A PROBABILITY DISTRIBUTION APPROACH TO DESCRIBE CUSTOMER COSTS DUE TO ELECTRIC SUPPLY INTERRUPTIONS

A Thesis
Submitted to the College of Graduate Studies and Research
in Partial Fulfilment of the Requirements
for the Degree of
Master of Science
in the
Department of Electrical Engineering
University of Saskatchewan

by

EVA K. Y. CHAN
Saskatoon, Saskatchewan
June 1993

COPYRIGHT

ACKNOWLEDGEMENTS

The author expresses her sincere appreciation and gratitude to Dr. R. Billinton and Prof. G. Wacker for their valuable guidance and assistance during the course of this work. Special thanks are extended to Dr. R. Billinton for his patience during the preparation of this thesis.

The author is grateful to Prof. W. H. Laverty of the University of Saskatchewan for the valuable discussions. The author would also like to acknowledge the help provided by G. Tollefson.

The author takes this opportunity to acknowledge the continuous encouragement and support provided by her family and best friend Victor during the course of her studies.

Financial assistance provided by the Natural Sciences and Engineering Research Council of Canada in the form of a research grant is thankfully acknowledged.

UNIVERSITY OF SASKATCHEWAN

Electrical Engineering Abstract 93A381

A PROBABILITY DISTRIBUTION APPROACH TO DESCRIBE CUSTOMER COSTS DUE TO ELECTRIC SUPPLY INTERRUPTIONS

Student: Eva K. Y. Chan

Supervisors: Prof. Roy Billinton and Prof. Garry Wacker

M.Sc. Thesis Submitted to the
College of Graduate Studies and Research
June 1993

ABSTRACT

Reliability worth assessment is an important aspect of power system planning and operation. An equally important issue is how to utilize customer costs of electric supply interruption as surrogates to appropriately quantify reliability worth. The objective of this research work was to develop a practical alternative to the conventional customer damage function (CDF) method, to describe the interruption cost data. The alternate technique is designated in this thesis as the probability distribution approach. The fundamental cost data utilized in this thesis comes from the 1991 cost of interruption study performed by the Power Systems Research Group at the University of Saskatchewan which surveyed residential, agricultural, commercial and small industrial customers. This project was sponsored by the Natural Sciences and Engineering Research Council of Canada and eight participating utilities.

The probability distribution cost model developed in this thesis is capable of recognizing the dispersed nature of the outage cost data and can be used in a wide range of studies in each electric power system functional zone and hierarchical level. The generation of a three dimensional sector customer damage function, which describes the cost distribution patterns as a function of outage duration, is illustrated. A Monte Carlo simulation approach is utilized to estimate the Interrupted Energy Assessment Rate (IEAR) using the two different cost modeling approaches. The analysis clearly shows that the IEAR estimated using the distribution model is considerably larger than the value obtained using the CDF method.

This thesis also illustrates the application of the IEAR in conjunction with the various system operating and reliability data to conduct HLI cost / benefit assessments. The procedure is demonstrated using a small hypothetical test system.

Table of Contents

COPYRIGHT	i
ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	xi
1. INTRODUCTION	1
1.1. Power System Reliability Concepts	1
1.2. Reliability Cost / Reliability Worth	4
1.2.1. Cost of Interruption	6
1.2.2. University of Saskatchewan Database	6
1.2.3. Utilization of Interruption Cost Data	7
1.2.4. Applications in Power System Generating Capacity Planning	8
1.3. Scope of Work	9
1.4. Thesis Outline	9
2. MEASUREMENT OF RELIABILITY WORTH	11
2.1. Introduction	11
2.2. Cost of Interruption Data	12
2.2.1. Overview of Methodologies	13
2.2.2. The 1991 NSERC Cost of Interruption Study	14
2.2.3. Selection of Usable Data	14
2.3. Customer Damage Function	17
2.3.1. Economic Sectors and Groups Identification	17
2.3.2. Cost Normalization Process	18
2.3.3. Sector Customer Damage Function	19
2.3.4. Composite Customer Damage Function	20
2.3.5. An Illustrative Example Using the RBTS	21
2.4. Summary	24
3. AN ALTERNATE APPROACH TO DESCRIBE CUSTOMER COSTS OF INTERRUPTION	25
3.1. Introduction	25
3.2. Need for Investigating Customer Cost Distributions	26
3.3. Basic Probability Distribution Concepts	27

3.3.1. Random Sampling Process Assumptions 27
3.3.2. Normal Probability Distribution 28
3.4. Interruption Cost Calculation 29
3.5. Normality Transformation Technique 31
3.5.1. Advantages of Normally Distributed Cost Data 32
3.5.2. Transformation Equation 33
3.5.3. Transformation to Symmetry and the Third Moment Statistic 34
3.5.3.1. Definitions 34
3.5.3.2. Algorithm 34
3.5.4. Fourth Moment Statistic 35
3.5.4.1. Definitions 36
3.5.5. Hypothesis Testing 37
3.5.5.1. Test Criteria 39
3.6. Cost Distribution at Intermediate Durations 41
3.6.1. Cost Characteristic Parameters 42
3.6.2. Regression Analysis 42
3.6.2.1. Basic Concepts 42
3.6.2.2. Least-Squares Method and Goodness-of-fit Measurement 44
3.7. Conversion to True Cost Values 45
3.8. Summary 45

4. INTERRUPTION COST MODELING OF THE 1991 NSERC CUSTOMER RESPONSES 48

4.1. Introduction 48
4.2. Usable Peak Normalized Cost Data 48
4.3. Cost Modeling using Customer Damage Functions 51
4.3.1. Sector Customer Damage Functions 51
4.3.2. Composite Customer Damage Function 53
4.4. Cost Modeling using Probability Distributions 58
4.4.1. Data Transformation Results 58
4.4.1.1. Histogram 60
4.4.1.2. Normal Probability Plot 60
4.4.1.3. Transformation Without Including Zero-Valued Data 70
4.4.2. Cost Distribution at Intermediate Durations 71
4.4.3. Three-Dimensional Customer Damage Function 84
4.4.3.1. Generation Of A Normal Cost Curve 84
4.4.3.2. Three-Dimensional SCDF 86
4.5. Summary 88

5. GENERATION OF AN INTERRUPTED ENERGY ASSESSMENT RATE 91

5.1. Introduction 91
5.2. IEAR Estimation by Monte Carlo Simulation 92
5.2.1. Basic Models in HLI Adequacy Assessment Using Monte Carlo Simulation 93
5.2.2. Mathematical Formulation 95

5.2.3. Outage Cost Estimation Using a CCDF 96
5.2.4. Outage Cost Estimation Using Probability Distributions 96
5.2.4.1. Standardized Variable 97
5.2.4.2. Selection of a Normal Variate 97
5.3. Numerical Examples Using the RBTS 99
5.3.1. Imposing Maximum and Minimum Bounds 100
5.3.2. Simulation Convergence 101
5.4. Summary 104
6. COST / BENEFIT APPROACH TO GENERATING SYSTEM PLANNING 106
6.1. Introduction 106
6.2. Basic Concepts of the Cost / Benefit Approach 107
6.3. The Roy Billinton Test System 108
6.4. Cost of System Adequacy 110
6.4.1. System Fixed Cost 111
6.4.2. System Production Cost 112
6.4.2.1. Load Modification Technique 112
6.5. Customer Cost of Unserved Energy 114
6.6. Optimum Planning Reserve Margin for the RBTS 114
6.6.1. Effects of the IEAR Value On the Optimum Reserve Margin 117
6.7. Capacity Expansion Analysis in the RBTS 117
6.7.1. Cost Minimization Approach 119
6.7.2. Constant Risk Approach 121
6.8. Summary 123
7. SUMMARY AND CONCLUSIONS 125
REFERENCES 128
Appendix A. CRITERIA FOR SELECTING AND CODING USABLE COMMERCIAL SECTOR RESPONSES 135
Appendix B. BREAKDOWN OF ECONOMIC SECTOR BY SIC GROUP 137
Appendix C. NORMAL DISTRIBUTION 138
Appendix D. 1991 SECTOR $/KW PEAK NORMALIZED COSTS 140
Appendix E. RESULTS OF DISTRIBUTION ANALYSIS USING 1991 $/KW COST OF INTERRUPTION DATA 143

List of Figures

Figure 1.1:	Subdivision of System Reliability	2
Figure 1.2:	Hierarchical Levels of Power System Evaluation	3
Figure 1.3:	Consumer, Utility and Total Costs as a Function of System Reliability	5
Figure 2.1:	CCDF Generated from the RBTS $/KW Cost Data	23
Figure 3.1:	Normal Distribution, $N(\mu,\sigma^2)$	29
Figure 3.2:	Skewness and Kurtosis (a) Distribution differing in skewness (b) Distribution differing in kurtosis	36
Figure 3.3:	Critical Regions for (a) One-Sided Test and (b) Two-Sided Test	38
Figure 4.1:	SCDFs Generated from the 1991 $/KW Cost Data	55
Figure 4.2:	CCDF of an Example Service Area : 1991 $/KW Cost	57
Figure 4.3:	Results of $/KW Cost Transformation for the Commercial Sector (with "0" data) (a) 2-Sec $/KW	62
Figure 4.4:	Results of $/KW Cost Transformation for the Commercial Sector (without "0" data) (a) 2-Sec $/KW	72
Figure 4.5:	Scatter Diagrams of the Cost Distribution Parameters as a Function of Interruption Duration : 1991 Commercial $/KW Cost	81
Figure 4.6:	Regression Results : 1991 Commercial $/KW Cost Distribution Parameters as a Function of Interruption Duration	83
Figure 4.7:	Normal Curves for the 1991 Commercial Sector's Transformed $/KW Interruption Cost	85
Figure 4.8:	Commercial 3-D SCDF using the 1991 $/KW Interruption Cost	87
Figure 5.1:	Block Diagram for HLI Adequacy Assessment	93
Figure 5.2:	System Capacity Model : Combination of Unit Generating Cycles	94
Figure 5.3:	Risk Model : Superimposition of the Generation Model on the Load Model	94
Figure 5.4:	Standard Normal Curve, Z-Score and Variate	98
Figure 6.1:	Optimum Reliability Level as Determined by the Cost / Benefit Approach	107

Figure 6.2: Variation of Costs that Society faces directly with the Planning Reserve Margin : RBTS (using 1991 customer damage function) 116
Figure 6.3: Effect of IEAR upon the Optimum Reserve Margin of the RBTS 118
Figure 6.4: System Expansion Study of the RBTS using a Minimum Cost Criterion (IEAR = 4.46 $/KWh) 120
Figure 6.5: System Expansion Study of the RBTS using a Constant Risk Criterion 122

List of Tables

Table 2.1:	Cost of Interruption in $/KW extracted from the RBTS	21
Table 2.2:	RBTS Distribution of Energy Consumption and Peak Demand	22
Table 2.3:	Composite Customer Damage Function : RBTS	22
Table 3.1:	Basic Statistics for the 1991 Commercial Sector Cost Data ($/interruption)	27
Table 3.2:	Usable $/interruption and $/KW Responses : 1991 Commercial Sector	31
Table 3.3:	Breakdown of 1991 Commercial Responses by 2-digit SIC Group : $/interruption data	40
Table 4.1:	Number of Usable 1991 $/KW Responses	49
Table 4.2:	Aggregated $/KW Peak Normalized SIC Group Costs and Weighted Total for the Commercial Sector	52
Table 4.3:	SIC Major Groups Weighting Factors	54
Table 4.4:	SCDF's and CCDF of an Example Service Area : 1991 $/KW Cost of Interruption Data	56
Table 4.5:	Customer Mix of an Example Service Area	56
Table 4.6:	Commercial Sector Moment Test Results (with "0" data) : 1991 $/KW Cost	59
Table 4.7:	Commercial Sector Moment Tests Results (without "0" data) : 1991 $/KW Cost	71
Table 4.8:	Commercial Cost Distribution Parameters : 1991 $/KW Cost	80
Table 5.1:	RBTS Generating Unit Rated Capacity and Reliability Data	100
Table 5.2:	IEAR values from a Simulation Period of 3000 Years	102
Table 5.3:	IEAR Convergence Study by Varying Initial Seeds and Simulation Time	103
Table 6.1:	100 Points Load Data for the RBTS	109
Table 6.2:	Generating Unit Capacity and Reliability Data for the RBTS	110
Table 6.3:	Additional Generating Units for the RBTS	110
Table 6.4:	Priority Loading Order and Generating Unit Cost Data for the RBTS	111
Table 6.5:	Unit Expected Energy Output and Energy Cost for the RBTS	113

Table 6.6: Capacity Expansion Study of the RBTS using a Cost Minimization Approach (IEAR = 4.46 $/KWh) 119
Table 6.7: Capacity Expansion Study of the RBTS using a Constant Risk Approach 121
Table C.1: Cumulative Distribution Function of the Standard Normal Distribution 138
Table D.1: Average $/KW Peak Peak Normalized Costs for the Residential Sector 140
Table D.2: Aggregated $/KW Peak Normalized SIC Group Costs and Weighted Total for the Small Industrial Sector 141
Table E.1: Moment Test Results : without zero-valued data 143
Table E.2: Distribution Parameters and Results of Regression Analysis 144

LIST OF ABBREVIATIONS

CCDF	Composite Customer Damage Function
CDF	Customer Damage Function
EENS	Expected Energy Not Supplied
EES	Expected Energy Supplied
EIR	Energy Index of Reliability
F & D	Frequency and Duration
HLI	Hierarchical Level One
HLII	Hierarchical Level Two
HLIII	Hierarchical Level Three
IEAR	Interrupted Energy Assessment Rate
LDC	Load Duration Curve
LM	Load Modification
LOEE	Loss of Energy Expectation
LOLE	Loss of Load Expectation
MCS	Monte Carlo Simulation
MTTF	Mean Time To Failure
MTTR	Mean Time To Repair
NSERC	Natural Sciences and Engineering Research Council
RBTS	Roy Billinton Test System
SCDF	Sector Customer Damage Function
SIC	Standard Industrial Classification
pdf	Probability Distribution Function
C	Customer Cost of Unserved Energy
$C_i(d_i)$	Interruption Cost in \$/KW for a Duration d_i in Hours of Load Loss Event i
d_i	Duration in Hours of Load Loss Event i
x	Original Interruption Cost
y	Normal-transformed Interruption Cost

1. INTRODUCTION

1.1. Power System Reliability Concepts

A modern power system serves one primary function and that is to supply its customers with electrical energy as economically as possible and with a reasonable degree of reliability and quality. Customer expectations regarding quality of service are rising because of the high degree of dependence on electrical energy in todays social and working environment. Society has come to expect that the supply of electrical energy should be continuously available on demand. Realization of this expectation, however, is not technically feasible due to the random failure nature of the system, which is generally outside the control of power system engineers. Customers also expect to receive quality electrical service at the lowest possible cost. From a service industry standpoint, consumers' expectations of high quality and low cost can be balanced and hopefully achieved using three primary considerations: risk, cost and benefit.

The risk considerations include all aspects of the ability of a power system to provide an adequate supply of electric energy to its customers. The simplest way to diminish the likelihood of energy deficiency is to increase the investment during the planning, design and operating phases. Historically, criteria and techniques used in these phases were all deterministically based. The essential weakness of these deterministic criteria is their inability to respond to and reflect the probabilistic nature of system behaviour and of component failures [1].

Recognition of the need for probabilistic evaluation of system behaviour dates back to the early 1930s [1] and a wide range of probabilistic techniques have been developed since that time. The earliest application of probabilistic techniques was in the area of generating capacity reliability evaluation. Many papers and other contributions in this area have been published over the last fifty years [1].

Concern regarding the ability of a power system to provide an adequate supply of electrical energy is usually referred to by the term "reliability". The word, reliability, when used in power system evaluation has a wide range of meaning. Power system reliability assessment can be divided into the two basic concepts of system security and system adequacy [2] as shown in Figure 1.1. System security relates to the ability of the

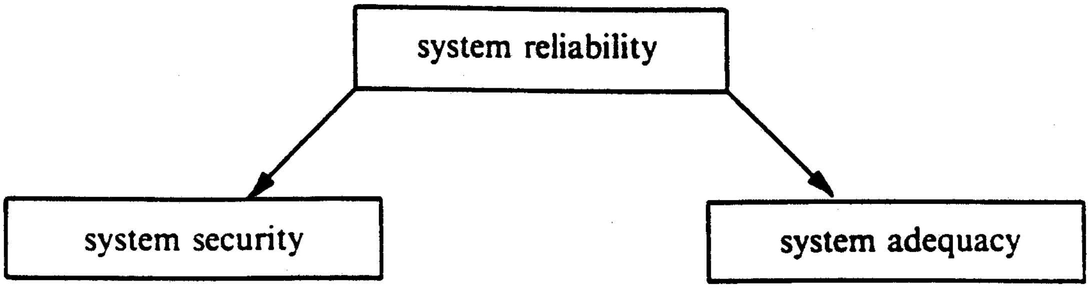


Figure 1.1: Subdivision of System Reliability

system to withstand disturbances arising within the system. These include conditions associated with both local and widespread disturbances and the loss of major generation and transmission facilities. System adequacy, on the other hand, relates to the existence of sufficient facilities within the system to satisfy the customer load requirements. These include the facilities necessary to generate sufficient energy and to transport the energy to the actual load points. The concept of adequacy is therefore associated with static conditions which do not include disturbances such as those considered under system security. It is important to realize that most of the probabilistic methods presently available for reliability evaluation are in the domain of adequacy assessment [3, 4, 5, 6]. The work presented in this thesis deals only with adequacy assessment in electric power systems.

The basic techniques utilized for power system adequacy assessment can be categorized in terms of their application to segments of a complete power system. These segments or functional zones are the areas of generation, transmission and distribution.

Adequacy studies can be conducted in each of these functional zones. The segments can be combined to create the hierarchical levels [2] shown in Figure 1.2. Hierarchical Level

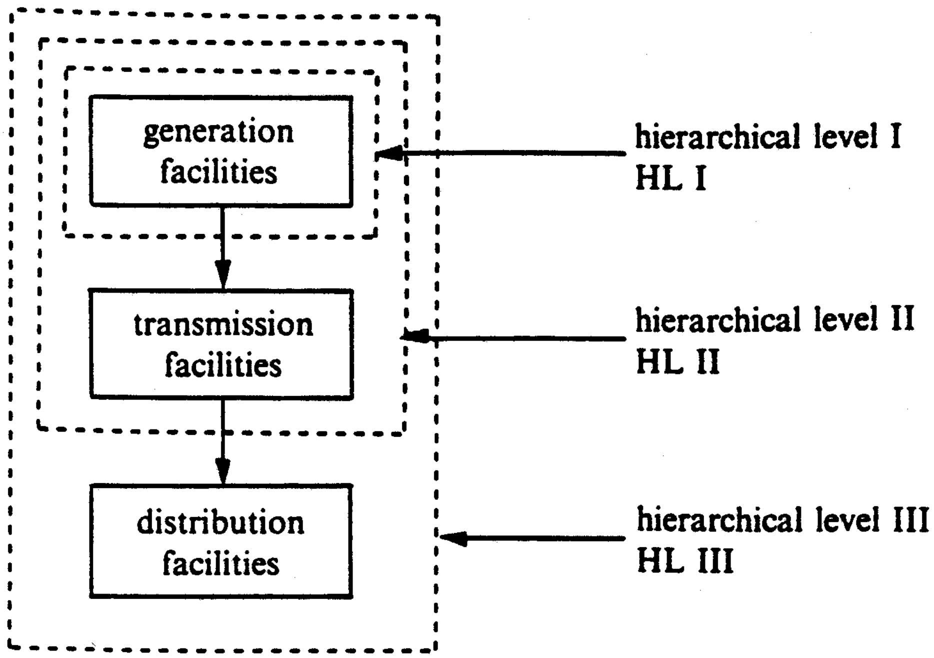


Figure 1.2: Hierarchical Levels of Power System Evaluation

I (HLI) is concerned only with the generation facilities. At this level, the total system generation is examined to determine its ability to meet the total system load requirements. This activity is usually termed as "generating capacity reliability evaluation" [1]. Hierarchical Level II (HLII) includes both generation and transmission facilities. Adequacy analysis at this level is usually termed as "composite system evaluation". Hierarchical Level III (HLIII) includes all three functional zones in an assessment of consumer load point adequacy. An HLIII evaluation can become very complex and is not usually done directly. For this reason, the distribution functional zone is usually analyzed as a separate entity [2]. The research work presented in this thesis is concerned only with HLI studies.

1.2. Reliability Cost / Reliability Worth

Predictive reliability assessment provides the means of quantifying and incorporating the reliability constraints in an expansion analysis of the system. These studies, however, are only part of an overall planning assessment. The economic constraints of the power system also play a major role in utility supply and demand-side planning. It was noted earlier that the simplest way to diminish the likelihood of a power deficiency is to increase the system reliability level. Such action without examining the cost considerations could lead to excessive operating costs. It is evident therefore that the economic and the reliability constraints can conflict and this leads to difficult managerial decisions.

In order to resolve and satisfy the dilemma between economic and reliability issues, utilities have incorporated both reliability criteria and certain cost considerations in their decision making process. The basic approach, however, is to select, based on experience and judgement, a reliability criterion, such as a Loss of Load Expectation [2] of 0.1 days/yr. Alternative expansion plans are then examined based on meeting this criterion. The plan with the lowest present worth over the planning horizon is then presumed to be the optimum expansion plan. The present worth of an expansion scheme refers to the system investment cost needed to achieve the preselected reliability level and is often termed as "reliability cost". A basic question in this approach is "what is an acceptable level of reliability?" It has been shown that most Canadian utilities base their selection of a reliability criterion on experience and judgement, and do not explicitly incorporate cost considerations [7]. In order to determine what is an acceptable level of reliability, utilities must recognize the real benefit or the perceived value of electrical energy to their customers. An awareness of the benefit can provide valuable input to balancing the economic and reliability constraints in power system planning. An approach which is receiving considerable attention at the present time is to compare the "reliability cost" with the "reliability worth".

The term, reliability worth, refers to the benefit derived by the customer in receiving electrical energy. Reliability worth assessment provides the means of improving the cost effectiveness of utility supply and demand-side planning by integrating customer perspectives in the assessment process. This approach allows power system planners to conduct generating capacity planning using customer-driven requirements in addition to the conventional system reliability indices.

The basic concept of reliability cost / reliability worth evaluation is relatively simple and can be illustrated using Figure 1.3. The utility curve in Figure 1.3 shows that the

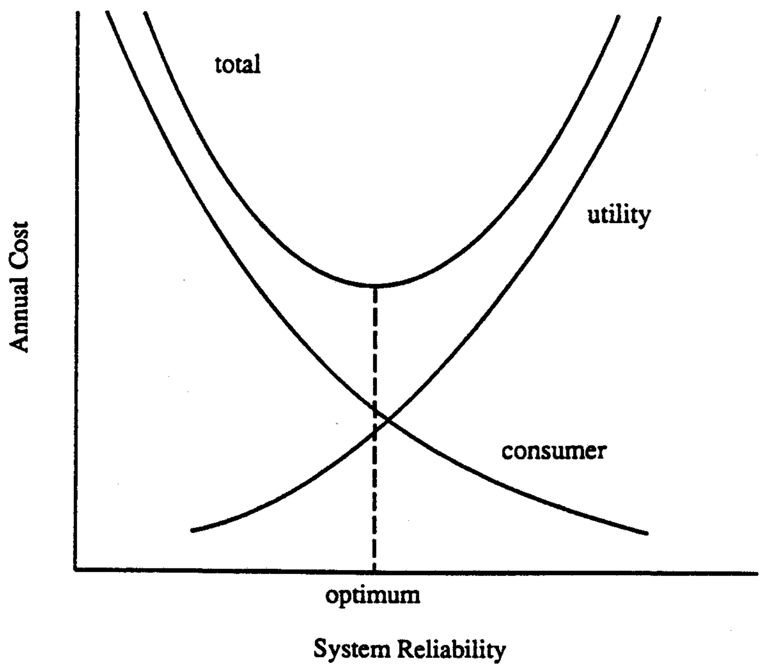


Figure 1.3: Consumer, Utility and Total Costs as a Function of System Reliability

system cost will generally increase with higher investment cost in equipment and facilities as consumers are provided with higher reliability. On the other hand, the consumer costs associated with supply interruptions decrease with fewer occurrences in power failures as the reliability increases. The total cost to society is the sum of these two individual costs. The resulting total cost curve exhibits a minimum which corresponds to the optimum or target level of reliability.

Two comprehensive bibliographies [8, 9] concerning the concepts of reliability worth and its measurement have been published. These bibliographies contain the bulk of the major contributions in the area of power system reliability worth assessment.

1.2.1. Cost of Interruption

Assessing the worth of reliability is a very difficult task to conduct directly. The difficulty comes from the fact that there are many intangibles involved in the evaluation process, which are not always amenable to quantification in dollar values. Many approaches have been attempted over the past two decades to quantify reliability worth. Most of these methods are based on the assessment of the effects and impacts of unreliability [10]. It is believed that quantifying the costs and losses incurred by electrical consumers as a result of power deficiencies is an easier task. The unreliability costs or the costs of interruption are not identical to the worth of reliability but are considered to be reasonably representative measures.

The various methods that have been used to obtain cost of interruption data can be conveniently grouped into three categories, namely, indirect analytical evaluations, studies of actual blackouts, and customer surveys. The major contributions made since 1980 to obtaining outage costs are shown in [9]. The three alternatives were recently reviewed by Sanghvi [11] and the pros and cons of each method are discussed in [10]. Among the three methods, customer surveys have the distinct advantage in that the customer is in the best position to assess the effects of power cessation and therefore best able to estimate the associated costs. The customer survey approach, therefore, is rapidly gaining acceptance as the preferred methodology by many electric utilities.

1.2.2. University of Saskatchewan Database

Over the last twelve years, the Power Systems Research Group at the University of Saskatchewan has conducted several customer surveys to determine the impacts of power interruptions on Canadian electrical users. The Group conducted its first extensive study of interruption costs in the residential, commercial and small industrial sectors in 1981 [12]. The second survey, which covered the agricultural sector, was conducted a few years later in 1985 [13]. Both of these studies were sponsored by the Canadian Electrical Association (CEA). The Group is currently conducting a study funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada in conjunction with seven Canadian electrical utilities [14]. The customer survey approach was again chosen as the data collection scheme for this study. This project was performed by five members of the Group, including the author, under the supervision of Dr. R. Billinton and Professor G. Wacker. The author's primary task as a project member was to conduct data analysis (both cost and non-cost related) in the commercial and small industrial sectors. The cost of interruption data from this latest survey provides the essential elements of the reliability worth assessment presented in this thesis.

1.2.3. Utilization of Interruption Cost Data

The outage cost estimates obtained from survey respondents vary widely for different customers and with different interruption related characteristics, such as interruption duration, frequency, time of occurrence, etc. The interruption duration is usually considered to be the primary variable. The outage cost for a given type of customer, as a function of interruption duration, is referred to as a Customer Damage Function (CDF). The CDF for a specific economic sector is designated as a Sector Customer Damage Function (SCDF). A Composite Customer Damage Function (CCDF) can be obtained by weighting the SCDF's for all sectors in a studied service area by their relative energy consumption. The generation of a CCDF for a service area is an attempt to define the total customer costs for that area as a function of interruption duration. The procedures for developing a CCDF are presented in [15] and [16]. The major shortcoming of the CCDF representation is that the function considers only the average monetary losses at various interruption scenarios. It does not, and can not, reflect the dispersed nature of the data within a specific customer group.

A CCDF can be expressed in terms of cost per respondent ($/interruption), cost normalized by the respondent's annual peak demand ($/KW) or cost normalized by the respondent's annual consumption ($/KWh). The most common form is the $/KW as it can be used directly for subsequent calculations leading to a single value which quantifies the worth of reliability.

Customer interruption costs must be related to a predictable system adequacy index in order to provide a practical tool for application in system design, planning and operation. A factor designated as the Interrupted Energy Assessment Rate (IEAR) [17] can be used to link the cost to the system reliability. Using an IEAR in conjunction with the system Expected Energy Not Supplied (EENS) or the Loss of Energy Expectation (LOEE) gives the cost of unserved energy in dollars. This unserved energy cost quantifies the worth of reliability as a single representative value which reflects the expected damage to the economy of a specific service area when power supply cessation occurs. It is the primary ingredient in forming the customer cost curve in Figure 1.3.

1.2.4. Applications in Power System Generating Capacity Planning

The concept of reliability cost / reliability worth (sometimes referred to as the cost / benefit approach) can be effectively utilized in generating capacity (HLI) planning. One of the primary applications is to evaluate the optimum level of system adequacy, which is often expressed in terms of a percent reserve margin [18]. The reserve margin can be defined as the additional generating capacity required in excess of the load. Such a margin is necessary and must be planned in advance to safeguard against non-scheduled outages, ageing of equipment etc. The cost / benefit method can be used to determine the future required reserve margin for which the sum of system and customer costs is minimum. Another important application of the cost / benefit criterion is in the area of capacity expansion analysis. Future increases in electrical energy demand requires designing, building and commissioning additional generating units to maintain acceptable levels of system adequacy. The conventional approach in HLI system expansion planning is to preselect a target reliability index such that the number of new units required to meet this target can be determined. If a system has a calculated index which exceeds the critical value, additional units are added until the calculated index satisfies the desired value. Using the cost / benefit approach, the number of new units required can be determined by evaluating the total societal cost associated with the system after each unit addition. The system configuration which corresponds to the lowest cost is selected. This method therefore considers the benefit of capacity expansion as seen by the customers in addition to the costs incurred by the power system.

1.3. Scope of Work

The primary objectives of the research work described in this thesis were to:

1. Investigate an alternative technique to the CCDF approach, to describe the cost of interruption data collected from the 1991 customer survey conducted at the University of Saskatchewan.
2. Present the 1991 $/KW interruption cost analysis using both CCDF and the newly developed approach.
3. Develop IEAR's from both the CCDF and the new approach using Monte Carlo simulation.
4. Illustrate the application of IEAR in an HLI reliability cost / benefit study.

1.4. Thesis Outline

The basic concepts of reliability worth and its measurement are presented in Chapter 2. The chapter includes a brief overview of the various methods of obtaining interruption costs. The 1991 customer survey conducted at the University of Saskatchewan is discussed together with the author's contributions. This chapter also illustrates the procedures used in developing a CCDF together with some applications using a hypothetical test system.

Chapter 3 begins with a discussion of the weaknesses of the conventional CCDF. The chapter then proceeds to introduce the idea of probability distributions of the cost estimates. A detailed procedure for determining the distribution pattern of a group of data is given. The chapter concludes by stating concerns regarding the distribution patterns of non-studied durations and proposes a solution to the problem using regression analysis.

The outage cost data collected from the 1991 Canadian electrical customer survey are analyzed and the results presented in Chapter 4. Both the CCDF and the probability distribution approaches are utilized to describe the data.

Chapter 5 illustrates the role of the two different cost models in the generation of an IEAR using an HLI Monte Carlo simulation technique. Numerical examples using the hypothetical test system are presented and discussed.

Chapter 6 illustrates the application of an IEAR for cost / benefit assessment in generating capacity planning. The study includes the evaluation of an optimum planning reserve margin and a capacity expansion analysis. All the studies are illustrated using the 1991 cost of interruption data applied to the hypothetical test system.

The summary and conclusions of the thesis are presented in Chapter 7.

2. MEASUREMENT OF RELIABILITY WORTH

2.1. Introduction

The price consumers are willing to pay for the benefit they associate with any product or service is referred to as its perceived value [19]. In today's increasing competitive marketplace, this concept of perceived value is most frequently applied in manufacturing and services industries with regard to the development of product features, pricing, packaging and promotion strategies. Quality of service therefore becomes a customer-driven requirement. Until recently, this perceived value concept has had little place in electric power utility planning. The major reason behind this is that the industry has had total market dominance and has not found itself in competition similar to that faced by most other service industries. Application of the perceived value concept, however, has been shown to be a useful tool to integrate supply and demand-side planning and to help to identify cost effective market strategies, rate options, and services which are beneficial to both utilities and their customers [19]. The perceived value of service is commonly referred to in the electric power supply industry as the benefit accruing to customers by receiving a given level of service reliability and is simply termed as the reliability worth.

The basic function of a modern electric power system is to satisfy the system load requirement at the lowest possible cost, while maintaining a reasonable level of reliability. It is, however, difficult to define just what is a "reasonable" level of reliability. The justification is traditionally based on past experience and judgement. There is increased concern that a more extensive economic justification of the selected reliability level is required. This concern is heightened by the fact that utilities are faced with increasing energy costs, environmental concerns and the need to conserve resources [14]. The concern has demanded that increased emphasis be placed on customer perspectives in utility planning and operation.

In order to appropriately examine system reliability from an economic viewpoint, it is necessary to first associate the reliability level with the investment cost required to achieve it and secondly, with the benefit accruing to customers in receiving it. The ability to assess the level of reliability within a system and the cost associated with it is well established. In comparison, the perceived value, benefit or worth of electrical service reliability is not particularly easy to define and more difficult to evaluate. The essential problem comes from the fact that there are many intangibles involved in the evaluation process which are not always amenable to quantification in dollar values. An approach that can be used to solve this problem is to approximate the reliability worth by the customer costs or losses resulting from electrical supply interruptions. These measures of unreliability are not identical to the reliability worth but are considered to be acceptable surrogates. The substitution is based on a theoretically sound assumption that the sum of all losses experienced by a customer as a result of power cessation should be equal to the amount the customer would be willing to pay for avoiding it. This chapter provides an overview of the methodologies used to estimate costs of interruption and illustrates the application of these data to develop customer damage functions (CDF).

2.2. Cost of Interruption Data

Interruption costs can be broadly classified into direct and indirect costs. Direct costs are those arising directly from a power outage and include such impacts as lost production, spoilage of food or raw materials, lost personal leisure time, or paid staff unable to work. Indirect costs are related to secondary consequences of power failures such as crime during a blackout (short term) and business relocation (long term) [10]. Most of the direct impacts are relatively easy to identify and quantify, while indirect effects are typically quite difficult to quantify in monetary terms. This section provides an overview of the range of methods used to evaluate reliability worth or costs of unreliability. Some details on the 1991 NSERC cost of interruption study in which the author participated are also presented.

2.2.1. Overview of Methodologies

The various methods that have been used to evaluate costs of interruption can be conveniently grouped into three categories, namely, analytical methods, case studies, and customer surveys.

Analytical methods, in a general sense, analyse the interruption costs from primarily a theoretical economic perspective. Perhaps the best known approach among these is to quantify reliability worth by relating the use of electricity to the Gross National Product (GNP) [20]. This method, though very simple and easy to use, has several shortcomings. The inability to provide assessments other than on a large geo-political scale is considered its weakest aspect. Another major disadvantage relates to its inability to address the non-production sectors, i.e. the residential sector.

The case studies approach attempts to estimate losses associated with actual power interruptions. This approach has been limited to major, large-scale disturbances such as the 1977 New York blackout [21]. The study considered a wide range of societal impacts together with the direct and indirect consequences of the extensive outage. A very important finding of this study was that the indirect costs were much higher than the direct costs. Such results and other conclusions drawn from a blackout case study, while valuable to the particular incident, cannot be easily generalized.

The customer survey method is based on the assumption that the customer is in the best position to estimate the losses due to the unavailability of electricity. This approach is considered to be the most customer specific method of obtaining cost outage information. It can easily include the effect of variables such as timing, duration, and frequency of interruptions, and can be tailored to suit a utility's needs. There is no doubt that the expense and effort of conducting customer postal surveys is significantly higher than the other two methods, nevertheless, it appears to be the method favoured by most utility planners.

2.2.2. The 1991 NSERC Cost of Interruption Study

The Power Systems Research Group at the University of Saskatchewan is currently conducting a Canadian customer survey to assess power system reliability worth. This study is funded by NSERC in conjunction with seven Canadian power companies. The survey, upon its completion, will cover the residential, commercial, small industrial and agricultural sectors. Several papers and reports based solely on data collected from this activity have been published. A general overview of the survey methodology used is given in [14]. The reader is referred to some of the earlier publications of the Group [22, 23] for more detailed explanations of survey rationale, questionnaire development and content. Analyses of the cost related questions are reported in [14] and [24]. The responses to non-cost related questions are presented in [24] and [25]. More complete details and results will be presented in the final project report. The author began her participation in this project in the Fall of 1991 with the primary responsibility of conducting data analysis (both cost and non-cost related) in the commercial and small industrial sectors. The findings from the cost related analyses are the primary focus of this thesis.

2.2.3. Selection of Usable Data

The results of any statistical analysis are very dependent on the sample data. A good set of data is therefore required in order to yield meaningful, representative and convincing results. It should be appreciated that the data selection procedures and criteria described following are tailored to suit the 1991 NSERC study and may not be applicable to other surveys.

The number of usable responses is question specific due to the fact that respondents do not necessarily answer all questions and/or some questions are poorly answered. The complete list of criteria for selecting the usable responses for each question in the commercial sector are given in Appendix A. Details on the other studied sectors are given in the NSERC final project report [26]. Unless all possible errors or ambiguities are identified and cleared in the first place, they may contaminate all subsequent analyses. A statistical software package called the SPSS-X Data Entry [27] system was used to perform the data "cleaning" function. The SPSS-X system also allows verification of errors introduced by human mistakes during data entry. This is facilitated by giving the operator the opportunity to enter the data twice. If any mismatch occurs between the two entries, the system will prompt the operator for the correct entry.

Since the analysis of cost related responses is of primary interest in this study, the selection of usable cost estimates is discussed explicitly in this section. The selection is particularly important in analysis based on average values because of the high dependence of the mean value on outliers. An extreme sample of data on the high end will grossly inflate the sample mean while an extreme value on the low end may cause significant deflation.

Prior to checking the abnormality of the cost data, the factors affecting their magnitudes must first be identified. It is understood that an outage cost is largely influenced by its associated interruption duration. Generally, the longer the duration, the higher the cost. Based upon this, a check for consistency is performed to exclude those responses with higher losses for shorter durations than for longer durations. The consistency check does not apply to the 2-sec, 1-min and 20-min scenarios since a large group of respondents indicated that momentary loss of electricity causes more serious damage to their equipment (mainly computers and torque-driven machines) than a longer power failure.

Another major factor which affects the magnitude of an outage cost is the type of customer. In each economic sector, respondents sampled from the same utility service area are expected to conduct similar activities and therefore their outage costs are expected to fall within a relatively small range. Based upon this, for each interruption scenario, cost estimates were checked for abnormal values in each sector. The ten highest and lowest values were identified and the respondents were contacted to verify the values. In most cases, explanations were provided and the extreme costs either remained unchanged or were changed in accordance with the clarification provided. The outliers were excluded in only very few cases because of their abnormality or because the respective respondents could not be contacted.

If a zero value is reported for a particular interruption scenario, it is important to ask if the respondent really means absolutely no monetary loss or is he/she implying that there is difficulty in estimating the loss. The determination can be made based on the cost estimates given at adjacent durations. For example, if a "0" is given for the 2-hr duration while a non-zero value is given for the 1-hr duration, it is logical to assume that the reported "0" cost is not a valid estimate. Zero values with ambiguity about their meanings were excluded from the analysis.

