1999

Framework for power system annual risk assessment

Youjie Dai
Iowa State University

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Framework for power system annual risk assessment

by

Youjie Dai

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering (Electric Power)

Major Professor: Dr. James D. McCalley

Iowa State University

Ames, Iowa

1999

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of

Youjie Dai

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

To my wife

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The deterministic method has been the primary means of performing power system security assessment for a long time. This is partly because it is easy to understand and implement, and partly because it is usually quite conservative. In the past where monopoly was prevalent, the conservativeness resulted in a high degree of reliability in most power systems, while the investment and operational costs rose without the pressure of competition. However, now because of the deregulation and practical difficuties to obtain authorizations from regulatory bodies to build power plants and transmission lines, people are more and more willing to operate power systems with lower security margins. This demands more accurate and comprehensive risk assessment tools. On the other hand, because of the fast development of the computer and of computational mathematics, probabilistic risk assessment becomes more and more practical. This kind of risk assessment can deal with both operational and planning problems. Although planning and operations are normally regarded as different catagories, this paper is aimed at building a framework for power system risk assessment in the planning stage such that it is developed naturally from the operational stage. The framework is modular so that it is relatively easy to implement, and each module can be improved individually without influencing other parts of the framework. Compared with Monte Carlo simulation where possible system trajectories are sampled, our framework employs the expected trajectory, while accounting for the load uncertainty. One of the most prominent advantages of our proposed technique is that it can provide us decomposible and assignable risk among system components. The IEEE RTS' 96 is used as the test power system for our

proposed framework. Various calculation results are listed and analyzed. Some facility planning decisions are suggested based on our calculations. Our proposed framework is shown to be valid and efficient by these calculations.

CHAPTER 1 INTRODUCTION

1.1 Background

1.1.1 The Importance of Risk Assessment

"No matter how good the deregulation is claimed to be, there must be something wrong if light bulbs stop shining." This concern was expressed by a power engineer when we discussed about deregulation. Unfortunately, this concern was strengthened by the severe power outage on December 7, 1998 in San Francisco, just 8 months after the formal beginning of the power system restructuring there. Is there anyting wrong with "deregulation", perhaps better known as "restructuring"?

In its simplest form, restructuring is the move to introduce competition among electrical energy suppliers. It has been developed for more than two decades [1], and is still developing now [2] [3]. Deregulation has been tried in other industries such as the airlines, gas, and telecommunication industries and achieved significant success. Furthermore, power system restructuring was also introduced into many other countries including United Kingdom, Sweden, Norway, New Zealand, Australia, Spain, etc. Even Japan [4] and Korea [5] enacted appropriate acts to facilitate competition. In the United States, PJM and California have already built their own market. It seems deregulation is sweeping the whole world.

An important issue within deregulation is reliability and whether it will decrease. Competition will inevitably lead to larger uncertainties in load and generation and in-

evitably drive competitors to operate their components or systems closer to their security limits. Without accurate risk assessment and appropriate control, system reliability can not be guaranteed. Therefore, the world wide deregulation trend asks for more comprehensive and accurate risk assessment tools.

As is mentioned by reference [6], today's risk assessment is like charting the unexplored oceans and imaginary lands for the medieval European people. The safe area identified by traditional deterministic criteria no longer satisfies a system operator who is either a competitor or an Independent System Operator (ISO) under the pressure of competition. What if we go outside the so-called safe area? Unlike the medieval analogy, we may suffer a great loss if we happen to choose a wrong direction. However, in our case, we are capable of obtaining abundant knowledge about the power system using computers, in contrast to the medieval's ignorance to the outside world and relatively crude measurers. Unlike the brave Renaissance explorers, we can perform risk assessment without having to actually risk our dangers.

1.1.2 Project Background

From the beginning, power systems are inherently uncertain. To name a few uncertainties, we have

- Load.
- Generation.
- Topology of the network.
- Measurement errors.
- Component failures.
- Weather conditions.

- Reservior constraints if there are some hydro plants.
- Mistakes of operators.

Now deregulation brings us market uncertainties. Despite these uncertainties, we have to maintain the system reliability, otherwise the benefit of deregulation will be lost. In recognizing the importance and urgency of constructing the risk assessment framework under the new deregulated environment, the Electrical Power Research Institute (EPRI) funded a project (contract W08604-01) to deal with framework construction of power system risk assessment. The project is also funded by National Science Foundation Grand ELS9502790. This dissertation reports on a part of the work in this project. Our task is to build a framework to evaluate risk under all these uncertainties, so that the evaluated risk can be compared with benefit to facilitate decision-making.

1.1.3 The Concept of Risk

The term "risk" is not a new concept in engineering circles; there is a large body of literature on risk analysis that has been generated by the process control industry, the airline industries, NASA, and the nuclear power industry. In addition, the insurance, financial, and casino industries all use risk analysis tools to deal with various problems. A representative sample of references include [7], [8], [9]. The main idea of the risk analysis in these areas is to research on the stochastic behaviours of loss.

In power systems, risk means different things to different people. Some people refer risk to the probability, some people refer it to the impact, while others refer it to the expected impact. We think that the IEEE Standard Dictionary provides a clear definition that conforms quite well to the main precepts on which power system and security assessment is based. In this dictionary [10], risk associated with an event is defined as "the simple product of probability and consequence" of the event. Risk can also be thought of as the expected value of the consequence. In this dissertation, risk is defined

as the expected cost of operating under stressed conditions suffered by power system facilities under a reasonably likely scenario. According to this definition, we do not calculate financial or insurance risk suffered in the deregulated power market, although our calculation results can provide the basis for computing these risks. In addition, we do not calculate risk due to very low probability events such as independent, simultaneous outage of multiple components.

1.1.4 Reliability vs. Risk Assessment

Traditionally, electric utilities were more concerned about reliability than economy. With the developement of restructuring, it is gradually accepted that economy should be as important as reliability, if utilities are to be competitive. With the development of risk assessment tools, it is now possible to assess economy and reliability in a common framework. Therefore, the risk assessment, which combines reliability and economic analysis, becomes more and more attractive [11]. Under the new deregulated environment, we think risk assessment is more useful than reliability evaluation because of the following reasons.

- Risk emphasizes the economic cost/benefit comparison.
- Risk assessment provides a simple index that is relatively easy to understand and calculate compared with reliabilty evaluation where a large number of indices have been proposed and used. Only expected cost need to be calculated, frequency indices and many other indices are not required.
- Risk is decomposable and therefore can be related to individual market players, while reliability is more system oriented and thus less decomposable.
- Risk assessment can take factors from other markets (e.g., oil, gas, coal markets, etc.) into account. Because risk is measured in dollars, it is familiar to non-

engineers employed in electricity markets. However, traditional reliability indices such as LOLP (loss of load probability) are not as meaningful to people without the appropriate expertise on reliability indices.

1.1.5 Hierarchical Level of Risk

It has been suggested that reliability evaluation can be divided into three hierarchical levels [12].

- Hierarchical level 1: only includes generation facilities.
- Hierarchical level 2: includes generation and transmission facilities.
- Hierarchical level 3: includes generation, transmission and distribution.

The problem we are addressing in this dissertation belongs to hierarchical level 2. However, we discuss risk here instead of reliability.

1.1.6 Operation vs. Planning

During the early days of electric power systems, the functions of planning and operating a power system were often considered quite distinct [13]. The problems planners and operators had to address were different, and so were tools employed to solve the problems. Such a division was formed due to computational difficulties rather than physical essence. Physically, the planning analysis should be an accumulation of the operational analyses.

This problem has been recognized by CIGRE amongst others for some considerable time, and indicated by the CIGRE "Power System Reliability Analysis Application Guide" [14]. However, the scope of those original considerations were limited only to the area of adequacy of single systems. Since then, the GIGRE Working-Group 38.03

are making effort to extend this previously limited scope to interconnected systems from an adequacy viewpoint, and to security aspects of single systems.

One objective of this dissertation is aimed at developing planning risk assessment which is naturally extended from operational risk assessment, so that risk can be estimated at every hour. Due to the developement of computers and compuation techniques, it is time now to "close the gap between the thought processes of system planners and of operators" [13].

1.1.7 Adequacy vs. Security

Power system reliability assessment has long been divided into two distinct aspects: adequacy and security. Adequacy has been widely researched due to its simplicity [12]. However, security is not well addressed due to the complexity. The basis for security-related considerations was laid out by T. Dy Liacco in [15][16][17]. Generally accepted definitions of adequacy and security are given by the North American Reliability Council (NERC) Planning Standards [18]. According to the NERC standards, adequacy is defined as "the ability of the electric systems to supply the aggregate electrical demand and energy requirements of their customers at all times, taking into account scheduled and reasonably expected unscheduled outage of system elements". Security is defined as "the ability of the electric systems to withstand sudden disturbances such as electric short circuits or unanticipated loss of system elements".

In the past, some persons have interpreted security as the ability of electric systems to withstand sudden disturbances in terms of the short term, so-called transient effects, whereas adequacy has been interpreted as the ability of the system to supply the load without violation of circuit or bus voltage constraints. Under this "traditional" interpretation, a secure system was one that would maintain uninterrupted supply to all loads, and all bus voltages and flows would be within defined ratings for all contingencies in the credible contingency set. This division conveniently corresponded to the

way in which the two were studied: security was studied using dynamic analysis (time domain simulation) and adequacy was studied using steady-state analysis (power flow simulation). The division is also somewhat mathematical rather than physical. Due to the complexity of security assessment, traditional reliability evaluators are happy to calculate only adequacy indices, and leave security problems unanswered or partially addressed by crude and conservative deterministic approaches. However, the physical boundary of the two are somewhat smeared. If the system can not be kept stable, how can it provide load?

This dissertation addresses both adequacy and security without making a clear division between them. Instead, the total risk is functionally divided, as shown in the next subsection.

1.1.8 Composition of Risk

The risk of a power system can be functionally divided into 3 types: circuit overload, power flow infeasibility problems, and dynamic problems. We are motivated to include these three types of problems under the same umbrella because our intent is to develop a single assessment framework to encompass all of them. This conceptualization of risk is illustrated in Figure 1.1. Figure 1.1 also illustrates a decomposition of each of the three forms of security. The reason for making this further decomposition is because each of the six types of security at the lowest level of the hierarchy differ in terms of the specific calculation procedures we have developed to assess them.

1.1.9 Time Horizons for Risk Assessment

Three types of time horizons are normally adopted for risk assessment.

- Real time operation risk assessment, used to provide guidelines for real time system operation. The risk assessment is often carried out less than one hour before the

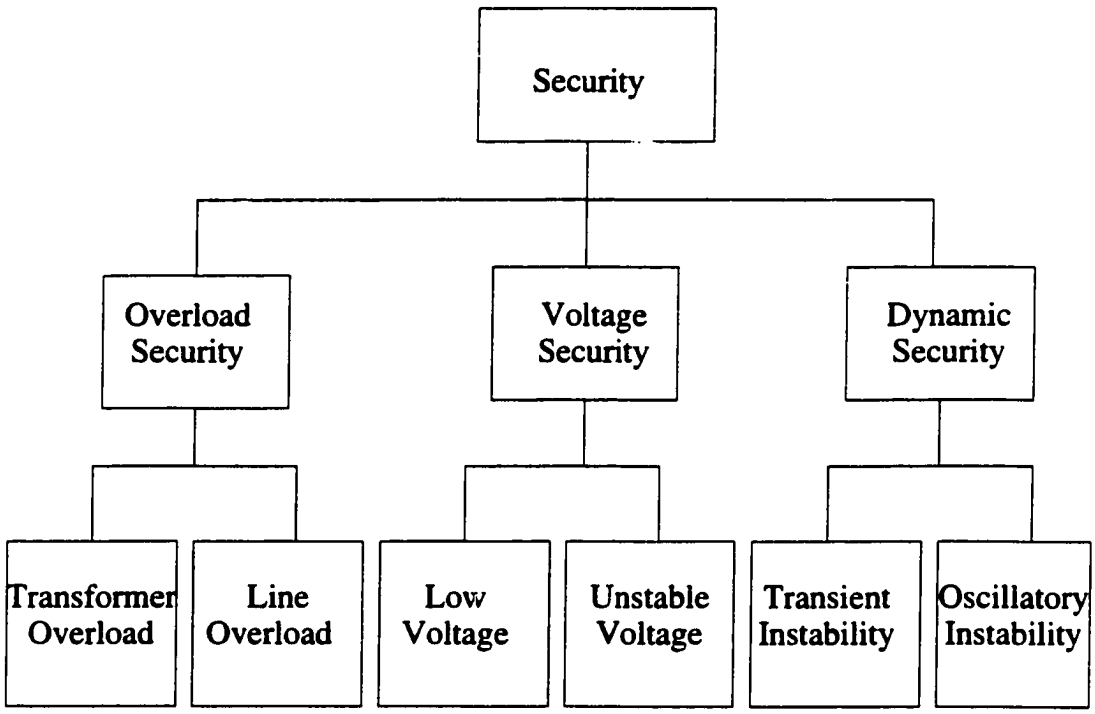

Figure 1.1 Composition of security

operation time.

- Short term operation risk assessment, used to provide cost and benefit analysis for short term economic decision making. The time horizon is about one day to one week.

- Long term planning risk assessment, used to provide information for facility planning, long term economic decision making, determination of service charges, etc. The time horizon is one year or more.

The shorter the time, the easier it is to identify the system trajectory, however, more modeling details of the system can be considered. On the other hand, the longer the time, the more attention has to be paid for system trajectory identification, so that fewer modeling details are considered. This dissertation mainly deals with planning risk assessment. One year risk assessment is performed on a test system, the IEEE RTS' 96.

However the proposed fundamental framework can also be applied to other systems for other long time horizons as well.

1.2 Research Objectives

The objectvie of this dissertation is to develop an effective framework and a set of algorithms to calculate cumulative risk assessment for one year. Only thermal overload risk assessment and power flow infeasiblity risk assessment are addressed to limit the scope. Both adequacy and security risk are considered and combined together to get the composite risk. The purpose of the risk assessment is to answer the following practical questions.

- When will the system suffer severe risk?
- Where will the system suffer severe risk?
- Which contingency plays the key role in causing risk?
- What is the main cause of risk? Thermal overload, generation limit, or voltage stability?
- If we make alternative decisions, will the risk be significantly reduced? By how much?

Answers to these questions are useful in making decisions about alternatives having varying degrees of risk and expected benefit.

In the era of deregulation, risk assessment has to conform to the needs of the deregulated environment. This means risk assessment has to have the following features.

- It should accommodate appropriate uncertainties. Due to the competition in the power market, we believe load and generation will become more and more uncer-

tain. Thus load forecasting error and the control effect of generators should be included in our model.

- It should be decomposable and assignable. The power system is no longer controlled by a single vertical company in the deregulated market, it is a complicated system owned by various companies. Therefore, identifying the allocation of risk among different companies becomes very important. Knowing a single risk assessment number for the system is not enough. We have to provide tools to calculate who suffers how much risk and who is responsible for this risk.

1.3 Structure of the Dissertation

Chapter 2 proposes a new system trajectory model and compares it with two other models, then proposes a framework to perform cumulative risk assessment. Chapter 3 illustrates techniques we adopt to identify system trajectory and its uncertainty, including load forecasting error identification, maintenance scheduling, and unit commitment arrangement. Chapter 4 elaborates how to use the interior point algorithm to calculate power flow infeasibility risk. Chapter 5 explains our proposed method for thermal overload risk assessment based on convolution and DC power flow. Chapter 6 combines thermal overload risk and power flow infeasibility risk to get the composite risk, identifies the weak time period and weak area of the IEEE RTS'96, proposes alternative facility plans and unit commitment patterns, and performs cost and benefit comparison. Chapter 7 summarizes the work we have done and proposes possible future research directions.In order to show our philosophy and keep the consistency, we adopt the IEEE RTS 96 system as the example system throughout this dissertation. However, this does not prevent readers from applying our proposed philosophy on any other power systems.

CHAPTER 2 OVERVIEW OF CUMULATIVE RISK ASSESSMENT FRAMEWORK

2.1 Introduction

One of the main features of our risk assessment framework is to first identify the power system trajectory over a period of time, then perform risk assessment based on the identified trajectory model. There are several classes of system trajectory representation models available currently. This chapter first addresses comparisons between different models, then introduces our cumulative risk assessment framework based on our own system trajectory model.

2.2 Comparison of Three System Trajectory Models

The system trajectory model is very important in annual risk assessment. If the model is chosen too simple, accuracy can not be guaranteed. Yet if the model is overly-complicated, computation is too burdensome. One has to balance between accuracy and complexity requirements. According to different accuracy and complexity, there are at least two classes of models available currently.

2.2.1 Snapshot Models

The earliest and most common class of models we call the snapshot models, employed by well known software products such as PROCOSE [19], TRELSS [20], and others. In

this class, several typical snapshots of loadings are chosen according to experience. Then for every snapshot, unit commitment is arranged and reliability indices are computed. The cumulative risk is just a weighted average of the indices accross the various snapshots, where weights are relative probabilities assigned to the snapshots. Because of the simplicity of this model, risk assessment or reliability evaluation can be performed fairly quickly. Typically the number of the snapshots are quite limited, and engineering judgement is used to choose them. However, this method does not account for intertemporal dependencies such as those that occur due to start up and shut down of generators. It is hard to capture peak risk time periods accurately because they may not occur at extreme loading conditions. As we will show in chapters 5 and 6, peak risk frequently occurs when a load increment is just below that necessary to cause a unit commitment change, or a load decrement is just at the point to cause a unit commitment change.

2.2.2 Sequential Trajectory Models

The second class of models is called sequential trajectory models. It simulates the system trajectory hour by hour over a whole year. Therefore it has to calculate maintenance schedule and unit commitment for generators. Then the risk assessment or reliability evaluation is calculated based on the simulated trajectory (or trajectories). One such model called sequential Monte Carlo (MC) simulation model has already been proposed in [21]. The model simulates the power system sequentially hour by hour until the whole year is finished, then repeats the simulation using randomly chosen trajectories until risk assessment (or reliability evaluation) becomes numerically stable. This model can provide the most accurate result. However, it is very computationally intensive. This is of particular significance for risk decomposition purposes where we desire to identify who causes and incurs risk because early termination of the simulations can result in erroneous zero risk assignment to some agents. In order to obtain accurate risk assessment for every hour and every component, several million simulations must be

performed. Currently, such a calculation requirement can not be handled by a standard work station within a reasonable time period for typical size power systems. Therefore, this model is normally used to calculate system indices rather than individual component risk or risk with respect to certain hours.

2.2.3 Sequential Mean-Variance Model

Based on the analysis and comparison of the snapshot model and the sequential MC simulation model, we propose a new model called the sequential mean-variance (MV) model. It belongs to the second class. However, it is completely different from the sequential MC simulation model in that it has only one expected system trajectory instead of thousands of possible system trajectories. In this model, we assume we know that we have the expected load profile over the next year. Then we can use it to arrange maintenance schedule and unit commitment. The expected load profile can be obtained from load forecasting, or it can be the load profile of previous year, expanded by, say, a 3–5% increment. It is reasonable for us to assume that the system trajectory obtained by this method is the expected system trajectory. Then we should also include variances, so that at every time point, every load is modeled according to a Gaussian distribution. Furthermore, we assume that the variances of the loads are equal at any time point. We adopt one day ahead load forecasting error as the relative error of a load. We use one day ahead error because

- Most costly situations occur within 24 hours of a contingency.
- Any cost having potential to occur after a 24 hour period can be largely eliminated via operation action.
- There exists one day ahead market in many restructured power systems.

The total load forecasting error is then proportionally distributed among all loads. If we have enough data, the covariance matrix of loads can also be identified via statistical methods. In this dissertation, we assume this data is available. Our model considers the sequence of the unit commitment, therefore it can capture peak risk periods easily. It includes load variances, therefore it can perform probabilistic analysis. For a reasonable time period like one year, one can decompose risk among components and hours accurately.

2.2.4 Comparison of Three Trajectory Models

A comparison of these three trjactory models is shown in Table 2.1.

From this table, we can see that our sequential MV model

- is faster than the sequential MC model;
- is more accurate than the snapshot model;
- is risk decomposable;
- can capture peak risk time period accurately.

Table 2.1 The comparison of three trajectory models

Model Name	snapshot	sequential MC	sequential MV
speed	fast	slow	fast
accuracy	low	high	high
risk decomposable?	yes	almost impractical	yes
ability to capture peak risk	low	high	high

2.3 Framework of Annual Risk Assessment

In [22] the risk-based security assessment approach is proposed. The main idea of this approach is to use risk measured in dollars instead of reliability indices to evaluate the security level of components and systems. The definition of risk is explained in Chapter 1. In this dissertation, we extend the risk-based security assessment (RBSA) approach to the facility planning problem. Here, we desire to compute risk over a time period, say one year, for the purpose of deciding among different planning alternatives. Our approach is to simulate the system on an hour by hour basis, compute risk for each hour, and then accumulate the risk over the time period. We call the result a cumulative risk, given by the expression below

$$Risk(Im|T,\Omega) = \int_T \int_{E_i} Pr(E_i|t,\Omega) \cdot Im(E_i|t,\Omega) dE_i dt \tag{2.1}$$

where T is the time duration of the study period, Ω is the system trajectory during T, E_i is contingency state i, Pr is probability, and Im is some kind of impact(e.g., voltage collapse, load curtailment, thermal overload).

Our framework of annual risk assessment is shown in Figure 2.1. It is decomposed into two modules. One module is used to identify system trajectory model. The other module (included within a dotted line) is used to perform various kinds of risk assessment. The risk assessment module is enclosed by a dotted line because we want to emphasize the submodules in this module. After each cumulative risk assessment, the facility plan and unit commitment is adjusted based on our risk assessment results. Then risk of alternative plans can be calculated and compared.

In the system trajectory identification module, the expected system trajectory and the load forecasting error are identified by

1. Developing an hour by hour load forecast.

2. Identifying and modeling the error associated with the load forecast.

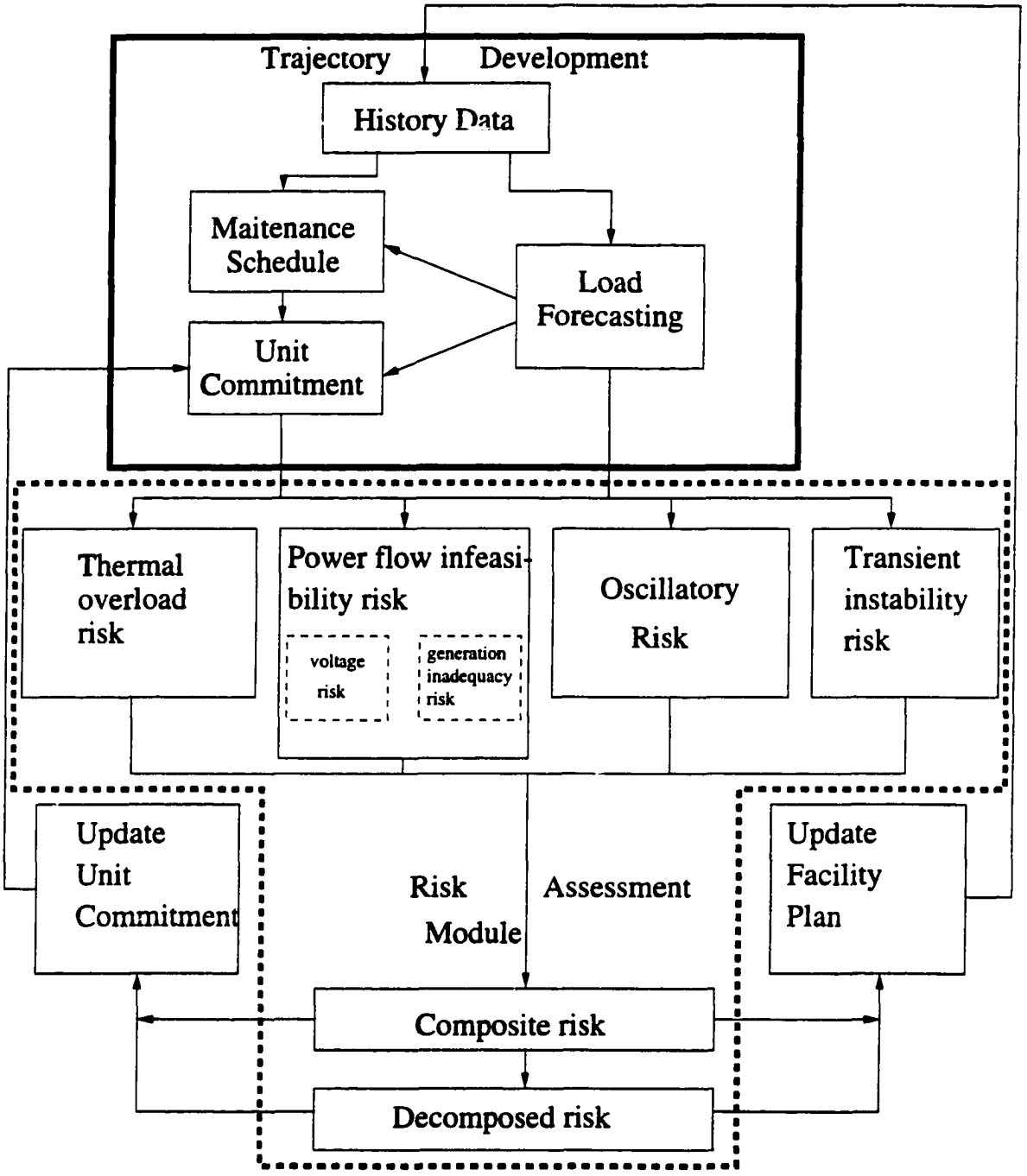

Figure 2.1 Cumulative risk assessment framework

3. Forming a maintenance schedule for all generation units.

4. Developing a unit commitment plan based on the load forecast and the maintenance schedule.

In order to show the effectiveness of our framework, we adopt ARIMA model to identify the load forecasting error, equal LOLP criterion to determine the maintenance schedule, and priority list method to determine the unit commitment. However, we do not purpose that these methods are best nor do we claim that our implementation of them can not be improved. We do claim and illustrate that an expected trajectory can be obtained and given that, we may perform a cumulative risk assessment over an extended time period such as a year. Our current research results point out several research directions that are useful for cumulative risk assessment.

Given that we have the trajectory of operating conditions, we accumulate the composite risk at each hour to get annual composite risk in the risk assessment module. As shown in Figure 2.1, this module can include thermal overload risk assessment, power flow infeasibility risk assessment, steady state instability risk assessment and transient instability risk assessment. Power flow infeasibility risk assessment is a new concept proposed in this dissertation to assess the risk associated with power flow infeasibility. Power flow infeasibility is caused by voltage risk and/or generation inadequacy risk. The two types of risk can be calculated together and presented in a decomposed fashion. We will provide a detailed explanation of this concept in Chapter 4. Thermal overload risk is the risk suffered due to life shortening or sagging when a component is thermally overloaded. Oscillatory instability risk is a risk suffered when the system is oscillatorily unstable. Transient instability risk is a risk suffered when a credible contingency causes transient instability problem. After the risk assessment of the four modules, the composite risk can be calculated by summing them together. The risk decomposition among hours, components and contingencies can also be obtained and used to adjust facility

plan and unit commitment. In our work on the planning problem, however, we have only included two submodules: power flow infeasibility submodule and thermal overload risk assessment submodule. The composite risk and decomposed risk are also calculated only based on these two submodules. Therefore we have not included the ability to compute cumulative risk for transient instability or oscillatory instability. Our decision to exclude assessment of these risks was made based only on our ability to complete the work during the project time frame.

Component outages and load uncertainties are considered at each hour. However, this increases the complexity of risk assessment. Normally, Monte Carlo simulation is adopted by other researchers to solve similar calculation problems, whereas we design clever algorithms to avoid MC simulation in this dissertation. Compared with MC method, our algorithms can identify the contribution of each outage, the risk suffered by each component, and risk due to different reasons (whether it is due to generation limit, voltage limit or voltage stability), with much less computation time.

The results of the risk calculation may be provided in a "decomposed" form, by security problem type, by electrical facility type, or by electrical facility owner. The results of the composite and decomposed risk are used to modify the facility plan or to identify problems in the short-term operating strategies such as the unit commitment.

CHAPTER 3 IDENTIFICATION OF SYSTEM TRAJECTORY MODEL

3.1 Introduction

For our sequential mean variance model, we assume that the system trajectory is completely specified by the expected load profile, the load forecasting error, the maintenance schedule and the unit commitment. We assume we know the expected load profile. The main purpose of this chapter is to show how we determine the load forecasting error, the maintenance schedule and the unit commitment. The three problems can be solved independently, or they can be solved in an integrated fashion. The latter method typically results in a more optimal trajectory, but it is more complicate and computationally expensive. Rather, we choose to solve them independently in order to maintain model simplicity and enhance computational efficiency.

In each of the three modules, there are several alternative methods. We do not claim that the method adopted by us in each module is the best one. Our main contribution in this dissertation is not to propose the best method to identify load forecasting error, or maintenance schedule, or unit commitment. Our main contribution is to illustrate that we can combine them together to obtain a system trajectory, and based on this trajectory we can perform risk assessment. Therefore, in the remaining sections of this chapter, we provide the overview of the methods we adopt; the analytical details of each method is relegated to the appendices.

3.2 Load Forecasting Error Identification

The forecasted load, or the expected load profile of the IEEE RTS'96 is shown in Figure 3.1. This forecasted load profile may be developed from any desirable method. However, there always exist a load forecasting error so that the actual load at a time is not exactly equal to the corresponding forecasted load. One merit of our risk assessment framework is that we can account for such an uncertainty in our risk assessment. The introduction of load forecasting error in our load model makes our calculation more reasonable and accurate. However, it also brings the problem of how to identify the load forecasting error. This section is aimed at explaining the method we adopted to identify the load forecasting error.

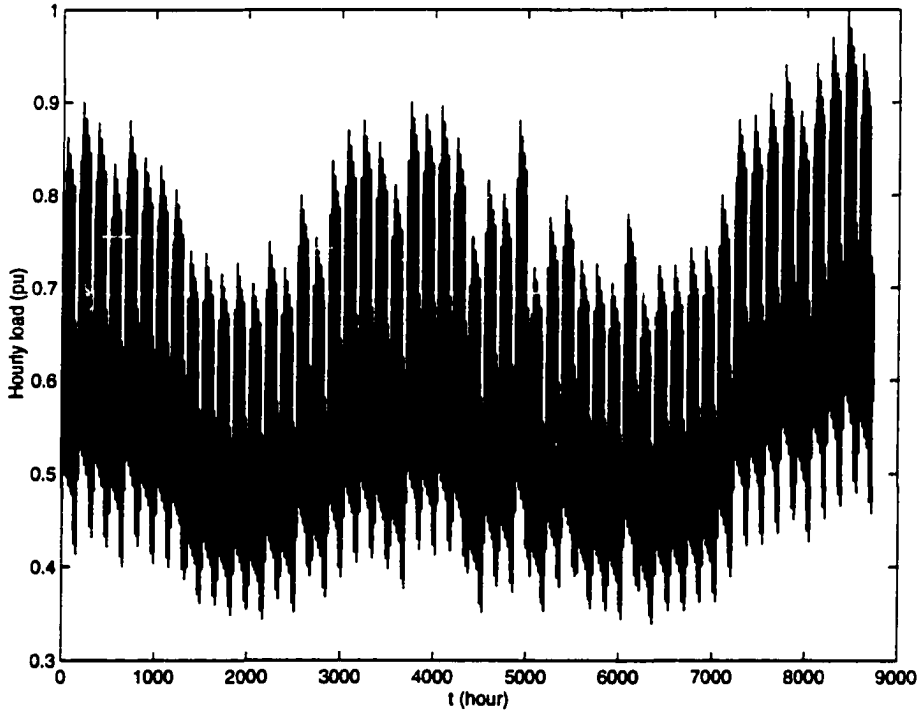

Figure 3.1 The expected load profile

There are many techniques available today to forecast load: time series, artificial neural network (ANN), fuzzy logic, chaos, etc. Time series forecasting theory has developed very fast during the last three decades, largely due to the requirement of describing economic and some specific physical behaviours. Box and Jenkins developed a systematic theory named after them in 1960's [23] [24] [25] [26]. The emergence of FFT in 1965 [27] brought us a powerful tool to handle stochastic digital time series signals in frequency domain. The widely used Kalman filter techniques [28] enable us to estimate a set of loads instead of one load point. A long list of literature reveals people's sustained interest in applying artifial neural network (ANN) in time series forecasting [29], [30], [31], [32], [33], [34], [35], [36]. Fuzzy logic theory [37] and chaos [38] theory have become competitive alternatives to time series and ANN. Methods combining ANN and fuzzy logic have been proposed [39] [40] [41] [42]. Some commercial programs such as NPREDICT are on sale [29], which enable determination of the point estimate and the variance at each time step. However, no forecast tool is perfect. Errors are inevitable. And the forecasting results of different models can not be simply compared by the magnitude of estimated errors (because error estimation comes from history data, not future data). However, a good estimator's forecasting error series should approximate a white noise, otherwise the error series can be further represented by time series models and the overall load forecasting model can be improved by doing so. Therefore, it is natural for us to decompose the load forecasting into two parts: one is for load model identification and the other is for error series analysis. In load model identification, any kind of model mentioned previously can be applied. In error series analysis, time series techniques can be applied to improve the forecasting by whitening the errors. Since time series theories are well developed and can identify the load model and whiten the error series simultaneously, we choose a typical time series model, i.e., the autoregressive integrated moving average (ARIMA) model to fit the data. Of course, we can also use ANN or fuzzy or chaos models to represent load series. However, the obtained error series may

not be a white noise and we have to use the time series model again on the error series to improve the forecasting accuracy.

We take the following steps to identify load forecasting error.

1. Perform spectrum analysis to detect the seasonality.
2. Detrend the seasonality.
3. For the detrended signal, use an autocorrelation function curve or partial autocorrelation function curve to determine the order of this signal.
4. Estimate paramters of AR and MA parts independently.
5. Use the identified model to forecast load and identify the one-day-ahead load forecasting error.

The forecasted load and 95% confidence interval for the next week is shown in Figure 3.2. From this figure we can see that the error is small during the next week. Therefore our forecast is accurate. The estimated square root load forecasting error vs. time curve is shown in Figure 3.3. From this figure we can see that the further away the future is, the more uncertain the forecast becomes.

We have assumed that each load value used in our trajectory has associated with it some error. This error characterizes the potential for deviation away from the load forecast for which the system coordinator (perhaps the independent system operator) is unable to make effective and economically efficient adjustments. We assume that such adjustments would be possible given more than a one-day advance warning by using the day-ahead electricity market, but they would not be possible for advance warning less than one day. Therefore, we use the estimated error of a one-day forecast in our work. This estimated error was computed by averaging the errors of a day-ahead load forecast as compared with historical data, over one year. The estimated one day ahead load forecasting error is 1.92% by our program. The details are shown in Appendix B.

