DEVELOPMENT OF A ROADWAY WEATHER SEVERITY INDEX

Showcase Evaluation #16

Final Technical Report

By

Christopher Strong, Yurii Shvetsov and Julia Sharp
Western Transportation Institute
College of Engineering
Montana State University

Prepared for the

U.S. Department of Transportation
Research and Innovative Technology Administration

August 2005

DISCLAIMER

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the U.S. Department of Transportation or Montana State University.

Alternative accessible formats of this document will be provided upon request. Persons with disabilities who need an alternative accessible format of this information, or who require some other reasonable accommodation to participate, should contact Kate Heidkamp, Communications and Information Systems Manager, Western Transportation Institute, Montana State University, PO Box 174250, Bozeman, MT 59717-4250, telephone number 406-994-7018, e-mail: KateL@coe.montana.edu.

ACKNOWLEDGMENTS

The authors are indebted to those who provided data for this project. For Oregon, this includes the staff of the ITS Unit in the Oregon Department of Transportation's (ODOT) Traffic Management Section, along with Sylvia Vogel from the Crash Analysis and Reporting Unit within ODOT's Transportation Data Section. For Montana, this includes Chandra Roth in the Maintenance Division and Pierre Jomini within the Engineering Division. For California data, thanks go to the Highway Safety Information System program managed through the Turner Fairbanks Highway Research Center.

We are also appreciative of the statistical support provided by Dr. John Borkowski of the Mathematics Department at Montana State University.

The authors would also like to thank the University Transportation Center program of the Research and Innovative Technology Administration for providing funding for this project. We are also thankful for the support of the Rural California/Oregon Advanced Transportation Systems (COATS) Steering Committee throughout this project, and are hopeful that the results will be useful as they work to improve the safety and efficiency of the rural transportation system in northern California and southern Oregon.

GLOSSARY OF ABBREVIATIONS

AADT	Annual Average Daily Traffic
ATR	Automatic Traffic Recorder
CRTMP	Crew Time Management Program
FHWA	Federal Highway Administration
INDOT	Indiana Department of Transportation
LRS	Linear Referencing System
MAF	Monthly Adjustment Factor
MDT	Montana Department of Transportation
Mn/DOT	Minnesota Department of Transportation
MTO	Ministry of Transportation of Ontario
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
ODOT	Oregon Department of Transportation
PennDOT	Pennsylvania Department of Transportation
RWIS	Road Weather Information System
SHRP	Strategic Highway Research Program
VIF	Variance Inflation Factor

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Winter weather conditions play a significant role in the operation of the surface transportation system. When water freezes on the road surface or blowing snow obscures visibility, motorist safety may be compromised. Moreover, the presence of snow or ice on the roadway can reduce travel speeds until appropriate winter maintenance activities are undertaken.

These effects are often thought of in causative ways: for example, an increase in bad winter weather conditions will result in an increase in crashes and an increase in winter maintenance costs. If these causative relationships could be quantified, this could have far-reaching impacts, including facilitating proactive approaches to winter roadway safety, assessing the cost-effectiveness of various winter maintenance resources, evaluating the effectiveness of safety projects while taking into account the impacts of unusual weather, and others.

The Western Transportation Institute, with funding from the U.S. Department of Transportation's University Transportation Center program, in partnership with the California Department of Transportation, undertook a research project to quantify the relationship between winter weather severity and highway safety. This report summarizes the findings of this research effort.

Chapter 2 provides a literature review of some of the more common indices that have been previously developed for winter weather severity. Chapter 3 presents the methodology that was used in this project, with a review of project goals, data sources, and modeling assumptions. Chapter 4 presents model forms that were developed for three states. Chapter 5 develops a method for translating these models into a weather severity index. Chapter 6 discusses the implications of the results, and Chapter 7 provides conclusions regarding future research directions.

2. LITERATURE REVIEW

At the beginning of this research project, it was unclear whether weather severity indices had been developed specifically for transportation. Therefore, it was assumed that models outside of transportation, for example, winter impacts on wildlife (1, 2, 3), icebreaking (4), and degree days for home heating and cooling (5), would be useful. However, a number of indices were found that provide interesting examples and approaches for developing a roadway weather severity index related to safety. This chapter provides a detailed write-up of each index, its purpose, its form, statistical validity, and any pertinent strengths and weaknesses[1].

2.1. Hulme (1982)

In the early 1980s, several weather indices had been developed for the summer months in England, as summer weather affected the tourism and agricultural industries. Hulme initiated work on a winter weather index, as winter weather also has some effect on industrial output in England (6). The index, WI_{Hulme}, was constructed based on data between the winters of 1929-1930 and 1980-1981 for four stations in England: Durham Observatory, Edgbaston Observatory (Birmingham), Victoria Park (Swansea) and Leuchars (Fife). The winter index for the four stations was:

$$WI_{Hulme} = 10T_{avg} - (18.5\ D_{snow})^{1/3} - D_{frost1} + 200 \qquad (2\text{-}1)$$

where T_{avg} = mean daily temperature

D_{snow} = number of days with snow lying at 9:00 (GMT)

D_{frost1} = number of night ground frosts (grass minimum temperatures below 32° F [0° C])

Appropriate weights had to be assigned to each element, which ensured that the contribution to the index was equivalent for each parameter. To do so, the distribution of each element had to be normalized, which made a cube root transformation for the number of days with snow lying at 9:00 necessary. Normal equations were solved to find the coefficients.

A low index indicated a severe winter and a high value, a mild winter, where winter is defined to be from the first of December to the end of March. The author notes that if D_{snow} = 0, the index is large, so the author suggests that when D_{snow} = 0, a value of 0.60 is assigned to the cube root of zero.

When the index was applied to the four stations, three winters showed great severity in each of the individual stations. These three winters had the lowest index values for these four stations, except at Swansea in the 1978-1979 winter where there were five other winters more severe. Correlation coefficients suggest that all winter weather factors between the four weather stations

[1] Variable labels may be altered from the original document to maximize ease of comparability between different models.

are strongly correlated, and all are statistically significant at $\alpha = 0.01$. There does appear to be some variation in the correlation coefficients however, especially between Swansea and the other stations. The correlation between Swansea and the other stations tends to decrease with the distance the station is from Swansea. This type of correlation decay may be expected, but it is also expected to occur between the other stations as well, and the correlations between other stations remain relatively large. This may indicate that winter weather in Swansea is not typical in the rest of the United Kingdom.

Trends in averages were evaluated from decade to decade for the four stations and summarized. The summary statistics reveal that the mildest winters were in the 1930s, while the most severe were in the 1960s for all of the weather stations except Swansea. Swansea statistics indicate that the 1970s were marginally mild whereas the 1940s were the most extreme. The trends from the congruous three weather stations are analogous to northern hemisphere climate trends.

Monthly analysis on the winter index was not feasible, as the winter index is based on an entire season. In addition, had different parameters been selected for utilization in the winter index, winter index values would have been different. The author suggested that another parameter of interest would be freeze/thaw cycles, which have effects upon biological and physical processes in soil and contribute to roadway deterioration. Prediction using the winter index was weak, as extrapolation of the curves formed from previous winter indices to future curves is not justified.

2.2. Rissel and Scott (1985)

The Pennsylvania Department of Transportation (PennDOT) conducted a study to determine whether maintenance manpower was effectively used in the state of Pennsylvania (7). The main objectives of the study were to determine the cost-effectiveness of single- and dual-shift staffing during the winter months; to identify maintenance activities that are not snow-related that can be performed during cold weather; to estimate the amount of work that can be accomplished with single- and dual-shifts; and to ascertain optimum winter staffing patterns. PennDOT is concerned with keeping costs down while providing a safe and effective winter maintenance system.

Research data was collected by interviewing maintenance personnel to establish staffing schedules, obtaining data from the National Oceanic and Atmospheric Administration (NOAA) to establish weather patterns, inspecting departmental records, and performing a statistical analysis.

Computer programs were written to refine the weather data provided by a weather data form collected for various weather stations and districts. The programs were also used to calculate winter maintenance work hours. The Modified Weather Program defined the duration of ice storms in order to account for instances that precipitation is observed when below-freezing temperatures were recorded for the time period, but there was no snow present. The Storm Clearing Program estimated the time to clear the roadway following a storm. The Daylight Work Program determined when there was adequate light to work on site in morning and evening hours. Finally, the Crew Time Program (CRTMP) calculated regular time and premium time required for winter procedures.

Severity of weather was previously identified by the amount of snowfall that had fallen in a given winter. However, managers realized that the amount of snow fallen may not be indicative of winter severity in terms of maintenance costs. After establishing the CRTMP, managers noted that maintenance premium time incurred in the winter did not occur during the corresponding winter classification (winters are classified as light, average, or heavy, depending upon total snowfall amounts in a winter).

The severity index was developed based on meteorological data using information based upon the programs written for the project. The severity index (SI_{Penn}) is

$$SI_{Penn} = S_{season} + 2D_{med} + D_{hvy} + D_{frost} - \frac{D_{freeze}}{2} + H_{si} \tag{2-2}$$

where S_{season} = total inches of snowfall in the period,
D_{med} = number of days with snowfall of 1 to 6 inches,
D_{hvy} = number of days with snowfall great than 6 inches,
D_{frost} = number of days with a maximum temperature (T_{max}) above 32° F and a minimum temperature (T_{min}) below 32° F,
D_{freeze} = number of days with temperatures below 32° F, and,
H_{si} = total hours in the period when snow or ice occurs.

The severity index was calculated for each weather station from November 1, 1981 to March 31, 1982. The severity index was used to determine if there was a correlation between the index and the CRTMP results. For each weather station, the total single-shift premium hours were calculated and then plotted against the severity index for the same weather station. A linear regression procedure was run and the least-squares regression line for the fit was

$$H_{prem} = -37.9 + 0.8SI_{Penn} \tag{2-3}$$

where H_{prem} = total premium hours.

The correlation between the premium hours and severity index is 0.94, which indicates that the severity index is very strongly correlated with premium hours.

The regression was run on the severity index for the whole winter period. Similar procedures could be used for a portion of the winter using similar shift hour data. A more refined severity index could also be developed statistically, as taking some of the components out of the severity index that was established reduced the correlation.

2.3. Strategic Highway Research Program (1993)

High or low winter maintenance costs cannot be based on the higher or lower efficiency of snow and ice control alone. Winter weather patterns may also explain some expenses; therefore, a winter severity index should also be capable of assessing efficiency. The authors and researchers of the Strategic Highway Research Program (SHRP) derived a new index, because they believed other indices were inadequate models for winter maintenance purposes (8). They wanted to develop a model that helped reflect the causes of frost formation, and that used few parameters, so it could be employed in other countries.

Weather data was summed from daily records from the National Weather Service (NWS) and then averaged for each month from November 1 to March 31 to eliminate the influence of the month length (the number of days could vary from month to month). The Winter Index (WI_{SHRP}) was established to be the following equation:

$$WI_{SHRP} = a\sqrt{t_{seasonindex}} + b\ln\left(\frac{S_{daily}}{10}+1\right) + c\sqrt{\left(\frac{d_{freeze1}}{T_{range1}+10}\right)} + d \tag{2-4}$$

where $t_{seasonindex}$ = average value of $t_{dayindex}$ over season ($0 \le t_{seasonindex} \le 1$)
$t_{dayindex}$ = 0, if minimum air temperature (T_{min}) is above 32° F (0° C)
1, if maximum air temperature ($T_{\max}$) > 32° F (0° C) while $T_{\min} \le$ 32° F (0° C)
2, if $T_{\max} \le$ 32° F (0° C)
S_{daily} = Mean daily values of snowfall (millimeters)
$d_{freeze1}$ = Mean daily values of the number of days with minimum air temperature at or below 32° F (0° C) ($0 \le d_{freeze1} \le 1$)
T_{range1} = Mean monthly maximum air temperature minus the mean monthly minimum air temperature (° C)

Based on other studies, T_{range1} was determined to have a similar but inverse distribution to relative humidity in the United States. Because T_{range1} is used as an effective indication of atmospheric humidity, the $c\sqrt{\left(\frac{d_{freeze1}}{T_{range1}+10}\right)}$ term was used to reflect the likelihood of frost.

To attain the coefficients for the winter index, it was necessary to assign weights to each term according to the significance of the term to maintenance costs. The weights were established to be 35 percent for both T_{index} (temperature index) and S_{daily} (snowfall), and 30 percent for frost likelihood. The coefficients were then found by taking into account the critical levels of each parameter to winter maintenance costs – values that may be indicative of typical or average storm conditions leading to winter maintenance activities in a specific area: 1.87 for T_{index}, 16.5 for S_{daily}, and 1 for $d_{freeze1}$. Assuming the other parameters are constant, values for each coefficient can be obtained by solving a system of equations. The resulting Winter Index is:

$$WI_{SHRP} = -25.58\sqrt{T_{index}} - 35.68\ln\left(\frac{S_{daily}}{10}+1\right) - 99.5\sqrt{\left(\frac{d_{freeze1}}{T_{range1}+10}\right)} + 50 \tag{2-5}$$

Index values range from –50 (most severe and a maximum level of snow and ice control) to 0 (not too severe and an average level of snow and ice control) to +50 (warm and snow and ice control is not necessary). The index reaches its maximum level when *TI*, *S*, and *N* are all zero. The index reaches its minimum level when *TI*, *S*, and *N* are all at their maximums, which is also the point at which these terms have the greatest contribution to the index.

The SHRP winter index was examined spatially and temporally to observe weather patterns in the United States. The spatial distribution of the index shows that WI_{SHRP} increases from north to south and is larger in northern latitudes in the eastern United States than in similar latitudes in the western United States. WI_{SHRP} in the Great Plains was lower than in the eastern and western portions of the United States at the middle latitudes, and in mountainous areas was lower than in similar latitudes with less mountainous areas.

Temporally, nine stations were selected to observe the variability in WI_{SHRP} from year to year from the winter of 1950 through the winter of 1988. No warm or cold trend can be seen from figures showing correlations between WI_{SHRP} and minimum temperature, snowfall, and frost, although WI_{SHRP} varies from year to year. WI_{SHRP} is shown to vary with snowfall, minimum temperature and the number of frosts; however, WI_{SHRP} is more strongly correlated with snowfall than with the other variables.

Most importantly, the WI_{SHRP} may be used to evaluate snow and ice control costs. In a study involving 40 states, the cost of average annual snow and ice control per centerline mile was plotted versus the WI_{SHRP} averaged over all available sites for each state. The paper showed that a smaller WI_{SHRP} value is associated with higher costs and that there is a strong log-linear relationship between cost and index.

Snow and ice control costs can also be viewed as a function of population density, as roads in more densely populated areas have a heavier flow of traffic and may be given higher priority for winter maintenance than those more sparsely populated areas. The following equation was obtained by using stepwise regression procedures using a significance level of 0.01:

$$C = 632.3 + 7.3P^{(-0.09WI_{SHRP})} - \left(\frac{0.19WI_{SHRP}{}^{3}}{1+P} \right) \tag{2-6}$$

where C = cost of snow and ice control ($), and
P = persons per km^2.

The regression equation explained 84 percent of the variability in the costs; therefore, the equation is a good fit to the data and may be used as a tool in cost-reduction analysis at levels including state or district, as well as at a particular maintenance location. This was done in the state of Washington for six districts and 56 weather stations; for each district, it was shown that WI_{SHRP} has an inverse relationship with expenditures. The same is shown for two particular maintenance locations in Washington. It is interesting to note that higher correlations for all stations are possible when smaller-area costs are utilized rather than state averages.

Also of interest are the amounts of sand and salt utilized and the correlation of these with WI_{SHRP}. There was not enough data available for an appropriate statistical analysis, but an analysis was done to give an indication of the relationship between sand and salt consumption and WI_{SHRP} in Duluth, Minnesota, where index values were calculated from the Duluth Airport, and Routes 302 and 303 were combined to give a relative amount of sand and salt consumption on an urban commuter route for the 1986 and 1988 winters. The relationships found between WI_{SHRP} and sand and salt consumption are:

$$Sand = 571.66 + 12.247WI - 0.347(WI + 50)^2 \quad (2\text{-}7)$$

$$Salt = 83.15 + 1.245WI - 0.054(WI + 50)^2 \quad (2\text{-}8)$$

Both appear to be related to WI_{SHRP}, as correlation values were 0.977 and 0.96 for these two equations, respectively.

Main characteristics of the climate in the United States appear to be accounted for in WI_{SHRP}. The index shows the necessity of snow and ice control and provides a cost analysis on the Winter Index and snow and ice control expenditures. Historical costs and weather data can be utilized to evaluate the efficiency and effectiveness of snow and ice control at national, state, district, or county levels.

Results of the SHRP report were preliminary, but the paper concluded that the index could be utilized in evaluating the need for winter maintenance and the efficiency of winter maintenance practices.

2.4. Knudsen (1994)

Based on the measurements of an integrated system for ice detection and weather observations, a winter index was calculated for every day and every county in Denmark (9). Special emphasis was put on changes in weather which would require winter maintenance activities (e.g. salting, snow clearing). The purpose of this paper was to compare winter index, winter activity, salt consumption, and expenses based on the salt consumption and the total expenditure for a season or a year respectively.

Six winter seasons of data, from 1986/87 to 1991/92, were taken from several sources.

The weather index (WI_{Dane}) is based on meteorological data continuously available from a large number of road weather stations covering 13 of the 14 Danish counties. WI is computed as follows:

$$WI_{Dane} = \sum_{Oct\,15}^{Apr\,15} WI_{DaneDay} \quad (2\text{-}9a)$$

$$WI_{DaneDay} = x_{freeze}\left(1 + x_{frost} + x_{refreeze} + x_{snow} + x_{drift}\right) \quad (2\text{-}9b)$$

where x_{freeze} = 1, if the road temperature is below 0.5° C at any moment within a 24-hour period, otherwise 0.

x_{frost} = the number of times the road temperature drops below 0° C, provided that it is at the same time lower than the air dew point for at least 3 hours, with an interval of at least 12 hours.

$x_{refreeze}$ = the number of times the road temperature drops below 0° C (from at least 0.5° C to –0.5° C) within a 24-hour period.

x_{snow} = 1, if a snowfall of at least 1 cm is reported within a 24-hour period, otherwise 0.

x_{drift} = 1, if some noteworthy snowdrift has occurred, otherwise 0.

The quantities x_{freeze}, x_{frost} and $x_{refreeze}$ are calculated as an average over all road weather stations within a county. The values of WI_{Dane} are calculated for every county.

The level of activity is determined by the recorded turn-outs for salting and snow clearing:

$$AI_{Dane} = \frac{r_{saltturns}}{r_{salt}} + \frac{r_{clearsnow}}{r_{snow}} \tag{2-10}$$

where AI_{Dane} = activity index for a given day
r_{salt} = number of salt routes
$r_{saltturns}$ = number of turn-outs for salting
r_{snow} = number of snow routes
$r_{clearsnow}$ = number of snow routes cleared

The salt consumption was recorded for every winter season in every county. Before it was compared with other indices, it was divided by the road area in square meters. The expenses were recorded on an annual basis and also divided by the road area.

Multiple linear regression was used to compare winter index, winter activity, salt consumption, and expenses. Regression of AI_{Dane} over WI_{Dane} on a seasonal basis resulted in $R^2 = 0.95$, but this was based on only six data points. Regression of AI_{Dane} against WI_{Dane} for every year and county, together with year/county indicator variables, showed that WI_{Dane} was the most significant variable. No R^2 value was given for this kind of analysis.

The relationship between WI_{Dane} and salt consumption was almost the same as between WI_{Dane} and AI_{Dane}, with $R^2 = 0.96$ on a seasonal basis, and $R^2 = 0.38$ for every year and every county. Comparison of WI_{Dane} and expenses showed almost no correlation. The data revealed that the expenses differ greatly by county and by year, while there is no relationship between the years. The authors suggest that the lack of correlation between WI_{Dane} and expenses may be due to the fact that every county has its own way of organizing expenses, and due to some local differences in price levels.

