Estimating Wet-Pavement Exposure with Precipitation Data: Final Report

Deliverable for the Caltrans Research Project

Entitled "Validate Percent Wet Time Statewide" (Contract 65A0226)

Prepared by

Shaowei Wang, P.E., Co-Principal Investigator

David Veneziano, Ph.D., Research Scientist

Jiang Huang, M.Sc., Research Associate

Xianming Shi, Ph.D., P.E., Principal Investigator

of the

Western Transportation Institute

College of Engineering

Montana State University – Bozeman

Submitted to the

California Department of Transportation (Caltrans)

Division of Research and Innovation, MS-83

1227 "O" Street, PO Box 942873

Sacramento, CA 94273-0001

Attention: Hassan M. AbouKhadijeh

March 2008

Disclaimer

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the California Department of Transportation. 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, Western Transportation Institute, PO Box 174250, Montana State University, Bozeman, MT 59717-4250, telephone: (406) 994-6114, fax: (406) 994-1697. For the hearing impaired, call (406) 994-4331 TDD.

Acknowledgments

The authors would like to thank the Caltrans Project Manager Hassan M. AbouKhadijeh and the Caltrans technical panel for their guidance and assistance throughout the project. In addition, the authors thank Bill Huang, Jaime Garcia, Brad Boehm and Thomas Schriber of Caltrans for their assistance in providing the Wet Table C lists used in later stages of this work, as well as providing an understanding of the process itself.

The authors thank the student researchers who assisted with various aspects of this project, Monika Akbar and Rex Jones.

The authors also thank the California Data Exchange Center (CDEC), the California Irrigation Management Information System (CIMIS), MESOWEST, the National Climatic Data Center (NCDC) and the National Weather Service (NWS) for their provision of the weather data used in this work.

Finally, we would also like to extend our appreciation to Manjunathan Kumar for his leadership in the early stage of this research.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Achieving the best safety record in the nation is a primary goal for Caltrans. To this end, Caltrans develops a list of high collision concentration locations (Table C) every quarter using the Traffic Accident Surveillance and Analysis System (TASAS) database. Table C identifies the ramps, intersections and highway segments with accident rates that are significantly higher than the statewide average in 36-, 24-, 12-, 6-, and 3-month periods. The identified locations in Table C are then investigated individually to evaluate collision risk based on observed frequency. Caltrans also develops a Wet Table C annually that analyzes updated lists of wet accidents alone using a similar methodology as Table C.

The existing table of percent wet time (i.e., wet pavement factors) was developed in 1972 using eleven years of data from 1957 to 1967. A Caltrans task force that investigated the methodology used to develop the Table C and Wet Table C recommended that the table of percentage wet time be updated. This research is the update of that table. The use of more recent data is expected to better reflect current climatic trends being taken into account in identifying significant wet-pavement crash locations in Wet Table C.

In addition to updating wet-percent factors, this research examined current views toward Wet Table C within Caltrans, as well as the practices of other states in identifying wet pavement crash locations. It also examined the preparation and use of precipitation data in generating wet percent factors. Finally, the research examined what, if any, differences arose when using wet percent factors of a finer resolution than the countywide scale employed in the current Wet Table C. A singular value for an entire county may not reflect the climatic variation that occurs along the length of a particular roadway.

Results from updating the wet percent factors, indicated that the new factors that were generated were reasonably similar in value to the older factors. This would suggest that, although changes have occurred over time in terms of precipitation received by county, these changes have not been radical. Given the quantity and accuracy of the data that were available to the researchers in generating the new factors, a factor that excluded the contribution of snow (i.e., included only rainfall data) was also produced. This new type of factor may prove to be useful in areas where substantial rain and snow precipitation both occur throughout the year and thus present inherent problems for measurement accuracy and have differing impacts on safety. As a result, the development of such a factor allows for new avenues of analysis to be made within the Wet Table C process.

McNemar tests were employed to evaluate what, if any, differences existed between the Wet Table C locations identified as having wet-accident significance using the 1972 countywide factors and the updated 2008 countywide factors. The same evaluation was performed comparing these factors against factors generated at a finer resolution. The finer resolution factors were generated for quarter-mile sections along the study roadways, with consecutive identical values subsequently combined to form segments of varying length. Based on the results of the McNemar tests, two conclusions were drawn. First, no significant differences were observed between lists developed using the 1972 and the 2008 countywide factors, indicating that the sites identified for further investigation were similar despite the use of newer data. Second, based on the statistical

evaluation performed on a limited sampling of highways, no difference was found between the lists produced using a singular wet percent time factor and one produced using finer resolution factors. Therefore, the research suggests that Caltrans can continue its use of the countywide average when producing Wet Table C lists.

To generate the new wet percent factors, the dataset required activities to ensure its gaps in the data were filled. Gaps were the result of a number of different causes, including equipment malfunctions, deletion through quality control checks, and others. Results of the processes employed to address missing data indicated that the adopted infill procedures, specifically revision and Nearest Neighbor Frequency Assignment (NNFA), functioned well in addressing the gaps that existed in the data. A simulation test was used to determine the effectiveness of the infill procedures. A total of 384 hours of missing data was simulated, of which 30 hours were originally rainfall hours for one station. A neighboring station located 10 miles away had complete data for the same month. NNFA was employed as the infilling procedure, with the results indicating that the method effectively infilled 24 of the 30 rainfall hours, along with 360 of the 384 non-rainfall hours. Using these figures, the error percentage of infilling was calculated as (30-24)/384=1.6%.

An examination of the various steps and processes employed by the researchers in updating the wet percent factors indicated that that there are some aspects that could be automated through the use of computer programs. These included activities such as data collection and reprocessing. However, the central tasks of data quality control and missing data handling primarily involved human intervention, which was time consuming. At present, it is not possible to develop an automatic data-quality-control algorithm to handle these critical steps by code. As a result, the need remains for some human intervention in the process, at least for the foreseeable future. Additionally, the processes identified as candidates for automation still require further investigation before a conclusion can be drawn regarding the practicality and utility of an automated updating process.

While no significant differences were noted between the 1972 and 2008 factors or the Wet Table C lists produced using them, there were some minor increases and decreases in the factors produced for the new (2008) table. The overall recommendation that can be made as a result of this research is that Caltrans may proceed with phasing out the use of the 1972 factors as soon as it is deemed practical. This recommendation is based on the evidence provided both through the statistical tests performed and the direct comparison of individual county factors. The processes and procedures employed to generate the new wet percent factors appear to have successfully produced new factors that did not significantly deviate from those currently employed. This was primarily evidenced by the similarity in factors that were developed for each county compared to the original factors. Additionally, the Wet Table C lists of site significance developed with each of these factors showed no significant statistical differences.

1. INTRODUCTION

Achieving the best safety record in the nation is a primary goal for Caltrans. To this end, Caltrans develops a list of high collision concentration locations (Table C) every quarter using the Traffic Accident Surveillance and Analysis System (TASAS) database. Table C identifies the ramps, intersections and highway segments with accident rates that are significantly higher than the statewide average in 36-, 24-, 12-, 6-, and 3-month periods. The identified locations in Table C are then investigated individually to evaluate collision risk based on observed frequency. Caltrans also develops a Wet Table C annually that analyzes updated lists of wet accidents alone using a similar methodology as Table C.

The existing table of percentage wet time (i.e., wet pavement factors) was developed in 1972 using eleven years of data from 1957 to 1967. A Caltrans task force that investigated the methodology used to develop the Table C and Wet Table C recommended that the table of percentage wet time be updated. This task force also surveyed 44 Caltrans personnel on their perception and use of Table C and Wet Table C. This survey revealed that more than half of the respondents felt the existing Wet Table C may not accurately identify all locations that require safety improvements.

It can be surmised that outdated wet percent time factors may misidentify locations as being significant and requiring site investigation. Conversely, it can also be surmised that incorrect wet percent time factors used in the development of Wet Table C could result in locations needing safety investigations not being identified. Data has shown that wet pavement has been a crucial factor in California traffic accidents. In 2003, for instance, about 9 percent of fatal accidents occurred on wet pavement. Therefore it is critical that all high-frequency wet-collision locations are identified for further study. Within this context, there is a need to review and update the wet pavement factors used to identify high-frequency wet collision concentration locations in California using TASAS, with the use of the updated factors phased in by Caltrans as soon as deemed practical.

This research develops an updated wet percent time table to replace the old table, which was developed with data that is 40 years old. The new table will better reflect current precipitation trends throughout the state. Taking such updated trends into account will assist Caltrans in efforts to identify and treat high-frequency wet-collision concentration locations in California. Subsequently, this could result in further reductions of wet-pavement collisions, saving the lives of the traveling public and reducing the financial impacts of crashes occurring on wet pavements overall.

Statistics indicate that rain and wet pavement have more significant impacts on road safety than snow and ice. In 2001, nearly 79 percent of weather-related crashes in passenger vehicles in the nation occurred on wet pavement, and nearly 49 percent happened during rainfall [1]. Each fatal crash leads to substantial costs in terms of human suffering and financial loss. According to Caltrans estimates, a 1 percent reduction in wet-pavement collisions in California would save about $13 million, based on 2001 collision statistics. Furthermore, the indirect benefits of safer highways and reduced crashes are significant in terms of relieving congestion, improving overall highway efficiency and, potentially, saving energy. An improved identification process will also lead to higher confidence among district personnel in the versions Wet Table C provided to them annually.

1.1. Background

There is a need to update the wet percentage factors that are used to identify high-frequency wet-collision concentration locations in California (Wet Table C) using TASAS. The existing table of percentage wet time was developed in 1972 using 11 years of data from 1957 to 1967. Figure 1.1 shows changes in precipitation levels in the southwestern United States from 1961 to 1990. The data show variations of up to 50 percent in that time period. Figure 1.1 also shows the predicted changes in precipitation for the 21st Century which may be expected. Climatic change over time as portrayed in Figure 1.1 indicates that the percent-wet-time table developed in 1972 may not reflect the current percent wet time and needs to be updated. The National Assessment Synthesis Team 2000 that examined the climate changes observed that the pattern for regional precipitation changes is more mixed; precipitation levels in some areas are up and they are down in some areas.

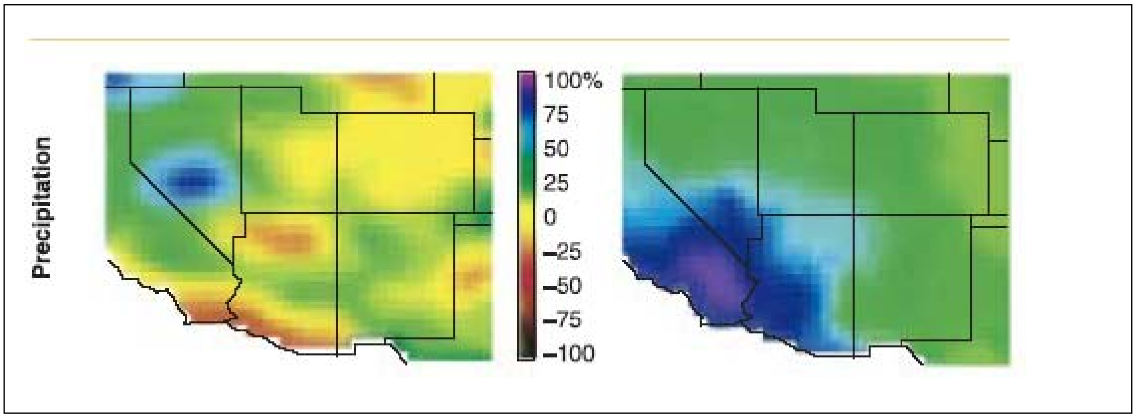

Figure 1.1: Percent Changes observed in Precipitation from 1961 to 1990 (left) and Predicted 21st Century Changes (right)
(Source: http://www.usgcrp.gov/usgcrp/Library/nationalassessment/overviewlooking.htm)

The research presented here will result in validation and updating of the wet percent time table used in the TASAS. This research also looks at other ways to improve the wet percent time table, such as examining the feasibility of annual, automated updates and the use of finer geographical units (e.g., percent-wet-time values for every one-mile highway section instead of the current unit of a single average value for each county).

1.2. Research Objectives

As previously indicated, the current wet-percent factors utilized by Caltrans were developed in 1972. The data used to develop these factors is now well over 40 years old. The strong likelihood of climatic change over time suggests that the 1972 percent-wet-time table does not reflect current conditions and needs updating. The primary objective of this research was to update the countywide wet percent factors using recent precipitation data from 1995 through 2005. The use of more recent data is expected to better reflect current climatic trends, resulting in wet-percent factors that more accurately identify significant wet-pavement crash locations in Wet Table C.

In addition to updating wet-percent factors, an additional objective of this research is to examine other aspects related to the preparation and use of precipitation data in generating wet percent factors. For example, the number of weather stations collecting data in California has grown exponentially since the original wet percent factors were developed. The acquisition, storage and processing of these data require entirely new strategies compared to those employed when the original factors were developed. This research addresses these issues and examines how the development of future wet-percent factors may evolve as a result.

A final objective of this research is to examine what, if any, differences arise when using wet percent factors of a finer resolution than the countywide scale employed in the current Wet Table C. A singular value for an entire county may not reflect the climatic variation that occurs along the length of a particular roadway. This research examines to what extent such differences may exist.

1.3. Report Organization

The research conducted in this report is presented in ten chapters, plus corresponding appendices. This chapter introduces the research problem and discusses why it is important. Chapter 2 presents the results of a thorough search and review of available literature related to the determination of wet-pavement exposures, existing methods in quality control of precipitation data, methods for interpolating missing data and, most importantly, how the National Weather Service (NWS) and other weather institutes forecast precipitation levels and develop isohyetal lines or precipitation contour maps. Chapter 2 also presents the results of an online survey of Wet Table C users in various Caltrans districts, as well as results of another online survey of state departments of transportation. These surveys document current practices in measuring traffic exposure to wet-pavement conditions, both within California and nationally.

Chapter 3 presents the methodology employed in updating the wet percent time table. This includes a high-level overview of the precipitation data sources and acquisition, the quality control checks employed to verify the consistency and accuracy of that data, and the investigation, analysis and selection of a method to handle missing data. The thrust of this chapter is the presentation of research steps in the form of flowcharts, which illustrate the processes employed and presented in subsequent chapters.

Chapter 4 discusses the data sources and activities related to the collection and processing of precipitation data. The research used precipitation data from five sources: the California Data Exchange Center (CDEC), the California Irrigation Management Information System (CIMIS), MESOWEST, the National Climatic Data Center (NCDC), and the National Weather Service (NWS). As each of these sources collects and saves its data in proprietary format(s), data reprocessing was required. Chapter 4 describes the efforts made by the researchers to develop a single database containing the precipitation elements of interest in a common, unified format.

Chapter 5 describes the quality control activities undertaken to ensure that the unified dataset developed previously was of the highest quality, consistency and completeness. As data were acquired over a lengthy period of time from a variety of independent sources that used different sensors and measurement techniques, it was not surprising that quality issues were present throughout the data. To address this, the researchers employed three levels of quality control, which are discussed in detail in Chapter 5. Level 1 sought to identify obvious data errors (e.g., negative precipitation values), Level 2 compared data between stations to ensure reasonable values were present between neighboring sites. Level 3 employed manual checks to examine and address quality control issues raised by the previous two levels.

Chapter 6 discusses the procedures employed to address the problem of missing data for various sites. Given the dense network of collection sites and the length of time over which data were collected and employed in this research, instances of gaps in the data were to be expected. To address these data gaps, procedures had to be identified and tested to determine how to best in-fill missing data to provide a continuous dataset for a given site. To this end, the research employed two strategies. The first, a revision strategy, assumed that if the missing period was less than 48 hours long, the value of the missing data should have remained the same as at neighboring stations. Subsequently, the missing values were in-filled by using the values of neighboring stations. The second strategy, employed when data were missing for more than 48 continuous hours, is referred to as the Nearest Neighbor Frequency Assignment. This strategy in-filled missing data at one station by assigning to it the observed values of its nearest neighboring stations that had data available for the missing period. Each of these strategies, including the decision-making criteria associated with them, is discussed in greater detail in Chapter 6.

Chapter 7 discusses the research steps involved in the development of the new wet percent table. This portion of the research utilized the previously sanitized data in developing raster maps based on the annual average wet percent time calculated for each station. This work made use of the Zonal Statistics package provided in ArcGIS's Spatial Analyst. The generated raster maps were used to interpolate the wet percent time for every location with known latitude and longitude in California. Maps were developed for every year of the data time period (1995-2005) based on the annual average of percent wet time for each weather station. Based on these maps, an 11-year average value of annual wet percent time was calculated for each county to produce the new county-level percent wet time factor.

Chapter 8 examines the feasibility of updating the wet percent time annually using data from selected weather stations, to be implemented in conjunction with TASAS. Due to climate change, the amount of precipitation may vary significantly from year to year, and an average percentage wet time may not represent current conditions for any given year at specific locations. It should be noted that the Wet Table C contains recommendations for 0.2-mile sections of highways and calls for spot improvements for improving safety.

Chapter 9 looks at the use of finer geographical units - i.e., an individual percent-wet-time value for every one-mile highway section instead of the current unit of a single average value for each county. The single, countywide number currently employed (and updated in this research) fails to take into account California's varying geography and the different microclimates that may exist within a county. The use of one wet-percent-time value for a county may lead to a situation where wet-accident locations in a given area are being incorrectly identified as being significant based on a countywide wet percent factor whose calculation was skewed by excessive or minimal precipitation in another area. This chapter examines how the lists developed using a singular factor compare to those developed through the use of a segment-based factor.

Chapter 10 ties together the results of the work performed in the previous chapters. This chapter provides a brief overview of the steps that were employed to complete each research activity, as well as the end results of that activity. In addition to summarizing the completed work, recommendations for future research efforts are also presented.

2. LITERATURE REVIEW

This chapter details the findings in literature related to the major topics of this research. As shown in Table 2.1 the topics included the methods of measuring wet-pavement exposure and developing isohyetal lines, quality control of precipitation data, handling missing data, and previously conducted studies that have established percent wet-time factors for other regions in the nation.

Table 2.1: Topics Covered in Literature Review

	Area of Interest	Sub-Areas
Topic 1	Wet-Pavement Exposure Measurement and Isohyetal Lines	Measurement of wet-pavement hours
		Wet-pavement definition
		Wetness due to fog, evaporation time, etc.
		Techniques to establish isohyetal lines
		Alternatives to isohyetal lines
Topic 2	Precipitation Data Quality Control	Issues with hourly precipitation data
		Quality control measures for precipitation data
Topic 3	Handling Missing Data	Methods to handle missing data
Topic 4	Similar or Related Studies	Studies that developed percent wet time factors or related to wet-pavement accident reduction

2.1. Measuring Vehicle Exposure to Wet Pavement and Developing Isohyetal Lines

Adequate pavement skid resistance is critical to ensure highway safety. The various factors that affect pavement skid resistance include pavement surface condition, traffic volume and speed, locational attributes (slopes, sharp curves and intersections), and, of interest in this research, pavement wetness [2, 3].

To identify traffic accidents that occur due to pavement wetness, measurements of wet time and vehicle exposure to wet pavement are necessary. Wet time refers to the proportion of time during which pavement is damp enough to cause traffic accidents and is measured on an hourly or a daily basis. Measurement is critical to programs established to reduce wet-pavement accidents, since the product of wet time and vehicle-miles traveled reflects the rate of wet-pavement accidents in a region. Previous wet-pavement exposure estimation methods have focused on estimating the proportion of time during which the pavement is wet, usually on an hourly basis, and then linking vehicle exposure to that wet pavement and the occurrence accidents [4].

Frederick and Miller (1979) defined the geographic and frequency variation in short-duration rainfall over the Eastern and Central United States [5]. The Department of Water Resources of California extracted annual maximum values for short-duration rainfall from recording rain gauges distributed throughout California. The state was divided into 14 meteorologically and topographically homogenous regions. For each of the 250 stations having at least 15 years of annual maximum 5- to 60-minute amounts, frequency tables were constructed using *Fisher-Tippett Type I distribution*. All subsequent analysis was conducted using ratios of N-minute precipitation (N=5, 10, 15, 30 and 60) for return periods of 2, 5, 10, 25, 50, and 100 years. This study supported the hypothesis that as the duration decreases, the N- to 60-minute ratios are less dependent on elevation. Another approach of examining the variation of ratios by elevation was to group stations by elevation modes. The result indicated that for areas below 500 feet, the average ratio was the highest and decreased as the elevation increased. Analysis was conducted for 30- to 60-minute ratios for the 2- and 100-year recurrence intervals. The lowest ratios were found in the coastal and orographic regions. In contrast, the highest ratios were found in the Central Valley and desert areas.

In the Eastern and Central United States, the N- to 60-minute ratios decreased with an increase in return periods. In California (San Francisco area), the 5- to 60-minute ratio at the 100-year return period was found to be 116% of the same ratio at the 2-year return period. The study computed the variation of ratio of 10- to 60-minute values for four regions in California. Results suggested that the ratio increased as the recurrence interval increased. The increase was found to be not completely linear, but rather asymptotic at some point. Of the 250 stations, only 18 had 5-, 10-, 15-, and 30- to 60-minute ratios that decreased with increasing return period. Forty other stations had one or more of N-minute that decreased with increasing return period. Remaining stations had N-minute to 60-minute ratios that increased for all durations. Another duration investigated in the study was the 15-minute duration. The highest ratios were found in the Central Valley and the southeast desert. The lowest ratios were observed in the coastal areas with the Sierra Nevada being slightly higher.

Karr and Guillory (1972) described a method for computing *wet exposure* (i.e., vehicles exposed to a wet-pavement condition) for the state of California as documented in a study entitled “A Method to Determine the Exposure of Vehicles to Wet Pavements [6].” The study related accident data to actual traffic counts and precipitation data. The National Weather Service provided the hourly precipitation values accumulated at its 350 continuous weather recording stations located throughout California for the period 1957-1967. In the Caltrans study, Part 8, “Grooved Pavements” and Part 9, “Open Graded Asphalt Concrete Overlays” provided a good database of wet-pavement accidents before and after the minor improvements. The study covered a total of 3 billion vehicle miles of traffic that produced 16,385 accidents. Nine traffic count stations were selected at random to measure traffic. Eight stations counted traffic for one week of every month and the ninth station counted traffic for every hour of the year. Annual and monthly average daily traffic figures were also available from “Traffic Volume on California State Highways.” Based on the precipitation, traffic, and accident figures, a wet exposure methodology was developed. This methodology is presented in Appendix A, Note 1. The result of this methodology was the so-called “R” Method of Caltrans, which provided a relatively more accurate wet exposure estimate compared to the Actual Wet Exposure. The estimated values varied between –1% and +7%. The Caltrans report further stated the R method provided the best estimate of actual wet exposure, although traffic counts and weather data were not sufficiently accurate.

Blackburn et al. (1978) argued that one of the essential factors in developing an accident rate-skid number relationship was the wet-pavement accident rate [7]. This factor was considered important because wet-pavement conditions vary greatly with geographic location. For the before-and-after studies, wet-pavement condition might change with time at the same location. In this research, the wet-pavement accident rate for each study section was dependent on the exposures to wet-pavement conditions. The exposure estimate for each study section was made for the same period as the accident date for that section. Weather records from 70 Local Climatological Data (LCD) reports were used to estimate the exposure time for all geographic areas where the test and control sections were located. The methodology employed in this research is presented in Appendix A, Note 2. The research found that the passage of traffic tends to minimize the regional differences in pavement drying times. Based on the study findings, a time of ½ hour for pavement drying was proposed.

Peters et al. (1980) described a methodology that estimated the wet-pavement exposure for arid climates, specifically the state of Arizona. In an effort to approximate the amount of wet-pavement traffic volume, or wet AADT, three methods were examined [8]. A discussion of these is presented in Appendix A, Note 3.

Harwood et al. described a FHWA study undertaken by MRI and the Pennsylvania Transportation Institute to improve the ability of state agencies to estimate the wet weather exposure measures [3]. Based on the research findings, an improved method of estimating wet-pavement exposure from available weather records was developed by Harwood et al. [4]. This method included explicit consideration of the drying period during which pavements remain wet after rainfall ceases and the period that pavements are wet due to melting of snow and ice. The original MRI technique also considered wet time due to trace amounts of rainfall (less than 0.01″ per hour) that were part of a longer period during which measurable rainfall occurs, but ignored periods of rainfall composed entirely of trace amounts. The development of the technique included field observations of pavement drying times, and the technique was validated using wet-pavement exposure data from a moisture sensor implanted in an Interstate highway bridge near Iowa City, Iowa.

2.2. Quality Control of Precipitation Data

Quality control of precipitation data is an important procedure before the data are used for estimating wet percent time at any weather station. The absence of quality control procedure can result in poor quality data, or noise that severely limits their usefulness for this project. Often such problematic data appear as outliers in precipitation and other weather observations collected by various weather networks. Quality control (QC) is essential to identify and flag those data that are potentially in error. The following sections first synthesize the information on the types and causes of data errors, then looks at information on procedures taken to control the quality of weather data (including precipitation data).

2.2.1. Previous Studies

Common errors in hourly precipitation data may be derived from the observational station or during the data entry process. For instance, observer error occurs when incorrect measurements were recorded as a result of careless reading of gauges. To identify such errors, four methods are generally used for quality checks of precipitation data, including extreme value check, internal consistency check, temporal consistency check, and spatial consistency check. The outcomes of such checks are typically of a "yes" or "no" nature; when a problem with the data is observed, it is flagged with notations.

Eischeid et al. (1995) described temporal tests and spatial quality-control interpolation methods in their study of raw data from the Global Historical Climate Network (GHCN), where each station selected was required to have at least 20 years of records within the 1951-1980 period [9]. The temporal test assumes that the individual monthly value should be "similar" to values for the same month for other years. Outliers are identified based on limits determined from a multiple of the inter-quartile range (IR, 75th percentile minus the 25th percentile) obtained by sample distribution of each month for each station. Outlier is flagged when $Xi – q50 > f\ IR$ where Xi is the monthly mean of year i, q50 is the median and f is the multiplier. The temporal check is designed to determine whether or not the month in question is consistent with the sample population of other such months for the same station. The author also listed six different spatial interpolation techniques to estimate each monthly time series. For both temperature and precipitation, a multiple regression scheme was found to be the best estimator for the majority of records.

Schmidlin et al. (1995) developed and automated QC procedures for daily snow water equivalent data [10]. Two approaches were commonly used in QC of climatic data (Reek et al. 1992): 1) comparisons with nearby stations to detect inconsistencies; or 2) a determination whether a datum is outside of reasonable ranges, or is not logically consistent with observations from adjacent periods or with simultaneous ancillary measurements. The former is appropriate only if the stations are not too far apart to allow reliable comparisons among neighboring stations. These two approaches were implemented in the study through conducting recording error and limit checks, as well as a day-to-day consistency check/micro-scale variability check. The authors also pointed out that any flagged value by the automated QC procedure must be evaluated by a human analyst. Not only was the automated procedure not perfect, but a QC flag might be associated with a valid measurement that was inconsistent with a previous, erroneous one. When potential errors were identified by the automated procedure, subsequent days were compared to the most recent measurement that was apparently correct.

Feng et al. (2004) summarized QC methods used specially for precipitation and other weather data [11]. The study examined the daily meteorological data from 726 stations in China from 1951 to 2000, and developed an unprecedented climatic dataset that contains 10 daily variables including precipitation. The characteristics of the original stations' data and QC methods used in developing this dataset are described in Appendix A, Note 5.

Numerous factors that could affect the consistency of the record at a given station were also identified by Feng et al., including: a) damage and replacement of a rain gauge; b) change in the gauge location or elevation; c) growth of high vegetation or construction of a building; d) change in measurement procedure; or e) human, mechanical, or electrical error in taking readings [11]. A method called *double mass analysis* is used for adjusting inconsistent data.

Hubbard et al. (2004) illustrated a QC approach that allowed tailoring to regions and sub-regions and introduced a new spatial regression test [12]. Threshold testing, step change, persistence, and spatial regression were included in a test of three decades of temperature and precipitation data at six weather stations representing different climate regimes. The study underscores the fact that precipitation is more difficult to QC than temperature. The new spatial regression test presented in this document outperformed all the other tests. The study incorporated four procedures as described in Appendix A, Note 6.

Eching and Snyder (2004) established statistical criteria to assess quality and reasonableness of hourly and daily weather data for the California Irrigation Management Information System (CIMIS) weather stations [13]. The QC criteria, based on means (m) and standard deviations (s), were developed from historical CIMIS weather station data. Two statistical QC limits (3s and 2s for upper and lower control limits) were developed. A new version of the control chart, a time-variant control chart was introduced.

Urzen et al. (2004) described a multi-sensor approach to the real-time QC of precipitation data used in the *PrecipVal* system by the National Climate Data Center (NCDC) [14]. The data layers used for QC included station data, radar and satellite data, and rapid update cycle (RUC) model output, whenever available. Based on the number of independent data layers agreeing with the observation, observation confidence was generally assigned. The authors argued that adding poor-quality data could increase, rather than decrease, uncertainties in the QC system.

You et al. (2005) developed a spatial regression test in their research [15]. Measurements from neighboring stations were used in a spatial regression test to provide preliminary estimates of the measured data points. The new method assigned the weights according to the standard error of estimate, not the distance between the target station and nearby stations. As such, the spatial test was employed to study patterns in flagged data in the extreme events.

Daly et al. (2005) discussed opportunities for improvements in the QC of climate observations in the context of increased supply of and demand for climate data [16]. They argued that technological advances provided “opportunities for qualitatively different approaches to QC, methods that are sophisticated, largely automated, data-rich, updatable, and capable of furnishing quantitative error and confidence information.” The ultimate goal was a QC approach that was “self-consistent and physically plausible, in accord with known principles of how the atmosphere works,” and that could be “updated to reflect changes in our knowledge base.” The authors suggested that there was much benefit in estimates of observational validity and estimation uncertainty that are quantitative and continuous, rather than categorical (such is the outcome of traditional QC methods). As an example, they presented the first generation of a spatial QC system developed for the USDA Natural Resources Conservation Service (NRCS) SNOTEL temperature data, termed Probabilistic-Spatial Quality Control (PSQC) system. The system was spatially oriented and used a knowledge-based algorithm to make predictions. It operated on the assumption that spatial consistency, if assessed accurately, was a useful indicator of data validity. The predictive tools were adopted from a knowledge-based climate mapping system developed at Oregon State University, termed PRISM (Parameter-elevation Regressions on Independent Slopes Model). Experience indicated that PRISM provided a relatively high degree of skill to the spatial interpolation process, especially in complex regions, involving a climatologically aided interpolation (CAI) technique.

2.2.2. Key Findings from Previous Studies

Generally, QC measures applicable for weather data include the extreme value check and internal consistency check, designed to review the data from a single station to detect potential outliers [9]. Recently, the use of multiple stations in QC procedures has proven useful. For example, spatial consistency tests were successfully utilized to identify outliers by comparing a station's data against those from neighboring stations [9, 11, 12, 17]. Such tests involve the use of data from neighboring stations to evaluate the measurement at the station of interest, by weighting according to the inverse of the distance to the location [17. 18], or through other statistical approaches (e.g., multiple regression, [9]; bivariate linear regression test, [12]; spatial regression test [12]).

The spatial regression test (Hubbard et al. 2005) did not assign the largest weight to the nearest neighbor but, instead, assigned the weights according to the standard error of estimate between the station of interest and each neighboring station [12]. Hubbard et al. (2005) used seeded errors to test the performance of the threshold method, the step change method, the persistence test, and the spatial regression test [12]. It was found that the spatial regression test outperformed the other three methods, which missed many of the errors identified by the spatial regression.

For this project, the QC of the precipitation data in the state of California will follow the guidelines established by the National Weather Service in the Technique Specification Package 88-21-R1 [19]. For each weather station, only the data that pass both QC procedures will be used for calculating the wet percent time at the weather station location.

1) Validity checks: for an accumulated precipitation of 24 hours, if the value falls outside of the lower and higher control limits, the observation is flagged as suspicious and is discarded.

2) Internal consistency checks: examine reasonable, physically possible meteorological relationships among elements within an observation. For the precipitation observations, the following rules may apply, depending on the data availability.

- Rule 1 Present weather vs. AP<24: If Present Weather = Rain, Snow, Other Frozen Precipitation Types, Showers, or Thunderstorm and AP<24 = 0, Then "Fail"
- Rule 2 Past weather vs. AP<24: If Past Weather = Drizzle, Rain, Snow, or Showers, and AP<24 = 0, Then "Fail"

- <u>Rule 3</u> Snow depth vs. AP<24: If Snow Depth Increases and AP<24 = 0, Then "Fail"
- <u>Rule 4</u> Snowfall vs. AP<24: If Snowfall >0 and AP<24 = 0, Then "Fail"
- <u>Rule 5</u> AP24 vs. AP<24: If AP<24 Added Over a 24-hour Period is Not Equal to AP24, Then "Fail"

2.3. Handling Missing Data

A complete dataset of hourly precipitation is crucial to determining the wet time as well as to measuring the vehicle exposure to wet pavements. One issue with the historical data is that some weather stations stop operating during the time period of interest, and others became operational. As such, the set of weather stations included in the data will vary from year to year and from month to month. The result is a significant amount of missing data, mainly from weather stations that were not operational during certain time periods. Another source of missing data involves errors in observation reporting and recording, equipment faults, and values outside the tolerance limits. Such faulty data may be identified through QC procedures and will be treated as missing data once they fail the QC tests.

Methods of handling missing or incomplete data depend upon how data points became missing. There are three unique types of data missing mechanisms. The data is called "missing completely at random" if the probability that a variable is missing in an observation is unrelated to the value of the variable itself. When this probability can be predicted by another variable in the dataset, the data is called "missing at random." In both cases, missing data points are ignorable; that is, we can simply delete the observations that contain missing values and run the analysis on what remains. If, however, a dataset does not meet either of the above two conditions, then the pattern of missing data is non-random and is explainable only by the variable on which values are missing. In this case, the missing data is non-ignorable, and any analysis would have to include a model that accounts for missing data. The following sections synthesize the information on the methods of handling missing data, with a focus on those that may be applicable for this project.

2.3.1. Previous Studies

Commonly used interpolation methods for meteorological applications include nearest-station assignment, inverse-distance weighting, inverse-distance-square weighting, Thiessen-polygon method, orthogonal-polynomial approximation, Lagrange method, interpolation by splines, kriging, and interpolation by empirical orthogonal functions. Each method has its merits and is applicable according to temporal length scale, spatial length scale, stationarity, and variability of the field under consideration. Statistical methods of interpolating missing precipitation data in the multivariate study, in which more than one variable is studied, include single imputation (mean imputation, conditional mean imputation), multiple imputation, (full information) maximum likelihood and expectation maximization. Because there is only one variable studied in this project, we will focus on the univariate methods for missing data only.

Paulhus and Kohler (1952) described three methods used specially for interpolating missing precipitation and other weather data [20]. A detailed presentation of procedures is included in Appendix A, Note 7. Among the three interpolation methods, two of them, namely, the normal-ratio and 3-station-average, were selected for use by the Weather Bureau.

Samuel et al. (2001) developed a method coupling of the methods of 1) Inverse-Distance Weight, and 2) Nearest-Station Assignment for interpolating the precipitation data, based on the fact that the precipitation field can have a short spatial correlation length scale and large variability [21]. The interpolation used the observed daily data for the period of January 1, 1961, to December 31, 1997, (13,514 days) within the latitude–longitude box (458–648N, 1168–1248W). This method was able to reliably calculate not only the number of precipitation days per month, but also the precipitation amount for a day. The temperature field had a long spatial correlation scale, and its data were interpolated by the inverse-distance-weight method. Cross-validation showed that the interpolated results on polygons were accurate and appropriate for soil quality models. The computing algorithm used all daily observed climatic data, even though some stations had records for a very short time or only summer records. A few relevant methods were briefly reviewed and assessed for their suitability for processing a long time series (37 years) of daily climatic data for use by soil quality models. When the observational stations were very sparse and the climatic conditions were complex, this method would result in substantial spatial errors for a climatic parameter that varied over short length scales.

As described in Section 2.2.1, the nearest station method was used by Feng et al. (2004) to detect the outliers by comparing the data of neighboring stations [11]. The same regression approach was used to interpolate missing data. This approach is detailed in Appendix A, Note 8. The number of neighboring stations used in the applied equations was not fixed. Instead, the number varied depending on the availability of station data for the year/month in question. Accordingly, the regression models also changed in time. Moreover, the surrounding stations that might be optimal for a particular calendar month (e.g. January) might not be optimal for a different month (e.g. July). Thus, the spatial outlier check and estimation of missing data were applied for individual months.

2.3.2. Key Findings from Previous Studies

The methods listed in the project proposal such as *listwise deletion* (delete observations with missing values), *expectation maximization* and *multiple imputation* are appropriate for estimating missing data in general, e.g., the multivariate study. For this project, since the precipitation data have their own statistical characteristics (large variability and short spatial correlation scale), they can be better handled using various univariate methods.

The selected interpolation methods must satisfy the following:

- To provide the best fit for the annual wet percent time;
- To dynamically adapt to the number of stations in order to use all the hourly precipitation data available; and

- To provide realistic estimates of hourly precipitation data and eliminate gaps in the dataset.

Some thoughts on the common approaches for interpolating missing precipitation data are described as follows.

(1) Inverse-distance-weighting: It is based on the assumption that the influence of the nearby observations on data at an interpolated point solely depend on, and decreases with, the distance between the stations. This method is most suitable for temperature data.

(2) Nearest-neighbor method (similar to the Three-Station-Average Method or Normal Ratio Method): It takes full advantage of the nearby observations. Each station with the missing data is assigned the observed value of the nearest station that had data for the hour. It will not be practical when the stations are sparse and the climatic conditions are complex. It is recommended when priority is given to the computational speed.

(3) Kriging: It minimizes the mean-squared error between the estimated field and the true field, when the covariance field is known (in the form of a variogram). Kriging preserves more variance than the inverse-distance-weighting method and is a spatial interpolation tool often implemented with geographic information systems (GIS). Kriging requires a field that is relatively stationary in time and homogeneous in space. Such a requirement makes it a poor choice for the handling of missing hourly precipitation data (Daly et al. [22]).

(4) Empirical orthogonal function method: It is the most effective tool in dealing with spatially inhomogeneous climate fields but might yield unreasonable results when the field is highly non-stationary, as in the case of hourly precipitation data.

2.4. Wet-Pavement Accident Reduction Studies

Hankins et al. (1971) investigated the influence of vehicle and pavement factors on wet-pavement accidents [23]. By examining 501 wet-weather vehicular accidents, the study analyzed five variables potentially related to the friction available at the tire-pavement interface, including the tire pressure, trend depth, and speed of the accident vehicle, as well as friction and macrotexture of the pavement surface at the accident site. The study concluded that the lack of pavement texture, low pavement friction, high vehicular speed, worn tires, and high vehicle tire pressures all contributed to wet-pavement accidents. These variables were found more significant for certain accident types than for others.

Smith and Elliott (1975) performed a two-year before-and-after study of grooved Portland Cement Concrete (PCC) pavement in California, the goal of which was to validate the findings of a report on the same subject published in 1972 [24]. The results suggested that wet-pavement accidents (fatal and injury) decreased an average of 70 percent on grooved payment compared with a 2 percent reduction on the control sections. The dry-pavement accident rates decreased 21 percent and 24 percent on the grooved and control sections respectively. The largest reductions in wet-pavement accidents were in sideswipe and hit-objects followed by rear-end and miscellaneous accidents. In short, results of the study showed that grooving led to significant reduction in wet-pavement accident rates. However, no effect of grooving was found on dry-pavement accident rates. The study also showed that grooving did not impair motorcycle safety during either wet or dry conditions. Though lighter motorcycles were more sensitive to grooves, grooved pavements did not present a serious control problem to the riders. As part of the study, the authors provided a linear equation to predict the reductions in wet-pavement accidents from future projects. This equation is presented in Appendix A, Note 9.