Continuity of cost data is particularly important in the generation of a cost model [15]. It is understood that not all respondents will answer the questionnaire completely or even if they do, the answers may not be usable for analysis in accordance with the criteria described above. In the previous studies involving generation of CDF [16, 28], the practice was to include data from only those respondents who answered the cost question completely. Respondents who gave one or more bad answers were excluded from the CDF generation process. The resulting average estimates at various studied scenarios are considered to be continuous since they were calculated from values given by a consistent set of respondents. This concept of continuity allows the loss associated with an intermediate duration (duration that has not been studied in the questionnaire) to be confidently and meaningfully interpolated. The same approach was initially applied, in this study, to the commercial and small industrial sectors. In both analyses, only respondents who reported usable answers continuously from the 2-sec to 1-day scenarios were included. The resultant number of data in each scenario when compared to the original responses drops significantly as less than 60% of the respondents provided complete answers.

While continuity is important to the damage function, the function also has to be representative. If less than 60% of the usable responses are used, the resulting average costs may not be considered reliable and representative estimates of the sampled population. The dramatic reduction of the sample size resulting from the continuity selection approach therefore calls for a better selection method. Several different approaches have been tried by the author to improve the usable sample size. It was finally decided to use a similar criteria, to that described above, which includes respondents who gave consistent answers continuously from the 1-hr to 1-day scenarios. Respondents who reported usable values from 1-hr to 1-day completely and left some or all of the values at 2-sec, 1-min and 20-min unfilled or poorly answered were considered acceptable under the new criteria. This new approach yields a significant increase in sample size. The percentage of usable data when compared to the original criteria increases to 80%. All the cost estimates used in the generation of interruption cost models presented from this point are based on data selected using this scheme.

2.3. Customer Damage Function

Outage cost data must be capable of conversion into a form which can be used in subsequent calculations of power system reliability worth. The Customer Damage Function, which describes the cost of interruption for a given type of customer as a function of duration, can be used to serve this purpose. The remainder of this chapter focuses on the generation of customer damage functions.

2.3.1. Economic Sectors and Groups Identification

There are many factors which contribute to the magnitude of an interruption cost. These factors can be either customer related or interruption related. One of the most influential customer related variables is the type of customer. The Power Systems Research Group has adopted the following four economic divisions to categorize the Canadian electrical users from whom the 1991 NSERC cost of interruption data were collected,

1. Small Industrial Sector,
2. Commercial Sector,
3. Agricultural Sector, and
4. Residential Sector.

These major sector divisions can be further divided into various functional groups according to the products or services they provide. Statistics Canada's Standard Industrial Classification (SIC) protocol is normally used for this purpose because of its wide acceptance and availability. The SIC breakdown is not used in the residential sector. Appendix B gives a list of SIC class descriptions for the other three sectors.

The CDF which describes the interruption costs as a function of duration for a given economic sector is often referred to as the Sector Customer Damage Function (SCDF). Depending on the availability of SIC breakdowns in a particular sector, the SCDF can be obtained using either the weighted sum of the SIC group costs or as a simple average of the entire sector. Each economic sector is expected to experience a different cost for a particular outage duration and the procedure for combining the individual sector costs utilizes a weighting process. Summation of the weighted SCDF's yields the Composite Customer Damage Function (CCDF). This function represents the total costs associated with power interruptions as a function of duration for all users in the studied service area.

2.3.2. Cost Normalization Process

It has been stated previously that customers having similar characteristics have been categorized into groups in this interruption cost study. The individual costs in each group often exhibit a large variation. One of the reasons for this variation is the great diversity in the energy requirements of the consumers within each group. Dollar or absolute cost values therefore are often normalized, assuming that the cost variations between respondents from the same category will be significantly reduced [12]. The consumers energy requirements can be determined from their annual consumption or annual peak demand. These two values are often used as the normalizing variables applied to the costs per interruption. When the cost per interruption is normalized by the user's annual energy consumption, the ratio is designated as the cost per kilowatt-hour ($/KWh). The cost per interruption normalized by the user's annual peak demand is designated as the cost per kilowatt ($/KW). It should be clearly appreciated that $/KWh estimates are not obtained by dividing by the unsupplied energy due to interruptions. The normalized $/KW values are of primary importance in studies pertaining to system cost and benefit as they can be conveniently used in subsequent calculations leading to the overall cost of unsupplied energy.

2.3.3. Sector Customer Damage Function

A conventional CDF describes the monetary losses for a given interruption scenario in terms of expected or average values which can be evaluated using a bottom-up approach [12, 15, 16, 28]. Depending on the available information on the SIC groups in a particular sector, the first step of the bottom-up approach is either the calculation of SIC group costs or the calculation of a simple overall sector average. These calculations can be expressed mathematically as follows:

$$\begin{array}{l}\textit{Average Cost}\\ \textit{in the } j^{th} \textit{ SIC category}\end{array} = \frac{\sum_{i=1}^{k} cost_i}{k} \quad (\$/interruption), \tag{2.1}$$

$$\begin{array}{l}\textit{Aggregated Peak–normalized Cost}\\ \textit{in the } j^{th} \textit{ SIC category}\end{array} = \frac{\sum_{i=1}^{n} cost_i}{\sum_{i=1}^{n} peak_i} \quad (\$/KW), \tag{2.2}$$

$$\begin{array}{l}\textit{Aggregated Consumption–normalized}\\ \textit{Cost in the } j^{th} \textit{ SIC category}\end{array} = \frac{\sum_{i=1}^{m} cost_i}{\sum_{i=1}^{m} cons_i} \quad (\$/KWh), \tag{2.3}$$

where:

$cost_i$	is cost estimate in \$ of respondent i,
$peak_i$	is annual peak demand in KW of respondent i,
$cons_i$	is annual consumption in KWh of respondent i,
j	is the SIC class index of a given sector,
k	is the number of usable cost estimates (\$) in the j^{th} group,
n	is the number of respondents for which both usable cost estimates (\$) and peak demand (KW) values are available, and,
m	is the number of respondents for which both usable cost estimates (\$) and consumption (KWh) values are available.

Equations 2.1 to 2.3 are applicable to sectors with SIC breakdowns, i.e. the small industrial, commercial and agricultural. The same equations, however, can be applied to the residential sector but with the exception that the respondents in the entire sector are all considered at the same time without SIC divisions.

Equation 2.1 is a simple calculation of the average dollar interruption costs in a given group. The two normalized costs were calculated similarly with one exception and that is aggregated averages are computed instead of the simple expected values. Calculations of the aggregated averages using Equations 2.2 and 2.3 are performed by summing the dollars costs for the respondents in each SIC category and dividing this total cost by the total of the energy consumption (or peak demand) associated with the group. The aggregating process reduces the effect of respondents who have fairly low consumptions but high interruption cost estimates, as discussed in [12]. It must be appreciated that the aggregated cost calculations include only those respondents from whom both dollar cost estimates and annual consumption values (or peak demand values) are available.

The second step of the bottom-up approach is to weigh and sum the average or aggregated SIC group costs in a given sector to create a Sector Customer Damage Function (SCDF). The SIC weighting factors for a particular sector are the relative amounts of energy consumed by the respective user groups within that sector. This weighting process is not used to obtain the residential SCDF.

2.3.4. Composite Customer Damage Function

The final step of the bottom-up approach is to combine the various SCDF's to yield a CCDF for the entire studied area. The generation of a CCDF for a given service area is an attempt to define the total average customer costs for that area as a function of interruption duration. Each economic sector in the studied area is expected to experience a different cost for a particular interruption duration. The approach for combining the individual sector costs is to create a weighted average. In order to obtain this average, the customer mix for the area must be known so that various sector costs can be proportionally weighted by their respective annual energy consumption or peak demand.

Weighted sector costs at a given interruption scenario are then summed to give the overall average monetary interruption loss for the entire service area.

2.3.5. An Illustrative Example Using the RBTS

The interruption cost data from a hypothetical test system designated as the RBTS [29] are used to illustrate the procedure for combining various SCDF's to form a CCDF. The cost data used in the test system were obtained from studies conducted by the Group and by Ontario Hydro. The tables and figures presented in the remainder of this chapter are based on extractions from the RBTS data.

Table 2.1 gives cost of interruption data ($/KW) by sector using a 1987 Canadian dollar base. These sector $/KW values were obtained using Equation 2.2. The customer

Table 2.1: Cost of Interruption in $/KW extracted from the RBTS

Sector	1 min	20 min	1 hr	4 hr	8 hr
Large Users	1.005	1.508	2.225	3.968	8.240
Industrial	1.625	3.868	9.085	25.163	55.808
Commercial	0.381	2.969	8.552	31.317	83.008
Residential	0.001	0.093	0.482	4.914	15.690
Government	0.044	0.369	1.492	6.558	26.040
Office Space	4.778	9.898	21.065	68.830	119.160
Agriculture	0.060	0.343	0.649	2.064	4.120

mix for the studied area must be known to combine the seven sector costs at any given duration. Table 2.2 shows the energy consumption distribution and the peak demands for each sector. In this study, weighting by annual peak demand was used for short duration interruptions (1-min and 20-min) and weighting by energy consumption used for interruptions longer than one half an hour [12]. The resulting CCDF is tabulated in Table 2.3 and depicted in Figure 2.1. Both the cost ($/KW) and the duration (min) are plotted on a logarithmic scale in order to yield the best visual display of all the five data points. Each point represents an average interruption cost at a particular studied duration. These points are joined together by straight line segments to form a piece-wise linearly increasing function. This simple model allows easy calculation of any average intermediate cost using linear interpolation between two adjacent known averages. Extrapolation can be used to obtain costs associated with interruptions lasting less than one minute. Interruptions which last longer than eight hours will be assigned the same cost as the average 8-hr loss.

Table 2.2: RBTS Distribution of Energy Consumption and Peak Demand

Sector	Energy (%)	Peak Demand (%)
Large User	31.0	30.0
Industrial	19.0	14.0
Commercial	9.0	10.0
Agricultural	2.5	4.0
Residential	31.0	34.0
Government	5.5	6.0
Office Space	2.0	2.0

Table 2.3: Composite Customer Damage Function : RBTS

Interruption duration	Interruption Cost (1987 $/KW)
1 minute	0.67
20 minutes	1.56
1 hour	3.85
4 hours	12.14
8 hours	29.41

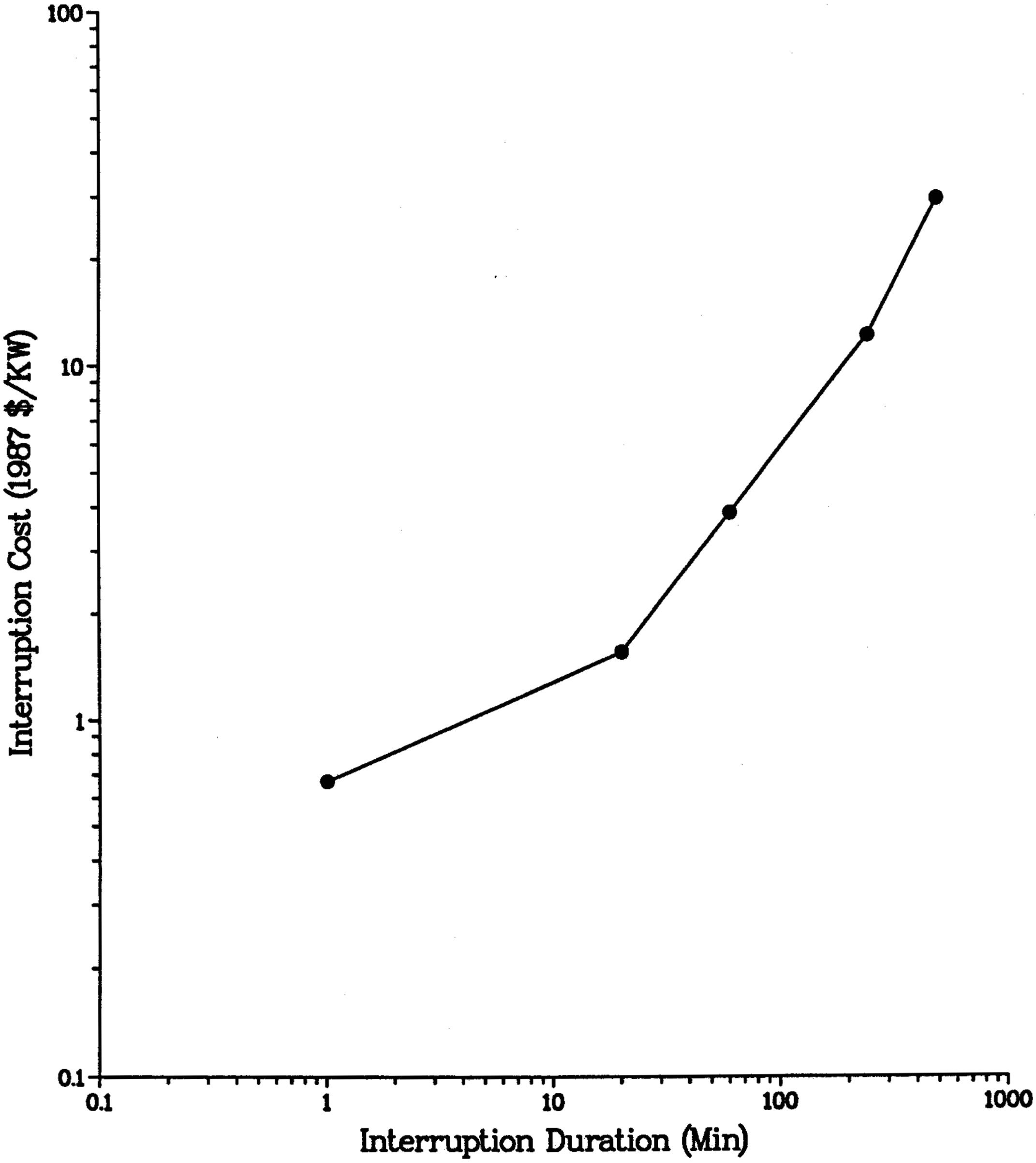


Figure 2.1: CCDF Generated from the RBTS $/KW Cost Data

2.4. Summary

This chapter presents a brief philosophical discussion on reliability worth evaluation. The problem of evaluating reliability worth has been approached by using a surrogate, namely the customer cost of an interruption. An overview of the methodologies used in estimating interruption costs is presented together with a very brief discussion of the merits and demerits of each. The customer survey approach, though expensive, appears to be the method favoured by most power utility companies since it can easily incorporate customer effects and interruption related variables in the evaluation process. The 1991 NSERC cost of interruption study is briefly discussed and the data selection criteria designed particularly for the study is presented.

A procedure commonly used to generate the CCDF is also described in this chapter. It should be noted that the SCDF for a non-SIC sector is obtained by simple sector averages while the SCDF for a sector with SIC breakdowns is obtained by summing the weighted SIC group costs. The sector weighting process is then utilized to combine the various SCDF's to form a CCDF for the entire studied area. This weighting procedure is illustrated using a hypothetical test system (RBTS).

3. AN ALTERNATE APPROACH TO DESCRIBE CUSTOMER COSTS OF INTERRUPTION

3.1. Introduction

Customer costs of interruption are believed to provide a representative indirect measure of reliability worth. This belief is based on the assumption that the monetary loss resulting from a power failure is equal to the price customers are willing to pay for avoiding it. The customer survey approach has three distinct advantages over the other interruption cost estimation methods presented in Chapter 2. It is based on the premise that the customer is in the best position to assess his/her monetary losses associated with a power failure. This approach can also easily include effects of customer and interruption related variables and as it is quite flexible, it can be tailored to fit a utility's unique requirements.

Using the survey approach, assessment of reliability worth is based solely on the usable outage cost data gathered from the respondents. The assessment, however, cannot be performed until the data are converted into a practical representation or cost model. The traditional cost model is known as a Composite Customer Damage Function (CCDF). A CCDF defines the overall *average* costs of interruption as a function of duration in a given study area. An important question to ask in this approach is how well do the expected values represent the entire response? Interruption cost analyses presented in [13], [16] and [28] show that the dollar values exhibit a very large variation. In some cases, the standard deviation is more than four times the mean. The cost normalization process was introduced to reduce this problem intending that cost variations between respondents from the same sector will be significantly reduced through this process. It was observed, however, that the normalized values still have considerable variation. Some aspects of this issue are addressed in [28].

Relatively little consideration has been given to the variation of cost values about their means. This information, however, is crucial to a complete understanding of the cost profile. A major objective of this research work therefore was to investigate the dispersed nature of the interruption cost data. This chapter introduces a new concept, designated as the probability distribution approach, to represent the data. This is an attempt to describe the variation in the cost estimates commonly disregarded by the conventional CCDF representation.

3.2. Need for Investigating Customer Cost Distributions

Cost of interruption data is the fundamental ingredient used in quantifying the worth of reliability. A representative and reliable cost model therefore is essential in both demand and supply-side utility planning. The conventional CCDF representation, though easy and simple to use, contains no information on the scattered nature of the outage loss estimates. The simple function may therefore not be considered to provide a reasonable representation when most data display large differences from the means.

The interruption costs described by a CCDF are essentially expected values, average values or in statistical terms, the mean values. The mean is a measure of central tendency, a numerical value that tends to locate in some sense the *middle* of a set of data. The variability and spread among the data can be described by basic statistics such as the range, variance and standard deviation. These statistics are shown in Table 3.1 for the $/interruption data obtained from the 1991 commercial customer survey. The fact that the dollar values display a very large variation (i.e. standard deviation), which in most cases is over three times the mean, emphasizes the need to have some indication of how good or useful are the mean values. While the mean value represents a good approximation of the interruption cost encountered by some customers, many customers experience much different losses. For example, consider the statistics associated with the 1-sec interruption scenario given in Table 3.1. Although the average cost is low ($140.7/interruption), some users have significantly larger losses (e.g. $7500.00/interruption). Conversely, the average cost for a 1-day interruption is large ($17138.7/interruption) but there are also some users who indicate that they have negligible losses. The positive skewness value associated with each data group implies

Table 3.1: Basic Statistics for the 1991 Commercial Sector Cost Data ($/interruption)

Interruption Duration	Mean	Standard Deviation	Minimum	Maximum	Skewness
1 second	140.7	787.4	0.00	7500.0	7.72
1 minute	170.9	810.6	0.00	7500.0	7.03
20 minutes	400.5	1187.0	0.00	9158.0	5.06
1 hour	1182.6	4798.5	0.00	61375.0	10.17
2 hours	2087.5	6499.8	0.00	62250.00	6.94
4 hours	4352.9	15000.0	0.00	158908.00	7.52
8 hours	7806.7	23385.8	0.00	220600.00	6.65
1 day	17138.7	116875.8	0.00	1685000.00	13.71

that the number of respondents with costs higher than the mean value greatly exceeds those with smaller costs. It is evident therefore that the conventional CCDF fails to provide complete information on the cost profile. The basic statistics given in Table 3.1 clearly illustrate the need to develop a new approach to describe customer interruption costs.

3.3. Basic Probability Distribution Concepts

The variability of a set of widely scattered data is most commonly measured and described by a probability distribution [30]. In this section, basic random variable and probability distribution concepts are addressed. Special emphasis is given to the normal random variable and its probability distribution since it is the most important continuous random variable encountered in practice.

3.3.1. Random Sampling Process Assumptions

Good data collection or sampling techniques are essential in a survey study as any inferences concerning the population will ultimately be made based on the statistics calculated from the sample. The cost of interruption surveys conducted at the University of Saskatchewan assume that the sample is selected such that their opinions and characteristics are representative of the entire population of electrical users in the studied area [12]. The primary purpose of this assumption is to enable the use of inferential statistics (e.g. tests of hypothesis) to interpret results from the sample and then use them to draw inferences regarding the population. A valid sampling procedure provides each electrical user in a studied sector with the same opportunity of being selected. The reader is referred to [12] for a detailed description of the sampling procedure.

A characteristic of interest from each individual element of a random sample is defined as a random variable [30]. In this case, the outage costs experienced by the randomly selected respondents are the random variables under study. Reported outage cost estimates are observed values of the random variables. The probability pattern that gives the relative frequencies associated with all the possible values of a random variable in the sample is known as the probability distribution of the random variable [30]. The shape of a distribution defined by a mathematical formula can provide a formal and yet simple indication of the data behaviour.

3.3.2. Normal Probability Distribution

There are two distinct types of probability distribution, namely discrete distributions whose domain is the set of all discrete numbers, and continuous distribution whose domain is the set of all real numbers. It is reasonable to assume that cost estimates follow a continuous distribution since the data are given in dollars which conceptually can take any possible real value.

There are many different types of probability distribution available to describe a continuous random variable. The most common ones are the normal distribution, the exponential distribution and the Weibul distribution. The normal distribution is the most common distribution in use and is a continuous distribution with a symmetrical bell-shape contour. Its position and shape are solely determined by the mean (μ) and standard deviation (σ). The parameter, μ, corresponds to the central position of the distribution; and the parameter, σ, determines the spread about μ as illustrated in Figure 3.1. A normal distribution is commonly denoted by $N(\mu,\sigma^2)$ where σ^2 is defined as the variance of the distribution. As shown in Figure 3.1, 95% of the observations fall

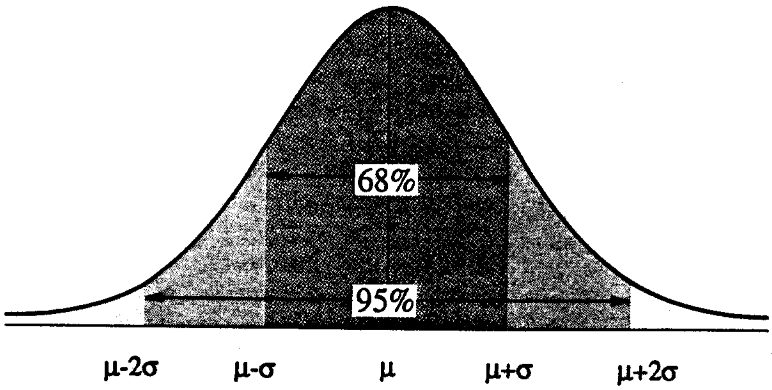


Figure 3.1: Normal Distribution, N(μ,σ^2)

within two standard deviations of the mean, and, 68% within one standard deviation. The exact theoretical proportion of cases falling into various regions of the normal curve can be found from Table C.1. Deviations from normality can be characterized by the standard third and fourth moments of a distribution [31]. The third standardized moment, $\sqrt{b_1}$, characterizes the skewness of a distribution while the fourth standardized moment, b_2, gives the kurtosis. These two standard moments are zero for a perfect normal distribution.

3.4. Interruption Cost Calculation

Before performing distribution analysis, special attention must be paid to the conditions under which data are provided. These conditions or factors can affect the magnitude of an outage cost estimate. The primary factors are the duration of interruption and the type of customer. It is expected that for any specific interruption duration, the set of data associated with the residential sector will be quite different from that associated with the commercial or industrial sectors. In order to conduct a customer and interruption specific study, the data therefore are grouped such that each group has a unique and distinctive distribution. The data collected for the 1991 cost study were divided into four major groups: residential, commercial, small industrial and farm sectors. Each of these groups was then subdivided into smaller categories using the interruption scenarios developed for the sector. The commercial sector study, for instance, consists of eight individual analyses containing data for 2-sec, 1-min, 20-min, 1-hr, 2-hr, 4-hr, 8-hr and 1-day interruptions. The cost data were not grouped according to SIC category due to the lack of sufficient responses in some of the classes. This issue is explained later in this chapter.

The individual respondent outage costs are calculated using the following equations in order to conduct the distribution analysis.

$$\textit{Dollar Cost of the } i^{th} \textit{ respondent} = cost_i \quad i=1,\cdot k \quad (\$/interruption), \tag{3.1}$$

$$\textit{Peak–normalized Cost of the } i^{th} \textit{ respondent} = \frac{cost_i}{peak_i} \quad i=1,\ m \quad (\$/KW), \tag{3.2}$$

$$\textit{Consumption–normalized Cost of the } i^{th} \textit{ respondent} = \frac{cost_i}{cons_i} \quad i=1,\ n \quad (\$/KWh), \tag{3.3}$$

where:

$cost_i$	is the cost estimate in \$ for respondent i,
$peak_i$	is the annual peak demand in KW for respondent i,
$cons_i$	is the annual consumption in KWh for respondent i,
k	is the number of usable cost estimates (\$),
n	is the number of respondents for which both usable cost estimates (\$) and peak demand (KW) values are available, and,
m	is the number of respondents for which both usable cost estimates (\$) and consumption (KWh) values are available.

It should be realized that k, m and n do not necessary have the same value. The \$/interruption cost distribution model is determined using k pieces of data where the data selection is based on the criteria described in Section 2.2.3. The $/KW and $/KWh cost models are determined using *m* and *n* pieces of data respectively. The number of $/KW normalized cost data *m* depends simultaneously upon the number of usable cost estimates ($) and the availability of the annual peak demand data (KW). A respondent's outage cost will be normalized and included in the analysis if his/her dollar estimate is usable and at the same time a non-zero peak demand value is supplied. The same rule is applied to the *n* pieces of $/KWh normalized cost data. Table 3.2 presents a comparison between the amount of $/interruption and $/KW data in the 1991 commercial sector survey. Due to insufficient peak demand information, the usable $/KW data are largely reduced.

Table 3.2: Usable $/interruption and $/KW Responses : 1991 Commercial Sector

	Interruption Durations							
	2 sec	1 min	20 min	1 hr	2 hr	4 hr	8 hr	1 day
$/int	143	145	176	216	216	216	216	216
$/KW	49	53	65	72	72	72	72	72

3.5. Normality Transformation Technique

Unless a formal examination is conducted, it is difficult if not impossible to identify the distribution model which best fits a given set of cost data. Conceptually, the data can follow any possible continuous distribution. Rather than arbitrarily selecting a probability distribution and examining its appropriateness to describe the data, a more systematic procedure designated as Normality Transformation was used to conduct the analysis. The idea is to transform a set of cost values (calculated using Equation 3.1, 3.2 or 3.3) in such that they can be represented by a normal distribution. This section focuses on the transformation procedure and emphasizes the advantages of such an approach.

3.5.1. Advantages of Normally Distributed Cost Data

The primary advantages in having the various groups of interruption cost normally distributed are as follows. There are more tools available to test normality than any other distributions. While the goodness-of-fit techniques are not yet available for testing all existing continuous distributions, they are relatively well established in regard to testing the appropriateness of a normal distribution as the underlying phenomenon from which the data arise. The best known goodness-of-fit technique in testing normality perhaps is the Moment Test [31].

The normal distribution is also a very crucial representation employed in many statistical analyses. It is by far the most important theoretical distribution in statistics and serves as a reference for describing the form of many other distributions. It is also an essential requirement for inferential statistics [32].

It must be realized that cost data are not available at all possible durations since only a limited number of interruption scenarios can be included in a survey questionnaire. In order to establish the distribution model at an unquestioned or intermediate duration, the analysis will have to utilize results obtained from adjacent questioned durations. This task will be virtually impossible if the data at each studied duration have a different distribution model as the correlation between various models is extremely difficult to define. The third advantage of the normality transformation therefore is that the process permits the inference of intermediate cost distributions. Using this approach, data groups are transformed into various normal distributions characterized by different means and variances. It is believed that finding the relationship between various normal parameters is a more realistic task than attempting to determine the correlation between various distinct models.

3.5.2. Transformation Equation

A power transformation can be used to stabilize the variability about the mean [32]. This is done by raising each data value to a specified power. For example, a power transformation of 2, squares all of the data values and a transformation of 1/2 uses the square root of all the values. Box and Cox [33] suggested the following family of power transformations:

$$y = \begin{cases} \frac{x^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \log(x) & \text{if } \lambda = 0, \end{cases} \tag{3.4}$$

where x refers to the original data, λ is the power exponent and y is the transformed value. There are two limitations to this family of equations. It applies only to continuous variables and it does not apply to zero-valued data. All "0" estimates, which involve in the transformation process presented from this point, remain unchanged.

Box and Cox introduced Equation 3.4 to ease the process of linear regression by reducing the non-normality of the dependent variable [34]. They also developed a maximum likelihood estimator of λ for this purpose. The family of transformations, however, employed in this study does not relate to regression analysis. It is instead utilized to reduce non-normality in a *single variable* study which is, in this case, the customer interruption cost. The maximum likelihood estimation of λ hence is not applicable and is not used. A new approach was developed to find the λ value that will best satisfy normality tests of a single variable distribution. The approach involves two major steps:

1. The original data groups are transformed into various sets of symmetrical distributions using the Box and Cox power transformation.
2. The best normal transformation is chosen from the set of symmetrical distributions using a hypothesis test.

The statistics involved in these procedures are the third ($\sqrt{b_1}$) and fourth (b_2) standardized moments. Details on their functions in the distribution analysis are presented in the following sections.

3.5.3. Transformation to Symmetry and the Third Moment Statistic

It has been stated previously that a perfect normal distribution is characterized by both b_2 and $\sqrt{b_1}$ being zero. It is, however, difficult to consider both elements simultaneously while conducting a normality transformation and it was therefore decided to focus first on the skewness of the transformed data.

3.5.3.1. Definitions

The first step in the normality transformation process is to perform a symmetrical transformation. An *iteration* approach is utilized with the skewness $\sqrt{b_1}$ as the stopping rule parameter. The $\sqrt{b_1}$ statistic is defined as [31]:

$$\sqrt{b_1} = \frac{m_3}{m_2^{3/2}}, \tag{3.5}$$

where

$$m_K = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^{-K}}{n}, \qquad K = 2 \text{ or } 3, \tag{3.6}$$

and

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}, \tag{3.7}$$

where y refers to the transformed cost in this particular study and n is the sample size. If a distribution is symmetrical about its mean $\bar{y}$, as is the normal distribution, $\sqrt{b_1}=0$. Values of $\sqrt{b_1} \neq 0$ indicates skewness and therefore non-symmetry. Figure 3.2 (a) is an illustration of distributions with different $\sqrt{b_1}$ values.

3.5.3.2. Algorithm

The closer the $\sqrt{b_1}$ value is to zero, the more symmetrical is the distribution. The measure of skewness will typically never be exactly zero for a given set of sampling data, but will fluctuate about zero because of sampling variations or an imperfect random process. The approach used to find the best symmetrical transformation searches for a minimum value of $|\sqrt{b_1}|$. The iterative technique for finding this value is described as follows:

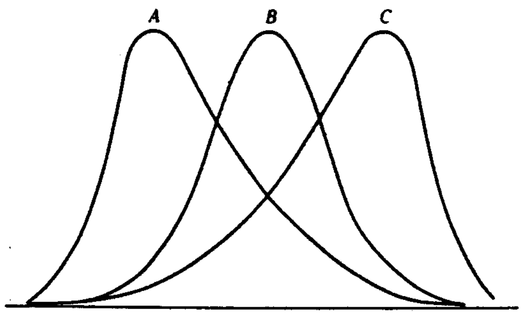
1. Set the $\sqrt{b_1}$ stopping rule value to 0.001.
2. Set the initial $\lambda = 0$.
3. Calculate a set of transformed data, y, using Equation 3.4.
4. Calculate the corresponding $\sqrt{b_1}$ statistic for the transformed cost using Equations 3.5 to 3.7.
5. If the absolute value of $\sqrt{b_1}$ calculated in Step 4 is equal to or less than the value set in Step 1, the iteration process terminates.
6. If the $\sqrt{b_1}$ value calculated from Step 4 is greater than zero and its absolute value is greater than the stopping rule value set in Step 1, λ should be decreased by a small value ($\lambda_{new} = \lambda_{old} - 0.05$). The iteration process continues by going back to Step 3.

   If the $\sqrt{b_1}$ value calculated from Step 4 is less than zero and its absolute value is greater than the stopping rule value set in Step 1, λ should be increased by a small value ($\lambda_{new} = \lambda_{old} + 0.05$). The iteration process continues by going back to Step 3.

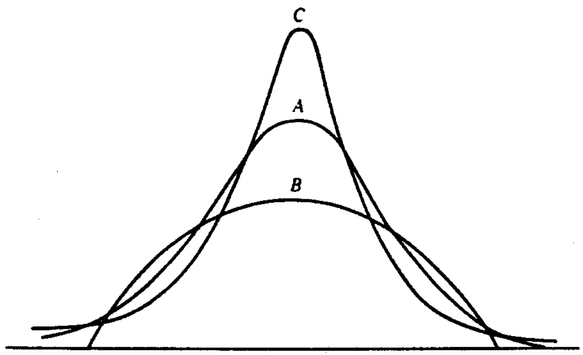
This iteration process generates a number of transformed data sets before the $\sqrt{b_1}$ stopping value is reached. The last value of λ generated in the process gives the best symmetrical transformation. This transformation, however, does not necessarily produce an appropriate normal distribution as a normal distribution must also satisfy the kurtosis condition.

3.5.4. Fourth Moment Statistic

The iteration process described above creates for each group of data a set of approximate symmetrical distributions using different λ values. In order to select the most adequate set of normally distributed data from these distributions, each transformation must be checked against the goodness-of-fit tests of normality using the fourth standardized moment statistic b_2 jointly with the skewness criterion.



(a)
A. $\sqrt{b_1} > 0$, B. $\sqrt{b_1} = 0$, C. $\sqrt{b_1} < 0$.



(b)
A. $b_2 = 0$, B. $b_2 < 0$, C. $b_2 > 0$.

Figure 3.2: Skewness and Kurtosis (a) Distribution differing in skewness (b) Distribution differing in kurtosis

3.5.4.1. Definitions

The mathematical formulation of the b_2 statistic is given by [31]

$$b_2 = \frac{m_4}{m_2^{\,2}} - 3, \tag{3.8}$$

where

$$m_K = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^{-K}}{n}, \qquad K = 2 \text{ or } 4, \tag{3.9}$$

and

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}, \tag{3.10}$$

where y is the random variable under study and n is the sample size. The b_2 statistic characterizes the kurtosis, tail-thickness or peakedness of a distribution [35] and its value is zero for a perfect normal distribution. Values of $b_2 > 0$ indicate distributions with "thicker" tails than normal, and values of $b_2 < 0$ indicate distributions with "thinner" tails. Figure 3.2 (b) gives a graphical illustration of symmetrical distributions at different b_2 values.

3.5.5. Hypothesis Testing

Selection of the best λ value is based on the idea that one of the symmetrical distributions achieved using the iteration algorithm developed in Section 3.5.3 will provide the best fit to a normal distribution. This selection was performed using a *hypothesis testing* [30] procedure. The test consists of two major steps. The first step is to formulate two hypotheses. The null hypothesis, H_o, is the hypothesis upon which attention is focused. Generally this is a statement that something is true. For example, the statement "the transformation with $\lambda = 0$ results in a normal distribution" is a null hypothesis. In contrast, an alternative hypothesis, H_a, is a statement that the null hypothesis is not true. For example, an alternate hypothesis can be a statement such as "the transformation with $\lambda = 0$ does not give a normal distribution".

The second step is to determine the test criteria which consist of (i) selecting a test statistic, (ii) specifying a level of significant, α, and (iii) determining the critical region. The value of the test statistic is used to make the decision to "fail to reject H_o" or to "reject H_o". The test statistic values which lead to the rejection of the null hypothesis defines the critical region and the location of the critical region is determined by the alternate hypothesis statement. If the alternate hypothesis contains information regarding the direction in which the critical region is located, namely less than (<) or greater than (>), then a one-sided test should be performed. On the other hand, if H_a is a statement that something is NOT true (≠), then a two-sided test should be performed [30]. Assuming the test statistic is a normal variable, this relationship can be illustrated using Figure 3.3. If the test statistic falls within the critical region (the shaded region of Figure 3.3), the null hypothesis H_o will be rejected. In contrast, the H_o will not be rejected when the test statistic falls outside the critical region.

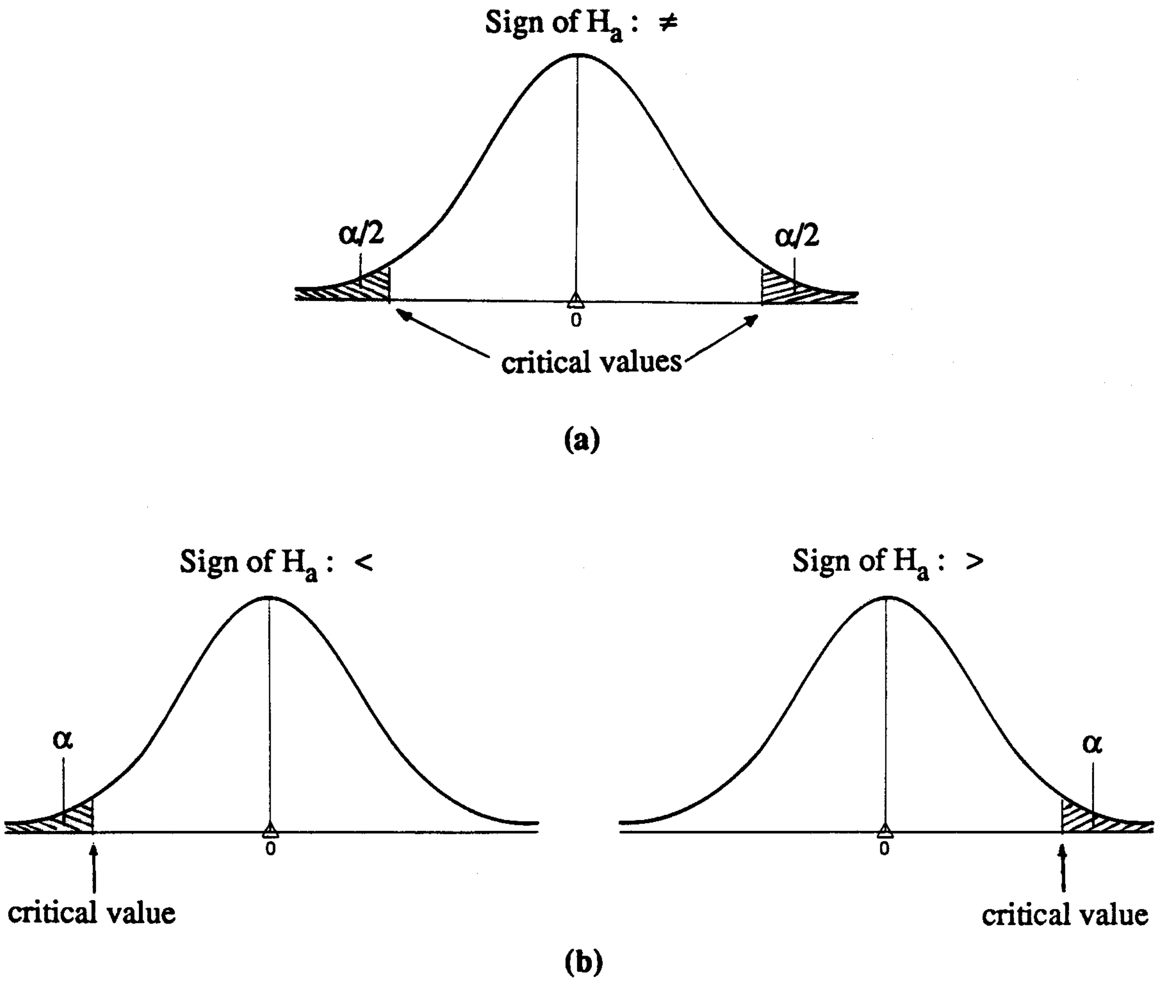


Figure 3.3: Critical Regions for (a) Two-Sided Test and (b) One-Sided Test

The critical value of the test statistic is dictated by the probability of committing a type I error, α [35]. This error results when the null hypothesis is rejected when it is, in fact, true. The value assigned to α will depend on the seriousness or impact of the error. A relatively small probability should be used if the error could have serious effects. The most frequently used probability values for α are 0.01 and 0.05. If the test statistic is an approximately normal variable, α represents the area of the critical region under a normal curve as shown in Figure 3.3.

