Pergamon

Accid. Anal. and Prev., Vol. 27, No. 3, pp. 371–389, 1995

0001-4575/95 $9.50 + .00

0001-4575(94)00078-6

EFFECT OF ROADWAY GEOMETRICS AND ENVIRONMENTAL FACTORS ON RURAL FREEWAY ACCIDENT FREQUENCIES

VENKATARAMAN SHANKAR,[1] FRED MANNERING,[1] and WOODROW BARFIELD[2]

[1]Department of Civil Engineering, 121 More Hall FX-10, [2]Department of Industrial Engineering, Mechanical Engineering Building FU-20, University of Washington, Seattle, WA 98195, U.S.A.

(*Accepted* 29 *October* 1994)

Abstract—This paper explores the frequency of occurrence of highway accidents on the basis of a multivariate analysis of roadway geometrics (e.g. horizontal and vertical alignments), weather, and other seasonal effects. Based on accident data collected in the field, a negative binomial model of overall accident frequencies is estimated along with models of the frequency of specific accident types. Interactions between weather and geometric variables are proposed as part of the model specifications. The results of the analysis uncover important determinants of accident frequency. By studying the relationship between weather and geometric elements, this paper offers insight into potential measures to counter the adverse effects of weather on highway sections with challenging geometrics.

Keywords—Negative binomial, Accident frequency

INTRODUCTION

Beginning in winter 1996, the Washington State Department of Transportation proposes to implement an Intelligent Transportation System (ITS) on Interstate 90 (I-90) to reduce the likelihood of vehicular accidents in the vicinity of Snoqualmie Pass (located about 50 kilometers east of Seattle; WA; see Fig. 1). This ITS system is to consist of variable message signs, variable speed-limit signs, and in-vehicle signing. The intent of the signing is to reduce accident risk by warning drivers of adverse weather conditions, ice on the roadway, and lane blockages caused by vehicle accidents and/or disablements. However, to evaluate the effectiveness of such a signing system, it is important first to develop an understanding of the factors that have historically contributed to the likelihood of an accident in the study area (i.e. before the signing system is implemented). With this in hand, a before-and-after statistical comparison of the data can be conducted and the effectiveness of the signing system can be assessed. The intent of this paper is to investigate and report the "before" conditions to isolate the factors that are contributing to accident likelihoods.*

*The study of historic data is important even if the evaluation of some future accident-reduction strategy is not being undertaken. This is because an analysis of historic accident data can provide valuable information relating to the nature of accidents and factors affecting their frequency and the likely effectiveness of alternative accident-mitigation strategies.

The section of Interstate 90 shown in Fig. 1 experiences a high number of vehicular accidents as a result of challenging roadway geometrics (i.e. small horizontal curve radii and steep grades) and adverse weather conditions. The climate in the vicinity of the Snoqualmie Pass summit is severe. At an elevation of over 900 m above sea level, the area receives an average of over 100 cm of rainfall and over 1700 cm of snowfall, annually. Snowfall occurs during every month except July and August. During a large portion of the year, residual snow and ice accumulated on the ground contribute to adverse driving conditions. Factors that contribute to the accidents include driver behavior, geometric characteristics (e.g. grade and curve radii), weather-related variables (e.g. rainfall and snowfall, intensity of snowfall and rainfall), interactions between geometrics and weather elements, and seasonal effects such as traffic volume, precipitation, and ambient temperature-related variations.

The intent of this paper is to focus on the nonbehavioral determinants of accident risk, specifically roadway geometrics and weather conditions. The paper begins with a review of previous research on accident frequencies in terms of their relationship to geometric and weather-related elements. On the basis of this review, we present an appropriate methodology to establish an explicit relationship between geometric and weather-related elements and accidents. This is followed by a description of available

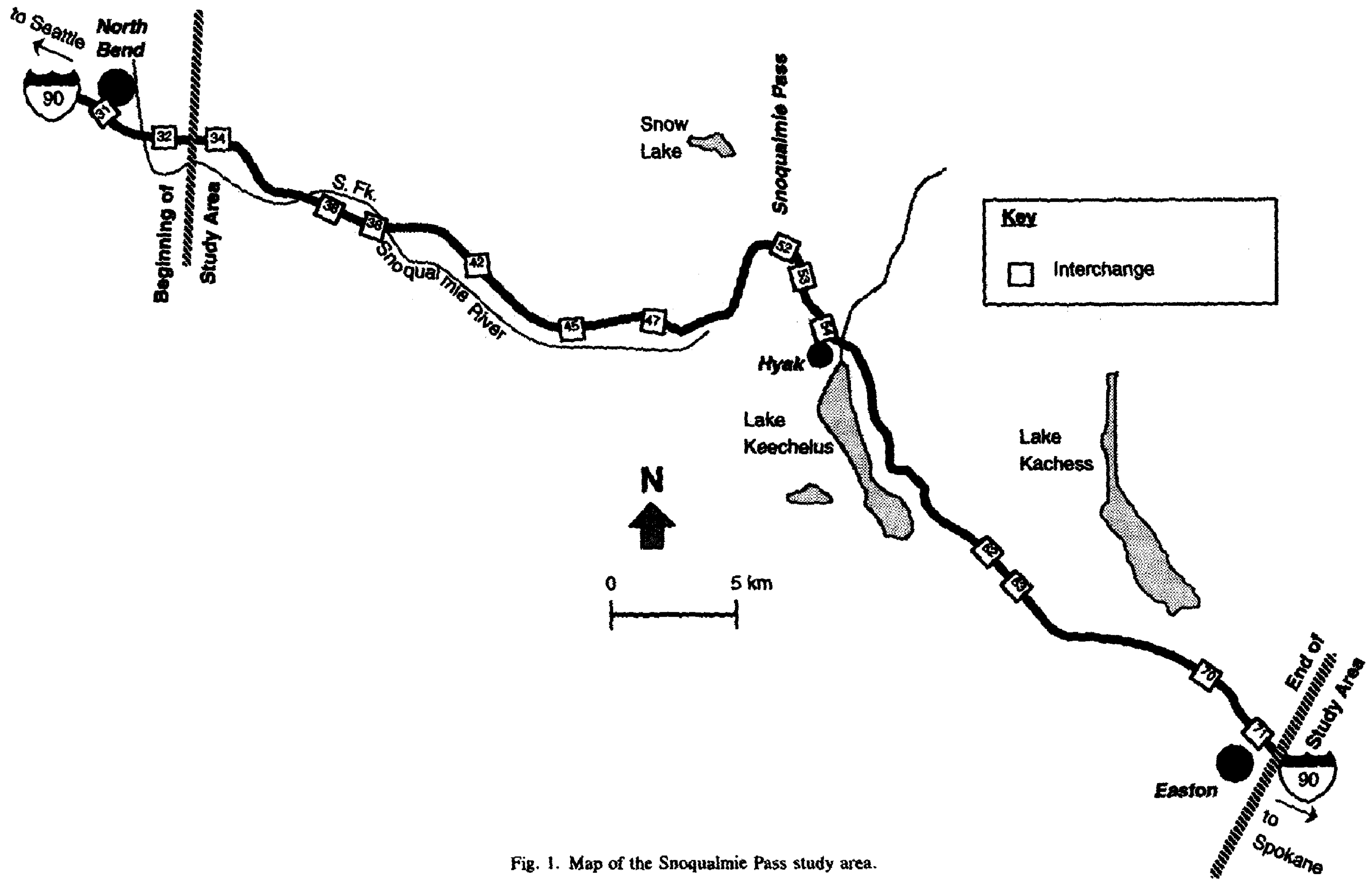

Fig. 1. Map of the Snoqualmie Pass study area.

data and a discussion of model estimation results. Finally, a summary of model findings and implications is presented.

PREVIOUS RESEARCH

Previous research, for the most part, has dealt with modeling relationships between accident occurrences and geometric elements. Examples of this include the work of Wong and Nicholson (1992). They observed that modifications to roadway geometrics were important because of the strong association between adverse geometric elements and high-accident locations. This association has been confirmed in studies by Boughton (1975), National Cooperative Highway Research Program (1978), and the Federal Highway Administration (1982). Other empirical relationships between vehicle accidents and highway geometrics have been studied through the use of statistical models to investigate accident involvement rate, accident probability, geometric design variables critical to safety, and the accident-reduction potential of geometric improvements (National Cooperative Highway Research Program 1978; Hammerslag, Roos, and Kwakernaak 1982; Okamoto and Koshi 1981; Miaou et al. 1991).