Dean (1976) investigated the relationship between accident experience and average wet pavement skid resistance, using data from rural highway sections in New Brunswick, Saskatchewan, Ontario and Kentucky [25]. Two classes of highway were analyzed: two-lane undivided rural arterials having posted speeds of 60-65 mph (96-104 km/hr) and AADT volumes of 0-8400 vpd; four-lane divided rural freeways and parkways having posted speeds of 70 mph (112 km/hr) and AADT volumes of 1100-34000 vpd. The data exhibited strong nonlinear variation and considerable scatter. Ten-point moving average plots of wet-pavement accident experience vs. SN40 served to subdue scatter and identify SN40 levels at which significant increases in wet-pavement accident experience occurred. Plots of wet-pavement accidents per mile vs. SN40 appeared to have less scatter than those in which the significant relative frequency of wet-pavement accidents was used as the wet-pavement accident variable. In addition, wet-pavement accident experience averaged over two or more years gave less scatter than plots of yearly wet-pavement accident experience. Two-lane undivided rural highway data exhibited a higher level of wet-pavement skid resistance demand than that for four-lane divided rural highways. Wet-pavement skid resistance ranges at which marked increases in wet-pavement accident experience occurred were, for two-lane rural arterials: SN40 = 55-60; for four-lane rural freeways and parkways: SN40 = 43-50.

Holbrook (1976) developed quantitative relationships between accident occurrence and variables such as surface, weather, and seasonal factors in order to build a rational maintenance program in the state of Michigan [26]. Using almost 40,000 accidents recorded at 2,000 intersections, a wet surface model was developed taking skid number, wet time, and seasonal weather effects into account. The model is presented in Appendix A, Note 10. A key finding of the model's development was that the effect of monthly wet time was an important variable. Its effect on wet-pavement accident percentages was approximately logarithmic for all months and surface friction conditions. The study also indicated that variations in the monthly surface wet time occurred in Michigan on a predictable yearly basis. To the extent that traffic volumes also had seasonal variations, monthly wet time should be included in resurfacing decision. The author suggested that a resurfacing policy taking account of regional and monthly wetness patterns would be valuable. It was suggested that where seasonal and regional wetness patterns exist, consideration of surface friction improvements should include both wet time and skid number to achieve full potential of wet-pavement accident reduction.

Runkle and Mahone (1976) outlined a systematic program for identifying and treating high or potentially high wet-pavement accident sites in Virginia [27]. There were two databases used in the wet-pavement accident reduction program—the state accident database, and the skid resistance number database maintained by Virginia DOT. Potential high wet-pavement accident sites were selected in 0.5-mile (0.8-km) segments incremented by 0.1-mile (0.16-km) lengths, based on high accident occurrence and low pavement friction value. The developed process to identify potential sites is outlined in Appendix A, Note 11.

Levy (1977) investigated the effect of pavement skid resistance on wet-pavement accidents in Indiana [28]. A wet-pavement accident index was formulated and used as an indicator of the relative safety in comparing sections of highway when wet. Data analyses to correlate the wet pavement accident index with average skid number were conducted on Interstate, four-lane, and two-lane road sections. There was no single value of minimum skid number found to be applicable to all road sections. The type of road, its volume, geometry and amount of access control should all be considered in determining minimum skid number standards.

Dierstein (1977) described a strategy for reducing wet-pavement accidents in Illinois, where a disproportionate number of wet-pavement accidents were associated mostly with intersections, curves, hills, railroad crossings, and interchange areas [29]. A long-term strategy for reducing wet-pavement accidents involved upgrading and prolonging friction characteristics in new or existing pavements. Specification changes limited the use of crushed stone and required either slag or a 50-50 blend of slag and crushed dolomite or slag and crushed gravel in bituminous surface courses depending on highway class and traffic volume. A Portland cement concrete special provision required that the final finish be obtained by use of an artificial turf drag immediately followed by a mechanically operated metal-comb transverse grooving device. In existing surfaces, friction could be improved by bituminous resurfacings containing coarse aggregates with high friction characteristics or by grooving, planing, milling, profiling, repaving, and acid etching.

Blackburn et al. (1978) argued that one of the essential factors in developing the accident rate-skid number relationship was the wet-pavement accident rate [7]. In this study, the raw data of 100,000 80-characters were processed using a number of computer programs. Results suggested that a small but significant relationship between wet-pavement accident rate (AR) and skid number (SN40) existed. Highway type, area type, and ADT were found to have significant effects on this relationship. Second, the slope of the AR-SN40 relationship was found to be sensitive to the dry-pavement accident rate. Third, the influence of pavement texture, exposure to high intensity rainfall, and geometric variables on the AR-SN40 relationship was found insignificant. Finally, there were strong correlations between wet-pavement and dry-pavement accident rates. Such correlations were higher for urban than for rural sections, and higher in the after period than in the before period. The correlation for urban, multilane, uncontrolled access sections was found significantly higher than for the other sections and the correlation for rural, multilane, controlled access was found significantly lower.

The National Transportation Safety Board (NTSB 1980) undertook a special study to determine the magnitude of the wet-pavement accident problem nationwide, the significance of the wet-pavement accident locations, and the characteristics of these accidents [30]. Data developed by NTSB showed that during 1976 and 1977, 13.5 percent of all fatal accidents occurred on wet pavement, while precipitation occurred only about 3.0 to 3.5 percent of the time nationwide. This indicated that fatal accidents on wet pavement occur 3.9 to 4.5 times more often than might be expected, and that the wet-pavement accident problem should be of concern to all states. To measure the performance of the activities of states aimed at reducing wet-pavement accidents, NTSB developed a Wet Fatal Accident Index (WFAI) for each state. This method indicated an area in the United States with good performance and a belt with poorer-than-average performance.

Kamel and Gartshore (1982) presented a program for identifying and treating high or potentially high wet-pavement accident sites in Ontario [31]. Highway locations with an excessive rate of wet-pavement accidents were identified and ranked utilizing computerized accident data files from the Ontario Ministry of Transportation and Communications. Criteria for site selection, procedures for subsequent site investigation, and selection of appropriate remedial measures were outlined and discussed. Modified bituminous surface course mixes maintained better surface textures and provided longer lasting skid resistance characteristics. These mixes were used for black spot treatments, and in new surface construction on main highways. Collision data indicated that rehabilitating pavements with low friction levels and high wet-pavement accident rates had resulted in substantial reductions in accidents.

Dahir and Gramling (1990) conducted a NCHRP synthesis that provided information on the programs used by a number of agencies in gathering data and correcting areas of potential wet-pavement accidents [32]. This report summarized agencies' programs in areas such as accident reporting, vehicle testing, friction testing, corrective actions for problem areas, and tort liability and gave some general guidelines for the content of a wet-pavement safety program.

Collins and Pietrzyk (2001) described the evaluation of a fully automated motorist warning system for wet-pavement conditions [33]. The demonstration took place on one expressway interchange ramp where 69 percent of total recorded crashes had been classified as "run-off" crashes during wet-pavement conditions. Since less than one-half of the wet-pavement crashes occurred during rain, the dynamic motorist warning system was developed as a potential solution to wet-pavement accident reduction by attracting attention to the advisory speed limit signs and thus encouraging motorists to reduce vehicle speed. A pavement sensor embedded in the roadway activated two flashing beacons located above the signs whenever moisture was detected. Infrared radar recorded vehicle speed at the site. Speeds and volumes were grouped into a matrix according to weather conditions and time periods (sunlight visibility and peak traffic hours). In total, more than 27,000 wet-pavement vehicle speeds were compared before and after system activation. The average reduction in travel speed was 10mph (16 km/h) during heavy rain and 5mph (8 km/h) during light rain; the standard deviation for vehicle speed also was reduced after system activation. No run-off crashes were reported at the site after the first week of the evaluation period.

2.5. User Surveys

In addition to the literature review presented in the previous sections, two surveys were conducted to determine the state of the practice with respect to wet weather. A state survey was conducted within California to document the concerns of the Wet Table C users, with the intention of using these findings to guide the present project. In addition, a national survey was conducted to determine what, if any, activities other states were performing with respect to wet-pavement accident analysis. The key findings of these surveys are presented in the following sections. Additional information of interest, such as tables and graphs, are presented in Appendix B (California Wet Table C Survey) and Appendix C (National Survey).

2.5.1. Survey of Wet Table C Users

A follow-up survey to that previously conducted by a Caltrans task force that investigated the methodology used to develop Wet Table C Wet Table C was conducted to determine the latest views toward the process. The survey included users in various Caltrans Districts who were identified through the help of the Technical Advisory Panel (TAP) for this project. A list of thirty Wet Table C users, which included at least one representative for each district, was assembled. An online survey with five questions was developed to document the concerns of the Wet Table C users. A copy of the survey presented to participants is provided in Appendix B. A total of 23 Wet Table C users participated in the survey. In some cases, participants did not answer all of the questions, resulting in instances where no response was provided. As a result, the summaries of some questions have information provided by less than 23 respondents.

It should be noted that, due to staff turnover, transfers and other issues, the group of 30 personnel targeted for the survey differed from those previously surveyed by the Caltrans task force. As a result, some of the conclusions reached by the survey presented in this research may differ from those previously obtained. For example, the Caltrans task force survey found that respondents did not believe Wet Table C accurately identified all locations that required safety improvements. However, in the results presented here, respondents indicated that Wet Table C was adequate in identifying only the locations that were in need of safety improvement. This does not indicate that the need to update the wet percent factors was not necessary; rather, it indicates that two different sample populations held different views at different times regarding the Wet Table C process.

2.5.1.1. Survey of Wet Table C Users

The Wet Table C survey consisted of a series of five questions. The first question posed to survey participants sought feedback with respect to the reason(s) that no action might be taken after investigating a "required" location in Wet Table C. A detailed analysis of responses is presented in Appendix B, Note 1. The overall conclusion to be drawn from the responses provided to this question is that problems not related to wet pavement or those which are related to peak hour congestion are the most frequent reasons action is not being taken after investigating a required location in Wet Table C. Each of these selections was cited by the majority of respondents as being those that happened most frequently, as evidenced by their shared position as the top ranked problem. For those who did not select it as the first ranked problem, peak-hour congestion was the clear leader for second most frequent occurrence. In general, it would appear that analysts encounter cases with seasonal traffic peaks or overlapping safety problems from previous years far less frequently, as evidenced by these options occupying the third and fourth ranking slots.

2.5.1.2. Effectiveness of Wet Table C

The second questions respondents were presented asked how well Wet Table C identified only the locations that were in need of improvement based on the experience in investigating “required” locations. A detailed analysis of responses is presented in Appendix B, Note 2. Results indicated that the majority of respondents believed Wet Table C was adequate in identifying only the locations that were in need of safety improvement. Despite responses saying that Wet Table C is not adequate, it can be concluded that overall, the table is meeting the needs of most users in its present form. Note that this finding differs from that of the Caltrans task force; this is the result of differing sample populations holding different views regarding Wet Table C.

2.5.1.3. Provision of One Factor

The third question posed to respondents asked whether they agreed with the use of just this one factor (i.e. percent wet time factor) for an entire county. A detailed analysis of responses is presented in Appendix B, Note 3. Results indicated that nearly half of respondents believed that one percent wet time factor per county was adequate. Nearly one-third of respondents were neutral as to whether such a singular factor should be employed for an entire county. The remaining respondents believed that the use of only one factor per county was not advisable. Based on the responses obtained to this question, it would appear that there are essentially two schools of thought with respect to the issue. There is the school that believes the use of one factor per county is entirely appropriate. Then there is the school (likely including some of those who were neutral to the issue) that believes more than one factor is needed, based on the characteristics of the county itself.

2.5.1.4. Perceived Improvements

The fourth question asked respondents to rank a series of possible improvements to Wet Table C. The improvements to be ranked included:

- Update the current wet percent time table (updating its percentages, which are now 30 years old);
- Modify the wet percent time table to a better geographical unit than a county-based table (i.e. one value for one county);
- Update the wet percent time table every year to better represent the time the pavement was wet;
- None of the above.

A detailed analysis of responses to this question is presented in Appendix B, Note 4. Results indicated that respondents see the need for an update to the wet table percentages. Such an update was the rationale for the research conducted here. There was some question as to how frequently updates should occur however. One side sees a less frequent update as being necessary, while the other sees a yearly update as being the best improvement. In the middle of these two groups fall those who believe that a new geographic unit is most necessary. Overall, when looking at the rankings of the inclusion of a new geographic unit, it is clear that the majority of respondents view this as a necessary improvement.

2.5.1.5. Geographic Unit

The fifth question presented asked what an appropriate geographic unit would be, other than a county. A detailed analysis of responses to this question is presented in Appendix B, Note 5. Of the available options, most respondents ranked the creation of a geographical zone based on precipitation and traffic volumes as being the best available improvement option. Creating percentages based on highway milepost was the second most favored option, followed by the creation of a geographical-based zone using solely precipitation data.

2.5.1.6. Additional Information

In addition to the responses summarized in the preceding sections, respondents were asked if there were any additional concerns or suggestions that they had with respect to the percent wet time being used in the development of Wet Table C. The responses provided are presented in Appendix B, Note 6.

2.5.1.7. Conclusions

In summary, Caltrans personnel who responded to the survey indicated that they felt the current Wet Table C adequately identifies the locations needing improvement. Respondents did view a modification of the wet percent time table to a better geographical unit than county-based as the highest among potential improvements that could be made. Expounding on this, most respondents viewed a geographical zone based on precipitation data alone as the preferred alternative to the current county-based system. The broad conclusion that can be drawn from the survey of Caltrans users of the Wet Table C is that it is adequately meeting their needs, but they recognize that specific updates and improvements would be beneficial.

2.5.2. National Survey on Wet-Pavement Exposure Measurement

A second survey was conducted to obtain background on the current practices different transportation agencies employ in tackling wet-pavement accidents. The survey consisted of twenty questions designed to obtain information regarding each agency's state of the practice with respect to measuring wet-pavement exposure and factoring this into accident analysis. The survey presented to participants is provided at the end of Appendix C, as is a list of the agencies that responded.

2.5.2.1. Reduction of Wet-Pavement Accidents

The first question posed to survey participants asked whether their state was specifically focused on systematically reducing the number of wet-pavement accidents. This would include activities such as wet safety or skid accident reduction programs. Detailed results for this question are presented in Appendix C, Note 1. Results indicated that there was a fairly even split between respondent states who did not have a specific focus on reducing wet-pavement accidents and those who did. Respondent states that said they did have a systematic focus on wet-pavement accident reduction were asked to provide further details of their programs. Activities related to such programs ranged from simply confirming that wet weather crash analysis was conducted to making use of skid test pavement friction data.

2.5.2.2. Other Programs Including Wet-Pavement Accident Analysis

Expanding on the previous question, participants were asked whether their state had any other programs that might include periodic wet-pavement accident analysis. Detailed results for this question are presented in Appendix C, Note 3. Nearly half of the states who responded to this question had no additional analysis programs that identified wet-pavement accidents. Despite this, at least one-third of those who responded to the question did report some program(s) where such accidents were identified. Programs ranged from project-specific identification (3R projects for example), to spot location analysis performed on request, to analysis conducted specifically on winter crashes. In essence, these programs are accomplishing the same goal as specific wet-pavement accident analysis programs; the primary difference is that they fall under a different umbrella that is sometimes project or event specific.

2.5.2.3. Accident Report Classification

The key to identifying wet-pavement accidents is the availability of accurate accident data. To this end, the survey asked participants if their state classified wet-pavement accidents from accident reports (also referred to as police reports). The results of this question are presented in Appendix C, Note 3. Results indicated that most respondent states have some type of pavement condition record in the accident report file collected by the police. Regardless of the response provided, it is clear that the majority of the participant states have at least some knowledge with respect to pavement conditions at the time of an accident available to them.

2.5.2.4. Number of Wet-Pavement Accidents

Next, respondents were asked for their opinion as to whether the state they represented experienced as significant number of wet-weather accidents. The results of this question are presented in Appendix C, Note 4. A fair percentage of respondents believed their state did have a wet-weather accident problem. Those who did not believe their state experienced a problem with wet-weather accidents were typically from states where little precipitation occurs (e.g. the Southwest).

2.5.2.5. Annual Traffic Safety Report

Respondents were asked whether their state’s annual traffic safety report included a list of high concentration wet-pavement accident locations. The results of this question are presented in Appendix C, Note 5. Few respondent states included such a list in their publications. This is understandable, as annual safety reports typically focus on the big picture, rather than events that are often site specific.

2.5.2.6. Additional Reports Containing Wet-Weather Accident Information

Survey respondents were also asked if any additional reports produced by their state contained listings of high concentration wet-pavement accident locations. The results of this question are presented in Appendix C, Note 6. The majority of states who responded did not include information on high wet-weather accident locations in any reports aside from an annual safety report (if they do in fact include such information in that report). This was not surprising given the specific nature of such crashes and the time required to reduce accident data down to such a level of detail.

2.5.2.7. Wet-Pavement Accident Identification Tools

In terms of the tools used to identify wet-weather accidents, respondents were asked what was being utilized by their state. States that responded to this question confirmed that their analysis involved simple database queries. It was interesting that GIS did not play a larger role in the data analysis of the states surveyed. This may stem from the format of the crash data not lending itself to GIS analysis.

2.5.2.8. Measurement of Wet-Pavement Exposure

With respect to wet-pavement exposure, respondents were asked whether their state currently measured such data and made use of it in wet-pavement accident analysis. Only ten states responded to the question, thus the results obtained should be viewed with caution. Based on the limited number of responses obtained to this question, it would appear that states are not collecting wet-pavement information. In terms of measuring wet-pavement exposure, responses from two states indicated that such information was collected, while five states did not collect this information. Remaining states were unsure if such information was collected.

2.5.2.9. Wet-Pavement Exposure Data

In terms of the data used in measuring wet-pavement exposure, respondents were asked if their state used hourly precipitation data, other precipitation data, other weather data than precipitation data, don't know, or data sets other than weather data. Given the low number of responses to the question posed previously, it was not surprising that only four states responded to this question. One state listed that they used data sets other than weather data, namely friction data, in measuring wet-pavement exposure. The method by which such data were used in measuring wet-pavement exposure was not elaborated however. The three remaining respondents did not know what dataset their state used to measure wet-pavement exposure.

2.5.2.10. Sources of Weather Data

Despite the low number of states measuring wet-pavement exposure, it was surprising that no states responded when asked what sources their weather data came from. One possible reason for this is that the safety analysts completing this survey might not have had access to such data for their activities. Background research, making use of documentation such as ITS deployment statistics from the U.S. DOT[1] revealed that at least 32 of the states who participated in the survey have RWIS networks throughout their state. In addition, six other agencies had links to the National Weather Service (NWS) on their websites, suggesting that in the absence of a Road Weather Information System (RWIS) network, they were still obtaining weather information from an outside source. As a result, it would be reasonable to conclude that most of the states who participated in the survey obtained their weather data from NWS or RWIS, despite these sources not being explicitly cited in the survey responses.

2.5.2.11. Methods for Measuring Wet-Pavement Exposure

When asked what methods were used to measure wet-pavement exposure at wet-pavement accident locations, only four states provided responses. One state used a factor representing the wet-pavement time developed on a county basis (i.e., one factor estimating the wet-pavement hours for one county in the jurisdiction). The remaining three respondents were unsure as to how their state measured exposure for a specific location. Given the lack of states that appear to be measuring exposure for wet-weather accidents, the lack of responses to this question was expected.

2.5.2.12. Tools to Measure Wet-Pavement Exposure

Respondents were next asked what tools their state used to measure wet-pavement exposure. Once again, only four states provided responses to the question. One state used a custom tool for measuring exposure, while the remaining three state responders did not know what tool was used.

2.5.2.13. Exposure Estimate Updates

When asked how often their state updated measurements of wet-pavement exposure, two respondents answered yearly, while two were unsure. The failure of other states to reply to this question might stem from the lack of analysis being performed specifically related to wet-weather crashes.

2.5.2.14. Documentation

Four participant states responded when questioned whether they had documentation of the methodology used in determining wet-pavement exposure available. One state did have such documentation available for internal use, but this could not be released to the WTI project team. Two states responded that no such documentation was available, while the final respondent was unsure whether their state had such documentation.

2.5.2.15. Additional Information

In addition to the responses summarized in the preceding sections, respondents were asked if there was any additional information they could provide with respect to their

[1] http://www.itsdeployment.its.dot.gov/SurveyOutline1.asp?SID=swstw

state's practices pertaining to wet-pavement accident analysis. To that end, 18 states provided additional information. The responses provided by states are presented in Appendix C, Note 7.

2.5.2.16. Conclusions

Results of the state survey suggest that a number of states have a focus on reducing wet-weather accidents. Those states that do not have such a focus are those where wet weather is not as common (e.g., the Southwest). Some states that did not have a primary focus on such accidents did have a periodic review as part of a related project (e.g., 3R projects). However, a far greater number of states that did not have a primary focus also lacked a secondary one.

When asked if some form of information regarding pavement conditions (wet, icy, etc.) was recorded in police accident reports, the majority of respondent states confirmed that they did in fact have such a measure. There was an even balance between those respondents who believed there was a wet-weather accident problem in their state and those who did not. Once again, this is likely a function of geographical location.

Only a limited number of states presented any wet-weather accident information in annual safety reports or other departmental publications. Similarly, when identifying wet-weather accidents, respondent states only used simple database queries, rather than more recent applications (such as GIS) to query out information related to accidents.

Few states measured wet-pavement exposure, with only one state specifying the source of its data (friction data). Only one state provided its method for measuring wet-pavement exposure, stating that this figure was developed at the county level of geographic coverage. Similarly, states did not specify or elaborate on what tools they used to measure wet-pavement exposure or how often such updates were made.

3. METHODOLOGY

3.1. Defining Wet Percent Time

Wet percent time refers to the proportion of time during the year that pavement is damp. Wet time is usually measured on an hourly or a daily basis and expressed as a percentage. For example, a wet percent time of 5 percent would indicate that the pavements in a particular county were wet for 5 percent of the entire year. Expressed in terms of hours, this would equal 365 x 24 x .05 = 438 total hours per year of wetness. The measurement of wet-pavement exposure, which is the product of wet time and vehicle-miles traveled, is critical to programs established to reduce wet-pavement accidents.

This research adopted the definition of wet time based on Caltrans' 1972 study [6], i.e., the total number of hours during which a measurable amount (≥0.01 inches or 0.25mm) of rainfall occurred. This total time is subsequently used to calculate the percentage of time per year during which measurable rainfall events occur. The percentage of rainfall-event time per year is calculated as:

$$P_{Rain} = \frac{\text{number of hours with HPD >= 0.01in}}{\text{number of hours in a year}}$$

Where HPD = Hourly Precipitation Data

The value of 0.01 inches (0.25mm) per hour is used as this is viewed as the minimum amount of rainfall necessary to keep a pavement damp for one hour. As such, trace amounts of rainfall were not considered by this research. The previous (1972) Caltrans Wet Percent Time table, however, did not explicitly consider the distinction between frozen and non-frozen precipitation. While California has a considerable amount of desert area, snowfall is not uncommon in the high country and mountains. For example, much of northern California, including the Truckee/Lake Tahoe/Reno/Carson City area, which is generally thought of as a desert region, may receive snowfall any time between early November and late May. Therefore, this project developed two categories of wet percent time table (including snow and excluding snow). The category that includes snow data is compatible with the old wet percent time table, which did not differentiate rain and snow precipitation types. The category that excludes snow data was based on the new methodology, which selected only precipitation data characterizing rainfall, determined by using the average monthly air temperature. Chapter 4 provides details on how precipitation type was determined from the collected precipitation data.

3.2. Data Flow and Manipulations

Figure 3.1 is a data flow chart that illustrates the process employed in developing the wet percent time of each station. The detailed procedures are described in the following text.

First, hourly precipitation data was downloaded from the various available network sources. In total, the most recent 11 years (1995-2005) of hourly precipitation data from all rain gauge stations throughout California were downloaded from five different sources. These sources included the California Data Exchange Center (CDEC), the California Irrigation Management Information System (CIMIS), MESOWEST, the

National Climatic Data Center (NCDC) and the National Weather Service (NWS). Once the archived hourly precipitation data was downloaded from the upstream weather data providers, it was uploaded to a database developed by the research team. This was followed by three processes to check and improve the quality of the data: *Reprocessing*, *Quality Control*, and *Missing Data In-Filling*.

The initial step, reprocessing, involved checks to identify errors in the meta-data, to standardize the data format, and to create a common, standardized database. This was followed by the quality control (QC) step, which consisted of three levels. First, the reprocessed data went through baseline data quality control—recording-error checks and gross-limit checks—to identify obvious data problems. The second-level of quality control included statistical spatial consistency checks, temporal consistency checks, and multi-sources checks. After these steps were completed, the missing data (data missing from raw datasets and questionable data that failed the QC tests) were in-filled using the methodology described in Chapter 6. Visual expert QC and interactive correction with graphical assistance were the third-level QC steps, after which the data was ready for calculating the annual wet percent time for each station. All of the procedures discussed here will be presented in further detail in subsequent chapters.

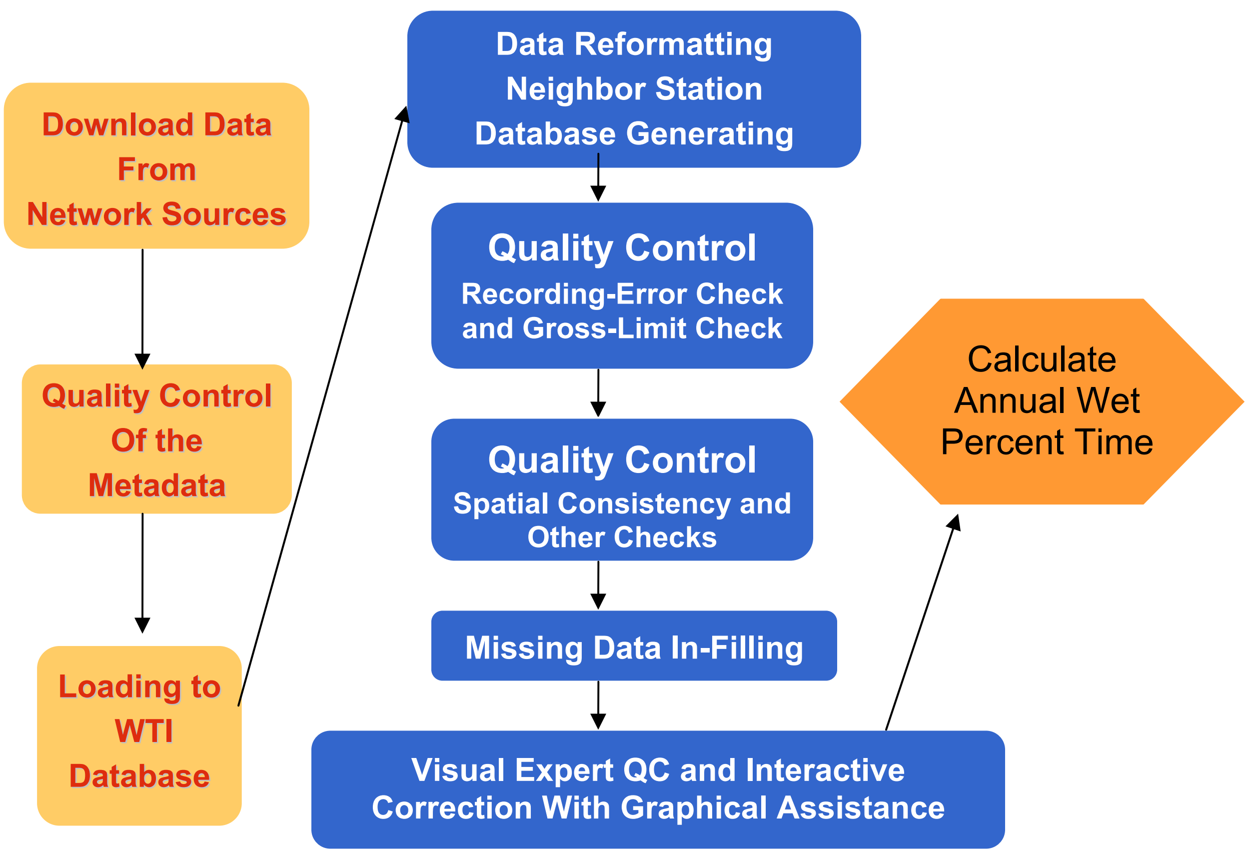

Figure 3.1: Data Flow Chart Illustrating Procedures Used in Developing the Wet Percent Time of Each Station

3.3. Wet Percent Time Calculation

The Caltrans Wet Percent Time table currently provides a single average value for each county in the state of California. These were developed in 1972 through the use of isohyetal lines, using annual precipitation data collected between 1957 and 1967. Given advances in software capabilities and computing power, new methods are now available to calculate wet percent time. The research team investigated these methods through literature review and GIS analysis.

An improved method using the Zonal Statistics provided in ArcGIS's Spatial Analyst was tested and chosen for this project to produce an updated County-Average Wet Percent Time table. Researchers used this method to produce a wet percent time raster map based on the annual average wet percent time that was calculated for each station, as illustrated in Figure 3.2. Such a map was then used to interpolate the wet percent time for every location with known latitude and longitude in California. Based on the raster map and the county boundary shape file, the average wet percent time for the entire county could be calculated through the ArcGIS's Zonal Statistical Analysis feature. This county-average value was developed on the conical system to account for the curvature of the earth. The belief is that the GIS methodology employed was an improvement on past methods and on other methods currently available. Chapter 7 discusses the detailed procedures related to the development of the new county-wide Wet Percent Time table using ArcGIS Zonal Statistical Analysis.

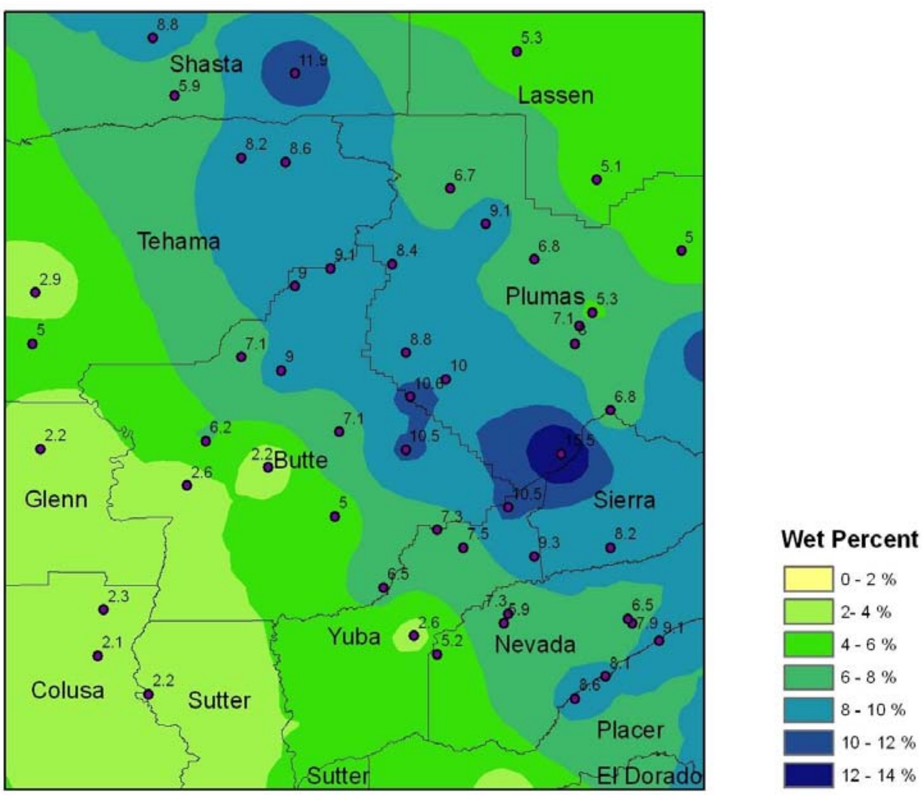

Figure 3.2: Sample Wet Percent Time Raster Map

4. DATA COLLECTION

4.1. Data Acquisition Methods

Historical hourly precipitation data of California reported by rain gauges were obtained from five network data sources, including:

- CDEC: California Data Exchange Center,
- CIMIS: California Irrigation Management Information System
- MESOWEST: University of Utah
- NCDC: National Climatic Data Center
- NWS: National Weather Service.

Historical weather data was acquired from these sources via a number of approaches (completion of data request forms, computer graphical interfaces, FTP downloads, etc.). Data from CIMIS, NCDC and NWS, were downloaded directly through their web interface. The large amount of data available and acquired from CDEC and MESOWEST were obtained by compact disks provided by the vendor and via download through an FTP portal, respectively.

4.2. Data Spatial and Temporal Characteristics

Based on the density and distribution of available precipitation data stations, an 11-year period was chosen for this project, beginning on January 1, 1995 and extending through September 30, 2005. A total of 1,718 weather stations with hourly precipitation data were available from the previously listed sources. All metadata for each station were collected to assist in the subsequent tasks related to data quality control and missing data handling. Table 4.1 shows the number of stations available for each source and the corresponding years of archived data. The spatial distribution of these stations is shown in Figure 4.1. Note that this map was generated using the raw metadata provided by each source and may contain errors in latitude and longitude (as indicated by the points displayed outside of the state's borders).

Table 4.1: Data Network Sources of Hourly Precipitation Data of California

Data Source	Number of Stations	Hourly Precipitation Data	Years of Archived Data	Full Name of Data Source
CDEC	507	YES	1984-2006	California Data Exchange Center
CIMIS	153	YES	1982-2006	California Irrigation Management Information System
MESOWEST	1058	YES	1999-2006	University of Utah
NCDC	397	YES	1993-2006	National Climatic Data Center
NWS	512	YES	1990-2006	National Weather Service

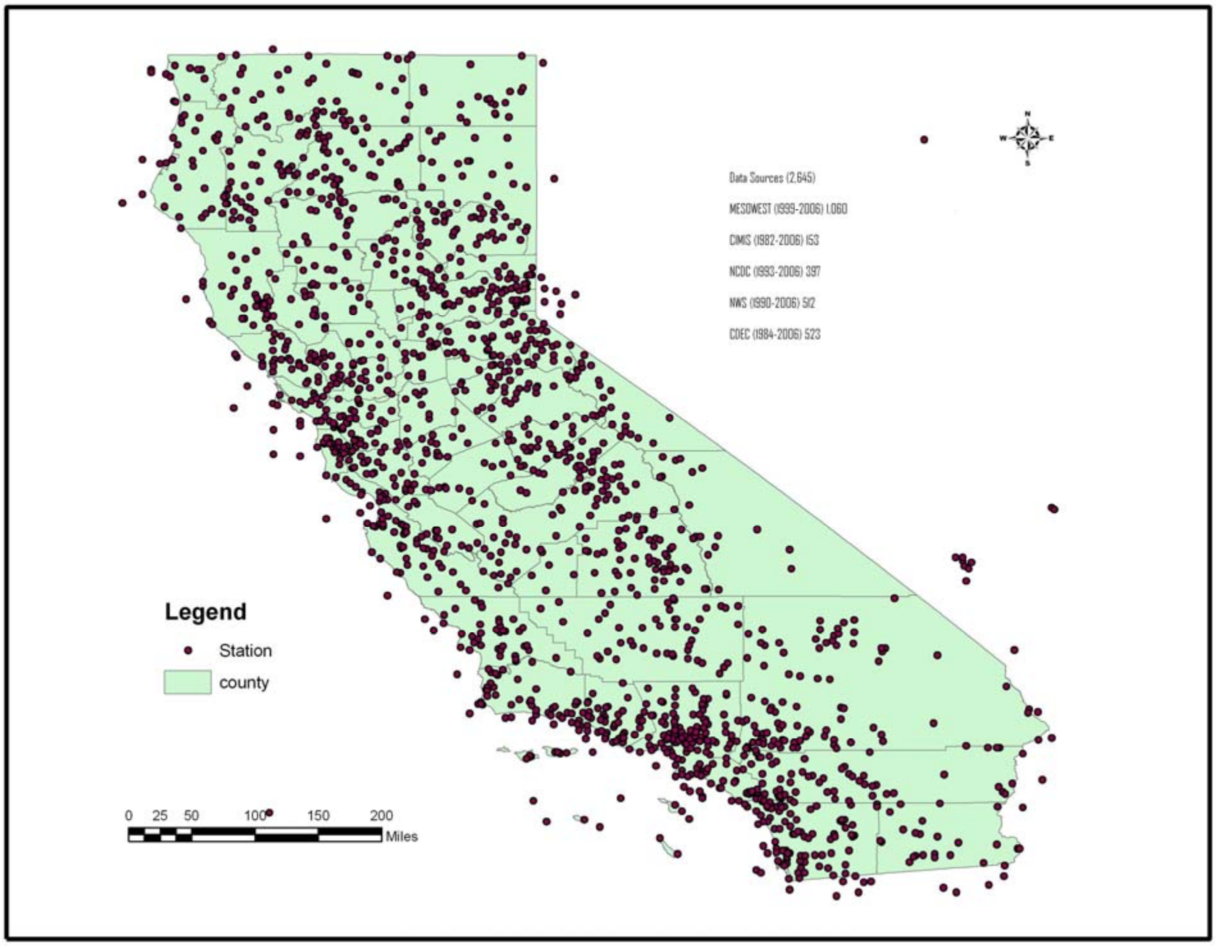

Figure 4.1: Spatial distribution of stations

4.3. Data Description and Sample Data Records

As stated previously, the data collected from the networks included metadata and precipitation data. Metadata descriptions included the station name, county location, station elevation, latitude, longitude, actual start and end recording date, precipitation type and data source (i.e. the vendor). Table 4.2 provides a sample of metadata for four CDEC stations. Following download the precipitation data were reformatted (i.e. standardized) and stored in the WTI database. These data included observations of eight variables, including station ID, Date, Time, precipitation, Precipitation_QC, QC_flag, Wet_flag, buddy, precip2. Table 4.3 shows the sample data set from station C0677 with all the eight variables acquired from the network.

Table 4.2: Sample Meta data of four stations from CDEC

ID	Name	County	Elev	LATITUDE	LONGITUDE	s_date	e_date	P_TYPE	DATASOURCE
STP	STAMPEDE	SIERRA	5956	39.47	-120.1	1989-07-01	2006-10-01	RAIN	CDEC
MNT	MONITOR PASS	ALPINE	8350	38.67	-119.61	1999-06-30	2006-10-01	RAIN	CDEC
HOR	HORSE MEADOW (NRCS)	ALPINE	8557	38.83	-119.88	2005-01-25	2006-10-01	RAIN	CDEC
RKC	ROCK CREEK NEAR MAMMOTH LAKES	MONO	7040	37.55	-118.66	1999-07-19	2006-10-01	RAIN	CDEC

Table 4.3: Station BMW reformatted data from CDEC network

ID	DateTime	Precip	Precip_QC	QC_Flag	Wet_Flag	buddy	precip2
BMW	1999-09-14 05:00:00	*NULL*	37.8	m	*NULL*	*NULL*	*NULL*
BMW	1999-09-14 06:00:00	*NULL*	37.8	m	*NULL*	*NULL*	*NULL*
BMW	1999-09-14 07:00:00	*NULL*	37.8	m	*NULL*	*NULL*	*NULL*
BMW	1999-09-14 08:00:00	37.8	*NULL*	PREC	*NULL*	*NULL*	*NULL*
BMW	1999-09-14 09:00:00	*NULL*	*NULL*	m	*NULL*	*NULL*	*NULL*
BMW	1999-09-14 10:00:00	*NULL*	*NULL*	m	*NULL*	*NULL*	*NULL*

4.4. Data Format

As stated previously, the acquired data acquired from the five different vendors were recorded in different formats, so reformatting data was necessary before quality control was undertaken. All station data were reformatted into an input format developed by the research team. The reformatting included: unifying the data type (i.e. accumulative vs. incremental precipitation), unifying data units (i.e. inches of precipitation to hundredths), unifying time (i.e. time zones), unifying symbols (number vs. string for missing values), and unifying date formats.

Some vendor networks collected accumulative precipitation data (i.e. a running total throughout an event) while others collected incremental data (i.e. accumulation throughout an hour). The researchers unified these different data types into incremental hourly precipitation data since the research interest was in incremental data equal to or greater than 0.01 inch/hour. As one would expect, the units (i.e. depth) of recording were different between vendors. The researchers unified these units to the hundredth inch since 0.01 inch was utilized previously by Caltrans as the cut off value of wet time calculation. The time zone employed between vendors also differed and included Pacific Standard Time (PST) vs. Greenwich Mean Time (GMT); as California is located in PST, this was selected as the database time zone. When missing values occurred, numbers were used to indicate the values that are missing. All character strings were also deleted and replaced by numbers. Finally, all dates were reformatted into a common scheme in the database. A sample data file of each network is provided in Appendix D.