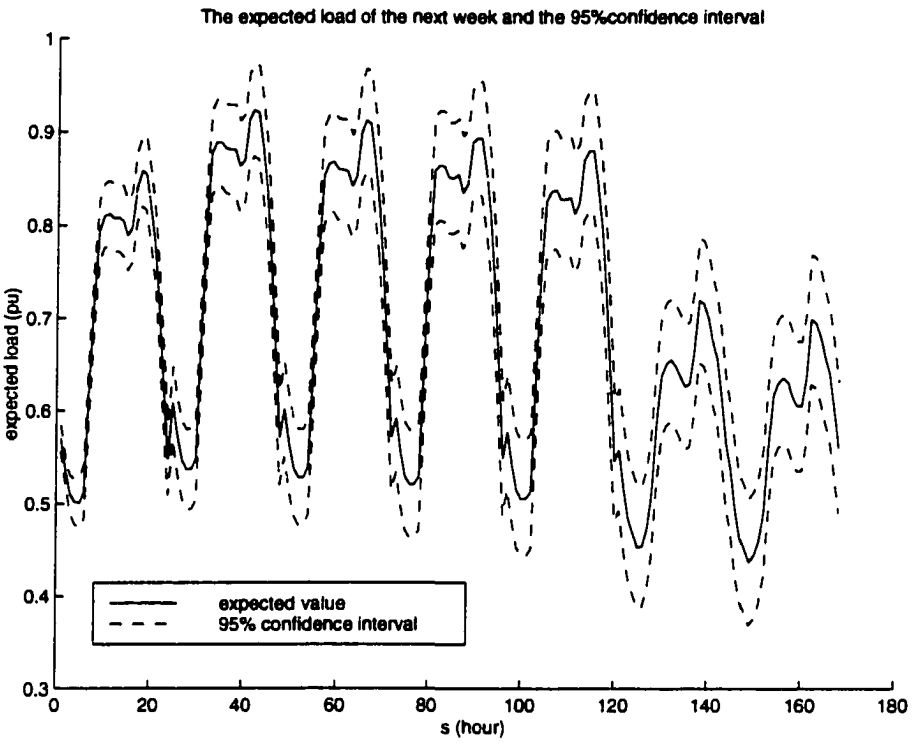

Figure 3.2 The forecasted load and 95% confidence interval for the next week

3.3 Maintenance Scheduling

Our maintenance schedule problem is: given the expected load profile, how to arrange maintenance schedules of generation units in order to minimize the economic cost? In order to narrow our problem, we assume only generation units need maintenance work, i.e., we do not consider maintenance of transmission circuits.

Maintenance scheduling for critical system components is a traditional problem investigated not only for power systems, but as well for other industries, where reliability and security are essential. It is still an important research issue today. Researchers try to solve two basic problems currently:

1. Optimization problem: given history data, how do we minimize a utility function, which is usually an economic cost?

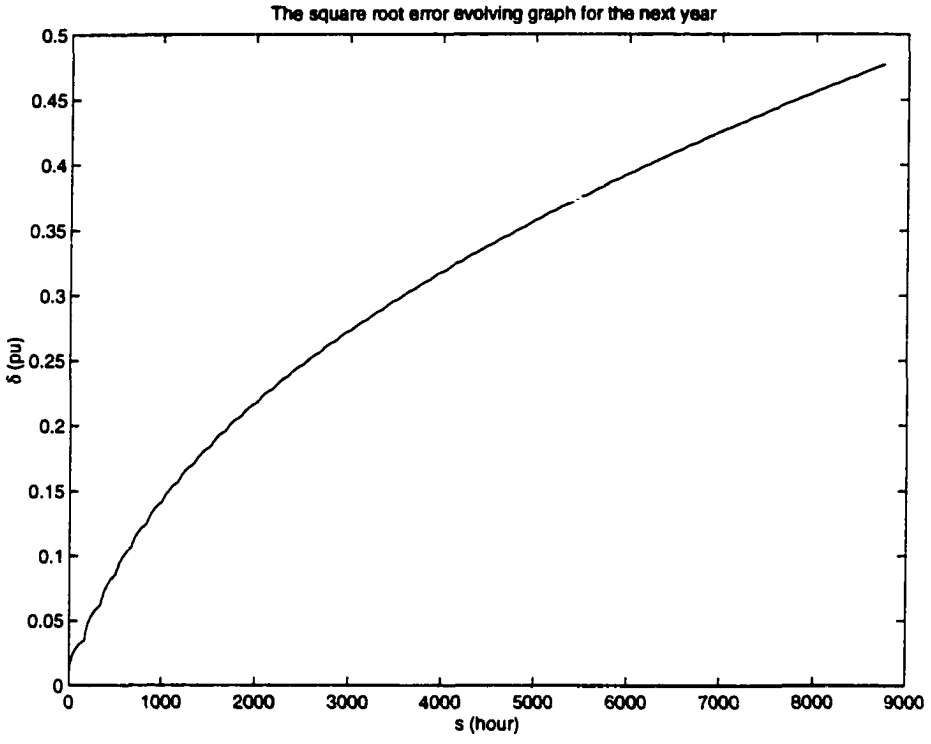

Figure 3.3 The estimated square root error vs. time curve

2. Adaptive adjustment: if there is an on-line monitor for the aging of materials, how do we adjust or rearrange maintenance schedule according to monitored data?

The optimization model is typically posed in terms of a stochastic programming problem [?], however, it can also be solved by a deterministic method [44]. Since the problem is usually very complicated, traditionally it is solved by applying equal reserve criterion or equal Loss Of Load Probability (LOLP) criterion [45]. Equal reserve criterion requires us to make the total reserve curve over a year as level as possible, on the other hand, equal LOLP criterion requires us to make the LOLP curve over a year as level as possible. There are several ways to make use of the equal LOLP criterion. The most accurate one is to form the integer 0-1 programming model, whose objective function is to minimize the summation of LOLP, or the variance of the LOLP curve. However, this method is complex and time consuming. Another traditionally used method is to make

use of the effective load carrying capacities of units to flatten the equivalent load profile. This is the one we used in our program. In order to avoid computational burden, the transmission constraints are ignored.

Recently, with the development of computers and computation techniques, more and more researchers take transmission constraints into account and solve the optimization problem directly. They solve the optimization problem by various methods. Some use iterative programming [44], some use simulated annealing [46], some use the Hopfield neural network [47]. Some researchers, on the other hand, investigate the influence of on-line monitoring on the maintenance scheduling [48] [49].

We prefer to apply equal LOLP criterion by using effective load carrying capacities of units because of the following reasons.

- Our main task is to develop and illustrate the whole planning stage risk assessment framework, not a single maintenance problem. Therefore traditional methods with simple algorithms and stable behavior are preferable than modern sophisticated algorithms.

- The method we adopt has been used for years. Its effectiveness is proven by time.

Other methods may be attractive and could be considered in the future.

We use the daily peak load curve here to arrange maintenance schedule because it is normally adopted in maintenance scheduling currently. Other methods based on hourly peak load curve may be attractive and could be considered in the future.

In order to introduce the basic idea of our method, we first introduce the concept of effective load carrying capacity. When a generation unit is started, not all of its capacity can be used as load reserve. Because it has forced outage rate, some of its capacity must be reserved to compensate for its outage and keep the system at the same LOLP level from a statistical point of view. What remains of its capacity is the effective capacity

to support load, called the effective load carrying capacity. The effective load carrying capacity of each unit can be calculated based on the system LOLP vs. reserve curve. The influence of a unit in maintenance to the power system is tantamont to increasing the load by its load carrying capacity during its maintenance time period. Under the equal LOLP criterion, we desire the equivalent load curve over a year to be as level as possible. Therefore, we take the following steps to arrange the maintenance schedule.

1. Reorder the effective load carrying capacities (ELCC) of generators from the largest to the smallest.

2. Pick up one ELCC from the list in order, look up for the required number of maintenance days (denoted by m).

3. Calculate all summations of the m successive load, find the smallest one, arrange unit there.

4. If this is the last unit, stop; else move to the next unit, go to 2.

Figure 3.4 shows the original daily load curve over a year. Figure 3.5 shows the equivalent daily load curve over a year after we have arranged the maintenance schedule. We see that the equivalent daily load curve is more level than the original daily load curve. It proves that our technique to levelize the system LOLP is effective.

3.4 Unit Commitment Arrangement

When load demand is low, it is not economic to turn all generators on. Therefore, the start up and shut down schedule of generation units should be carefully arranged to minimize the fuel consumption cost. This problem is called unit commitment problem. It can be expressed as a nonlinear optimization model with integer 0-1 variables.

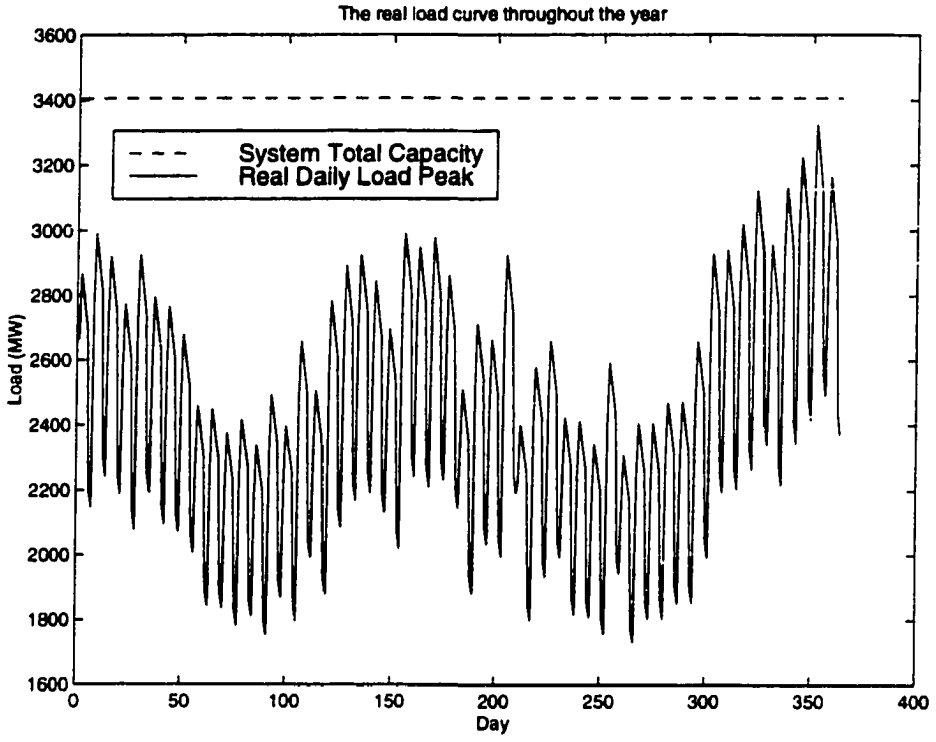

Figure 3.4 The daily peak load curve

In operations, unit commitment (UC) schedules are normally computed for only a few weeks (or less) in advance because of load uncertainty and because of the computational burden. However, the operational goal is to set a specific UC schedules, and therefore the assessment must accurately reflect the actual condition of the short term time interval. In using our one-year assessment, we are more interested in obtaining a trajectory that reflects expected conditions and is reasonable, for the purpose of probing the system's weaknesses. As a consequence, we use a very efficient, but approximate UC scheduling calculation to determine a full year's UC schedule. This schedule would of course serve as a long term plan which can be adjusted by more exact, weekly UC calculations.

Traditional unit commitment calculation algorithms such as Lagrange relaxation algorithm [50] and genetic algorithm [51] are short term oriented algorithms. When one year is considered, a dynamic programming calculation for one unit and one year will

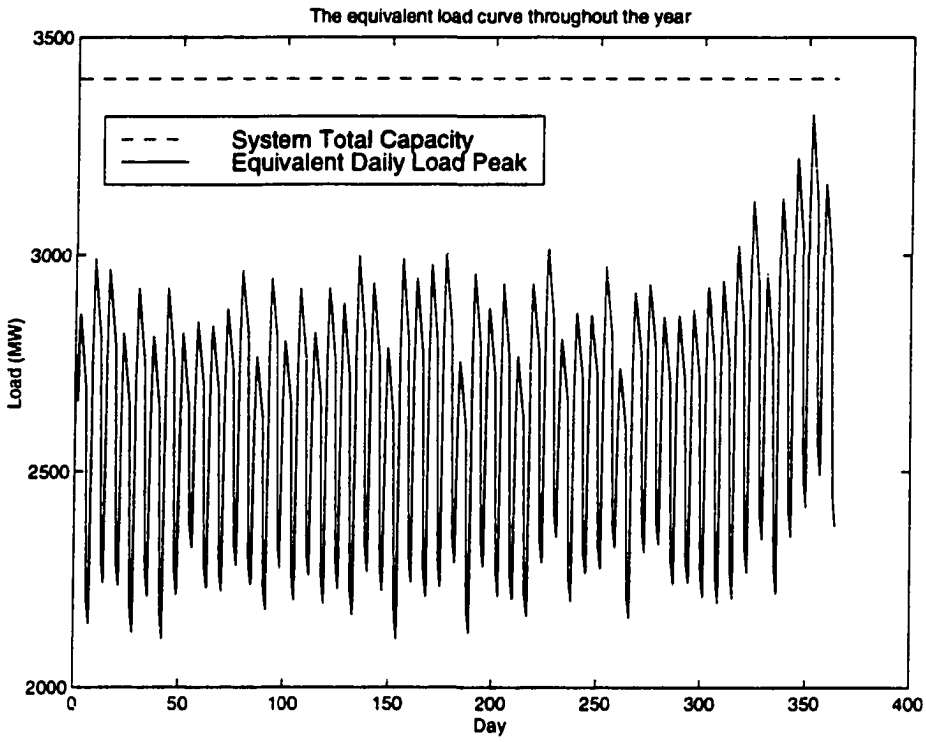

Figure 3.5 The daily equivalent load curve after maintenance arrangement

require 8 to 10 minites on a 166 MHz HP station. In each iteration of the Lagrange relaxation algorithm, the dynamic programming calculation has to be applied to each generation unit. There are tens of units in a normal power system. Furthermore, the Lagrange relaxation algorithm normally requires tens or even hundreds of iterations. Therefore, the Lagrange relaxation algorithm has to do thousands or even more than tens of thousands of the dynamic programming calculations. It is extremely time consuming to use Lagrange relaxation method. Genetic algorithm (GA) and dynamic programming are also computational for such a big problem. In this situation, the priority list method is very attractive [52]. It is simple, fast, and it can give us a feasible but suboptimal solution. There are no reported instances in the literature of unit commitment solved by priority list with hydro-thermal coordination. In this dissertation, we propose a technique to deal with hydro-thermal coordination. Furthermore, we adopt a piecewise

linear model for fuel consumption curve, which is more suitable in a deregulated power market. Our basic idea for the priority list method is as follows.

1. Since all fuel consumption curves of thermal units are linearized, the marginal costs of thermal units are within a limited set.

2. There is a system marginal price at each hour. It is the lowest acceptable price for generation companies to increase the total generation of the system.

3. The allocation of hydro energy can be adjusted by a psuedo price. If the hydro energy price is larger than the system marginal price at some hour, hydro energy should not be consumed at that hour. If the hydro energy price is smaller than the system marginal price at some hour, hydro energy should be fully used to substitute thermal energy at that hour. If the hydro energy price is the same as the marginal thermal energy price at some hour, we can allocate hydro power at will between its minimum and its maximum at that hour.

4. The psudo hydro energy price should be within the limited marginal cost set. We can pick up one possible psudo hydro energy price randomly from the limited marginal cost set first. Then the reasonable hydro energy consumption of a season should be a range covering the maximum hydro energy constraint. If the hydro energy consumption range is below the maximum hydro energy constraint, the chosen psudo hydro energy price is too expensive, we should lower it; if the hydro energy consumption range is above the maximum hydro energy constraint, the chosen psudo hydro energy price is too cheap, we should raise it. By adjusting the psudo hydro energy price, we can allocate hydro-thermal coordination and arrange unit commitment further by thermal-only priority list method.

As far as the IEEE RTS'96 system is concerned, according to our unit commitment arrangement, the system total fuel cost for one year is $\$ 2.54 \times 10^8$. The duality gap is

0.43. It means the real optimal fuel cost should be between $\$ 1.78 \times 10^8$ and $\$ 2.54 \times 10^8$. We are satisfied with our suboptimal solution. Moreover, the power flow calculations for all hours converge. The details of our unit commitment calculation are provided in Appendix D. In addition, we provide a review of Lagrange relaxation method in order to illuminate the reason for choosing the priority list method.

CHAPTER 4 POWER FLOW INFEASIBILITY RISK ASSESSMENT

4.1 Introduction

The concept of power flow infeasibilty risk is proposed in this dissertation. We describe this concept and then provide algorithmic details in the later sections. The solution of the power flow equations is a fundamental problem in power system analysis. We use the feasible region to denote the set of points where the power flow equations have a solution and all system parameters (e.g., line flows, bus voltages, generations) are within their limits. We use the infeasible region to denote the set of points where the power flow equations have a solution, but where one or more limits are violated. We use the unsolvable region to denote the set of points where the power flow equations have no real solution. Figure 4.1, proposed in reference [53], illustrates the concepts of the three security regions. Therefore, if we can not find a feasible power flow solution for a system state, we call the system state infeasible; if we can not find a power flow solution at all, we call the system state unsolvable.

Infeasibility has various causes, including

1. The load exceeds the available generation capacity.

2. The load exceeds the voltage instability level.

3. A bus voltage limit is violated.

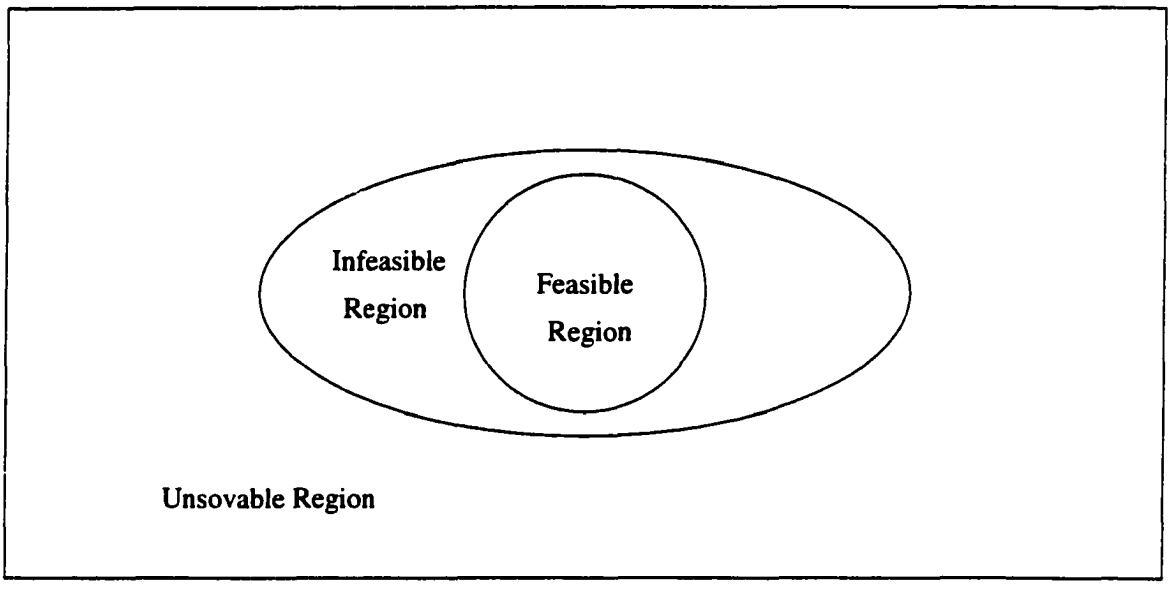

Figure 4.1 Power flow security regions

4. A branch flow limit is violated.

Of these, the first three can be alleviated by taking action to reduce the system demand. Because this action incurs a direct cost, it is an approximate measure of the impact corresponding to the condition which led to it. This comprises the fundamental idea of this section.

Given a trajectory of operating condition over a time interval, for each hour and each outage condition, we compute the maximum loadability of the system, a value that primarily depends on the unit commitment. The amount by which the hour's load exceeds the system maximum loadability is a measure of the impact corresponding to that loading and outage condition.

There are several related ideas required to complete the picture of the approach presented in this section, as follows.

Maximum Loadability Found By Optimization

We obtain maximum loadability for a given network and unit commitment by solving

a constrained optimization problem. In this problem, the objective function is the total demand, which is maximized, and constraints include the power flow equations, real and reactive generation limits, and bus voltage limits.

System Based or Zone Based

Load increase occurs at each bus in the system in proportion to the percentage of total load initially located at the bus. In some cases, however, it is more realistic to group load buses into zones, so that load changes in each zone are dictated by a specified zone weight between 0 and 1, where higher weights cause heavier load increases.

Reduced Number of Loadability Calculations

The maximum loadability must be found for each outage at each hour of the study period. However, since the calculation depends on only the network configuration (which includes the topology and the unit commitment), we need perform the calculation only for each outage and for each unit commitment. This greatly reduces the computational burden because the number of different unit commitment patterns is typically much less than 8760 hours.

Risk Calculation

We model each hour's total load as a normally distributed random variable having a mean equal to the forecasted value for the hour and standard deviation equal to one day ahead forecast error. Given the hour's maximum loadability, we can compute cost of the expected curtailed load. The expected curtailed load is found according to

$$E[L] = \int_{L_{max}}^{\infty} x f_L(x) dx \tag{4.1}$$

where L_{max} is the maximum loadability and $f_L(x)$ is the probability distribution for the load of the hour, as illustrated in Figure 4.2.

Overload Risk Not Included

The risk associated with overload is not included in the power flow infeasibility risk calculation. The reason is that overload risk is not directly alleviated by load curtailment.

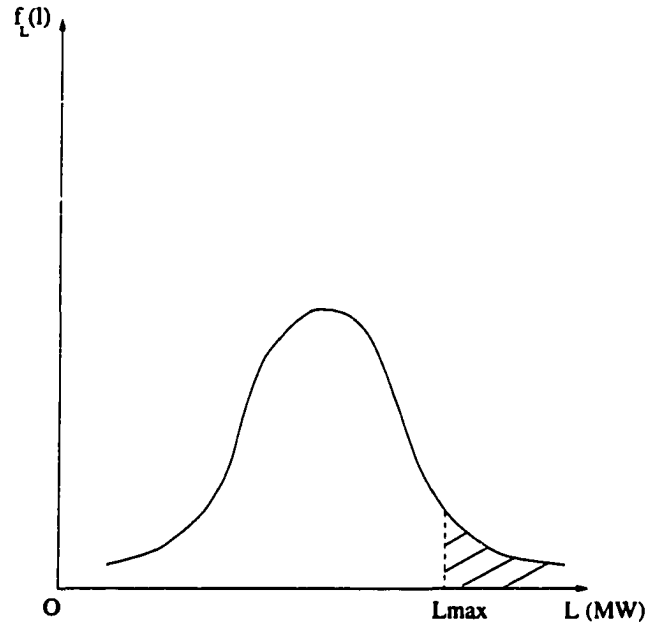

Figure 4.2 The expected curtailed load

Therefore, the load curtailment required to eliminate an overload may significantly overestimate the associated risk. The algorithm which we use to compute overload risk is described in the next chapter.

Risk Is Assignable

We can decompose the risk calculation to identify risk contributed from and incurred by each component and therefore each facility owner as well. In addition, shadow prices obtained from the optimization solution can be used to predict changes in risk caused by changes in system conditions.

4.2 Previous Work

Voltage stability reliability evaluation has been done by Aboreshaid and Billinton in [54], and Melo, Mello, and Granville in [55] with two different methods. Reference [54] uses two indices, one is a stability indicator and the other is a load multiplication factor. The main task is to calculate the indices, so that if the stability indicator

exceeds a threshold, the system is regarded as voltage unstable. The meaning of load multiplication factor is explained thoroughly in [58].

Reference [55] considers an entire year's load scenario with load uncertainty, and it uses the Monte Carlo simulation to generate system states. The interior point algorithm is employed to calculate the minimum load curtailment. Several reliability indices are calculated. However, for hourly risk assessment over a year, we believe that this approach requires prohibitive computation in order to achieve an acceptable accuracy.

As far as the interior point algorithm is concerned, references [56] and [57] proposes two different methods. We simplified their methods by eliminating some equations. This makes our equations about one forth the number of their equations.

4.3 Outage Model

Our outage model, which is used for both power flow infeasibility risk assessment and overload risk assessment, include single component outage for all branches and generators and also double line outage (i.e., "common mode") of transmission lines in the same corridor. Each outage constitutes a unique "state". The normal condition for which there is no outage, is also a state.

Denote N_s to be the total number of states considered, with the normal state numbered 1. Also, denote Pl_j as component j's outage probability in an hour. The probability of having state i occur in the next hour is denoted as P_i , given by

$$P_i = Pl_{i-1} \prod_{\substack{j=1 \\ j \neq i-1}}^{N_s} (1 - Pl_j) \tag{4.2}$$

We note that state i corresponds to component outage $i-1$ because here, we have numbered the normal state as 1. The probability of having the normal state in the next

hour is

$$P_1 = \prod_{j=2}^{N_s}(1 - Pl_j) \tag{4.3}$$

The common mode Markov chain model is shown in Figure 4.3 [12].

Denote λ_i, λ_j to be the failure rate of line i and j respectively. For thermal overload, permanent outage rates should be adopted because thermal overload is a relatively long time scale problem. Assuming the proportion for one line outage to develop to common mode outage is $\alpha = 0.075$ for all common mode failures, then the common mode outage rate is

$$\lambda_{ij} = \alpha(\lambda_i + \lambda_j) \tag{4.4}$$

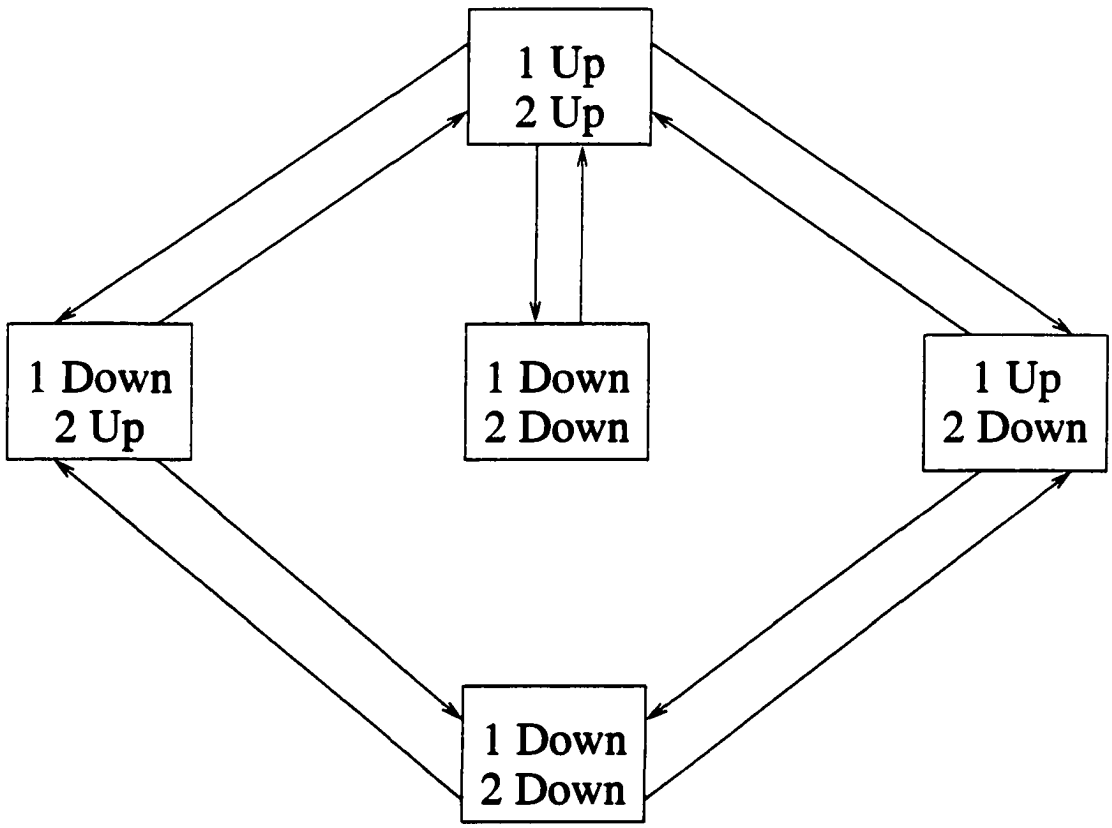

Figure 4.3 Common model Markov chain model

when the failure rates for the two lines are very small. The failures of a common mode outage often occur at the same place on the same tower. Assuming we have enough crews to repair failures, we can choose the repairing time as the maximum repairing time of line i and line j, or

$$\mu_{ij} = \min(\mu_i, \mu_j) \tag{4.5}$$

where μ is the repairing rate, it is the inverse of the repairing time. Thus we can calculate the forced outage rate (FOR) for the common mode as follows.

$$FOR = \frac{\lambda_{ij}}{\lambda_{ij} + \mu_{ij}} \tag{4.6}$$

Some contingencies may cause isolated islands. In our program, those islands are individually dealt with.

4.4 One Zone Model for Power System Maximum Loadability

Let us begin from a maximum loadability problem. Suppose all loads vary proportionally at any time. That means the whole power system is considered as a whole zone. The optimization problem is as follows

$$\max g(P_{D\Sigma}) \tag{4.7}$$

where $g(P_{D\Sigma}) \triangleq P_{D\Sigma} = \sum_{k=1}^{N} P_{Dk}$, subject to

$$P_{gk} - p_{dk}P_{D\Sigma} - P_{ek} = 0, \quad k = 1, \cdots, N \tag{4.8}$$

$$Q_{gk} - q_{dk}P_{D\Sigma} - Q_{ek} = 0, \quad k = 1, \cdots, N \tag{4.9}$$

$$P_{gkmin} \leq P_{gk} \leq P_{gkmax}, \qquad k \in \Omega_p \tag{4.10}$$

$$Q_{gkmin} \leq Q_{gk} \leq Q_{gkmax}, \qquad k \in \Omega_q \tag{4.11}$$

$$V_{kmin} \leq V_k \leq V_{kmax}, \qquad k \in \Omega_v \tag{4.12}$$

where

- $g(P_{D\Sigma})$ is the objective function of loads. In this case, it is the total summation of loads, which is also represented by $P_{D\Sigma}$.
- P_{gk} is the active generation power at bus k.
- Q_{gk} is the reactive generation power at bus k.
- p_{dk} is the ratio constant for $P_{Dk}/P_{D\Sigma}$, where P_{Dk} is the active load at bus k. We always have $\sum_{k=1}^{N} p_{dk} = 1$.
- q_{dk} is the ratio constant for $Q_{Dk}/P_{D\Sigma}$, where Q_{Dk} is the reactive load at bus k.
- P_{gkmin}, P_{gkmax}, Q_{gkmin}, Q_{gkmax}, V_{kmin}, V_{kmax} are lower and upper bounds for P_{gk}, Q_{gk} and V_k respectively.
- Ω_p is a set of buses which has P_g as a variable. We do not use the generation bus set because when $P_{gkmin} = P_{gk} = P_{gkmax}$, P_{gk} becomes a constant rather than a variable, bus k is then unnecessary to be included into Ω_p.
- Ω_q is a set of buses which has Q_g as a variable.
- Ω_v is a set of buses which has V as a variable.
- N is the number of buses.
- P_{ek}, Q_{ek} are the real and reactive power injections obtained from the power flow equations:

$$P_{ek} = V_k \sum_{l=1}^{N} Y_{kl} V_l \cos(\theta_k - \theta_l - \phi_{kl}) \tag{4.13}$$

$$Q_{ek} = V_k \sum_{l=1}^{N} Y_{kl} V_l \sin(theta_k - \theta_l - \phi_{kl}) \tag{4.14}$$

where Y is the admittance matrix of the network, θ is the bus voltage angle, $\phi_{kl} = \angle Y_{kl}$.

The model can be written in a generalized form as follows.

$$\min g(z, y) \tag{4.15}$$

subject to

$$h(z, y) = 0$$

$$z_l \leq z \leq z_u$$

where z is a variable vector with boundaries, y is a variable vector without boundaries. Instead of solving this problem directly, we solve the following alternative problem

$$\min g(z, y) - \mu \sum_j \ln s_{uj} - \mu \sum_j \ln s_{lj} \tag{4.16}$$

subject to

$$h(z, y) = 0 \tag{4.17}$$

$$z + s_u - z_u = 0 \tag{4.18}$$

$$z - s_l - z_l = 0 \tag{4.19}$$

This formulation results in solution by the direct interior point algorithm. The barrier items added in the objective function keep the solution point within the range. Denote (z^*, y^*) and (z^a, y^a) to be the optimal solutions of the two problems 4.15 and 4.16, respectively. Then if z^* is on the boundary, when μ is large, z^a tends to be deeper in the interior of the region, but the optimization function tends to be smoother, which makes the algorithm less prone to oscillate. When μ is small, z^a tends to be more approximate to z^* , but the optimization function has sharp turns near boundaries, which makes it hard to get the z^a solution. Therefore, a typical interior point algorithm always begins with a large μ , then makes μ smaller and smaller. Theoretically when $\mu \to 0$, $z^a \to z^*$. In practice, when μ is small enough, we can terminate the calculation. The Lagrange

function for problem 4.16 is as follows.

$$L = g(z,y) - \mu \sum_j \ln s_{uj} - \mu \sum_j s_{lj} - \lambda^T h(z,y) - \pi_l^T (z - s_l - z_l) - \pi_u^T (z + s_u - z_u) \tag{4.20}$$

The KKT first order optimality condition gives the following equations.

$$\nabla_z L = \nabla_z g - J_z^T \lambda - \pi_l - \pi_u = 0 \tag{4.21}$$

$$\nabla_y L = \nabla_y g - J_y^T \lambda - J_y^T \lambda = 0 \tag{4.22}$$

$$\nabla_{s_u} L = -\mu S_u^{-1} e - \pi_u = 0 \tag{4.23}$$

$$\nabla_{s_l} L = -\mu S_l^{-1} e + \pi_l = 0 \tag{4.24}$$

$$\nabla_{\pi_u} L = -(z + s_u - z_u) = 0 \tag{4.25}$$

$$\nabla_{\pi_l} L = -(z - s_l - z_l) = 0 \tag{4.26}$$

$$\nabla_{\lambda} L = -h(z,y) = 0 \tag{4.27}$$

where $J_z = \frac{\partial h(z,y)}{\partial z}$, $J_y = \frac{\partial h(z,y)}{\partial y}$, $e = [1, 1, \cdots, 1]^T$, $S_u \triangleq diag(s_{u1}, \cdots, s_{um})$, $S_l \triangleq diag(s_{l1}, \cdots, s_{lm})$. Substituting expressions from our model into these equations, we get

$$-1 + \sum_{k=1}^{N} \lambda_{pk} p_{dk} + \sum_{k=1}^{N} \lambda_{qk} q_{dk} = 0 \tag{4.28}$$

$$-\lambda_{pk} - \pi_{pk}^l - \pi_{pk}^u = 0, \; k \in \Omega_p \tag{4.29}$$

$$-\lambda_{qk} - \pi_{qk}^l - \pi_{qk}^u = 0, \; k \in \Omega_q \tag{4.30}$$

$$\sum_{l=1}^{N} \lambda_{pl} \frac{\partial P_{el}}{\partial \theta_k} + \sum_{l=1}^{N} \lambda_{ql} \frac{\partial Q_{el}}{\partial \theta_k} = 0, \; k \in \Omega_\theta \tag{4.31}$$

$$\sum_{l=1}^{N} \lambda_{pl} \frac{\partial P_{el}}{\partial V_k} + \sum_{l=1}^{N} \lambda_{ql} \frac{\partial Q_{el}}{\partial V_k} - \pi_{vk}^l - \pi_{vk}^u = 0, \; k \in \Omega_v \tag{4.32}$$

$$P_{gk} - p_{dk}P_{D\Sigma} - P_{ek} = 0,\ k = 1, \cdots, N \tag{4.33}$$

$$Q_{gk} - q_{dk}P_{D\Sigma} - Q_{ek} = 0,\ k = 1, \cdots, N \tag{4.34}$$

$$\pi_{uk} = -\frac{\mu}{s_{uk}} \tag{4.35}$$

$$\pi_{lk} = -\frac{\mu}{s_{lk}} \tag{4.36}$$

$$z - s_l - z_l = 0 \tag{4.37}$$

$$z + s_u - z_u = 0 \tag{4.38}$$

where $\pi_u = [\pi_p^u, \pi_q^u, \pi_v^u]^T$, $\pi_l = [\pi_p^l, \pi_q^l, \pi_v^l]^T$, $z = [P_g, Q_g, V]^T$, $\lambda = [\lambda_p, \lambda_q]^T$. Ω_θ is the set of all buses except the swing bus.

