University of Alberta

Statistical Analysis on Adverse Weather Related Power Transmission Line Outages in Alberta

By

Ming Wu ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Statistics

Department of Mathematical & Statistical Sciences

Edmonton, Alberta

Spring 2005

0-494-08177-5

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN:
Our file Notre reterence
ISBN:

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Dedication

This thesis is dedicated to my parents.

Abstract

This thesis analyzes the spatial and temporal distributions of lightning and explores the unique characteristics of adverse weather caused line-related transmission line outages for several voltage level transmission lines in Alberta. The Lasso regression variable selection procedure and Cp criterion are applied to express the duration of the lightning-caused line-related transmission line outage as a function of weather and lightning elements. The major results and conclusions are as follows:

(1) The spatial and temporal distributions of the lightning flashes are not uniform. The mean and variance of lightning flashes reach maxima in central Alberta.

(2) The range of the weather elements, which cause the serious outages to the power transmission lines are documented. and

(3) The major lightning and weather elements that cause the power transmission lines outages are identified for three transmission line voltage levels. Future lightning-caused line-related transmission line outages can be predicted by the resulting functions.

Acknowledgement

I would like to express my gratitude to my two supervisors, Drs. Don Koval and Samuel Shen, for their remarkable support and guidance in this research. Their stimulating suggestions and encouragement helped me in all the time of research for and writing of this thesis. I would also like to thank Dr. Hee-Seok Oh, who has influenced my views on statistical science, for invaluable discussions on the model selection methods. Special thanks to Bin Shen and Mark Duller for the assistance they have given to this thesis. Last, I am thankful to the Department of Mathematical and Statistical Sciences of University of Alberta for giving me this opportunity to study towards to this degree.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Today, electricity consumers are becoming more sensitive to electrical disturbances and demanding a highly reliable power transmission system. Because of the high cost of the power transmission system outages, electrical utilities are forced to investigate power transmission system outage patterns and setup a reliability evaluation system for system designs and maintenance activities. A wide range of research regarding power transmission system reliability has been conducted and many reliability evaluation methodologies and equations have been developed for the different utilities' applications. The IEEE Std. 493 (i.e., IEEE Gold Book) is a standard (i.e., recommended practice) for power engineers to plan and design reliable industrial and commercial electric power distribution system.

Due to the variance of power transmission system configurations and environments, there are no superb transmission system reliability evaluation methodologies presently available. However, historical power system reliability data analysis is the common crucial starting point to develop any new reliability evaluation methods. By defining the needs and purposes of the data, each utility can design and operate its own data collection system. Suitable reliability statistics and indices can be organized into an appropriate data structure, which is necessary for the application of the reliability methodologies.

Power transmission system configurations can be altered by many outage conditions which will interrupt power delivery to customers. The identified primary causes of transmission line forced outages by the Canadian Electricity Association (CEA) are defective components, adverse weather, adverse environment, system conditions, human elements, foreign interference, and the other unknown elements (Shen, 1999).

It is well known that adverse weather is one of the major causes of power system outages. In North America, adverse weather has caused power utilities outages to become a nightmare for both customers and utilities. People still remember the 1998 ice storm disaster in Québec, which caused about 1.4 million Canadians in Québec and 230,000 Canadians in Ontario to be without power for many days. This catastrophe cost Hydro-Québec $525 million and took weeks to fully restore the damaged transmission and distribution system to their operational states. In the same year, Virginia Power reported that over 400,000 customers were affected by a Christmas Eve ice storm. About 75% of its customers around Williamsburg were out of power for several days. A vast ice storm happened in January 12, 1998 in New England resulting in about 750,000 residential customer interruptions. The recent blackout in Nova Scotia, from November 14, 2004, was a power outage event caused by a pre-winter storm. The 45 centimeters of heavy snow and 90-kilometre-an-hour wind brought down six major transmission towers in Dartmouth and Sackville area. Approximately 100,000 homes and businesses lost electricity initially and 35,000 homes were still without power after two days. These disasters require significant research efforts to reveal the patterns of adverse weather in order to assess the impact on power system network configurations and assess the costs of these interruptions to society.

A previous study on Alberta power transmission line outages has been conducted by Shen et al. (1999) by using the 19-year Alberta Power System Reliability Database and the 33-year Alberta Environmental Service Weather Data. The main results of this work can be summarized as follows:

Adverse weather caused line-related transmission outages accounted for 33% of the total transmission line-related outages in Alberta. A correlation analysis of the transmission line outage data revealed that wind and precipitation-caused outages were not proportional to the physical length of the transmission lines. However, lightning-caused outages were strongly correlated with the physical length of the transmission lines. This result contradicts some published methodologies which found that the statistics of the transmission line outages are different between the voltage classes and different geographical locations.

Shen et al. (1999) also completed a case study based on this outage data set using Monte Carlo Simulation method. One important result of this simulation is that an independent power supply lowers the frequency and duration of load point interruptions.

It is pointed out by Shen (1999) that the analysis on the 33-year Alberta Environmental Service Weather Data is not accurate enough to help power and utility engineers to have a good sense of the various characteristics of Alberta's climate. This leads to an important question: what is the climate in Alberta?

Considerable research on the Alberta climate has been conducted to date. The Agroclimatic Atlas of Alberta (2001) published by Alberta Agriculture, Food and Rural Development includes a detailed description of Alberta climate conditions. However, the atlas is not designed for the purpose of system analysis of power transmission systems as the climate variables provided in the atlas are not suitable for, and cannot been treated as an applicable industrial guide for power system design, operation and maintenance. A quantitative description from a power-engineering point of view of climate showing the characteristic values of climate variables over the Alberta region is needed. In other word, the climatology of Alberta for power system reliability analysis of power transmission systems has to be established.

Probabilistic analysis and simulation methods with consideration of weather effects have been extensively applied in the power system reliability assessment over the last several decades. Some of the common methologies are found in Billinton et al. (1986), Billinton et al. (1991) and Bhuiyan et al. (1994). These methodologies focus on studying the stochastic procedure underneath the power system behaviour. The probability distribution function estimation and a thorough understanding of the power system to which an evaluation model is applied are the crucial parts to establish a relatively stable and effective evaluation model. Various assumptions have to be made to apply the reliability methodologies, especially for interruptions of complex power systems.

An alternative approach for power system reliability assessment modeling is the regression approach. Linear or non-linear regression models can provide a utility with the interpretable relationship between weather-related power outages and the weather elements and reliable prediction. The Lasso regression shrinkage method proposed by Tibshirani (1996) is considered a good method in regression model selection. By imposing a penalty function on estimating the regression coefficients, the Lasso methods enjoy a continuous coefficients selection procedure and produce better results than the subset selection method (Tibshirani, 1996) The detailed discussion of this method and its application in expressing the duration of the lightning related power transmission line outages as the function of the weather and lightning elements in Alberta is presented later in this thesis.

1.2 Objectives

The objectives of this thesis are to conduct a thorough analysis on a 19-year, historical Alberta power transmission system reliability database (January 1977 to December, 1995), and to identify the various features of the temporal and spatial patterns and characteristics of adverse weather caused line-related power transmission line outages. Various statistics for the performance of the transmission line-related adverse weather caused outages will be organized and stored in a Microsoft Access database system. All the statistics can be used as input to reliability methodologies used for evaluating system design and operation of the Alberta transmission network.

The weather and climate of Alberta are reviewed and the lightning patterns of Alberta will be analyzed. A clear picture of the climatology in Alberta will be drawn and the characteristics of lightning will be fully investigated. These results provide the reference for the power transmission system reliability analysis when there is a need to include weather and lightning elements in assessing the reliability of a power transmission system.

The regression approach instead of the traditional probabilistic approach is applied to form a practical model for the purpose of finding the main factors of the weather and lightning elements and future prediction of transmission lines failures. This regression approach study provides the power engineer with an important new viewpoint of implementing the weather elements into power reliability modeling and improves a utility's ability to be prepared for the weather-related power transmission line outages by prediction.

1.3 Outline of Thesis

Chapter 2 reviews the weather and climate patterns and characteristics in Alberta of three weather elements, which are air temperature, precipitation and wind. The main objective for this chapter is to reveal the unique patterns and characteristic of lightning in Alberta. Various statistics of lightning are investigated.

Chapter 3 covers the analysis on the Alberta power transmission system outage data especially on adverse weather related sustained power transmission line forced outage analysis. Various statistics for the adverse weather related power transmission line outages stratified by the power line voltage levels are presented here. This chapter also includes correlation analysis between weather elements and the duration of the adverse weather related power transmission line outages.

The major model selection procedures are discussed in the Chapter 4. The focus is on the Lasso method and its implementation. The models between the duration of the lightning-caused power line-related outages and the weather and lightning elements are established by Lasso method and Cp criterion.

Chapter 5 presents the major results and conclusions of this thesis.

Chapter 2

Weather, Climate and Lightning in Alberta

2.1 Weather and Climate in Alberta

Weather is the specific condition of the atmosphere at a particular place and usually over a short period of time. It is measured in terms of such factors as wind, temperature, humidity, atmospheric pressure, cloudiness, and precipitation. In most places, weather can change from hour-to-hour, day-to-day, and season-to-season due the geographic features and the interactions of different types of air masses.

Alberta is in western Canada with the Rocky Mountains located at its Western border. The rest of the province is on a great plain, shared with Saskatchewan on the east, and with the state of Montana to the south.

Generally, the frontal precipitation, convective precipitation and topographic precipitation are the three types of precipitation received in Alberta. It is pointed out by Chetner et al. (2003) that in southernwest Alberta the moist, southeasterly winds and increasing elevation can create ground fogs several thousand meters deep. And this form of weather is common at Lethbridge.

Climate in a narrow sense is usually defined as the "average weather", or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over period of time ranging from months to thousands or millions of years. The classical period is 30 years, as defined by the World Meteorological Organization (WMO). These quantities are most often surface variables such as temperature, precipitation, and wind. Climate in a wider sense is the state, including a statistical description, of the climate system (Intergovernmental Panel on Climate Change, 2001).

Alberta has a continental climate with warm summers and cold winters. Northern Alberta has many fewer frost-free days and lower temperatures than Southern Alberta, which is almost desert-like in its summer heat and lack of rain. The reason of this phenomenon is that south receives more solar radiation than the north does. It is found that in July, the average daily temperature ranges from warmer than 18°C in the south to cooler than 13°C in the Rocky Mountains and the north. In January, the average daily temperature ranges from cooler than -24°C in the far north to warmer than -10°C in the south and the mountains (Chetner et al., 2003). Western Alberta near the mountains enjoys the warmth brought by winter Chinooks. The range of average temperatures across the province in winter is greater than in summer because of the reduced heating from the sun at northern latitudes. Central and southern Alberta are the most likely places in Canada to experience tornados because of the summer heat. And violent summer thunderstorms are common in the eastern half of the province.

The extreme minimum (in the winter) or maximum (in the summer) temperatures are also the important element to describe the temperature status in a certain area. For Alberta, in the winter, the extreme minimum temperatures varies from cooler than -54°C in northern Alberta to warmer than -46°C in southern Alberta. In the summer, the maximum temperatures are observed from 32°C in the mountains to warmer than 40°C in southern Alberta.

Precipitation is generally highest along the mountains and into west central Alberta due to the combined effects of frontal and topographic precipitation. Precipitation from May 1 to August 31 varies from slightly below 200 millimeters (mm) in the driest prairie areas to more than 325 mm in the mountains. From September 1 to April 30, precipitation ranges from less than 150 mm in the driest prairie region to more than 275 mm in the mountains.

The spatial distribution of the wind in Alberta is not uniform. The average wind speeds in each monsoon season gradually increase from northern Alberta to southern Alberta. High wind speeds are usually found in southern Alberta near the border of Montana and the center of low wind speeds is observed in northwestern Alberta. In winter (December, January and February), the wind prevailing wind directions in northern Alberta are north or northeast and in central and southern Alberta, they become west or southwest. In spring (March, April and May), the prevailing wind directions in northern Alberta are southeast or southwest and they are south or southwest in central and southern Alberta. In summer (June, July and August), the prevailing wind directions in the whole Alberta are almost west or northwest. And in fall (September, October and November), the prevailing wind directions in northwest Alberta are northwest; the prevailing wind directions in northeast Alberta are east or southeast; the prevailing wind directions in central Alberta are west and the prevailing wind directions in southern Alberta are southwest.

With predicted global warming of 1.4 to 5.8 °C between 1900 and 2100 CE (Intergovernmental Panel on Climate Change, 2001), the impact of this global warming within Alberta should be evaluated. Shen (1999) reports that the maximum temperature in Alberta does not exhibit an upward trend, but the minimum temperature has clearly been increasing since approximately 1920 CE. This rise in Summer minimum temperatures is therefore an indicator that the atmosphere is more prone to thunderstorms and hail activity (Francis, 1998). Francis (1998) showed that tornado frequency in the Prairie Provinces has shown a tendency to increase in the Spring and early Summer in step with increases in average monthly temperature. However, there still needs to be more investigations to reveal other impacts of a warmer climate on the climate of Alberta.

2.2 Lightning in Alberta

Lightning is one of the major adverse weather conditions that can cause frequent power system outages and is the single largest cause of power outages in many lightning-prone regions in the world. Lightning is the primary cause of transmission line outages in Alberta (Shen et al., 1999). Burrows et al. (2002) indicated that Alberta is one of the highest lightning areas in Canada. Studies to investigate the spatial and temporal distributions of lightning from a power utilities perspective and the patterns of lightning-caused outages have been conducted for some areas in North American. For example, Lopez et al. (1997) examined the spatial and temporal distributions of lightning over Arizona from a power engineering point of view. Shen et al. (1999) pointed out that the duration of the lightning-caused outages is proportional to the physical length of the transmission lines in Alberta. Orville et al. (2002) found the mean positive lightning multiplicity is observed to have maximum values in Alberta. However, the characteristics of lightning and the relationship between lightning and power transmission system outages are still not fully understood by power system planning engineers in Alberta.

The spatial distribution of lightning flashes is not uniform in Alberta due to geographical variations. This study is presented to reveal hot spots of lightning flashes and spatial patterns of lightning distribution in Alberta to provide information for utilities to maintain or improve their power system reliability performance. The lightning patterns also varied over time. From historical lightning data, two questions will be addressed: (1) whether there are certain annual or monthly lightning cycles; (2) is the occurrence of lightning predictable? If these questions can be fully answered, then a power outage alert system could be developed according to the lightning patterns and power system reliability improved.

2.2.1 Data

A 20-year (1983-2002) daily lightning data set used in this study was provided by Alberta Lands and Forest Service. The data were collected by the operational systems manufactured by Lightning Location and Protection, Inc. (LLP) from 1983 (Nimchuk 1989) until the Canadian Lightning Detection Network (CLDN) was completed in February 1998 (Burrows et al., 2002). The CLDN consists of 82 sensors that detect lightning over most of Canada (Burrows et al., 2002). A figure to show the location of Canadian lightning sensors can be found in Burrows et al. (2002). Two types of the sensors are involved in the CLDN. One is the LPATS-IV (Lighting Positioning and Tracking Sensors, Series 4) which detect time-of-arrival of radio pulses generated by lightning and another one is the IMPACT/ES sensors (Improved

Accuracy from Combined Technology, ES version) which combine the magnetic direction-finding method with time-of-arrival technology (Burrows et al., 2002). The detailed technical information on sensors and detection methods is in Cummins et al. (1998).

There are varying degrees of uncertainty associated with the detection efficiency of the network (Kochtubajda et al., 2002). The study in Alberta reported detection efficiencies of approximately 70% - 80% within 300km by the LLP system (Nimchuk, 1989). An analysis of the CLDN by Global Atmospherics, Inc. estimated that the flash detection efficiency is 85% - 90% out to 200km from the network periphery, decreasing to 80% at the periphery, and to 10% - 30% at a distance of 300km beyond the periphery (Orville et al., 2002). Because the high detection efficiency of both networks, there are no corrections for detection efficiency applied in this study.

Every lightning event during 1983 to 2002 period was recorded by its occurrence date, the longitude, the latitude, the peak current amplitude in kA, the multiplicity (number of strokes) and the polarities. In this thesis, only the lightning flashes are analyzed. Cummins et al. (1998) define a flash as a series of separate strokes, to a maximum of 15, which occur for a period of one second within 10km of the first stroke detected, with the time interval between strokes being less than 0.5s. The peak current amplitude in the data is for the first stroke, which is generally the strongest.

There are two types of lightning flashes, the Cloud-to-ground (CG) flashes and Cloud-to-cloud (CC) flashes. It is generally believed that the CG flashes are more destructive to the power transmission system than the CC flashes. But the dataset used in this study does not separate the two types of flashes. Due to the fact that the CC flashes account for very small portion (approximately 1%) of total flashes that can be measured by the lightning detective network. In this study, we assumed that all the lightning flashes are the CG flashes.

2.2.2 Spatial Distribution of Lightning Flashes

Alberta is located within 49°N to 60°N latitude and 110°W to 120°W longitude. The lightning data within Alberta was retrieved from the original lightning data set. Approximately 6.5 million lightning flashes were detected in this 20-year period in Alberta. All the peak current amplitudes of these lightning flashes are larger than 0.4 kA either in positive polarity or in negative polarity. Generally, the positive CG flashes occurred much less frequently than negative CG flashes. There are around 6.5%~15% positive CG flashes out of the total lighting flashes for each year. However, it is generally believed that the positive CG flashes do more damage to the electrical components than the negative CG flashes due to their larger peak current amplitude and causing longer outage durations. The patterns of positive lightning should be investigated from a power utility's perspective.

To obtain the lightning patterns, the Alberta province is divided into a 110 by 100 grid, that is 0.1° by 0.1° in latitude-longitude scale. All the lightning variables were calculated for each cell. The grid cell areas are not equal due to the ellipsoid shape of the earth surface. The areas vary from 62 km^2 to 81 km^2 from the northern border to the southern border of Alberta. The difference of the grid cell areas is so large that they cannot be treated as equal-area cells. Otherwise, the final results (e.g., lightning frequency per km^2 per year) would lead to a different conclusion. Therefore, in this study, all the lightning variables for each cell are averaged based on their areas.

The 20-year mean frequency of overall CG lightning flashes and the physical locations of the power transmission lines imposed are plotted in Figure 2-1. It is shown that the spatial distribution of overall CG lightning has a fan-shape pattern. The high-frequency lightning areas are in central Alberta, from Swan Hills in the north to Banff in the south along the Rocky Mountains. The mean lightning frequency of this area is above 1.5 flashes/ km^2/year. This fan-shape pattern agrees with the results of Burrows et al, 2002. For central Alberta as a whole, the average lightning frequency is over 0.6 flashes/ km^2/year. The number is smaller than that of Burrows et al, 2002 who only used 1998 to 2000 lightning data. In 1986, a large increase in the number of annual lightning flashes was detected and this measurement was doubled after this year. This is why our result, a 1983-2002 mean, is smaller than that of Burrows et al.'s (2002), i.e. the 1998-2000 mean. From Figure 2-1, three hotspots with high overall lightning frequency over 2.1 flashes/km^2/year were found. They were Edson, Rocky Mountain House and Waiparous.

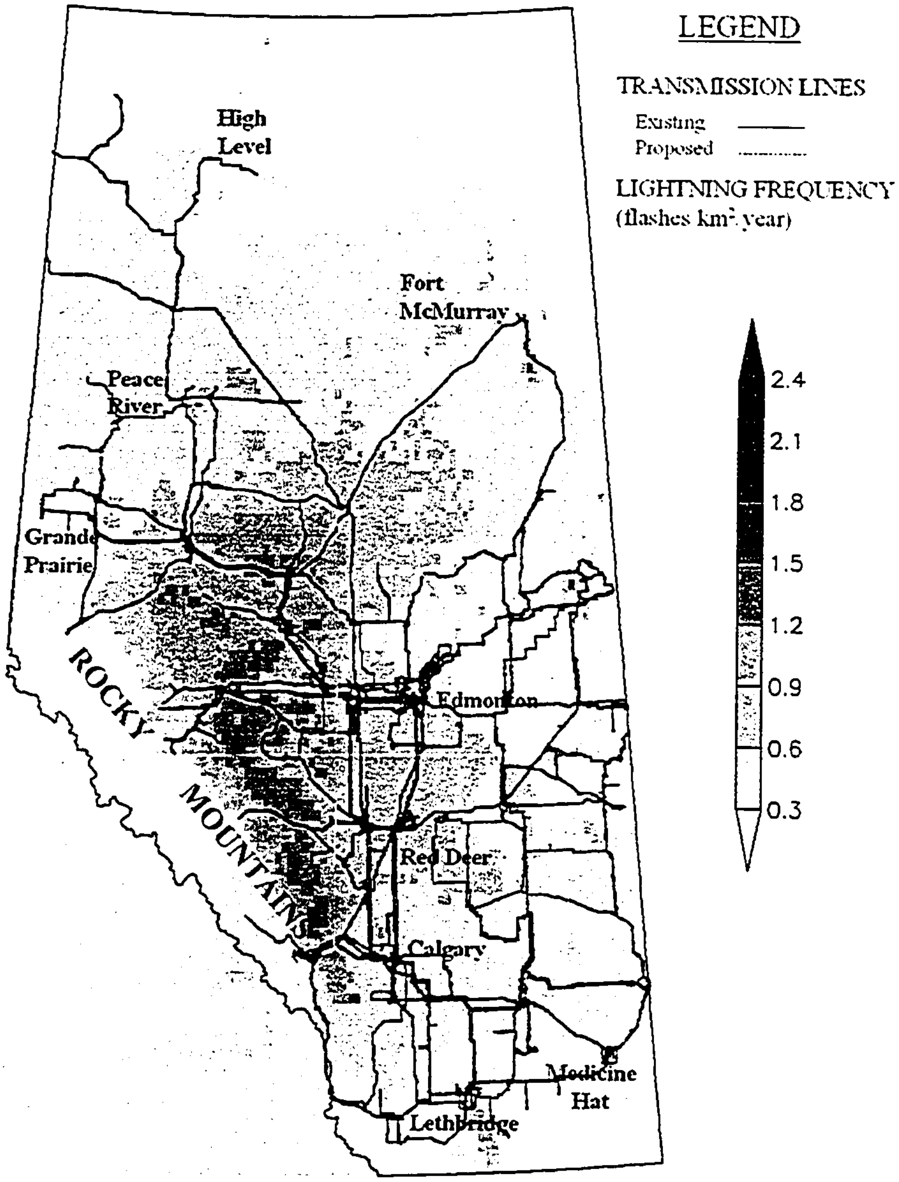


Figure 2-1. The 20-year mean overall lightning flashes per km^2 per year and the power transmission lines in Alberta

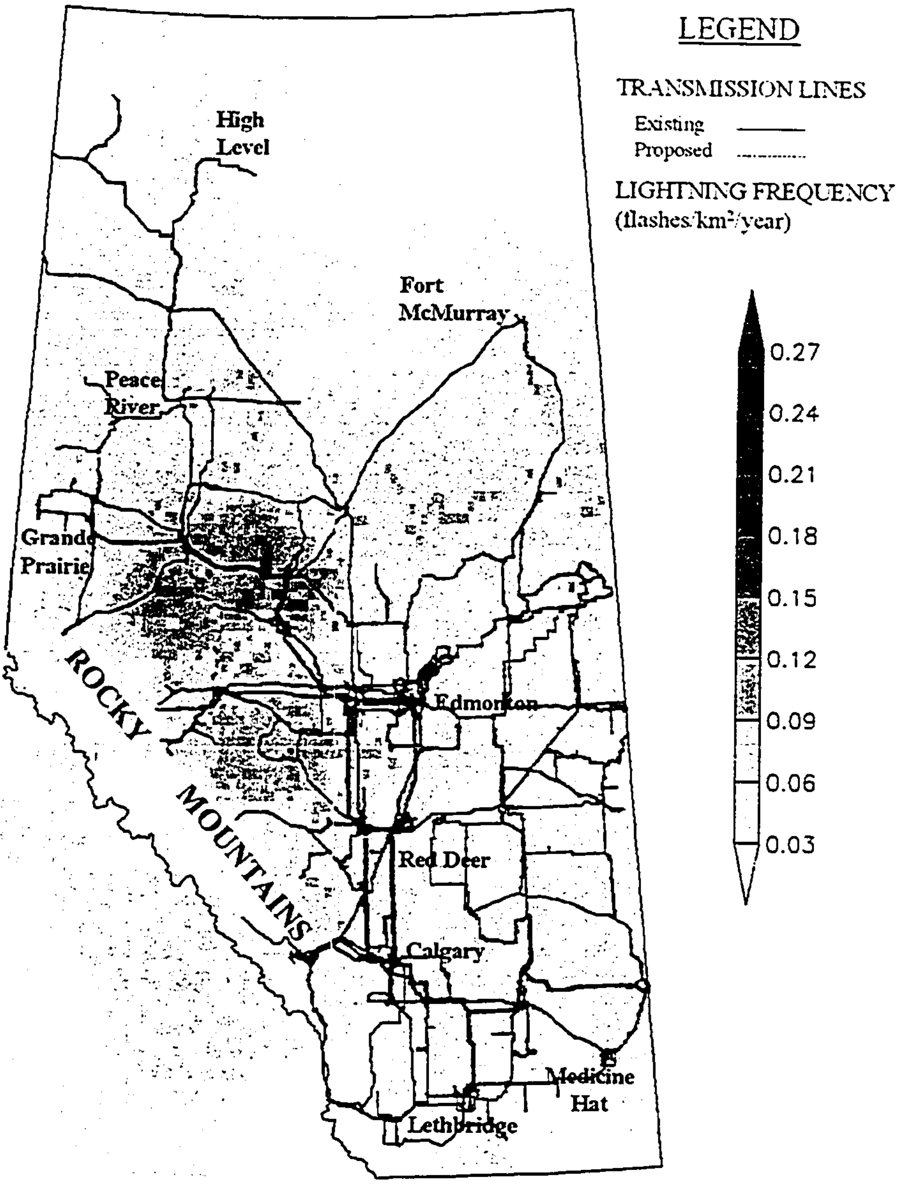


Figure 2-2. The 20-year mean positive lightning flashes per km^2 per year and the power transmission lines in Alberta

The spatial distribution of the positive polarity CG lightning, represented by 20-year mean, is shown in Figure 2-2. A high frequency lightning area can also be found in central Alberta. Compared to the overall lightning pattern, this high frequency lightning area for positive polarity CG lightning is smaller than that for overall lightning and is shifting to the north. The maximum positive lightning frequency, which is over 0.15 flashes/km^2/year, is located around Swan Hills. For other parts of Alberta, except for the far north and far south area, the positive lightning frequency is almost uniformly distributed. In the southern Alberta, where half of the power transmission lines are located, the positive lightning frequencies are mostly under 0.06 flashes/km^2/year. It is interesting to notice that there is another high frequency area around Lethbridge.

The ratio of positive polarity CG flashes to total CG flashes was investigated. The uniformly distributed ratio pattern confirms that the positive lightning frequency and overall lightning frequency have similar spatial distribution. In other words, the positive lightning frequency is proportional to the overall lightning frequency in Alberta.

