

Portland State University

PDXScholar

---

Civil and Environmental Engineering Faculty  
Publications and Presentations

Civil and Environmental Engineering

---

2008

# Combining climate, crash, and highway data for improved ranking of speed and winter-weather related crash locations in Oregon

Christopher M. Monsere

*Portland State University*, [monsere@pdx.edu](mailto:monsere@pdx.edu)

Peter G. Bosa

*Portland State University*

Robert L. Bertini

*Portland State University*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/cengin\\_fac](https://pdxscholar.library.pdx.edu/cengin_fac)



Part of the [Civil and Environmental Engineering Commons](#), and the [Urban Studies and Planning Commons](#)

Let us know how access to this document benefits you.

---

## Citation Details

Monsere, Christopher M.; Bosa, Peter G.; and Bertini, Robert L., "Combining climate, crash, and highway data for improved ranking of speed and winter-weather related crash locations in Oregon" (2008). *Civil and Environmental Engineering Faculty Publications and Presentations*. 78.

[https://pdxscholar.library.pdx.edu/cengin\\_fac/78](https://pdxscholar.library.pdx.edu/cengin_fac/78)

This Post-Print is brought to you for free and open access. It has been accepted for inclusion in Civil and Environmental Engineering Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

## **Combining Climate, Crash, and Highway Data for Improved Ranking of Speed and Winter-Weather Related Crash Locations in Oregon**

Christopher M. Monsere, P.E.<sup>1</sup>, Peter G. Bosa<sup>2</sup>, Robert L. Bertini, P.E.<sup>3</sup>

**Abstract:** In recent years, the techniques for screening transportation networks to identify high crash locations have become more sophisticated. Many transportation agencies, however, lack sufficient data, either in timeliness, completeness, or accuracy to implement many of the recent advances. This paper presents the results of an empirical analysis of screening and ranking for specific crash type (speed and ice related crashes) on rural 1.6 km (1 mi) highway sections of Oregon highways. The analysis includes data generated with the extensive use of spatial techniques and incorporates climate data to enhance environmental considerations. The paper compares the results of five ranking methods—critical rate (by functional class), critical rate (by functional class and climate zone), potential for crash reduction, expected frequency (adjusted by empirical-Bayes), and frequency. For the empirical-Bayes (EB) methods, safety performance functions were generated using negative binomial regression techniques. The twenty top 1.6 km (1 mi) sections were identified for each method and compared. The results reveal that the frequency and expected frequency methods identified the most sites in common, followed by the rate-based methods. The potential for crash reduction method identified the most unique ranked list. The results highlight the differences in ranking methods and confirm that even with significant aggregation to improve the rate-based methods they did not identify segments similar to the more sophisticated EB-techniques.

CE Database subject headings:

<sup>1</sup> Assistant Professor, Department of Civil & Environmental Engineering, Portland State University, P.O. Box 751, Portland, OR 97207-0751, E-mail: [monsere@pdx.edu](mailto:monsere@pdx.edu)

<sup>2</sup> Graduate Research Assistant, School of Urban Studies and Planning, Portland State University, P.O. Box 751, Portland, OR 97207-0751, E-mail: [pbosa@pdx.edu](mailto:pbosa@pdx.edu)

<sup>3</sup> Associate Professor, Department of Civil & Environmental Engineering and Nohad A. Toulon School of Urban Studies & Planning, Portland State University, P.O. Box 751, Portland, OR 97207-0751, E-mail: [bertini@pdx.edu](mailto:bertini@pdx.edu)

## INTRODUCTION

For state highway safety improvement programs, the identification of high crash locations (HCL) is a critical tool for safety engineers and planners. In general, networks can be screened and locations ranked based on frequency, severity, or rate of crashes, some combination of these factors in indices, trend or pattern analysis, or an estimate of the excess crashes with respect to the expected average for similar sites (i.e., potential for improvement, potential accident reduction). In recent years, techniques for screening transportation networks to identify high crash locations have become more sophisticated. These methods have more data requirements than the traditional methods and state departments of transportation (DOTs) have been slow to adopt them. Many agencies still rely on traditional methods to identify candidate locations for safety improvement (frequency, rate, severity). For example, Hallmark, et al. (2001) recently surveyed 17 other state DOTs for the Iowa DOT and found that they all used frequency, rate, severity or a combination of methods.

Each method has relative advantages and disadvantages but the final criterion for how well a screening method does is, ultimately, how effectively it identifies locations where correctable crashes are found (*Cheng and Washington, 2005a*). Once an HCL has been identified, the typical procedure is to perform detailed engineering studies to diagnose potential countermeasures based on site conditions, crash history, and other factors. Following a benefit-cost analysis, the most promising sites are usually programmed for improvement. Thus, it is important to identify sites with the most “promise” for improvement since engineering studies are expensive, agencies have limited budgets, and if a site with potential is not identified, an opportunity to substantially improve safety are missed (*Hauer, 2002*).

While networks are often screened for total crashes, another approach is to screen for unusual occurrences of a particular crash type with some countermeasure in mind. Analysis of recent data in Oregon reveal that speeding (defined as a driver charged with a speeding, racing, driving too fast for conditions as indicated by police officer, or exceeding the posted limit) is a contributing factor in approximately 31% of fatal crashes (*Monseré et al., 2004*). Speed has a strong correlation with highway crashes for numerous reasons: it reduces a driver’s ability to safely navigate around curves or objects, increases the stopping distance of a vehicle, and decreases the time available for collision avoidance in dangerous situations. Safety professionals at the Oregon DOT are strongly interested in managing vehicle speeds to improve road safety through engineering, enforcement, and education strategies. The current Oregon DOT HCL method is a sliding window approach that calculates a safety index as a function of

frequency, rate, and severity for all crashes and does not prioritize high crash locations by a specific crash type (*ODOT, 2006*).

The focus of this paper is to compare five screening and ranking methods for one specific crash type—speed and ice related crashes. The paper compares critical rate (by functional class), critical rate (by functional class and climate zone), potential for crash reduction, and empirical-bayes adjusted frequency and frequency screening methods. In the sections that follow, the paper describes the data that were assembled for analysis and spatial techniques used to create additional data. Each of the ranking methods is then described, including the development process of safety performance functions for use in the empirical-bayes methods. Finally, the results of the analysis are presented, followed by conclusions.

### **Data Assembly**

Two primary sources were used in generating the data for analysis—crash and highway data. The Statewide Crash Data System (CDS) includes all legally reported crashes that occur on public roads in Oregon. For this analysis, three years of data (2000-2002) on state highways were used. The crash data contains details about the crash, environmental and roadway conditions, and driver information. For state highways, crashes are located by a route-milepost system. Crashes were selected from the crash data base where both icy road conditions were present and at least one vehicle was coded with a speed involvement flag. The speed-involved flag is generated by the system and includes crashes where speed was too fast for conditions and where the vehicle was exceeding the posted limit. Both of these fields are based on judgment and are known to be somewhat biased (especially in icy conditions), however, it is the only current way to identify speed-involved crashes.

Other highway data were available in a spatial format such as number of lanes, total roadway width, shoulder width, average daily volumes, presence of medians, and functional class. A geographic information systems (GIS) representation of the state highway system contains a linear referencing system which allows these data to easily be attributed to any point on the highway network. City boundaries were also used to attribute sections as rural or urban and as shown later, climate data was used to attribute environmental factors. How these data were used for network screening is described in the following sections.

