Estimation of Safety at Two-Way Stop-Controlled Intersections on Rural Highways

JAMES A. BONNESON AND PATRICK T. MCCOY

The application of the generalized linear modeling approach to the development of a model relating unsignalized intersection traffic demands to accident frequency is described. Several techniques for assessing model fit have been described and any inherent limitations noted. The model was based on the product of the intersection traffic demands raised to a power. This model form was found to explain a large portion of the variability in accidents among intersections of similar geometry and traffic control. The analysis of accident data for 125 two-way stop-controlled intersections supports the theory that the distribution of accident counts can be described by the negative binomial distribution. Also supported is the assertion that the mean accident frequency for the group of similar intersections is gamma distributed. Knowledge of these distributions and their parametric values can be used to identify hazardous locations and the true effect of safety treatments on accident frequency.

The intersection of a minor roadway with a rural highway presents several safety problems. Accident rates may be higher at intersections on rural highways than at intersections on urban highways. This trend may be attributed to the higher speeds on the rural highway and a reduced driver expectancy for intersecting traffic movements on rural highways. The combination of high-speed rural highway traffic with low-speed intersecting traffic (i.e., crossing or turning vehicles) can lead to significant speed differentials and an increased potential for accidents. If the highway has a multilane (or expressway) cross section with wide median, the intersection will have a large conflict area. This characteristic, combined with high expressway speeds, can further degrade the safety of the crossing or turning maneuver.

The frequency of access points, the amount of roadway lighting, and the magnitude of traffic demands are typically lower in rural than in urban areas. These characteristics tend to make drivers on rural highways more relaxed and potentially less attentive. As a result, the highway driver's expectancy for turning or crossing vehicles may be relatively low, which can also increase accident potential at intersections on rural highways.

Unfortunately, little is known about the safety of unsignalized intersections on rural highways. As a result, it is difficult to accurately identify intersections that are truly hazardous. Moreover, it is difficult to fully assess the safety benefits of any corrective measures (e.g., advance signing, signalization) applied to hazardous intersections.

Civil Engineering Department, University of Nebraska-Lincoln, Lincoln, Nebr. 68588-0531.

The objective of this research was to develop a methodology for assessing the safety and efficiency of both intersections and interchanges on rural highways. This paper describes the development of a safety-prediction model for rural intersections. The safety of an intersection is defined as the expected number of accidents per year. The research described is part of a more comprehensive analysis of the economic benefits and costs of interchanges, relative to intersections, on rural expressways conducted by the authors for the Nebraska Department of Roads (*1*).

MODELING CONSIDERATIONS

Many factors affect the number of accidents that occur at an intersection. The factors can be categorized into those representing exposure to potential accident events and those affecting the probability that a given potential accident event will result in an accident. Factors that represent exposure include time period and traffic demand. Factors that affect accident probability include urban/rural environment, traffic control, frequency of access points, speed limit, shoulder width, median type, median width, lighting level, availability of left-turn bays, number of legs, and number of traffic lanes.

Several modeling issues were considered in determining the form of the safety-prediction model and the types of data needed to calibrate it. One of the more critical issues is sample size. The high variability of accident data tends to translate into the need for large sample sizes to establish sound statistical relationships. One way to overcome the uncertainties of high variability is to increase the size of the data base. An increase could be accomplished by including more intersections in the data base (if available) or by increasing the duration of the time period that brackets the data (provided that environmental or driver behavior changes do not become significant).

Similarity among intersections in the data base is another issue. The calibrated model must be able to predict accident frequency for the "typical" type of intersection being considered. Thus, a large data base containing many different types of intersections should be subset to yield a smaller data base having intersections more consistent with the type of intersection being considered. For example, in this research, we initially subset the statewide data base to remove all urban intersections.

Although subsetting has the beneficial effect of increasing similarity in the resultant data base, it is achieved by eliminating a portion of the data base. Thus, it has the disadvantage of reducing the available sample size. The trade-offs between sample size and sample similarity must be carefully considered and a proper balance achieved. In general, the amount of subsetting is generally limited to that which will allow the questions of the research to be answered without compromising the analyst's predetermined sample size requirement.