4.4.1. Missing Data Codes

The alphanumeric characters "m", “mm”, "M", “MM”, and "-9999" were coded as the data values to indicate missing data.

4.4.2. Precipitation Data

The "T" and "t" characters were coded as a data value used to indicate a trace of precipitation (physical elements PC, PP, and PY) and were decoded as 0.001 inch. If the decimal point was omitted, it was assumed that the value was observed in hundredths of an inch, with the value divided by one hundred.

4.4.3. Semi-Annual Time Changes

According to the Uniform Time Act of 1966, daylight saving time begins on the last Sunday in April at 0200 (2:00 a.m.) local standard time. That time was amended by Public Law 99-359 (July 8, 1986) to the first Sunday in April, beginning in 1987. Therefore, 0201 local standard time would become 0301 local daylight saving time. Standard time begins on the last Sunday in October at 0200 local daylight saving time. Therefore, 0201 local daylight saving time becomes 0101 local standard time. Downloaded data were adjusted accordingly, with local time becoming Pacific Standard Time.

4.5. Determining Precipitation Type

Precipitation contains 3 major types: rainfall, snowfall and mixture of both. Wet percent time is defined as the percent of wet time (percent of hours when hourly rainfall>=0.01 inch); therefore, this research must distinguish between precipitation data and rainfall data. Rainfall happens when the surface temperature is beyond 3.3 ºC, as described in the literature previously presented in Chapter 2. This research used 3.3 ºC as a threshold value to distinguish precipitation data that was not rainfall. The PRISM (Parameter-elevation Regressions on Independent Slopes Model) monthly minimum temperature data (see Figure 4.2 for an example map of such data) were used to identify rainfall months which met the criteria of having temperatures above 3.3 ºC. When monthly minimum temperature was above this value, the precipitation data were considered as rainfall data; all data below this threshold were considered snowfall or non-rainfall data.

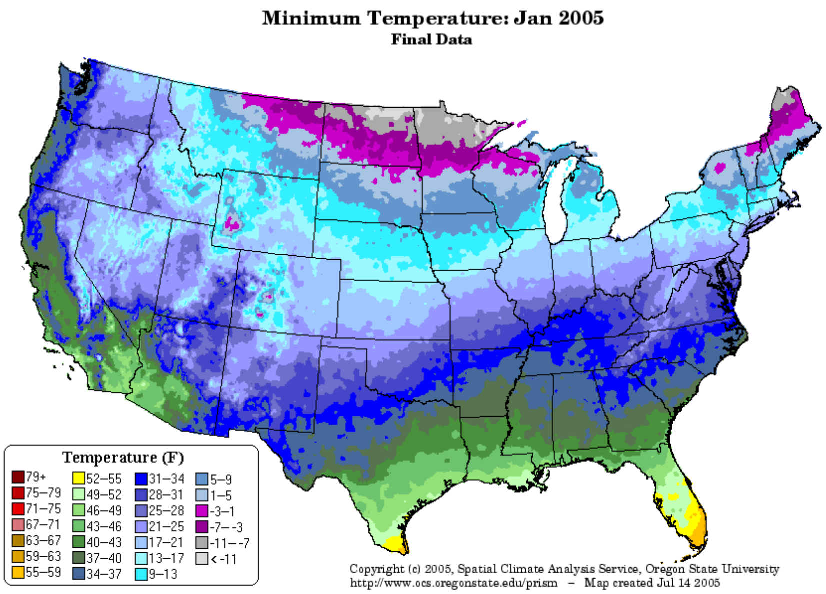

Figure 4.2: Example of Monthly PRISM map
(Source: Spatial Climate Analysis Service – Oregon State University)

4.6. Challenges and Problems

The accuracy of wet percent time estimates are dependent on the accuracy of the measurements of hourly precipitation data collected by the vendors. This presents two issues: first, there is the question of absolute accuracy, i.e., the fact that precipitation gauges are not 100% efficient at collecting precipitation (especially snow). This is likely to induce errors in the wet percent time values to be calculated. Second, the relative accuracy and/or data compatibility between networks (and countries) using different gauges for measuring hourly precipitation amount must be considered. This is an important issue because hourly precipitation data from many networks is being merged into single database and used in updating the wet percent time table update. The factors from this table are being used for a wide variety of purposes, from addressing safety issues to updating long range transportation plans. As a result, any significant errors in the precipitation data employed can have far ranging consequences.

Errors in measurement are generally classified as either random or systematic. The former include instrument failures, observer errors and variations associated with local site or topographic conditions. Standardized observation procedures and site inspections, superior quality control and assurance procedures, and inter-comparison of data from nearby stations help to identify such errors. However, corrections to address such problems are not necessary for the purposes of this research, particularly given that the minimum amount of hourly rainfall required for determining wet time is specified as 0.01 inch. Once this threshold is met, the hour is classified as rainfall no matter what the cumulative amount of precipitation is. In other words, the amount of hourly rainfall does not have to be accurate, eliminating the need for corrections. Identifying when a rainfall event occurred while taking into account these potential errors remain a concern to this research. As a result, the following chapter will examine quality control procedures employed to address these errors and ensure the completeness or the developed database.

5. DATA REPROCESSING AND QUALITY ASSURANCE

Given the dense network of weather stations with available data for the state of California, a large data set was developed. This consisted of more than 2000 stations identified by the researchers that provided over 11 years of data. Quality control of these data sets is an important procedure that must be undertaken before the data can be used for any application. The absence of quality control procedures may result in data with poor quality (outliers, missing records) or noise (inconsistent or unrepresentative records) that severely limit their usefulness for the research pursued here. Quality control is essential to identifying and flagging potentially erroneous data, and allowing issues to be addressed to produce a complete and representative data set. Data preprocessing and quality control procedures are described in the following sections of this chapter.

5.1. Data Preprocessing

Data preprocessing activities include metadata checking, data reformatting and neighborhood database generation.

5.1.1. Metadata Checking

Metadata include information specific to a station including the station name, station id, latitude, longitude, elevation, and recording date. Errors in the metadata of a station are common and could result in serious problems in a combined data set. For instance, stations that are incorrectly located due to latitude and/or longitude errors will affect results drawn from the data for the entire period covered, both in terms of the site location itself and when performing comparisons of data between sites. Therefore, a check for consistency of variables pertaining to each station is necessary to ensure the accuracy and quality of precipitation data. An example of such a consistency check was the flagging of metadata where a discrepancy of more than about 20 km horizontally or 100 m vertically occurred. Such flagged data were subsequently checked manually for accuracy.

An example of this is presented in Table 5.1. In this instance, the metadata for three sites sharing a similar location are compared. Note that while the three sites share a similar latitude and longitude, the elevation of Station MUD was significantly higher compared to the other two sites. As a result, this metadata record was flagged for further manual checks

Table 5.1: Example of metadata error (elevation)

Station ID	BGB	MUD	BGBC1
River Basin	Trinity	Trinity R	Trinity R
Hydrologic Area	North Coast	North Coast	North Coast
Latitude	40.7330N	40.7170N	40.7433N
Longitude	123.200W	123.283W	123.250W
Elevation	1270’ft	3400’ft	1500’ft
County	Trinity	Trinity	Trinity
Operator	NWS	CA DWR	MESOWEST

5.1.2. Data Reformatting

Various data sources use different formats and time zones to record their data. As a result, the raw data collected by the researchers required reprocessing and reformatting as follows. Date and time were converted into Pacific Standard Time. Raw data were decoded for two precipitation elements (PC—Precipitation Accumulated, and PP—Incremental Precipitation). Sub-hourly PP data were subsequently converted to hourly PP data. Following the conversions, all hourly precipitation data and its metadata were recorded in a central database. An example of reformatted data ready for recording in the central database is presented in Table 5.2.

Table 5.2: Sample reformatted data set for station C0677

ID	DateTime	Precip	Precip_QC	QC_Flag	Wet_Flag	buddy	precip2
C0677	2005-08-25 21:50:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-25 22:00:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-25 22:20:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-25 22:30:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-25 22:50:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-25 23:00:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-25 23:20:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-25 23:30:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*

C0677	2005-08-25 23:50:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*
C0677	2005-08-26 00:00:00	0	*NULL*	P24I	*NULL*	*NULL*	*NULL*

5.1.3. Neighborhood Database Generation

To facilitate inter-station comparisons, a table of neighboring stations was developed. This table identified stations that were within 15 miles of one another and whose elevations were similar. The table was generated to support future data processing where the need arose to quickly extract a file from a neighboring target station when data were being quality-controlled, and for missing data infill activities (discussed in Chapter 6).

5.2. Data Quality Control

As mentioned previously, the quality issues of hourly precipitation data stem from the fact that data were collected from a variety of independent sources that use different sensors and measurement techniques. Observational data such as precipitation are also subject to systematic and random errors. Such data tend to be more sensitive to measurement errors because factors of influence like the catch efficiency of a rain gauge, are extremely sensitive to environmental conditions [37, 38]. In addition, transmission errors from remote stations may also produce errors. Some of the data used in this research were collected at such locations, with data transmission performed electronically through several ports (both from the field stations and for the acquisition activities by the researchers). Finally, some stations receive infrequent maintenance, resulting in data which are less reliable due to equipment breakdowns.

A quality control diagram is presented in Figure 5.1, illustrating the steps taken for quality control of hourly precipitation data. The process consisted of a combination of manual and automatic quality assurance processes, and was employed in the processing of all the collected data sets. Summary statistics of data from each station were generated, allowing for identification of problematic sites. Based on the quality control process, 22 stations from CDEC, 6 stations from CIMIS, and 45 stations from MESOWEST were removed from the central database.

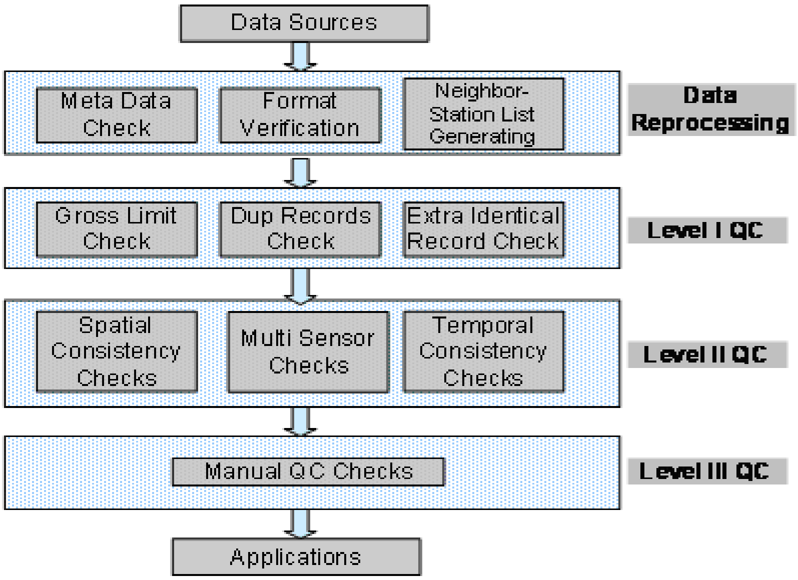

Figure 5.1: The Quality Control process

The quality control diagram data shows the three levels of quality control techniques that were used to examine the collected data. A more detailed discussion of these processes is presented in the following sections.

5.2.1. Level-I QC

These quality control processes may be considered "gross error checks." They were the most preliminary checks that could be performed on individual observations. If an observation failed the preliminary checks, it may be eliminated from the database or was manually corrected, depending on conditions and researcher judgment.

Duplicate record check: Multiple rows of identical records in which all variables shared the same values in the dataset were considered to be duplicates. In such cases, only one record was kept. Sample data showing duplicate records with the same time and date are shown in Table 5.3.

Table 5.3: Example of duplicate record

ID	DateTime	Precip	Precip_QC	QC_Flag	Wet_Flag	buddy	precip2
ASC	2000-01-02 06:00:00	*NULL*	*NULL*	Q1	*NULL*	*NULL*	*NULL*
ASC	2000-01-02 07:00:00	*NULL*	*NULL*	Q1	*NULL*	*NULL*	*NULL*
ASC	2000-01-02 07:00:00	*NULL*	*NULL*	Q1	*NULL*	*NULL*	*NULL*
ASC	2000-01-02 08:00:00	*NULL*	*NULL*	Q1	*NULL*	*NULL*	*NULL*
ASC	2000-01-02 09:00:00	*NULL*	*NULL*	Q1	*NULL*	*NULL*	*NULL*

Gross range error check: According to the literature reviewed in Chapter 2, it is reasonable to expect that hourly precipitation totals should be greater than or equal to zero but less than or equal to 5 inches per hour. As a result, any precipitation in the database that had a negative value or unusually large values were flagged and checked further, unless the negative values occurred during the resetting time[2]. Note that due to evaporative losses, at some times hourly rainfall total will be a negative number (-1 in/hr < 0 in/hr). In such cases, the researchers manually checked these values and determined if the flagging was legitimate. If the negative value was greater than -1, it was flagged as questionable data. Extreme observations above or below these thresholds were extracted for further quality control.

Extra identical values check: If an instrument malfunctioned, it is likely that the data would be identical for a long time period (with the exception of zero values over dry periods). Non-zero values that were identical for over 48 hours were flagged and checked. Neighboring station data were used to provide the spatial comparison for the station in question during the time period of interest.

5.2.2. Level-II QC

Spatial consistency check (Buddy check): The Spatial Consistency Check was used to identify outliers that were not spatially consistent with the neighboring gauges [39]. This process is also referred to as the "buddy check." Its steps are as follows. First, a list of neighboring stations within 0.25 degrees latitude and longitude was generated. Then the data from neighboring stations was compared using this database. Hourly precipitation data were considered outliers if they were greater than 2.2 times the standard deviation of all the values from the neighboring stations for that hour. In such cases, the value was flagged. If the hourly precipitation totals at one station indicated a heavy rainfall (>0.1 in/hr) during a short time period when the other neighbor values were all zero, the record was also flagged for further checking. Finally, hourly precipitation totals of zero for a long time period compared to neighboring values reporting heavy rainfall were flagged.

Additionally, monthly/daily rainfall time series around a suspect station were checked and compared to the time series of its neighborhood stations. If the data from a station were erratic or inconsistent with the data from another nearby station, the erratic data were also removed from the dataset. However, the data from its nearby stations were kept in the database. Each annual maximum value was checked by comparing their values with those from the nearby stations (unless the gauge was under the influence of an intense thunderstorm). Note that if a gauge was under the influence of an intense storm, it did not have to be spatially consistent with its neighbors. Such gauges identified as

---

[2] Some field measuring devices accumulate precipitation cumulatively during the year. Generally, this sensor type is used for real-time collection duration of hourly or event data. These stations' accumulation tanks periodically dump the accumulated precipitation to make room for more precipitation. This may cause the value transmitted to jump backward several inches. As a result, the cumulative value usually gets larger until it is reset. A reset may occur if a technician visits the site or it is near the beginning of the season. The dates that designate a season vary according to different agencies (i.e., July-June, October-September).

outliers from the spatial consistency check were compared against lightning data. If there was at least one lightning strike within a 10km radius of the gauge during the past one hour in question, then the data from that gauge were considered valid.

Wet hour frequency data spatial consistency check: As this research was interested in the number of measurable rainfall hours (>=0.01 in/hr) occurring during each year for each station, the time series of monthly/yearly wet hour frequency of all nearby stations were compared to detect any spatial inconsistency. Previous literature indicated that precipitation frequency should vary less with horizontal and vertical distance and exhibit greater spatial coherence than total amounts [40].

Figure 5.2 provides an example of a spatial consistency check for nearby stations. The time series plot of monthly total rainfall of four nearby stations is used to determine any significant discrepancy among all these stations' data. The monthly total rainfall of nearby stations was calculated with the hourly total rainfall data for the month from April through October of 2005. Station PR3 displayed an unusually higher trend for April and May than neighboring stations. As a result, the data from these two months at station PR3 were flagged and checked by other quality control methods.

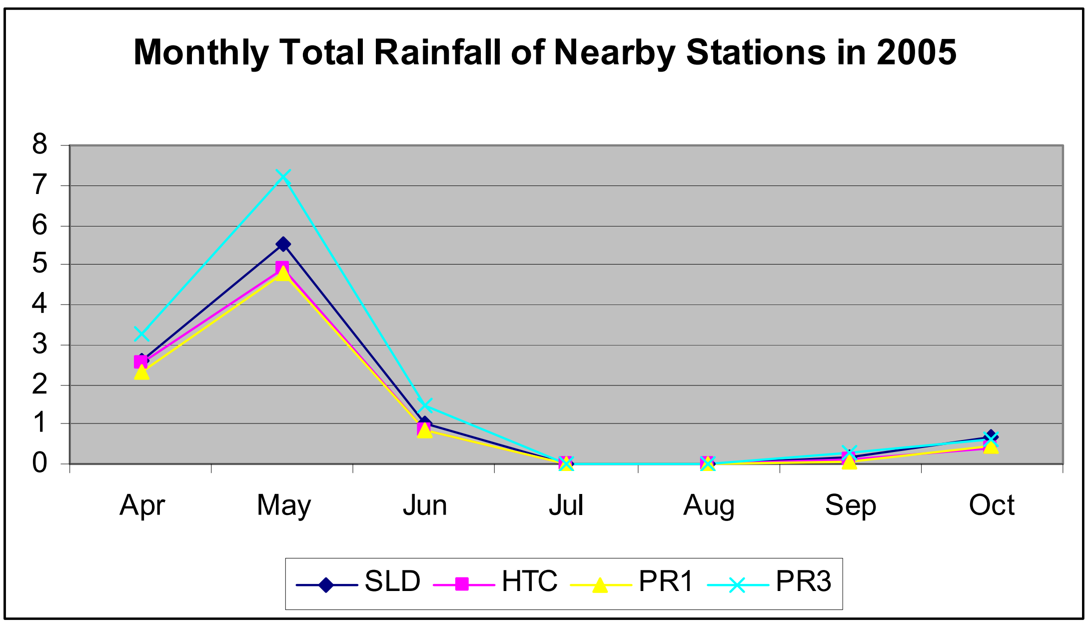

Figure 5.2: Time series plot of summer monthly total precipitation

Next, the monthly wet percent data of these nearby stations from April to October were compared, with no significantly different trends being found. This comparison is presented in Figure 5.3.

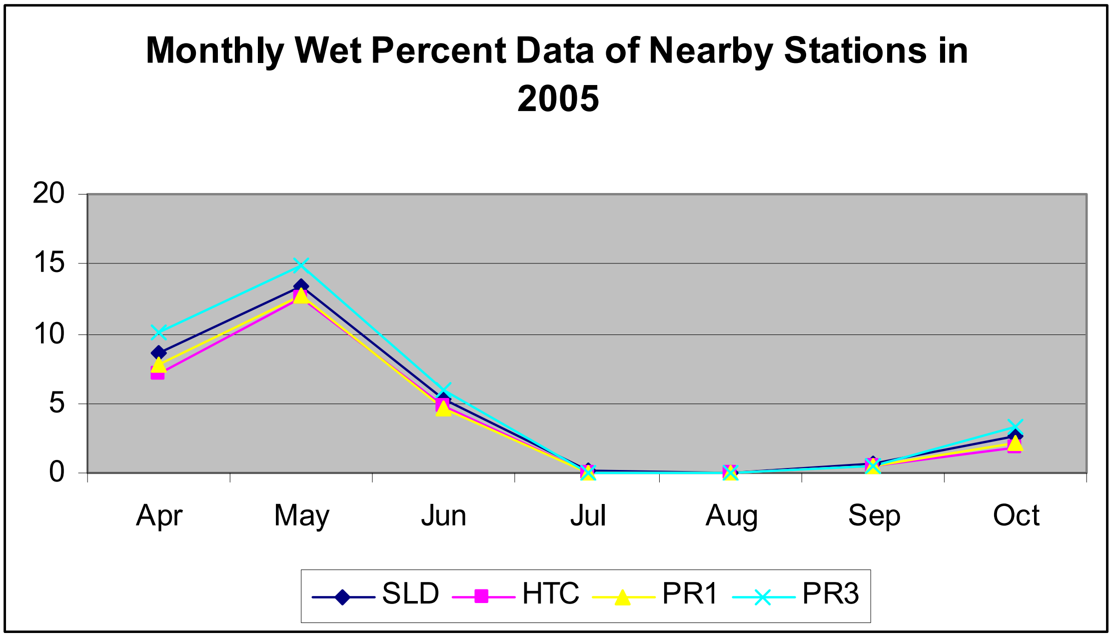

Figure 5.3: Time series plot of summer monthly wet percent variations

The monthly wet percent data that were flagged by the spatial quality check required manual checks and subsequent decisions to determine if these data were valid. Manual checks were performed by consulting extreme-event records, which identified weather events that may have contributed to the readings in question. As a result of the spatial consistency check, 65 stations were flagged as questionable. These stations' data were sent for manual checking and were excluded if they failed that procedure.

Temporal consistency check: The temporal consistency check was designed to detect problematic data or stations by looking at continuous station measurements over time. This check was computationally more intensive than the spatial consistency check. The researchers performed the temporal consistency check only if it was deemed necessary. The process used a 5-year, 31-day window to calculate the mean and standard deviation of the daily precipitation total. This total should fall within the limits of the mean of daily total precipitation by +/- 3 times the standard deviation. The 5-year, 31-day moving window represented a good compromise between including enough days to produce a stable mean and standard deviation, while not including so many as to dilute seasonal and inter-annual trends in spatial climate patterns.

Multiple sensors check: The data from different sources (e.g., vendors) were compared to check for discrepancies. Using daily precipitation total data or monthly average precipitation data, possible errors were identified and flagged, with the extreme-event record consulted to identify and confirm the errors. Since values from other observed variables like radiation are closely related negatively to rainfall, hourly radiation data could also be used to identify possible errors.

5.2.3. Level-III QC

Manual quality control check: Data flagged as incorrect or questionable were checked manually by two processes. These included visual expert quality control and interactive correction with graphical assistance (comparisons with climate maps, extreme-event catalogs). Most suspected errors required manual verification and correction.

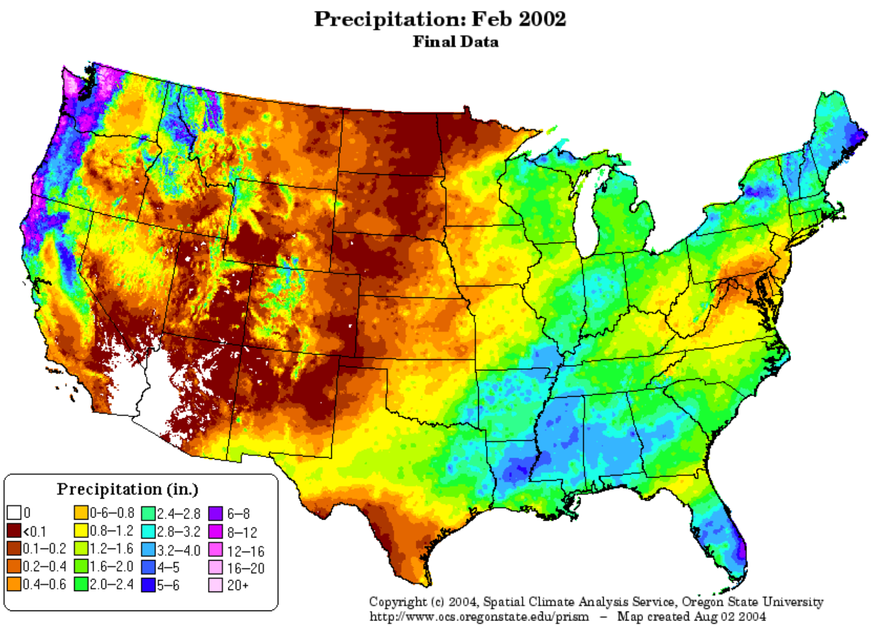

Figure 5.4: Example graphical data to support interactive correction

5.3. Challenges and Problems

Given the quantity of data comprising the central data set, some compromises had to be made for the sake of processing time. As the element of interest in this research was hours that received precipitation of at least 0.01 inch/hour and not the actual cumulative amount of precipitation, the data set was not corrected for systematic gauge-measuring errors. Additionally, because rainfall is highly variable in space and time, traditional QC methods were not entirely appropriate for the data used in this research and were revised accordingly to meet the needs of the project. Finally, it was not feasible to develop a fully automated quality control procedure; therefore, manual review of questionable data was utilized. Examination of large data sets required many weeks, and may have introduced some inconsistencies due to the requirement of human visual inspection. When possible errors were found, they were individually checked and verified. However, it is possible that introducing a human component to this process may have resulted in the misidentification or miscorrection of some errors.

6. MISSING DATA HANDLING

Once the overall quality of the data set had been checked, gaps where data had been eliminated during the QC process or simply had never been collected due to short duration station failures were addressed. This was done through the Missing Data Handling process. This process is described in the following sections.

6.1. Missing Data Handling

A complete dataset of hourly precipitation is crucial to determining wet percent time, as well as subsequently measuring vehicle exposure to wet pavements. However, due to a variety of issues, the data collected and processed for this work was not entirely complete. To address the issue of completeness, techniques were employed to replace missing data.

6.1.1. Sources of Missing Data

There are basically three sources of missing data: a significant amount of missing data is mainly from weather stations that were not operational during certain time periods. Also, errors in data due to observational reporting and recording, equipment faults, and values outside the tolerance limits identified through QC procedures result in missing data. Finally, data may be reported intermittently by a station resulting in missing values.

6.1.2. Revision of Missing Data

A large amount of data was found to be missing from stations in cases where the values for the periods immediately before and after the missing data period remained the same. If such a period was 48 hours or less in length, it was reasonable to assume the value would have remained the same. In such cases, the missing data were simply replaced with the data observed prior to the gap occurring. This kind of revision was not employed when the period of consecutive missing values was over 48 hours in length. This was done to avoid confusion with a possible stuck rain gauge. When employed, revision reduced the number of occurrences of missing values among each station, as illustrated in Figure 6.1.

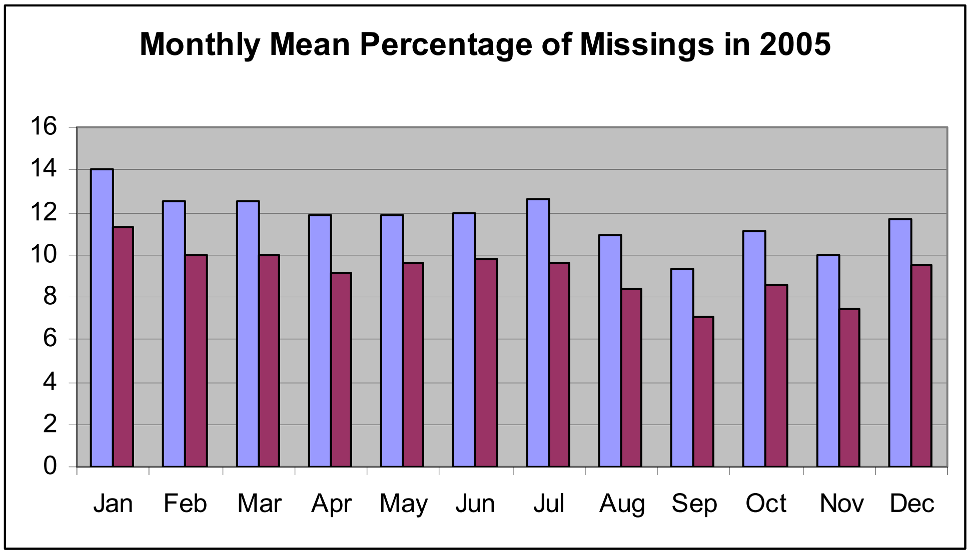

Figure 6.1: Mean percentage of missing CDEC data before and after revision.
Note: Lighter bars stand for the percent of missing data from CDEC archive, while dark bars represent missing data following infill.

6.1.3. Traditional Methods for Infilling Missing Data

Methods for estimating missing data reported in previous literature varied, depending upon the number of variables present in a respective data set. For multivariate cases, single/multiple imputation, maximum likelihood estimate, or nonparametric approximation, lag-1 Markov Process procedures were applicable. For uni-variate cases, list-wise deletion, nearest neighbor regression, ARIMA models, and neural network model were available as applicable procedures. These methods have their own advantages and disadvantages, but they are mostly applied to data that are normally distributed and spatially and temporally stationary.

Hourly Precipitation Total Data (HPTD) is highly spatial and temporally uncorrelated, as well as discontinuous. The discrete nature of precipitation in time and space has always posed unique problems for meteorologists and climatologists compared to more continuous variables such as temperature and pressure. Therefore, most traditional approaches of infill are not suitable for hourly precipitation data. As a result, the estimation of hourly precipitation missing data has always been a challenge, but has rarely been reported or addressed in literature.

6.1.3.1. Infilling Missing HPFD vs. Infilling Missing HPTD

For this research, Hourly Precipitation Frequency Data (HPFD) derived from the HPTD was used to calculate the annual wet percent time. Recall that wet percent time was defined as the percentage of hours of measurable rainfall greater than or equal to 0.01 inches during each year. In the central data set, an indicator variable named wet_flag was created to indicate when hourly rainfall total fell at or above 0.01 inch. If the hourly rainfall total equaled or exceeded 0.01 inch, the variable wet_flag took on the value of 1;

otherwise it was assigned a 0. If the raw hourly precipitation total data were missing, the wet_flag value was missing as well. As a result, the handling of missing hourly total precipitation data is equivalent to the handling of missing hourly precipitation frequency data, which exhibits greater spatial coherence than total precipitation data [40]. To address this, Nearest Neighbor Frequency Assignment (NNFA) was selected to infill missing hourly precipitation frequency data in this project, as this method was frequently cited in literature as being appropriate in such cases.

6.1.3.2. NNFA Method to Infill Missing HPFD

Nearest Neighbor Assignment (NNA) method was frequently cited in literature for the estimation of missing short-time climate data. This method takes full advantage of neighboring observations to fill gaps in data. Each station with missing data is assigned the observed values of the nearest neighboring station(s) that had data available for the hour. However, most of the NNA methods employed are for the non-precipitation data, or for the monthly total or yearly total precipitation data.

Pesonen et al. (1998) compared NNA with three other methods, substituting means, random values and neural network, to determine the most effective strategy for replacing missing data [41]. Their results indicated that NNA performed as well as the neural network in the study. Additionally, Toth et al. (2000) developed and compared the accuracy of short-term rainfall forecasts using NNA, artificial neural networks and auto-regressive moving average (ARM) models [42]. Results indicated that NNA performed better than the ARM model, but neural network performed slightly better than NNA.

6.1.4. NNFA Method

The Nearest Neighbor Frequency Assignment (NNFA) method is based on the previously discussed Nearest Neighbor Assignment method and has been shown to be suitable for replacing missing hourly precipitation frequency data. It was deemed the best applicable method for this research as it is intuitive and shown to be accurate. Its data needs are met by the large network of stations available in this research being sufficiently dense, with the nearest neighbors close enough to share most similar meteorological features that govern rainfall. Finally, this method has been shown to handle large volumes of data in a timely fashion. The rules employed by the researchers when applying the NNFA method are as follows:

1. If missing hours fall inside a consecutive heavy rainfall (HPD>=1.0in/hr), infill wet_flag = 1 for the missing hours;

2. If missing hours fall inside consecutive rainfall (0<HPD<1.0in/hr) and period lasts for more than 6 hours, infill wet_flag = 1.

Otherwise,

3. If only one hour is missing, use the majority vote method, where the majority of nearby stations report;

4. If several hours (<=24 hr) are missing, use the average hourly rainfall frequency counts from the same time period from nearby stations;

5. If several days (<=7 days) are missing, use the average hourly rainfall frequency counts from the same days from nearby stations;

6. If more than one week but less than one month is missing, use the average hourly rainfall frequency counts from the same month from nearby stations;

7. If more than one month is missing, use the average hourly rainfall frequency counts for each missing month from nearby stations.

6.2. NNFA Method Validation/Performance Evaluation

Data sets were extracted from the central database to test the validity and applicability of the NNFA method in different hydrological regions. This was accomplished by creating a sufficient number of gaps (missing data) of various durations with the intent to simulate raw missing frequency data recovery. The NNFA method was adjusted to infill the frequency of wet-hour (frequency of hours of rainfall>=0.01 inch) by using data from neighboring stations instead of infilling the total amount of rainfall when the data are missing. The majority vote method was also implemented as a complementary aspect of the NNFA method. Results indicated that the NNFA method worked better when infilling the frequency of rainfall for missing data compared to infilling the actual amount of rainfall received. Therefore, the frequency infilling NNFA approach was chosen to generate the final results.

The NNFA method was next employed on a small, experimental portion of the precipitation dataset. Test results demonstrated the feasibility and reliability of the NNFA method, allowing it to be applied more comprehensively to infilling operations on the central database.

In the next section, we will use simulated missing precipitation frequency data from a randomly chosen station that in reality had complete data available during a specific month. The NNFA method will be applied to this simulated data set to accomplish infilling, with an error percentage index calculated and time series plots developed to check for data discrepancies.

6.3. Example of Sample Test Data for NNFA Method Validation

To test the validity of the NNFA method, a station named ATW in Tulare County was selected from the CDEC database, with missing data simulated for a given period of time. May of 2005 was selected as the month for which data would be simulated as missing. A total of 384 hours of missing data were simulated, of which 30 hours were originally rainfall hours (>=0.01 in/hr). A neighboring station to ATW named GNF was located 10 miles away and had complete data for the same month. The two respective stations are depicted in Figure 6.2, with their specific characteristics presented in Table 6.1. The NNFA method was next used to infill the missing data of ATW, with the requisite summary statistics calculated for before and after the missing data were infilled. An error percentage index was also computed, with the results indicating that the NNFA method worked well for infilling the missing hourly precipitation data. These results are presented in Table 6.2.

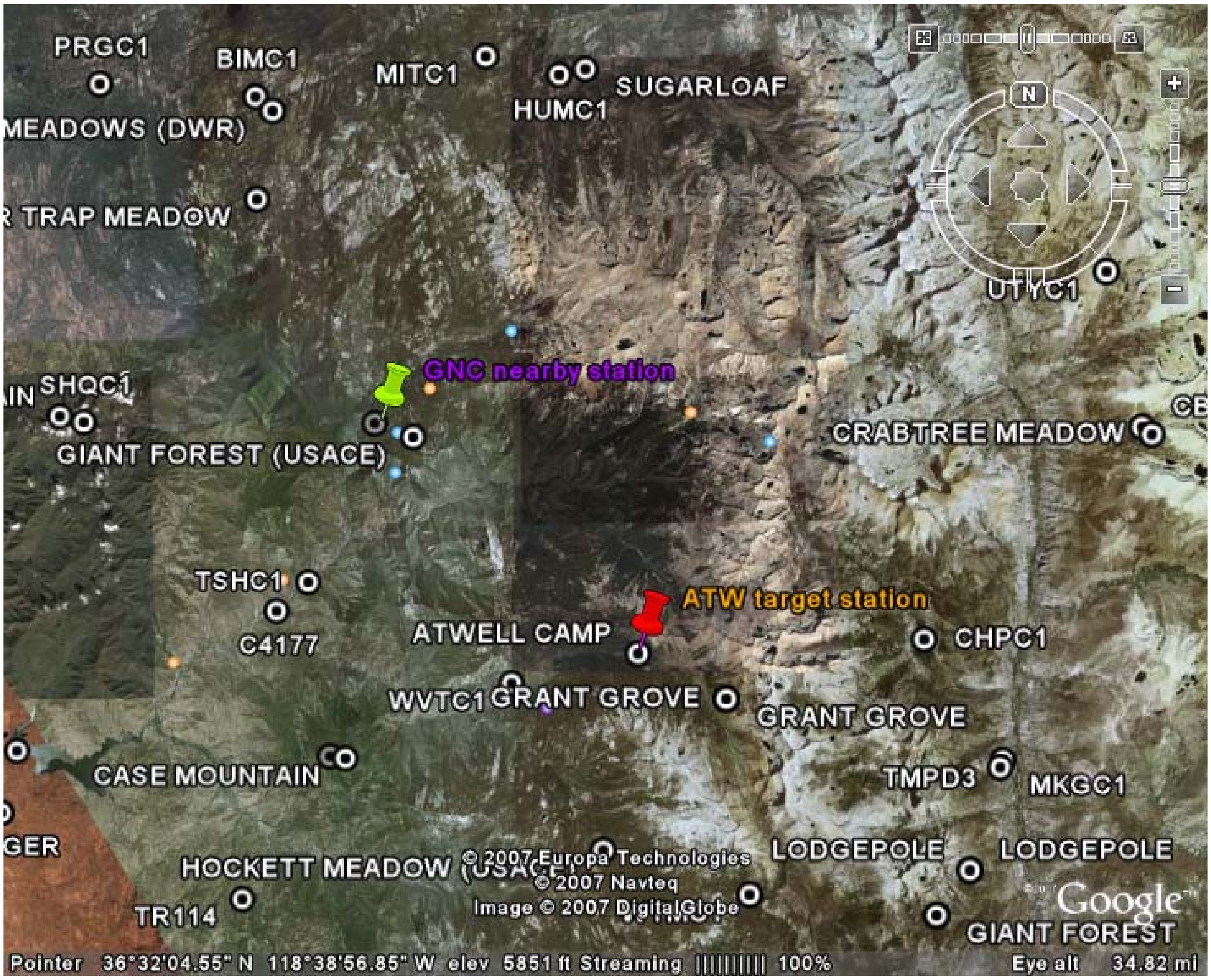

Figure 6.2: Location of station GNC with respect to station ATW

Table 6.1: Characteristics of stations ATW and GNC

Station ID	**ATW**	**GNF**
River Basin	KAWEAH R	KAWEAH R
Hydrologic Area	TULARE LAKE	TULARE LAKE
Latitude	36.4640°N	36.5620°N
Longitude	118.6310°W	118.7650°W
Elevation	6400' ft	6650' ft
County	TULARE	TULARE
Network	CDEC	CDEC

Table 6.2: Statistical results of NNFA infill

Summary Statistics (Before infill)

Station ID	Total hours	Total missing	Total missing rainfall hours	Total month
ATW	744	384	30	1
GNF	744	0	0	1

Summary Statistics (After infill)

Station ID	Total hours	Total missing	Total missing rainfall hours	Total month
ATW	744	384	24	1
GNF	744	0	0	1

Statistical Summary of Daily Wet Percent (31 days)

Daily Wet Percent (hr)	ATW	GNF
Avg.	5.1	6.11
S.D.	8.06	9.13

Correlation Coefficient of Daily Total and Daily Wet Percent

Between 2 stations	Distance	Correlation Coefficient (Daily Total)	Correlation Coefficient (Daily Wet Percent)
ATW and GNF	10 mi	0.9813	0.9918

In evaluating the performance of NNFA in the simulated example, a comparison between the infilled prediction and the actual observation can be used to assess the accuracy of the method. As the results presented in Table 6.2 indicate, the NNFA method effectively infilled 24 of the possible 30 rainfall hours, along with 360 non-rainfall hours of the possible 384. Taking these figures, the error percentage of infilling was calculated as (30-24)/384=1.6%.

Figures 6.3 and 6.4 illustrate the daily wet percentage time series and the daily total time series in May of 2005 for the two neighboring stations after the infill of missing hours.