In terms of the relationship between accidents and weather elements, a number of important studies have been conducted (Ivey et al. 1981; Jovanis and Delleur 1981; Mori and Uematsu 1967; Snyder 1974). This past work studied the effect of rainfall and snowfall on accident occurrences and attempted to quantify the contribution of these environmental elements to increasing accident likelihoods. Other types of methodologies have also been applied to the problem of accident analysis. For example, risk-based approaches, applied to the prediction of wet-weather accidents, have also been documented (Brodsky and Hakkert 1988), and recently, the effect of winter pavement maintenance on accident rates has been investigated (Hanbali 1992). Finally, seasonal variations in weather elements coupled with corresponding variations in traffic volumes have been examined (Jones, Janssen, and Mannering 1991) in a multivariate framework.

Although past research work has provided insight into the effect of weather on accident rates and frequencies, efforts to investigate the interaction of weather and geometric elements and their consequent impact on accident likelihoods have been minimal. The study of such interactions is important, because it could shed light on the impact of weather on critical geometric design elements and serve as a guide in the design of roadway geometrics so as to minimize accident likelihoods in the presence of varying climatic conditions. This concept contrasts with present roadway geometric design practice, which applies a uniform nationwide standard in terms of assumed weather impacts on geometric design (Mannering and Kilareski 1990). Presumably, much could be gained by adjusting this standard to account for weather conditions that deviate greatly from the norm.

In addition to weather and geometrics, it may be argued that human factors contribute significantly to accident occurrences and hence warrant inclusion in the modeling effort. Previous research (Treat 1980; Sabey and Taylor 1980) indicates that in 95% of all traffic accidents human factors are involved either alone or in combination with other factors. However, other research (Massie, Campbell, and Blower 1993) tempers the criticism of research excluding human factors by pointing out that the human-factors approach ignores the problem associated with classifying collisions and their related causes, be they human or otherwise. The authors add that such an approach fails to address the issue of helping drivers avoid collisions. Identification of geometric and weather-related factors and their interrelationship can be used to assist the driver in reducing the chances of a collision by offsetting the ignorance factor caused by unanticipated changes in roadway geometrics and their interrelation with adverse weather conditions.

From a methodological perspective, attempts to model accident frequencies have varied from the use of least squares regression techniques to methods involving exponential distribution families including the Poisson and negative binomial models. Previous research on Poisson and least squares (Jovanis and Chang 1986; Joshua and Garber 1990; Miaou and Lum 1993) indicates the inappropriateness of least squares techniques to modeling of accident frequencies, and recommends the employment of the Poisson distribution. The Poisson distribution, however, suffers from an important limitation, namely that the mean and variance are constrained to be equal. Overdispersion (variance greater than the mean) or underdispersion (variance less than the mean) of data violates this constraint and leads to biased coefficient estimates. A more general distribution, such as the negative binomial, has been employed in such situations (Engel 1984; Lawless 1987; Manton, Woodbury and Stallard 1981) to relax this constraint. Negative binomial distributions have been employed frequently in physics, medical sciences, and marketing. Documented use of the negative binomial distribution in the field of traffic engineering includes applications in trip generation (Frisbie 1980) and transportation economics (Hellerstein 1991). In terms of applying the negative binomial distribution to model accident occurrences, research has been conducted on accident proneness (Bates and Neyman 1952), accident migration (Maher 1987; 1990), accident "blackspots" identification (Senn and Collie 1988) and accident frequencies (Miaou 1994; Maher 1991; Poch and Mannering 1994).

The body of extant literature provides important methodological direction for our study of the interrelationship between roadway geometrics and weather and accident frequencies. Details of this methodological direction are discussed in the following section.

METHODOLOGY

Count data are often modeled by assuming Poisson distributions (Cameron and Trivedi 1986). The Poisson distribution is a useful starting point because (i) it lends itself well to the modeling of count data by virtue of its discrete, nonnegative integer-distribution characteristics and (ii) can be generalized to more flexible distributional forms. In terms of accident frequencies, in this study we will focus on modeling the number of accidents occurring on a specified section of roadway in a one-month time period. In such a case, the Poisson distribution gives

$$P(n_{ij}) = e^{-\lambda_{ij}}\lambda_{ij}^{n_{ij}}/n_{ij}! \tag{1}$$

where $P(n_{ij})$ is the probability of n accidents occurring on roadway section i in month j and λ_{ij} is the expected number of accidents on roadway section i in month j. Given a vector of geometric, traffic and weather data, λ_{ij} can be estimated by the equation,

$$\ln \lambda_{ij} = \mathbf{X}_{ij}\beta \tag{2}$$

where $\mathbf{X}$ is a vector of geometric, traffic, and weather data for roadway section i in month j and β is a vector of estimable coefficients. As mentioned in our review of previous research, the Poisson distribution constrains the mean and variance to be equal, (i.e. $E[n_j] = \mathrm{Var}[n_{ij}]$). As previously mentioned, estimation using a Poisson distribution violating this assumption (i.e. when data are overdispersed or underdispersed) results in biased estimates of β. It is well known, based on the findings of many previous research efforts, that accident frequency data tend to be overdispersed, with the variance being significantly greater than the mean. Consequently, the Poisson distribution can lead to erroneous coefficient estimates and erroneous inferences can be drawn. To overcome this, the negative binomial distribution, which includes a gamma-distributed error term, is appropiate because it relaxes the Poisson's mean-variance equality constraint. The negative binomial model is derived by rewriting eqn 2 as,

$$\ln \lambda_{ij} = \mathbf{X}_{ij}\beta + \varepsilon_{ij} \tag{3}$$

where $\exp(\varepsilon_{ij})$ is a gamma-distributed error term. This results in the mean-variance relationship,

$$\mathrm{Var}[n_{ij}] = E[n_{ij}][1 + \alpha E[n_{ij}]] \tag{4}$$

If α is significantly different from zero, the data are over-dispersed or underdispersed. If α is equal to zero the negative binomial reduces to the Poisson distribution.

The resulting probability distribution under the negative binomial assumption is,

$$P(n_{ij}) = \frac{\Gamma(\theta + n_{ij})}{\Gamma(\theta)n_{ij}!}\, u_{ij}^{\theta}(1 - u_{ij})n_{ij} \tag{5}$$

where $u_{ij} = \theta/(\theta + \lambda_{ij})$, $\theta = 1/\alpha$, and $\Gamma(.)$ is a value of the gamma function. Estimation of λ_{ij} can be conducted through standard maximum likelihood (ML) procedures (see Greene 1993). Using eqn 5, the likelihood function (the product of probabilities) for the negative binomial is,

$$L(\lambda_{ij}) = \prod_{i=1}^{N}\prod_{j=1}^{T}\frac{\Gamma(\theta + n_{ij})}{\Gamma(\theta)n_{ij}!}\left[\frac{\theta}{\theta + \lambda_{ij}}\right]^{\theta}\left[\frac{\lambda_{ij}}{\theta + \lambda_{ij}}\right]^{n_{ij}} \tag{6}$$

where T is the last month of accident data and N is the total number of roadway sections. This function is maximized to obtain coefficient estimates for β and α.

Careful attention must be paid to the appropriateness of the negative binomial distribution in the case of overdispersed data. For example, eqn 3 may hold while the distribution of n_{ij} conditioned on $\mathbf{X}_{ij}$ may not be negative-binomial distributed. In such a case, the coefficient estimates will be consistent though less efficient than those for the correct distribution. Importantly, the asymptotic variance-covariance matrix will be incorrect and likely underestimated. However, in practice, this underestimation is not likely to affect substantive conclusions drawn from model estimation (see Lawless 1987).

In addition to maximum likelihood estimation procedures, other methods such as quasilikelihood, weighted least squares (McCullagh and Nelder 1983), moment estimation techniques (Breslow 1984), and regression-based estimation (Cameron and Trivedi 1986, 1990) are available. Examples of the application of the moment method and regression-based estimation in accident modeling indicates that these methods should be used with caution (Miaou 1994). Indications from statistical research on the estimation of α, the dispersion coefficient, suggest that for large samples ($N > 20$) the quasilikelihood and maximum likelihood methods perform best (Piegorsch 1990).

EMPIRICAL SETTING

The study area consists of a 61 km portion of I-90 located some 48 km east of Seattle (see Fig. 1). This portion of Interstate 90 generally consists of a three-lane (3.66 m lanes) cross-section, in each direction, with 3.05 m shoulders and a 104.6 kph speed limit. Virtually no variation in travel lane and shoulder widths exists in the study area.

Data from a number of sources were gathered over the period from January 1988 to May 1993. Precipitation data were assembled from the Desert Research Institute and Western Regional Climate Center and the geometric attributes of the roadway and accident data were obtained from the Washington State Department of Transportation. The available precipitation data consisted of information relating to monthly rainfall and snowfall including average monthly snowfall and rainfall, maximum daily snowfall and rainfall and number of snowy and rainy days per month. Three weather stations located at Snoqualmie Falls, Stevens Pass, and Cle Elum (all in Washington state near the study area) were used as the sources of climatic data. Weather data were assigned to sections based on their geographic proximity and elevation levels.*

Geometric characteristics included: number of horizontal curves, number of horizontal curves underdesigned (those curves with design speeds less than 112.6 kph, less than 96.5 kph, and less than 80.45 kph), maximum and minimum horizontal radii, number of vertical curves, and maximum and minimum grades.

