplest form of the Poisson regression model, which includes only truck exposure (v_i). That is, Y_i is assumed to be Poisson distributed with mean $\mu_i = v_i \exp(\beta_1)$. This model served as a baseline for the measurement of model improvement as additional explanatory variables were included.

Model 1: This model includes only truck exposure and yearly dummy variables (x_{ij}, $j = 2, \ldots, 5$) to capture year-to-year changes in the overall truck accident involvement rate.

Model 2: This model includes truck exposure, yearly dummy variables, and traffic variables, including AADT per lane (x_{i6}) and percent trucks ($x_{i,12}$).

Models 3–5: These models include truck exposure, yearly dummy variables, traffic variables, and geometric design variables, including horizontal curvature, length of original curve, vertical grade, length of original grade, and paved inside shoulder width (x_{ij}, $j = 7, 8, 9, 10, 11$). The interactions between horizontal curvature and length of curve ($x_{i,13}$) and between vertical grade and length of grade ($x_{i,14}$) were also tested. (Note that the interaction between HC and VG was not found to be significant at a 20 percent α level.)

Model 6: This model uses the same explanatory variables as in Model 5. It was intended for examining the effect of short road sections on the estimation of model coefficients. Only road sections with length greater than 0.08 km (0.05 mi) were used to estimate model coefficients. There were 7,004 road sections and 1,603 reported truck accident involvements.

Model 7: This model has the same explanatory variables as in Model 5. It was used for checking the effect of road sections with large model residuals on the estimation of model coefficients. Based on Model 5, there were 53 road sections with large standardized residual values [defined as road sections with $|y_i - \hat{\mu}_i|/(\hat{\tau}\,\hat{\mu}_i)^{1/2} > 5$]. These road sections were first removed and Model 5 was then recalibrated to obtain Model 7.

The following observations can be made from these eight models:

1. The AIC value continues to decrease and $L(\hat{\beta})$ continues to increase from Model 0 to Model 3, as yearly dummy variables, traffic variables, and geometric design variables are included in the model.
2. By comparing Model 3 with Model 4, it can be observed that it is through the interaction with horizontal curvature that length of curve becomes a significant factor in affecting truck accident involvement rate. This is shown by the unadjusted *t*-statistic of 0.02 for $\hat{\beta}_8$ in Model 3 and of 2.76 for $\hat{\beta}_{13}$ in Model 4.
3. It is suggested from Model 4 that length of grade by itself is a significant determinant for truck accident involvement rate. [The adjusted *t*-statistics for $\hat{\beta}_{10}$ is $1.99/(1.57)^{1/2} = 1.59$.] As can be seen from Model 5 through the interaction with vertical grade, the effect of length of grade becomes more significant. [The adjusted *t*-statistics for $\hat{\beta}_{14}$ is $2.26/(1.57)^{1/2} = 1.80$.] In this study, Model 5, which had the lowest AIC value, was considered for further analyses and illustrations. The asymptotic correlation matrix, $\hat{\rho}_{ij}$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, k$, for the estimated regression coefficients in Model 5 is shown in Table 3.
4. The comparison of the estimated coefficients of Model 5 and Model 6 suggested not only that the conclusions reached regarding the significance level of the relationships between truck accidents and the examined traffic and highway geometric variables were consistent, but also that the estimated coefficient values were very close. This suggests that the Poisson regression model is not sensitive to the length of road sections.
5. The comparison of the estimated regression coefficients for the traffic and geometric design variables (i.e., β_6 through β_{14}) between Model 5 and Model 7 suggested that the deletion

TABLE 2 Estimated Regression Coefficients of Some Tested Poisson Regression Models and Associated Statistics