4.5 Model Reduction

We can use a Newton-Raphson algorithm to solve these equations. However, the model is of very large dimension. In the worst case, it has 16N-3 variables and the same number of equations. It is possible to reduce the dimensionality by substituting some variables with others and eliminating some equations.

From 4.35 to 4.38 we get

$$\pi_k \triangleq \pi_k^u + \pi_k^l = \mu\left(\frac{1}{z_k - z_k^l} - \frac{1}{z_k^u - z_k}\right) \tag{4.39}$$

Substituting it into 4.29, 4.30 we get

$$\begin{cases} \lambda_{pk} = -\pi_{pk}, & k \in \Omega_p \\ \lambda_{qk} = -\lambda_{qk}, & k \in \Omega_q \end{cases} \tag{4.40}$$

We can always find π_{pk} and π_{qk} such that

$$\begin{cases} \lambda_{pk} = -\pi_{pk}, & k \notin \Omega_p \\ \lambda_{qk} = -\pi_{qk}, & k \notin \Omega_q \end{cases} \tag{4.41}$$

Eliminating λ in 4.28, 4.31, 4.32 we get

$$\sum_{l=1}^{N} \pi_{pl} \frac{\partial P_{el}}{\partial \theta_k} + \sum_{l=1}^{N} \pi_{ql} \frac{\partial Q_{el}}{\partial \theta_k} = 0, \qquad k \in \Omega_\theta \tag{4.42}$$

$$\sum_{l=1}^{N} \pi_{pl} \frac{\partial P_{el}}{\partial V_k} + \sum_{l=1}^{N} \pi_{ql} \frac{\partial Q_{el}}{\partial V_k} + \pi_{vk} = 0, \qquad k \in \Omega_\theta \tag{4.43}$$

$$P_{gk} - p_{dk} P_{D\Sigma} - P_{ek} = 0, \quad k = 1, \cdots, N \tag{4.44}$$

$$Q_{gk} - q_{dk} P_{D\Sigma} - Q_{ek} = 0, \quad k = 1, \cdots, N \tag{4.45}$$

$$1 + \sum_{k=1}^{N} \pi_{pk} p_{dk} + \sum_{k=1}^{N} \pi_{qk} q_{dk} = 0 \tag{4.46}$$

where

$$\pi_{pl} = \begin{cases} \mu\left(\frac{1}{P_{gl} - P_{glmin}} - \frac{1}{P_{glmax} - P_{gl}}\right) & , \ l \in \Omega_p \\ a\ variable & , \ l \notin \Omega_p \end{cases} \tag{4.47}$$

$$\pi_{ql} = \begin{cases} \mu\left(\frac{1}{Q_{gl} - Q_{glmin}} - \frac{1}{Q_{glmax} - Q_{gl}}\right) & , \ l \in \Omega_q \\ a\ variable & , \ l \notin \Omega_q \end{cases} \tag{4.48}$$

This set of equations only has $4N-1$ variables at the most, about 1/4 of the original set of equations. If we sum all equations in expression 4.44 together we get

$$P_{D\Sigma} = P_{g\Sigma} - \sum_{k=1}^{N} P_{ek} \tag{4.49}$$

Substitute it into 4.42 and 4.43, eliminate the equation in 4.42 corresponding to the swing bus, then we further reduce the number of equations to be at most $4N-2$.

4.6 Solution Algorithm and Shadow Prices

Denote J_b to be the big Jacobian matrix for 4.42-4.46, then we have

$$\begin{bmatrix} \Delta z^k \\ \Delta y^k \end{bmatrix} = -J_b^{-1} f(z^k, y^k) \tag{4.50}$$

at every step, where $f(\cdots,\cdots)$ is the left hand side function vector of the equations. However, we have to keep z^{k+1} within the feasible region. Therefore, every time we have to choose maximum α such that $z^k + \alpha\Delta z^k \in S$, where S is the feasible region. Once we get α, we choose z^{k+1} to be between z^k and $z^k + \beta\alpha\Delta z^k$, where $\beta = 0.9995$. We desire to choose z^{k+1} such that $\|f(z^{k+1}, y^{k+1})\|_\infty$ is minimized, or at least less than $\|f(z^k, y^k)\|_\infty$. The algorithm we adopt is as follows.

1. Pick up a starting value $\mu = \mu_0$, calculate $f_0 = \|f_{\mu_0}(z_0, y_0)\|_\infty$, set $f_1 = f_0\|$.
2. If after 3 iterations $\|f\|_\infty$ reduces less than $10f_0 = \|f_\mu(z, y)\|_\infty$.
3. If $\|f\|_\infty \leq \frac{1}{10}f_1$, or $\|f\|_\infty \leq 0.01$, set $\mu = \frac{1}{10}\mu$.
4. If $\|f\|_\infty \leq \varepsilon$ and $\mu \leq \mu_{min}$, stop, where ε and μ_{min} are small positive numbers chosen by the user; otherwise go to 2.

Our approach calculates maximum loadability due to voltage stability, voltage limits, or generation limits. We also desire to identify binding constraints on the objective. The "shadow price" in our optimization model provides us full information for such discrimination.

The "shadow price" refers to $-\pi_p, -\pi_q, -\pi_v$ in our model. From previous derivation, we know for a variable z_k,

$$\pi_{zk} = \mu\left(\frac{1}{z^k - z^{kmin}} - \frac{1}{z_{kmax} - z_k}\right) \tag{4.51}$$

Therefore, when z_k hits the upper bound, $\pi_{zk} < 0$; when z_k hits the lower bound, $\pi_{zk} > 0$; when z_k is far away from both bounds, π_{zk} is very near 0 as μ is near zero. On the other hand, according to optimization theories, π_{zk} can be interpreted as the decrement of the Lagrange function due to the per unit increment of the z_{kmax} when z^k hits the upper bound, or the decrement of the Lagrange function due to the per unit decrement of the z_{kmin} when z_k hits the lower bound. Therefore it reflects the

price of the confined boundary. When z_k hits the upper bound, as z_{kmax} increases Δz_k, the total load will increase $-\pi_{zk}\Delta z_k$, provided Δz_k is sufficiently small; when Δz_k hits the lower bound, as z_{kmin} increase Δz_k, the total load will decrease $\pi_{zk}\Delta z_k$, provided Δz_k is sufficiently small. Therefore $-\pi_{zk}$ always reflects a "price" of the corresponding boundary.

In our program, we use the following algorithm to identify the binding constraints.

1. If $\pi_{pk} < -\varepsilon_p$ for all $k \in \Omega_p$, where ε_p is a positive threshold value, set type=1, the maximum loadability (ML) is confined by generation constraints.

2. Otherwise, find the bus with the lowest voltage, denote the bus as bus m.

   - If $\pi_{vm} < \varepsilon_v$, where ε_v is a positive threshold value, set type=2, the ML is confined by voltage stability.
   - Otherwise, set type=3, the ML is confined by voltage lower bound.

4.7 Illustration of One Zone Model on IEEE RTS'96

VSTAB contains a standard continuation power flow calculation program developed by Powertech Labs Incorparated [59]. We compare results from our software with those from VSTAB in order to provide validation evidence.

For the typical case of the IEEE RTS'96 system given in A.1, we choose bus 13 as the swing bus, relieve the generation limits at this bus to be ineffective, and fix all other active generation power. We also choose $V_{max} = 1.05pu$, $V_{min} = 0.9pu$ for all generation buses, and $V_{max} = 1.15pu$, $V_{min} = 0.85pu$ for all load buses. The maximum loadability calculated by our program is 4048 MW, while the VSTAB's calculation result is 4025 MW. This shows that our software for maximum loadability calculation is correct.

Our approach is very attractive because, for a specified unit commitment, it can also identify the dispatch that maximizes loadability. As an illustration, we repeat the

previous calculation with real power injection at each generation bus defined as a decision parameter. This time, the maximum loadability increases to 4208 MW as a result of the optimized dispatch. Furthermore, we can activate the generation limits on the swing bus, which is required for a realistic solution, to identify maximum loadability of 3358 MW. This value is 47 MW less than the installed capacity. This difference is due to losses in the transmission network.

Now suppose we change the unit commitment to be as shown in Table 4.1, and change the swing bus to bus 23. V_{max} is set to 1.10pu for generation buses, and 1.15pu for load buses. V_{min} is set to 0.90pu for generation buses, and 0.85pu for load buses. The swing bus's voltage is set to 1.05pu. The $P_g, Q_g, V, \pi_p, \pi_q, \pi_v$ for all buses are shown in Table 4.2.

Table 4.1 Unit commitment pattern for analysis I(in MW or MVar)

Gen. bus	P_{gmax}	P_{gmin}	Q_{gmax}	Q_{gmin}
15	155	54.25	80	-50
16	155	54.25	80	-50
18	400	100.00	200	-50
21	400	100.00	200	-50
22	250	0.00	80	-50
23	660	248.50	310	-125

From this table, we can see π_p is near 0 for every generation bus. It means we can not increase the maximum loadability simply by increasing generation capacity at any generation bus. The ML is not confined by generation capacity. Now look at bus voltages. The minimum bus voltage is 0.85pu, corresponding to bus 7. The voltage hit the lower bound, therefore the corresponding π_v is as high as 12.62. It means the ML is confined by the voltage lower bound. If we reduce the lower bound to 0.8pu, the ML increases from 1870 MW to 1905MW. Bus 7 remains to have the lowest voltage

Table 4.2 Shadow prices

bus	P_g	Q_g	V(pu)	π_p	π_q	π_v
1	0	0	0.9574	-1.3654	-1.7302	0.0009
2	0	0	0.9573	-1.3849	-1.7385	0.0009
3	0	0	0.9828	-0.5778	-1.2639	0.0003
4	0	0	0.9507	-1.2107	-1.7204	0.0010
5	0	0	0.9638	-1.3028	-1.7642	0.0007
6	0	0	0.9742	-1.3236	-1.7868	0.0005
7	0	0	0.8500	-3.4305	-4.5938	12.6217
8	0	0	0.8786	-2.5304	-3.5117	0.0066
9	0	0	0.9712	-0.8525	-1.6249	0.0006
10	0	0	0.9896	-1.0572	-1.7493	0.0002
11	0	0	0.9914	-0.5326	-1.1177	0.0002
12	0	0	0.9691	-0.5114	-1.1127	0.0006
13	0	0	0.9808	-0.4275	-0.9025	0.0004
14	0	200.0	1.0566	-0.2927	-0.6958	-0.0035
15	154.5	79.9	1.0746	-0.0392	-0.1939	-0.0071
16	154.6	79.9	1.0741	-0.0516	-0.1963	-0.0069
17	0	0	1.0910	-0.0144	-0.0632	-0.0027
18	390.6	83.6	1.1000	-0.0022	0.0000	-4.9412
19	0	0	1.0596	-0.0551	-0.1146	-0.0013
20	0	0	1.0524	-0.0516	-0.0410	-0.0011
21	374.7	64.0	1.1000	-0.0008	0.0000	-8.6093
22	208.8	-33.1	1.0997	-0.0004	0.0011	-0.6943
23	659.6	110.7	1.0500	-0.0486	0.0000	—
24	0	0	1.0122	-0.1729	-0.6187	-0.0002

0.8197pu. This time, we come to the voltage stability saddle node. Now we investigate π_q for generation buses, we find bus 14's $\pi_q = -0.6864$, which is absolutely largest. It means bus 14's reactive power is not enough. If we increase bus 14's reactive power upper limit to 300 MVar, the ML increases to 1940MW. Q_{g14} becomes 299.52 MVar, very near the upper bound. However, π_{q14} is only -0.0438, very near zero. Therefore, further increment in Q_{g14}'s upper limit will not cause too much increment in the ML. In fact, when the upper limit increases to 400MVar, the ML becomes 1941MW, only 1MW bigger than the 300 MVA's case. From this example, we can see that shadow prices are much better measurements for the ML type discrimination than distances from boundaries.

From equations 4.47 and 4.48, it seems different μ may cause different "shadow prices". However, as $\mu \to 0$, the "shadow prices" keeps constant when they are not near zero. For the previous example, as the final μ_{min} varies from 0.01 to 1×10^{-6}, the corresponding ML and π_p values are shown in Table 4.3. From the table, we can see when $\mu_{min} \leq 0.0001$, the ML result does not change too much when μ_{min} reduces, all π_p values almost keep constant except for some absolutely very small values, which vanish very quickly(e.g., bus 21, 22). Therefore, choosing small enough μ_{min} will guarantee sufficient accuracy. In our program, we choose

$$\mu_{min} = \frac{0.02}{n} \tag{4.52}$$

where n is the number of variables.

4.8 Zone-Based Maximum Loadability

In many applications, it is of interest to study the maximum loadability for only a subarea in a larger model. Also, it may be of interest to weight one area more heavily than another regarding load increase, to reflect actual load increase differences in the

Table 4.3 The ML and π_p for different μ_{min}

bus	μ_{min}				
	0.01	0.001	0.0001	0.00001	0.000001
1	-1.3520	-1.3646	-1.3655	-1.3656	-1.3656
2	-1.3706	-1.3841	-1.3850	-1.3851	-1.3852
3	-0.5980	-0.5789	-0.5776	-0.5775	-0.5775
4	-1.2043	-1.2104	-1.2108	-1.2108	-1.2108
5	-1.2913	-1.3021	-1.3029	-1.3029	-1.3029
6	-1.3101	-1.3228	-1.3237	-1.3238	-1.3238
7	-3.2954	-3.4221	-3.4315	-3.4324	-3.4325
8	-2.4514	-2.5256	-2.5310	-2.5315	-2.5316
9	-0.8611	-0.8530	-0.8524	-0.8524	-0.8524
10	-1.0556	-1.0571	-1.0572	-1.0572	-1.0572
11	-0.5578	-0.5341	-0.5324	-0.5322	-0.5322
12	-0.5391	-0.5132	-0.5111	-0.5110	-0.5110
13	-0.4591	-0.4295	-0.4272	-0.4270	-0.4270
14	-0.3280	-0.2948	-0.2925	-0.2923	-0.2922
15	-0.0875	-0.0420	-0.0388	-0.0385	-0.0384
16	-0.0991	-0.0544	-0.0512	-0.0509	-0.0509
17	-0.0631	-0.0173	-0.0141	-0.0138	-0.0137
18	-0.0527	-0.0052	-0.0018	-0.0015	-0.0014
19	-0.1032	-0.0581	-0.0547	-0.0545	-0.0544
20	-0.1005	-0.0547	-0.0512	-0.0509	-0.0509
21	-0.0513	-0.0037	-0.0004	-0.0001	0.0000
22	-0.0461	-0.0030	-0.0002	0.0000	0.0000
23	-0.0980	-0.0518	-0.0483	-0.0480	-0.0479
24	-0.2147	-0.1753	-0.1727	-0.1724	-0.1723
$P_{D\Sigma}$	1861.68	1869.14	1870.02	1870.12	1870.13

system. Finally, in the case where the maximum loadability optimization algorithm does not converge, we would like to curtail in order to find a loading that does converge. In doing so, it is attractive if we can identify specific areas of the system where load curtailment is most effective. This also reflects that the loads of different areas can have different interruption priority, as indicated by the prices they pay. The optimization model for the zone-based maximum loadability problem is

$$\max P_{D\Sigma} = \sum_{i=1}^{d} w_i P_{D\Sigma i} \tag{4.53}$$

subject to

$$P_{gk} - P_{Dk}^0 - p_{dk} P_{D\Sigma i} - P_{ek} = 0, \qquad k \in D_i, \tag{4.54}$$

$$Q_{gk} - Q_{Dk}^0 - q_{dk} P_{D\Sigma i} - Q_{ek} = 0, \qquad k \in D_i, \tag{4.55}$$

$$i = 1, \cdots, d$$

$$P_{gkmin} \le P_{gk} \le P_{gkmax}, \qquad k \in \Omega_p \tag{4.56}$$

$$Q_{gkmin} \le Q_{gk} \le Q_{gkmax}, \qquad k \in \Omega_q \tag{4.57}$$

$$V_{kmin} \le V_k \le V_{kmax}, \qquad k \in \Omega_v \tag{4.58}$$

where $i = 1, \cdots, d$, D_i is the ith zone bus set, P_{Dk}^0, Q_{Dk}^0 are the constant active and reactive parts of the load at bus k. The simplified KKT first order equations are as follows.

$$\sum_{l=1}^{N} \pi_{pl} \frac{\partial P_{el}}{\partial \theta_k} + \sum_{l=1}^{N} \pi_{ql} \frac{\partial Q_{el}}{\partial \theta_k} = 0, \; l \in \Omega_\theta \tag{4.59}$$

$$\sum_{l=1}^{N} \pi_{pl} \frac{\partial P_{el}}{\partial V_k} + \sum_{l=1}^{N} \pi_{ql} \frac{\partial Q_{el}}{\partial V_k} + \pi_{vk} = 0, \tag{4.60}$$

$$k \in \Omega_v$$

$$P_{gk} - P_{Dk}^0 - p_{dk}P_{D\Sigma i} - P_{ek} = 0, \quad k = 1, \cdots, N \tag{4.61}$$

$$Q_{gk} - Q_{Dk}^0 - q_{dk}P_{D\Sigma i} - Q_{ek} = 0, \quad k = 1, \cdots, N \tag{4.62}$$

$$-\frac{\partial g}{\partial P_{D\Sigma i}} + \sum_{k \in D_i} \pi_{pk} p_{dk} + \sum_{k \in D_i} \pi_{qk} q_{dk} = 0, \; i = 1, \cdots, d \tag{4.63}$$

When upper and lower bounds are applied to $P_{D\Sigma_i}$ such that

$$0 \leq P_{D\Sigma_i} \leq P_{D\Sigma imax} \tag{4.64}$$

then the problem becomes the minimum load curtailment problem, i.e., for the given load level, we want to calculate the minimum load curtailment required to meet all constraints.

$$g(P_{D\Sigma}) = -\sum_{i=1}^{d} w_i (P_{D\Sigma imax} - P_{D\Sigma i}) \tag{4.65}$$

For the minimum load curtailment problem, only 4.63 needs to be changed in the KKT equations as follows.

$$-\frac{\partial g}{\partial P_{D\Sigma i}} + \sum_{k \in D_i} \pi_{pk} P_{dk} + \sum_{k \in D_i} \pi_{qk} q_{dk} + \pi_{Di} = 0, \; i = 1, \cdots, d \tag{4.66}$$

where $\pi_{Di} = \mu(\frac{1}{P_{D\Sigma i}} - \frac{1}{P_{D\Sigma imax} - P_{D\Sigma i}})$.

4.9 Illustration of Zone Based ML Algorithm on IEEE RTS'96

Supppose the unit commitment pattern for the IEEE RTS'96 is as shown in Table 4.4. When only line 6–10 is outaged, the program fails to converge if we assume that all loads change proportionally. Now we divide the system into 4 zones.

- Zone 1 includes buses 11–24. It is mainly a generation provider.

Table 4.4 Unit commitment pattern for analysis II(in MW or MVar)

Gen. bus	P_{gmax}	P_{gmin}	Q_{gmax}	Q_{gmin}
14	0	0.00	200	-50
15	155	54.25	80	-50
16	155	54.25	80	-50
18	400	100.00	200	-50
21	400	100.00	200	-50
22	250	0.00	80	-50
23	660	248.50	310	-125

- Zone 2 includes buses 1–5,9,10. It is mainly a load consumer.
- Zone 3 includes buses 7,8. It is a heavy load center loosely connected with the system.
- Zone 4 includes bus 6. Since the shunt reactor is not removed after line 6–10's outage, bus 6's load is expected to suffer low voltage.

After the division, we use our program to calculate the maximum loadability by setting $w_i \triangleq -\frac{\partial g}{\partial P_{D\Sigma i}} = 1$ for all $i = 1, \cdots, 4$. The calculation result is shown in Table 4.5.

Zone 2 and zone 3 has negative load, which is not reasonable. When $w = [1, 10, 10, 20]$, zone 1–4 have loads [856.1, 744.6, 104.9, 4.1]MW respectively, the total load becomes 1709.7 MW.

Table 4.5 The maximum loadability under the unit commitment of Table 4.4

	zone 1	zone 2	zone 3	zone 4	total
$P_{D\Sigma i}$(MW)	2097.3	-90.5	-19.3	15.4	1998.4

Now suppose the required total load level is 1797.4 MW. Assigning the load proportionally to each zone we get $P_{D\Sigma max} = [957.3, 567.6, 186.7, 85.8]$ MW. Set $w = [1, 1, 1, 1]$ and solve the minimum load curtailment problem, we get $P_{D\Sigma} = [957.3, 219.3, 132.9, 0]$ MW. The total load is 1309.5 MW. Zone 1's load has no curtailment, and zone 4 has no load. It means we must remove the shunt reactor simultaneously when the cable 6–10 is outaged; otherwise bus 6 will suffer low voltage. If we do this, $P_{D\Sigma} = [957.3, 567.6, 149.6, 42.2]$ MW. The total load now becomes 1716.6 MW, which is 407.1 MW more than when we did not remove the shunt reactor.

Previously, we introduced the meaning of shadow prices. Now π_D has the same meaning. For example, for the unit commitment in Table 4.4, we have

$$\pi_D = [-0.9294, 0, -0.0002, 4.1260]$$

Since $\pi_{D1} < 0$ and is large in magnitude, we expect $P_{D\Sigma 1}$ to hit the upper limit 957.3 MW. Since $\pi_{D4} > 0$ and is large in magnitude, we expect $P_{D\Sigma 4}$ to hit the lower limit 0. These are confirmed by our calculation. For example, if we raise $P_{D\Sigma 1max}$ by 500 MW to 1767.8 MW. $P_{D\Sigma} = [1457.3, 204.2, 106.3, 0.1]$ MW, and the total load becomes 1767.8 MW, 458.3 MW more than the original one. However, if we raise $P_{D\Sigma 2max}$ to be 500 MW more, then the total load is 1390.5 MW, almost unchanged compared with the original result. This illustrates that the shadow prices tell us which limits are more important than others.

Now we can explain the meaning of π_{pk} when $k \notin \Omega_p$, and the meaning of π_{qk} when $k \notin \Omega_q$. They are the shadow prices of P^0_{Dk} and Q^0_{Dk}, the constant parts of load. It means when $\pi_{pk} < 0$, decreasing P^0_{Dk} will increase the maximum loadability; when $\pi_{pk} > 0$, increasing P^0_{Dk} will increase the maximum loadability. The same comment is applied to π_{qk} and Q^0_{Dk}.

This time we change our unit commitment pattern to be as shown in Table 4.6. Assume the total required load is 0.95 times total generation, and assign loads propor-

tionally in zones, that is $P_{D\Sigma max} = [1571.1, 931.5, 306.4, 140.8]$ MW. Our calculation result is $P_{D\Sigma} = [1571.1, 931.5, 285.0, 140.7]$ MW. The total load curtailment is 21.4 MW. Suppose there are tie lines connected between buses 6,7,8,13,23 and the outer power systems respectively. Table 4.7 shows the shadow prices for tie lines.

Table 4.6 Unit commitment pattern for analysis III (in MW or MVAR)

Gen. bus	P_{gmax}	P_{gmin}	Q_{gmax}	Q_{gmin}
1	192	62.00	80	-50
2	192	62.00	80	-50
13	591	206.85	240	0
14	0	0.00	200	-50
15	215	66.25	110	-50
16	155	54.25	80	-50
18	400	400.00	200	-50
21	400	400.00	200	-50
22	300	0.00	96	-60
23	660	248.50	310	-125

Table 4.7 Shadow prices for tie line flows

Tie line bus	π_p	π_q
6	-0.0816	-0.1589
7	-0.9846	-1.4436
8	-0.5999	-0.9711
13	-0.0036	-0.0004
23	-0.0003	0.0001

If the price for MW or Mvar input at each tie line bus is the same, we should buy reactive power from the tie line at bus 7, because it is the most valuable power (shadow price -1.4436, the negative sign means increasing the input tie line flow will increase the ML). Now decrease Q^0_{D7} by 10 MVar and recalculate. We get $P_{D\Sigma} =$

$[1571.1, 931.5, 299.1, 140.7]$ MW, the total load curtailment is 7.3 MW, 14.1 MW less than the oringinal case. However, if we buy active power 10 MW from bus 23's tie line, $P_{D\Sigma} = [1571.1, 931.5, 285.0, 140.7]$ MW, the total load curtailment is 21.4 MW, almost unchanged compared with the oringinal one.

4.10 Expansion of Our Algorithm to ATC Calculation

Our algorithm can also be applied to available transfer capability (ATC) calculation from the power flow feasibility point of view, either including or excluding thermal constraints. This application is not directly employed in our power flow infeasibility risk assessment, because in our risk assessment case only maximum loadability should be identified for each zone in every hour. This section is written here to show the power of our algorithm and for the purpose of completeness.

Suppose now the system is running at the typical case described in reference [60], what is the maximum load the system can provide if only bus 13 and bus 23 are allowed to increase generation? Here we assume bus 13 and bus 23 belong to the same generation company.

The problem can be explained in this way.

1. The whole system is a zone.

2. P_{gk}, V_{gk} are fixed for generation buses except buses 13 and 23.

3. Q_{gk} are allowed to adjust within ranges.

Now we can use our zone-based maximum loadability subroutine to calculate. The TTC (total transfer capability) at bus 8 is 418.6 MW. Subtracting the base case power flow solution 171 MW at bus 8, the ATC (available transfer capability) is 247.6 MW. However, in order to afford this power, $P_{g13} = 585.7$ MW, $P_{g23} = 541.0$ MW, subtracting

the base case solution where $P_{g13} = 186.7$ MW, $P_{g23} = 660$ MW, we get the generation company has to increase 280 MW in order to provide the ATC 247.6 MW at bus 8. This is 32.4 MW more than the ATC power. This extra 32.4 MW is used to balance the extra transmission loss imposed by the increased power transfer.

However, here thermal overload constraints are ignored. If we want to include the thermal overload constraints at each line, what shall we do? We should include the thermal overload constraints into the optimization problem 4.53, as follows.

Define

$$f_l = V_{l_i}^2 + V_{l_j}^2 - 2V_{l_i}V_{l_j}cos(\theta_{l_i} - \theta_{l_j}),\ l = 1, \cdots, L \tag{4.67}$$

where L is the number of lines, l_i and l_j are the bus numbers of line l's two buses. Suppose line l's admittance is $y_l = g_l + jb_l$, and its thermal overload current rating is I_l, then

$$f_l \leq \frac{I_l^2}{g_l^2 + b_l^2} \tag{4.68}$$

Define $\tau_{lmax} = \frac{I_l^2}{g_l^2+b_l^2}$. The optimization problem can be generalized as

$$\min g(z, y) \tag{4.69}$$

subject to

$$h(z, y) = 0 \tag{4.70}$$

$$-\tau_{lmax} \leq f_l(z, y) \leq \tau_{lmax} \tag{4.71}$$

$$z_l \leq z \leq z_u \tag{4.72}$$

where $g(z, y)$ is the objective function of ATC. z includes bounded variables such as P_g, Q_g and V, while y includes unbounded variables θ. By adding barrier items, building Lagrange function and applying KKT first order conditions, we can simplify the equations

to be as follows (see [61] for details).

$$\sum_{l=1}^{N} \pi_{pl} \frac{\partial P_{el}}{\partial \theta_k} + \sum_{l=1}^{N} \pi_{ql} \frac{\partial Q_{el}}{\partial \theta_k} + \sum_{l=1}^{L} \pi_{rl} \frac{\partial f_l}{\partial \theta_k} = 0, \quad k \in \Omega_\theta \tag{4.73}$$

$$\sum_{l=1}^{N} \pi_{pl} \frac{\partial P_{el}}{\partial V_k} + \sum_{l=1}^{N} \pi_{ql} \frac{\partial Q_{el}}{\partial V_k} + \pi_{vk} + \sum_{l=1}^{L} \pi_{rl} \frac{\partial f_l}{\partial V_k} = 0, \quad k \in \Omega_v \tag{4.74}$$

$$P_{gk} - P_{Dk}^0 - p_{dk} P_{D\Sigma i} - P_{ek} = 0, \quad k = 1, \cdots, N \tag{4.75}$$

$$Q_{gk} - Q_{Dk}^0 - q_{dk} P_{D\Sigma i} - Q_{ek} = 0, \quad k = 1, \cdots, N \tag{4.76}$$

$$-\frac{\partial g}{\partial P_{D\Sigma i}} + \sum_{k \in D_i} \pi_{pk} p_{dk} + \sum_{k \in D_i} \pi_{qk} q_{dk} = 0, \quad i = 1, \cdots, d \tag{4.77}$$

$$f_l - r_l = 0, \quad l = 1, \cdots, L \tag{4.78}$$

By using this model, we recalculate the ATC for the previous example. When the continuous ratings are applied to lines, the ATC becomes 393.6 MW, line 7–8 is constrained. When the long-term emergency ratings or short term emergency ratings are applied to lines, the ATC becomes 418.6 MW. No line is constrained.

4.11 Factors That Influences the Algorithm's Speed and Convergence

According to our experience, μ adjustment scheme, initial point, and power flow equation multiplier are the three main factors that influences the algorithm's speed and convergnce.

Consider a simple optimization problem as follows.

$$\min z = (x+1)^2 + y^2 \tag{4.79}$$

subject to

$$0 \leq x \leq 1$$

$$0 \leq y \leq 1$$

We know the optimal solution is $(x^*, y^*) = (0, 0)$, and $z^* = 1$. The unconstrained optimal solution is $(x^u, y^u) = (-1, 0)$, $z^u = 0$. The alternative interior point problem is as follows.

$$\min z = (x+1)^2 + y^2 - \mu \ln(1-x) - \mu \ln x - \mu \ln(1-y) - \mu \ln y \tag{4.80}$$

subject to

$$0 \leq x \leq 1$$

$$0 \leq y \leq 1$$

For the alternative problem, when $\mu = 1$, the z-surface within the feasible region is shown in Figure 4.4; when $\mu = 0.1$, the z-surface within the feasible region is shown in Figure 4.5. From the two figures, we can see that when μ is large, the z-surface is smooth, therefore we can easily get the optimal solution. When μ is small, the z-surface has a sharp fold near the boundary. When we come to a point just before this trench, since from the algorithm point of view, we do not recogonize the upcoming boundary wall at this point, and still head toward the unconstrained optimal solution (-1,0). Once we move, we immediately hit the boundary and have to stop somewhere before the boundary to keep feasibility. Maybe this time we are still before the trench, maybe after the trench. If we are after the trench, we recognize a steep fall and move back toward the trench. On the whole, we did not do much effort on traveling along the trench to the optimal

(0,0). Unless we find the exact bottom of the trench, we can not recognize the correct moving direction. Therefore, when μ is small, the algorithm is easy to get oscillation. For this example, when $(x, y) = (0.05, 0.35)$, the required moving direction is (-0.05,-0.35). However, when $\mu = 0.1$, the calculated moving direction $d = (-0.0049, -0.1861)$. When $\mu = 0.0001$, $d = (-1.0284, -0.3497)$. Therefore, when $\mu = 0.1$, we can move toward the optimization easily, while when $\mu = 0.0001$, we are expected to oscillate.

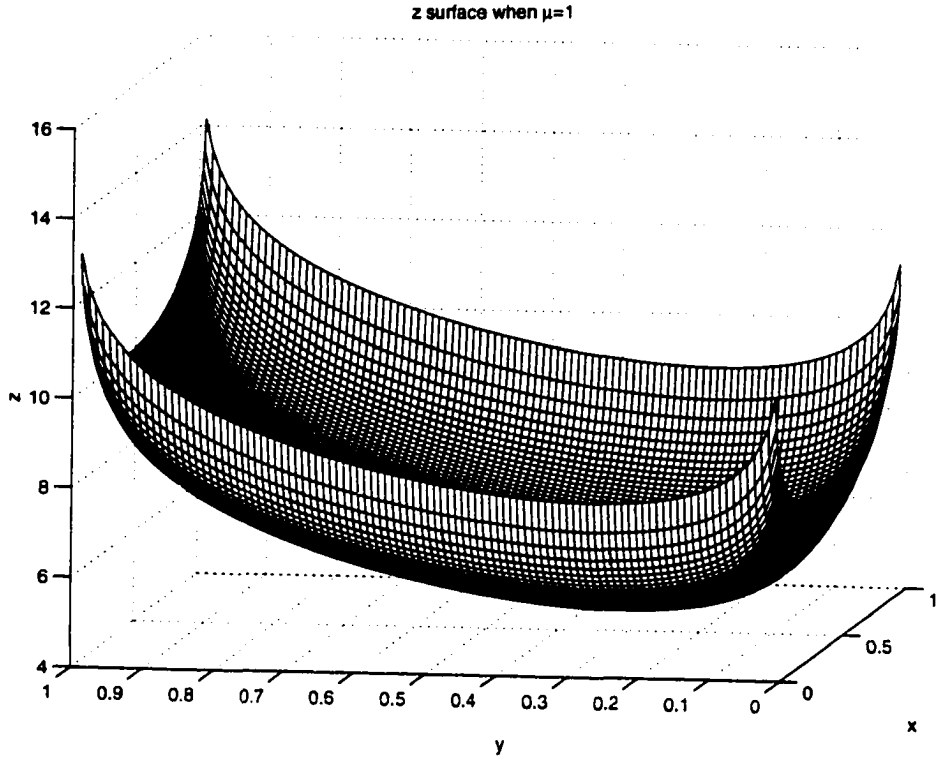

Figure 4.4 z-surface when $\mu = 1$

Choosing the initial point in the interior rather than near the boundary and starting μ with a large value are the measures to avoid oscillation. Let us use the minimum load curtailment subroutine mentioned before as the example. We still use the typical case for IEEE RTS'96. Suppose we divide the system into 4 zones as mentioned previously. Suppose we start from the exact middle point for every variable with upper and lower

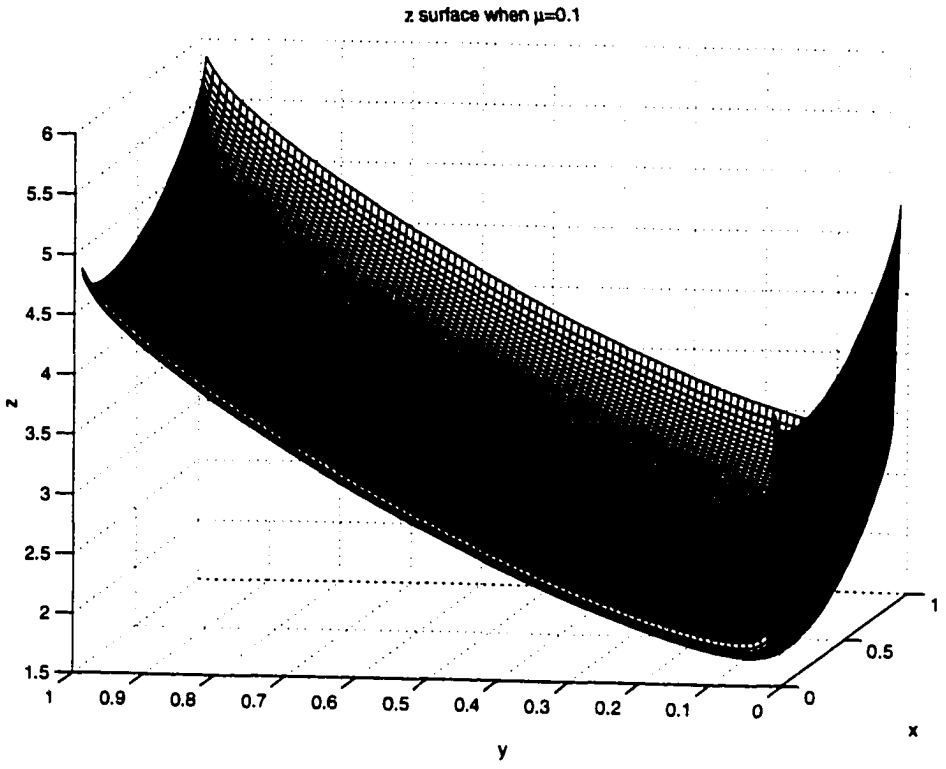

Figure 4.5 z-surface when $\mu = 0.1$

limits, and 0 for other variables. When we keep $\mu = 1.0$, we spend 11 iterations to make equation errors less than 0.001, the calculated total load is 2447 MW. When we keep $\mu = 0.01$, we spend 38 iterations, and the calculated total load is 2914 MW. The result is more accurate because it is larger. When $\mu = 0.0001$, the subroutine fails to converge after 100 iterations. Therefore, in our subroutine, μ starts from a large value, gets adjusted to a smaller value after each convergence, if oscillation is encountered, μ is increased.