With these data in hand, the issue of dividing the study area into manageable sections of roadway must be addressed. The existing literature addresses several important issues relating to roadway section length determination in a linear regression context (Okamoto and Koshi 1989). The findings of these studies show that great care must be taken in determining roadway section lengths because of two model estimation concerns: (i) the possibility of heteroskedasticity (i.e. error terms are not identically distributed) and (ii) the possibility of biased model coefficients. Heteroskedasticity, especially in the context of a negative binomial specification (as opposed to a Poisson specification), is an important issue due to the incorporation of the gamma-distributed error term.

The most popular alternatives for determining roadway section lengths are the use of fixed-length sections or homogeneous sections (i.e. sections with homogeneous geometric characteristics; see Miaou et al. 1991). With regard to homogeneous sections (both in terms of geometrics and weather), several important problems arise. One of these problems is that roadways with numerous horizontal curves and grades tend to produce sections that are less than 1 km in length (i.e. to ensure homogeneity in geometrics). This can result in locational error problems because accidents, in most states, are locationally reported to the nearest milepost (1.609 km). Potential bias resulting from such accident-reporting locational error is clearly undesirable.

Homogeneity of weather data presents a different problem. Weather data, by virtue of their geographic characteristics, usually encompass much larger areas, and if allowed to govern section lengths, are likely to result in long, geometrically diverse sections, thus violating geometric homogeneity.

Finally, the unequal length of sections that will result from the homogeneity requirement may exacerbate potential heteroskedasticity problems (i.e. unequal sample sizes; see Mannering 1995) and lead to a loss in estimation efficiency. The resulting increase in the standard errors of model coefficients could lead the analyst to draw erroneous inferences with regard to the effects of model covariates.

The disadvantages of using fixed-length sections, relative to homogeneous sections, are far less severe. In fact, most potential disadvantages can be overcome by accounting for the nonhomogeneity of geometric and weather-related variables by including detailed measures of the variability across sections in the model specification (e.g. number of curves, maximum grade, and number of underdesigned curves, and so on). If such data are available, there is little need to constrain the analysis to homogeneous sections. Moreover, fixed-length sections

*As a result, several contiguous sections shared the same weather information (this can be seen in Table 1). Shared weather data raises the issue of serial correlation of model error terms. Weather information shared by contiguous sections causes any shocks in data to propagate through sections common to those data, thereby causing spatial correlation. To date, the effect of such spatial correlation has not been specifically investigated in a count data model context. However, based on experiences in linear regression contexts, it can be reasonably assumed that spatial correlation could cause some loss of efficiency of parameter estimates. Studies have shown that, in most practical contexts, this is not a major concern (Mannering 1995).

may offer other advantages such as mitigating the effects of accident migration, which is a phenomenon involving the migration of accidents to a different portion of a hazardous roadway section after corrective measures have been taken on some other portion of the roadway (see Boyle and Wright 1984; McGuigan 1985; Maher 1987). If one were to use geometrically homogeneous sections, it would be exceedingly difficult to account for the effect that changes in accident likelihoods on one section would have on others (due to accident migration). However, the use of fixed-length, nonhomogeneous sections accounts for the possibility of accident migration, to some extent, because the migration across the homogeneous "subsections" that constitute the fixed length section is internalized.

As a result of the above discussion, the sections considered in this study were determined to be fixed, equal-length sections. Thus, accident frequencies and associated geometric and weather data were compiled along ten sections of equal length over the 61 km study area (i.e. each section is 6.1 kilometers in length). Accident frequencies and roadway geometrics for both roadway directions (eastbound and westbound) were used.* A total of 2,225 reported accidents occurred in the study area between January 1988 and May 1993.† Accidents were sorted by year and month and integrated with geometric and monthly weather data into one database. The consolidated database, after accounting for some missing weather data (which resulted when weather stations were not functioning due to mechanical failures) consisted of 464 observations with some sections experiencing zero accidents in some months.‡ The implicit specification of accident frequency per month as the dependent variable allows the modeling of seasonal variations in traffic volumes, ambient temperature, and other environmental data such as daylight duration.

Table 1 summarizes the averages of the variables measured in this study. Mean section accident frequency per month was 3.26 (Fig. 2 shows average per-month accident frequencies by section) with an observed monthly minimum of zero and maximum of 28 (the observed monthly variance was 16.32). Other values worthy of note include the high number of horizontal curves on the 10 sections. The twelve horizontal curves in section 6 (sections 6 and 7 are near the summit) suggests complex geometrics in the area (i.e. about two horizontal curves per km). Also the average monthly snowfall, observed to be 145.78 cm in sections 5–8, is quite high and reflects the severe climate resulting from the relatively high elevation.

MODEL ESTIMATION

The negative binomial estimation of section-accident frequencies is presented in Table 2. This table shows that all variables are of plausible sign with reasonably high statistical significance.

Table 2 shows that the majority of independent variables specified in this model positively affect accident frequency, indicating a likelihood in increase in frequency with increasing variable values. The number of curves variable provides some insight into potential geometric hazards. The number of curves with design speeds between 96.5 kph and 128.7 kph appears to have a greater effect (0.117) than those designed under 96.5 kph (0.046). The higher coefficient value for higher design-speed curves is likely capturing the tendency of drivers to slow down for curves with low design speeds due to a combination of the visual effect of the curve and speed reduction signs usually found in those locations.

Grade appears to have a strong positive effect on accident frequency, although in a stepwise, as opposed to a continuous, manner. In comparison to those sections with grades less than 2%, those with maximum grades exceeding 2% will experience a significant increase in accident frequency. Intuitively, this captures the effect of speed differentials that play a significant role in accidental occurrences, although, to some extent, the presence of climbing lanes offsets the detrimental impact of grades, especially those impacts caused by slow-moving heavy vehicles. In the present context, however, a geometric variable accounting for climbing-lane effects was not found to be significant because there is little variation in this variable across sections. A review of the data showed that any vertical grade reason-

*Interstate 90 has divided cross-sections with different grades and horizontal curve attributes in 3 of the 10 study sections. By combining both east and west directions, we constrain the βs to be the same. An empirical test of this assumption revealed that this constraint is statistically valid.

†In this analysis we include only those accidents reported to the Washington State Highway Patrol (WSP). Although this section of highway is heavily patrolled by WSP, it is likely that some minor accidents are never reported.

‡Note that our data have repeated observations from the same section of roadway. That is, each section produces as many as 12 observations (corresponding to 12 months) per year. Such data raise the possibility of error-term correlation among observations produced by the same section, with observations from the same section sharing unobserved factors that may impact accident likelihoods (e.g. a scenic distraction). A likely consequence of such correlation is some loss in efficiency of coefficient estimates. However, research by Mannering and Winston (1991) indicates that the efficiency loss from this source is small, particularly if section-specific constants are included in the model specification (as will be the case in this study).

Table 1. Sample summary statistics (section averages)

Variable	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7	Section 8	Section 9	Section 10
Accident frequency (per month)	1.80	2.25	1.66	2.49	7.81	8.35	5.92	4.15	2.81	2.86
Number of horizontal curves with a design speed less than 112.6 kph	1	3	1	3	1	5	9	2	8	2
Number of horizontal curves with a design speed less than 96.5 kph	0	1	1	0	1	5	7	2	7	1
Number of horizontal curves with a design speed less than 80.5 kph	0	0	0	0	0	1	1	0	0	0
Number of horizontal curves in section	8	8	10	9	10	12	10	9	10	4
Maximum horizontal curve radius in section (m)	3030	3636	909	3030	1515	3030	695	1736	1736	1818
Minimum horizontal curve radius in section (m)	636	595	595	606	333	333	347	347	347	788
Number of vertical curves in section	7	8	9	10	8	5	16	7	15	5
Maximum grade in section (%)	5.00	3.00	1.76	3.63	5.29	4.22	2.00	2.60	3.83	5.00
Minimum grade in section (%)	0.03	0.27	0.14	0.46	3.29	0.67	0.43	0.08	0.20	0.74
Average monthly rainfall (cm)	12.53	12.53	12.53	12.53	21.75	21.75	21.75	21.75	4.83	4.83
Maximum daily rainfall in the month (cm)	2.73	2.73	2.73	2.73	5.28	5.28	5.28	5.28	1.40	1.40
Number of rainy days in the month	15.95	15.95	15.95	15.95	17.00	17.00	17.00	17.00	9.52	9.52
Average monthly snowfall (cm)	1.70	1.70	1.70	1.70	145.78	145.78	145.78	145.78	8.50	8.50
Maximum daily snowfall in the month (cm)	1.28	1.28	1.28	1.28	27.65	27.65	27.65	27.65	3.30	3.30
Number of snowy days in the month	0.20	0.20	0.20	0.20	10.31	10.31	10.31	10.31	1.24	1.24