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Section length and number of road sections	≥0.01 mi 8,263	≥0.01 mi 8,263	≥0.01 mi 8,263	≥0.01 mi 8,263	≥0.01 mi 8,263	≥0.01 mi 8,263	>0.05 mi 7,004	≥0.01 mi 8,210
β_1 Dummy intercept	-0.212230 (±0.025;-8.60)	0.121263 (±0.058;2.08)	0.570141 (±0.112;5.07)	-0.472330 (±0.287;-1.65)	-0.472494 (±0.287;-1.65)	-0.431762 (±0.288;-1.50)	-0.526103 (±0.290;-1.81)	-0.500055 (±0.298;-1.68)
β_2 Dummy variable for 1986	-----	-0.363320 (±0.082;-4.44)	-0.163271 (±0.086;-1.90)	-0.182576 (±0.086;-2.12)	-0.185384 (±0.086;-2.15)	-0.183853 (±0.086;-2.14)	-0.171759 (±0.087;-1.97)	-0.252277 (±0.089;-2.83)
β_3 Dummy variable for 1987	-----	-0.340802 (±0.080;-4.24)	-0.139415 (±0.085;-1.64)	-0.160249 (±0.085;-1.89)	-0.162656 (±0.085;-1.91)	-0.161461 (±0.085;-1.90)	-0.160869 (±0.086;-1.86)	-0.186897 (±0.087;-2.14)
β_4 Dummy variable for 1988	-----	-0.327909 (±0.078;-4.21)	-0.090187 (±0.085;-1.06)	-0.114524 (±0.085;-1.35)	-0.112753 (±0.085;-1.33)	-0.111511 (±0.085;-1.31)	-0.096243 (±0.086;-1.12)	-0.166516 (±0.088;-1.90)
β_5 Dummy variable for 1989	-----	-0.518223 (±0.079;-6.54)	-0.289009 (±0.088;-3.29)	-0.315484 (±0.088;-3.57)	-0.313863 (±0.088;-3.57)	-0.311155 (±0.088;-3.54)	-0.299701 (±0.089;-3.36)	-0.355124 (±0.091;-3.92)
β_6 AADT per lane (10^3)	-----	-----	0.027600 (±0.015;1.85)	0.026710 (±0.015;1.73)	0.022138 (±0.015;1.38)	0.024400 (±0.015;1.59)	0.025220 (±0.015;1.63)	0.030559 (±0.016;1.94)
β_7 Horizontal curvature	-----	-----	-----	0.147259 (±0.022;6.85)	0.089178 (±0.028;3.15)	0.088861 (±0.028;3.14)	0.096170 (±0.029;3.27)	0.081928 (±0.030;2.75)
β_8 Length of original curve	-----	-----	-----	0.004148 (±0.232;0.02)	-----	-----	-----	-----
β_{13} (Horizontal curvature)× (Length of original curve)	-----	-----	-----	-----	0.232377 (±0.084;2.76)	0.234209 (±0.084;2.78)	0.221877 (±0.087;2.56)	0.239432 (±0.088;2.73)
β_9 Vertical grade	-----	-----	-----	0.083423 (±0.027;3.06)	0.084194 (±0.027;3.09)	0.077815 (±0.028;2.81)	0.078218 (±0.028;2.78)	0.050211 (±0.028;1.77)
β_{10} Length of original grade	-----	-----	-----	0.165342 (±0.078;2.11)	0.156212 (±0.078;1.99)	-----	-----	-----
β_{14} (Vertical grade)× (Length of original grade)	-----	-----	-----	-----	-----	0.033973 (±0.01;2.26)	0.031085 (±0.015;2.03)	0.044749 (±0.015;2.89)
β_{11} Deviation of paved inside shoulder width from 12 ft	-----	-----	-----	0.088652 (±0.036;2.46)	0.091478 (±0.036;2.54)	0.085763 (±0.036;2.37)	0.094814 (±0.036;2.60)	0.088546 (±0.037;2.36)
β_{12} Percent trucks (e.g., 15)	-----	-----	-0.028940 (±0.004;-6.96)	-0.025260 (±0.004;-5.91)	-0.025738 (±0.004;-6.01)	-0.025233 (±0.004;-5.88)	-0.025308 (±0.004;-5.82)	-0.022769 (±0.004;-5.11)
†	1.90	1.84	1.76	1.57	1.57	1.57	1.32	0.97
$L(\hat{\beta})$	-3916.4	-3895.0	-3845.5	-3775.3	-3771.7	-3771.0		
AIC Value	7834.7	7800.0	7705.0	7574.5	7567.3	7566.0		
Expected vs. Observed Total Truck Accident Involvements	1,641.8 1,643.0	1,645.6 1,643.0	1,641.6 1,643.0	1,642.3 1,643.0	1,644.2 1,643.0	1,644.3 1,643.0	1,604.5 1,603.0	1,540.8 1,539.0

Notes: (1) Values in parentheses are (unadjusted) asymptotic standard deviation and t-statistics of the coefficients above.
(2) ----- Not included in the model.
(3) 1 mile = 1.61 km, 1 ft = 0.3048 m.