To measure the year-to-year lightning frequency variability, the standard deviation of the lightning frequency of each grid was calculated. In Figure 2-3, the standard deviation of the 20-year overall lightning frequency was presented. The standard deviation of the overall lightning frequency shows a similar pattern to the mean of overall lightning frequency. A high variation of the overall lightning frequency occurred in central Alberta. This variation indicates that the regions with high mean lightning frequency have larger year-to-year change than the low frequency

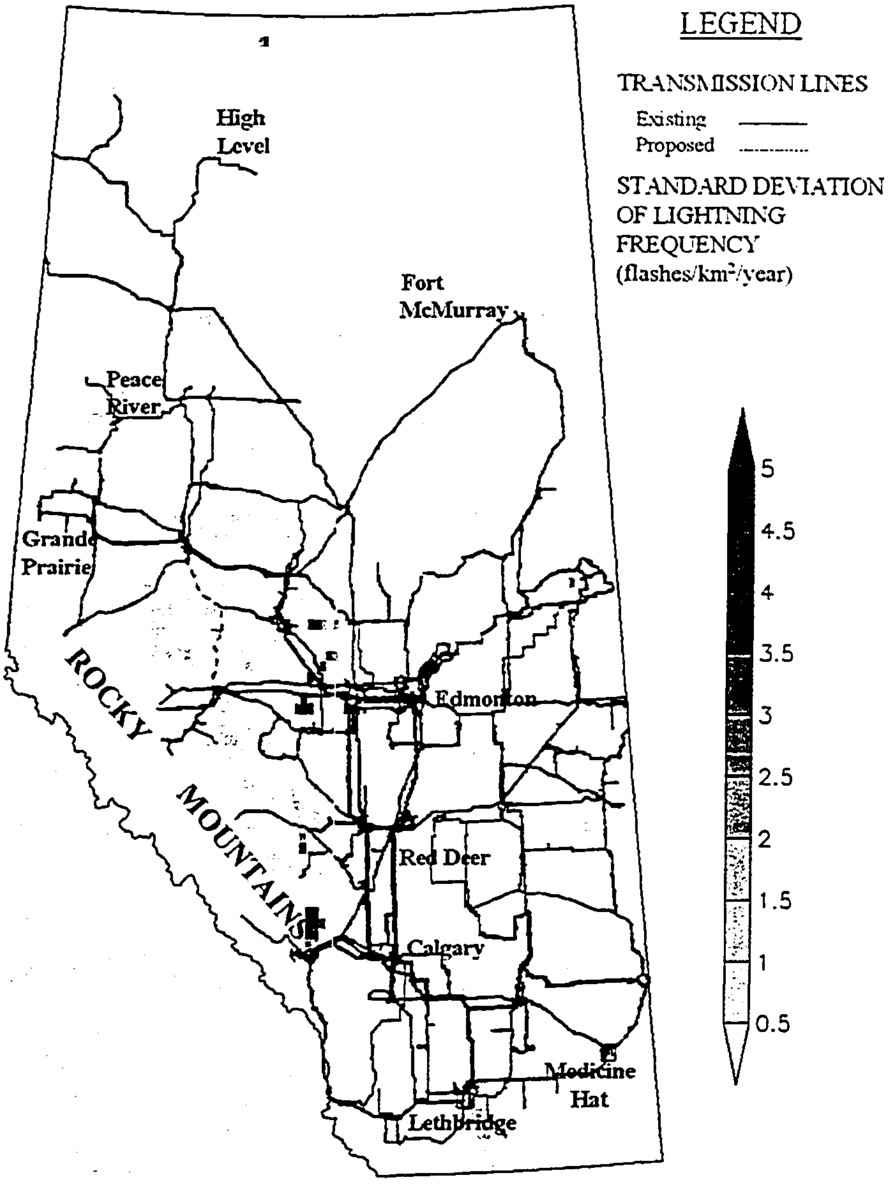


Figure 2-3. The standard deviation of the overall lightning frequency and the power transmission lines during 1983-2002 in Alberta

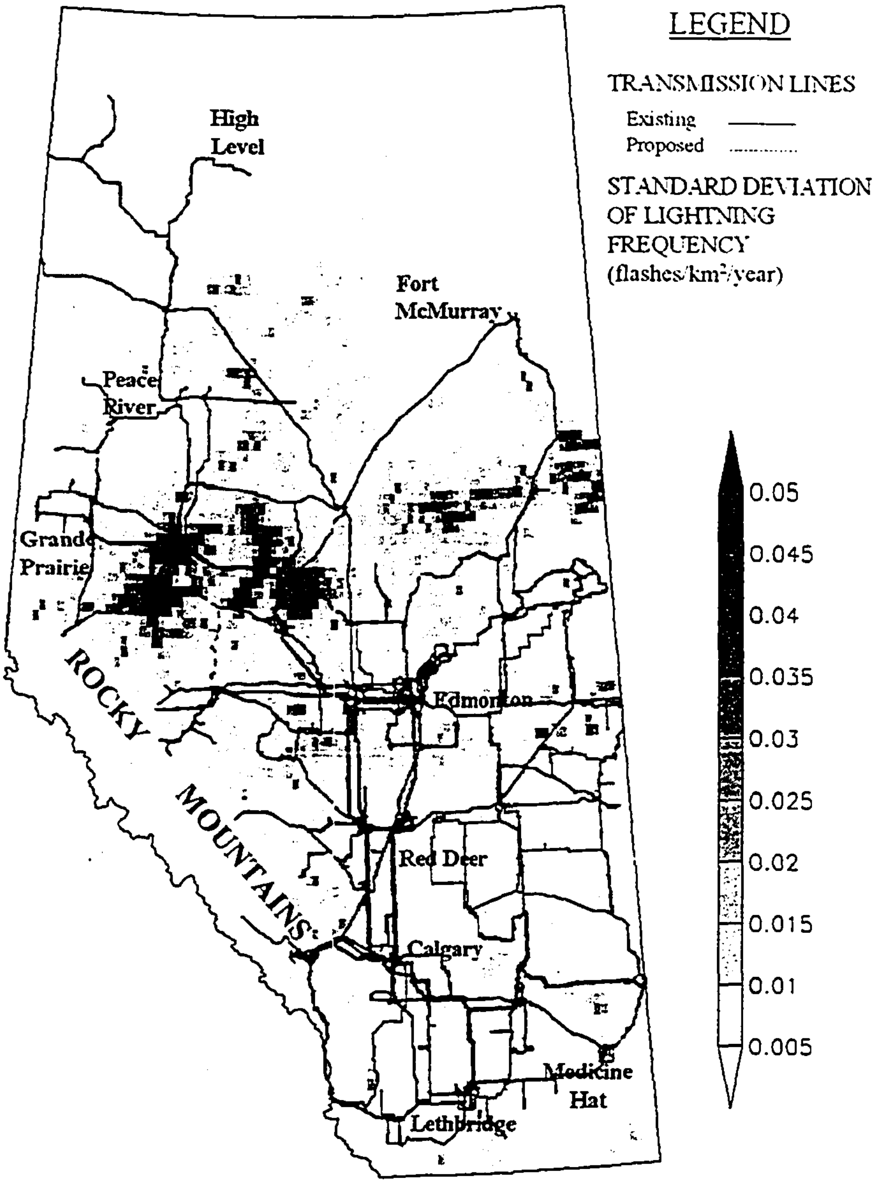


Figure 2-4. The standard deviation of the positive polarity lightning frequency and the power transmission lines during 1983-2002 in Alberta

regions. For the positive polarity lightning, two centers with large variations were found. The standard deviation of 20-year positive polarity lightning (Figure2-4) shows that one center coincides with the high positive lightning frequency region, the other is located to the west of the first center and near the border of Alberta and Saskatchewan. The latter has a large variation in the frequency of positive lightning occurrences but a small 20-year mean of positive lightning frequency. This means there were some years with abnormal positive lightning frequencies in this area.

2.2.3 Temporal Distribution of Lightning Flashes

The annual time series of the number of overall lightning flashes and the number of positive polarity flashes are shown in Figure 2-5. It is found that both time series have increasing trends in the study period. The numbers of both overall lightning flashes and positive lightning flashes are almost doubled after 1986. The ratio of the number of the positive lightning flashes to the number of the overall lightning flashes varies each year during 1983-2002 from 6% to 15% except in 2002, when that ratio jumped up to 39%. However, to find out the reasons for this phenomenon is beyond the scope of this thesis. In 1994, the maxima of the number of both annual overall lightning flashes and the annual positive lightning flashes were found. It is unknown if this phenomenon is due to error in the data or due to weather physics.

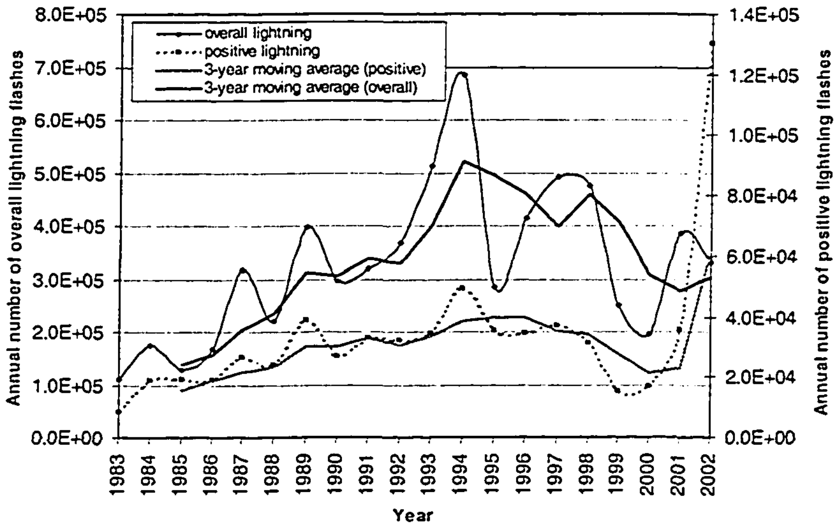


Figure 2-5. The annual time series of the number of overall lightning flashes and the number of positive polarity lightning flashes with moving average trend during 1983-2002 in Alberta

A parameter that can display the temporal distribution of the lightning is days with lightning occurrence. It is the number of days per km^2 per year when one or more flashes were detected in a grid cell. The 20-year mean days with lightning occurrence for the overall lightning are presented in Figure 2-6. It is shown that the pattern is pretty close to the overall lightning frequency pattern. In the central Alberta along the Rocky Mountains, the days with lightning is above 0.25 days/km^2/year. In most parts of northern Alberta and southwestern Alberta, this amount is under 0.15 days/km^2/year. In Figure 2-7 it is shown that the days with lightning occurrence pattern for the positive lightning is not as smooth as that for the total lightning, but the rough outline shows that the two patterns are similar. The maximum days with positive lightning occurrence (over 0.06days/km^2/year) is located in central Alberta,

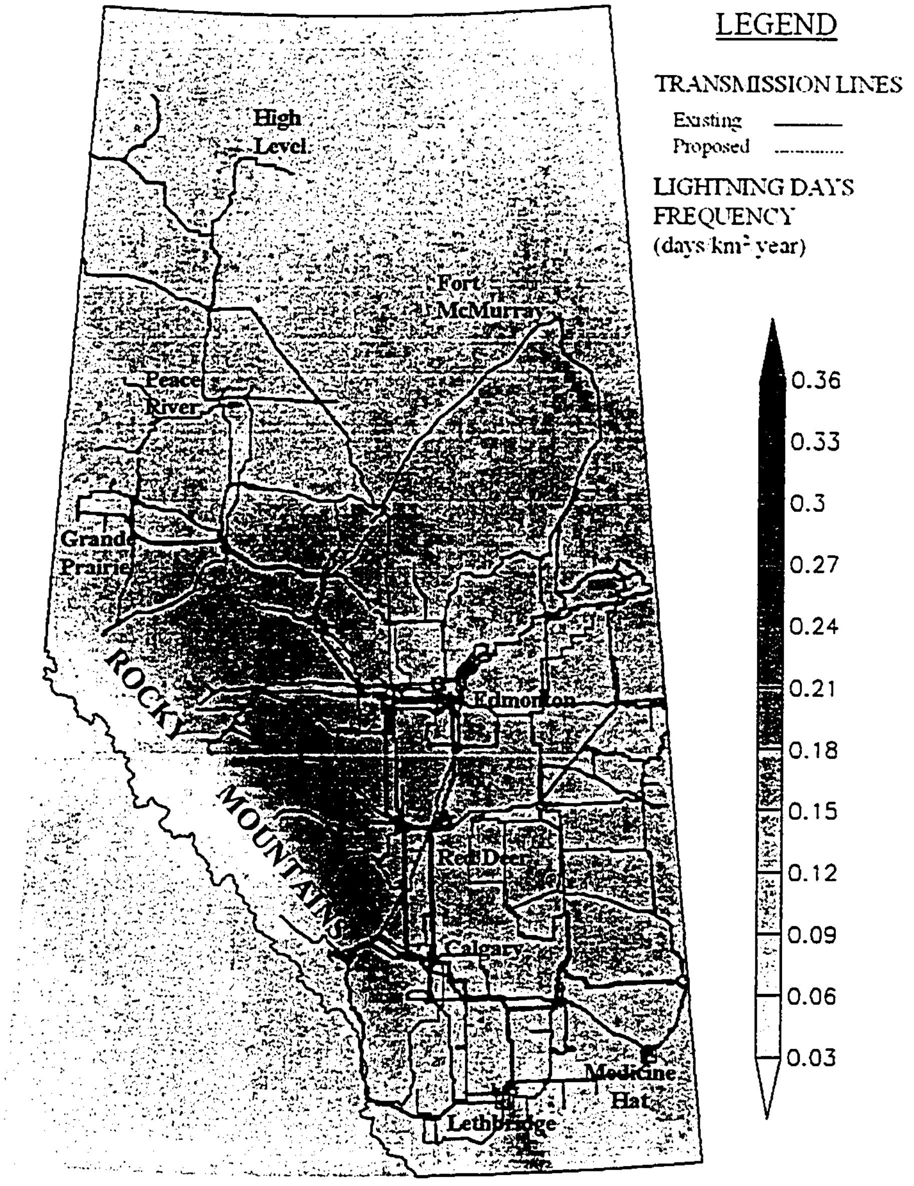


Figure 2-6. The 20-year mean days with overall lightning occurrence and the power transmission lines during 1983-2002 in Alberta

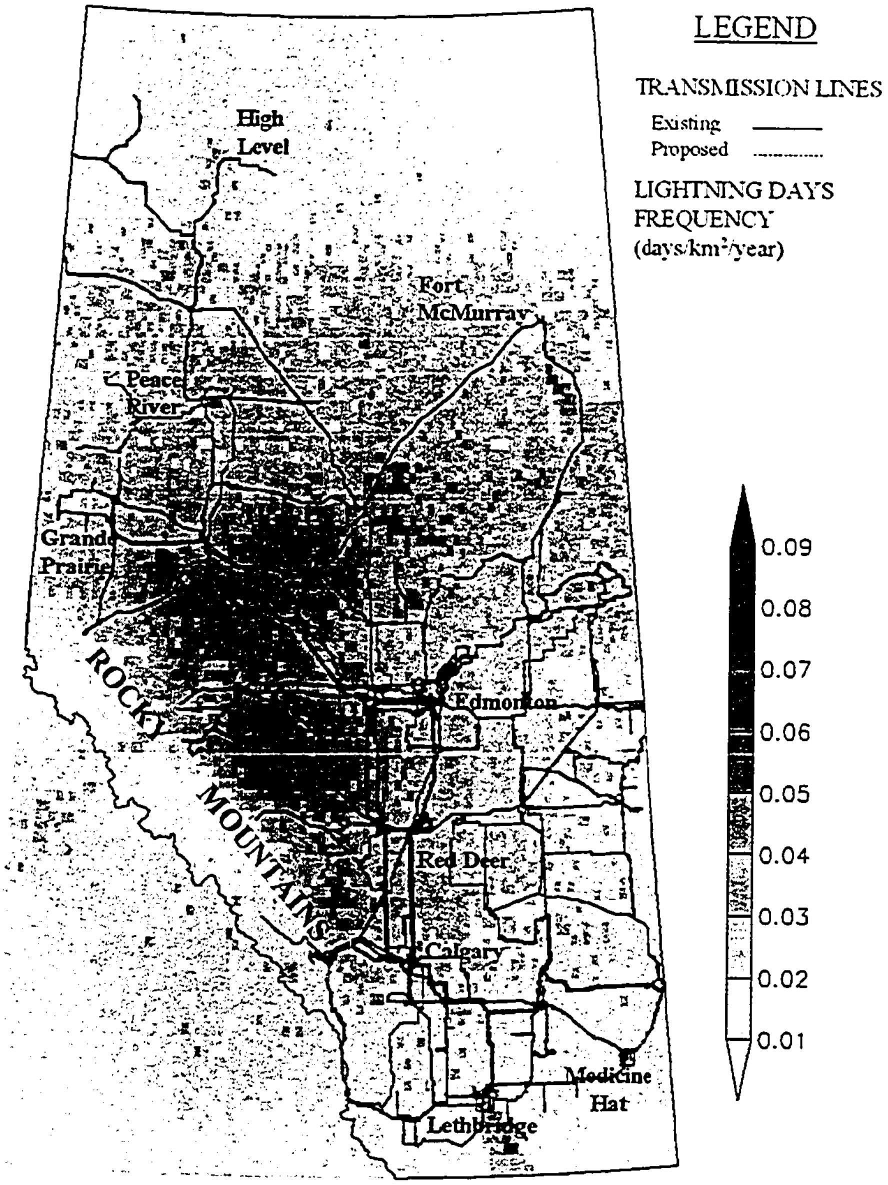


Figure 2-7. The 20-year mean days with positive polarity lightning occurrence and power transmission lines during 1983-2002 in Alberta

with smaller north-south span compared to the days with overall lightning occurrence pattern.

The distribution of the number of monthly total days of CG lightning occurrence in Alberta was found to be non-uniform. The lightning flashes recorded mainly occurred during the summer season and there are only a very small amount of lightning was detected during the winter season.

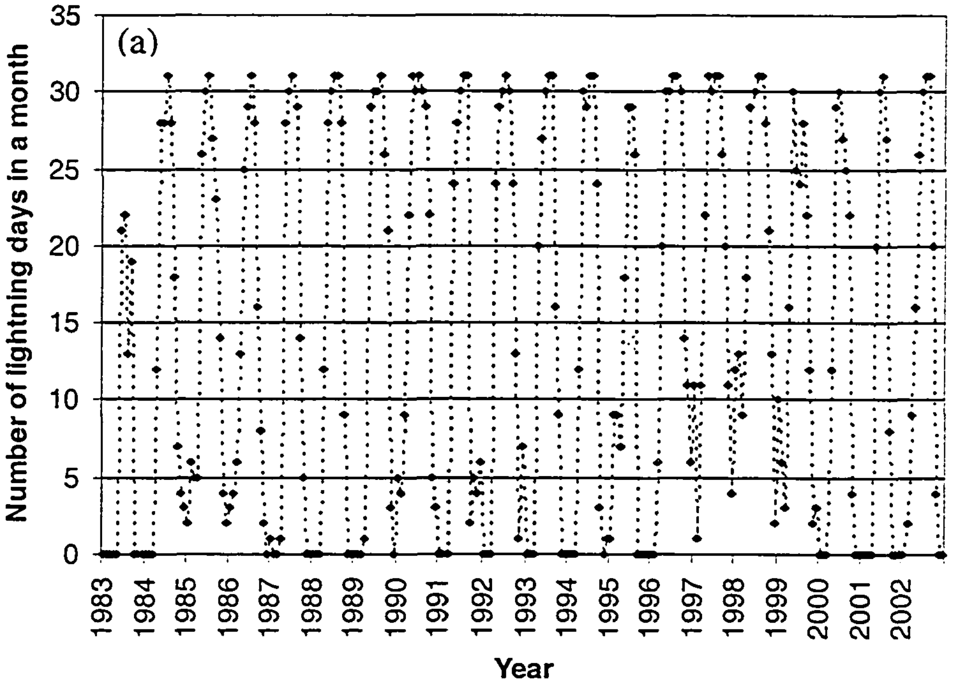
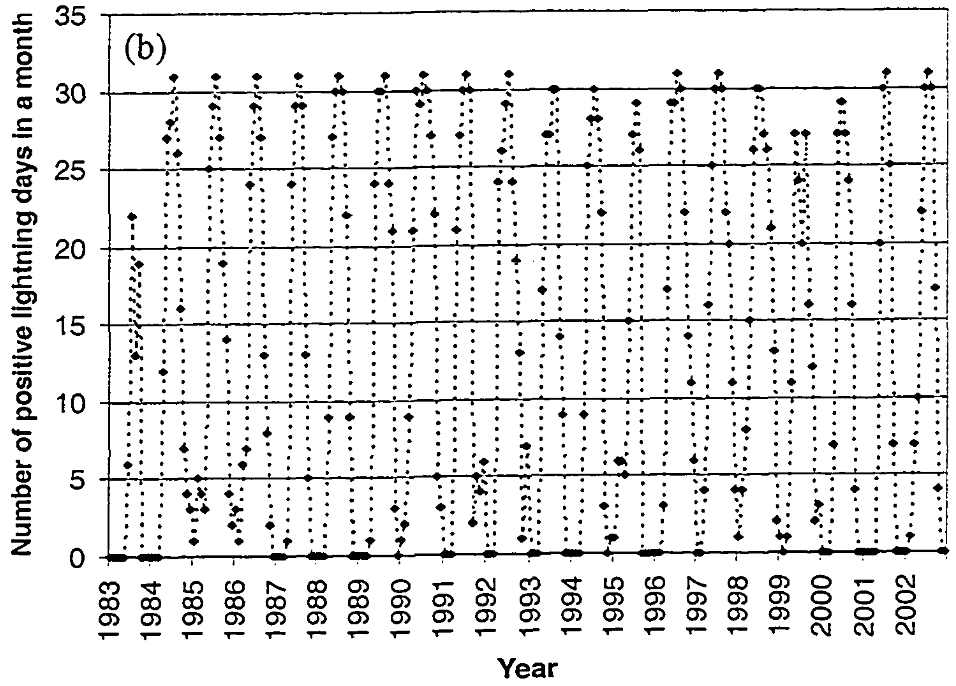


Figure 2-8. Panels (a) and (b) are the time series of the 20-year monthly total days with lightning occurrence for overall lightning and positive lightning. Panels (c) and (d) are the time series of 20-year mean of the monthly total days with lightning occurrence for overall lightning and positive lightning during 1983-2002



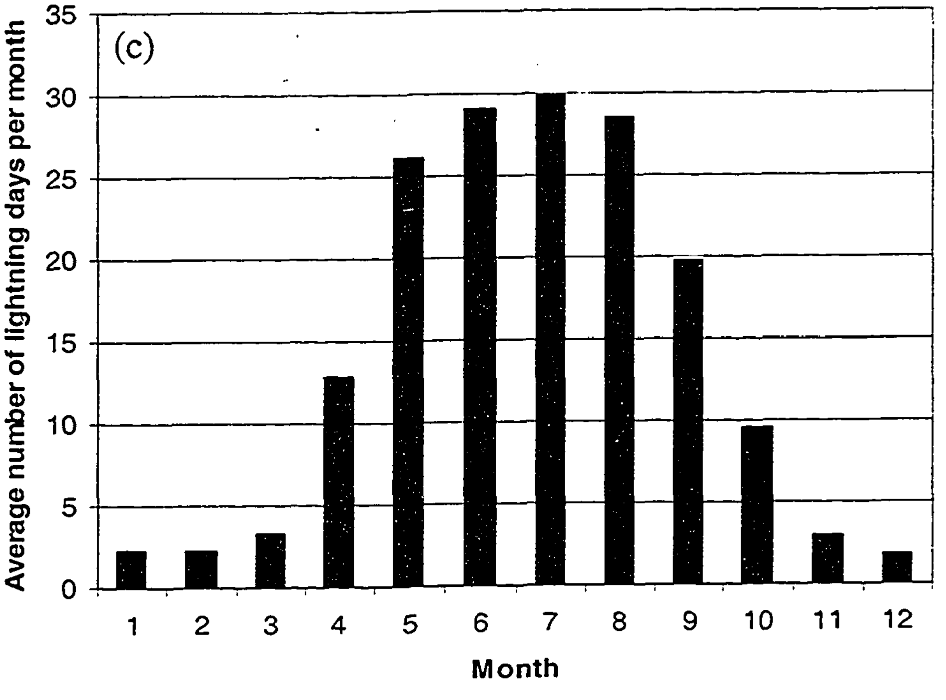


Figure 2-8 continued

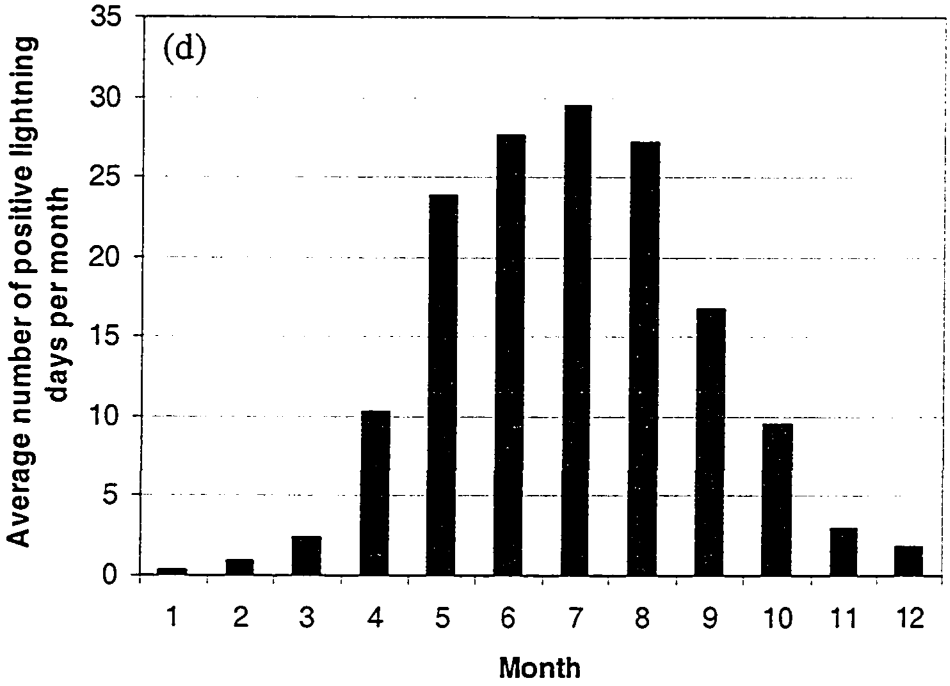


Figure 2-8 continued

The 20-year monthly total days of the lightning occurrence for the overall lightning and positive lightning 1983-2002 are shown in Figures 2-8a and 2-8b and the time series of 20-year monthly average days with the lightning occurrence for both overall lightning and positive lightning during 1983-2000 are shown in Figures 2-8c and 2-8d, respectively. An obvious annual cycle was detected both for the days with overall lightning occurrence and for the days with positive lightning occurrence. During the May to September period there are over or slightly below 20 days of overall lightning occurrences, and it can be defined as the main lightning season in Alberta. Especially in June and July, the days with lightning is over 29, which means that lightning happens somewhere in Alberta almost everyday in these two months.

The standard deviation of the days with lightning occurrence for overall lightning and positive lightning were also analyzed (not shown in this thesis). Similar results as the standard deviation of the frequency of the lightning were obtained. The regions with large number of days with lightning occurrence also have a large variability.

2.2.4 Summary

In summary, the spatial and temporal distributions of the lightning flashes are strongly correlated with geographical features and influenced by elevated terrain features in Alberta. The maxima of both the mean of the density of the annual overall lightning flashes and the mean of the days with overall lightning occurrence are found in central Alberta, along the Rocky Mountains. The regions where the mean of the density of the overall lightning flashes and the mean of the days with the overall lightning occurrence are high also have larger year-to-year variation than other regions. The mean and variance of the frequency of overall lightning flashes are almost equal to each other. Therefore, the Poisson distribution can be used to model the frequency of the overall lightning flashes. Positive lightning flashes are proportional to the overall lightning flashes and have the similar spatial and temporal distributions as that of the overall lightning flashes, but only in smaller magnitude. Increasing trends are detected for both the number of annual overall lightning flashes and the number of annual positive lightning flashes. Days with lightning flashes occurrence reaches a maximum in June and July.

Chapter 3

Forced Power Transmission Line Outages Analysis

3.1 Basics of Power System and Reliability

A power system is a continuous dynamic system involving generators, thousands of transmission lines and distribution networks. Any impromptu response to the utility equipment or component failures may result in the whole power system being out of service even though the rest of the system is still in the normal working state. It is crucial to understand how the power system and its components perform in order to apply the suitable reliability method to establish the applicable reliability evaluation system.

Electrical power is produced by spinning electrical generators driven by various fuel sources, such as nuclear, coal, oil, natural gas, hydro power, geothermal, , wind, etc. No matter what it is that spins the generator, commercial electrical generators of any size generate what is called 3-phase Alternating Current (AC) power.

Bulk power system is dominantly an AC system, as opposed to a Direct Current (DC) system, because of the following reasons:

1. Large electrical generators can economically generate AC power
2. Transformers must have alternating current to operate in order to transform the transmission line voltage levels from one level to another and finally down to the customer's utilization voltage level (e.g., 120V).