### **Ranking Methods Considered**

This empirical analysis of ranking speed and ice crashes compares five methods: critical rate (by functional class), critical rate (by functional class and climate zone), potential for crash reduction, expected frequency (adjusted

by empirical-Bayes), and frequency. While severity data were available, rankings that consider severity were not conducted. The rural state highway system was the focus of this analysis which was categorized according to functional class (interstate, principal arterial, and minor arterial or collector). For use in this network screening, the highways were divided in to 1.6 kilometer (1 mile) segments. A sliding window approach, which can improve screening methods used in practice, was not used in this comparison. There are 530 1.6 km (1 mi) mile rural interstate sections, 2,224 rural principal arterial sections, and 2,343 rural minor arterial sections. The following sections describe the methods used.

### **Rate Quality Control by Functional Class**

The rate quality control (RQC) method determines a critical rate for a particular segment using the exposure (vehicle-kilometers traveled—VKT), a probability constant, and the average rate for similar sections. The critical rate is calculated for each 1.6 km (1 mi) section  $i$  as shown in equation 1:

$$R_{Ci} = R_A + K \sqrt{\frac{R_A}{M_i} + \frac{1}{2M_i}} \quad (1)$$

where  $R_{Ci}$  = critical rate for each 1.6 km (1 mi) section  $i$ ,  $R_A$  = the average rate for similar facility,  $K$  = probability constant based on desired level of significance (1.645 for 95%) and  $M_i$  = millions of vehicle kilometers traveled (VKT or entering vehicles =  $(V \cdot D \cdot L) / 1,000,000$ ). In this analysis, the average rate is for all roadway segments in the database sharing a similar functional classification. The first column (statewide) in Table 1 lists the mean crash rates for each of the three functional classifications. All rates were calculated for the total speed and ice crashes that occurred in the three year period. The 2001 volumes were used for exposure and were assumed to be constant over the three year period. If the observed crash rate for a segment exceeds the critical rate, the section is considered hazardous and then ranked by the percent that it exceeds the critical rate.

### **Rate Quality Control by Functional Class with Climate Zones**

An improvement in the rate quality control method is to better define peer sites for the crash type being screened. In this analysis, it is appropriate to consider environmental factors related to snow and ice when aggregating roadway segments. Oregon's geography produces a diverse climate which means that winter weather related crashes do not occur with the same frequency in all regions of the state. Temperate regions such as the coast and central Willamette valley tend to have much fewer ice related crashes than the mountainous regions of the Cascades and eastern Oregon. To capture this, the state was partitioned into nine climate regions.

Based on data availability, two variables were considered as contributing factors to icy road conditions—the mean annual days of measurable snowfall exceeding 0.25 cm (0.1 in) and the mean annual days when the temperature fell below freezing. These data were obtained from the National Oceanic and Atmospheric Administration’s (NOAA) Climate Atlas of the United States (NOAA, 2002). Coverage areas were presented as polygons, with each polygon representing the mean numbers of days in each category over a 30 year period. As shown in the upper left of Figure 1, there are nine categories for mean annual days with a temperature below 0° C (32° F). Note that the legend is not provided in the figure but darker shades indicate a greater mean number of days (the maximum was 300 in the highest elevations, but the mean is 106 days). As shown in the upper right of Figure 1, there are nine categories for 29-year mean annual days with measurable snowfall exceeding 0.25 cm (0.1 in) shown. The darker shades again represent more days (the darkest color represents a maximum of 90 days). Though this resolution is a bit coarse (0.8 to 1.6 km, or 0.5 to 1 mile) it was sufficient to estimate unique climate zones. By making one of the layers transparent, it was possible to visually merge the two layers and hand-trace the zones based on the snow and freezing temperature layers and an understanding of the geographic features of the state. The result is a set of nine individual climate zones shown in the bottom half of Figure 1.

Using these zones, highways in them were attributed with the climate data using a spatial query. The average crash rate for speed-ice related crashes for all roadway sections in each climate zone was determined by functional classification. It should be noted that the climate zones 1 and 8 do not have any interstate highways. The results of this analysis are shown in Table 1. The table reveals that, as expected, the climate zones with more winter weather conditions have higher average crash rates in each zone (i.e. 5, 7, 8, and 9) with the exception of minor arterials in the North Coast mountains. In fact, the crash rate for rural principal arterials sections in the Cascade Mountains is 0.28 crashes per MVKT (0.45 per MVMT), more than 2 times the statewide average of 0.12 crashes per MVKT (0.20 per MVMT) for speed and ice crashes. Other mountainous climate zones such the Northeast and Central High Desert have similarly high crash rates. The Oregon Coast and Valley & Foothills climate zones as well as the South Coast Mountains all have low winter weather crash rates, due in large part to the temperate natures of these regions.

In general, using more aggregation should improve critical rate methods since sites are compared more directly to peer sites. Using the average rate for each climate zone and functional class, a critical rate was calculated for each segment in the analysis region as described previously. Sections were again ranked in descending order by the percent that the observed crash rates for speed and ice crashes exceeded the critical rate.

### Potential for Crash Reduction

To apply the potential for crash reduction method, safety performance functions were estimated for the three functional classes of roadways. The best of these models were then combined with observed crash counts for three years in an empirical-Bayes approach to estimating the “excess” crashes on each 1.6 km (1 mi) section. These models were also used to generate a fourth ranking based on an adjusted frequency described in the next section.

Predictive crash models are used to estimate the frequency of crashes based on a set of explanatory variables. In most recent research, safety performance functions have been estimated with Poisson and negative binomial (NB) regression models. Safety performance functions (SPFs) developed using these techniques have been advanced by Hauer (1997), Poch and Mannering (1996), Shankar et al. (1995), and Miaou and Lum (1993) and Hauer (1998), Vogt and Bared (1998), and Persaud (1993), Harwood et al. (2002) and many others. As suggested in Lord et al. (2005), the Poisson and NB models are theoretically appealing representations of the crash occurrence process. The Poisson model is suited for crash count data since the distribution approximates rare events; however, it requires the mean of the count process equal its variance. Evidence in a large body of literature suggests that most crash data will likely be overdispersed (i.e., the variance will be significantly greater than the mean) and the restriction can be relaxed in the negative binomial regression method. The Poisson model in terms of the probability of having  $y_i$  number of crashes per year at location is given by equation (2) as:

$$P(y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \quad (2)$$

where  $P(y_i)$  = probability of a crash at location  $i$  having  $y$  crashes per year,  $\lambda_i$  is the expected number of crashes per year for location  $i$ , i.e.,  $E(y_i)$ . Poisson regression models estimate  $\lambda_i$  using standard maximum likelihood methods as a function of  $X_i$  explanatory variables using the log-linear model in equation (3) (Washington et al., 2003).

$$\ln \lambda_i = \beta X_i \quad (3)$$

The NB model adds an independent distributed error term  $\varepsilon$  in the parameter equation (3) which allows relaxes the assumption the mean equal the variance as shown in (4). The NB model is estimated from equation (5) using standard maximum likelihood techniques where  $\alpha$  = a measure of dispersion (Washington et al., 2003).

$$\ln \lambda_i = \beta X_i + \varepsilon_i \quad (4)$$

$$L(\lambda_i) = \prod_i \frac{\Gamma((1/\alpha) + y_i)}{\Gamma((1/\alpha) y_i!)} \left( \frac{1/\alpha}{1/\alpha + \lambda_i} \right)^{1/\alpha} \left( \frac{\lambda_i}{1/\alpha + \lambda_i} \right)^{y_i} \quad (5)$$

In this research, crash models were developed using 2001 data. As one would expect given that the more populated areas of the state do not experience significant winter weather, there is a large percentage of 1.6 km (1 mi) segments with zero crashes in the section for the year 2001 (72% of interstate, 85% principal arterial, 94% minor arterial). The mean number of crashes per km per year for interstates (0.35, or 0.56 crashes per mile), and principal arterials (0.14, or 0.23 crashes per mile), and rural minor (0.04, or 0.07 crashes per mile) are low.