Now we choose $\alpha = 0.999$, and the initial value

$$P_{gk}^0 = P_{gkmin} + \alpha(P_{gkmax} - P_{gkmin}),\ k \in \Omega_p \tag{4.81}$$

By using our subroutine, we spend 42 iterations to get the solution. Now change back to choosing $\alpha = 0.5$. The number of iterations is reduced to 25. However, only

when the initial point is near the effective constraints will the algorithm oscillate. If at the first step, the point leaves the constraints and never come back, oscillation will not happen. For example, when we choose $\alpha = 0.1$, the number of iterations is 23, even a bit better than $\alpha = 0.5$.

The better performance of $\alpha = 0.1$ compared with $\alpha = 0.5$ can be explained by Figure 4.6. The upper and lower limits on variables constitute a bounded space. Power flow equations constitute a hypercurve. The intersection of the bounded space and the hyperplance is the feasible region. When the initial point 1 (IP1) is near the upper bound, it heads toward the unconstrained optimal solution and hits the upper bound. However, the power flow equations are not balanced and the algorithm is prone to oscillation near the boudary. On the other hand, if we start from IP2, which is far away from the boundary, it will recognize the existence of the power flow hypercurve and merge into it somewhere before the boundary. Then it moves along the hypercurve until it hits the boundary. Since the solution now is feasible, further iterations are not needed.

Another method is to multiply the power flow equations with a large multiplier c, say, 100. Then equation 4.75 and 4.76 become

$$c \cdot (P_{gk} - P^0_{Dk} - p_{dk} P_{D\Sigma i} - P_{ek}) = 0, \tag{4.82}$$

$$c \cdot (Q_{gk} - Q^0_{Dk} - q_{dk} P_{D\Sigma i} - Q_{ek}) = 0, \tag{4.83}$$

where $k = 1, \cdots, N$.

The solution has not been changed. However, since we emphasize the importance of power flow equations, we are expected to merge into the hypecurve faster, the number of iterations can then be reduced. The process is as shown in Figure 4.7. Table 4.8 shows different numbers of iterations under different values of c for the previous example by using our subroutine. It seems c=100 is an appropriate value. When c is too large, it moves slowly along the hypercurve, and the number of iterations will increase.

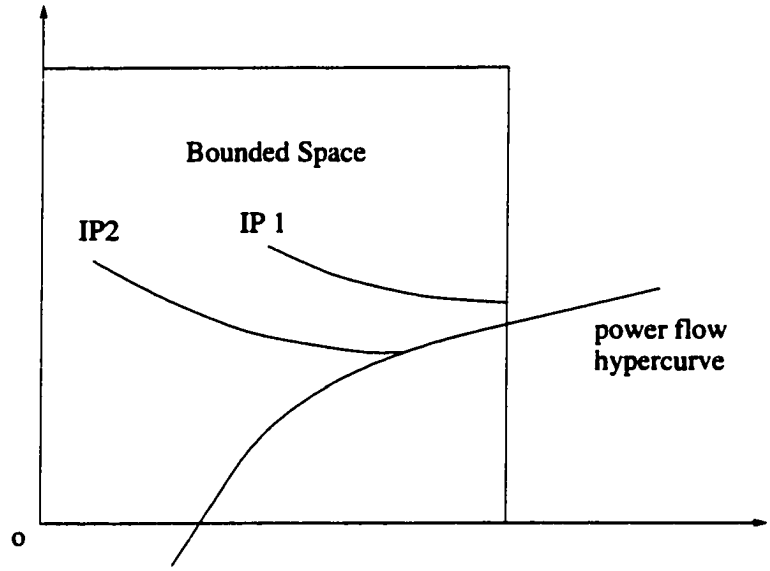

Figure 4.6 The influence of initial point position

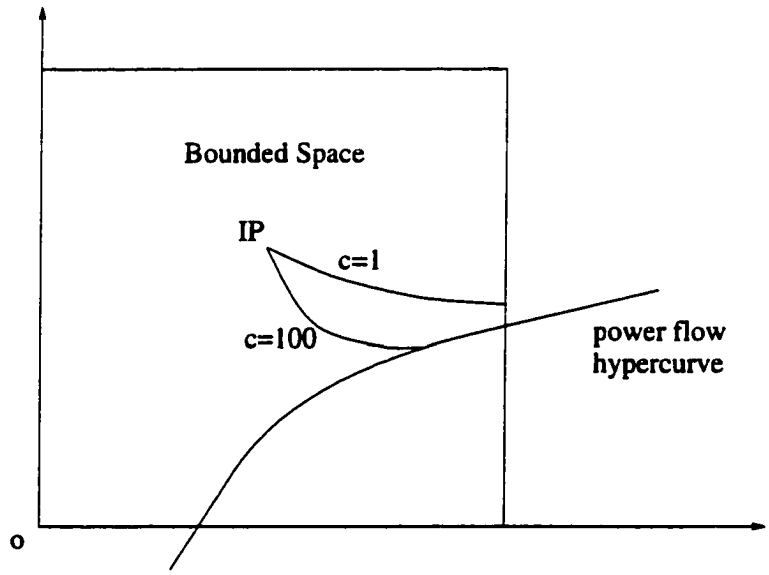

Figure 4.7 The influence of power flow multiplier

Table 4.8 c .vs. number of iterations

c	10	100	1000	10000
iter	21	17	17	26

Someone may suggest finding the exact trench bottom when a boundary is encountered to avoid oscillation. However, when two or more boudaries are encountered, finding the bottom of the trench will do little to help since the trench itself is bent. It will take many iterations to make a turn. Therefore, enlarging μ and adjusting the initial point are more reliable measures.

4.12 Illustration of Annual Power Flow Infeasibility Risk Assessment

We divide the IEEE RTS system into zones according to our experience, so that in each zone all loads vary proportionally.

- Zone 1 includes bus 11-4. It is mainly a generation provider.
- Zone 2 includes bus 16, 9, 10. It is mainly an electricity consumer..
- Zone 3 includes bus 7, 8. It is heavy load center connected loosely with the system.

There are usually quite a few hours which have the same unit commitment. We pick up the peak load among these hours, denote it P_{zmax}. Assuming the percentage load forecasting error is σ_r, we choose

$$P_{zu} = (1 + 4\sigma_r)P_{zmax} \tag{4.84}$$

where P_{zu} is the upper limit of the load level. Then we can apply our minimum load curtailment subroutine to calculate the load unbalance and the corresponding risk. We

use the term "load unbalance" instead of "load curtailment" because there are many ways to compensate for load unbalance. We can buy power from neighbour systems, we can start up emergency units, or we can curtail load. Load curtailment is normally the last means and rarely happens. Therefore, the concept of load unbalance is broader and more appropriate than load curtailment.

In our program, we assume the risk corresponding to one MWh load unbalance is \$100. We also make use of shadow prices to discern load unbalance due to different reasons. We use the following scheme to discern different types of load unbalance.

1. If $\pi_{zi} < -0.5$ for all i, set type =0, the system has no load unbalance.
2. Otherwise, if all π_{pk} are negative and large in magnitude when $k \in \Omega_p$, set type=1, the system is confined by active generation capacity.
3. Otherwise, find the bus with the lowest voltage, denote it as bus m.
4. If π_{vm} is small in magnitude, set type=2, the system is confined by voltage instability; otherwise, set type=3, the system is confined by voltage lower limit.

The steps for annual load unbalance risk estimation are as follows.

1. Combine hourly unit commitment patterns of the year. Reorder them from the highest total capacity to the lowest total capacity.
2. For each unit commitment pattern,
   (a) Start from a flat start point, use interior poinnt algorithm to get the maximum loadability for the normal state, record maximum load levels and types for load unbalance.
   (b) For each contingency,

i. If this is the first unit commitment, start from the solution of the normal state.

ii. Otherwise, look up the record. If in the history a flat start sign was set for this contingency, start from a flat start point;otherwise start from the normal state solution point.

iii. If at this time, oscillation is encountered and a flat start process is called by our interior point subroutine, or if the number of iterations exceeds 30, set the flat start sign. The following unit commitment patterns will begin from a flat start point for this contingency.

iv. Record maximum load levels and types of load unbalance.

(c) Calculate load unbalance risk and probability for different zones and different types for hours which have this unit commitment pattern, add them into annual risk assessment records.

3. Output Calculation results.

First we choose $V_{gmax} = 1.10pu$, $V_{gmin} = 0.90pu$, $V_{lmax} = 1.15pu$, $V_{lmin} = 0.85pu$. Here V_{gmax}, V_{gmin} are the generation bus voltage upper limit and lower limit respectively. V_{lmax}, V_{lmin} are the load bus voltage upper limit and lower limit respectively. The total load unbalance risk curve is shown in Figure 4.8. The load unbalance risk in three different zones are shown in Figures 4.9, 4.10, 4.11 respectively. The load unbalance risks due to different reasons are shown in Figure 4.12, 4.13, 4.14 respectively. Comparison of one figure to another should be done keeping in mind the different scales used on the vertical axis. From these figures we make the following observations.

- Most of the total load unbalance risk is focused on the first half of the year.
- Zone 3 is the weakest part of the system. It suffers the largest risk. Zone 1, which is mainly a generation provider, suffers very little risk.

- Zone 3 suffers most of its risk in the first half year. Zone 2 suffers most of its risk at the end of the year. Zone 1 suffers most of its risk from the 5000th hour to the 7000th hour, about from the beginning of July to the end of September.
- Most of the risk is due to active generation constraint. The risk due to voltage lower limit is also significant. There is also some risk caused by voltage instability.

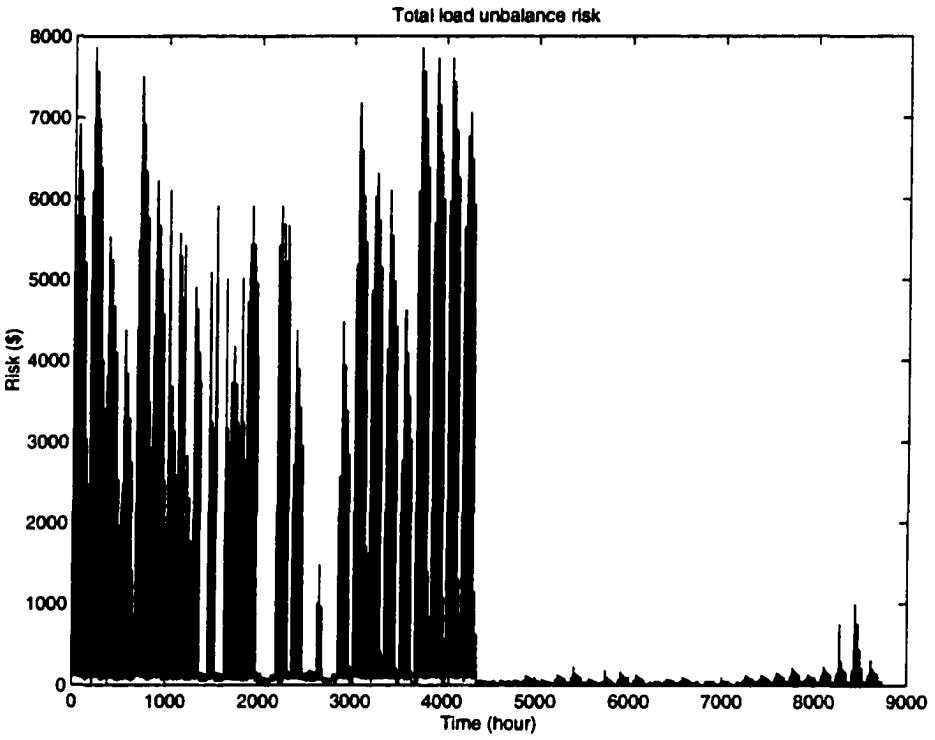

Figure 4.8 Load unbalance risk over a year

Table 4.9 shows the risk in different zones and due to different contingencies. Only the contingencies with risk greater than $10,000 are listed. From the table, we have the following observations.

- The normal state contributes most of the risk. It means our unit commitment arrangement may have some drawbacks.
- Loss of a line normally has less contribution to the total risk than loss of a unit.

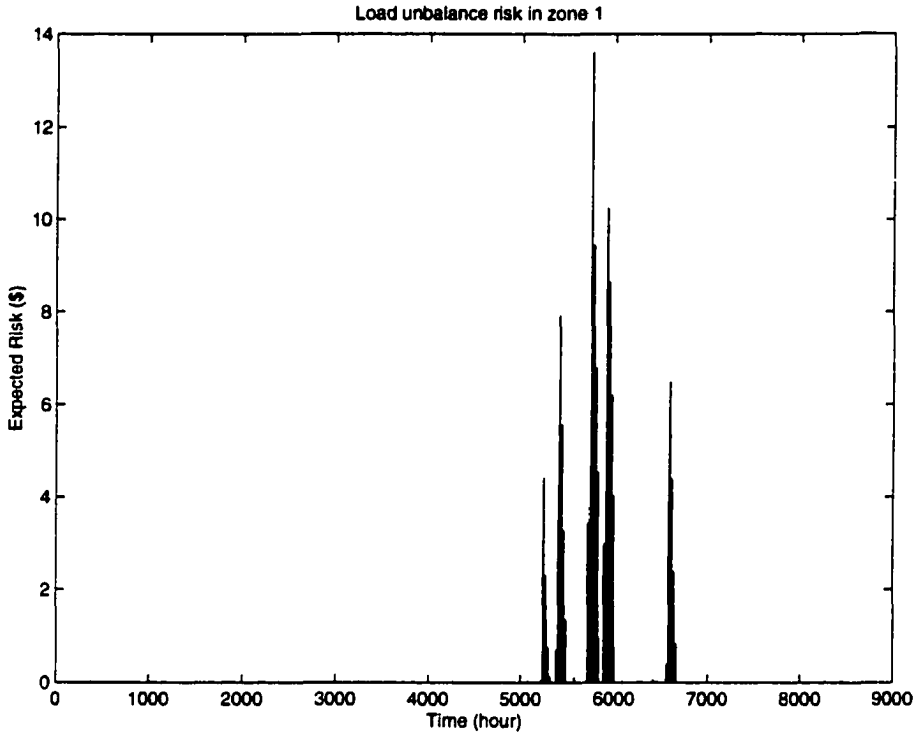

Figure 4.9 Zone 1 risk over a year

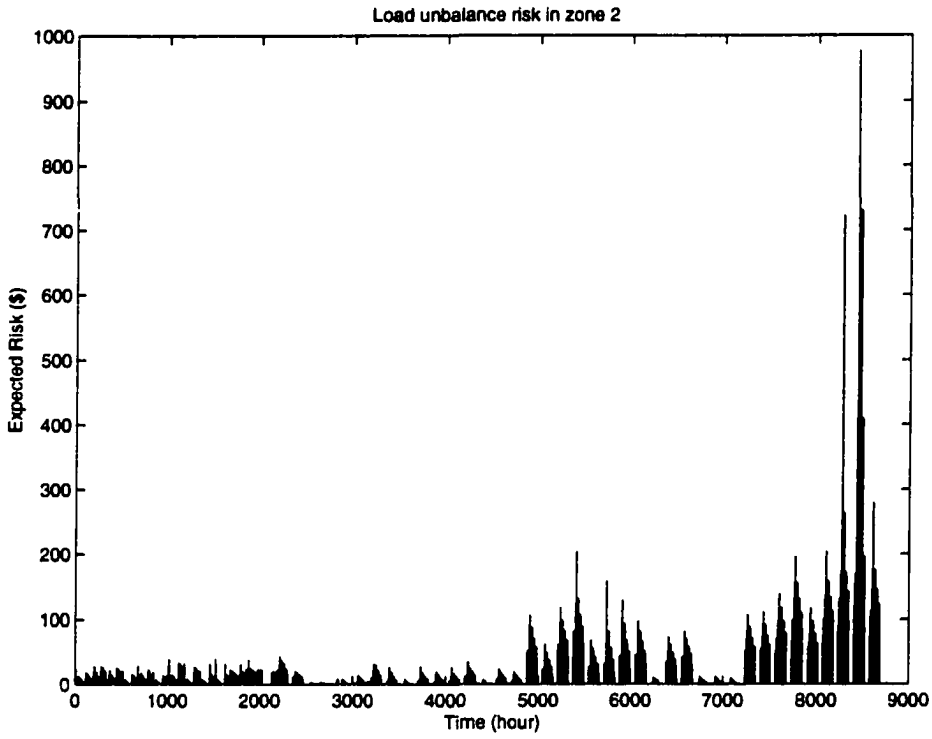

Figure 4.10 Zone 2 risk over a year

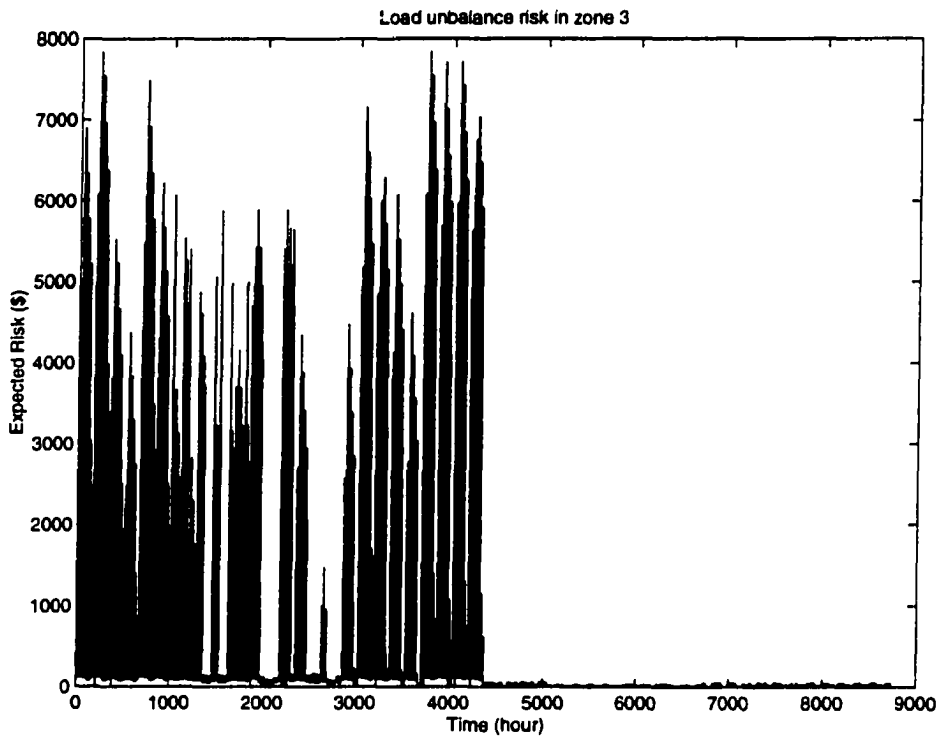

Figure 4.11 Zone 3 risk over a year

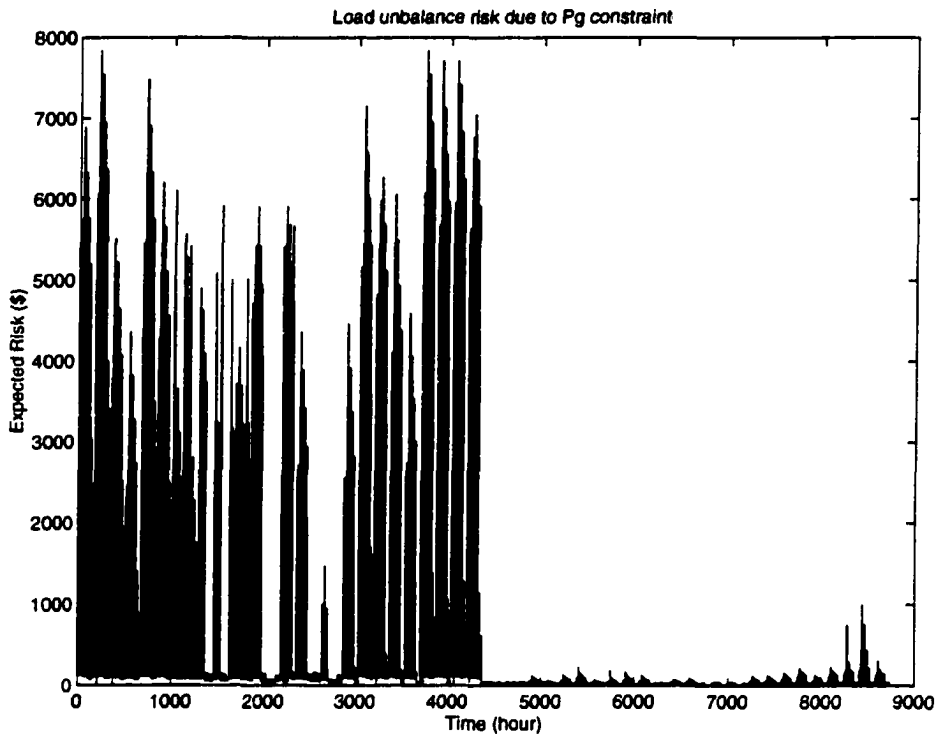

Figure 4.12 Load unbalance risk due to generation limit

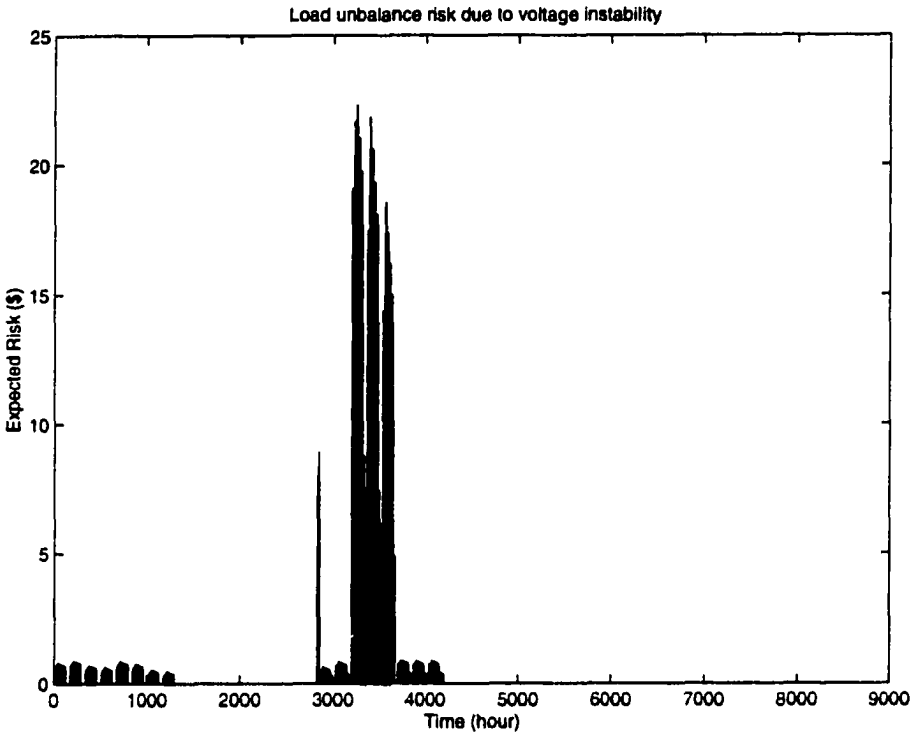

Figure 4.13 Load unbalance risk due to voltage instability

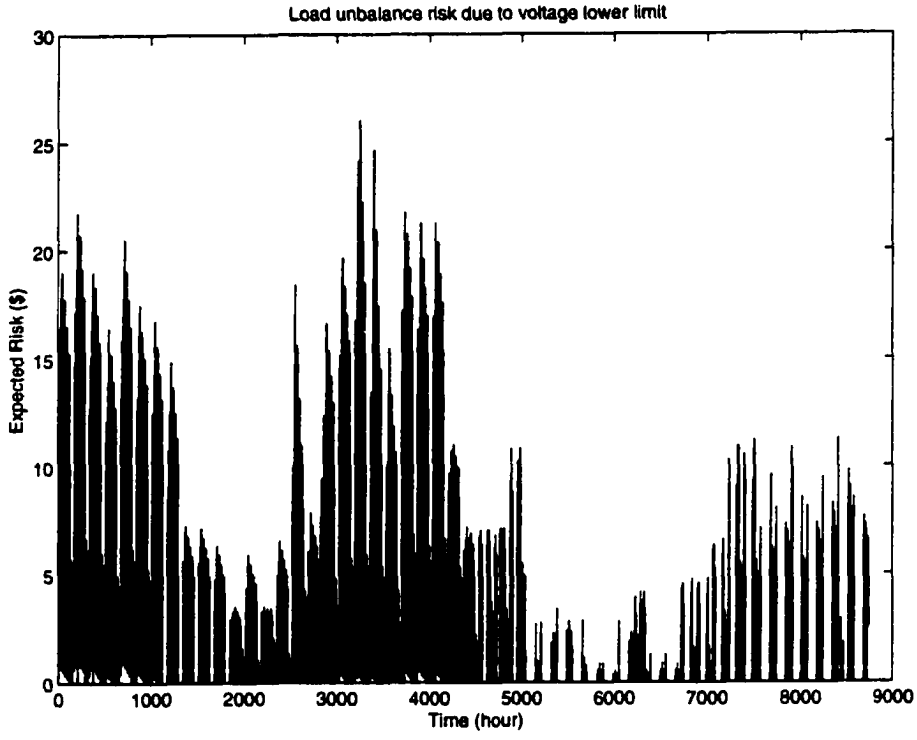

Figure 4.14 Load unbalance risk due to voltage lower limit

- Common mode outages are not on the list. It is due to their low probabilities.
- Nuclear units (unit 23 and unit 24) have high risk.

Therefore, if we want to reduce risk, we can either adjust the unit commitment or upgrade the reliability of nuclear plants. We think the former is more effective.

Table 4.9 Load unbalance risk in different zones ($)

State	Zone 1	Zone 2	Zone 3	Total
Normal	0.65	4981	5513771	5518753
Line 3-24	0.00	29.02	15926	15955
Line 7- 8	0.00	6.84	12870	12877
Line 8-10	0.00	1.16	12465	12466
Line 9-11	0.00	9.15	10590	10599
Line 9-12	0.00	10.11	10659	10669
Line 10-11	0.00	11.77	11525	11537
Line 10-12	0.00	14.54	11855	11870
Unit 3	0.00	203.5	11983	12186
Unit 4	0.00	203.5	11983	12186
Unit 7	0.00	204.0	11316	11520
Unit 8	0.00	204.0	11316	11520
Unit 12	0.04	5909	48507	54415
Unit 13	0.04	5909	48507	54415
Unit 14	0.04	5909	48507	54415
Unit 21	4.03	2521	48460	50986
Unit 22	2.95	2509	43483	45995
Unit 23	89.04	20713	84376	105178
Unit 24	49.58	21912	84575	106537
Unit 31	0.08	2273	50532	52805
Unit 32	0.08	2273	50532	52805
Unit 33	264.94	18673	66150	85088
Total	458.46	96382	6390746	6487586

Table 4.10 gives the expected load unbalance hours for each zone. From this table, we can also see zone 3 suffers load unbalance very seriously.

Table 4.11 and 4.12 provides the risk due to different constraints and the expected load unbalance hours due to different constraints respectively. From these tables we

Table 4.10 Expected load unbalance hours in different zones

State	Zone 1	Zone 2	Zone 3
Total	0.11	15.64	1839.56

can see that generation inadequacy is the major problem of this system. Therefore, increasing the generation in zone 3 may reduce the risk. Figure 4.15 shows zone 3's generation capacity over one year. It seems the lack of generation capacity in zone 3 during the first half of the year is the major reason for high risk. Table 4.13 shows the buses that hit their lower voltage limit and the expected number of hours for these buses to suffer low voltage. Only bus 6 and bus 7 suffer the lowest voltage.

Table 4.11 Load unbalance risk due to different constraints ($)

State	P_{gmax}	Voltage instability	Voltage lower limit	Total
Total	6456985	4494	26107	6487586

Table 4.12 Expected load unbalance hours due to different constraints (hours)

State	P_{gmax}	Voltage instability	Voltage lower limit	Total
Total	1846.29	0.82	5.63	1852.74

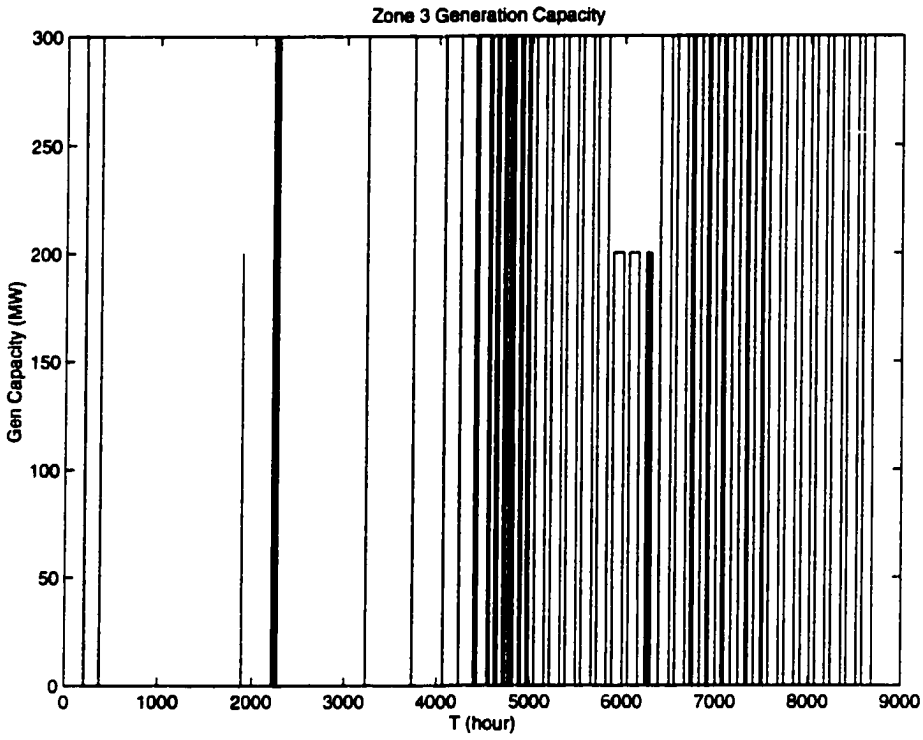

Figure 4.15 Zone 3 capacity over one year

Table 4.13 Low voltage bus and expected low voltage hours

bus no.	voltage instability hours	voltage lower limit hours
6	0.00	1.43
7	0.82	4.20

4.13 Summary of this Chapter

Power flow infeasibility risk is the risk suffered when the power flow is infeasible. It can be caused either by generation constraints, or by voltage instability, or by voltage lower limit constraints. A new interior point algorithm is proposed to calculate power flow infeasibility risk due to load unbalance, and shadow prices are employed to identify the reason for power flow infeasibility. This method is then expanded to deal with multi-zones. The multi-zone model is then adopted for every hour and every outage to accumulate risk over a year. It is also shown that our proposed method can even be used to calculate ATC. The risk assessment results for IEEE RTS'96 are shown to be decomposable and informative.

CHAPTER 5 THERMAL OVERLOAD RISK ASSESSMENT

5.1 Introduction

Many researchers have developed probabilistic techniques for assessing power system overload security. For example, Billinton [12] proposed enumerating contingency states for a specific loading condition, investigating the overload security of each state, and combining the corresponding probabilities of the violated states. However, several other researchers account for load uncertainty. Their research resulted in the probabilistic load flow (PLF) [62]. Early in 1973, the concept of probabilistic load flow was first proposed by Borkowska [63]. In that paper, DC power flow and convolution techniques of random variables were employed to calculate probabilistic power flow. In 1975, O.A. Klitin from American Electric Power (AEP) proposed a completely different method (widely known as the AEP method) independently [64]. The concept of stochastic load flow proposed by him was based on Gaussian distributed random variables, while the probabilistic load flow has no assumptions on the distributions of those random variables. Further development of the probabilistic power flow included FFT based discrete convolution method to facilitate convolution calculation [65]; a multilinearisation technique to linearize the power system at several points instead of one point to overcome the nonlinearity of power flow equations [66]; a method to calculate probabilistic load flow when dependences between loads are considered [67]; and an application of the least square method with the number of unknown variables greater than the number of equations in probabilistic

power flow calculation [62]. Further development of the stochastic power flow included a method to represent non-Gaussian probability distributions by the method of Gaussian sum approximations [68], and an optimal stochastic load flow model [69]. There are also other methods to solve the probabilistic power flow problem, such as the Monte Carlo simulation [70] and the probabilistic collocation method [71].

The PLF method is mainly proposed for probabilistic analysis of operating conditions. In addition, efforts have been made to perform risk assessment for power system planning in [72]. Paper [72] uses the daily load curve (its basic time unit is one day) to predict the future load and load uncertainty, then perform cost and benefit comparison among different system expansion plans based on PLF techniques. However, it does not perform sequential unit commitment and does not calculate risk hour by hour. In this dissertation, we apply the sequential mean-variance model to generate a system trajectory, as described in Chapter 3. Furthermore, our risk assessment result is decomposable and assignable according to different hours, different components and different contingencies. These two aspects make our thermal overload risk assessment unique relative to the work of other researchers. However, for the risk assessment of a specific hour, our method is similar to a typical PLF. We employ a convolution technique based on a linearized model to calculate overload probability, which is very similar to the method proposed in [63]. In order to facilitate our calculation, we propose a new method called segmentwise cluster based convolution method to deal with convolutions between random variables [73]. This provides an increase in computational speed sufficient to allow the computation to be done hourly over a year's trajectory of operating conditions rather than for a single snapshot as was done in [63]. In addition, we include impact assessment through a "component risk table" to enable calculation of risk, which extends the probability calculation of [63]. We illustrate the risk assessment results on the IEEE RTS'96.

5.2 Some Assumptions

5.2.1 Load Model

We have assumed a standard deviation on total load as a percentage of the total load. We assume this solution also applies to each individual load as the same percentage of their expected values. Denote P_{Di} to be the real part of the load at bus i. For different buses (when $i \neq j$), it is normally incorrect to assume that loads P_{Di} and P_{Dj} are independent. Therefore, we assume there is a covariance matrix for the random load vector. This covariance matrix can be estimated by statistical methods. However, in our risk assessment for the IEEE RTS'96, since we lack real data to estimate the covariance matrix, for the purpose of illustration, we assume the correlations between loads within the same voltage level are all 0.5, while the correlations between loads of different voltage levels are all zero. Once we have the estimated covariance matrix from history data, we could easily substitute it into our software.