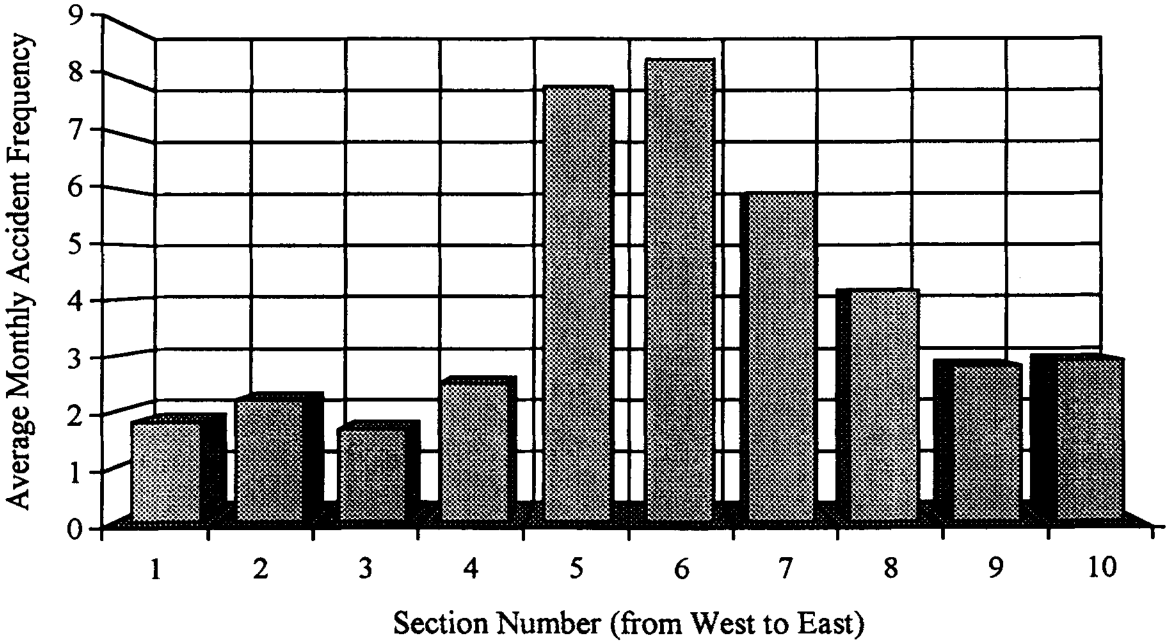

Fig. 2. Average monthly accident frequencies on the ten 6.1 km sections. (Sections numbered sequentially from west to east in the study area.).

Table 2. Negative binomial estimation results (total section accident frequency)

Variable	Estimated coefficient	t-statistic	p-value
Number of horizontal curves designed between 96.5 kph and 112.6 kph	0.117	2.437	0.015
Number of horizontal curves designed below 96.5 kph	0.046	2.205	0.027
Maximum grade in section indicator (1 if greater than 2%, 0 otherwise)	0.133	2.748	0.006
Maximum rainfall indicator (1 if greater than 2.54 centimeters on any given day in the month; 0 otherwise)	0.209	1.401	0.161
Number of rainy days in the month	0.018	1.975	0.048
Rainfall-curve interaction indicator (1 if maximum rainfall greater than 2.54 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	0.184	1.239	0.215
Maximum daily snowfall in the month	0.033	2.231	0.026
Snowfall-grade interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and grade greater than 2%; 0 otherwise)	0.291	1.930	0.053
Snowfall-curve interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	0.387	2.137	0.032
Section location indicator (1 if section number is 5, 6, 7 or 8; 0 otherwise)	−0.466	−1.812	0.070
Year of occurrence indicator (1 if 1988, 0 otherwise)	0.273	2.330	0.020
Year of occurrence indicator (1 if 1990; 0 otherwise)	−0.167	−1.410	0.159
α (dispersion coefficient)	0.418	8.463	0.000
Number of observations	464		
Log-likelihood at zero	−2193.39		
Log-likelihood at convergence	−970.93		
ρ^2	0.56		

ably long (longer than 2 km) and exceeding 2% had a climbing lane.

Maximum rainfall played a significant, positive role in accident occurrences. Employed as an indicator variable, it captures not only the effect of intensity of rainfall and potential hydroplaning of vehicles but also may be capturing the effects of exposure and pavement condition. For example, the pavement surface is likely to remain wet or icy during the night or early morning when daily rainfall exceeds 2.54 cm.

The number of rainy days played a significant, positive role in accident occurrences. This variable appears to capture exposure effects such as exposure to wet pavements and lower visibility effects. More interesting, given the fact that the Seattle area generally experiences intermittent rainfall throughout the year, drivers may be inclined to pay less attention to the risk of an accident during rainy weather. The number of rainy days variable could possibly be playing a surrogate role for increased accident risk arising from driver complacency.

Maximum daily snowfall intuitively captures the positive effect that snow plays in accident occurrences. Maximum snowfall exceeding 5.08 centimeters, employed as an indicator variable, appears to account for traction and lane-marking-related problems caused by increasing snow depth on the pavement. In combination with grades, as evidenced by the interaction term, it positively impacts accident frequency. This illustrates the dangerous combination of traction, lane-markings, and speed differentials. In addition, it also suggests that the effect of climbing lanes could likely be annulled by the obliteration of lane markings on snow-covered pavements. In the presence of underdesigned horizontal curves, the snowfall variable portrays a stronger effect than the grade interaction by virtue of its higher coefficient.

The section location indicator variable shows that the middle portion of the study corridor (sections 5, 6, 7, and 8, which include the summit and the immediate area surrounding it) is associated with lower accident rates with all other factors held constant.* This is likely the result of changes in driver behavior, with drivers becoming more cautious as they gain elevation and approach/depart from the Snoqualmie Pass Summit.

The year 1988 was found to affect accident frequencies positively. Although normal precipitation levels were observed during this year, the positive coefficient value captures unobserved effects such as unusually cold ambient temperatures resulting in ice-covered pavements and construction-related effects such as lane closures.†

The year 1990, in a similar manner, was specified as an indicator variable. The negative effect of this variable seems to account for some decrease in traffic volumes as well as extra caution used by drivers in the presence of adverse driving conditions created by abnormally high levels of precipitation that occurred during the year.

Finally, an examination of ρ^2 (0.52) for the model indicates a good statistical fit, while the dispersion coefficient, α, was estimated to be significantly different from zero ($t = 8.463$) indicating overdispersion of data, a phenomenon that cannot be handled by a Poisson distribution.‡

Other issues worthy of note in the estimation context pertain to the impact of weather-related variables on pertinent variables such as traffic volume, temperature, and daylight time. The high level of significance of the weather-related variables coupled with their interaction with geometric variables suggests that they capture seasonal trends in traffic volume, temperature, and daylight time as well.§ The significance of weather-related variables and their use as surrogates for traffic volumes is corroborated in previous research (Jones, Janssen, and Mannering 1991).

It should also be noted that we tried to include a variety of other interactions between two variables and among three or more variables in our model (e.g. rainfall exceeding 2.54 cm on any given day in the month, at least one horizontal curve with less than 96.5 kph design speed, and a grade greater than

*Note that this does not imply that these sections of the study area have lower overall accident rates (see Fig. 2). It indicates only that these sections have lower than expected accident rates when the accumulated effects of geometrics and weather have been taken into account.

†It should be noted that since I-90 is a captive corridor with few alternate routes to/from eastern Washington, construction activities did not cause significant decreases in traffic volumes in the study corridor.

‡It is interesting to note that several variables that were found to be significant in a Poisson specification of our model turned out to be insignificant under the negative binomial assumption. This occurred because the Poisson specification underestimated coefficient variances due to the inherent overdispersion of data. Variables found significant in the Poisson but insignificant in the negative binomial included average daily rainfall for the month, number of snowy days in the month, average snowfall in the month and curve radii.

§The absence of traffic volume, temperature, and other variables in the model raises the possibility of a model specification error (i.e. an omitted variables bias). To test for this we used a series of month indicator variables (e.g. January, February, etc.) and time of year variables (e.g. Winter, Summer, Spring, and Autumn). These variables are highly correlated with traffic volumes (and their seasonal variation), temperature variations, and other possible omitted variables. These indicator variables were all statistically insignificant, suggesting the possible omitted variables bias is not playing a significant role in our model.