LITERATURE REVIEW

A comprehensive review of the literature regarding accident frequency prediction models before 1981 has been provided by Satterthwaite (*2*). A more recent review has been provided by Hauer et al. (*3*). In the latter work, the authors cite considerable evidence supporting the "product-of-flows-to-power" model, wherein the expected accident frequency is a function of the product of traffic demands entering the junction. In most instances, the traffic demands are raised to a power less than unity, indicating a nonlinear relationship between demand and accident frequency. Some researchers [e.g., Hauer et al. (*3*)] have considered only the flows for the conflicting traffic patterns. Others [e.g., Van Every (*4*)] relate expected accident frequency to the product of the average daily traffic demands on the major and minor roadways at the junction. The latter approach was particularly appealing because it does not require detailed information about travel patterns through the intersection. The model used by Van Every (*4*) is shown in Equation 1.

$$A = 0.00042 T_m^{0.5} T_c^{0.5} \tag{1}$$

where

A = expected annual accident frequency,
T_m = major road traffic demand (veh/day), and
T_c = minor (cross) road traffic demand (veh/day).

On the basis of this review of the literature, it was determined that both exposure measures (i.e., time period and traffic demand) were essential components of the data base. It was also determined that environment (urban versus rural), traffic control (signal, sign), and geometry (number of legs) were among the most important factors to be considered when subsetting the accident data base.

MODEL DEVELOPMENT

Data Base

The data base used for this study was obtained from FHWA via the Highway Safety Research Center at the University of North Carolina. This data base was a subset of FHWA's Highway Safety Information System (HSIS) (*5*). HSIS integrates the accident, roadway design, and traffic volume data from the highway departments of Utah, Minnesota, Illinois, Maine, and Michigan. The inclusion of roadway design and average daily traffic volume (ADT) data in the accident data base is one of HSIS's key features; it is a feature not found in most state accident record systems. This type of comprehensive data base was essential to the calibration of a safety-prediction model because it allowed the relationship between geometry, traffic demand, and accidents to be fully explored.

The data base includes all accidents during 1985, 1986, and 1987 for Minnesota. The other states in HSIS were not included because they did not recognize junctions as entities or explicitly differentiate between intersections and interchanges.

The data base can be described as a junction-based file because all data subsetting is related to the type of junction. The subset factors were selected such that the data base included intersections that are similar with regard to nonurban environment, four legs only, two-way stop-control, and intersection geometry (i.e., not an interchange ramp-junction). Once the appropriate intersections were identified, traffic demand and accident data files were scanned to find the corresponding demands and accidents (if any) for these locations. Accidents were assumed to be intersection related if they occurred within 153 m (500 ft) of the junction.

The resulting data base contained major and minor road ADTs for 125 intersections, which experienced 250 accidents in the 3-year period. Further subsetting based on accident pattern, number of lanes, median width, and so forth was not considered because the sample size was determined to be at its smallest acceptable value. A summary of the data base used in this study is given in Table 1.

The data base included two general types of major road geometry. Included were 108 intersections with a two-lane major road and no median and 17 intersections with a four-lane major road and a median of 10.4 m (34 ft) or more. The two-lane roads generally had major road ADTs under 4,000,

TABLE 1 Accident Data Base Statistics[1]

	Statistic			
Variable	Minimum	Maximum	Mean	Std. Dev.
Major Road ADT	430	37,900	4,030	6,140
Minor Road ADT	45	8,850	680	1,060
Accidents/year[2]	0	7	0.67	1.20

[1]Statistics based on 125 rural highway intersections with two-way stop-control.

[2]Average annual accident frequency (based on data for three years).

whereas the four-lane roads had ADTs over 6,000; there was some overlap in the ADT range of 4,000 to 6,000.

This dichotomy in the data base was considered desirable because it related to the objectives of the research. In this regard, the question posed by the research was, What are the expected number of accidents at a given highway junction if an unsignalized intersection or an interchange were built there? The answer would be used to make "planning-level" decisions regarding the most appropriate junction type. To answer this question, safety models were developed for each junction type (only the intersection model is described herein) for a common range of traffic demands. With the dichotomized data base, we were able to analyze intersections with ADTs well into the range found at interchanges. In application, it is understood that the model's predictions would reflect the likelihood that a two-lane major road exists for lower-demand situations and a four-lane road exists for higher-demand situations.