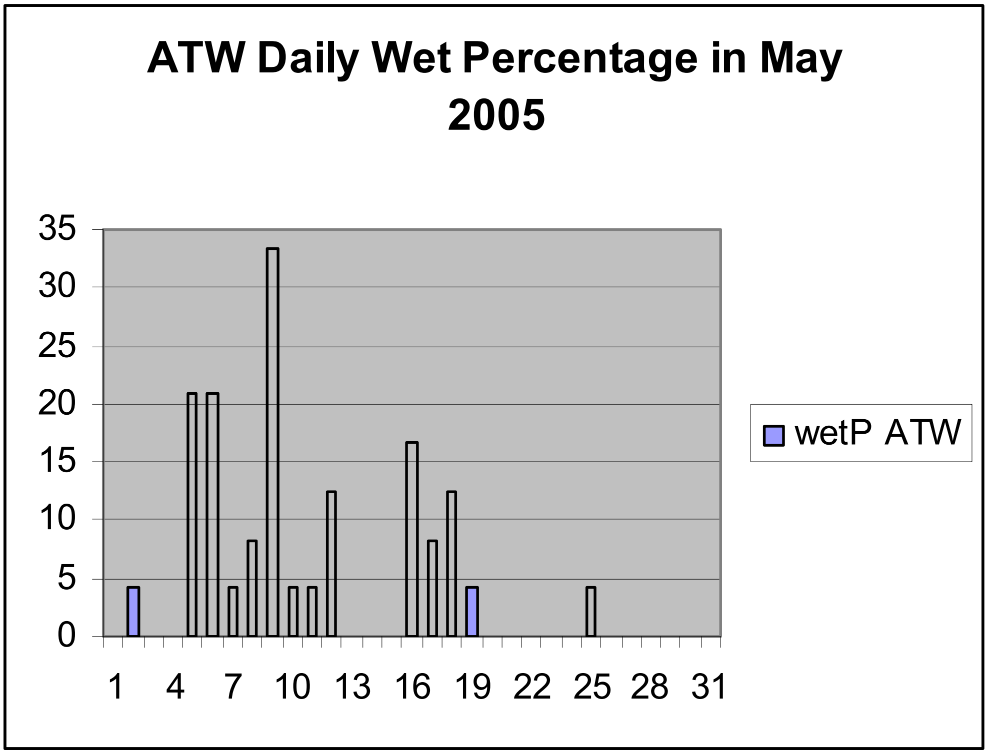

Figure 6.3: Daily wet percentage time series, station ATW

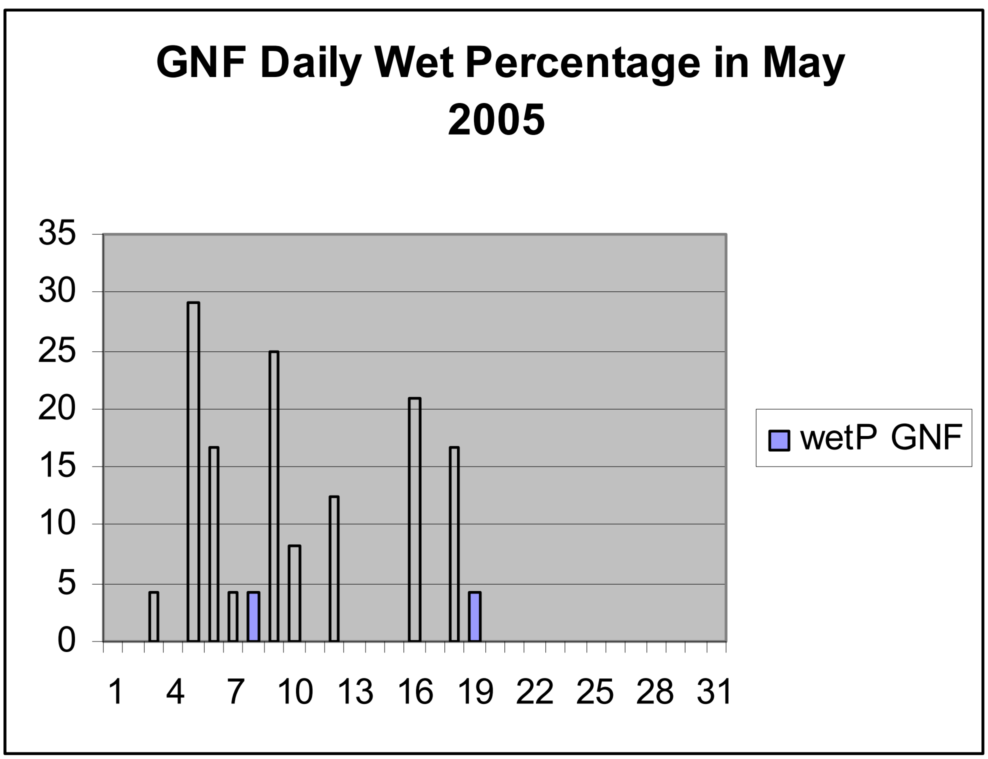

Figure 6.4: Daily wet percentage time series, station GNF

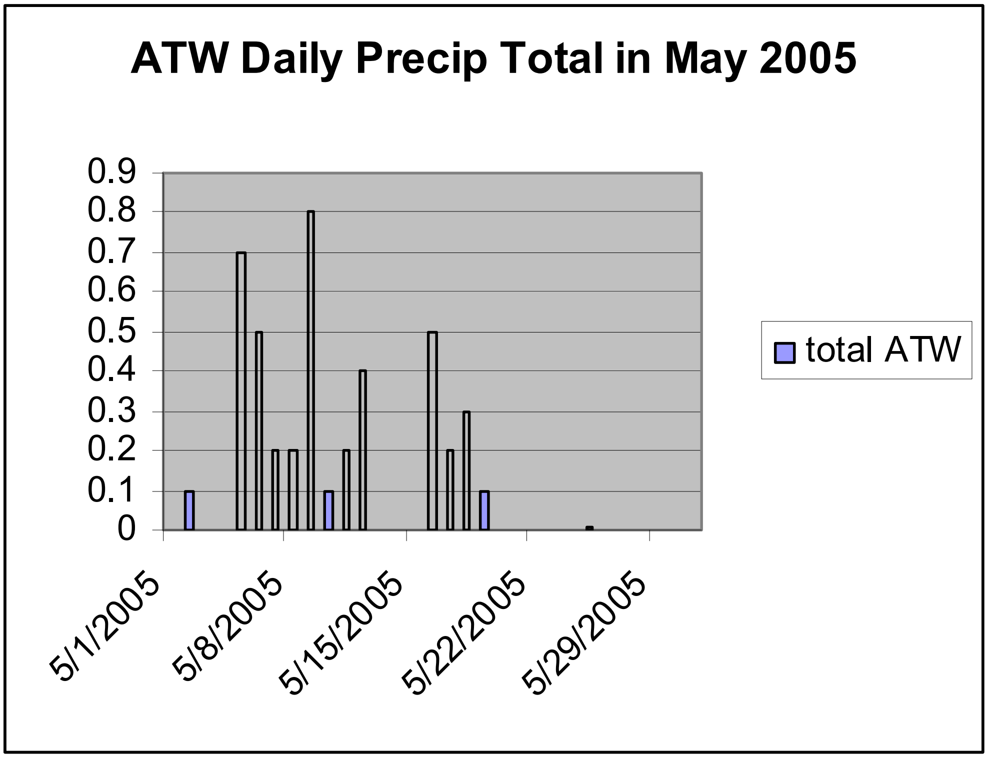

Figure 6.5: Daily total time series, station ATW

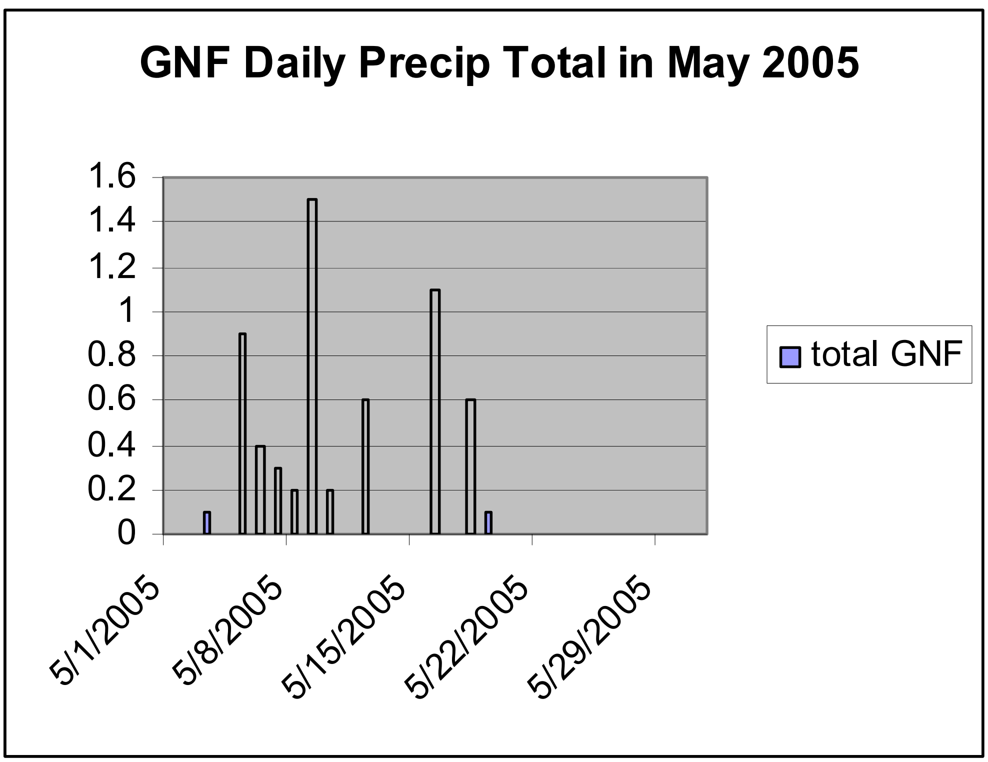

Figure 6.6: Daily total time series, station GNF

6.4. Statistical Summary after Data Manipulations

Overall, 422 of the 1,718 stations available for constructing the wet percent table from 1995 to 2005 were removed from our database. This left 1,296 stations where complete data was available. The eliminated data sets were removed for a number of reasons, including a short duration of available data for revision and infilling procedures, as well as large percentages of errors in data for over 15 percent of missing data existing in the data sets.

6.5. Challenges and Problems

In western California, where mountainous terrain and cold ocean airflows produce sharply different precipitation and temperature conditions, neighboring stations may not have been the best suited for infilling one another. For example, precipitation at a station on the windward side of a mountain range is likely to be more highly correlated with another station on the windward side, compared to a station that is closer but on the leeward side.

In the modified infilling scheme applied in this research, in addition to distance, a second component was used in weighting stations: a historical regressions relationship between stations. That is, a station was given higher weight if it exhibited a historically strong data relationship with the station to be in-filled. However, this strategy only works if both stations have valid data that overlap in time for at least five years; otherwise, the regression function becomes unstable and cannot be relied upon to accurately estimate historical relationships. To determine the extent of relationships between stations, assessments were made separately for each month of the year.

7. UPDATING WET PERCENT FACTOR TABLE

The sections presented in this chapter detail the procedures employed for updating the countywide average wet percent time table. These procedures involved three basic steps, which are discussed here in detail.

7.1. Calculate Individual Station Wet Percent Time

The first step of the wet percent table development process was to calculate the wet percent time for each station for each individual year. The results of these calculations were stored in a summary table similar to that presented in Table 7.1. The wet percent time results for all CIMIS, CDEC and MESOWEST stations that were missing data for less than 15 percent of the total hours for each year were calculated using quality-controlled data available from 1995 to 2005. The summary table also included station coordinate data (latitude, longitude) and elevation data. A quality control process was followed to remove outliers of the wet percent time for each station. The stations located at elevations higher than 6000 feet with wet percent time greater than 25 percent were removed to reduce bias in the calculation of the countywide average wet percent time.

Table 7.1: Example Summary Table of Wet Percent Time by Station

ID	Lat.	Long.	Elev. (feet)	Wet Percent Time											
				Avg	05	04	03	02	01	00	99	98	97	96	95
183	36.49	-117.92	3684	0.3	0.5	0.2	0.3								
189	36.36	-117.95	3682	0.4	0.5	0.3									
117	34.478	-117.261	2890	0.5	0.8	0.6	0.3	0.3	0.2	0.2	0.5		0.3		1.2
186	32.493	-114.826	48	0.6		0.8	0.3								
91	41.959	-121.471	4035	0.7	0.7		0.4			0.6		1.1	0.7	0.6	0.7
BUT	32.74	-114.88	320	0.7	0.8	0.9	0.5								
SQK	32.9	-114.49	300	0.7	0.7	1.0	0.6		0.7						
CAU	32.97	-115.17	278	0.8	0.8	1.1	0.5								
172	35.734	-119.749	225	0.8				0.4	1.2						
RIC	34.06	-114.7	820	0.9		1.0	0.6		0.7	0.5	0.8	1.3	0.9	0.8	1.1
151	33.532	-114.634	251	0.9	1.0	1.2			0.8		0.5				
BPT	38.27	-119.28	6650	0.9	0.9	1.0	0.7								
135	33.557	-114.666	275	0.9	1.0	1.1	0.9		1.0		0.5	1.0			
FIS	32.98	-116.05	760	0.9	1.1	1.1			0.7			0.9	1.0	0.6	1.3

PIC	32.95	-114.73	840	0.9										0.6	1.2
88	34.932	-119.605	2290	0.9	0.5		0.9					0.5		1.7	1.1
175	33.389	-114.726	230	1.0	0.7	1.4	1.1		0.7						
7	36.851	-120.59	185	1.0	1.1	0.9	0.7	0.7	1.2	1.2	0.8	1.6			
192	35.26	-117.22	5148	1.1	1.1										
MOJ	35.05	-116.08	950	1.1	1.1	1.1	1.0		1.1		0.8	1.5	0.8		

7.2. Interpolate Raster Grid

Once summary tables of wet percent time by station and year were developed, an interpolation of wet percent time for an entire surface of grid cells was performed. This was based on a limited number of sample points (i.e., weather stations within a given area).

The first step was to develop a model that provided insight into the relationship between wet percent time factors, latitude and longitude coordinates, and elevation. A multiple linear regression model was applied to the summary table data to check the possibility of interpreting wet percent time based on latitude, longitude and elevation. However, this model did not reveal any patterns that indicated a linear relationship between those factors.

Next, the research team explored current literature related to the use of spatial interpolation in the interpretation of precipitation data. The assumption underlying spatial interpolation is that points closer together in space are more likely to have similar values than points more distant. The literature review revealed several different interpolation methods (Natural Neighbor, IDW, Spline, and Kriging) were available to employ through ArcGIS. Kriging is a stochastic method that has been widely used for estimation of hourly precipitation [43, 44, 45]. In this research project, the Ordinary Kriging method was selected to generate the wet percent time raster grid using ArcGIS.

7.3. Generate Countywide Average Wet Percent Time

The classical arithmetic average method fails to quantify the accuracy of the estimated countywide average wet percent time in this project. The weather stations with hourly precipitation data used in this research were unevenly distributed throughout the state of California. As a consequence, the number of stations located in every county was not consistent, and the area covered by each station varied. If simple averaging of the wet percent time for all stations within the county were employed to generate the new countywide average, this would put less weight on the stations that cover larger areas, producing biased results. The wet percent time table developed in 1972 used isohyetal lines produced by the annual average wet percent time from 1957 to 1967, shown in Figure 7.1. An improvement over this method was developed using ArcGIS. The "zonal statistics" function of the Spatial Analyst extension to ESRI's ArcMap 9.2 program was used to calculate the countywide average, which averaged the wet percent time value at each pixel of the wet percent raster grid contained within each county.

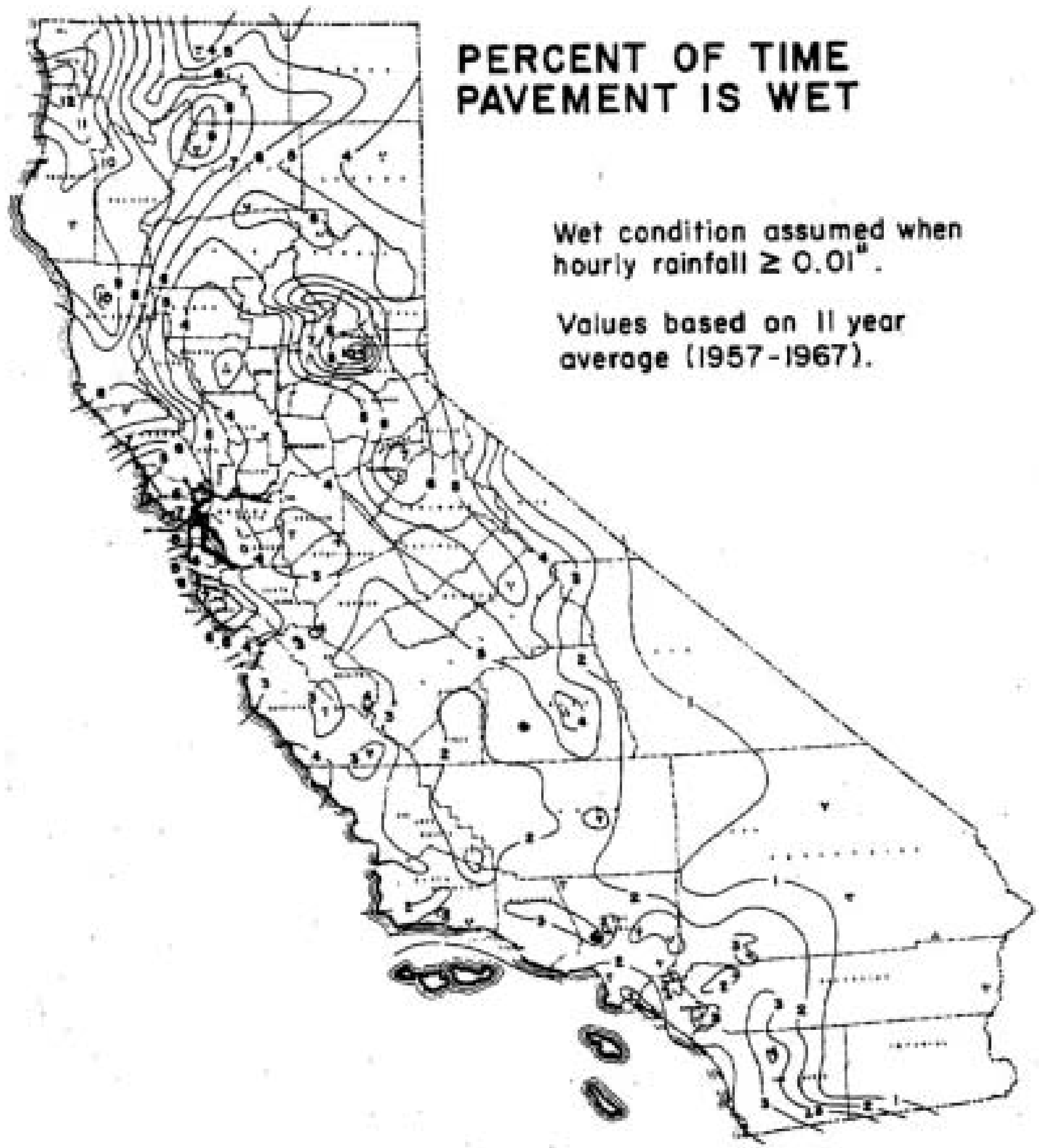

Figure 7.1: Isohyetal Lines Developed in 1972

Based on this procedure, two sets of county-based wet percent time data were generated for the updated wet percent time table. These results are presented in Table 7.2 and Figure 7.2. Table 7.2 presents the updated wet percent time factors, as well as the original factors from 1972. As this table indicates, the 1972 factor and the new factor developed (including snow) are reasonably similar. This would suggest that, although changes have occurred in terms of precipitation received by county, no radical changes have occurred over time. The additional factor provided in Table 7.2, "excluding snow," is a calculation of what the countywide factor would be if only precipitation in the form of rainfall were taken into account. This new type of factor may prove to be useful in areas where rain and snow precipitation vary greatly throughout the year and thus present inherent problems for accuracy and have differing impacts on safety. Such areas include those of Northern California.

Table 7.2: Updated Wet Percent Time Table

County Name	Old Wet %	Include Snow	Exclude Snow	County Name	Old Wet %	Include Snow	Exclude Snow
Alameda	4	4	4	Orange	2	2	2
Alpine	5	6	3	Placer	6	6	2
Amador	5	5	4	Plumas	6	6	1
Butte	5	5	3	Riverside	2	2	2
Calaveras	6	5	2	Sacramento	4	4	3
Colusa	3	3	2	San Benito	3	3	3
Contra Costa	4	4	4	San Bernardino	2	2	1
Del Norte	11	9	5	San Diego	3	3	3
El Dorado	6	7	3	San Francisco	5	5	5
Fresno	3	3	1	San Joaquin	3	2	2
Glenn	4	3	2	San Luis Obispo	3	3	3
Humboldt	10	8	5	San Mateo	5	5	5
Imperial	1	1	1	Santa Barbara	2	3	3
Inyo	1	2	1	Santa Clara	4	4	3
Kern	2	2	1	Santa Cruz	6	6	6
Kings	2	2	1	Shasta	7	8	2
Lake	6	6	3	Sierra	7	8	3
Lassen	4	5	1	Siskiyou	7	7	2

Los Angeles	2	3	2
Madera	4	4	2
Marin	6	5	5
Mariposa	5	5	2
Mendocino	8	7	3
Merced	3	2	2
Modoc	5	5	2
Mono	2	3	2
Monterey	3	3	3
Napa	5	6	4
Nevada	8	8	3

Solano	4	6	5
Sonoma	7	6	5
Stanislaus	3	2	2
Sutter	4	4	3
Tehama	6	6	3
Trinity	9	8	2
Tulare	3	3	1
Tuolumne	5	5	1
Ventura	3	3	3
Yolo	4	4	3
Yuba	6	6	3

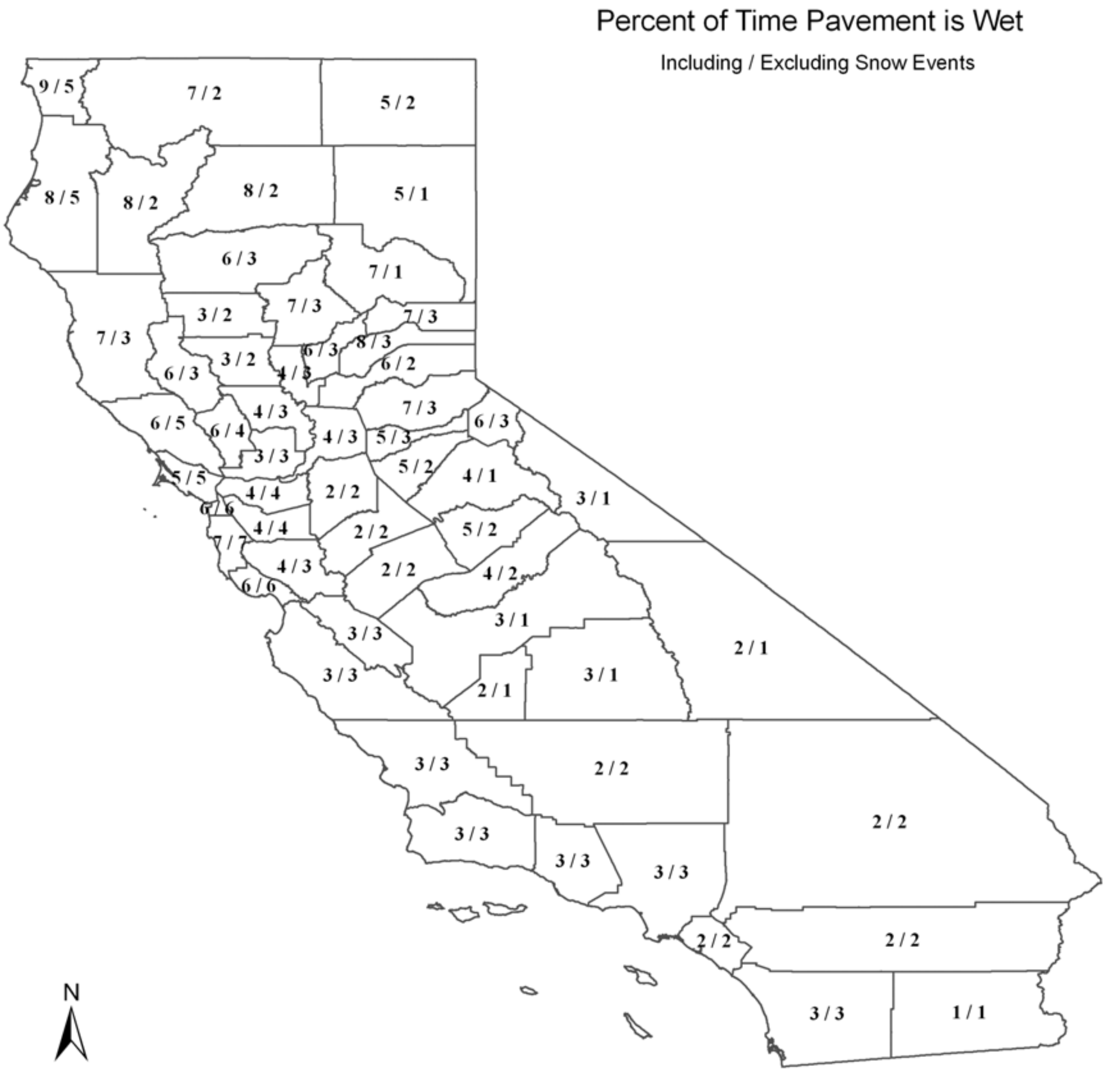

Figure 7.2: Updated Wet Percent Time Map for California

8. ANNUAL UPDATING

Due to climate change, it is reasonable to anticipate that in the future, the amount of precipitation an area receives may vary significantly from year to year. In such a case, an average percentage wet time may not represent the current conditions for any given year at specific locations. Therefore, it would be better to update the wet percent time table annually. This chapter examines the feasibility of developing a framework that can be implemented along with TASAS and/or future Caltrans systems to update the wet percentage time annually using the data from selected weather stations.

8.1. MYSQL/Linux/Python Open Source Platform

This research involved managing and processing large amounts of hourly precipitation data. As a consequence, a high-performance database system was required for data management, data quality control and missing data handling. A MYSQL/LINUX/Python Open Source platform was developed during the project to meet these needs. It was based on Debian LINUX 4.0 running in VMWARE and hosted on an Intel Pentium D 3 GHz Desktop. This platform also provided a phpMyAdmin control panel interface to handle the administration of MySQL database over the Web. All programming codes for data import and processing were programmed using Python 2.4. This open sourced platform provides more flexibility and can be easily ported to other computers.

Table 8.1: Data Platform Used in This Project

Item	Value
Host Hardware	Intel ® Pentium D 3 GHz-2GB MB - 250GB ATA drive
Database	MYSQL 4.0.24
Database Management System (DBMS)	MyAdminPHP control panel
Database Server Operating System	Debian Linux(2.4.27) running on VMware 5.5.2
Web Server Operating System	Apache 1.3
Data Process Language	Python 2.4
Precipitation Data Records/Size	133,248,069 / 7 GB

As shown in Figure 8.1, the current architecture is based on the client/server model, which provides usability and flexibility. It contains a centralized MySQL database server for storing all the precipitation data and station metadata. The database can be remotely accessed and processed by client workstations through a phpMyAdmin web interface and Python programs using MySQL ODBC Driver. For data safety, a firewall was set up on the Linux server and routing data backup was performed every week.

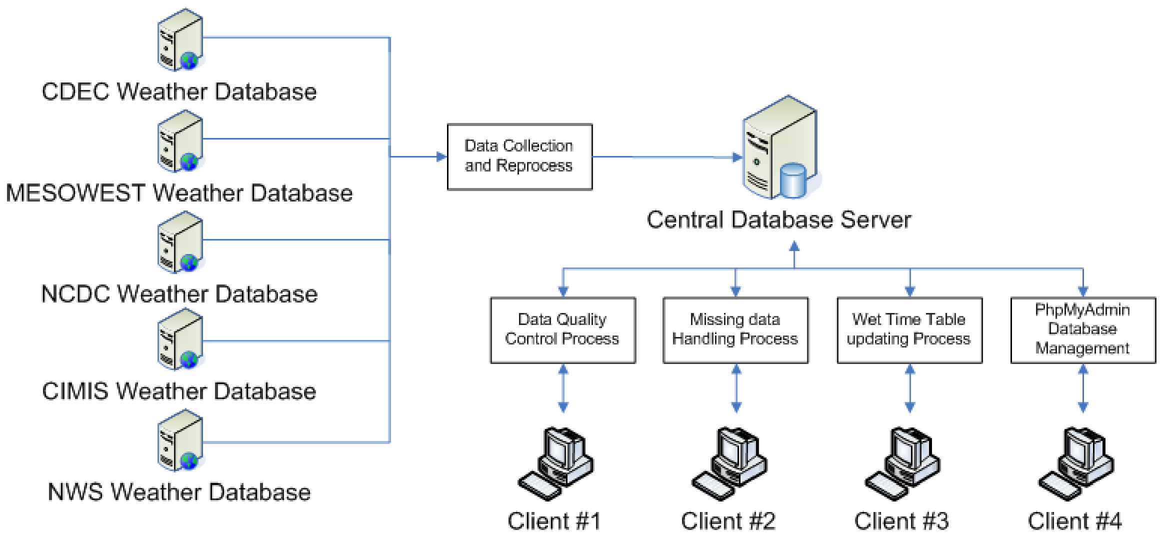

Figure 8.1: Data Flow Diagram with Client/Server Architecture

Based on the current platform and following the data flow in updating the wet percent time table, the following sections will examine the feasibility of annually updating the wet percent time in an automated manner.

8.2. Data Collection and Reprocess

For annual updating, all new hourly precipitation data for an entire year from selected stations within California will need to be downloaded. This process can be automated by coding. The present research utilized five network data sources: the California Data Exchange Center (CDEC), the California Irrigation Management Information System (CIMIS), MESOWEST, the National Climatic Data Center (NCDC), and the National Weather Service (NWS). The CDEC, CIMIS and MESOWEST provide both real-time and archived historical hourly precipitation figures, while the NCDC and NWS provide only historical hourly precipitation data for California. All of these data can be downloaded through the respective source's web site download interface.

Once downloaded, all hourly precipitation raw data, which are recorded in different data and time zone formats (GMT, PST or PDT), need go through meta-data checking and data reformatting before their inclusion in the central database. Since all the data sources and data formats are fixed, this process can be automated by software.

8.3. Data Quality Control

The data quality control process presents a challenge in terms of the programming required for automatic updating. As described in Chapter 4, three levels of quality control steps need to be followed to ensure the quality and completeness of the collected hourly precipitation data. Currently, the Level 1 quality control step can be handled by codes with predefined logic. However, for the Level 2 and Level 3 quality control steps, manual processing and human intervention are necessary. These steps involve visual expert quality control checks and interactive correction with graphical assistance to compare the spatial and temporal distribution of wet percent time with climatic maps, orographical data and extreme-event logs.

8.4. Missing Data Handling Process

Another challenge to automated updating is presented by the process of missing data handling. As discussed in previous chapters, there are several reasons for missing data. Significant amounts of hourly precipitation data may be missing from weather stations that were not operational during certain time periods during a year. Additionally, data involving observation errors, reporting and recording errors, or questionable values identified through Level 1 and Level 2 quality control procedures need to be treated as missing data. Other instances of missing data occur when values are reported intermittently for some stations or sensors. The missing data handling process involves revision of such data, as well as employing the Nearest Neighbor Frequency Assignment method (NNFA) to in-fill missing hourly precipitation frequency data (refer to Chapter 5 for details). In this research project, all of these different types of missing data were handled on a case-by-case basis and in-filled either by codes or by hand, depending on the circumstances. As a result, it may not be entirely feasible to automate this entire process, as some human intervention is often required.

8.5. Wet Time Table Updating

Once the hourly precipitation data have gone through the quality control and missing data handling processes, the wet percent time for each station for each year can be calculated. These calculations are subsequently stored in a summary table, which includes station name, station ID, station latitude, longitude, elevation and wet percent time data. The resulting table can be accessed by ArcGIS through the MySQL ODBC Driver to produce wet percent time raster maps. Such maps may then be used to interpolate the wet percent time for the entire surface area of California. Based on the raster maps and the county boundary shapefile, the average wet percent time for the entire county can be calculated through ArcGIS's Zonal Statistical Analysis feature. In this project, all processes utilizing ArcGIS were handled manually; however, these can be programmed in Python through ArcGIS's Python interface to facilitate automated updating.

8.6. Conclusions

From the earlier discussion, it is evident that there are some steps in the process of updating the wet percent time table that could be automated through the use of computer programs. These include activities such as data collection and reprocessing. However, the central tasks of data quality control and missing data handling primarily involve human intervention, which is time consuming. At present, it is not possible to develop an automatic data-quality-control algorithm to handle these critical steps by code. In addition to these technological challenges, there are also other issues that must be taken into account, like budgeting for annual updates, the hosting and personnel for managing the database system, and so forth. All of these components require further investigation before a conclusion can be drawn regarding the practicality and utility of an automated updating process.

9. FINER RESOLUTION FACTORS

9.1. Introduction

Previous chapters have discussed the development of updated wet percent time factors, which will subsequently be used in producing future Caltrans Wet Table C lists. These updated figures are expected to be an improvement over those currently in use, whose development was based on data over 35 years old. However, the new wet percent figures still consist of only one number per county. This single number fails to take into account California's varying geography, which can produce different microclimates within a county. The use of one wet percent time factor in a county may lead to a situation where wet-accident locations in an area are falsely being identified as significant based on a countywide wet percent factor whose calculation was skewed by these different microclimates.

In light of the potential for singular countywide wet percent time factors to identify what could be termed "false positive" significant wet-accident locations, the research examined alternative factors for developing Wet Table C lists. The most logical approach was to examine the use of more specific wet percent time factors throughout a county, at a localized level. Given the wealth of weather station data acquired for the update of wet percent time factors, it was possible to develop localized factors which correspond to highway segments.

The approach taken for this portion of the research was to develop new Wet Table C lists using localized wet percent time factors for highway segments along three routes. These lists were then compared to lists developed for the same routes using the singular 1972 and 2008 countywide factors. In addition, a comparison between the lists developed using the singular 1972 and 2008 factors was also undertaken.

The research question which seeks to be answered through the generation of these lists is this: Is there a statistically significant difference in the Wet Table C lists developed using a countywide factor versus lists which employ more localized factors? In other words, does the use of a localized factor produce a list containing different "significant" sites from those identified when one factor was used for an entire county? If this is proven to be true (i.e. the lists are significantly different) this would warrant a reexamination of the use of a single, countywide factor in the development of Wet Table C lists.

9.2. Study Routes

Consultation with Caltrans district staff identified three routes for evaluation at a finer geographic resolution: Interstate 5 in Tehama, Shasta, and Siskiyou Counties, U.S. Highway 395 in Lassen County, and California 299 in Shasta County. Each of these routes passes through unique geographic regions/terrains that see differing levels of precipitation, both throughout the year and cumulatively. The routes, which are all located in northern California, are presented in Figure 9.1.

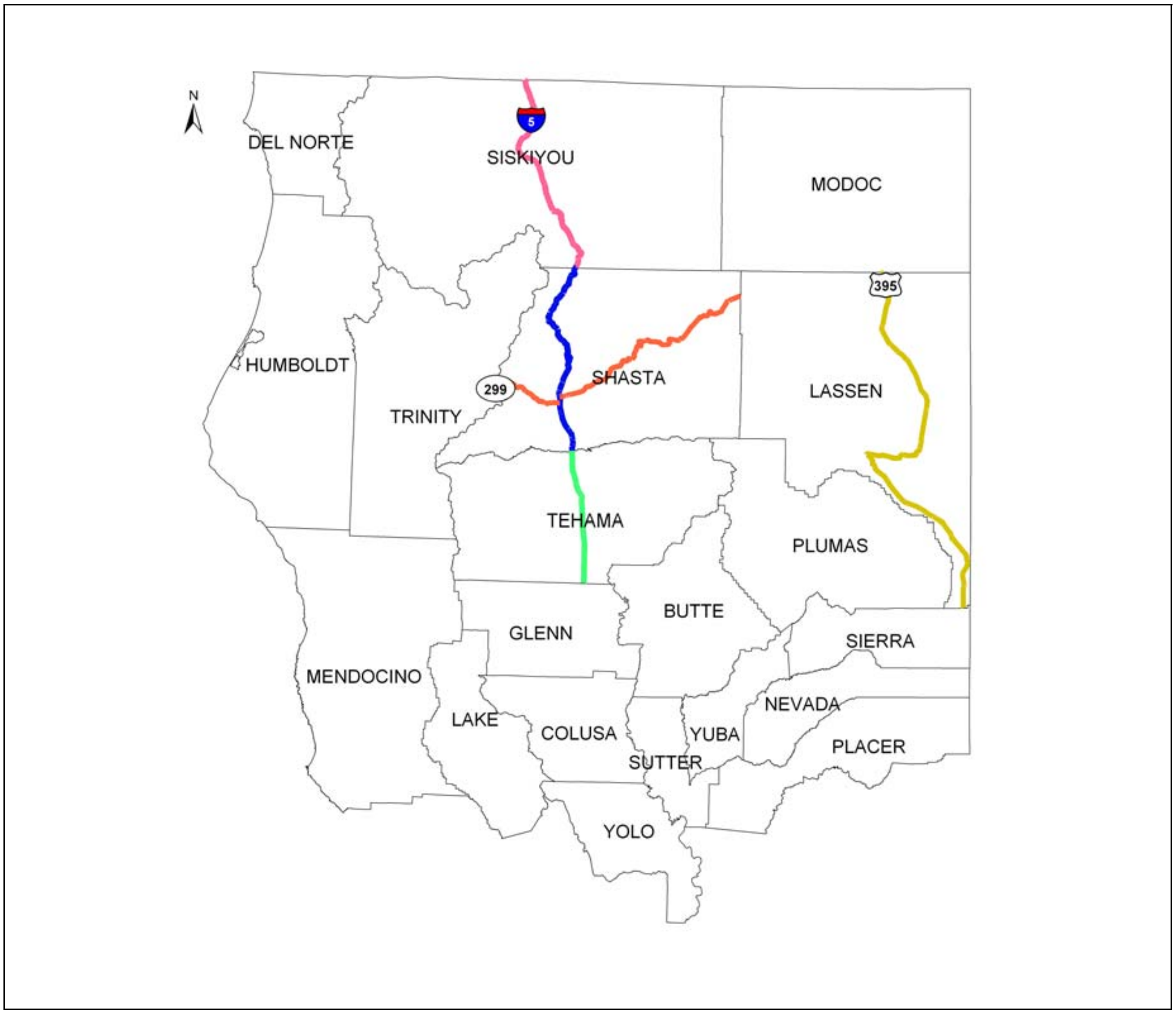

Figure 9.1: Interstate 5, U.S. 395 and CA 299 study segments

9.3. Methodology

The previous section touched upon the generation of Wet Table C lists using a finer resolution without defining what was meant by this term. In the past, limited data and computing power have limited the generation of wet percent time factors to one number for an entire county. Large amounts of detailed data, streamlined analysis packages (e.g., ArcGIS) and enhanced computing power now available allowed the researchers to produce wet percent time factors at a much finer geographic scale. This finer geographic scale can consist of a number of different permutations, ranging from a factor produced for a many square-mile area down to a resolution of a quarter mile-square area along a roadway segment. The ability to work in this more precise scale allows for an examination of the difference between lists generated using a countywide factor versus one of a finer resolution.

For this research, the finer geographic scale produced by the researchers was at the quarter-mile level. This was selected as it was the finest geographic resolution that could be produced by the software utilized (ArcGIS). The generation of these finer resolution factors followed the processes discussed in Chapter 7. Once generated, factors along a length of roadway that shared the same value were combined. In other words, all quarter-mile segments previously generated that shared a common factor were combined into one continuous segment. This was done in order to simplify subsequent Wet Table C production.

In many areas, particularly in the mountains, the total yearly precipitation along a given length of roadway can vary dramatically. This is illustrated in Figure 9.2. As this figure shows, there may be a problem with using one wet percent factor per county. The countywide factor produced using recent data is 5.0, which itself is an increase over the 1972 factor. However, when looking at finer resolution factors at certain points along the length of the roadway, they are considerably greater or less than these singular values. The end result is that there is a potential for a given segment of roadway to be misidentified as having a significant wet-crash problem when, in fact, it may not based on local meteorological conditions. Graphs similar to Figure 9.2 produced for the additional study segments are presented in Appendix E.