Table 3. Accident frequency elasticity estimates

Elasticity with respect to	Value
Number of rainy days in the month	0.2624
Maximum daily snowfall in the month	0.1012
Number of horizontal curves designed between 96.5 kph and 112.6 kph	0.1346
Number of horizontal curves designed below 96.5 kph	0.0968

2%). However, all such variables produced statistically insignificant coefficients and were thus excluded from our final specification.

Elasticities of independent variables were estimated to determine the impact of those variables on accident frequency. Elasticities can be roughly interpreted as the percentage change in the average frequency of accidents λ_{ij} due to a one-percent change in the independent variable. Elasticity of accident frequency λ_{ij}, with respect to x_{ijk} (the kth independent variable for section i in month j) is defined as,

$$E^{\lambda_{ij}}_{x_{ijk}} = \frac{\partial \lambda_{ij}}{\lambda_{ij}} \times \frac{x_{ijk}}{\partial x_{ijk}} \tag{7}$$

Using eqn 3, eqn 7 gives,

$$E^{\lambda_{ij}}_{x_{ijk}} = \beta x_{ijk} \tag{8}$$

where β is the coefficient corresponding to covariate x_{ijk}.

With eqn 8, elasticities of λ_{ij} for each section observation were computed and sample averages were then estimated. Note that the elasticities of indicator variables are not meaningful, so only the elasticities of continuous variables are presented in Table 3.

Table 3 provides some interesting insights. For example, a one percent increase in the number of rainy days in a month causes a .26% increase in accident frequencies. Similarly, a one percent increase in the maximum daily snowfall in a month results in a .10% increase in accident frequencies. This suggests that, at least for these two variables, accident likelihoods may be more sensitive to rain than snow. However, these are not the only snow/rain variables in the model (i.e. indicator variables are not included in Table 3) and, as will be shown, indicator variables that show an interaction between climatic conditions and roadway geometrics have a large impact on accident frequencies.

Finally, it is also important to point out that all variables shown in Table 3 are inelastic (elasticity less than unity). This suggests that, while the effect of these variables on accident frequencies is statistically significant, they may be nearing thresholds where accident frequencies have relatively low sensitivity to any changes in the explanatory variables.

To gather some understanding of the relative importance of the indicator variables included in the model, a numerical computation can be performed to provide an idea of the relative effect of indicator variables on average accident frequency. This is accomplished by using a ratio of coefficients. For example, the average accident frequency λ_{ij} for section i in month j can be said to increase 14.0% ($e^{0.133}/e^{0}$), if the maximum grade on the section is raised to exceed 2%, assuming the error terms are independent of x_{ij} and remain unchanged. Table 4 presents the change in the average accident frequency caused by threshold changes in the indicator variables.

This table shows that snowfall-horizontal curve and snowfall-grade indicators have a large effect on accident frequencies (47.3% and 33.8%, respectively). Rainfall indicators also strongly impact accident frequencies. These findings underscore the importance of accounting for weather/geometric interactions when assessing accident likelihoods.

IMPLICATIONS OF FINDINGS

The proposed model accounts for plausible and intuitive geometric and weather-related factors that influence accident frequencies. Specifically, the model offers insight into the combined effect of weather and geometric elements through interaction variables. The employment of indicator-type interaction variables allows designers to determine thresholds of geometric variables, such as maximum grade, beyond which their interaction begins significantly to affect accident frequencies.

The findings of this paper have significant implications for highway design standards. Current standards establish geometric design criteria on the basis of pavement-tire interactions on wet pavements. Our findings show that, in order to reduce accident likelihoods in areas that frequently experience adverse weather, the basis of establishing design criteria should be expanded beyond wet pavements. Specifically, great effort should be expended to avoid steep grades and horizontal curves with low design speeds in areas with adverse weather. Intuitively, this seems obvious, but our model provides a method of quantifying the impacts of these geometric characteristics. For example, for our study area, eliminating all horizontal curves with a design speed less than 96.5 kph on a roadway section that experiences at least 5.1 cm of snowfall one or more days

Table 4. Percentage change in accident frequencies due to indicator variables

Variable	Percentage change in mean accident frequency, λ_{ij}
Maximum grade in section indicator (1 if greater than 2%; 0 otherwise)	14.2
Maximum rainfall indicator (1 if greater than 2.54 centimeters on any given day in the month; 0 otherwise)	23.2 to 48.1*
Rainfall-curve interaction indicator (1 if maximum rainfall greater than 2.54 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	20.2
Snowfall-grade interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and grade greater than 2%; 0 otherwise)	33.8
Snowfall-curve interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	47.3
Section location indicator (1 if section number is 5, 6, 7 or 8; 0 otherwise)	−32.3
Year of occurrence indicator (1 if 1988, 0 otherwise)	31.4
Year of occurrence indicator (1 if 1990, 0 otherwise)	−15.4

*It is assumed that the change in one indicator variable will not be accompanied by a simultaneous change in any other variable with the exception of the interaction variables. For example, a change in the maximum monthly rainfall variable (to greater than 2.54 cm) does not affect the location dummy. However, by virtue of the monthly rainfall's interaction with horizontal curves, maximum rainfall could have an additive effect because it could influence two variables. This explains the percentage range shown for maximum rainfall.

in a month, can reduce the monthly accident frequency by 47.3% (see Table 4). Although our model results are site-specific, a more global application of our approach could serve as a basis for a cost-benefit analysis that could guide geometric design policy more effectively than the current wet-pavement approach.*

In terms of using our model results to evaluate the proposed use of variable message signs, variable speed-limit signs, and in-vehicle signing to warn drivers of weather- and traffic-related dangers in the study corridor, a comparison of our "before" estimation results (as shown in Table 2) can be made with similar estimations conducted using data collected after the signing system has been implemented. A series of likelihood ratio tests can be conducted to test for overall coefficient stability (between before-and-after data) and individual coefficient stability can be evaluated on estimated coefficients such as grade, snowfall, snow-grade interactions, snow-horizontal curve interactions, and the various rainfall variables (see Mannering, Murakami, and Kim 1994 for an application of such coefficient stability tests). The finding of statistically significant instability in coefficients could then be attributed to the variable message/speed-limit signing and the in-vehicle signing systems. Such an analysis is important because we are not simply testing for differences in before-and-after accident frequencies, but isolating the true causality of these differences by controlling for the complex interaction between geometrics and weather conditions. A more simplistic comparison of before-and-after data could easily lead to erroneous conclusions. For example, one could conclude that the signing system was ineffective in reducing accidents, but slight variations in weather between before-and-after data could be masking the system's effectiveness.

Finally, in addition to being able to determine whether or not the proposed signing system was effective in reducing accident frequencies, an analysis of changes in coefficient elasticities and the magnitudes of indicator variables will allow us more precisely to isolate the effectiveness of the signing system. For example, we may be able to state specifically that the signing system mitigated the adverse effects of high snowfall on grades greater than 2%. Such specificity is needed to make definitive statements regarding IVHS technologies.

*A more global application would be a negative binomial accident frequency model estimated with roadway sections that had widely varying geometric, traffic flow, and access-control characteristics (as opposed to the comparatively homogeneous sections used in this study). Such an application would likely find that variables such as lane widths, shoulder widths, peak-hour traffic volumes, daily traffic volumes, percentage of trucks, type of road indicators (i.e. urban freeway, rural arterial, etc.), and type of interchange/intersection indicators, play a role in the frequency of accident occurrence.

ACCIDENT FREQUENCY MODELS BY TYPE OF ACCIDENT

In addition to modeling overall accident frequency on highway sections (i.e. as demonstrated above) separate regressions of specific accident types will also provide valuable information. Separate regression models have the potential for providing greater explanatory power relative to a single, overall frequency model because separate models allow coefficient estimates to vary by the type of accident. Intuitively, such variation seems reasonable. For example, we would expect a steep grade to have a different effect on the likelihood of an overturn accident than it would on a rear-end accident.

To evaluate the impacts of geometrics and weather on specific accident types, models were estimated for accidents classified as sideswipes, rear-end, parked vehicles, fixed object, overturns, and same direction (all others). Estimation results for these models are presented in the Appendix. All models were negative binomial regressions except the overturn accident frequency model, which was a Poisson regression (i.e. statistically the overturn data were not overdispersed).

Interpretations of the results shown in Tables A1 through A6 are presented below. These interpretations are presented by type of variables.