3.5.5.1. Test Criteria

The criteria used in hypothesis testing to select the best normality power transformation factor are listed as follows:

1. $\sqrt{b_1}$ and b_2 are the test statistics.
2. H_0: The data transformed using λ_i are normally distributed;
   H_a: The data transformed using λ_i are NOT normally distributed,

   where λ_i refers to the power transformation factor associated with iteration i.
3. The two-sided significant level is 0.05. The area of the critical region on each side of the normal curve therefore is 0.025.
4. The $\sqrt{b_1}$ and b_2 statistics are computed by the S_u approximation, and the Anscombe and Glynn approximation respectively [31]. The resulting statistics, Z_1 and Z_2, are approximately standard normal variables with zero mean and unity variance.
5. Given Z_1 and Z_2, the decision to "reject" or "fail to reject" H_0 is made by reference to any standard normal distribution table (i.e. Table C.1).

   When a two-sided significant level of 0.05 is in place, H_0 will be rejected if $|Z_k| \geq 1.96$ where k = 1 or 2. (Note: 1.96 is the value of z obtained from Table C.1 when the area is 0.025)

The $\sqrt{b_1}$ and b_2 statistics are not used directly in the decision making process as they are not normally distributed. The hypothesis test described above requires every test statistic to be an approximately normal variable. Consequently, Z_1 and Z_2 are used because of their standard normal properties. This hypothesis test requires a minimum sample size of seven [31]. Table 3.3 shows that some of the commercial sector responses in the 1991 study do not have a usable SIC group size of seven or greater. Data transformation therefore was not conducted by SIC category.

Table 3.3: Breakdown of 1991 Commercial Responses by 2-digit SIC Group : $/interruption data

SIC Group	Interruption Durations 2 sec	1 min	20 min	1 hr	2 hr	4 hr	8 hr	1 day
60	26	28	39	45	45	45	45	45
61	6	6	6	6	6	6	6	6
62	11	11	11	13	13	13	13	13
63	27	25	32	39	39	39	39	39
64	5	5	6	7	7	7	7	7
65	20	22	25	32	32	32	32	32
69	4	4	4	5	5	5	5	5
91	7	7	7	11	11	11	11	11
92	10	11	12	15	15	15	15	15
96	12	12	14	16	16	16	16	16
97	4	3	6	9	9	9	9	9
99	11	11	14	18	18	18	18	18

A computer program COSNOR written in Fortran-77 was developed to combine the iteration process of symmetrical transformation and the hypothesis test selection procedures. The values of λ, $\sqrt{b_1}$, Z_1 and Z_2 are computed during each iteration. These values from iteration i are stored if both absolute values of Z_1 and Z_2 are less than 1.96. The iteration process is continued until $\sqrt{b_1}$ is less than or equal to 0.001. Following the iteration process, only those transformations which satisfy the hypothesis test are retained by the computer. If more than one set of iteration results are stored, the group with the smallest $\sqrt{b_1}$ value is selected. The associated λ value is the best normality power transformation factor.

3.6. Cost Distribution at Intermediate Durations

It is quite obvious that in reality the duration of a power failure is not limited to one of the questioned durations and it can take any possible real value. Unfortunately, a survey questionnaire can only include a very limited number of interruption scenarios and therefore cost estimates are not readily available for every possible case. The problem then becomes one of how to infer intermediate costs using statistics calculated from the known durations.

Finding an average cost at any intermediate or non-questioned duration is simple and straight forward when a conventional CCDF is used to describe the interruption costs. As discussed in Section 2.3.5, linear interpolation between two calculated averages from the adjacent studied durations is used in most cases. The same problem however is more difficult if the costs at each studied duration are described by a probability distribution. The interpolation method is not applicable as no single average cost is calculated for any particular studied duration.

It is noted in Section 3.5.1 that the utilization of a unique distribution model to describe the various groups of cost data has definite advantages. It is very difficult, if not impossible, to find an appropriate intermediate cost model if the various studied durations have quite different distributions (i.e. normal, exponential or Weibul etc.). The problem becomes easier to attack if the data at all studied durations can be transformed into a unique model. The normal distribution was selected as the common model in this study mainly because of its popularity and simplicity. When the normality transformation approach is applied to a group of cost data, the results can be characterized by a distinctive set of parameters. The following sections include the discussion of these parameters and the method for using them to establish a distribution pattern at an intermediate duration.

3.6.1. Cost Characteristic Parameters

The distributed nature of the interruption cost data for a particular customer sector and a specific outage scenario can be characterized by the following parameters:

1. Proportion of zero-valued data, P_z,
2. Normality power transformation factor, λ,
3. Mean of the normal-transformed distribution, μ and
4. Variance of the normal-transformed distribution, σ^2.

This set of parameters is unique to each data group. P_z gives the proportion of original zero-valued data. λ converts the original data such that the transformed distribution satisfies the goodness-of-fit test of normality. μ defines the central location of the resulted normal distribution while σ^2 describes its dispersion.

3.6.2. Regression Analysis

The problem addressed in this section is how to predict the parameters P_z, λ, μ and σ^2 at a new non-questioned duration based on the known values at the questioned durations.

3.6.2.1. Basic Concepts

Regression analysis is a statistical tool for evaluating the relationship of one or more independent variables A_1, A_2, ..., A_k to a single dependent variable B. One major application is to provide an equation to describe (e.g. predict) the dependent variable as a function of the independent variables. The simplest form of regression problem deals with one dependent variable and one independent variable. Four separate and independent regression analyses examining the relationship of the outage duration with each of the P_z, λ, μ and σ^2 were conducted. The interruption duration is the independent continuous variable and the four parameters are the dependent variables in their respective studies.

More information is generally sought in a regression analysis than a simple description of the observed data. Inferences are often drawn about the relationship between the dependent and independent variables in the population from which the sample was taken. Several assumptions must be made in order to perform this function. These assumptions are not listed here and interested readers are strongly encouraged to consult the related material given in [32] and [36]. It is assumed in this study that all assumptions are satisfied and that observations drawn from the sample can be used as inferences of the entire studied population.

Regression analysis starts with finding the curve or mathematical model (e.g. straight line, parabola, etc.) that best fits the data in such a way as to closely approximate the true (but unknown) relationship of the dependent and independent variables. Given a sample of n observations, each value of A will have a value of B. These n pairs of observations or observed points can be denoted by (A_1, B_1), (A_2, B_2), ..., (A_n, B_n) and are considered as data points on a two-dimensional plot known as a scatter graph. Once a scatter graph is generated, the next task is to choose the most appropriate mathematical model to describe it. There are no clear rules for finding an appropriate model and it is more or less a trial and error process. An intelligent choice, however, involves recognizing the "shape" of the observed data and being aware of the variety of mathematical functions which might provide an acceptable fit. In practice, the analyst usually starts with the simplest possible model, namely the straight line. Once a model is selected, the next step is to determine the "best-fitting" representation. There would certainly be no problem in deciding what is "best-fitting" if the data permitted a single mathematical formula to describe each and every point in the graph. Unfortunately, this will never happen with real-life data (e.g. survey data). It should therefore be realized that any model is an approximation to the true underlying phenomenon and cannot be expected to predict precisely each value of B from a given value of A. In most cases more than one curve can be drawn through a set of observed points, the last and the most important task of a regression analysis is to select the most appropriate equation.

3.6.2.2. Least-Squares Method and Goodness-of-fit Measurement

The most commonly used approach to select a best fitting curve is the *least-squares* method. This procedure determines the best fitting curve as that curve which minimizes the sum of the squared vertical distances from the observed points to the curve. The least-squares method can be described as follows. Let $\hat{B}_i$ denote the estimated response at A_i based on a fitted regression curve. The vertical distance between the observed point (A_i, B_i) and the corresponding point $(A_i, \hat{B}_i)$ on the fitted curve is given by the absolute value $|B_i - \hat{B}_i.|$ The sum of squares of all such distances is then given by

$$SSE = \sum_{i=1}^{n} (B_i - \hat{B}_i)^2, \tag{3.11}$$

where *SSE* is called the *residual sum of squares* [36]. The least-squared curve is obtained by minimizing the *SSE* value.

After having established the least-squared curve, it is important to ask how well does the curve actually fit? The *coefficient of determination* R^2 [36] is a commonly used measure of fit and can be written as

$$R^2 = 1 - \frac{SSE}{SSY}, \tag{3.12}$$

where

$$SSY = \sum_{i=1}^{n} (B_i - \bar{B})^2. \tag{3.13}$$

SSY is often termed the *total sum of squares* [36] since it represents the total squared deviations of the observed B's from the mean $\bar{B}$. The larger the R^2 value the better the fit. In this study, a critical value of 80% was selected meaning any curve of R^2 less than 80% will not be considered as having an acceptable fit.

Once the best fitting equations for each of the four relationships are found, the set of parameters associated with any particular interruption duration can be predicted simply by substituting the duration value into the equations. The cost distribution model at the intermediate duration can be easily recovered using these parameters. Application of this regression technique is illustrated in Chapter 4 using the 1991 interruption cost data.

3.7. Conversion to True Cost Values

Section 3.4 clearly notes that cost distribution analysis must be performed independently on groups of data segregated by interruption duration and by major customer sector. In order to determine the total interruption cost associated with a particular duration for the entire studied service area, all the resulting contributions from each sector must be combined. This process is very similar to the weighting process used to achieve a conventional CCDF (Section 2.3).

Previous sections deal primarily with the study of normally transformed cost. It is important to realize that before a transformed cost is used in further calculations, it must be converted back to its true or original value. The conversion is simply an inverse function of Equation 3.4:

$$x = \begin{cases} [1+\lambda \cdot y]^{1/\lambda} & \text{if } \lambda \neq 0, \\ \log^{-1}(y) & \text{if } \lambda = 0, \end{cases} \tag{3.14}$$

where x and y denote the original cost and transformed cost respectively. Once all the true sector interruption cost values at a given duration are calculated, the total cost for the studied area can be obtained. The sector costs are proportionally weighted by their respective energy consumption (or peak demand), and the weighted values are summed to provide the total cost. This process is discussed further in Chapter 5 where it is used in conjunction with Monte Carlo simulation in an HLI reliability study.

3.8. Summary

The fact that interruption cost estimates collected from customer mail surveys display a significant degree of variability indicated the need for a better representation than one containing only the mean values. From a statistical point of view, the cost variation can be described by a continuous probability distribution. One of the most common continuous distributions is the Normal Distribution and this chapter addresses three of the many merits associated with normally distributed data. Due to these distinct advantages, the outage cost estimates were transformed and fitted to normal distributions. Prior to conducting the transformation, the data were grouped in accordance with the durations for which they were given and with the economic sector which the respondents belong. This grouping scheme permits a duration and customer specific analysis of interruption cost distributions.

In order to clearly illustrate the procedures used to transform the outage costs into normally distributed representations, the two basic steps are repeated as follows:

1. Transforming the original data to symmetrical distributions using the Box and Cox power transformation, and
2. Choosing the symmetrical distribution which best describes a normal distribution using a hypothesis test.

The primary statistics involved in these procedures are the standardized third ($\sqrt{b_1}$) and fourth (b_2) moments.

It should be realized that cost data are not readily available at every possible duration due to the limited number of interruption scenarios allowed in a survey questionnaire. Regression analysis can be used to estimate the probability distribution model for intermediate durations. In this approach, equations describing the relationship between interruption duration and each of the four distribution parameters (P_z, λ, μ and σ^2) are obtained using the least-squares method. The set of parameters associated with an intermediate duration are then calculated from these equations. Once the parameters are known, the intermediate distribution can be easily developed. This chapter also notes that the transformed costs must ultimately be converted back to their true values which are then used in the overall weighting and summing process to generate the total cost of an actual interruption.

It is believed that the probability distribution modeling technique proposed in this chapter recognizes the distributed nature of the monetary interruption losses which is overlooked in the conventional customer damage function approach. The distribution technique when utilized in a reliability worth assessment study should provide a realistic and effective assessment of the losses incurred by electrical users due to power failures. The techniques described in this chapter are applied to the 1991 NSERC cost of interruption data in the following chapters.

4. INTERRUPTION COST MODELING OF THE 1991 NSERC CUSTOMER RESPONSES

4.1. Introduction

Reliability worth assessment is an integral part of an overall HLI study. As noted in the previous chapters, one practical way to assess the benefit or worth of power system reliability is to utilize the customer costs of power interruptions. A basic procedure for investigating these costs, which yields acceptable results, is to survey the electrical users. Data gathered from the survey respondents therefore are the key ingredients in the worth assessment process. The raw cost data, however, must be described by mathematical functions or cost models in such a way that they can be used in practical applications. Chapter 2 and Chapter 3 introduce two different modeling techniques which have been designated as the customer damage function approach and the probability distribution method respectively. The procedures utilized in each technique are given in the respective chapter.

The objective of this chapter is to illustrate each modeling approach using the 1991 NSERC customer interruption cost data. The peak demand normalized cost ($/KW) analysis is presented and the associated results are used to quantify reliability worth in the subsequent chapters.

4.2. Usable Peak Normalized Cost Data

The amount of data available to generate a cost model is dictated by the selection criteria discussed in Section 2.2.3. Both cost models illustrated in the subsequent sections of this chapter employ data selected using the same criteria. Tables 4.1 (a) to (c) give the number of usable 1991 $/KW responses in the residential, commercial and small industrial sectors respectively. The agriculture sector is not included in this analysis as

Table 4.1: Number of Usable 1991 $/KW Responses

(a) Residential : "Preparatory Action" Responses

20 min/month in winter	1351
1 hour/month in winter	1345
4 hour/month in winter	1343
8 hour/year in winter	1328
24 hour/year in winter	1313
48 hour/month in winter	1318
4 hour/month in summer	1338
48 hour/year in summer	1334
24 hr twice/year in summer	1332

Note : Only the first five interruption scenarios are used to generate various cost models.

(b) Commercial : "Worst Case" Responses

SIC Group	Interruption Durations							
	2 sec	1 min	20 min	1 hr	2 hr	4 hr	8 hr	1 day
60	7	9	14	16	16	16	16	16
61	1	1	1	1	1	1	1	1
62	1	1	1	2	2	2	2	2
63	15	15	17	17	17	17	17	17
64	2	2	2	2	2	2	2	2
65	6	7	8	9	9	9	9	9
69	0	0	0	0	0	0	0	0
91	4	4	4	4	4	4	4	4
92	3	4	5	5	5	5	5	5
96	7	7	9	10	10	10	10	10
97	1	0	1	2	2	2	2	2
99	2	3	3	4	4	4	4	4
TOTAL	49	53	65	72	72	72	72	72

Table 4.1, continued

(c) Small Industrial : "Worst Case" Responses

SIC Group	2 sec	1 min	20 min	1 hr	2 hr	4 hr	8 hr	1 day
	Interruption Durations							
04	0	0	1	1	1	1	1	1
06	3	3	4	4	4	4	4	4
07	9	9	12	16	16	16	16	16
08	1	2	2	2	2	2	2	2
09	1	1	1	2	2	2	2	2
10	16	17	21	22	22	22	22	22
11	2	2	2	2	2	2	2	2
15	3	3	4	4	4	4	4	4
16	14	16	17	17	17	17	17	17
17	3	3	3	3	3	3	3	3
18	1	1	1	1	1	1	1	1
19	2	2	2	3	3	3	3	3
24	1	1	2	2	2	2	2	2
25	11	13	13	13	13	13	13	13
26	4	4	4	4	4	4	4	4
27	3	3	3	4	4	4	4	4
28	14	14	15	17	17	17	17	17
29	2	3	4	4	4	4	4	4
30	30	33	36	36	36	36	36	36
31	9	9	10	10	10	10	10	10
32	5	5	6	7	7	7	7	7
33	5	5	7	7	7	7	7	7
35	11	10	11	14	14	14	14	14
36	0	0	0	0	0	0	0	0
37	8	8	7	10	10	10	10	10
39	5	5	6	7	7	7	7	7
TOTAL	163	172	196	212	212	212	212	212

the data were not complete at the time of preparing the results. It should be noted that the interruption cost studies of the commercial and industrial sectors are based on responses given to the worst-case cost question while the residential sector study uses the preparatory action cost responses [12]. The responses in the commercial and industrial sectors have been subdivided into SIC classes.

4.3. Cost Modeling using Customer Damage Functions

Interruption cost data collected from electrical customers are duration specific as the customers are asked to provide their best estimates of monetary losses under several different outage scenarios. Studies of this type provide data which can be conveniently used to create a customer damage function (CDF). The CDF is the conventional approach to modeling interruption costs. The basic idea is to determine the average outage cost as a function of duration. Although average costs are readily available at the questioned durations, the averages at other possible durations must be evaluated by linear interpolation from the damage function. The detailed procedures involved in developing various CDF's are discussed in Chapter 2. The major shortcoming of this representation is that only the average or expected values of the interruption estimates are considered while basic statistical analysis shows that costs under every studied scenario display a very large variation. A CDF is, however, very easy to develop and use despite its inability to recognize the scattered nature of the data. In short, the procedure for developing a CDF is to use the average costs associated with each studied interruption scenario in each major sector to create major sector customer damage functions (SCDF). The SCDF's are then weighted proportionally to their respective energy consumption to form the composite customer damage function (CCDF) for the entire studied area.

4.3.1. Sector Customer Damage Functions

In order to determine the CDF of a sector containing SIC categories, the SIC group costs must first be determined. Table 4.2 gives the aggregated $/KW peak normalized SIC group costs in the commercial sector. The respective sample size is shown in the parentheses. These costs were computed using Equation 2.2. Two similar tables summarizing the results of the residential and small industrial studies are presented in

Table 4.2: Aggregated $/KW Peak Normalized SIC Group Costs and Weighted Total for the Commercial Sector

SIC Group	Interruption Duration 2- sec		1- min		20- min		1- hr	
60	0.3354	(7)	0.1131	(9)	4.4444	(14)	14.3921	(16)
61	0.0000	(1)	0.0000	(1)	4.9020	(1)	14.7059	(1)
62	0.0000	(1)	0.0000	(1)	7.0000	(1)	31.0000	(2)
63	0.0323	(15)	0.4779	(15)	10.8868	(17)	32.9444	(17)
64	0.0000	(2)	7.0506	(2)	14.6449	(2)	23.3945	(2)
65	0.1603	(6)	0.5671	(7)	0.5754	(8)	4.6237	(9)
69	0.0000	(0)	0.0000	(0)	0.0000	(0)	0.0000	(0)
91	0.0382	(4)	0.0382	(4)	0.3725	(4)	1.0315	(4)
92	0.0000	(3)	7.2067	(4)	9.8699	(5)	15.4664	(5)
96	0.1176	(7)	0.1312	(7)	0.3079	(9)	18.5081	(10)
97	7.5758	(1)	0.0000	(0)	0.1212	(1)	0.3060	(2)
99	0.0000	(2)	0.2359	(3)	4.7185	(3)	2.7331	(4)
Weighted Total	0.4752	(49)	1.7790	(53)	5.2580	(65)	13.7910	(72)

SIC Group	Interruption Duration 2- hr		4- hr		8- hr		1- day	
60	49.5543	(16)	74.9343	(16)	152.6789	(16)	195.0251	(16)
61	39.2157	(1)	98.0392	(1)	196.0784	(1)	245.0980	(1)
62	67.0000	(2)	154.0000	(2)	291.7500	(2)	334.0000	(2)
63	79.5083	(17)	148.2050	(17)	281.6807	(17)	335.1512	(17)
64	43.1532	(2)	252.4635	(2)	322.1203	(2)	356.0992	(2)
65	6.8306	(9)	14.3662	(9)	24.9882	(9)	79.3125	(9)
69	0.0000	(0)	0.0000	(0)	0.0000	(0)	0.0000	(0)
91	1.4136	(4)	2.6982	(4)	8.3333	(4)	14.6132	(4)
92	23.8583	(5)	34.9433	(5)	47.6474	(5)	61.8323	(5)
96	19.5049	(10)	22.6185	(10)	31.5109	(10)	36.3884	(10)
97	0.6120	(2)	1.6902	(2)	3.3804	(2)	6.8775	(2)
99	5.4661	(4)	11.2098	(4)	22.6970	(4)	28.2464	(4)
Weighted Total	30.1775	(72)	64.9784	(72)	109.7704	(72)	136.0388	(72)

Appendix D. The overall sector costs are then calculated by weighting and summing the SIC group costs. The weighting factors used in this study were adopted from the relative amount of energy consumed by users in the respective SIC groups in the British Columbia Hydro (BCH) service area. These factors are listed in Table 4.3 and were used to weight the cost data given in Table 4.2 to produce the overall sector costs shown in the row labelled "Weighted Total".

Each sector cost and its respective interruption duration can be considered as a pair of observations where the duration is the independent variable and the cost is the dependent variable. All such pairs when plotted on a paper visually portray the commercial SCDF as shown in Figure 4.1. In this context, both axes have been transformed to a logarithmic scale in order to cover the wide range of values. The construction of the small industrial SCDF is the same as the commercial one and this result is also depicted in Figure 4.1. The formation of the residential SCDF, however, is slightly different as no SIC grouping is available for this sector. The sector costs are obtained directly by averaging all the individual costs. This result is also shown in Figure 4.1. All three SCDF's are summerized in Table 4.4.

4.3.2. Composite Customer Damage Function

In order to build the CCDF for the entire service area from the SCDF's shown in Table 4.4, the sector costs have to be proportionally weighted by their respective sector energy consumptions within the area. An example service area sector composition modified from the RBTS was used for this particular study. Since only three major economic sectors are considered, the percentages of energy consumption in the non-studied sectors were distributed and added to the sectors under study. The resulting customer mix is given in Table 4.5 and was used to weight the sector costs given in Table 4.4. The resultant CCDF is tabulated in the last column of Table 4.4 and displayed in Figure 4.2.

Table 4.3: SIC Major Groups Weighting Factors

SIC	Description	Weighting Factor
	COMMERCIAL SECTOR	
60	Food, Beverage and Drug Retail	0.206117
61	Shoe, Apparel, Fabric and Yarn Stores	0.023738
62	Household Furniture, Appliances and Furnishings	0.022031
63	Automotive Vehicles, Parts and Accessories	0.096064
64	General Merchandising	0.079489
65	Other Retail Sales	0.065418
69	Non-Store Retail	0.000234
91	Hotels and Accommodations	0.156040
92	Food and Beverage Services	0.149989
96	Entertainment Services	0.098637
97	Personal Services	0.049491
99	Other Stores	0.052752
	SMALL INDUSTRIAL SECTOR	
04	Logging	0.003413
05	Forestry Services	0.003617
06	Mining Industries	0.186795
07	Crude Petroleum	0.008201
08	Quarry and Sand Pit	0.001593
09	Services Incidental to Mineral Extraction	0.000267
10	Food Industries	0.020365
11	Beverage Industries	0.003371
12	Tobacco Industries	0.000003
15	Rubber Products	0.000257
16	Plastic Products	0.003811
17	Leather and Allied Products	0.000062
18	Primary Textile	0.000108
19	Textile Products	0.000387
24	Clothing Industries	0.000762
25	Wood Industries	0.118137
26	Furniture and Fixture Industries	0.001097
27	Paper and Allied Products	0.496420
28	Printing and Publishing and Allied Products	0.004598
29	Primary Metal Industries	0.003411
30	Fabricated Metal Products	0.010617
31	Machinery Industries	0.002671
32	Transportation Equipment	0.004204
33	Electrical and Electronic Products	0.003111
35	Non-Metallic Mineral Products	0.019163
36	Refined Petroleum and Coal Products	0.020395
37	Chemical and Chemical Products	0.079250
39	Other Mining Industries	0.003914

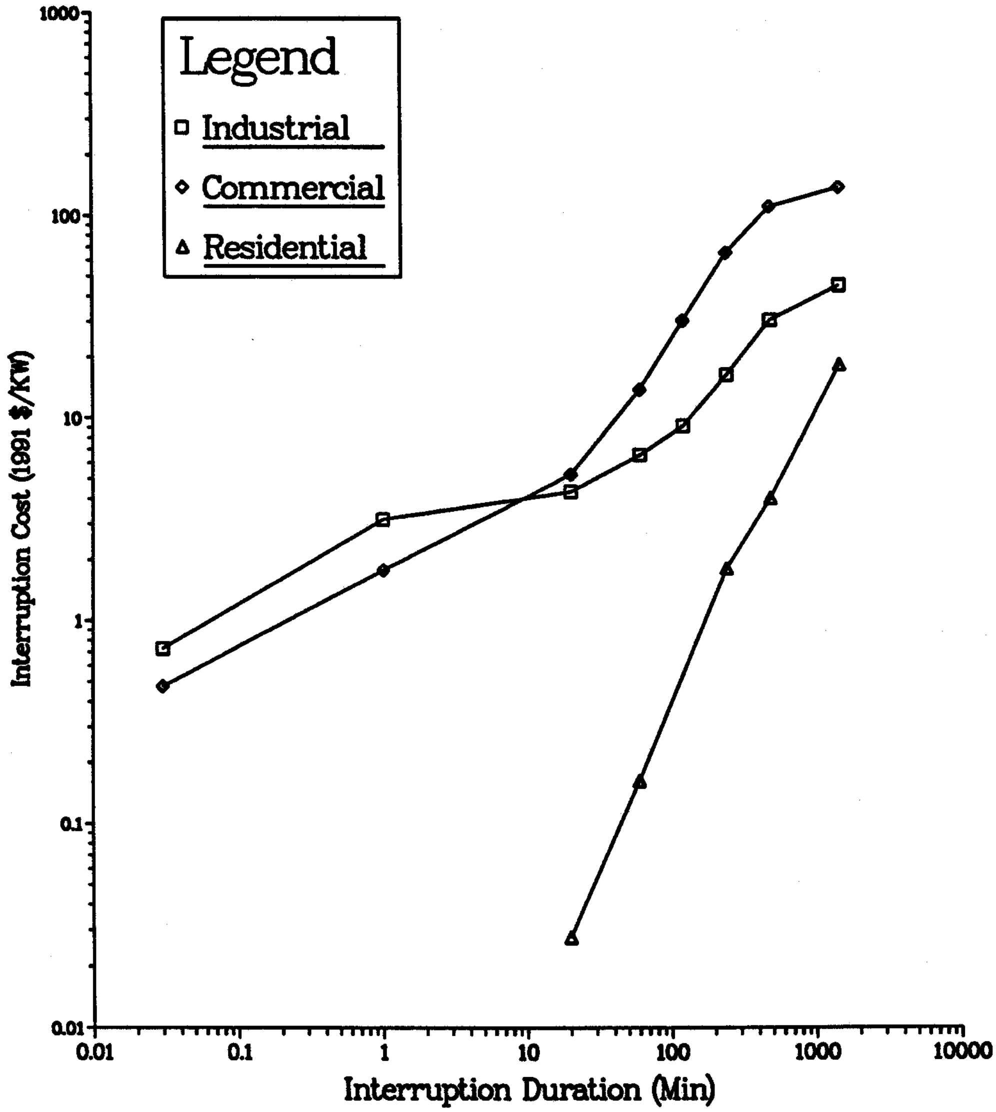


Figure 4.1: SCDFs Generated from the 1991 $/KW Cost Data

Table 4.4: SCDF's and CCDF of an Example Service Area : 1991 $/KW Cost of Interruption Data

Interruption Duration	Industrial	User Sector Commercial	Residential	Weighted Total
2 sec	0.7291	0.4752	0.0000*	0.2614
1 min	3.1663	1.7790	0.0002*	1.0443
20 min	4.3217	5.2580	0.0278	2.3455
1 hr	6.5508	13.7910	0.1626	5.4819
2 hr	9.1189	30.1775	0.5428**	11.3123
4 hr	16.2679	64.9784	1.8126	24.1940
8 hr	30.3254	109.7704	4.0006	41.7957
1 day	44.7320	136.0388	18.2491	59.9140

* value obtained using extrapolation
** value obtained using interpolation

Table 4.5: Customer Mix of an Example Service Area

User Sector	Energy(%)
Industrial	15.00
Commercial	32.00
Residential	53.00

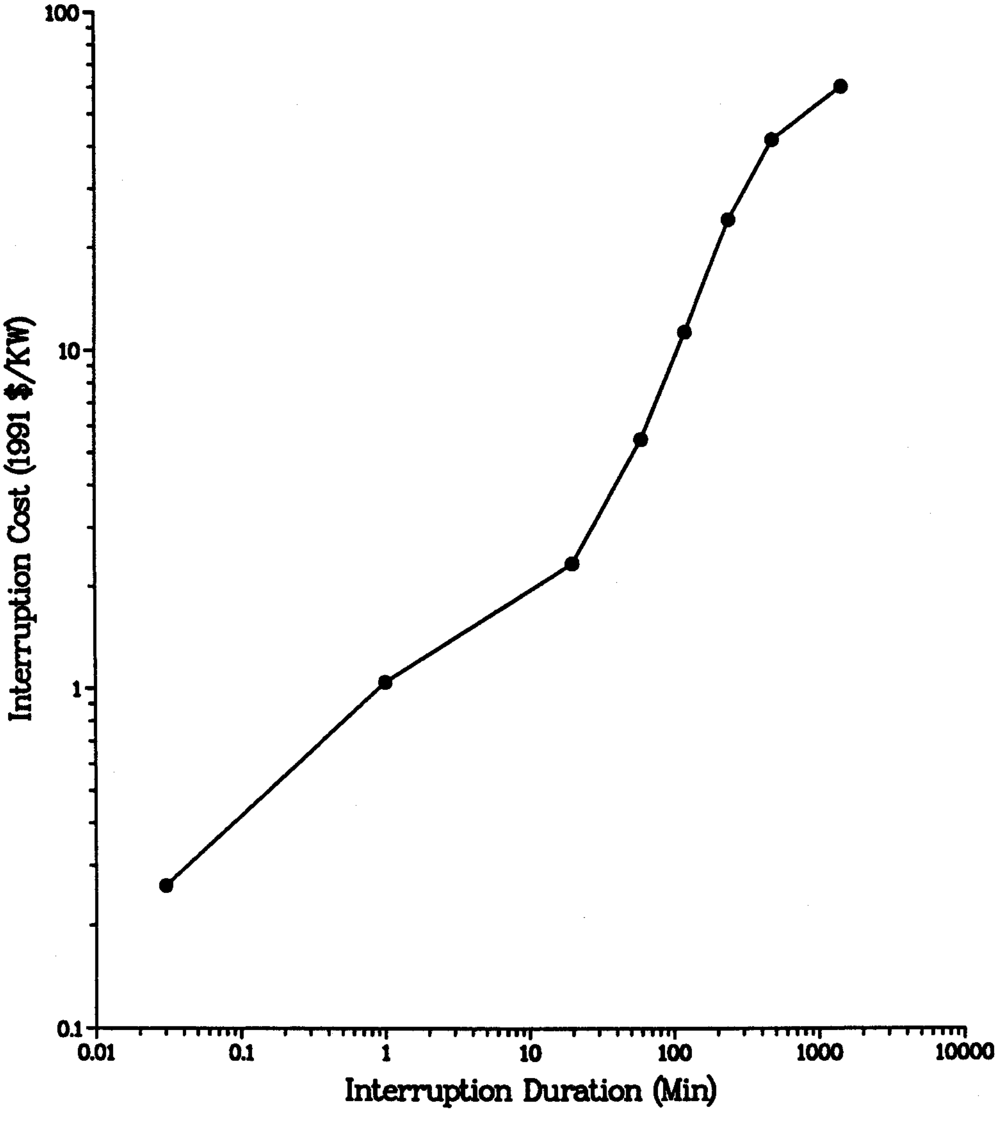


Figure 4.2: CCDF of an Example Service Area : 1991 $/KW Cost

The customer damage functions illustrated in Figures 4.1 and 4.2 utilize piece-wise linearly increasing relationships where a segment between any two successive studied durations is described by a linear or straight line equation. Once the formulation of each segment is known, the average interruption cost at any possible intermediate duration can be easily determined using linear interpolation between the two adjacent studied durations.

4.4. Cost Modeling using Probability Distributions

The probability distribution method of modeling the interruption cost data is a new concept introduced in this research work. The basic idea is to describe the data in such a way that their behaviour can be recognized and incorporated in a wide range of applications. A major assumption in this technique is that the customer survey sampling procedure provides truly random interruption cost data. The pattern that gives the relative frequencies associated with all possible values of the outage cost is known as its probability distribution. Cost of interruption modeling using the distribution approach is more complicated and difficult to use compared to the conventional CDF. It does, however, represent the entire data rather than consolidating these data into a single value. Every single estimate given by a respondent contributes to defining the shape of the distribution. The likelihood of a particular cost value being used to represent the outage loss associated with a given duration therefore depends on its relative frequency.

4.4.1. Data Transformation Results

The basic objective of the distribution approach is to transform a given group of cost data into a distribution which will satisfy the normality conditions. The transformation process is illustrated using the 1991 commercial $/KW responses. The number of usable data in each data category is given in Table 4.1. It can be seen from this table, as an example, that the probability distribution of the commercial $/KW responses for a 1-hr interruption is estimated using 72 data points.

As discussed in Section 3.5.3.2, an iterative process is employed to perform the transformation. For each category of data, the process generates as many sets of symmetrically transformed data as required using Equation 3.4 until the preselected $\sqrt{b_1}$ value of 0.001 or less is achieved. Values of kurtosis b_2, Z_1 and Z_2 approximations are also calculated at each iteration. Following each iteration, a hypothesis test is utilized to determine whether or not the resulting transformed data satisfy the normality conditions. As discussed in Section 3.5.5, if a Type I Error of 0.05 is deemed acceptable and a two-sided test is performed, both Z approximations should have values less than 1.96 in order to accept the null hypothesis that the transformed data are normally distributed. Only those transformations which fulfil the hypothesis test are retained after the iteration process terminates. The set of transformed data which has the smallest $\sqrt{b_1}$ value is selected as this distribution pattern has the best fit to a normal curve. Table 4.6 gives the results of this transformation and selection process for the commercial sector.

Table 4.6: Commercial Sector Moment Test Results (with "0" data) : 1991 $/KW Cost

Interruption Duration	λ	Z_1	Z_2	$\sqrt{b_1}$	b_2	Normality
2 sec	0.2838	-0.0002	3.4979	-0.0001	4.5511	NO
1 min	-0.0229	-0.0003	3.4224	-0.0001	4.1520	NO
20 min	-0.1605	-0.0020	0.3168	-0.0008	-0.0345	YES
1 hr	-0.0252	0.0004	1.5826	0.0001	0.8321	YES
2 hr	-0.0455	-0.0010	1.1023	-0.0004	0.4500	YES
4 hr	-0.0389	0.0002	0.8820	0.0001	0.3003	YES
8 hr	-0.0102	0.0026	1.2594	0.0010	0.5660	YES
1 day	0.0277	-0.0012	1.6300	-0.0005	0.8745	YES

It can be seen in Table 4.6 that all eight groups of transformed costs satisfy the condition of skewness at a Type I Error of 0.05 as all the Z_1 values are less than 1.96. The Z_2 values, on the other hand, are too high in some cases. It is evident therefore, for any given group of data, the set of symmetrical transformations produced by the iteration process does not necessarily contain an acceptable normal distribution. The last column of Table 4.6 indicates the conclusions of the analysis. The transformed commercial $/KW interruption costs at all durations other than the 2-sec and 1-min scenarios satisfy both the third and fourth moment tests. Histograms and normal probability plots have been used to portray the results.

4.4.1.1. Histogram

The easiest and most convenient way to display a distribution is by using a *histogram* in which a set of data is represented by a simple bar graph. The observed values are grouped into appropriate intervals and the occurrence of data in each interval is tabulated. If a variable is normally distributed, the shape or contour of the histogram should be very similar to that of a normal distribution.

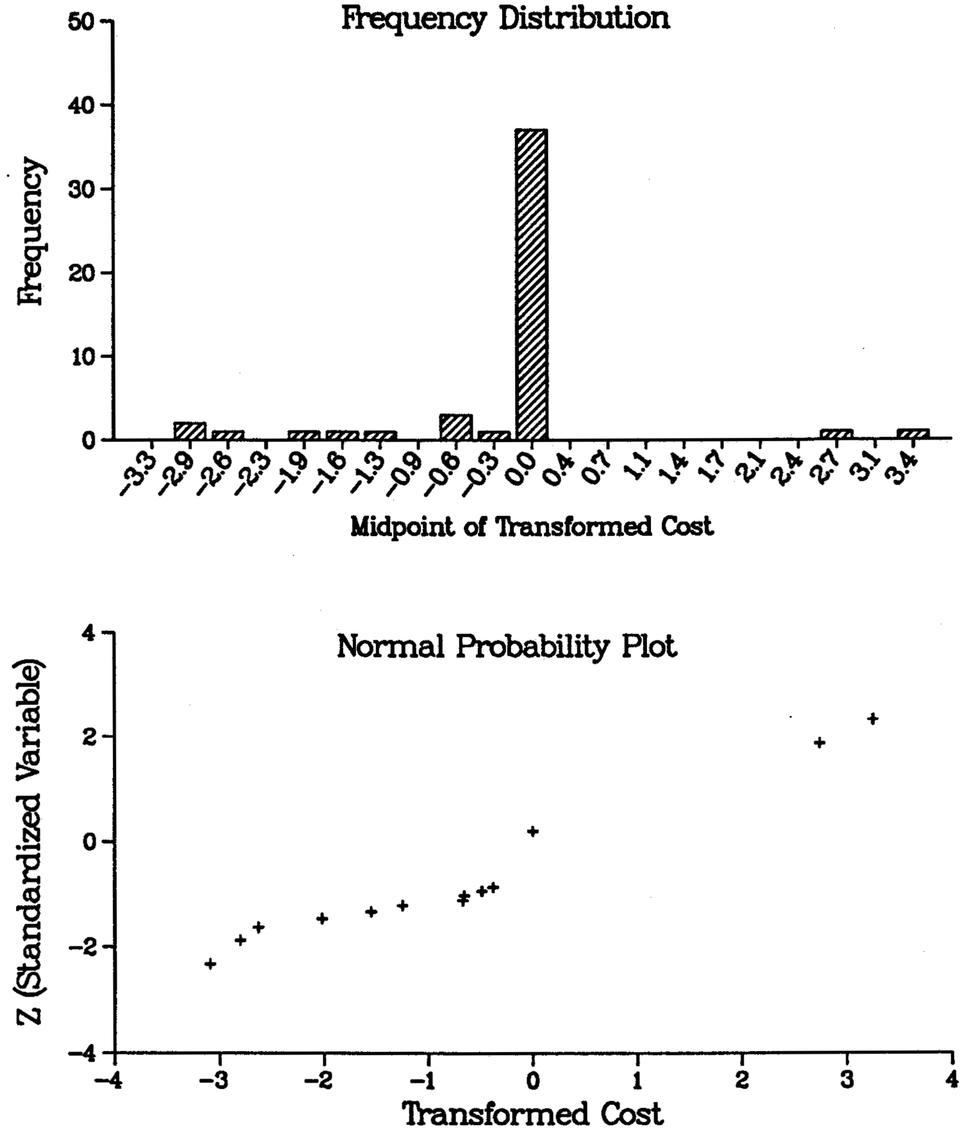
The transformed 1991 commercial $/KW cost histograms are presented in the upper regions of Figures 4.3 (a) to (h). It can be seen that the histograms for the 2-sec and 1-min interruptions bear little relationship to the expected bell-shape associated with a normal distribution. In comparison, histograms of longer interruptions show more association with a normal curve. This visual judgement, however, is very subjective and sometimes inaccurate. One difficulty is associated with selecting an appropriate interval width. If the intervals are too small relative to the available sample size, the resulting histogram is ragged, and in contrast, if they are too large, then some information is lost [32].