As the electricity enters the transmission system it is converted to extremely high voltages (i.e. 230,000 to 765,000 volts) by using large transformers for long-distance transmission on the transmission grid. The purpose of doing so is to reduce the losses of energy from conductor heating and allows power to be shipped economically over long distances. A typical maximum transmission distance is about 300 miles (483 km).

Transmission lines are interconnected at switching stations and substations to form a network of lines and stations. In North America this is commonly called the "grid". And there are three distinct power grids or "interconnections" existed. These three grids are the Eastern Interconnection, the Western interconnection and the interconnection that comprises most of the state of Texas. It was pointed out by the U.S.- Canada Power System Outage Task Force, 2003, that the three interconnections are electrically independent from each other except for a few small direct current (DC) ties that link them.

When electricity is delivered to the remote load points, it is ready to be used in a home or business. The electricity is stepped-down from the transmission grid to the distribution grid. The conversion from the transmission grid to the distribution grid occurs in a power substation. A power substation typically does two or three things:

1. It has transformers that step transmission voltages down to distribution voltages (typically less than 35,000 volts).
2. It has a "bus" that can split the distribution power off in multiple directions.

3. It has circuit breakers and switches so that the substation can be disconnected from the transmission grid or separate distribution lines can be disconnected from the substation when necessary.

North American is operating and maintaining the largest power system in world. It is essential for modern society to have a secure and reliable power system.

Reliability of the power system is defined as the overall ability of the system to perform its function and is usually divided into adequacy and security. Adequacy relates to the existence of sufficient facilities within the system to satisfy consumer load demand. These facilities include the those necessary to generate sufficient energy and those associated transmission and distribution facilities required to transport the energy to the actual consumer load points. Adequacy is, therefore, associated with static conditions, which do not include system disturbances. Security related to the ability of the system to respond to disturbances arising within that system. Security is, therefore, associated with the response of the system to whatever perturbations it is subjected to (Billinton et al., 1996). The perturbations include the conditions associated with both local and widespread disturbances and the loss of major generation and transmission facilities. Security assessment in many cases involves the dynamic behavior of the power system.

In reality, the division is not intended to indicate that there are two distinct processes involved in power system reliability, but is intended to ensure that reliability can be calculated in a simply structured and logical fashion (Shen, 1999). A lot of efficient probabilistic techniques have been developed for the adequacy assessment due to its nature with static conditions. It is easy to define the concept of adequacy in a utility and to calculate accurate indices as an input to the power system reliability assessment system. For the security assessment, it is still a wide-open research area and much more work will be needed in the future.

Power system adequacy indices, which are obtained by assessing past system performance, encompass the effect of all system faults and failures irrespective of cause and, therefore, include the effect of insecurity as well as those due to inadequacy (Shen, 1999). In this thesis, the basic adequacy indices to evaluate weather-related transmission line outages in Alberta are presented.

3.2 The Forced Outage Data of the Alberta Power Transmission Lines

In 1975, the Canadian Electrical Association (CEA) adopted a proposal to create a facility for centralized collection, processing and reporting of reliability and outage statistics for electrical generation, transmission and distribution equipment. To coordinate the development for this Equipment Reliability Information CEA constituted the consultative Committee on Outage Statistics. In 1978, the transmission stage of the information system was implemented when Canadian utilities began supplying data on transmission equipment in accordance with the Instruction Manual for Reporting Component Forced Outages of Transmission Equipment (Koval, 1996).

The Instruction Manual for Reporting Component Forced Outages of Transmission Equipment also outlines the scheme of the procedure of force transmission outages collection and basic data structure. It provides a guide for the Canadian utilities to develop a unique scheme of transmission force outages data collection and storage to fulfill their own special needs for the transmission system reliability evaluation. A 19-year (1977-1995) forced power transmission equipment outage data used by this thesis is the power transmission equipment forced outage raw data set provided by ATCO Electric Limited, which is a data set reporting the transmission equipment forced outages in Alberta during the period of 1977 to 1995.

A forced outage can be described as an outage that results from emergency conditions directly associated with faulty components, requiring that they be taken out of service immediately. This is done automatically or as soon as switching operations can be performed, or an outage caused by improper operation of equipment or human error (Shen, 1999). The major classification criterion for transmission forced outages used by utilities is outage duration. By this criterion two types of forced outages, “sustained” and “transient” are defined. A “sustained” forced outage refers to a transmission forced outage, the duration of which is one minute or more. It does, therefore, not include automatic reclosing events. A “transient” forced outage refers to a transmission outage, the duration of which is less than one minute (Electrical Power Research Institute, May 1985). Sustained forced outages are further divided into “line-related” and “terminal- related” forced outages by the component where the forced outages occur while transient forced outages are usually only defined in terms of “line-related” forced outages.

The ATCO Electric Limited power transmission equipment forced outages raw data set recorded the performance of the power transmission equipments failure for three main voltage levels (72kV, 144kV and 240kV) from January 01, 1977 to December 31, 1995 in Alberta. There are four-voltage levels transmission system in

Alberta, 72kV, 144kV, 240kV and 500kV. The equipment failure data for 500kV transmission lines were not available for this thesis.

The transmission outages are recorded chronologically according to different transmission components (e.g., Line, Circuit Breaker, Transformer or DC Station, etc.). Every outage event has a detailed description by date, time, line number, duration of the outage, the end substation of the transmission line, line-related or terminal-related, voltage level, count, tower, CEA primary cause, mode and origin. This detailed description makes it possible to analysis the transmission line outages from many different perspectives so that the complex nature of the outages can be reflected by the indices calculated by the data.

To understand the data, it is necessary to define the data base structure of transmission line performance data. Shen (1999) categorized the data as shown in Figure 3-1. The dataset first was divided by three voltage levels in kV, which were 72kV, 144kV and 240kV. Then it was subdivided into sustained forced outages and transient forced outages. The "line-related" and "terminal-related" forced outages are further subdivided into primary caused and sub component categories.

The primary causes of "line-related" transmission line forced outages include system condition, human element, defective equipment, foreign interference, adverse environment, consequential and unknown by the CEA classification. One of the primary causes of the transmission line-related outage, according to CEA definition, is adverse weather, which includes wind, precipitation (ice, wet snow and frost) and lightning.

A Microsoft Access database for the transmission line outages data was established according to the above structure. The latitude and longitude location of the transmission lines are measured by using the Alberta Power System Map. The geographical location data was appended to the transmission line outages raw data.

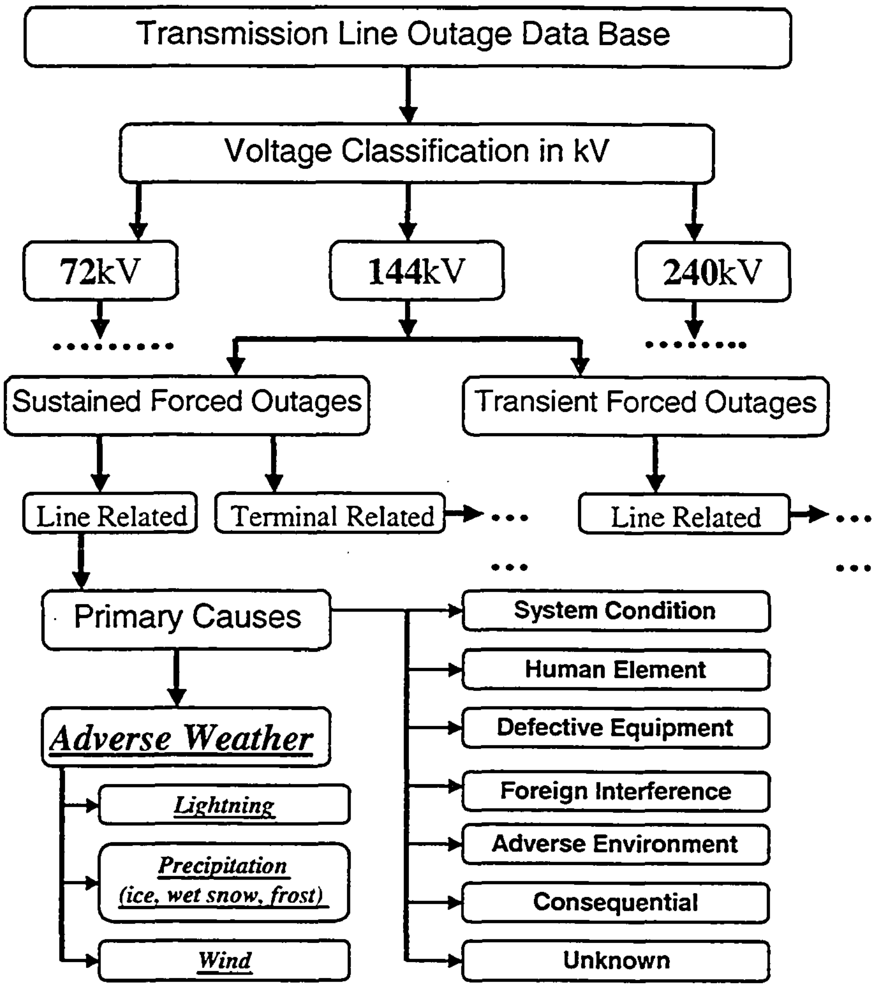


Figure 3-1 Alberta transmission line outage database structure

3.3 Summary Statistics of Weather-Related Transmission Line Outages

The primary causes of transmission line outages are adverse environment, adverse weather, consequential, defective equipment, foreign interference, human element, system condition and unknown. Each cause uniquely affects the transmission line outage state. Two parameters, frequency and duration of the outages, are generally used to measure to what extent these causes result in the power transmission line failure.

The frequency and duration of Alberta power transmission line outages stratified by the CEA primary causes are shown in Figures 3-2 and 3-3. The first three major causes of the transmission system outages based on the frequencies and duration of outages are defective equipment, human element and adverse weather. Adverse weather is the third primary cause of transmission system outages in term of frequency, which is 343 of the total 1936 times. Adverse weather caused 17.71% of all the transmission outages in Alberta. In terms of outages duration, the repair time of the adverse weather caused transmission outages is shorter than defective equipment caused transmission outages. There are 6,075.68 hours of adverse weather caused transmission outages, which represents 26.74% of the total transmission line-related outages duration, from 1977 to 1995 in Alberta.

Both in the frequency and duration domain, it has been observed that adverse weather has a large effect on the power transmission system operation. The study to reveal the unique features of adverse weather caused line-related transmission outage is based on this fact.

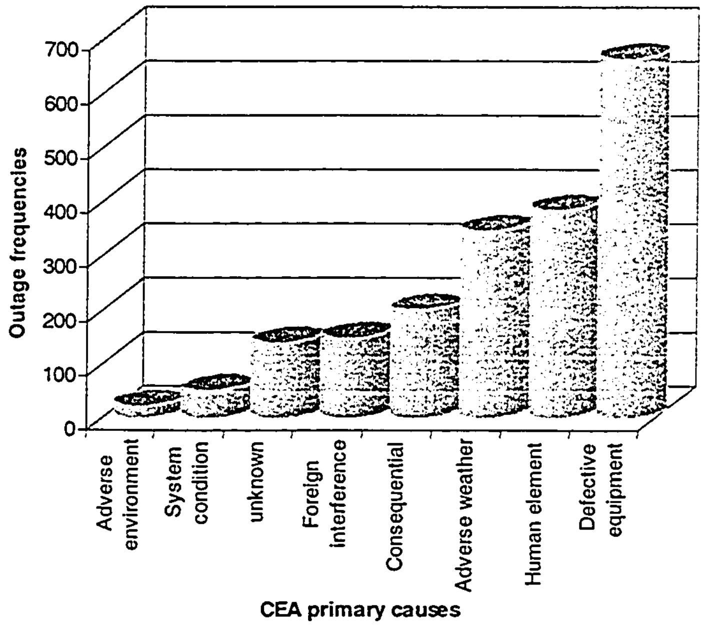


Figure 3-2. Alberta power transmission system outages frequencies stratified by CEA primary causes from 1977 to 1995

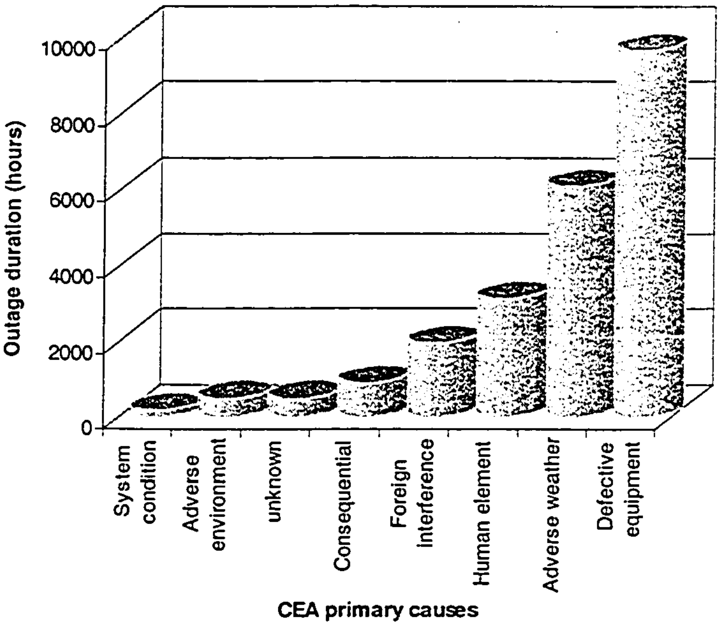


Figure 3-3. Alberta power transmission system outages duration stratified by CEA primary causes from 1977 to 1995

In the adverse weather caused transmission outages category, line-related outages are considered to be the most general case and this thesis will place an emphasis on this domain. The major transmission lines in the data set are in three voltage levels (i.e., 72kV, 144kV and 240kV). It is natural to view the transmission line outages from different voltage levels. The frequency of sustained and transient weather-related transmission line-related outages classified by voltage level elements is shown in Table 3-1.The frequency of sustained weather-related line-related outages of 72 kV lines is larger than the sum of the other two considered voltage levels for transmission lines (144 kV and 240 kV). The 240kV transmission lines are the lines with the least probability to be damaged by the adverse weather. Transient line-related outages were only found on 240kV transmission lines.

Table 3-1. The frequency of forced line-related adverse weather caused transmission line outages classified by voltage level

Voltages	Frequency – outages of 19 years	
	Sustained outages	Transient outages
72kV	163	0
144kV	90	0
240kV	66	2

Wind, lightning, precipitation (ice, wet snow and frost) are specified by CEA as the adverse weather conditions which caused transmission line outages. It is shown in Table 3-2 that the frequency of sustained line-related transmission line outage caused by wind is 110 times, which is two times more than that caused by lightning and nine times more than that caused by ice, wet snow and frost in the 19-year period. In the frequency domain, the outage frequencies under the three weather conditions are similar to each other. There is no weather condition that can be specified as the major dominant weather condition causing the transmission line outages. The two recorded transient outages were only caused by lightning. It is noted that the weather events are dependent on the geographical location of the power transmission lines. This result should only be used in power transmission system evaluation within Alberta and cannot be extended to other areas.

Table 3-2. The frequency of forced line-related adverse weather caused transmission line outages classified by weather elements

	Frequency – outages of 19 years	
Weather Elements	Sustained outage	Transient outage
Wind	110	0
Lightning	108	2
Ice, wet snow, frost	101	0

The weather events highly depend on the geographical feature so the adverse weather-related transmission line outages are highly correlated with the geographical distribution of the power transmission system. The Province of Alberta was divided into four distinctive geographical regions to investigate the different features of the adverse weather cause line-related transmission line outages in different areas. The results are shown in Table 3-3. Most of the outages occurred in Southeastern Alberta and there were no weather caused line-related transmission line outages in Southwestern Alberta. It took most of the time to repair the transmission lines damaged by ice, wet snow and frost in the southeast region. This is also the case in the northwest region. However, in northeast Alberta, the wind is the dominant weather element which caused the most outages.

Table 3-3. The geography distribution of sustained line-related transmission line outage duration (hours) stratified by adverse weather causes

Region	Ice, wet snow, frost	Lightning	Wind	Percentages
Northeast	1.7 h	37.05 h	54.57 h	3.29
Northwest	177.32 h	99.27 h	85.72 h	12.77
Southeast	1268.05 h	238.17 h	874.65 h	83.94

The adverse weather caused line-related transmission line outages are not uniformly distributed in the time domain. It is important to obtain the knowledge of the variance of the adverse weather caused line-related power transmission line outages over time so that it provides the base for the utilities for future outages prediction.

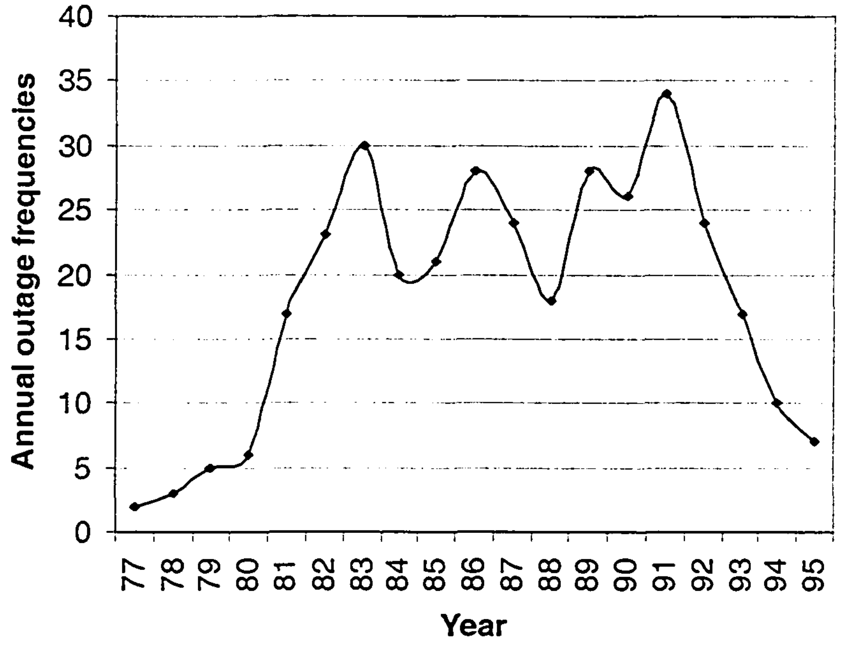


Figure 3-4. The annual frequency of adverse weather caused line-related transmission line outages from 1977 to 1995

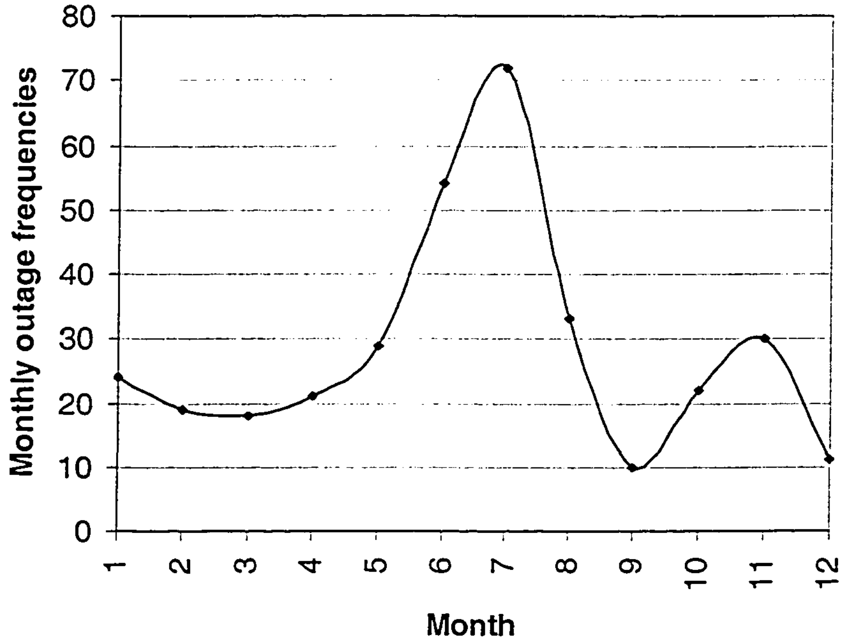


Figure 3-5. The monthly frequency of adverse weather caused line-related transmission line outages from 1977 to 1995

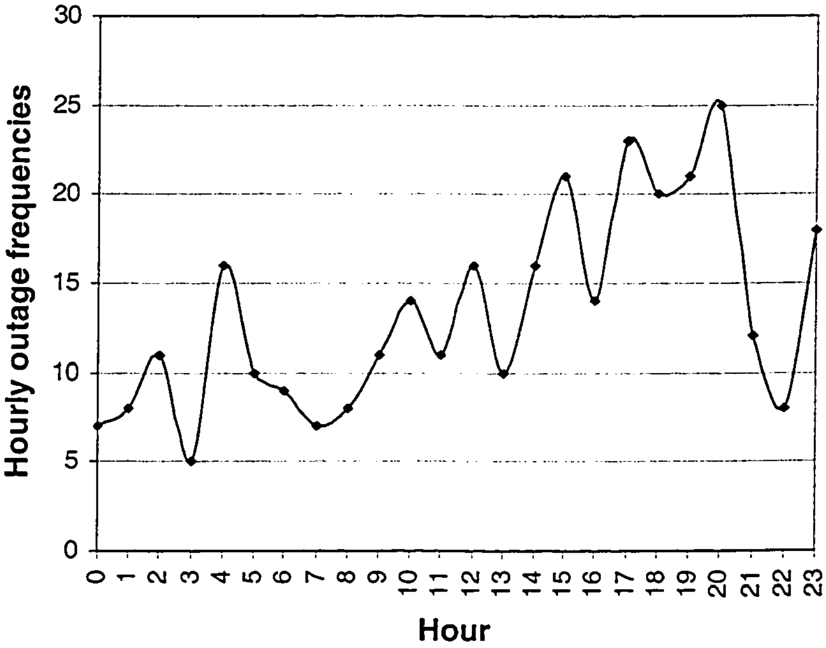


Figure 3-6. The hourly frequency of adverse weather caused line-related transmission line outages from 1977 to 1995

The time series of outage frequencies (e.g. annual, monthly, hourly) of line-related transmission line outages are shown in Figures 3-4 to 3-6. From 1977 to 1980, the adverse weather-related outages are in low frequency. The outage frequency is the lowest in 1977, which was 2 times. From 1981, the outages became high frequent events with small annual fluctuations and the frequency reached the peak in 1991, which was 34 times. The decreasing trend of outage occurrence was detected after 1991. There is a roughly increasing trend of weather related outages from the winter season to the summer season. July with 62 outages and June with 49 outages are the two months with the highest number of transmission line outages. After the summer season, the frequencies of the outages reach a minimum in September with 10 outages. From midnight to 9:00 AM, the frequencies of outages are symmetrically distributed with peaks occurring between 4:00 to 5:00 AM. After 9:00AM, the frequencies of outages go up until they reach the maximum between 8:00 to 9:00 PM. The period with low outage frequencies are between 3:00 to 4:00 AM. It is more likely that transmission line outages recorded in the afternoon than in the morning.

The correlation analysis of physical length of transmission line and adverse weather caused line-related transmission line sustained outages is presented in Table 3-4. Wind and precipitation-caused transmission line outages are not proportional to physical length of the transmission line. However, lightning-caused outages are correlated with the physical length of the transmission lines.

Table 3-4. Correlation of transmission line physical length and adverse caused transmission line sustained outage duration time

	Wind	Lightning	Ice, wet snow and frost
Line length	0.018	0.143	0.046

In the following sections, the detailed analysis of adverse weather caused line-related transmission line outages stratified by three adverse weather conditions will be presented. Transient outages are omitted in the analysis due to the fact that the duration of transient outages are very short (e.g., less than one minute) and the frequencies of transient outages are very low. The weather information around the outage occurrence areas was retrieved from the weather database to evaluate the effects of weather on the power transmission line outages. The weather information includes maximum daily air temperature, minimum daily air temperature, maximum daily precipitation and maximum hourly wind speed. The maximum daily air temperature is the maximum of the daily maximum air temperature observed in all the weather stations in the power outage area; the minimum daily air temperature is the minimum of the daily minimum air temperature observed in all the weather stations in the power outage area; the maximum daily precipitation is the maximum of the daily precipitation observed in all the weather stations in the power outage area; and the maximum hourly wind speed in the maximum of the maximum hourly wind speed observed in all the weather stations in the power outage area.

3.4 Precipitation (Ice, Wet Snow and Frost)-Caused Line-Related Sustained Power Outages

Most of the current power transmission system reliability evaluation systems are based on some assumption of the underlying probability distribution of the duration of the outage events. The assumptions have to be carefully checked before establishing a real time applicable reliability evaluation system in the power industry. It is important to estimate and analyze the probability distribution of the duration of the outage events by using the raw outage data so that the unique characteristic of the outage events can be well understood. Since there are different transmission line configurations for different voltage levels, the probability distributions of the duration of the power transmission line sustained outage in Alberta were analyzed according to their voltage levels.

The estimated probability density functions of the duration of precipitation-caused line-related sustained power transmission line outages classified by their voltage levels (i.e. 72kV, 144kV and 240kV) in Alberta are plotted in Figures 3-7 to 3-9. The number of data for estimating the probability density functions of the duration of precipitation-caused line-related sustained transmission line outages for the 72kV, 144kV and 240kV transmission lines are 39, 46 and 16 respectively. The little bumps of the estimated probability functions at the large end of the outage duration axis are due to some extreme long duration outages. The precipitation-caused outage, particularly in this thesis, refers to ice, wet-snow and frost caused outage. The shortest precipitation-caused transmission line outages, 1 minute, occurred on a 72kV transmission line and three 144kV transmission lines and the longest precipitation-caused transmission line outage occurred on a 240kV transmission line. There is a high probability that a long repair time may be needed for the outages on 240kV lines. The average duration of precipitation-caused outages on 72kV and 240kV lines are longer than that on 144kV lines. It is noted that the mean and median of the duration of the precipitation-caused transmission line outage on each voltage level lines are different and the underlying probability distributions are skewed to the right. The underlying distribution of the duration of the outages on 72kV voltage transmission lines has the largest skewness among that of the three voltage levels transmission lines. This result may challenge the reliability evaluation models which assume that the probability distribution of the duration of the outages are in the symmetric probability distribution family.