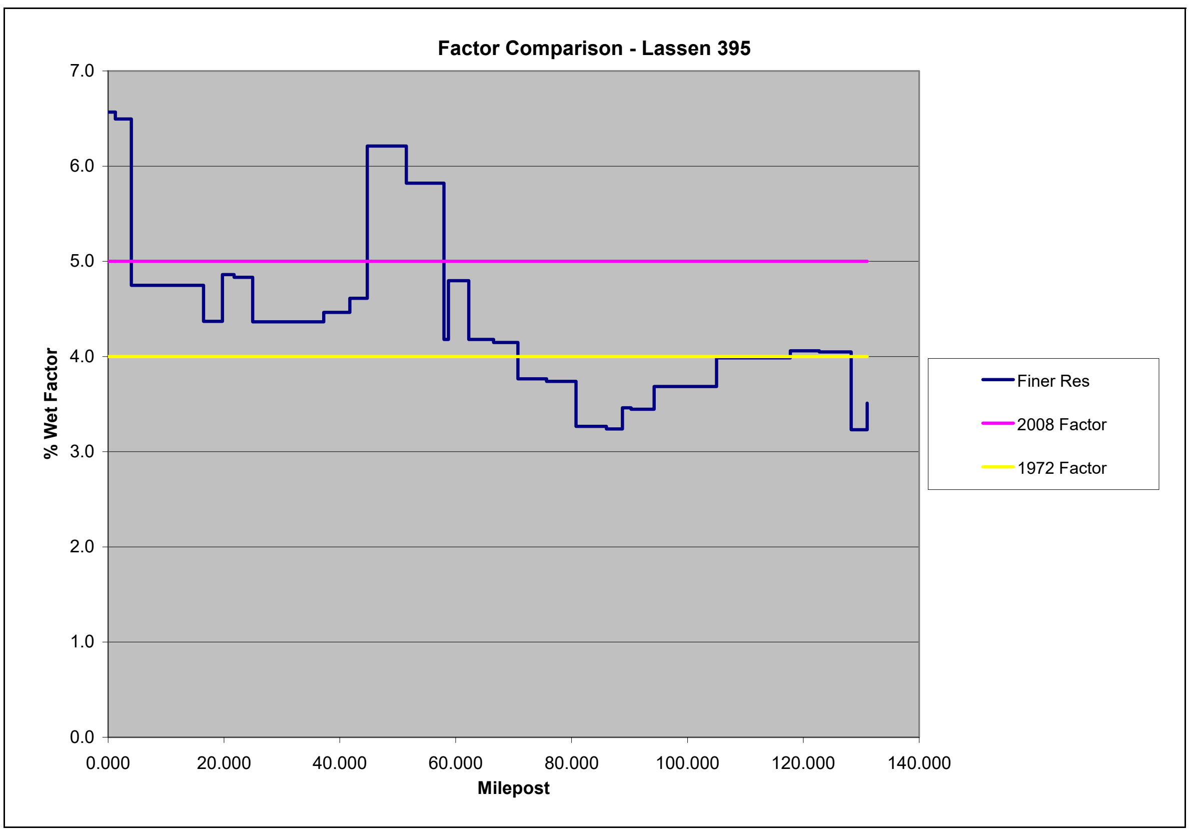

Figure 9.2: Differences between wet percent factors on U.S. 395 in Lassen County

Caltrans personnel generated Wet Table C lists through their automated process using the appropriate 1972 factors, as well as the newer 2008 factors for each route. These lists are presented in Appendix F, Figures F.1 and F.2. Due to database programming complexities, Caltrans personnel were not able to generate similar lists using the finer resolution factors, necessitating the researchers to compute these Wet Table C lists by hand/spreadsheet. This hand generation followed the same procedures and equations employed by Caltrans's automated process. The data to support this effort, namely wet-pavement crashes located by milepost, were provided by Caltrans. These lists are presented in Appendix F, Figures F.3 and F.4. Both strategies employed in generating Wet Table C lists examined the study routes between July 1, 2003, and June 30, 2006.

In total, four evaluations per scenario were conducted. The Wet Table C process identifies locations with significant wet-accident problems for three time periods: 36, 24 and 12 months. In addition, if a location is found to be significant during any of these time periods *and* experiences 9, 6 or 3 crashes during 36, 24 or 12 months, respectively, site investigation by Caltrans personnel is required. These four outcomes were the focus of the statistical evaluation.

Note that the starting mileposts of segments differed between the Caltrans and researcher-produced lists, as the result of differences in the crash data employed (e.g., ramp-related crashes being employed by Caltrans but not by the researchers). This was not of consequence, as the lists generated by each group were not used for comparisons. Rather, Caltrans lists were compared to themselves, as were researcher lists. Beyond this, the only element that differed between the "before" and "after" periods was the wet percent factor being employed.

9.4. Statistical Procedure

The development of alternative Wet Table C lists by Caltrans and the researchers allowed for a statistical evaluation to be performed. This consisted of McNemar tests (an agreement test), which examined whether there were statistically significant differences between the alternative lists. The McNemar tests sought to establish whether the changes observed to occur between lists during the direct comparison procedure were statistically significant. Before moving on to the presentation of the results of this test, some background on the procedure and the rationale for why it was applied to this research are in order.

Sheskin (2000) defines the McNemar test as a nonparametric procedure for categorical data employed in a hypothesis testing situation involving a design with two dependent samples [46]. In essence, the test is employed to evaluate an experiment in which a sample of n subjects or n pairs of matched subjects (in this research, roadway segments) is evaluated on a dichotomous dependent variable (a variable taking on the form of 0 or 1, in this research, a segment being Significant = 1 or Not Significant = 0). The test is particularly applicable in evaluating before/after designs, such as that being examined in the present research (i.e., comparing the changes using an older method (countywide factor) to changes in the later method (finer resolution)).

As Sheskin describes, n subjects are administered a pretest on a dichotomous variable [46]. In the context of the present research, this pretest is a quantification of whether or not an nth segment was significant in terms of wet-pavement accidents. Following this, all subjects (i.e., segments) are exposed to an experimental treatment (i.e., the new, finer resolution wet percent time factor) followed by a post-test on the same dichotomous variable [46]. For the present research, this post-test would consist of a quantification of whether an nth segment changed in terms of wet-pavement accident significance.

To be applied properly, the McNemar test has four assumptions that must be met. First, the subjects must be randomly sampled from the population they represent. In this study, the highways selected for this review have been randomly selected from a larger available sample of roadways throughout the state of California. Second, each of the *n* samples must be independent of other observations. For the present research, each of the segments is independent of neighboring segments in that there is no overlap between segments; rather, there is a clear point where one segment ends and the next begins. Third, the scores of the subjects are in the form of a dichotomous categorical measure involving two mutually exclusive categories [46]. This requirement is met by the present research as the dependent variable can only take on two forms when considering the significance of wet-pavement accidents in the lists: Significant = 1; Not Significant = 0. Finally, the McNemar test should not be applied to small sample sizes. Once again, this is not an issue in the present research as the Wet Table C lists produced a total of between 32 and 47 individual segments for individual analysis, depending on the scenario under evaluation.

A 2 x 2 table best summarizes the McNemar test, and is presented in Table 9.1. As shown, the entries for cells *a*, *b*, *c* and *d* represent the number of observations that occur in each of the four possible categories that can be employed to summarize the two responses of a subject on a dichotomous variable [46]. Cell *a* will represent the total number of segments that were found to have wet-pavement accident significance using a countywide factor (old factor) and a finer resolution factor (new factor). Cell *b* will represent the total number of segments that were found significant using the old factor but were not significant using the new factor. Cell *c* represents the total number of segments that were not significant using the old factor but were found to be significant using the new factor. Finally, cell *d* represents the total number of segments that were not significant using the old factor or the new factor. If there is an increase in the number of significant sites being identified by a finer resolution factor, one could expect the proportion of segments in cell *c* to be larger than the proportion in cell *b*.

Table 9.1: 2 x 2 table for McNemar test

		Condition 2: Finer Resolution Factor		
		Significant	**Not Significant**	**Row Sums**
Condition 1: Countywide Factor	**Significant**	a	b	$a + b = n_1$
	Not Significant	c	d	$c + d = n_2$
	Column Sums	$a + c$	$b + d$	n

The hypothesis evaluated in this research (and by the McNemar test in the case of a before/after design) is whether there is a significant difference between the Condition 1 and Condition 2 scores of subjects on the dependent variable [46]. Specifically, the hypothesis being tested is on the underlying populations π_b and π_c, represented by the proportions:

$\pi_b = b/(b + c)$, and

$\pi_c = c/(b + c)$.

If no difference exists between these proportions, then $\pi_b = \pi_c = 0.5$. Using this information, the null and alternative hypotheses may be stated as:

H_0: $\pi_b = \pi_c$, the proportion of segments in cell *b* equals the proportion in cell *c*;

H_1: $\pi_b \neq \pi_c$, the proportion of segments in cell *b* does not equal the proportion in cell *c*.

As this is a nondirectional hypothesis (i.e., a test is not being performed to determine if one of the proportions is greater or less than another), it is evaluated using a two-tailed test. The chi-squared distribution is employed in evaluating the McNemar test statistic, although it is actually used to provide an approximation of the exact sampling distribution, which is the binomial distribution.

The test statistic for the McNemar test is computed as:

$$\chi^2 = \frac{(b - c)^2}{b + c}$$

Where:

b and *c* represent the number of observations occurring in cells *b* and *c* of the summary test table (see Table 9.1).

The obtained χ^2 value is interpreted by referencing a table of the chi-square distribution. One degree of freedom is employed in the analysis. For a nondirectional alternative hypothesis, as is being tested in this research, the null hypothesis may be rejected if the obtained chi-square value is equal to or greater than the critical two-tailed value obtained from the table at the pre-specified level of significance. For this research, a level of significance of 0.990 was employed to ensure that the results were being interpreted with the highest level of accuracy possible. In all cases, the critical value associated with a 0.990 level of significance for 1 degree of freedom was 6.63.

9.5. Results

9.5.1. 1972 vs. 2008 Countywide Factor

Before an evaluation of what differences may have existed between lists produced using a countywide factor compared to a finer resolution factor, it was necessary to determine what, if any, differences existed between the 1972 and 2008 countywide factor lists. These lists were developed by Caltrans using the established, automated process in which the singular factor is entered as an input. The lists generated using each of these factors included 32 segments that were found to either have wet-pavement accident significance or required further investigation by Caltrans personnel. The results of the McNemar test are presented in Table 9.2.

As the results in Table 9.2 indicate, none of the test conditions were found to reject the null hypothesis. Instead, the data suggest that there may be no significant difference between the lists generated using the 1972 factors and the 2008 factors. Indeed, none of the computed test statistics approached being comparable to the critical value at a 0.990 level of significance.

These are not necessarily negative results, as they reinforce the fact that a changeover to the updated 2008 factors in the Wet Table C generation process is likely to produce lists that are comparable to those currently being produced. Additionally, the two factors produced similar results with respect to requiring further site investigations be performed. This is important in that limited personnel are available to conduct such investigations, and ensuring they are utilized efficiently by not misidentifying investigation sites is of the utmost importance. Still, these conclusions are not entirely certain, as the statistical evaluations were performed on highway segments located in four northern California counties. It is possible that different results may have been obtained for segments evaluated in different counties and in other portions of the state.

Table 9.2: McNemar test results, 1972 vs. 2008 Countywide Factor

36 Months

		Condition 2: 2008 Factor		
		Significant	**Not Significant**	**Row Sums**
Condition 1: 1972 Factor	**Significant**	*26*	*0*	*26*
	Not Significant	*0*	*6*	*6*
	Column Sums	*26*	*6*	*32*

	Test Statistic	Critical value	Result
πb = b/(b+c) = 0	0	6.63	**fail to reject**
πc = c/(b+c) = 1			

24 Months

		Condition 2: 2008 Factor		
		Significant	**Not Significant**	**Row Sums**
Condition 1: 1972 Factor	**Significant**	*18*	*1*	*19*
	Not Significant	*0*	*13*	*13*
	Column Sums	*18*	*14*	*32*

	Test Statistic	Critical value	Result
πb = b/(b+c) = 1	1	6.63	**fail to reject**
πc = c/(b+c) = 0			

12 Months

		Condition 2: 2008 Factor		
		Significant	**Not Significant**	**Row Sums**
Condition 1: 1972 Factor	**Significant**	*6*	*2*	*8*
	Not Significant	*0*	*24*	*24*
	Column Sums	*6*	*26*	*32*

	Test Statistic	Critical value	Result
πb = b/(b+c) = 1	2	6.63	**fail to reject**
πc = c/(b+c) = 0			

Investigation Required

		Condition 2: 2008 Factor		
		Significant	**Not Significant**	**Row Sums**
Condition 1: 1972 Factor	**Significant**	*9*	*2*	*11*
	Not Significant	*0*	*21*	*21*
	Column Sums	*9*	*23*	*32*

	Test Statistic	Critical value	Result
πb = b/(b+c) = 1	2	6.63	**fail to reject**
πc = c/(b+c) = 0			

9.5.2. 1972 vs. 2008 Finer Resolution Factor

Although the results of the previous section indicated that there is no significant difference between the lists generated using the 1972 and 2008 countywide factors, it was still of interest to the research to see how the 1972 factor lists compared to those generated using the 2008 finer resolution factors. These lists were developed by the researchers following the established Caltrans methodology. The only difference between the 1972 and 2008 calculations was the use of the finer resolution factor in the latter list. The lists generated using the differing factors included 47 segments that were found to either have wet-pavement accident significance or required further investigation by Caltrans personnel. The results of the McNemar test are presented in Table 9.3.

As the results indicate, the null hypothesis failed to be rejected for all test cases. Rather, the data suggest that there may be no significant difference between the lists generated using the 1972 factors and the finer resolution factors. Indeed, none of the computed test statistics approached being comparable to the critical value for 0.990. These results should be viewed with some caution, as once again the statistical evaluations were performed on a limited sample of highway segments. It is possible that different results may have been obtained for segments evaluated in different counties and in other portions of the state.

Still, the results are somewhat surprising in that the trend observed in the graph of Figure 9.2, as well as those in Appendix E, showed a good deal of variation in the finer resolution factors that were computed along each study roadway. These graphs would lead one to assume that some difference would be observed, given the radical deviations that the finer resolution factors showed from the singular countywide factors. However, one must take into consideration that the routes examined, while showing precipitation variations, were generally in rural areas. Such routes see lesser traffic volumes daily and, as a consequence, a reduced number of crashes occurring along their length. The end result is that, while variations in precipitation may be observed along a segment, if no crashes occurred along that segment, no differences will be observed in Wet Table C. The locations where a sufficient number of crashes occur that might lead to a difference between lists are ironically those where precipitation is not likely to vary greatly: urban areas.

As a result of this, the results generated through the Wet Table C process were largely the same between the countywide factor and a finer resolution one. Some instances were observed where a segment shifted between significance and non-significance and vice versa, but these were quite limited. Based on these results, it was reasonable to assume that no significant differences would be observed between the 2008 countywide factor list and one generated using a finer resolution factor. This is confirmed in the next section.

Table 9.3: McNemar test results, 1972 vs. 2008 Finer Resolution Factor

36 Months		**Condition 2: Finer Resolution Factor**		
		Significant	**Not Significant**	**Row Sums**
Condition 1: 1972 Factor	**Significant**	*35*	*4*	*39*
	Not Significant	*4*	*4*	*8*
	Column Sums	*39*	*8*	*47*

	Test Statistic	Critical value	Result
$\pi b = b/(b+c) = 0.5$	0	6.63	**fail to reject**
$\pi c = c/(b+c) = 0.5$			

24 Months		**Condition 2: Finer Resolution Factor**		
		Significant	**Not Significant**	**Row Sums**
Condition 1: Countywide Factor	**Significant**	*26*	*4*	*30*
	Not Significant	*4*	*13*	*17*
	Column Sums	*30*	*17*	*47*

	Test Statistic	Critical value	Result
$\pi b = b/(b+c) = 0.5$	0	6.63	**fail to reject**
$\pi c = c/(b+c) = 0.5$			

12 Months		**Condition 2: Finer Resolution Factor**		
		Significant	**Not Significant**	**Row Sums**
Condition 1: Countywide Factor	**Significant**	*9*	*4*	*13*
	Not Significant	*4*	*30*	*34*
	Column Sums	*13*	*34*	*47*

	Test Statistic	Critical value	Result
$\pi b = b/(b+c) = 0.5$	0	6.63	**fail to reject**
$\pi c = c/(b+c) = 0.5$			

Investigation Required		**Condition 2:**		
		Significant	**Not Significant**	**Row Sums**
Condition 1: Countywide Factor	**Significant**	*14*	*0*	*14*
	Not Significant	*1*	*32*	*33*
	Column Sums	*15*	*32*	*47*

	Test Statistic	Critical value	Result
$\pi b = b/(b+c) = 0$	1	6.63	**fail to reject**
$\pi c = c/(b+c) = 1$			

9.5.3. 2008 vs. 2008 Finer Resolution Factor

As indicated in the previous section, the lack of significant differences between the 1972 and 2008 single county factor lists and the 1972 single factor and 2008 finer resolution factor lists indicates that no difference can be expected when the 2008 singular factor is employed. Once again, the lists tested in this section were developed by the researchers following the established Caltrans methodology. The lists generated using the differing factors included 44 segments that were found to either have wet-pavement accident significance or required further investigation by Caltrans personnel. The results of the McNemar test are presented in Table 9.3.

As suspected, the null hypothesis failed to be rejected for all test cases. The data suggest that there may be no significant difference between the lists generated using the 2008 factors and the finer resolution factors. As with previous tests, none of the computed test statistics approached being comparable to the critical value for 0.990. Just as before, these results should be viewed with some caution, as once again the statistical evaluations were performed on a limited sample of highway segments. However, once again, the rural nature of the roadways examined contributed to the lack of differences through the sparse number of crashes occurring over long distances. Different results may have been obtained for segments evaluated in different counties and in other portions of the state. However, as stated previously, the locations where such shifts are likely to occur because of numerous crashes, urban areas, are less likely to see significant variations in precipitation.

Based on the results obtained from this final evaluation, it is reasonable to conclude that the 2008 single county factors may be employed in the Wet Table C process without concern for the varying factors that may in reality occur along the length of a roadway. This is a favorable result, as it eliminates the need for extensive coding work to be done to incorporate varying factors on the legacy system currently employed.

Table 9.4: McNemar test results, 2008 vs. 2008 Finer Resolution Factor

36 Months		Condition 2: Finer Resolution Factor				
		Significant	**Not Significant**			**Row Sums**
Condition 1: 2008 Factor	**Significant**	*39*	*1*			*40*
	Not Significant	*0*	*4*			*4*
	Column Sums	*39*	*5*			*44*
			Test Statistic	Critical value	Result	
	π_b = b/(b+c)= 1		1	6.63	**fail to reject**	
	π_c = c/(b+c)= 0					

24 Months		Condition 2: Finer Resolution Factor				
		Significant	**Not Significant**			**Row Sums**
Condition 1: 2008 Factor	**Significant**	*30*	*0*			*30*
	Not Significant	*0*	*14*			*14*
	Column Sums	*30*	*14*			*44*
			Test Statistic	Critical value	Result	
	π_b = b/(b+c)= 0		0	6.63	**fail to reject**	
	π_c = c/(b+c)= 0					

12 Months		Condition 2: Finer Resolution Factor				
		Significant	**Not Significant**			**Row Sums**
Condition 1: 2008 Factor	**Significant**	*13*	*0*			*13*
	Not Significant	*0*	*31*			*31*
	Column Sums	*13*	*31*			*44*
			Test Statistic	Critical value	Result	
	π_b = b/(b+c)= 0		0	6.63	**fail to reject**	
	π_c = c/(b+c)= 0					

Investigation Required		Condition 2:				
		Significant	**Not Significant**			**Row Sums**
Condition 1: 2008 Factor	**Significant**	*15*	*0*			*15*
	Not Significant	*0*	*29*			*29*
	Column Sums	*15*	*29*			*44*
			Test Statistic	Critical value	Result	
	π_b = b/(b+c)= 0		0	6.63	**fail to reject**	
	π_c = c/(b+c)= 0					

9.6. Conclusions

Based on the results of the McNemar tests, two conclusions may be drawn. First, the lack of any significant difference between lists developed using the single, countywide factors developed in 1972 and 2008 indicates that the transition between the older and newer factor will likely be seamless in terms of the results generated. That no significant differences observed between lists generated using these two factors suggests that Caltrans may proceed with phasing out the use of the 1972 factor as soon as is deemed practical.

The second conclusion that may be drawn is that, at least based on the statistical evaluation performed on a limited sampling of highways, there is no difference between the lists produced using a singular wet percent time factor and one produced using finer resolution factors. The primary result of this is that Caltrans can continue its use of the countywide average when producing Wet Table C lists. Aside from this, it is interesting that the varying factor did not produce greater differences between lists. This appeared to be the result of the low number of crashes that were observed over long distances. However, it might prove of interest for Caltrans to revisit the means by which wet percent time is included in the Wet Table C process in light of this result.

As stated in previous sections, the evaluations presented here were performed on a limited number of highways. It may still prove advisable to make a more widespread evaluation of the impacts that factors (either singular or finer resolution) may have in different portions of the state, as well as for longer segments of roadway.

10. CONCLUSIONS, RECOMMENDATIONS AND FUTURE RESEARCH

The research presented in this report has led to the development of an updated set of wet percent time factors for California highways. These new figures can be used by Caltrans to replace the old set, which was developed with data that is 40 years old. The new table more accurately represents current precipitation trends and should assist in improved identification of high-frequency wet-collision locations in California. This improvement should result in a reduction of wet-pavement collisions, saving lives among the traveling public and reducing the financial impacts of crashes occurring on wet pavements overall.

The primary objective of this research was to update countywide wet percent factors using recent precipitation data from 1995 through 2005. In addition to updating wet percent factors, another objective of this research was to determine the state of the practice with respect to wet-pavement crash location identification nationally, as well as user perceptions of the Wet Table C process and wet percent factors within California. Additionally, the research examined other aspects of the preparation and use of precipitation data in generating wet percent factors. A final objective of this research was to examine what, if any, differences arose between the old (1972) and new (2008) wet percent factors, as well as the differences between these sets of factors and those of a finer resolution (i.e., data representing mile-long sections of highways vs. data representing quarter-mile sections).

The following sections summarize the key conclusions and findings of the research. Note that the conclusions are presented in the same order as the chapters in which they were covered. As a result, conclusions related to the surveys conducted in Chapter 2 will be presented first, followed by missing data handling conclusions, and so forth. This section will be followed by recommendations on how Caltrans should proceed. This chapter concludes by looking at tracks for future research.

10.1. Conclusions

Results of the survey of Caltrans personnel indicated that they thought the current Wet Table C process adequately identified the locations needing improvement. Respondents did view a modification of the wet percent time table to a geographical unit other than county-based as the highest among potential improvements that could be made. In general, Caltrans personnel believed the Wet Table C process was adequately meeting their needs, but recognized that specific data updates and process improvements would be beneficial.

Results of the state survey indicated that a number of states have a focus on reducing wet-weather accidents. Those states that do not have such a focus are those where wet weather is not as common (e.g., the Southwest). Some states that did not have a primary focus on such accidents did have a periodic review as part of a related project, such as 3R projects. A majority of states record information regarding pavement conditions (wet, icy, etc.) in police accident reports. A limited number of states presented wet weather accident information in annual safety reports or other departmental publications. Only a few states indicated that they measured wet pavement exposure.

Results of the processes employed to address missing data indicated that the adopted infill procedures, specifically revision and Nearest Neighbor Frequency Assignment (NNFA), functioned well in addressing the gaps that existed in the data. These gaps were the result of a number of different causes, including equipment malfunctions, deletion through quality control checks, and others. The revision process simply used station data that was available before and after the gap occurred to replace the missing figures, as long as the gap was less than 48 hours in length. For longer periods of missing data, infilling via NNFA was employed, which made use of data from neighboring stations to fill in gaps. Researchers used a simulation test to determine the effectiveness of the infill procedures. A total of 384 hours of missing data was simulated, of which 30 hours were originally rainfall hours for one station. A neighboring station located 10 miles away had complete data for the same month. NNFA was employed as the infilling procedure, with the results indicating that the method effectively infilled 24 of the 30 rainfall hours, along with 360 of the 384 non-rainfall hours. Using these figures, the error percentage of infilling was calculated as (30-24)/384=1.6%.

With respect to the actual update of the wet percent factors, which was the primary objective of this research, the new factors that were generated were reasonably similar in value to the older factors. This would suggest that, although changes have occurred over time in terms of precipitation received by county, these changes have not been radical. Given the quantity and accuracy of the data that were available to the researchers in generating the new factors, it was also possible to derive a factor that excluded the contribution of snow (i.e., included only rainfall data). This new type of factor may prove to be useful in areas where substantial rain and snow precipitation both occur throughout the year and thus present inherent problems for measurement accuracy and have differing impacts on safety. As a result, the development of such a factor allows for new avenues of analysis to be made within the Wet Table C process.

An examination of the various steps and processes employed by the researchers in updating the wet percent factors indicated that that there are some aspects that could be automated through the use of computer programs. These included activities such as data collection and reprocessing. However, the central tasks of data quality control and missing data handling primarily involved human intervention, which was time consuming. At present, it is not possible to develop an automatic data-quality-control algorithm to handle these critical steps by code. As a result, the need remains for some human intervention in the process, at least for the foreseeable future. Additionally, the processes identified as candidates for automation still require further investigation before a conclusion can be drawn regarding the practicality and utility of an automated updating process.

McNemar tests were employed to evaluate what, if any, differences existed between the Wet Table C locations identified as having wet-accident significance using the 1972 countywide factors and the updated 2008 countywide factors. The same evaluation was performed comparing these factors against factors generated at a finer resolution. The finer resolution factors were generated for quarter-mile sections along the study roadways, with consecutive identical values subsequently combined to form segments of varying length. Based on the results of the McNemar tests, two conclusions were drawn.

First, no significant differences were observed between lists developed using the 1972 and the 2008 countywide factors, indicating that the sites identified for further investigation were similar despite the use of newer data. Second, based on the statistical evaluation performed on a limited sampling of highways, no difference was found between the lists produced using a singular wet percent time factor and one produced using finer resolution factors. Therefore, the research suggests that Caltrans can continue its use of the countywide average when producing Wet Table C lists.

10.2. Recommendations

The overall recommendation that can be made as a result of this research is that Caltrans may proceed with phasing out the use of the 1972 factors as soon as it is deemed practical. This recommendation is based on the evidence provided both through the statistical tests performed and the direct comparison of individual county factors. The processes and procedures employed to generate the new wet percent factors appear to have successfully produced new factors that did not significantly deviate from those currently employed. This was primarily evidenced by the similarity in factors that were developed for each county compared to the original factors. Additionally, the Wet Table C lists of site significance developed with each of these factors showed no significant statistical differences.

Aside from the primary recommendation for Caltrans to replace the 1972 countywide factors with the new 2008 factors, it is recommended that consideration be given to future research into different aspects of wet-pavement crashes and the processes employed to identify locations of significance. The final section of this chapter discusses such recommendations in more detail.

10.3. Future Research

During the course of this research and as a result of its findings, several courses of future research were identified. Many of these are broad ideas at present, requiring further refinement should Caltrans show interest in pursuing them. They are provided to give the reader an idea of how the work presented in this text can be leveraged to exploit other avenues of research pertaining to wet-pavement crashes.

The work required to produce the new wet percent factors required a good deal of manual intervention. This intervention was required in many aspects of the work, ranging from data acquisition to quality control checks to the raster map development process. A number of these tasks can be automated; however, the efforts required to automate many of the processes, such as some of the manual quality control checks employed, was beyond the scope and funding of this research. As such, the need remains to investigate the development of more automated procedures so that the list of percent wet factors may be updated with a higher frequency (e.g., yearly or every few years rather than decades).

In speaking with Caltrans personnel, it was interesting to learn that as part of the Wet Table C process, safety personnel are sent into the field to visit each location requiring investigation. While it is heartening that each site is investigated, and thus treated equally, it seems that a ranking scheme to prioritize the order in which sites are visited might be beneficial. Required site visits entail personnel traveling to a site on a rainy day to observe what problems may be present in the wet environment that contribute to crashes. However, there are a finite number of days each year that some areas see rain, and it would seem that a ranking hierarchy allowing safety personnel to make visits to sites deemed most critical would make better use of available resources and weather conditions. Therefore, research into developing such a ranking strategy is strongly encouraged.

Traditionally, precipitation data and processing limitations have limited the extent to which wet-pavement crash locations could be identified. Given the wealth of precipitation data that was collected and refined in this work, it was possible to develop monthly (vs. annual) wet percent time factors, and to develop separate factors that included or excluded the contribution of snow. These more refined factors make it possible to examine wet-pavement crashes at a more micro level, as opposed to the macro level currently employed by California and other states. This opens up the potential for identifying and addressing sites and issues by taking into account the timing and type of precipitation involved. For example, the factor developed excluding snow would help Caltrans identify sites that experienced a significant number of wet-pavement crashes during rain events. This may allow for implementation of specific treatments that apply to rain-related crashes (as opposed to snow-related crashes).

Discussions with Caltrans personnel also indicated that a shift is being considered in the performance of safety evaluations that will include the use of new tools that are or will be coming on-line, such as SafetyAnalyst. The incorporation of wet percent factors into such tools, and the outputs resulting from their use, will require investigation. In such instances, an examination of whether the results generated using a countywide wet percent factor versus a finer resolution factor are different would once again be advisable.

Finally, while the Wet Table C process is doing an effective job of identifying sites with wet-pavement crash problems, it would be useful to examine and apply the procedures of other states to California data. Of specific interest would be the means by which other states include wet pavement exposure factors into their calculations. At the same time, an evaluation would allow for a comparison between respective lists to be made. Such a comparison would allow for validation of the Wet Table C process (if no differences were noted), or the identification of possible issues that need to be addressed or equations that need to be adjusted in the process.

11. REFERENCES

1. Shi, X., Pan, T., Veneziano, D., Huang, J., Ghosh, M., and Kumar, M., “Estimating the Wet Pavement Exposure with Precipitation Data: Literature Review and Current Practices,” A report prepared for the California Department of Transportation, Sacramento, CA, December 2006.
2. American Association of State Highway and Transportation Officials (AASHTO), “Guidelines for Skid Resistant Pavement Design,” 1976.
3. Harwood D.W., R.B. Blackburn, B. Kulakowski, and D. Kibler, “Wet Weather Exposure Measures,” FHWA/RD-87/105, Washington, DC, 1988a.
4. Harwood D. W., R.R. Blackburn, D.F. Kibler, and B.T. Kulakowski, “Estimation of Wet Pavement Exposure from Available Weather Records,” *Transportation Research Record* No. 1172, 1988.
5. Frederick R.H. and J.F. Miller, “Short Duration Rainfall Frequency Relations for California,” *Third Conference on Hydro Meteorology*, Bogota, Columbia, August 20-24, 1979.
6. Karr J.I. and M. Guillory, “A Method to Determine the Exposure of Vehicles to Wet Pavements,” California Division Of Highways, Department of Public Works, January 1972.
7. Blackburn R.R., D.W. Harwood, A.D. St., John, and M.C. Sharp, “Effectiveness of Alternate Skid Reduction Measures—Volume 1: Evaluation of Accident Rate-Skid Number Relationships,” FHWA-RD-79-22, FHWA, Washington, DC, November 1978.
8. Peters R.J., G.J. Allen, and G.R. Morris, “Wet Pavement Accident Analysis for Arid Climates,” Arizona Department of Transportation, 1980.
9. Eischeid J.K., C.B. Baker, T. Karl, and H.F. Diaz, “The Quality Control of Long-term Climatological data using objective data analysis,” *J. Appl. Meteor.*, 34, 2787-2795, 1995.
10. Schmidlin T.W., S.W. Daniel, M. Mckay, and R.P. Cember, “Automated Quality Control Procedure for the ‘Water Equivalent of Snow on the Ground’ Measurement,” *J. Appl. Meteor.*, 34: 143-151, 1995.
11. Feng S., S.Q. Hu, and W.H. Qian, “Quality control of daily meteorological data in China, 1951–2000: a new dataset,” *Int. J. Climatol.* 24, 853-870, 2004.
12. Hubbard K.G., S. Goddard, W.D. Sorensen, N. Wells, and T.T. Osugi, “Performance of Quality Assurance Procedures for an Applied Climate Information System,” *J. Atmos. Oceanic Technol*, 22: 105-112, 2004.
13. Eching S.O. and R.L. Snyder, “Statistical Control Charts for Quality Control of Weather Data for Reference Evapotranspiration Estimation,” *Acta Horticulturae.* 664: 189-196, 2004.

14. Urzen M., S. Anzari, and S. Del Greco, "Automated Spatial Precipitation Estimator (PrecipVal)," Proc. 14th AMS Conf. On Applied Climatology, Amer. Meteorological Soc., Seattle, WA, January 13-16 [Available at http://ams.confex.com/ams/pdfpapers/70681.pdf].

15. You J.S. and K.G. Hubbard, "Quality Control of Weather Data during Extreme Events," *J. Atmos. Oceanic Technol.* 23, 184-197, 2005.

16. Daly C., K. Redmond, W. Gibson, M. Doggett, J. Smith, and G. Taylor, "Opportunities for Improvements in the Quality Control of Climate Observations," *15th AMS Conf. On Applied Climatology*, Amer. Meteorological Soc., Savannah, GA, June 20-23, 2005.

17. Wade C.G., "A Quality Control Program for Surface Mesometeorological Data,". *J. Atmos. Oceanic Technol*., 4, 435–453, 1987.

18. Guttman C.R.G., F.A. Jolesz, R. Kikinis, R.J. Killiany, M.B. Moss, T. Sandor, and M.S. Albert, "White matter changes with normal aging," *Neurology* 50, 972-978, 1998.

19. National Weather Service. "Technique Specification Package 88-21-R1 For AWIPS-90 RFP Appendix G Requirements Numbers: Quality Control Incoming Data," AWIPS Document Number TSP-032-1992R1, NOAA, National Weather Service, Office of Systems Development, 1993.

20. Paulhus J.L.H. and M.A. Kohler, "Interpolation of Missing Precipitation Records," *Mon. Wea. Rev.,* 80**,** 129–133, 1952.

21. Samuel S. and P. Shen, "Interpolation of 1961-97 Daily Temperature and Precipitation Data onto Alberta Polygons of Ecodistrict and Soil Landscapes of Canada," *J. Appl. Meteorol.* 40, 2162-2177, 2001.

22. Daly C., R.P. Neilson, and D.L. Phillips, "A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain," *J. Appl. Meteor.* 33(2), 140-158, 1994 [Available at http://ams.allenpress.com/archive/1520-0450/33/2/pdf/i1520-0450-33-2-140.pdf].

23. Hankins K.D., R.B. Morgan, B. Ashkar, and P.R. Tutt, "Influence of Vehicle and Pavement Factors on Wet-Pavement Accidents," *Highway Research Record* No. 376, 1971. Highway Research Board.

24. Smith R.N. and L.E Elliott, "Evaluation of Minor Improvements (Part 8): Grooved Pavement Supplemental Report," California Division of Highways, Traffic Department, CA-DOT-TR-2152-11-75-01, 1975.

25. Dean J.P., "The Relationship between Accident Experience and Wet Pavement Skid Resistance," *Transportation Research Record* No. 624, pp.129-135. Transportation Research Board, 1976.

26. Holbrook L.F., "Prediction of Wet Surface Intersection Accidents from Weather and Skid Test Data," *Transportation Research Record* No.623, pp.29-39, Transportation Research Board, Washington, DC, 1976.

27. Runkle S.N. and D.C. Mahone, “Virginia’s Wet Pavement Accident Reduction Program,” *Transportation Research Record* No.622, pp. 91-99, Transportation Research Board, Washington D.C., 1976.

28. Levy J.L., “The Effect of Pavement Skid Resistance on Wet Pavement Accidents in Indiana,” Purdue and Indiana State Highway Commission JHRP, 136pp, 1977.

29. Dierstein P.G., “A Strategy for Reducing Wet Pavement Accidents in Illinois,” Illinois Department of Transportation, 41pp, 1977.

30. NTSB, “Fatal Highway Accidents on Wet Pavement–The Magnitude, Location, and Characteristics. Special Study,” National Transportation Safety Board, 45pp, 1980.

31. Kamel N. and T. Gartshore, “Ontario’s Wet Pavement Accident Reduction Program,” ASTM Special Technical Publications, (Ed.) Hayden C.M. pp.98-117, American Society for Testing and Materials, 1982.

32. Dahir S.H.M. and W.L. Gramling, “Wet-Pavement Safety Programs,” *NCHRP Synthesis of Highway Practice* No. 158, Transportation Research Board, 1990.

33. Collins J.S. and M.C. Pietrzyk, “Wet and Wild: Developing and Evaluating An Automated Wet-Pavement Motorist Warning System,” *Transportation Research Record* No. 1759, pp. 19-27, Transportation Research Board, 2001.

34. Meek D.W. and J.L. Hatfield, “Data quality checking for single station meteorological databases,” *Agric. For. Meteor.*, 69, 85–109, 1994.

35. Reek T., S.E. Doty, and T.W. Owen, “A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network,” *Bull. Amer. Meteor. Soc.*, 73,753-762, 1992.

36. Allen R.J. and A.T. Degaetano, “Estimating Missing Daily Temperature Extremes Using an Optimized Regression Approach,” *Int. J. Climatol.* 21, 1305-1319, 2001.

37. Groisman, P. Y., and D. R. Legates, “Documenting and Detecting Long-term Precipitation Trends: Where We Are and What Should be Done,” *Climate Change*, Vol. 31, No. 2-4, pp.601-622, December 1995.

38. DeFelice, T.P. *Introduction to Meteorological Instrumentation and Measurement*. Prentice Hall, Saddle Brook, New Jersey, 1998.

39. Kondragunta, C. R., “An Outlier Detection Technique to Quality Control Rain Gauge Measurements,” *Eos Trans. Amer. Geophys. Union*, 82 (Spring Meeting Suppl.), Abstract H22A-07A, 2001.

40. Englehart, P. J., and A. V. Douglas, “A Statistical Analysis of Precipitation Frequency in the Conterminous United States, Including Comparisons with Precipitation Totals,” *Journal of Applied Meteorology*, Vol. 24, No. 4, pp. 350-362, 1985.

41. Pesonen, E., Eskelinen, M., and M. Juhola, “Treatment of Missing Data Values in A Neural Network Based Decision Support System for Acute Abdominal Pain,” *Artificial Intelligence in Medicine*, Vol. 13, No.3, pp.139-146, 1998.

42. Toth, E., Brath, A., and A. Montanari, "Comparison of Short-term Rainfall Prediction Models for Real-time Flood Forecasting," *Journal of Hydrology*, Vol. 239, No. 1-4, pp. 132-147, 2000.

43. Tsintikidis, D., K.P. Georgakakos, J.A.Sperfslage, D.E. Smith, and T.M. Carpenter. "Precipitation Uncertainty and Raingauge Network Design within Folsom Lake Watershed," *Journal of Hydrologic Engineering*, Vol. 7, Issue 2, pp. 175-184, 2002.

44. U.S. Environmental Protection Agency, "Developing Spatially Interpolated Surfaces and Estimating Uncertainty," EPA-454/R-04-004.

45. Karnieli, A. "Application of Kriging Technique to Areal Precipitation Mapping in Arizona," *GeoJournal*, Vol. 22, No. 4, pp. 391-398, 1990.

46. Sheskin, D. *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman & Hall/CRC, 2000.

APPENDIX A: NOTES ON LITERATURE REVIEW

Note 1: Carr and Guillory's Wet Exposure Methodology [6]

Wet Time: First, define the wet condition. The value of 0.01 inches (0.25 mm) per hour was chosen as the minimal amount of rainfall necessary to keep pavement damp for one hour. Wet time based on 11 years of data was plotted for 350 weather stations throughout California and an isohyetal map (with contours of equal wet time) was developed.

The wet time was defined as percent of hours during which a measurable amount (0.01″ or more) of rainfall occurred. A similar definition of wet-pavement exposure was used in studies by the National Transportation Safety Board (NTSB) and the states of Arizona and Michigan. This method, however, has two weaknesses. First of all, it did not explicitly consider the distinction between frozen and non-frozen precipitation. This distinction is important because accidents classified by road surface conditions have separate categories for wet pavements and for ice- and snow-covered pavements. Thus, this method is only applicable to snow-free areas or to data from which winter months had been excluded. This limitation might not be critical in California, where snowfall is rare in most populated areas. In general, the snow and ice precipitations should be considered separately from rainfall, and be excluded from the calculation of wet-pavement hours. Secondly, this California/NTSB method did not explicitly consider possible variations in the duration of rainfall within each hour, possible variations in pavement drying time, pavement wetness due to melting ice and snow, and pavement wetness due to fog.