4.4.1.2. Normal Probability Plot

A normal probability plot or a normal plot is a special tool for presenting the observed values in order to examine how well the data fit a normal distribution. In this representation, each transformed cost is paired with its corresponding value from the normal distribution. (The corresponding value from the normal distribution is based on the number of responses in the sample and the rank order of the sample [31].) If the data are normally distributed, the points will fall, more or less, on a straight line. The lower regions in Figures 4.3 (a) to (h) show the normal plots of the transformed $/KW data at various interruption scenarios. It is easier to visually judge the departures from normality in a normal plot compared to a histogram since the normal plot is a straight line when the data are normally distributed. The set of data points in the 2-Sec and 1-Min normal plots do not suggest a straight line which supports the observations drawn from the corresponding histograms that the transformed cost at these durations are not normally distributed. The normal plots for longer durations indicate quite reasonable straight line relationships.

Although normal probability plots are very helpful in observing the departures from normality, there are a number of shortcomings in this technique, especially for small samples. The normal plots for samples smaller than 25 can show substantial variation and nonlinearity even if the underlying distribution is normal [31]. In general, graphs should only be used for informal preliminary judgements and as adjuncts to more formal techniques. Numerical values should always be used to quantify the information and evidence contained in the graphs and should serve to verify the inference suggested in the visual display. Numerical techniques are extremely important in order to avoid making possible spurious conclusions from graphical analysis.

Moment tests which numerically quantify the normality of a variable are considered to be the most powerful goodness-of-fit testing tool in terms of sensitivity to skewed and non-normal kurtosis distributions [37]. These tests were employed in this research work. The results are shown in Table 4.6. The same information is repeated in Figures 4.3 (a) to (h) for the sake of completeness. In each figure, the total number of responses n, and the number of original zero-valued data n_z are also indicated. The value of n for any particular duration is the sum of the frequencies encountered in the respective histogram. The value of n_z is the absolute number of zero-valued data reported. This value although very close to the "0" bar frequency, is not necessary the same. Data which falls within the "0" midpoint interval can be composed of non-transformed zero-valued data, and also transformed low original costs. The latter category usually occupies a very small percentage of the whole.



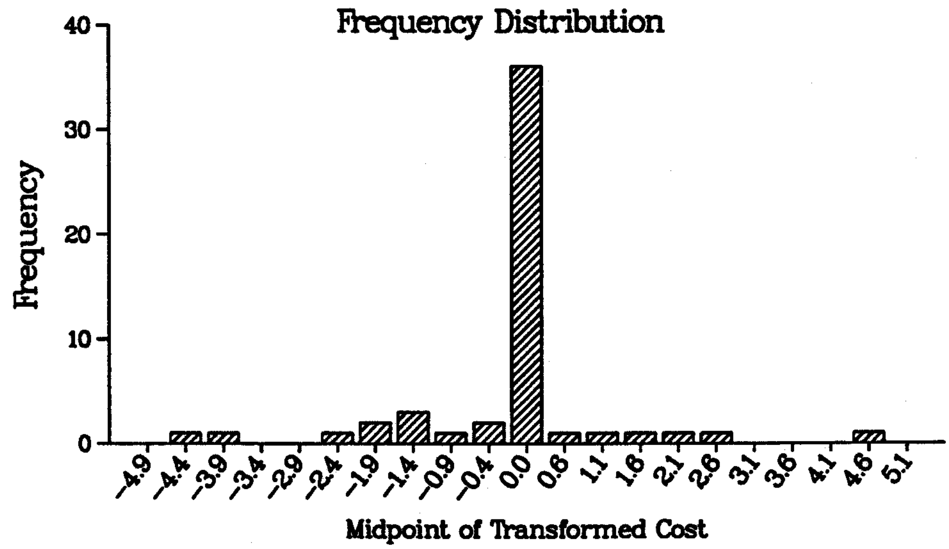
$n = 49$ $\quad$ $n_z = 37$

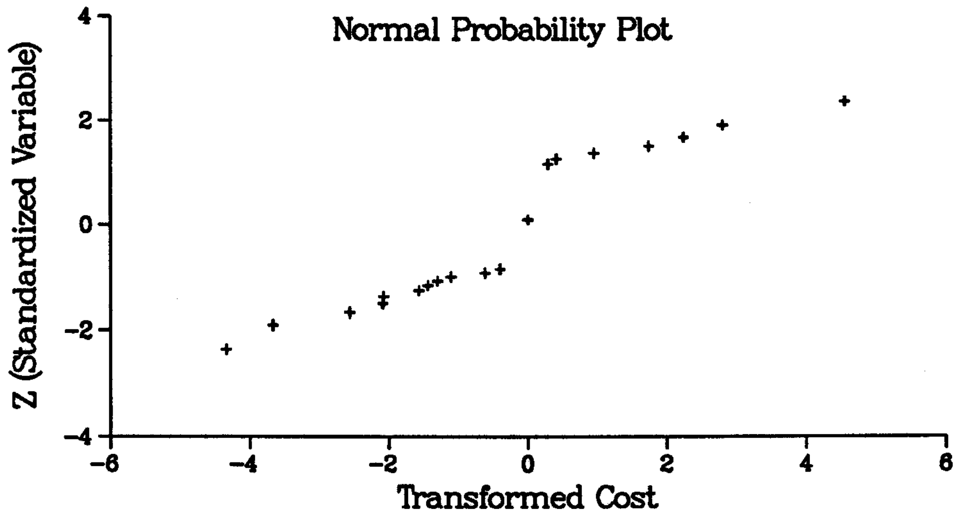
$Z_1 = -0.0002$ $\quad$ $Z_2 = 3.3979$

$\sqrt{b_1} = -0.0001$ $\quad$ $b_2 = 4.5511$

$\lambda = 0.2838$

Figure 4.3: Results of $/KW Cost Transformation for the Commercial Sector (with "0" data) (a) 2-Sec $/KW





$n = 53$

$n_z = 35$

$Z_1 = -0.0003$

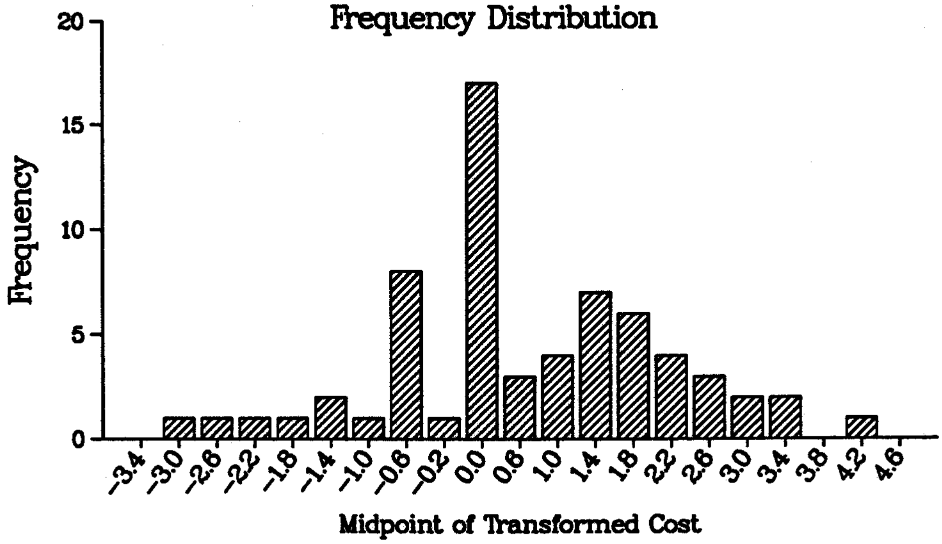
$Z_2 = -3.4224$

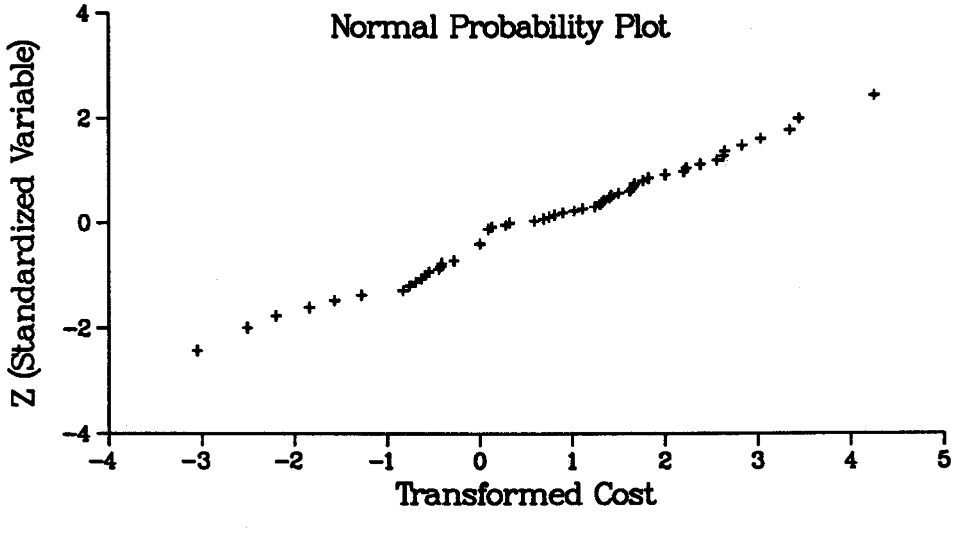
$\sqrt{b_1} = -0.0001$

$b_2 = -4.1520$

$\lambda = -0.0229$

Figure 4.3, continued (b) 1-Min $/KW





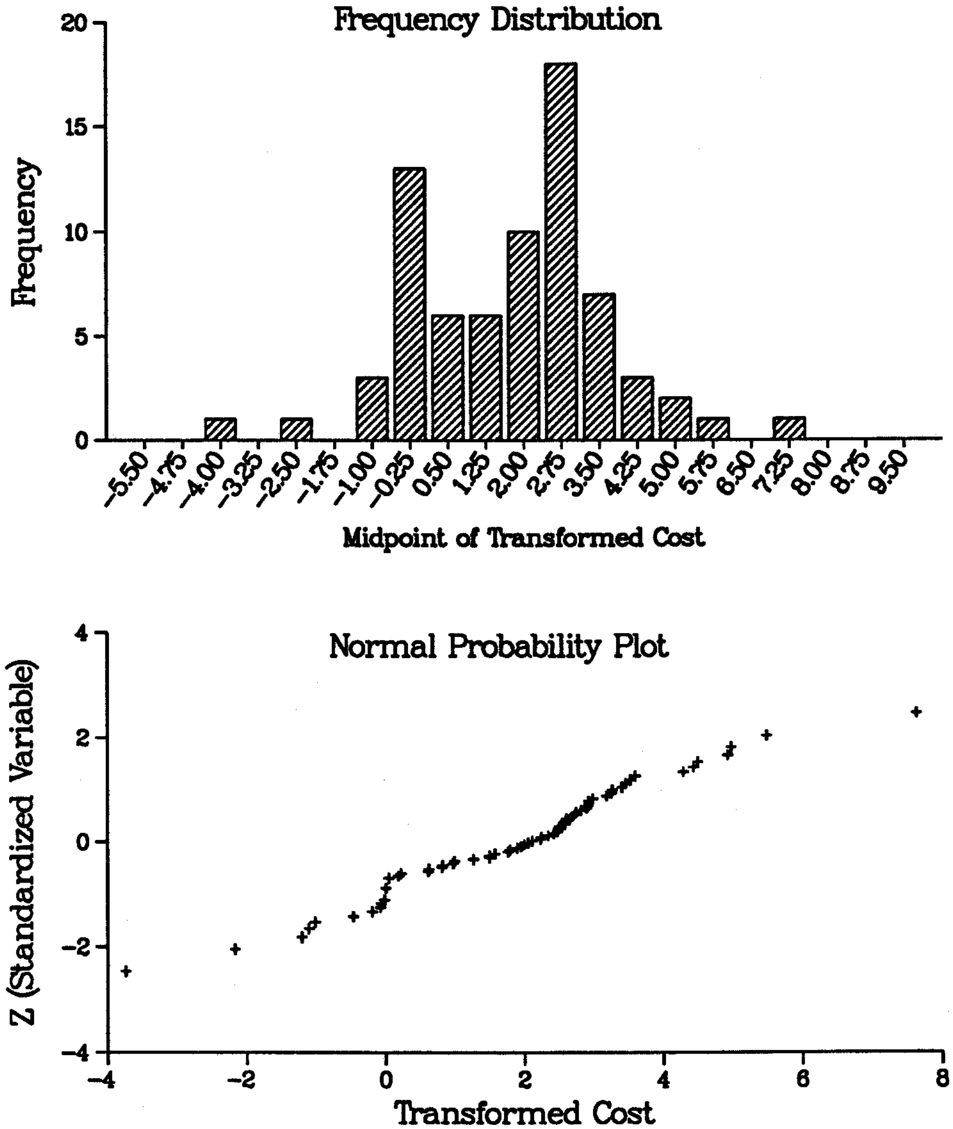
$n = 65$ $n_z = 13$

$Z_1 = -0.0002$ $Z_2 = 0.3168$

$\sqrt{b_1} = -0.0008$ $b_2 = -0.0345$

$\lambda = -0.1605$

Figure 4.3, continued (c) 20-Min $/KW



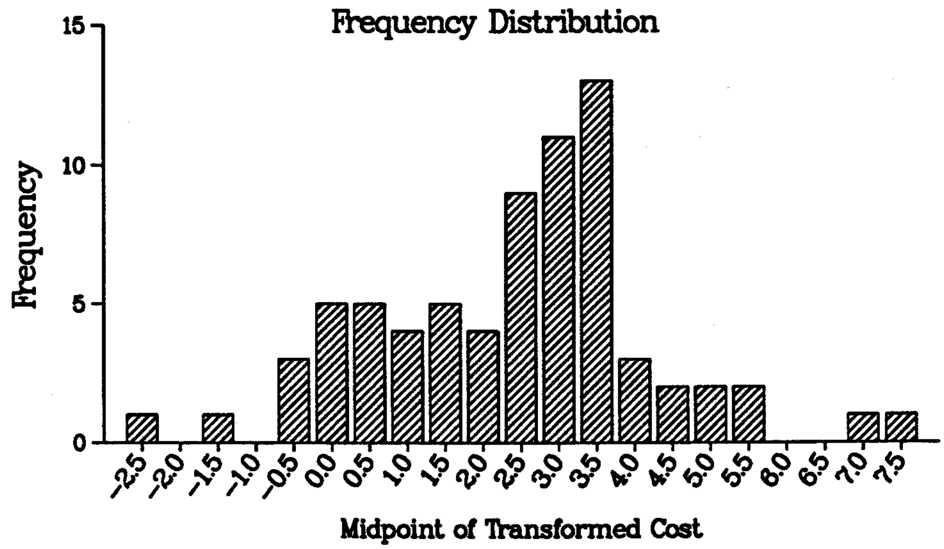
$n = 72$ $\qquad$ $n_z = 7$

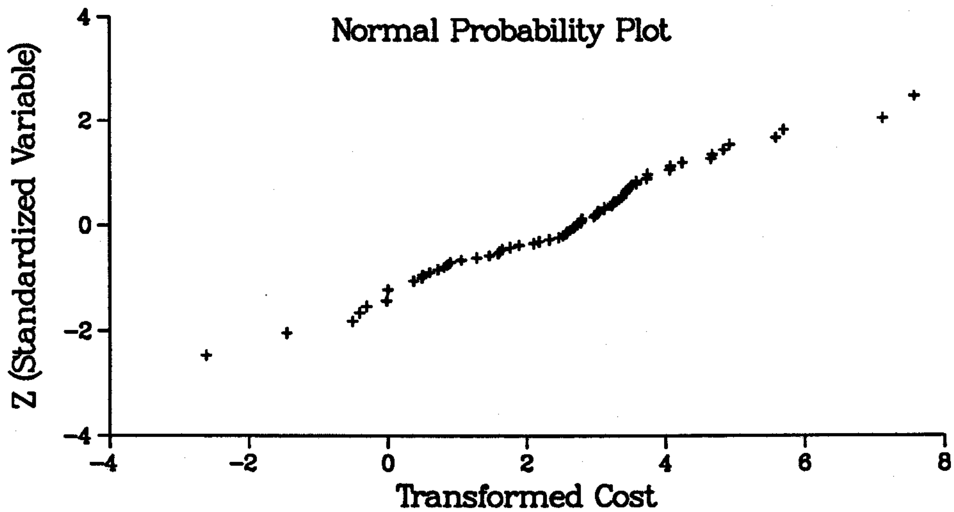
$Z_1 = 0.0004$ $\qquad$ $Z_2 = 1.5826$

$\sqrt{b_1} = 0.0001$ $\qquad$ $b_2 = 0.8321$

$\lambda = -0.0252$

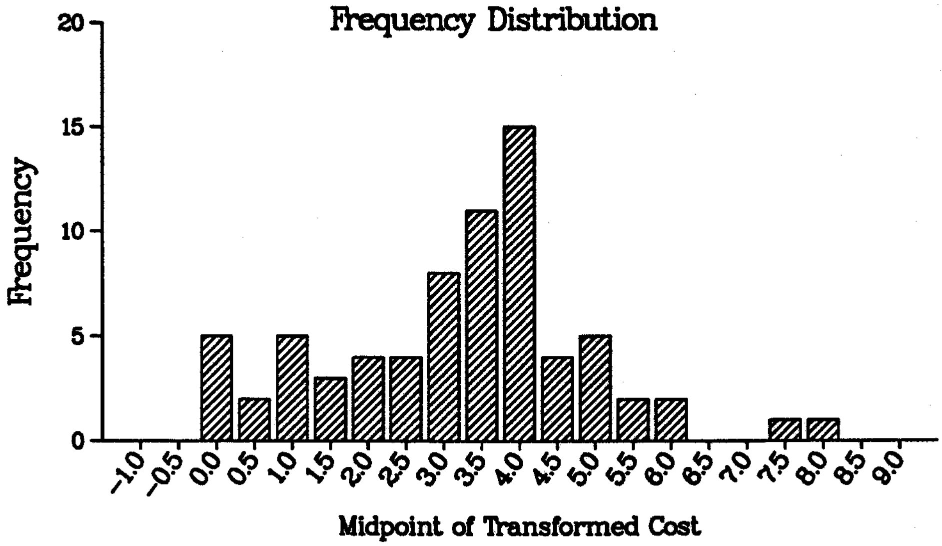
Figure 4.3, continued (d) 1-Hr $/KW

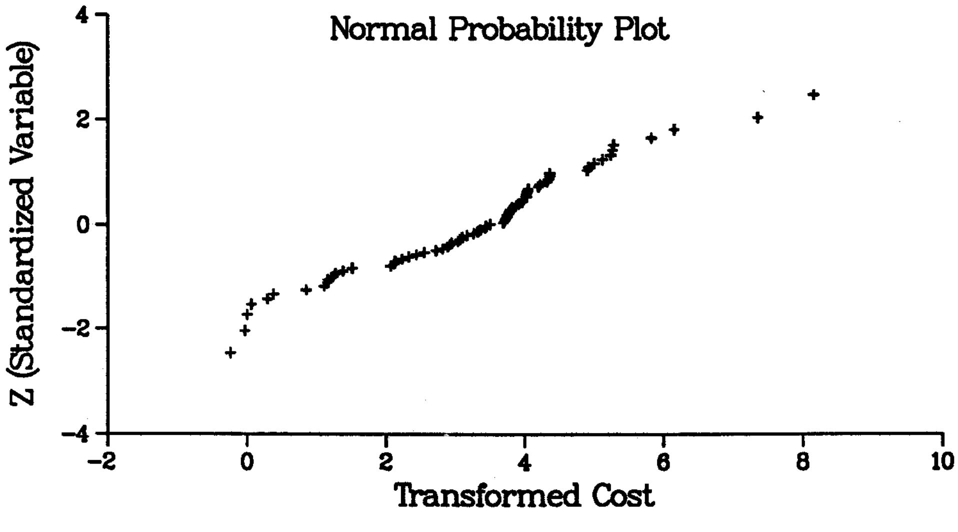




$n = 72$	$n_z = 4$
$Z_1 = -0.0010$	$Z_2 = 1.1023$
$\sqrt{b_1} = -0.0004$	$b_2 = 0.4500$
$\lambda = -0.0455$	

Figure 4.3, continued (e) 2-Hr \$/KW





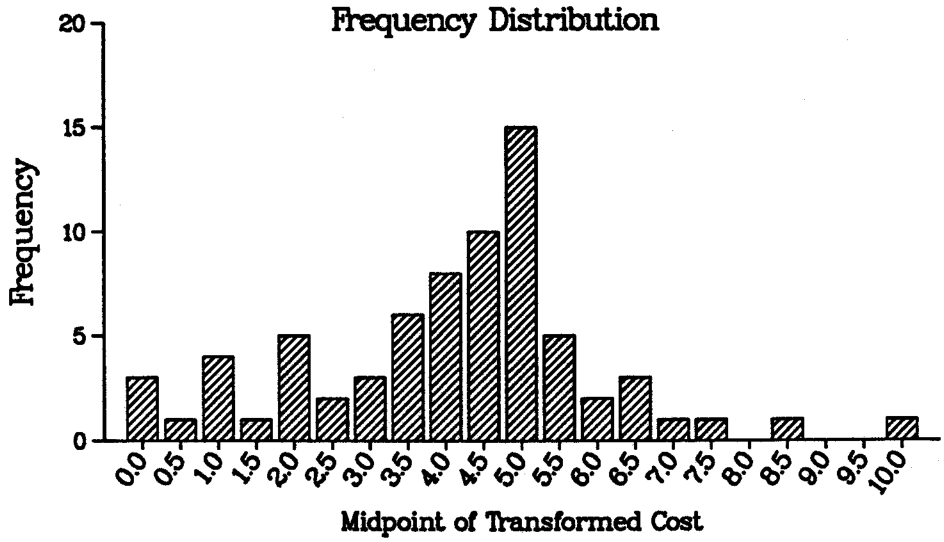
$n = 72$ $n_z = 2$

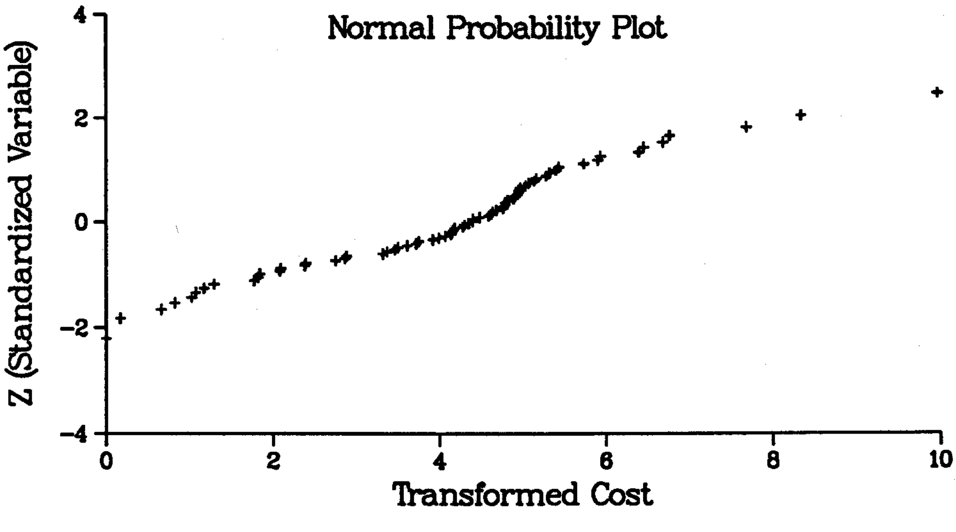
$Z_1 = 0.0002$ $Z_2 = 0.8820$

$\sqrt{b_1} = 0.0001$ $b_2 = 0.3003$

$\lambda = -0.0389$

Figure 4.3, continued (f) 4-Hr $/KW





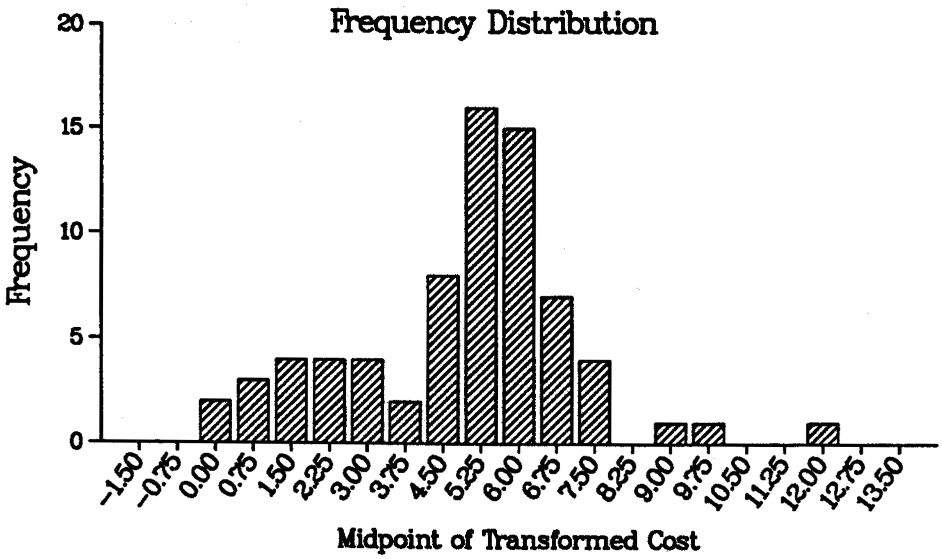
$n = 72$ $n_z = 2$

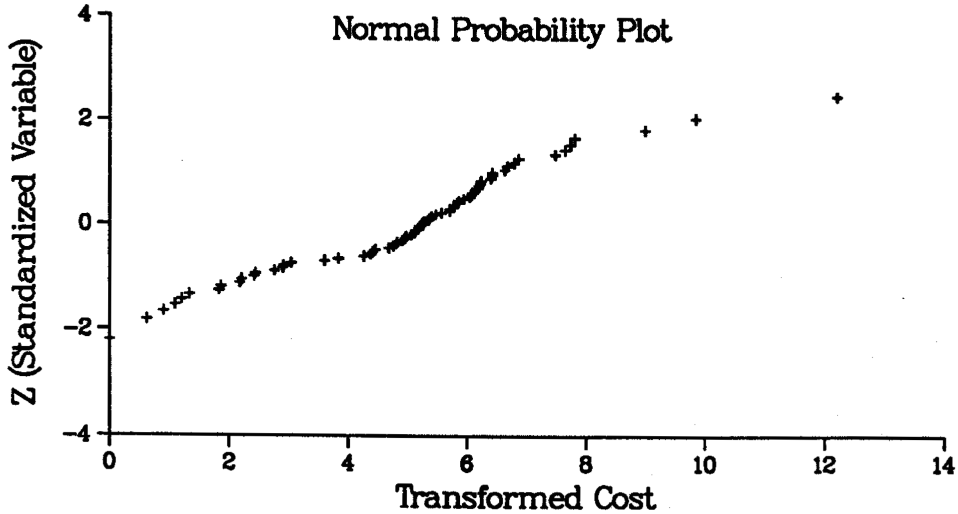
$Z_1 = 0.0026$ $Z_2 = 1.2594$

$\sqrt{b_1} = 0.0010$ $b_2 = 0.5660$

$\lambda = -0.0102$

Figure 4.3, continued (g) 8-Hr $/KW





$n = 72$ $n_z = 2$

$Z_1 = -0.0012$ $Z_2 = 1.6300$

$\sqrt{b_1} = -0.0005$ $b_2 = 0.8745$

$\lambda = 0.0277$

Figure 4.3, continued (h) 1-Day $/KW

4.4.1.3. Transformation Without Including Zero-Valued Data

The transformation technique proposed in Chapter 3 is not effective in some cases in terms of satisfying the moment test. After investigating several possible reasons for this, it was found that the technique fails simply because inappropriate data are used in the transformation. Section 3.5.2 notes that the Box and Cox power transformation works only with continuous variables. The distribution analysis presented in the previous sections, therefore, assumes that the outage costs are continuously distributed. There is, however, strong evidence which shows that the costs may come from a somewhat different distribution phenomenon.

A pictorial display often reveals something that the analyst may not have anticipated. This is the case when considering the histograms in Figures 4.3 (a) and (b). These pictorial representations suggest that the outage costs at the 2-sec and 1-min durations may have been sampled from both continuous and discrete distributions. This observation is based on the fact that the "0" bar in each histogram has an extremely high frequency compared to other intervals. For example, in the 2-sec case, there are 37 occurrences in the zero midpoint interval which is 75.5% of the total data points. This "0" bar is not in harmony with and is isolated from the other transformed values. The values in this interval may have experienced a distribution of their own apart from the other continuous values. A close examination of these "0" bar values revealed that they are composed of primarily original zero-valued data. If these zero values are extracted from the whole, the shape of the resulting histogram may perhaps more closely follow a normal curve. Based upon this, the transformation procedure was repeated with the "0" data excluded. The new results are given in Figure 4.4. The success of this approach are quantified by the moment test statistics summerized in Table 4.7. The 2-sec and 1-min groups now successfully satisfy the tests. Their kurtosis values are improved by at least a factor of six. The other data categories previously satisfied the tests also show better normal distributions. This analysis supports the proposition that the transformation technique works well on continuous data. It failed in some of the previous analyses because the outage cost data came from a mixture of continuous and discrete distributions. All subsequent analyses described in this thesis have used the new method to perform the Box and Cox transformation. Two similar studies conducted in the residential and industrial sectors are summarized in Tables E.1 (a) and (b) respectively.

Table 4.7: Commercial Sector Moment Tests Results (without "0" data) : 1991 $/KW Cost

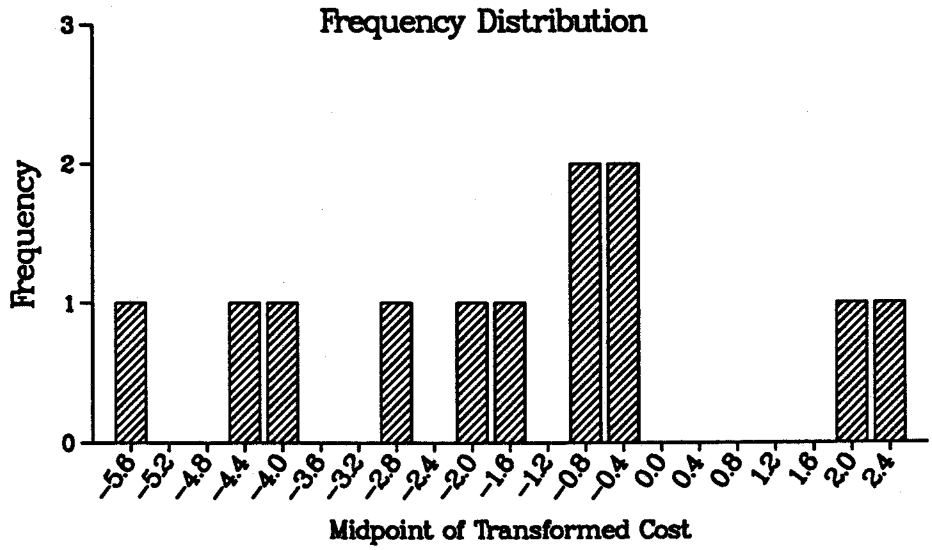
Interruption Duration	λ	Z_1	Z_2	$\sqrt{b_1}$	b_2	Normality
2 sec	0.0781	-0.0006	-0.1394	-0.0004	-0.6992	YES
1 min	-0.1043	-0.0011	-0.0230	-0.0007	-0.4863	YES
20 min	-0.0992	-0.0022	-0.1502	-0.0010	-0.2996	YES
1 hr	-0.0023	0.0010	1.9589	0.0004	1.2902	YES
2 hr	-0.0350	-0.0012	1.6094	-0.0004	0.8727	YES
4 hr	-0.0461	-0.0017	1.1101	-0.0007	0.4586	YES
8 hr	-0.0277	0.0009	1.2810	0.0004	0.5872	YES
1 day	0.0063	0.0007	1.5810	0.0003	0.8386	YES

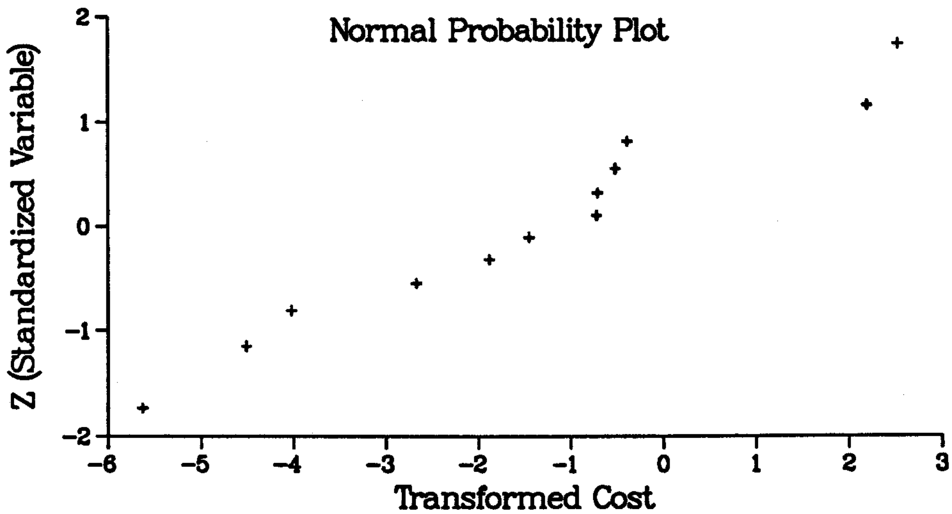
Although zero-valued data are not utilized to build the cost distribution model, these data represent a special group of respondents who believe that power failures have absolutely no monetary impact on their functions and activities. These "0" data, if neglected completely in the worth assessment process, will cause the resulting value of unserved energy to be overestimated. The quantity of "0" data therefore must be known and retained so that the data can be used at a later stage in the analysis. The frequency of "0" data is shown following in terms of the proportion of these data in the total number of usable data. This proportion is denoted as P_z and can be obtained by dividing the value of n_z by n, given in Figure 4.3.

4.4.2. Cost Distribution at Intermediate Durations

Section 3.6 notes that the interruption cost data given by any selected category of users at a given duration can be described by the following parameters:

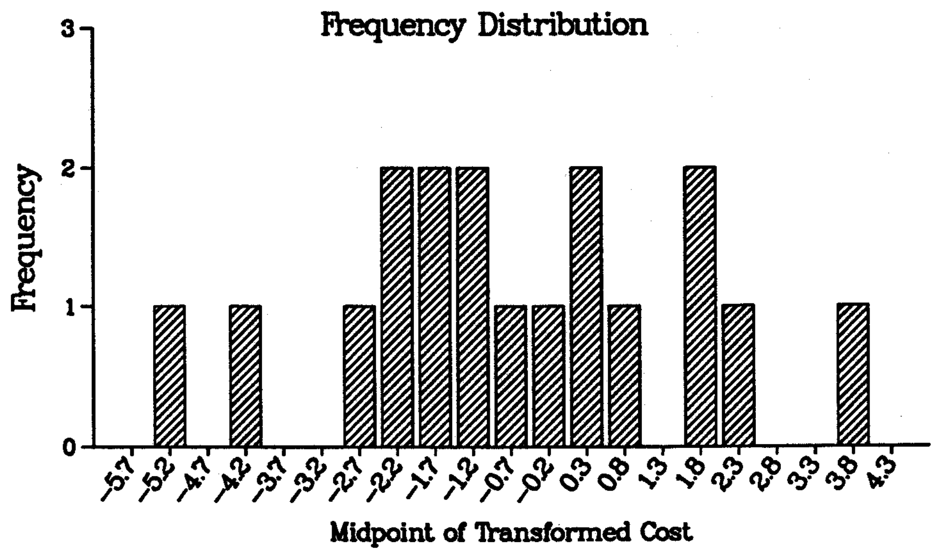
1. Proportion of zero-valued data, P_z,
2. Power transformation factor, λ,
3. Mean of the normal-transformed cost distribution, μ and
4. Variance of the normal-transformed cost distribution, σ^2.