2.5. Decker et al (2001)

Issues arise in which the efficiency of winter maintenance practices must be measured to evaluate the cost-effectiveness of current practices, as well as improvements in winter maintenance practices with new technology and techniques. Efficiency and effectiveness of winter maintenance practices is measured based on labor, equipment, and material costs used for snow removal and a winter storm severity index. This index was developed primarily to establish a means of quality control for external contractors hired for winter maintenance purposes (10).

The winter storm severity index is an adaptation of the index adopted for the Strategic Highway Research Program (8). The formula for the winter index is the same as was used in by SHRP; however, the variable definitions were slightly different:

$$WI_{Utah} = a\sqrt{t_{seasonindex}} + b\ln\left(\frac{S_{daily}}{10} + 1\right) + c\sqrt{\left(\frac{d_{freeze2}}{T_{range2} + 10}\right)} + d \tag{2-4}$$

where $t_{seasonindex}$ = average value of $t_{dayindex}$ over season ($0 \leq t_{seasonindex} \leq 1$) [same as SHRP]

$t_{dayindex}$ = 0, if minimum air temperature (T_{min}) is above 32° F (0° C)
1, if maximum air temperature ($T_{\max}$) > 32° F (0° C) while $T_{\min} \leq$ 32° F (0° C)
2, if $T_{\max} \leq$ 32° F (0° C) [same as SHRP]

S_{daily} = Mean daily values of snowfall (millimeters) [same as SHRP]

$d_{freeze2}$ = t_{freeze} averaged over all days in study period

t_{freeze} = 0, if average daily temperature ($T_{avg} = [T_{min} + T_{\max}] \div 2$) > 32° F (0° C)
1, if $T_{avg} \leq$ 32° F (0° C)

T_{range2} = Difference between maximum and minimum daily air temperatures, averaged over period

The coefficients in the winter index are determined by particular weights and critical values of the parameters in each term that are indicative of typical weather conditions at a given location (8).

The winter maintenance metric (WMM_{Utah}) is a measure of winter maintenance efficiency that normalizes labor, equipment and material costs by the product of lane kilometers of a roadway for the responsible maintenance facility and the winter storm index for a particular storm:

$$WMM_{Utah} = \frac{C}{WI_{Utah}} \times r \tag{2-11}$$

where C = cost of snow and ice control ($)

r = number of lane-kilometers of roadway

The costs for fighting a snowstorm include labor, snow- and ice-control material costs, and capital equipment costs and are collected and archived daily at many maintenance facilities. Roadway service categories of levels of winter service in a facility's area largely contribute to the cost of snow fighting; higher roadway service categories are expected to have higher gross costs. Therefore, the lane-kilometers for higher service categories were given a greater weight than the lane-kilometers of a lower service area. The winter maintenance metric can be calculated as a daily measure utilizing daily snow-fighting costs and daily winter climate data; however, because storms may last longer than one day, WMM_{Utah} would ideally be calculated based on the average snow fighting costs and dividing by the average winter index for a particular storm.

The winter maintenance metric was tested for winter storms from the winter of 1996 through the spring of 1999 in three of the Utah Department of Transportation's facilities: Tooele, West Jordan, and Kimball Junction. Tooele and West Jordan are valley locations and Kimball Junction is a mountain location, where the frequency and severity of winter storm events is expected to be greater than in the two valley locations. Costs at the mountainous location are generally more than the two valley locations, but the unit cost efficiency is considered comparable if the gross costs are normalized by the lane kilometers and winter index.

The winter index critical values (values that may be indicative of typical or average storm conditions leading to winter maintenance activities in a specific area) that represented the storm activity in these Utah regions were 1.87 and 200, respectively. Using the same weights as used in the SHRP Index, the calculated coefficients for the winter index were a = -25.59, b = -11.50, c = -99.50 and d = 50.00. No lane-kilometer weighting or equivalence length considerations were given to any particular location, because managers at each location agreed that no one section of roadway was receiving more resources than any other. Both the daily and average WMM were calculated for each winter storm. During the 1996-97 winter season, storm duration, snowfall, temperature and total costs of snow fighting varied, but the daily and storm average WMM_{Utah} remained relatively the same with a nominal level of 0.5 to large values of more than 6.0. The study showed for the 1996-97 winter that fighting small amounts of snow leads to a smaller WMM than larger amounts, especially in late winter months when special winter crews have been released from their jobs for the season; thus, maintenance crews are more efficient with larger amounts of snow.

The study concludes that a winter maintenance metric is indeed important to maintenance managers in efficiently fighting a snowstorm. The metric can indicate when the systems used are running efficiently and under what circumstances, so managers can make wise roadway winter maintenance decisions.

2.6. Andrey et al (2001)

The goal of this study was to assess the suitability of existing and modified winter weather indices in explaining the temporal and spatial variability of salt use on highways in the Province of Ontario, Canada (11).

The performance of three indices was evaluated relative to each other and various individual weather variables. The indices considered were SHRP, Hulme and Salt Days. The salt day indicator was developed by the Illinois State Water Survey and applied in short-term forecasts of snow removal budget allocations for the Illinois Department of Transportation (12). The Illinois index is simply a count of a number of days in a month that meet certain criteria:

$$WI_{Illinois} = D_{snow} + D_{cold} \tag{2-12}$$

where D_{snow} = Daily snowfall accumulation is greater than or equal to 0.5 in (1.3 cm), and
D_{cold} = Number of days where mean daily temperature is between 15° and 30° F (-9° to -1° C).

The study timeframe consisted of five winter seasons (November 1 to March 31) from 1993 to 1998. Daily and monthly temperature (daily minimum and maximum, daily and monthly means) and precipitation (daily and monthly snowfall, number of days per month with freezing rain/drizzle, snow and blowing snow) data were obtained from the Meteorological Service of Canada. Twelve climate stations were chosen for analysis, each representing one winter road maintenance district of the Ministry of Transportation of Ontario (MTO). Road salt use was chosen as the dependent variable in the analysis. Road salt usage data were obtained from MTO. The length of the road network by road class for each MTO district was used to standardize the salt use data.

The relative performance of the SHRP, Hulme and Salt Days indices was evaluated through linear regression using provincial salt use data, standardized by road length, and climate information for winter seasons between 1993 and1998. SPSS version 8.0 was used for analysis.

Virtually no relationship was found between the Hulme index and the Salt Days index (r = 0.05). The association between the SHRP index and each of the other two indices is modest (r = 0.64 and r = -0.54, respectively). This suggests that care must be taken in choosing a winter index for a particular purpose.

The SHRP index explained more of the variation in standardized salt use than either of the other two indices. It performed adequately in all districts. However, it was only marginally better than either snowfall amount or snow days alone, and accounted for less than half of the variability in monthly salt use across 12 districts. This is much lower than in its original U.S. application.

The following modifications were made to the SHRP equation weights and constants to account for the unique Ontario climate and to assess model sensitivity. First, Ontario data were used to define $t_{seasonindex}$ and S_{daily} maxima. $t_{seasonindex}$ values as low as 1.9 and S_{daily} values exceeding 25 mm (1.4 in) occurred in Ontario during the study period. These values were used to derive new coefficients for the Ontario-based SHRP model. This resulted in the following equation:

$$WI_{Ontario} = -25.39\sqrt{t_{seasonindex}} - 23.27\ln\left(\frac{S_{daily}}{10}+1\right) - 99.5\sqrt{\left(\frac{d_{freeze1}}{T_{freeze1}+10}\right)} + 50 \qquad (2\text{-}13)$$

The second modification included changing the weighting for the three terms in the model. The weight of S_{daily} was systematically increased, while the balance between the other two terms was altered. A total of 28 different weighting schemes were tested. Correlation coefficients between salt use and modified SHRP index showed that these two modifications had virtually no effect on the ability of the SHRP model to explain the spatial and temporal variability in road salt use on Ontario highways.

The third modification involved replacing the third term of the equation, which deals with frost likelihood with a freezing rain term. Again, 28 different weighting schemes were used. The results were similar to earlier models, with correlation ranging from r = -0.723 to r = -0.531.

All iterations of the SHRP index performed better in most southern stations than northern ones. This suggests that the SHRP model may be less applicable in regions with very cold climates, because of the temperature threshold beyond which salt is neither applied not effective.

The modified SHRP indices were then compared with a standard linear regression of three individual, unweighted climate variables: snowfall amount (S_{season}), freezing rain days (D_{frrain}), and temperature index ($t_{seasonindex}$). A multiple linear regression model was developed for each of the twelve districts. The number of weather variables was restricted to S_{season}, D_{frrain} and $t_{seasonindex}$ in order to avoid multicollinearity. To account for non-weather factors, two additional models were developed based on roadway characteristics. The first model includes $t_{seasonindex}$, S_{season}, D_{frrain} and the proportion of roads that are Class 3 or 4 ($x_{roadclass}$). The second model includes $x_{roadclass}$ with the best SHRP index (r = -.723; Ontario thresholds used for $t_{seasonindex}$ and S_{season}, weights of 5 percent, 75 percent and 20 percent for $t_{seasonindex}$, S_{season} and D_{frrain}, respectively). The resulting equations are:

$$WI_{Ontario2} = -0.901 + 0.172\, S_{season} + 0.197\, D_{frrain} + 1.362\, t_{seasonindex} - 2.032\, x_{roadclass} \quad (2\text{-}14)$$

$$WI_{Ontario3} = 5.400 - 0.101\, WI_{Ontario} - 1.525\, x_{roadclass} \quad (2\text{-}15)$$

The overall r-values were nearly identical (r = -.743 and r = -.745, respectively), and only marginally higher than earlier models. The R^2 table for individual terms showed that road class information contributed very little to the performance of the models.

One interesting aspect of the study was that the authors obtained data from every climatic zone in the province. Results of the study suggest that separate models should be developed for each climatic zone. At the same time, a weakness of the study was in the fact that only one weather station was used to represent each climatic district.

It is also noted that, as suggested in the discussion section of the article, a unique model can be developed for every part of the winter as the contribution of factors varies in the early, mid and late season.

2.7. Andrey et al (2003)

A study was conducted by the University of Waterloo and Meteorological Service of Canada to examine temporal variations in weather-related collision and injury risk using collision and weather data for Ottawa, Canada over the period 1990-1998 (13). The objectives of the study were to estimate and compare the risk of collision and injury during precipitation relative to normal seasonal conditions – both overall and disaggregated by time period. Temporal comparisons included: weekend versus weekday, daytime versus nighttime, daytime peak period versus daytime non-peak period, and early-winter season versus late-winter season.

This study was based on analysis of weather and collision records for a nine-year period (1990-1998) for Ottawa, Canada. Weather data were obtained from the Meteorological Service of Canada for MacDonald-Cartier International Airport. Hourly and six-hour precipitation amounts, and hourly observations of both weather conditions (e.g., rain, snow, fog, blowing snow) and visibility, were used to define events and controls. Collision data were obtained from Transport Canada's national collision database (TRAID3). Data were extracted for all reportable collisions that occurred within Ottawa for the years 1990-1998.

Matched-pair analysis was used to examine the relationship between collision statistics and weather. Each time period during which precipitation occurred was paired with a control time period where precipitation and other inclement weather conditions did not occur. The event and control were spaced exactly one week apart and matched on duration, time of day and day of week. They were of variable length, reflecting the nature of the storms that occurred. Events and controls were selected based on a specified set of criteria, producing a set of 771 matched pairs, with between 61 and 100 event-control pairs for each of the nine calendar years.

The relative risk of collision was calculated by dividing the total number of collisions in the event periods by the total number of collisions in the control periods. Injury risk was calculated similarly, except each injured person was counted separately.

No adjustment for traffic volume was made, since a study by Doherty *et al.* (14) showed that, on average, traffic volumes in another mid-sized Canadian city dropped only two percent during rainfall relative to fair-weather conditions.

Relative risk was modeled with the binomial distribution in GLIM (Generalized Linear Interactive Models) software. Risk estimates were produced using 95 percent confidence intervals. Four different models were developed: the relative risk of injury during rainfall, the relative risk of injury during winter precipitation, the relative risk of collision during rainfall and the relative risk of collision during winter precipitation. Summary data on the weather conditions and collision experiences during the events and controls are provided in the report.

Overall, collision risks more than doubled during rainfall events and increased by approximately 50 percent during snow events relative to normal seasonal conditions. The corresponding increases for injury risks were approximately 70 percent and 20 percent. In all cases, confidence intervals were quite narrow, and risk ratio estimates were significantly above 1.0.

The greater increases for rainfall relative to snowfall may be explained by the fact that snowfall accumulations were available for six-hour periods only (rather than hourly). Thus snowfall events may include periods of time when precipitation was absent. In contrast, defining rainfall events on a variable hourly basis seemed to isolate and concentrate the effects of weather.

Risk estimates were broken down by time of day, traffic condition, day of the week and time of year. In most cases, differences were not statistically significant. There are two exceptions. First, winter-precipitation collision rates were found to be significantly higher on weekends (Friday to Sunday) than on weekdays (Monday to Thursday). Second, winter precipitation risk ratios for collisions were higher in November-December (early winter) than in January-April. Although not statistically significant, weekend rates were also higher for rainfall collision rates, rainfall injury rates and winter-precipitation injury rates. The results by time of day were variable.

In summary, weather does appear to interact with some situational factors, resulting in especially high risk levels during precipitation that occurs on weekends and at the beginning of the snow season.

As the authors point out, an advantage of the matched-pair design is the fact that this design controls for many non-weather related variables that are time-dependent and that affect collision risk. A disadvantage of this approach is that in ignoring traffic volumes, the study's results may be skewed, as several studies cited in the paper indicated that volume reductions on highways can be substantial (from 10 to 50 percent) during heavy snowfalls.

2.8. McCullouch et al (2004)

The Indiana Department of Transportation (INDOT) conducted a study to develop a winter severity index (15). The index could be used to compare the efforts of snow and ice removal between the different climatic zones in Indiana, compare and analyze mild and severe winters, and to provide a quantitative method for determining what relationships exist between different weather events and snow and ice removal. The goal of the study was to derive an index that did not require cumbersome and time-consuming data collection.

A number of existing weather indices were analyzed to determine if one or a combination of indices could be used for INDOT. The first index considered by the study group was Wisconsin Index, used to measure the type of winter in Wisconsin and to evaluate performances and expenditures in snow and ice removal by county (16). The Wisconsin Index has the form

$$WI_{Wisc} = 10\frac{E_{snow}}{63} + 5.9\frac{E_{frrain}}{21} + 8.5\frac{S_{season}}{314} + 9.4\frac{H_{si}}{1125} + 9.2\frac{E_{incidents}}{50} \quad (2\text{-}16)$$

where E_{snow} = Number of snow events,
E_{frrain} = Number of freezing rain events,
S_{season} = Snow amount,
H_{si} = Total storm duration,
$E_{incidents}$ = Number of incidents (drifting, cleanup, and frost runs).

Next, the modified Hulme index was analyzed (17). The equation is

$$WI_{HulmeMod} = 10\,T_{\max} - D_{frost1} - (18.5\,D_{snow})^{1/3} + a \quad (2\text{-}17)$$

where T_{max} = Mean maximum air temperature,
D_{frost1} = Total number of ground frosts,
D_{snow} = Number of days with snow cover at 9:00 AM,
a = Constant.

Two more indices, the SHRP index and the modified SHRP index for the Ontario Province, were considered. Plots of these indices against the cost/mile of snow and ice removal for one of Indiana's maintenance units showed no correlation between the two. Therefore, none of the existing indices were found useful, and a new index for INDOT was developed.

For this study, Indiana was divided into four different winter climatic zones. NOAA weather data was collected for one city in each zone: Evansville, Fort Wayne, Indianapolis, and South Bend. Weather data was obtained for five winter months – November through March – for the winter seasons of 1999-2000, 2000-01, 2001-02 and 2002-03.

The authors conducted a survey of field crews and employees involved in snow and ice removal, in order to identify weather factors with the most influence on snow and ice removal. Four such factors were identified, and were incorporated into the following model (which was calibrated using multiple regression analysis in SAS):

$$WI_{Indiana} = 100 * 0.06\ D_{frost2} + 0.29\ D_{frrain1} + 0.38\ D_{drift} + 0.27\ (D_{snowevent}) \tag{2-18}$$

where D_{frost2} = Number of frost days (i.e. $T_{min} \leq 32°$ F and minimum dew point ≤ 32° F).

$D_{frrain1}$ = Number of freezing rain days (i.e. number of days with freezing rain and/or drizzle and $T_{min} \leq 32$ ° F)

$D_{snowevent}$ = Number of snow event days, where a snow event day is defined as the number of days with $T_{min} \leq 32°$ F times the snowfall intensity divided by the average temperature during the event

D_{drift} = Number of drifting days (i.e. number of days with wind speeds > 15 mph and snow on ground or a snow event)

Lane mile snow removal costs were used to validate the $WI_{Indiana}$ by analyzing the correlation between $WI_{Indiana}$ and snow removal costs. The above four-factor equation was not satisfactory, because of suspected bias in field crew responses and skewed results. Thus a multiple regression analysis was performed based on regional weather data, and three additional factors were included in the model: snow depth (S_{depth}); storm intensity (H_{storm}), which in fact measures storm duration; and average temperature (T_{avg}). A separate equation was established for each weather climatic zone in Indiana. Also, a statewide equation was generated.

$$\begin{aligned} WI_{SouthBend} = {} & -5.98483\ D_{frost2} + 13.73518\ D_{frrain1} + 12.57288\ D_{drift} \\ & - 25.18103\ D_{snowevent} + 28.78145\ S_{depth} + 4.29121\ H_{storm} + 6.77877\ T_{avg} \end{aligned} \tag{2-19a}$$

$$\begin{aligned} WI_{FortWayne} = {} & 7.05832\ D_{frost2} - 16.21024\ D_{frrain1} + 6.31394\ D_{drift} \\ & + 31.24970\ D_{snowevent} + 25.36240\ S_{depth} + 1.23828\ H_{storm} - 6.95440\ T_{avg} \end{aligned} \tag{2-19b}$$

$$\begin{aligned} WI_{Indianapolis} = {} & 3.42152\ D_{frost2} + 7.96888\ D_{frrain1} + 7.24260\ D_{drift} \\ & + 14.044284\ D_{snowevent} + 16.63333\ S_{depth} + 1.50251\ H_{storm} - 3.90486\ T_{avg} \end{aligned} \tag{2-19c}$$

$$\begin{aligned} WI_{Evansville} = {} & 0.01116\ D_{frost2} + 23.6838318\ D_{frrain1} + 43.46891\ D_{drift} \\ & - 18.77938\ D_{snowevent} + 63.02214\ S_{depth} + 0.23399\ H_{storm} - 0.32291\ T_{avg} \end{aligned} \tag{2-19d}$$

$$\begin{aligned} WI_{Indiana2} = {} & 0.71839\ D_{frost2} + 16.87634\ D_{frrain1} + 12.90112\ D_{drift} \\ & - 0.32281\ D_{snowevent} + 25.72981\ S_{depth} + 3.23541\ H_{storm} - 2.80668\ T_{avg} \end{aligned} \tag{2-19e}$$

The correlation between the above weather indices and the cost/mile of snow and ice removal was determined by means of graphically comparing the temporal variation of the costs and each weather index. The authors conclude that the correlation between these models and the snow removal costs was good, with the exception of the statewide model. One significant shortcoming of the report is that the authors did not include any quantitative measures of correlation they obtained, such as R^2. The graphs presented in the report may be interesting, but from a statistical viewpoint, a residual plot together with some numeric measurements of the goodness of fit would be more appropriate.