TABLE 3 Asymptotic Correlation Matrix, ($\hat{\rho}_{ij}$), of the Estimated Regression Coefficients, $\hat{\beta}$, for Model 5

	1	2	3	4	5	6	7	13	9	14	11	12
1	1.000	-0.095	-0.087	-0.083	-0.054	-0.193	-0.059	0.034	-0.080	0.099	-0.905	-0.058
2	-0.095	1.000	0.564	0.586	0.578	-0.169	-0.026	0.017	-0.001	-0.029	0.059	-0.305
3	-0.087	0.564	1.000	0.601	0.594	-0.194	-0.029	0.018	0.004	-0.034	0.058	-0.324
4	-0.083	0.586	0.601	1.000	0.628	-0.256	-0.037	0.025	0.012	-0.048	0.080	-0.401
5	-0.054	0.578	0.594	0.628	1.000	-0.314	-0.039	0.029	0.013	-0.053	0.070	-0.425
6	-0.193	-0.169	-0.194	-0.256	-0.314	1.000	0.047	-0.076	-0.070	0.127	-0.076	0.590
7	-0.059	-0.026	-0.029	-0.037	-0.039	0.047	1.000	-0.792	-0.050	0.024	0.009	0.106
13	0.034	0.017	0.018	0.025	0.029	-0.076	-0.792	1.000	-0.007	-0.020	-0.003	-0.069
9	-0.080	-0.001	0.004	0.012	0.013	-0.070	-0.050	-0.007	1.000	-0.783	-0.068	0.001
14	0.099	-0.029	-0.034	-0.048	-0.053	0.127	0.024	-0.020	-0.783	1.000	-0.041	0.113
11	-0.905	0.059	0.058	0.080	0.070	-0.076	0.009	-0.003	-0.068	-0.041	1.000	-0.281
12	-0.058	-0.305	-0.324	-0.401	-0.425	0.590	0.106	-0.069	0.001	0.113	-0.281	1.000

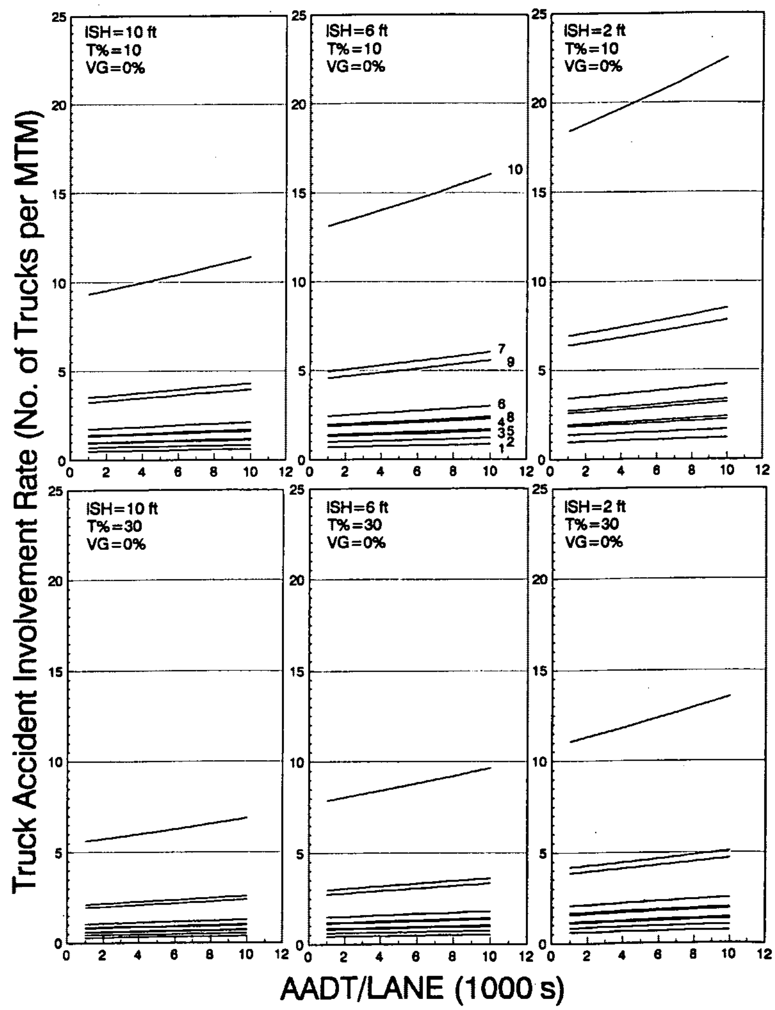

FIGURE 1 The relationship between truck accident involvement rate and key highway geometric design variables for rural Interstate highways (*continued on next page*).