5.2.2 Generation Model

We assume every generation bus is an independent generation company. That means if a unit is outaged, other units at the same bus will increase power to make up for the generation loss. Only when the total capacity is below the required power at the operation state is there generation loss on this bus. Now we express this idea analytically. Suppose for bus i, the generation capacity at operation state is $C_i^t(0)$, where 0 denotes no unit is outage, and the real generation at operation state is P_{Gi}. When one or more units at a bus fails, the bus is transfered from the normal state to some other state, say, state n. If we denote $C_i(n)$ to be the corresponding generation capacity at bus i and state n, then the generation inadequacy is

$$-\Delta P_{Gi}(n) = (P_{Gi} - C_i(n))u(P_{Gi} - C_i(n)) \tag{5.1}$$

where

$$u(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$$

and $\Delta P_{Gi}(n)$ is defined as the increment of generation at bus i and state n.

5.2.3 Consideration of Reactive Power

It is the current flowing through a line that causes the thermal overload problem. Therefore, not only active power influences the thermal overload, but also the reactive power. In order to enhance computational speed, we use an approximate model to include this effect. Since the disturbances of the system are normally small, we assume that voltage at every bus remains constant before and after the contingency. Therefore, for small disturbances, the reactive power will remain constant for every line. However, there are two PQ pairs corresponding to two ends of a line respectively. We choose the one which makes the value $\sqrt{P^2 + Q^2}/V$ larger, where all variables are in per unit.

5.2.4 Consideration of Corrective Actions and Protection devices

A thermal overload on a line can not be tolerated for a long time without taking corrective action. Here we assume all thermal overload can be alleviated within the time interval of 1 hour. This implies that the duration of overloads are considered to be one hour.

5.3 Flow Distribution

5.3.1 Linearization Around Operating Point

We employ variable s to identify the outage state, while we use state 1 to represent the normal state without outages. From the DC power flow formulation, we can obtain

the following expressions for branch flows corresponding to any outage state s and the normal state 1.

$$P_l(s) = A(s)^T B_{ps}^{-1}(P_G - P_D) \tag{5.2}$$

$$P_l(1) = A(1)^T B_{p1}^{-1}(P_G - P_D) \tag{5.3}$$

where

- P_G is the vector of real power generator levels at each bus.
- P_D is the vector of real power load levels at each bus.
- B_{ps} is the B-matrix for outage state s.
- $A(s)$ is the connection matrix of the network for outage state s, having rows corresponding to buses (excluding the swing bus) and columns corresponding to branches.
- $P_l(s)$ is the vector of branch power flows for outage state s.

Subtracting equations 5.2 from 5.3 we get

$$P_l(s) = P_l(1) + [X_l(s) - X_l(1)](P_G - P_D) \tag{5.4}$$

where $X_l(s) = A^T(s)B_{ps}^{-1}$. Since $X_l(s)$ are independent of generation and load level, we can calculate and store them beforehand to save computation time.

If we set P_G and P_D to be their expected values for the hour, and use a full AC power flow solution to obtain $P_l(1)$, then equation 5.4 provides the expected flows for all branches at the hour assuming the system is in state s. We will use this below in obtaining the distribution of flows due to uncertainty in generation.

Now we define ΔP_G and ΔP_D as the vectors of random variables corresponding to generation and load levels, respectively. We describe each component of ΔP_G with a

two state probability mass function. We describe each element of ΔP_D with a normal distribution having a mean equal to 0 and a standard deviation derived from our load model assumption. That means the standard deviation of ΔP_D is the same as the standard deviation of load P_D, which is obtained from the load forecasting error as described in Section 2, Chapter 3. The vector of random variables corresponding to variations in branch flows at outage state s are then given as

$$\Delta P_L(s) = X_l(s)(\Delta P_G - \Delta P_D) \tag{5.5}$$

and the vector of random variables corresponding to the branch flows

$$P_L(s) = P_l(s) + \Delta P_L(s) \tag{5.6}$$

Substitution of equation 5.5 into 5.6 yields

$$\begin{aligned} P_L(s) &= (P_l(s) + X_l(s)\Delta P_G) - X_l(s)\Delta P_D \\ &\triangleq P_{LG}(s) + P_{LD}(s) \end{aligned} \tag{5.7}$$

Here $P_{LG}(s)$ is the vector of random variables corresponding to branch flows due to variation in generation, given by

$$P_{LG}(s) = P_l(s) + X_l(s)\Delta P_G \tag{5.8}$$

Also, $P_{LD}(s)$ is the vector of random variables corresponding to branch flows due to variations in loads, given by

$$P_{LD}(s) = -X_l(s)\Delta P_D \tag{5.9}$$

5.3.2 Obtaining the distribution of $P_{LG}(s)$ and $P_{LD}(s)$

Because the product term in equations 5.8 represent vector multiplications, they involve summations of random variables, and therefore the distributions of $\Delta P_{LG}(s)$ can be obtained by convolution of the component distributions. The convolution algorithm

we use was developed explicitly for this purpose, and it is quite efficient. It is fully described in Section 5.4.

We know that if a vector of random variable Y is given by

$$Y = AZ$$

where A is a constant matrix and Z is a vector of random variables, then

$$\begin{aligned} E(Y) &= AE(Z) \\ cov(Y) &= Acov(Z)A^T \end{aligned}$$

By assumption, the vector of load variations ΔP_D are normally distributed and they have 0 means. We also assume that the covariance matrix for ΔP_D is available, as we explained in subsection 5.2.1. Therefore, from equation 5.9, the covariance matrix for $P_{LD}(s)$ can be easily obtained from

$$cov(P_{LD}(s)) = X_l(s)cov(\Delta P_D)X_l^T(s) \tag{5.10}$$

The diagonal elements of equation 5.10 are the line flow variances, which completely describe the zero mean distributions for each element in $P_{LD}(s)$.

5.4 Segmentwise Cluster Based Convolution Method

5.4.1 Segmentwise Cluster Based Random Variable Description

A discrete random variable can be easily represented by two vectors: one is for possible values and the other is for probabilities at those values. However, this kind of representation has some drawbacks. The number of possible states for the summation of random variables grows exponentially fast with the number of random variables. Some states may be relatively close in term of their values so that combining them as a single state results in little loss in accuracy but significant gain in computation speed. This

is conceptually appealing, yet a practical problem remains: how to select the value and probability of the combined state?

Segmentwise convolution is one possible solution which has been proposed and used before. From the lowest value of the variable to the highest value, we can use a grid to divide possible states and cluster states within an interval of the grid into one equivalent state. For example, a random variable has possbile values (1,2,7), and probabilities (0.2,0.2,0.6). We can use a grid with start point 0, and interval gap 5 to divide states, i.e., we can define (0,5], (5,10] to be two cluster intervals. Then as $1, 2 \in (0, 5]$, they should be combined into one state, and 7 is another state. However, previous implementations of the convolution algorithm uses the midpoint of each segment to represent the segment's value. This is not accurate because the values of the states within the segment interval may be far away from the midpoint. In our program, we adopt a cluster based description. For segment (0,5], the expected value is $\frac{1*0.2+2*0.2}{0.2+0.2} = 1.5$, the probability is $0.2 + 0.2 = 0.4$. Therefore, the value vector (1.5,7), and the probability vector (0.4,0.6) can be use to represent the random variable. We call those values which are used to represent segments representative values.

In our program, we use a structure to represent a random variable, say X. The structure includes the smallest value (or the start value) of X (denoted by StartX), the interval gap between segments (denoted by Δ), the sparse vector for representative incremental values of all segments (denoted by dX), and the sparse vector for probabilities of all segments (denoted by probX). Here dX of a segment is defined as the representative value subtracted by the left hand side boundary limit of this segment. The structure is shown in the following pseudo code.

```
Struct random_variable{
    double StartX;
    double Δ;
    sparse_vector dX;
```

sparse_vector probX;

};

Denote the representative value of the ith segment of X to be $X(i)$, then we can calculate $X(i)$ from our data structure

$$X(i) = (i-1) * \Delta + dX(i) \tag{5.11}$$

Using this data structure, we can compile subroutines for convolution, deconvolution, and multiplication with a coefficient.

5.4.2 Convolution

Given random variables X and Y, what is the distribution of random variable $Z = X + Y$, if the two random variables are independent of each other? The calculation is defined as convolution. It can be expressed as $Z = X \oplus Y$, where the sign $\oplus$ indicates that X and Y are independent. If X and Y are both represented in our standard structure, how do we design an algorithm to get Z with the same structure? Here Δ is the same for every random variable. Of course $StartZ = StartX + StartY$, and Z represented by our proposed structure can be obtained by convoluting all possible combinations of X's segments and Y's segments. When the ith segment of X and the jth segment of Y is convoluted, the representive value of the corresponding segment in Z is $z = X(i) + Y(j)$, where $X(i)$ and $Y(j)$ can be computed by 5.11. The corresponding segment number of z (denoted by k) is

$$k = \lceil \frac{z - StartZ}{\Delta} \rceil \tag{5.12}$$

The equation can be simplified as

$$k = i + j - 2 + \lceil \frac{dX(i) + dY(j)}{\Delta} \rceil \tag{5.13}$$

For any combination of $X(i)$ and $Y(j)$, we can always find the appropriate index k of the interval to accommodate the value $z = X(i) + Y(j)$ by using 5.13. As long as k

and z are known, we can use them to update previous $Z(k)$ and $probZ(k)$.

$$Z(k) = \frac{probZ(k)Z(k) + probX(i)probY(j)z}{probZ(k) + probX(i)probY(j)} \tag{5.14}$$

$$probZ(k) = probZ(k) + probX(i)probY(j) \tag{5.15}$$

The $dZ(k)$ can be obtained by

$$dZ(k) = Z(k) - (k-1)\Delta - startZ \tag{5.16}$$

Once all combinations are processed, the random variable Z has been established.

5.4.3 Deconvolution

Assume we know $Z = X \oplus Y$, and Z and X are known, however, we do not know Y. Then how do we calculate Y from Z and X? This is just like an inverse calculation of convolution. It is called deconvolution. Can we just calculate $Z \oplus (-X)$ so that the convolution subroutine can be used? No, because Z and X are correlated. So how to find Y by deconvolution? Here we propose a recursive subtracting technique. It differs from other deconvolution algorithms in that it fully employs our proposed random variable structure.

Denote the final element index for X, Y and Z as n_x,n_y,n_z respectively, where n_y is unknown now. Since we have $X(n_x) + Y(n_y) = Z(n_z)$, we have

$$Y(n_y) = Z(n_z) - X(n_x) \tag{5.17}$$

And we also have

$$StartY = StartZ - StartX \tag{5.18}$$

From $Y(n_y)$ and StartY, the index n_y can be easily determined. And

$$probY(n_y) = \frac{probZ(n_z)}{probX(n_x)} \tag{5.19}$$

Once we get the final element of Y, we can eliminate its contribution, i.e., for every possible value of $z = X(i)+Y(n_y)$, we can find the appropriate index k for z by applying 5.13, and then update previous $Z(k)$ and $probZ(k)$ by

$$Z(k) = \frac{Z(k)probZ(k) - z \cdot probX(i)probY(n_y)}{probZ(k) - probX(i)probY(n_y)} \tag{5.20}$$

$$probZ(k) = probZ(k) - probX(i)probY(n_y) \tag{5.21}$$

After adjustment, the final element of Y is eliminated. Then we can use the same procedure to calculate the second largest state of Y. We can repeat this procedure until we obtain all elements of Y. The flowchart of the algorithm is shown in Figure 5.1.

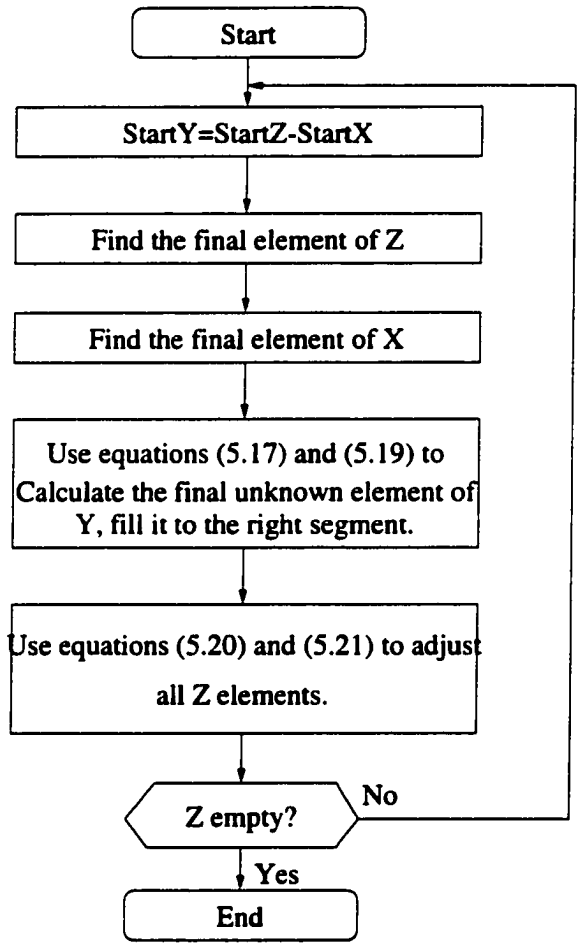

Figure 5.1 Deconvolution subroutine flowchart

5.4.4 Multiplication with A Coefficient

This subroutine is used to calculate $Z = c \cdot X$, where c is a real number, X is a random variable represented by our proposed random variable structure. In our program, we choose 5 MW as the segment width. Therefore, for every $X(i)$, we calculate $cX(i)$, then find an appropriate index k to place it by applying the expression

$$k = \lceil \frac{cX(i) - startZ}{\Delta} \rceil \tag{5.22}$$

We choose the starting base $StartZ$ to be the largest multiple of the segment width which is less than the smallest possible cX value.

5.5 Obtaining the Distribution of $P_{LG}(s)$

Based on the segmentwise cluster based convolution method, we can calculate the probabilistic distribution of $P_{LG}(s)$ easily. From equation 5.8, since $X_l(s)$ is a constant matrix, and ΔP_G is a random vector, we can apply the multiplication and convolution subroutines directly to get the distribution of $P_{LG}(s)$. Here ΔP_G is calculated from equation 5.1. According to equation 5.1, bus i's ΔP_{Gi} is a function of the generation capacity C_i. Since C_i is a random variable with multiple states, ΔP_{Gi} is also a random variable with multiple states. Its magnitude with respect to each state can be obtained from equation 5.1; its probability with respect to each state is the summation of the probabilities of all possible C_i's that result in ΔP_{Gi}.

The probability mass function of C_i can also be calculated by our convolution algorithm. At bus i, if the number of operational generators is n_i, then the Capacity at generation bus i is

$$C_i = \sum_{j=1}^{n_i} C_{ij} \tag{5.23}$$

where C_{ij} is the capacity of generator j at bus i. Every C_{ij} is a two state random variable, with the probability (1-FOR) at the full generation, and the probability FOR at

0. Therefore, C_i is the convolution result of all C_{ij}'s at bus i. We can use our convolution algorithm to calculate the distribution of the random vector $C = [C_1, C_2, \cdots C_{N_g}]^T$ from scratch, where N_g is the number of generation buses.

However, the unit commitment does not change very often, and when it does, not by very much. Therefore, constructing random vector C for each hour will waste a large amount of time. Therefore, we only construct C the first hour. After that, we check if the unit commitment is different from the previous time, if it is not, we do not have to change C. Otherwise, we compare unit states for every unit, and if the unit is started up currently, we apply the convolution algorithm to update C_i at the corresponding bus i. If the unit is shut down, we use the deconvolution subroutine to update C_i at the corresponding bus i.

5.6 Screening Technique

In summing up the possible states for a line, we may ignore the very low probability states. For line i under state s, we denote the line flow $P_{Li}(s)$'s mean and standard deviation to be $p(s,i)$ and $\sigma(s,i)$ respectively. We further define $\bar{p}(s,i)$ as the number of standard deviations from the mean for the flow $P_{Li}(s)$ on line i under state s to the upper limit $P_{Limi}(s)$ when the overload risk becomes obvious, given by

$$\bar{p}(s,i) \triangleq \frac{P_{Limi}(s) - p(s,i)}{\sigma(s,i)} \tag{5.24}$$

If the line flow $P_{Li}(s)$ is normally distributed, when $\bar{p}(s,i) \geq 4.0$, $Prob(P_{Limi}(s) > \bar{p}(s,i)) < 0.000033$. Even if we consider the influence of generators which makes the distribution deviate from normal, such a small probability can be ignored. The screening condition that only those cases with $\bar{p}(s,i) < 4$ are considered saves a large amount of calculation time. Use of this technique together with our segmentwise cluster based convolution method leads to the final realization of our program that can calculate

thermal overload risk using such a thorough model within a reasonable computation time.

5.7 Risk Calculation

Having determined the probabilistic distributions of the flows through lines and transformers, we can combine these probabilistic distributions with component thermal overload "risk curves" to perform system-wide cumulative thermal overload risk assessment. These curves are computed and stored in advance of the trajectory-based cumulative risk calculation. They provide the expected impact given the flow for all possible flows and all circuits.

5.7.1 Component Risk

Let us denote the flow distribution of line i under state s as $f_i(P_{li}|s)$ for the random flow $P_{Li}(s)$ on circuit i. Then we may obtain the total risk for the hour, for the given circuit i, and for the given system state s, as

$$Risk_i(P_{Li}(s)) = \int_{-\infty}^{\infty} f_i(P_{li}|s) Risk_i(Im_i|P_{li}) dP_{li} \tag{5.25}$$

Here, the term $Risk_i(Im_i|P_{li})$ is the component risk for circuit i. It provides the risk, given the flow for the next hour, under the uncertainty of the impact of this flow on this circuit (denoted by Im_i).

5.7.1.1 Component Risk for Lines

When the circuit is a transmission line, the impact depends on the conductor temperature θ_i, and the component risk is cmputed as

$$Risk_i(\theta_i|P_{li}) = \int_{-\infty}^{\infty} Pr(\theta_i|I_i)\{Im_{sag}(\theta_i) + Im_{annealing}(\theta_i)\} d\theta_i \tag{5.26}$$

where I_i is the conductor current corresponding to the flow P_{li}. The function $Pr(\theta_i|I_i)$ provides the distribution of the conductor temperature under the uncertainty of ambient temperature and the component of wind speed normal to the conductor. The term $Im_{sag}(\theta_i)$ is the economic impact of conductor clearance loss due to sag. The term $Im_{annealing}$ is the economic impact of conductor strength loss due to annealing. Development and use of equation 5.26 is from [74].

Use of equation 5.26 requires that the distribution $Pr(\theta_i|I_i)$, which depends on the ambient temperature and wind speed, match the season and time of day for which they are used. We generally recommend using at least eight distributions: one pair for each of the four seasons such that each pair consists of a day time and a night time distribution. Of course, one may use as many distributions as are necessary for the particular application.

5.7.1.2 Component Risk for Transformers

When the circuit is a transformer, the impact depends on the hottest-spot temperature (HST) of oil. Under a specified operating condition X (typically characterized by transformer loading in terms of current), the risk of a period of T is defined as

$$Risk(X) = Risk_1(X) + Risk_2(X) \tag{5.27}$$

where $Risk_1(X)$ corresponds to loss of life:

$$Risk_1(X) = \int_0^T \int_{\theta_0}^{\theta} Pr(\theta|X) \times Im_{loss\_of\_life} d\theta dt \tag{5.28}$$

and $Risk_2(X)$ corresponds to transformer dielectric failure:

$$Risk_2(X) = \int_0^T \int_{\theta_0}^{\theta} Pr(\theta|X) \times H(t|\theta) \times Im_{failure} d\theta dt \tag{5.29}$$

Here $Pr(\theta|X)$ is the probability density function of the HST θ under the condition X; $H(t|\theta)$ is the hazard function of the transformer dielectric failure given θ at time t;

$Im_{loss\_of\_life}$ is the impact cost of the transformer's loss of life; $Im_{failure}$ is the impact cost of the transformer's dielectric failure [75].

Use of equations 5.28 and 5.29 requires the distribution of $Pr(\theta|X)$, which depends on the ambient temperature. In order to keep consistency, we use the same ambient temperature model here as the model used in the line risk calculation.

5.7.2 Risk Curves

Use of equation 5.26 and 5.27 is computationally expensive, and therefore they are done in separate procedures before the risk calculation. This procedure results in a set of "risk curves", one for each circuit, such as that shown in Figure 5.2. This makes solving equation 5.25 much more efficient.

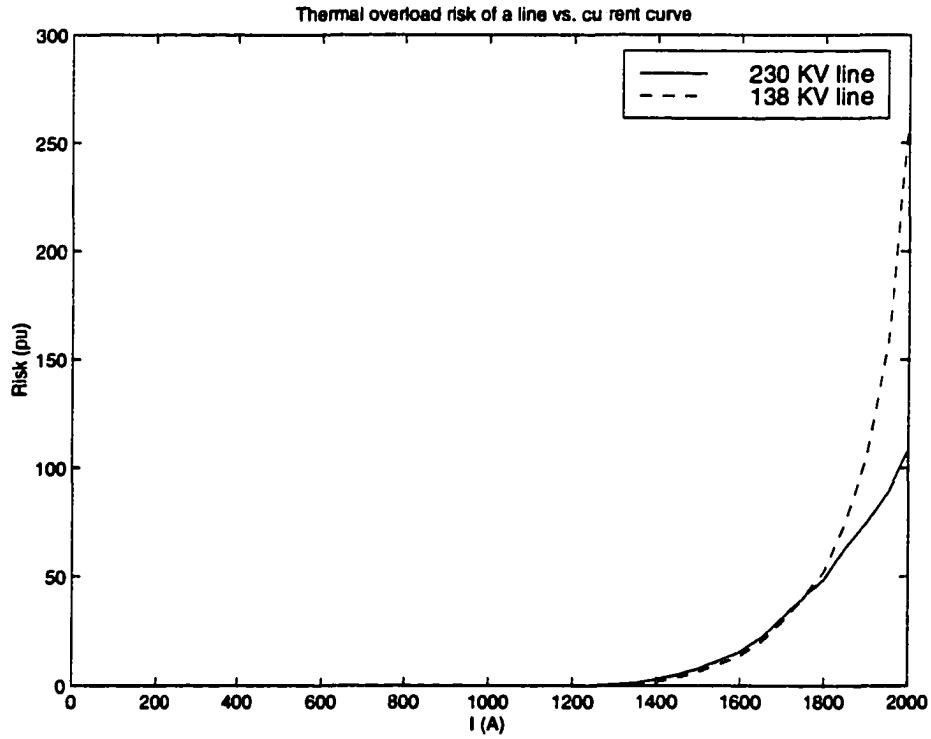

Figure 5.2 The thermal overload risk vs. line current curve

When the distribution of ambient temparature and the average wind speed are given, [74] has already shown the curve of thermal overload risk vs. line flow current can be drawn. According to [74], when the probabilistic distributions of the ambient temperature and the wind speed are chosen to be typical, the thermal overload risk curves for a typical 230 KV line and a typical 138 KV line are shown in Figure 5.2. Here risk is in pu value. One pu risk equals the money required to reconductor the whole line. We choose it to be \$108,000/mile for a 230 KV line, and \$666,667/mile for a 138 KV line according to [11].

Similarly, thermal overload of a transformer could result in aging of materials such as paper and oil [75]. Under the same typical probabilistic distribution of ambient temperature, according to [75], the thermal overload risk vs. load curve for a typcial transformer is shown in Figure 5.3. Here risk is also expressed in pu value. One pu risk equals the money required to rebuild the transformer. We choose it to be \$1,000,000.

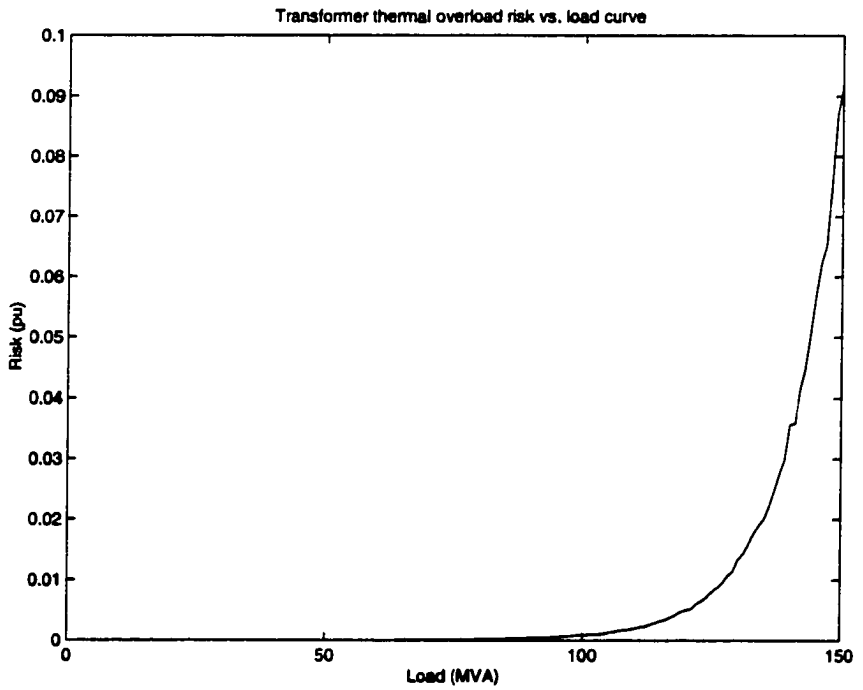

Figure 5.3 The thermal overload risk vs. transformer load curve

5.8 Flowchart of Annual Thermal Overload Risk Assessment

In summary, the flowchart of the annual thermal overload risk assessment module is shown in Figure 5.4.

5.9 Analysis of Calculation Results

For the IEEE RTS'96, the thermal overload risk during a whole year is shown in Figure 5.5. From the figure we can see the 1798th hour has a peak risk which is much higher than other hours. Figure 3.1 shows the peak risk hour is not at peak load at

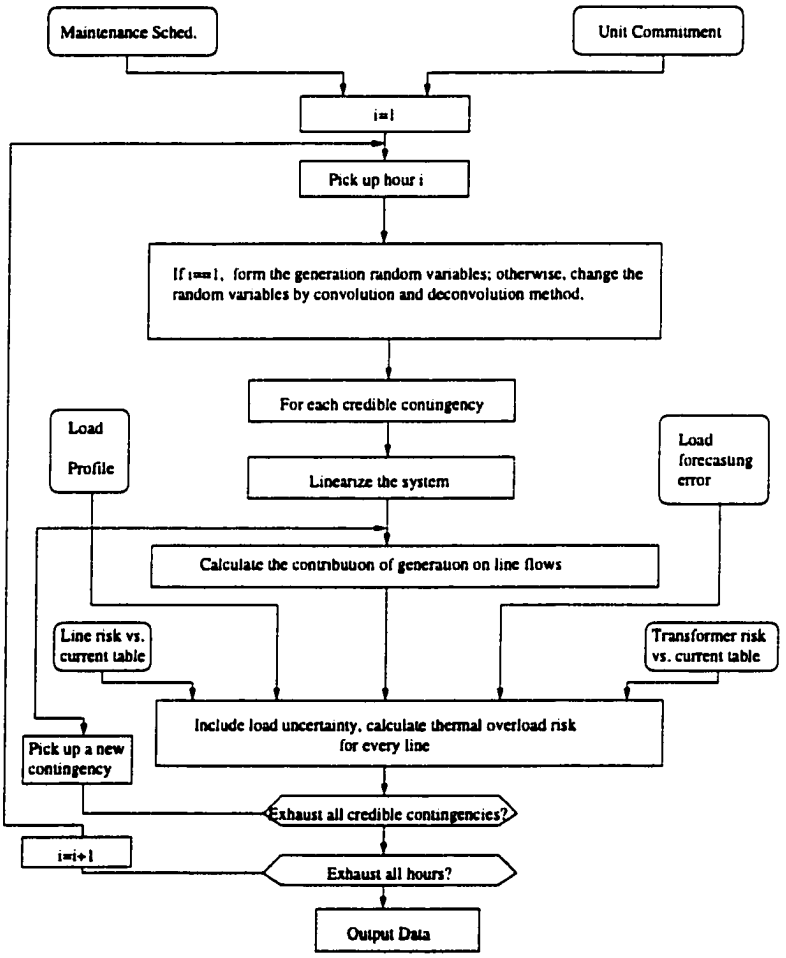

Figure 5.4 The flowchart for thermal overload risk assessment

all. The risk curve of the day including the peak risk hour is shown in Figure 5.6. The corresponding load profile within the same period is shown in Figure 5.7. It is shown in Figure 5.6 that the load decreases a little from hour 1797 to 1798. We checked the unit commitment and found that this little decrement triggers 2 generation units with capacity 197 MW each at bus 13 to be shut down. However, at this time the load is not small enough. This results in long distance power transfer, and causes high thermal overload risk ultimately. If we use snapshot models to represent system trajectory, we are unlikely to capture such a partial peak risk.

There is a period which has almost no risk between hour 5401 and hour 6668. This is because during this time period, generation is distributed more evenly so that customers do not have to require power transfer from a long distance.

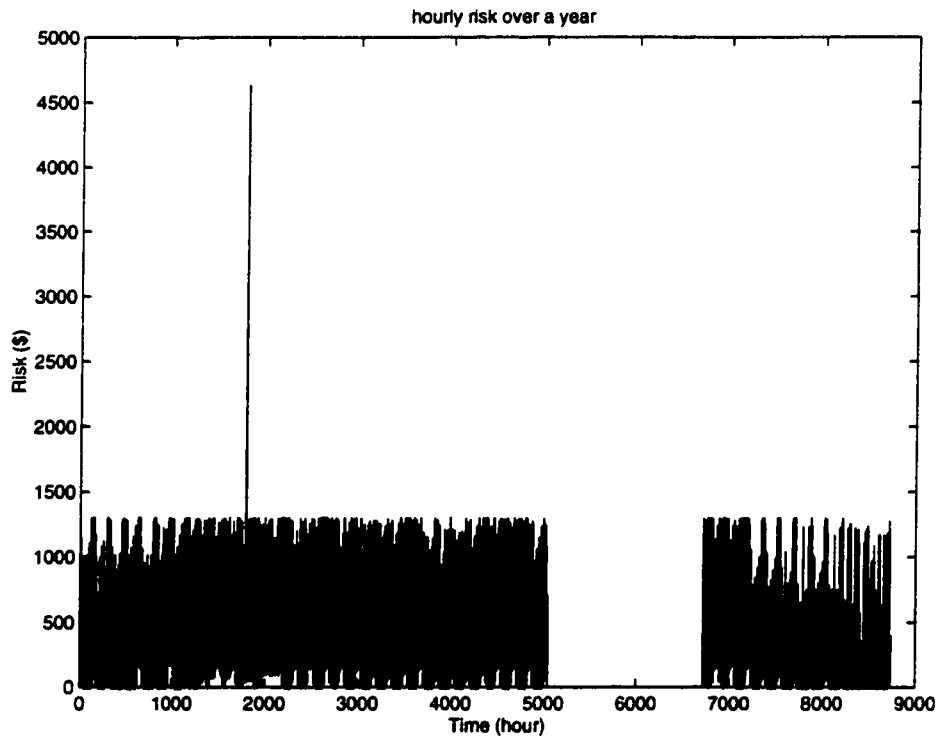

Figure 5.5 Hourly thermal overload risk over a year

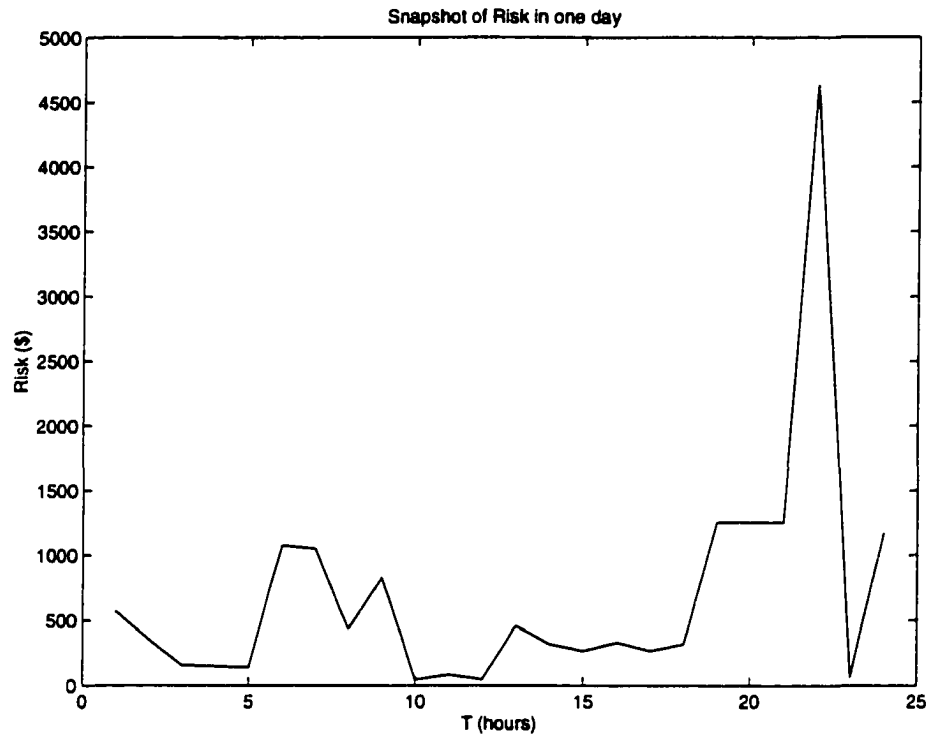

Figure 5.6 Risk curve of the day with the peak risk hour

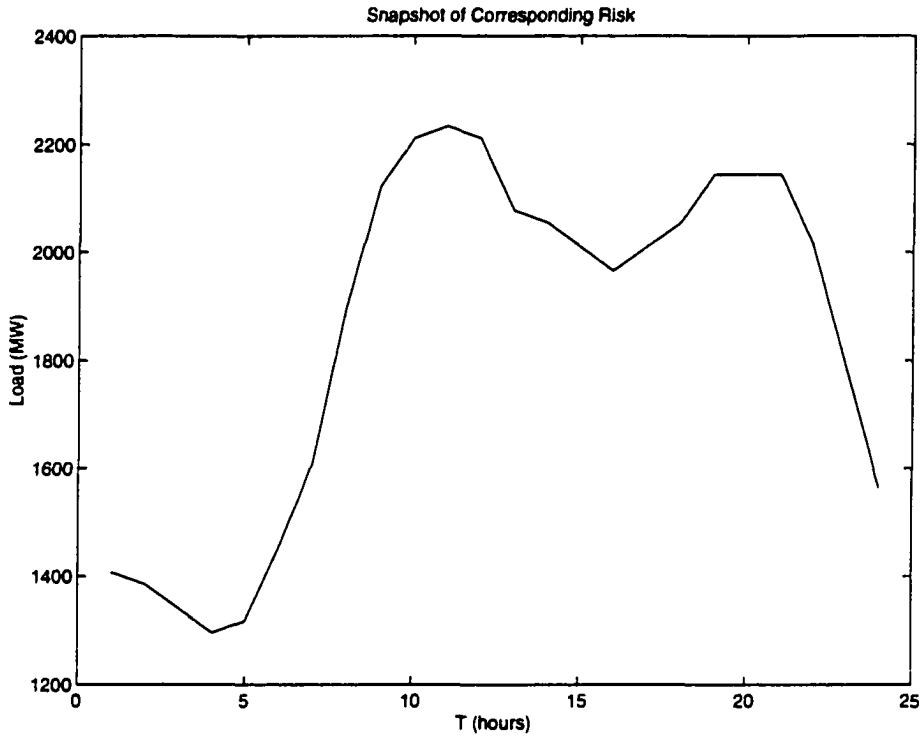

Figure 5.7 Load profile of the day with the peak risk hour

The annual risk suffered by severe lines is listed in Table 5.1. All other lines have less than $1 risk.

There are 46 network states, including normal state, one-line outages, common mode outages. The states which suffer high risk are listed in Table 5.2.