The relationship between average annual accident frequency and traffic demand is given in Table 2 for the two-way stop-controlled intersections included in this study. Examination of the row and column summaries indicates a positive correlation between traffic demand and average annual accident frequency. The traffic demand ranges in Table 2 were selected such that approximately one-fifth of the total number of accidents are located in each row and column. The intent of distributing the accidents in this manner is to obtain an equal weight, in terms of observations, underlying the average annual accident frequency provided in the row and column summaries.

Modeling Techniques

Model Structure

The nonlinear relationship between accident frequency and traffic demand evidenced in Table 2 is consistent with the nonlinear product-of-flows-to-power formulation, as advocated by others (*2–4*). As a result, the following model form was considered:

$$E(m) = b_0 T_{\mathrm{m}}^{b_1} T_{\mathrm{c}}^{b_2} \qquad (2)$$

where

$E(m)$ = expected accident frequency,
b_i = regression constants,
T_{m} = major road traffic demand (veh/day), and
T_{c} = minor (cross) road traffic demand (veh/day).

The approach taken in calibrating the accident prediction model was based on procedures described by Hauer et al. (*3*), who have argued against using traditional least-squares regression of accident data because of violations in two assumptions (i.e., normally distributed error structure and constant variance) on which this type of analysis is based. Instead, Hauer et al. (*3*) advocate the use of a generalized linear model [e.g., GLIM (*6*)] wherein these assumptions are avoided, thereby yielding a better predictor of accident frequency as influenced by other factors.

Before proceeding it is important to define the safety m of an intersection as its mean accident frequency. This quantity can be estimated by taking the average of the m's [$E(m)$] for a large number of similar intersections, each having identical traffic demands. In this context, similar intersections have one or more geometric or traffic control characteristics in common. The estimate becomes more stable as the intersections become more similar (i.e., as the number of characteristics that they have in common increases).

In the last few years, Hauer et al. (*3*) and others have convincingly argued that the distribution of accident counts for a group of similar sites (e.g., intersections, road sections) can be described by the family of compound Poisson distributions. In this context, there are two sources of variability underlying the count distribution. One source stems from the

TABLE 2 Cross-Tabulation of Accidents by Daily Traffic Demand for Two-Way Stop-Controlled Intersections

Major Road (veh/day)	Minor Road (veh/day)					Row Summary
	45-517	518-1070	1071-1682	1683-3137	3138-8850	
430-2037	0.17[a] 9.67/57[b]	0.81 5.67/7	1.00 1/1	0 0/0	0 0/0	0.25 16.33/65
2038-3887	0.17 2.66/16	0.61 3.67/6	1.13 5.67/5	1.67 1.67/1	2.67 2.67/1	0.56 16.33/29
3888-7675	0.46 3.67/8	1.50 3/2	1.84 3.67/2	1.17 2.33/2	5.00 5/1	1.18 17.67/15
7676-17150	0 0/0	1.00 5/5	3.34 6.67/2	0 0/0	4.00 8/2	2.19 19.67/9
17151-37900	0 0/0	1.33 1.33/1	0.67 1.33/2	2.67 10.67/4	0 0/0	1.90 13.33/7
Column Summary	0.20 16.0/81	0.89 18.67/21	1.53 18.33/12	2.10 14.67/7	3.92 15.67/4	0.67 83.33/125

[a]The top number in each cell represents the average annual accident frequency for the corresponding range of major and minor daily traffic demands (based on data for three years).

[b]The bottom numbers in each cell represent the average annual number of accidents and number of junctions (annual accidents / total junctions) for the range of daily traffic demands.

differences in the m's among the similar sites. The other stems from the randomness in accident frequency at any given site, which is traditionally described as Poisson.

In spite of being similar, each site has its own regional character and driver population, giving it a unique mean accident frequency, m. Thus, the distribution of m's within the group of similar sites can be described by a probability density function with mean $E(m)$ and variance $V(m)$. Hauer et al. (*3*) have shown this distribution to be adequately described by the gamma density function.