Correlation between Precipitation, Traffic, and Accidents: The National Weather Service completed a series of reports on the hourly occurrence of rainfall over a 10-year period in six California cities. The results suggested that Los Angeles, San Francisco, and Oakland tended to have above-normal precipitation before 8:00 a.m. and below-normal precipitation between 3.00 p.m. and 7.00 p.m. Sacramento had an average occurrence of rainfall during both morning and afternoon peaks.

The accident distribution followed the same profile as traffic volume except for 3.00 a.m., which had a disproportionately higher percentage of accidents. In other words, percent of accidents appeared to be low when the precipitation was below average and vice versa. However, rainfall probably caused a portion of these higher accidents at 3.00 a.m. According to highway accident records, a large percentage of accidents around 3.00 a.m. were attributed to driver conditions, i.e., "had been drinking or HBD" and "sleepy" drivers. Unlike hourly distribution, seasonal distribution of traffic and precipitations seemed not to coincide. In fact, rainfall and traffic distribution profiles opposed each other. As such, dry-accidents correlated to traffic distribution and wet-accidents correlated to rainfall distribution.

Wet Exposure: To develop the equation for estimating the wet exposure, the Actual Wet Exposure was calculated beforehand by matching up hourly traffic volume to the hours with 0.01-inch or more precipitation. The Actual Wet Exposure was calculated using traffic counts from nine stations, of which only one had the hourly traffic counts available for every hour of the year, and the traffic counts were "assumed" for the eight stations that had hourly traffic counts available for only one week out of every month of the year.

Intuitively, the prototype equation for estimating wet exposure was developed by adjusting the total vehicle miles with the percent of hours with 0.01-inch or more precipitation, $P_{w0.01}$, which led to the equation cited as follows:

$$W_E = (P_{w0.01}) \times (AADT) \times (days) \times (miles)$$

Where W_E = Estimated Wet Exposure for the year

$P_{W\,0.01}$ = Percent of hours with 0.01" or more precipitation

AADT = Annual average daily traffic

The estimated wet exposures of the nine stations using the equation above were found overestimated as much as 32%, as compared with the Actual Wet Exposure. This error was attributed to incorporation of the AADT in the equation. Instead of AADT, an adjusted ADT that approximated an average winter month's ADT was used. A new term 0.86 defined as the ratio of the average rainy monthly ADT to the AADT was added to the prototype equation, which changed to the format as follows:

$$W_E = 0.86 \times (P_{w0.01}) \times (AADT) \times (days) \times (miles)$$

Where W_E = Estimated Wet Exposure for the year

$P_{W\,0.01}$ = Percent of hours with 0.01" or more precipitation

AADT = Annual average daily traffic

0.86 = Adjusting factor: ratio of the average rainy monthly ADT to the AADT

This equation to develop wet exposures was defined by Caltrans as the P_w method. Compared with the Actual Wet Exposures, the estimated wet exposure values by the P_w method were within 14%. The overall wet exposure of the nine selected stations was within 3% of the Actual Wet Exposures, and seven of the nine locations also yielded values within 3% of the Actual Wet Exposures.

It was further identified that the P_w method did reflect the local traffic patterns that varied considerably in California since the statewide AADT was used in the method, and the 0.86 value in the equation reflected a statewide distribution of traffic. Hence, the equation was further modified as follows:

$$W_E = K \times (Pw_{\,0.01}) \times (AADT) \times (days) \times (miles)$$

Where K could be developed for any area.

The Division of Highways' Traffic Volumes and Analysis Section further developed three factors (L, I, R) for calculating monthly ADTs from annual ADTs. Its equation for MADTs on California state highways was:

$$MADT = AADT \times (L + I\ R)$$

Where

L = 1.00 for an average (statewide) day

I = Monthly Distribution Factor

MADT = Monthly average daily traffic

I averaged (– 0.44) for seven months (October to April) that had substantial precipitation and the equation became: MADT = AADT×(1-0.44R).

The seasonal factor "R" was calculated as the average summer months minus the average winter month; that result is divided by AADT:

$$R = \frac{\sum C_{sm} - \sum C_{wm}}{6 * AADT}$$

Where

ΣC_{sm} = Average summer months ADT

ΣC_{wm} = Average winter months ADT

Then the prototype equation changed to:

$W_E = 0.98 \times (1-0.44R) \times (Pw_{\,0.01}) \times AADT \times (days) \times (miles)$

Note 2: Blackburn et al.'s Accident Rate – Skid Number Relationship [7]

A method developed by the Midwest Research Institute (MRI) was adapted to define wet time. The method classified an hour as wet time if:

1. A measurable amount (0.01 ″ or greater) of non-frozen precipitation occurred during the hour;
2. A trace amount (less than 0.01″) of non-frozen precipitation occurred during the hour, except that periods of precipitation composed entirely of trace amounts are not counted;
3. Fog occurred during the hour; or
4. The hour immediately followed a period of precipitation (frozen or non-frozen) in which a measurable amount of precipitation occurred in at least 1 hour).

Rules 1 through 3 accounted for wet time due to rain or fog. Rule 4 accounted the period of time after precipitation had stopped, when the pavement was still wet. This study focused on estimation of wet-pavement drying time as the hours of rain and fog were already well documented in the literature. The evaporation rate is influenced by many factors (temperature, relative humidity, wind velocity, cloud cover, traffic volume).

A comparison was made between highway drying times in Ohio and Louisiana against the evaporation rate at the nearest Class A pan location, assuming the depth of water on the pavement to be 0.02 inch. The study found that highway pavements dried faster than the pan evaporation rates. Furthermore, the passage of traffic tends to minimize the regional differences in pavement drying times. Based on the study findings, ½ hour for pavement drying time was proposed. The sensitivity of the wet-pavement exposure time to the proposed change was evaluated by applying the original and modified technique at five different locations (Detroit, Mich., Columbia, S.C., Worcester, Mass., West Palm Beach, Fla., and Astoria, Ore.) and the differences were found to be insignificant. As a result, Rule 1 through 3 remained same, whereas Rule 4 was modified as follows:

A half hour is classified as wet time if it immediately follows a period of precipitation (frozen or non-frozen) in which a measurable amount did occur in at least one hour.

Note 3: Peters et al. Wet-Pavement Exposure Methodology [8]

Actual Rainfall and Traffic Volume Method: Reported precipitation for each given hour of the year was noted and the traffic volume, as recorded for the corresponding hour of precipitation plus 1 hour, was recorded as the hours of wet pavement. The traffic volume divided by 365 yielded the value of wet AADT. However, the authors argued that this method was impractical because it required constant recording of 24-hour precipitation and traffic.

Average Hourly Rainfall and Average Hourly Traffic: Reported precipitations adjusted to the number of wet-pavement hours were averaged for the time period to derive a percent of a 24-hour day each hour is wet. Also hourly traffic volume counts were averaged and an Average Annual Hourly Traffic (AAHT) was derived. The hourly percent of wet pavement multiplied by the AAHT yielded the wet AAHT. The wet AAHT summed for 24 hours was the wet AADT. The authors argued that this method yielded high wet-pavement rates and was not considered reliable.

Ratio of Wet Hours of Pavement to Total Hours and AADT: The number of hours of wet pavement divided by total hours, yielded percent of wet-pavement time. This multiplied by the AADT yielded a wet AADT. According to the authors of this study, this method seemed to provide the best approximation of wet percent of total AADT.

An isohyetal map (with contours of equal wet time) was developed using the Caltrans method as described on page 7 (similar to the third method above), which provided a fair approximation of wet-pavement time for any year. However, when the percent wet time taken from this map was applied to years where there were abnormal precipitation times, durations, or magnitude, a large error occurred in approximating wet AADT. To circumvent this difficulty, an improved isohyetal map based on climatological data was proposed in this study.

Accident reports related to wet pavement for a 72-mile section over a three-year period were checked against the tabulation of dates and hours in which wet pavement was assumed possible, using the reported climatological data. The results showed that 77 percent of reported wet pavement accidents occurred within the assumed hours of wet pavement. Additionally 9 percent of the wet pavement occurred one hour before the assumed hours of wet pavement and 9 percent occurred one hour after. These led to an accuracy of 95 percent in determining the hours of wet pavement.

Five 24-hour precipitation-recording stations in Arizona were used to determine actual percentage of wet pavement. The state was then divided into five sectors by use of the precipitation intensity maps to identify locations of similar precipitation characteristics near the five recording stations. With 42 locations, an isohyetal map was created. Using the "percent of time pavement is wet" and the AADTs for the state system, wet-pavement accident rates were calculated for each route or location. These rates can be compared by road category to determine high wet-pavement accident locations, based on which necessary pavement surface treatment can be undertaken to increase friction and reduce skid potential.

The following equations were used:

$$\text{Wet-accident Rate} = A_w \times 10^6 / (L \times ADT_w \times T_w)$$

Where

A_w	=	Number of accidents reported occurring on wet pavement in a particular period
L	=	Length of section considered in miles
ADT_w	=	Average daily traffic volume on wet pavement for the same period, obtained by multiplying percent of time pavement is wet from map by AADT
T_w	=	Number of wet days in period

$$\text{Dry-accident Rate} = A_d \times 10^6 / (L \times ADT_d \times T_d)$$

Where

A_d	=	Number of accidents reported occurring on dry pavement in a particular period
L	=	Length of section considered in miles
ADT_d	=	Average daily traffic volume on dry pavement for the same period
T_d	=	Number of dry days in period

The authors suggested a faster method of obtaining a general indication of possible wet-pavement impact using the wet accident index (WAI) determined by the following equation.

$$WAI = (A_w \times P_d) / (A_d \times P_w)$$

Where

P_d	=	$100 - P_w$
A_w	=	Number of accidents on wet pavement
A_d	=	Number of accidents on dry pavement
P_w	=	percent time pavement is wet

P_d = percent time pavement is dry

If WAI = 1, there is no difference between wet and dry accidents;

If WAI < 1, the accident risk on wet pavement is less than on dry pavements; and

If WAI > 1, the accident risk on wet pavement is greater than on dry pavements.

Note 4: Harwood et al.'s Estimate Wet Weather Exposure Measures [3]

The primary objectives of this study were to:

- Establish the minimum levels of wetness at which tire-pavement friction was substantially reduced;
- Develop a model for predicting how many hours per year this minimum wetness level was exceeded, as a function of regional meteorological and pavement characteristics; and
- Demonstrate use of the model by developing wetness maps for representative regions.

This study involved a series of laboratory and field investigations to develop the basic building blocks of a wet-pavement exposure estimation model, including:

- The minimum water film thickness on a pavement surface that substantially reduces tire-pavement friction; and
- The time required for a wet pavement to dry following the end of rainfall, as a function of ambient environmental conditions.

Analytical and observational studies were also conducted to investigate:

- The relationship between the total amount of rainfall in a given period and the duration of rainfall within that period;
- The conditions under which pavement drying cannot occur due to saturated or nearly saturated atmospheric conditions;
- The conditions under which pavement wetness may result from condensation on the pavement during fog; and
- The conditions under which pavement wetness may result from ice and snow conditions.

The research concluded that the minimum level of wetness that substantially reduced pavement surface friction was between 0.001 and 0.009 in. (0.25 and 0.23 mm) of water on the pavement. This minimum level of wetness was likely to be exceeded in any hour that has at least 0.01 in. (0.25 mm) of rainfall.

The improved MRI method was incorporated into a computer model, known as the WETTIME model, for application by highway agencies. The model includes a unique algorithm to estimate the number of hours with wet-pavement conditions on monthly and annual bases. The model incorporated the following elements in the wet-time estimation:

- Minimum level of wetness that reduced pavement surface friction
- Rainfall intensity and duration
- Runoff period following rainfall
- Pavement drying period following rainfall and runoff
- Pavement wetness due to fog
- Estimation of exposure to ice and snow conditions

The following rules form the basis for the WETTIME model.

1. An hour with no precipitation is counted as DRY, unless there is still pavement drying underway from the previous hour.
2. If non-frozen precipitation of 0.01 in. or more occurs during an hour, then the time while the rain is falling and the subsequent drying time is counted as WET.
    a. For an isolated hour of precipitation (no precipitation in either the previous or the following hour), the duration of pavement wetness due to the rainfall is determined as follows:

Table 11.1: Duration of Wetness and Total Rainfall

Total Amount of Rainfall During the Hour (in.)	Duration of Wetness
0.01	15 min.+ runoff + drying time
0.02	30 min. + runoff + drying time
0.03-0.04	45 min. + runoff + drying time
0.05 or more	60 min. + runoff + drying time

    b. For the first hour of two or more consecutive hours of precipitation, the duration of wetness is determined as described in the table. Whatever is the duration of the rainfall period, it is assumed to occur at the end of the hour.
    c. For the last hour of two or more consecutive hours of precipitation, the duration of wetness is also determined as described in the table. Whatever the duration of the rainfall period, it is assumed to occur at the beginning of the hour.
    d. For a middle hour of a period of three or more consecutive hours of precipitation, the rainfall is assumed to last for the entire hour.
    e. The runoff period following the end of rainfall is assumed to be five minutes.

f. Pavement drying usually begins at the end of rainfall and runoff, and continues until the pavement is dry or a new storm begins. If the pavement is still wet at the end of an hour, pavement drying continues into the next hour.

g. The start of pavement drying may be delayed if the ambient air is nearly saturated (as indicated by a dew point temperature within 2° F of the ambient air temperature). During the day time, the delay in the start of drying will last a maximum of two hours or until the air is no longer saturated. At night, the delay in the start of drying will last until the air is no longer saturated or until the drying due to solar radiation begins shortly after dawn.

h. The duration of pavement drying is determined from a statistical model that predicts drying time. The factors used to predict pavement drying time are: solar radiation, wind speed, air temperature, relative humidity, and pavement type. The predicted pavement drying time is rounded to the nearest five minutes. The program user can specify the pavement type.

i. The environmental factors in the pavement drying model are determined from weather data.

3. If fog occurs during an hour where the air is nearly saturated and the wind speed is 3 mph or less, then the hour is counted as WET. Pavement drying following a period of fog follows the same rules as following a period of non-frozen precipitation.

4. If frozen precipitation of 0.01 in. or more occurs during an hour, then the hour is counted as ICE and SNOW.

   a. If a trace of frozen precipitation occurs during an hour and the temperature remains below 32° F, the hour is counted as ICE and SNOW. When these conditions apply continuously for several hours, subsequent hours may also be counted as ICE and SNOW.

   b. The pavement drying time following a period of frozen precipitation is determined by the same rules as for non-frozen precipitation and is counted as WET.

Procedures were also developed making use of the WETTIME model output to prepare isoexposure contour maps for entire states or regions. Isoexposure contours were lines connecting locations of equal wet-pavement exposure.

The number of first-order weather stations available in each state is very limited. Therefore, the study developed two alternative methods to include more weather stations (i.e., NOT first-order weather stations) to adequately cover the whole state. The first approach estimated annual wet-pavement exposure at minor stations from annual exposure estimates for first-order stations by proportioning on annual number of hours with measurable precipitation. The quality of hourly precipitation data from National Climatic Data Center (NCDC) for many minor stations was found too poor to produce accurate estimates. A brief comparative analysis of minor and first-order weather station data near Kansas City concluded that the hourly precipitation data from minor weather stations could not be relied upon. A major reason for this sub-quality of data from minor stations was that some of these minor stations had rainfall gauges that were not, in fact, read each hour and that several hours of rainfall results were accumulated into a single hour without this being indicated in the NCDC file.

The second approach estimated annual wet-pavement exposure at minor stations from annual exposure estimates for first-order stations by proportioning on total annual rainfall. The long-term estimates of annual precipitation totals based on 30 years of records published by NCDC for numerous weather stations in each state were used to provide a reliable basis for estimating wet-pavement exposure at minor stations. To obtain consistent results, it was necessary to use an average of the wet-pavement exposure estimates for the two closest first-order weather stations, weighed inversely by their distance from the minor station. This weighing procedure avoided discontinuities on the contour map near the boundaries of the area of influence of different first-order weather stations and provided a smooth transition between their areas of influence.

When the wet-pavement exposure estimate is not available, a highway segment would have the same wet-pavement accident rate for a given number of wet-accidents regardless of where the segment is located in the state. This unrealistic assumption is made implicitly by any highway agency that does not incorporate a measure of wet-pavement exposure in its wet-pavement accident surveillance program. This method would miss a highway section with high wet-pavement accident experience that happens to be located in a relatively dry part of the state or, to identify a highway section as a problem location merely because it happens to be located in a relatively wet part of the state. The accident rate estimation based on wet-pavement exposure estimates is thus necessary for a realistic comparison of wet-accident risks. The improved MRI model for the estimation wet-pavement exposure helps improve the wet-accident rate estimations.

Note 5: Feng et al.'s Quality Control Process [11]

QC of the data included visual inspection of graphs of all station time series, tests for precipitation digitized six months out of phase, tests for different stations having identical data, and other tests.

High-Low Extreme Check for Daily Values: This method compared daily values of a number of variables from individual stations with established extreme values. Data values greater than the highest values or lower than the lowest values were listed, flagged and excluded from subsequent quality-control calculations.

Internal Consistency Check: Reek et al. (1992) outlined eight rules to identify erroneous data of temperature and precipitation. It was concluded that the errors were usually due to data reporting and digitizing, typos, unit differences and the use of different base values in data reporting [35]. In the study, the authors used three of their rules to check the daily data: a) internal inconsistency; b) excess diurnal range check, which identified errors with extraordinarily large daily range of maximum value – minimum value, while both the values were within their reasonable ranges; and c) a "flatline" check, which identified data of the same value for at least seven consecutive days (not applied to zero-precipitation data).

Temporal Outliers Check: When data values were much larger (or smaller) than neighboring values but did not exceed the threshold for being detected by the internal consistency check, the data created a large step change from the previous daily value(s) [36]. To identify these outliers, the *Lanzante's biweight mean and biweight standard deviation method* was used.

Spatial Outliers Check: This method detected the outliers by comparing the data of neighboring stations. Correlation coefficients R were computed for each month between daily data at a station (candidate station) and the 10 nearest stations. The minimum criterion was that R be significant at the 95 percent confidence level. Stations with large positive R were used to create their linear regression for the same variable between neighboring stations and the candidate station. The root-mean-square error (RMSE) of the regressions was also computed. If more than five neighboring stations had significant correlation with the candidate station in a specific month, the five neighboring stations with the lowest RMSE were chosen. After having N (N ≤ 5) regression equations, a daily value V_i of a variable was assigned to be suspicious if it fell outside the specified confidence intervals for all N pairs of stations [12]:

$$VF_{ij} - F \times RMSE_j < V_i < VF_{ij} + F \times RMSE_j$$

Where j = 1, … N; N was the number of neighboring stations; i = 1, … , m; m was the specific day in a month and m was the total number of days of that month; V_i was the data at the candidate station for day i; VF_{ij} was the fitted value by linear regression of neighboring station j for day i; and F defined the desired confidence limit. $F = 5$ was chosen for precipitation by Feng et al. [11].

Inhomogeneity Check: After the consistency checks mentioned above, three statistical methods were used to check the inhomogeneity in the data series that may be caused by relocation of stations and/or sensor problems. Data discontinuities resulting from such interruptions were identified by these methods, and were confirmed using the stations' metadata where available.

Note 6: Hubbard et al.'s Quality Control Process [12]

"Upper and Lower" Threshold Test: It checked whether or not a given variable fell in a specific range for the month in question. When the limits were determined based on the statistics of the distribution, it was called the *sigma test* [19].

Step Change (SC) Test: It checked whether or not the change in consecutive values of the variable fell within the climatologically expected lower and upper limits on daily rate of change for the month in question.

Persistence Test: It checked the variability of the measurements. When a sensor fails it will often report a constant value; thus the standard deviation (S) will become smaller. If the sensor is out for an entire reporting period, the standard deviation will be zero. In other cases the station may work intermittently and produce reasonable values interspersed with zero values, thereby greatly increasing the variability for the period. Thus, when the variability is too high or too low the data should be flagged for further checking. The first step was to calculate the standard deviation from daily values for each month (j) and year (k) of the 30-year record, S_{jk}. Then the mean standard deviation was calculated for each month j by averaging S_{jk} over the years. Likewise, the standard deviation of these monthly values was calculated over all years. The persistence test compared the standard deviation for the time period being tested.

Spatial Weighted Regression Test: It checked whether or not the variable fell inside the confidence interval formed from estimates based on N ''best fit'' neighboring stations during a time period of length n. Hubbard et al.[12] chose the values of N and n to be 5 and 24, respectively. The surrounding stations were selected by specifying a radius around the station and finding those stations with the closest statistical agreement to the target station.

Note 7: Paulhus and Kohler's Methods for Interpolating Missing Precipitation and Weather Data [20]

Regression Equation Method: A least-squares regression equation in the format of $P_X = b_1P_1 + b_2P_2 + b_3P_3$ is known to be satisfactory for estimating the precipitation P_x at Station X from the precipitation (P_1, P_2 and P_3) at three index stations [34]. However, the least-squares regression has limited use in that not only it is applicable only when reports are available from all the selected index stations but also the required analysis is quite time consuming. These disadvantages appear to make this method undesirable for handling large-scale missing data, as in the case of this project.

To obtain a set of coefficients in the above equation (b_n) by a simpler procedure, the coefficients were defined as a function of the intercepted angle between rays from station X bisecting the angles between lines from X to index stations 1, 2, and 3 (θ_n) and relative distance $(d_n/\Sigma d)$ between X and 1, 2, and 3, respectively, i.e., $b_n = f\left(\theta_n, \frac{d_n}{\Sigma d}\right)$. The method was tested using sample data, but no satisfactory results were obtained. As such, the method needs further study.

Three-Station-Average Method: This method averages the precipitation of all three nearby stations to estimate the precipitation of station X, i.e., $P_X = \frac{1}{3}(P_1 + P_2 + P_3)$. Preferably the stations are of similar hydrometeorologic characteristics. This method has a disadvantage in that the three stations have to be evenly distributed around the X station and the straight average of precipitation at surrounding stations would not always yield accurate estimates in mountainous regions. However, the three-station-average method was found to be reasonably satisfactory. A slight improvement of this arithmetic averaging method was to weight each of the nearby station records by its distance, L_j, from the station in question.

Normal Ratio Method: This method uses the ratio of the normal annual precipitation (N_x) at the interpolation station X to that at the index stations as a weighting factor, i.e.,

$P_X = \frac{1}{3}\left[\left(\frac{N_X}{N_1}\right) P_1 + \left(\frac{N_X}{N_2}\right) P_2 + \left(\frac{N_X}{N_3}\right) P_3\right]$. The study showed the normal-ratio method yielded better results than the three-station-average method.

The authors pointed out that the normal-ratio method would be used by the Weather Bureau whenever the normal annual precipitation at any of the index stations differed from that of the interpolation station by more than 10 percent. But the simpler three-station-average method might be used. Also, the principles to govern the application of these two methods and principles to selecting the index stations for making interpolations were listed.

Finally, the selection of period for estimates was determined by the portion of records missing at a station, and they fell into two categories:

- When the record for Station X was missing for an entire month, the monthly amounts at Stations 1, 2, and 3 were used as the basis of the estimate. When the record was missing for only a portion of a month, the amounts used for estimating were for the period actually missed or for a few days longer, depending on whether the missing period began and ended between storms or during a storm or storms.
- When the missing record extended into two months, first an estimate was made of the missing amount for the total storm, without regard to month. Next, an estimate of the portion of the missing record in one of the two months was made using only precipitation for the corresponding days at stations having the same observation time as the missing station. Finally, the portion for the other month was obtained by subtracting the two values obtained in the preceding steps.

Note 8: Feng et al.'s Interpolation Procedure [11]

Correlation coefficients R were computed for each month between daily data at a station (candidate station) and the 10 nearest stations. The minimum criterion was that R be significant at the 95 percent confidence level. Stations with large positive R were used to create their linear regression for the same variable between neighboring stations and the candidate station. The root-mean-square error (RMSE) of the regressions was also computed. If more than five neighboring stations had significant correlation with the candidate station in a specific month, the five neighboring stations with the lowest RMSE were chosen. After having N (N ≤ 5) regression equations, a daily value V_i of a variable was assigned to be suspicious if it fell outside the specified confidence intervals for all N pairs of stations [12]:

$$VF_{ij} - F \times RMSE_j < V_i < VF_{ij} + F \times RMSE_j$$

Where j = 1, … N; N was the number of neighboring stations; i = 1, … , m; m was the specific day in a month and m was the total number of days of that month; V_i was the data at the candidate station for day i; VF_{ij} was the fitted value by linear regression of neighboring station j for day i; and F defined the desired confidence limit. $F = 5$ was chosen for precipitation by Feng et al. [11].

For the interpolation of missing data, the following equation was used:

$$v_{ei} = \sum_{j=1}^{N} \left[\mathrm{VF}_{ij} \times \mathrm{RMSE}_j^{-2} \right] \Big/ \sum_{j=1}^{N} \mathrm{RMSE}_j^{-2}$$

where V_{ei} is the estimated value and the other symbols are the same as in the first equation.

Note 9: Smith and Elliott's Equation to Predict Wet-Accident Reductions from Future Projects [24]

$WAR_A = 0.30 + 3.17\,(DAR_A)$

Where

WAR_A = Wet-accident rate after grooving

DAR_A = Dry-accident rate after grooving

The authors noted that the results strictly apply to high-volume urban freeways (60,000 – 200,000 ADT) and where all roads had a history of wet-accidents. As a rule, grooving is justified only when the wet-accident rate is four times greater than the dry-pavement accident rate. Therefore, grooving is not warranted unless this condition is met. Finally, the estimated savings in accident costs should exceed the cost of grooving to be economically viable.

Note 10: Holbrook's Wet Surface Model [26]

$(WA/TA) = [M(i)*(WH/TH)]^{g(\mu)}$

Where WA/TA represents the proportion of wet-accidents;

WH/TH represents the proportion of wet time;

$g(\mu) = \theta_4[\mu^3 - \mu] + \theta_5[\mu^2 - \mu] + \mu$

The model was fitted to the actual data. The non-linear least squares procedure provided predicted wet percentages values that were in close agreement with the actual wet-accident percentages. The study suggested that estimated surface wet time and skid number were important factors in wet-accidents. However, no critical skid number emerged as a point above which wet-accidents disappeared. Rather, the study indicated that wet-accidents appeared to be a continuously decreasing function of the surface friction. Below a skid number of approximately 30, wet-accidents increased at a slightly increasing rate with declining surface friction. This was found to be true for all months and wetness categories.

Note 11: Runkle and Mahone's Wet Accident Site Benchmark Procedure [27]

The benchmark was defined as follows to identify the potential sites. When wet-accidents accounted for at least 30 percent of the total number of accidents, or SN40 values were below 30, the sites were considered potentially problematic and needed attention. Once such potential sites were selected, the data on surface mix type were obtained as it is an important factor in providing adequate skid resistance.

Initial economic analysis was then conducted to compare the savings from estimated wet-accident reduction against the probable cost of making improvements (site treatments such as resurfacing) necessary to reduce wet-accidents. The authors provided a benchmark for such improvements, i.e., wet-accidents should account for at least 20 percent of the total number of accidents.

Following the initial economic analysis, sites were ranked based on their breakeven value, which is the total cost divided by the projected savings. For sites with high priorities, a field review was then conducted to investigate the geometrics, traffic turbulence, sight distance, roadside development, traffic control, posted speed limit, and general pavement condition at the sites. If SN values were found to be of little importance, and other variables such as traffic control or roadside congestion were found to be prime factors, then the matter was referred to the Traffic and Planning Division. In contrast, if the evidence indicated higher SN would be of value, a second economic analysis was performed to take into account the accident data, skid test data, and field review data, in order to identify sites where the greatest benefit/cost ratio could be realized.

APPENDIX B: CALIFORNIA USERS SURVEY

Note 1: Reason for No Action Taken

The first question posed to survey participants sought feedback with respect to the reason(s) that no action might be taken after investigating a "required" location in Wet Table C. Respondents were given the following options and asked to rank them from 1 to 5, with 1 being the most frequent reason that no action was taken and 5 being the least:

- Not a wet pavement/grooving related safety problem;
- Peak-hour congestion related collisions;
- Increased traffic volume during seasonal peaks;
- The location is repeated or overlaps with another location from last year's Wet Table C;
- None of the above (Please describe in the comment section at the end of the survey).

Depending on their use of Wet Table C, not all respondents ranked each of these choices. Rather, some respondents only ranked those conditions which they had direct experience with. As a result, some options (such as "None of the above" selection) received a low number of ranked responses.

The results in the chart below indicate the ranking breakdowns of the primary reasons cited by respondents for not taking action after investigating required locations. The responses indicate that the most common reasons for no action being taken were that the safety problem was either not wet-pavement related (10) or was the result of congestion occurring during the peak hour (10).

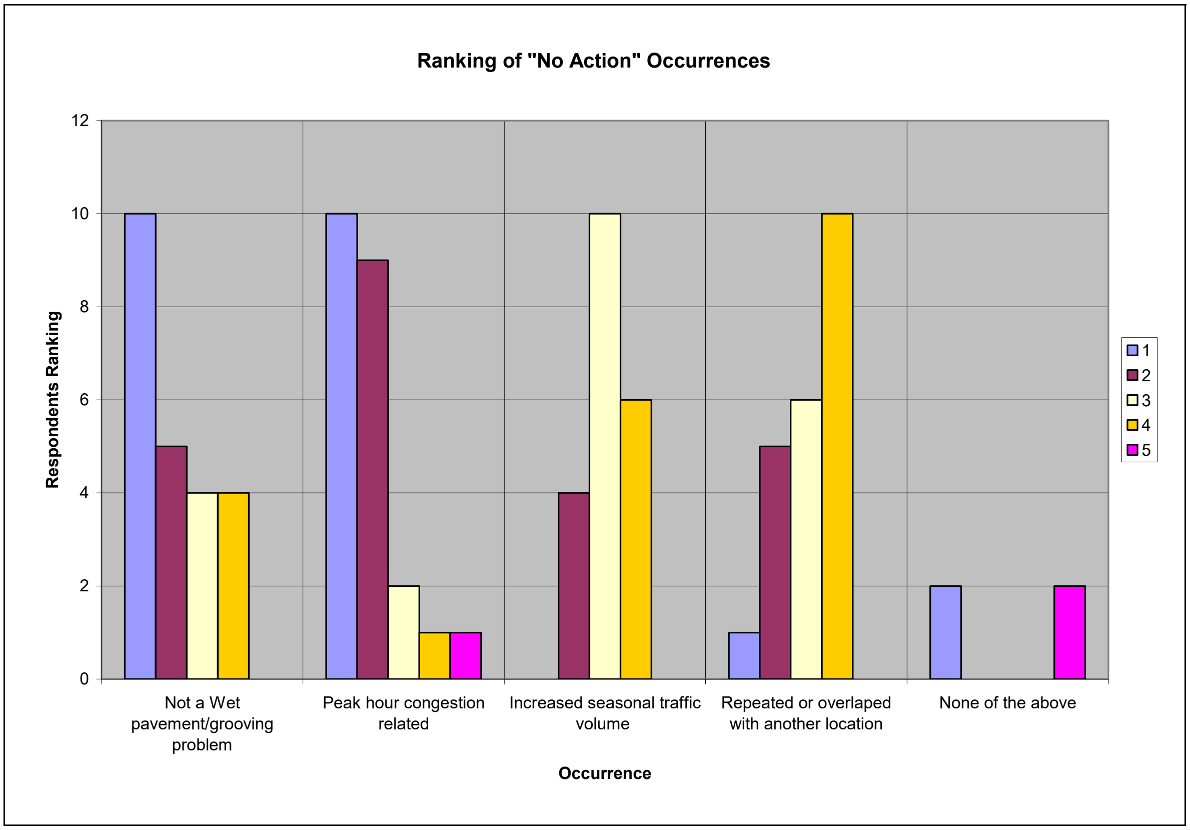

Figure 11.1: Reasons for "No Action" being taken

As stated, one of the primary reasons for no action being taken was that the safety problem was not wet pavement or grooving related. Ten respondents ranked this as their number one reason for no action being taken. Five respondents ranked this as their second most frequent reason for no action being taken, while four ranked it as their third and fourth problem.

Similar to the problem not being wet pavement or grooving related, ten respondents ranked peak-hour congestion related collisions as their top reason for why no action was taken for a required location in Wet Table C. Nine respondents ranked this as their second most frequent reason for taking no action, while two ranked it as their third. Two remaining respondents ranked this as their fourth and fifth reasons for no action taken.

No respondent ranked increases in seasonal traffic volume as the primary reason for why no action was taken at Wet Table C locations. Only four ranked this as their second most frequent reason. Ten respondents cited this as their third most frequent reason for not taking action, which was the majority for this occurrence. Finally, six respondents cited this problem as the fourth most frequent occurrence they encountered.

Only one respondent ranked locations which were repeated or overlapped those from the prior year as their top reason for why no action was taken. Five respondents ranked this as their second most frequent occurrence, while six ranked it as their third. Ten respondents ranked this as their fourth most frequent occurrence resulting in no action being taken.

Four respondents ranked none of the previously cited reasons as being the reason for which no action was taken. Two respondents ranked this as their most frequent occurrence, while two listed it as their fifth. The low number of respondents who supplied a ranking for this option should be viewed with caution. While two respondents do place this as their primary reason for not taking any action, it must be noted that twenty other respondents ranked problems such as not wet pavement or grooving related or peak-hour congestion related as their most frequent reason. Each of these is clearly a greater problem than an unidentified problem which a small number of analysts encountered.

Note 2: Effectiveness of Wet Table C

The second question respondents were presented asked how well Wet Table C identified only the locations that were in need of improvement based on the experience in investigating "required" locations. The responses to this question were requested in the format of a Likert scale between 1 (more than adequate) and 3 (Not adequate) and an option to select "Other" was also provided. All 23 survey participants answered this question, with the following breakdown in their responses:

Table 11.2: Effectiveness of Wet Table C

Identification Performance	No. of Responses	Response Percentage
More than Adequate	1	4.35%
Adequate	17	73.91
Not Adequate	4	17.39%
Other (Please Describe)	1	4.35%

Note 3: Provision of One Factor

Currently, the wet-percent-time table utilized by Caltrans provides one estimate for the percentage of time that pavement is wet for each county throughout the state. Respondents were asked whether they agreed with the use of just this one factor (i.e. percent wet time factor) for an entire county. The respondents were asked to rank their agreement with this statement in a Likert scale of 1 (Strongly agree) to 5 (Strongly disagree). An option to select "Other" was also provided. The 23 participants who answered the question provide the following responses:

Table 11.3: User views toward use of one factor

Response	No. of Responses	Response Percentage
Strongly Agree	0	0.00%
Agree	11	47.83 %
Neutral	7	30.43%
Disagree	1	4.35%
Strongly Disagree	3	13.04%
Other	1	4.35%

As the results indicate, nearly half of respondents believed that 1 wet percent time factor per county was adequate. Nearly one-third of respondents were neutral as to whether such a singular factor should be employed for an entire county. Five of the remaining respondents believed that the use of only one factor per county was not advisable. The respondent who fell into the "Other" category expressed that "San Bernardino County is so different that one estimate cannot cover the entire county adequately."

Based on the responses obtained to this question, it would appear that there are essentially two schools of thought with respect to the issue. There is the school that believes the use of one factor per county is entirely appropriate. Then there is the school (including some of those who were neutral to the issue) that believes more than one factor is needed, based on the characteristics of the county itself.

Note 4: Perceived Improvements

To obtain the views of Wet Table users with respect to updating the table (either to bring its percentages up to date or to provide more than one percentage for a county), respondents were asked to rank a series of possible improvements. The improvements to be ranked included:

- Update the current wet percent time table (updating its percentages which are now 30 years old);
- Modify the wet percent time table to a better geographical unit than a county based table (i.e. one value for one county);
- Update the wet percent time table every year to better represent the time the pavement was wet;
- None of the above.

Respondents were asked to rank these improvements from 1 to 4, with 1 representing the most desirable option. Twenty two respondents ranked the first three options, with only seven ranking the final option (None of the above). This would suggest that all respondents believed some change was necessary to the wet table percentages; however, their opinion as to the importance of the various possible improvements varied, as seen in the chart below.

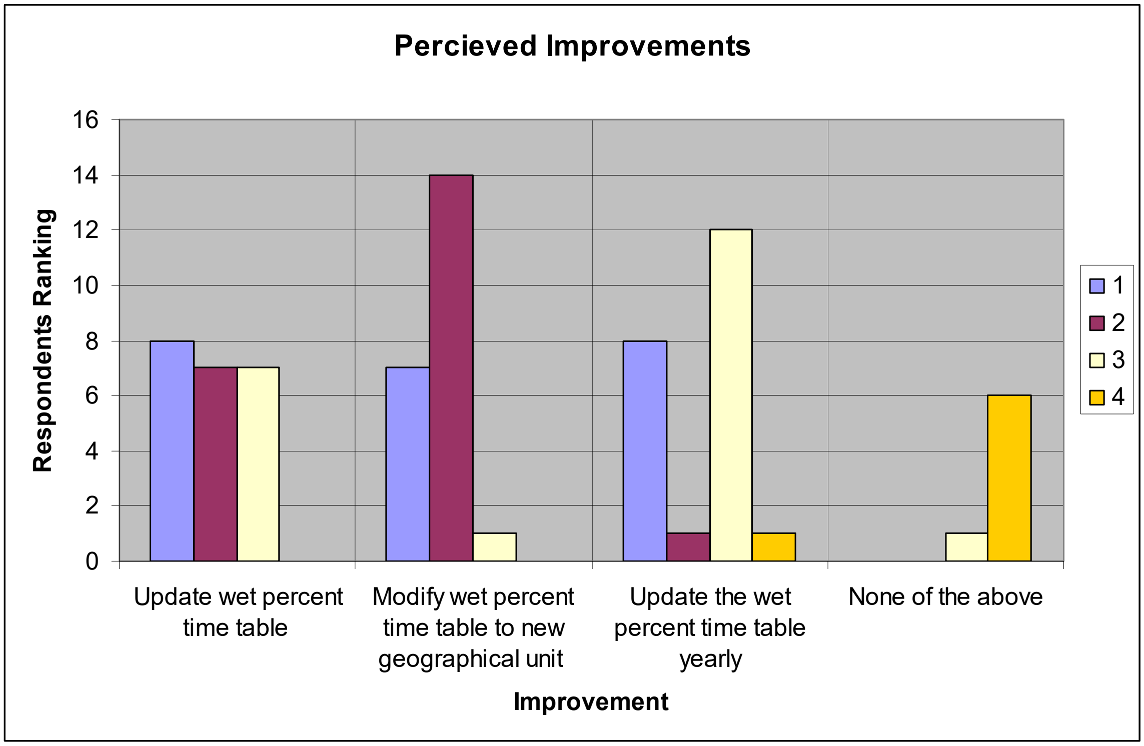

Figure 11.2: Rankings of perceived improvements to Wet Table

As the chart indicates, updating the wet percent time table was viewed to have varying importance to users, with approximately one-third of respondents ranking this improvement as first, second or third most important. This would suggest that users see the need for an update to the table, but most users don't view a one-time update as being the most important improvement which can be made.

Modification of the wet percent time table to cover a new geographic unit instead of a county was viewed by seven respondents as being the most important improvement that could be made. However, 14 respondents viewed this as being the second most important improvement which could be made, suggesting that the inclusion of a geographic component is something a majority of wet table users see value in.

Updating the wet percent table on a yearly basis produced mixed opinions in terms of its importance. Eight respondents ranked this improvement as first, while twelve ranked it as third. Once again, this would suggest that there are those who recognize the need for an update to the table, there is just not a consensus as to whether it should be a once-a-year occurrence or something updated over longer intervals.

Only seven respondents said none of the previous selections were important. Not surprisingly, these rankings were in the third and fourth positions only. It is believed that the responses, primarily those ranking the option as fourth, were simply made to complete the series. As such, the results of the "None" category should be considered with caution in comparison to the rankings of the other three improvement options available.

Note 5: Geographic Unit

As a follow up to the previous questions, respondents were asked what an appropriate geographic unit would be, aside from a county. Geographic units included:

- Highway milepost;
- Geographical zone based on precipitation data;
- Geographical zone based on precipitation and traffic volume data;
- None of the above.