The following variables were selected from the analysis data set for inclusion in the model. A summary of the response variable and the five predictive variables used in the model is shown in Table 2.

- TOTALSI\_01: total speed and ice crashes on the 1.6 km (1 mi) segment for 2001
- AADT: average daily traffic for 2001 (vehicles per day)
- SNWDYS: 30 year mean number of days with snowfall greater than 0.25 cm (0.1 in), (days)
- TOT\_SURF\_W: total width of the roadway (including shoulders), (feet)
- WIDTH\_LANE: total width of the roadway (including shoulders) per lane, m (ft).
- TOT\_SURF\_W: total width of the roadway (including shoulders), m (ft)
- SP\_CD: posted speed limit, km/h (mi/h)

The counts are clearly over-dispersed and the negative binomial regression models were chosen to model speed and ice crashes. Predictive variables for horizontal curvature were also desired but were not available. For the TOT\_SURF\_W variable on interstates, the maximum value for both roadways was used. The WIDTH\_LANE variable was created and intended to capture the amount of roadway allocated per lane. For example, a two-lane road with 3.6 m lanes and 1.8 m shoulders would have a 5.4 m WIDTH\_LANE variable.

Six negative binomial regression models were estimated for 2001 speed and ice crashes (TOTALSI\_01) as a function of exposure, road geometry, and weather variables in Table 2 using the *glm.nb* function in the statistical software R (R, 2005). The following models were specified (note that exposure was transformed with the natural logarithm consistent with other approaches to model counts).

- Model 1: TOTALSI\_01 ~ LOGAADT
- Model 2: TOTALSI\_01 ~ LOGAADT + SNWDYS
- Model 3: TOTALSI\_01 ~ LOGAADT + SNWDYS + WIDTH\_LANE
- Model 4: TOTALSI\_01 ~ LOGAADT + SNWDYS + TOT\_SURF\_W
- Model 5: TOTALSI\_01 ~ LOGAADT + SNWDYS + TOT\_SURF\_W + SP\_CD
- Model 6: TOTALSI\_01 ~ LOGAADT + SNWDYS + WIDTH\_LANE + SP\_CD



For each model, estimates of the variable coefficients were generated, diagnostic plots were created, correlations of variables were determined, and measures of model fit generated. The estimated coefficients and significance for each of the models are shown in Table 3. In all models, the exposure variable is significant. For all models of interstate functional class the coefficient for exposure is counterintuitive in that the sign is negative. This is because the higher volume sections are in climate zones with very limited winter weather conditions sections (Valley & Foothills). The weather variable (SNWDYS) is also significant in all models and has the expected sign in all models. The geometry variables are only significant in about half of the models. The posted speed code is significant in the interstate and principal arterial models (but has the opposite sign) and in model 6 for the minor arterials. In Table 3, measures of model fit are presented.  $R^2$  values (as in least squares linear regression) are not available for negative binomial regression but a number of pseudo- $R^2$  values have been proposed. In Table 3, the pseudo- $R^2$   $R^2_{DEV}$  presented in Cameron and Trivedi (1998) is shown (1-Residual Deviance/Null Deviance). In addition, Akaike's Information Criteria (AIC) are shown. Models with lower AIC criteria are preferred and it does not necessarily increase with additional regressors.

Using the diagnostics, the significance of the coefficients for each model, and judgment, a model was chosen for each functional class. The selected models are shaded in Table 3. For interstates, model 5 was chosen and model 6 was chosen for principal arterials. In both cases, both the  $R^2_{DEV}$  and AIC were the best for these models. For minor arterials, model 2 was chosen since the additional predictive variables in model 4, 5, 6 were not significant. Two dimensional plots (volume and crash) of the chosen models are shown on left of Figure 2 with plots of observed and fitted values on the right.

Using these models, the expected number of crashes for similar segments ( $\lambda_i$ ) can be made. The predictive equation takes the form in (6) where  $\beta$  is the coefficient values estimated by the model and  $X_n$  is the predictive variable.

$$\lambda_i = EXP(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_n X_n) \quad (6)$$

Using an empirical-Bayes approach discussed in Hauer et al. (2002), the expected number of crashes can be combined with the observed count of crashes at a location to produce an improved estimate of the expected number of crashes. The observed crash counts ( $K$ ) can be combined with the expected value of  $\lambda$  from equation (6) in the following equation:

$$\lambda = E(\lambda)\alpha + (1 - \alpha)K \quad (7)$$

where  $\alpha$  is the weight and is calculated in equation (8)

$$\alpha = \frac{1}{1 + (E(\lambda) * Y) / \phi} \quad (8)$$

where  $Y$  is the years of crash counts  $K$  and  $\phi$  is the overdispersion parameter estimated from the SPFs and shown in Table 3. For a full derivation and justification, see the discussion by Hauer (1997, pp. 193–194). When subtracting the expected number of crashes estimated in equation (7) from the observed number of crashes, values resulting in a positive value indicate the “excess” crashes above the average. The segments are then ranked by excess crashes in descending order. Figure 3 shows graphically the results of the screening and ranking method. The predicted 3 year snow and ice crashes for each segment (from equation 7) are plotted against the observed number of crashes. An equivalent line is also shown in the Figure. Points that are above the equivalent line have excess crashes, those below have fewer crashes than expected and the magnitude of the excess was used to rank these segments.

### Frequency and Adjusted Frequency

The final ranking method involves a simple rank of the three-year counts (frequency) and the rank based on the EB estimate in equation (7). These ranking are included for additional comparison to the other ranking methods.

## RESULTS

The ranking results of all five methods are shown in Table 4. Any segment that was identified in the “Top 20” has been included with rankings higher than 20 being shaded for better visual identification. The list is sorted in descending order for the “Potential for Crash Reduction” method. The table includes values for all predictive variables. A total of 54 segments were identified as top 20 by at least one method. The top 20 sites by each method identified segments with total speed and ice crashes of 224, 198, 163, 268, and 287 crashes, respectively. For reference there were 3,026 crashes on all 5,097 1.6 km (1 mi) segments in the database.

A simple visual inspection of Table 4 reveals that the ranking methods identify different sites. A tabulation was made of the number of common segments identified by each paired comparison and shown in Table 5. There are no sites that were identified by all five methods. From the table one can see that the critical rate methods identify 9 common segments. However, the sites ranked 1 and 3 by the critical rate with climate zones (also ranked by the potential for crash reduction) were ranked 287 and 31 by the critical rate with only functional class aggregation.