The strength of the INDOT project is the idea of obtaining data from every climatic zone in the state. This approach appears very promising, especially in light of the good results obtained by INDOT. Surveying the field crew or a similar technique could also be used in developing a weather index for other purposes, such as for predicting highway accident rates. However, the results of the INDOT project are not in favor of using this approach, due to a significant bias and skewed results of such a survey.

Another interesting idea is defining weather *events*, rather than concentrating purely on recorded weather data. Such weather events may be defined using simultaneous or consecutive weather data.

There were also several weaknesses in the INDOT project. First, the resulting models are only valid for the state of Indiana; no attempt was made to generalize them to other locations. Next, no general model was obtained even for Indiana, but rather separate equations for each climatic zone. The statewide equation above does not offer good fit of the data. To construct a general equation with a good fit, a climatic zone classification variable could be included in the model.

In addition, only linear multiple regression was used. Higher-order effects and interactions were ignored. However, in this project, good correlation with the snow and ice removal data justifies this limited approach.

2.9. Haider (2004)

The Minnesota Department of Transportation (Mn/DOT) has been seeking to develop a winter severity index based on the index developed in Wisconsin. Their formula is as follows:

$$WI_{MinnRaw} = 10 \times \frac{E_{snowevent}}{57} + 5.9 \times \frac{E_{frrain}}{19} + 8.5 \times \frac{S_{season}}{106} + 9.4 \times \frac{H_{si}}{29.05} \tag{2-20}$$

where E_{snow} = Number of snow events,
E_{frrain} = Number of freezing rain events,
S_{season} = Snow amount (inches), and
H_{si} = Total storm duration (hours)

The raw index was scaled to develop a maximum severity index value of 100:

$$WI_{Minn} = \frac{WI_{MinnRaw}}{0.338} \tag{2-21}$$

Although Mn/DOT has calculated index values for several winters, the index is still in its early stages of development (18).

2.10. Nixon and Qiu (2005)

A study was conducted at the University of Iowa to develop a storm severity index that evaluates to what extent an individual storm creates challenges for maintenance activities (19). This relates to the Indiana model which considered the effects of snow events on costs.

Three steps were taken in the development of the index. First, various storms were classified by six factors, including pre- and post-storm conditions, surface temperature, wind speed, and precipitation type. Second, a multiple regression model was built to produce a storm severity index between 0 and 1. Third, representative storms were ranked in severity by winter maintenance personnel and the model was modified to reflect this ranking.

The first step in developing an index for individual storms was to develop a method of describing storms. A simplified version of the description developed by Nixon and Stowe (20) was adopted. Nixon and Stowe describe storms in terms of the six factors mentioned earlier. For this study, the following modifications were made:

1. Only four possible storm event types were considered: heavy snow (>6 inches in 24 hours), medium snow (2 to 6 inches), light snow (<2 inches), and freezing rain. It was decided to remove frost as a possible event type.
2. The in-storm wind condition was incorporated as another factor.
3. The number of temperature ranges was reduced from four to three: cold (< 25°F), mid (25° to 32°F), and warm (> 32°).
4. The post-storm conditions were simplified into two categories – light wind (<15 mph) and strong wind (>15 mph) – rather than four.

The equation format for the storm severity index was based on the SHRP index. The general form of the equation was taken as:

$$SSI_{Iowa} = \left[\frac{1}{b} \left[(E_{storm} \times T_{roadduring} \times W_{during}) + B_{before} + T_{roadafter} + W_{after} - a \right] \right]^{0.5} \tag{2-22}$$

where SSI_{Iowa} = storm severity index,
E_{storm} = storm type (1 = freezing rain, 2 = light snow, 3 = medium snow, 4 = heavy snow),
$T_{roadduring}$ = in-storm road surface temperature (cooling, same or warming),
W_{during} = in-storm wind condition,
B_{before} = early storm behavior (rain or no rain),
$T_{roadafter}$ = post-storm temperature (cooling, same or warming),
W_{after} = post-storm wind condition,
a, b = parameters to normalize the index from 0 to 1.

The factor coefficients were then estimated. Values that were used as an initial approximation of these values were obtained by studying the Federal Highway Administration's (FHWA) Manual of Practice recommended treatments (21) and comparing how (for example) road temperature impacts treatment amounts and frequency.

The storm severity scores were calculated for 252 different storms based on the initial equation form and coefficients. Then the initial scores were modified (by changing a and b) so that the computed storm severity index values have an approximately normal distribution.

To test the accuracy and reliability of the model, ten representative storm scenarios were selected out of 252 possible storm events and described in a survey form. Over three dozen maintenance supervisors in Iowa ranked these ten scenarios according to the level of difficulty that these events would pose to them in their maintenance activities. Since there was less than perfect agreement between the initial storm severity index and the supervisor rankings, the index was adjusted according to the supervisor rankings. The resulting factor coefficients for the index are shown in Table 2-1.

Table 2-1: Factor Coefficients for Storm Severity Index

Factors	Values	Scores
Storm type	Freezing rain	0.72
	Light snow	0.35
	Medium snow	0.52
	Heavy snow	1
Storm temperature	Warm	0.25
	Mid range	0.4
	Cold	1
Wind conditions in storm	Light	1
	Strong	1.2
Early storm behavior	Starts as snow	0
	Starts as rain	0.1
Post storm temperature	Same	0
	Warming	-0.087
	Cooling	0.15
Post storm wind conditions	Light	0
	Strong	0.25

(Source: 19)

Plots of the distribution and density function of the storm severity index are given in the article. The plots show that the distribution of the index is approximately normal. In its final form, the storm severity index provides a measure of the severity of any given storm based solely on the meteorological description of that storm.

The approach to developing a severity index taken by the authors has several advantages. First, the idea of using a classification of factors by several value ranges and assigning a weight to every range makes the model easy to interpret and easy to use. The inclusion of interactions between factors and the near normality of the resulting storm severity index are also among the strengths of this model.

It is particularly noteworthy that the final model is largely based on the opinion of the experts in the field, which was assessed by means of a survey. In other words, the model takes into account the experience of road maintenance crews, which makes it more useful for practical purposes.

Among the shortcomings of this model, the model was not tested for applicability in other states and other climatic regions, or how it can be adjusted for use in those regions. In addition, the exact methods of obtaining parameter estimates are not discussed in the article. Insufficient argument was given in favor of the selected model form; however, the authors do point out that this model is just one of the possible ways to quantify storm severity.

2.11. Road Sense Index (1994)

This project, funded by the Insurance Corporation of British Columbia, sought to develop a driving risk index similar to an ultraviolet exposure index which had been developed for the Greater Vancouver area (22). The index would be developed to coincide with peak commute hours, and was intended to alert motorists about hazardous conditions and adjust their behavior – including not making a trip – accordingly. This project differs from other reviewed studies in that it sought to correlate weather parameters with safety – a similar objective to the present study.

To develop the index, the researchers used hourly weather data for Vancouver Airport from 1986-1991 as provided in an Environment Canada climatological data base. Crash data was summarized over hourly intervals during the entire period. A multiple linear regression was used with all weather variables, time of day, day of week, month and holiday set as independent variables. The dependent variable was the natural logarithmic transformation of the number of crashes plus one. Other transformations were explored, but the natural log yielded the best R^2 value.

The SAS generalized linear model procedure was used to conduct the analysis. Of approximately 30 variables in the weather data set, the variables shown in Table 2-2 were identified as significant. Coefficient estimates are provided in the paper. The R^2 for the model was 0.5949, with traffic conditions (represented by time of day, day of week, month and holiday status) accounting for most of the variations. When the model was applied only to casualty accidents, the R^2 declined to 0.3761, so the road sense index was developed upon the total number of crashes.

Table 2-2: Significant Variables in Road Sense Index

Continuous	Dummy	Categorical	Interaction
• Relative humidity • Sea level pressure • Visibility • Wind speed	• Fog • Holiday	• Day of Week • Month • Rain • Rain Showers • Snow • Snow Showers • Time of Day • Thunderstorm	• Temperature × Season

To develop the index, the SAS FASTCLUS procedure was used to develop 20 clusters, which were manually adjusted to form a simpler 10 point index. In general, higher index values correlate with afternoon peak commute traffic. A holiday would reduce the index by 2 values.

Storm activity would tend to raise the index value. Light rain could raise the index by 2 values, heavy rain or thunderstorms by 3 values, and snow by 4 values.

The index was implemented by forecasting values for morning peak and afternoon peak hours, and setting a value 90 minutes before the start of the peak period. While index values were developed for three-hour periods, one value was selected for the entire period for simplicity in communicating information to the public. The worst case impact of traffic over each three-hour period, along with the worst case impact of weather, were combined to develop the index value (even if those hours were different). The index model was transferred to Microsoft Excel for ease of application.

After initial application, there were comments that the afternoon peak index (because of higher traffic) seemed relatively high compared to the morning index value. In addition, there was concern about differences in road conditions based on localized different weather conditions in the Greater Vancouver area. Finally, weather seemed to have little effect on index values from day to day, which seemed to undermine the public's confidence in the index.

To address these challenges, public outreach was used to explain that the index reflected both weather and traffic conditions. Second, there was a public perception that worsening weather should increase crash risk, but the data showed a leveling off of crash rates in worsening weather as drivers adjust their behavior. To compensate for that, weather and temperature factors were used to amplify the index values to better match public perceptions of risk. Weather factors ranged from 1 to 1.4, and temperature factors ranged from 1 to 1.13. The product of both factors could not be greater than 1.45. An "extraordinary condition" adjustment factor of 1.4 was set for conditions which were without precedent in the model's development. However, either the temperature and weather factors or the "extraordinary condition" factor could be included, but not both. To account for local variability, additional options were added in the Excel application for higher index areas. Additional differentiation was added in the index values using a half-point scale to help boost credibility.

The final index values, with frequency of observations and accident frequency, are shown in Table 2-3.

Table 2-3: Road Sense Index Values

Index Value	Frequency	Average Reported Accidents
1	625	0.93
1.5	4,589	1.18
2	2,149	2.19
2.5	4,889	3.07
3	4,090	3.96
3.5	3,016	5.15
4	6,272	6.73
4.5	3,526	7.49
5	5,055	8.72
5.5	2,529	9.47
6	4,020	10.40
6.5	2,023	11.32
7	2,922	12.27
7.5	1,473	13.92
8	1,829	14.85
8.5	1,583	15.95
9	1,218	19.38
9.5	626	25.26
10	125	34.5

2.12. Summary

The indices in this chapter show a variety of different approaches to developing winter indices. There are a couple of key observations that were used in helping to direct further work on this research project.

First, the weather indices focused primarily on correlations with maintenance cost and activity. As decision support tools for winter maintenance have improved resulting in increased standardization of winter maintenance practices, such correlations would be expected. The relationship between winter weather and safety, however, appears to be relatively unexplored.

Second, the review models used a variety of approaches regarding the definition of variables. Some models used continuous variables, others used categorical or dummy variables, while some used both. Some models developed variables on the basis of events or storms, while others used aggregate information over an entire season. Most models used a simple multiple linear form, while others used interaction terms or transformations of either the independent or response variables. Given that the multitude of approaches appeared to be successful, there does not appear to be a universal relationship between specific weather variables and other factors.

Third, models have been developed over a variety of geographic scales, with some analyses finding that a single model was not appropriate for application over larger geographic areas. While it would be desirable to identify a universal relationship, this may not stand up to statistical scrutiny.

Fourth, the models were developed using a variety of levels of statistical rigor. Documentation of some indices did not indicate whether common statistical techniques were used to assess the validity of the concluding model forms (for example, significance of variables, multicollinearity, etc.). If models are to be implemented to help in planning or advisory roles – like the Road Sense Index – rigorous statistical analysis is required to ensure that the resulting recommendations may be interpreted correctly.

3. METHODOLOGY

Based on the findings of the literature review, there is room for the development of a winter severity index that focuses on relationships between crash data and winter weather. This chapter summarizes the methodology that was used in this research project to quantify the relationships, focusing on project goals, data sources and modeling assumptions.

3.1. Project Goals

The initial objective in this research was to develop a weather severity index that could correlate with safety or maintenance costs, and that could reflect unusual weather conditions throughout the year. As the project advanced, this goal was trimmed back to focus on winter weather. Because of the number of studies which have correlated winter weather with winter maintenance costs, this research project focused on the relationship between winter weather and safety. The purpose of this project, then, was to develop an index for roadway weather severity during the winter months that corresponds to crash likelihood.

3.1.1. Intended Applications

This research is more basic in nature, and it was unclear at the outset who would be the users of the resulting index. Examples of potential applications include the following:

- Use to support before-after safety studies and correct for the effects of variable winter weather conditions
- Use to support pre-trip traveler information and provide motorist warnings when conditions may tend to increase safety risk
- Use as a decision support tool in winter maintenance operations by highlighting circumstances where motorist safety may be most compromised

With the possible exception of the first case, ease of interpretation of the final index value is critical.

3.1.2. Statistical Approach

Later sections in this chapter will discuss the specific statistical approach in greater detail. However, there are certain broad parameters that governed the research team's approach to this project.

Isolate Weather from Other Factors

As shown in Figure 3-1, crashes may occur from many factors acting independently or cooperatively. A vehicle may have poor braking characteristics or a distracted driver, which can increase the risk of crashes. A roadway may have narrow clear zones, unexpectedly tight turns, poor visibility locations or other factors that tend to increase crash rates. Similarly, weather may have a variety of effects on safety, such as reducing visibility, roadway friction, or the ability to control a vehicle. These factors may have compounding effects as well, such as when there is ice on a curved section of highway. This research seeks to isolate the effects of weather from other factors in how it influences highway safety.

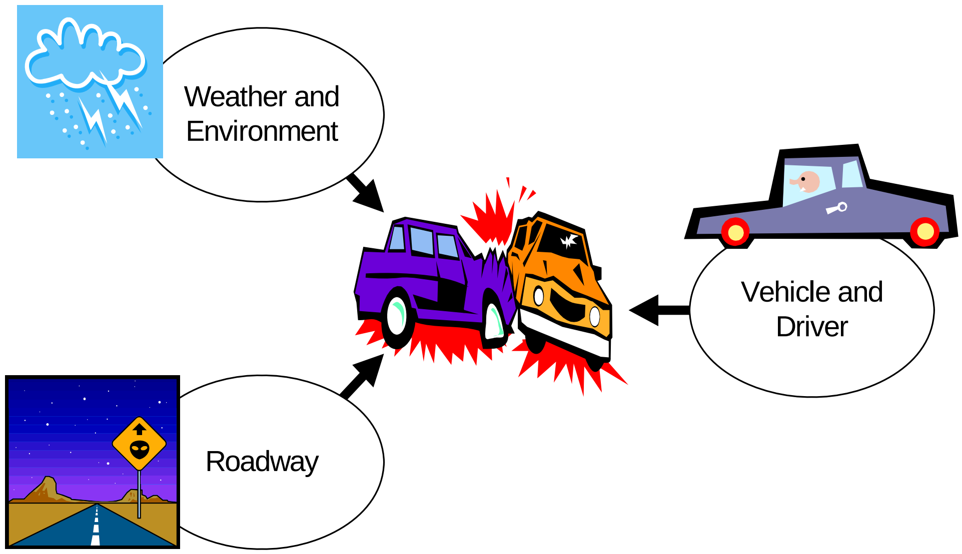


Figure 3-1: Contributing Factors to Crashes

To isolate the effects of weather, it is important to compare crash statistics where the other factors – vehicles/drivers and roadway – are held fixed. This can be accomplished through looking at specific highway segments or using short durations of time. However, even this may be complicated if, as was found in earlier cited studies, drivers alter their behavior during bad weather, or perhaps even decline to make a trip.

Use Readily Available Data

To maximize its utility, it would be important for the model to use weather data which is easily accessible to potential users. This would include data sources such as NWS or RWIS data (described in Section 3.2).

Identify Significant, Explanatory Variables

As was described in Chapter 2, many statistical approaches seek to maximize statistical correlation (R^2) between weather parameters and the response variable. The research team felt that this approach may lead to models which show successful correlation between data but not causation. To this end, models were selected not simply to maximize R^2, but were instead chosen to highlight significant variables that seemed to have explanatory power and the potential for transferability.

Use Simple Functional Form

In general, the models presented in Chapter 2 have simple functional forms that lend to easier comprehension and interpretation. Our approach is the same. While we are open to complex interactions between variables and transformations of different variables, a substantial positive effect on correlation must be demonstrated before inclusion in the model. The preference will be to use linear or multiplicative forms, with simple transformations of the response variable or individual independent variables.

3.2. Data Sources

For this project, the research team obtained and analyzed data from three states: California, Oregon and Montana. For each of these states, the following types of data were used: 1) weather data for selected locations, either from road weather information system (RWIS) sensors or from the NWS weather stations; 2) accident counts for selected segments of highways; 3) annual average daily traffic (AADT) counts for selected locations or, if unavailable, for other locations in close proximity; and 4) monthly adjustment factors (MAF) to average daily traffic counts, for locations where traffic counts are available, or for other locations in proximity of the traffic count locations.

3.2.1. California

As an initial proof of concept for this research, and because the motivation for this research project was borne out of rural areas in northern California and southern Oregon, the research team selected California State Route 299 for analysis. Weather data for California were obtained from stations with archived data available through the NWS website (23). The extracted weather files were in text format and corresponded to winter months (October through March) from October 1991 until March 2000. Weather records were obtained for the fifteen weather stations that are closest to State Route 299; these are listed in Table 3-1. It should be noted that most of these weather stations were not immediately adjacent to the roadside.

Table 3-1: California Weather Stations

Weather Station	County
Eureka	Humboldt
Willow Creek	Humboldt, Trinity
Big Bar E	Trinity
Weaverville	Trinity
Trinity River	Trinity
Whiskeytown	Shasta
Redding	Shasta
Round Mountain	Shasta

Weather Station	County
Buckhorn	Shasta
Hat Creek	Shasta
Pit River P H 1	Shasta, Lassen
Adin RS	Lassen, Modoc
Canby 3 SW	Modoc
Alturas	Modoc
Cedarville	Modoc

The data contained daily measurements, taken once a day, of six weather parameters, listed in Table 3-2.

Table 3-2: California Weather Variables

Variable Name	Description	Format	Units
PRCP	Daily precipitation	Numeric	in to .01
SNOW	Daily snowfall	Numeric	in to .1
TMIN	Daily minimum temperature	Numeric	° F
TMAX	Daily maximum temperature	Numeric	° F
TOBS	Temperature at observation time	Numeric	° F
SNWD	Snow depth at observation time	Numeric	in to 1.0

Crash data for California State Route 299, obtained from FHWA's Highway Safety Information System, contained accident records for all of State Route 299, and covered the period of time from January 1991 until December 1999. The variables included in the crash data are described in Table 3-3.

Table 3-3: California Crash Variables

Variable Name	Description	Format	Units
CNTY_RTE	County and route codes	String	
MILEPOST	Ramp milepost	Numeric	mile
RODWYCLS	Roadway classification	Character	
RTE_NBR	Route number	Numeric	
HOUR	Time of the accident	Numeric	hhmm
COUNTY	County code	Numeric	
HWY_GRP	Highway group (divided/undivided)	Character	
FILETYP	File type	Character	
SDE_HWY	Side of highway	Character	
WEATHER	Weather condition (rain, snow, fog, etc.)	Character	
LIGHT	Light condition	Character	
RDSURF	Road surface (slippery/muddy)	Character	
ACCTYPE	Type of collision	Character	
NUMVEHS	Total number of vehicles	Numeric	
POP_GRP	Population group (city/rural group)	Character	
SEVERITY	Collision severity	Character	
VEH_INVL	Number of vehicles involved	Numeric	
CASENO	Case number (includes date of crash)	String	yyyymmdd + case_number

For each weather station, a location along State Route 299 close to that station was selected. AADT volumes for 1991 through 2000 were requested from Caltrans for all such locations. Monthly adjustment factors for traffic counts, to account for seasonal variability in traffic volumes, were available for only one location on State Route 299, in Humboldt County, and these factors were calculated from traffic data for 2001-2003. Based on the available data, the research team assumed that seasonal variation in traffic volumes is the same at all locations on State Route 299 through the 10-year period of interest.