Respondents were asked to rank these options from 1 to 4 (with 1 being the most desirable). The number of respondents who ranked each option varied, but most respondents ranked at least three of the choices (most frequently leaving "None of the above" unranked). The chart below illustrates the responses obtained in terms of rankings.

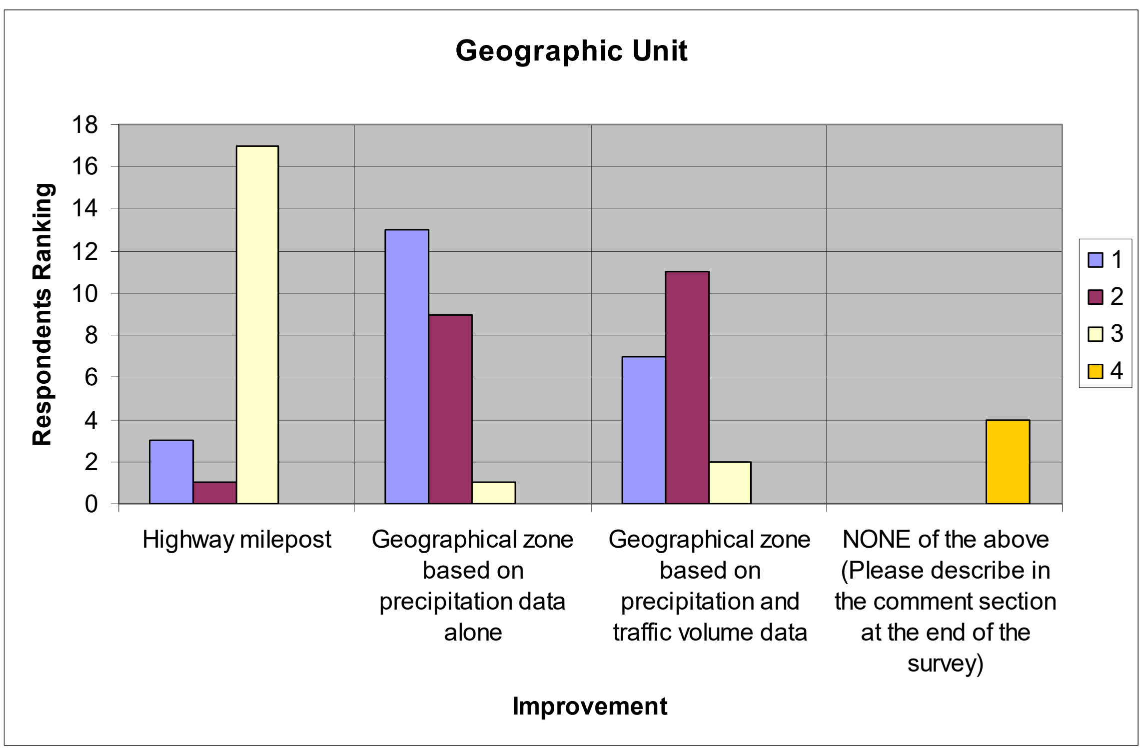

Figure 11.3: Rankings of geographic improvements to Wet Table

In ranking the highway milepost option, most respondents (17) viewed this as the third best improvement that could be made. Only three respondents viewed this as being the best option, while one viewed it as the second best improvement. When examining the rankings of the geographical zone based on precipitation data option, respondents again viewed this as being most desirable, with thirteen ranking this improvement first. An additional nine respondents ranked this as the second best improvement. Only one respondent ranked this improvement third. Clearly respondents saw some value in adding a geographic component to the Wet Table.

The creation of a geographical zone based on precipitation and traffic data received a top ranking from seven respondents. Eleven respondents viewed this option as the second best improvement available, while two ranked it as the third As an aside, four respondents ranked the "None of the above" option as the least desirable option. This result should be viewed with caution, as it is believed that the respondents simply ranked this option to complete the series.

Note 6: Additional Information

Ten respondents provided additional comments and information. These included the following:

- There should be a correlation between the geographical annual precipitation for an area and the route. Some routes in San Diego County begin near the ocean, continue east over the mountains, and terminate in the desert. The volume of vehicles on a specific route should also be taken into consideration.
- As a table C user for more than 10 years, I can identify locations from the table C all that require wet related improvements. The table C wet is not needed.
- How about Fog and Fog Drip? A redwood tree can drip 6 inches of water in a night during heavy fog but this would not show up as precipitation.
- Wet table C should always be one of the factors considered in all investigation.
- Wet table C should be considered at all times.
- The percent wet time is currently based on 30 year old data and may not reflect changes in weather patterns. The current data also does not reflect the fact that some counties extend from the valley floor to the top of the Sierra Nevada Mountains and have a significant variation in percent wet time within a particular county.
- The data the table is based on should be as up to date as possible.
- Having the percent wet time validated and changed based on location should help with identifying locations that are currently not being picked up and on the flip side, eliminate locations that should not be on the wet table c.
- San Bernardino County is so vast that one unit of rain calculation cannot possibly cover the entire county adequately. We have such variables as mountainous

terrain, desert terrain, inland valley terrain, all of which have different weather patterns and different patterns of precipitation.

- Need a justified number of accidents (concentrations) than the required within a shorter distance.

Wet Table C Users Survey

Introduction

Dear Wet Table C User: This survey is being conducted by Western Transportation Institute (WTI) / Montana State University for Caltrans. WTI has been contracted by Caltrans to update the percent wet time table used in the annual development of Wet Table C. The answers you provide will only serve as guidance for our updating the percent wet time table and will not be used for any other purpose. You have been identified as a Wet Table C user with the help of Division of Traffic Operations at Caltrans Head Quarters. Please feel free to call (406) 994-7909 or email me at mkumar@coe.montana.edu. If you choose none of the above as an answer to any of the following questions, please describe why chose none of the above in the comment and suggestions section at the end of the survey. Thanks.

1. Please order the following reasons for “No Action” after investigating “Required” location in Wet Table C starting with the most frequent reason and using a number between 1 and 4.

   ☐ NOT a Wet pavement / grooving related safety problem

   ☐ Peak hour congestion related collisions

   ☐ Increased traffic volume during seasonal peaks

   ☐ The location is repeated or overlaps with another location from last year's Wet Table C

   ☐ NONE of the above (Please describe in the comment section at the end of the survey)

   Rank values must be between 1 and 5

2. How well does Wet Table C identify only the locations that need improvement based on your experience in investigating the “Required” locations in Wet Table C?

   1. More than adequate
   2. Adequate
   3. Not adequate
   4. Other (Please Describe) ____

3. Currently the wet percent time table provides one estimate of percent wet time (i.e. percent wet time factor) for each county. Do you agree with the use of one factor for one county?

   1. Strongly agree
   2. Agree
   3. Neutral
   4. Disagree

5. Strongly disagree
6. Other (please describe) ____

4. Please order the following improvements that need to be made with wet percent time table, in your opinion using a number between 1 and 4. Currently, the wet percent time table provides an estimate of percentage of time the pavement was wet in a year for each county (i.e. one estimated value for one county)

☐ Update wet percent time table (current wet percent time table was developed using data that are 30 years old)

☐ Modify the wet percent time table to a better geographical unit than a county based table (i.e. on value for one county)

☐ Update the wet percent time table every year to better represent the time the pavement was wet

☐ NONE of the above (Please describe in the comment section at the end of the survey)

Rank values must be between 1 and 4

5. If it is decided to change the basis for the percent wet time factor from County to another geographical unit, which of the following you think will be appropriate (please order the following starting with the highest priority using a number between 1 and 3)?

☐ Highway milepost

☐ Geographical zone based on precipitation data alone

☐ Geographical zone based on precipitation and traffic volume data

☐ NONE of the above (Please describe in the comment section at the end of the survey)

Rank values must be between 1 and 4

6. Please write down your concerns and suggestions below as related to the percent wet time used in the development of Wet Table C

APPENDIX C: NATIONAL SURVEY

Note 1: Reduction of Wet Accidents

As shown in the chart below, there was a fairly even split between respondents from states that did not have a specific focus on reducing wet-accidents (18) and those that did (17). Only four of the respondents were unsure whether their state had such a focus.

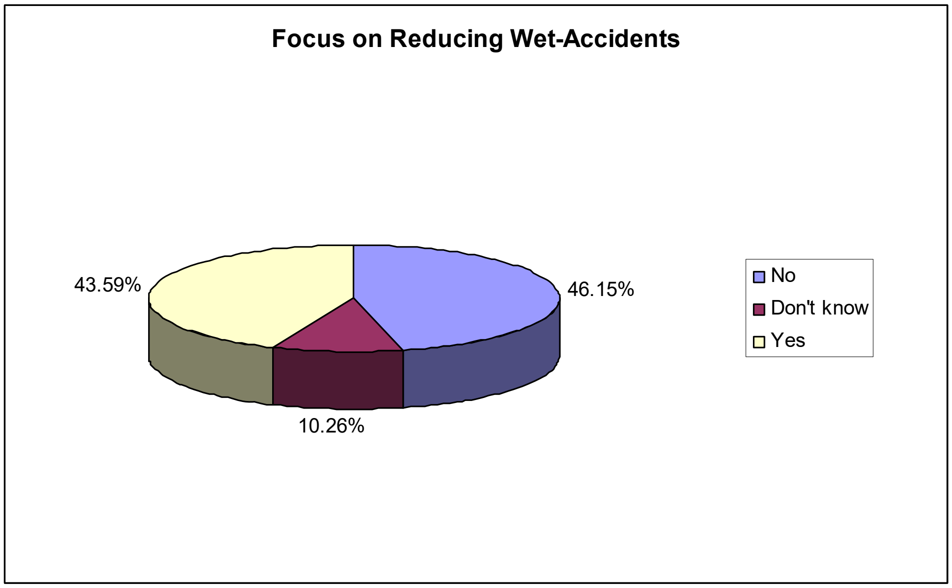

Figure 11.4: Focus of states on wet-accident reduction

Respondent states that replied they did have a systematic focus on wet-accident reduction were asked to provide further details of their programs. Activities related to such programs ranged from simply confirming that wet weather crash analysis was conducted to making use of skid test pavement friction data (nine states reported making use of such data). Specific responses included:

- Wet Weather Crash analysis (performed).
- We perform ASTM standard skid tests on all state maintained roadways on an annual basis and as requested. We review crashes for locations with notable high percentages of wet pavement or rear end crashes. By comparing the two lists of locations we can identify locations where improvements that address the increase in the friction number may be beneficial. Depending on the magnitude of the problem and project cost, any funding category may be used to make the improvement. Generally these are funded through Resurfacing, Restoration, or Rehabilitation (RRR or 3R) projects or the Highway Safety Improvement Program.
- Our Highway Safety Improvement Program annually identifies spot locations around the state that meet or exceed criteria for wet road condition crashes.
- We have a Friction program that lets us know where locations appear with Friction numbers below 30. When spots are identified, our staff review wet

crashes at these spots. If there is a problem, safety funding is available to fix the spot.

- Identify locations on the state highway system were the percentage of wet road crashes exceeds the expected occurrence by facility type. Identified locations are then skid tested and those exhibiting low friction numbers are addressed through various surface treatments.
- Wet pavement crash reduction policy letter and wet pavement crash clusters.
- Blowing Snow Team to ID slick road areas and solutions.
- We are currently developing this program.
- We merge our wet weather crash data with below threshold skid coefficient numbers.
- We select the 15 to 20 locations with the highest wet/total accident ratio. We visually review the locations for signage, pavement condition, geometrics, etc. to see what might be suggested to reduce the number of accidents.
- Wet weather crash cluster sites are determined via accident data and high friction asphalt is installed as well as other counter measures considered.
- Annual skid testing to improve wet crash locations.
- We identify clusters of 10 or more wet surface crashes within a half-mile section. We then merge the data from these sections with pavement data which contains friction number. Any of these sections with a wet surface crash experience of double our statewide average for wet surface crashes and a friction number of 35 or less is considered a priority location.
- The program is based on a regular inventory of skid testing and requests in areas where staff suspect problems.
- Locations identified through our HES (Hazard Elimination Safety) program.
- Use skid resistant pavement on all roadways with Average Daily Traffic (ADT) greater than or equal to 3000. Roadways with lower volumes are reviewed.
- We have a program of collecting friction numbers, which is done by the Research Division.

As the responses indicate, states take a number of different approaches to examining wet accidents. The most common response shows that some use of friction data is made, either during the crash analysis itself or after as the result of high crash locations during wet weather being identified. For those states that do not use friction data, the analysis appears to focus on identifying crashes which have occurred during wet weather and establishing whether these exceed expected thresholds.

A limited number of the responses to the question elaborated on specific treatments which were utilized to address wet-accidents once high concentrations had been identified. One respondent state was more proactive and used high-friction pavements when ADT exceeded a given threshold. Other states addressed the problem retroactively by programming 3R work or specific safety treatments to be applied once a site had shown a vulnerability to wet-accidents.

Note 2: Other Programs Including Wet-Accident Analysis

Expanding on the previous question, participants were asked whether their state had any other programs that might include periodic wet-accident analysis. As the chart below indicates, nearly half of the states that responded to this question had no additional analysis programs which identified wet-accidents. Despite this, at least one-third of those who responded to the question did have some program(s) where such accidents were identified.

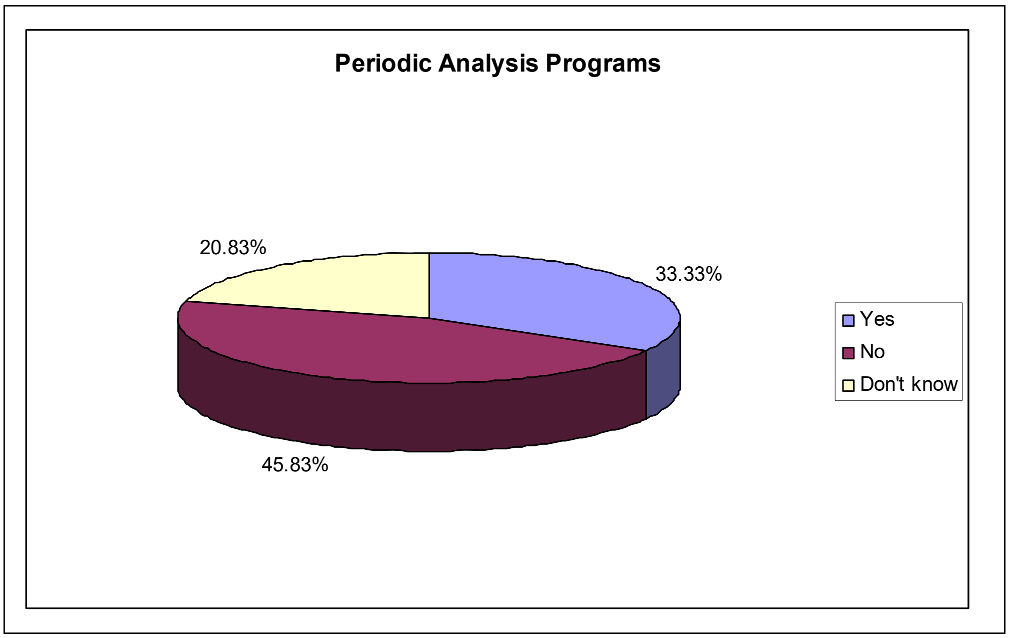

Figure 11.5: Existence of additional wet-accident analysis

The states with additional programs were asked to specifically list these, and their responses included the following:

- Analysis of traffic crash data.
- As part of overall crash analysis of a corridor or specific location.
- Project specific when data identify a location.
- Standard in-house collision concentration listings.
- Occasional analysis on request.
- Safety assessments on 3R projects.
- Run crash queries involving wet weather crashes.
- Winter crash analysis.

As these responses indicate, the programs that identify wet-accidents widely vary. These range from project specific identification (3R projects, for example), to spot location analysis performed on request, to analysis conducted specifically on winter crashes. In essence, these programs are accomplishing the same goal as specific wet-accident analysis programs; the primary difference is that they fall under a different umbrella which is sometimes project or event specific.

Note 3: Accident Report Classification

The key to identifying wet-accidents is the availability of accurate accident data. To this end, the survey asked participants if their state classified wet-accidents from accident reports (also referred to as police reports). As the table below indicates, all but two of the respondent states have some type of pavement condition record in the accident report file collected by the police. When the recorded wet condition becomes more specific (e.g. when the condition is refined to state snow or ice conditions) the number of states with such information available decreased by over half, to 15. However, this does not mean that other responding states do not break their wet pavement condition requirements down further; rather, these responses reflect the best knowledge of the personnel completing the survey.

Table 11.4: State accident reporting conditions

Reporting Condition	No. of Respondents	Response Percentage
Accidents that occurred when the pavement conditions were wet	37	94.87%
Accidents that occurred when the pavement was snowy or icy	15	38.46%
Don’t know	2	5.13%
Other (please describe)	3	7.69

Three respondent states recorded “other” wet pavement related information. Descriptions of what they record included:

- We have several selections under the category Roadway Surface Condition that the reporting officer selects.
- We have five categories: Wet, Snow/Slush (a: sanded, b: unsanded), Ice/Packed Snow (a: sanded, b: unsanded).
- Roadway flooded.

The first two responses illustrate that the respondent states use a number of different classifications for the condition of pavements at the time of an accident. The listing of roadway flooded by the third responding state indicates the type of specific event which other states include in their reporting criteria.

Note 4: Number of Wet Accidents

When asked for their opinion as to whether the state they represented experienced a significant number of wet weather accidents, responses varied widely. A fair percentage of respondents believed their state did have a wet weather accident problem. Such respondents typically represented states that experience significant amounts of precipitation (e.g., the Pacific Northwest). Similarly, those who did not believe their state experienced a problem with wet weather accidents were from states where little precipitation occurs (e.g. the Southwest). Those respondents who were unsure as to whether their state had such a problem were from areas where weather tended to vary, such as the Midwest. One respondent provided information in a previous question stating that of the approximately 320,000 crashes in their state per year, roughly 52,000 were weather related.

Table 11.5: Respondent perception of wet accident problem

Significant number of crashes	No. of Responses	Response Percentage
Yes	14	35.90%
No	15	38.46%
Don’t know	10	25.64%

Note 5: Annual Traffic Safety Report

Respondents were asked whether their state’s annual traffic safety report included a list of high concentration wet-accident locations. As the table below illustrates, few states include such a list in their publications. This is understandable, as annual safety reports typically focus on the big picture, rather than events that are often site specific. Interestingly, the six respondents who reported their state does compile such lists were located in the Southeast and Northeast, rather than areas like the Northwest where wet weather accidents may be more common.

Table 11.6: Inclusion of wet accident information in safety reports

Inclusion in annual safety report	No. of Responses	Response Percentage
Yes	6	15.38%
No	29	74.36%
Don't know	4	10.26%

Note 6: Additional Reports Containing Wet Weather Accident Information

Survey respondents were also asked if any additional reports produced by their state contained listings of high concentration wet-accident locations. Thirty-three states responded to the question, with the results presented in the table below.

Table 11.7: Inclusion of wet accident information in non-safety reports

Inclusion in other annual reports	No. of Responses	Response Percentage
Yes	4	12.12%
No	24	72.73%
Don't know	5	15.15%

As the results indicate, the majority of states that responded did not include information on high wet-weather accident locations in any reports aside from an annual safety report (if they do in fact include such information in that report). This once again is not surprising given the specific nature of such crashes and the time required to reduce accident data down to such a level of detail. The four states that responded "yes" with respect to including locational lists in additional reports mainly included such information in publications related to specific corridors or sections throughout their state.

Note 7: Additional Information

Respondent comments included the following:

- Low friction areas are cross referenced with wet weather crash locations above 30%.
- Isobar charts of historic percent of time that the pavement is wet are used to help identify over representations of wet crashes.
- We can retrieve wet pavement data fairly easily [Response from state that did not respond to questions related to wet pavement exposure].
- We recently discontinued our Wet Weather Accident Analysis. Our previous criteria was a minimum of 10 wet crashes in a 3 year period and a Wet/Dry ratio >= 0.33 We ran two reports, one for intersections and one for 1 mile segments. We discontinued the report because we shifted our safety program to locations with fatal and disabling crashes. We are considering re-establishing our Wet/Dry program (but a final decision has not been made).
- We do friction testing in our State in a three year cycle.
- Our biggest issue is wintry roads. The first onset of real winter weather in December is when safety performance is at its worst and gets slightly better in January and February.

- Wet-accident data is available in our crash database. This data is considered during our highway safety studies. There does not appear to be any systematic use of crash data for wet weather studies.
- We are mostly concerned with icy/snowy road conditions. Most of the current efforts are directed toward this area of weather related problems.
- We are currently developing a comprehensive plan for all weather related crashes, including the development of Road Weather Safety Audit Program.
- Our methodology is basic and is as follows:
    1. a) Identify location of wet weather crashes - at & between intersections
       b) Identify Skid numbers locations
    2. Rank merge skid data and crash data (table and GIS)
    3. Rank sites
       a) Assess dollar costs to crash severity
       b) Cost to fix – reduction of 50%
- We have a skid trailer and do skid testing when requested by our engineers in locations suspected of having a low wet-weather skid resistance. We have the capability of creating reports showing where collisions occur when the roadway is wet, but this has not been a high priority in recent years.
- We receive relatively little rainfall, so wet weather crashes are not our biggest issue. About 7.8% of all crashes in 2005 were on wet pavement, including about 5.7% of fatal crashes.
- We conduct a statewide pavement surface friction inventory using a locked wheel friction tester conforming to ASTM E-274. We use the E-501 ribbed test tire and conduct the test wet by spraying water on the pavement during the test cycle. Thus the surface friction inventory values are “wet pavement” values.
- We get reports from District Maintenance personnel and occasionally from law enforcement about problem areas. Generally our response is to install “Slippery When Wet” signs and alert Pavement Management so that they can consider the area for resurfacing.
- The Pavement Management group looks at skid testing results and when there are low numbers in an area they will look at crash data to see if the wet weather or skidding crashes are over represented in the data. In addition, they monitor known problem areas, before and after pavement treatments.
- The accident reports have two categories—surface conditions and Weather. Reports show surface conditions as “wet” for accidents that happened during wet weather. Also, under weather information it is noted if it was snowing or rain. No specific analysis has been done relative to wet weather accidents.

- We analyze all accident types in our Safety Assessment Process on 3R projects which include wet pavement analysis by looking for patterns of accidents.
- We just began analysis of snow and ice crashes in the state.

APPENDIX D: NETWORK SAMPLE DATE FILES

Sample data from NCDC

722920 23191 200403030851 290 10 *** 722 CLR * * * 10.0 00 ** ** * 48 39 1011.8 29.87 ****** *** *** ***** ***** ***** ***** **

722920 23191 200403030951 290 10 *** 722 CLR * * * 10.0 00 ** ** * 50 37 1011.6 29.87 ****** *** *** ***** ***** ***** ***** **

722920 23191 200403031051 300 9 *** 722 CLR * * * 10.0 00 ** ** * 50 39 1011.3 29.87 ****** *** *** ***** ***** ***** ***** **

722920 23191 200403031151 290 9 *** 722 CLR * * * 10.0 00 ** ** * 50 39 1011.3 29.87 ****** 51 49 ***** ***** ***** ***** **

722920 23191 200403031251 *** 0 *** 722 CLR * * * 10.0 00 ** ** * 52 37 1011.5 29.87 ****** *** *** ***** 0.00 ***** ***** **

722920 23191 200403031351 *** 0 *** 722 CLR * * * 10.0 00 ** ** * 54 34 1012.0 29.89 ****** *** *** ***** ***** ***** ***** **

722920 23191 200403031451 160 7 *** 722 CLR * * * 10.0 00 ** ** * 54 36 1012.4 29.90 ****** *** *** ***** ***** ***** ***** **

722920 23191 200403031551 120 3 *** 722 CLR * * * 10.0 00 ** ** * 57 36 1013.0 29.92 ****** *** *** ***** ***** ***** ***** **

722920 23191 200403031651 280 3 *** 722 CLR * * * 10.0 00 ** ** * 59 36 1013.1 29.93 ****** *** *** ***** ***** ***** ***** **

722920 23191 200403031751 270 7 *** 722 CLR * * * 10.0 00 ** ** * 61 41 1013.2 29.94 ****** *** 50 ***** ***** ***** ***** **

722920 23191 200403031851 260 7 *** 722 CLR * * * 10.0 00 ** ** * 61 37 1013.3 29.94 ****** *** *** ***** 0.00 ***** ***** **

722920 23191 200403031951 270 9 *** 722 CLR * * * 10.0 00 ** ** * 61 37 1012.5 29.92 ****** *** *** ***** ***** ***** ***** **

Sample data from CDEC

Title: "ACN.csv"

1781	PST	'PRECIPITATION	ACCUMULATED (inches)'
19940701	0	6.6	
19940701	100	6.6	
19940701	200	6.6	
19940701	300	6.6	
19940701	400	6.6	
19940701	500	6.6	
19940701	600	6.6	
19940701	700	6.6	
19940701	800	6.6	
19940701	900	6.6	
19940701	1000	6.6	
19940701	1100	6.6	
19940701	1200	6.6	
19940701	1300	6.6	
19940701	1400	6.6	
19940701	1500	6.6	
19940701	1600	6.6	
19940701	1700	6.6	
19940701	1800	6.6	

Sample data from MESOWEST

CMNC1	991013/1000	0
CMNC1	991013/1100	0
CMNC1	991015/1000	0
CMNC1	991015/1100	0
CMNC1	991015/2000	0
CMNC1	991016/1100	0
CMNC1	991017/0800	0
CMNC1	991017/0900	0
CMNC1	991017/1000	0
CMNC1	991017/1100	0
CMNC1	991017/1200	0
CMNC1	991017/1300	0
CMNC1	991017/2000	0
CMNC1	991020/0700	0
CMNC1	991020/0800	0
CMNC1	991020/1700	0
CMNC1	991022/0800	0
CMNC1	991022/0900	0
CMNC1	991023/0900	0
CMNC1	991023/1000	0
CMNC1	991023/1100	0
CMNC1	991024/1100	0

Sample data from CIMIS

Stn Id	Station	Region	Date	Hour	Jul	qc	Precip (in)
79	Angwin	North Coast Valleys	5/11/1989	100	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	200	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	300	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	400	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	500	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	600	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	700	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	800	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	900	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1000	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1100	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1200	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1300	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1400	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1500	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1600	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1700	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1800	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	1900	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	2000	131	*	0
79	Angwin	North Coast Valleys	5/11/1989	2100	131	*	0

Sample data from NWS

COOPID,STATIONNAME,CD,ELEM,UN,YEAR,MO,DA,TIME,HOUR01,F,F,TIME,
HOUR02,F,F,TIME,HOUR03,F,F,TIME,HOUR04,F,F,TIME,HOUR05,F,F,TIME,HOU
R06,F,F,TIME,HOUR07,F,F,TIME,HOUR08,F,F,TIME,HOUR09,F,F,TIME,HOUR10,
F,F,TIME,HOUR11,F,F,TIME,HOUR12,F,F,TIME,HOUR13,F,F,TIME,HOUR14,F,F,T
IME,HOUR15,F,F,TIME,HOUR16,F,F,TIME,HOUR17,F,F,TIME,HOUR18,F,F,TIME,
HOUR19,F,F,TIME,HOUR20,F,F,TIME,HOUR21,F,F,TIME,HOUR22,F,F,TIME,HOU
R23,F,F,TIME,HOUR24,F,F,TIME, TOTAL,F,F

------,------------------------------,--,----,--,----,--,--,----,------,-,-,----,------,-,-,----,------,-,-,----
,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,--
----,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,-----
-,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-,-,----,------,-
,-

047633,SACRAMENTO WSO CITY ,00,HPCP,HI,1936,01,01,0100, 00000,g,
,0200, 00000, , ,0300, 00000, , ,0400, 00000, , ,0500, 00000, , ,0600, 00000, , ,0700,
00000, , ,0800, 00000, , ,0900, 00000, , ,1000, 00000, , ,1100, 00000, , ,1200, 00000, ,
,1300, 00000, , ,1400, 00000, , ,1500, 00000, , ,1600, 00000, , ,1700, 00000, , ,1800,
00000, , ,1900, 00000, , ,2000, 00000, , ,2100, 00000, , ,2200, 00000, , ,2300, 00001, ,
,2400, 00000, , ,2500, 00001, ,

047633,SACRAMENTO WSO CITY ,00,HPCP,HI,1936,01,02,0100, 00003, ,
,0200, 00008, , ,0300, 00009, , ,0400, 00004, , ,0500, 00005, , ,0600, 00000, , ,0700,
00001, , ,0800, 00000, , ,0900, 00001, , ,1000, 00002, , ,1100, 00008, , ,1200, 00008, ,
,1300, 00011, , ,1400, 00008, , ,1500, 00008, , ,1600, 00003, , ,1700, 00001, , ,1800,
00000, , ,1900, 00001, , ,2000, 00001, , ,2100, 00002, , ,2200, 00001, , ,2300, 00000, ,
,2400, 00001, , ,2500, 00086, ,

047633,SACRAMENTO WSO CITY ,00,HPCP,HI,1936,01,04,0100, 00000, ,
,0200, 00000, , ,0300, 00003, , ,0400, 00002, , ,0500, 00002, , ,0600, 00000, , ,0700,
00000, , ,0800, 00000, , ,0900, 00000, , ,1000, 00000, , ,1100, 00000, , ,1200, 00000, ,
,1300, 00000, , ,1400, 00000, , ,1500, 00000, , ,1600, 00000, , ,1700, 00000, , ,1800,
00000, , ,1900, 00000, , ,2000, 00000, , ,2100, 00000, , ,2200, 00000, , ,2300, 00000, ,
,2400, 00000, , ,2500, 00007, ,

047633,SACRAMENTO WSO CITY ,00,HPCP,HI,1936,01,07,0100, 00000, ,
,0200, 00000, , ,0300, 00000, , ,0400, 00000, , ,0500, 00000, , ,0600, 00000, , ,0700,
00000, , ,0800, 00000, , ,0900, 00000, , ,1000, 00000, , ,1100, 00000, , ,1200, 00000, ,
,1300, 00000, , ,1400, 00000, , ,1500, 00000, , ,1600, 00000, , ,1700, 00000, , ,1800,
00004, , ,1900, 00000, , ,2000, 00000, , ,2100, 00000, , ,2200, 00000, , ,2300, 00000, ,
,2400, 00000, , ,2500, 00004, ,