These two sites are outliers for their climate zones but not when compared against the statewide average. The remaining sites identified by both methods are fairly close in rank. The potential crash reduction identified only 3 sites in common with the simple frequency and no segments in common with EB adjusted frequency. The EB adjusted frequency and simple frequency produced the most similar rank ordered list with 16 segments in common. This is perhaps as expected, as the count of crashes is included in the EB adjustment. The segment ranked first by the potential for crash reduction method, is ranked 5 by the critical rate with climate zones and frequency, but 287 by the critical rate with functional class and 169 by the adjusted EB method. This segment is a rural freeway (I-5) south of Salem that experiences relatively few snow days (4) but recorded 15 speed and ice crashes in the three year period. The predictive models estimate a mean crash occurrence of essentially 0 (0.09 crashes) and the critical rate for other interstate section in its climate zone is 0.02/MVKT (0.01/MVMT). Only 1 crash occurred in 2001 (used in the modeling effort). That 15 ice related crashes occurred here over 3 years appears unusual but could be a result of the weather-related conditions that occurred in 2000 (5 crashes) and 2002 (9 crashes). A modeling approach that used annual weather data rather than 30-year averages might be able to better capture these unusual weather events.

Without having knowledge of the true safety of segments, it is difficult to say which ranking method is superior. In most recent work, however, rate-based methods have been discouraged in favor of EB-based methods. It was thought that better aggregation might better approximate the potential for crash reduction methods but this appears to not be the case. This analysis shows that these methods will identify different segments than the EB-based method. The frequency method appears to identify similar segments as the EB-based methods. Recent research by Cheng and Washington (2005) suggest that the EB approach produces a superior identification of truly unsafe sites by limiting the number of falsely identified sites. The soon to be released *Safety Analyst* software will use the potential for crash reduction method as one network screening method. In practice, a number of states use rate, or critical rate. What is useful from this analysis is that it is clear that rate-based methods will produce different list even with more detailed aggregations.

One issue that deserves further analysis is the safety performance models generated for use in the analysis. A simplification was made in adopting consistent 1.6 km (1 mi) sections for modeling rather than varying section lengths to capture consistent variables (i.e. breaking segments by changes in lane width). There were highway data that were developed in Oregon by Strathman et al. (2001) but the segment lengths were very short. In Strathman's models of mainline crashes, zero-inflated models were fit. With speed and ice crashes being a subset of total crashes,

the negative binomial models did not fit well. An alternative approach that was considered was to model total crashes, then estimate the expected number of speed and ice crashes by assuming an average percentage of total crashes for each of the functional classes and climate zones. While the models of total crashes had better fits, the usefulness of the modeling approach to capture other variables' influences on speed and ice crashes would be diminished. Finally, none of these SPF models included any measure of intersections or driveways. These variables should be added to future models. The resolution and comparability of the weather data (30 year averages) might highlight segments as "above average" when they may have just been misclassified.

## **CONCLUSIONS**

This empirical analysis of ranking speed and ice crashes compared five ranking methods: critical rate (by functional class), critical rate (by functional class and climate zone), potential for crash reduction, expected frequency (adjusted by empirical-Bayes), and frequency. For each method, the "top 20" 1.6 km (1 mi) sections were identified. A comparison of the methods showed that rate-based methods identified similar rank-order segments and that the EB-adjusted frequency compared well with the simple frequency method. The potential for crash reduction screening method (with EB-adjusted expected crash frequencies) identified the most unique list. A definitive answer on which method is superior is difficult, since that answer would require a detailed analysis of each site identified. However, with other research confirming the advantages of EB-based methods, one should consider these results to support the use of count-based methods since even with significant aggregation to improve the rate-based methods, a different set of segments were identified. The SPFs used in this screening and ranking method could be improved. Other predictive variables could be included in the models and additional years of data could be included in the modeled data. While this research effort focused on identifying segments related to speed and ice conditions, the methodology could be applied to any number or combinations of crash variables. For example, the analysis of wet weather crashes could include precipitation, pavement roughness, and the usual highway geometry in generating a network screening approach. Finally, the treatment of severity or crash type in network screening should be explored further.

## **ACKNOWLEDGMENT**

The authors gratefully acknowledge the Oregon Department of Transportation for sponsoring this research as well as the Department of Civil & Environmental Engineering in the Maseeh College of Engineering & Computer Science at Portland State University.

## REFERENCES

- Cameron, C.A. and Trivedi .P. K. (1998). *Regression Analysis of Count Data*. Econometric Society Monographs No. 30, Cambridge University Press, Cambridge, United Kingdom.
- Cheng, W. and S. Washington. (2005) Experimental evaluation of hotspot identification methods. *Accident Analysis and Prevention*. Vol. 37. pp 870-881
- Climate Atlas of the United States (2002), National Oceanic Atmospheric Administration. Accessed <http://www.ncdc.noaa.gov/oa/about/cdrom/climatls2/datadoc.html>
- Hallmark, S.L., R. Basavaraju, and M. Pawlovich. (2002) Evaluation of the Iowa DOT's Safety Improvement Candidate List Process. Center for Transportation Research & Education. Iowa State University. Ames, IA. 2002.
- Harwood, D.W., Bauer, K.M., Potts, I.B., Torbic, K.R., Richard, K.R., Kohlman Rabbani, E.R., Hauer, E., and Elefteriadous, L. (2002) *Safety Effectiveness of Intersection Left- and Right-Turn Lanes*. FHWA, U.S. Department of Transportation.
- Hauer, E. (1996). Identification of sites with promise. In *Transportation Research Record: Journal of the Transportation Research Board* 1542, TRB, National Research Council. Washington, D.C., 1996. pp. 54-60.
- Hauer, E. (1997) *Observational Before-After Studies in Road Safety*. Pergamon, Oxford, U.K., 1997.
- Hauer, E., Harwood, W.D., Council, F.M., and Griffith, M.S. (2002) Estimating Safety by the Empirical Bayes Method. In *Transportation Research Record: Journal of the Transportation Research Board* 1784, National Research Council, Washington, D.C., pp. 126-131.
- Hauer, E., J. C. N. Ng, and J. Lovell. Estimation of Safety at Signalized Intersections. In *Transportation Research Record: Journal of the Transportation Research Board* 1185, TRB, National Research Council. Washington, D.C., 1988, pp. 48-61.
- Hauer, E., J. Kononov, B. Allery, and M. S. Griffith. (2002) Screening the Road Network for Sites with Promise. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1784, TRB, National Research Council, Washington, D.C., 2002, pp. 27-32.
- Lord, D., Washington. S., Ivan. J. (2005). Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis and Prevention*. Vol. 37. pp 35-46