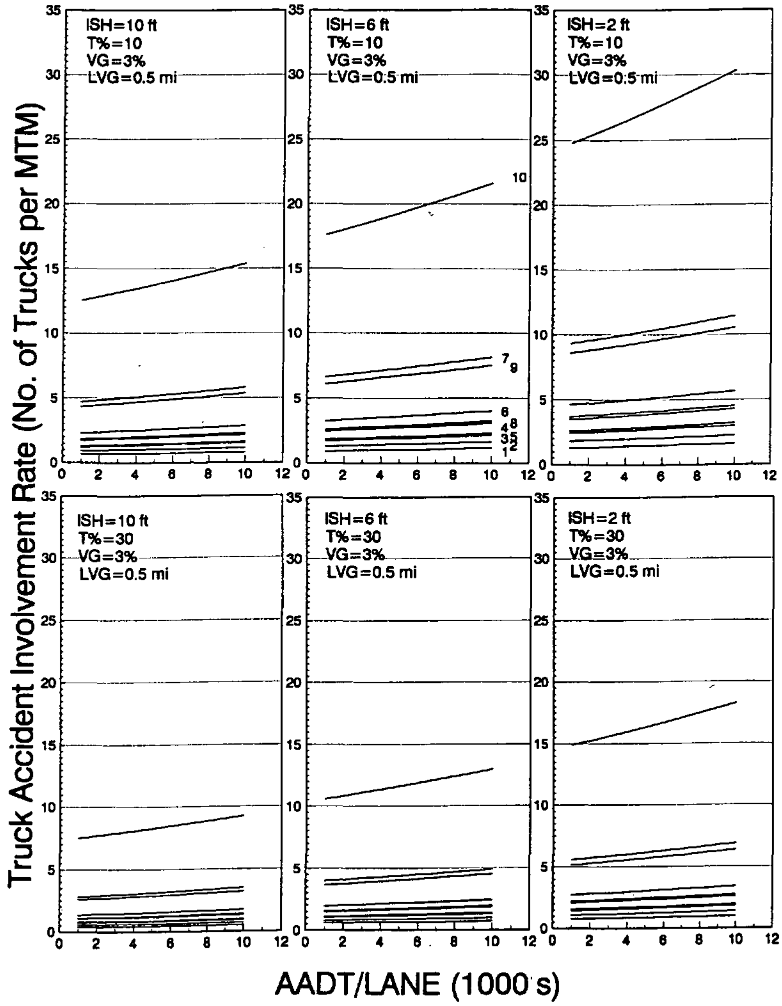

FIGURE 1 *(continued).*

of the 53 road sections with high standardized residuals only slightly altered the coefficient estimates. This also meant that no particular road section had unusually high influence on the estimates.

6. All of the estimated coefficients for the traffic and geometric variables are consistent among different models and have expected algebraic signs.

7. Based on Model 5, truck accident involvement rates for different combinations of AADT/lane, horizontal curvature, length of original curve, vertical grade, length of original grade, paved inside shoulder width, and percent trucks are illustrated in Figure 1. These rates are computed using the average estimated coefficients for 1987–1989 dummy variables as a base rate: $\hat{\lambda}_i = \exp[\hat{\beta}_1 + (\hat{\beta}_3 + \hat{\beta}_4 + \hat{\beta}_5)/3 + x_{i,6}\hat{\beta}_6 + x_{i,7}\hat{\beta}_7 + x_{i,13}\hat{\beta}_{13} + x_{i,9}\hat{\beta}_9 + x_{i,14}\hat{\beta}_{14} + x_{i,11}\hat{\beta}_{11} + x_{i,12}\hat{\beta}_{12}] = \exp[-0.626471 + 0.02440x_{i,6} + 0.088861x_{i,7} + 0.234209x_{i,13} + 0.077815x_{i,9} + 0.033973x_{i,14} + 0.085763x_{i,11} - 0.025233x_{i,12}]$
[In Figure 1, Lines 1 through 10 in each part of the figure show truck accidents–geometric design relationships for different combinations of horizontal curvature (HC) in degrees per 30.48-m (100-ft arc) and length of original curve (LHC) in mi: Line 1: HC = 0; Line 2: HC = 3, LHC = 0.1; Line 3: HC = 3, LHC = 0.5; Line 4: HC = 3, LHC = 1.0; Line 5: HC = 6, LHC = 0.1; Line 6: HC = 6, LHC = 0.5; Line 7: HC = 6, LHC = 1.0; Line 8: HC = 9, LHC = 0.1; Line 9: HC = 9, LHC = 0.5; Line 10: HC = 9, LHC = 1.0. Note that this figure applies mainly to road sections with 3.66-m (12-ft) lane width and 3.05-m (10-ft) paved outside shoulder width. Also, in each part of the figure, the line numbers from the bottom to the top are: 1, 2, 3, 5, 4, 8, 6, 9, 7, and 10.]