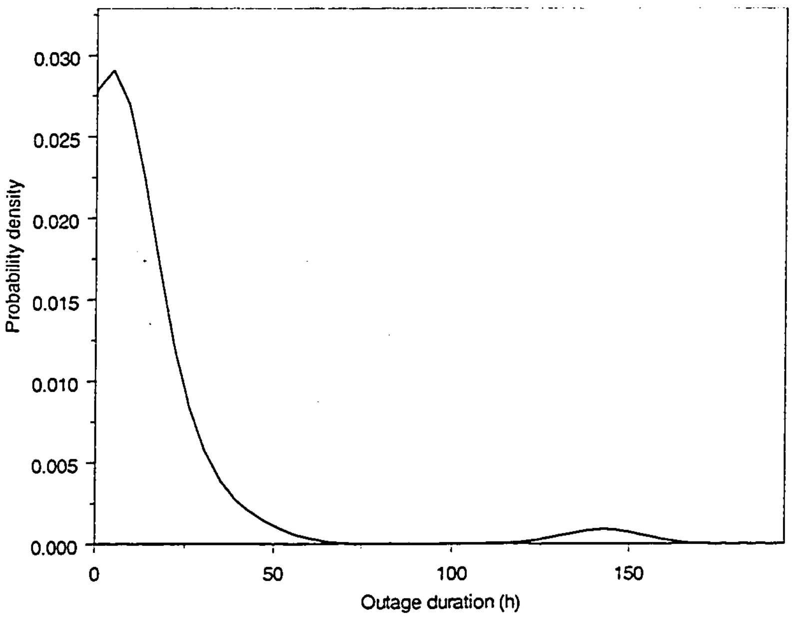


Figure 3-7. The probability density function of outage duration of precipitation-caused line-related sustained power transmission line outages on 72kV voltage lines estimated by 1977 – 1995 ATCO power outages data

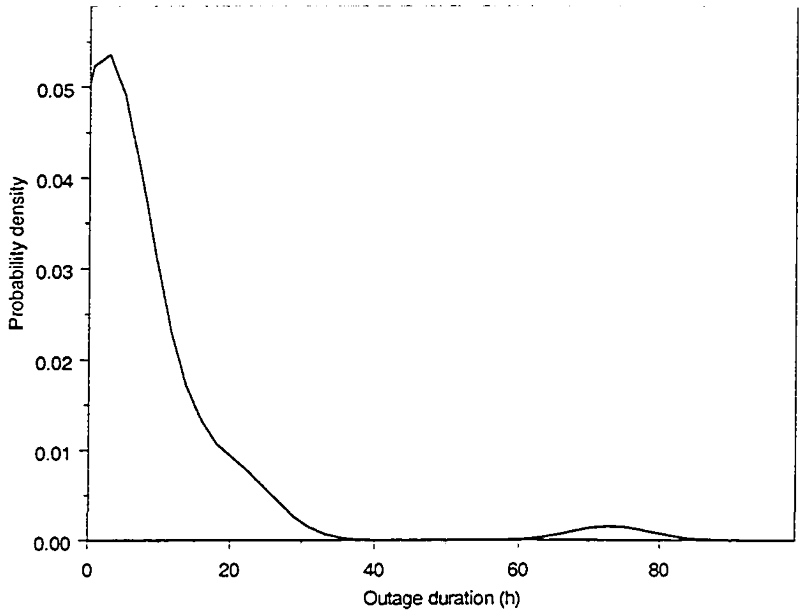


Figure 3-8. The probability density function of outage duration of precipitation-caused line-related sustained power transmission line outages on 144kV voltage lines estimated by 1977 – 1995 ATCO power outages data

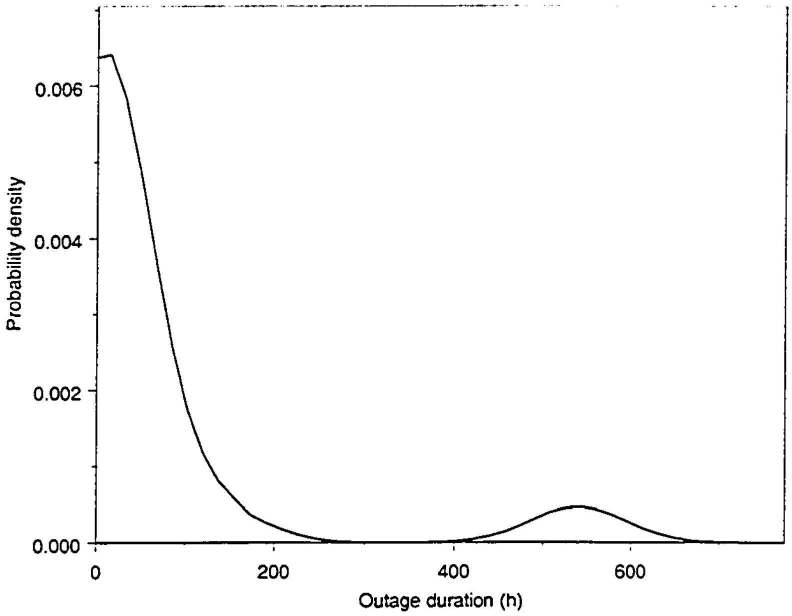


Figure 3-9. The probability density function of outage duration of precipitation-caused line-related sustained power transmission line outages on 240kV voltage lines estimated by 1977 – 1995 ATCO power outages

The annual total duration of the precipitation-caused line-related sustained outages classified by three voltage levels are shown in Figure 3-10. The precipitation-caused outages on 72kV and 144kV voltage transmission lines occurred in high frequency during the study period. In over 50% of the years, there were precipitation-caused outages that took place on these two voltage levels lines. For 240kV voltage transmission lines, the precipitation-caused outages were only found in 6 continuous years starting from 1983. The maximum, 163 hours, of the annual total duration of the precipitation-caused outages of the 72kV transmission lines was found in 1992. In the same year, the long precipitation-caused outages, 84 hours, also occurred on the 144kV lines. In 1983, the annual total duration of the precipitation-caused outages reached its peak, 96 hours, on the 144kV transmission lines. For the 240kV transmission lines, two long outages, 540 hours and 133 hours, were recorded in 1986. Without these two outage events, the average of the annul total duration of the precipitation-caused outages computed by average the annual total duration of the outages by the number of years with outage occurrence for the three voltages transmission lines were of the same order of magnitude. There are no precipitation-caused outages recorded before 1980.

Figure 3-11 is the monthly time series of the monthly total duration of the precipitation-caused outages classified by three voltage levels. The precipitation-caused outages of 72kV transmission lines were found in every month of a year except June, July and August. For 72kV transmission lines, the longest monthly total duration, 154 hours, of precipitation-caused outages occurred in February and the outage durations in March, April, May and November also exceeded 40 hours.

Compared to the outage durations in these months, the durations of outages found in the other month were short.

For the 144kV transmission lines, June and July are the two months without any precipitation-caused outages occurred. It is in November that the monthly total duration of the precipitation-caused outages reached a maximum of 119 hours. Two other outages with long duration were found in February and October. Other than these three months, the monthly total outage durations are short.

There were only six months recorded with precipitation-caused outages on the 240kV transmission lines. For this voltage level, the precipitation-caused outages are clustered in the winter and spring seasons and there are no precipitation-caused outages during the summer. It was noted that in May there were 700 hours of outages on 240kV transmission lines. This duration is the longest monthly total outage duration among the three voltage level transmission lines.

The hourly time series of the hourly total durations of the precipitation-caused outages is presented in Figure 3-12. For the 72kV transmission lines, the one-hour period with the longest duration of the outages is between 14:00 to 15:00. Midnight (i.e. from 23:00 to 24:00) is the period with the second longest duration of outages. There are 6 one-hour periods out of 24 without any precipitation-caused outages of 72kV transmission lines.

For the 144kV transmission lines, the one-hour period with the longest duration of the outages was from 8:00 to 9:00. It is interesting to point out that this period was also the period with long outage durations for the 72kV transmission lines. There are 8 one-hour periods out of 24 without any precipitation-caused outages of

144kV transmission lines. Most of the durations of the outages in the one-hour period with outages occurrence were in the same order of magnitude for 144kV transmission lines.

The occurrence of the precipitation-caused outages with long duration of the 240kV transmission lines were clustered between 23:00 to 1:00. The magnitude of the outage durations is larger than the outage duration of the other two voltage level transmission lines. There were only 12 one-hour periods out of 24 when the precipitation-caused outages were observed on 240kV transmission lines.

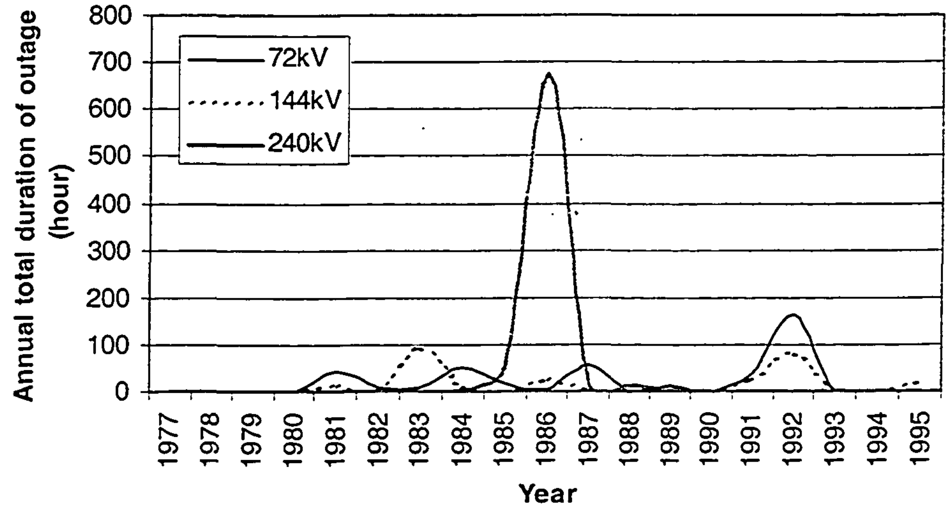


Figure 3-10. The annual time series of the annual total duration of the precipitation-caused line-related sustained power transmission line outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

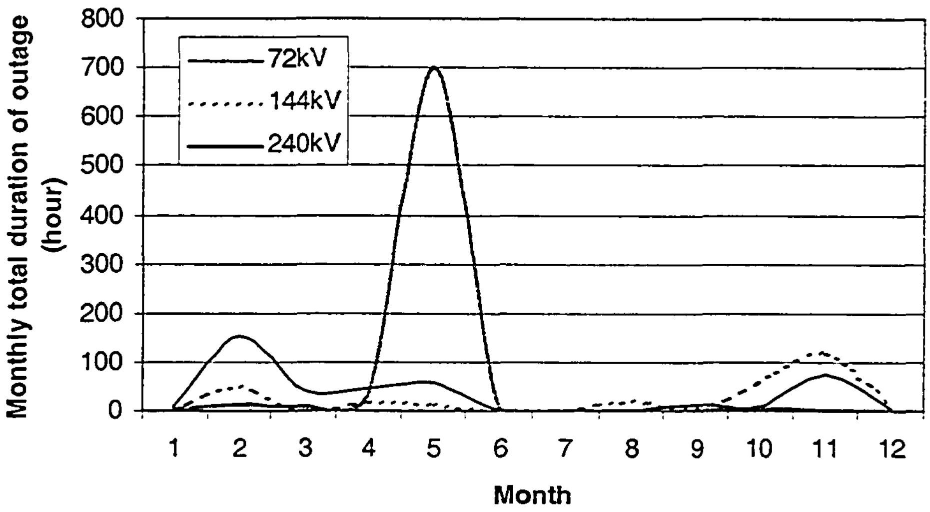


Figure 3-11. The monthly time series of the monthly total duration of the precipitation-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

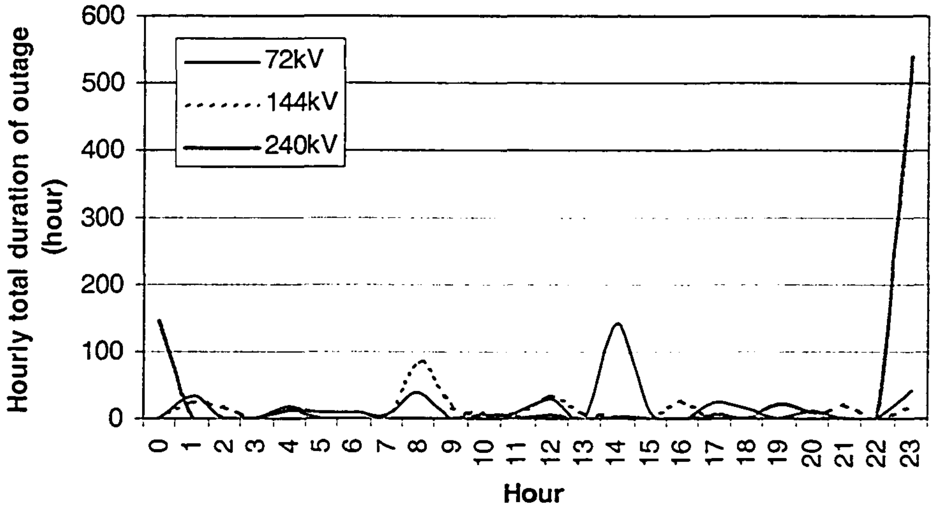


Figure 3-12. The hourly time series of the hourly total duration of the precipitation-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

The daily maximum and daily minimum air temperature data in the precipitation-caused outages occurrence area were retrieved from the weather database. The daily mean air temperature is the simple average of the daily maximum and daily minimum air temperature. The daily mean air temperature versus cumulative proportion of the total duration of the precipitation-caused transmission line outages classified by three voltage levels are presented in Figures 3-13 to 3-15.

For the 72kV transmission lines, there were 72% of the precipitation-caused outages in terms of outage duration occurred under mean daily air temperature zero °C. There is no outage recorded with mean daily air temperature below –20 °C. In the Figure 3-13, a steep slope is shown in the curve and this indicates that there were some outages with long outage duration that occurred under the air temperature between –11.5 °C and –10.8 °C. The maximal mean daily air temperature with precipitation-caused outage occurrence at this voltage level lines is about 8 °C.

For the 144kV transmission lines, the mean air temperature with precipitation-caused outage occurrence has the similar range (i.e. –20 °C to 10 °C) as that for the 72kV transmission lines. The only precipitation-caused outage with mean air temperature out of this range was an outage with mean air temperature around 13 °C. The total duration of precipitation-caused outage with mean daily air temperature below 0 °C count for about 75% of the total outage duration on this level lines. The total outage duration with mean air temperature around –18.5 °C accounted for 25% of the total outage duration at this voltage level and indicates that the 144kV voltage lines can be severely damaged under weather conditions with precipitation and the mean daily air temperature around –18.5 °C. And another range of mean daily air temperature with long durations of precipitation-caused outage is between –8.3°C to –4.8 °C

For the 240kV transmission lines, the range of mean daily air temperature with precipitation-caused outage occurrence was from –10 °C to 11 °C and this range was narrower than that of the other two voltage levels transmission lines. The long durations of precipitation-caused outage was found under the mean daily air temperature between 9.3°C to 11.0 °C and the total duration of the outage under this mean air temperature accounted for almost 72% of the total duration of the outage for this voltage transmission lines. The reason for this phenomenon is interesting and will be further investigated.

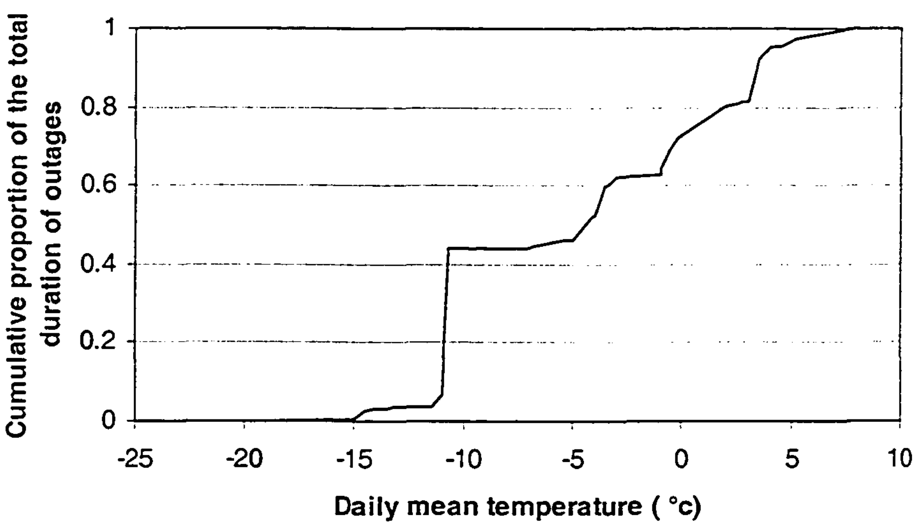


Figure 3-13. The daily mean air temperature vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outage on 72kV transmission lines in Alberta from 1977 to 1995

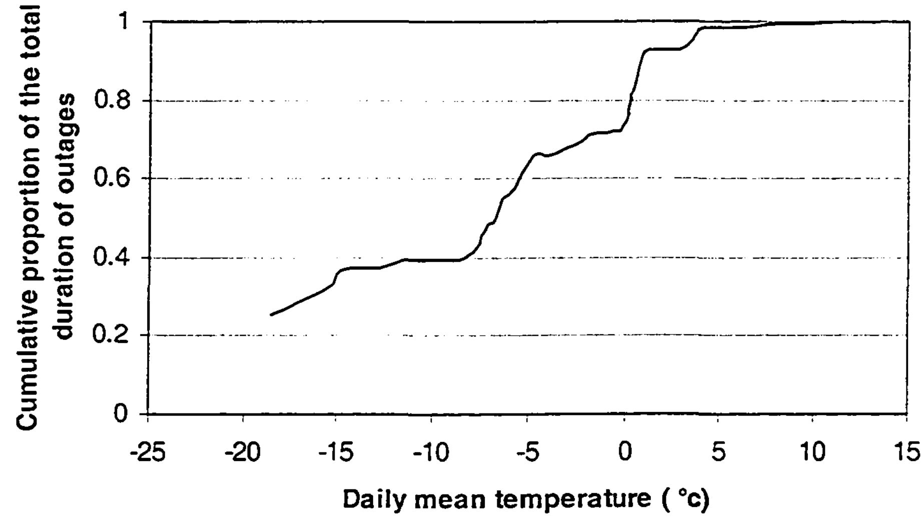


Figure 3-14. The daily mean air temperature vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outage on 144kV transmission lines in Alberta from 1977 to 1995

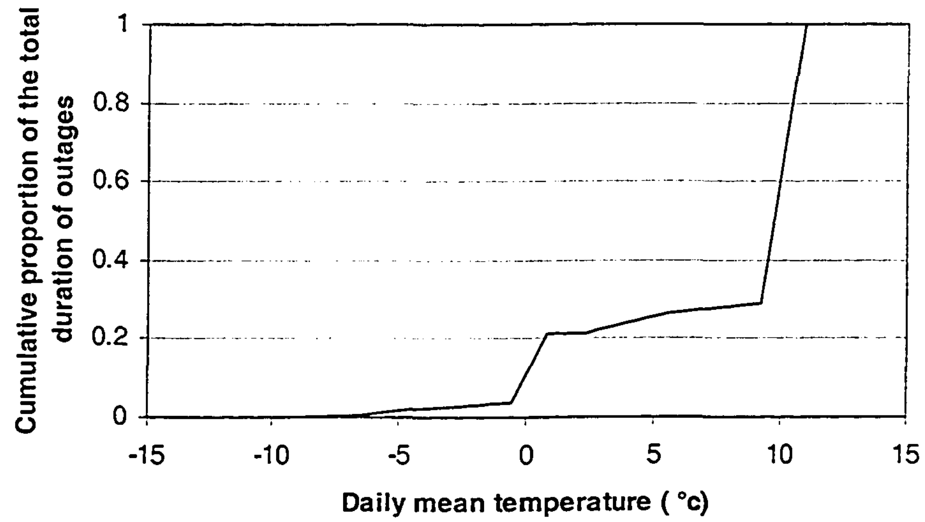


Figure 3-15. The daily mean air temperature vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outage on 240kV transmission lines in Alberta from 1977 to 1995

The maximum daily precipitation versus cumulative proportion of the total duration of precipitation-caused transmission line outages stratified by three voltage levels are presented in Figures 3-16 to 3-18.

For the 72kV transmission lines, the maximum daily precipitation with precipitation-caused outages varied from 0mm to 56mm. In terms of the outage duration, the most damage on these lines caused by precipitation was caused under weather condition with maximum daily precipitation between 11.2 and 11.4 mm. Twenty-two percent of the total duration of outages occurred with 0 mm maximum daily precipitation. This indicates that humid weather even without precipitation recorded (e.g. fog) can still cause serious problems for the 72kV transmission lines. In a small probability, precipitation-caused outages on 72kV lines occurred when the maximum daily precipitation was above 20mm.

For the 144kV transmission lines, the range of maximum daily precipitation with outage occurrence is from 0mm to 56mm. The weather with 0.0mm to 1.4mm precipitation is in high probability to cause outages with long durations. With 1mm daily maximum daily precipitation, the maximal duration of precipitation-caused outage was observed on 144kV transmission lines.

The range of maximum daily precipitation for precipitation-caused outages on the 240kV transmission lines is from 0 mm to 38 mm and this range is narrower than that for the other two voltage levels. In terms of outage duration, almost 88% of the outages occurred when there was 11.0 mm to 18.0 mm of maximum daily precipitation. There is a small probability that precipitation-caused outages would occur on 240kV lines with over 20 mm precipitation.

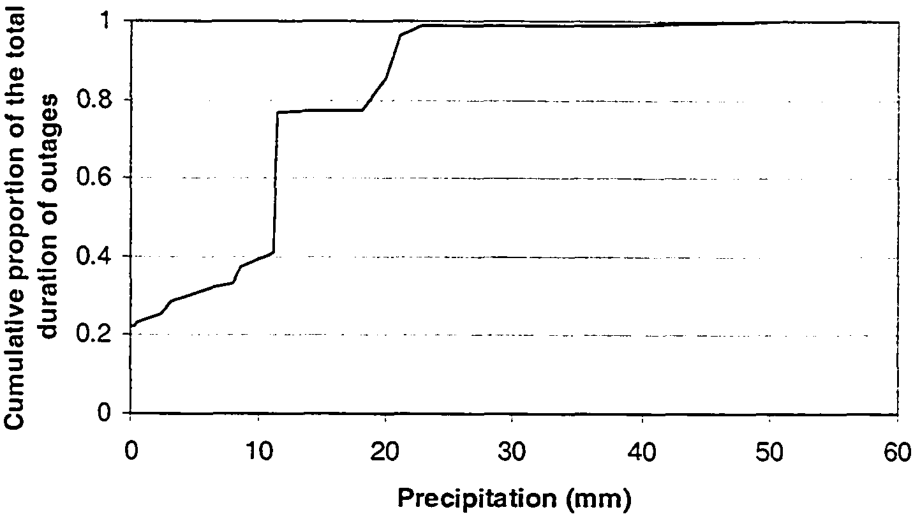


Figure 3-16. The maximum daily precipitation vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

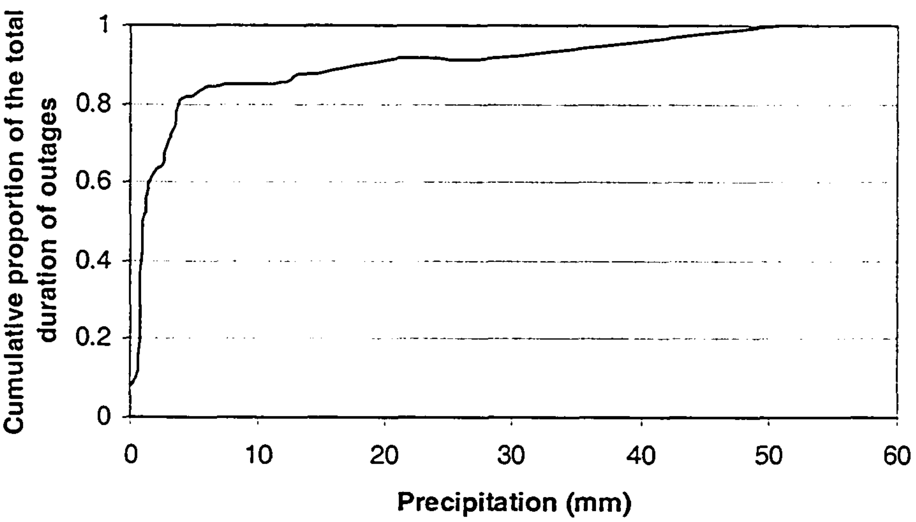


Figure 3-17. The maximum daily precipitation vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outages on 144kV transmission lines in Alberta from 1977 to 1995

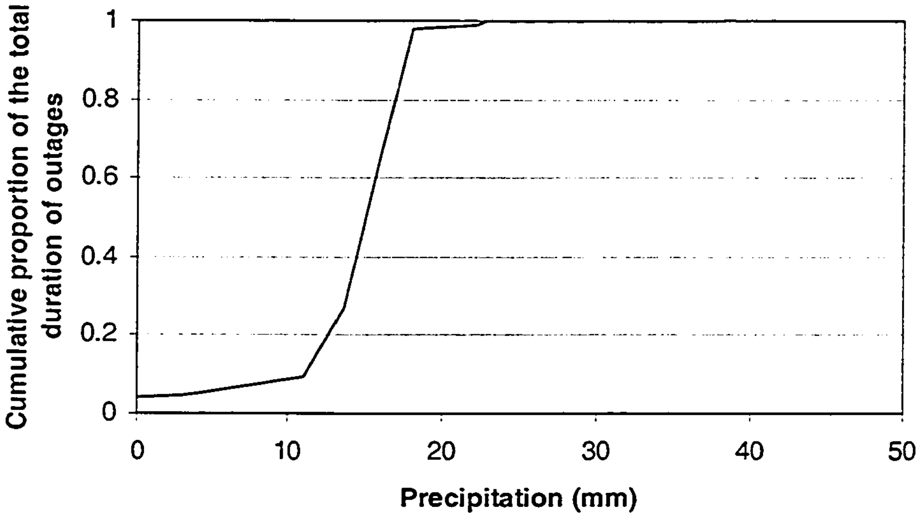


Figure 3-18. The maximum daily precipitation vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

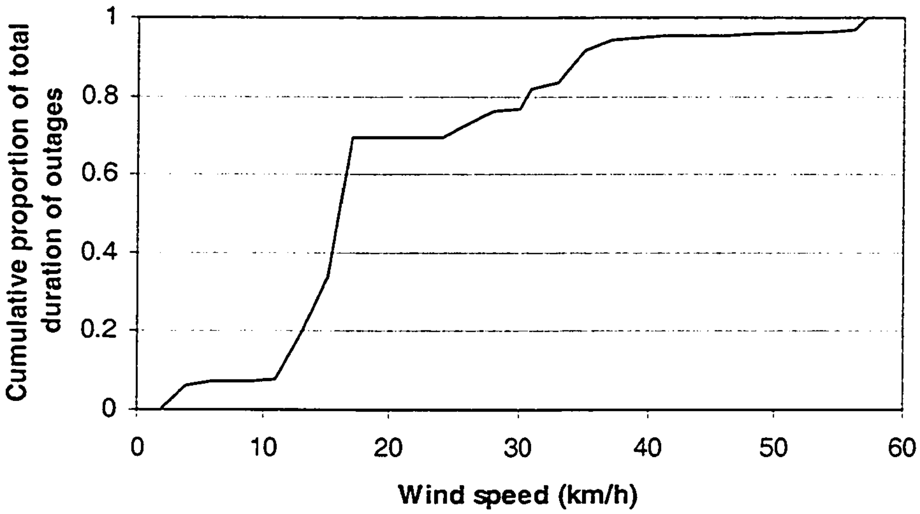


Figure 3-19. The maximum hourly wind speed vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

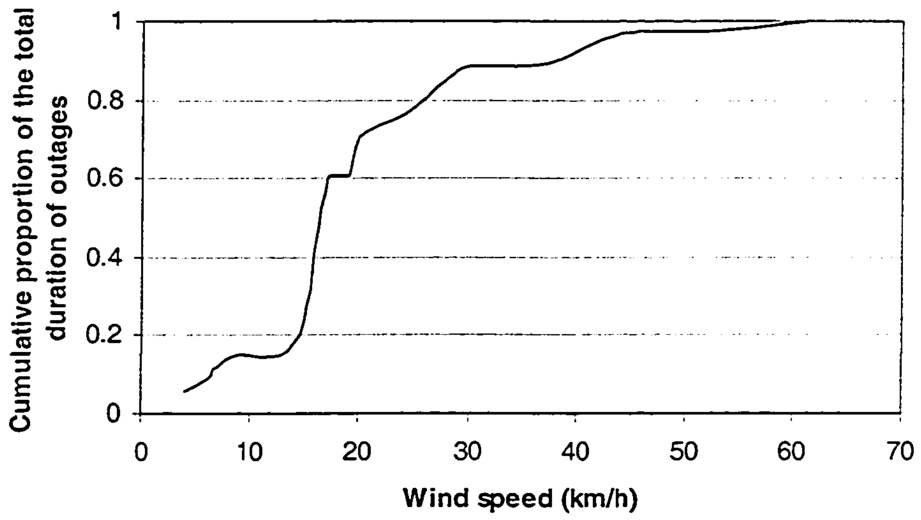


Figure 3-20. The maximum hourly wind speed vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outages on 144kV transmission lines in Alberta from 1977 to 1995

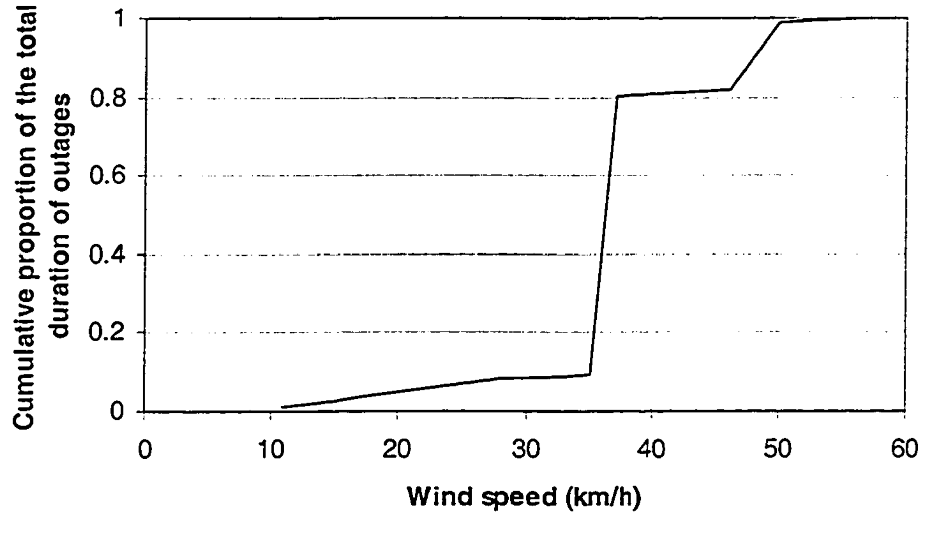


Figure 3-21. The maximum hourly wind speed vs. cumulative proportion of the total duration of precipitation-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

The hourly wind information (i.e. maximum hourly wind speed and wind direction) around areas with precipitation-caused outage was retrieved from a wind database. The results are shown in Figures 3-19 to 3-21. For precipitation-caused outages on 72kV transmission lines, the associated maximum hourly wind speed varied from 2km/h to 57km/h. In terms of the duration of the outages, outages occurred with a high probability when the maximum hourly wind speed was between 11km/h and 17km/h. The similar distribution of the maximal hourly wind speed is shown for the outage on 144kV transmission lines. For precipitation-caused outages on 240kV transmission lines, the range of the associated maximum hourly wind speed is narrower than that for precipitation-caused outages on the other two voltage lines. The long duration of precipitation-caused outages on 240kV transmission lines was associated with maximum hourly wind speeds of between 35 and 37km/h.