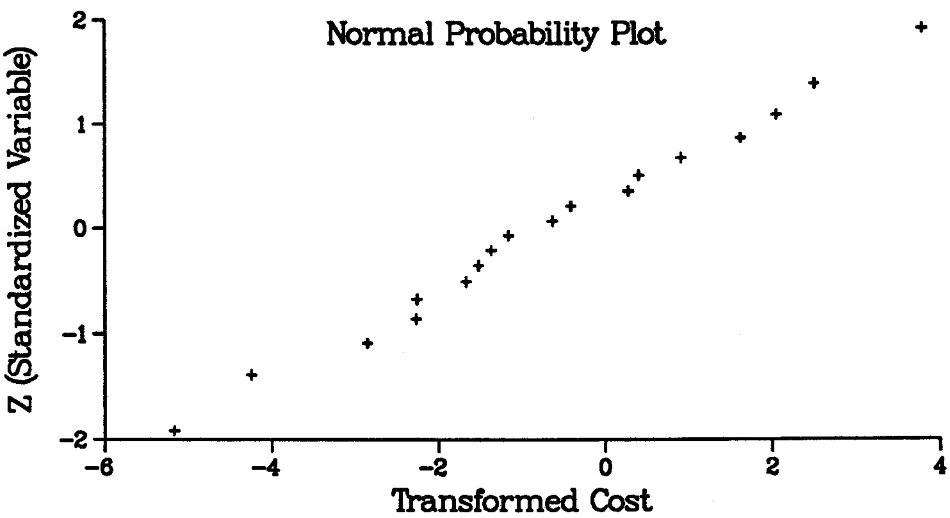




$n = 12$	$\lambda = 0.0781$
$Z_1 = -0.0006$	$Z_2 = -0.1394$
$\sqrt{b_1} = -0.0004$	$b_2 = -0.6992$

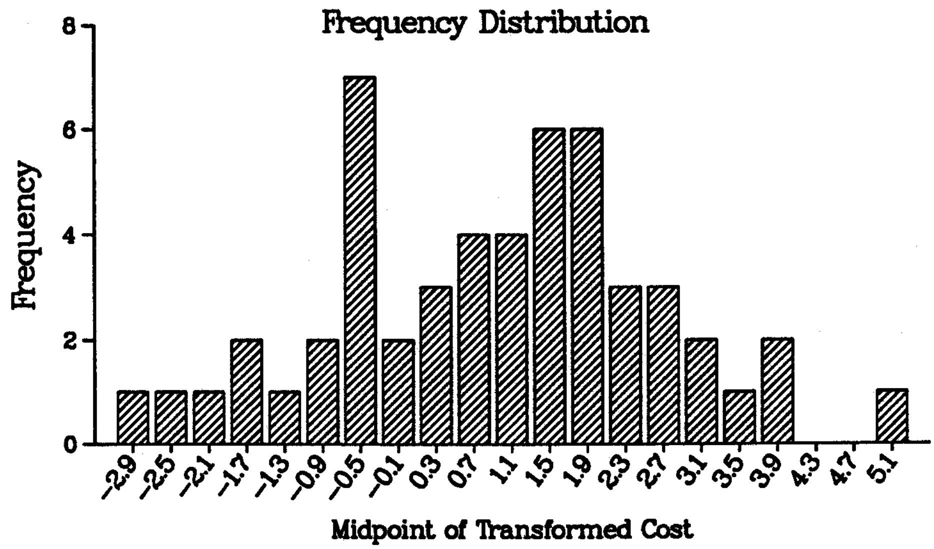
Figure 4.4: Results of $/KW Cost Transformation for the Commercial Sector (without "0" data) (a) 2-Sec $/KW

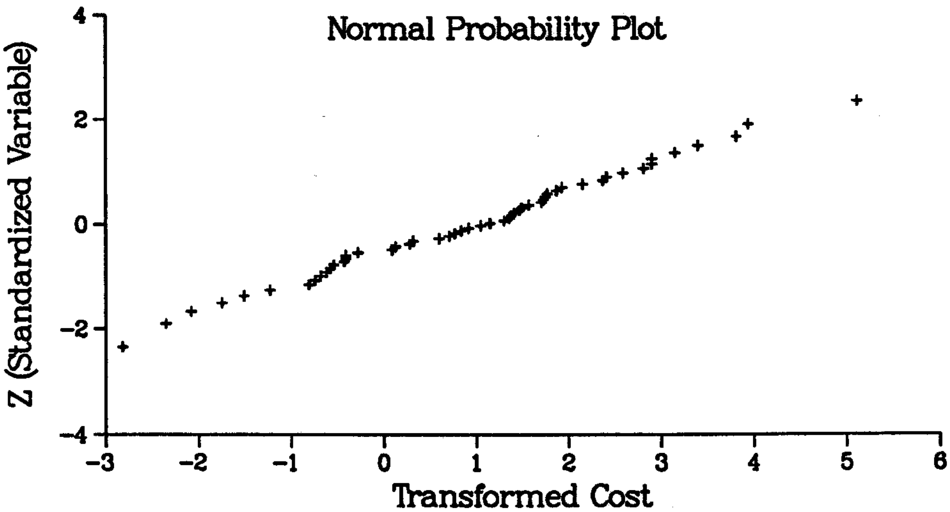




$n = 18$	$\lambda = -0.1043$
$Z_1 = -0.0011$	$Z_2 = -0.0230$
$\sqrt{b_1} = -0.0007$	$b_2 = -0.4863$

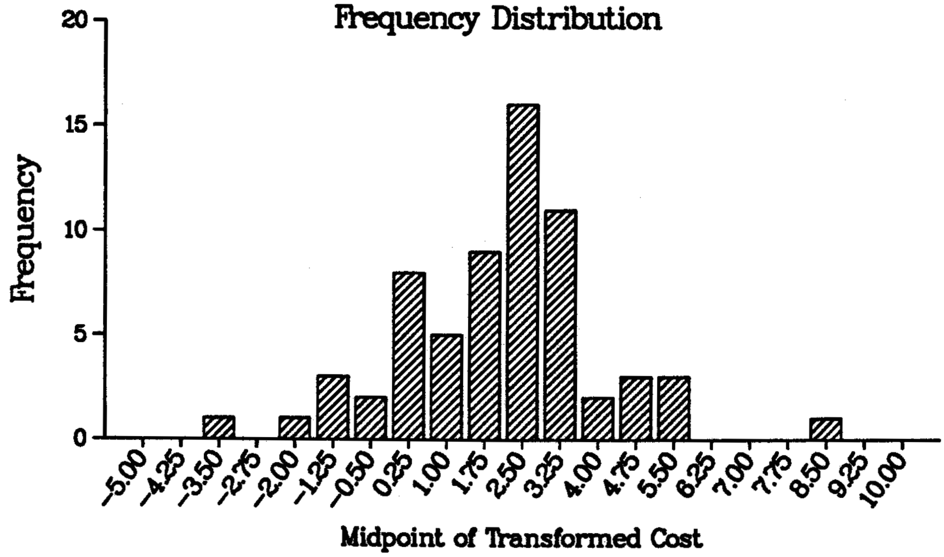
Figure 4.4, continued (b) 1-Min $/KW

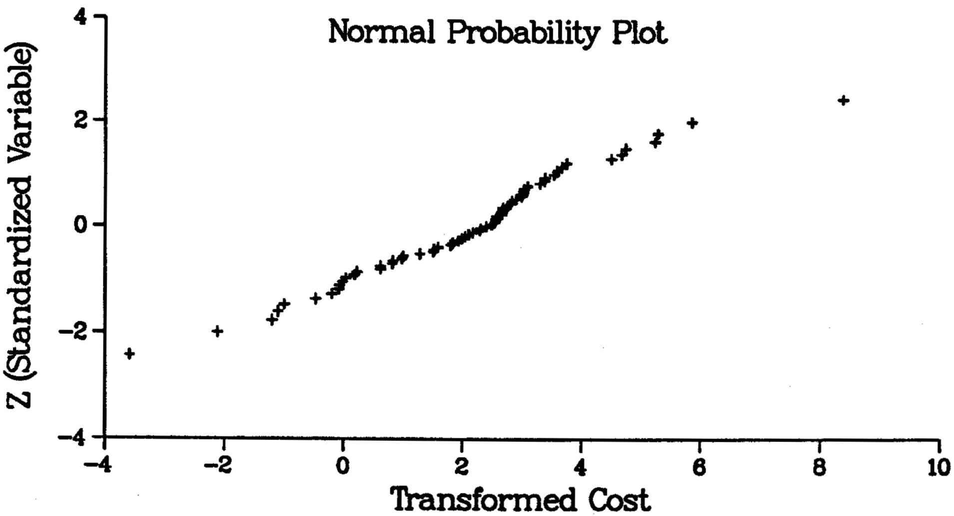




n = 52	$\lambda = -0.0992$
$Z_1 = 0.0022$	$Z_2 = -0.1502$
$\sqrt{b_1} = -0.0010$	$b_2 = -0.2996$

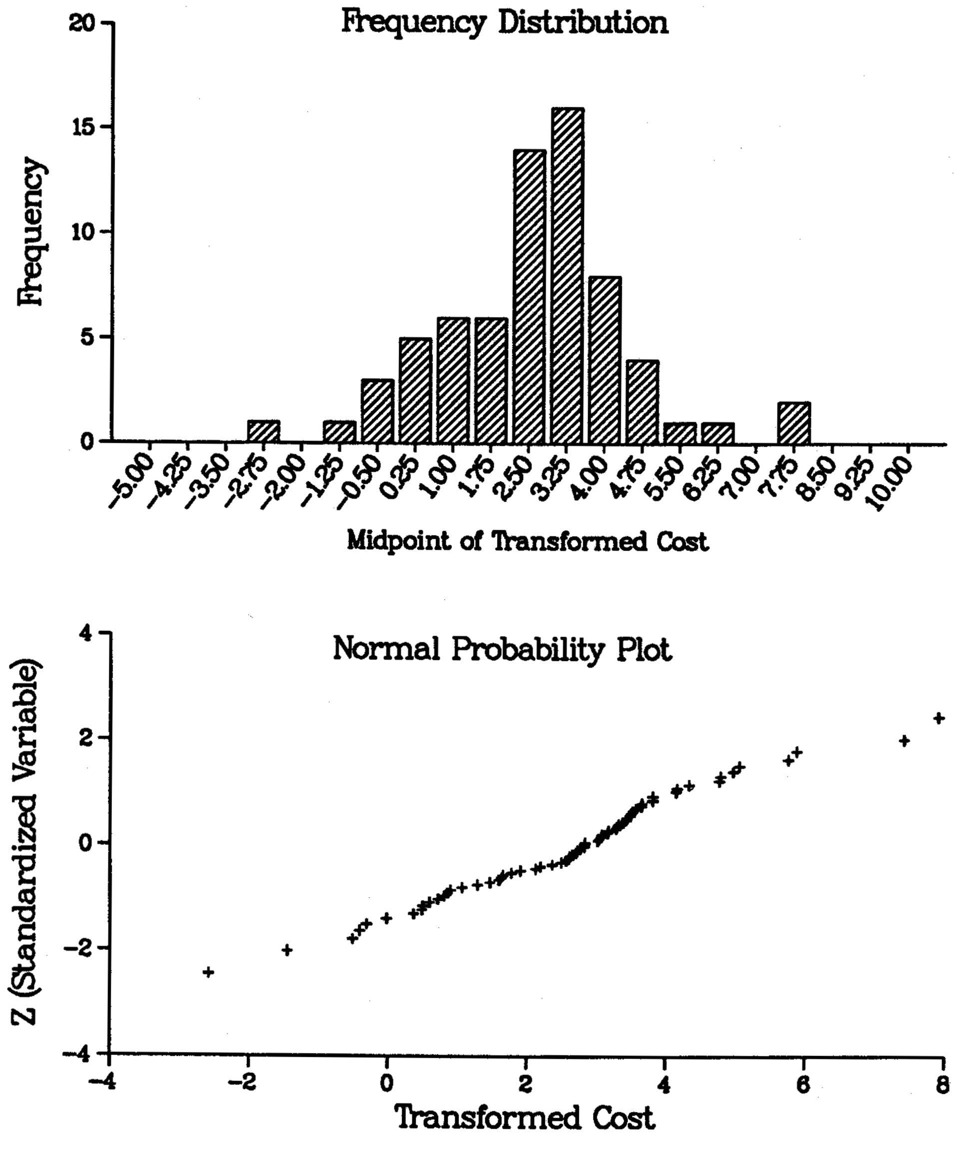
Figure 4.4, continued (c) 20-Min $/KW





$n = 65$	$\lambda = -0.0023$
$Z_1 = 0.0010$	$Z_2 = 1.9589$
$\sqrt{b_1} = 0.0004$	$b_2 = 1.2901$

Figure 4.4, continued (d) 1-Hr $/KW

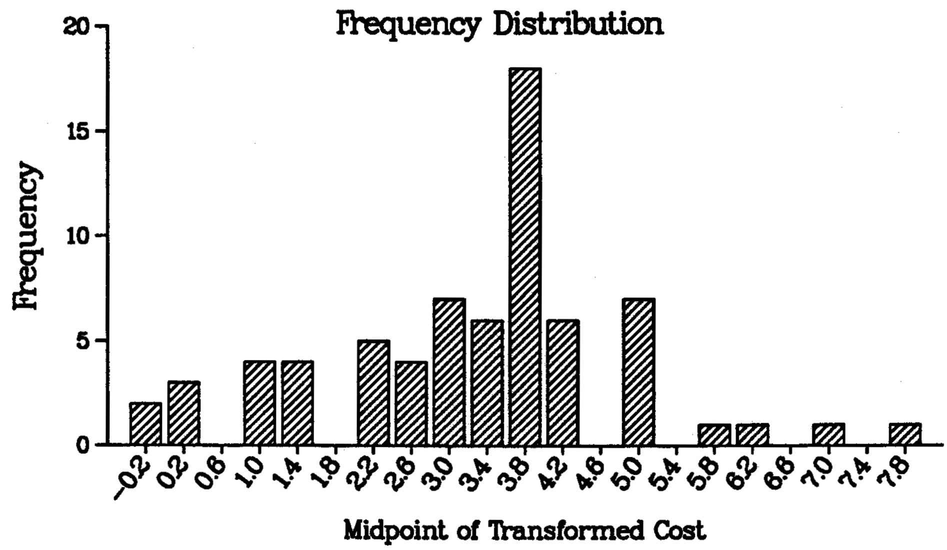


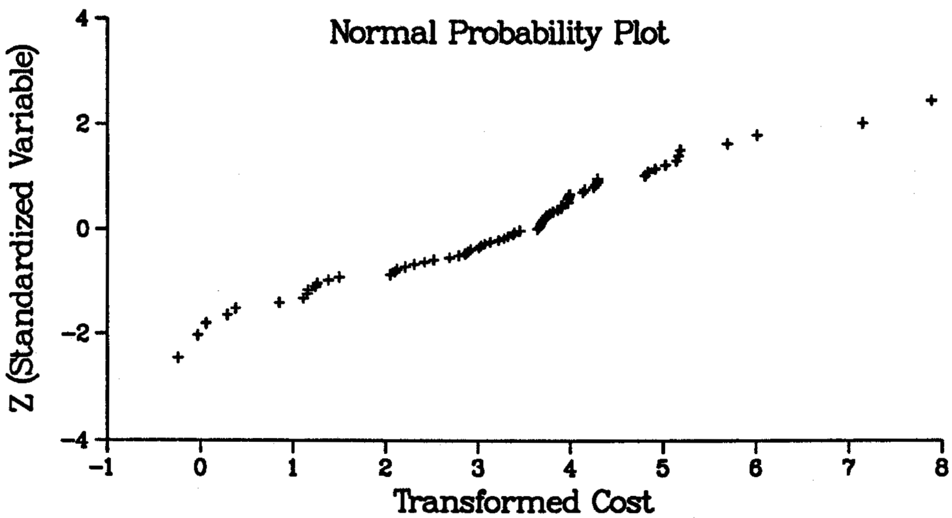
$n = 68$ $\lambda = -0.0350$

$Z_1 = -0.0012$ $Z_2 = 1.6094$

$\sqrt{b_1} = -0.0004$ $b_2 = 0.8727$

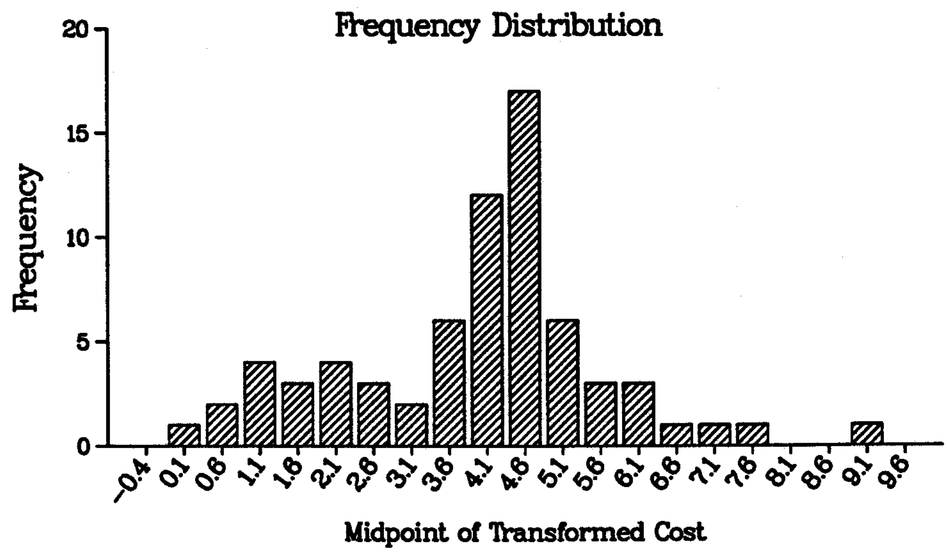
Figure 4.4, continued (e) 2-Hr $/KW

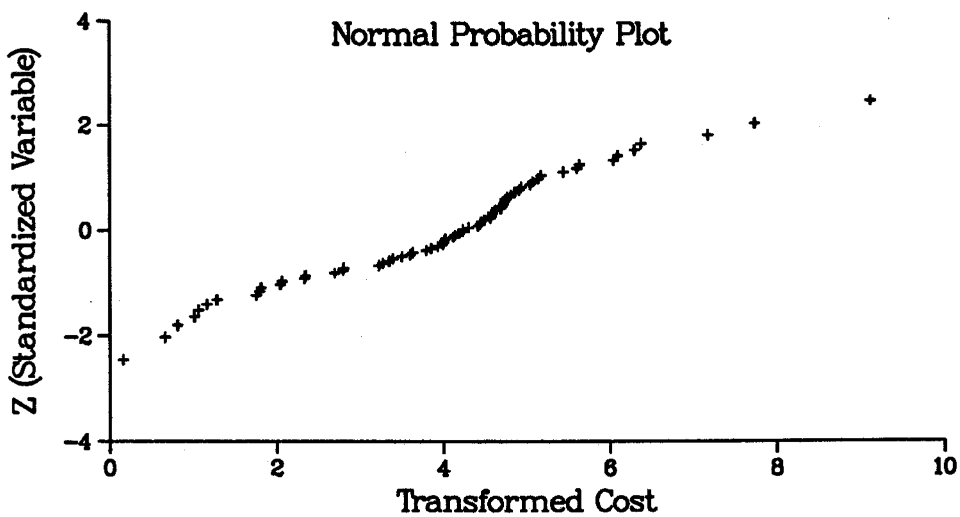




$n = 70$	$\lambda = -0.0461$
$Z_1 = -0.0017$	$Z_2 = 1.1101$
$\sqrt{b_1} = -0.0007$	$b_2 = 0.4586$

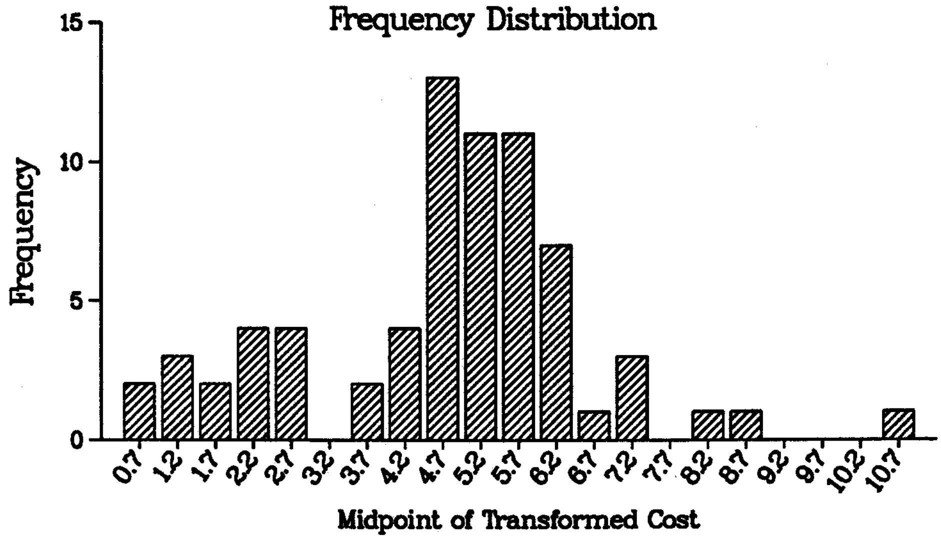
Figure 4.4, continued (f) 4-Hr $/KW

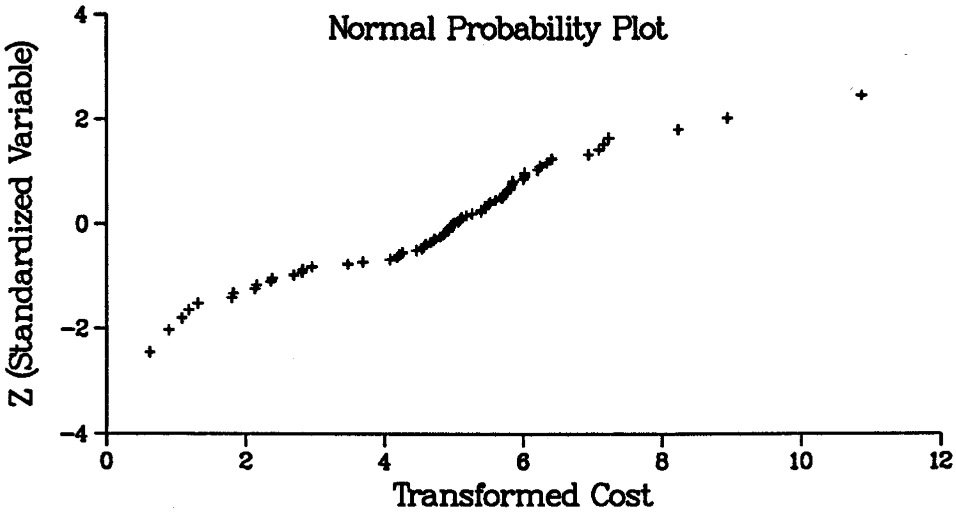




n = 70	$\lambda = -0.0277$
$Z_1 = 0.0009$	$Z_2 = 1.2810$
$\sqrt{b_1} = 0.0004$	$b_2 = 0.5872$

Figure 4.4, continued (g) 8-Hr $/KW





$n = 70$	$\lambda = 0.0063$
$Z_1 = 0.0007$	$Z_2 = 1.5810$
$\sqrt{b_1} = 0.0003$	$b_2 = 0.8386$

Figure 4.4, continued (h) 1-Day $/KW

The value of P_z for any studied scenario can be easily assessed from the original data groups. The other three parameters can be obtained through the normality transformation process. All these parameters cannot be obtained in the same manner for non-studied durations since cost data are not collected at these points. Regression analysis was used to predict the intermediate parameters using known values at the studied durations. The basic approach is to determine mathematical formulations to describe the known parameter values as a function of duration. Given the best fitting equation for each of the four relationships, a particular parameter at a non-studied duration can be predicted by substituting the duration value into the respective equation. The probability distribution for a group of transformed costs at an intermediate duration can be easily recovered when the parameters are known. Table 4.8 summarizes the values of these parameters for the commercial sector. Two similar tables are given in Appendix E for the residential and small industrial sectors.

Table 4.8: Commercial Cost Distribution Parameters : 1991 $/KW Cost

Duration	λ	Mean (μ)	Variance (σ^2)	P_Z
2 sec	0.0781	-1.4827	6.1106	0.7551
1 min	-0.1043	-0.6686	5.4441	0.6604
20 min	-0.0992	0.9219	2.9045	0.2000
1 hr	-0.0023	2.0813	3.8423	0.0972
2 hr	-0.0350	2.6400	3.4916	0.0556
4 hr	-0.0461	3.2577	2.5495	0.0278
8 hr	-0.0277	3.9965	2.8251	0.0278
1 day	0.0063	4.7753	3.5965	0.0278

The simplest way to portray the possible models which fit the relationships tabulated in Table 4.8 is by using scatter diagrams as presented in Figure 4.5. It can be seen from these scatter plots that the relationship between the questioned outage duration and each of the four parameters is non-linear. None of the relationships shown in the given forms can be approximated by a simple straight line equation and therefore it is necessary to try to model these data with suitable curves. Before applying any particular curve, a close examination of the observed points can sometimes help to resolve

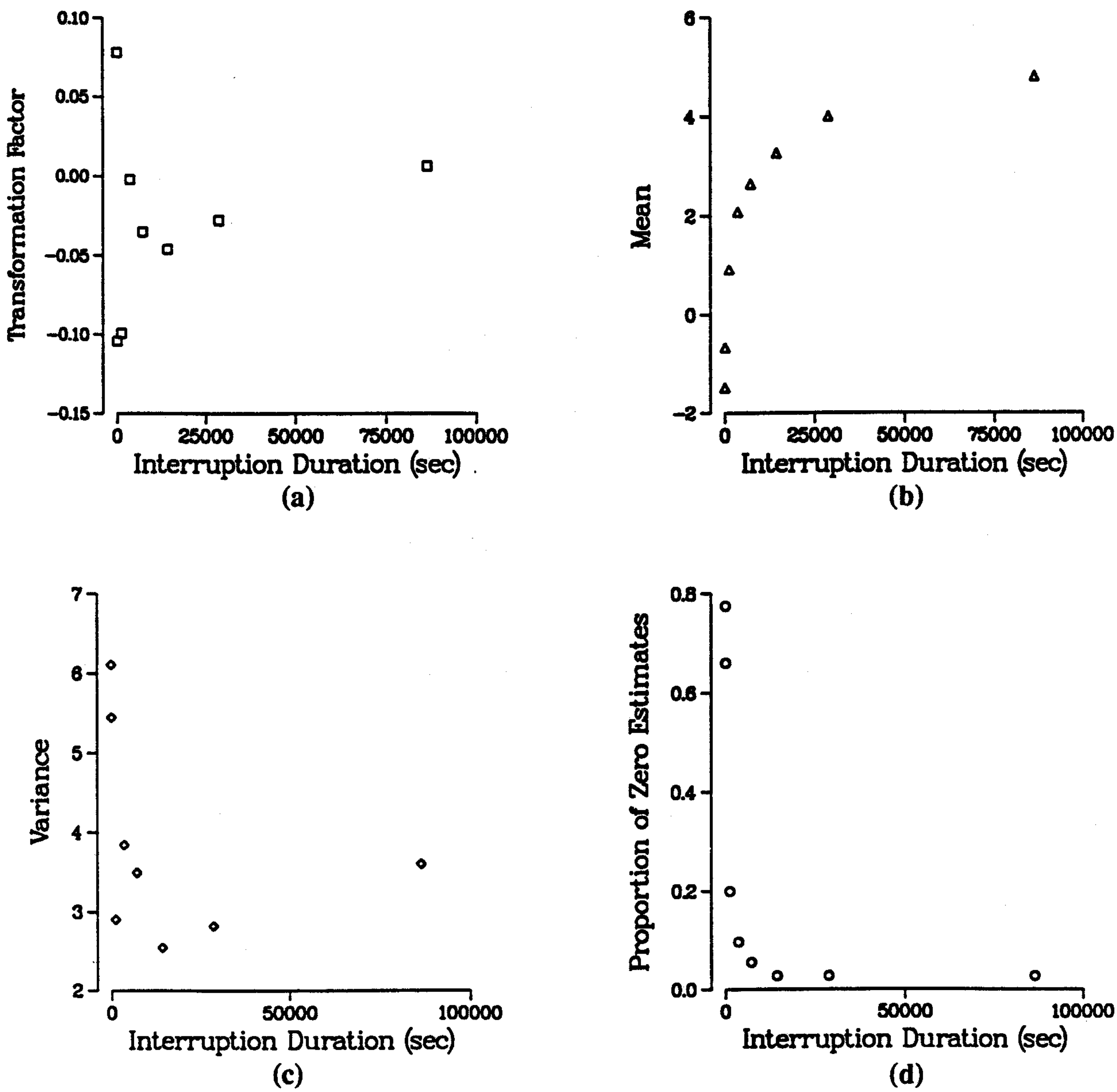


Figure 4.5: Scatter Diagrams of the Cost Distribution Parameters as a Function of Interruption Duration : 1991 Commercial $/KW Cost

problem. The last data point in each of the scatter diagrams is well removed from the others. This lack of data between the seventh and the eighth observed points will greatly influence the regression result as points with an even spread and continuous pattern generally result in a more closely fitting equation. One common practice in dealing with largely scattered points is to change the form of one or both variables in the hope that the new data points exhibit a better shape. The most common approach to transforming the independent variable axis by using a logarithmic function is employed in this analysis.

The regression procedure started with the straight line model and then proceeded to higher order polynomials until a R^2 of 80% or more was achieved. Figures 4.6 (a) to (d) show the scatter plots of the data on a logarithmic scale with the respective regression curve superimposed on each graph. The values of R^2 are also indicated. The mathematical equations for the four fitted curves are as follows:

$$\lambda = 0.1567 - 0.2933 \cdot \log(d) + 0.1001 \cdot [\log(d)]^2 - 0.0096 \cdot [\log(d)]^3, \qquad (4.1)$$

$$\mu = -1.6006 + 0.1424 \cdot \log(d) + 0.2417 \cdot [\log(d)]^2, \qquad (4.2)$$

$$\sigma^2 = 6.8094 - 1.4619 \cdot \log(d) + 0.1423 \cdot [\log(d)]^2, \qquad (4.3)$$

$$P_z = \begin{cases} 0.8915 - 0.2124 \cdot \log(d), & d < 4 \text{ hours}, \\ 0.0278, & d \geq 4 \text{ hours}. \end{cases} \qquad (4.4)$$

It can be seen that the proportion of zero estimate, P_z, has a slightly different representation than the others. It is approximated by a straight line model for interruption durations below 4 hours and remains constant for all longer interruptions. If the inversely-proportional linear representation is applied to interruptions longer than 4 hours, a negative value of P_z will occur which violates its minimum allowable value of zero. A small positive proportion of "0" estimates is therefore assigned for an electrical outage which lasts longer than 4 hours. The set of fitted curve equations for the residential and industrial sectors are given in Tables E.2 (a) and (b) respectively.

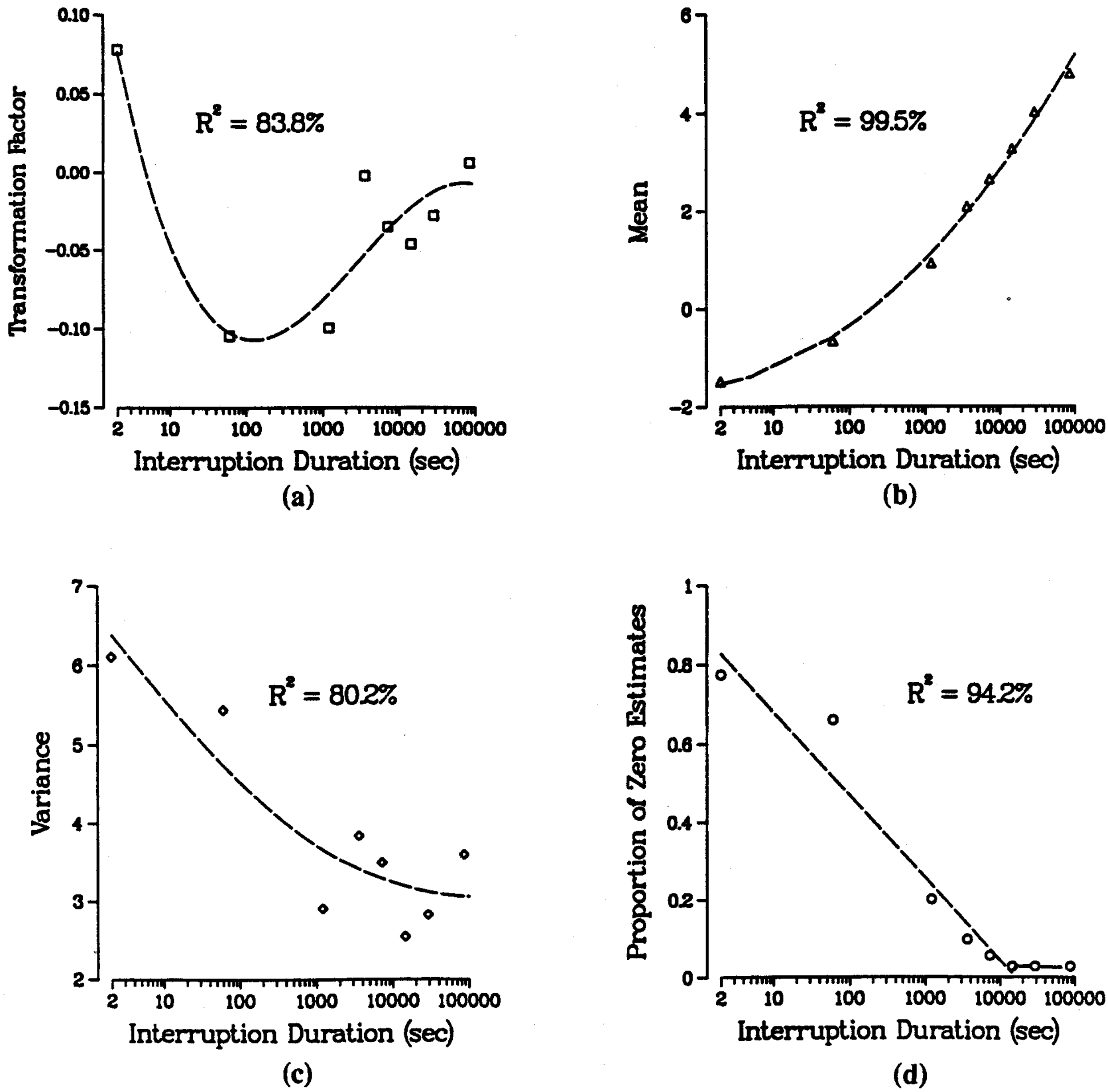


Figure 4.6: Regression Results : 1991 Commercial $/KW Cost Distribution Parameters as a Function of Interruption Duration

4.4.3. Three-Dimensional Customer Damage Function

Figure 4.1 shows the sector CDF's using simple average cost values. In this type of graph, interruption duration is the independent variable and the average cost is the dependent variable and there is one single cost estimate corresponding to every specified duration. In contrast, when outage cost probability distributions are created, the resulting model is more difficult to illustrate. The difficulty comes from the fact that, at any particular duration, the cost data is no longer represented by a single value, but is now characterized by a distribution pattern which requires a two-dimensional (2-D) representation. If this 2-D distribution is to be described as a function of duration, a third axis is required and the resulting display becomes a three-dimensional (3-D) plot. This type of representation is designated, in this study, as the 3-D customer damage function. This section briefly discusses the generation of a 3-D Commercial SCDF.

4.4.3.1. Generation Of A Normal Cost Curve

The first step in building a 3-D SCDF is to create a set of normal sector-cost curves at the selected interruption durations. The moment test results given in Table 4.7 show that the transformed $/KW costs at all eight durations satisfy the normality tests. The distribution patterns of these transformed data are described by the normal parameters (μ, σ^2) given in Table 4.8, i.e. the transformed costs at the 1-hr scenario can be represented by a normal curve, N(2.0813, 3.8423). It must be appreciated that the resultant 1-hr normal curve is not necessarily identical to the contour of the frequency distribution pattern given in Figure 4.4 (d). It is, however, a valid approximation of the actual underlying distribution.

The normal probability density function (pdf) [30] denoted as *f(y)* is expressed mathematically by the following equation:

$$f(y)=\frac{1}{\sigma\sqrt{2\pi}}\cdot\exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right),\qquad \text{for all real } y, \tag{4.5}$$

where y is the normal random variable under study. The commercial sector-cost models at various durations obtained using this formula are depicted in Figure 4.7. In this picture, the horizontal scale uses the values of transformed cost y and the vertical scale

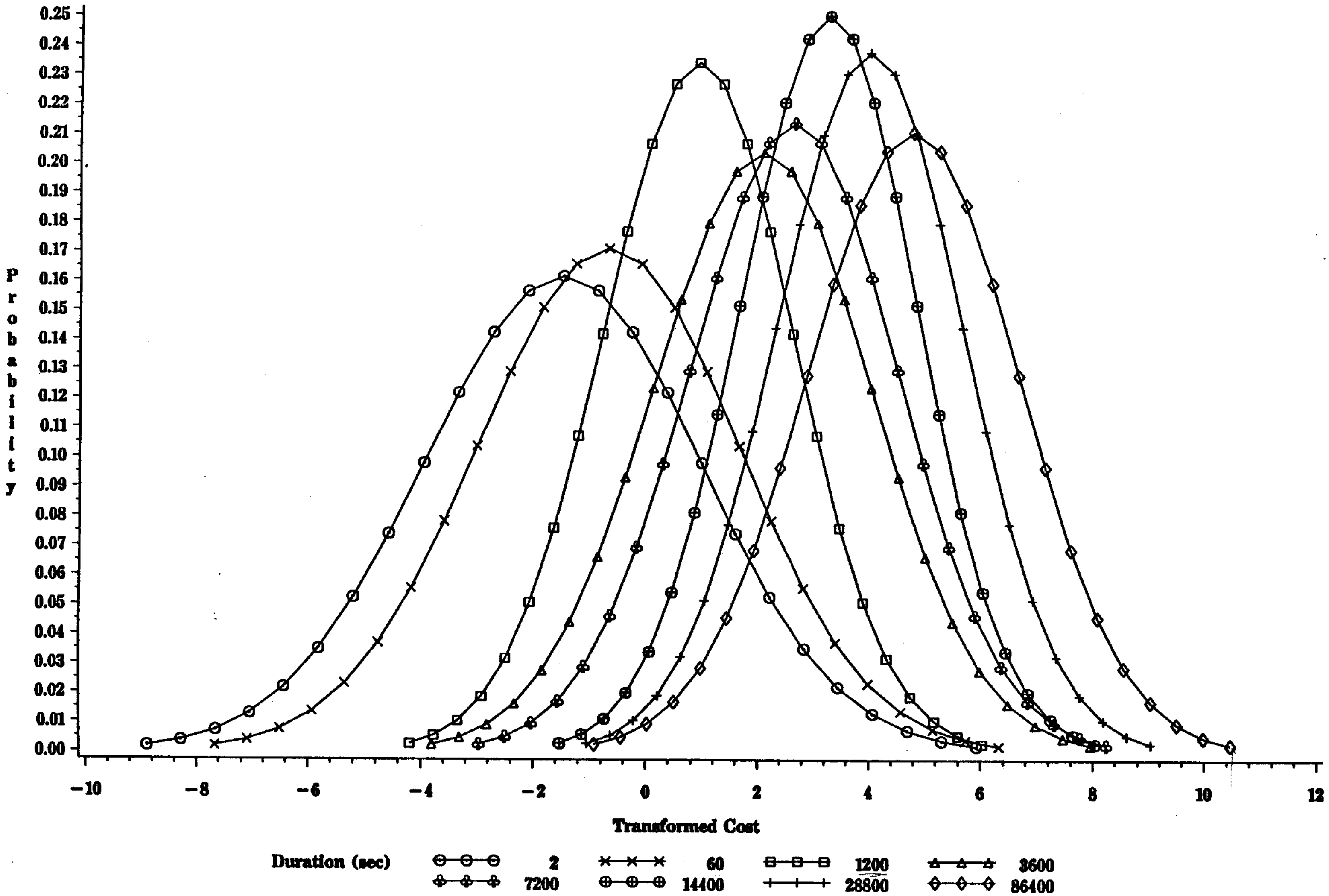


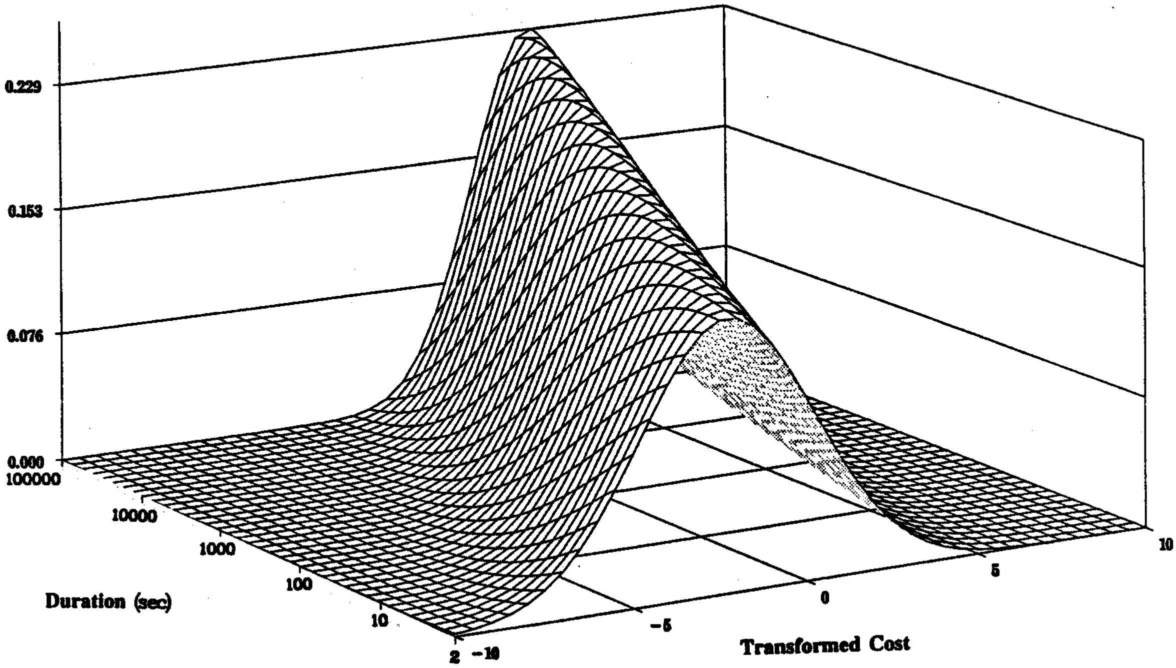
Figure 4.7: Normal Curves for the 1991 Commercial Sector's Transformed $/KW Interruption Cost

uses the values of *f(y)*. The eight normal curves are not themselves sufficient to generate a 3-D commercial SCDF as more data are required to fill the gaps between these studied durations. A valid 3-D plot cannot be generated if there is insufficient data points or if the points are very irregularly spaced [38]. The results of the regression analysis presented in Section 4.4.2 can be utilized to resolve this problem as they permit generation of intermediate models to fill the gaps. The number of models, however, has to be kept within a reasonable number so that the grid size on the resulting 3-D graph will be large enough to see. The normal curve parameters μ and σ^2 for any specific non-questioned interruption duration can be calculated using Equations 4.2 and 4.3 respectively. The normal curve can be created using Equation 4.5 once the parameters are known.