8. For the ranges of covariates indicated in Table 1, Model 5 suggests the following relationships between geometric design elements and truck accident involvement rates:

1. As AADT/lane increases by 1,000 vehicles/lane, truck accident involvement rate increases by about 2.5 percent.

2. As horizontal curvature increases, truck accident involvement rate increases. However, the increase depends on the length of curve. For example, for a curve with 0.1 mi in length and with curvature greater than 1 degree/30.48-m (100-ft) arc, as horizontal curvature increases by 1 degree, truck accident involvement rate increases by about 11.9 percent.

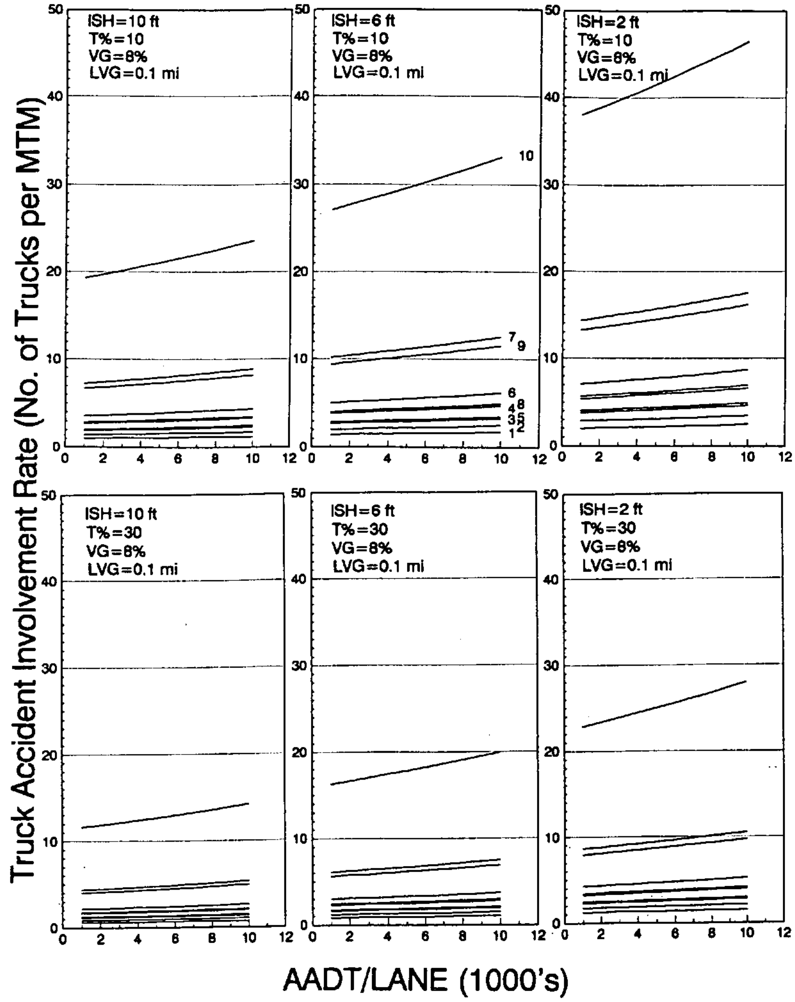

FIGURE 1 *(continued)*.

3. As vertical grade increases, truck accident involvement rate increases. The increase, however, depends on the length of grade. For example, for a grade with 0.8 km (0.5 mi) in length and with vertical grade greater than 2 percent, as grade increases by 1 percent, truck accident involvement rate increases by about 9.9 percent.

4. As the length of curve increases, truck accident involvement rate increases. The increase, however, depends on the curvature degree. For example, for a 5 degree curve, as the length of curve increases by 0.16 km (0.1 mi), truck accident involvement rate increases by about 7.3 percent.

5. As the length of grade increases, truck accident involvement rate increases. The increase depends on the steepness of vertical grade. For example, for a 5 percent grade, as the length of grade increases by 0.8 km (0.5 mi), truck accident involvement rate increases by about 5.2 percent.

6. As paved inside shoulder width per direction increases by 0.3048 m (1 ft), truck accident involvement rate decreases by about 8.2 percent.

7. For a constant vehicle density, as percent trucks in the traffic stream increases by 5 percent, truck accident involvement rate decreases by about 11.9 percent.

Example Applications

Based on Model 5, the reduction in the expected number of truck accident involvements and its estimated one-standard deviation (from Equations 8 and 9) caused by improvements in horizontal curvature, vertical grade, and paved inside shoulder width of a road section, are illustrated in Tables 4 and 5. These illustrations assume no changes in truck travel after the improvements. The expected reductions caused by an improvement in one geometric design element are shown in Table 4, and the expected reductions caused by improvements in two geometric design elements are shown in Table 5. Note that Equations 8 and 9 can be used to estimate the expected reductions of a road section caused by improvements in any combination of geometric design elements.