Variable Number of horizontal curves designed below 96.5 kph
Finding Tendency to increase sideswipe and rear-end collisions but decrease parked vehicle as well as fixed object collisions

This finding suggests that drivers tend to slow down on underdesigned curves with design speeds less than 96.5 kph because of signing and visual perception and thus avoid severe accidents such as fixed-object collisions. However, this possible speed reduction does not appear to be enough to avoid sideswipe or rear-end accidents caused by lane violations or speed differentials due to braking on the curve. In addition, it must be noted that parked-vehicle collisions tend to decrease because drivers are likely to avoid parking on the shoulder (for purposes such as chaining) on a tight curve because they perceive greater safety hazards at such locations.

Variable Number of horizontal curves designed between 96.5 kph and 112.6 kph
Finding Tendency to increase same direction (all others) and fixed object collisions

This finding suggests that curves underdesigned between 96.5 kph and 112.6 kph do not create the visual impact on drivers to decrease speeds. The result is an increase in both lane violations (resulting in vehicular collisions) and vehicles running off the roadway and colliding with fixed objects. From a severity viewpoint, fixed-object collisions are more likely to result in serious injuries than vehicular collisions in the same direction. Consideration should then be given to upgrading marginally underdesigned curves (96.5 kph to 112.6 kph) if fixed-object collisions show increasing trends at certain locations.

Variable Number of horizontal curves in section
Finding Tendency to increase same direction (all others) and fixed-object collisions

This finding suggests two separate phenomena. Vehicular collisions in the same direction tend to increase on sections as the number of horizontal curves increase, because speeds on curves do not decrease enough to avoid lane violations. The fact that fixed-object collisions tend to increase with the total number of curves in a section indicates the increased likelihood of fixed objects, such as guardrails, being present on sections with more curves. The presence of such objects prevents a more severe type of accident, such as a vehicle overturn, from occurring. The caveat stemming from this finding is that it is preferable to design longer but fewer horizontal curves where the terrain makes construction of straight sections impossible.

Variable Average spacing of horizontal curves in section
Finding Tendency to increase overturn collisions

This finding uncovers an effect of roadway geometrics that was not distinguishable in the overall accident frequency model. The very significant t-statistic (4.636) indicates the significant effect that spacing of horizontal curves in a section has on driver speeds. Intuitively, if curves are spaced farther apart, vehicular speeds are likely to climb as a result of lower caution exhibited by drivers. Consequently, there is a greater risk of an overturn if curves are spaced farther apart in a section. Careful attention should be paid to the application of corrective measures in this regard. The obvious interpretation is to decrease the spacing of curves to decrease the frequency of overturn accidents; however, it would appear counterintuitive to locate curves physically nearer as a countermeasure. The surrogate action is to place more advance warning signs in sections with longer curve spacing. By placing more advance warning signs and strategically locating them, the spacing of curves in the driver's mind is subliminally altered.

Variable Lowest horizontal curve radius in section
Finding Tendency to increase sideswipe collisions and decrease overturn collisions
The lowest horizontal radius suggests the type of terrain where the section is located. Sections with longer minimum radii could lull the driver into lane violations that result in sideswipes. However, low-radii curves are usually associated with winding sections of highway, which decrease the likelihood of gathering sufficient speed for an overturn accident.

Variable Maximum grade in section
Finding Tendency to increase rear-end and same-direction (all others) collisions
This finding suggests several processes stemming from the presence of grades in a section. Between any two sections, the section with the steeper maximum upgrade will experience a greater number of rear-end and other same-direction accidents. In addition, rear-end accidents will increase substantially if the maximum grade exceeds 2% in that section. Both effects are explained by speed differentials occurring due to the impact of grades. The impact of grades is reversed in the presence of downgrades. Between any two sections, the section with the steeper maximum downgrade will experience fewer rear-end and other same-direction accidents, presumably from lower speed differentials. Much of the effect of the higher braking distance on downgrades appears to be offset by the visual impact of brake lights warning drivers of the potential slowing of vehicles ahead. In contrast, drivers are unlikely to use brakes on an upgrade which eliminates a critical warning sign of speed reductions.

Variable Maximum rainfall on any given day in the month
Finding Tendency to increase sideswipe, parked-vehicle, fixed-object, and overturn collisions and decrease rear-end collisions
This result reflects some interesting phenomena affecting driver behavior and the driving task. Accidents resulting from the loss of steering control, such as lane violations and running-off-the-roadway, are expected to increase in occurrence with increases in maximum daily rainfall. Maximum rainfall indicates the intensity of rainfall and how that results in water puddles forming in wheel ruts in the pavement. The presence of such puddles contributes to vehicle hydroplaning and also excessive lateral drag resulting in lane violations and off-roadway accidents. On the other hand, as intensity of rainfall increases, visibility decreases and drivers maintain greater headways paying more attention to the driving task. Much of this attention is focused on the vehicle ahead and quite likely much less is paid to the area of peripheral vision. This overcompensation on vehicle headways reduces rear-end accident risks but increases other accident types.

Variable Average daily rainfall in the month
Finding Tendency to increase rear-end collisions
This likely is an outgrowth of a seasonal effect that is descriptive of pavement condition. As opposed to maximum rainfall on any given day in the month, this variable captures the loss of traction due to wet pavements. An increase in average daily rainfall is indicative of a more prolonged wet-month weather effect. Drivers are less likely to pay attention to prolonged effects as opposed to short-term effects such as thunderstorms. In addition, as mentioned in the discussion of the overall model, the Seattle area receives rainfall for a large portion of the year on an intermittent basis. Drivers in the region may be less likely to acknowledge the hazards of wet pavements.

Variable Number of rainy days in the month
Finding Tendency to decrease sideswipe and rear-end collisions and increased fixed-object collisions
It is speculated that the findings relating to this variable can be partially attributed to drivers reducing their speed in the presence of other vehicles during rainy periods (so much so that some types of vehicle collisions actually decrease). The positive coefficient for fixed-object collisions suggests that this possibility of cautious driving behavior during rainy conditions does not transfer to all driving situations.

Variable Maximum snowfall on any given day in the month
Finding Tendency to increase rear-end, same-direction (all others), parked-vehicle, and fixed-object collisions
This finding illustrates a number of consequences associated with intensity of snowfall. Loss of traction, reduced visibility, and obliteration of lane markings act individually or in combination to increase the likelihood of accident types such as rear-end and other collisions in the same direction as well as collisions with parked vehicles and fixed objects.

Variable Number of snowy days in the month
Finding Tendency to increase sideswipe, fixed-object, and overturn collisions

This variable could be capturing a number of effects. For example, the loss of traction, obliteration of lane markings, and seasonal trends in weather and temperature could all be reflected in the significance of this variable.

Variable Snowfall-grade interaction
Finding Tendency to decrease sideswipe, rear-end, other collisions in the same and parked-vehicle collisions

The coefficients of the interaction between snowfall and grade on rear-end accidents, other vehicular collisions in the same direction, and parked-vehicle accidents (as shown in tables A2, A3, and A4) appear counterintuitive. However, a closer examination of the estimation results shown in tables A2, A3, and A4 indicates that the net effect of snowfall and grade is to increase the frequency of these accident types, a conclusion also drawn from the positive coefficient of the snowfall-grade interaction variable for sideswipe collisions. In order to illustrate the net positive effect of snowfall-grade interaction on rear-end collisions, it is observed in Table A2 that the coefficient of the maximum daily snowfall variable is 3.468, which implies that when the daily maximum in any given month exceeds 5.1 cm rear-end accident frequencies are expected to increase 32-fold ($e^{3.468}$). In the presence of a significant interaction with grade on the section, (i.e. when maximum grade in the section exceeds 2%) this 32-fold compounding effect is tempered to 4.5-fold ($e^{(3.468-1.964)}$) by the negative coefficient (-1.964) of the interaction between snowfall and grade. This is a very intuitive occurrence indicating that the compounding effect of snowfall is not as severe when grades exceed 2%, quite possibly due to the presence of climbing lanes on upgrades and driver caution on downgrades. Examination of the snowfall-grade coefficient in other relevant accident types indicates a similar pattern. Corrective action then appears to be the construction of climbing lanes in areas where snowfall intensity is severe (exceeding 5.1 cm a day) and grades exceed 2%.

Variable Snowfall-curve interaction
Finding Tendency to increase rear-end, other collisions in the same direction, and overturn collisions

This finding is similar to those described previously for the interaction between snowfall and grade. The net effect of snowfall on curves is tempered in the presence of underdesigned curves (<112.6 kph), presumably due to warning signs and driver caution. Corrective action to mitigate the impact of snowfall appears to be the installation of warning signs in advance of underdesigned curves advising drivers of poor traction and slower speeds.

Variable Rainfall-curve interaction
Finding Tendency to increase rear-end and other collisions in the same direction and decrease fixed-object and overturn collisions

It is likely that this variable is capturing complex interactions among roadway and geometric conditions and driver behavior. To be able to speculate further on the nature of these findings, additional data on other roadway types (e.g. nonfreeways) is necessary. This would allow us to isolate the effect of rainfall-curve interactions by providing greater variance in the data.