Abbess et al. (*7*) have shown that if accident occurrence at a particular site is Poisson distributed, the distribution of accidents around the $E(m)$ of a group of sites can be described by the negative binomial distribution. The variance of this distribution is

$$V(x) = E(m) + \frac{E(m)^2}{k} \tag{3}$$

where x is the observed accident count. Since the variance of the Poisson distribution is $E(m)$, it is apparent that the variance of the negative binomial distribution exceeds that of the Poisson by $[E(m)^2]/k$. Hauer et al. (*3*) have shown that this latter quantity is equivalent to the variance of the mean accident frequency for the group of similar sites, $V(m)$. Hauer et al. (*3*) have also shown that the parameter k can be estimated by fitting Equation 3 to $V(x)$ and $E(m)$ estimates for the group of similar sites. The $V(x)$ is estimated as the squared difference between the accident count and the corresponding $E(m)$ for each site in the group.

Generalized Linear Model

The analysis tool used to estimate the model coefficients was the nonlinear regression procedure (NLIN) in the SAS statistical software package (*8*). This procedure is sufficiently general that it can be modified to accommodate error structures that are not normally distributed. It can also be easily modified to yield maximum-likelihood model coefficients. With these modifications, the NLIN procedure can be used as a generalized linear model similar to the GLIM package (*6*). An example application of NLIN to generalized linear modeling is described in the SAS documentation (*8*, Example 6). The SAS code described in this documentation was modified (due to some errors in printing) and enhanced to include the negative binomial and gamma distributions.

The generalized linear modeling approach relates a linear predictive equation to the expected value of an observation via a "link function." The link function equates this linear predictive relationship to a nonlinear, and perhaps bounded, dependent variable. One link function is theoretically related to the error structure of the data. This link function is sometimes referred to as the "natural" (or canonical) link. The selection of the appropriate link function is often based on the distribution of the error structure; however, as noted by McCullagh and Nelder (*9*), this is not a requirement. The natural link functions for the Poisson distribution and negative binomial distributions are given by Equations 4 and 5, respectively:

$$\eta = \ln[E(m)] \tag{4}$$

$$\eta = \ln\left[\frac{E(m)}{k + E(m)}\right] \tag{5}$$

where the linear predictive equation is

$$\eta = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n \tag{6}$$

To implement proposed model form, it was necessary to take the inverse of the link function [i.e., $E(m) = f^{-1}(\eta)$], equate it to the right-hand side of Equation 2, and solve for η. For the Poisson link function, the resulting linear predictive model takes the following form:

$$\eta = \ln(n) + \ln(b_0) + b_1\ln(T_m) + b_2\ln(T_c) \tag{7}$$

where $\ln(n)$ is termed the offset variable with an implied coefficient of 1.0. For this type of analysis, the offset variable is equivalent to the number of years underlying the observed count (in this study, $n = 3$ years for all observations). This linear predictive model form lends itself to further expansion if additional regression parameters are desired in the model. For example, a study by Pickering et al. (*10*), where additional parameters were included to examine the effects of various geometric elements on accident frequency at unsignalized T-intersections, illustrates the use of this model.

A similar calculation of the linear predictive model form was not as successful for the negative binomial link function. In fact, it was not possible to implement the proposed model (i.e., Equation 2) in its intended form using the natural link for the negative binomial structure. Because of this loss of generality, and in recognition that it is not a requirement to use the natural link, the Poisson link was used for all analyses in this study.

Quality of Fit

Several statistics are available from the NLIN procedure for assessing the model fit and the significance of model coefficients. One measure of model fit provided by NLIN is the generalized Pearson X^2 statistic. This statistic is calculated as

$$X^2 = \sum \frac{[x - \hat{E}(m)]^2}{\hat{V}(x)} \tag{8}$$

where $\hat{V}(x)$ is estimated from Equation 3 by substituting $\hat{E}(m)$ for $E(m)$. This statistic is available from NLIN as the "Weighted Sum of Squares" for the Residual. McCullagh and Nelder (*9*) indicate that this statistic follows the χ^2 distribution with $n - p - 1$ degrees of freedom, where n is the number of observations and p is the number of model parameters. This statistic is asymptotic to the χ^2 distribution for larger sample sizes and exact for normally distributed error structures. As noted by McCullagh and Nelder (*9*), this statistic is not well defined in terms of minimum sample size when applied to nonnormal distributions; therefore, it probably should not be used as an absolute measure of model significance.