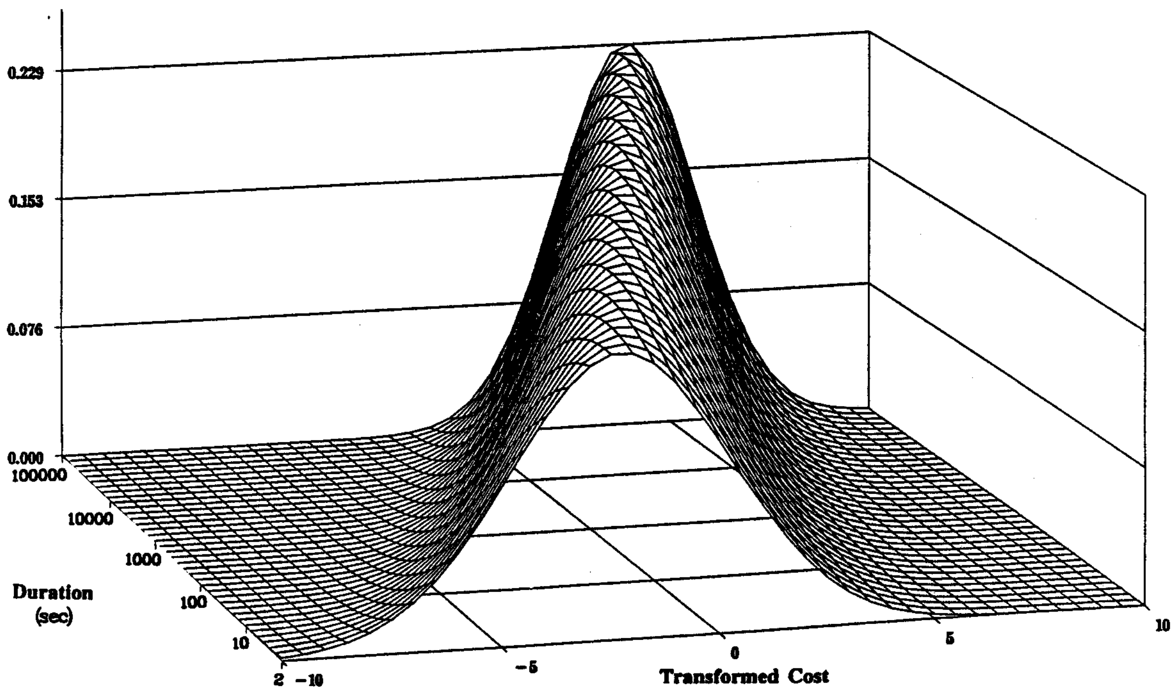
4.4.3.2. Three-Dimensional SCDF

The normal curves presented in Figure 4.7 do not show the dimensional separation of the models for the various durations. This simple representation has the advantage of clearly illustrating the width and height comparisons between different curves. An additional axis is required in order to establish a SCDF from these 2-D distribution models. In this thesis, the axes of a 3-D SCDF are assigned in the following manner: the y-axis identifies the duration in seconds, *d*; the x-axis identifies the transformed cost, *y*; the z-axis identifies the pdf of the costs, *f(y)*. A SCDF in this form describes the distribution of the normal-transformed $/KW costs as a function of duration. Given any duration of power outage, the SCDF is now capable of revealing the associated outage cost distribution (the z-x plane of the 3-D plot). If the power transformation factor at that duration is known, any actual outage cost can be easily recovered from its transformed value.

Figures 4.8 (a) and (b) show the commercial 3-D SCDF at two different angles. Four major observations can be drawn from the pictures. These findings are all well within expectations. First, the pdf of small transformed costs decreases as the duration increases meaning respondents reported fewer small losses at longer durations. Second, the pdf of high transformed costs increases as the duration increases implying that there are more severe losses reported at long durations. Third, the cost models at short



(a)



(b)

Figure 4.8: Commercial 3-D SCDF using the 1991 $/KW Interruption Cost

durations have greater width than at the high durations which implies that respondents reported a large range of outage costs at short durations, while estimates given at long durations concentrate more around the peak. Finally, it can be seen that the mean transformed cost (peak of the normal curve) increases as the duration increases.

4.5. Summary

Two interruption cost modelling techniques designated as the customer damage function method and the probability distribution approach were utilized in this chapter to describe the 1991 NSERC customer interruption cost data. The conventional CDF represents the outage costs using expected values. These averages are relatively easy to compute and to use. They do not, however, reflect the entire cost profile. This difficulty was alleviated by the probability distribution approach.

This chapter utilizes the 1991 $/KW data to conduct a user and duration specific customer interruption cost study. The outage estimates were grouped according to (i) interruption scenarios under which they were given, (ii) SIC category and (iii) economic sector of the respondents. Simple averages were first calculated at the lowest data group level and then combined to create the sector or the entire service area expected interruption costs using weighting factors designed for the respective purpose. The three sector customer damage functions and the system composite customer damage function generated using the 1991 $/KW responses are given in Figures 4.1 and 4.2 respectively. The functions, in their given forms, can be represented by piece-wise linearly increasing relationships. These simple representations permit easy calculation of the average outage costs at intermediate durations.

A probability distribution method was developed to describe the dispersed nature of the outage estimates, which is overlooked by the conventional CDF approach. This chapter illustrates the procedures used to develop a distribution pattern using the 1991 commercial $/KW responses. Prior to the analysis, the responses were grouped into representative categories similar to those employed in the CDF method, except that SIC breakdown was not applied. The data groups were symmetrically transformed and then tested for normality using a hypothesis testing method. The results are presented in Table

4.6 and Figure 4.3. The primary finding from these results is that the data sets belonging to the 2-sec and 1-min scenarios could not be transformed into acceptable normal distributions.

The Box and Cox power transformation should only be applied to continuous variables and therefore the assumption was made that the customer interruption costs at every studied scenario contained only continuous values. This assumption is crucial to the distribution approach, but there is strong evidence that the 2-sec and 1-min data groups also contain values sampled from a discrete distribution. It was found that the discrete distribution primarily consists of zero-valued data. These data were removed and the power transformation was performed on the non-zero cost values. The new results, presented in Table 4.7 and Figure 4.4, show that the modified method permits acceptable normality transformations in all data groups.

The last step of the distribution approach is to perform regression analysis in order to determine the models associated with intermediate durations. The basic objective is to develop a mathematical relationship between each of the distribution parameters (P_z, λ, μ and σ^2) and the outage duration. The results of the commercial sector study are presented in Figure 4.6 and expressed mathematically in Equations 4.1 to 4.4. The value of any intermediate parameter can be easily calculated from the respective equations.

The conventional CDF can be portrayed by a simple 2-D plot in which the interruption cost is the dependent variable and the duration is the independent variable. This representation allows the analyst to easily obtain the outage cost value associated with a given duration. The same philosophy was applied to the distribution method. In this approach, the normal-transformed costs at each interruption scenario are portrayed by a normal curve. In order to describe the individual cost curves as a function of duration, a third axis is required. The resulting representation is designated, in this study, as the 3-D Customer Damage Function. The commercial 3-D SCDF is given in Figure 4.8. This function gives a pictorial display of the entire commercial $/KW cost profile. The representation, however, is not as practical as its 2-D counterpart given in Figure 4.1 in terms of use in an analytical HLI study. It is difficult, in an analytical study, to incorporate the distributed nature of the outage cost data. Chapter 5 illustrates the role of cost distributions in an overall HLI reliability evaluation using Monte Carlo simulation.

5. GENERATION OF AN INTERRUPTED ENERGY ASSESSMENT RATE

5.1. Introduction

Users are the ultimate recipients of the service provided by the electric utility industry therefore their requirements and opinions should be an important input factor in utility supply and demand-side planning. In recent years, there has been substantial interest in attempts to relate power system planning to a customer driven issue, namely the worth of electric service reliability. The term reliability worth refers to the benefit derived by the users in receiving electrical energy and can be approximated by the customer costs of interruption.

The cost of interruption data collected from the customer surveys have been described by a cost model. This model serves one primary function and that is to relate the monetary estimates to the respondent type and to the duration of interruption. A model in a form such as this is essential if the cost data is to be utilized to quantify the worth of service reliability. A customer damage function is a traditional form of representing the cost of interruption data and the formation of various CDF's is discussed in Chapter 2. A new method of characterizing the data, namely the probability distribution approach, is introduced in this thesis. The concepts, the requirements and the formation of this representation are presented in Chapter 3. In order to create a practical tool for subsequent quantitative reliability worth calculations, an interruption cost model must be related to the reliability indices used in system planning and operation. A factor designated as the Interrupted Energy Assessment Rate (IEAR) [17] has been developed to serve this purpose. An IEAR can be used to link the interruption cost models to a wide range of reliability cost / reliability worth assessments. This chapter presents the procedures utilized to create an IEAR in an HLI study using each of the two 1991 cost models developed in Chapter 4.

5.2. IEAR Estimation by Monte Carlo Simulation

The basic factors involved in IEAR estimation are the customer interruption costs, the amount of unsupplied energy and the duration of all load loss events in a given studied period. The customer interruption costs can be calculated by using either one of the two cost models developed in Chapter 4. The expected energy not supplied and the duration of power failures can be readily assessed by both analytical and simulation techniques. These methods can be briefly described as follows:

1. *Analytical* methods represent the system by a mathematical model and evaluate the adequacy indices by solving the model, and,
2. *Monte Carlo simulation* (MCS) methods involve the generation of an artificial history of the power system. The adequacy indices are the inferences drawn from this simulated history of the system behaviour.

There are both merits and demerits in either method. Generally, analytical approaches have the advantage of directness where suitable approximations and assumptions are often made to simplify the complexity. This approach, however, does not explicitly consider various unit functions such as complicated component failure and repair distributions, and system operating policies such as variation in reserve requirements. Monte Carlo simulation, on the other hand, provides the greatest capacity to include such complex functions and policies [39, 40]. Another major limitation of the analytical approach is that it usually predicts only the average reliability indices. From the sensitivity studies conducted in [41], it is believed that the utilization of average indices to estimate the IEAR results in large errors. The difficulty was alleviated by using individual load loss event data which are available in the MCS approach. Due to this distinct advantage, a Monte Carlo simulation approach has been used to conduct the HLI reliability evaluation described in this thesis. The major shortcoming of Monte Carlo simulation methods is that they often require considerable computing time.

5.2.1. Basic Models in HLI Adequacy Assessment Using Monte Carlo Simulation

The basic approach in an HLI adequacy study includes three different models, the generation model, the load model and the risk model as shown in Figure 5.1. In MCS,

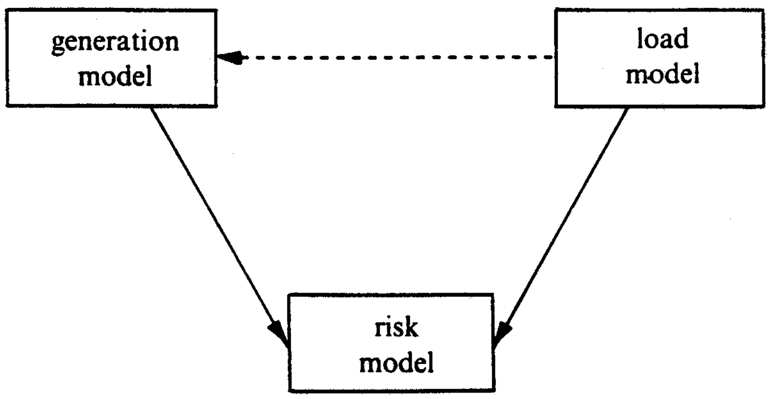


Figure 5.1: Block Diagram for HLI Adequacy Assessment

the generation model is an artificial history of the system up, down or derated states. These states can be obtained by combining the operating cycles of individual generating units where each cycle is determined from the unit Time-To-Failure (TTF) and Time-To-Repair (TTR) [42] distributions. Consider a system consisting of two simple up-down-up generating units. The generation model or the available capacity of the system can be obtained by adding the two individual unit operating cycles. Figure 5.2 illustrates the combining process [43]. The load model used in MCS is generally assumed to have a discrete change every hour and to be constant throughout the hour. In order to create the risk model, the system capacity model is superimposed on the load model to evaluate the system capacity reserves or deficiencies depending on the difference between the available capacity and the load demand. Figure 5.3 illustrates the superimposition process [43]. Each of the shaded areas in the picture represents the amount of energy curtailed when an power outage occurs. This unsupplied energy is denoted in this study as *E*, while the corresponding duration is designated as *d*.

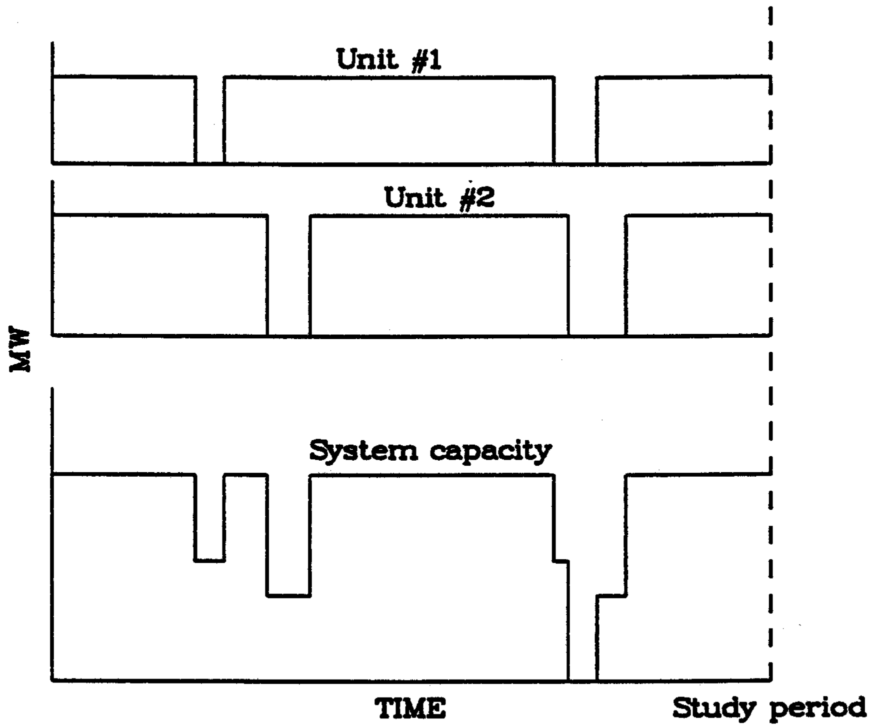


Figure 5.2: System Capacity Model : Combination of Unit Generating Cycles

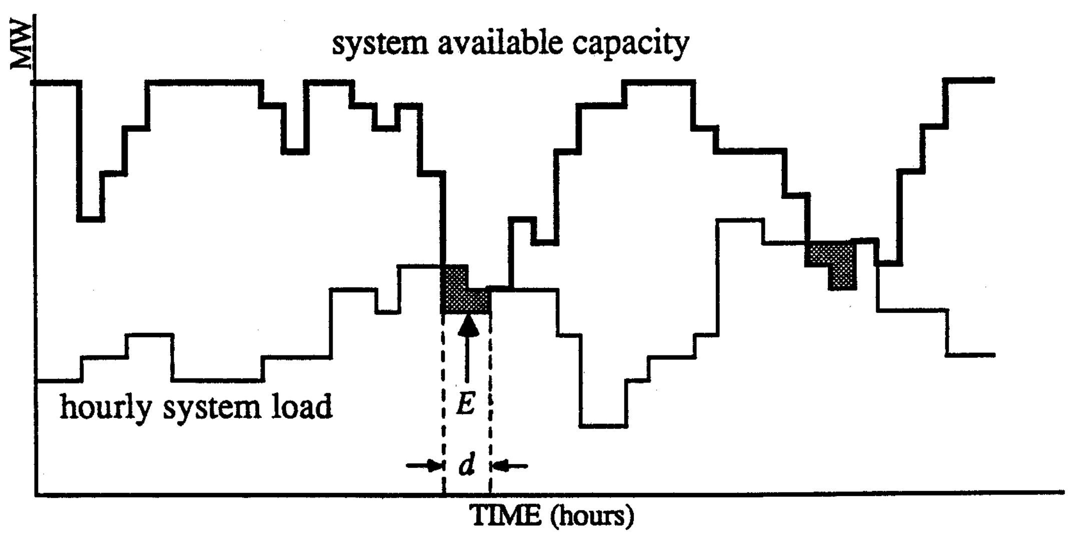


Figure 5.3: Risk Model : Superimposition of the Generation Model on the Load Model

The Monte Carlo simulation approach uses a minimum time unit of one hour. This implies that the simulation process takes place every hour in the entire sampling period. Since most of the reliability indices are yearly based, the sampling period is generally a multiple of one year. In each sampling year, the interruption duration in hours, the amount of energy not supplied in KWh and the frequency of each load loss event can be determined by observing the simulated system behaviour as shown in Figure 5.3. These three values can be conveniently used to calculate various yearly reliability indices [41, 43].

5.2.2. Mathematical Formulation

It was noted that the utilization of average indices results in a significant error in an IEAR estimation. An alternate approach is to use the individual load loss event information from the entire sampling period using Equation 5.1 [43].

$$IEAR = \frac{\sum_{i=1}^{n} C_i(d_i) \cdot E_i / d_i}{\sum_{i=1}^{n} E_i} \qquad (\$/KWh), \qquad (5.1)$$

where:

d_i is the duration in hours of interruption i,

$C_i(d_i)$ is the interruption cost in $/KW at duration d_i,

E_i is the energy not supplied in MWh of interruption i,

n is the total number of interruption experienced during the simulation per

The peak normalized $/KW cost values are used as they can be applied in direct manner. The values of E_i and d_i are obtained from a simulated history of the system. The cost element $C_i(d_i)$ is obtained from an interruption cost model which can be portrayed either by a customer damage function or by a probability distribution. The remainder of this section is devoted to the procedures used to estimate an IEAR from these two models.

5.2.3. Outage Cost Estimation Using a CCDF

The CCDF cost model is relatively easy to use and can be graphically represented as shown in Figure 4.2 where the weighted average outage costs are plotted as a function of interruption duration. The resulting CCDF is approximated by a piece-wise linear relationship in which every segment between two studied durations is described by a straight line equation. The interruption loss C_i corresponding to a simulated outage duration d_i was computed by linear interpolation from the damage function. Interruptions which last less than two seconds were approximated by linear extrapolation. Interruptions which last longer than one day were assigned a cost equal to the average one-day cost. Utilization of a CCDF is relatively straight forward since there is only one calculated average cost value associated with each simulated duration.

5.2.4. Outage Cost Estimation Using Probability Distributions

An interruption cost model described by a probability distribution is more difficult to use than the conventional CCDF since the costs are not described in the form of pre-calculated single averages. The \$/KW outage costs in their transformed representations are characterized by normal probability distributions. A practical way to select a value, i.e. the variate, from a distribution is by using random numbers.

By definition, random numbers must be independent of each other and should have equal probability of assuming one of the possible values. There are many ways to generate random numbers. One way is to use a digital computer to generate a sequence of random numbers using specific formulae developed for this purpose. The numbers generated in this manner have practically the same behaviour as random numbers, and are called pseudorandom numbers. The random numbers utilized in this application are uniformly distributed pseudorandom real numbers between zero and one. Their functions regarding the IEAR estimation are as follows. A (0,1) random number is used to make a decision regarding whether a zero-valued outage cost will be assigned to the simulated duration d_i or whether to proceed further to generate a random variate from the normal probability distribution which describes the group of transformed data at d_i.

A different (0,1) random number x_1 is generated in each studied sector for every load loss event which occurs with an outage of d_i. The x_1 value is first compared with the proportion of zero-valued data P_z. A zero sector cost is assigned and the evaluation process terminates if x_1 is less than or equal to P_z. If x_1 is greater than P_z, the process continues and x_1 is used to generate another random number x_2 in order to select a transformed outage cost y from the normal curve predicted at d_i. The concepts and procedures for using x_2 to select a normal random variate y are described in the following sections.

5.2.4.1. Standardized Variable

The key to working with a normal distribution is the standard score, Z, which is the number of standard deviations from which a value is removed from the mean. The value of Z associated with the variate y is given by

$$Z=\frac{y-\mu}{\sigma}, \tag{5.2}$$

where σ is the standard deviation and μ is the mean of the normal distribution from which y is sampled. The Z score is also known as a "standardized" variable because its unit are standard deviations [30]. The normal probability distribution associated with this standardized score is called the standard normal distribution. The probability that a value picked at random falls between two values of Z can be represented by the area within the interval under a standard normal curve. The random number x_2 is utilized in this application to represent an area as shown in Figure 5.4. This figure also shows the relationship of x_2 with the standard score Z_1 and with the normal variate y.

5.2.4.2. Selection of a Normal Variate

The first step in sampling a non-zero cost in any particular economic sector for a simulated duration of d_i is to determine the value of Z_1 corresponding to the random number x_2 generated in the sector. The Z_1 value can be looked up in a standard normal table as given in Appendix C when a manual calculation is performed. When a digital computer program is used, the value can be determined using a set of mathematical formulae provided for this purpose. Once the value of Z_1 is known, it can be used in

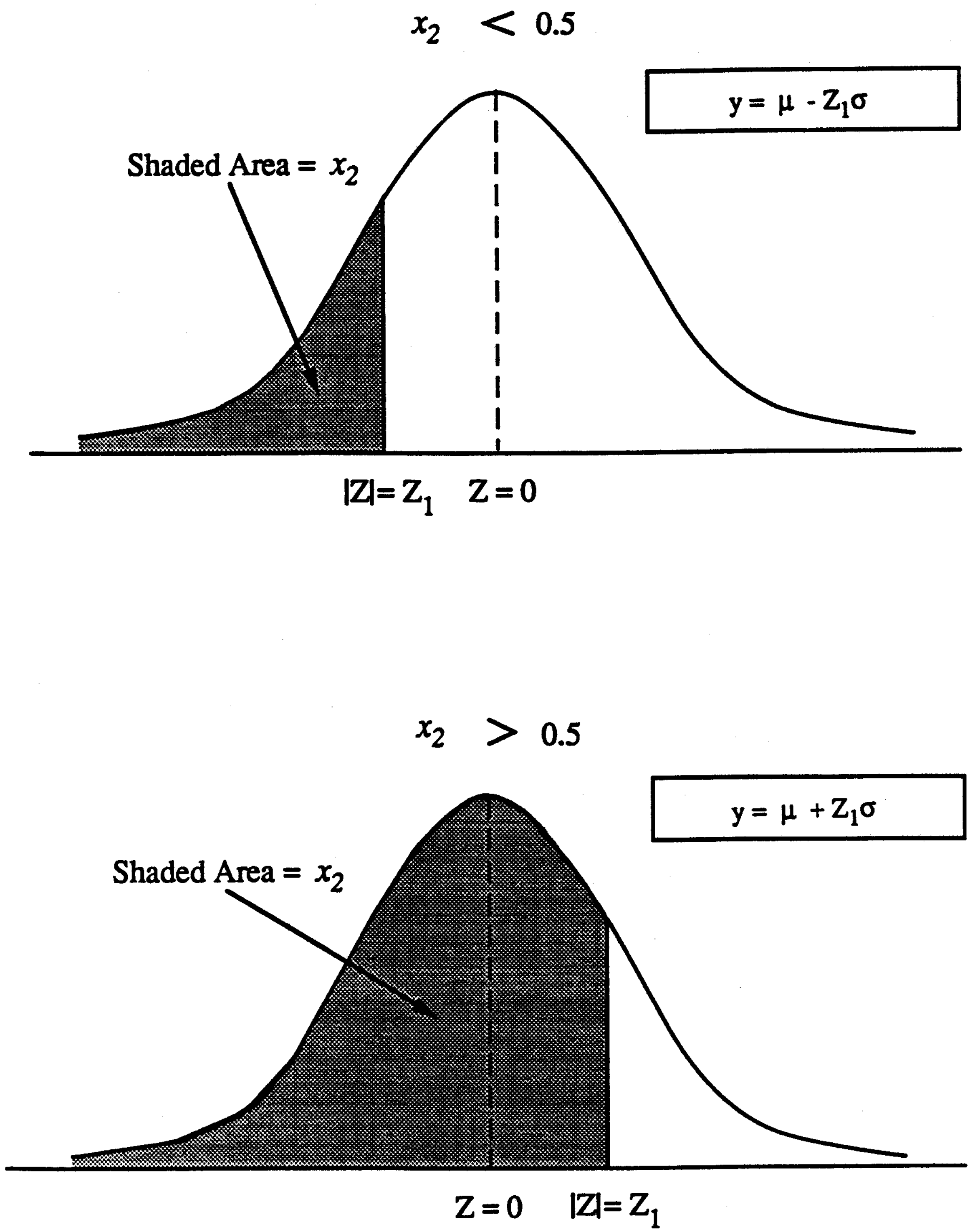


Figure 5.4: Standard Normal Curve, Z-Score and Variate

conjunction with x_2 to determine an associated transformed cost y. There are three possible domains in which the y value can lie depending on the value or position of x_2. It can be larger than, less than or equal to the mean of the normal distribution from which it is sampled. If x_2 is larger than 0.5, the corresponding y value is equal to μ plus the product of Z_1 and the standard deviation σ. If x_2 is smaller than 0.5, the y value is equal to μ minus the product of Z_1 and σ. If x_2 equals 0.5, y will have the same value as the mean. Given the sector transformation factor λ at d_i, y can be converted to its actual \$/KW value. After having estimated all actual sector monetary losses, the total outage cost $C_i(d_i)$ for the entire service area can be evaluated by weighting the sector costs by their relative energy consumptions in the area.

5.3. Numerical Examples Using the RBTS

The Monte Carlo simulation approach was utilized to calculate the system IEAR of the RBTS using the interruption cost data collected from the 1991 survey. A digital computer program GRAP [43] written in Fortran-77 was used to perform the simulation. The RBTS generation model used in the simulation is given in Table 5.1 where MTTF and MTTR denote the unit mean-time-to-failure and mean-time-to-repair in hours respectively. The hourly load cycle data is identical to that of the IEEE-RTS given in [44]. This model specifies the system load values for every hour of the year (8736 data points). The annual peak demand is considered to be 185 MW. The interruption cost models described in both CCDF and probability distribution are given in Chapter 4. The program GRAP includes a subroutine to perform the IEAR estimation using a CCDF. The alternate technique of using the probability distribution approach was added to GRAP in order to perform the additional calculations.

The study began with a 3000-year simulation period and an initial seed of 1340983. Using the CCDF cost model, the estimated IEAR is 4.42 \$/KWh. When the normal probability distribution method is used, the IEAR becomes 15.74 \$/KWh. This value is approximately three times larger than the value obtained using the CCDF model. In order to examine the confidence associated with both cases, a simulation convergence study was conducted.

Table 5.1: RBTS Generating Unit Rated Capacity and Reliability Data

Unit No.	Capacity (MW)	MTTF (hours)	MTTR (hours)
1	40.0 (thermal)	1460	45
2	40.0 (thermal)	1460	45
3	10.0 (thermal)	2190	45
4	20.0 (thermal)	1752	45
5	5.0 (hydro)	4380	45
6	5.0 (hydro)	4380	45
7	40.0 (hydro)	2920	60
8	20.0 (hydro)	3650	55
9	20.0 (hydro)	3650	55
10	20.0 (hydro)	3650	55
11	20.0 (hydro)	3650	55

5.3.1. Imposing Maximum and Minimum Bounds

Prior to examining the solution convergence, the question of upper and lower limits on the normally distributed variates was considered. The fact that an interruption cost model in the form of a normal-transformed cost distribution is statistically sound does not guarantee that every value sampled from the model will give a "reasonable" actual value. The sampled cost should be a value which after conversion has a minimum possible value of zero and a realistic maximum value. One way of selecting the maximum bound value at a given duration for a specific sector is to use the actual maximum estimate reported by that group. The maximum values at all studied scenarios can then be used to infer the maximum values at other possible durations. Any generated value greater than the maximum boundary will be assigned a value equal to the maximum. This boundary value restriction approach eliminates the likelihood of a randomly generated outage cost having an unreasonable monetary value. The IEAR estimated from the distribution approach dropped slightly to 15.29 $/KWh when the boundary limitations were imposed. The decrease suggests that there are indeed some sampled costs which after conversion give actual losses higher than the maximum reported estimates.

5.3.2. Simulation Convergence

It is difficult to make a general statement regarding the choice of seed and the amount of simulation time required to achieve reasonable confidence in a MCS analysis. Past studies have indicated that the simulation time required is a direct function of the system size and the reliability of the system [41]. The RBTS is a small system with a relatively high system reliability. The determination of a desired simulation time is basically a trial and error process.

The IEAR convergence was evaluated as follows. A selected simulation period was deemed acceptable only if the IEAR's generated from different seeds for this period do not show significant variation as measured by percentage difference. In this study, a critical value of 10% was used which means that any simulation period giving a difference exceeding 10% will not be accepted. Based on this, the IEAR values calculated by the two cost models using a simulation period of 3000 years were verified. The simulation was repeated using five different initial seeds and the results are shown in Table 5.2. The ten possible pairs of IEAR values obtained from the CCDF exhibit a maximum percentage difference of 6.58. It can be therefore concluded that when a CCDF is used and a preselected maximum allowable difference of 10% is imposed, a 3000-year simulation time is long enough to reach an acceptable level of IEAR convergence. In contrast, the IEAR values resulting from the probability distribution model have a larger variation with values ranging from 11.5 to 19.9 $/KWh which corresponds to a maximum difference of 72.6%. It can be concluded that the IEAR values determined from the distribution approach do not satisfy the desired level of convergence by using a simulation time of 3000 years.

It should be appreciated that the number of customer outage events which contribute to the IEAR estimation is not equal to the number of simulation years. In a highly reliable system, there will be many years during which customers will not suffer any power failures. The outage cost estimation process receives a contribution only if a power deficiency occurs. The number of load loss events that contribute to the calculation of an IEAR is therefore much smaller than the number of sampling years. In

Table 5.2: IEAR values from a Simulation Period of 3000 Years

Initial Seed	IEAR ($/KWh) CCDF	Prob. Distribution
1	4.298766	19.86309
16807	4.471748	17.44538
1340983	4.424137	15.29099
3333335	4.396407	11.50754
4444447	4.581708	17.45414

the previous runs conducted, the total number of load loss events and hence the number of outage cost calculations is in the neighbourhood of 650. The duration of the sampling period should be increased in order that more load loss events are simulated and a higher confidence in the IEAR value is attained. Tables 5.3 (a) to (c) show the IEAR values generated from different random number streams using 3000-year, 10000-year and 99000-year periods respectively.

When the sampling period increases to 10000 years, the number of interruptions which occur increases to about 2100. The maximum variation in IEAR values drops to 17.4% from the 72.6% found in the 3000-year simulation. The variation, however, is still larger than 10%. It was therefore concluded that a simulation time of 10000 years is not long enough to achieve the desired level of IEAR convergence. As a consequence, the simulation time was further increased to 99000 years which results in approximately 21000 interruptions. The largest percentage difference which now exists is 4.03. From the studies conducted, it is believed that an IEAR value in the range of 14.27 to 14.84 $/KWh is a reasonable estimate using the probability distribution method. The IEAR variation also decreases as the simulation time increases in the case of using a CCDF. At a simulation time of 99000 years, the IEAR value is approximately 4.46 $/KWh.

Table 5.3: IEAR Convergence Study by Varying Initial Seeds and Simulation Time

(a) IEAR values from a 3000-year simulation

Initial Seed	No. of Outages	IEAR ($/KWh) CCDF	IEAR ($/KWh) Prob. Distribution
1	657	4.298766	19.86309
16807	664	4.471748	17.44538
1340983	676	4.424137	15.29099
3333335	660	4.396407	11.50734
4444447	564	4.581708	17.45414

(b) IEAR values from a 10000-year simulation

Initial Seed	No. of Outages	IEAR ($/KWh) CCDF	IEAR ($/KWh) Prob. Distribution
1	2163	4.370974	17.190388
16807	2114	4.453438	17.522074
1340983	2080	4.416052	14.922331
3333335	2187	4.469794	16.111372
4444447	1978	4.415451	15.020037

(c) IEAR values from a 99000-year simulation

Initial Seed	No. of Outages	IEAR ($/KWh) CCDF	IEAR ($/KWh) Prob. Distribution
1	21614	4.453482	14.314417
16807	20860	4.468817	14.776078
1340983	21141	4.465606	14.843599
3333335	21280	4.478129	14.268806
4444447	21297	4.458311	14.650346

5.4. Summary

Monte Carlo simulation was used in this study to conduct the HLI reliability assessment because of its ability to provide individual load loss event information such as the duration of failure and the energy not supplied. Estimation of an IEAR using the outage duration and curtailed energy distributions provides a more accurate result compared to that obtained using average values.

The three elements of an IEAR evaluation using MCS are the cost C_i, the energy not supplied E_i and the duration d_i of every simulated outage during a given studied period. The values of E_i and d_i were obtained from the simulated system operating history. The value of C_i was estimated by using either the customer damage function or the probability distribution. The CCDF cost model is relatively easy to use. Under most circumstances, the cost C_i corresponding to a simulated outage duration d_i is computed using linear interpolation. The evaluation process is more complex when the distribution method is used. The procedure developed in this research work is briefly summarized as follows.

The first step is to generate a different uniform (0,1) random number x_1 in every studied economic sector. When an outage of d_i hours occurs, the value of x_1 is used to (i) determine whether or not a zero interruption cost should be assigned to the outage or (ii) proceed to generate another random number x_2 in order to sample a randomly distributed transformed cost y from the appropriate normal curve. If the value of x_1 is less than or equal to the proportion of zero-valued data P_z, d_i is assigned a zero outage cost and the cost calculation terminates for that particular sector. On the other hand, if x_1 is higher than P_z, the process continues to select a y value. The value of y can be any possible transformed cost under the normal cost curve built at d_i. The basic idea is to convert the x_2 value to a standard Z score, Z_1. Depending on the position of Z_1, the value of y can be determined using the mean and standard deviation of its corresponding normal distribution. The selected y value for each sector is converted to its actual \$/KW cost using the transformation factor λ before the total outage cost C_i for the studied area is determined.

The MCS approach was utilized to calculate the system IEAR of the RBTS. The IEAR generated from the 1991 CCDF is 4.42 $/KWh using a simulation period of 3000 years and an initial seed of 1340983. The probability distribution method produced a much larger estimated IEAR of 15.74 $/KWh. This value drops slightly to 15.29 $/KWh when boundary limitations are imposed. The difference between the results obtained from the CCDF and distribution method is quite significant and further simulation convergence studies were conducted to verify the accuracy in both cases. A selected simulation period was deemed acceptable only if the IEAR's generated from different seeds do not show a percentage difference larger than 10%. Based upon this, the IEAR values generated from the CCDF were found to be satisfactory with a simulation period of 3000 years. In comparison, the distribution model did not give satisfactory results until the simulation time was increased to over 10000 years.

The simulation studies presented in this chapter suggest that the IEAR estimated using the probability distribution cost model is considerably larger than the value obtained using the CCDF approach. It is reasonable therefore to suggest that for each simulated load loss event, the outage cost sampled from the associated normal distribution representation is generally greater than the average cost calculated from the corresponding CCDF. This finding is supported by the preliminary statistical interruption cost data studies. Table 3.1 shows that in the commercial sector the original untransformed cost exhibits a very large *positive* skewness for every interruption scenario which implies that the number of respondents with costs higher than the mean value greatly exceeds those with smaller costs.

In the studies described in the following chapter, the IEAR values of 4.46 and 14.57 $/KWh obtained using the CCDF approach and the distribution method respectively are used in a generating capacity context.