Variable Spring/summer month indicator
Finding Tendency to decrease same-direction (all others) parked-vehicle and fixed-object collisions

This indicates primarily the effects of seasonal trends such as daylight duration and ambient temperature. These are important determinants of accidents such as vehicular collisions in the same direction and collisions with parked vehicles and fixed objects. The finding illustrates the impact of pavement conditions, such as black ice, as well as visibility. It should be noted that the "spring/summer" indicator was found significantly to decrease accident frequencies in spite of increased exposure due to the higher traffic volumes typically observed during the spring and summer months.

Variable Year of occurrence indicator
Finding Tendency to increase all accident types

This finding indicates that some unobserved effects (e.g. ice accumulation on the pavement and within-day temperature variations) were more severe during the subject year than usual, thus tending to increase the likelihood of an accident.

Variable Section location indicator
Finding Tendency to increase rear-end and overturn collisions, but decrease fixed object collisions

Section location indicators capture unobserved factors attributable to specific locations within the corridor. Such unobserved factors could include visual distractions and other attributes of the highway section that are difficult to quantify.

In summary, note that the coefficient estimates presented in these tables show that there are significant differences in the magnitudes of the coefficient estimates (and in some cases the signs of the coefficient estimates) among different accident types. The results of these separate accident frequency models can be used in the same way as the overall accident frequency model. That is, to evaluate the effectiveness of highway design improvements and IVHS systems in reducing specific types of accidents.

CONCLUDING REMARKS

This paper presents an appropriate model to explore the frequency of occurrence of accidents on the basis of a multivariate analysis of geometrics and weather-related effects. A negative binomial model of overall accident frequency is estimated along with models of the frequency of specific accident types. Interactions between weather and geometric variables are proposed as part of the model specifications and the results of the analysis uncover important determinants of accident frequency. By accounting for interactions between weather and geometric elements, this paper offers insight into possible strategies that could be undertaken to counter the adverse effects of weather. This paper also presents an important basis for a comprehensive before-and-after analysis of the effectiveness of safety improvements (e.g. ITS). In particular, the approach presented herein can be used to evaluate thoroughly the safety impacts of variable-message/speed-limit signs, in-vehicle units, and other ITS technologies. Such evaluations will serve as a cornerstone to justify future ITS expenditures.

REFERENCES

Bates, G. E.; Neyman, J. Contributions to the theory of accident proneness: I. An optimistic model of the correlation between light and severe accidents. University of California Public Statistics 1:215–254; 1952.

Breslow, N. Extra-Poisson variation in log-linear models. Applied Statistics 33:38–44; 1984.

Boughton, C. J. Accidents and geometric design research report ARR44. Vermont South, Victoria, Australia: Australian Road Research Board; 1975.

Brodsky, H.; Hakkert, A. S. Risk of a road accident in rainy weather. Accident Analysis and Prevention 20:161–176; 1988.

Boyle, A. J.; Wright, C. C. Accident "migration" after remedial treatment at accident blackspots. Traffic Engineering and Control 25:260–267; 1984.

Cameron, A. C.; Trivedi, P. K. Econometric models based on count data: Comparison and applications of some estimators and tests. Journal of Applied Econometrics 1:29–53; 1986.

Cameron, A. C.; Trivedi, P. K. Regression-based tests for overdispersion in the Poisson model. Journal of Econometrics 46:347–364; 1990.

Engel, J. Models for response data showing extra-Poisson variation. Statistical Neerlandica 38:159–167; 1984.

Federal Highway Administration. Synthesis of safety research related to traffic control and roadway elements. Report TS-82-232. Washington, DC: Federal Highway Administration, Office of Research, Development and Technology; 1982.

Frisbie, G. A. Ehrenberg's negative binomial model applied to grocery store trips. Journal of Marketing Research 17:385–390; 1980.

Greene, W. Econometric analysis. New York, NY: Macmillan Publishing; 1993.

Hammerslag, R.; Roos, J. P.; Kwakernaak, M. Analysis of accidents in traffic situations by means of multiproportional weighted Poisson model. Transportation Research Record 847:29–36; 1982.

Hanbali, R. M. Influence of winter road maintenance on traffic accident rates. Ph.D. dissertation. Milwaukee, WI: Marquette University; 1992.

Hellerstein, D. M. Using count data models in travel cost analysis with aggregate data. American Journal of Agricultural Economics 73:860–867; 1991.

Ivey, D.; Griffin, L. I.; Newton, T. H.; Lytten, R. L.; Hankins, K. C. Predicting wet weather accidents. Accident Analysis and Prevention 13:83–99; 1981.

Jones, B.; Janssen, L.; Mannering, F. Analysis of the frequency and duration of freeway accidents in Seattle. Accident Analysis and Prevention 23:239–255; 1991.

Joshua, S. C.; Garber, N. J. Estimating truck accident rate and involvements using linear and Poisson regression models. Transportation Planning and Technology 15:41–58; 1990.

Jovanis, P.; Chang, H. L. Modeling the relationship of accidents to miles traveled. Transportation Research Record 1068:42–51; 1986.

Jovanis, P.; Delleur, J. Exposure-based analysis of motor vehicle accidents. Accident Analysis and Prevention 13:83–99; 1981.

Lawless, J. F. Negative binomial and mixed Poisson regression. The Canadian Journal of Statistics 15:209–225; 1987.

Maher, M. J. Accident migration—a statistical explanation? Traffic Engineering and Control 28:480–483; 1987.

Maher, M. J. A bivariate negative binomial model to explain traffic accident migration. Accident Analysis and Prevention 22:487–498; 1990.

Maher, M. J. A new bivariate negative binomial model for accident frequencies. Traffic Engineering and Control 32:422–425; 1991.

Mannering, F. L. Modeling driver decision making: A review of methodological alternatives. Forthcoming in: Woodrow Barfield and Tom Dingus, eds. Human factors in intelligent vehicle highway systems. Hillsdale, NJ: Lawrence Erlbaum; 1995.

Mannering, F. L.; Kilareski, W. Principles of highway engineering and traffic analysis, New York, NY: John Wiley & Sons; 1990.

Mannering, F. L.; Murakami, E.; Kim, S. G. Models of travelers' activity choice and home-stay duration: Analysis of functional form and temporal stability. Transportation 21:371–392; 1994.

Mannering, F. L.; Winston, C. Brand loyalty and the decline of American automobile firms. Brookings Papers on Economic Activity: Microeconomics 3:67–114; 1991.

Manton, K. G.; Woodbury, M. A.; Stallard, E. A variance components approach to categorical data models with heterogeneous cell population: Analysis of spatial gradients in lung cancer mortality rates in North Carolina counties. Biometrics 37:259–269; 1981.
Massie, D. L.; Campbell, K. L.; Blower, D. F. Development of a collision typology for evaluation of collision avoidance strategies. Accident Analysis and Prevention 25:241–257; 1993.
McCullagh, P.; Nelder, J. A. Generalized linear models. London: Chapman and Hall; 1983.
McGuigan, D. R. D. Accident "migration"—or a flight of fancy? Traffic Engineering and Control 26:229–233; 1985.
Miaou, S. P. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis and Prevention 26:471–482; 1994.
Miaou, S. P.; Lum, H. Modeling vehicle accidents and highway geometric design relationships. Accident Analysis and Prevention 25:689–709; 1993.
Miaou, S. P.; Hu, P. S.; Wright, T.; Davis, S. C.; Rathi, A. K. Development of relationships between truck accidents and highway geometric design: Phase I. Technical memorandum prepared by the Oak Ridge National Laboratory. Washington, DC: Federal Highway Administration; March and November 1991.
Mori, H.; Uematsu, T. Statistical analysis of road accidents as related to traffic environmental factors. NRIPS Report 1:13–29; 1967 (in Japanese).
National Cooperative Highway Research Program. Cost and safety effectiveness of highway design elements. NCHRP report 197. Washington, DC: Transportation Research Board; 1978.
Okamoto, H.; Koshi, M. A method to cope with the random errors of observed accident rates in regression analysis. Accident Analysis and Prevention 21:317–332; 1989.
Piegorsch, W. W. Maximum likelihood estimation of the negative binomial dispersion coefficient. Biometrics 46:863–867; 1990.
Poch, M.; Mannering, F. L. Analysis of the effect of roadway geometrics and traffic related elements on the frequency of intersection accidents. Working Paper. Seattle, WA: University of Washington; 1994.
Sabey, B.; Taylor, H. The known risks we run: The highway. Supplementary report SR 567. Crowthorne, United Kingdom: Transport and Road Research Laboratory; 1980.
Senn, S. J.; Collie, G. S. Accident blackspots and the bivariate negative binomial. Traffic Engineering and Control 29:168–169; 1988.
Snyder, J. F. Environmental determinants of traffic accidents: An alternate model. Transportation Research Record 486:11–18; 1974.
Treat, J. R. A study of precrash factors involved in traffic accidents. Research review. October 6–November 1. Ann Arbor, MI: University of Michigan, Highway Safety Research Institute; 1980.
Wong, Y. D.; Nicholson, A. Driver behavior at horizontal curves: risk compensation and the margin of safety. Accident Analysis and Prevention 24:425–436; 1992.