Another, more subjective, measure of model fit can be obtained from a plot of the prediction ratio versus the estimate of the expected accident frequency [i.e., $\hat{E}(m)$]. In this context, the prediction ratio is the ratio of the observed accident frequency to the expected accident frequency. This type of plot yields a visual assessment of the predictive capability of the model over the full range of $\hat{E}(m)$. A well-fitted model would have the prediction ratios symmetric about 1.0 over the range of $\hat{E}(m)$. This technique was applied by Hauer et al. (*3*) in a recent study of safety at signalized intersections.

The significance of the parameter coefficients (with respect to the hypothesis that they equal zero) is also helpful in assessing the relevance of model factors. In this regard, NLIN provides the standard error and 95 percent confidence interval for each coefficient. Because the Pearson X^2 statistic (i.e., Equation 8) has some limitations, the significance of the individual parameter coefficients may represent a more realistic measure of model fit.

Finally, the dispersion parameter σ_d is noted by McCullagh and Nelder (*9*) to be a useful statistic for assessing the amount of variation in the observed data. This statistic can be calculated by dividing Equation 8 by the quantity $n - p$. It is also available from NLIN as the "Weighted Mean Square" for the Residual. A dispersion parameter near 1.0 indicates that the assumed error structure is approximately equivalent to that found in the data. For example, if a Poisson error structure is assumed [i.e., $V(x) = E(m)$] and the dispersion parameter is 1.68, the data have greater dispersion than is explained by the Poisson distribution. In this situation, the negative binomial distribution might be considered, since it has a larger variance than does the Poisson (see Equation 3).

Analysis Procedure

Coefficient estimation for the proposed model was a multistep process. First, the data were analyzed using a Poisson error structure. Then NLIN was used to fit Equation 3 to the squared residuals from the first analysis [Hauer et al. (*3*) indicate that the squared residuals can be used as an estimate of $V(x)$]. This second step yielded an estimate of the k parameter and a measure of its statistical significance. The need for a third analysis step was based on an assessment of the dispersion parameter and the k parameter significance. If the dispersion parameter was more than 1.0 and the k parameter was statistically significant, a third analysis step was conducted using the negative binomial error structure with k from Step 2 as an estimate of the shape parameter. The residuals from this analysis were analyzed in a fourth step to determine a new k parameter. The third and fourth steps were repeated until convergence on a value of k. This procedure is consistent with that described by Hauer et al. (*3*).

Calibrated Models

The results of the first analysis step (Poisson error) are given in Table 3. As indicated in the table, there was sufficient dispersion (i.e., $\sigma_d > 1.0$) to justify further analysis in terms of the third and fourth steps (negative binomial error). The second step indicated that a k parameter of 3.9 was significant. The Pearson X^2 statistic indicated that the assumed Poisson error structure did not account for a significant portion of the dispersion in the observations. Thus, further analysis via Steps 3 and 4 appeared warranted.

After two iterations of Steps 3 and 4 using a negative binomial error structure, the k parameter converged to 4.0. The fit of Equation 3 with the squared residuals is shown in Figure 1. Each circle represents the expected accident frequency averaged for several intersections. These intersections were ranked in ascending order of accident frequency before averaging. Average values eliminate the visual "noise" associated with a plot of 125 individual data points and thereby better illustrate the underlying curvilinear trend. The number of intersections included in each average was based on the desire to plot 9 or 10 points that have weight as nearly equal as possible (8 points at a weight of 14 intersections and 1 at a weight of 13 worked best). Although most of the observations are clustered near the origin, there appears to be a definite correlation with the variance function.

On the basis of this analysis, the following model was developed for predicting the annual expected accident frequency for two-way stop-controlled intersections on rural highways:

$$E(m) = 0.692\left(\frac{T_m}{1{,}000}\right)^{0.256}\left(\frac{T_c}{1{,}000}\right)^{0.831} \qquad (9)$$

The coefficients in this model are significant at a 95 percent level of confidence (i.e., 5 percent chance of false rejection). The Pearson x^2 statistic was significant, indicating that the

TABLE 3 Model Statistics

Statistic		Analysis Steps 1-2	Analysis Steps 3-4
Model	Error Distribution	Poisson	Neg. Binomial (k=4.0)
	Pearson X^2	215	136*
	$\chi^2_{0.05,n-p-1}$ (n = 125, p = 3)	147	147
	Dispersion Parameter, σ_d	1.76	1.11
Parameter[1]	Intercept, b_0	0.650 (0.080)*	0.692 (0.121)*
	Major Road, b_1	0.292 (0.065)*	0.256 (0.099)*
	Minor Road, b_2	0.791 (0.066)*	0.831 (0.102)*

* Denotes significance at a 95-percent confidence level (5 percent chance of false rejection).