- Miaou, S.P. and Lum, H. (1993) Modeling Vehicle Accidents and Highway Geometric Design Relationship. *Accident Analysis and Prevention*, Vol. 25, No. 6, p. 689-709.
- Monsere, C. Bertini, R.L., Bosa, P., Chi, Delia. (2006). *Comparison of Identification and Ranking Methodologies for Speed-Related Crash Locations*. Oregon Department of Transportation, FHWA-OR-RD-06-14, Salem Oregon, June 2006.
- Monsere, C., Dill, J., Newgard, C., Rufolo, T., Wemple, E., Milliken, C. and Bertini, R.L. (2004) *Impacts and Issues Related to Proposed Changes in Oregon's Interstate Speed Limits*. Portland State University, Center for Transportation Studies, Research Report, September 2004.
- Oregon Department of Transportation (ODOT). (2006) *Traffic Manual*. Salem, Oregon. May 2006.
- Persaud, B. N., and L. Dzbik. Accident Prediction Models for Freeways. In *Transportation Research Record: Journal of the Transportation Research Board* 1401, TRB, National Research Council. Washington, D.C., 1993, pp. 55–60.
- Poch, M. and Mannering, F. (1996) Negative Binomial Analysis of Intersection-Accident Frequencies. *Journal of Transportation Engineering*, Vol. 122, No. 2, p. 105-113.
- R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Shankar, V., Mannering F., and Barfield, W. (1995) Effects of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies. *Accident Analysis and Prevention*, Vol. 27, No. 3, p. 371-389.
- Strathman, J.G., Duecker, K.J., Shang, J., and Williams, T. (2001) *Analysis of Design Attributes and Crashes on the Oregon Highway System*. Oregon Department of Transportation Research Group and FHWA, U.S. Department of Transportation
- Vogt, A., and J. Bared. Accident Models for Two-Lane Rural Segments and Intersections. In *Transportation Research Record: Journal of the Transportation Research Board* 1635, TRB, National Research Council. Washington, D.C., 1998, pp. 18–29.
- Washington, S.P., Karlaftis, .M.G., and Mannering, F.L., (2003) *Statistical and Econometric Methods for Transportation Data Analysis*. Chapman & Hall/CRC New York.
- Washington. S. and Cheng, (2005) *W. High Risk Crash Analysis - Final Report 558*. Arizona Department of Transportation. FHWA-AZ-05-558. Phoenix, Arizona.

**Table 1 Average Crash Rate per MVKT (MVMT) by Climate Zone and Functional Classification for Speed and Ice Related Crashes on Rural Highways, 2000-2002.**

<b>Functional Class</b>	<b>Statewide</b>	<b>Oregon Coast</b> 1	<b>Valley &amp; Foothills</b> 2	<b>N. Coast Mountains</b> 3	<b>S. Coast Mountains</b> 4	<b>Cascade Mountains</b> 5	<b>Columbia Basin</b> 6	<b>Southeast</b> 7	<b>Central High Desert</b> 8	<b>Northeast</b> 9
Interstate	0.11 (0.18)		0.01 (0.02)		0.03 (0.04)	0.04 (0.06)	0.08 (0.12)	0.27 (0.44)		0.37 (0.59)
Principal Arterial	0.12 (0.20)	0.03 (0.05)	0.05 (0.08)	0.07 (0.11)	0.03 (0.05)	0.28 (0.45)	0.07 (0.12)	0.11 (0.18)	0.08 (0.13)	0.14 (0.23)
Minor Arterial	0.09 (0.15)	0.07 (0.11)	0.05 (0.08)	0.11 (0.17)	0.03 (0.04)	0.17 (0.27)	0.08 (0.13)	0.06 (0.1)	0.11 (0.17)	0.16 (0.25)

**Table 2 Summary of Variables Used in Model Fitting.**

Variable	Interstate (n=530)				Principal (n=2,224)				Minor (n= 2,343)			
	Min	Max	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max	Mean	Std Dev
TOTALSI_01 (crashes)	0	13	0.67	1.36	0	9	0.23	0.70	0	4	0.07	0.3
AADT (vpd)	7,700	137,800	22,859	19,740	220	110,300	5,224	6,944	30	29,700	1,994	3,340
SNWDYS (days)	0	46	12.58	13.59	0	90	17.37	18.89	0	90	14.19	16.07
WIDTH_LANE m (ft)	3.66 (12)	8.23 (27)	5.92 (19.43)	0.38 (1.26)	2.44 (8)	21.95 (72)	5.45 (17.88)	1.82 (5.96)	2.74 (9)	13.72 (45)	4.33 (14.19)	1.37 (4.49)
TOT_SURF_W m (ft)	7.92 (26)	20.12 (66)	12.11 (39.72)	1.2 (3.94)	6.71 (22)	29.87 (98)	11.69 (38.36)	4.02 (13.2)	5.18 (17)	25.6 (84)	8.43 (27.65)	2.62 (8.59)
SP_CD km/h (mi/h)	80.47 (50)	104.6 (65)	103.9 (64.56)	3.36 (2.09)	40.23 (25)	104.6 (65)	87.48 (54.36)	4.55 (2.83)	32.19 (20)	88.51 (55)	85.02 (52.83)	11.68 (7.26)



**Table 3 Estimates of Coefficients for Models of Speed and Ice Crashes on Rural Highways, 2001.**

		LOG(AADT)		SNWDYS		WIDTH_LN	TOT_SURF_W		SP_CD				
	Intercept	Average daily traffic, vpd		Number of days with > 0.25 cm (0.1in) snow accumulation		Average road width allocated per lane, m (ft)	Total surface width of roadway, m (ft)		Posted speed limit kph (mph)	Pseudo R2	AIC	$\phi$ Overdispersion <sup>1</sup>	
Interstate	1	14.95 ***	-1.62 ***							0.24	1050.2	1.26	
	2	12.99 ***	-1.44 ***	0.015 **						0.26	1046.7	1.19	
	3	10.65 ***	-1.46 ***	0.016 **		0.430 (0.131) *				0.26	1045.0	1.14	
	4	10.73 ***	-1.71 ***	0.018 **			0.404 (0.123) ***			0.29	1033.1	1.07	
	5	18.27 ***	-1.65 ***	0.020 ***			0.266 (0.081) *	-0.063 (-0.101) **		0.30	1028.3	1.04	
	6	23.48 ***	-1.51 ***	0.020 ***		-0.061 (-0.018)		-0.092 (-0.147) ***		0.29	1033.3	1.07	
Principal Arterial	1	-4.45 ***	0.36 ***							0.04	2470.3	4.22	
	2	-7.11 ***	0.59 ***	0.033 ***						0.20	2281.9	1.83	
	3	-7.09 ***	0.64 ***	0.034 ***		-0.084 (-0.026) *				0.21	2278.7	1.80	
	4	-7.17 ***	0.60 ***	0.033 ***			-0.006 (-0.002)			0.20	2283.8	1.83	
	5	-12.72 ***	0.63 ***	0.033 ***			-0.002 (-0.004)	0.060 (0.096) ***		0.21	2272.0	1.75	
	6	-12.69 ***	0.68 ***	0.033 ***		-0.083 (-0.025) *		0.060 (0.096) ***		0.22	2267.0	1.73	
Minor Arterial	1	-6.72 ***	0.56 ***							0.09	1098.5	3.47	
	2	-7.78 ***	0.63 ***	0.028 ***						0.15	1055.8	1.84	
	3	-7.77 ***	0.69 ***	0.028 ***		-0.096 (-0.029)				0.16	1055.6	1.85	
	4	-7.82 ***	0.66 ***	0.028 ***			-0.017 (-0.005)			0.16	1057.5	1.88	
	5	-6.56 ***	0.66 ***	0.026 ***			-0.015 (-0.004)	-0.011 (-0.024)		0.16	1056.2	1.63	
	6	-6.45 ***	0.69 ***	0.028 ***		-0.102 (-0.032)		-0.011 (-0.025) *		0.16	1053.7	1.60	

\*\*\* Significant > 0.001, \*\* Significant at 0.01, \* Significant at 0.05, # Significant at 0.1

<sup>1</sup> Estimates of overdispersion parameter produced by R is the reciprocal of the parameter in Hauer et al. (2002)  
The values in the table are shown consistent with Hauer's.