The correlations between duration of precipitation-caused line-related sustained power outage on three voltage (i.e. 72kV, 144kV and 240kV) transmission lines and the weather elements (i.e. daily mean air temperature, maximum daily precipitation and maximum hourly wind speed) are presented in table 3-5. The weather elements, which are highly correlated with the duration of the transmission line outages, are highlighted by a gray background.

Table 3-5. The correlations between the duration of precipitation-caused line-related sustained power transmission outage on three voltage (i.e. 72kV, 144kV and 240kV) transmission lines and the weather elements (i.e. daily mean air temperature, maximum daily precipitation and maximum hourly wind speed)

Voltage level of the outage lines	Correlation coefficients		
	Temperature	Precipitation	Wind
72kV	-0.005	0.07	-0.11
144kV	-0.05	-0.36	0.005
240kV	0.40	0.20	0.20

Some of the main results of analyzing precipitation-caused sustained power transmission line outages are summarized in Table 3-6.

Table 3-6. The summary of the main results for precipitation-caused sustained power transmission line outages

		Voltage levels		
		72kV	144kV	240kV
Probability distribution of duration of outage	Mean	10.30	6.25	47.55
	Median	2.42	1.57	6.46
	Skewness	4.88	4.24	3.67
Time when long outages occurred	Year	1992	1983	1986
	Month	Feb.	Nov.	May
	Hour	14~15	8~9	23~24
Weather condition when long outages occurred	Temp (°C)	-11.5~ -10.8	-18.5 -8.3~ -4.8	9.3~11.0
	Pcpn (mm)	11.2~11.4	0.0~1.4	11.0~18.0
	WS (km/h)	11~17	13~17	35~37

Temp: the daily mean air temperature
Pcpn: the maximum daily precipitation
WS: the maximum hourly wind speed

3.5 Wind-Caused Line-Related Sustained Power Outages

The estimated probability density functions of the outage duration of wind-caused outages for three voltage transmission lines are plotted in Figures 3-22 to 3-24. The mean and median of the duration of wind-caused outage are different for the three voltage transmission lines and the probability distribution of the duration of wind-caused transmission line outages are skewed to the right. The number of data for estimating the probability density functions of the duration of wind-caused line-related sustained transmission line outages for the 72kV, 144kV and 240kV transmission lines are 68, 29 and 15 respectively. The probability distribution of the outage duration on 144kV transmission lines has the largest skewness among three voltage levels. The maximal duration, 191.6 hours, of wind-caused outages was observed on 240kV transmission line. The average duration of wind-caused outages of 72kV transmission lines are shortest among the three voltage levels and wind-caused outages on 240kV transmission lines tend to have longer outage durations.

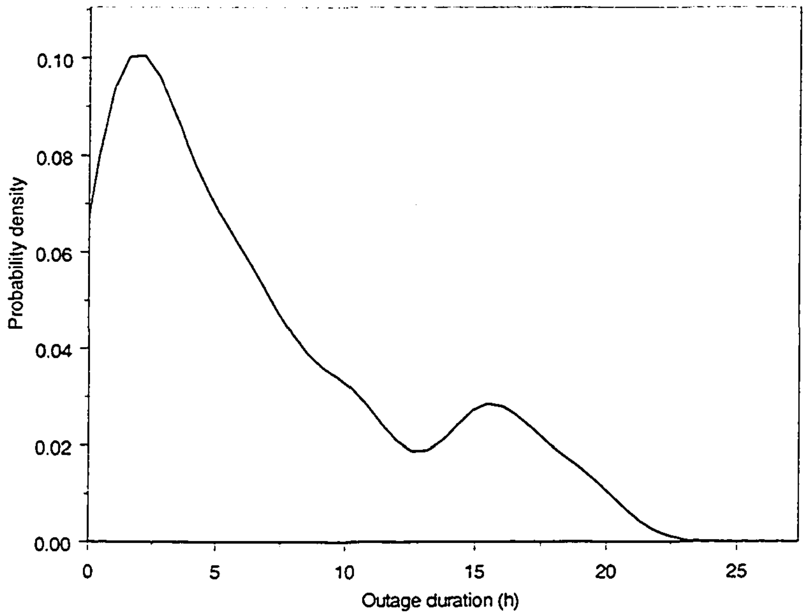


Figure 3-22. The probability density function of outage duration of wind-caused line-related sustained power transmission line outages on 72kV voltage lines estimated by 1977 – 1995 ATCO power outages data

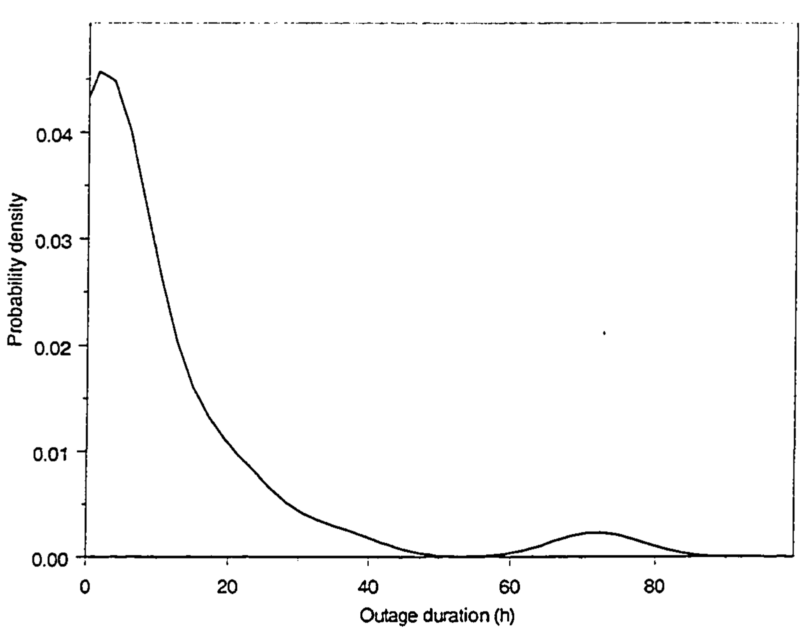


Figure 3-23. The probability density function of outage duration of wind-caused line-related sustained power transmission line outages on 144kV voltage lines estimated by 1977 – 1995 ATCO power outages data

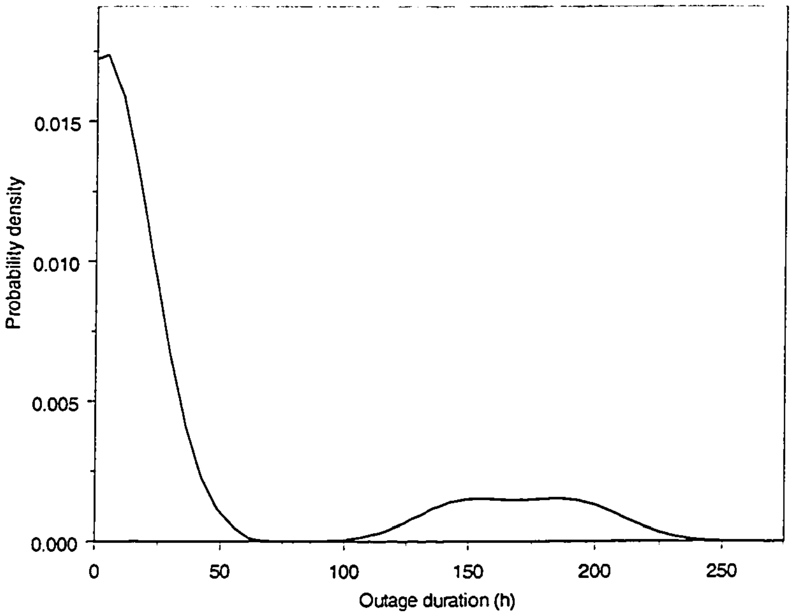


Figure 3-24. The probability density function of outage duration of wind-caused line-related sustained power transmission line outages on 240kV voltage lines estimated by 1977 – 1995 ATCO power outages

The annual time series of the annual total duration of wind-caused outages classified by three voltage levels are shown in Figure 3-25.

For 72kV transmission lines, wind-caused outages were observed in every year in the 19-year period except 1977,1978, 1980,1983 and 1995. It is also found that wind-caused outages with long outage duration on 72kV transmission lines occurred in 1984,1985 and 1986 and the annual total duration of the outages for these three years are 47.4 hours, 52.9 hours and 68.6 hours respectively. After 1984, the annual total duration of wind-caused outages on 72kV voltage lines for each year is over 10 hours, except for 1995.

For the 144kV transmission lines, wind-caused outages started in 1982 and ended in 1993. The maximal duration, 72 hours, of wind-caused outages at 72kV level transmission lines was found in 1993. Two clusters of years with the duration of wind-caused outages over 10 hours were observed in the period 1984 to 1986 and 1991 to 1993. The duration of wind-caused outages occurred in other years were under 10 hours, except for 1988.

For 240kV transmission lines, there are only 8 out of 19 years with occurrences of wind-caused outage. In 1987, the annual total duration of wind-caused outage on this voltage lines reach the maximum, 339.6 hours. The duration of wind-caused outage is under 13 hours in years when the outages occurred, except 1987.

The monthly time series of the monthly total duration of wind-caused line-related sustained power transmission outages classified by three voltage levels is presented in Figure 3-26. October is the month with the maximal monthly total duration, 90.2 hours, of wind-caused outages for the 72kV transmission lines. Other than this month, the monthly total duration of wind-caused outages in May, June and July are also long, and each of them is about 50 hours per month. Wind-caused outages for 72kV transmission lines almost occurred in every month except September.

It is June and July that are the months with monthly total duration of wind-caused outages over 50 hours per month for 144kV transmission lines. And especially in July, the monthly total duration of wind-caused outages on 144kV transmission lines reached 109.0 hours. The monthly total duration of wind-caused outages for 144 kV transmission lines in all months other than June and July are small and there were no wind-caused outages observed in November and December.

There were only six months in which wind-caused outages were observed for 240kV transmission lines. The month with extremely high monthly total duration, 340.0 hours, of wind-caused outages was found to be September. The monthly total duration of wind-caused outages in the other months of the year are smaller than 10 hours per month, except in March, which had 19.0 hours.

Figure 3-27 shows the hourly time series of the hourly total duration of wind-caused line-related sustained power transmission outages classified by three voltage levels. For 72kV transmission lines, the highest hourly total duration of wind-caused outages was observed from 8:00 PM to 9:00 PM. In terms of outage duration, there were more wind-caused outages occurring in the afternoon than that in the morning for 72kV transmission lines. Periods without wind-caused outages for 72kV transmission lines were between midnight to 2:00 AM, 3:00 AM to 4:00 AM and 8:00 AM to 9:00 AM.

For the 144kV transmission lines, there are 8 one-hour periods without any wind-caused outages occurrence. Between 4:00 AM to 5:00 AM and 9:00 PM to 11:00 PM were periods with high hourly total duration of wind-caused outages occurrence. And especially in the period 9:00 PM to 10:00 PM, the hourly total duration of wind-caused outages reach the maximum, 72.0 hours.

For the 240kV transmission lines, wind-caused outages were rarely occurring events. There are only 10 one-hour periods out of 24 with wind-caused outages. Except for the period of 4:00 AM to 5:00 AM and period of 6:00 PM to 7:00 PM, the hourly total duration of wind-caused outages were below 7 hours per hour. The period with extremely high hourly total duration, 339.6 hours, of wind-caused outages was observed between 6:00 PM to 7:00 PM.

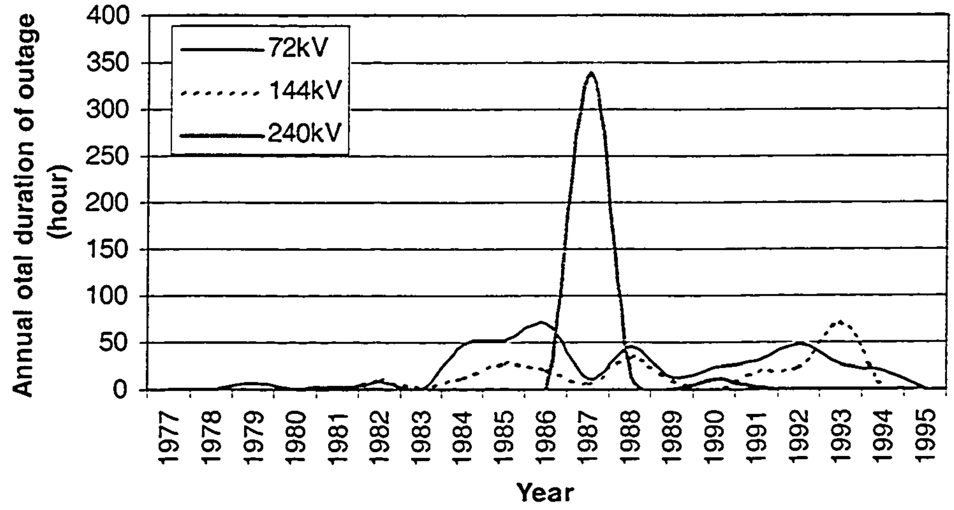


Figure 3-25. The annual time series of the annual total duration of wind-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

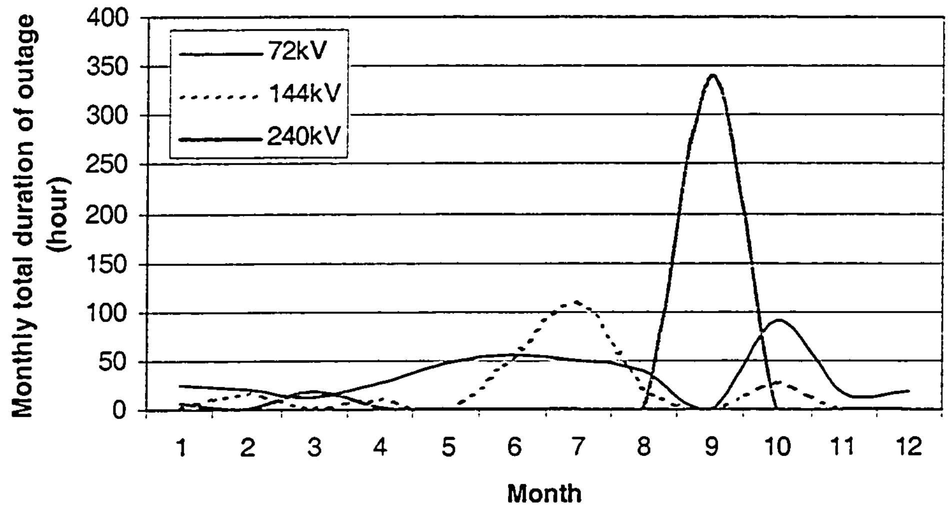


Figure 3-26. The monthly time series of the monthly total duration of wind-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

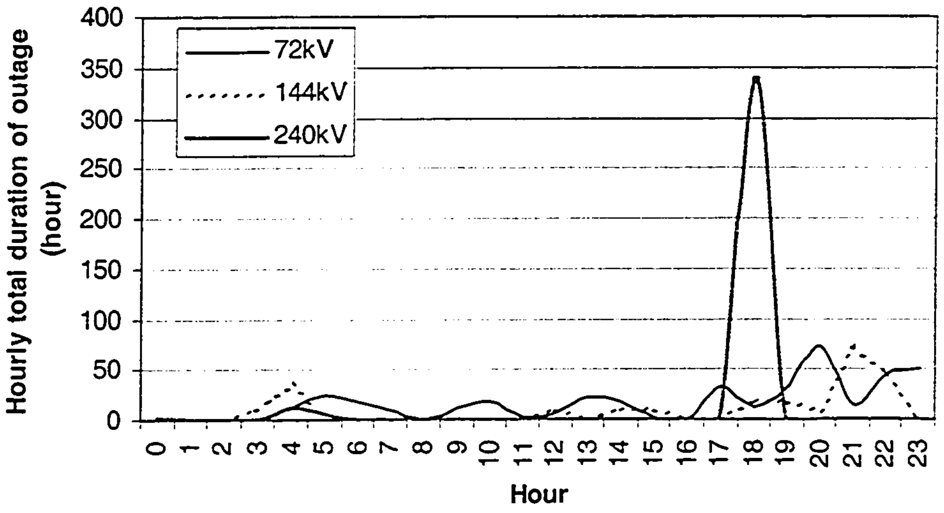


Figure 3-27. The hourly time series of the hourly total duration of wind-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

The daily mean air temperature versus cumulative proportion of the total duration of wind-caused line-related sustained power transmission line outages on three voltage transmission lines are plotted in Figures 3-28 to 3-30. For wind-caused outages on 72kV transmission lines, the mean air temperature varied from –19.4 °C to 23.8 °C. In terms of outage duration, wind-caused outages under weather conditions with a mean air temperature of between 6.3 °C and 20°C account for about 66.3% of the total outages on 72kV transmission lines. For wind-caused outages on 144kV transmission lines, the mean air temperature varied from –9.3 °C to 27.8 °C. In terms

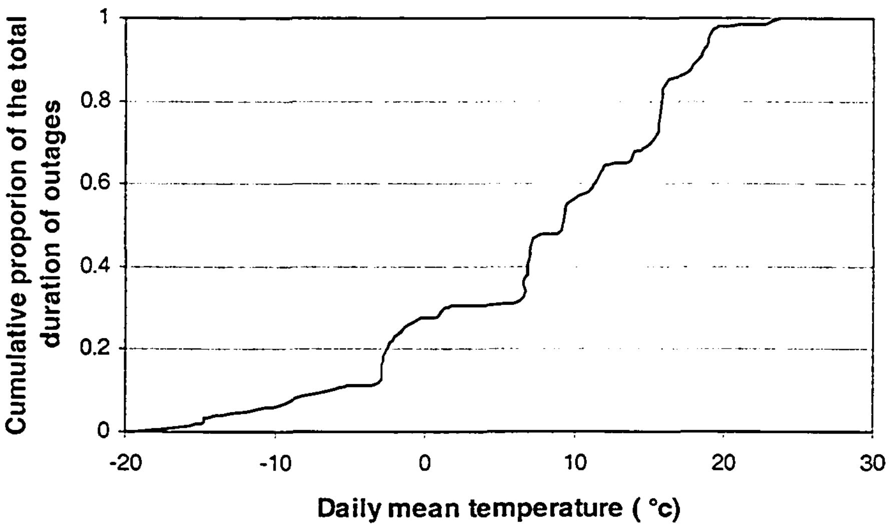


Figure 3-28. The daily mean air temperature vs. cumulative proportion of total duration of wind-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

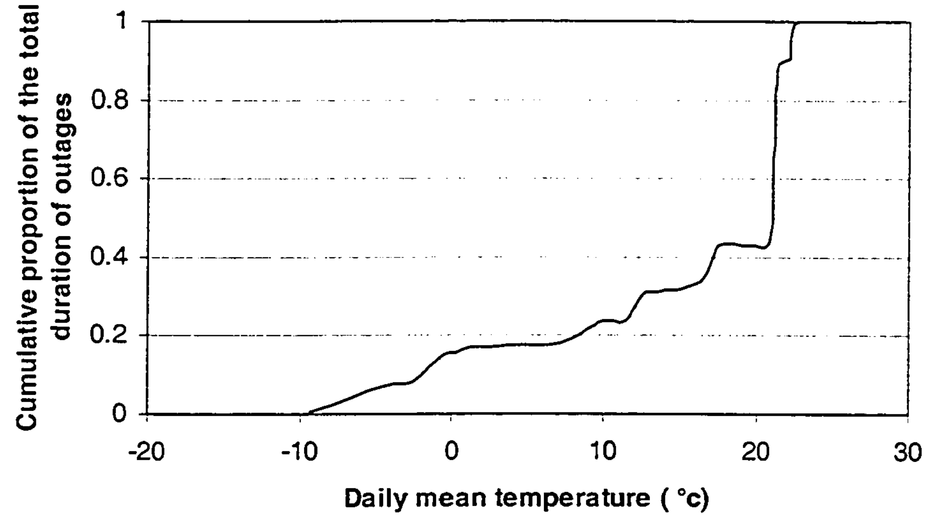


Figure 3-29. The daily mean air temperature vs. cumulative proportion of total duration of wind-caused line-related sustained power transmission line outages on 144kV transmission lines in Alberta from 1977 to 1995

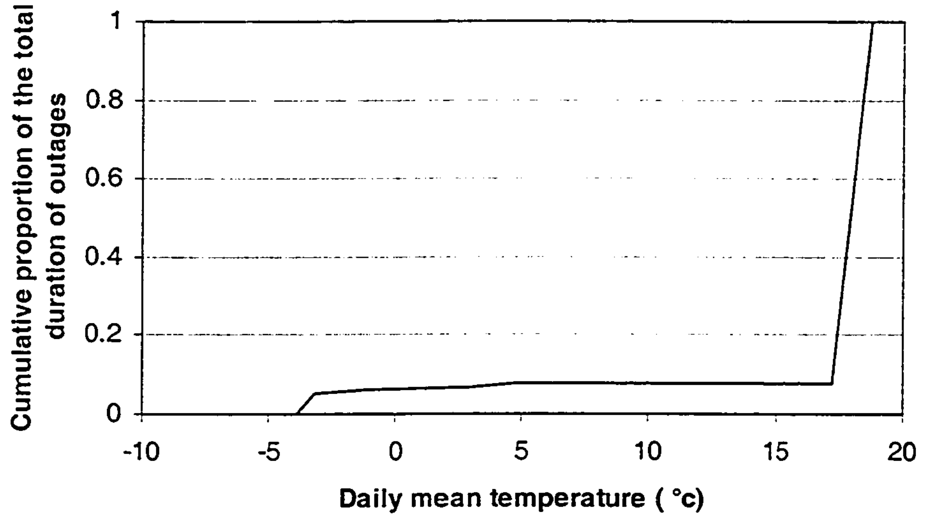


Figure 3-30. The daily mean air temperature vs. cumulative proportion of total duration of wind-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

of outage duration, wind-caused outages occurring under weather conditions with a mean air temperature of between 20.8 °C and 22.8°C account for about 56.6% of the total outages on the 144kV transmission lines. The range of the mean air temperature with wind-caused outage occurrence on the 240kV transmission lines are much narrower that that on the other two voltage lines. The extreme high duration of wind-caused outages were observed when the mean air temperature is between 17.3°C and 18.8°C.

The maximum daily precipitation versus cumulative proportion of the total duration of wind-caused line-related sustained power transmission line outages on three voltages transmission lines are shown in Figures 3-31 to 3-33. For the 72kV transmission lines, the duration of wind-caused outages without precipitation account for 10.5% of the total duration of wind-caused outages. In terms of outage duration, 21.9% of the outages occurred with maximum daily precipitation below 1mm. Outages occurring with maximum daily precipitation below 10mm accounted for 63.9% of outages. Wind-caused transmission line outages under weather conditions with extremely high maximum daily precipitation (i.e., 30mm) were rare and only accounted for 4.5% of the total duration of wind-caused outages on 72kV transmission lines.

For the 144kV transmission lines, 90.9% of outages occurred with maximum daily precipitation below 7mm. And 13.0% of outages were observed without any precipitation recorded. Wind-caused transmission line outages under weather conditions with extremely high maximum daily precipitation (i.e. 30mm) are rare and they only accounted for 4.4% of the total duration of wind-caused outages on 144kV transmission lines.

For the 240kV transmission lines, almost all wind-caused outages occurred under weather conditions with a maximum daily precipitation of between 8.1mm and 36mm. Only 3.1% of wind-caused outages occurred without any precipitation observed.

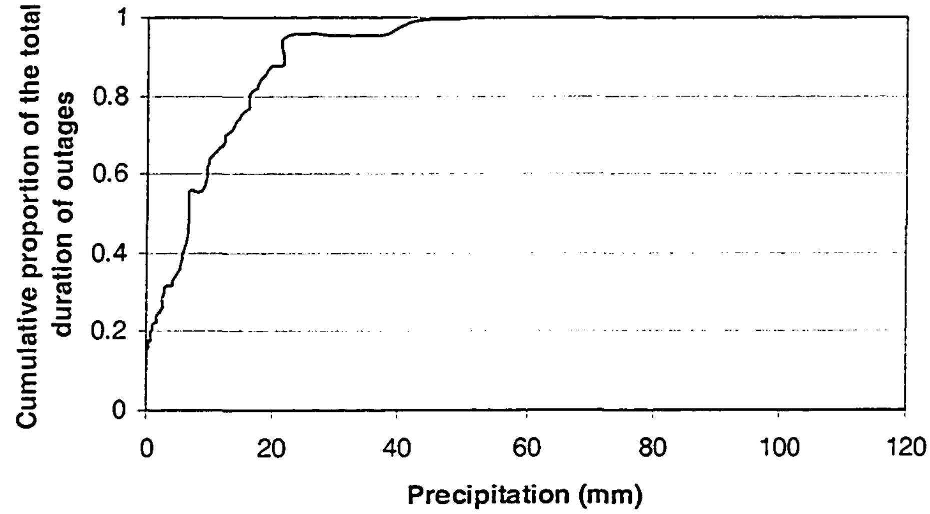


Figure 3-31. The maximum daily precipitation vs. cumulative proportion of the total duration of wind-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

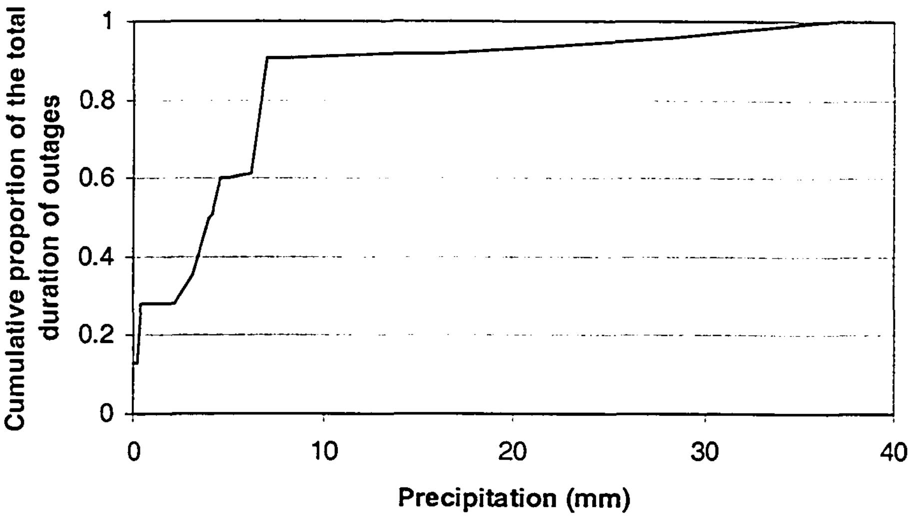


Figure 3-32. The maximum daily precipitation vs. cumulative proportion of the total duration of wind-caused line-related sustained power transmission lien outages on 144kV transmission lines in Alberta from 1977 to 1995

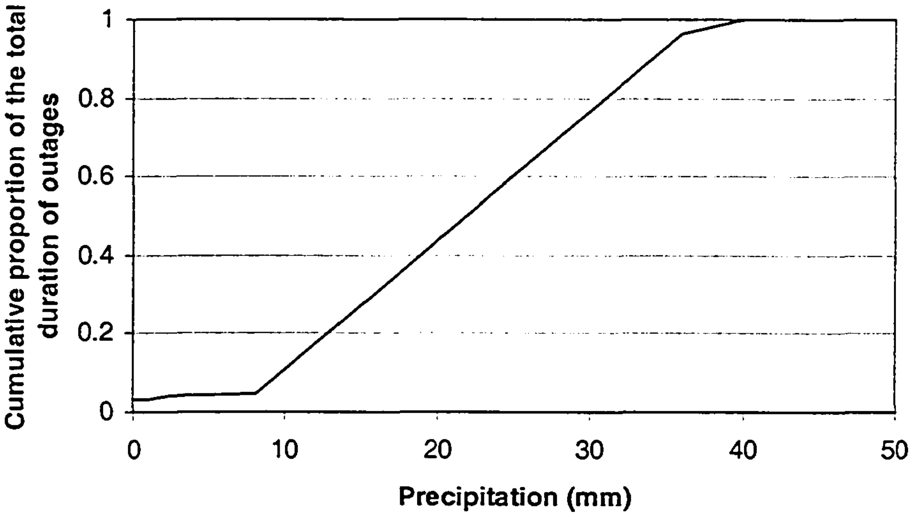


Figure 3-33. The maximum daily precipitation vs. cumulative proportion of the total duration of wind-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

The maximum hourly wind speed versus cumulative proportion of the total duration of wind-caused line-related sustained power transmission line outages on three voltages transmission lines are shown in Figures 3-33 to 3-35. For wind-caused outages on the 72kV transmission lines, the associated maximum hourly wind speed varied from 4km/h to 74km/h. The extremely high duration of the outages for 72kV transmission lines occurred when the maximum hourly wind speed is between 24km/h and 31km/h. For wind-caused outage on 144kV transmission lines, the maximum hourly speed of the wind between 9km/h to13km/h and the range of the maximum hourly wind speed of the wind which caused the outages is similar to that for 72kV transmission lines. For wind-caused outage on 240kV transmission lines, the maximal hourly speed of the wind is 46km/h to 50km/h.