Table 5.1 The annual risk suffered by severe lines

Severe line	Risk ($)
11–13	145
14–16	2333
15–16	979
16–17	2796207
16–19	1300
17–18	139006

Table 5.2 The annual risk for different outages

Outage	Risk ($)
Line 14–16	1300
Line 15–16	217
Line 16–19	979
Line 17–18	2333
Common mode B	145
Common mode D	2934996

The total annual risk is $2,939,971. We can see from Table 5.1 that line 16–17 and line 17–18 suffer severe risk. The owner of these two lines should receive some compensation for their high risk. We can also see from Table 5.2 that common mode D contributes the most to the annual risk. Therefore, if we can build the double lines from bus 15–21 on separated towers, the risk will be significantly reduced.

Figure 5.5 shows risk distribution among hours. Table 5.1 shows decomposition of risk suffered by different components. Table 5.2 shows decomposition of risk caused by different components. Figure 5.5, Table 5.1 and Table 5.2 shows the essence of decomposable and assignable risk assessment.

5.10 Summary of This Chapter

When the whole year's system trajectory is identified, the annual thermal overload risk assessment is just the cumulation of risk over all hours. For every hour, the thermal overload risk assessment is in essence a probabilistic power flow problem and can be solved by a probabilistic power flow algorithm. The fundamental idea of the method proposed in [63] is borrowed to design our own thermal overload risk assessment algorithm. The power system is first linearized around operating point, then decomposed into two parts. The distribution of the part which is the linear combination of loads can be computed directly by employing Gaussian distribution properties. The distribution of the part which is the linear combination of generations can be computed by convolution methods. In order to facilitate our calculation, we develop a new convolution method called segmentwise cluster based convolution method. Annual thermal overload risk assessment performed on the IEEE RTS'96 shows that our proposed method can specify when and where a power system is going to suffer severe risk, and who causes the risk.

CHAPTER 6 COMPOSITE CUMULATIVE RISK ASSESSMENT AND COST BENIFIT ANALYSIS

6.1 Introduction

Having designed methods to perform annual risk assessment for power system infeasibility and thermal overload, we can now identify and compare alternative unit commitment scenarios and facility plans. The basic idea is to identify risk and cost as a single economic measure and then perform cost-benefit comparison between all alternative plans. The IEEE RTS'96 is used to illustrate this idea. There are many other decision making methods based on different ideas according to different preferences [76]. We choose a simple one only for an illustration purpose because the main thrust of this dissertation is on risk calculation rather than decision making. Future researchers will be able to combine this work with decision making theories.

In the process of illustrating how our cumulative risk assessment aids decision making, we will also observe that our risk assessment method offers the advantage of enabling composition or decomposition of risk. One may compose risk into a single system or regional index that is useful for comparing various alternative plans. One may also decompose risk to identify the particular causes of risk and those entities that are incurring the risk as well.

6.2 Composite Cumulative Risk Assessment

The composite cumulative risk assessment can be obtained simply by adding cumulative thermal overload risk and cumulative power flow infeasibility risk. Figure 6.1 illustrates the hourly composite risk variation over a year. We can see that the first half year suffers severe risk, while the second half year is relatively secure.

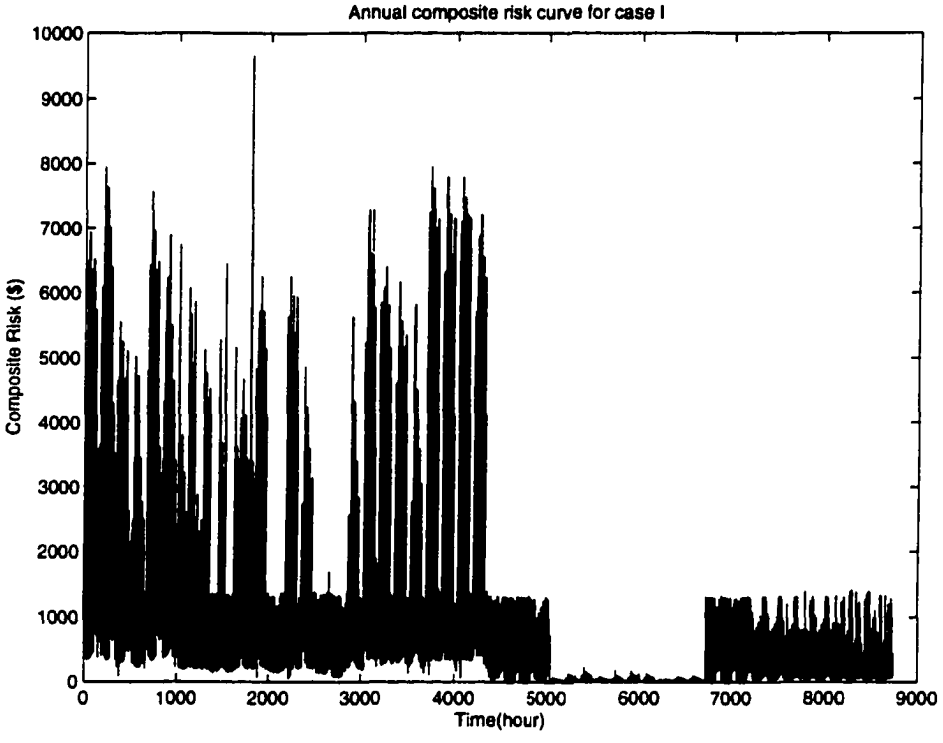

Figure 6.1 Hourly composite risk variation over a year for the original plan

The peak risk occurs at hour 1798, the same hour when the thermal overload risk also reaches its peak. The period from hour 5401 to hour 6668 still has little risk, very similar to the thermal overload risk.

Table 6.1 illustrates how we may report risk associated with each outage state. All states with the risk greater than $10,000 are listed. From this table, first we notice the normal state suffers the largest part of the risk, $5,518,754. This implies that our unit commitment scenario may be inappropriate when transmission constraints are

considered. Second, common mode D outage causes the second largest risk, $2,935,657. If we build the two parallel lines from bus 15 to bus 21 on different towers, because of the very low probability of the concurrent two line outage contingency, this part of risk can be largely eliminated. Third, all transformer outages cause large risk. This shows the reliability importance of transformers in this system. Finally, the outages of both line 7–8 and line 8–10 cause large risk. This illustrates that the subarea constituted by buses 7, 8, 9 and 10 together with the lines between them is a weak area with high risk.

Table 6.1 The annual composite risk contributed by different system states

System State	Risk ($)
normal	5518754
common mode D	2935657
transformer 3-24 outage	15955
line 7-8 outage	12877
line 8-10 outage	12466
transformer 10-12 outage	11870
transformer 10-11 outage	11537
transformer 9-12 outage	10669
transformer 9-11 outage	10599

The annual power flow infeasibility risk is $6,487,586. The annual thermal overload risk is $2,939,971. The annual composite risk is $9,427,557. The composition of risk is shown in Table 6.2. In this table, we use PFI to represent Power Flow Infeasibility, TO to represent Thermal overload. We also use

- Type 1 to represent the PFI risk due to generation inadequacy.
- Type 2 to represent the PFI risk due to voltage instability.
- Type 3 to represent the PFI risk due to voltage exceeding lower limit.

These acronyms will be used throughout this chapter.

Table 6.2 The composition of risk for the original plan ($)

PFI Risk				TO Risk	Total Risk
Decomposition			Total		
Zone 1	Zone 2	Zone 3	6487586	2939971	9427557
458	96382	6390746			
Type 1	Type 2	Type 3			
6456985	4494	26107			

6.3 Alternative Plan I–Loosening the Acceptable Load Voltage Limits

Originally we set the load voltage limit $V_{lmin} = 0.85$, $V_{lmax} = 1.15$, this may cause power flow infeasibility risk when a load bus's voltage reaches the lower bound and can not be lower. If we loosen the load voltage limit to be $V_{lmin} = 0.80$, $V_{lmax} = 1.20$, the composition of risk is shown in Table 6.3, the hourly composite risk variation over a year is shown in Figure 6.2.

Table 6.3 The composition of risk for Alternative Plan I ($)

PFI Risk				TO Risk	Total Risk
Decomposition			Total		
Zone 1	Zone 2	Zone 3	6483226	2939971	9423197
458	96135	6386633			
Type 1	Type 2	Type 3			
6457926	4418	20882			

As expected, the power flow infeasibility risk due to voltage lower limit reduces from $26107 to $20882. However, the total power flow infeasibility risk only drops 0.067%. The thermal overload risk will not be changed when voltage limits are changed. Therefore, lowering voltage limit has little effect on reducing the total risk.

Intuitively, the voltage instability risk should not decrease when the voltage limits

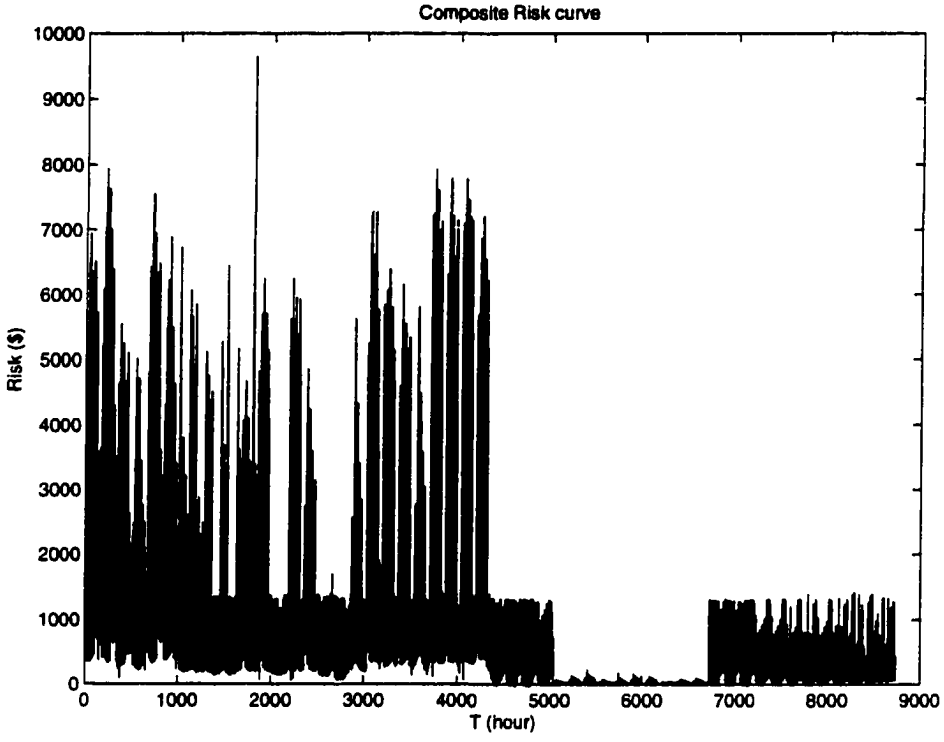

Figure 6.2 Hourly composite risk variation over a year for Alternative Plan I

are loosened. However, our calculation indicates that the voltage instability risk (type 2 risk) reduces a little from \$4494 to \$4418. This may be because we use the "shadow price" of the lowest bus voltage (in pu value) to discriminate whether the risk is due to voltage instability (type 2) or voltage exceeding the lower limit (type 3), but other buses may be comfined by voltage limits as well. When the voltage limits of other buses are loosened, the type 2 risk may be reduced due to the larger space for optimization.

6.4 Alternative Plan II–Tightening the Acceptable Load Voltage Limit

If we tighten the load voltage limit to be $V_{lmin} = 0.90$, $V_{lmax} = 1.10$, the hourly composite risk variation over a year is shown in Figure 6.3, the composition of risk is

shown in Table 6.4. From this table, we can see the total risk increases, as expected. Although the voltage instability risk reduces significantly, it does not disappear. It means there exist some extreme cases which have very normal bus voltages, but suffer voltage stability problem. One must be careful to detect these cases in operations since they are not readily identified by voltage magnitudes.

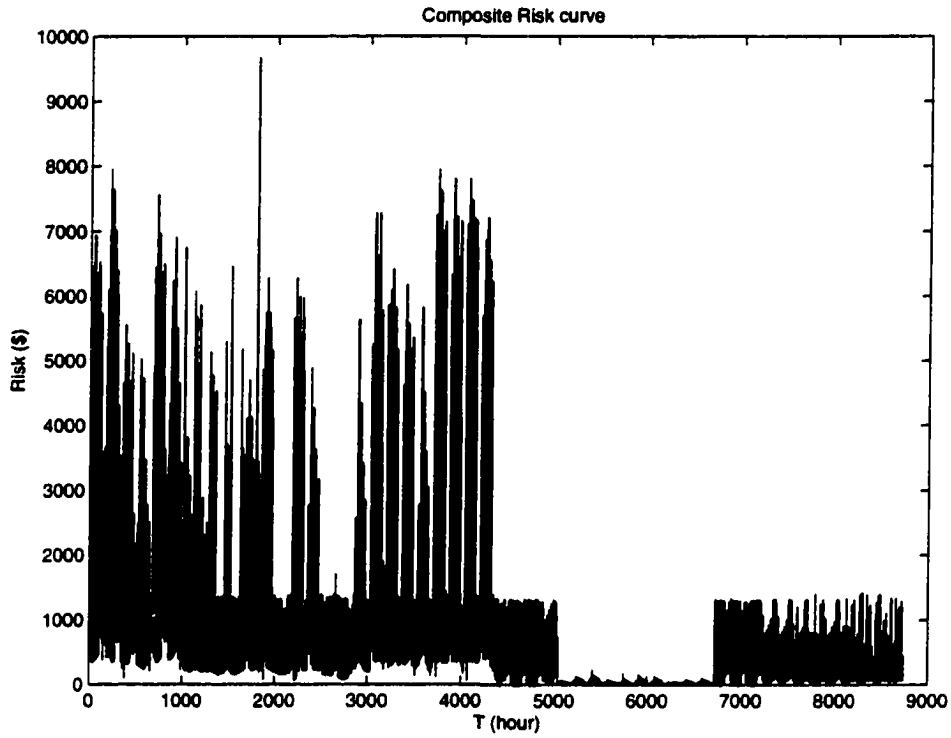

Figure 6.3 Hourly composite risk variation over a year for Alternative Plan II

6.5 Alternative Plan III–Modifying Unit Commitment

Since the major power flow infeasibility risk is due to the generation constraint and zone 3 suffers most of the risk, it is natural for us to constrain on all available units in Zone 3 (which consist only of bus 7) all the time except when they are under maintenance. Under such a scenario, the composition of risk is shown in Table 6.5. The

Table 6.4 The composition of risk for Alternative Plan II ($)

PFI Risk				TO Risk	Total Risk
Decomposition			Total		
Zone 1	Zone 2	Zone 3	6516779	2939971	9456750
528	96104	6420147			
Type 1	Type 2	Type 3			
6462457	421	53901			

hourly composite risk variation over a year is shown in Figure 6.4. Comparison of this figure with Figure 6.1 indicates a large risk reduction has resulted from this change in unit commitment plan (note the difference in scale).

From Table 6.5 we have the following observations.

- Zone 1's power flow infeasibility risk remains almost unchanged, zone 2's increases a little, but zone 3's reduces tremendously.
- Load unbalance risk due to generation constraints reduces tremendously, voltage instability risk disappears altogether, voltage lower limit risk becomes almost zero.
- The total power flow infeasibility risk reduces from $6,487,586 to $128,320, $6,359,266 less.

The corresponding thermal overload risk is $2,904,812, $35,159 less than the original

Table 6.5 The composition of risk for Alternative Plan III ($)

PFI Risk				TO Risk	Total Risk
Decomposition			Total		
Zone 1	Zone 2	Zone 3	128320	2904812	5855423
458	100892	26970			
Type 1	Type 2	Type 3			
128310	0	9			

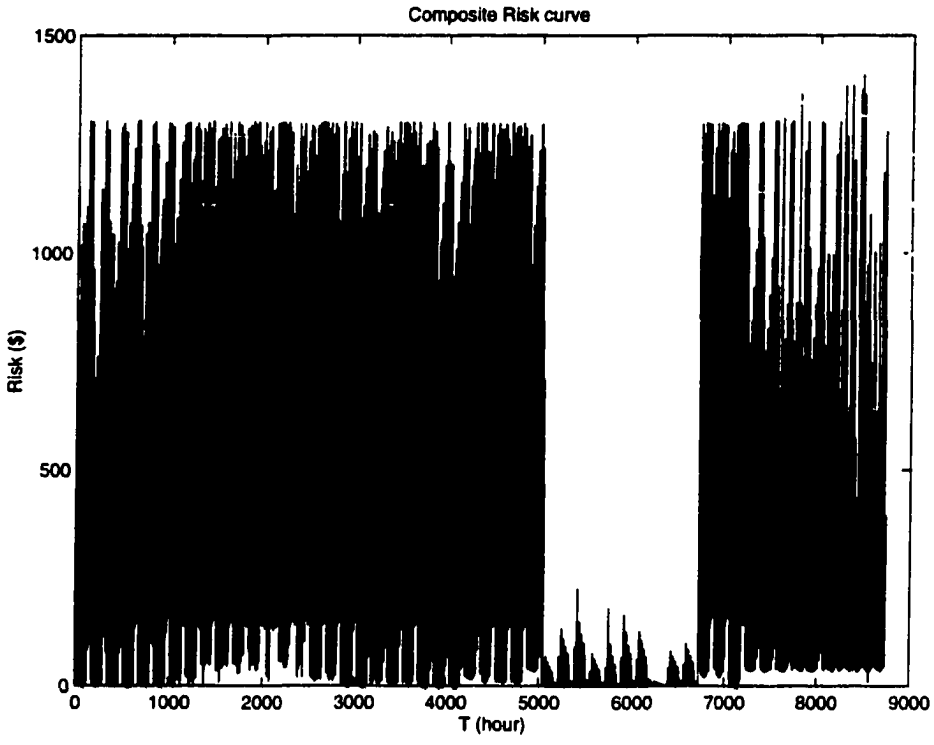

Figure 6.4 Hourly composite risk variation over a year for Alternative Plan III

one. The total composite risk is \$3,033,132, \$6,394,425 less than the original case. Rerunning the unit commitment program for the alternative case we get the total fuel consumption is \$259,832,774, \$5,855,423 more than the original case. Compared with the risk reduction, the benefit of the unit commitment adjustment is \$539,003. It means the generation company at bus 7 can make use of its unique position to make a profit up to \$539,000, if there is no contract between it and the ISO, or between it and electricity customers to put a price cap upon it. The location monopoly prevents the power market from a complete free market. Some regulations such as price cap may be indispensable to guarantee a fair market.

6.6 Alternative Plan IV–Building Double Line 15–21 onto Seperate Towers

Since common mode D causes high thermal overload risk, if we build double line 15–21 onto separate towers, we expect to reduce risk significantly. We call this plan Alternative Plan IV. Under this plan, the composition of risk is shown in Table 6.6, the hourly composite risk variation over a year is shown in Figure 6.5.

Table 6.6 The composition of risk for Alternative Plan IV ($)

PFI Risk				TO Risk	Total Risk
Decomposition			Total		
Zone 1	Zone 2	Zone 3	6486924	4975	6491900
458	96381	6390085			
Type 1	Type 2	Type 3			
6456323	4494	26107			

By adopting this plan, we reduce risk by $2,934,996 for thermal overload risk, $662 for power flow infeasibility risk, and $2,935,657 for composite risk annually. The total thermal overload risk is $4,975, the total power infeasibility risk is $6,486,924, the total composite risk is $6,491,900. According to [11], the construction fee to build another line from bus 15 to bus 21 is $108,000/mile × 33 miles = $3,672,000. If the second year annual composite risk reduction is at least as great as that of the last year, we would earn back our investment within 2 years.

6.7 Alternative Plan V–Combination of Plan III and Plan IV

If we both adjust unit commitment as plan III does and build double line 15–21 onto seperate towers as plan IV does, we get plan V. Under this plan, the composition of risk is shown in Table 6.7, the hourly composite risk variation over a year is shown in Figure

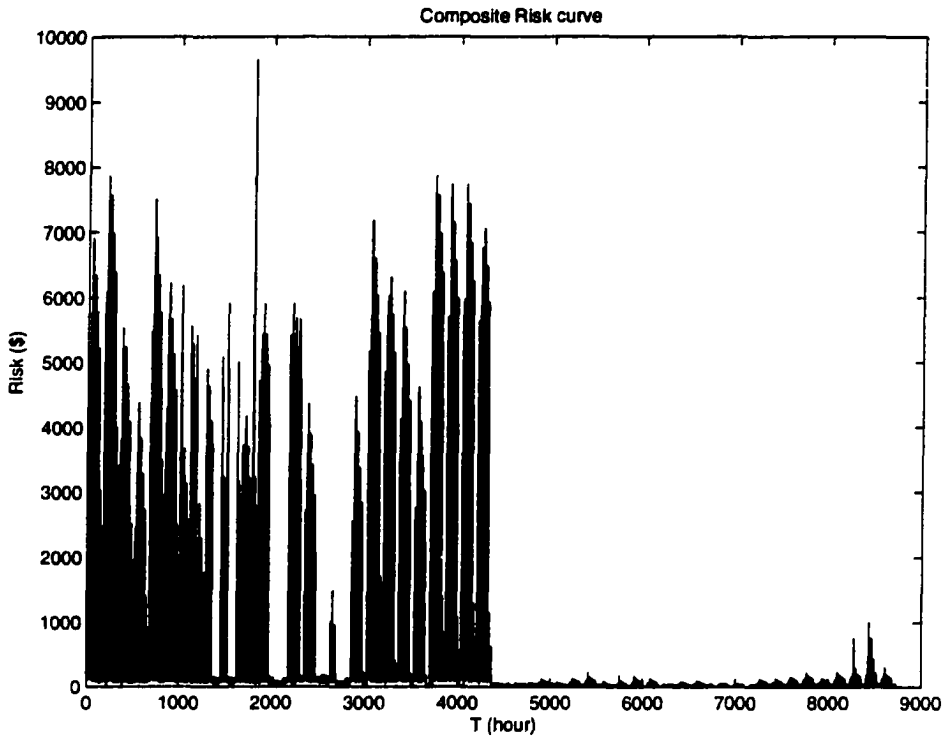

Figure 6.5 Hourly composite risk variation over a year for Alternative Plan IV

6.6.

The total power flow infeasibility risk is \$128,319, the total thermal overload risk is \$251, the total composite risk is \$128,570. The one time construction cost is \$3,672,000. The cost of fuel consumption is increased by \$5,855,423 compared with the original case.

6.8 Comparison Between Plans

Choosing the best plan from several alternative plans belongs to the theory of decision making. According to different criteria [76], one may choose different plans. Although the main goal of this dissertation is to develop the cumulative risk calculation procedure, it is important to point out that this procedure provides information useful in making planning decision. However, the subject of decision making under risk is a rich one [76]

Table 6.7 The composition of risk for Alternative Plan V ($)

PFI Risk				TO Risk	Total Risk
Decomposition			Total		
Zone 1	Zone 2	Zone 3	128319	251	128570
458	100891	26970			
Type 1	Type 2	Type 3			
128309	0	9			

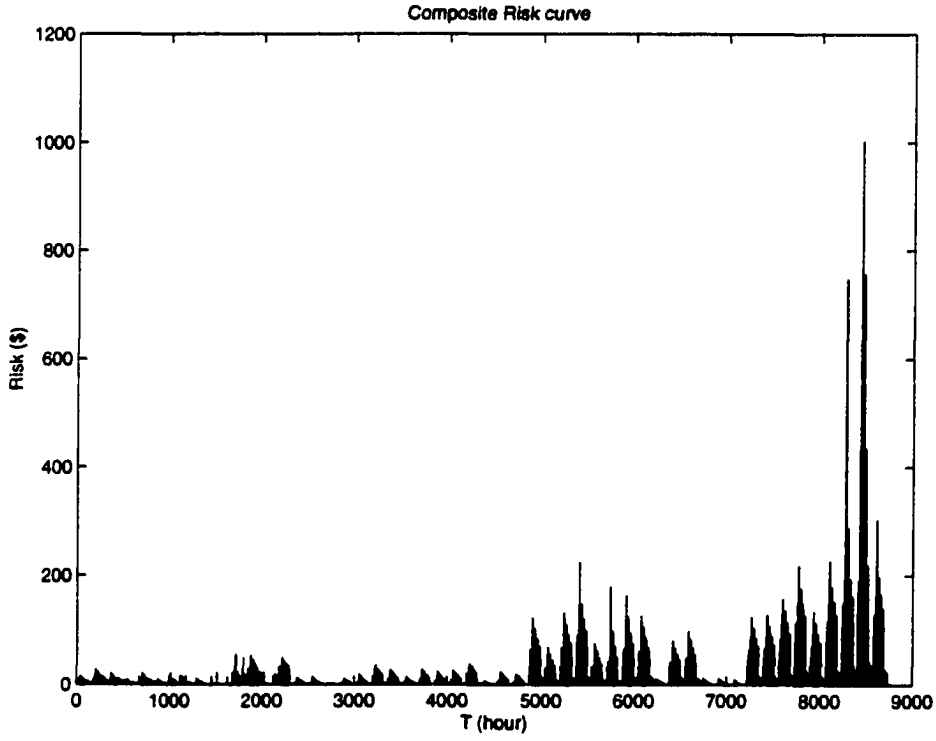

Figure 6.6 Hourly composite risk variation over a year for Alternative Plan V

and deserves much more attention than we can give it here. To make the point, however, we choose a very simple decision paradigm and illustrate its use with respect to choosing among the alternatives studied. This decision paradigm is to select the alternative which has the lowest total annual cost. The cost and benefit comparison between the original plan and the five alternative plans are shown in Table 6.8. Here the total annual cost includes the composite risk, the equivalent annual investment, and the annual fuel cost. The equivalent annual investment comes from the allocation of one time investment among its life cycle. The only investment involved here is rebuilding line 15–21 along a different set of towers. We assume line 15–21 has 20 years of life time, and ignore the time value of money. Then the equivalent annual investment is the one time investment divided by 20. For plan V, its one time investment is $3672000. Then its equivalent annual investment is $3672000/20=$183600. Furthermore, we use fuel cost increment with respect to the original plan instead of the real fuel cost. Therefore, only plan III and plan V have positive fuel cost increment, other plans have zero cost increment. Based on this decision paradigm, we can calculate the savings compared with the original case by subtracting the total annual cost from that of the original case.

From Table 6.8, we can see plan V has the largest saving. Therefore, plan V is the best plan according to our decision paradigm.

Table 6.8 Comparison between different plans (all items are in dollars)

Plan	thermal risk	infeasibility risk	composite risk	annual investment cost	fuel cost increment	total annual cost	savings
original	2939971	6487586	9427557	0	0	9427557	0
plan I	2939971	6483226	9423197	0	0	9423197	4360
plan II	2939971	6516779	9456750	0	0	9456750	-29193
plan III	2904812	128320	3033132	0	5855423	8888555	539002
plan IV	4975	6486924	6491900	183600	0	6665400	2752057
plan V	251	128319	128570	183600	5855423	6167593	3259964

6.9 Summary

Cumulative power flow infeasibility risk and thermal overload risk assessment can be combined to get composite cumulative risk assessment. Based on the risk assessment results, we can design several alternative facility plans or unit commitment scenarios. Then cost and benefit analysis can be performed by our program. For the IEEE RTS'96, 5 alternative plans are proposed. The most beneficial one is to make bus 7's generation units always on as long as they are not at maintenance, and separate the double line 15–21 onto seperate towers. It may be possible for the generation company at bus 7 to take advantage of its market power, depending on its contracts.

CHAPTER 7 SUMMARY OF THE DISSERTATION AND FUTURE RESEARCH DIRECTIONS

7.1 Summary of the Dissertation

Under the pressure of power system restructuring, power system risk assessment becomes more and more important. Risk assessment has some advantages over reliability evaluation in that risk assessment can easily be included into economic cost and benefit analysis. Because different market participants have different objectives in risk assessment, the assessed risk should be decomposable and assignable.

The introduction of competitive market into the power industry has increased uncertainty in operating and planning power system. One consequence of this is that utilities are paying more and more attention to operational planning rather than long term planning. For operational planning, the system trajectory should be considered in order to guarantee appropriate accuracy. Two types of models are proposed to represent system trajectory: the snapshot model and the sequential model. The snapshot model has already been employed by developers of some commercial packages. However, it is not very accurate, and it can not capture high risk time periods when load level is low and unit commitment is not appropriately arranged. The sequential model is a promising model in that it accounts for sequential unit commitment variation and its influence on risk assessment. The sequential Monte Carlo simulation model is accurate but too time consuming. Therefore, we propose a new model called the sequential mean-variance model. It has one expected system trajectory, together with a uniform variation at any

time point.

Based on the proposed model, a cumulative risk assessment framework is built. The framework is highly modularized. It includes a system trajectory identification module and a cumulative risk assessment module. The system trajectory identification module is decomposed to three submodules: load forecasting error identification submodule, maintenance scheduling submodule, and unit commitment submodule. The cumulative risk assessment module includes power flow infeasibility risk assessment submodule and thermal overload risk assessment submodule currently. It can be expanded to include steady state risk assessment submodule and transient risk assessment submodule, both of which belong to dynamic risk assessment.

Various methods are investigated to develop each module required for the trajectory, and one method is chosen for each. The methods we chose are shown to be easy to implement and valid for cumulative risk assessment. However, they are not necessarily the best methods. Due to our modularized framework structure, these methods can be easily upgraded or substituted without modifying the overall program structure.

The interior point algorithm we use in power flow infeasibility risk assessment and the convolution method we use in thermal overload risk assessment are introduced. The potential applications of the interior point algorithm are not limited to risk assessment and are illustrated to some extent in this dissertation. They include maximum loadability calculation, ATC calculation, and applications of shadow prices.

The IEEE RTS'96 is used as the example system throughout the dissertation to show the effectiveness of our framework and algorithms. This choice is for convenience and does not mean our framework can not be applied to other power systems. The risk assessment results show our assessed risk can be decomposed to answer the following questions:

- When will the system suffer severe risk?

- Where will the system suffer severe risk?
- How much risk is suffered by each component?
- How much risk is caused by each contingency?
- What is(are) the reason(s) for a high risk?

Furthermore, based on our risk assessment, we can propose alternative plans and perform cost and benefit comparison between these plans. Five alternative plans are proposed for the IEEE RTS'96 and compared. The most economic one is chosen. Our cumulative risk assessment framework is shown to be effective and useful.

7.2 Contributions of the Dissertation

The dissertation is part of an EPRI project (contract W08604-01). In the report of this project [22], we proposed the concept of risk, and we measure risk directly in dollars. This makes risk easy to understand and easy to include within other economic analysis. Furthermore in [22], we proposed frameworks for operational risk assessment and for cumulative risk assessment for planning. This dissertation is aimed at building the framework of cumulative risk assessment for planning. It has the following contributions:

- It proposed and developed the framework of cumulative risk assessment that identifies the system trajectory and performs various kinds of risk assessment based on this trajectory.
- It decomposed system trajectory identification into load forecasting error identification, maintenance scheduling, and unit commitment arrangement, and then proposed one feasible method for each module.
- It proposed the idea of power flow infeasibility risk, which includes generation inadequacy risk, voltage instability risk and voltage exceeding the lower limit risk.

By using the "shadow prices" in our calculation model, different types of risk can be discriminated.

- It proposed a method to calculate cumulative power flow infeasibility risk based on a zone-based maximum loadability subroutine. This zone-based maximum loadability subroutine can also be used to calculate maximum loadability, ATC (available transfer capability), and provide sensitivity information for generation limits, voltage limits, and tie lie flows.

- It developed a simplified direct interior point algorithm to calculate zone-based maximum loadability. This simplified direct interior point algorithm has only about one-fourth the number of equations and variables compared with the original interior point algorithm in each iteration.

- It explained the reason of oscillation sometimes encountered in the simplified direct interior point algorithm and provided various techniques to overcome it.

- It proposed a method to perform cumulative thermal overload risk assessment by combining the probabilistic power flow calculation with the component thermal overload risk assessment. Component thermal overload risk tables are built as an interface between probabilistic power flow calculation and the component thermal overload risk assessment.

- It developed a probabilistic power flow calculation method based on linearization around the operating point and the decomposition of contributions from the variation of generations and the variation of loads. A new method called the segmentwise cluster based convolution method is developed to calculate the contribution from the variation of generations. The linear transformation of Gaussian distributions is employed to calculate the contribution from the variation of loads.

- Most other approaches measure overload impact in terms of the curtailed load necessary to sufficiently reduce the overload, or in terms of the increased cost of redispatch necessary to reduce the overload, or both. However, we measure the overload impact in terms of the amount of sag and the damage to the conductor. This is because other researchers use deterministic thermal overload limits for lines, while we disregard these hard limits and use the risk in dollars as the unique measure for thermal overload problem. Our method eliminates heterogeneous criteria applied among different lines, and shows the potential risk and benefit to exceed the traditional thermal overload limits without taking corrective actions. It is suitable for planning under restructured power market where generation companies are willing to pick up more power, transmission companies are willing to push the thermal limits of lines and transformers, and distribution companies are unwilling to curtail load. Trying to avoid high thermal overload risk in the planning stage without considering corrective actions can provide clues for choosing alternative plans. The best plan chosen from our cost and benefit analysis can avoid the difficulty of corrective coordination between different companies, while it reduces the system thermal overload risk under an acceptable level.

- It combined the thermal overload risk and the power flow infeasibility risk together to get the composite cumulative risk assessment.

- It advocated decomposible and assignable risk assessment so that the following questions can be answered by our risk assessment.

  - When will the system suffer severe risk?
  - Where will the system suffer severe risk?
  - Which contingency plays the key role in causing risk?

  - What is the main cause of risk? Thermal overload, generation limit, or voltage stability?

- It performed risk assessment for one original plan and five alternative plans for the IEEE-RTS96. It applied a simple decision paradigm to compare costs and benefits of these plans and choose the best plan from them. It illustrated that our proposed risk assessment framework can be used for decision making analysis.

7.3 Conclusions

Power systems are undergoing structural changes due to "deregulation". This brings the necessity and urgency of decomposable and assignable cumulative risk assessment whose results are expressed directly in dollars. However, current power industry is still accustomed to traditional reliability evaluation, which usually provides system-wide indices, only captures limited number of snapshots, is normally conservative and sometimes inaccurate, and is also difficult to understand and difficult to connect with economic analysis. Under this environment, we developed a systematic framework to identify system trajectory over a period of time and perform decomposable risk assessment based on the identified system trajectory. Our program based on the proposed framework can identify when the system will suffer severe risk, where the system will suffer severe risk, who plays the key role in causing risk, and what is the main cause of risk. It can also provide clues for alternative risk reduction plans, and then perform cost and benefit comparison between different plans. This comparison is useful for long term facility planning and long term contract economic analysis as well. Our program can also be combined with contract analysis to determine who is going to charge who for what and by how much to insure the fair power market. It seems there is a wide range of use for the decomposable cumulative risk assessment framework.

7.4 Future Research Directions

In the future, this research can be extended along the following directions.

- Design modules for steady state and transient risk assessment.
- Improve methods employed in all submodules.
- Consider transmission constraints in maintenance scheduling and unit commitment.
- Consider corrective actions and power system models in more details.
- Design programs to calculate the variance of assessed risk.
- Include decision making theories into cost and benefit comparison between alternative plans.
- Include economic competitive market models and game theories into risk assessment and decision making.
- Consider parallel computation.
- Consider sequential linear programming, quadratic programming or some other classical algorithms for maximum loadability calculation and compare their robustness and efficiency.

APPENDIX A IEEE RTS'96

The data of this system is listed in [60].