TABLE 4 Expected Reductions in Truck Accident Involvements on a Rural Interstate Road Section After an Improvement in One Geometric Design Element

Length of Original Curve (mi)	Horizontal Curvature (HC) in degrees/100-ft arc: for 2° ≤ HC ≤ 12°				
	Reduce 1°	Reduce 2°	Reduce 3°	Reduce 4°	Reduce 5°
0.10	10.6% (±2.5%)	20.1% (±4.5%)	28.6% (±6.0%)	36.2% (±7.2%)	43.0% (±8.1%)
0.25	13.7% (±1.9%)	25.5% (±3.3%)	35.7% (±4.2%)	44.5% (±4.9%)	52.1% (±5.3%)
0.50	18.6% (±2.7%)	33.8% (±4.4%)	46.1% (±5.4%)	56.1% (±5.8%)	64.3% (±6.0%)
0.75	23.2% (±4.3%)	41.1% (±6.6%)	54.8% (±7.7%)	65.3% (±8.0%)	73.4% (±7.8%)
≥1.00	27.6% (±5.8%)	47.6% (±8.6%)	62.1% (±9.6%)	72.5% (±9.5%)	80.1% (±9.0%)

Length of Original Grade (mi)	Vertical Grade (VG): for 2% < VG <9%				
	Reduce 1%	Reduce 2%	Reduce 3%	Reduce 4%	Reduce 5%
0.10	7.8% (±3.1%)	15.0% (±5.7%)	21.6% (±7.9%)	27.7% (±9.7%)	33.4% (±11.3%)
0.50	9.0% (±2.5%)	17.3% (±4.6%)	24.7% (±6.3%)	31.5% (±7.7%)	37.7% (±8.8%)
1.00	10.6% (±2.1%)	20.0% (±3.7%)	28.5% (±5.0%)	36.0% (±5.9%)	42.8% (±6.7%)
≥2.00	13.5% (±2.1%)	25.3% (±3.6%)	35.4% (±4.6%)	44.2% (±5.4%)	51.7% (±5.8%)

	Paved Inside Shoulder Width (ISH) per Direction: for ISH ≤ 12 ft				
	Increase 1 ft	Increase 2 ft	Increase 3 ft	Increase 4 ft	Increase 5 ft
	8.2% (±4.2%)	15.7% (±7.7%)	22.7% (±10.7%)	29.0% (±13.2%)	34.9% (±15.4%)

Notes: (1) Values in parentheses are one standard deviation of the expected reductions above.
(2) 1 ft = 0.3048 m; 1 mi = 1.61 km.

TABLE 5 Expected Reductions in Truck Accident Involvements on a Rural Interstate Road Section After an Improvement in Two Geometric Design Elements

Length of Original Curve (LHC) = 0.10 mi and Length of Original Grade (LVG) = 0.50 mi					
Vertical Grade (VG): for 2% < VG < 9%	Horizontal Curvature (HC) in degrees/100-ft arc: for 2° ≤ HC ≤ 12°				
	Reduce 1°	Reduce 2°	Reduce 3°	Reduce 4°	Reduce 5°
Reduce 1%	18.7% (±3.1%)	27.3% (±4.4%)	35.0% (±5.6%)	42.0% (±6.6%)	48.1% (±7.4%)
Reduce 2%	26.0% (±4.5%)	33.9% (±5.0%)	40.9% (±5.8%)	47.2% (±6.4%)	52.8% (±7.0%)
Reduce 3%	32.7% (±5.8%)	39.9% (±5.9%)	46.3% (±6.2%)	52.0% (±6.5%)	57.1% (±6.9%)
Reduce 4%	38.8% (±7.0%)	45.3% (±6.7%)	51.1% (±6.7%)	56.3% (±6.7%)	61.0% (±6.9%)
Reduce 5%	44.3% (±7.9%)	50.3% (±7.4%)	55.5% (±7.1%)	60.3% (±7.0%)	64.5% (±6.9%)