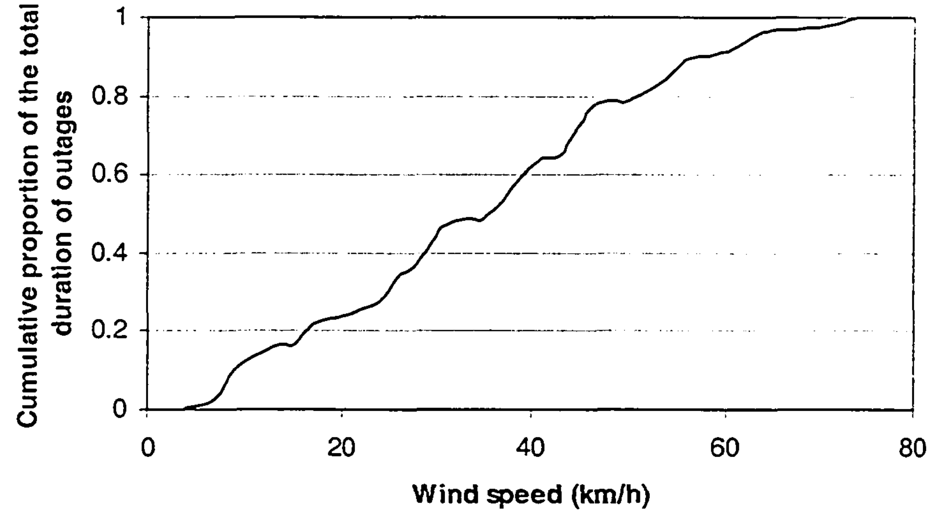


Figure 3-34. The maximum hourly wind speed vs. cumulative proportion of total duration of wind-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

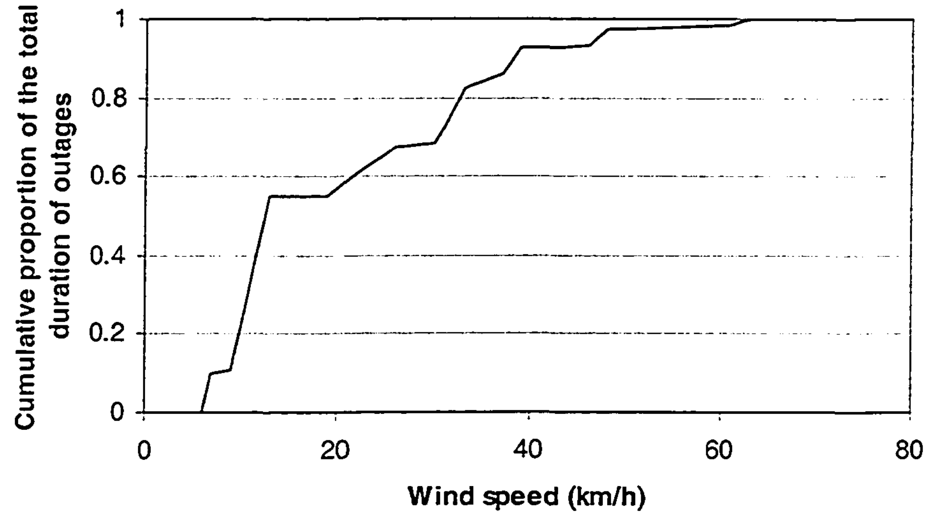


Figure 3-35. The maximum hourly wind speed vs. cumulative proportion of the total duration of wind-caused line-related sustained power transmission line outages on 144kV transmission lines in Alberta from 1977 to 1995

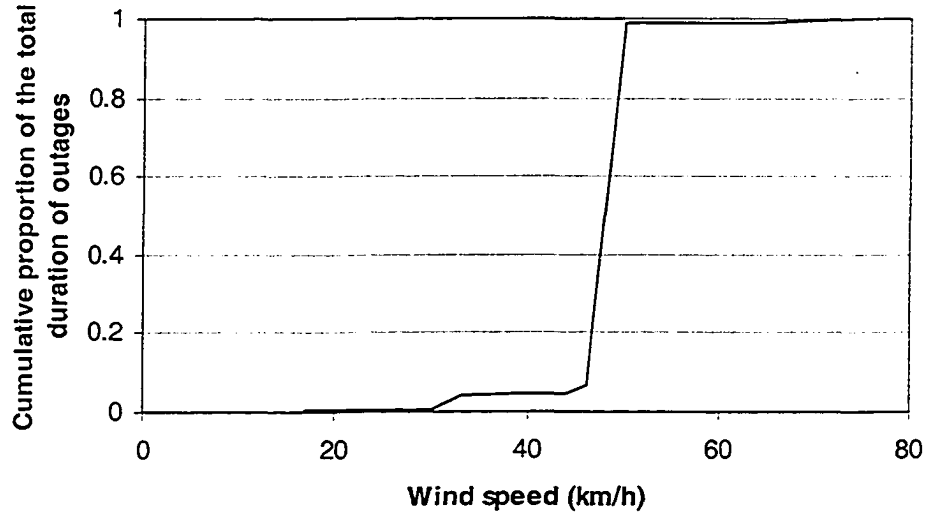


Figure 3-36. The maximum hourly wind speed vs. cumulative proportion of the total duration of wind-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

The correlations between duration of wind-caused line-related sustained power outages on three voltage (i.e. 72kV, 144kV and 240kV) transmission lines and the weather elements (i.e. daily mean air temperature, maximum daily precipitation and maximum hourly wind speed) are presented in table 3-7. The weather elements which are highly correlated with the duration of wind-caused transmission line outages are highlighted by a grey background.

Table 3-7. The correlations between the duration of wind-caused line-related sustained power transmission line outages on three voltage (i.e. 72kV, 144kV and 240kV) transmission lines and the weather elements (i.e. daily mean air temperature, maximum daily precipitation and maximum hourly wind speed)

Voltage level of the outage lines	Correlation coefficients		
	Temperature	Precipitation	Wind
72kV	0.13	0.08	-0.05
144kV	0.30	-0.08	-0.34
240kV	0.66	0.76	0.06

Table 3-8. The summary of the main results for wind-caused sustained power transmission line outages

		Voltage levels		
		72kV	144kV	240kV
Probability distribution of duration of outage	Mean	6.16	8.38	24.58
	Median	4.48	1.97	1.33
	Skewness	0.92	3.11	2.48
Time when long outages occurred	Year	1986	1993	1987
	Month	Oct.	July	Sept.
	Hour	20~21	21~22	18~19
Weather condition when long outages occurred	Temp (°C)	6.3~20.0	20.8~22.8	17.3~18.8
	Pcpn (mm)	0~8.0	0~0.6 2.2~5.0 6.2~7.0	8.1~36.0
	WS (km/h)	24~31	9~13	46~50

Temp: the daily mean air temperature
Pcpn: the maximum daily precipitation
WS: the maximum hourly wind speed

Some of the main results of analyzing wind-caused sustained power transmission line outages are summarized in Table 3-8.

3.6 Lightning-Caused Line-Related Sustained Power Outages

The estimated probability distribution density functions of the duration of lightning-caused line-related sustained power transmission line outages classified by their voltage levels (i.e. 72kV, 144kV and 240kV) in Alberta are plotted in Figures 3-37 to 3-39. The number of data for estimating the probability density functions of the duration of lightning-caused line-related sustained transmission line outages for the 72kV, 144kV and 240kV transmission lines are 58, 15 and 35 respectively.

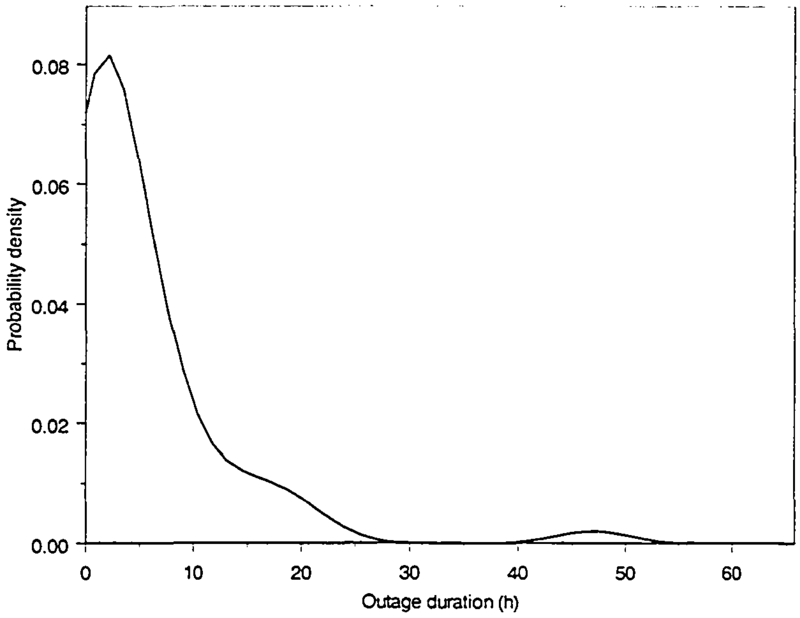


Figure 3-37. The probability density function of outage duration of lightning-caused line-related sustained power transmission line outages on 72kV transmission lines estimated by 1977 – 1995 ATCO power outages data

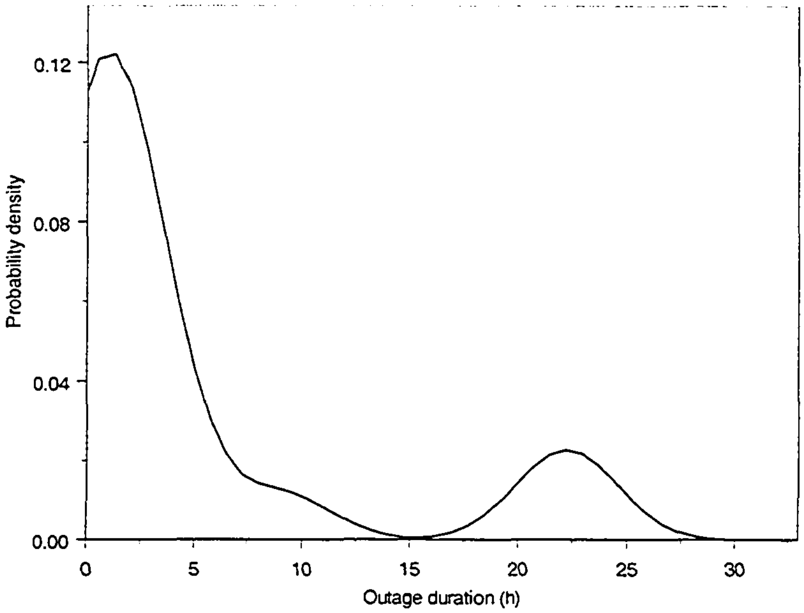


Figure 3-38. The probability density function of outage duration of lightning-caused line-related sustained power transmission line outages on 144kV transmission lines estimated by 1977 – 1995 ATCO power outages data

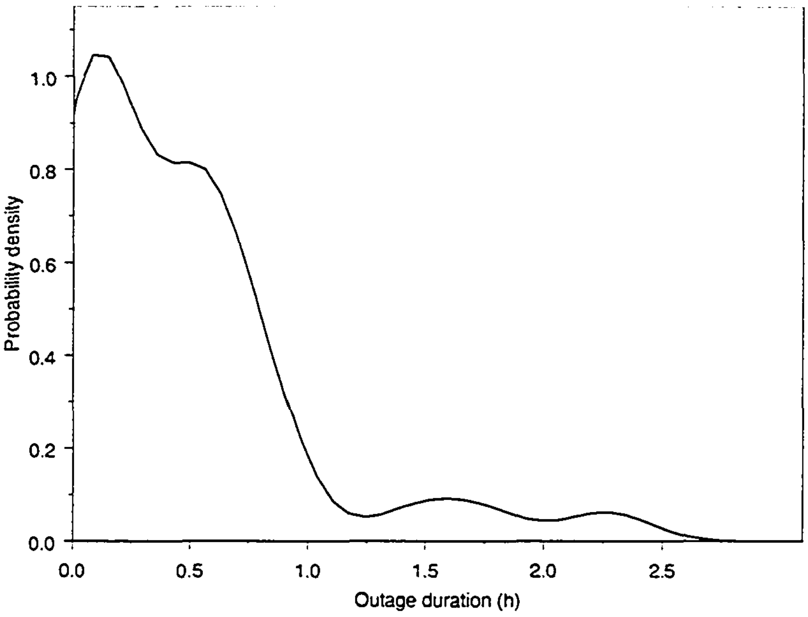


Figure 3-39. The probability density function of outage duration of lightning-caused line-related sustained power transmission line outages on 240kV transmission lines estimated by 1977 – 1995 ATCO power outages data

The shortest lightning-caused transmission line outage occurrence on the three voltage lines was 1 minute and the longest lightning-caused transmission line outage occurred on a 72kV transmission line. In terms of probability, the lightning-caused transmission line outages with high duration would highly likely occurred on 72kV voltage transmission lines and least likely occurred on 240kV transmission lines. The average duration of lightning-caused transmission line outages on 72kV and 144kV transmission lines are longer than that on 240kV transmission lines.

It is noted that the mean and median of the outages duration on each of the voltage level transmission lines studied are different and the underlying probability distributions are skewed to the right. The probability distribution of the outage duration for 72kV transmission lines has the largest skewness among three voltage levels. This result may challenge reliability evaluation models which assume that the probability distribution functions of the duration of the outage are in the symmetric probability distribution function family.

The annual time series of the annual total duration of lightning-caused transmission line outages for different voltage level transmission lines are plotted in Figure 3-40. It can be clearly found that lightning-caused transmission outages have temporal variability. For the 72kV transmission lines, the maximal annual total duration, 51.7 hours, of lightning-caused transmission line outages was observed in 1989. During the 19-year study period, lightning-caused transmission line outages occurred most of the time on this voltage level line. The average annual total duration of lightning-caused transmission line outages on 72kV transmission lines is much higher than that on the other two voltage levels. This result indicates that among the three voltage levels, lightning seriously affect 72kV transmission lines and can cause serious damage to 72kV transmission lines. There is no obvious annual trend found in this figure. In 1983-1984 and in 1988-1989, dual peak patterns were observed.

In 1991, the maximal annual total duration, 44.7 hours, of lightning-caused outages for 144kV transmission lines was observed. The outages for this voltage level transmission lines took place only in 7 out of the total 19 years. This pattern reveals that 144kV transmission lines are not sensitive to the impact of lightning.

Lightning-caused transmission line outages for 240kV transmission lines happened almost in every year during the 19-year period, although the average annual outage durations are small. It was found that among the three voltage levels that the 240kV and 72kV transmission lines are more sensitive to lightning than the 144kV transmission lines.

The monthly total lightning-caused transmission outages are shown in Figure 3-41. Almost all lightning-caused outages are found in the period of May to September, the primary lightning season in Alberta. In July, the outages caused by lightning reached the maximum for three voltage lines. For the 72 kV transmission lines, the distribution of the monthly total duration of outages show a mount-shape feature centered at July with a little skew to the right. Lightning-caused transmission line outages on 72kV transmission lines have the longest monthly total outage duration in every month from May to September compared to that of other voltage levels.

There were no lightning-caused outages observed in September for the 144kV transmission lines. A 4-minute lightning-caused outage on the 240kV transmission level lines happened in January. However, related lightning data cannot be found in the lightning data set for this lightning-caused outage.

The distribution of monthly total duration of lightning-caused outages on 240kV transmission lines shows that the maximal monthly total duration of lightning-caused transmission outages occurred in May, with another peak occurring in July. In June, August and September, the outage duration on the 240kV transmission lines was no more than 2 hours every month and was much smaller compared to that for the other two voltage levels. In summary, lightning-caused outages in Alberta were strongly monthly correlated.

Figure 3-42 shows the hourly time series of the hourly total duration of lightning-caused transmission line outages. More outages can be seen from noon to the midnight than that in the morning for all the voltage levels. The maximal total hourly outage, 47 hours, on 72kV transmission level occurred between noon to 1:00 PM. There were no lightning-caused outages that occurred on these transmission lines from 11:00 AM to noon.

The maximal total hourly duration of lightning-caused transmission line outages on the 144kV transmission lines was found between 1:00 AM to 2:00 AM. At this level, lightning-caused transmission line outages are clustered between 3:00 PM to 6:00 PM and midnight to 2:00 AM. Other than these periods, lightning-caused transmission line outages only occurred in the period between noon to 1:00 PM.

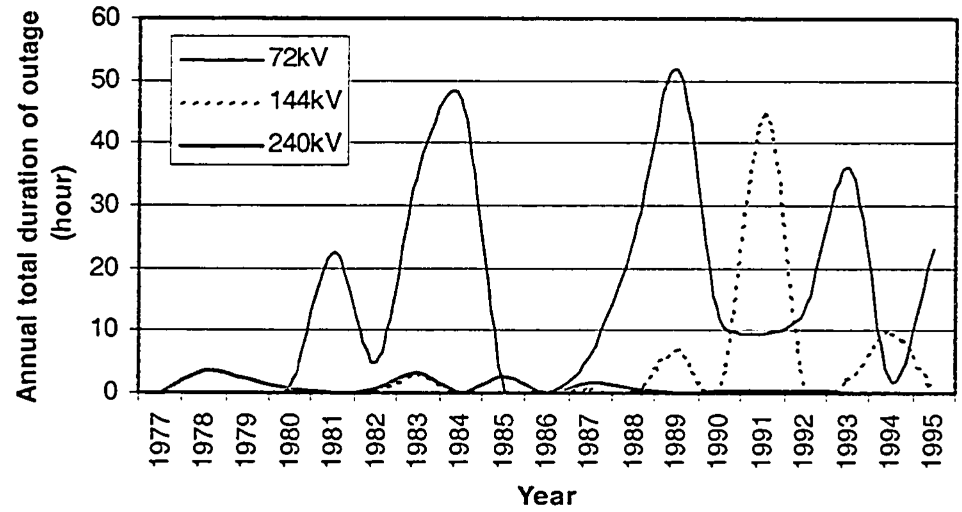


Figure 3-40. The annual time series of the annual total duration of lightning-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

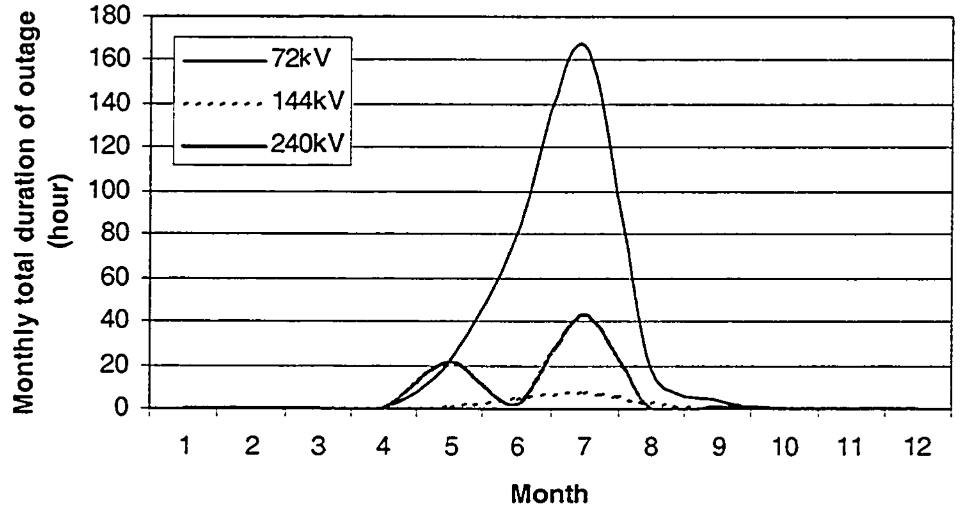


Figure 3-41. The monthly time series of the monthly total duration of lightning-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

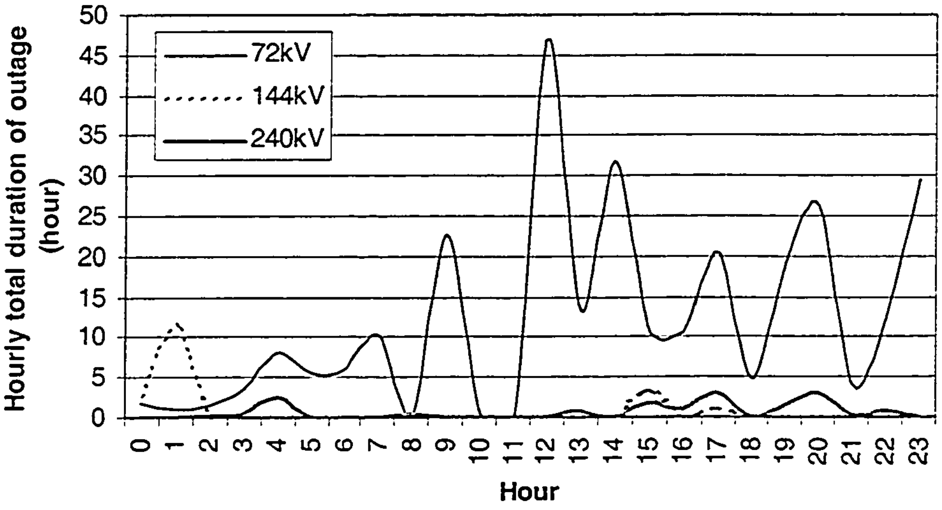


Figure 3-42. The hourly time series of the hourly total duration of lightning-caused line-related sustained power transmission outages classified by three voltage levels (i.e., 72kV, 144kV and 240kv) from 1977 to 1995 in Alberta

The maximal total hourly duration of the outage on the 240kV lines was found in the period between 8:00 PM to 9:00 PM.

The daily mean temperature versus cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outage on three voltage levels are presented in Figures 3-43 to 3-45. All lightning-caused transmission line outages were observed with daily mean air temperatures over 2.0 °C. And there were no lightning-caused transmission line outages that occurred with daily mean air temperatures over 25.0 °C.

For the 72kV transmission lines, the daily mean air temperature with lightning-caused transmission line outages occurrence varied from 8.9 °C to 24.0 °C. In terms of outage duration, the major lightning-caused transmission line outages occurred when the daily mean air temperature was in range of 10.3 °C to 11.5 °C. For the 144kV transmission lines, the range of the daily mean air temperature with lightning-caused transmission line outage occurrence was 11.0 °C to 24.3 °C. In terms of outage duration, the major lightning-caused transmission line outages occurred when the daily mean air temperature was in the range of 15.0 °C to 16.3°C. For the 240kV transmission lines, the range of the daily mean air temperature with lightning-caused transmission line outage occurrence was 2.0 °C to 20.8 °C. This range is wider than that for the other two transmission line voltage levels and both the upper bound and the lower bound of this range are smaller than that for the other two voltage levels. In terms of outage duration, most lightning-caused transmission line outages on 244kV transmission lines occurred when the daily mean air temperature was in the range of 14.3°C to 20.8°C.

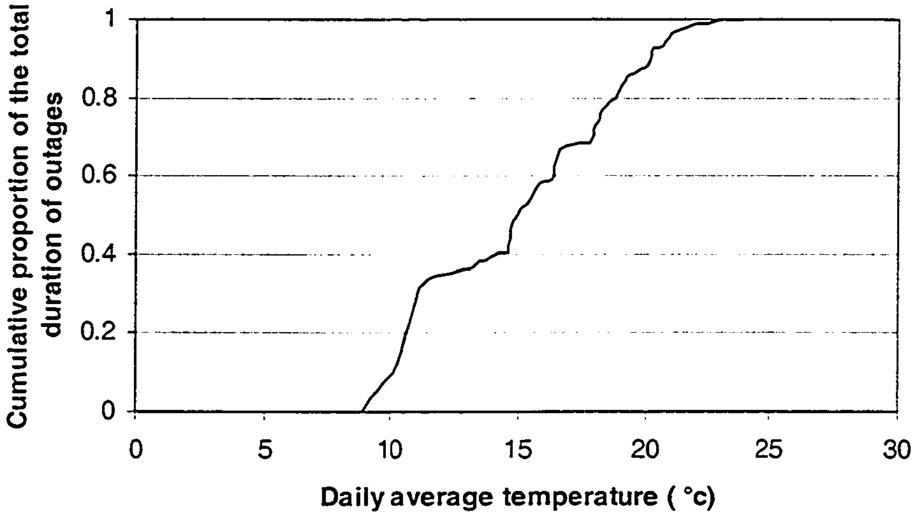


Figure 3-43. The daily mean air temperature vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

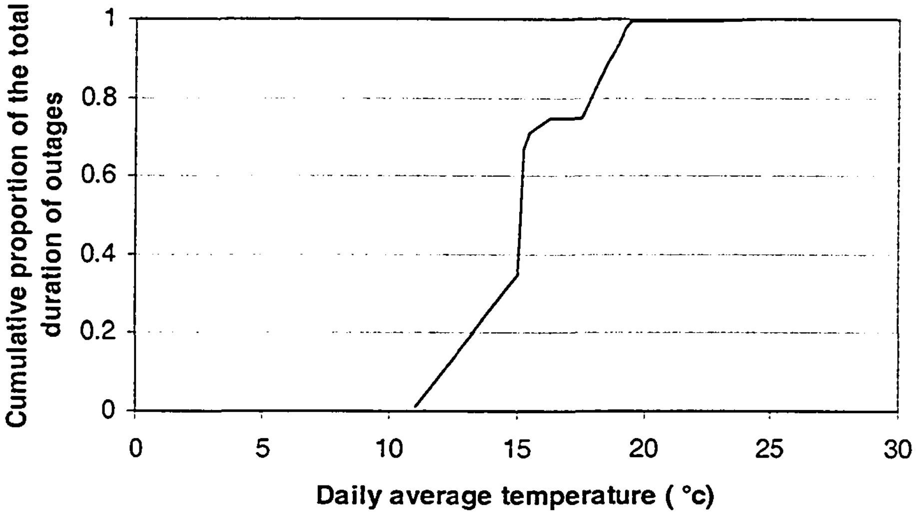


Figure 3-44. The daily mean air temperature vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 144kV transmission lines in Alberta from 1977 to 1995

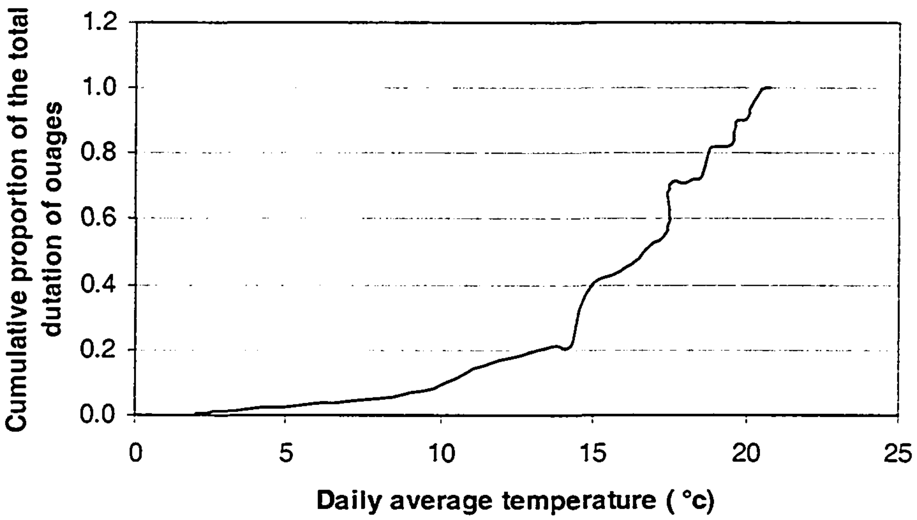


Figure 3-45. The daily mean temperature vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

Dry lightning is considered the primary cause of forest fires. It is commonly accepted by power engineers that dry lightning also plays an important role in lightning-caused transmission line outages. The rough definition of dry lightning is the lightning produced by a thunderstorm without rain. However, this definition is not precise and the thunderstorm without rain actually does have rain; it's only that the rain dries up before it reaches the ground. The precipitation data of outages lines located area were examined to reveal to what extent dry lightning was the cause of transmission line outages in Alberta.