3.2.2. Oregon

Oregon weather data were received from Oregon Department of Transportation (ODOT) in XML format, arranged by month and year, and contained weather records from RWIS stations in Oregon. The included RWIS stations for Oregon are listed in Table 3-4. RWIS is a term that encompasses the sensing and collecting of on-site weather and road-condition information, the processing and dissemination of the information, and the creation and dissemination of forecasts of road and weather conditions (24). RWIS deployments typically include pavement sensors which discern pavement temperature, the water phase on the pavement's surface (water, ice, or dry pavement) and the concentration of chemical deicer applied to the road; and atmospheric sensors, which can detect air temperature, relative humidity, wind speed and direction, precipitation type, intensity and rate, and visibility perceived by the driver (25). Some sensors are also used to measure barometric pressure and solar radiation (26). Unlike NWS collection sites, RWIS are located on the highway, and are thus expected to better represent conditions experienced by drivers. The RWIS data covered the period from July 1997 to May 2004. For the purposes of the project, records that correspond to winter months, from October 1997 to December 2003, were selected. The remainder of the data was discarded.

Table 3-4: Oregon RWIS Stations

SysID	StnID	Name	Latitude	Longitude	Elev (Ft)
24	0	Bend Pkwy / North Canal (US 97 MP 135)	44.078	-121.304	3,550
24	2	Willamette Pass (Ore 58 MP 62.3)	43.598	-122.035	5,080
24	3	Bly Mountain (Ore 140 MP 30.6)	42.343	-121.399	4,920
68	0	Interstate Bridge South Span (I-5 MP 308.3)	45.617	-122.676	75
68	1	Interstate Bridge North Span (I-5 MP 308.8)	45.621	-122.673	75
68	3	Glen Jackson Bridge North Channel (I-205 MP 26.4)	45.591	-122.547	55
68	8	Timber Junction (US 26 MP 37.7)	45.760	-123.301	780
68	9	I-205 at Division (I-205 MP 19.6)	45.502	-122.565	270
68	10	Brightwood Weigh Station (US 26 MP 36.5)	45.376	-122.035	1,070
100	0	Bend (US 97 MP 134)	44.090	-121.307	3,547
100	1	Lava Butte (US 97 MP 999)	43.914	-121.351	4,508
100	5	Shaniko (US 97 MP 56)	45.003	-120.753	3,340
110	0	Yaquina Bay Bridge (US 101 MP 141.5)	44.624	-124.058	120
281	0	Celilo West of The Dalles (I-84 MP 96)	45.643	-120.979	225
281	1	John Day River Bridge (I-84 MP 114.41)	45.729	-120.652	305
281	2	Wasco Junction (US 97 MP 7.5)	45.609	-120.725	1,100
281	3	Kent (US 97 MP 40.9)	45.194	-120.696	2,705
281	4	Mount Identifier (US 97 MP 61.6)	44.990	-120.849	3,480
282	1	Ladd Canyon at the Summit (I-84 MP 273.4)	45.188	-117.992	3,630
282	2	Weatherby Rest Area (I-84 MP 336)	44.496	-117.368	2,390
282	3	Battle Mountain (US 395 MP 39.7)	45.271	-118.978	4,250
283	0	Siskiyou Summit (I-5 MP 4.5)	42.063	-122.603	4,300
283	1	Hayes Hill (US 199 MP 16.4)	42.330	-123.591	1,640
283	2	Sexton Summit (I-5 MP 69)	42.601	-123.384	1,980
283	3	OBrien (US 199 MP 41)	42.001	-123.723	1,700
283	4	Medford Viaduct (I-5 MP 28.9)	42.331	-122.871	1,360
283	5	Port Orford (US 101 MP 301.5)	42.743	-124.490	90
283	6	Highway 42 at Coos County Line (Ore 42 MP 42)	42.965	-123.868	500
283	7	Diamond Lake (Ore 138 MP 83)	43.127	-122.132	5,260
285	1	Enchanted Way/South Salem (I-5 MP 247)	44.824	-123.019	500
285	2	North Albany (I-5 MP 236)	44.670	-123.059	240
285	3	Wards Butte / Cottage Grove (I-5 MP 170)	43.751	-123.115	720

Weather observations were recorded by RWIS every few minutes. The interval between observations ranges from 1 to 15 minutes. Table 3-5 lists the variables included in every RWIS observation.

Table 3-5: Oregon Weather Variables

Variable Name	Description	Format	Units
SystemID	Region code	integer	
StationID	RWIS station code	integer	
Location	Description of location	string	
Latitude	Latitude	integer	degrees
Longitude	Longitude	integer	degrees
Elevation	Elevation	integer	ft
AirTemp	Air temperature	numeric	° F
DewPtTemp	Dew point temperature	integer	° F
RelHumdy	Relative humidity	integer	%
BaroPrs	Barometric pressure	numeric	
WindSpdAvg	Average wind speed	numeric	mph
WindSpdGust	Wind gust speed	numeric	mph
WindDirMin	Minimum wind direction	character	
WindDirAvg	Average wind direction	character	
WindDirMax	Maximum wind direction	character	
PrecipType	Precipitation type	string	
Visibility	Visibility	numeric	mi
PrecipRate	Precipitation rate	numeric	in
PrecipAccum	Precipitation accumulation	numeric	in
PrecipIntense	Precipitation intensity	string	
DateTime	Date and time	string	

Accident records for the state of Oregon, also requested from ODOT, were obtained in text comma-delimited format. Only crashes within five miles of an RWIS station along the same highway were requested. Thus, the accident data contained records for ten-mile highway sections, and covered the time period from 1997 to 2003. The data factors recorded for every accident are given in Table 3-6.

Table 3-6: Oregon Crash Variables

Variable Name	Description	Format	Units
DATE	Date of accident	string	mm/dd/yyyy
HWY	Route number	integer	
LOCATION	Description of location	string	
MP	Milepost	numeric	
PAVEMT	Pavement condition	character	wet or dry
VEHNUM	Number of vehicles involved	integer	
FATAL	Fatalities recorded (yes/no)	character	yes or no
INJURY	Injuries recorded (yes/no)	character	yes or no

Both AADT and MAF counts were taken from Oregon automatic traffic recorder (ATR) tables, available on the ODOT Web site (27). For every RWIS location, the closest location with available traffic data was selected on the same highway, and traffic counts extracted from ATR tables. Table 3-7 lists the ATRs that represent traffic data for every RWIS location.

Table 3-7: Correspondence between Oregon RWIS and ATR Locations

SysID	StnID	Highway		Milepost	County	Location for Traffic Counts
		Posted	ODOT			
24	0	US 97	4	135	Deschutes	
24	2	OR 58	18	62.3	Lane	Oakridge
24	3	OR 140	20	30.6	Klamath	Beatty
68	0	I-5	1	308.3	Multnomah	
68	1	I-5	1	308.8	Multnomah	
68	3	I-205	64	26.4	Multnomah	
68	8	US 26	47	37.7	Washington	Sunset Tunnel
68	9	I-205	64	19.6	Multnomah	Glen Jackson
68	10	US 26	26	36.5	Clackamas	Rhododendron
100	0	US 97	4	134	Deschutes	
100	1	US 97	4	151	Deschutes	
100	5	US 97	4	56	Wasco	RWIS 281-2
110	0	US 101	9	141.5	Lincoln	N. Newport
281	0	I-84	2	96	Wasco	Arlington (Gilliam Co)
281	1	I-84	2	114.41	Sherman	Arlington (Gilliam Co)
281	2	US 97	42	7.5	Sherman	
281	3	US 97	42	40.9	Sherman	RWIS 281-2
281	4	US 97	42	61.6	Wasco	RWIS 281-2
282	1	I-84	6	273.4	Union	Baker Valley (Baker Co)
282	2	I-84	6	336	Baker	Baker Valley (Baker Co)
282	3	US 395	28	39.7	Umatilla	Pilot Rock
283	0	I-5	1	4.5	Jackson	RWIS 283-4
283	1	US 199	25	16.4	Josephine	O'Brien
283	2	I-5	1	69	Josephine	Grave Creek
283	3	US 199	25	41	Josephine	
283	4	I-5	1	28.9	Jackson	
283	5	US 101	9	301.5	Curry	
283	6	OR 42	35	42	Coos	Brockway (Douglas Co)
283	7	OR 138	73	83	Douglas	Glide
285	1	I-5	1	247	Marion	N. Albany (Linn Co) Lake Creek (Linn Co)
285	2	I-5	1	236	Linn	N. Albany (Linn Co) Oakland (Douglas Co)
285	3	I-5	1	170	Lane	Oakland/Martin Creek (Douglas Co) Winchuck (Curry Co)

Whenever a location for traffic counts is not listed, traffic data was available for the RWIS location. Highway route numbers are given both for ODOT's linear referencing system (LRS) classification, as well as posted route numbers. Since some ATRs entered service or were decommissioned during the seven-year period of interest, it was necessary to estimate traffic at some RWIS through two or more ATR. Whenever an AADT was unavailable for a particular year at a certain ATR recorder, it was approximated by averaging AADT for the preceding and the following year. If an AADT value for the last year (2003) was unavailable, the closest available value was taken to approximate it.

3.2.3. Montana

Montana RWIS weather data were obtained from the Montana Department of Transportation (MDT) in comma-delimited text format, with a separate file for every RWIS location. Table 3-8 lists RWIS stations in the state of Montana. Because several stations had the same region and station code, the station code was changed for some stations. The data covered the period from November 1996 or later, depending on location, to September 2003. As with the other two states, only records that correspond to winter months were used in the analysis.

Table 3-8: Montana RWIS Locations

SYSID	RPUID	District	Location	Lat	Long
267	1	Butte	Monida Pass	44.558	-112.314
267	3	Butte	Boulder Hill	46.323	-112.069
301	5	Great Falls	Sieben	46.888	-112.111
301	6	Great Falls	Prickley Pear	46.914	-112.117
301	4	Great Falls	Gary Cooper Bridge	47.140	-111.859
629	3	Great Falls	Sweetgrass	48.960	-111.941
269	0	Missoula	Lookout Pass	47.454	-115.695
150	5	Missoula	Ninemile	47.023	-114.389
150	2	Missoula	Bearmouth	46.719	-113.295
267	4	Butte	Garrison	46.524	-112.808
564	2	Butte	Bozeman Pass	45.667	-110.804
564	3	Butte	East Livingston	45.686	-110.504
263	1	Billings	Reedpoint	45.710	-109.578
263	0	Billings	Yellowstone River Bridge	45.794	-108.468
263	3	Billings	Arrow Creek Hill	45.780	-108.163
263	4	Billings	Aberdeen Hill	45.028	-107.317
263	5	Billings	Hysham Hills	46.165	-107.309
563	1	Glendive	Sweeney Creek	46.267	-106.308
302	3	Glendive	Beaver Hill	47.022	-104.330
628	3	Missoula	Yaak Hill	48.583	-115.984
628	4	Missoula	Crystal Creek	48.118	-115.420
628	0	Missoula	Essex	48.282	-113.607

Table 3-8: Montana RWIS Locations (cont.)

SYSID	RPUID	District	Location	Lat	Long
629	2	Great Falls	Two Medicine Bridge	48.453	-113.195
629	0	Great Falls	Inverness	48.553	-110.648
150	8	Glendive	US2 @Stateline	48.138	-104.047
301	7	Missoula	Pablo	47.602	-114.113
628	1	Missoula	Dickey Lake	48.695	-114.784
267	2	Butte	MacDonald Pass	46.562	-112.309
312	0	Helena	Euclid Avenue	46.600	112.067
302	4	Great Falls	Loma	47.951	-110.505
268	3	Billings	Bull Mountain Divide	46.247	-108.461
563	4	Glendive	Alzada	45.003	-104.374
150	4	Missoula	Greenough Hill	46.903	-113.423
301	3	Great Falls	Helmville	46.968	-112.975
301	2	Great Falls	Bowmans	47.292	-112.152
563	0	Glendive	Lame Deer Divide	45.629	-106.511
564	1	Butte	Karst	45.345	-111.173
268	5	Billings	Lewistown Divide	47.062	-109.184
302	0	Glendive	Lufborough Hill	47.078	-107.572
302	2	Glendive	Lindsay Divide	47.283	-105.294
302	5	Glendive	Malta South	47.961	-108.306
628	5	Glendive	Sioux Pass	47.920	-104.326
268	1	Billings	Judith Gap	46.688	-109.750
150	1	Missoula	Lolo Pass	46.636	-114.580
629	1	Great Falls	Pendroy	48.073	-112.334
564	0	Butte	Raynolds Pass	44.728	-111.470
564	4	Butte	Norris Hill	45.500	-111.696
563	2	Glendive	Hillside	46.834	-106.271
269	2	Glendive	Redstone Hill	48.819	-104.997
150	7	Glendive	Cow Creek	47.688	-105.492
563	3	Glendive	Ekalaka	46.084	-104.436
269	3	Glendive	Comerstown Turn-off	48.810	-104.253
269	1	Glendive	McDonald's	48.427	-105.442
628	2	Missoula	Flathead River	48.219	-114.238
301	0	Great Falls	Monarch Canyon	47.143	-110.823
268	4	Billings	Hays Site	47.919	-108.726
263	2	Billings	Roscoe Hill	45.339	-109.494
268	0	Billings	East of Denton	47.300	-109.851
269	5	Missoula	Swan Lake South	47.591	-113.756
267	5	Butte	Big Hole Pass	45.315	-113.310

Weather observations were recorded by RWIS a few times every hour. The interval between observations is irregular for some RWIS stations, ranging from less than 1 minute to about 20 minutes; however, there were always at least four observations within an hour. Table 3-9 lists the variables included in every RWIS observation.

Table 3-9: Montana Weather Variables

Variable Name	Description	Format	Units
RA_SYSID	MDT Server number	integer	
RA_RPUID	MDT RPU number	integer	
RA_SENID	MDT Sensor number	integer	
RA_DATM	Date and Greenwich Mean Time	string	mm/dd/yyyy hh:mm
RA_WN_AVGSPD	Average wind speed	numeric	kph
RA_WN_GSTSPD	Wind gust speed	numeric	kph
RA_WN_DIRAVG	Average wind direction (compass degrees)	integer	degrees
RA_SF_COND	Surface condition (pavement)	integer	code 0 to 250
RA_SF_TEMP	Surface temperature (pavement)	numeric	0.01 ° C
RA_SF_BKTEMP	Back temperature	numeric	0.01 ° C
RA_SF_BOTEMP	Bottom temperature	numeric	0.01 ° C
RA_SF_FRZTEMP	Freeze temperature	numeric	0.01 ° C
RA_SF_REFTEMP	Reference temperature	numeric	0.01 ° C
RA_SF_CHEMFCT	Chemical factor	integer	0 to 95
RA_SF_CHEMPCT	Chemical percent	numeric	%
RA_SF_DEPTH	Water depth	numeric	0.1 mm
RA_SF_ICEPT	Percentage of sensor covered with ice	numeric	%
RA_SF_SPSTATUS	System on/off	character	
RA_AP_TEMP	Atmospheric temperature	numeric	0.01 ° C
RA_AP_DEWP	Atmospheric dewpoint	numeric	0.01 ° C
RA_AP_HUMID	Atmospheric humidity	numeric	%
RA_PR_TYPE	Precipitation type	integer	code 0 to 250
RA_PR_INTENS	Precipitation intensity	integer	code 0 to 250
RA_PR_RATE	Precipitation rate	numeric	0.025 mm/hr
RA_PR_ACCUM	Precipitation accumulation	numeric	0.025 mm/hr
RA_SSF_TEMP	Subsurface temperature	numeric	0.01 ° C

Accident records for the state of Montana were also obtained from MDT, password-protected in Microsoft Excel format. As was done in Oregon, only those crashes which occurred within five miles of an RWIS station along the same highway were requested; therefore, the accident data contained records for 10-mile highway sections. The crash records were split into separate files by RWIS location, and covered the time period from January 1996 through September 2003. The data factors recorded for every accident are given in Table 3-10.

Table 3-10: Montana Crash Variables

Variable Name	Description	Format	Units
DATE	Date of accident	string	mm/dd/yyyy
TIME	Time of accident	string	hh:mm
LOCATION	Highway route number and milepost	string	
WEATH_COND	Weather condition	character	cond. code
ROAD_COND	Pavement condition	character	cond. code
NUM_VEH	Number of vehicles involved	integer	
PEDESTR	Pedestrians involved (yes/no)	character	
FATAL	Number of fatalities	integer	
INJURY	Number of injuries	integer	

Montana AADT traffic counts were received from MDT electronic documents. Files that contain AADT counts for various locations in Montana are available in PDF format for the period from 1996 to 2003. For every RWIS location, the closest location on either side with available traffic data was selected on the same highway, and the AADT values interpolated to approximate the traffic count at that RWIS location. Table 3-11 lists the AADT locations used for linear interpolation for every RWIS location in Montana. Both posted and Montana highway route numbers are listed.

Monthly adjustment factors for traffic volumes in Montana were taken from the MDT ATR yearbooks, which were available from the MDT Web site in PDF format for 1998-2003 (28). Since ATR locations do not exactly correspond to AADT or RWIS locations in Montana, the closest ATR location was chosen for every RWIS station. As some ATRs entered service or were decommissioned during the period from 1998 to 2003, such choices were made separately for 1998-2000 and 2001-2003.

Whenever an AADT or MAF value was unavailable for a particular year at a certain ATR recorder, it was approximated by averaging AADT or MAF for the preceding and the following year. If an AADT or MAF value for the last (2003) or the first (1996-1998) year was unavailable, the closest available value was taken to approximate it.

Table 3-11: Montana RWIS and Traffic Count Locations

SysID	StnID	Highway		MP	AADT Location		MAF Location			
					1996-2003		1998-2000		2001-2003	
		Posted	MDT		MP1	MP2	HWY	MP	HWY	MP
150	5	I-90	I-90	81.8	77.38	82.62	I-90	114.3	I-90	114.3
150	2	I-90	I-90	145.8	138.09	153.32	I-90	114.3	I-90	114.3
267	4	I-90	I-90	174.4	174.36	175.53	I-90	114.3	I-90	114.3
150	7	MT-13	P-25	19.9	0.01	23.89	US-2	592.7	US-2	592.7
269	1	MT-13	P-32	25.5	16.62	35.23	US-2	592.7	US-2	592.7
150	8	US-2	N-01	667.1	645.01	659.26	US-2	592.7	US-2	592.7
628	2	MT-35	P-52	49.0	45.02	49.86	US-93	22.7	US-93	22.7
301	7	US-93	N-05	52.4	47.37	59	US-93	22.7	US-93	22.7
301	0	US-89	P-60	53.5	53.37	71.02	US-87	35.1	US-87	35.1
302	4	US-87	N-10	53.1	43.15	78.94	US-87	31.3	US-87	31.3
268	3	US-87	N-16	33.3	25.15	46.2	US-87	76.3	US-87	76.3
268	5	MT-200	N-57	90.3	83.4	112.98	MT-200	76.3	MT-200	76.3
268	0	MT-81	P-81	17.0	5.52	14.88	US-191	9.5	US-191	9.5
268	1	US-191	N-63	18.6	18.46	19.22	US-191	9.5	US-191	9.5
267	5	S-278	S-278	32.0	0	43.03	S-278	3.4	S-278	3.4
150	4	MT-200	N-24	22.1	1.07	31.86	MT-200	5.1	MT-200	5.1
267	1	I-15	I-15	0.3	0	0.5	I-15	22.3	I-15	15.2
564	0	MT-87	P-13	1.2	0	8.63	US-20	4.2	US-20	4.2
564	4	US-287	P-13	60.0	49.25	65.33	US-287	7.8	US-287	7.8
563	2	MT-59	P-18	36.5	34.9	38.09	MT-59	14	MT-59	14
268	4	MT-66	P-66	10.5	0	10.61	S-241	1.9	S-241	1.9
628	3	US-2	N-01	4.8	3.72	13.91	US-2	118.8	US-2	118.8
628	4	US-2	N-01	53.1	42.01	80.55	US-2	118.8	US-2	118.8
302	3	I-94	I-94	234.8	231.4	234.91	I-94	242.8	I-94	242.8
302	5	US-191	N-61	122.5	102.31	155.53	US-191	69.6	US-191	9.5
267	3	I-15	I-15	170.9	164.92	176.1	I-15	191.8	I-15	191.8
269	0	I-90	I-90	0.2	0	10.53	I-90	46.8	I-90	46.8
263	5	I-94	I-94	59.9	53.14	67.98	I-94	130.85	I-94	130.85
563	1	I-94	I-94	112.6	103.96	120.14	I-94	130.85	I-94	130.85
628	0	US-2	N-01	179.9	153.25	197.82	US-2	221.5	US-2	221.5
629	2	US-2	N-01	210.7	209.27	221.79	US-2	221.5	US-2	221.5
302	0	MT-200	N-57	175.3	158.36	212.08	MT-200	209.8	MT-200	209.8
302	2	MT-200	N-57	296.5	289.83	303.94	MT-200	209.8	MT-200	209.8
629	1	US-89	P-03	62.6	62.6	62.6	US-89	81.5	US-89	81.5

Table 3-11: Montana RWIS and Traffic Count Locations (cont.)