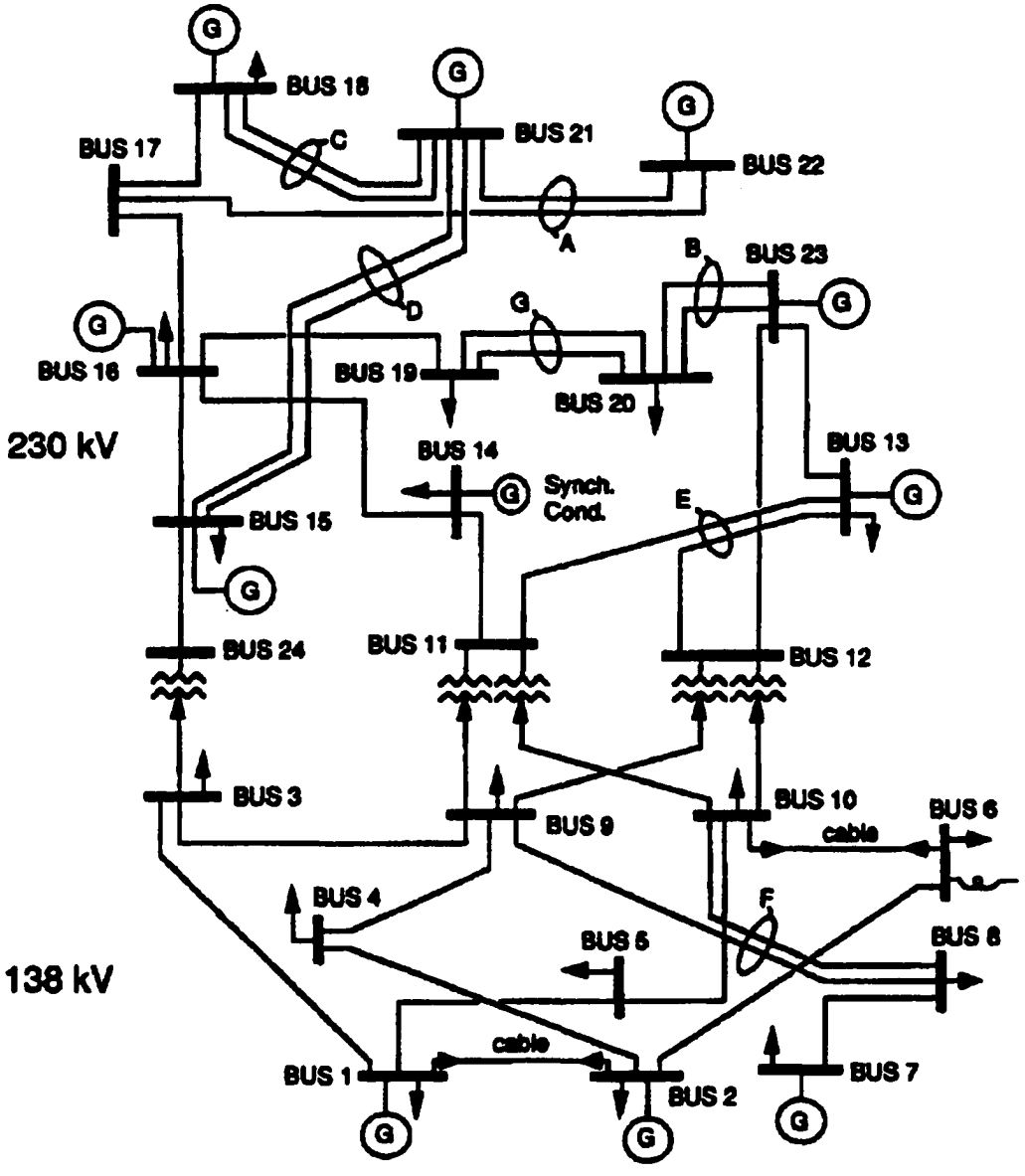

Figure A.1 The network of IEEE RTS'96

APPENDIX B DETAILS FOR LOAD FORECASTING ERROR IDENTIFICATION

Theory of Time Series Analysis

Time Series Models

There are many kinds of stochastic models. The widely used one is the autoregressive moving average (ARMA) model. The general expression is as follows.

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \cdots + \phi_p(x_{t-p} - \mu) + \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} \tag{B.1}$$

where

- x_t is the stochastic signal at time t.
- μ is the expected value of x_t. For a stationary process, μ remains the same at any time, and it can be estimated by the average of sample observations according to the weak law of large number.
- $\phi_1, \cdots, \phi_p$ are the AR part coefficients.
- p is the order of the AR part.
- ε_t is a Gaussian white noise with variance σ^2, that means ε_t is described by $N(0, \sigma^2)$, and $\varepsilon_i, \varepsilon_j$ are independent for all $i \neq j$.
- q is the order of the MA part.

- $\theta_1, \cdots, \theta_q$ are the MA part coefficients.

We define $\tilde{x}_t = x_t - \mu$ and denote the model as ARMA(p,q). When $\theta_1, \cdots, \theta_q = 0$, the model becomes an autoregressive (AR) model; when $\phi_1, \cdots, \phi_q = 0$, the model becomes moving average (MA) model.

If we use B to denote the backward shift operator, we can get

$$\phi(B)x_t = \theta(B)\varepsilon_t \tag{B.2}$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \tag{B.3}$$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B_q \tag{B.4}$$

When we choose $B = z^{-1}$, we can get the z transformation form of equation 1 as follows.

$$\phi(z^{-1})X(z) = \theta(z^{-1})\epsilon(z) \tag{B.5}$$

or

$$X(z) = H(z)\epsilon(z) \tag{B.6}$$

where $H(z) = \frac{\theta(z^{-1})}{\phi(z^{-1})}$ is the transfer function, $X(z)$ and $\epsilon(z)$ are the transformation of x_t and ε_t respectivley. Therefore, $\tilde{x}_t$ can be regarded as an output of a linear discrete system with transfer function $H(z)$ driven by a white noise ε_t. To keep the process stationary, all poles of $H(z)$ (or all zeros of $\phi(z^{-1})$) should lie inside the unit circle. If one pole is outside the unit circle, $\lim_{t\to\infty} x_t$ is unbounded. Such a system is to be avoided.

If one pole is on the unit circle, a seasonal model, generally called the autoregressive integrated moving average (ARIMA) model, is often used. The general form of $\phi(B)$ is

$$\phi(B) = (1 - B^r)^d x_t \tag{B.7}$$

where r, d are positive integers. If we denote

$$y_t = (1 - B^r)^d x_t \tag{B.8}$$

then $\tilde{y}_t = y_t - \mu_y$ is according to ARMA model. For a strict ARIMA model, $\mu_y = 0$. However, μ_y can be nonzeros to include a linear evolving trend.

This model can be denoted as ARIMA(p,d,r,q). Due to error stability consideration, d seldom exceeds 2.

When $\theta(z^{-1})$ has zeros lying in the unit circle, from x_t, ε_t can be observed. Such a property is called invertibility. However, when one zeros lies outside the unit circle, the system is not invertible and ε_t can not be observed. Therefore, such a system model should be avoided.

Decomposition of Model Building and Error Analysis

Although the maximum likelihood estimations for the ARIMA model parameters are very good estimations theoretically, they are difficult to calculate. On the other hand, our algorithm should accommodate different models such as artificial neural network, chaos, or fuzzy logic. Therefore, we decompose the process of estimating the AR part parameters and the error analysis for MA part parameters. The AR part parameters can be easily obtained by the least square method. Suppose the time series begins from $t = 1$. From the data, we have the following equations.

$$\begin{aligned} e_{p+1} &= x_{p+1} - \phi_1 x_p - \cdots - \phi_p x_1 \\ &\vdots \\ e_N &= x_N - \phi_1 x_{N-1} - \cdots - \phi_p x_{N-p} \end{aligned}$$

where e_t is the error series. The foregoing equations can be expressed as

$$\underline{x} - U\underline{\phi} = \underline{e} \tag{B.9}$$

where $\underline{x} = \begin{bmatrix} x_{p+1} \\ \vdots \\ x_N \end{bmatrix}$, $U = \begin{bmatrix} x_p & \cdots & x_1 \\ \vdots & & \vdots \\ x_{N-1} & \cdots & x_{N-p} \end{bmatrix}$, $\underline{\phi} = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_p \end{bmatrix}$.

According to the least square method [77], the least square estimation

$$\hat{\underline{\phi}} = (U^T U)^{-1} U T \underline{x} \tag{B.10}$$

Then we can get

$$\hat{\underline{x}} = U\hat{\underline{\phi}} = U(U^T U)^{-1} U^T \underline{x} \tag{B.11}$$

However, any other model to get $\hat{\underline{x}}$ is acceptable. Once we get $\hat{\underline{x}}$, we can calculate $\underline{e}$. Then $\underline{e}$ is ready for further time series analysis.

Time Domain Analysis: Autocorrelation and Partial Autocorrelation coefficients

The autocorrelation of a stationary stochastic process is defined as the normalized autocovariance.

$$\rho_m = \frac{\gamma_m}{\gamma_0} = \frac{E(\tilde{x}_t \tilde{x}_{t-m})}{E(\tilde{x}_t^2)} \tag{B.12}$$

where ρ_m is the autocorrelation at time m, γ_m is the autocovariance at time m.

Another useful measure is the partial autocorrelation. The mth partial autocorrelation (denoted α_{mm}) is defined as the last coefficient in a linear projection of x_t on its m most recent values.

$$\begin{bmatrix} \alpha_{m1} \\ \alpha_{m2} \\ \cdots \\ \alpha_{mm} \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{m-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{m-2} \\ \vdots & & & \vdots \\ \gamma_{m-1} & \gamma_{m-2} \cdots \gamma_0 & & \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_m \end{bmatrix} \tag{B.13}$$

Note that $\gamma_0, \cdots, \gamma_m$ are all expected values. Theoretically, they should be obtained from an unlimited number of data. However, they can also be estimated from a limited sample data.

One estimation method for γ_m is to use the following formulus[77].

$$\hat{\gamma}_m = \frac{1}{N}\sum_{n=0} N - m - 1 x_N(n) x_N(n+m) \tag{B.14}$$

where $\hat{\gamma}_m$ is the estimated data of γ_m, $x_N(n)$ is the periodical expansion of $x(n)$ with period N. However, when N is extremely large, this method is slow. In such a case, we can use the power spectrum to estimate all γ_m at the same time. The algorithm is as follows [78][79].

1. Add N zeros after observed $x_t(t = 0, 1, \cdots, N-1)$ to get $x_{2N}(n)$, then use FFT to get $X_{2N}(k)$, $k = 0, \cdots, 2N-1$.
2. Calcualte $\frac{1}{N}|X_{2N}(k)|^2$.
3. Apply IFFT to $\frac{1}{N}|X_{2N}(k)|^2$ to get $\hat{\gamma}_m$, $m = 0, 1, \cdots, 2N-1$. Only $\hat{\gamma}_0, \cdots, \hat{\gamma}_{N-1}$ is required because of the symmetry.

α_{mm} can be estimated by the Levinson-Durbin recursive algorithm, which is described in details in [77],[78] and [79].

For AR(p) process, $\hat{\alpha}_{mm}$ are approximately independently and normally distributed with zero mean and variance $\frac{1}{N}$. Therefore, the probability of $\hat{\alpha}_{mm}$ located within $\pm\frac{2}{\sqrt{N}}$ is approximately 95% [77].

For MA(q) process, we have

$$Var(\hat{\rho}_m) = \frac{1}{N}[1 + 2\sum_{i=1}^{q} \rho_i^2] \tag{B.15}$$

for $m = q+1, q+2, \cdots$.

Thus in particular, if we suspect that the data is an MA(q) process, we can check whether $\hat{\rho}_m$ lies between $\pm\frac{2}{\sqrt{N}}$ about 95% of the time when $m > q$.

Frequency Domain Analysis

For an unlimited length signal, the power spectrum $S(e^{j\omega})$ and the autocovariance γ_m has the following relationship.

$$S(e^{j\omega}) = \sum_{m=-\infty}^{\infty} \gamma_m e^{-j\omega m} \gamma_m = \frac{1}{2\pi} \int_0^{2\pi} S(e^{j\omega}) e^{j\omega m} d\omega \tag{B.16}$$

For the limited length signal, the period density spectrum $\hat{S}_{per}^{2N}(k)$, is just the DFT of $\hat{\gamma}_m$.

$$\hat{S}_{per}^{2N}(K) = \sum_{m=-(N-1)}^{N-1} \hat{\gamma}_m e^{-j\omega m} \tag{B.17}$$

$$\hat{\gamma}_m = \frac{1}{2N} \sum_{k=-(N-1)}^{N-1} \hat{S}_{per}^{2N}(k) e^{j\frac{2\pi kt}{2N}} \tag{B.18}$$

For the linear system in ARIMA model described in B.7, we have

$$S_x(e^{j\omega}) = S_\varepsilon(e^{j\omega}) |H(e^{j\omega})|^2 = \sigma^2 |H(e^{j\omega})|^2 \tag{B.19}$$

Therefore, if $H(z)$ has a pole near $e^{j\omega_0}$, then $S_x(e^{j\omega})$ should be extremely high at frequency ω_0. So $S_x(e^{j\omega})$ is a good measure to find the seasonality of the ARIMA model.

Furthermore, it is proved that minimizing the error in the time domain is the same as minimizing the error in the frequency domain [80].

For the ARIMA model, when $d = 1$, $y_t = x_t - x_{t-r}$, or in the frequency domain $Y(z) = (1 - z^{-r})X(z)$. Denote $F(z) = 1 - z^{-r}$, then $F(z)$ is a transfer function of a filter.

$$F(e^{j\omega}) = 1 - e^{-jr\omega}) = (1 - \cos(r\omega)) + j \sin(r\omega) = 2 \sin \frac{r\omega}{2} e^{j(\frac{\pi - r\omega}{2})} \tag{B.20}$$

when $r = 24$, the magnitude-frequency curve is shown in Figure B.1.

From Figure B.1 we can see that the filter will damp $\omega = \frac{\pi}{12}, \frac{2\pi}{12}, \cdots, \pi$ parts of signal. Therefore, if there are peaks at these points in the raw signal's spectrum, the filter is desirable to flatten the spectrum.

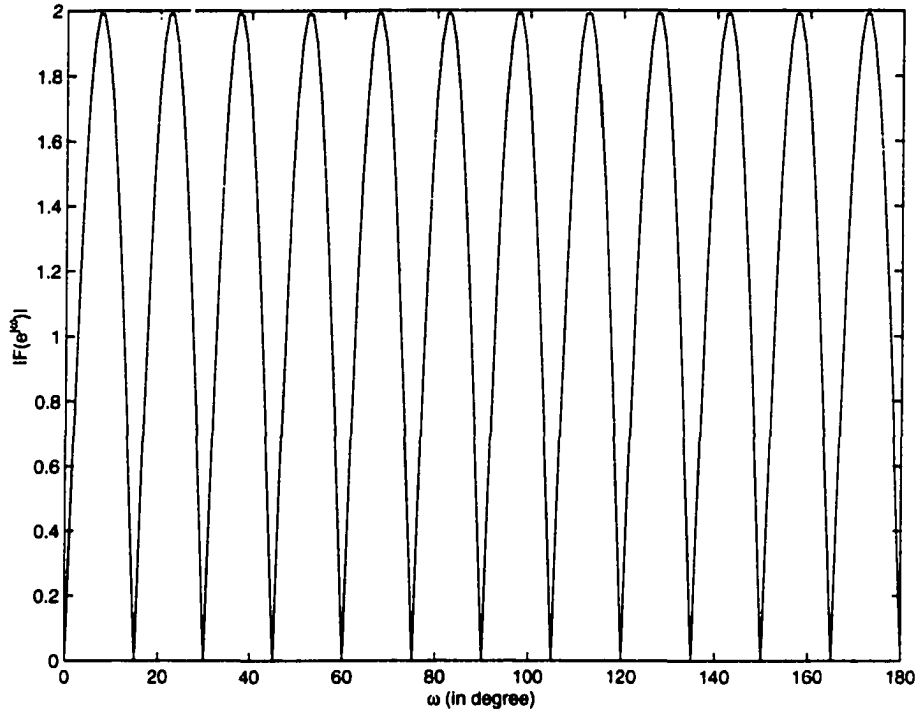

Figure B.1 The magnitude-frequency curve for $F(z) = 1 - z^{-r}$

In the next section, these theories will be used to identify the load forecasting error of the IEEE RTS'96.

Applications to the IEEE RTS'96

Order Determination for the load of the IEEE RTS'96

The IEEE RTS'96 is shown in Appendix A. The hourly load during one year is shown in Figure B.2. The estimated spectrum is shown in Figure B.3. It has obvious frequency peaks once a day, once a week, and twice a day respectively. It is reasonable to apply differentiation by choosing $r = 24$ first to get rid of day cycling trend, then use AR(168) model for the differentiated signal to count on the week's influence. The described model is model ARIMA(168,1,24,1), where q is the order of the MA part to be determined.

Let $y_t = x_t - x_{t-24}$. The signal is shown in Figure B.4. The AR part coefficients for

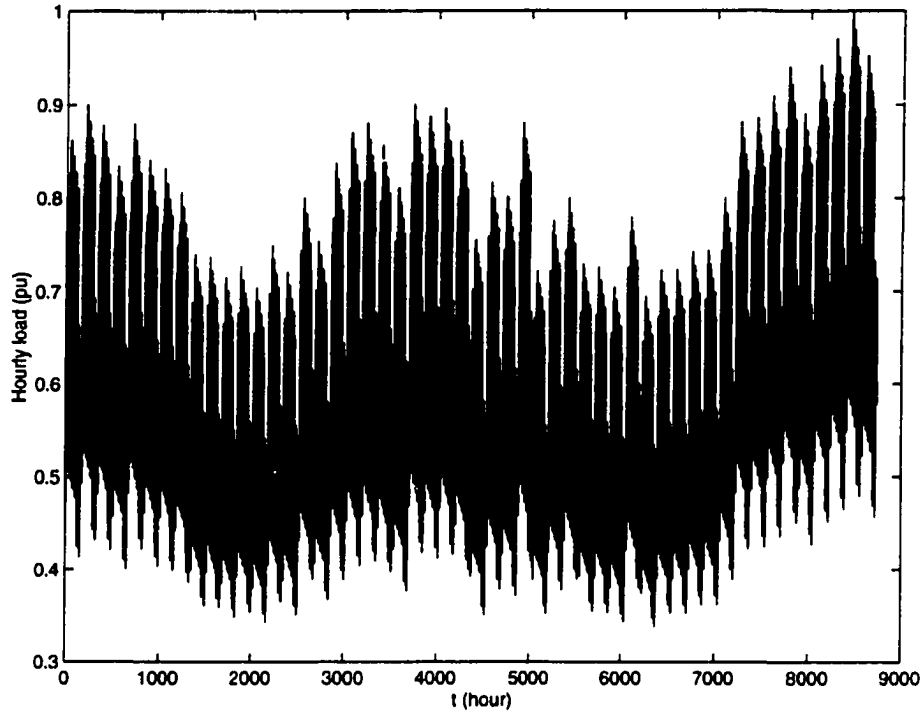

Figure B.2 The hourly load during one year

$\tilde{y}_t = y_t - \hat{\mu}_y$ are shown in Figure B.5. According to our calculation, $\hat{\mu}_y = -2.0124 \times 10^{-4}$, very close to 0. It means the average of the load will keep constant in the near future except for some oscillations. The error series after subtracting the AR part estimation is shown in Figure B.6. The estimated spectrum of the error series is shown in Figure B.7. The estimated autocorrelation coefficients (ACF) of the error series are shown in Figure B.8. The average ratio to exceed the bound $\pm\frac{2}{\sqrt{N}}$ is 4.82%, just below 5%. Therefore, it is reasonable to assume the MA part error series is a white noise. The partial autocorrelation coefficients (PACF) of $\tilde{y}_t$ are shown in Figure B.9. It cuts off after 168, with the ratio 1.86% to exceed the bound, well below 5%. Therefore, our model is correct. In our model, the estimated σ is 6.86×10^{-3}.

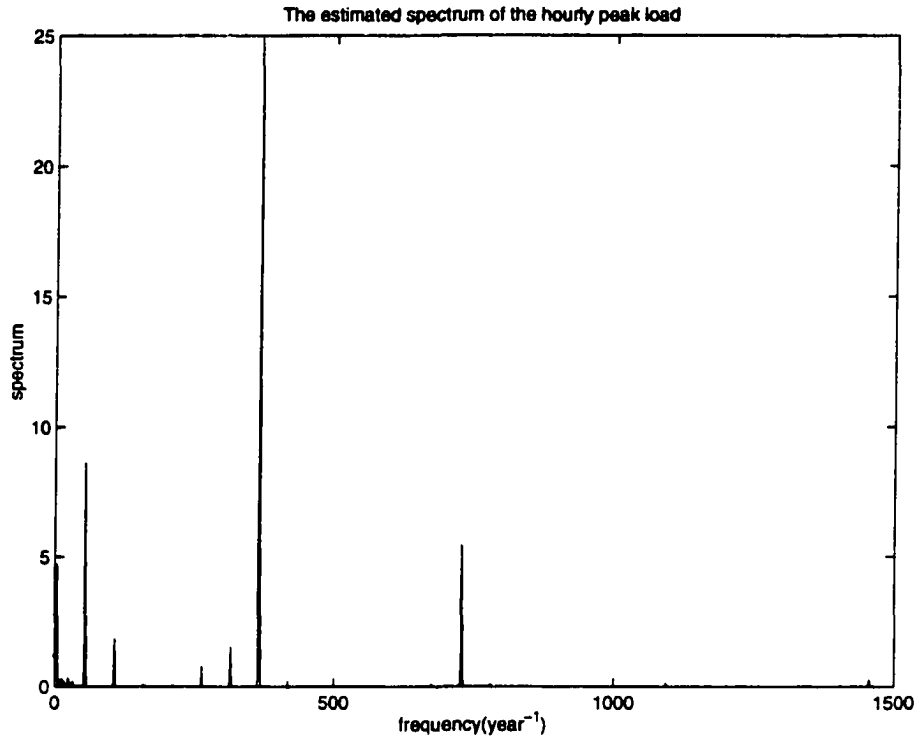

Figure B.3 The spectrum of the hourly load

Load Forecasting

Any ARIMA model can be written in the form

$$\xi_t = F\xi_{t-1} + C + V_t \tag{B.21}$$

where

$$\xi_t = \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-p-rd+1} \end{bmatrix}, F = \begin{bmatrix} \phi_1 & \cdots & \phi_{p+rd-1} & \phi_{p+rd} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, C = \begin{bmatrix} \phi_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, V_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Then

$$\xi_{t+s} = F^s\xi_t + (I + F + \cdots + F^{s-1})C + (V_{t+s} + FV_{t+s-1} + \cdots + F^{s-1}V_{t+1}) \tag{B.22}$$

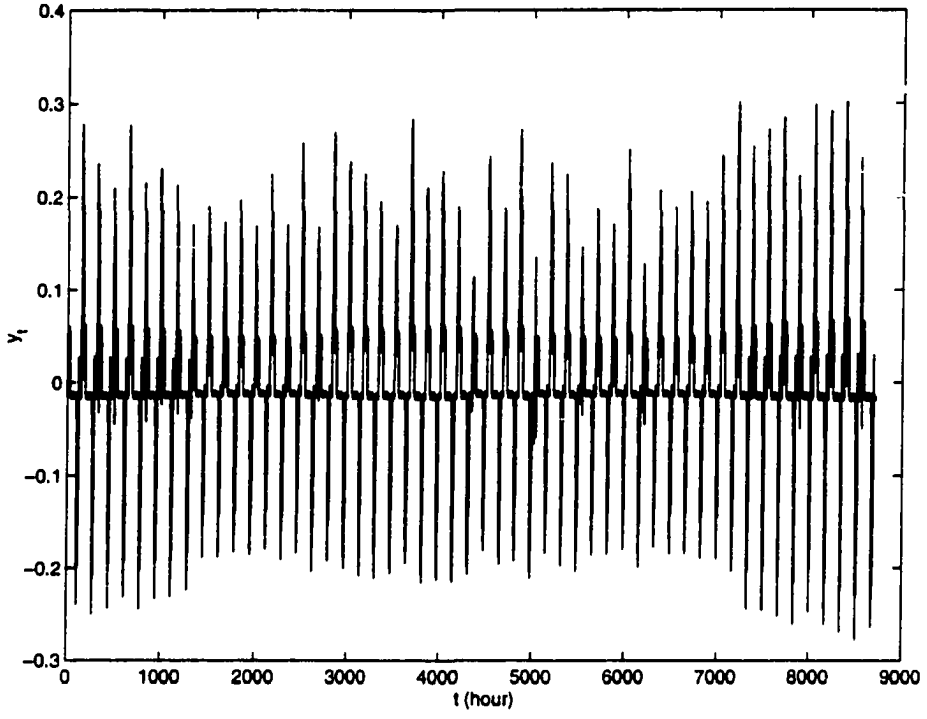

Figure B.4 The signal y_t's curve

Denote f_j to be the first element of the matrix F^j, then the error item

$$e_{t+s}|\xi_t = \varepsilon_{t+s} + f_1\varepsilon_{t+s-1} + \cdots + f_{s-1}\varepsilon_{t+1} \tag{B.23}$$

Therefore

$$Var(x_{t+s}|\xi_t) = (1 + f_1^2 + \cdots + f_{s-1}^2)\sigma^2 \tag{B.24}$$

Denote $\delta_s = \sqrt{Var(x_{t+s}|\xi_t)}$. Then x_{t+s} should lie between $\hat{x}_{t+s} \pm 2\delta_s$ with probability 95%. The forecast load and 95% confidence interval for every hour in the next week is shown in Figure B.10. From this figure we can see that the error is small during the next week. Therefore, our forecast is accurate. Figure B.11 shows the $\delta - s$ curve. It seems that δ increases steadily and there is no trend to stop. It means the ARIMA model is only appropriate for short term load forecast. For one year load forecast, the error is so large that the forecast values are meaningless. In fact, this property is also a common

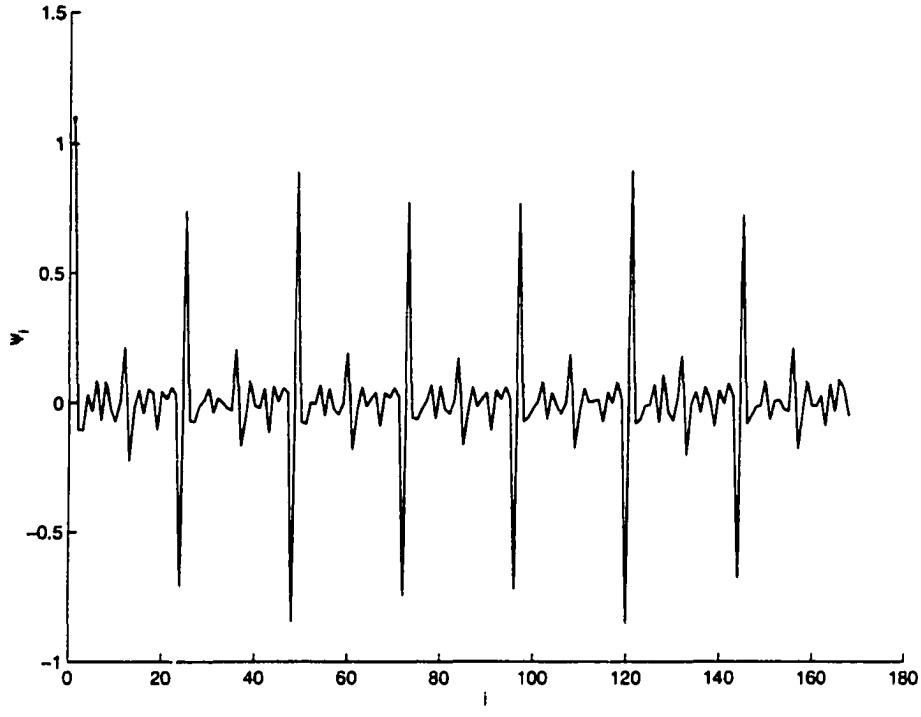

Figure B.5 The AR part coefficients

property shared by most other forecasting tools. It is just a mathematical statement of the common law that near future can be predicted well, but the far future can not.

According to our calculation, the one-day ahead forecasting error

$$\delta_{24} = 1.92\%$$

We will use this error as the error in our load model. The reasons are explained before.

Expansion to Multivariate Case

In the power system, there are usually several load buses and the load in each bus varies differently. In this case, the ARIMA model expressed in B.7 can also be applied, except that x_t now becomes a vector, $\phi_1, \cdots, \phi_p$ now becomes matrices, and ε_t becomes a random vector with a mean vector and a covariance matrix. The model parameters

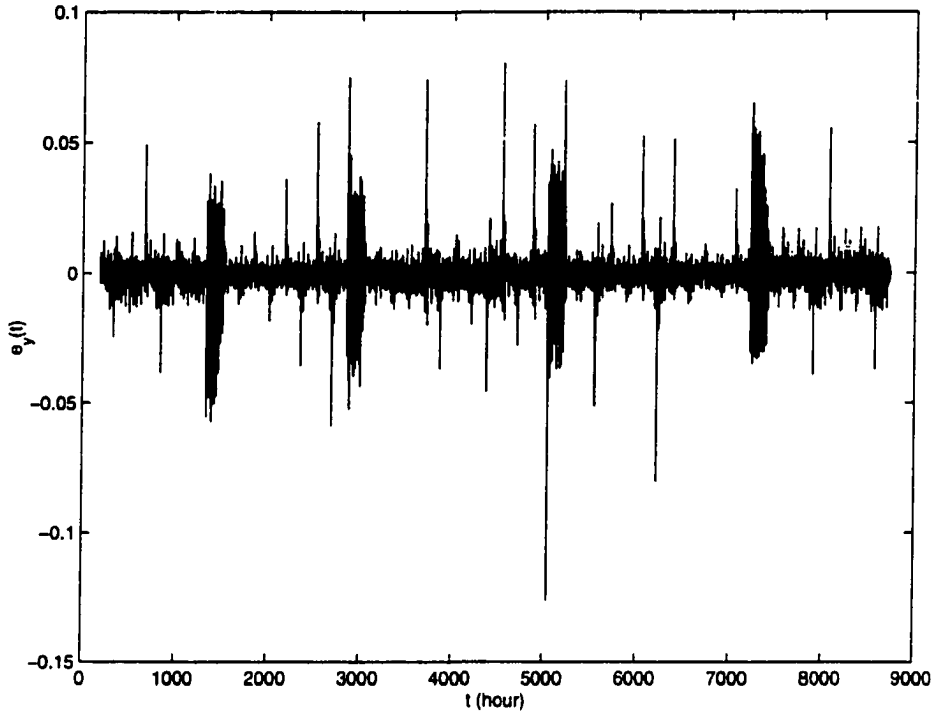

Figure B.6 The error series

can be identified by least square method. The model can then be expressed in state space form, and the Karlman recursions can be applied for load forecast [77].

In power flow feasibility risk assessment, we do not need to consider the correlation between loads, simply regard them varying proportionally. In thermal overload risk assessment however, our algorithm allows to consider the covariance matrix among loads. Since we do not have data, we assume the covariances among loads are as follows.

$Var(P_{Dk}) = \delta_{24}^2 P_{Dk}^2$,

$Cov(P_{Dk}, P_{Dl}) = 0.5\delta_{24}^2 P_{Dk} P_{Dl}$, when $k, l \in \Omega_{138}$, or $k, l \in \Omega_{230}$.

$Cov(P_{Dk}, P_{Dl}) = 0$, when $k \in \Omega_{138}$, $l \in \Omega_{230}$.

Here k, l are bus numbers, P_{Dk} is the load at bus k, Ω_{138} and Ω_{230} are bus sets for 138 KV and 230 KV networks respectively.