In the study period, 8 out of 108 lightning-caused sustained transmission line-related outages have zero precipitation, with 5 of these 8 occurring on 72kV transmission lines, 2 of them occurring on 144kV transmission lines and 1 occurring on a 240kV transmission line. In terms of duration of the outages, 12.5 % lightning-caused transmission line outages were caused by dry lightning for the 72kV transmission lines; 33.7% lightning-caused transmission line outages were caused by dry lighting for 144kV transmission lines and 0.5% lightning-caused transmission line outages were caused by dry lightning for 240kV transmission lines.

The maximum daily precipitation versus cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on three voltage lines are plotted in Figures 3-46 to 3-48. For 72kV transmission lines, outages caused by lightning with 0mm to 2mm maximum daily precipitation and lightning with 16.5mm to 18.6mm maximum daily precipitation took the longest time to repair. In terms of outage durations, lightning-caused transmission line outages with maximum daily precipitation less than 1mm account for 19.2% of the total outage for

72kV transmission lines. In terms of probability, lightning-caused transmission line outages tend to occur with small amounts of maximum daily precipitation.

For the 144kV transmission lines, the duration of the transmission line outages caused by lightning with maximum daily precipitation less than 1mm account for 36.9% of the total outage duration .In terms of outage duration, lightning with maximum daily precipitation in range of 4mm to 9mm caused the most damage on the transmission lines. The probability of lightning-caused transmission line outage with over 25mm maximum daily precipitation is small.

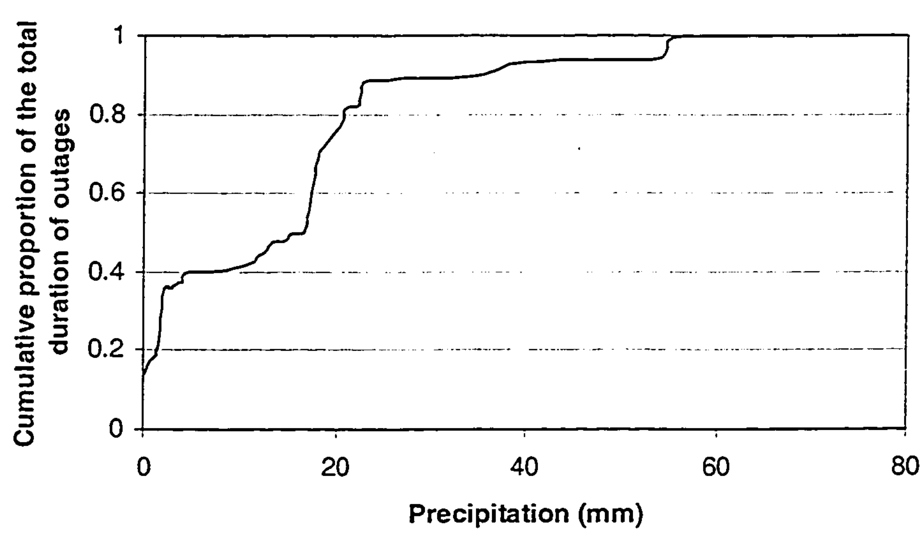


Figure 3-46. The maximum daily precipitation vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

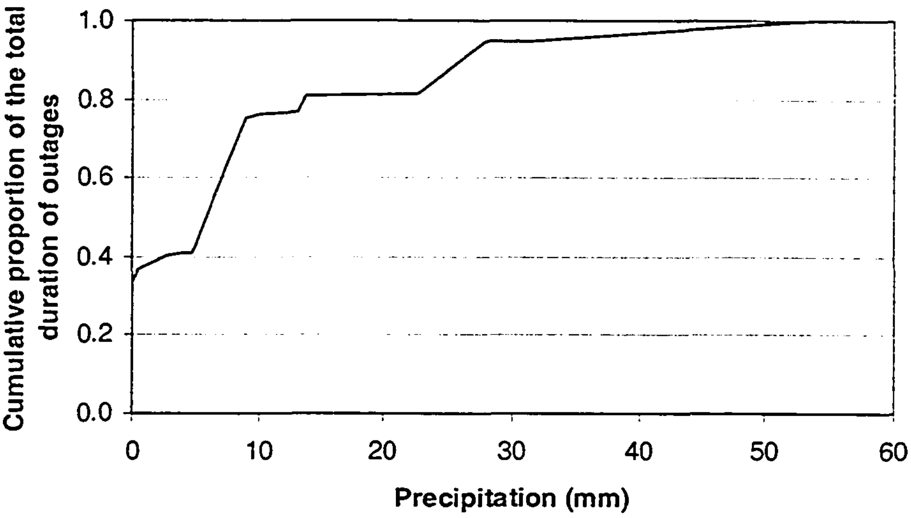


Figure 3-47. The maximum daily precipitation vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 144kV transmission lines in Alberta from 1977 to 1995

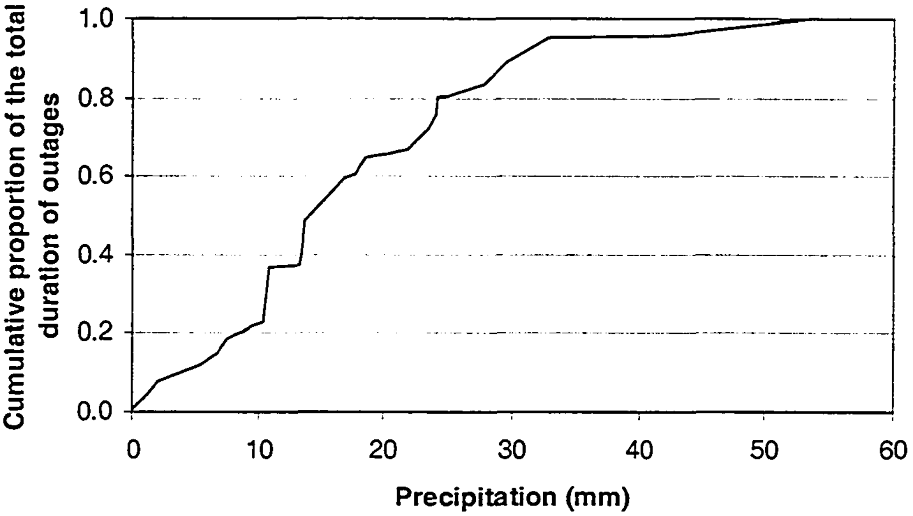


Figure 3-48. The maximum daily precipitation vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

For the 240kV transmission lines, dry lightning did not cause a serious problem. The duration of the transmission line outages caused by lightning with maximum daily precipitation less than 1mm only accounted for 0.9% of the total duration of the outages. Lightning weather with 10.4mm to 13.6mm of maximum daily precipitation caused the most damage to this voltage level transmission lines.

The maximum hourly wind speed versus cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on three voltage transmission lines are shown in Figures 3-49 to 3-51. For lightning-caused transmission line outages on the 72kV transmission lines, the associated maximum hourly wind speed varied from 4km/h to 52km/h. The maximum hourly

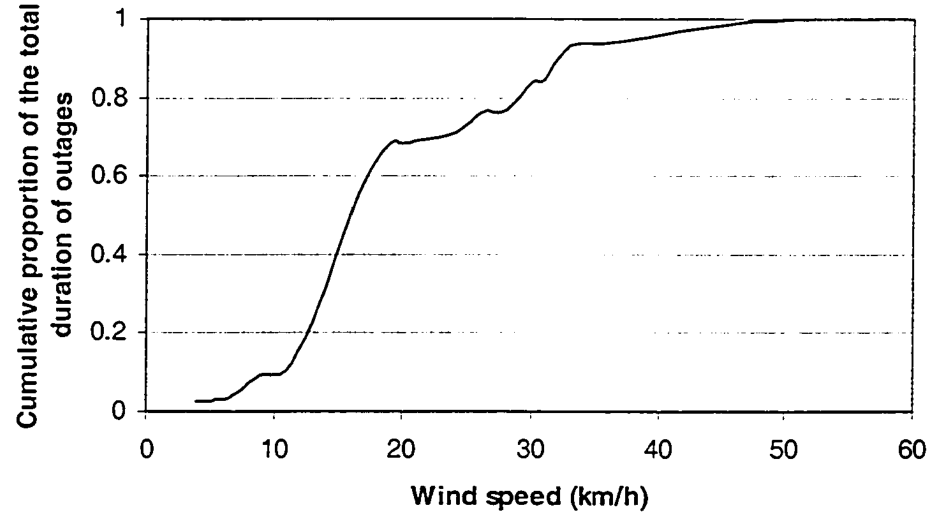


Figure 3-49. The maximum hourly wind speed vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 72kV transmission lines in Alberta from 1977 to 1995

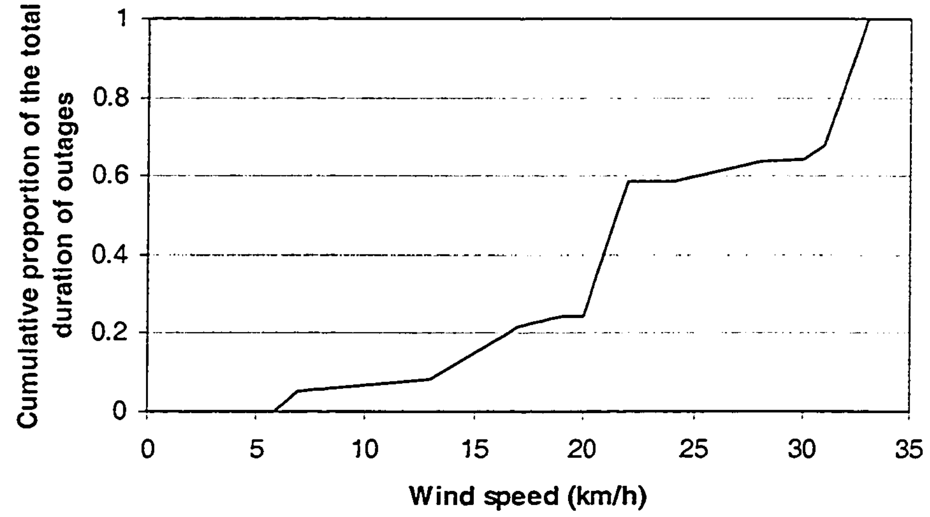


Figure 3-50. The maximum hourly wind speed vs. cumulative proportion of the total duration lightning-caused line-related sustained power transmission line outages on 144kV transmission lines in Alberta from 1977 to 1995

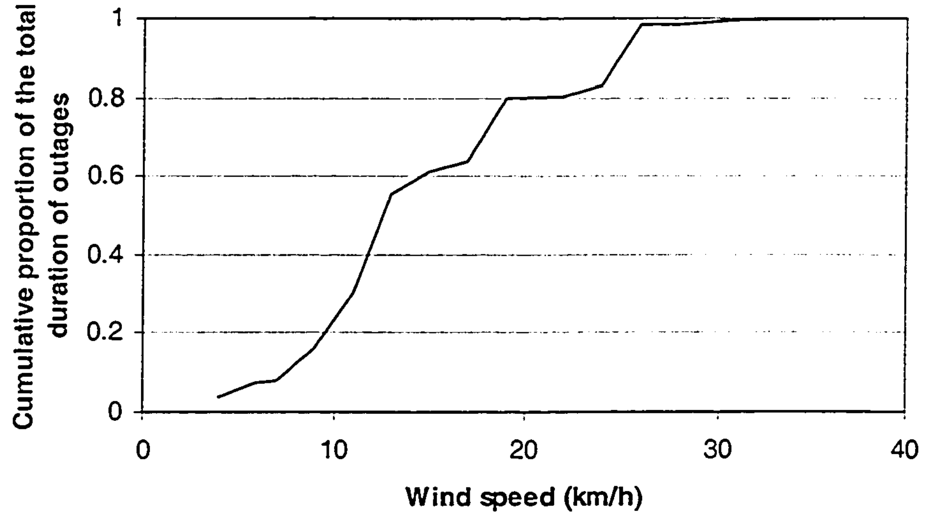


Figure 3-51. The maximum hourly wind speed vs. cumulative proportion of the total duration of lightning-caused line-related sustained power transmission line outages on 240kV transmission lines in Alberta from 1977 to 1995

speed of the wind which caused extreme high duration of the outage was in the range of 11km/h to 19km/h .For lightning-caused transmission line outages on 144kV transmission lines, the range of the maximum hourly wind speed was from 6km/h to 33km/h and this range is smaller than that for 72kV transmission lines. The maximum hourly speeds of the wind which caused extreme high duration of the outage were in the range of 20km/h to 22km/h and 31km/h to 33km/h. For lightning-caused transmission line outages on 240kV transmission lines, the range of the maximal hourly wind speed was from 4km/h to 37km/h. The maximum hourly speeds of the wind which caused extreme high duration of the outage were in the range of 7km/h to 13km/h.

The correlations between duration of lightning-caused line-related sustained power transmission line outage on three voltage (i.e. 72kV, 144kV and 240kV) levels transmission lines and the weather elements (i.e. daily mean air temperature, maximum daily precipitation and maximum hourly wind speed) are present in table 3-9. The weather elements, which are highly correlated with the durations of lightning-caused transmission line outages, are highlighted by a grey background in the table.

Table 3-9. The correlations between the duration of lightning-caused line-related sustained power outage duration on three voltage (i.e. 72kV, 144kV and 240kV) transmission lines and the weather elements (i.e. daily mean air temperature, maximum daily precipitation and maximum hourly wind speed)

Voltage level of the outage lines	Correlation coefficients		
	Temperature	Precipitation	Wind
72kV	-0.31	0.05	0.07
144kV	-0.35	-0.11	0.24
240kV	0.06	-0.05	-0.15

Some of the main results of analyzing lightning-caused sustained power transmission line outages are summarized in Table 3-10.

Table 3-10. The summary of the main results for lightning-caused sustained power transmission line outages

		Voltage levels		
		72kV	144kV	240kV
Probability distribution of duration of outage	Mean	5.02	4.49	0.46
	Median	1.85	1.25	0.40
	Skewness	3.41	2.03	1.93
Time when long outages occurred	Year	1989	1991	1978
	Month	July	July	July
	Hour	11~12	1~2	20~21
Weather condition when long outages occurred	Temp (°C)	10.3~11.5	15.0~16.3	14.3~20.8
	Pcpn (mm)	0~2 16.5~18.6	4~9	10.4~13.6
	WS (km/h)	11~19	20~22 31~33	7~13

Temp: the daily mean air temperature
Pcpn: the maximum daily precipitation
WS: the maximum hourly wind speed

3.4 Summary

Adverse weather has been shown to be the major cause for power transmission system outages in Alberta. For line-related power transmission line outages, most were sustained outages, and the transient outages, caused by lightning, were only found on 240kV transmission lines. Most of the line-related power transmission outages were recorded in southeastern Alberta, where more than half of the power transmission lines were located.

The probability density functions of the duration of adverse-weather-caused line-related sustained power transmission line outages are skewed to the right. This means the mean of the distribution is larger than the median. The median is an appropriate measurement to describe the center of the distribution.

Precipitation-caused transmission line outages with long duration were more likely to occur on the 240kV transmission lines and less likely to occur on the 144kV transmission lines. For 72kV and 144kV transmission lines, most outages were observed when the daily mean air temperature is below 0°C. In contrast, for 240kV transmission line, outages with long duration tended to occur when the daily mean air temperature is above 0°C. Precipitation-caused line-related power transmission line outages were rarely found when the daily air temperature is below –20°C. Humid weather with small amounts of daily precipitation or without daily precipitation caused serious damage on the 72kV and 144kV transmission lines. High maximum daily precipitation (e.g., 30mm) did not cause much of the precipitation-caused line-related power transmission line outages. Winds with speeds in the range of 11km/h to 17km/h were found when the most of the precipitation-caused line-related power outages were recorded for 72kV and 144kV transmission lines. For the 240kV transmission lines, they were in the range of 35km/h to 37km/h.

Wind-caused transmission line outages with long duration were more likely to occur on the 240kV transmission lines and were less likely to occur on the 72kV transmission lines. Most of the wind-caused transmission line outages occurred when the daily mean air temperature was above 0°C for three voltage transmission lines. For 144kV transmission lines, wind-caused outages tended to occur when the maximum hourly wind speed was below 20km/h.

Lightning-caused transmission line outages with long duration were more likely to occur on the 72kV transmission lines and were less likely to occur on the 144kV transmission lines. Lightning-caused transmission line outages are strongly month correlated. Dry lightning can seriously damage 72kV and 144kV transmission lines. But it had little effect on the 240kV transmission lines.

The results for this chapter are based on the unique geographic structure of the power transmission system in Alberta. They may not be applicable to other areas.

Chapter 4

Regression Models

4.1 Introduction

The purpose of regression is to reveal the relationship between response variables and predicting variables. The linear model is the most widely used model in the practice and the theories for establishing the linear models have been well developed. Even though the linear models are simple and were largely developed in the precomputer age of statistics, there are still good reasons to study and use them. They often provide an adequate and interpretable description of how the inputs affect the output. For prediction purposes they can sometimes outperform fancier nonlinear models, especially in situations with small numbers of training cases, low signal-to-noise ratio or sparse data. Furthermore, linear methods can be applied to transformations of the inputs and this, considerably, expands their scopes (Hastie et al., 2001).

In system reliability assessment, biostatistics, economics, sociology, and many other fields it is seldom the case that subject matter knowledge exists that would allow the analyst to prespecify a model, a transformation for the response variable, and a structure for how predictors appear in the model (Harrell, 2001). Some people even question whether the notion of a model exists in many cases. However, we are forced to develop a model for the purpose of finding the major elements which contribute the most to producing the response or to predict the value of a response variable based on a number of predicting variables. Fortunately, by various regressions approach an accurate, interpretable model can be set up.

In the practice, since regression models are often developed using a "convenience sample", that is, a dataset that was not collected with such predictions in mind. The resulting models are often fraught with difficulties such as the following:

1. The most important predictor or response variables may not have been collected and this may mislead people in their understanding of the real underlying processes.
2. Key variables are missing in large numbers of subjects.
3. Data are not missing at random.
4. Operational definitions of some of the key variables were never made.
5. Observer variability studies may not have been performed, so that the reliability of measurements is unknown.

Some of the above problems can be avoided by well-planed data collection procedures. One can design data collection instruments containing the necessary variables, and all terms can be given standard definitions for all data collection sits. Also, steps can be taken to minimize the amount of mission data. However, few data collection procedure can overcome all the difficulties. This makes a need to add other techniques (i.e., techniques for dealing with missing data) to set up a good regression model.

4.2 Model Variables Selection Methods

4.2.1 Ordinary Least Squares Method

Consider the use of usual linear regression model which has the form

$$f(X) = \beta_0 + \sum_{j=1}^{p} X_j \beta_j \tag{4.1}$$

to predict a real-valued output Y, where $X = (X_1, X_2, ..., X_p)$ is a vector of inputs or observations, and p is the number of predictors. The basic assumptions for the linear model are that either the conditional mean $E(Y \mid X)$ is linear or the linear model is a reasonable approximation of the relationship between Y and X.

Our aim is to estimate the parameter β for the purpose of prediction or determine the effect elements of X on Y. The estimators can be obtained by Ordinary Least Squares (OLS) method which is the most popular estimation method. The idea of the OLS method is to find the β which can minimize the residual sum of squares defined as follow

$$\begin{aligned} RSS(\beta) &= \sum_{i=1}^{N} (y_i - f(x_i))^2 \\ &= \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 . \end{aligned} \tag{4.2}$$

If we denote X as the $N \times (p+1)$ design matrix with each row an input vector (with a 1 in the first position), and let y be the N-vector of outputs. The residual sum of squares can be written as

$$RSS(\beta) = (y - X\beta)^T (y - X\beta) . \tag{4.3}$$

Differentiating with respect to β we obtain

$$\frac{\partial RSS}{\partial \beta} = -2X^T(y - X\beta)$$
$$\frac{\partial^2 RSS}{\partial \beta \partial \beta^T} = -2X^T X. \tag{4.4}$$

Assuming that X is nonsingular and hence $X^T X$ is positive definite, we set the first derivative to zero

$$X^T(y - X\beta) = 0 \tag{4.5}$$

to get the *unique solution*

$$\hat{\beta} = (X^T X)^{-1} X^T y, \tag{4.6}$$

and the predicted values of y is

$$\hat{y} = X\hat{\beta} = X(X^T X)^{-1} X^T y. \tag{4.7}$$

One problem in practice is that the columns of X might be not linearly independent. This means that X is not of full rank. Then the least squares coefficients $\hat{\beta}$ are not uniquely defined. The natural way to deal with this problem is recoding, or dropping redundant columns in X, or both.

In order to analyze the properties of $\hat{\beta}$ and draw inferences about $\hat{\beta}$, some assumptions have to be made. Suppose the observation y_i is uncorrelated and has constant variance σ^2, and that the y_i is conditional on x_i. The variance-covariance matrix of the least squares parameter estimates is obtained as

$$Var(\hat{\beta}) = (X^T X)^{-1} \sigma^2. \tag{4.8}$$

In practice, the variance σ^2 can be replaced by

$$\hat{\sigma}^2 = \frac{1}{N - p - 1} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2, \tag{4.9}$$

an unbiased estimator of σ^2 and $N > p$.

If the deviations of Y around its expectation are additive and Gaussian,

$$\begin{aligned} Y &= E(Y \mid X_1, \ldots, X_p) + \varepsilon \\ &= \beta_0 + \sum_{j=1}^{p} X_j \beta_j + \varepsilon, \end{aligned} \tag{4.10}$$

where the error ε is a Gaussian random variable with expectation zero and variance σ^2, then $\hat{\beta}$ follows the normal distribution

$$\hat{\beta} \sim N(\beta, (X^T X)^{-1} \sigma^2). \tag{4.11}$$

The hypothesis test and the confidence intervals of the estimated coefficients can be conducted based on above results

Even though the OLS method is popular for its simple solution and well-developed theories there are two reasons why the data analyst is often not satisfied with the OLS estimates. The first is prediction accuracy, the OLS estimates often have low bias but large variance. Prediction accuracy can sometimes be improved by shrinking or setting some coefficients to zero. The second reason is interpretation. With a large number of predictors, we often would like to determine a smaller subset that exhibits the strongest effects.

The two standard techniques for improving the OLS estimates, subset selection and Lasso regression will be discussed in the following sections.

4.2.2 Subset Selection

Least squares regression is used to estimate the coefficients of the inputs that are retained. But in most of the case, we only want a subset of the variables, and eliminate the rest from the model. There are a number of different strategies for choosing the subset.

Best subset regression finds, for each $k \in \{0, 1, 2, \ldots, p\}$, the subset of size k that gives smallest residual sum of squares. The question of how to choose k involves the tradeoff between bias and variance, and there are a number of criteria that one may use. Typically the model which minimizes the estimate of the expected prediction error will be chosen.

Forward stepwise selection starts with the intercept, and then sequentially adds into the model the predictor that most improves the fit. Suppose the current model has k inputs, represented by parameter estimates $\hat{\beta}$, and then one predictor is added, resulting in estimates $\tilde{\beta}$. The improvement in fit is often based on the F statistic

$$F = \frac{RSS(\hat{\beta}) - RSS(\tilde{\beta})}{RSS(\tilde{\beta})/(N-k-2)}, \tag{4.12}$$

where $RSS(\tilde{\beta})$ is the residual sum of squares for the least squares fit of the bigger model with $k+1$ parameters, and $RSS(\hat{\beta})$ is the residual sum of squares for the least squares fit of the smaller model with k parameters. The F statistic measures the change in residual sum of squares per additional parameter in the bigger model. A predictor producing the largest value of F will be added in the model, stopping when no predictor produces and F-ratio greater than the 90th or 95th percentile of the $F_{1,N-k-2}$ distribution.

Backward stepwise selection starts with the full model, and sequentially deletes predictors. Like forward selection, it typically uses an F-ratio to choose the predictor to delete. The predictor producing the smallest value of F is deleted at each stage, stopping when each predictor in the model produces a value of F greater than the 90th or 95th percentile when deleted. Backward selection can only be used when $N > p$, while forward stepwise can always be used. There are also hybrid stepwise selection strategies that consider both forward and backward moves at each stage, and make the best move; these require a parameter to set the threshold between when an addition if chosen over a deletion.

The F-ratio stopping rule provides only local control of the model search, and does not attempt to find the best model along the sequence of models that it examines. As with all-subsets selection, the model from the sequence that minimizes an estimate of expected prediction error can be chosen.

4.2.3 Lasso Selection

Subset selection is a discrete process, which means a variable is either retained or discarded in the model. This often causes high variance, and does not reduce the prediction error of the full model. To overcome this problem, shrinkage methods were proposed. The idea for the shrinkage methods is to continuously shrink the regression coefficients by using some penalty function. And by doing this, the variance of the estimated regression coefficients is reduced. Ridge regression and Lasso selection method are the two main shrinkage methods. Ridge regression is a continuous process that shrinks coefficients and hence is stable; however, it does not set any coefficients to 0 and hence does not give an easily interpretable model. Lasso is shorten for "least absolute shrinkage and selection operator" and proposed by Tibshirani, 1996. It shrinks some coefficients and sets others to 0, and tries to retain the good features of both subset selection and ridge regression.

The Lasso estimate is defined by

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2$$

$$\text{subject to } \sum_{j=1}^{p} |\beta_j| \leq t \ . \tag{4.13}$$

The L_1 Lasso penalty $\sum_1^p |\beta_j|$ is used here. This constraint makes the solutions nonlinear in the y_i, and a quadratic programming algorithm is used to compute them. Because of the nature of the constraint, making t sufficiently small will cause some of the coefficients to be exactly zero. The lasso does a continuous subset selection. If t is chosen larger than $t_0 = \sum_1^p |\hat{\beta}_j|$ (where $\hat{\beta}_j$ are the least squares estimates), then the lasso estimates are the $\hat{\beta}_j$'s. On the other hand, for $t = t_0 / 2$ say, then the least squares coefficients are shrunk by about 50% on average. The tuning parameter t should be adaptively chosen to minimize an estimate of expected prediction error. The generally used method to find the best t is the cross-validation.