Length of Original Curve (LHC) = 0.10 mi					
Paved Inside Shoulder Width per Direction (ISH): for ISH ≤ 12 ft	Horizontal Curvature (HC) in degrees/100-ft arc: for 2° ≤ HC ≤ 12°				
	Reduce 1°	Reduce 2°	Reduce 3°	Reduce 4°	Reduce 5°
Increase 1 ft	18.0% (±4.4%)	26.7% (±5.3%)	34.5% (±6.3%)	41.4% (±7.1%)	47.6% (±7.8%)
Increase 2 ft	24.7% (±7.2%)	32.7% (±7.3%)	39.9% (±7.5%)	46.2% (±7.9%)	52.0% (±8.2%)
Increase 3 ft	30.9% (±9.8%)	38.2% (±9.3%)	44.8% (±9.0%)	50.7% (±8.9%)	55.9% (±8.9%)
Increase 4 ft	36.6% (±12.0%)	43.3% (±11.1%)	49.3% (±10.5%)	54.7% (±10.0%)	59.5% (±9.7%)
Increase 5 ft	41.8% (±13.8%)	48.0% (±12.7%)	53.5% (±11.8%)	58.4% (±11.1%)	62.8% (±10.5%)

Notes: (1) Values in parentheses are one standard deviation of the expected reductions above.
(2) 1 ft = 0.3048 m; 1 mi = 1.61 km.

To give a simple illustration of the computations involved, consider a curved road section i with 0.16 km (0.10 mi) in length. By reducing 1 degree (per 30.48-m or 100-ft arc) of the curve and all else equal, the expected truck accident involvement reduction percentage is calculated as

$$\begin{aligned}\hat{R}_i &= \{1 - \exp[(x^a_{i,7} - x^b_{i,7})\hat{\beta}_7 + (x^a_{i,13} - x^b_{i,13})\hat{\beta}_{13}]\} \times 100 \\ &= \{1 - \exp[(x^a_{i,7} - x^b_{i,7})\hat{\beta}_7 + (x^a_{i,7} \times x^a_{i,8} - x^b_{i,7} \times x^b_{i,8})\hat{\beta}_{13}]\} \\ &\quad \times 100 \\ &= \{1 - \exp[(-1) \times 0.088861 + (-1 \times 0.1) \times 0.234209]\} \\ &\quad \times 100 \\ &= [1 - \exp(-0.1123)] \times 100 \\ &= 10.6\end{aligned}$$

The standard error of this expected reduction percentage is computed using Equation 9 as

$$\begin{aligned} s.d.(\hat{R}_i) \approx & \left\{\exp\left[(x^a_{i,7} - x^b_{i,7})\hat{\beta}_7 + (x^a_{i,13} - x^b_{i,13})\hat{\beta}_{13} + \frac{\hat{\tau}}{2}\right.\right. \\ & \times \left((x^a_{i,7} - x^b_{i,7})^2 s_{7,7} + (x^a_{i,13} - x^b_{i,13})^2 s_{13,13}\right. \\ & \left.\left.\left. + 2(x^a_{i,7} - x^b_{i,7})(x^a_{i,13} - x^b_{i,13})\hat{\rho}_{7,13}(s_{7,7}s_{13,13})^{1/2}\right)\right]\right\} \\ & \times \left\{\exp\left[\hat{\tau}\left((x^a_{i,7} - x^b_{i,7})^2 s_{7,7} + (x^a_{i,13} - x^b_{i,13})^2 s_{13,13}\right.\right.\right. \\ & \left.\left.\left. + 2(x^a_{i,7} - x^b_{i,7})(x^a_{i,13} - x^b_{i,13})\hat{\rho}_{7,13}(s_{7,7}s_{13,13})^{1/2}\right)\right] - 1\right\}^{1/2} \\ & \times 100 \\ = & \left\{\exp\left[-0.1123 + \frac{1.57}{2}\left((-1)^2(0.028)^2\right.\right.\right. \\ & + (-1 \times 0.1)^2(0.084)^2 \\ & \left.\left.\left. + 2 \times (-1)(-1 \times 0.1)(-0.792)(0.028)(0.084)\right)\right]\right\} \\ & \times \left\{\exp\left[1.57\left((-1)^2(0.028)^2 + (-1 \times 0.1)^2(0.084)^2\right.\right.\right. \\ & \left.\left. + 2 \times (-1)(-1 \times 0.1)(-0.792)(0.028)(0.084)\right)\right] \\ & \left. - 1\right\}^{1/2} \times 100 \\ = & \ 2.5 \end{aligned}$$

where $(s_{7,7})^{1/2}$ and $(s_{13,13})^{1/2}$ are the standard deviations of the estimated regression coefficients $\hat{\beta}_7$ and $\hat{\beta}_{13}$, respectively, and are available in Table 2.