SysID	StnID	Highway		MP	AADT Location		MAF Location			
					1996-2003		1998-2000		2001-2003	
		Posted	MDT		MP1	MP2	HWY	MP	HWY	MP
564	2	I-90	I-90	321.8	321.74	330.99	I-90	342.3	I-90	284.4
564	3	I-90	I-90	337.7	333.53	337.88	I-90	342.3	I-90	284.4
564	1	US-191	N-50	55.3	47.93	66.83	US-191	49.8	US-191	49.8
629	0	US-2	N-01	337.6	332.62	341.58	US-2	372.9	US-2	372.9
150	1	US-12	N-93	0.0	0	0	US-93	72.8	US-93	72.8
628	1	US-93	N-05	160.2	156.9	177.94	US-93	78	US-93	78
563	3	MT-7	P-27	14.2	13.54	22.5	MT-7	79.2	MT-7	79.2
263	4	I-90	I-90	552.3	530.68	544.01	I-90	533.1	I-90	533.1
263	2	MT-78	P-78	18.8	0.47	19.69	US-212	72.1	US-212	72.1
263	1	I-90	I-90	390.8	378.04	391.12	I-90	447.8	I-90	447.8
263	0	I-90	I-90	452.3	450.1	452.4	I-90	447.8	I-90	447.8
263	3	I-90	I-90	468.6	462.29	473.24	I-90	447.8	I-90	447.8
629	3	I-15	I-15	395.3	390.4	397.9	I-15	359.6	I-15	359.6
563	0	US-212	N-37	50.1	42.08	61.48	US-212	76.7	US-212	76.7
563	4	US-212	N-23	139.4	109.47	137	US-212	76.7	US-212	76.7
269	5	MT-83	P-83	47.0	0	47.76	MT-83	0.8	MT-83	0.8
301	4	I-15	I-15	242.0	239.96	244.06	I-15	231.8	I-15	269
301	5	I-15	I-15	216.0	209.1	226.76	I-15	231.8	I-15	191.8 231.9
301	6	I-15	I-15	218.6	209.1	226.76	I-15	231.8	I-15	191.8 231.9
301	3	MT-200	N-24	53.3	35.68	55.65	MT-200	0.73	MT-200	0.73
301	2	MT-200	N-24	110.4	109.05	116.88	MT-200	0.73	MT-200	0.73
267	2	US-12	N-08	27.9	27.63	39.83	MT-200	0.73	MT-200	0.73
312	0	US-12	N-08	42.2	41.8	42.54	MT-200	0.73	MT-200	0.73
628	5	MT-16	N-62	21.0	3.92	24.96	MT-16	82	MT-16	82
269	3	MT-5	P-30	13.4	0	23.4	MT-16	82	MT-16	82
269	2	MT-5	P-22	19.8	17.9	30.27	MT-16	82	MT-16	82

3.3. Data Modification

Once the data for the three states were obtained, they had to be modified and converted to a form suitable for analysis. The first part of the process was to convert weather data from all three states to the same units, so as to make the analysis results comparable across different states. Since California and Oregon weather was recorded in English units (inches and feet), the research team converted Montana weather data from metric to English units. Montana weather variables were also given the same names as in Oregon.

Next, weather, crash and traffic data had to be combined into one dataset. Highway route numbers were listed according to the national classification in some parts of data, and state classification in other parts. This created an obstacle to combining the data. Therefore, for the states where this problem arose (Oregon and Montana), all highway route numbers were converted to the state classification.

Because of the topographic changes from one RWIS location to another, weather observations taken at an RWIS location can reasonably describe weather patterns only along a small portion of a highway close to the RWIS station. As weather generally does not change much within a five-mile radius, it was decided that only the accidents which occurred within five miles of an RWIS station should be selected for analysis. For California data, every such ten-mile section was chosen based on the point along State Route 299 closest to a NWS weather station, and crash data were filtered according to this criterion. (Oregon and Montana crash data were originally requested for ten-mile sections.)

It was believed that locations with unusually high or low traffic could distort the analysis results. At locations with high traffic volumes, weather is likely to be a smaller factor in crash frequency, whereas at locations with low traffic volumes, crash rates can fluctuate wildly based on a single crash. Consequently, such locations were identified and removed from the dataset. This includes RWIS placed in urban areas, such as Redding in California, Portland in Oregon, and Helena in Montana. A location was labeled as high-traffic whenever AADT > 60,000 vehicles/day for all years; and a location was labeled as low-traffic whenever AADT < 800 vehicles/day.

To account for a variety of possible weather events, a number of new variables were created. Many of these variables mimic those used by other researchers, in projects summarized in Chapter 2. The values of the new variables were computed based on the available data. Since the type of weather data available varied state-by-state, some new variables could be computed only for certain states. Table 3-12 summarizes new variables defined for all three states, and for every state lists new variables that are unique for that state.

Most of the variables in Table 3-12 are defined to describe a daily period, rather than the time of observation. For consistency and because the time interval between observations varied by state, RWIS recorder and year/month, all original weather variables were averaged by day. Thus the resulting dataset contained only daily measurements of every weather factor. A separate dataset was also created, which contained monthly averages of every weather variable. As we discuss in Section 3.4, such monthly data were used in subsequent statistical analysis.

Table 3-12: New Variables

Variable Name	Variable Description
All states:	
Frost	Frost occurs during the day: 1, if Tmax > 32 and Tmin < 32; 0, otherwise
Nprecip	Number of days in a month with precipitation
Nsnow	Number of days in a month with snowfall
Templow	1, if temperature stays below freezing point through the day
Tindex	0, if Tmin > 32; 1, if Frost = 1; 2, if Tmax < 32
frrain	Freezing rain occurs: 1, if temperature > 32 and PRCP or Rain > 0
California:	
Temprang	Tmax - Tmin
Oregon and Montana:	
AvTemp	Average air temperature during the day
Tmin	Minimum air temperature during the day
Tmax	Maximum air temperature during the day
AvPrecRate	Average rate of precipitation during the day; includes rain and snow
Snow	Average rate of snowfall during the day
Rain	Average rate of rainfall during the day
SnowFreq	Frequency of snowfall during the day
RainFreq	Frequency of rainfall during the day
Dew	Frequency of dew during the day
Hoar	Frequency of hoar during the day
Montana:	
frrainnew	Freezing rain recorded by RWIS (based on PrecipType value)
Mixed	Mixed precipitation recorded by RWIS (based on PrecipType value)
Sleet	Sleet recorded by RWIS (based on PrecipType value)
Hail	Hail recorded by RWIS (based on PrecipType value)
Frostnew	Frost conditions recorded by RWIS (based on SurfCond value)
Dewnew	Dew recorded by RWIS (based on SurfCond value)
Hoarnew	Hoar recorded by RWIS (based on SurfCond value)

RWIS collects and stores data automatically; therefore, occasionally erroneous information is entered due to sensor malfunction, power failure or adverse external conditions. Quite frequently, a zero value was entered when in fact no measurement was taken. Such erroneous data values were identified, and a "missing" value was entered. The following criteria were used to label a recorded value as erroneous:

- a value of "No Data" is recorded;
- TMAX >130 °F;

- TMIN >100 °F;
- AvTemp >110 °F;
- DewPtTemp >100 °F; or
- TMAX – TMIN > 70 °F.

Zero values of a certain variable were converted to "missing" if for an extended period of time (up to a year), there were no other values recorded for this variable at the same RWIS location. A few RWIS locations were removed from the dataset due to lack of data recorded for these locations.

As the effect of weather phenomena on highway safety may vary in different climatic zones, it was decided to introduce a simple but efficient climatic zone classification into the data. All weather station locations in three states were classified as mountains (climatic zone 1), valleys (zone 2) or plains (zone 3). The classification was conducted manually for every location, according to the following criteria.

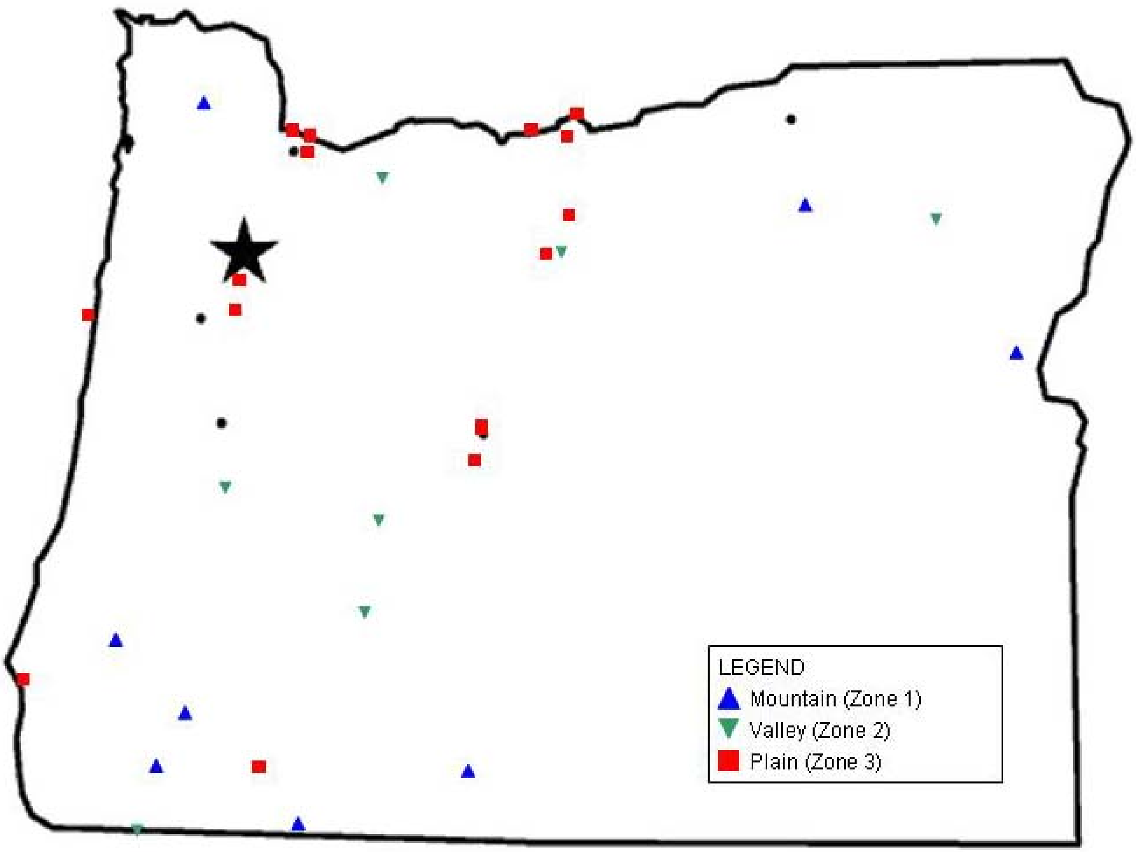
1. A location is classified as mountain, provided there is significant decrease in altitude in any direction from the road, within 5-10 miles from the location, and a significant change (increase or decrease) in the opposite direction.

2. A location is classified as valley whenever a minor change in altitude is observed in the direction of the road, and a significant increase in altitude on either side of the road, within 5-10 miles. This zone also includes locations with significant increase or no change in altitude on the opposite side of the road, that is, mountain front locations.

3. A location is classified as plain if a change of altitude is insignificant around the location up to a distance of 5-10 miles in every direction.

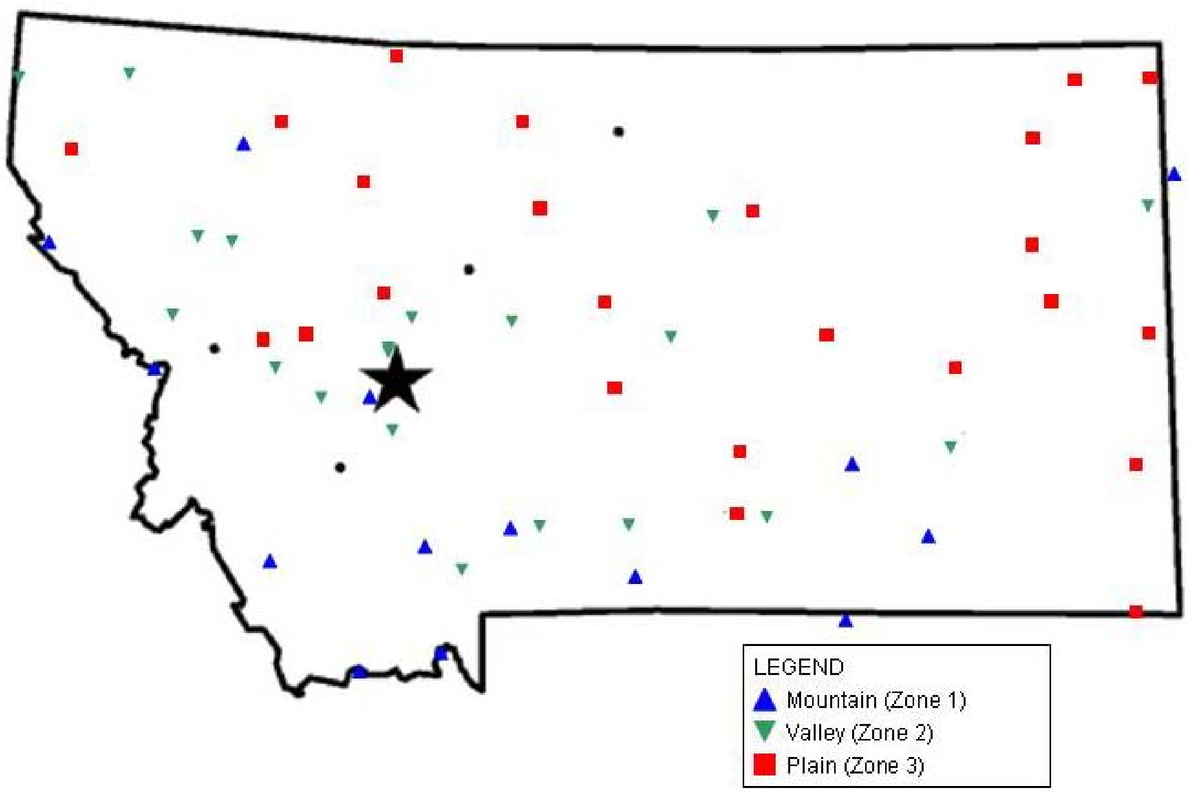
Figures 3-2, 3-3 and 3-4 show approximate distribution of climatic zone assignments in every state. One can clearly see from the figures that locations assigned to any one climatic zone are spread throughout Oregon and Montana, and there is no particular clustering of climatic zones along State Route 299 in California.

Figure 3-2: Climatic Zone Distribution in Oregon



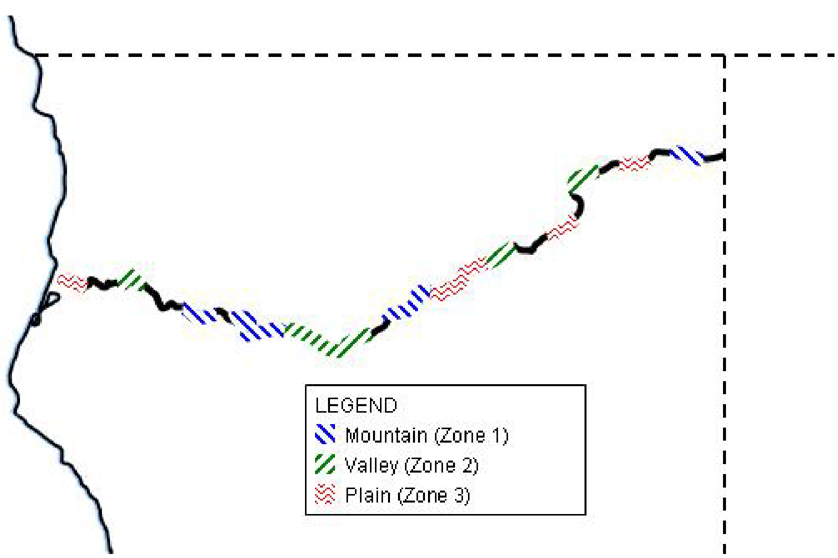
(Base Map: http://www.50states.com)

Figure 3-3: Climatic Zone Distribution in Montana



(Base Map: http://www.50states.com)

Figure 3-4: Climatic Zone Distribution along State Route 299 in California



After the initial analysis was run and preliminary models constructed, residual plots for every factor were studied in order to locate monthly observations with unusually high numbers of accidents. Such observations were individually investigated; some were removed from the data if it was found that the abnormality was not attributable to weather. In addition, separate clusters of points on residual plots were studied. They were matched to locations that exhibit unusual weather or traffic patterns (such as coastal locations in Oregon, and Cedarville in California) and also removed from consideration. In California, a few high-accident one-mile sections of State Route 299 were identified. As the high number of accidents at these locations is most likely the result of road condition or geometry and not weather, these locations were not included in the analysis.

3.4. Modeling Assumptions

One of the goals of the project was to construct models that would establish a relationship between highway crash rates and weather phenomena, and thereby would help to understand how the accident rates depend on weather. It is widely known that the occurrence of highway accidents is influenced by a variety of factors of different nature. It would be unrealistic to try to include all these factors in the analysis. For this reason, the research team decided to concentrate only on weather factors and disregard the others (with the exception of traffic volumes). As a consequence, the resulting models would offer a less-than-perfect fit of the data. However, the relationships exhibited by the models would be exclusively due to the weather factors.

The accident data can be modeled with a number of distributions. While hourly, daily and weekly accident counts are best approximated by discrete distributions, such as Poisson or negative binomial, averages over longer periods tend to be more normally distributed. To determine which distribution to use, a test run of the analysis was conducted, and Poisson and negative binomial models were built for weekly, monthly and seasonal crash rates. Between the two distributions, the negative binomial model offered a poorer fit of the data and was discarded. While the Poisson model provided a relatively adequate fit, it failed in comparison with the normal model. It did not identify any significant weather factors, instead concentrating on location parameters. Therefore, it was decided that the normal distribution should be used to model the accident data.