[1] Parameter values include coefficient estimates and standard error. Standard error is in parentheses.

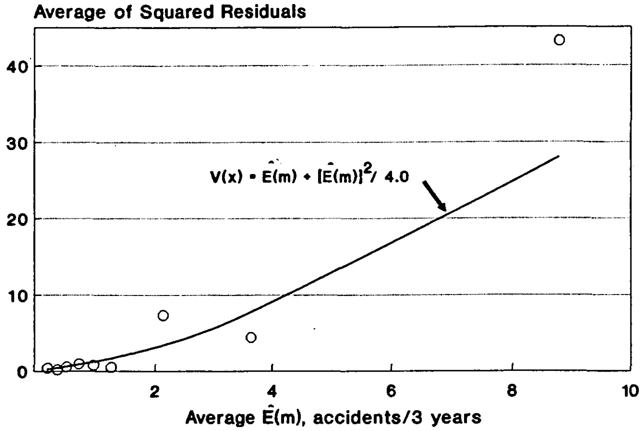

FIGURE 1 Relationship between expected accident frequency and estimated accident variance.

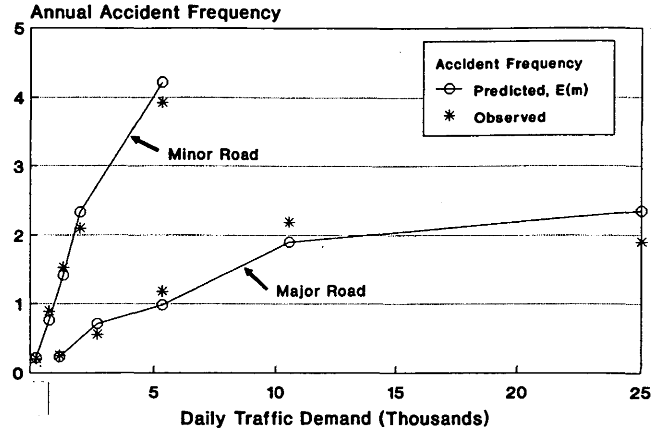

FIGURE 3 Accident frequency as a function of daily traffic demand.

negative binomial error structure is able to explain a significant portion of the deviation from the predicted values. The dispersion parameter was much nearer to 1.0 than for the first analysis, which further supports the selection of the negative binomial structure.

The fit of the model to the data can be assessed using the open circles in Figure 2. The ratio of the observed accident count to the expected accident frequency for a 3-year period is shown. The desired symmetry of observations around 1.0 is not as apparent as desired; however, the large black dots (averages of 14 ratios previously ranked in ascending order) indicate that a symmetry around 1.0 exists.

The curvilinear trend suggested by some combinations of the open circles stems from the reciprocal nature of the plotted quantities [i.e., $E(m)$ versus $1/E(m)$]. For example, the most distinct curvilinear combination in Figure 2 represents those intersections having an observed accident count of one accident per 3 years.

The predictive capability of the model is shown in Figure 3. The starred data points represent the average annual accident frequencies from the row and column summaries of Table 2. Each open circle represents the average of the predicted annual accident frequency for the intersections that are included in the starred data points. The number of intersections that underlies each pair of data points is also provided in the row and column summaries of Table 2.

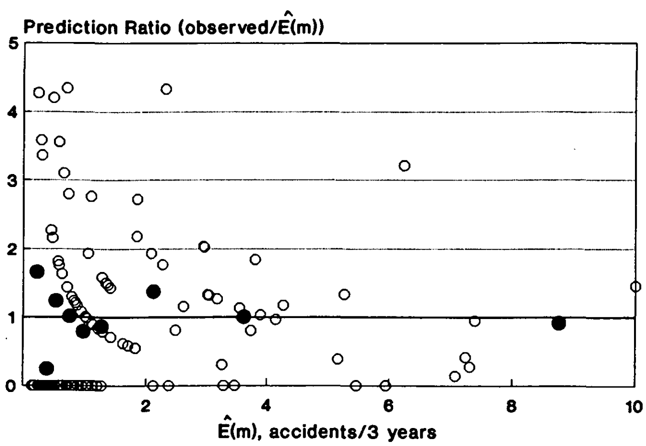

FIGURE 2 Evaluation of model fit using prediction ratios.