The Poisson regression model introduced in this paper can be developed and tested for other states in a similar manner. For those states in which detailed rural Interstate roadway and accident data are not available for conducting such an analysis, it is recommended that Model 5 be used with a slight modification, as follows:

$$\begin{aligned}\hat{\mu}_i = \left(\frac{\mathrm{AR}}{0.81}\right) v_i \exp(&-0.626471 + 0.02440x_{i,6} \\ &+ 0.088861x_{i,7} + 0.234209x_{i,13} + 0.077815x_{i,9} \\ &+ 0.033973x_{i,14} + 0.085763x_{i,11} - 0.025233x_{i,12}) \qquad (10)\end{aligned}$$

where AR represents the overall truck accident involvement rate/MTM in recent years for the rural Interstate Highways in another state of interest, and 0.81 is the overall truck accident involvement rate/MTM for the road sections examined in this study. This modification is intended to adjust for the differences between Utah and the state of interest in, for example, weather and socioeconomic conditions, as well as the differences in accident reporting practices for nonfatal accidents and in the criteria used for classifying roadways. Under this modified model, the expected percentage reductions in truck accident involvements and associated standard deviations can still be computed from Equations 8 and 9 without any changes.

ACKNOWLEDGMENTS

This paper is based on the results of a research project sponsored by the Office of Safety and Traffic Operations Research and Development, Federal Highway Administration, the U.S. Department of Transportation. The opinions expressed in this paper are, however, solely those of the authors.

REFERENCES

1. *Fatal and Injury Rates on Public Roads in the United States.* FHWA-SA-90-029, Federal Highway Administration, U.S. Department of Transportation, June 1990.
2. L. Evans. *Traffic Safety and the Driver.* Van Nostrand Reinhold, New York, N.Y., 1991.
3. Cost and Safety Effectiveness of Highway Design Elements. In *National Cooperative Highway Research Program, Report 197,* Washington, D.C., 1978.
4. K. Rumar. The Role of Perceptual and Cognitive Filters in Observed Behavior. *Human Behavior and Traffic Safety* (L. Evans and R. C. Schwing, eds.) Plenum Press, New York, N.Y., 1985, pp. 151–165.
5. P. P. Jovanis and H. L. Chang. Modeling the Relationship of Accidents to Miles Traveled. In *Transportation Research Record 1068*, TRB, National Research Council, Washington, D.C., 1986, pp. 42–51.
6. S.-P. Miaou and H. Lum. Modeling Vehicle Accidents and Highway Geometric Design Relationships. *Accident Analysis and Prevention* (in press).
7. S. C. Joshua and N. J. Garber. Estimating Truck Accident Rate and Involvements Using Linear and Poisson Regression Models. *Transportation Planning and Technology*, Vol. 15, 1990, pp. 41–58.
8. S.-P. Miaou et al. *Development of Relationships Between Truck Accidents and Highway Geometric Design: Phase I.* Technical Memorandum, Oak Ridge National Laboratory, March and Nov. 1991.
9. F. F. Saccomanno and C. Buyco. Generalized Loglinear Models of Truck Accident Rates. Paper presented at Transportation Research Board 67th Annual Meeting, Washington, D.C., January 1988.
10. F. M. Council and J. F. Paniati. The Highway Safety Information System. *Public Roads*, Vol. 54, No. 3, Dec. 1990.
11. D. R. Cox and P. A. W. Lewis. *The Statistical Analysis of Series of Events.* Chapman and Hall, London, 1966.

12. A. C. Cameron and P. K. Trivedi. Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests. *Journal of Applied Econometrics*, Vol. 1, 1986, pp. 29–53.
13. J. S. Cramer. *Econometric Applications of Maximum Likelihood Methods*. Cambridge University Press, New York, N.Y., 1986.
14. D. R. Cox. Some Remarks on Overdispersion. *Biometrika*, 70(1), 1983, pp. 269–274.
15. C. Dean and J. F. Lawless. Tests for Detecting Overdispersion in Poisson Regression Models. *Journal of the American Statistical Association*, Vol. 84, No. 406, 1989, pp. 467–472.
16. A. C. Cameron and P. K. Trivedi. Regression-Based Tests for Overdispersion in the Poisson Model. *Journal of Econometrics*, Vol. 46, 1990, pp. 347–364.
17. R. W. M. Wedderburn. Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss-Newton Method. *Biometrika*, 54, 1974, pp. 439–447.
18. A. Agresti. *Categorical Data Analysis*. John Wiley, New York, N.Y., 1990.
19. B. W. Lindgren. *Statistical Theory*. Third Ed. Macmillian, New York, N.Y., 1976.
20. *A Policy on Geometric Design of Highways and Streets*. American Association of State Highway and Transportation Officials, Washington, D.C., 1990.
21. C. V. Zegeer et al. *Cost-Effective Geometric Improvements for Safety Upgrading of Horizontal Curves*. University of North Carolina, Chapel Hill, May 1990.
22. H. Bozdogan. Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions. *Psychometrika*, 52(3), 1987, pp. 345–370.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.