It follows from the Central Limit Theorem that averages over longer periods of time tend to be distributed more normally. Due to this fact, daily and weekly counts were eliminated from a list of possible forms of data representation, since their distributions cannot be adequately approximated by the normal distribution. At this point, a selection had to be made between monthly and seasonal (i.e. over an entire winter weather season) averages as a form to represent crash data for further analysis.

There is a trade-off between monthly and seasonal averages: while seasonal data are more normal, fewer observations are available for analysis. On the other hand, monthly data offer more observations, but are not as normally distributed. Another preliminary run of the analysis was conducted to compare the performance of models built on monthly and seasonal data. From this run, it became apparent that the number of observations for seasonal data is insufficient for meaningful analysis. As monthly averages are sufficiently close to normal, the research team decided to build models based on monthly average accident rates. The fact that traffic volumes are available on a monthly basis also supports this choice.

It is well known that traffic volumes influence highway accident rates. To account for this effect, average daily traffic counts were incorporated into the response variable. The response variable used in all models is accident rate, defined as

$$Accident\ rate = \frac{Number\ of\ accidents\ per\ month}{AADT} \times 1000 \tag{3-1}$$

The monthly accident rate is measured in accidents per month per 1,000 daily vehicles on the road. This choice of response variable is consistent with previously obtained results, described in Chapter 2.

It should be noted that Oregon weather data provided the best picture of statewide weather patterns. Because of this and for consistency reasons, Oregon weather factors were taken as a basis for the models in all three states. For Montana models, only those factors were considered that have an equivalent in the Oregon data; the rest of the Montana weather variables were ignored. For California models, only those Oregon factors were used that were measured by NWS stations along State Route 299. These factors, along with the new variables described in Section 3.3, comprised the set of factors used for building the models.

The goal of model construction in this project was to obtain models that not only provide an insight into the relationship between crash rate and weather factors, but are also easy to interpret. This consideration effectively rules out the use of complex interaction terms in the analysis. To determine whether the use of higher-order terms (quadratic and/or cubic) was justified, another run of the analysis was conducted. First quadratic, then both quadratic and cubic terms were included in model selection procedures described in Section 3.5. It was discovered that the following two phenomena occurred most frequently: the higher-order terms were either highly collinear with the corresponding linear terms (and simply replaced the linear terms in the models), or they did not enter the models. In the few instances where quadratic or cubic terms entered the model, the value of R^2 was increased marginally, by about one percent. Moreover, including the higher-order terms in the final models, obtained through model selection and described in Chapter 4, did not improve the fit. Since the inclusion of such terms does not significantly improve the fit, but makes the models harder to interpret, it was decided to drop these terms from consideration.

Added-variable plots for every combination of weather factor/state/zone were studied to check for the possibility of other functional forms for some weather factors. However, these plots did not reveal any patterns that would indicate nonlinear functional form for any weather factor. With all the above considerations taken into account, the research team restricted attention to the linear form of the model, with no interactions and no higher-order terms.

Finally, for every preliminary model obtained through model selection, residual plots were constructed and analyzed, in order to determine the appropriate transformation of the response. The desired transformation would result in homogeneity of variance, and at the same time improve the model fit. The following functional forms of the response were considered: linear (i.e., no transformation), square and cubic root, and logarithm transformation. Through all the models in Oregon and Montana, the cubic root and logarithm transformations performed better than linear and square root forms. They resulted in the virtual elimination of heteroscedasticity. At the same time, the cubic root transformation provided the values of R^2 that were 4 to 10 percent higher than the R^2 for the corresponding logarithm models. For California models, the cubic root and logarithm performed equally well. For consistency, the cubic root transformation of the response was adopted for all three states.

3.5. Model Selection

This section describes the process of model selection that was used to obtain candidate models for subsequent validation. In every state, separate models were constructed for the three climatic zones described earlier, as well as a statewide model based on all available data from that state.

As we discussed in Section 3.4, the model selection process was focused not only on maximizing R^2, which measures how well the model fits the data, but also on identifying significant weather factors that influence highway accident rates. Because of this focus on significance rather than fit, the candidate models obtained through model selection may exhibit the values of R^2 that differ by a substantial amount, while their order of magnitude remains comparable. This approach appears less restrictive and results in a larger number of candidate models. At the validation stage, this approach in turn allows for selection of the final model from a larger pool of candidates.

At the model selection stage, all weather factors were standardized, by transforming their range of values to the interval from –1 to 1. This standardization made the parameter coefficients comparable to each other, and thus helped to determine which factors have greater influence on the accident rate. While this standardization proved useful for model development, the final results in Chapter 4 are presented in non-standardized forms for ease of understanding.

To identify the candidate models for subsequent validation and comparison, the following model selection procedures were implemented using SAS.

- Stepwise Selection. The stepwise selection procedure starts with an empty model, where no factors are present. For all factors not in the model, it then calculates the F statistic, which measures the contribution of the factor if it were added to the model. The factor with the highest contribution (i.e. the smallest P-value for the F statistic) is added to the model. Then the F statistic is calculated for the factors in the model, and the factor with the smallest contribution is dropped. These two steps (adding a factor, dropping a factor) are repeated until the following two conditions are satisfied:

  1. All factors not in the model produce the F statistic with the P-values greater than the first cut-off value (typically equal to 0.15), and

  2. All factors in the model produce the F statistic with the P-values less than the second cut-off value (typically equal to 0.10).

The stepwise selection procedure is a modification of the forward selection procedure, which involves only Step 1. Whereas in the forward selection procedure the variables are added to the model one-by-one, the stepwise selection alternates iterations of forward and backward selection.

- Backward Elimination. The backward elimination technique starts with the full model, where all factors are fitted in the model. For every factor in the model, it then calculates the F statistic, which measures the contribution of the factor to the model. The factor with the least contribution is dropped from the model, and the F statistic is recalculated for the remaining factors. The process is repeated until all remaining factors in the model produce the F statistic with P-values less than the cut-off value (typically equal to 0.10).

- Heuristic Selection. The heuristic selection was used where automated procedures were ineffective and to fine-tune the results of such procedures. For example, in some cases, the above two procedures did not produce adequate results. In such cases, the research team utilized experience and prior knowledge to identify weather factors that should be included in the model. If this approach did not result in candidate models with reasonably good fit and significant terms, models that involve all possible combinations of weather factors were studied to identify those that offer the best combination of fit and significance.

The above three methods were applied to four distinct datasets for every zone in every state. The datasets were obtained by removing one year of data (from 2000 to 2003) that is reserved for model validation. Separate candidate models were developed for each dataset.

At every stage of the model selection process, collinearity characteristics were monitored to detect potential multicollinearity problems among the factors in the model. Whenever a combination of factors was detected that exhibited multicollinearity, one or more factors were removed from the model until the collinearity characteristics reached acceptable levels. In this project, the value of variance inflation factor (VIF) below 5 was considered acceptable. The model selection process was then re-run to check if other factors would enter the model and replace those that were dropped.

At the outset of the model selection process, several candidate models for each zone in every state were identified and subsequently compared. The comparison of the models involved testing how well the candidate models fit all data, checking residual plots for unwanted patterns and outliers, and testing predictive power of the models at the validation stage.

It should be noted that when fitting a model, SAS drops all observations that contain missing values for any of the variables in the model. The problem of missing data presented a great challenge in the project. A substantial number of observations involved missing values for one or more weather variables. The pattern of occurrence of missing values was not random and therefore presented a problem. It is generally believed that variables with a lot of missing values should not be included in the model, since it is not known exactly why those values are missing. It may be, for example, that the amount of precipitation was not measured in extremely low temperature conditions, which are outside the sensor's operational range. In this case, dropping all observations with missing precipitation measurements would change the range of other variables, such as temperature, and thereby introduce a bias into the model. In our case, precipitation variables were not recorded over an extended period of time at certain locations. Including these variables in the model would effectively drop such locations from consideration, which would also introduce bias.

For some state/zone combinations, locations where precipitation data were missing constituted a substantial part of the available data, and the research team used the following approach. If no precipitation variable entered the model during stepwise selection procedure, it was considered sufficient evidence that precipitation variables are not important in the model. Therefore, precipitation variables were ignored, and model selection procedures were run on the remaining factors.

The model selection procedures only consider observations that contain no missing values for any of the variables in the list of potential model factors. This list may include variables that contain a lot of missing values, and even if these variables do not enter the model, the resulting model will be based on fewer observations than are available for the factors in the model. As a consequence, the resulting model may not offer a very good fit for the entire dataset.

All candidate models identified at the model selection stage were fit to all available data, and their performance was analyzed. Only the models that offered reasonable fit were retained for further testing and comparison.

At the next step of comparing the candidate models, residual plots and response versus factor plots were constructing and analyzed. Close attention was paid to the scale and spread of residuals, to detect potential heteroscedasticity problems. A determination was also made whether the range of predicted values is reasonable in the context of the problem (i.e., negative predicted values would mean a negative accident rate, which makes no sense). Finally, potential outliers were identified and studied. Whenever an outlier was removed from the dataset, the model selection process was repeated, and adjusted candidate models obtained.

3.6. Model Validation

The candidate models that performed well at the preceding steps were then validated with new data. Every model was based on a dataset with one year of observations removed. The year removed varied from 2000 to 2003. This year of observations was now used to determine how well a model could predict the accident rate based on new data. Prediction plots were constructed and visually inspected.

It should be noted that over-prediction of the accident rate for the year 2003 was observed in a number of models. Predicted accident rates were consistently higher than observed accident rates. It can be argued that other, non-weather related factors could have contributed to the gradual decline in the highway accident rates. This suggests that the functional dependence of the accident rate on weather changes over time, and the models have to be recalibrated every few years. The research team also believes that conservative prediction, i.e. over-prediction, is better than under-prediction, since being over-cautious is better in dealing with factors, such as accident rate, that involve safety risks.

The research team also gave preference to model forms that exhibited consistency of factors across the four datasets for every state/zone combination, and that showed reasonably good prediction power. Sometimes this necessitated a selection between two models that include one of the two similar factors, such as Snow and SnowFreq. A selection was typically based on consistency, i.e. a decision was made in favor of the factor that appeared in other state/zone models.

Since the final models would most likely be used for prediction purposes, prediction testing was the most important criterion in selecting the final models. Other criteria for the final selection included the number of factors in the model, with preference given to parsimonious models (i.e. models with the least number of factors); the value of R^2; and ease of interpretation of the model factors and coefficients.

Once the final models for every state/zone were selected, these models were fit with all available data. As was mentioned in Section 3.5, no standardization was made to the weather variables in this final run; instead, they were used in their original scale for ease of interpretation.

4. MONTHLY MODEL

4.1. Oregon

This section summarizes the final models for every zone in Oregon, as well as the statewide model.

4.1.1. Zone 1 (Mountains)

The final model for Zone 1 has the following form:

$$\begin{aligned} \text{AccRate} = &1.56324 - 0.02219\ \text{TMIN} - 0.01734\ \text{WindSpdAvg} \\ &+ 6.1992\ \text{Snow} - 0.2208\ \text{Frost} \end{aligned} \tag{4-1}$$

The model offers a reasonably good fit (R^2 = 0.3167), and the factors are significant with P-values ranging from 0.0002 to 0.0873.

The model suggests that when the average daily minimum temperature increases by one degree, the accident rate drops by about 0.022 accidents per month per 1,000 daily vehicles on the road. Likewise, the accident rate decreases by 0.017 when the average wind speed increases by 1 mph; increases by 1.240 when the average daily snowfall increases by 0.2 in (which corresponds to an additional 6 inches of snowfall per month); and decreases by 0.007 when the average daily likelihood of frosts increases by 1/30 (which corresponds to one additional frost day per month).

It follows from the model that a decrease in accident rate occurs with the increase in the likelihood of frost. This result seems counterintuitive, but the data suggested this relationship at every stage of model development. Models for other state/zone combinations, described below, exhibit similar effects of frost. We also note that this negative effect is extremely small in comparison to the effects of other factors in the model, such as Snow. The occurrence of this effect can be explained by the presence of other, non-weather related confounding factors that affect the accident rate.

4.1.2. Zone 2 (Valleys)

The final model for Zone 2 has the following form:

$$\begin{aligned} \text{AccRate} = &1.70484 - 0.03049\ \text{TMIN} - 0.01719\ \text{WindSpdAvg} \\ &+ 1.61371\ \text{SnowFreq} \end{aligned} \tag{4-2}$$

The model offers an exceptionally good fit (R^2 = 0.6685) compared to other models, and the factors are significant with P-values ranging from less than 0.0001 to 0.0384.

The model suggests that when the average daily minimum temperature increases by one degree, the accident rate drops by about 0.030 accidents per month per 1,000 vehicles on the road. Likewise, the accident rate decreases by 0.017 when the average wind speed increases by 1 mph;

and increases by 0.067 when the average daily frequency of snowfall increases by 1/24 (i.e. the average duration of snowstorms increases by 1 hour).

We note that the parameter estimates for TMIN and WindSpdAvg are nearly identical for Zone 1 and Zone 2 models. Therefore, these factors are equally important in predicting the accident rate at both mountain and valley locations. However, Snow and Frost terms in Zone 1 model are replaced by SnowFreq for Zone 2. This suggests, in particular, that the frequency of snowfall is a better predictor of accident rate at valley locations than the amount of daily snowfall. At the model selection stage, a model for Zone 2 including Snow and Frost was fit, but did not produce satisfactory results.

4.1.3. Zone 3 (Plains)

The final model for Zone 3 has the following form:

$$\begin{aligned} \text{AccRate} = {} & 0.62354 - 0.0038\ \text{TMAX} - 1.14164\ \text{SnowFreq} \\ & + 3.38151\ \text{Templow} \end{aligned} \tag{4-3}$$

The model offers a reasonably good fit (R^2 = 0.3283), and the factors are significant with P-values ranging from less than 0.0001 to 0.0255.

The model suggests that when the average daily maximum temperature increases by one degree, the accident rate drops by about 0.004 accidents per month per 1,000 daily vehicles on the road. Likewise, the accident rate decreases by 0.047 when the average duration of snowstorms increases by 1 hour; and increases by 3.382 when the average number of days per month with temperature below freezing increases by 1.

It should be noted that the temperature term changes from TMIN to TMAX compared to Zone 1 and Zone 2 models. This suggests that the average daily high temperature has a greater (negative) influence on the accident rate at plain locations. The SnowFreq has the opposite effect on the accident rate, compared to Zone 2. One possible explanation for this may be that as the frequency of snowfall increases at plain locations, fewer people decide to travel and as a consequence, fewer vehicles are on the roads. A new term that enters Zone 3 model is Templow. It suggests that the number of days with temperature below freezing has a significant impact on the accident rate at plain locations. It is also noteworthy that a wind factor is absent from the model for Zone 3. Thus, at mountain and valley locations in Oregon, wind has a significantly stronger impact on the accident rates, whereas at plains locations its effect is insignificant and is therefore absent from the model.

A notable difference observed in this model is the range of the predicted values of the accident rate. While the predicted range was normally between 0.3 and 1.4 for Zones 1 and 2, nearly all predicted values in this model are located between 0.3 and 0.45. This phenomenon suggests that generally, the accident rate is lower in the plains than in the mountains or valleys.

4.1.4. Statewide Model

The final statewide model for Oregon has the following form:

$$\begin{aligned} \text{AccRate} = {} & 1.16504 - 0.01195\ \text{TMIN} - 0.02522\ \text{WindSpdAvg} \\ & + 9.36849\ \text{Snow} + 0.70134\ \text{Templow} \end{aligned} \tag{4-4}$$

The model offers a good fit ($R^2 = 0.4552$), and the factors are significant with P-values ranging from less than 0.0001 to 0.0024.

The model suggests that when the average daily minimum temperature increases by one degree, the accident rate drops by about 0.012 accidents per month per 1,000 daily vehicles on the road. Likewise, the accident rate decreases by 0.025 when the average wind speed increases by 1 mph; increases by 1.874 when an additional 6 inches of snowfall per month are observed; and increases by 0.701 when the average number of days per month with temperature below freezing increases by 1.

The factors present in the statewide model come from the models for all three zones: TMIN and WindSpdAvg from Zones 1, 2; Snow from Zone 1; Templow from Zone 3. The sign and magnitude of the terms are consistent with the respective models. One notable exception is the Templow term, which changed from about 3.3 in Zone 3 to 0.7 in this model. This suggests that the number of days with temperature below freezing has a lesser effect statewide than in Zone 3.

4.2. Montana

This section summarizes the final models for every zone in Montana, as well as the statewide model.

4.2.1. Zone 1 (Mountains)

The final model for Zone 1 has the following form:

$$\text{AccRate} = 0.88376 + 0.44804\ \text{SnowFreq} - 0.26409\ \text{Frost} \tag{4-5}$$

The model offers a reasonably good fit ($R^2 = 0.2932$), and the factors are significant with P-values ranging from 0.0018 to 0.0359.

The model suggests that the accident rate increases by 0.019 when the average duration of snowstorms increases by 1 hour; and decreases by 0.009 when one additional frost day per month is observed. We note that an increase in snow frequency has a positive effect on the accident rate, while an increase in the likelihood of frost has a negative effect. The last fact is consistent with Oregon models.

4.2.2. Zone 2 (Valleys)

The final model for Zone 2 has the following form:

$$\begin{aligned}\text{AccRate} = {} & 1.31229 - 0.0221\ \text{TMAX} + 0.03308\ \text{DewPtTemp} \\ & - 48.21987\ \text{Rain} - 0.40899\ \text{Frost}\end{aligned} \quad (4\text{-}6)$$

The model offers a good fit ($R^2 = 0.3242$), and the factors are significant with P-values ranging from less than 0.0001 to 0.0312.

The model suggests that when the average daily maximum temperature increases by one degree, the accident rate drops by about 0.022 accidents per month per 1,000 vehicles on the road. Likewise, the accident rate increases by 0.033 when the average daily dew point temperature increases by one degree; decreases by 1.607 when the average daily rainfall increases by 1/30 in (which corresponds to an additional 1 inch of rain per month); and decreases by 0.014 when one additional frost day per month is observed.

Compared to the model for Zone 1 in Montana, the Frost term stays in the model and has a comparable coefficient. However, SnowFreq is replaced by a combination of TMAX, DewPtTemp and Rain. Of these new terms, Rain has the largest magnitude (about 50 times larger) than the rest of the terms. This indicates that Rain has a significantly stronger impact on the accident rates than other factors. The coefficient for TMAX is identical to that of TMIN in Oregon Zone 1 model. Frost has a negative effect, which is consistent with other models.

4.2.3. Zone 3 (Plains)

The final model for Zone 3 has the following form:

$$\begin{aligned}\text{AccRate} = {} & 1.19234 - 0.02153\ \text{TMAX} + 0.0277\ \text{DewPtTemp} \\ & + 14.0862\ \text{Snow}\end{aligned} \quad (4\text{-}7)$$

The model offers a good fit ($R^2 = 0.3700$), and the factors are significant with P-values ranging from less than 0.0001 to 0.0078.

The model suggests that when the average daily maximum temperature increases by one degree, the accident rate drops by about 0.022 accidents per month per 1,000 vehicles on the road. Likewise, the accident rate increases by 0.028 when the average daily dew point temperature increases by one degree; and increases by 2.817 when an additional six inches of snow per month are observed.

We note that in comparison to Zones 1 and 2, the Frost term disappears from the model, indicating that the impact of frosts on the accident rate becomes insignificant at the plain locations. Of the rest of the terms in Zone 2, TMAX and DewPtTemp remain in the model. The coefficient for TMAX is essentially identical to that of TMIN in the Oregon Zone 1 model, and to TMAX in Montana Zone 2. The DewPtTemp parameter estimate is similar to that in Zone 2 model. However, Rain is replaced by Snow, which has a larger magnitude than the rest of the factors. This indicates that Snow has a significantly stronger impact on the accident rates than other factors in Zone 3.