Table 4 Results of Network Screening and Ranking

ID	Functional Class	Climate Zone	Model Variables							Ranks					
			Total Crashes (00-02)	Total S&I Crashes (00-02)	Total S & I Crashes (01)	AADT (2001)	Posted Speed km/h (mi/h)	Total Surface Width m (ft)	Snow Days	Total Surface Width/Lane m (ft)	Critical Rate (FC)	Critical Rate (FC w/CZ)	Potential Crash Reduction	Frequency EB Adjusted	Frequency
249	Interstate	Valley	44	15	1	58,500	104.7 (65)	14.6 (48)	4	4.9 (16)	287	3	1	169	5
4722	Pr. Arterial	Valley	14	11	2	6,100	88.6 (55)	9.1 (30)	4	4.6 (15)	31	1	2	181	14
4644	Pr. Arterial	Cascade Mts	13	11	4	5,000	88.6 (55)	9.8 (32)	13	4.9 (16)	20	54	3	162	14
4645	Pr. Arterial	Cascade Mts	13	8	2	5,000	88.6 (55)	19.5 (64)	13	9.8 (32)	45	113	4	288	42
5431	Pr. Arterial	Cascade Mts	11	7	0	4,600	64.4 (40)	14.9 (49)	26	4.9 (16)	56	141	5	566	59
6156	Pr. Arterial	N. Coast Mts	9	8	7	8,900	88.6 (55)	6.7 (22)	4	6.7 (22)	108	22	6	222	42
5372	Pr. Arterial	Valley	14	10	4	15,700	88.6 (55)	20.7 (68)	8	5.2 (17)	166	17	7	104	22
248	Interstate	Valley	28	7	1	58,500	104.7 (65)	14.6 (48)	4	4.9 (16)	287	52	8	338	59
5430	Pr. Arterial	Cascade Mts	7	6	0	4,200	64.4 (40)	14.9 (49)	26	4.9 (16)	69	185	9	715	81
5368	Pr. Arterial	Valley	15	9	2	16,600	88.6 (55)	20.7 (68)	4	5.2 (17)	206	27	10	133	29
4754	Pr. Arterial	Cascade Mts	15	9	6	2,900	88.6 (55)	13.4 (44)	46	6.7 (22)	13	28	11	118	29
4526	Pr. Arterial	Central High Dst	6	6	2	3,100	88.6 (55)	11 (36)	13	5.5 (18)	48	11	12	358	81
5414	Pr. Arterial	Cascade Mts	11	7	4	1,300	88.6 (55)	11 (36)	46	5.5 (18)	8	12	13	215	59
4753	Pr. Arterial	Cascade Mts	14	9	8	2,900	88.6 (55)	13.4 (44)	46	4.6 (15)	13	28	14	101	29
4752	Pr. Arterial	Cascade Mts	10	9	6	2,900	88.6 (55)	13.4 (44)	46	4.6 (15)	13	28	14	101	29
4646	Pr. Arterial	Cascade Mts	12	7	2	7,700	88.6 (55)	11 (36)	13	5.5 (18)	123	283	16	199	59
4435	Pr. Arterial	Valley	34	5	1	10,500	72.5 (45)	9.4 (31)	2	4.9 (16)	287	78	17	600	107
2047	Pr. Arterial	Cascade Mts	11	8	2	8,400	88.6 (55)	12.2 (40)	26	6.1 (20)	98	283	18	129	42
4643	Pr. Arterial	Cascade Mts	10	6	1	5,000	88.6 (55)	9.8 (32)	13	4.9 (16)	89	236	19	264	81
180	Interstate	Valley	24	5	0	37,400	104.7 (65)	11.6 (38)	2	5.8 (19)	287	69	20	493	107
5415	Pr. Arterial	Cascade Mts	6	6	3	1,300	88.6 (55)	11 (36)	46	5.5 (18)	16	23	22	244	81
5417	Pr. Arterial	Cascade Mts	8	6	2	1,500	88.6 (55)	11 (36)	46	5.5 (18)	19	35	23	230	81
4735	Pr. Arterial	Valley	9	5	2	4,300	88.6 (55)	9.8 (32)	8	4.9 (16)	117	16	25	371	107
2826	Interstate	Northeast	24	21	13	8,900	104.7 (65)	11.6 (38)	26	5.8 (19)	4	41	26	1	1
4738	Pr. Arterial	Valley	6	4	0	2,900	88.6 (55)	16.8 (55)	13	8.5 (28)	121	19	38	613	158
5401	Pr. Arterial	Cascade Mts	26	13	7	1,800	88.6 (55)	11.6 (38)	90	5.8 (19)	1	2	40	12	11
2810	Interstate	Northeast	23	14	7	10,400	104.7 (65)	11.3 (37)	26	5.8 (19)	33	179	42	9	7
2780	Interstate	Colb. Basin	24	14	3	9,900	104.7 (65)	11.3 (37)	26	5.8 (19)	30	4	45	8	7
2825	Interstate	Northeast	24	17	7	8,800	104.7 (65)	11.6 (38)	26	5.8 (19)	12	85	50	4	4
2822	Interstate	Northeast	32	20	7	8,900	104.7 (65)	12.5 (41)	26	6.4 (21)	6	57	56	2	2
2808	Interstate	Northeast	17	12	1	10,200	104.7 (65)	11.3 (37)	26	5.8 (19)	39	248	71	14	12
5400	Pr. Arterial	Cascade Mts	17	11	4	1,800	88.6 (55)	11.6 (38)	90	5.8 (19)	3	5	72	23	14
2807	Interstate	Northeast	19	11	3	10,200	104.7 (65)	11.3 (37)	26	5.8 (19)	51	283	94	22	14
4757	Pr. Arterial	Cascade Mts	21	14	2	2,900	88.6 (55)	9.8 (32)	90	4.9 (16)	2	6	99	7	7
2864	Interstate	Northeast	12	11	4	9,700	104.7 (65)	11.6 (38)	26	5.8 (19)	49	280	135	17	14