In 2002, B. Efron et al. proposed a new model selection algorithm, Least Angle Regression, LARS. This method is a useful version of traditional forward selection methods. With a simple modification, so called Lasso Modification, of LARS methods, all possible Lasso estimates for a given problem can be obtained in an order of magnitude less computer time than previous computing algorithm. The computing package of LARS written by R language can be found at http://wwwstat.stanford.edu/~hastie/Papers/LARS/. In this thesis, the LARS computing package is used to calculate the Lasso estimates.

4.3 Model Selection Criterions

Assessment of the model performance in terms of its prediction capability is important in practice. This provides a guide for the choice of the model. Many criteria have been developed for this purpose and some generally used criteria are discussed here.

4.3.1 C_p Statistics and Akaike Information Criterion (AIC)

The estimated prediction error can be defined as

$$\overline{err} = \frac{1}{N}\sum_{i=1}^{N} L(y_i, \hat{f}(x_i)) \tag{4.14}$$

where N is the sample size and L is the loss function. Generally, this estimate is less than the true error $Err = E[L(Y, \hat{f}(x_i)]$, because the same data is being used to fit the method and assess its error.

The Mallow's C_p statistic is defined as:

$$C_p = \frac{RSS_p}{\hat{\sigma}_\varepsilon^2} + 2d - N \tag{4.15}$$

where RSS_p is the residual sum of squares for the fitted model with p parameters, N is the sample size and $\hat{\sigma}_\varepsilon^2$ is an estimate of the noise variance, obtained from the mean-squared error of a low-bias model. The model with the smallest C_p statistics is the final model.

The Akaike Information Criterion (also known as AIC) is based on the maximum likelihood estimates of the model parameters. And it is generally used criterion for model selection in practice. Suppose $\Pr_\theta(Y)$ is a family of densities for Y and θ is the parameter. Let $\hat{\theta}$ is the maximum-likelihood estimate of θ, and "loglik" is the maximized log-likelihood:

$$\log lik = \sum_{i=1}^{N} \log \Pr_{\hat{\theta}}(y_i). \tag{4.16}$$

For the Gaussian linear model (with variance $\sigma_\varepsilon^2 = \hat{\sigma}_\varepsilon^2$ assumed known), the AIC statistics is

$$AIC = -2 \log lik + 2d. \tag{4.17}$$

To use AIC for model selection, the model with smallest AIC over the set of models considered can be chosen as the best model. For nonlinear and other complex models, the d in the formula should be replaced by some measure of model complexity. A full study of the above issue is given in Hastie (2001), section 7.6.

Given a set of models $f_\alpha(x)$ indexed by a tuning parameter α, denote by $\overline{err}(\alpha)$ and $d(\alpha)$ the estimated error and number of parameters for each model. Then for this set of models AIC can be defined as

$$AIC(\alpha) = \overline{err}(\alpha) + 2\frac{d(\alpha)}{N}\sigma_\varepsilon^2. \tag{4.18}$$

The function $AIC(\alpha)$ provides an estimate of the estimated error curve, and the tuning parameter $\hat{\alpha}$ that minimizes the curve can be found. Then the final chosen model is $f_{\hat{\alpha}}(x)$.

4.3.2 Bayesian Information Criterion (BIC)

The Bayesian Information Criterion (BIC), like AIC, is also based on the maximum likelihood estimates of the model parameters. The general form of BIC is

$$BIC = -2\log lik + (\log N)d\,. \tag{4.19}$$

BIC tends to penalize complex models more heavily, giving preference to simpler models in selection.

Despite its similarity with AIC, BIC is motivated in quite a different way. It arises in the Bayesian approach to model selection. Suppose there is a set of candidate models M_m, $m = 1,\ldots,M$ and corresponding model parameters θ_m, and the aim is to choose a best model from these models. Assuming the prior distribution for the parameters of each model M_m is $\Pr(\theta_m \mid \mathrm{M}_m)$, and then the posterior probability of the given model is

$$\Pr(\mathrm{M}_m \mid \mathrm{Z}) \propto \Pr(\mathrm{M}_m)\Pr(\mathrm{Z} \mid \mathrm{M}_m) \tag{4.20}$$

where Z represents the observed data $\{x_i, y_i\}_1^N$. To compare two models M_m and M_l, the posterior odds are defined as:

$$\frac{\Pr(\mathrm{M}_m|\mathrm{Z})}{\Pr(\mathrm{M}_l|\mathrm{Z})} = \frac{\Pr(\mathrm{M}_m)}{\Pr(\mathrm{M}_l)}\,\frac{\Pr(\mathrm{Z}|\mathrm{M}_m)}{\Pr(\mathrm{Z}|\mathrm{M}_l)}\,. \tag{4.21}$$

If the odds are greater than one, then the model m is chosen. Otherwise, model l is chosen.

Normally, the prior over models is assumed to be uniform, so that $\Pr(\mathrm{M}_m)$ is constant. The estimated $\Pr(Z \mid \mathrm{M}_m)$ can be simplified by so-called Laplace approximation as:

$$\log \Pr(Z \mid \mathrm{M}_m) = \log \Pr(Z \mid \hat{\theta}_m, \mathrm{M}_m) - \frac{d_m}{2} \log N + O(1), \qquad (4.22)$$

where $\hat{\theta}_m$ is a maximum likelihood estimate and d_m is the number of free parameters in model M_m. If BIC is defined as

$$-2 \log \Pr(Z \mid \hat{\theta}_m, \mathrm{M}_m) \qquad (4.23)$$

This is equivalent to the BIC criterion of equation (4.22). Therefore, choosing the model with minimum BIC is equivalent to choosing the model with largest posterior probability.

For model selection purposes, there is no clear choice between AIC and BIC. BIC is asymptotically consistent as a selection criterion. And this means that given a family of models, including the true model, the probability that BIC will select the correct model approaches one as the sample size N goes to infinity. This is not the case for AIC, which tends to choose models which are too complex as N goes to infinity. On the other hand, for finite samples, BIC often chooses models that are too simple, because of its heavy penalty on complexity.

4.3.3 Cross-Validation

In practice, the simplest and most widely used method for estimating prediction error is cross-validation. If enough data was collected, some of the data can be used as the validation data to assess the performance of the prediction model. Since data are often scarce, this is usually not possible. Usually, people choose to use K-fold cross-validation method. The idea of K-fold cross-validation is to use part of the data to fit the model, and a different part to test it. The data are split into K roughly equal-sized parts. For the k th part, the model is fitted by the other $K-1$ parts of the data, and calculate the prediction error of the fitted model when predicting the k th part of the data.

Let $k:\{1,\ldots,N\}\to\{1,\ldots,K\}$ be an indexing function that indicates the partition to which observation i is allocated by the randomization. The fitted function is denoted by $\hat{f}^{-k}(x)$ and computed with the k th part of the data removed. Then the cross-validation estimate of prediction error is

$$CV = \frac{1}{N}\sum_{i=1}^{N} L(y_i, \hat{f}^{-k(i)}(x_i)) \tag{4.24}$$

When $K = N$, the case is known as leave-one-out cross-validation. In this case $k(i) = i$, and for the ith observation the fit is computed using all the data except the i th.

Given a set of models $f(x,\alpha)$ indexed by a tuning parameter α, denote by $\hat{f}^{-k}(x,\alpha)$ the αth model fit with the k th part of the data removed. Then for this set of models the cross-validation is defined by

$$CV(\alpha)=\frac{1}{N}\sum_{i=1}^{N}L(y_i,\hat{f}^{-k(i)}(x_i,\alpha)) \qquad (4.25)$$

The function $CV(\alpha)$ provides an estimate of the test error curve. The final chosen model is $f(x,\hat{\alpha})$ with the tuning parameter $\hat{\alpha}$ which can minimize the test error curve.

How to choose the K ? With $K = N$, CV is approximately unbiased for the true prediction error, but can have high variance because the N "training sets" are so similar to one another. And it was proved by Shao (1993) that leave-one-out cross-validation is asymptotically inconsistent in the sense that the probability of selecting the model with the best predictive ability does not converge to 1 as the total number of observation $n \rightarrow \infty$. The computational burden is also considerable.

On the other hand, with 5-fold or 10-fold cross-validation which is often used in practice, CV has lower variance. But bias could be a problem, depending on how the performance of the model varies with the size of the training set.

Whether the bias is a drawback in practice depends on the objective. Leave-one-out cross-validation has low bias but can high variance and the 5-fold or 10-fold cross-validation has low variance but has bias. Overall, the 5-fold or 10-fold cross-validation methods are recommended as a good compromise.

4.4 Case Study – Modeling The Duration of Lightning-Caused Sustained Line-Related Transmission Line Outages

The Lasso model selection method and Cp statistics Criterion are used to provide the relationship between duration of lightning-caused line-related transmission line outages and various weather and lightning elements. The response variable and main factors are:

Y: duration of a lightning-caused sustained line-related transmission line outage in hours.

Tmax: daily maximum air temperature when the outages occurred in Celsius degrees.

Tmin: daily minimum air temperature when the outages occurred in Celsius degrees.

Pcpn: maximum daily precipitation when the outages occurred in millimeters.

Ws: maximum hourly wind speed when the outages occurred in kilometers per hour.

Ollt10: number of overall lightning whose current amplitude is larger than 10kA

Pllt10: number of positive lightning whose current amplitude is larger than 10kA.

Omc: maximum current amplitude of the overall lightning in kiloampere

Pmc: maximum current amplitude of the positive lightning in kiloampere

The predictors have been normalized to have mean 0 and Euclidean norm 1 and the response has the zero mean. The expected models are linear models developed according to voltage levels (i.e. 72kV, 144kV and 240kV). It only considers the linear combination of the considered weather and lightning elements, the square transforms of them and the cubic transforms of them.

For the 72kV transmission lines, the major results are shown in Figure 4-1. The Cp statistics are used to determine the best model, the model with the smallest Cp statistics. When a Lasso estimator is found, a Lasso estimation step is completed. And the Lasso estimation procedure is stopped when the OLS estimator of the full model is obtained. The Cp statistics obtained the minimum when the lasso procedure is in step 3. In this step, the prediction variables with none-zero coefficients are Tmax, Tmin and cubic transform of Pmc. So the final model can be expressed as:

$$y = -6.2190 * T\max - 5.3676 * T\min + 10.5868 * Pmc^3,$$

where y is the duration of a single lightning-caused sustained line-related transmission line outage in hours for the 72kV transmission lines.

The durations of lightning-caused outages for 72kV transmission lines are negatively related to the daily maximum and daily minimum air temperature when the outages occurred with the positive lightning having the same maximum current amplitude. And under the same daily maximum and minimum air the duration of lightning-caused outages for 72kV transmission lines are dependent on the maximum current of the positive lightning. Compared to these three elements, the effects of the rest of the predictors for lightning-caused outages.are small. The general belief is that positive lightning plays the most important role for lightning-caused outages. This is confirmed by the above result for lightning-caused line-related outages on 72kV transmission lines.

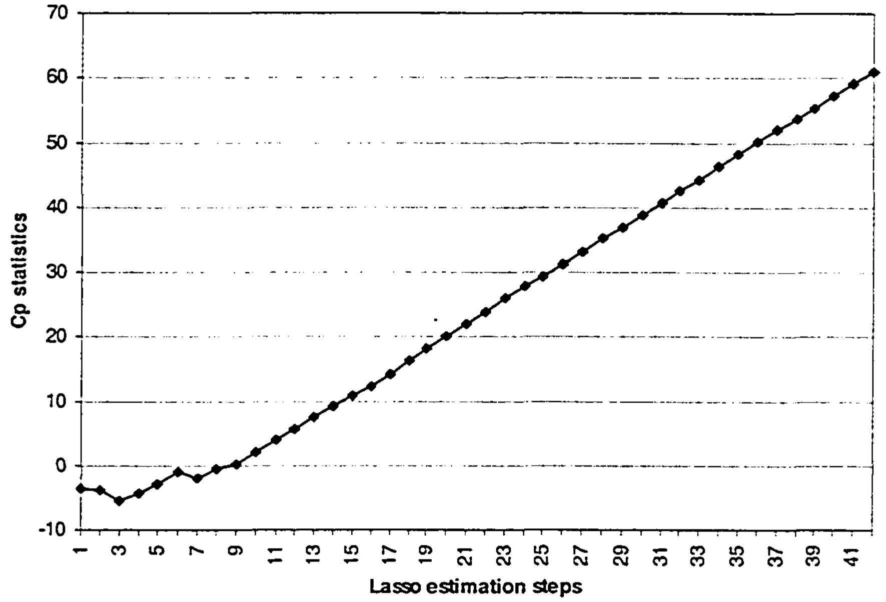


Figure 4-1. Cp statistics vs. Lasso estimation steps of estimating the regression coefficients for modeling lightning-caused line-related outages of 72kV transmission lines

Since the sample sizes of lightning-caused line-related outages for 144kV and 240kV transmission lines are relatively small and the software to complement the Lasso estimation procedure can only work well when the sample size is larger than the number of predictors, the full regression model cannot include all 24 predictors. The regression model selection procedure is applied to the 8 weather and lightning elements.

For the 144kV transmission lines, the results are shown in Figures 4-2. It is found that the major predictors in the 8 weather and lightning elements are daily maximum air temperature and hourly maximum wind speed. The coefficient of determination r^2 for the model is 0.3379. The model can be expressed as:

$$y = -14.0051 * T\max + 9.9393 * Ws \quad ,$$

where y is the duration of a single lightning-caused sustained line-related transmission line outage in hours for the 144kV transmission lines.

This result indicates that lightning-caused line-related transmission line outages are highly affected by the daily maximum air temperature and the hourly maximum wind speed. The duration of lightning-caused line-related outages of 144kV transmission lines is negatively related to the daily maximum air temperature when the maximum hourly wind speed is fixed. And under the same daily maximum air temperature, the duration of lightning-caused outages on 144kV transmission lines are only positively dependent on the maximum hourly wind speed.

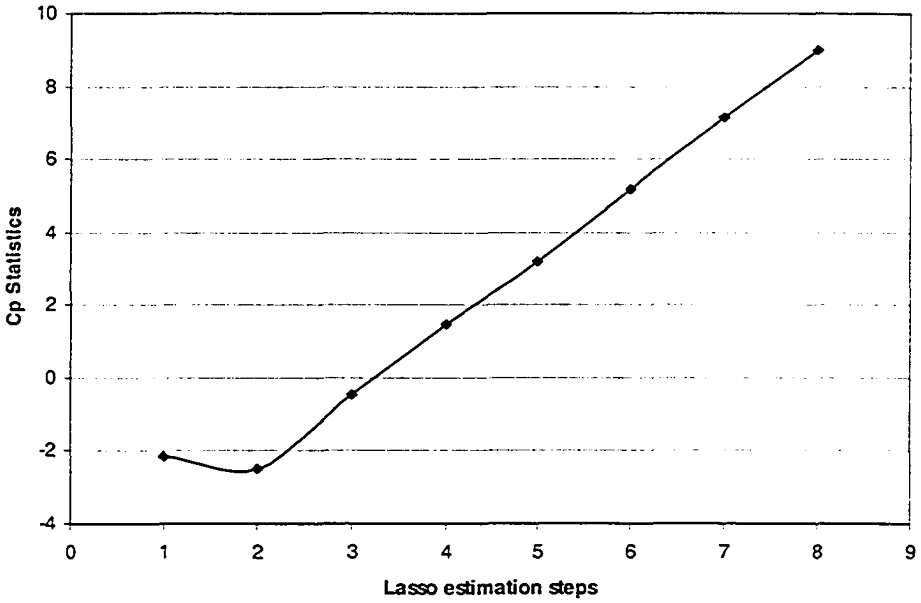


Figure 4-2. Cp statistics vs. Lasso estimation steps of estimating the regression coefficients for modeling lightning-caused line-related outages of 144kV transmission lines by using the weather and lightning elements as the predictors

For the 240kV transmission lines, the results are shown in Figure 4-3. It was found that the major predictor in the 8 weather and lightning elements is the maximum current amplitude of the overall lightning. The coefficient determination r^2 for the model is 0.4752. The model can be expressed as:

$$y = 1.2151 * Omc \ ,$$

where y is the duration of a single lightning-caused sustained line-related transmission line outage in hours for the 240kV transmission lines.

This result indicates that lightning-caused sustained line-related transmission line outages for 240kV transmission lines are highly affected by the maximum current amplitude of the overall lightning. The duration of lightning-caused line-related outages for 240kV transmission lines is positively related to the maximum current amplitude of the overall lightning.

The errors of using the above model for predicting the duration of lightning-caused outages on 240kV transmission lines are the quantitative measure of the model validity. Although the Lasso algorithm has selected this model, the linearity of the model is still in question, since the r^2 value is 0.4752 and the data points on the Omc-Duration plane are not clustered along a straight line. Further, the number of data used for estimating this model is only 22. Thus, selection of the other models may be possible when more data are available. Nonetheless, the current model still yields some useful predictions when OMC values are between –1 and 1. In this interval, the RSS value is 5.1664. For specific points of Omc = 0.1662, 0.0258, and 0.8387, the predicted and observed values are 0.2986, 0.5320, and 1.5197, respectively.

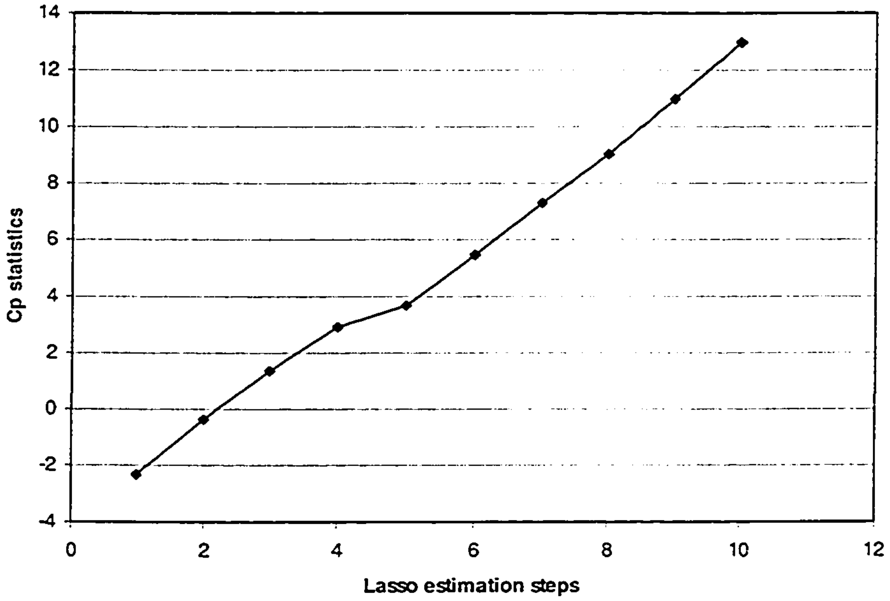


Figure 4-3. Cp statistics vs. Lasso estimation steps of estimating the regression coefficients for modeling lightning-caused line-related outages of 240kV lines by using the weather and lightning elements as the predictors

The assumption underlying the above models is that there exist a linear relationship between the response and predictors. It is worth further investigating whether this assumption holds and if the true relationship between response and predictors are non-linear. Not all the weather and lightning elements were included in the model selection. Some other weather and lightning elements besides those used here may still be of interest for further investigation.

When more data becomes available, research on analyzing the interaction among the predictors can be conducted and the estimated models may be different from the above models. Although the mechanisms of how the weather and lightning cause the outages of the power transmission lines is unknown, the models presented in this research can give power engineers useful information about which weather and lightning elements should be considered first when integrating weather and lightning elements to a power reliability evaluation model.

Chapter 5

Conclusions

This thesis used the 20-year (1983-2002) daily lightning data set provided by Alberta Lands and Forest Service, the Alberta Power Limited (i.e., now ATCO Electric) 19-year (1977-1995) outages observation data set and the Alberta Environmental Service weather data (1977 – 1996) to reveal the patterns and characteristics of weather, climate, lightning and adverse weather caused transmission lines outages.

A brief review of the weather and climate is presented in Chapter 2. The air temperature in northern Alberta is generally lower than it is in southern Alberta and the range of this difference varied in the different seasons. In winter, western Alberta, near the mountains, is warmer than other parts of Alberta because of winter Chinooks. The precipitation in Alberta is mainly in the Rocky Mountain area and more than half of the total annual precipitation was received in the summer (e.g. May to August). The spatial distribution of the wind in Alberta is not uniform. The average wind speeds in each monsoon season are gradually increased from northern Alberta to southern Alberta. The prevailing wind direction varied from season to season.

It was found that both the spatial and temporal distributions of lightning were not uniform. The high frequency center of overall lightning is located in central Alberta and the whole distribution shows the fan-shape pattern. The positive lightning is proportional (around 10%) to the overall lightning and it has similar spatial and temporal distribution only with the smaller order of the derived lightning variables (e.g. flashes/km^2/ year and days/km^2/year).

An increasing trend in the study period was detected both of annual total lightning flashes and annual positive lightning flashes. There were over 20 days of lightning occurrences in each month during the period from May to September. In June and July, lightning happened almost everyday. This period can be defined as the major lightning season in Alberta. The year-to-year days of lightning occurrences varied from place to place. The high variability of the days of lightning occurrence was found in the places with high mean days with lightning occurrences for both overall lightning flashes and positive lightning flashes.

For the adverse weather caused line-related outages in Alberta, almost all of them are sustained adverse weather caused line-related outages. Adverse weather caused line-related outages show strong geographical dependency and voltage dependency. The underlying probability distribution functions of sustained adverse weather caused line-related outages for three voltage levels (72kV, 144kV, 240kV) transmission lines are skewed to the right, which suggests that the underlying probability distribution functions are not coming from the symmetric distribution family. The probabilities of transmission line outages with long duration are different depending on the voltage level of the transmission lines.

Adverse weather caused line-related transmission line outages are not uniformly distributed by time. There is no trend for adverse-weather caused line-related transmission line outages detected. The high lightning season, from May to September, is also the high frequency season for lightning-caused line-related transmission line outages in Alberta. Lightning-caused line-related transmission line outages in Alberta are strongly monthly correlated. For the other adverse weather caused line-related transmission line outages, these relationships are not obvious. It is also of interest notice that lightning-caused line-related transmission line outages are also clustered during certain hours of a day.

Adverse weather caused outages only occurred in certain weather conditions. In terms of the outage duration, weather conditions with a high probability of lightning-caused outages have been identified. The correlation analysis between outage duration in hours and weather elements (i.e. daily mean air temperature, daily maximum precipitation and hourly maximum wind speed) has shown that there are some strong relationships between the outage durations and some of the weather elements.

The main model variable selection methods and model selection criterions were reviewed. The procedure of regression models selection by Lasso method and Cp criterion has been presented in this thesis. The main factors of the weather and lightning elements which affect the duration of lightning-caused line-related transmission line outages are identified. This information provides the starting point one needs to incorporate weather elements in a reliability evaluation system. Some applicable regression models for the purpose of estimating the main weather and lightning factors for lightning-caused line-related power transmission lines outages and prediction of the duration of the future lightning-caused line-related transmission line outage are established.

It is noted that some of the weather and lightning elements were defined as the main factors of lightning-caused line-related transmission line outages. It does not necessarily mean that the rest of the elements do not need to be considered when implementing the weather and lightning elements in the real time power transmission evaluation system. It only provides the starting points for power engineering to further investigate the physical process of how the weather and lightning affect the power transmission systems.

Further research is required to define the climatology from a power engineering point of view. Some new climate variables which will reflect the unique need of power utilities should be defined. In the future as more outage data is provided, more accurate regression models should be developed and some other model selection criterion like AIC and BIC could be applied.

References

Bhuiyan, M. R., and R. N. Allan, 1994: Inclusion of Weather Effects in Composite System Reliability Evaluation using Sequential Simulation, *IEE Proceedings – Generation, Transmission, Distribution*, **141**, 575-584.

Billinton, R., and L. Cheng, 1986: Incorporation of Weather Effects in Transmission System Models for Composite System Adequacy Evaluation, *IEE Proceedings*, **133**, 319-327.

Billinton, R., and W. Li, 1991: A Novel Method for Incorporating Weather Effects in Composite System Adequacy Evaluation, *IEEE Trans. on Power Systems*, **6**, 1154-1160.

Billinton, R, and R. N. Allan, 1996. Reliability Evaluation of Power Systems, Plenum, New York.

Burrows, W. R., P. King, P. J. Lewis, B. Kochtubajds, B. Snyder, and V. Tureotte, 2002: Lightning Occurrence Patterns Over Canada and Adjacent United States from Lightning Detection Network Observations. *Atmos.-Ocean*, **40**, 59-81.

Chetner, S. and Working group, 2003. Agroclimatic Atlas of Alberta, Alberta Agriculture, Food and Rural Development.

Cummins, K. L., M. J. Murphy, E.A. Bardo, W. L. Hiscox, R.B. Pyle, and A. E. Pifer, 1998: A Combined TOA/MDF Technology Upgrade of the U.S. National Lightning Detection Network. *J. Geophys. Res.* **103**,9035-9044.

Efron, B., T. Hastie, I. Joghstone and R. Tibshirani, 2002: Least Angle Regression. *Ann. of Statist.*, **32**, 407-451.

Electrical Power Research Institute, May 1985. Predicting Transmission Outages for System Reliability Evaluations, EPRI-3880, Project 1468-2.

Francis, D., and Hengeveld, H., 1998. Extreme Weather and Climate Change, Environment Canada.

Harrell, F. E. Jr., 2001. Regression Modeling Strategies; With Applications to Linear Models, Logistic Regression, and Survival Analysis, Springer Verlag, New York.

Hastie, T., R. Tibshirani, and J. Friedman 2001. The Elements of Statistical Learning; Datamining, Inference and Prediction, Springer Verlag, New York.

IEEE Gold Book, 1997. Recommended Practice for the Design of Reliable Industrial and Commercial Power System, IEEE Std. 493-1997.

Intergovernmental Panel on Climate Change. 2001. Climate Change 2001: The Scientific Basis.

Kochtubajda, B., R. E. Steward, J. R. Gyakum, and M. D. Flannigan, 2002: Summer Convection and Lightning over the Mackenzie River Basin and their Impacts during 1994 and 1995, *Atmos.-Ocean*, **40**, 199-220.

Koval, D. O., 1996: Transmission Equipment Reliability Data from Canadian Electrical Association. *IEEE Tran. Industry Application.* **32**,1431-1439.

Lopez, R. E., R. L. Holle, A. I. Watson, and J. Skindlov, 1997: Spatial and Temporal Distributions of Lightning over Arizona from a Power Utility Perspective. *J. Appl. Meteor.*, **36**, 825-831.

Nimchuk, N., 1989: Ground-Truthing of LLP Lightning Location Data in Alberta. *Proc. 10th Conference on Fire and Forest Meteorology*, 17-21.

Orville, R. E., G. R. Huffines, W. R. Burrows, R. L. Holle, and K. L. Cummins, 2002: The North American Lightning Detection Network (NALDN) – First Results: 1998-2000, *Mon. Wea. Rev.*, **130**, 2098-2109.

Shao, J, 1993. Linear Model Selection by Cross-Validation. *J. Am. Statist. Ass.*, **88**, 486-494.

Shen, B., Koval, D. O., and S. S. Shen, 1999: Modeling Extreme-Weather-Related Transmission Line Outages. Proc. of The 1999 IEEE Canadian Conference on Electrical and Computer Engineering, 1271-1276.

Shen, B., 1999: Analysis on Adverse-Weather-Related Power Transmission Line Outages, M Sc. Thesis. University of Alberta, Edmonton, Alberta.

Shen, S.S., 1999. An Assessment of The Change in Temperature and Precipitation in Alberta. Report prepared for Science and Technology Branch, Alberta Environment, No. 98-0282, Edmonton, Alberta.

Tibashirani, R, 1996: Regression Shrinkage and Selection via The Lasso. *J. Royal. Statist. Soc.* B. **58**, 267-288.

U.S.- Canada Power System Outage Task Force, 2003. Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations. http://www.iwar.org.uk/cip/resources/blackout-03/.