4.2.4. Statewide Model

The final statewide model for Montana has the following form:

$$\begin{aligned} AccRate = 1.09119 - 0.132\ TMAX + 0.01429\ DewPtTemp \\ + 0.00751\ WindSpdAvg + 5.08852\ Snow - 0.12366\ Frost \end{aligned} \tag{4-8}$$

The model offers a relatively good fit ($R^2 = 0.1988$), and the factors are significant with P-values ranging from less than 0.0001 for TMAX to 0.1045 for Frost.

The model suggests that when the average daily maximum temperature increases by one degree, the accident rate drops by about 0.132 accidents per month per 1,000 vehicles on the road. Likewise, the accident rate increases by 0.014 when the average daily dew point temperature increases by one degree; increases by 0.008 when the average wind speed increases by 1 mph; increases by 1.018 when an additional six inches of snow per month are observed; and decreases by 0.004 when one additional frost day per month is observed.

It is noted that all Zone 3 terms remain in the model, with Snow offering the largest parameter estimate. The signs of these three terms are also consistent with the earlier models. Frost reappears in the model, with a parameter estimate comparable to that in Zone 1. One new term observed in the statewide model is WindSpdAvg, with a very small positive coefficient. This is reminiscent of the Oregon models, where a wind term played an important role in all models.

The range of the predicted values of the accident rate is similar throughout all Montana models, with the exception of a few outliers, and is comparable to the predicted range in Oregon models. This suggests, in particular, that the accident rates are similar in all climatic zones throughout Montana.

Generally, the results obtained for Montana are poorer than the results for Oregon. This can be attributed to a variety of factors. First, the data collected and used for the analysis had a number of deficiencies. As different divisions within MDT are responsible for the placement of RWIS and ATR sensors, staff members use different considerations to determine the appropriate locations for placement of the sensors. This resulted in a mismatch between RWIS and ATR locations throughout Montana. In addition, the monthly traffic adjustment factors are available for locations that do not coincide with ATR counters. This mismatch between weather and traffic data, as well as between two types of traffic data, could have caused a significant bias in the models.

Another factor that could have caused bias is the sampling method. It is apparent from the data that placement of RWIS and ATR sensors gravitated more heavily toward unsafe or hazardous locations. Since such hazardous locations are over-represented, the data available do not offer a representative sample of locations throughout the state.

The difference in the results between Oregon and Montana could also have been caused by natural factors and differences in driver behavior. In Montana, winter weather conditions are observed more frequently throughout the year than in Oregon. As a result, Montana drivers may tend to be more skilled in winter driving techniques and therefore less prone to accidents. Road geometry and safety at specific locations may also have affected the model results. Finally, wildlife crashes may be an important factor that contributes to an increase in highway accident rates in Montana.

4.3. California

This section summarizes the final models for every zone, as well as the statewide model for the state of California.

4.3.1. Zone 1 (Mountains)

The final model for Zone 1 has the following form:

$$\text{AccRate} = 0.68288 + 0.03969\ \text{Frost} \tag{4-9}$$

The model offers a good fit (R^2 = 0.4096), and the factor is significant with P-value less than 0.0001.

The model suggests that the accident rate increases by 0.001 when one additional frost day per month is observed.

It is noteworthy that Frost is the only factor in the model, and its effect is small but statistically significant. No other weather effects were detected for Zone 1 in California.

4.3.2. Zone 2 (Valleys)

The final model for Zone 2 has the following form:

$$\text{AccRate} = 0.77838 - 0.00318\ \text{TMIN} + 0.03792\ \text{Nsnow} \tag{4-10}$$

The model offers a good fit (R^2 = 0.3679), and the factors are significant with P-values ranging from less than 0.0001 to 0.0401.

The model suggests that when the average daily minimum temperature increases by one degree, the accident rate drops by about 0.003 accidents per month per 1,000 vehicles on the road. Likewise, the accident rate increases by 0.001 when the average number of days per month with snowfalls occurring increases by 1.

The two terms in the model replace Frost in Zone 1. As in Zone 1, the effects of both terms are small but statistically significant.

4.3.3. Zone 3 (Plains)

The final model for Zone 3 has the following form:

$$\text{AccRate} = 1.02545 - 0.0061\ \text{TMIN} + 0.01502\ \text{Nsnow} \tag{4-11}$$

The model offers a relatively good fit ($R^2 = 0.1921$), and the factors are significant with P-values ranging from less than 0.0001 to 0.0256.

The model suggests that the accident rate drops by 0.006 when the average daily minimum temperature increases by one degree; and increases by less than 0.0001 when the average number of days per month with snowfall occurring increases by one.

The factors present in this model are the same as in the model for Zone 2. The sign and magnitude of the parameter estimates are also nearly identical to those in Zone 2. No other weather effects were detected for Zone 3 in California.

4.3.4. Statewide Model

The final statewide model for California has the following form:

$$AccRate = 1.02833 - 0.00432\ TMAX + 0.00809\ Frost \tag{4-12}$$

The model offers a relatively bad fit ($R^2 = 0.1231$), and the factors are significant with P-values ranging from less than 0.0001 to 0.0414.

The model suggests that the accident rate decreases by 0.004 when the average daily maximum temperature increases by one degree; and increases by 0.0002 when one additional frost day per month is observed.

It is noted that the two terms present in the model appear to relate to different zones: Frost to Zone 1, and temperature terms to Zones 2 and 3. Although TMIN is included in the zonal models whereas TMAX is included in the Statewide model, these variables are collinear, so the effects of both are similar. In this model, TMAX serves as a better predictor of the accident rate. The magnitude of the parameter estimates suggests a small but statistically significant, effect of weather on the accident rates.

This last phenomenon, observed consistently in all California models, leads one to a conclusion that weather has a very minor effect on the highway accident rates in California. However, results for Oregon and Montana, as well as the discussion below, suggest that such a conclusion may be misleading.

It is clear from the above discussion that the results obtained for California are poorer than the results for Oregon. As with Montana, this can be attributed to a variety of factors. First, the data used for the analysis was collected from only one highway, which may not be representative of the entire state. In addition, the NWS weather data, rather than RWIS data, were used. Since most NWS weather stations along State Route 299 are located at a considerable distance from the highway, weather conditions observed at these stations may not reflect weather conditions on the highway. The analysis involved only the sections of State Route 299 closest to the NWS stations, which may not be representative of all the sections of the highway. Some locations selected for analysis could be inherently unsafe due to road geometry and other non-weather factors.

As NWS stations do not measure the same weather variables as RWIS sensors, many weather variables were not available and could not be included in the models. This resulted in the truncated forms of the models being used; therefore, some important relationships between crash rates and weather could not be captured.

The fact that only one set of traffic monthly adjustment factors was available constituted another deficiency of the dataset and could have caused a significant bias in the models.

4.4. Summary

Table 4-1 summarizes the model coefficients for the models that were developed for each zone in each of the states. A cursory glance shows that while there are different variables included in different models, there are some common characteristics in relationship to the sign and order of magnitude of parameters. The intercept term, which reflects a base expected crash rate at each site, is always greater than zero. This suggests that crashes are dependent on factors other than weather, which agrees both with experience and the assumptions used in this analysis. Temperature terms – TMIN and TMAX – are always negative and relatively small, suggesting that increased temperatures would decrease crash rate, which is logical when examining winter travel months. Similarly, crash rates increase with Templow, which makes sense in that higher values of Templow indicate that the temperature is below freezing for a greater period of time. Increased snow – whether measured by Nsnow, Snow or SnowFreq – tends to lead to increased crash rates, with the earlier mentioned exception of the Oregon Zone 3 model.

As was alluded to earlier, there are some parameter estimates that are counterintuitive. Most prominent among these is Frost, which appears as a statistically significant factor in many of the models. In California, the variable showed a weak but positive correlation with crash rate, while in Oregon and Montana, increasing values of Frost result in decreased crash rates. For the Oregon Zone 1, Montana Zone 1, and Montana Statewide models, the weakly negative Frost term is countered by a strongly positive Snow or SnowFreq term. This suggests that Frost, as defined, tends to correct a potential exaggeration of the effects of snow on safety. The Montana Zone 2 model is the exception where Frost is negative and no snow-related term appears in the model. It is unclear what factors may be influencing crash rate behavior in this zone, but it is likely that other factors, such as site selection, could be important. WindSpdAvg has a fairly small parameter estimate, and is positive in Oregon models and negative in Montana models.

It should be noted that none of the variables included in any of these model forms are directly influenced by winter maintenance practices. None of the variables relate to snow or ice depth on the roadway, roadway friction, presence of chemicals on the road surface, or similar factors. While these factors would be expected to correlate strongly with roadway safety, there was insufficient data for them to be included in the models. The relative weakness of some model correlations may suggest that there are different levels of effectiveness in winter maintenance practices on different highways. Additional data would be necessary to establish a clearer connection between roadway safety and winter maintenance practice.

Table 4-1: Summary of Model Coefficients

	Oregon				Montana				California			
Variable	1	2	3	All	1	2	3	All	1	2	3	All
Intercept	1.56324	1.70484	0.62354	1.16504	0.88376	1.31229	1.19234	1.09119	0.68288	0.77838	1.02545	1.02833
DewPtTemp						0.03308	0.02770	0.01429				
Frost	-0.22080				-0.26409	-0.40899		-0.12366	0.03969			0.00809
Nsnow										0.03792	0.01502	
Rain						-48.21987						
Snow	6.19920			9.36849			14.08620	5.08852				
SnowFreq		1.61371	-1.14164		0.44804							
Templow			3.38151	0.70134								
TMAX			-0.00380				-0.02153	-0.13200				-0.00432
TMIN	-0.02219	-0.03049		-0.01195		-0.02210				-0.00318	-0.00610	
WindSpdAvg	-0.01734	-0.01719		-0.02522				0.00751				

5. WEATHER SEVERITY INDEX

5.1. Index Properties

Based on the models developed for the three states, a winter weather severity index was constructed. The index could be used to report the severity of winter weather conditions to the general pubic. In developing such an index, the research team adopted the following guidelines.

- Simplicity. The weather index should have a simple form. A scale of values from 1 to 10 was selected, where the value of 1 corresponds to the least severe weather, while the value of 10 represents the most severe weather conditions.

- Ease of Interpretation. The weather index should also be easy to interpret. The selected scale for the index seems to be a good choice in this respect. Every value of the index should have a clear meaning, and a unit change in the index (increase or decrease by 1) should easily translate into a certain change in the average accident rate.

- Consistency with Common Sense. The index must match the intuitive assumptions of the general public. For example, a mid-range value (5 or 6) should correspond to the average ("usual") crash hazard due to weather. Similarly, a value of 4 would be taken as "better than average" by the general public; therefore, it must reflect the same conditions as those perceived by the public.

- Ease of Computation. The weather index should be relatively easy to compute. It is likely that reporting of the index to the public would be bestowed upon the TV or radio stations. For this reason, computation of the index should be done with a simple algorithm. This algorithm could be provided to the weather reporting outlets, in the form of an Excel spreadsheet or a Java applet.

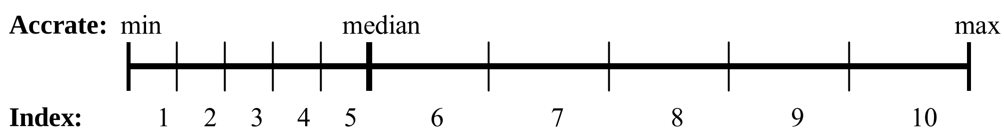
5.2. Calibration

The calibration of the weather severity index was based on the Oregon models, since they produced the best results among the three states. Separate indices for every climatic zone, as well as a statewide index, were constructed. Therefore, at a given time, there would be two index values for the same location. The statewide value could help travelers from other parts of the state interpret the winter weather relative to the state as whole, whereas the zone-based value could be used to show how a location compares to normal. So, for example, a short-term forecast for a mountain location may yield a statewide value of 7 and a mountain value of 4. This could be interpreted to mean that winter driving conditions in the mountains are a little worse than average for the state, but that the mountain driving is better than might normally be experienced. The index was based on the values of accident rate for the observed period of time. As the final models were fitted for the cubic root transformation of the accident rate, values of the transformed response were used to calibrate the index.

The following calibration procedure was implemented (see Figure 5-1). First, the observed range of values of the predicted accident rate was taken as the working overall range for accident rate.

A median of this range was then found, based on predicted accident rate for all locations in the same climatic zone (or all zones for statewide index) for the entire seven-year period, 1996 to 2003. The interval from the minimum observed accident rate to the median was split into 5 equal intervals, and the values from 1 to 5 assigned to these intervals. The interval from the median to the maximum observed accident rate was also split into 5 equal intervals, and the values from 6 to 10 assigned to these intervals. Finally, the transformation of the response was inverted, and a range of values of the untransformed accident rate was specified for every interval. The lowest interval is adjusted to cover all the values from zero to the minimum predicted response, while the highest interval would include all values above the maximum predicted response.

Figure 5-1: Calibration of Weather Index



The research team selected this approach among four different methods of constructing a weather index. Generally, a median is intuitively treated as corresponding to the average conditions. Since the fifth and the sixth interval are immediately adjacent to the median, the index values of 5 or 6 indicate average severity of weather, which is in line with public expectations. Thus, the key advantage of this approach is that it meets the third goal of index calibration (consistency with common sense), whereas the other three methods do not.

It should be noted that other values of the index indicate how much better (or worse) the weather conditions are, compared to the "average" severity of weather conditions.

5.3. Index Values

The final scale for the weather severity index is presented in Table 5-1. This scale describes what value of the index would be assigned to a certain month, if the accident rate during that month falls into a specific interval. As we mentioned earlier, the accident rate is measured in accidents per month per 1,000 daily vehicles on the road.

Table 5-1: Weather Severity Index Scale – Zones 1-3 and Statewide

Index Value	Accident Rate Interval							
	Zone 1		Zone 2		Zone 3		Statewide	
	From	To	From	To	From	To	From	To
1	0	0.1283	0	0.0557	0	0.0215	0	0.0829
2	0.1283	0.1615	0.0557	0.0884	0.0215	0.0294	0.0829	0.1235
3	0.1615	0.2001	0.0884	0.1319	0.0294	0.0390	0.1235	0.1756
4	0.2001	0.2443	0.1319	0.1877	0.0390	0.0505	0.1756	0.2406
5	0.2443	0.2947	0.1877	0.2573	0.0505	0.0641	0.2406	0.3200
6	0.2947	0.4482	0.2573	0.4849	0.0641	0.1088	0.3200	0.5302
7	0.4482	0.6474	0.4849	0.8179	0.1088	0.1704	0.5302	0.8166
8	0.6474	0.8984	0.8179	1.2764	0.1704	0.2517	0.8166	1.1911
9	0.8984	1.2071	1.2764	1.8805	0.2517	0.3555	1.1911	1.6656
10	1.2071	up	1.8805	up	0.3555	up	1.6656	up

The distribution of values of the observed accident rate within every value of the weather index is given in Figures 5-2 through 5-5, separately for Zones 1 – 3 and for the statewide index. The plots show that in general, there is a positive correlation between the accident rate and the weather index. This is particularly true for Zone 2 index and Statewide index.