6. COST / BENEFIT APPROACH TO GENERATING SYSTEM PLANNING

6.1. Introduction

A modern power system serves one primary function and that is to meet the customers energy requirements at an acceptable reliability level and at the lowest possible cost. The two stated requirements may conflict since it is generally true that the investment cost increases as the reliability level increases. From a service industry standpoint, an appropriate approach to balance the reliability and economic constraints is to simultaneously consider the effect of the reliability or risk level upon the system investment cost, and also the effect of the reliability level upon the benefit derived by the electrical energy recipients. In recent years, the utilization of this cost / benefit approach to utility supply and demand-sided planning has begun to receive considerable attention [45]. The procedures for evaluating the relationships between reliability cost and reliability level are reasonably well established and are used by a variety of utility planners. In comparison, the determination of what is an appropriate reliability level using a benefit or worth approach has not been extensively developed and applied. The major difficulty in this approach is that the worth of power system reliability as seen by the customers cannot be easily quantified and expressed in monetary values. One practical approach which yields acceptable results is to approximate the benefit in terms of the effects and impacts of unreliability. This concept is addressed in Chapter 2 in some detail.

The monetary losses that electrical users experience during power cessations are the basic ingredients in a reliability worth assessment. These costs when transformed into a customer factor designated as the interrupted energy assessment rate (IEAR) can be directly utilized to quantify the worth of electric service reliability in monetary terms.

The objective of this chapter is to show how an IEAR can be employed in a capacity adequacy study to determine the optimum level of reliability. This chapter illustrates a basic procedure for determining reliability cost and reliability worth at HLI. The demand for electrical energy tends to grow with time in our modern society and therefore it becomes necessary to add generating units to an existing system in order to meet this demand. This chapter also shows how the cost / benefit method can be used to perform capacity expansion planning. All these studies are illustrated using the RBTS.

6.2. Basic Concepts of the Cost / Benefit Approach

From an economic theory perspective, the selection of an optimum adequacy design level should depend on the cost of providing extra reliability versus the benefits accruing to society from the additional reliability. The application of this philosophy is known as the cost / benefit approach [18]. Power utilities use this approach to determine the target adequacy level by balancing the reliability cost and the reliability worth. This concept is portrayed graphically in Figure 6.1. In the utility or system cost curve, the cost increases

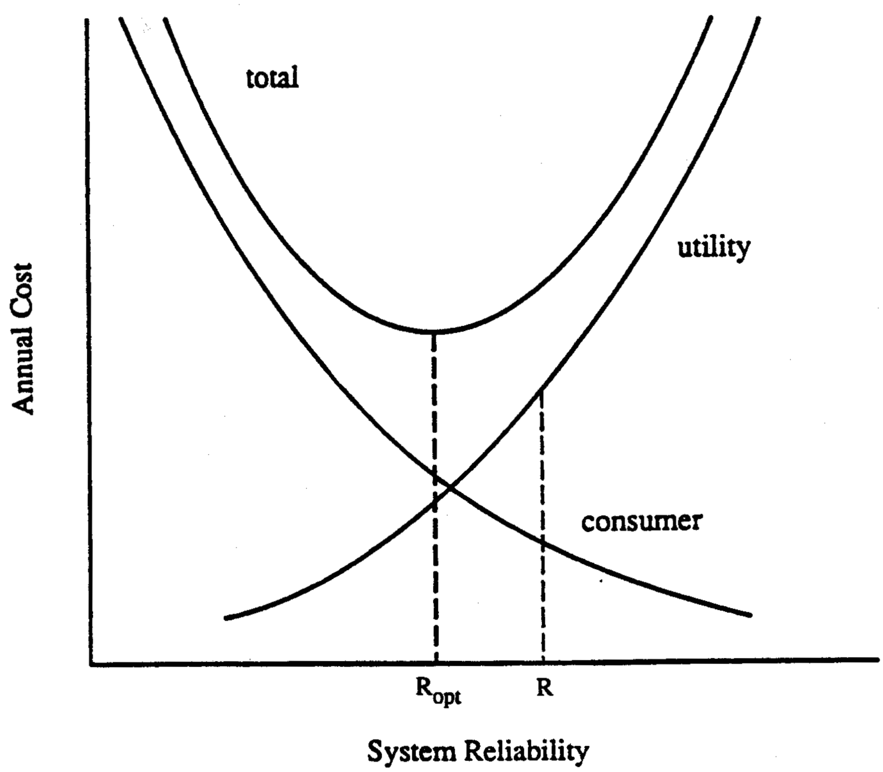


Figure 6.1: Optimum Reliability Level as Determined by the Cost / Benefit Approach

as consumers are provided with higher degrees of reliability. The common approach to power system planning looks strictly at this cost function. A level of reliability R is preselected and the system planning objective is to find a design which satisfies this reliability level at the lowest capital and operating costs. The selection of R is based entirely on past experience and judgement and does not incorporate any customer factors in the evaluation process. In contrast, the cost / benefit method takes into consideration both the cost of supplying power and the benefit or worth of the supplied power as seen by the customers. The consumer costs associated with supply interruptions, as seen from the customer cost curve in Figure 6.1, decreases as the reliability level increases. When the consumer cost is added to the system cost, the resultant value represents the total cost to society. This total societal cost curve exhibits a minimum which corresponds to the optimum or target level of reliability, R_{opt}. The value of R_{opt} therefore depends not only on the generating system and load data but also on the customer interruption costs.

6.3. The Roy Billinton Test System

The fundamental data utilized in this study to conduct the reliability cost / reliability worth assessment comes from the Roy Billinton Test System (RBTS) [29]. The interruption cost data given in [29] is replaced by the new 1991 NSERC $/KW values. The RBTS is a small system which can be analyzed without excessive computing time while involving practical system complexity. The determination of both system cost and customer cost depends largely on the unit operating and reliability data and therefore these data are extracted and shown in this chapter. The RBTS consists of eleven generating units with a total capacity of 240 MW. The minimum and the maximum ratings of these units are 5-MW and 40-MW respectively, and the annual system peak load is 185 MW. The hourly peak load model utilized in this particular study is approximated by a load duration curve described by the 100 data points provided in Table 6.1. The generating unit ratings and the reliability data are shown in Table 6.2 where FOR denotes the unit forced outage rate, MTTF is the mean time to failure and MTTR is the mean time to repair. The RBTS includes additional generation units for the purpose of system expansion. These are gas turbines with the specifications given in Table 6.3.

Table 6.1: 100 Points Load Data for the RBTS

Peak Load (p.u)	Study Period (p.u)	Peak Load (p.u)	Study Period (p.u)	Peak Load (p.u)	Study Period (p.u)	Peak Load (p.u)	Study Period (p.u)
1.0000	0.0000	0.9933	0.0002	0.9866	0.0003	0.9800	0.0004
0.9733	0.0006	0.9666	0.0008	0.9599	0.0010	0.9532	0.0015
0.9466	0.0024	0.9399	0.0034	0.9332	0.0040	0.9265	0.0058
0.9199	0.0076	0.9132	0.0081	0.9065	0.0100	0.8998	0.0137
0.8931	0.0160	0.8865	0.0189	0.8798	0.0239	0.8731	0.0290
0.8664	0.0333	0.8597	0.0401	0.8531	0.0464	0.8464	0.0517
0.8397	0.0614	0.8330	0.0718	0.8264	0.0823	0.8197	0.0906
0.8130	0.1004	0.8063	0.1122	0.7996	0.1254	0.7930	0.1353
0.7863	0.1452	0.7796	0.1574	0.7729	0.1704	0.7662	0.1823
0.7596	0.1918	0.7529	0.2005	0.7462	0.2114	0.7395	0.2232
0.7329	0.2339	0.7262	0.2436	0.7195	0.2561	0.7128	0.2670
0.7061	0.2773	0.6995	0.2909	0.6928	0.3030	0.6861	0.3163
0.6794	0.3300	0.6727	0.3448	0.6661	0.3616	0.6594	0.3769
0.6527	0.3934	0.6460	0.4094	0.6394	0.4260	0.6327	0.4420
0.6260	0.4591	0.6193	0.4771	0.6126	0.4932	0.6060	0.5089
0.5993	0.5242	0.5926	0.5380	0.5859	0.5501	0.5792	0.5625
0.5726	0.5742	0.5659	0.5869	0.5592	0.5992	0.5525	0.6134
0.5459	0.6265	0.5392	0.6415	0.5325	0.6544	0.5258	0.6706
0.5191	0.6881	0.5125	0.7043	0.5058	0.7218	0.4991	0.7410
0.4924	0.7603	0.4857	0.7810	0.4791	0.7992	0.4724	0.8158
0.4657	0.8302	0.4590	0.8473	0.4523	0.8599	0.4457	0.8758
0.4390	0.8880	0.4323	0.9029	0.4256	0.9159	0.4190	0.9293
0.4123	0.9420	0.4056	0.9549	0.3989	0.9647	0.3922	0.9721
0.3856	0.9783	0.3789	0.9827	0.3722	0.9867	0.3655	0.9905
0.3588	0.9949	0.3522	0.9977	0.3455	0.9991	0.3388	1.0000

Table 6.2: Generating Unit Capacity and Reliability Data for the RBTS

Unit No.	Capacity (MW)	FOR	MTTF (hours)	MTTR (hours)
1	40.0 (thermal)	0.030	1460	45
2	40.0 (thermal)	0.030	1460	45
3	10.0 (thermal)	0.020	2190	45
4	20.0 (thermal)	0.025	1752	45
5	5.0 (hydro)	0.010	4380	45
6	5.0 (hydro)	0.010	4380	45
7	40.0 (hydro)	0.020	2920	60
8	20.0 (hydro)	0.015	3650	55
9	20.0 (hydro)	0.015	3650	55
10	20.0 (hydro)	0.015	3650	55
11	20.0 (hydro)	0.015	3650	55

Table 6.3: Additional Generating Units for the RBTS

Capacity (MW)	FOR	MTTF (hours)	MTTR (hours)
10	0.120	550	75

6.4. Cost of System Adequacy

The cost of system adequacy or simply the system cost associated with constructing a generating system for any specified level of reliability can be evaluated relatively easily. The total system cost is made up of all the costs incurred by the utility in providing the consumers with electricity at a specified service reliability and does not include the cost of unserved energy. The two major components of the total system cost are the variable costs and fixed costs. The variable costs include operating costs and fuel costs. The operating cost which is relatively small includes payment for materials, supplies, power etc. The majority of the variable cost is the fuel cost, i.e. costs directly associated with energy production. The fixed costs are made up of the annual charges associated with the equipment regardless of whether or not it is operating. These charges are independent of the degree of usage; and they comprise primarily of interest, depreciation, rent, taxes, insurance and any other capital investment [29]. Table 6.4 shows the RBTS generating unit cost data and the priority loading order.

Table 6.4: Priority Loading Order and Generating Unit Cost Data for the RBTS

Priority Order	Rated Capacity (MW)	Fixed Costs ($/year)	Variable Costs ($/MWh)
1	40.0 (hydro)	100,000	0.50
2	20.0 (hydro)	50,000	0.50
3	20.0 (hydro)	50,000	0.50
4	20.0 (hydro)	50,000	0.50
5	20.0 (hydro)	50,000	0.50
6	5.0 (hydro)	12,500	0.50
7	5.0 (hydro)	12,500	0.50
8	40.0 (thermal)	790,000	12.00
9	40.0 (thermal)	790,000	12.00
10	20.0 (thermal)	680,000	12.25
11	10.0 (thermal)	600,000	12.50
Total		3,185,000	

6.4.1. System Fixed Cost

The unit fixed cost in $/yr can be obtained by multiplying the value given in $/KW/yr by the unit size in KW. The total system fixed cost is the sum of all such costs associated with the system generation. The annual system fixed cost of the base RBTS is $3,185,000. This cost value is independent of the loading order and reliability data of the units.

6.4.2. System Production Cost

The variable cost of a unit is the sum of the unit fuel and operating costs. The product of the unit variable cost and the expected energy supplied (EES) by the unit is the unit energy production cost expressed in "$". The summation of all committed unit production costs gives the system energy production cost. In this study, the load modification technique [46, 47, 48, 49] was utilized to evaluate each unit's EES and the system production cost.

6.4.2.1. Load Modification Technique

The load modification (LM) method is a unified probability technique which provides two important outcomes, namely generating reliability indices and the cost of energy production. The reliability indices include loss of load expectation (LOLE), energy index of reliability (EIR) [1], expected energy not supplied (EENS) of the system and the expected energy supplied (EES) by each unit. The method is essentially a sequential process of modifying a system load duration curve (LDC) with the capacity distribution of all committed generating units to give an equivalent load model [48]. The concept is to determine how the system load appears to the remainder of the system capacity when a given generating unit is committed to satisfy the demand. A prerequisite for this method is information on the priority loading order of the generating units.

The area under the original unaltered LDC is the expected load energy required by the system. The area under any capacity-modified LDC is the expected energy not supplied (EENS) by the system composed of all the generating units contributing to the modification process. The difference in area before and after a unit is added therefore is the expected energy output of that unit. The area under the equivalent LDC modified by the last unit on the priority list provides the various system adequacy indices such as the LOLE, LOEE and EIR. The formulation of these system reliability indices using the LM approach are given in [47] , [48] and [49].

The total energy demand in the RBTS is 992955.9 MWh/yr. The results obtained from the LM technique are presented in Table 6.5. Column 3 of the table shows the expected energy output by each unit with reference to its position on the given priority

Table 6.5: Unit Expected Energy Output and Energy Cost for the RBTS

Rate Capacity (MW)	Variable Energy Cost ($/MWh)	Expected Energy Output (MWh)	Expected Energy Cost ($)
40.0	0.50	340603.56	170301.78
20.0	0.50	173783.09	86891.55
20.0	0.50	167695.56	83847.78
20.0	0.50	134978.63	67489.32
20.0	0.50	94157.96	47078.98
5.0	0.50	17077.04	8538.52
5.0	0.50	14533.32	7266.66
40.0	12.00	46993.93	563927.16
40.0	12.00	2975.10	35701.20
20.0	12.25	130.01	1592.62
10.0	12.50	17.96	244.50
Total		992946.13	1,072,860.13

loading order. The individual unit energy production costs are obtained by multiplying the EES values by their corresponding variable costs. The sum of these individual costs gives a RBTS production cost of $1,072,860.

The total cost of system adequacy is the sum of the system fixed cost and the system production cost. The base RBTS therefore requires a system cost of $4,257,860 to satisfy the load demand of 185 MW using the loading order listed in Table 6.4. The total energy supplied by the base RBTS with reference to the given loading order is 992946 MWh and the system EENS, which is the total energy demanded minus the total energy supplied, is 9.77 MWh for a period of one year. The LOEE of the system is therefore 9.77 MWh/yr. The LOLE and EIR of the system are 1.0875 hrs/yr and 0.9999902 respectively.

6.5. Customer Cost of Unserved Energy

The benefit or worth of power system service reliability can be quantified in terms of a cost associated with generating capacity inadequacy. This cost is often referred to as the customer cost of unserved energy C, and can be expressed as:

$$C = IEAR \cdot EENS. \tag{6.1}$$

The IEAR factor provides a single numerical customer cost value which can be used in conjunction with the predicted energy not supplied (EENS) to link interruption costs to system adequacy. This energy method of quantifying the reliability worth assumes that the value of C increases linearly with energy curtailment due to load-supply deficiency. Numerical examples are given in this chapter using the RBTS and the 1991 outage cost values.

The EENS of the RBTS has a value of 9.77 MWh at a load demand of 185 MW. The IEAR can have two possible values depending on the cost modeling method utilized in the Monte Carlo simulation. The studies presented in Chapter 5 suggest that 4.46 $/KWh and 14.57 $/KWh are acceptable estimated IEAR values obtained from the CCDF and the distribution models respectively. Using the 4.46 $/KWh IEAR, the customer cost is $43,574. The customer cost increases dramatically to $142,349 when the IEAR becomes 14.57 $/KWh.

6.6. Optimum Planning Reserve Margin for the RBTS

The ultimate objective in using the cost / benefit method in HLI system planning is to establish an optimum adequacy level which satisfy both the reliability and economic constraints of the system. This planning process aims at determining an optimum reliability level by minimizing the total societal costs of electric power. The graphical illustration of this concept is shown in Figure 6.1. The system adequacy level is commonly measured in terms of the percentage planning reserve margin of the generating system. A reserve margin can be defined as the additional generating capacity above the demand. The amount of reserve must be planned in advance in order to safeguard against equipment failures, non-scheduled outages and excessive load growth [1]. The percentage reserve margin denoted in this thesis as PRM can be expressed as:

$$PRM = \frac{Total\ Generating\ Capacity - Peak\ Demand}{Peak\ Demand}. \tag{6.2}$$

The base RBTS has a total capacity of 240 MW which corresponds to a PRM of 29.73% at a peak load of 185 MW. Section 6.4 notes that this system configuration has a fixed cost of $3,185,000 and a production cost of $1,072,860. The total utility cost at the 29.73% reserve margin therefore is $4,257,860. The corresponding customer interruption costs are $43,574 and $142,349 determined from the 1991 CCDF and the distribution models respectively. Using the CCDF model, the total societal cost is $4,301,434. The total cost increases to $4,400,209 when the distribution model is used.

In order to determine the optimum reserve margin for the RBTS, it is necessary to examine the effects on the total societal cost of varying the reserve margin. Additional generating units can be added one at a time to the base RBTS to gradually increase the margin. The evaluation of fixed cost, production cost, customer interruption cost and the total societal cost is then repeated for each unit addition. The generating system for the RBTS is quite reliable and therefore the analysis starts at a lower reserve margin in order to clearly bring out the concepts illustrated in Figure 6.1.

The RBTS was modified by removing two of the original 20-MW hydro units. These hydro generators were then considered as two additional generating units and added prior to adding the gas turbines. After establishing the schedule for commissioning additional units, the associated fixed and production costs were determined. The estimated customer costs of unserved energy were then calculated by multiplying the IEAR by the expected unserved energy. The results are shown in Figure 6.2. It can be seen that the total cost curve exhibits a least-cost planning margin in the neighbourhood of 30% which corresponds to a total societal cost of $6.4 million per year. This total cost is composed of $3.2 million fixed cost, $3.2 million production cost and $43,574 customer cost obtained from the 1991 CCDF.

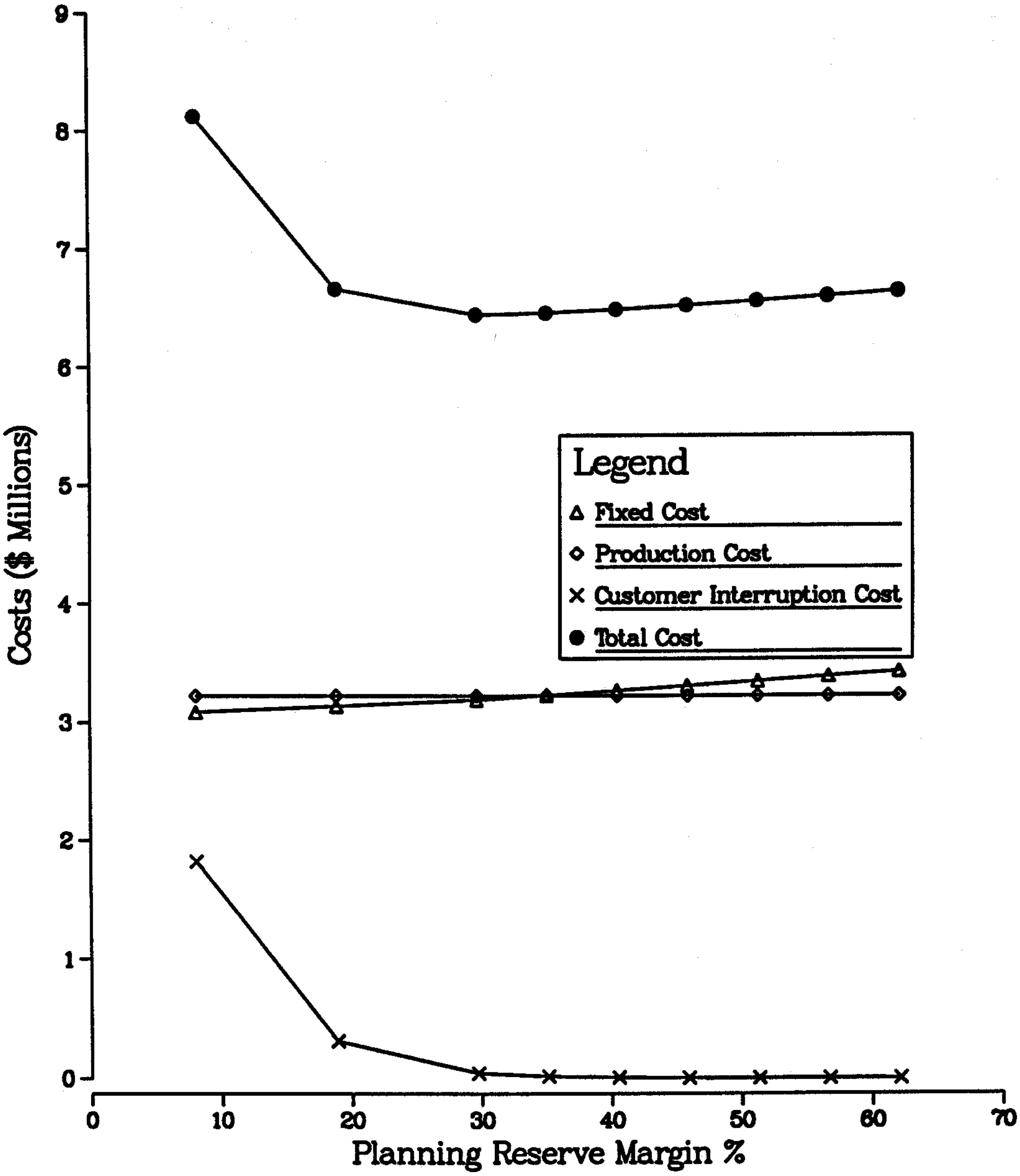


Figure 6.2: Variation of Costs that Society faces directly with the Planning Reserve Margin : RBTS (using 1991 customer damage function)

6.6.1. Effects of the IEAR Value On the Optimum Reserve Margin

The cost functions given in Figure 6.2 show that the total societal cost is critically dependent on the customer cost of unserved energy. Changes in customer cost can have great impacts regarding the position of the optimum PRM. Given that all system operating and reliability data remains unchanged, the consumer cost varies only if the IEAR value changes. In order to illustrate the effects on optimum PRM of different IEAR's, selected values in the range of 4.46 $/KWh to 14.57 $/KWh were used. A graphical representation of these results is shown in Figure 6.3. The base case IEAR for the RBTS is 4.46 $/KWh. If the IEAR value decreases to 2.23 $/KWh, the optimum PRM stays at 30% as in the base case. Increasing the IEAR value to 8.92 $/KWh increases the optimum point from 30 to 35%. If the IEAR value is further increased to 14.57 $/KWh, which corresponds to the IEAR estimated from the cost distribution model, the optimum reserve margin remains at 35%. These results suggest that the optimum reserve margin of the RBTS is sensitive to changes in the system IEAR. Using 10-MW unit additions, the optimum PRM increases by 20% when the IEAR value is doubled. The PRM remains at the 35% point when the IEAR is increased by 327%. It should be appreciated that the PRM will increase further if smaller units were used.

6.7. Capacity Expansion Analysis in the RBTS

A system peak load of 185 MW was used in the evaluation of the optimum planning reserve margin presented in the previous section. The load demand at a given time in the future however is generally expected to increase due to higher electricity usage. In order to meet this future load growth, new generating stations must be planned and constructed. The amount of time required to design, construct and commission such stations can be quite extensive (typically 5 to 10 years) depending on the environmental and regulatory requirements [42]. It is therefore essential to determine the future capacity expansion considerably in advance of the actual unit commitment date.

The concept of incorporating risk evaluation in capacity expansion analysis is illustrated using the RBTS. The basic objective is to determine how much additional capacity is required to meet a future forecast load. The problem can be analyzed using either the cost minimization approach or the constant risk technique.

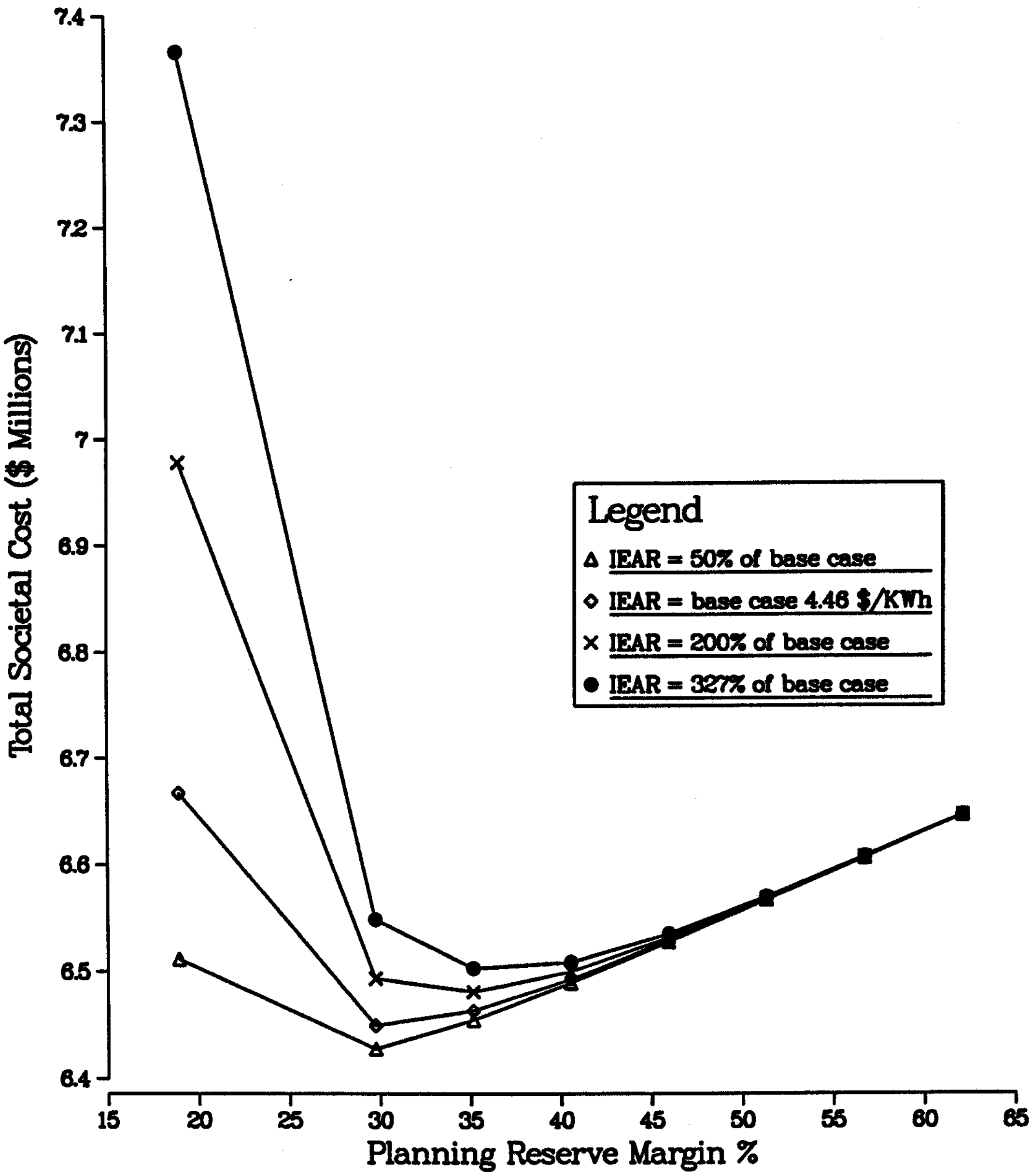


Figure 6.3: Effect of IEAR upon the Optimum Reserve Margin of the RBTS

6.7.1. Cost Minimization Approach

The cost minimization approach uses the cost / benefit technique to decide the number of additional units required to meet a given load level. It was noted in Section 6.6 that the RBTS is quite reliable at a peak load of 185 MW and no unit additions are required to achieve an optimum reserve margin of 30%. A capacity expansion study of the RBTS was performed by repeating the cost / benefit analysis at increasing peak load values and the results are presented in Figure 6.4. The same information is summarized in Table 6.6. The LOEE and LOLE indices are also shown in the same table.

Table 6.6: Capacity Expansion Study of the RBTS using a Cost Minimization Approach (IEAR = 4.46 $/KWh)

Peak Load (MW)	Units* Added	Reserve Margin (%)	LOEE (MWh/yr)	LOLE (Hrs/yr)	Total Cost (Million $)
185.0	0	29.73	9.77	1.087508	4.3014
190.0	1	31.58	6.41	0.716025	4.4682
195.0	1	28.21	10.39	1.143552	4.6397
200.0	2	30.00	6.87	0.756078	4.8311
205.0	2	26.83	11.01	1.184332	5.0294
210.0	3	28.57	7.35	0.780065	5.2451
215.0	3	25.58	11.67	1.214827	5.4682
220.0	4	27.27	7.86	0.810708	5.7056
225.0	4	24.44	12.39	1.248527	5.9490
230.0	5	26.09	8.41	0.856117	6.2031
235.0	5	23.40	13.19	1.321891	6.4652
240.0	6	25.00	9.00	0.915147	6.7372

* Number of units added to the 240-MW base RBTS

When the annual peak load increases by 5 MW, the minimum cost point shifts from 30 to 32%. This PRM corresponds to the addition of one 10-MW unit to the existing system at a total cost of $4.4682 millions. When the peak load is further increased to 195 MW, the optimum PRM shifts to 28%. The number of additional units remains at one.

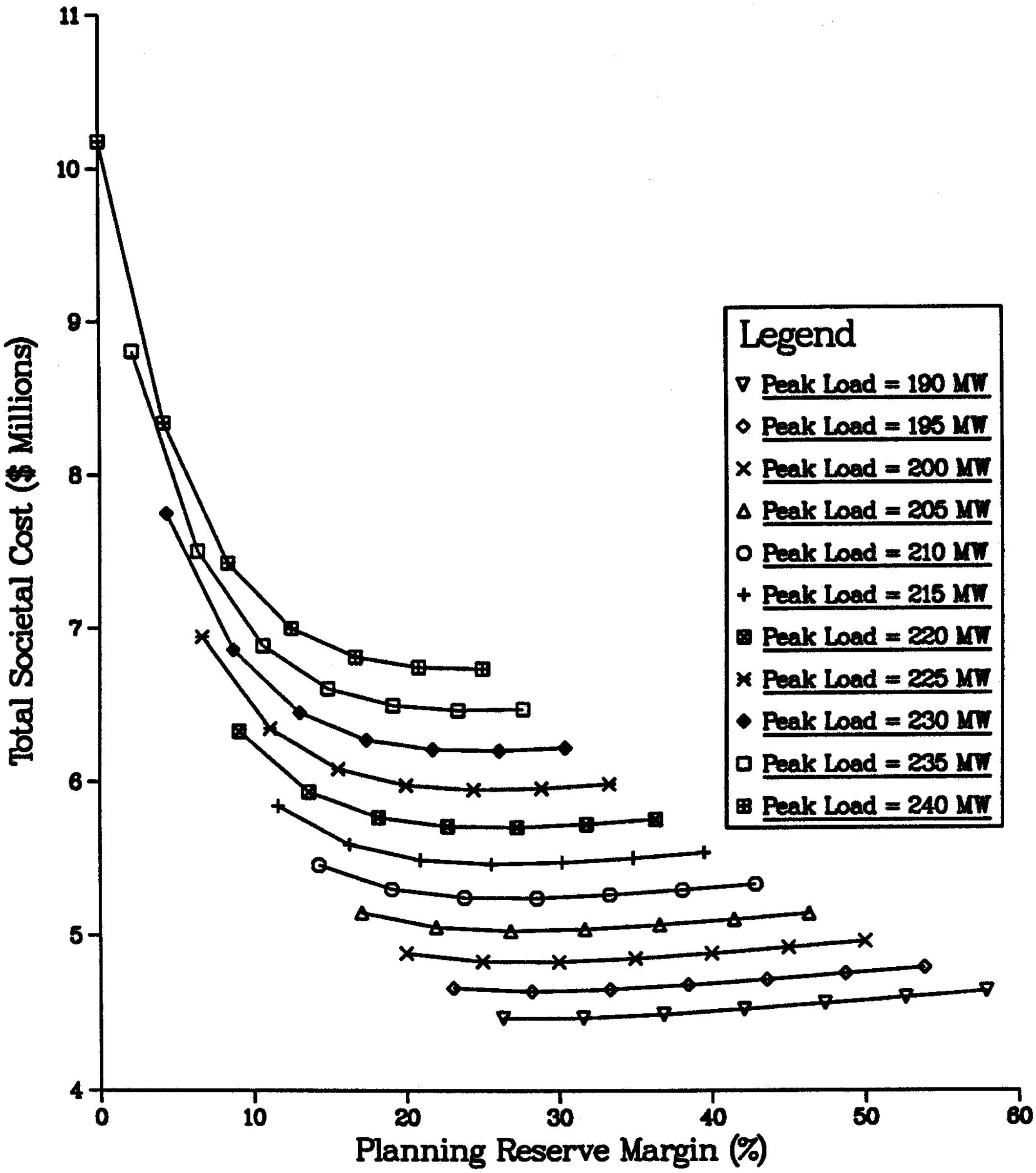


Figure 6.4: System Expansion Study of the RBTS using a Minimum Cost Criterion (IEAR = 4.46 $/KWh)

6.7.2. Constant Risk Approach

The constant risk approach is commonly used by utility planners in studies of future capacity expansion. This approach uses a preselected risk level, which is believed to be adequate for the given system. In this particular study, the risk criterion (LOEE) for the RBTS is 9.77 MWh/yr given the assumption that an installed capacity of 240 MW is adequate for a system peak load of 185 MW. The results of this expansion study are portrayed in Figure 6.5 and also summarized in Table 6.7.

Table 6.7: Capacity Expansion Study of the RBTS using a Constant Risk Approach

Peak Load (MW)	Units* Added	Reserve Margin (%)	Total Cost (Million $)	LOLE (Hrs/yr)	LOEE (MWh/yr)
185.0	0	29.73	4.3014	1.087548	9.77
190.0	1	31.58	4.4682	0.716025	6.41
195.0	2	33.33	4.6523	0.489406	4.23
200.0	2	30.00	4.8311	0.756079	6.87
205.0	3	31.71	5.0407	0.515520	4.55
210.0	3	28.57	5.2451	0.780065	7.35
215.0	4	30.23	5.4781	0.535348	4.90
220.0	4	27.27	5.7056	0.810707	7.86
225.0	5	28.89	5.9574	0.556076	5.28
230.0	5	26.09	6.2031	0.856116	8.41
235.0	6	27.66	6.4719	0.596573	5.70

* Number of units added to the 240-MW base RBTS

Unit additions are required when the risk level exceeds 9.77 MWh/yr. For example, a load growth of 5 MW (annual peak demand increases to 190 MW) requires the addition of one 10-MW unit to maintain the risk below the 9.77 MWh/yr decision line. This 12-unit system has a LOEE of 6.41 MWh/yr. Similarly, when the peak load increases by another 5 MW (annual peak demand of 195 MW), the system calls for another 10-MW unit addition which results in a LOEE of 4.23 MWh/yr. The new system now consists of

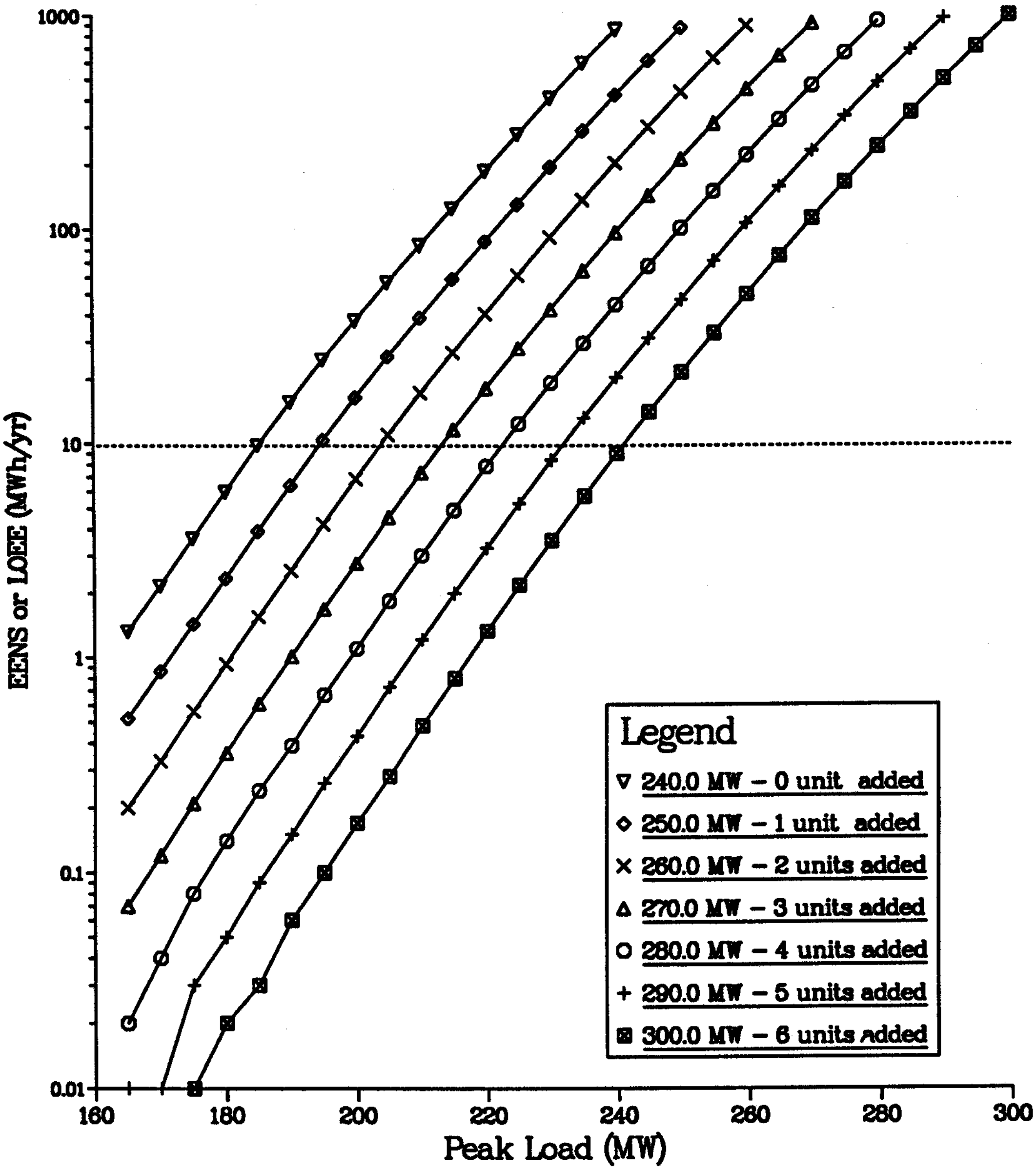


Figure 6.5: System Expansion Study of the RBTS using a Constant Risk Criterion

the 11 base units and two 10-MW additional units. If the peak load increases to 200 MW, further unit additions are not necessary since the new system has a LOEE of 6.87 MWh/yr which is less than the critical value of 9.77 MWh/yr.