APPENDIX A

Accident Frequency Model Estimation Results by Accident Type

Table A1. Negative binomial estimation results monthly section "sideswipe" accident frequency

Variable	Estimated coefficient	t-statistic	p-value
Constant	−2.772	−4.011	0.000
Number of horizontal curves designed below 96.5 kph	0.102	1.977	0.048
Lowest horizontal curve radius in section (meters)	0.01027	1.104	0.269
Number of rainy days in the month	−0.019	−1.132	0.258
Maximum rainfall indicator (1 if greater than 2.54 centimeters on any given day in the month; 0 otherwise)	0.959	3.910	0.000
Number of snowy days in the month	0.029	1.290	0.197
Snowfall-grade interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and grade greater than 2%; 0 otherwise)	0.930	3.869	0.000
Year of occurrence indicator (1 if 1988; 0 otherwise)	0.483	2.013	0.044
α (dispersion coefficient)	0.396	1.420	0.155

Number of observations	464
Log-likelihood at zero	−498.78
Log-likelihood at convergence	−298.14
ρ^2	0.40

Table A2. Negative binomial estimation results monthly section "rear-end" accident frequency

Variable	Estimated coefficient	t-statistic	p-value
Constant	−4.368	−5.346	0.000
Number of horizontal curves designed below 96.5 kph	0.080	1.679	0.093
Maximum grade in section	0.310	2.211	0.027
Maximum grade in section indicator (1 if greater than 2%; 0 otherwise)	1.271	1.732	0.083
Maximum rainfall on any given day in the month	−0.381	−1.902	0.057
Number of rainy days in the month	−0.048	−1.741	0.082
Average daily rainfall in any given month	0.149	2.324	0.020
Rainfall-curve interaction indicator (1 if maximum rainfall greater than 2.54 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	0.983	3.309	0.001
Maximum daily snowfall in the month (1 if maximum snowfall greater than on 5.1 centimeters on any given day in the month)	3.468	3.215	0.001
Snowfall-grade interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and grade greater than 2%; 0 otherwise)	−1.964	−2.382	0.017
Snowfall-curve interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	−1.707	−2.262	0.023
Section location indicator (1 if section number is 5, 6, 7 or 8, 0 otherwise)	0.815	2.408	0.016
Year of occurrence indicator (1 if 1988; 0 otherwise)	0.747	2.797	0.005
Year of occurrence indicator (1 if 1989; 0 otherwise)	0.762	2.908	0.004
α (dispersion coefficient)	0.910	3.023	0.002

Number of observations	464
Log-likelihood at zero	−544.59
Log-likelihood at convergence	−310.84
ρ^2	0.43

Table A3. Negative binomial estimation results monthly section "parked vehicle" accident frequency

Variable	Estimated coefficient	t-statistic	p-value
Constant	−3.290	−5.523	0.000
Number of horizontal curves designed below 96.5 kph	−0.167	−1.741	0.082
Maximum rainfall indicator (1 if greater than 2.54 centimeters on any given day in the month; 0 otherwise)	0.906	2.274	0.023
Maximum daily snowfall (1 if greater than 5.1 centimeters on any given day in the month; 0 otherwise)	2.500	3.705	0.000
Snowfall-grade interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and grade greater than 2%; 0 otherwise)	−1.381	−2.294	0.022
Year of occurrence indicator (1 if 1988; 0 otherwise)	0.629	1.385	0.166
Year of occurrence indicator (1 if 1989; 0 otherwise)	1.273	3.179	0.001
Spring/summer month indicator (1 if April, May, June, July or August; 0 otherwise)	−0.819	−1.607	0.108
α (dispersion coefficient)	1.505	2.375	0.018

Number of observations	464
Log-likelihood at zero	−487.54
Log-likelihood at convergence	−177.51
ρ^2	0.64

Table A4. Negative binomial estimation results monthly section "fixed object" accident frequency

Variable	Estimated coefficient	t-statistic	p-value
Constant	−2.156	−5.515	0.000
Number of horizontal curves designed between 96.5 kph and 112.6 kph	0.154	1.957	0.050
Number of horizontal curves designed below 96.5 kph	−0.130	−2.737	0.006
Number of horizontal curves in section	0.285	5.032	0.000
Maximum rainfall indicator (1 if greater than 2.54 centimeters on any given day in the month; 0 otherwise)	0.423	1.932	0.053
Number of rainy days in the month	0.023	1.821	0.069
Rainfall-curve interaction indicator (1 if maximum rainfall greater than 2.54 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	−0.507	−2.144	0.032
Maximum snowfall indicator (1 if greater than 5.1 centimeters on any given day in the month; 0 otherwise)	0.654	3.198	0.001
Number of snowy days in the month	0.050	2.604	0.009
Section location indicator (1 if section number is 1, 2, 3 or 4; 0 otherwise)	−1.431	−4.349	0.000
Section location indicator (1 if section number is 5, 6, 7, or 8; 0 otherwise)	−1.017	−3.257	0.001
Year of occurrence indicator (1 if 1988; 0 otherwise)	0.283	2.206	0.027
Spring/summer month indicator (1 if April, May, June, July or August; 0 otherwise)	−0.294	−2.081	0.037
α (dispersion coefficient)	0.282	3.078	0.002

Number of observations	464
Log-likelihood at zero	−845.42
Log-likelihood at convergence	−610.79
ρ^2	0.28

Table A5. Poisson estimation results monthly section "overturn" accident frequency

Variable	Estimated coefficient	t-statistic	p-value
Constant	−3.288	−7.961	0.000
Average spacing of horizontal curves in section (meters)	0.00784	4.636	0.000
Lowest horizontal curve radius in section (meters)	−0.00461	−2.857	0.004
Maximum rainfall indicator (1 if greater than 2.54 centimeters on any given day in the month; 0 otherwise)	0.692	3.030	0.002
Rainfall-curve interaction indicator (1 if maximum rainfall greater than 2.54 centimeters on any given day in the month and at least one horizontal curve has a design speed between 96.5 kph and 112.6 kilometers per hour; 0 otherwise)	−0.727	−2.952	0.003
Number of snowy days in the month	0.039	2.264	0.023
Snowfall-curve interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and at least one horizontal curve has a design speed between 96.5 kph and 112.6 kilometers per hour; 0 otherwise)	0.970	4.771	0.000
Section location indicator (1 if section number is 1, 2, 3 or 4; 0 otherwise)	2.260	4.119	0.000
Year of occurrence indicator (1 if 1988, 0 otherwise)	0.465	2.868	0.004

Number of observations	464
Log-likelihood at zero	−509.17
Log-likelihood at convergence	−368.75
ρ^2	0.28

Table A6. Negative binomial estimation results monthly section "same direction (all others)" accident frequency

Variable	Estimated coefficient	t-statistic	p-value
Constant	−4.007	−7.819	0.000
Number of horizontal curves designed between 96.5 kph and 112.6 kph	0.471	3.381	0.001
Maximum grade in section	0.344	2.939	0.003
Rainfall-curve interaction indicator (1 if maximum rainfall greater than 2.54 centimeters on any given day in the month and at least one horizontal curve has a design speed less than 96.5 kph; 0 otherwise)	0.787	3.857	0.000
Maximum daily snowfall (1 if greater than 5.1 centimeters on any given day in the month)	2.923	7.128	0.000
Snowfall-grade interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and grade greater than 2%; 0 otherwise)	−0.901	−2.218	0.027
Snowfall-curve interaction indicator (1 if maximum snowfall greater than 5.1 centimeters on any given day in the month and at least one horizontal curve has a design speed between 96.5 kph and 112.6 kph; 0 otherwise)	−1.232	−3.338	0.001
Spring/summer month indicator (1 if April, May, June, July or August; 0 otherwise)	−0.805	−2.881	0.004
Year of occurrence indicator (1 if 1988, 0 otherwise)	0.577	2.253	0.024
Year of occurrence indicator (1 if 1989, 0 otherwise)	0.412	1.647	0.100
α (dispersion coefficient)	0.562	2.524	0.011

Number of observations	464
Log-likelihood at zero	−540.84
Log-likelihood at convergence	−307.04
ρ^2	0.43