Figure 5-2: Distribution of Accident Rates by Index Value: Zone 1

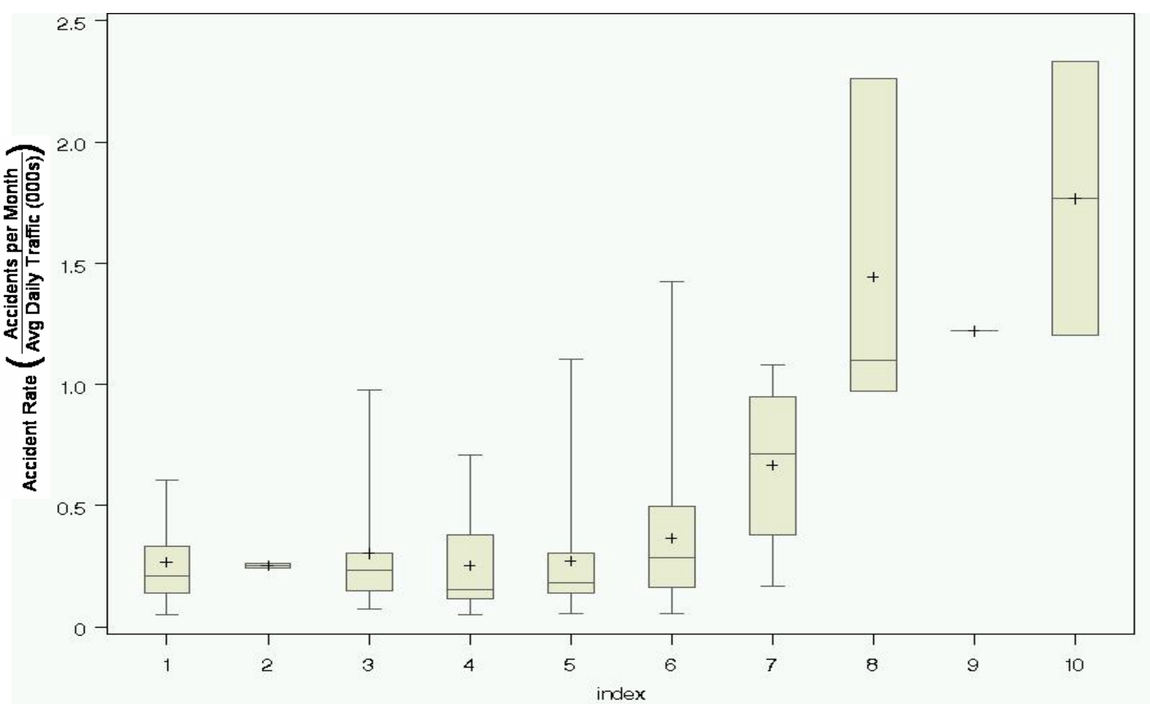


Figure 5-3: Distribution of Accident Rates by Index Value: Zone 2

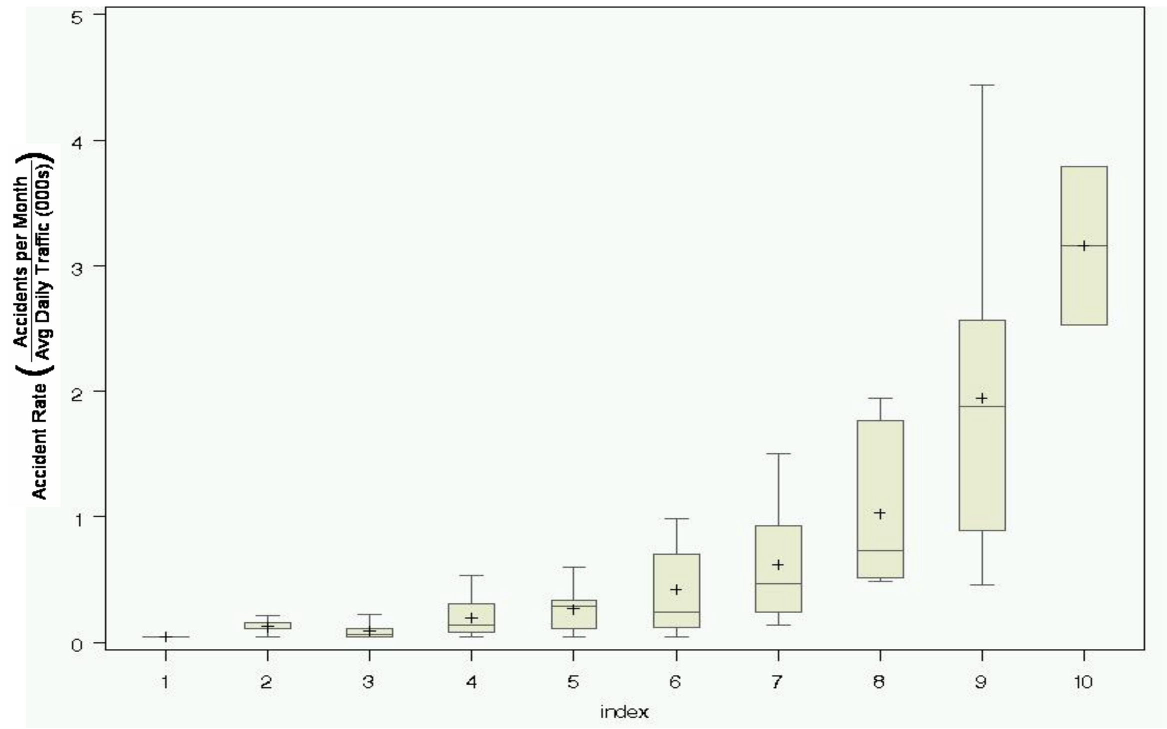


Figure 5-4: Distribution of Accident Rates by Index Value: Zone 3

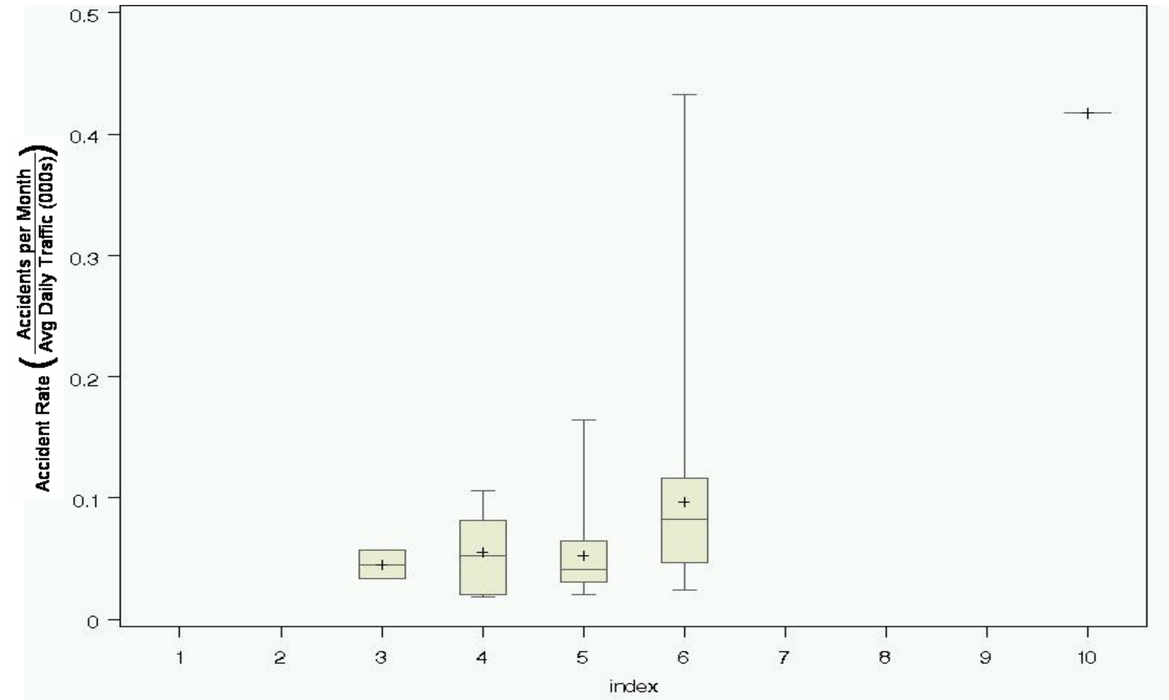


Figure 5-5: Distribution of Accident Rates by Index Value: Statewide

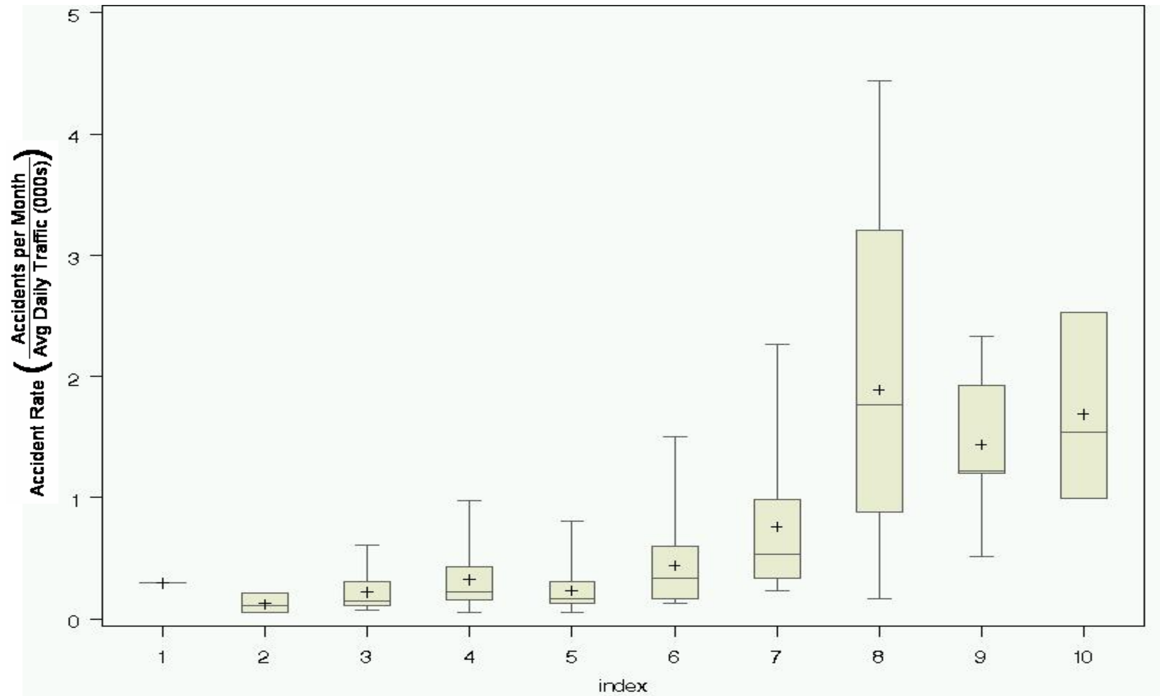
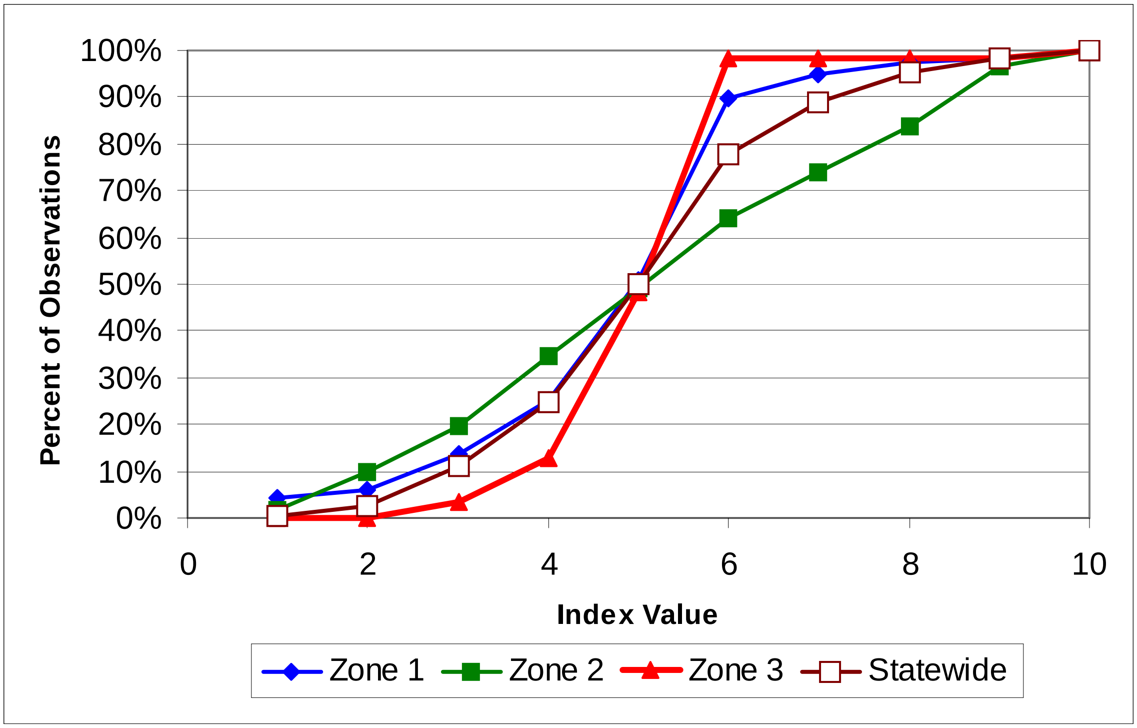


Figure 5-6 shows the cumulative distribution of the weather index for the Oregon data set used in the project. For every value of the weather index, the plot gives the percentage of observations for which the weather index is less than or equal to the specified value. Separate distributions are presented for Zone 1 – 3, as well as for the statewide index.

Figure 5-6: Cumulative Distribution of Weather Index, Zones 1-3 and Statewide



A similar calibration procedure could be implemented to construct weather severity indices for the other two states, California and Montana. However, because of the poor correlation between accident rate and weather for these two states, the resulting weather severity indices would appear questionable at best. Additional data collection and analysis is required to obtain meaningful indices for California and Montana, as well as other states.

5.4. Application

The winter weather severity index could be implemented with a simple procedure. A night-before weather forecast can be made for next day's weather. This forecast could include the average weather conditions across the state, as well as specific weather conditions in the mountains, valleys and plains. The forecast data would be then input into an Excel spreadsheet or a Java applet. The spreadsheet or applet would then calculate predicted accident rates, based on the models developed for each of the climatic zones and on the statewide model. It would then use the scale given in Table 5-1 to determine the value of the weather index for the appropriate climatic zone, and output that value on the screen. This value of the weather index could then be reported in the newspapers, on radio or on TV.

6. NEXT STEPS AND FUTURE RESEARCH

The purpose of this research project was to develop a roadway weather severity index that would correlate winter weather with roadway safety. This final report presented the results of modeling crash rate as a function of many weather-related parameters, and yielded different optimal model forms and coefficients for mountain, valley and plains locations, and for different states. These models showed that the weather parameters that were examined account for some, though not all, of the variability in crash rates experienced on different highway segments. Focusing on Oregon's data, the report presented index values which can be used to help characterize the relative safety risk associated with winter driving on a statewide level, as well as for different regions within a state. This may be valuable as a traveler information tool to encourage greater motorist awareness when hazardous conditions are present.

The authors believe that this report adds significant information to previously completed studies regarding weather severity. However, this investigation has raised numerous other questions that would merit further inquiry. These are discussed in the remainder of this chapter.

6.1. Methodology

One important area of future research would be to examine weekly or, better, daily model formulations, which could lead to improved forecasting and more proactive maintenance treatments. The index was developed using monthly data and was analyzed on a monthly basis. However, monthly data represent a significant aggregation of weather data, and may tend to de-emphasize the significance of certain weather events. On certain days during a month, the daily low temperature will be lower than the average of the daily low temperatures for the entire month. Consequently, there may be value in exploring a model formulation that uses daily data. This option was initially explored by the research team; however, since most days had no observed crashes, the predictive capability of the weather models was very poor. Undertaking similar research on highways where there are higher traffic volumes and a greater frequency of weather-related crashes could make a daily index more viable.

An additional area of improvement would be to include variables that may be more directly influenced by winter maintenance practices, such as snow depth or road surface condition. This investigation would clearly depend on data quality and availability, but again could provide some clear implications for winter maintenance practices that could best improve traveler safety.

Another important consideration is the effect of weather on traffic. The methodology used in this report adopted monthly adjustment factors, which reflected the seasonality of traffic. However, in some cases, motorists were electing not to make a trip when the weather was exceptionally bad. Even if this percentage is relatively small – for example, 10 percent – this could have a significant impact on the actual crash rates. Data sets which have daily traffic volumes observed at weather sites could have this daily fluctuation taken into account, perhaps resulting in greater model sensitivity.

Nixon et al (19), cited earlier, sought to develop a storm severity index, through the use of input from winter maintenance personnel. This approach, of looking at storm events as opposed to somewhat arbitrary 24-hour time periods, may be more valuable in evaluating winter maintenance costs and responses from the perspective of maintenance personnel, who are responsible for managing the roads during these times. This would require re-analysis of the weather data over time to classify when storms start and stop, and accordingly the development of new parameters to measure how a storm starts and stops, its intensity, and characteristics of the storm as it progresses.

6.2. Applications

6.2.1. Traveler Information

The focus of this investigation was on improving safety through enhanced traveler information. The idea is that motorists may tend to exercise greater caution when warned about the road conditions they may experience. However, the actual value or usage of this index by motorists in Oregon was not examined. If the index were to receive broad distribution and were widely heeded by motorists, there could be fewer crashes as a result, which would in turn reduce the measured effects of weather on safety. Demonstration of the index and measurement of the public's receptiveness to it would be valuable in determining the benefits of the current index, as well as the benefits of ongoing index validation and improvement.

6.2.2. Winter Maintenance

It was mentioned that the indices as developed do not lend themselves to a good understanding of how winter maintenance practices can best improve motorist safety. One question related to this would be to examine the temporal relationship between winter maintenance activities and safety. For example, is there a difference in the crash rate observed on a highway in the hour after a snow plow passes over the road, as opposed to two hours later? Does the use of anti-icing lower crash rates in the hour after a storm starts? These questions would require a significant amount of data regarding when winter maintenance activities occurred on different roadways, the level of surface of the roadway before and after maintenance, and hourly traffic volumes during these times. The data requirements are extensive, but if these data are available, they could provide significant insight into optimal winter maintenance practices, and potentially lead to a better understanding of the true benefits and costs associated with different practices.

Most winter severity indices examined in Chapter 2 focused on the relationship between winter maintenance activity costs and winter weather. This paper did not explore that relationship, and it would be interesting to see whether the indices produced for safety correlate with maintenance cost levels. If the correlations are not exact, this may suggest ways in which winter maintenance may be most cost-effective in improving safety.

REFERENCES

1. Leckenby, Donavin A. and Arthur W. Adams, "A Weather Severity Index on a Mule Deer Winter Range," *Journal of Range Management*, Vol. 39, No. 3 (1986), pp. 244-248.

2. McCaffery, Keith R., "Winter Severity Index: 2000-01," Accessed at http://www.dnr.state.wi.us/org/land/wildlife/harvest/Reports/severity.pdf on May 30, 2002.

3. Aprill, Dennis, "Severity Index Tells Tale," *North Country Press-Republican*, February 17, 2002.

4. United States Coast Guard, *Fiscal Year 2000 Performance Report and Fiscal Year 2002 Performance Plan*, pp. 35-36.

5. "Climate Prediction Center – Monitoring & Data Weekly and Monthly Degree Day Summary Explanations," National Weather Service Web Site, Accessed at http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/ddayexp.shtml on May 16, 2005.

6. Hulme, M., "A New Winter Index and Geographical Variations in Winter Weather," *Journal of Meteorology*, Vol. 7, No. 3 (1982), pp. 294-300.

7. Rissel, Martin C., and Douglas G. Scott, "Staffing of Maintenance Crews During Winter Months," *Transportation Research Record 1019*, Transportation Research Board, National Research Council, Washington [DC] (1985), pp. 12-21.

8. Boselly, S. Edward III, John E. Thornes, Cyrus Ulberg, and Donald D. Ernst, *Road Weather Information Systems – Volume 1: Research Report*, Report SHRP-H-350, Strategic Highway Research Program, National Research Council, Washington [DC] (1993).

9. Knudsen, Freddy, "A Winter Index Based on Measured and Observed Road Parameters." In *Proceedings of the 7th International Road Weather Conference* SIRWEC, Seefeld, Austria, March 21-22, 1994, pp. 175-185.

10. Decker, Rand, John L. Bignell, Cambria M. Lambertsen and Kristie L. Porter, "Measuring Efficiency of Winter Maintenance Practices," in *Transportation Research Record 1741*, Transportation Research Board, National Research Council, Washington [DC] (2001), pp. 167-175.

11. Andrey, Jean, Jianzhong Li and Brian Mills, "A Winter Index for Benchmarking Winter Road Maintenance Operations on Ontario Highways," Presented at 80th Annual Meeting of the Transportation Research Board, Washington [DC]: 2001.

12. Cohen, Stewart J., "User Oriented Climatic Information for Planning a Snow Removal Budget," *Journal of Applied Meteorology*, Vol. 20, No. 12 (1981), pp. 1420-1427.

13. Andrey, Jean C., Brian N. Mills and Jessica Valdermolen, "Temporal Analysis of Weather-Related Collision Risk for Ottawa, Canada: 1990-1998." Presented at 82nd Annual Meeting of the Transportation Research Board, Washington [DC]: 2003.

14. Doherty, Sean T., Jean Andrey and Jeffrey C. Marquis, "Driver Adjustments to Wet Weather Hazards," *Climatological Bulletin*, Vol. 27, No. 3 (1993), pp. 154-164.

15. McCullouch, Bob, Dennis Belter, Tom Konieczny and Tony McClellan, "Indiana Weather Severity Index." In *Proceedings of 6th International Symposium on Snow Removal and Ice Control Technology*, Spokane [WA], June 7-9, 2004, pp. 167-178.

16. Adams, Michael J., "Winter Severity Index," Wisconsin DOT internal document, 2001.

17. Thornes, John E., "Cost-Effective Snow and Ice Control for the 1990s," in *Transportation Research Record 1387*, Transportation Research Board, National Research Council, Washington [DC] (1993), pp. 185-190.

18. E-mail from Steve Haider, Minnesota Department of Transportation, June 22, 2004.

19. Nixon, Wilfrid A. and Lin Qui, "Developing a Storm Severity Index," Presented at 84th Annual Meeting of the Transportation Research Board, Washington [DC]: 2005.

20. Nixon, Wilfrid A. and Stowe, Robert, "Operational Use of Weather Forecasts in Winter Maintenance: A Matrix Based Approach." In *Proceedings of 12th International Road Weather Conference SIRWEC*, Bingen, Germany, June 16-18, 2004.

21. Ketcham, Stephen A., Minsk, L. David, Blackburn, Robert R., and Fleege, Edward J., *Manual of Practice for an Effective Anti-icing Program: A Guide For Highway Winter Maintenance Personnel*, Report No. FHWA-RD-95-202, Federal Highway Administration, Washington [DC]: 1996.

22. Chen, Wenjun, Peter Cooper, Stephen Bath, Mario Pinili and Bonnie Hamilton, "Road Sense Index." Insurance Corporation of British Columbia, January 1994.

23. National Climatic Data Center, National Weather Service Web Site, Accessed at http://cdo.ncdc.noaa.gov/CDO/cdo on August 24, 2005.

24. Boselly, S. Edward III, "Road Weather Information Systems: What Are They and What Can They Do for You?" *Transportation Research Record 1387*, Transportation Research Board, National Research Council, Washington [DC] (1993), pp. 191-195.

25. The University of California Transportation Center, "Road Weather Information Systems," Accessed at http://www.uctc.net/mainstream/ on August 24, 2005.

26. Ballard, Lisa, Andy Beddoe, Jason Ball, Eric Eidswick, and Kagan Rutz, *Assess Caltrans Road Weather Information Systems (RWIS) Devices and Related Sensors*, Montana State University, Western Transportation Institute, Bozeman [MT]: July 17, 2002.

27. "Transportation Data Traffic Counting Program," Oregon Department of Transportation Web Site, Accessed at http://www.oregon.gov/ODOT/TD/TDATA/tsm/tvt.shtml on August 24, 2005.

28. "Statistics and Data," Montana Department of Transportation Web Site, Accessed at http://www.mdt.mt.gov/publications/datastats.shtml on August 24, 2005.