Figure 3 provides further support for the exponential relationship between accident frequency and traffic demand. The close agreement between the observed and predicted frequencies suggests that Equation 9 is a good predictor of the expected number of accidents at a typical rural, two-way stop-controlled intersection. A comparison of Equation 9 with the model proposed by Van Every (*4*) (i.e., Equation 1) indicated good agreement between the models.

CONCLUSIONS

This paper describes the application of the generalized linear modeling approach to the development of a model relating intersection traffic demands to accident frequency. The general linear model was implemented using the nonlinear regression procedure (NLIN) of the SAS program (*8*) with appropriate modification.

Several techniques for assessing model fit have been described and any inherent limitations noted. Of the techniques, the authors place the most confidence in the plot of prediction ratio versus expected number of accidents. This plot indicates the amount of dispersion in the predicted values as well as the existence of any model bias over the range of accident frequencies considered.

The product-of-flows-to-power model formulation appears to explain a large portion of the variability in accidents among intersections of similar geometry and traffic control. The strength of this model format has been shown by others (*2–4,10*) and in this paper as applied to two-way stop-controlled intersections on rural highways. The form of this model suggests that mean accident frequency increases in a nonlinear fashion with increasing major or minor road demand.

The analysis of accident data for 125 two-way stop-controlled intersections supports the theory that the distribution of accident counts can be described by the negative binomial distribution. Also supported is the assertion that the mean accident frequency for the group of similar intersections is gamma distributed. Knowledge of these distributions and their parametric values can be used to identify hazardous locations and the true effect of safety treatments on accident frequency. A considerable amount of research in this area has been performed by Hauer and Persaud (*11*).

ACKNOWLEDGMENTS

The authors wish to express their appreciation to the individuals and agencies who assisted in making this project successful. These individuals include Duane Eitel of the Nebraska Department of Roads, Jeffrey F. Paniati of FHWA, and Forrest M. Council of the University of North Carolina. A special thank-you is extended to the latter two gentlemen for making the HSIS data base available. The authors would also like to acknowledge the technical assistance of Brian Moen and Jim Kollbaum, students at the University of Nebraska-Lincoln.

REFERENCES

1. J. A. Bonneson and P. T. McCoy. *Interchange vs. At-Grade Intersection on Rural Expressways.* Research Report TRP-02-25-91. Nebraska Department of Roads, Lincoln, May 1992.
2. S. P. Satterthwaite. *A Survey of Research into Relationships Between Traffic Accidents and Traffic Volumes.* TRRL SP 692. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1981.
3. E. Hauer, J. C. N. Ng, and J. Lovell. Estimation of Safety at Signalized Intersections. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988, pp. 48–61.
4. B. E. Van Every. A Guide to the Economic Justification of Rural Grade Separations. *Australian Road Research*, Vol. 12, No. 3, pp. 147–154.
5. F. M. Council and J. F. Paniati. The Highway Safety Information System. *Public Roads, A Journal of Highway Research and Development*, Vol. 54, No. 3, Dec. 1990, pp. 234–239.
6. R. J. Baker and J. A. Nelder. *The GLIM System: Release 3, Manual.* Rothhamsted Experimental Station, Harpenden, England, 1978.
7. C. Abbess, D. Jarrett, and C. C. Wright. Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment, with Special Reference to the "Regression-to-Mean" Effect. *Traffic Engineering and Control*, Vol. 22, No. 10, Oct. 1981, pp. 535–542.
8. *SAS/STAT User's Guide, Version 6* (fourth edition, Vol. 2, Chapter 29). SAS Institute, Inc., Cary, N.C., 1990.
9. P. McCullagh and J. A. Nelder. *Generalized Linear Models.* Chapman and Hall, New York, 1983.
10. D. Pickering, R. D. Hall, and M. Grimmer. *Accidents at Rural T-Junctions.* TRRL Research Report 65. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1986.
11. E. Hauer and B. N. Persaud. How To Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effects of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1987, pp. 131–140.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.