6.8. Summary

This chapter illustrates the role of customer interruption costs in an HLI adequacy assessment. The net societal cost associated with a given reliability level is the summation of system investment cost and customer interruption cost evaluated at that level. The cost / benefit approach provides a logical procedure for arriving at the optimum adequacy level by minimizing the total societal cost.

Sections 6.4 and 6.5 in this chapter illustrate the procedures for evaluating the system and customer costs of the RBTS respectively. There are two components of system cost, namely the fixed cost and the production cost. The planning reserve margin is 29.7% for the base RBTS with a total capacity of 240 MW and a peak load of 185 MW. The fixed and production costs for this system are $3,185,000 and $1,702,860 respectively. The resultant total RBTS cost is $4,257,860. The customer cost determined using the 1991 CCDF is $43,574. The total societal cost at the 29.7% reserve margin is therefore $4,301,434.

In order to determine the optimum PRM for the RBTS, the total societal cost associated with different planning reserve margins were evaluated and compared. The procedure is clearly illustrated in Section 6.6. The optimum reserve margin for the stated system configuration and loading order is approximately 30% with an EIR of 0.9999902. A sensitivity study of the optimum PRM shows that for the RBTS, this value is sensitive to changes in system IEAR.

The cost / benefit approach was also utilized to conduct capacity expansion analysis for the RBTS. The procedure is illustrated in Section 6.7. Assuming the system peak load grows at a 5 MW increment each year, the cost / benefit method was utilized to determine the number of units required at each load level to achieve a new optimum PRM. For example, a load growth of 10 MW requires one 10-MW unit to be added to the existing 240-MW base RBTS in order to give an optimum PRM of 28.2%. A capacity expansion study was also conducted using the more traditional constant risk approach. A preselected LOEE index of 9.77 MWh/yr was used as the critical value to determine whether or not unit additions are required under a given load condition. This alternate approach indicates that a load growth of 10 MW requires two 10-MW units to be added to the existing 240-MW base RBTS.

Several electric power utilities in Canada presently use a cost minimization approach in their capacity planning [7]. The remaining utilities use the constant risk technique. There are no clear rules regarding the selection of a particular method as it is largely a management decision. In most cases, the cost minimization approach results in a system which has lower cost and higher risk than the constant risk approach provides. For example, for an increase in peak load demand in the RBTS from 185 to 195 MW, the cost / benefit method provides a cost effective design with one 10-MW unit addition. The design results in a total cost of $4.6397 millions and a LOEE of 10.39 MWh/yr. In comparison, the same load growth requires two additional units when the constant risk criterion is used. The corresponding total cost and LOEE are $4.6523 millions and 4.23 MWh/yr respectively.

7. SUMMARY AND CONCLUSIONS

This research work extends the state of the art in reliability worth assessment by developing an alternate technique to the Composite Customer Damage Function (CCDF) approach to describe the customer costs due to electric supply interruptions. The new technique is designated, in this thesis, as the probability distribution approach. The primary objective of the research work described in this thesis was to illustrate this technique using the 1991 NSERC cost data. The thesis also shows how to estimate an Interrupted Energy Assessment Rate (IEAR) using the developed cost distributions.

Chapter 2 of this thesis briefly addresses the concepts associated with power system reliability worth assessment. In this approach, reliability worth is quantified using customer costs of interruption as a substitute. An overview of the methodologies used to estimate the interruption costs is presented together with a very brief discussion of the merits and demerits of each. The customer survey approach, although often expensive, appears to be the method favoured by power utility companies. This method was utilized by the Power Systems Research Group at the University of Saskatchewan to conduct the 1991 NSERC cost of interruption study. The data (both cost and non-cost related) screening process designed particularly for this study is described in some detail. This chapter also illustrates the procedures used to generate a conventional CCDF using a hypothetical test system (RBTS).

The conventional composite customer damage function describes the overall average interruption cost as a function of duration for the given studied area. Cost estimates collected from survey respondents, however, display a large degree of variability which indicates a need for a better representation than one containing only the mean values. Chapter 3 focuses on the development of a new cost modeling technique known as the probability distribution method. This new method is capable of recognizing the dispersed nature of the outage data and provides the ability to incorporate this behaviour in a wide range of applications. Due to the distinct advantages associated with normally distributed data, it was decided to attempt to fit the 1991 interruption data to normal distributions. The two basic steps utilized in this approach are (i) transforming the data into symmetrical distributions using the Box and Cox power transformation, and (ii) choosing the symmetrical distribution which best fits a normal curve using a hypothesis test. The primary statistics involved in these procedures are the standardized third and fourth moments. The developed technique was used to analyze and model the dispersed data at all the studied durations. Regression analysis was then utilized to determine the distribution models at non-studied or intermediate durations. The objective of the regression analysis was to determine the relationship between the studied interruption duration and each of the four distribution parameters (P_z, λ, μ and σ^2). Intermediate duration parameters can then be calculated from the respective equations.

Chapter 4 illustrates the generation of a conventional CCDF and the development of outage cost probability distribution models using the 1991 NSERC data. The various new CDF's were generated using the procedures developed in Chapter 2. A set of best normal curves describing the commercial $/KW responses were created using the technique described in Chapter 3. A three-dimensional (3-D) customer damage function was used to represent these sector cost curves as a function of duration. It can be seen from this chapter that more time and effort is required to generate a 3-D SCDF compared to that required by a 2-D representation.

Chapter 5 illustrates the procedures used to convert an interruption cost model into a customer-driven cost factor known as the interrupted energy assessment rate. A Monte Carlo simulation approach was utilized to calculate the system IEAR of the RBTS. The three elements in an IEAR evaluation are the customer interruption cost, the energy not supplied and the duration of every simulated load loss event. The latter two elements were obtained from the simulated system operating history. The IEAR cost factor was estimated using both the CCDF model and the probability approach. The CCDF model is relatively easy to use due to the assumed linear relationship between the cost and the interruption duration. The evaluation process is more complex using cost distribution models. In this process, random numbers are used to determine the customer costs associated with a simulated outage. IEAR values of 4.46 and 14.57 $/KWh were obtained for the CCDF and probability distribution approaches respectively. The difference between these values clearly illustrates the effect of recognizing the dispersed nature of interruption cost data rather than using only single point average values.

Chapter 6 illustrates the application of the derived IEAR values in HLI capacity planning and expansion studies. The technique used in this application is known as the cost / benefit approach. This chapter briefly illustrates the creation of a system adequacy target using a cost / benefit method in which the sum of the system and customer interruption costs is minimized. A major potential application domain for cost / benefit analysis is in system expansion assessment. This thesis illustrates how future generating capacity requirements can be determined using a process which recognizes both the utility costs and customer worth. The customer cost data representations developed in this research work can be used in a wide range of studies in each electric power system functional zone and hierarchical level.

This thesis compares two conceptually different techniques for modeling customer interruption costs. The conventional CCDF approach uses expected values to define the overall monetary losses incurred by electrical users due to power failures. This method is relatively easy to develop and to use in an evaluation of reliability worth. The basic two dimensional CCDF cannot reflect the dispersed nature of customer interruption costs and therefore provides a limited interpretation of the entire customer outage cost data base. This thesis presents a new approach designated as the probability distribution technique which provides a three dimensional representation of the sector customer outage costs. The thesis also shows how to estimate a utility service area IEAR using the distribution approach. The results obtained for the test system utilizing a Monte Carlo simulation approach are 4.46 $/KWh and 14.57 $/KWh for the CCDF and the distribution methods respectively. These values are significantly different and suggest that the utilization of the basic CCDF approach in reliability worth evaluation may significantly undervalue power system reliability worth. This observation needs to be tested in practical system applications.

REFERENCES

1. Billinton, R. and Allan, R.N., *Reliability Evaluation of Power Systems,* Plenum Publishing, New York, 1984.

2. Billinton, R. and Allan, R. N., ''Basic Power System Reliability Concepts'', *Reliability Engineering and System Safety, Vol. 27,* 1990, pp. 365-384.

3. Billinton, R., ''Bibliography on the Application of Probability Methods in Power System Reliability Evaluation'', *IEEE Transactions on Power System, Vol. PAS-91,* March/April 1972, pp. 649-660.

4. IEEE Committee Report (1971-1977), ''Bibliography on the Application of Probability Methods in Power System Reliability Evaluation'', *IEEE Transactions on Power System, Vol. PAS-97,* 1978, pp. 2235-2245.

5. Billinton, R., Lee, S. H. and Allan, R. N., ''Bibliography on the Application of Probability Methods in Power System Reliability Evaluation'', *IEEE Transactions on Power System, Vol. PAS-103,* February 1984, pp. 275-282.

6. Billinton, R., Shahidehpour, S. M., Singh, C. and Allan, R. N. (1977-1982), ''Bibliography on the Application of Probability Methods in Power System Reliability Evaluation'', *IEEE Paper No. 88 WM 172-9 PWRS,New York,* February 1988.

7. Billinton, R., ''Criteria used by Canadian Utilities in the Planning and Operation of Generating Capacity'', *IEEE Power Engineering Society Winter Meeting Paper No. 88 WM 150-5, New York,* February 1988.

8. Billinton, R., Wacker, G. and Wojczynski, E., ''Comprehensive Bibliography on Electrical Service Interruption Costs'', *IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No. 6,* June 1983, pp. 1831-1837.

9. Tollefson, G., Billinton, R. and Wacker, G., ''Comprehensive Bibliography on Reliability Worth and Electrical Service Consumer Interruption Costs: 1980 - 1990'', *IEEE Transaction on Power Systems, Vol. 6 No. 4,* Nov. 1991, pp. 1508-1514.

10. Billinton, R., Tollefson, G. and Wacker, G., ''Assessment of Electric Service Reliability Worth'', *IEE Conference Proceedings No. 338,* July 1991, pp. 9-14.

11. Sanghvi, A. P. (RCG/Hagler, Bailly, Inc.), ''Cost Benefit Analysis of Power System Reliability: Determination of Interruption Costs -- Measurement Methods and Potential Applications in Reliability Cost-Benefit Analysis'', *EPRI EL-6791, Vol. 1, Final Report,* April 1990.

12. Billinton, R., Wacker, G. and Wojczynski, E., *Customer Damage Resulting from Electric Service Interruptions,* Canadian Electrical Association, R&D Project 907U 131 Report, 1982.

13. Billinton, R., Wacker, G. and Brewer, R., *Farm Losses Resulting From Electrical Service Interruptions,* Canadian Electrical Association, R&D Project 309U 131 Report, 1987.

14. Tollefson, G., Billinton, R., Wacker, G., Chan, E., and Aweya, J., "A Canadian Customer Survey to Assess Power System Reliability Worth", *IEEE Paper No. 93 WM 175-0 PWRS*, 1993.

15. Wacker, G., Billinton, R. and Subramanian, R.N., "Using Cost of Electric Service Interruptions Survey in Determination of a Composite Customer Damage Function", *Proceedings of the International Association of Science and Technology for Development, IASTED, Energy Symposia, San Francisco, Calif.*, June 4-6 1984.

16. Subramaniam, R. K., "Development of a Composite Customer Damage Function", Master's thesis, University of Saskatchewan, April 1985.

17. Billinton, R., Oteng-Adjei, J. and Ghajar, R., "Comparison of Two Alternate Methods to Establish an Interrupted Energy Assessment Rate", *IEEE Transactions on Power Apparatus and Systems, Vol. PWRS-2*, August 1987, pp. 751-757.

18. Billinton, R., Oteng-Adjei, J., "Cost/Benefit Approach to establish Optimum Adequacy Level for Generating System Planning", *IEE Proceedings, Vol. 135, Part C, No. 2*, March 1988, pp. 81-87.

19. Levy, R. D. and Sanghvi, A. P., "Value-Based Utility Planning: Scoping Study", *ERPI Paper EM-4389, Final Report*, 1986.

20. Shipley, R. B., "Power Reliability Cost vs. Worth", *IEEE Transactions on Power Apparatus and Systems*, 1972, pp. 2204-2212.

21. Corwin, J. and Miles, W., "Impact Assessment of the 1977 New York City Blackout", *U. S. Department of Energy, Washington D. C.*, July 1977.

22. Wojczynski, E., Billinton, R. and Wacker, G., "Interruption Cost Methodology and Results-A Canadian Commercial and Small Industrial Survey", *IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103, No. 2*, Feb. 1984, pp. 437-444.

23. Wacker, G. and Billinton, R., "Customer Cost of Electric Interruptions", *IEEE Proceedings, Vol. 77, No. 6*, 1989.

24. Wacker, G., Tollefson, G. and Billinton, R., "A Canadian Cost of Interruption Study", *accepted for presentation to the Athens Power Tech Conference, Athens, Greece*, Sept. 1993.

25. Wacker, G., Tollefson, G. and Billinton, R., "Understanding the Customers' Role in Reliability Worth Assessment Using Customer Surveys", *18th Inter-RAM Conference, Philadelphia*, August 1992.

26. Billinton, R., Wacker, G., Tollefson, G., ''Assessment of Reliability Worth of Electric Power Systems : Final Report'', *Final Report for NSERC Strategic Project STR0045005*, 1993.

27. SPSS Inc., *SPSS-X Data Entry,* 1989.

28. Kos, P., ''Cost of Electricity Supply Interruptions to Agricultural Sector Customers'', Master's thesis, University of Saskatchewan, Oct. 1989.

29. Billinton, R., Oteng-Adjei, J., Kumer, S., Chowdhury, N., Chu, K., Debnath, K., Goel, L., Khan, E., Kos, P. and Nourbakhsh, G., ''A Reliability Test System for Educational Purposes - Base Data'', *IEEE Transactions on Power System, Vol.4, No.3*, August 1989, pp. 1238-1244.

30. Johnson, R., *Elementary Statistics,* Prindle Weber & Schmidt, Boston, 1984.

31. D'Agostino and Stephens, M. A., *Goodness-of-Fit Techniques,* Marcel Dekker, Inc., New York, 1986.

32. SPSS Inc., *SPSS Base System User's Guide,* 1990.

33. Box, G. E. P. and Cox, D. R., ''An Analysis of Transformations'', *J. R. Statist. Soc., Vol.26, No. B*1964, pp. 211-252.

34. Wetherill, G. B., *Regression Analysis with Applications,* Chapman and Hall, London, 1986.

35. Bethea, R. M., Duran, B. S. and Boullion, T. L., *Statistical Methods for Engineers and Scientists,* Marcel Dekker, Inc., New York and Basel, 1985.

36. Kleinbaum, D. G. and Kupper L. L., *Applied Regression Analysis and Other Multivariable Methods,* Duxbury Press, North Scituate, Massachusetts, 1978.

37. Pearson, E.S., D'Agostino, R. B. and Bowman, K. O., ''Tests for Departure from Normality. Comparison of Powers'', *Biometrika 64*, 1977, pp. 231-246.

38. SAS Institute Inc., *SAS/GRAPH User's Guide,* 1990.

39. Noferi, P.L., Paris, L. and Salvaderi, L., ''Monte Carlo Methods for Power System Evaluation in Transmission or Generation Planning'', *Proceedings 1975 Annual Reliability and Maintainability Symposium, Washington*, 1975.

40. ''Modeling of Unit Operation Considerations in Generating Capacity Reliability Evaluation. Volume 1: Mathematical Models, Computing Methods, and Results'', *Report EL-2519, Electric Power Research Institute, Palo Alto, Ca.*, July 1982.

41. Ghajar, R., ''Utilization of Monte Carlo Simulation in Generating Capacity Planning'', Master's thesis, University of Saskatchewan, Saskatoon, Canada, September 1986.

42. Billinton, R. and Allan, R.N., *Reliability Evaluation of Engineering Systems: Concepts and Techniques,* Plenum Publishing, New York, 1983.

43. Gan, L., ''Multi-Area Generating System Adequacy Assessment by Monte Carlo Simulation'', Master's thesis, University of Saskatchewan, Saskatoon, Canada, March 1991.

44. IEEE Task Force, ''IEEE Reliability Test System'', *IEEE Transaction, PAS-98*, December 1979, pp. 2047-2054.

45. Ruiu, D., Ye, C., Billinton, R. and Lakhanpal, D., ''Reliability Criteria Selection for Integrated Resource Planning'', *Canadian Electrical Association, Power System Reliability Subsection, Power System Planning and Operation Section, Engineering and Operation Division*, March 1993, Montreal.

46. Billinton, R. and Harrington, P.G., ''Reliability Evaluation in Energy Limited Generating Capacity Studies'', *IEEE Transactions on Power Apparatus and Systems, Vol. PAS-97, No. 6.*, Noc/Dec 1978, pp. 2076-2085.

47. Billinton, R., Cheung, L.C.H., ''An Energy Based Approach to Generating Capacity Reliability Evaluation'', *IEE Conference on Reliability of Power Supply Systems, Conference Publication 225*, September 1983.

48. Billinton, R., Cheung, L.C.H., ''Load Modification : A Unified Approach for Generating-Capacity Reliability Evaluation and Production-Cost Modelling'', *IEE Proceedings, Vol. 134, Part C, No. 4*, July 1987, pp. 273-280.

49. Billinton, R. and Oteng-Adjei, J., ''Modelling of Energy Limited Units in Generating Capacity Reliability of Small Hydro Systems'', *Small Hydro 88 Conference, Toronto, Canada*, July 1988.

A. CRITERIA FOR SELECTING AND CODING USABLE COMMERCIAL SECTOR RESPONSES

The following criteria pertain to the questionnaires developed and used by the Power Systems Research Group. A complete set of questionnaires are contained in [26].

Q1.1

1. The answer should be numeric
2. If a range of values is given, the mid-point should be selected. If this mid-point value is not a whole number, it should be rounded to the nearest whole number.
3. If no value is given, the answer should be considered a missing value and the computer field left blank

Q2

1. Only one box should be checked on each line.
2. If two or more non-adjacent boxes are checked, the answer should be rejected
3. If no box is checked, the answer should be considered a missing value and the computer field left blank
4. If both 1 and 2 are checked, then 1 should be selected
5. If adjacent boxes are both selected, alternate between higher and lower for selection

Q3

1. Two variables have been created for this question: type and duration
2. Only two computer fields are available for each variable. The first "type" should coincide with the first "duration"
2. If the respondent has indicated that any particular type of generator will operate indefinitely, the duration should be coded as "999"

Q4(a)

1. Only one box should be checked.
2. If two or more non-adjacent boxes are checked, the answer should be rejected
3. If no box is checked, the answer should be considered a missing value and the computer field left blank
4. If adjacent boxes are both selected, alternate between higher and lower for selection

Q4(b)

1. Only one box should be checked.
2. If two or more boxes are checked, the answer should be rejected
3. If no box is checked, the answer should be considered a missing value and the computer field left blank
4. An additional computer field has been added when OTHER has been checked, the duration should be entered in hours

Q5(a) WORST MONTH: eleven variables and computer fields have been created for this question

1. Each answer will be given a numerical value, JAN - 1 through to DEC - 12, ALL MONTHS THE SAME - 13
2. One or more boxes can be checked from JAN to DEC or only ALL MONTHS THE SAME. If one or more (but not all) months have been checked and ALL MONTHS THE SAME is also checked, correlate with answers to Q7(a). If Q7 has not been answered, then select ALL MONTHS THE SAME.
3. If all boxes from JAN to DEC have been checked, the answer should be considered as ALL MONTHS THE SAME
4. If the box ALL MONTHS THE SAME has been checked, only "13" should be entered into the first computer field with the other 10 computer fields left blank.
5. If no box has been checked, all the 11 computer fields should remain blank and the answer considered a missing value.

Q5(b) WORST DAY OF THE WEEK: six variables and computer fields have been created for this question

1. Each answer will be given a numerical value, MON - 1 through to SUN - 7, ALL SEVEN DAYS OF THE WEEK THE SAME - 8, ALL WEEKDAYS THE SAME - 9, WORST DAY IS IRREGULAR - 10
2. One or more boxes can be checked from MON to SUN or only ALL SEVEN DAYS THE SAME, or only ALL WEEKDAYS THE SAME, or only WORST DAY IS IRREGULAR. If one or more (but not all) days have been checked and ALL SEVEN DAYS THE SAME is also checked, correlate with answers to Q7(b). If Q7 has not been answered, then select ALL SEVEN DAYS THE SAME.
3. If one or more (but not all) weekdays days have been checked and ALL WEEKDAYS THE SAME is also checked, correlate with answers to Q7(b). If Q7 has not been answered, then select ALL WEEKDAYS THE SAME.
4. If one or more days have been checked and WORST DAY IS IRREGULAR is also checked, correlate with answers to Q7(b). If Q7 has not been answered, then select WORST DAY IS IRREGULAR.
5. If all boxes from MON to SUN have been checked, the answer should be considered as ALL SEVEN DAYS THE SAME
6. If the box ALL SEVEN DAYS THE SAME has been checked, only "8" should be entered into the first computer field with the other 5 computer fields left blank.
7. If no box has been checked, all the 6 computer fields should remain blank and the answer considered a missing value.

Q5(c) WORST TIME OF THE DAY: eight variables and computer fields have been created for this question

1. Each answer will be given a numerical value, EARLY MORNING - 1 through to OVERNIGHT - 9, ALL TIMES OF THE DAY AND NIGHT THE SAME - 10, ALL WORKING HOURS THE SAME - 11, WORST TIME IS IRREGULAR - 12
2. One or more boxes can be checked from EARLY MORNING to OVERNIGHT or only ALL TIMES OF THE DAY AND NIGHT THE SAME, or only ALL WORKING HOURS THE SAME, or only WORST TIME IS IRREGULAR. If one or more (but not all) days have been checked and ALL TIMES OF THE DAY AND NIGHT THE SAME is also checked, correlate with answers to Q7(c). If Q7 has not been answered, then select ALL TIMES OF THE DAY AND NIGHT THE SAME.
3. If one or more (but not all) times have been checked and ALL WORKING HOURS THE SAME is also checked, correlate with answers to Q7(c). If Q7 has not been answered, then select ALL WORKING HOURS THE SAME.

4. If one or more days have been checked and WORST TIME IS IRREGULAR is also checked, correlate with answers to Q7(c). If Q7 has not been answered, then select WORST TIME IS IRREGULAR.
5. If all boxes from EARLY MORNING to OVERNIGHT have been checked, the answer should be considered as ALL TIMES OF THE DAY AND NIGHT THE SAME
6. If the box ALL TIMES OF THE DAY AND NIGHT THE SAME has been checked, only "10" should be entered into the first computer field with the other 10 computer fields left blank.
7. If no box has been checked, all the 8 computer fields should remain blank and the answer considered a missing value.

Q6 Direct Costing Question:

1. The answer to each cell of the table should be numeric
2. If a cell has not been answered, it should be left blank and treated as a missing value.
3. "Zeros" should be entered only if a "zero" has been written in a cell by the respondent
4. If a respondent has indicated "minimal," "none," or something similar for the value in a cell, a "zero" should be entered
4. If the columns have not been totalled the coding person should total each column
5. If the respondent has provided a range for any cell and not totalled the column, the mid-point of the range should be entered for that cell and the mid-point should be used in the TOTAL. The low end of the range should be used to calculate the MIN TOTAL and the upper end of the range should be used to calculate the MAX TOTAL
6. If the respondent has provided a range across a row or a series of cells within a row, the low end of the range shall be entered for the shortest duration in the range and the upper end of the range should be entered in the largest duration. A linear relationship should be calculated for each other cell.

Q7 Variation of Interruption Cost

1. Two variables and computer fields have been created for each line of the question. The first variable is for the seven possible choices of estimates and the the second variable is for the value of the OTHER ESTIMATES
2. Each choice is given a numerical value from left to right: SAME AS WORST DAY - 1 to NEGLIGIBLE - 6, and OTHER ESTIMATES - 7
3. If any of the first 6 boxes have been checked, the corresponding numerical value should be entered into the first computer field with the second computer field left blank.
4. If two adjacent boxes have been checked, the lowest percentage should be selected.
5. If two or more non-adjacent boxes are selected, the answer should be disgared, the computer fields left blank, and considered as missing values.
6. If a line has been left unanswered, the two computer fields should be left blank and considered as missing values.

Q8 Cost Reduction

1. Seven computer fields are created for both parts of this question
2. If YES has been chosen, a numerical value of "1" should be entered in the first computer field. If NO has been chosen, a "2" should be entered.
3. If a NO has been selected, the remaining 6 computer fields [representing the cells of the table] should be left blank even if the respondent has provided values
4. The answer to each cell of the table should be numeric
5. If a cell has not been answered, it should be left blank and treated as a missing value.
6. "Zeros" should be entered only if a "zero" has been written in a cell by the respondent

Q9, Q10, Q12 and Q13

1. Only one box should be checked for each question.
2. If two or more boxes are checked, the answer should be rejected
3. If no box is checked, the answer should be considered a missing value and the computer field left blank

Q11

1. The answer for each part should be numeric
2. If a range of values is given, the mid-point should be selected. If this mid-point value is not a whole number, it should be rounded to the nearest whole number.
3. If no value is given, the answer should be considered a missing value and the computer field left blank

Q14(a): One computer field has been created for this question.

1. Each box has been given a numerical value starting with FOOD STORES-1 through to OTHER SERVICES-53. The numerical value representing the box selected by the respondent should be entered into the computer field.
2. If more than one box has been checked by the respondent, two main choices are available for coding: (1) select the one box that represents the main product as indicated by Q14(b) or (2) GENERAL MERCHANDIZE-18 can be selected if several choices indicating typical department store items have been checked.
3. The box selected by the respondent should be verified against the main products as indicated in Q14(b).

Please note: variables or computer fields have not been created for Q14(b): MAIN PRODUCTS and Q14(c): 4-DIGIT SIC CODE

PERMISSION

1. One variable is created
2. The value label 1 is assigned to YES permission
3. The value label 3 is assigned to NO permission
4. Only one box should be checked
5. If neither box is selected, the answer should be considered NO

B. BREAKDOWN OF ECONOMIC SECTOR BY SIC GROUP

INDUSTRIAL SIC GROUPS

SIC	DESCRIPTION
06	Mining Industries
07	Crude Petroleum
08	Quarry and Sand Pit
09	Services Incidental to Mineral Extraction
10	Food Industries
11	Beverage Industries
12	Tobacco Industries
15	Rubber Products
16	Plastic Products
17	Leather and Allied Products
18	Primary Textile
19	Textile Products
24	Clothing Industries
25	Wood Industries
26	Furniture and Fixture Industries
27	Paper and Allied Products
28	Printing and Publishing and Allied Products
29	Primary Metal Industries
30	Fabricated Metal Products
31	Machinery Industries
32	Transportation Equipment
33	Electrical and Electronic Products
35	Non-Metallic Mineral Products
36	Refined Petroleum and Coal Products
37	Chemical and Chemical Products
39	Other Manufacturing Industries

COMMERCIAL SIC GROUPS

SIC	DESCRIPTION
60	Food, Beverage and Drug Retail
92	Food and Beverage Services
61	Shoe, Apparel, Fabric and Yarn Stores
62	Household Furniture, Appliances and Furnishings
63	Automotive Vehicles, Parts and Accessories
64	General Merchandizing
65	Other Retail Sales
69	Non-store Retail
91	Hotels and Accommodations
96	Entertainment Services
97	Personal Services
99	Other Services

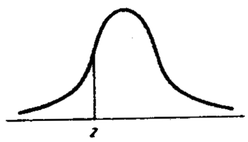
AGRICULTURAL SIC GROUPS

SIC	DESCRIPTION
111	Dairy Farms
112	Cattle Farms
113	Hog Farms
114	Poultry and Egg Farms
115	Sheep and Goat Farms
119	Livestock Combination Livestock
121	Honey and Other Apiary Products
122	Horse and Other Equine Farms
123	Furs and Skins, Ranch
129	Other Animal Specialty
131	Wheat Farms
132	Small Grain Farms (except wheat)
133	OilseedFarms (except corn)
134	Grain Corn Farms
135	Forage, Seed, Hay Farms
137	Tobacco Farms
138	Potato Farms
139	Other Field Crops
141	Combination Field Crop
151	Fruit
152	Other Vegetables
159	Combination Fruit & Veg
161	Mushrooms
162	Greenhouse Products
163	Nursery Products
169	Other Horticulture Specialties
171	Livestock, Field Crop and Horticultural Combination

C. NORMAL DISTRIBUTION

Table C.1: Cumulative Distribution Function of the Standard Normal Distribution

Areas under the standard normal curve (Areas to the left)



z	0	1	2	3	4	5	6	7	8	9
-3.0*	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0020	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3516	.3121
-.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

*For $z \leq -4$ the areas are 0 to four decimal places.

Table C.1, continued

z	0	1	2	3	4	5	6	7	8	9
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0†	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

†For $z \geq 4$ the areas are 1 to four decimal places.

Adapted from Probability with Statistical Applications, second edition, by F. Mosteller, R. E. K. Rourke, and G. B. Thomas, Jr. Reading, Mass.: Addison-Wesley, 1970, p. 473.

D. 1991 SECTOR $/KW PEAK NORMALIZED COSTS

Table D.1: Average $/KW Peak Peak Normalized Costs for the Residential Sector

Interruption Duration & Frequency	Average $ /KW Cost	
20 min/month in winter	0.0278	(1351)
1 hour/month in winter	0.1626	(1345)
4 hour/month in winter	1.8126	(1343)
8 hour/year in winter	4.0006	(1328)
24 hour/year in winter	18.2491	(1313)
48 hour/year in winter	44.4084	(1318)
4 hour/month in summer	0.6552	(1338)
48 hour/year in summer	23.0345	(1334)
24 hr twice/year summer	22.5679	(1332)

Note : A Load Factor of 23% was assumed

Table D.2: Aggregated $/KW Peak Normalized SIC Group Costs and Weighted Total for the Small Industrial Sector

SIC	Interruption Duration							
Group	2- Sec		1- min		20- min		1- hr	
04	--	(0)	--	(0)	3.2129	(1)	16.0643	(1)
06	1.6551	(3)	1.6643	(3)	1.4779	(4)	2.3266	(4)
07	3.3385	(9)	3.9805	(9)	90.1469	(12)	214.5381	(16)
08	7.9186	(1)	2.4642	(2)	3.2706	(2)	4.1443	(2)
09	0.0000	(1)	0.0000	(1)	0.0000	(1)	1.6482	(2)
10	0.9118	(16)	1.3430	(17)	4.9923	(21)	15.8998	(22)
11	0.2016	(2)	0.2016	(2)	0.4608	(2)	1.2097	(2)
15	0.0375	(3)	0.0911	(3)	0.8630	(4)	1.3908	(4)
16	0.3623	(14)	1.0454	(16)	1.4340	(17)	2.2571	(17)
17	0.0000	(3)	0.0000	(3)	0.0000	(3)	1.0634	(3)
18	13.4350	(1)	13.4350	(1)	13.4350	(1)	13.4350	(1)
19	0.0044	(2)	2.0650	(2)	3.5896	(2)	6.9401	(3)
24	0.0000	(1)	0.0000	(1)	2.2501	(2)	6.7502	(2)
25	1.0108	(11)	0.9155	(13)	1.2101	(13)	2.2766	(13)
26	0.0088	(4)	7.8052	(4)	8.7585	(4)	18.0347	(4)
27	0.1738	(3)	5.0159	(3)	5.2970	(3)	5.8529	(4)
28	1.0297	(14)	1.1057	(14)	2.4813	(15)	4.6707	(17)
29	1.2892	(2)	1.3891	(3)	1.7198	(4)	2.7516	(4)
30	1.0377	(30)	1.8955	(33)	3.2947	(36)	6.5981	(36)
31	1.6712	(9)	2.4377	(9)	3.3356	(10)	5.9783	(10)
32	0.7593	(5)	1.8982	(5)	13.9723	(6)	33.3920	(7)
33	2.5673	(5)	2.5673	(5)	3.5996	(7)	6.8203	(7)
35	0.5807	(11)	1.0507	(10)	2.9435	(11)	7.4569	(14)
36	--	(0)	--	(0)	--	(0)	0.0000	(0)
37	0.6352	(8)	0.6352	(8)	1.8822	(9)	3.6147	(10)
39	14.0293	(5)	14.0293	(5)	14.8042	(6)	11.9046	(7)
Weighted Total	0.7291	(163)	3.1663	(172)	4.3217	(196)	6.5508	(212)

Table D.2, continued

SIC Group	Interruption Duration 2- hr		4- hr		8- hr		1- day	
04	40.1606	(1)	120.4819	(1)	160.6426	(1)	160.6426	(1)
06	3.3079	(4)	5.1536	(4)	13.9467	(4)	32.1566	(4)
07	216.5901	(16)	227.7870	(16)	241.6942	(16)	304.2332	(16)
08	5.8244	(2)	9.1846	(2)	15.9050	(2)	15.9050	(2)
09	3.4063	(2)	6.8126	(2)	20.6575	(2)	40.4360	(2)
10	29.0981	(22)	53.1859	(22)	103.2057	(22)	259.4471	(22)
11	4.8963	(2)	7.0565	(2)	65.8122	(2)	65.8122	(2)
15	2.5153	(4)	5.0093	(4)	9.7594	(4)	21.6751	(4)
16	3.2491	(17)	6.9590	(17)	10.6697	(17)	15.2645	(17)
17	5.6891	(3)	13.5687	(3)	53.7218	(3)	134.9426	(3)
18	13.4350	(1)	26.1438	(1)	26.1438	(1)	26.1438	(1)
19	11.9967	(3)	26.2955	(3)	50.7074	(3)	50.9334	(3)
24	13.9023	(2)	27.8046	(2)	45.8052	(2)	45.8052	(2)
25	3.9417	(13)	7.4526	(13)	14.8352	(13)	17.1577	(13)
26	22.3330	(4)	44.0512	(4)	58.2730	(4)	59.1556	(4)
27	7.7024	(4)	14.0557	(4)	25.8050	(4)	30.8094	(4)
28	7.4306	(17)	14.1092	(17)	25.6635	(17)	43.7705	(17)
29	3.4150	(4)	5.8227	(4)	18.7455	(4)	30.6366	(4)
30	12.3182	(36)	30.2618	(36)	59.3590	(36)	85.4451	(36)
31	11.6436	(10)	20.4206	(10)	35.4620	(10)	37.9976	(10)
32	74.8910	(7)	261.3217	(7)	440.2326	(7)	844.9066	(7)
33	9.7845	(7)	20.5572	(7)	37.9091	(7)	52.1280	(7)
35	16.0036	(14)	29.7440	(14)	58.4458	(14)	71.7457	(14)
36	--	(0)	--	(0)	--	(0)	--	(0)
37	7.7812	(10)	19.8437	(10)	48.1409	(10)	74.4957	(10)
39	37.8625	(7)	42.8371	(7)	75.1187	(7)	81.0072	(7)
Weighted Total	9.1189	(212)	16.2679	(212)	30.3254	(212)	44.7320	(212)

E. RESULTS OF DISTRIBUTION ANALYSIS USING 1991 $/KW COST OF INTERRUPTION DATA

Table E.1: Moment Test Results : without zero-valued data

(a) Residential Sector

Interruption Duration	λ	Z_1	Z_2	$\sqrt{b_1}$	b_2	Normality
20 min	-0.2207	0.0012	1.8028	0.0001	0.8406	YES
1 hr	-0.1828	-0.0052	0.6028	-0.0005	0.0733	YES
4 hr	-0.0105	0.0042	-0.8253	0.0004	-0.1200	YES
8 hr	-0.0162	0.0020	1.9554	0.0002	0.4982	YES
1 day	0.0238	0.0004	1.9416	0.0000	0.7593	YES

(b) Small Industrial Sector

Interruption Duration	λ	Z_1	Z_2	$\sqrt{b_1}$	b_2	Normality
2 sec	0.0016	0.0012	1.4982	0.0004	0.7173	YES
1 min	-0.0488	0.0009	0.6371	0.0004	0.1513	YES
20 min	-0.0605	-0.0010	0.9255	-0.0003	0.2681	YES
1 hr	-0.0707	0.0019	0.6562	0.0005	0.1471	YES
2 hr	-0.0586	0.0011	0.4333	0.0003	0.0660	YES
4 hr	-0.0387	0.0015	-0.3535	0.0004	-0.1750	YES
8 hr	-0.0020	0.0014	-0.8199	0.0003	-0.2946	YES
1 day	-0.0105	0.0014	-1.3003	0.0003	-0.4030	YES

Table E.2: Distribution Parameters and Results of Regression Analysis

(a) Residential Sector

Duration	λ	Mean (μ)	Variance (σ^2)	P_Z (%)
20 min	-0.2207	-5.6618	4.8689	0.3295
1 hr	-0.1828	-2.7329	2.8790	0.0973
4 hr	-0.0105	0.2886	1.6551	0.0265
8 hr	-0.0160	1.1345	1.5725	0.0426
1 day	0.0238	2.8289	1.7337	0.0151

$\lambda = -0.6690 + 0.1456 \cdot \log(d)$ $\quad R^2 = 91.0\%$

$\mu = -19.23 + 4.5563 \cdot \log(d)$ $\quad R^2 = 98.2\%$

$\sigma^2 = 34.989 - 14.88 \cdot \log(d) + 1.6512 \cdot [\log(d)]^2$ $\quad R^2 = 99.8\%$

$P_z = 284.7 \cdot [\log(d)]^{-6.1758}$ $\quad R^2 = 91.9\%$

(b) Small Industrial Sector

Duration	λ	Mean (μ)	Variance (σ^2)	P_Z (%)
2 sec	0.0016	0.2418	4.4143	0.4785
1 min	-0.0488	0.3352	3.8770	0.3488
20 min	-0.6605	1.0487	2.7866	0.1513
1 hr	-0.0707	1.6327	2.3443	0.0613
2 hr	-0.0586	2.1500	2.2301	0.0283
4 hr	-0.0387	2.8272	2.3620	0.0047
8 hr	-0.0020	3.6939	2.9880	0.0047
1 day	-0.0105	3.9817	2.7908	0.0047

$\lambda = 0.0175 - 0.0515 \cdot \log(d) + 0.0053 \cdot [\log(d)]^2 + 0.0010[\log(d)]^3$ $\quad R^2 = 82.4\%$

$\mu = 0.148 \cdot 10.00^{0.2955 \cdot \log(d)}$ $\quad R^2 = 96.1\%$

$\sigma^2 = 4.2983 + 0.6234 \cdot \log(d) - 0.6779 \cdot [\log(d)]^2 + 0.1007[\log(d)]^3$ $\quad R^2 = 93.1\%$

$$P_z = \begin{cases} 0.5409 - 0.1297 \cdot \log(d), & d < 4 \text{ hours}, \\ 0.0047, & d \geq 4 \text{ hours}. \end{cases} \qquad R^2 = 98.6\%$$