5402	Pr. Arterial	Cascade Mts	18	9	5	1,800	88.6 (55)	11.6 (38)	90	5.8 (19)	5	10	145	38	29
2776	Interstate	Colb. Basin	13	10	3	10,100	104.7 (65)	11.6 (38)	26	5.8 (19)	64	18	163	27	22
2827	Interstate	Northeast	14	11	5	8,900	104.7 (65)	11.6 (38)	26	5.8 (19)	38	235	209	15	14
5399	Pr. Arterial	Cascade Mts	14	8	3	1,800	88.6 (55)	11.6 (38)	90	5.8 (19)	10	14	221	46	42
5405	Pr. Arterial	Cascade Mts	13	7	4	1,300	88.6 (55)	9.8 (32)	90	4.9 (16)	8	12	237	70	59
2884	Interstate	Southeast	12	11	7	8,000	104.7 (65)	12.2 (40)	13	6.1 (20)	35	111	256	13	14
5398	Pr. Arterial	Cascade Mts	12	8	1	1,800	88.6 (55)	13.4 (44)	90	4.6 (15)	10	14	276	42	42
5407	Pr. Arterial	Cascade Mts	7	6	1	1,300	88.6 (55)	12.2 (40)	90	6.1 (20)	16	23	277	92	81
5393	Pr. Arterial	Cascade Mts	27	18	9	7,300	88.6 (55)	15.2 (50)	90	5.2 (17)	7	20	278	3	3
2781	Interstate	Colb. Basin	15	12	4	9,500	104.7 (65)	11.6 (38)	46	5.8 (19)	36	7	310	10	12
2868	Interstate	Northeast	13	10	3	7,700	104.7 (65)	11.6 (38)	26	5.8 (19)	40	229	424	20	22
2821	Interstate	Northeast	21	10	3	8,900	104.7 (65)	12.5 (41)	26	6.4 (21)	54	283	430	19	22
2823	Interstate	Northeast	13	10	5	8,800	104.7 (65)	12.5 (41)	26	6.4 (21)	50	283	462	18	22
2786	Interstate	Northeast	12	10	5	9,200	104.7 (65)	11.6 (38)	46	5.8 (19)	57	283	599	16	22
5391	Pr. Arterial	Cascade Mts	25	15	5	7,700	88.6 (55)	20.1 (66)	90	5.2 (17)	18	51	606	5	5
5392	Pr. Arterial	Cascade Mts	28	14	6	7,700	88.6 (55)	17.4 (57)	90	5.8 (19)	22	65	667	6	7
7	Interstate	S. Coast Mts	46	8	3	14,700	88.6 (55)	13.4 (44)	4	6.7 (22)	193	9	771	35	42
6	Interstate	S. Coast Mts	26	8	5	13,900	88.6 (55)	13.4 (44)	4	6.7 (22)	185	8	878	33	42
2824	Interstate	Northeast	18	11	4	8,800	104.7 (65)	12.5 (41)	46	6.4 (21)	37	232	895	11	14

**Table 5 Number of Top 20 Segments Commonly Identified By Each Screening Method**

	<b>Critical Rate (FC)</b>	<b>Critical Rate (FC w/CZ)</b>	<b>Potential for Crash Reduction</b>	<b>Frequency EB Adjusted</b>	<b>Frequency</b>
Critical Rate (FC)	-	9	4	7	9
Critical Rate (FC w/CZ)	9	-	4	5	8
Potential for Crash Reduction	4	4	-	0	3
Frequency EB Adjusted	7	5	0	-	16
Frequency	9	8	3	16	-

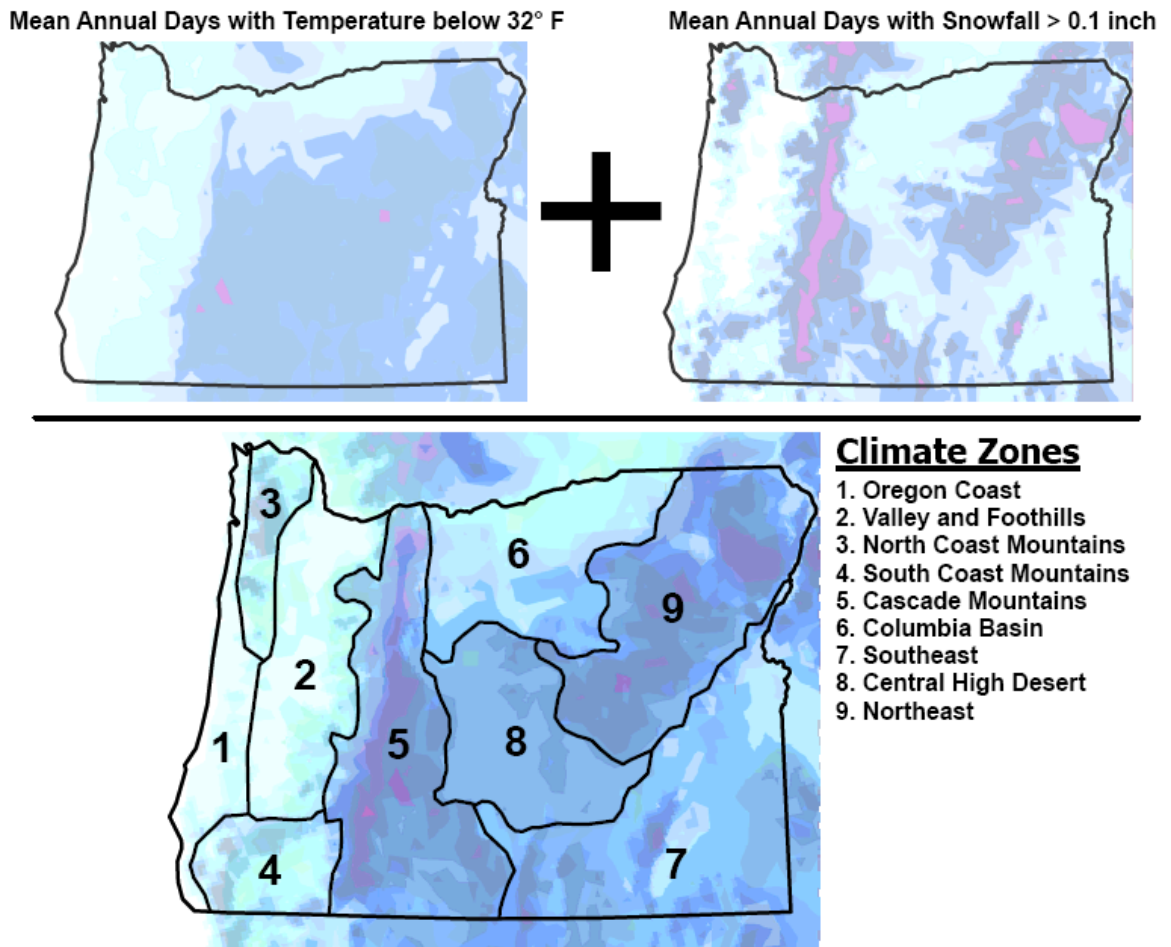
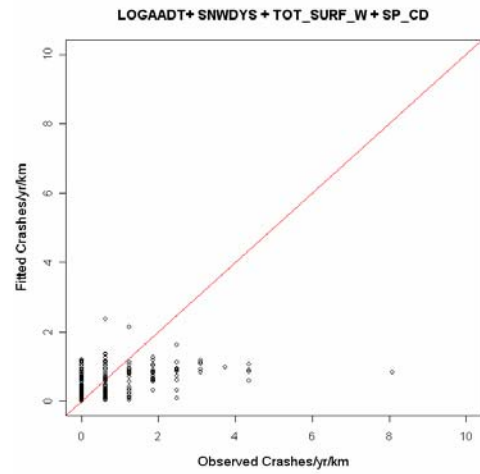
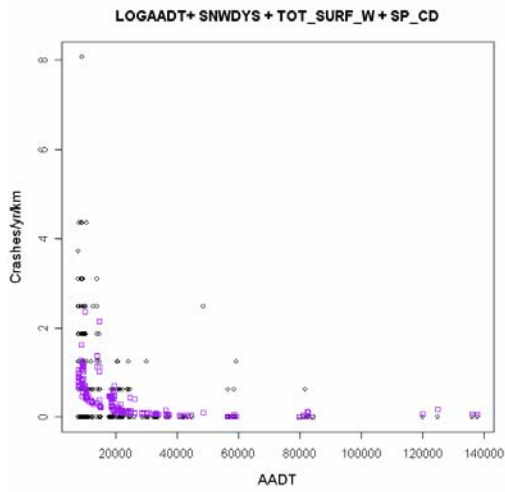
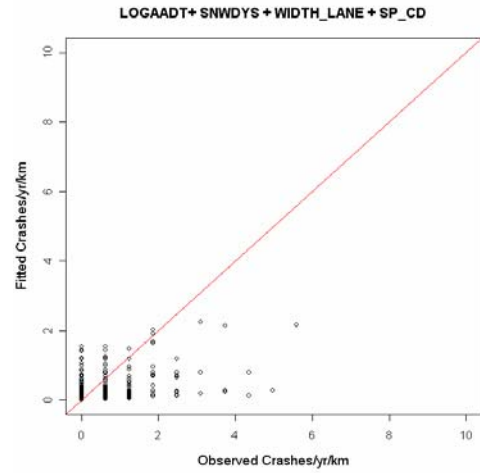
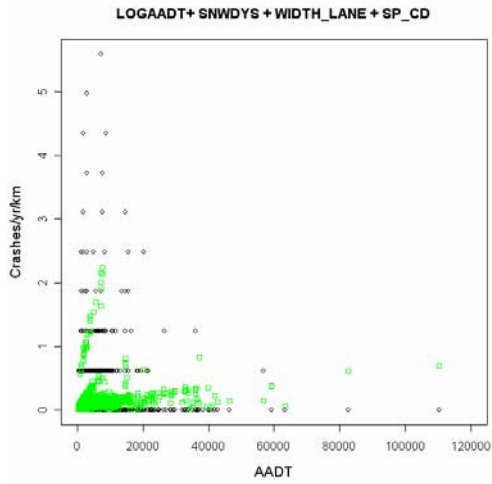