APPENDIX E: WET PERCENT FACTOR DIFFERENCES

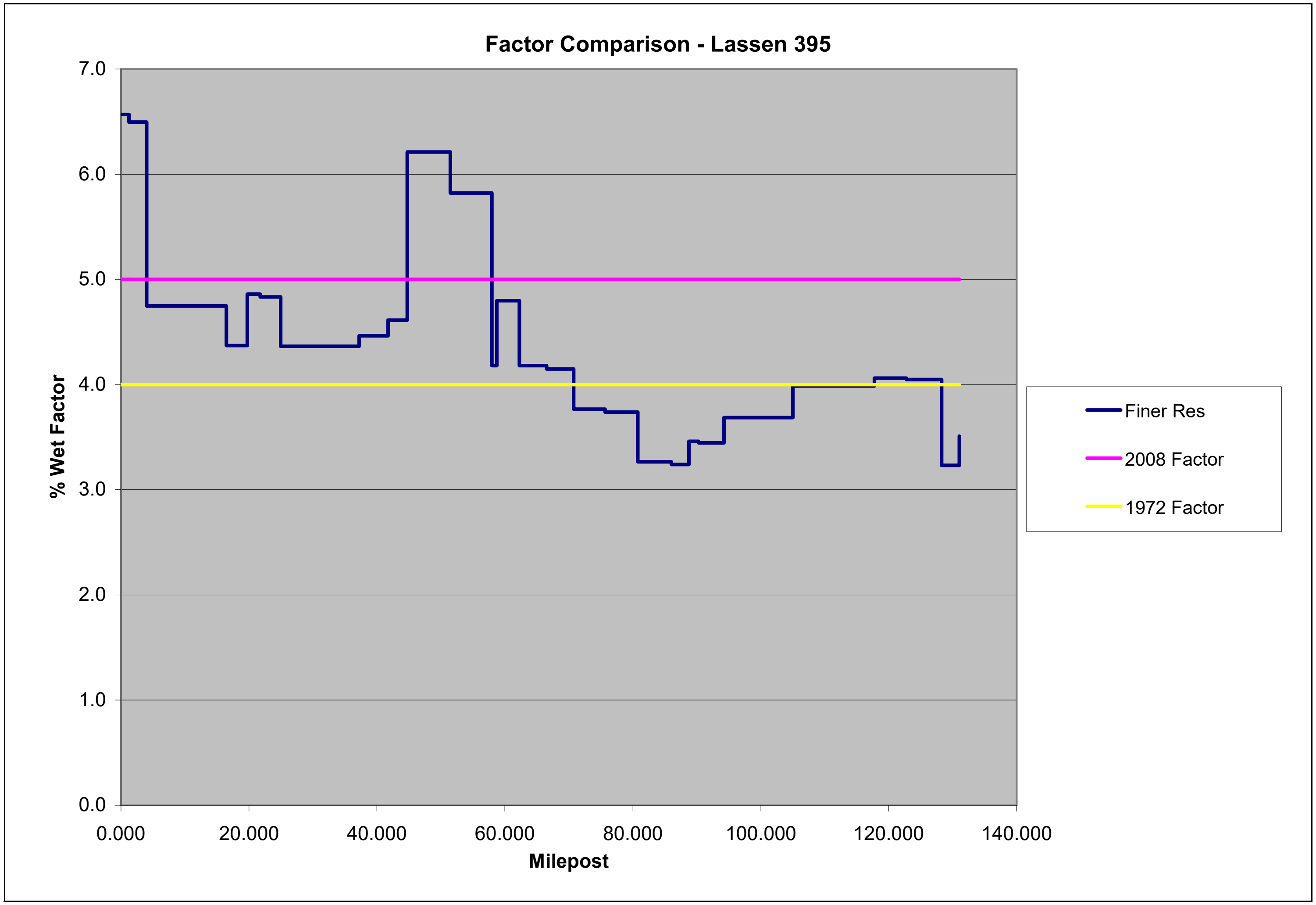

Figure 11.6: Differences between wet percent factors, Lassen 395

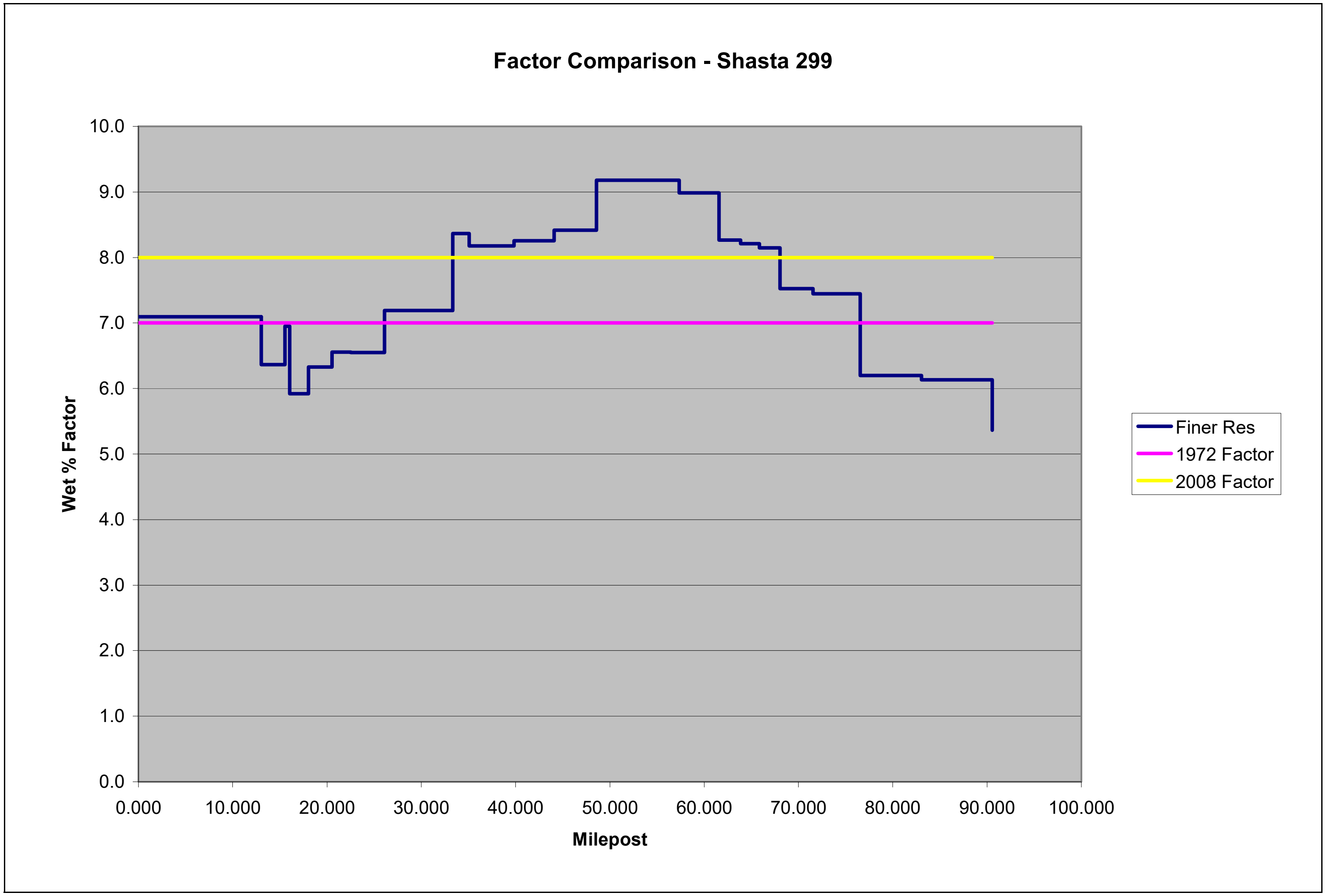

Figure 11.7: Differences between wet percent factors, Shasta 299

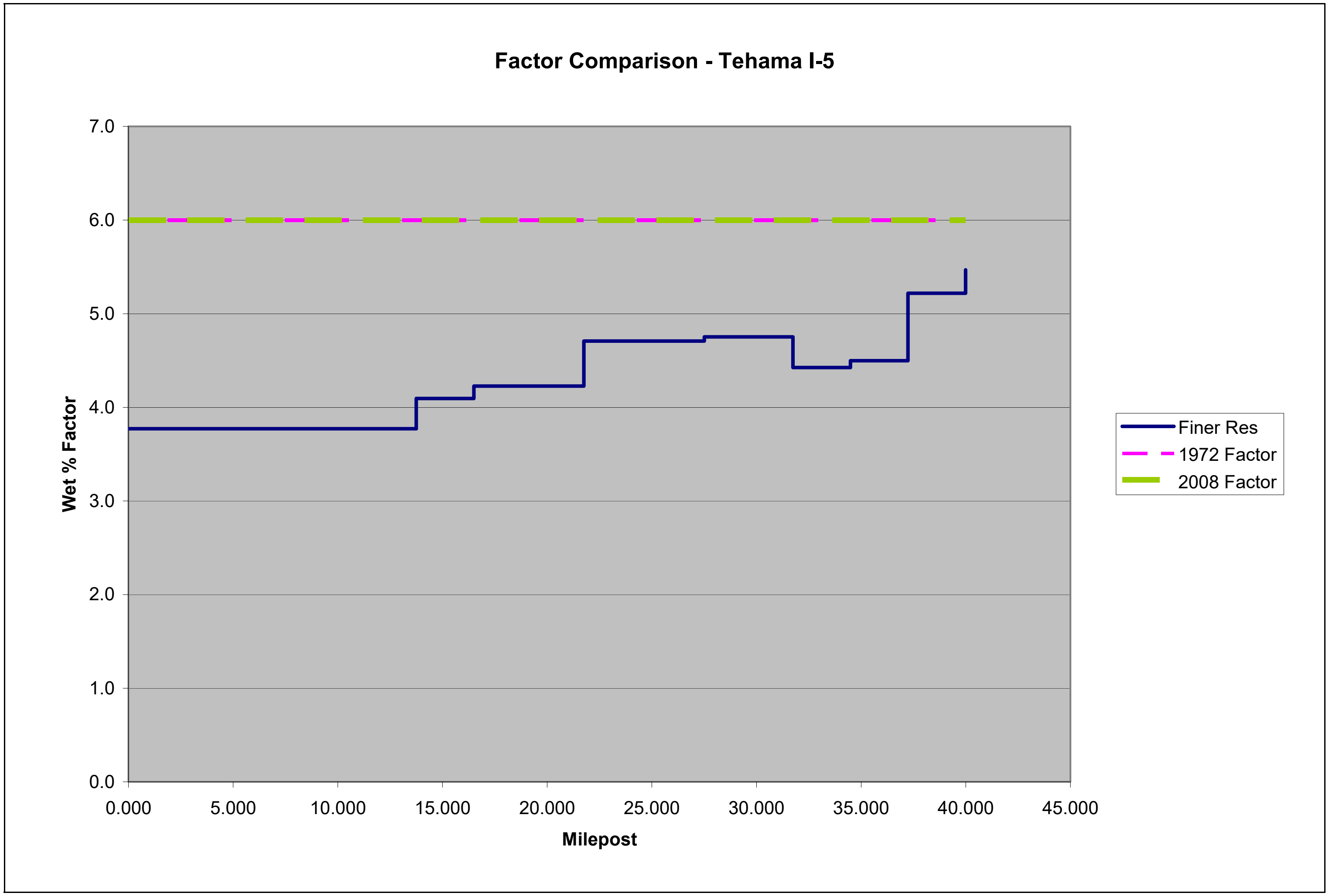

Figure 11.8: Differences between wet percent factors, I-5, Tehama County

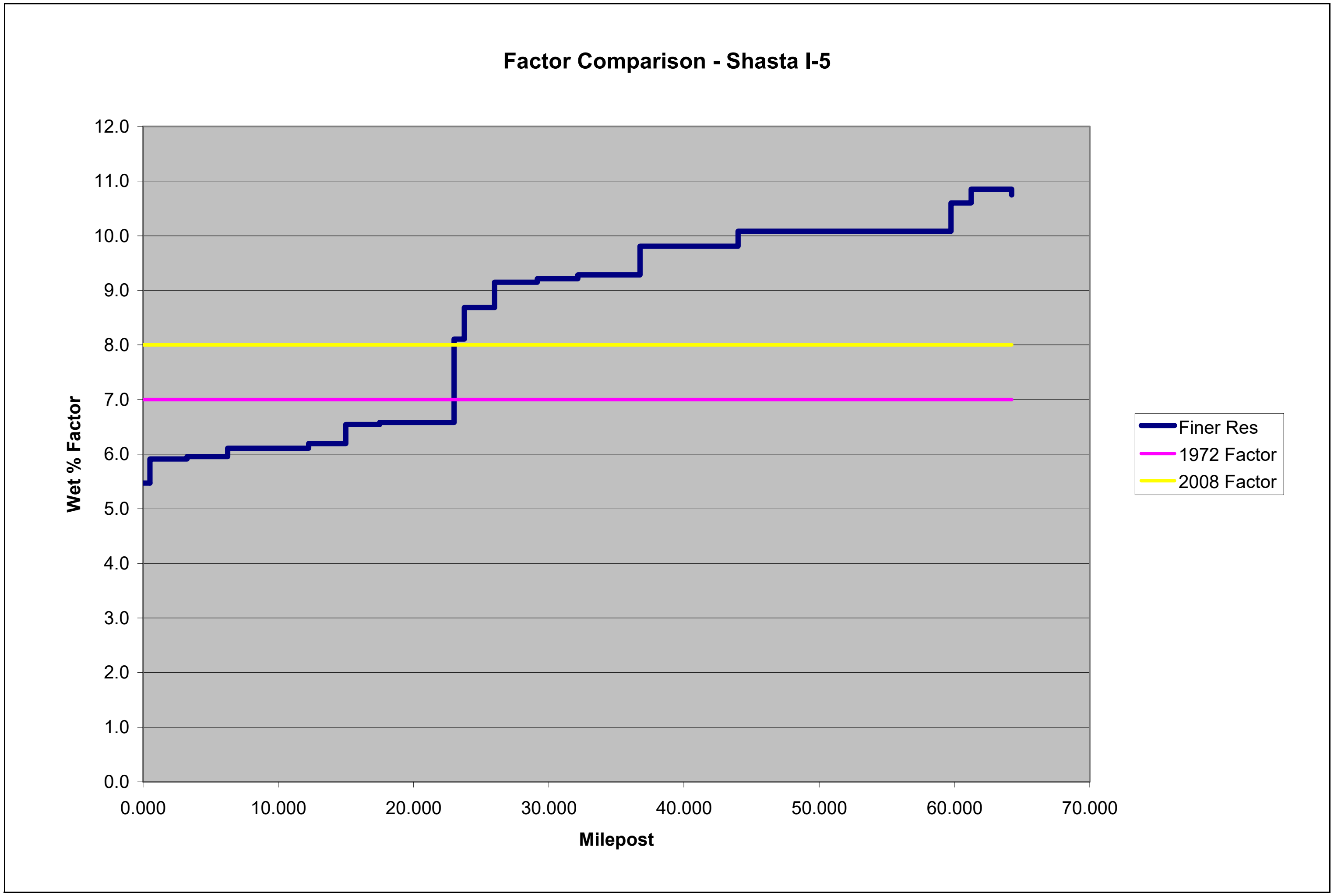

Figure 11.9: Differences between wet percent factors, I-5, Shasta County

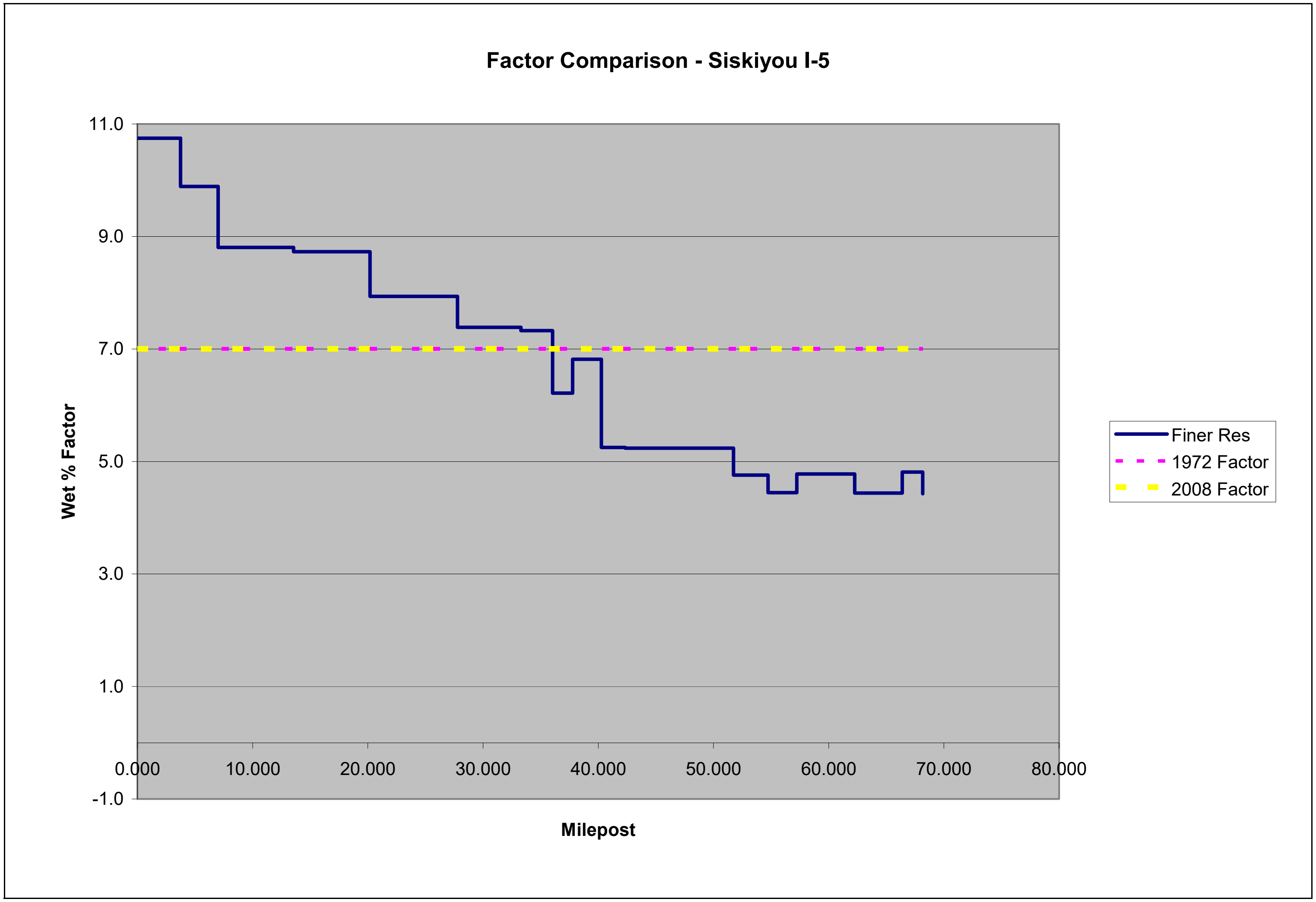

Figure 11.10: Differences between wet percent factors, I-5, Siskiyou County

APPENDIX F: WET TABLE C LISTS

Location Description	SCL RMP LNS	R U S	Rate Grp	Total Accidents / Significance: 36 mo. ACCS	24 mo. ACCS	12 mo. ACCS	6 mo. ACCS	3 mo. ACCS	AVE ADT: Main	X-St.	36 MOS AVERAGE ACC & RATES: NO OF ACC: F+I	TOT	RATE/MVM-MV: F+I	TOT	Req
003 TRI 60.19 SWIFT CR RD-LT (FS TAIL	---	R	I 17	2 N	2 Y	1 N	0 N	0 N	0.7	.1	0.03	0.06	0.36	0.82+	
005 TEH R 24.712 005/NB ON FR SO RDBL	O G	U	R 24	4 Y	3 N	0 N	0 N	0 N	2.8	-	0.18	0.49	0.99	2.65+	
005 SHA .781 005/SBON GAS POINT RD/4TH	O D	R	R 11	3 Y	2 N	1 N	0 N	0 N	2.2	-	0.11	0.30	0.66	1.77+	
005 SHA R003.320 TO R003.520 SOUTH	02D	S	H 60	4 N	4 Y	3 Y	2 N	2 N	25.6	-	0.31	0.84	0.79	2.13	Req
005 SHA R006.650 TO R006.850 SOUTH	04D	S	H 60	3 N	3 N	3 Y	3 Y	1 N	25.4	-	0.31	0.83	0.79	2.13	Req
005 SHA R 15.265 005/NB OFF TO HILLTOP	F H	U	R 26	9 Y	7 Y	2 N	1 N	1 N	5.2	-	0.47	1.37	1.17	3.43+	Req
005 SHA R 15.391 005/SB OFF TO EB 44	F L	U	R 70	6 Y	4 N	1 N	1 N	1 N	5.5	-	0.33	1.16	0.79	2.74+	
005 SHA R015.300 TO R015.500 SOUTH	02D	U	H 63	10 Y	5 Y	0 N	0 N	0 N	32.3	-	0.57	1.56	1.15	3.15	Req
005 SHA R015.500 TO R015.700 NORTH	02D	U	H 63	9 Y	5 Y	3 N	0 N	0 N	28.5	-	0.47	1.28	1.07	2.93	Req
005 SHA R016.820 TO R017.020 NORTH	02D	U	H 63	5 N	5 Y	0 N	0 N	0 N	29.0	-	0.48	1.32	1.08	2.96	
005 SHA R016.820 TO R017.020 SOUTH	02D	U	H 63	7 Y	4 N	0 N	0 N	0 N	29.0	-	0.48	1.32	1.08	2.96	
005 SHA R022.528 TO R022.728 SOUTH	02D	R	H 54	2 N	2 N	2 N	2 Y	2 Y	12.0	-	0.13	0.29	0.68	1.57	
005 SHA R028.906 TO R029.106 SOUTH	02L	R	H 51	17 Y	16 Y	4 Y	1 N	1 N	10.0	-	0.11	0.25	0.69	1.63	Req
005 SHA R029.106 TO R029.306 SOUTH	02L	R	H 51	3 Y	0 N	0 N	0 N	0 N	10.0	-	0.11	0.25	0.69	1.63	
005 SHA R029.506 TO R029.706 SOUTH	02L	R	H 51	4 Y	2 N	0 N	0 N	0 N	10.0	-	0.11	0.25	0.69	1.63	
005 SHA R029.706 TO R029.906 SOUTH	02L	R	H 51	13 Y	3 Y	0 N	0 N	0 N	10.0	-	0.11	0.25	0.69	1.63	Req
005 SHA R031.786 TO R031.986 SOUTH	03L	R	H 51	4 Y	1 N	0 N	0 N	0 N	10.0	-	0.11	0.25	0.69	1.63	
005 SHA R031.986 TO R032.186 SOUTH	02L	R	H 51	5 Y	4 Y	3 Y	0 N	0 N	9.8	-	0.10	0.25	0.69	1.63	Req

\+ denotes MV used in rates.

Req = investigation required (9, 6 or 3 or more accs. & significant in 36, 24 or 12 months, resp.)

Figure 11.11: Wet Table C list developed by TASAS using 1972 Factor

				Total Accidents / Significance					AVE ADT		36 MOS AVERAGE ACC & RATES NO OF ACC		RATE/MVM-MV		
Location Description	SCL RMP LNS	R U S	Rate Grp	36 mo. ACCS	24 mo. ACCS	12 mo. ACCS	6 mo. ACCS	3 mo. ACCS	Main	X-St.	F+I	TOT	F+I	TOT	Req
005 SHA R036.784 TO R036.984 NORTH	04D	R	H 53	8 Y	4 Y	1 N	1 N	1 N	9.6	-	0.11	0.25	0.73	1.68	
005 SHA R037.183 TO R037.383 SOUTH	05D	R	H 55	6 Y	6 Y	4 Y	1 N	1 N	9.6	-	0.08	0.18	0.51	1.21	Req
005 SHA 058.660 TO 058.860 SOUTH	02D	R	H 53	6 Y	2 N	1 N	1 N	0 N	9.0	-	0.10	0.23	0.72	1.66	
005 SIS R022.619 TO R022.819 NORTH	02D	R	H 54	3 Y	3 Y	0 N	0 N	0 N	8.4	-	0.09	0.20	0.66	1.52	
036 TRI 026.018 TO 026.218	02U	R	H 05	2 Y	2 Y	2 Y	0 N	0 N	0.3	-	0.03	0.05	3.96	7.87	
036 TEH 013.260 TO 013.460	02U	R	H 05	2 Y	2 Y	1 N	1 N	0 N	0.4	-	0.02	0.04	3.79	7.54	
036 TEH 040.570 TO 040.770	02U	S	H 07	3 N	3 Y	3 Y	1 N	1 N	3.3	-	0.14	0.37	3.32	8.59	Req
044 SHA L 1.738 044/WB OFF TO SB RTE 5	F L	U	R 70	12 Y	7 Y	0 N	0 N	0 N	6.3	-	0.38	1.32	0.79	2.74 +	Req
070 PLU 006.927 TO 007.127	02U	R	H 05	2 N	2 N	2 Y	1 N	1 N	1.5	-	0.06	0.11	2.84	5.65	
070 PLU 011.447 TO 011.647	02U	R	H 05	2 N	2 N	2 Y	1 N	1 N	1.5	-	0.06	0.11	2.84	5.65	
070 PLU 017.117 TO 017.317	02U	R	H 05	4 Y	4 Y	0 N	0 N	0 N	1.4	-	0.05	0.10	2.88	5.73	
096 SIS 037.227 TO 037.427	02U	R	H 05	2 Y	1 N	0 N	0 N	0 N	0.3	-	0.02	0.04	4.50	8.95	
096 SIS 063.387 TO 063.587	02U	R	H 05	2 N	2 Y	2 Y	2 Y	1 N	0.6	-	0.03	0.06	3.35	6.65	
096 SIS 068.737 TO 068.937	02U	R	H 05	2 Y	1 N	1 N	0 N	0 N	0.5	-	0.03	0.05	3.53	7.01	
096 SIS 075.817 TO 076.017	02U	R	H 05	2 N	2 Y	2 Y	1 N	0 N	0.5	-	0.03	0.06	3.44	6.84	
273 SHA 6.9 OX YOKE RD,RT	-XX	S	I 22	7 Y	3 N	2 N	2 N	1 N	11.9	3.1	0.37	0.88	0.32	0.77 +	
273 SHA 008.047 TO 008.247 SOUTH	02D	S	H 36	3 N	3 N	3 Y	3 Y	0 N	6.4	-	0.21	0.49	2.09	4.96	Req
299 TRI 002.177 TO 002.377	02U	R	H 05	5 Y	5 Y	2 N	1 N	0 N	3.1	-	0.15	0.30	2.51	4.98	

\+ denotes MV used in rates.

Req = investigation required (9, 6 or 3 or more accs. & significant in 36, 24 or 12 months, resp.)

Figure 11.12: Wet Table C list developed by TASAS using 1972 Factor cont'd

Location Description	SCL RMP LNS	R U S	Rate Grp	36 mo. ACCS	Total Accidents / Significance 24 mo. ACCS	12 mo. ACCS	6 mo. ACCS	3 mo. ACCS	AVE ADT Main	X-St.	36 MOS AVERAGE ACC & RATES NO OF ACC F+I	TOT	RATE/MVM-MV F+I	TOT	Req
299 TRI 012.367 TO 012.567	02U	R	H 05	3 Y	3 Y	2 N	2 Y	0 N	2.2	-	0.11	0.23	2.58	5.12	
299 TRI 013.197 TO 013.397	02U	R	H 05	7 Y	5 Y	3 Y	0 N	0 N	2.2	-	0.11	0.22	2.58	5.12	Req
299 TRI 020.007 TO 020.207	02U	R	H 05	3 Y	2 N	2 N	0 N	0 N	1.7	-	0.09	0.18	2.64	5.25	
299 TRI 023.962 TO 024.162	02U	R	H 05	3 Y	3 Y	3 Y	0 N	0 N	1.8	-	0.09	0.19	2.63	5.22	Req
299 SHA 003.082 TO 003.282	02U	R	H 05	4 Y	3 Y	2 N	1 N	0 N	4.0	-	0.16	0.31	2.56	5.09	
299 SHA 003.442 TO 003.642	02U	R	H 05	4 Y	3 Y	3 Y	1 N	1 N	4.0	-	0.16	0.31	2.56	5.09	Req
299 SHA 003.662 TO 003.862	02U	R	H 05	4 Y	4 Y	1 N	0 N	0 N	4.0	-	0.16	0.31	2.56	5.09	
299 SHA 004.006 TO 004.206	03U	R	H 12	8 Y	5 Y	1 N	1 N	1 N	4.0	-	0.09	0.18	1.45	3.01	
299 SHA 004.415 TO 004.615	02U	R	H 05	4 Y	3 Y	2 N	0 N	0 N	4.0	-	0.16	0.31	2.56	5.09	
299 SHA 4.85 RD TO GREENHORN MINE RT	---	R	I 16	3 Y	2 N	1 N	0 N	0 N	4.0	0	0.09	0.19	0.29	0.63 +	
299 SHA 004.897 TO 005.097	02U	R	H 05	7 Y	5 Y	0 N	0 N	0 N	4.0	-	0.16	0.31	2.56	5.09	
299 SHA 005.097 TO 005.297	02U	R	H 05	3 N	3 Y	1 N	0 N	0 N	4.0	-	0.16	0.31	2.56	5.09	
299 SHA 006.937 TO 007.137	02U	R	H 05	4 Y	2 N	2 N	0 N	0 N	4.0	-	0.16	0.31	2.56	5.09	
299 SHA 051.212 TO 051.412	02U	R	H 03	9 Y	8 Y	4 Y	1 N	0 N	4.8	-	0.14	0.30	1.95	4.04	Req
395 LAS 008.790 TO 008.990	02U	R	H 04	2 N	2 N	2 Y	0 N	0 N	5.5	-	0.07	0.13	1.39	2.79	

\+ denotes MV used in rates.

Req = investigation required (9, 6 or 3 or more accs. & significant in 36, 24 or 12 months, resp.)

Figure 11.13: Wet Table C list developed by TASAS using 1972 Factor cont'd

Location Description	SCL RMP LNS	R U S	Rate Grp	36 mo. ACCS		24 mo. ACCS		12 mo. ACCS		6 mo. ACCS		3 mo. ACCS		AVE ADT Main	AVE ADT X-St.	NO OF ACC F+I	NO OF ACC TOT	RATE/MVM-MV F+I	RATE/MVM-MV TOT	Req
003 TRI 60.19 SWIFT CR RD-LT (FS TAIL	---	R	I 17	2	Y	2	Y	1	N	0	N	0	N	0.7	.1	0.02	0.05	0.37	0.83+	
005 TEH R 24.712 005/NB ON FR SO RDBL	O G	U	R 24	4	Y	3	N	0	N	0	N	0	N	2.8	-	0.18	0.49	0.99	2.65+	
005 SHA .781 005/SBON GAS POINT RD/4TH	O D	R	R 11	3	Y	2	N	1	N	0	N	0	N	2.2	-	0.13	0.33	0.65	1.73+	
005 SHA R003.320 TO R003.520 SOUTH	02D	S	H 60	4	N	4	Y	3	N	2	N	2	N	25.6	-	0.35	0.94	0.77	2.09	
005 SHA R006.650 TO R006.850 SOUTH	04D	S	H 60	3	N	3	N	3	N	3	Y	1	N	25.4	-	0.34	0.93	0.77	2.09	
005 SHA R 15.265 005/NB OFF TO HILLTOP	F H	U	R 26	9	Y	7	Y	2	N	1	N	1	N	5.2	-	0.52	1.53	1.14	3.36+	Req
005 SHA R 15.391 005/SB OFF TO EB 44	F L	U	R 70	6	Y	4	N	1	N	1	N	1	N	5.5	-	0.37	1.29	0.77	2.68+	
005 SHA R015.300 TO R015.500 SOUTH	02D	U	H 63	10	Y	5	N	0	N	0	N	0	N	32.3	-	0.64	1.75	1.13	3.09	Req
005 SHA R015.500 TO R015.700 NORTH	02D	U	H 63	9	Y	5	Y	3	N	0	N	0	N	28.5	-	0.52	1.43	1.05	2.87	Req
005 SHA R016.820 TO R017.020 NORTH	02D	U	H 63	5	N	5	Y	0	N	0	N	0	N	29.0	-	0.54	1.47	1.06	2.90	
005 SHA R016.820 TO R017.020 SOUTH	02D	U	H 63	7	Y	4	N	0	N	0	N	0	N	29.0	-	0.54	1.47	1.06	2.90	
005 SHA R022.528 TO R022.728 SOUTH	02D	R	H 54	2	N	2	N	2	N	2	Y	2	Y	12.0	-	0.14	0.32	0.67	1.54	
005 SHA R028.906 TO R029.106 SOUTH	02L	R	H 51	17	Y	16	Y	4	Y	1	N	1	N	10.0	-	0.12	0.28	0.68	1.60	Req
005 SHA R029.106 TO R029.306 SOUTH	02L	R	H 51	3	Y	0	N	0	N	0	N	0	N	10.0	-	0.12	0.28	0.68	1.60	
005 SHA R029.506 TO R029.706 SOUTH	02L	R	H 51	4	Y	2	N	0	N	0	N	0	N	10.0	-	0.12	0.28	0.68	1.60	
005 SHA R029.706 TO R029.906 SOUTH	02L	R	H 51	13	Y	3	Y	0	N	0	N	0	N	10.0	-	0.12	0.28	0.68	1.60	Req
005 SHA R031.786 TO R031.986 SOUTH	03L	R	H 51	4	Y	1	N	0	N	0	N	0	N	10.0	-	0.12	0.28	0.68	1.60	
005 SHA R031.986 TO R032.186 SOUTH	02L	R	H 51	5	Y	4	Y	3	Y	0	N	0	N	9.8	-	0.12	0.28	0.68	1.60	Req

Column groups: Total Accidents / Significance; 36 MOS AVERAGE ACC & RATES.

\+ denotes MV used in rates.

Req = investigation required (9, 6 or 3 or more accs. & significant in 36, 24 or 12 months, resp.)

Figure 11.14: Wet Table C list developed by TASAS using 2008 Factor

Location Description	SCL RMP LNS	R U S	Rate Grp	36 mo. ACCS	Total Accidents / Significance 24 mo. ACCS	12 mo. ACCS	6 mo. ACCS	3 mo. ACCS	AVE ADT Main	X-St.	36 MOS AVERAGE ACC & RATES NO OF ACC F+I	TOT	RATE/MVM-MV F+I	TOT	Req
005 SHA R036.784 TO R036.984 NORTH	04D	R	H 53	8 Y	4 Y	1 N	1 N	1 N	9.6	-	0.12	0.28	0.72	1.65	
005 SHA R037.183 TO R037.383 SOUTH	05D	R	H 55	6 Y	6 Y	4 Y	1 N	1 N	9.6	-	0.08	0.20	0.50	1.19	Req
005 SHA 058.660 TO 058.860 SOUTH	02D	R	H 53	6 Y	2 N	1 N	1 N	0 N	9.0	-	0.11	0.26	0.70	1.62	
005 SIS R022.619 TO R022.819 NORTH	02D	R	H 54	3 Y	3 Y	0 N	0 N	0 N	8.4	-	0.09	0.20	0.66	1.52	
036 TRI 026.018 TO 026.218	02U	R	H 05	2 Y	2 Y	2 Y	0 N	0 N	0.3	-	0.02	0.05	4.03	8.02	
036 TEH 013.260 TO 013.460	02U	R	H 05	2 Y	2 Y	1 N	1 N	0 N	0.4	-	0.02	0.04	3.79	7.54	
036 TEH 040.570 TO 040.770	02U	S	H 07	3 N	3 Y	3 Y	1 N	1 N	3.3	-	0.14	0.37	3.32	8.59	Req
044 SHA L 1.738 044/WB OFF TO SB RTE 5	F L	U	R 70	12 Y	7 Y	0 N	0 N	0 N	6.3	-	0.42	1.48	0.77	2.68+	Req
070 PLU 006.927 TO 007.127	02U	R	H 05	2 N	2 N	2 Y	1 N	1 N	1.5	-	0.06	0.11	2.84	5.65	
070 PLU 011.447 TO 011.647	02U	R	H 05	2 N	2 N	2 Y	1 N	1 N	1.5	-	0.06	0.11	2.84	5.65	
070 PLU 017.117 TO 017.317	02U	R	H 05	4 Y	4 Y	0 N	0 N	0 N	1.4	-	0.05	0.10	2.88	5.73	
096 SIS 037.227 TO 037.427	02U	R	H 05	2 Y	1 N	0 N	0 N	0 N	0.3	-	0.02	0.04	4.50	8.95	
096 SIS 063.387 TO 063.587	02U	R	H 05	2 N	2 Y	2 Y	2 Y	1 N	0.6	-	0.03	0.06	3.35	6.65	
096 SIS 068.737 TO 068.937	02U	R	H 05	2 Y	1 N	1 N	0 N	0 N	0.5	-	0.03	0.05	3.53	7.01	
096 SIS 075.817 TO 076.017	02U	R	H 05	2 N	2 Y	2 Y	1 N	0 N	0.5	-	0.03	0.06	3.44	6.84	
273 SHA 6.9 OX YOKE RD,RT	-XX	S	I 22	7 Y	3 N	2 N	2 N	1 N	11.9	3.1	0.41	0.98	0.31	0.75+	
273 SHA 008.047 TO 008.247 SOUTH	02D	S	H 36	3 N	3 N	3 Y	3 Y	0 N	6.4	-	0.23	0.54	2.05	4.86	Req
299 TRI 002.177 TO 002.377	02U	R	H 05	5 Y	5 Y	2 N	1 N	0 N	3.1	-	0.14	0.27	2.56	5.08	

\+ denotes MV used in rates.

Req = investigation required (9, 6 or 3 or more accs. & significant in 36, 24 or 12 months, resp.)

Figure 11.15: Wet Table C list developed by TASAS using 2008 Factor cont'd

Location Description	SCL RMP LNS	R U S	Rate Grp	36 mo. ACCS		24 mo. ACCS		12 mo. ACCS		6 mo. ACCS		3 mo. ACCS		AVE ADT Main	AVE ADT X-St.	NO OF ACC F+I	NO OF ACC TOT	RATE/MVM-MV F+I	RATE/MVM-MV TOT	Req
299 TRI 012.367 TO 012.567	02U	R	H 05	3	Y	3	Y	2	N	2	Y	0	N	2.2	-	0.10	0.20	2.62	5.21	
299 TRI 013.197 TO 013.397	02U	R	H 05	7	Y	5	Y	3	Y	0	N	0	N	2.2	-	0.10	0.20	2.62	5.21	Req
299 TRI 020.007 TO 020.207	02U	R	H 05	3	Y	2	N	2	Y	0	N	0	N	1.7	-	0.08	0.16	2.69	5.35	
299 TRI 023.962 TO 024.162	02U	R	H 05	3	Y	3	Y	3	Y	0	N	0	N	1.8	-	0.09	0.17	2.68	5.32	Req
299 TRI 038.897 TO 039.097	02U	R	H 05	3	Y	0	N	0	N	0	N	0	N	3.0	-	0.13	0.26	2.57	5.11	
299 SHA 003.082 TO 003.282	02U	R	H 05	4	Y	3	Y	2	N	1	N	0	N	4.0	-	0.18	0.35	2.52	5.00	
299 SHA 003.442 TO 003.642	02U	R	H 05	4	Y	3	Y	3	Y	1	N	1	N	4.0	-	0.18	0.35	2.52	5.00	Req
299 SHA 003.662 TO 003.862	02U	R	H 05	4	Y	4	Y	1	N	0	N	0	N	4.0	-	0.18	0.35	2.52	5.00	
299 SHA 004.006 TO 004.206	03U	R	H 12	8	Y	5	Y	1	N	1	N	1	N	4.0	-	0.10	0.21	1.42	2.96	
299 SHA 004.415 TO 004.615	02U	R	H 05	4	Y	3	Y	2	N	0	N	0	N	4.0	-	0.18	0.35	2.52	5.00	
299 SHA 4.85 RD TO GREENHORN MINE RT	--	R	I 16	3	Y	2	N	1	N	0	N	0	N	4.0	0	0.10	0.22	0.28	0.62+	
299 SHA 004.897 TO 005.097	02U	R	H 05	7	Y	5	Y	0	N	0	N	0	N	4.0	-	0.18	0.35	2.52	5.00	
299 SHA 005.097 TO 005.297	02U	R	H 05	3	N	3	Y	1	N	0	N	0	N	4.0	-	0.18	0.35	2.52	5.00	
299 SHA 006.937 TO 007.137	02U	R	H 05	4	Y	2	N	2	N	0	N	0	N	4.0	-	0.18	0.35	2.52	5.00	
299 SHA 051.212 TO 051.412	02U	R	H 03	9	Y	8	Y	4	Y	1	N	0	N	4.8	-	0.16	0.33	1.91	3.96	Req
395 LAS 008.790 TO 008.990	02U	R	H 04	2	N	2	N	2	Y	0	N	0	N	5.5	-	0.08	0.16	1.37	2.74	

Total Accidents / Significance; 36 MOS AVERAGE ACC & RATES

+ denotes MV used in rates.

Req = investigation required (9, 6 or 3 or more accs. & significant in 36, 24 or 12 months, resp.)

Figure 11.16: Wet Table C list developed by TASAS using 2008 Factor cont'd

					1972 Single Factor				2008 Variable Factor			
Route	Segment (MP to MP)	36 Mo Acc	24 mo Acc	12 Mo Acc	36 Month	24 Month	12 Month	1972 Required	36 Month	24 Month	12 Month	2008 Required
02 SHA 005	0.69 to 0.89	6	5	3	Y	Y	N	N	Y	Y	N	N
02 SHA 005	0.9 to 1.10	4	2	1	Y	N	N	N	N	N	N	N
02 SHA 005	3.3 to 3.5	7	5	3	Y	Y	N	N	Y	Y	N	N
02 SHA 005	4.37 to 4.57	3	3	3	N	N	Y	Y	N	N	Y	Y
02 SHA 005	6.69 to 6.89	5	4	4	Y	N	Y	Y	Y	N	Y	Y
02 SHA 005	11.85 to 12.05	7	3	2	Y	N	N	N	Y	N	N	N
02 SHA 005	12.13 to 12.33	5	4	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	13.9 to 14.1	4	2	2	Y	N	N	N	Y	N	N	N
02 SHA 005	14.66 to 14.86	8	4	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	14.99 to 15.19	5	4	3	Y	Y	N	N	Y	Y	N	N
02 SHA 005	15.26 to 15.46	11	7	2	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	15.47 to 15.67	23	13	3	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	16.93 to 17.13	20	13	2	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	17.44 to 17.64	5	2	1	Y	N	N	N	Y	N	N	N
02 SHA 005	19.25 to 19.45	4	4	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	28.79 to 28.99	3	1	0	Y	N	N	N	Y	N	N	N
02 SHA 005	28.92R to 29.12R	2	2	2	N	N	Y	N	N	N	Y	N
02 SHA 005	28.93L to 29.13L	17	15	4	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	29.48 to 29.68	4	2	0	Y	Y	N	N	Y	N	N	N
02 SHA 005	29.69 to 29.89	15	3	0	Y	Y	Y	Y	Y	Y	N	Y
02 SHA 005	31.95 to 32.15	9	5	3	Y	Y	N	Y	Y	Y	Y	Y
02 SHA 005	36.09 to 36.29	3	2	1	Y	Y	N	N	Y	Y	N	N
02 SHA 005	36.78 to 36.98	9	5	1	Y	Y	Y	Y	Y	Y	N	Y
02 SHA 005	36.99 to 37.19	14	11	9	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	37.2 to 39.4	7	7	4	Y	Y	N	Y	Y	Y	Y	Y
02 SHA 005	39.14 to 39.34	3	3	2	N	N	N	N	Y	Y	N	N
02 SHA 005	39.79 to 39.99	2	2	2	Y	Y	Y	N	N	N	N	N
02 SHA 005	40.02 to 40.22	8	5	4	N	N	N	Y	Y	Y	Y	Y
02 SHA 005	58.13 to 58.33	2	1	1	Y	Y	N	N	N	N	N	N
02 SHA 005	58.84 to 59.04	8	3	2	N	N	N	N	Y	Y	N	N
02 SHA 005	60.4 to 60.6	2	1	0	Y	Y	Y	N	N	N	N	N
02 SHA 005	60.61 to 60.81	3	3	3	N	N	N	N	Y	Y	Y	Y
02 SHA 005	64.91 to 65..11	3	3	2	Y	Y	N	N	Y	Y	N	N
2 SIS 005	22.67 to 22.87	3	3	0	Y	Y	N	N	Y	Y	N	N
299 SHA	3.2 to 3.4	5	4	2	Y	Y	N	N	Y	Y	N	N
299 SHA	3.5 to 3.7	5	4	3	Y	Y	Y	Y	Y	Y	Y	Y
299 SHA	3.72 to 3.92	3	3	1	Y	Y	N	N	Y	Y	N	N
299 SHA	4.1 to 4.25	8	5	1	Y	Y	N	N	Y	Y	N	N
299 SHA	4.59 to 4.79	7	4	2	Y	Y	N	N	Y	Y	N	N
299 SHA	4.93 to 5.1	7	5	0	Y	Y	N	N	Y	Y	N	N
299 SHA	5.11 to 5.31	4	4	1	Y	Y	N	N	Y	Y	N	N
299 SHA	5.39 to 5.59	3	2	2	Y	N	N	N	Y	N	N	N
299 SHA	7.05 to 7.25	4	2	2	Y	N	N	N	Y	N	N	N
299 SHA	10.88 to 11.08	2	2	2	N	N	Y	N	N	N	Y	N
299 SHA	51.3 to 51.5	10	9	5	Y	Y	Y	Y	Y	Y	Y	Y
299 SHA	75.0 to 75.2	3	2	2	Y	N	N	N	Y	N	N	N
299 SHA	8.98 to 9.18	2	2	2	N	N	Y	N	N	N	Y	N

Figure 11.17: Wet Table C lists developed by WTI using 1972 and 2008 variable factors

					1972 Single Factor				2008 Variable Factor			
Route	Segment (MP to MP)	36 Mo Acc	24 mo Acc	12 Mo Acc	36 Month	24 Month	12 Month	1972 Required	36 Month	24 Month	12 Month	2008 Required
02 SHA 005	0.69 to 0.89	6	5	3	Y	Y	N	N	Y	Y	N	N
02 SHA 005	0.9 to 1.10	4	2	1	Y	N	N	N	N	N	N	N
02 SHA 005	3.3 to 3.5	7	5	3	Y	Y	N	N	Y	Y	N	N
02 SHA 005	4.37 to 4.57	3	3	3	N	N	Y	Y	N	N	Y	Y
02 SHA 005	6.69 to 6.89	5	4	4	Y	N	Y	Y	Y	N	Y	Y
02 SHA 005	11.85 to 12.05	7	3	2	Y	N	N	N	Y	N	N	N
02 SHA 005	12.13 to 12.33	5	4	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	13.9 to 14.1	4	2	2	Y	N	N	N	Y	N	N	N
02 SHA 005	14.66 to 14.86	8	4	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	14.99 to 15.19	5	4	3	Y	Y	N	N	Y	Y	N	N
02 SHA 005	15.26 to 15.46	11	7	2	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	15.47 to 15.67	23	13	3	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	16.93 to 17.13	20	13	2	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	17.44 to 17.64	5	2	1	Y	N	N	N	Y	N	N	N
02 SHA 005	19.25 to 19.45	4	4	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	28.79 to 28.99	3	1	0	Y	N	N	N	Y	N	N	N
02 SHA 005	28.92R to 29.12R	2	2	2	N	N	Y	N	N	N	Y	N
02 SHA 005	28.93L to 29.13L	17	15	4	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	29.48 to 29.68	4	2	0	Y	N	N	N	Y	N	N	N
02 SHA 005	29.69 to 29.89	15	3	0	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	31.95 to 32.15	9	5	3	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	36.09 to 36.29	3	2	1	Y	Y	N	N	Y	Y	N	N
02 SHA 005	36.78 to 36.98	9	5	1	Y	Y	N	Y	Y	Y	N	Y
02 SHA 005	36.99 to 37.19	14	11	9	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	37.2 to 39.4	7	7	4	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	39.14 to 39.34	3	3	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	40.02 to 40.22	8	5	4	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	58.84 to 59.04	8	3	2	Y	Y	N	N	Y	Y	N	N
02 SHA 005	60.61 to 60.81	3	3	3	Y	Y	Y	Y	Y	Y	Y	Y
02 SHA 005	64.91 to 65..11	3	3	2	Y	Y	N	N	Y	Y	N	N
2 SIS 005	22.67 to 22.87	3	3	0	Y	Y	N	N	Y	Y	N	N
299 SHA	3.2 to 3.4	5	4	2	Y	Y	N	N	Y	Y	N	N
299 SHA	3.5 to 3.7	5	4	3	Y	Y	Y	Y	Y	Y	Y	Y
299 SHA	3.72 to 3.92	3	3	1	Y	Y	N	N	Y	Y	N	N
299 SHA	4.1 to 4.25	8	5	1	Y	Y	N	N	Y	Y	N	N
299 SHA	4.59 to 4.79	7	4	2	Y	Y	N	N	Y	Y	N	N
299 SHA	4.93 to 5.1	7	5	0	Y	Y	N	N	Y	Y	N	N
299 SHA	5.11 to 5.31	4	4	1	Y	Y	N	N	Y	Y	N	N
299 SHA	5.39 to 5.59	3	2	2	Y	N	N	N	Y	N	N	N
299 SHA	7.05 to 7.25	4	2	2	Y	N	N	N	Y	N	N	N
299 SHA	10.88 to 11.08	2	2	2	N	N	Y	N	N	N	Y	N
299 SHA	51.3 to 51.5	10	9	5	Y	Y	Y	Y	Y	Y	Y	Y
299 SHA	75.0 to 75.2	3	2	2	Y	N	N	N	Y	N	N	N
395 LAS	8.98 to 9.18	2	2	2	N	N	Y	N	N	N	Y	N

Figure 11.18: Wet Table C lists developed by WTI using 2008 fixed and 2008 variable factors