Figure 1 Spatial Climate Data Combined to Define Climate Zones.

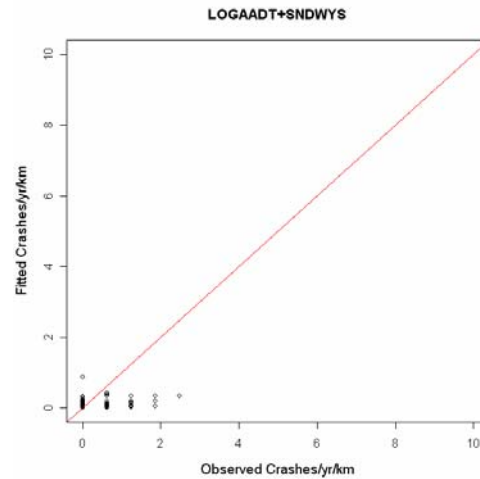
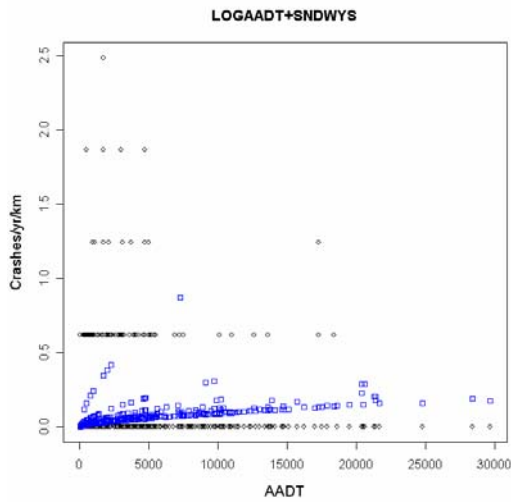
**RURAL INTERSTATE**



**RURAL PRINCIPAL ARTERIAL**



**RURAL MINOR ARTERIAL**



**Figure 2 Two-dimensional Plots and Fitted vs. Observed Values of Selected Models for Each Functional Class.**

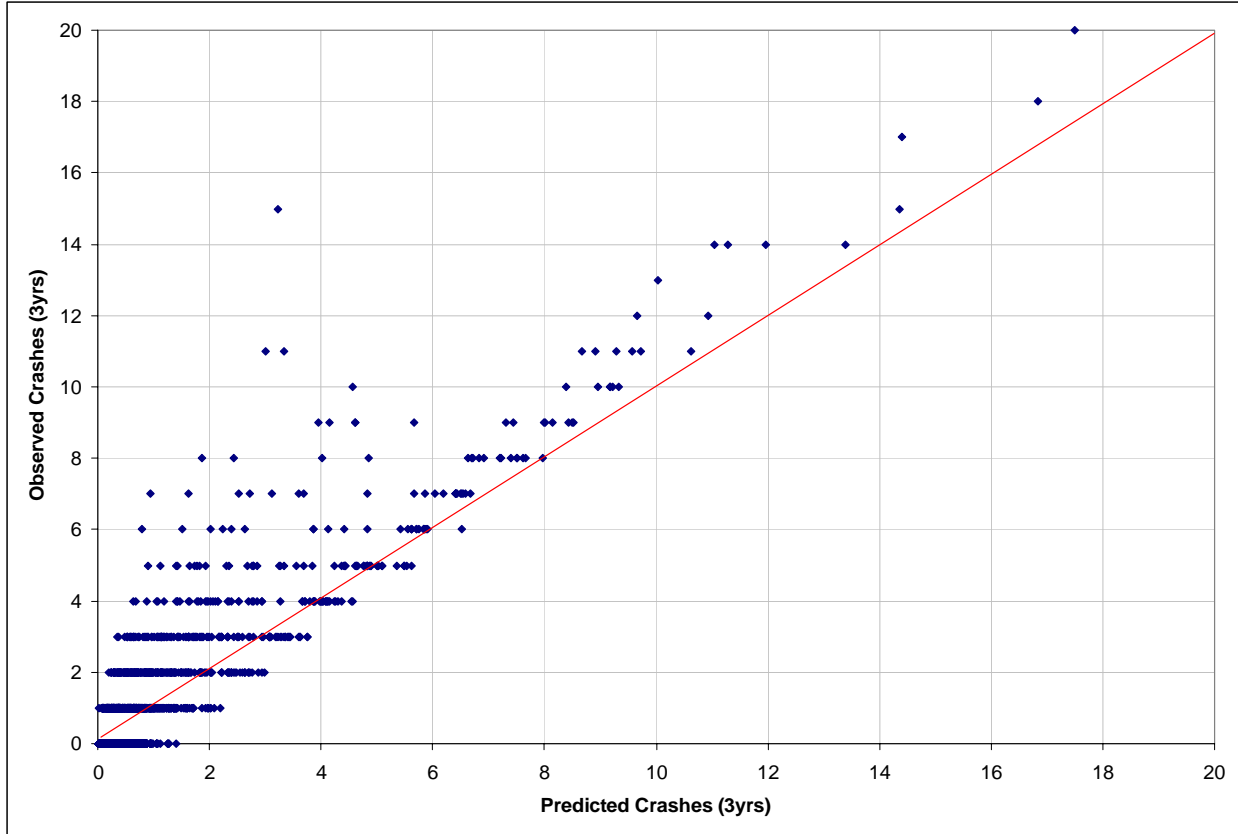


Figure 3 Results of the Potential for Crash Reduction Screening.