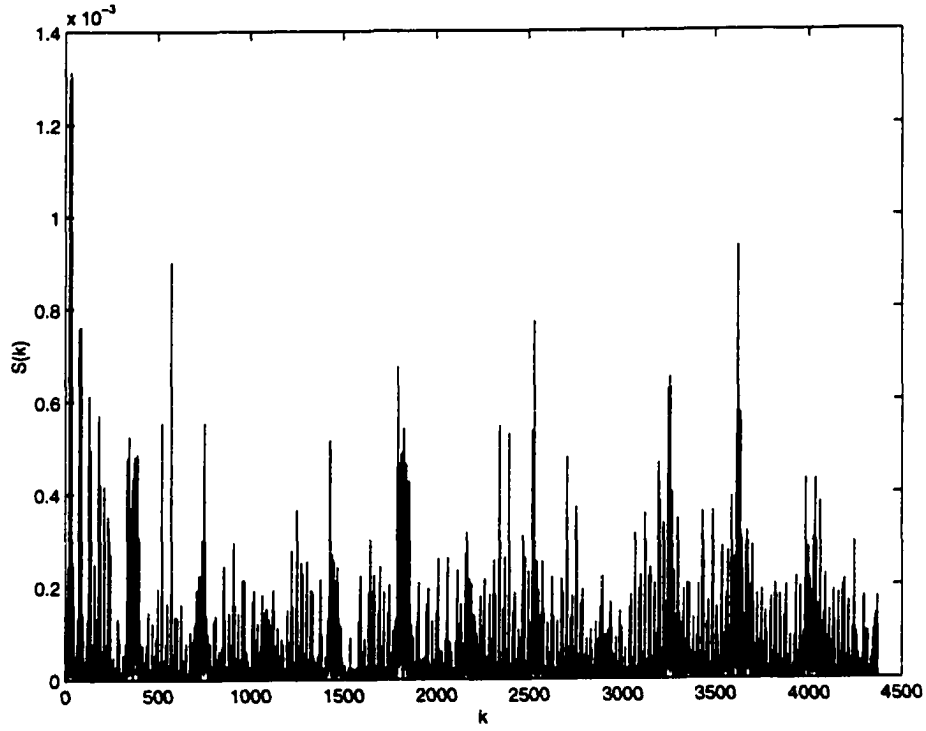

Figure B.7 The estimated spectrum of the error series

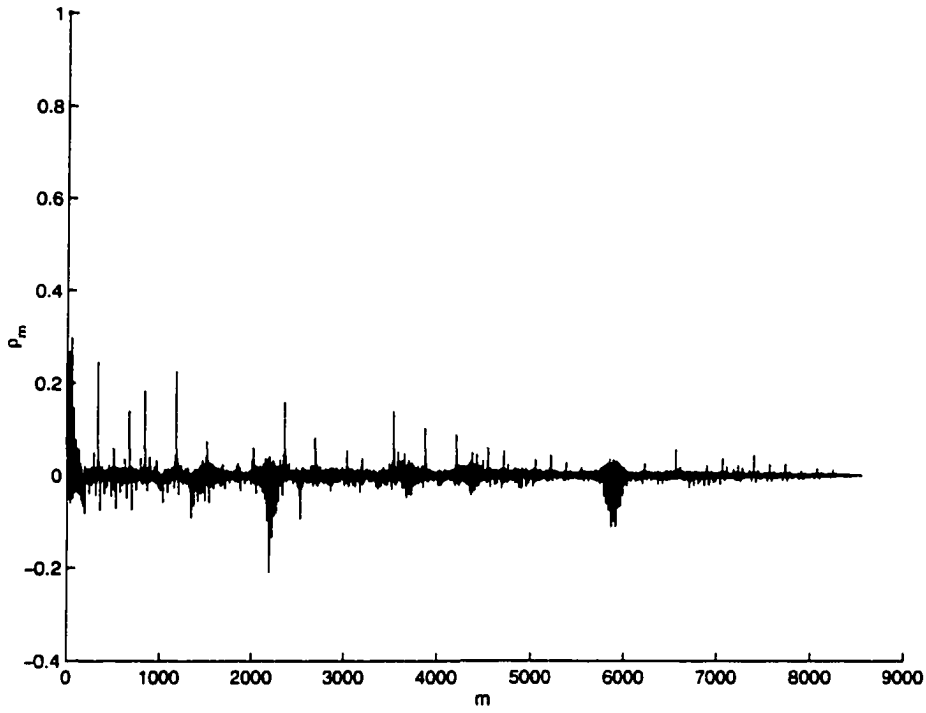

Figure B.8 The ACF of $e_{\tilde{y}}(t)$

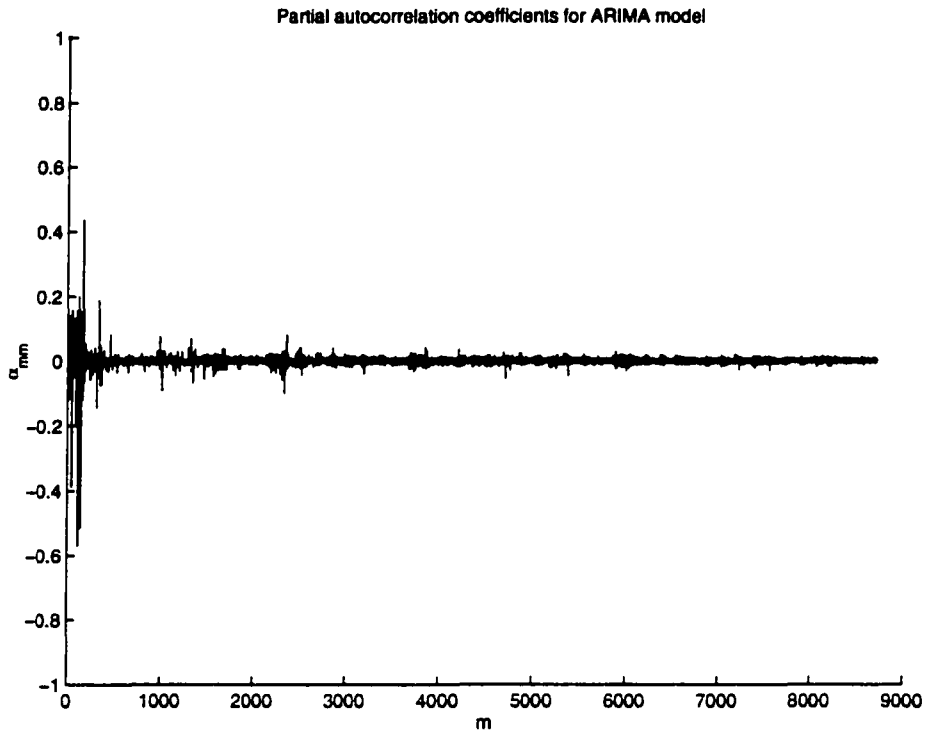

Figure B.9 The PACF of $\tilde{y}_t$

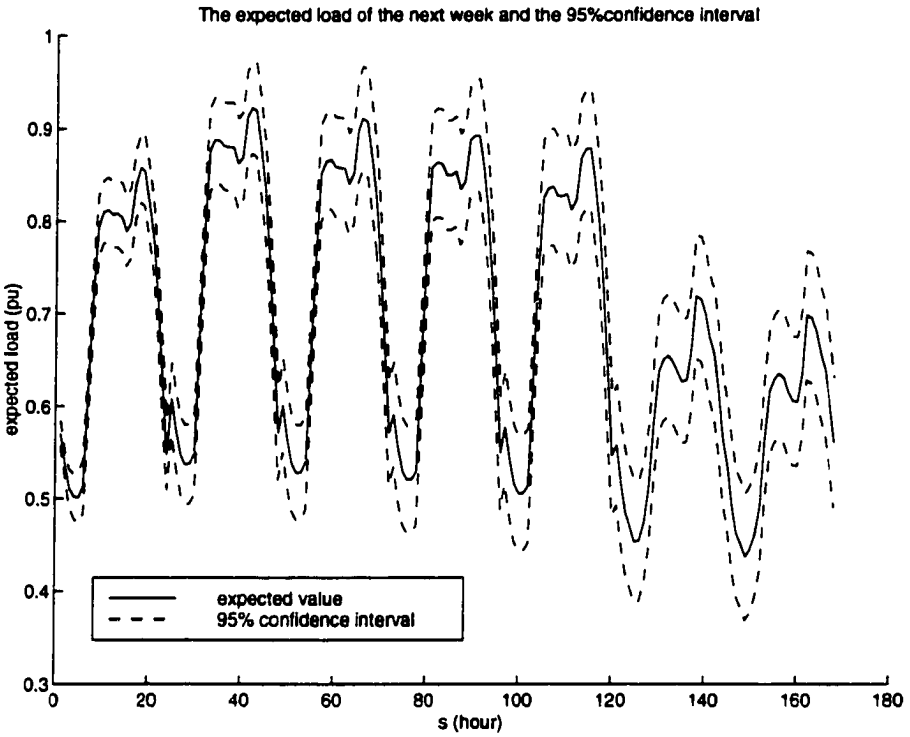

Figure B.10 The forecast load and 95% confidence interval for the next week

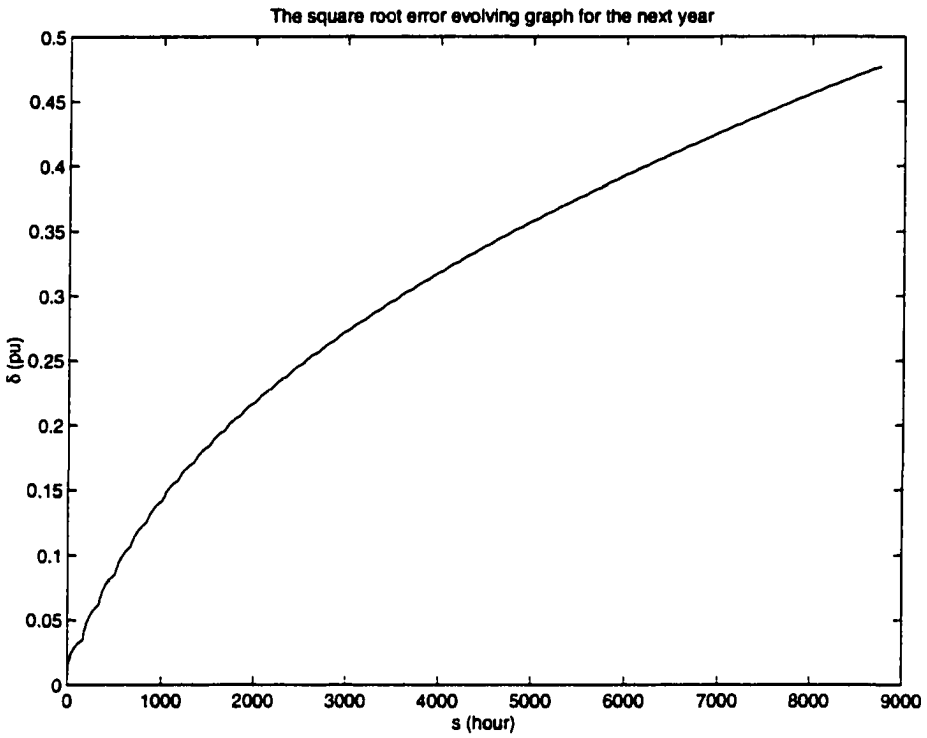

Figure B.11 The $\delta - s$ curve

APPENDIX C DETAILS FOR MAINTENANCE SCHEDULING

Principles of Our Method

Equal LOLP Method

Suppose originally the total capacity of the system is C_t. Due to generator outages, in order to keep the system LOLP no more than a certain level p (e.g., $p = 0.01$), the maximum load the system can provide should be less than the total capacity C_t. We denote the maximum load level A. Or we can say, under the LOLP level p, the system available capacity is A. Now consider increasing a generator with capacity C. In order to keep the same LOLP level, not all the capacity C can be used to increase the available capacity A. Therefore C is divided into two parts.

$$C = C_r + C_{ef} \tag{C.1}$$

where C_r is used to keep system risk constant, while C_{ef} is effective capacity increment, called effective load carrying capacity. Garver [81] derived a formula to calculate the effective load carrying capacity C_{ef} for a generation unit, as follows.

$$C_{ef} = C - M \log[(1 - r) + re^{C/M}] \tag{C.2}$$

where r is the forced outage rate of the generation unit, and M is the system's characteristic slope. Consider Figure C.1. It is drawn by our software for the IEEE RTS'96. The horizontal axis is for the logarithm of LOLP (loss of load probability). The curve shows

when the reserve capacity increases, the LOLP decreases. In one year, the system load varies between points A and B, as shown in the figure. Although the curve is apparently not a line, the tangent slope will not change much within the range between A and B. Therefore, we can use the following formula to calculate M.

$$M = \frac{R_B - R_A}{\log(P_A - P_B)} \tag{C.3}$$

where R refers to reserve and P refers to probability.

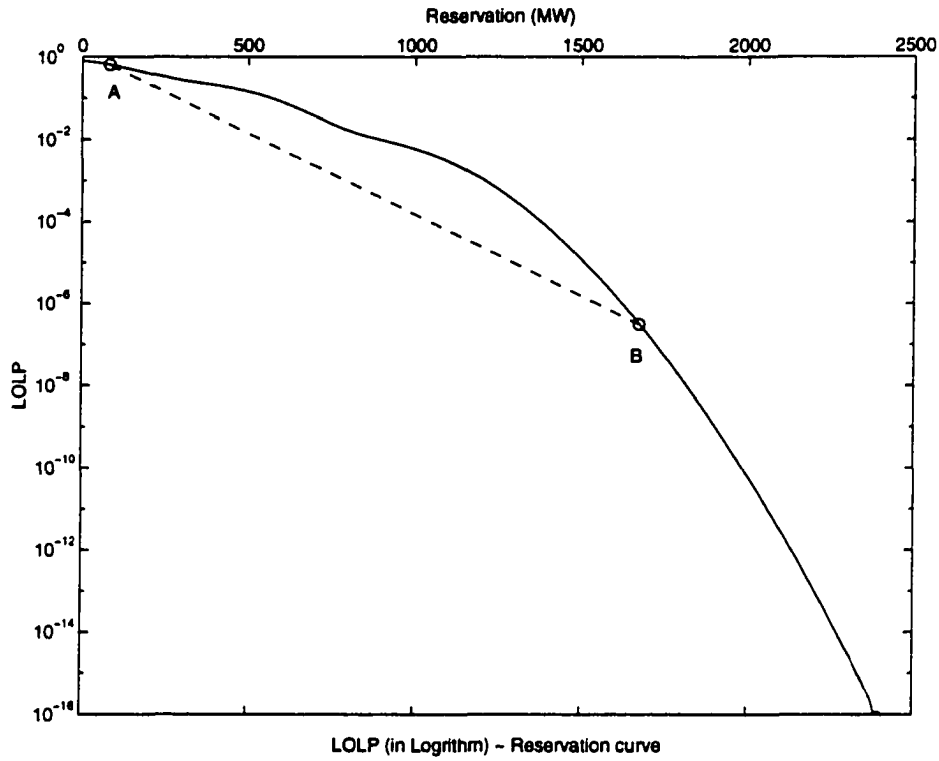

Figure C.1 Log LOLP vs. Reserve Capacity curve

Once we know M, we can use equation C.2 to calculate the effective load carrying capacity for every generation unit. Arranging a unit to maintenance is equivalent to adding the effective load carrying capacity on the load to get equivalent load curve. To make the LOLP curve level is tantamont to levelizing the equivalent load curve. Here we use a method called "minimum cumulative energy method" [45]. The steps of this method are as follows.

1. Reorder C_{ef} of generators from the largest to the smallest.
2. Pick up one C_{ef} from the list in order, look up for the required number of maintenance days (denoted by m).
3. Calculate all summations of the m successive load, find the smallest one, arrange unit there.
4. If this is the last unit, stop; else move to the next unit, go to 2.

There are several ways to calculate the LOLP vs. reserve capacity curve. We have used the cumulant method, but other methods should be considered as well.

Calculating LOLP by Cumulant Method

A linear system can be expressed as $Y = AX + B$, where

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

If $X_1, \cdots, X_n$ are independent random variables, and their distributions are known, then the distributions of $Y_1, \cdots, Y_m$ can be easily obtained by Fourier transformation method or cumulant method. Let us introduce several concepts first.

Characteristic function: For a random variable X $f(x)$, $\phi(t) = Ee^{jtX} = \int_{-\infty}^{+\infty} e^{jtX} f(x)dx$ is defined as the characteristic function (often abbreviated to c.f.) of X.

From the equation, we see that $\phi(t)$ can be regarded as the Fourier inverse transformation of $f(x)$. So we have

$$\phi(t) = \int_{-\infty}^{+\infty} e^{jtx} f(x)dx \tag{C.4}$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-jtx} \phi(t)dt \tag{C.5}$$

Another concept is

Moment: EX^r is defined as the moment of order r, $r = 1, 2, \cdots$. $\mu'_r = E(X-a)^r$ is defined as the moment of order r about point a. $\mu_r = E(X-EX)^r$ is defined as the moment of the mean, or the central moment of order r.

$EX^r, \mu'_r and \mu_r$ have the following relations.

$$EX^r = \mu'_r|_{a=0} \tag{C.6}$$

$$\mu'_r = \sum_{j=0}^{r} \begin{pmatrix} r \\ j \end{pmatrix} \mu_{r-j} \mu_1'^{\,j} \tag{C.7}$$

$$\mu_r = \sum_{j=0}^{r} \begin{pmatrix} r \\ j \end{pmatrix} \mu'_{r-j} (-\mu'_1)^j \tag{C.8}$$

Therefore, if EX^r is known, μ_r, μ'_r can be obtained with ease. On the other hand,

$$\frac{d^r \phi(t)}{dt^r}\bigg|_{t=0} = E[(jX)^r e^{jtX}]|_{t=0} = j^r E(X^r e^{jtX})|_{t=0} = j^r EX^r \Rightarrow$$

$$EX^r = \frac{1}{j^r} \frac{d^r \phi(t)}{dt^r} \tag{C.9}$$

Expression C.9 means if $\phi(t)$ is known, EX^r and its corresponding μ_r can be easily obained. Specifically the mean $\mu_1 = EX$ and the variance $\mu_2 = EX^2 - (EX)^2$ can be easily obtained.

Now consider about $Y_i = \sum_{j=1}^{N} a_{ij} X_j + b_i$, $i = 1, 2, \cdots, m$. It can be derived that

$$\phi_{Y_i}(t) = E e^{jt(\sum_{j=1}^{n} a_{ij} X_j + b_i)} = e^{jb_i t} \prod_{j=1}^{n} \phi_{X_j}(a_{ij} t) \tag{C.10}$$

Then $f_{Y_i}(y)$ can be obtained by Fourier transformation. The steps of this method are as follows.

1. Use inverse Fourier transformation, calculate $\phi_{X_j}(t)$ from $f_{X_j}(x)$, $j = 1, \cdots, n$. Here IFFT algorithm can be adopted [78].

2. Calculate $\phi_{X_j}(a_{ij}t)$ from $\phi_{X_j}(t)$.

3. Use expression C.10 to calculate $\phi_{Y_i}(t)$, $i = 1, \cdots, m$.

4. Use Fourier transformation to calculate $f_{Y_i}(y)$. Here FFT algorithm can be adopted.

However, the calculation amount is still large. So the cumulant method is developed by some researchers to calculate risk. The concept of cumulant κ_r is defined as follows.

Cumulant: If $\phi(t) = Ee^{jtX}$ is the characteristic function of X, then

$$L(t) \triangleq Log(Ee^{jtX}) \triangleq \kappa_1 \cdots (jt) + \kappa_2 \cdots \frac{(jt)^2}{2} + \cdots + \kappa_r \cdot \frac{(jt)^r}{r!} + \cdots \qquad \text{(C.11)}$$

where κ_r is defined as the cumulant of order r, $r = 1, 2, \cdots$.

κ_r and μ_r have the following relations.

$$\begin{aligned}
\mu_2 &= \kappa_2 \\
\mu_3 &= \kappa_3 \\
\mu_4 &= \kappa_4 + 3\kappa 2^2 \\
\mu_5 &= \kappa_5 + 10\kappa_3\kappa_2 \\
\mu_6 &= \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3 \\
\mu_7 &= \kappa_7 + 21\kappa_5\kappa_2 + 35\kappa_4\kappa_3 + 105\kappa_3\kappa_2^2 \\
\mu_8 &= \kappa_8 + 28\kappa_6\kappa_2 + 56\kappa_5\kappa_3 + 35\kappa_4^2 + 210\kappa_4\kappa_2^2 + 280\kappa_3^2\kappa_2 + 105\kappa_4^4 \qquad \text{(C.12)}
\end{aligned}$$

and

$$\begin{aligned}
\kappa_2 &= \mu_2 \\
\kappa_3 &= \mu_3 \\
\kappa_4 &= \mu_4 - 3\mu_2^2 \\
\kappa_5 &= \mu_5 - 10\mu_3\mu_2 \\
\kappa_6 &= \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3 \\
\kappa_7 &= \mu_7 - 21\mu_5\mu_2 - 35\mu_4\mu_3 + 210\mu_3\mu_2^2 \\
\kappa_8 &= \mu_8 - 28\mu_6\mu_2 - 56\mu_5\mu_3 - 35\mu_4^2 + 420\mu_4\mu_2^2 + 560\mu_3^2\mu_2 - 630\mu_2^4
\end{aligned} \tag{C.13}$$

From the definition of κ_r we can easily get the following expression

$$\kappa_r = \left. \frac{1}{j^r} \frac{d^r L(t)}{dt^r} \right|_{t=0}.$$

Therefore, we have the following formula

$$\kappa_{r(aX)} = \frac{1}{j^r} \frac{d^r LogEe^{jtaX}}{dt^r} = \frac{1}{j^r} \frac{d^r L(at)}{d(at)^r}|_{at=0} = a^r \kappa_{rX} \tag{C.14}$$

$$\kappa_{r(X_1+X_2)} = \frac{1}{j^r} \frac{d^r LogEe^{jt(X_1+X_2)}}{dt^r}|_{t=0} = \frac{1}{j^r} \frac{d^r LogEe^{jtX_1}}{dt^r}|_{t=0} + \frac{1}{j^r} \frac{d^r LogEe^{jtX_2}}{dt^r}|_{t=0} = \kappa_{rX_1} + \kappa_{rX_2} \tag{C.15}$$

$$\kappa_{rY_i} = \begin{cases} \sum_{j=1}^{n} a_{ij}\kappa_{rX_j} + b_i, & r = 1 \\ \sum_{j=1} n a_{ij}\kappa_{rX_j}, & r > 1 \end{cases} \tag{C.16}$$

Once κ_{rY_i} are obtained, the distribution of Y_i can be approximated by Gram-Charlier expansion

$$\begin{aligned}
\bar{F}(y) = P(Y \geq y) = P(\frac{Y-\mu}{\sigma} \geq \bar{y}) = \int_{\bar{y}}^{+\infty} N(y)dy + N(\bar{y})[\frac{g_3}{3!}H_2(\bar{y}) + \frac{g_4}{4!}H_3(\bar{y}) + \frac{g_5}{5!}H_4(\bar{y}) \\
+ \frac{g_6 + 10g_3^2}{6!}H_5(\bar{y}) + \frac{g_7 + 35g_3g_4}{7!}H_6(\bar{y}) + \frac{g_8 + 56g_3g_5 + 35g_4^2}{8!}H_7(\bar{y}) + \cdots]
\end{aligned} \tag{C.17}$$

or by Edgeworth expansion

$$\bar{F}(y) = P(Y \geq y) = P(\frac{Y-\mu}{\sigma} \geq \bar{y}) = \int_{\bar{y}}^{+\infty} N(y)dy + N(\bar{y})[\frac{g_3}{3!}H_2(\bar{y}) + \frac{g_4}{4!}H_3(\bar{y}) + \frac{10g_3^2}{6!}H_5(\bar{y}) + \frac{g_5}{5!}H_4(\bar{y}) + \frac{35g_3g_4}{7!}H_6(\bar{y}) + \frac{280g_3^3}{9!}H_8(\bar{y}) + \frac{g_6}{6!}H_5(\bar{y}) + \frac{56g_3g_5}{8!}H_7(\bar{y}) + \frac{2100g_3^2g_4}{10!}H_9(\bar{y}) + \frac{15400g_3^4}{12!}H_{11}(\bar{y}) + \cdots] \tag{C.18}$$

where

$$g_r = \frac{\kappa_r}{\sigma^r} = \frac{\kappa_r}{\kappa_2^{\frac{r}{2}}},\ r = 2, 3, \cdots. \tag{C.19}$$

is the normalized cumulant of the order r;

$$N(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} \tag{C.20}$$

is the probability density function of the standard normal distribution.

$H_r(y)$ is the Hermite polynomial of order r. The first ten polynomials are

$$\begin{aligned}
H_0(x) &= 1 \\
H_1(x) &= x \\
H_2(x) &= x^2 - 1 \\
H_3(x) &= x^3 - 3x \\
H_4(x) &= x^4 - 6x^2 + 3 \\
H_5(x) &= x^5 - 10x^3 + 15x \\
H_6(x) &= x^6 - 15x^4 + 45x^2 - 15 \\
H_7(x) &= x^7 - 21x^5 + 105x^3 - 105x \\
H_8(x) &= x^8 - 28x^6 + 210x^4 - 420x^2 + 105 \\
H_9(x) &= x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x \\
H_{10}(x) &= x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945
\end{aligned} \tag{C.21}$$

$$\bar{y} = \frac{y - \mu}{\sigma} \tag{C.22}$$

is the normalized given value.

The total calculation steps are as follows

1. Calculate κ_{rX_j}, $j = 1, 2, \cdots, n$, $r = 2, 3, \cdots$.
2. $i = 1$.
3. Use expression C.16 to calculate κ_{rY_i}.
4. USe expression C.19 to calculate g_{rY_i}, use expression C.22 to get a normalized $\bar{y}_i$ for a given y_i.
5. Use expression C.17 or expression C.18 to get the probability $barF_{Y_i}(y) = P(Y_i \geq y_i)$.
6. If $i = m$ then stop; else $i = i + 1$, go to 3.

However, when we use this method, we have to pay attention to the following points:

1. The convergence of the two series expansions can only be guaranteed when $f(x)$ has a derivative function that is continuous at every point, and some other conditions are met. This is not true in LOLP calculation, where the distribution function for the generation is obviously a discrete distribution functions [82]. However, when there are many generators, the distribution is approximate to continuous distribution.
2. The method is originally used for adjustments when a distribution is near normal distribution. From our experience, when the number of generators is large, the generation's distribution is near normal distribution. So we can use this method

in LOLP calculation. However, this is not always true in probabilistic power flow convolution.

3. We can always use pu value in X_i and Y in calculation. Appropriate choice of calculation base can reduce the computer's interception error. In theory, no matter how large you choose the base, the magnitude of κ_r may be different, but the magnitude of g_r remains the same.

Calculation Results for the IEEE RTS'96

The daily peak load curve of the IEEE RTS'96 is shown in Figure C.2. Every unit can be regarded as a two-state component, with the probability $1-\rho_i$ to be at the normal state, and ρ_i to be at the forced outage state. Therefore, the generation of one unit is a discrete random variable C_i, the moment items μ_{ir} of it can be calculated by the following formula

$$\begin{aligned} \mu_{i1} &= (1-\rho_i)C_i \\ \mu_{ir} &= (1-\rho_i)(C_i-\mu_{i1})^r + \rho_i(-\mu_{i1})^r,\ r=2,3,\cdots. \end{aligned} \tag{C.23}$$

Then we can calculate κ_{ir} and the total generation's κ_r by using C.13, if we assume all generators are independent with each other. After that, we use Gram-Charlier expansion to calculate the LOLP curve, as shown in Figure C.3. We can see that when the load is near the generation capacity, a little increment in load will cause a dramatic increment in LOLP. Figure C.1 shows the relationship between the LOLP and the reserve capacity. It can be used to calculate M, as we mentioned before. For this system, M=109.3 MW. Table C.1 shows the load carrying capacity C_{ef} of each type of unit. We can see that C_{ef} tends to be larger when the unit capacity is larger, and smaller when the forced outage rate is larger.

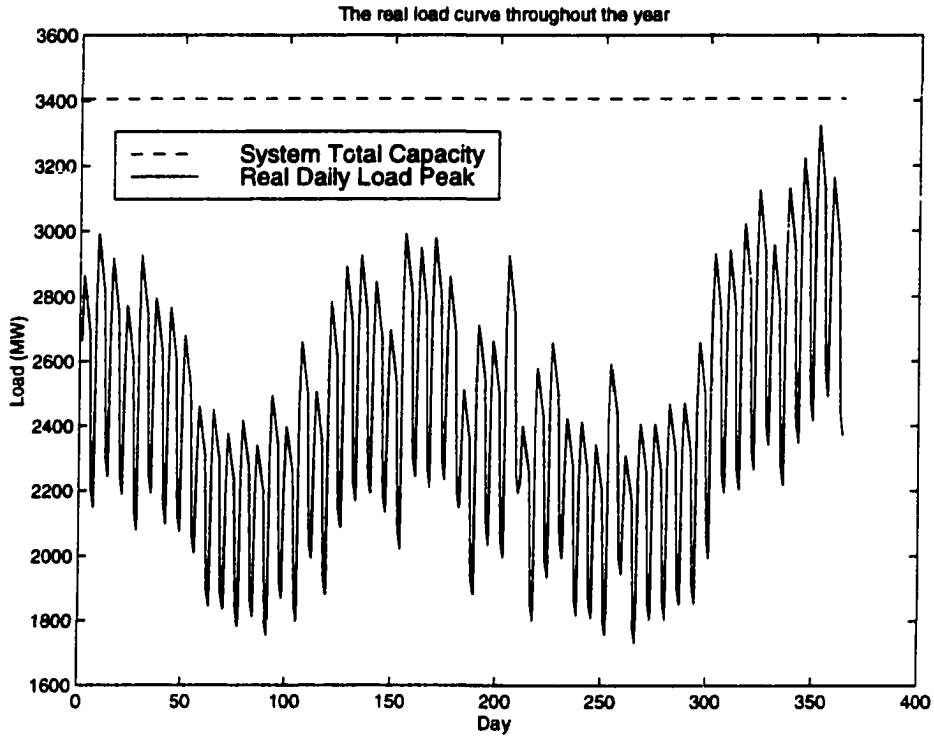

Figure C.2 The daily peak load curve

The equivalent daily load curve after maintenance schedule arrangement is shown in Figure C.4. We can see that it is much smoother than Figure C.2. It proves that our technique to levelize the system LOLP is effective.

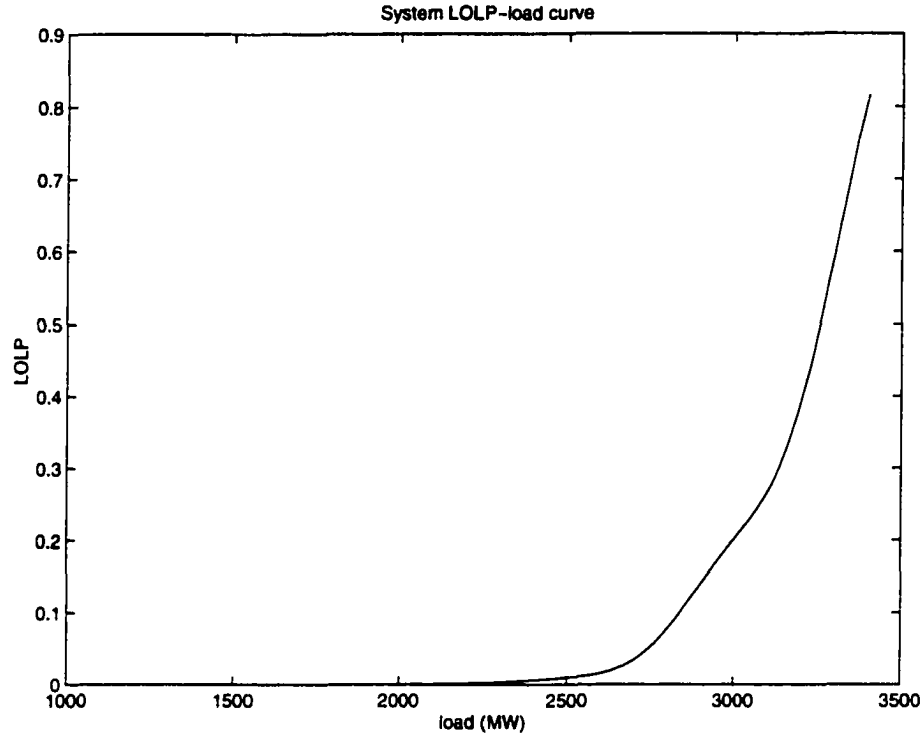

Figure C.3 The LOLP vs. Load curve

Table C.1 The effective load carrying capacities of different unit types

Unit type	Description	Capacity (MW)	FOR	C_{ef}(MW)
1	Fossil Steam	12	0.02	11.75
2	Combustion Turbine	20	0.10	17.83
3	Fossil Steam	76	0.02	73.83
4	Fossil Steam	100	0.04	93.65
5	Fossil Steam	155	0.04	142.11
6	Fossil Steam	197	0.05	172.33
7	Fossil Steam	350	0.08	234.11
8	Nuclear	400	0.12	212.82
9	Hydro	50	0.01	49.37

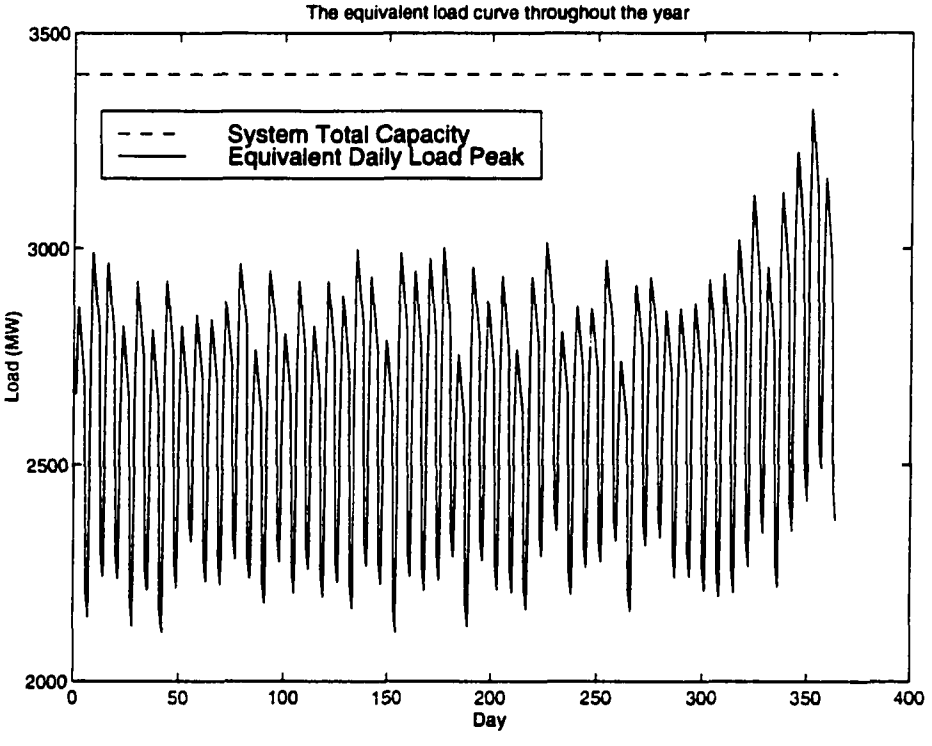

Figure C.4 The daily equivalent load curve after maintenance arrangement

APPENDIX D DETAILS FOR UNIT COMMITMENT ARRANGEMENT

Problem Discription in Mathematical Form

The unit commitment problem can be described as the following mathematical optimization problem.

$$\min \sum_{i=1}^{n_t} \sum_{t=1}^{N} (F(P_{ti}^t, U_i^t) + S_i^t) \tag{D.1}$$

subject to

$$\sum_{i=1}^{n_t} P_{ti}^t + \sum_{i=1}^{n_h} P_{hi}^t + \sum_{i=1}^{n_n} P_{ni}^t = P_D^t,\ t = 1, 2, \cdots, N \tag{D.2}$$

$$\sum_{t=1}^{0.25N} \sum_{i=1}^{n_h} P_{hi}^t = E_1 \tag{D.3}$$

$$\sum_{t=0.25N+1}^{0.5N} \sum_{i=1}^{n_h} P_{hi}^t = E_2 \tag{D.4}$$

$$\sum_{t=0.5N+1}^{0.75N} \sum_{i=1}^{n_h} P_{hi}^t = E_3 \tag{D.5}$$

$$\sum_{t=0.75N+1}^{N} \sum_{i=1}^{n_h} P_{hi}^t = E_4 \tag{D.6}$$

$$P_{hi}^t \le C_{hs},\ t = 0.25(s-1)N + 1, \cdots, 0.25sN,\ s = 1, \cdots, 4. \tag{D.7}$$

and a minimum gap constraint, which can be described as follows:

if there exist t_1, t_2 such that $U_i^{t_1} = U_i^{t_2} = 3$ and for all $t_1 < t < t_2$, $U_i^t \neq 3$, then if $t_2 - t_1 < min\_gap_i$, $U_i^t = 2$ for all such t, $i = 1, \cdots, n_t$.

In this model, the symbols are explained as follows.

- P_{ti}^t is the generation of thermal unit i at time t.
- U_i^t is the unit state of thermal unit i at time t. There are four states:

  0 denotes the unit is in maintanance.

  1 denotes the unit is in cold reserve.

  2 denotes the unit is in hot reserve.

  3 denotes the unit is in operation.

- P_{hi}^t denotes the generation of hydro unit i at time t.
- P_{ni}^t denotes the generation of nuclear unit i at time t.
- P_D^t denotes the total load at time t. However, it is a load used for calculation purposes. We choose it to be the actual load multiply 1.06 in order to accommondate transmission loss.
- S_i^t denotes the start up cost of unit i at time t.
- $F(\cdot,\cdot)$ is the operational cost for thermal unit i at time t, it is a function of P_{ti}^t and U_i^t.
- n_t, n_h, n_n are the numbers of thermal units, hydro units and nuclear units at time t that is not in maintanance respectively.
- N is the number of hours for one calculation year. Here only 52 weeks of data are available, therefore $N = 8736$.
- E_1, E_2, E_3, E_4 are total availabe hydro energy of spring, summer, autumn and winter respectively.

- C_{hs} is the available capacity of season s, $s = 1$ means spring, $s = 2$ means summer, $s = 3$ means autumn, $s = 4$ means winter.

- min_gap_i is the minimum duration for unit i in cold reserve between two non-consecutive operation states. It is the sumation of minimum down time and the minimum up time.

The first constraint reveals the power balance. The second through fifth represent hydro power energy constraints. The sixth expresses hydro unit capacity constraint. The last constraint reveals minimum cold reserve hours between two nonconsecutive operation states. If a gap is shorter than the minimum gap, we should keep the unit in hot reserve. However, the minimum shut down time and start up time between an operation state and a hot reserve state are ignored, and the generation variations during the transition process are also ignored in order to simplify our calculation. Furthermore, we assume there is no cost for using hydro power, and nuclear units are always in full capacity as long as they are not in maintenance. These assumptions are reasonable as hydro power is always very cheap compared with thermal units, and most companies are reluctant to adjust nuclear units.

Review of Lagrange Relaxation Method

The Lagrange relaxation method constructs the Lagrange function first. The Lagrange function for our problem can be expressed as follows.

$$\begin{aligned}
& L(P_{ti}^t, U_i^t, \lambda^t, \mu^s) \\
= \; & \textstyle\sum_{i=1}^{n_t} \sum_{t=1}^{N} (F(P_{ti}^t, U_i^t) + S_i^t) \\
+ \; & \textstyle\sum_{t=1}^{N} \lambda^t (P_D^t - \sum_{i=1}^{n_t} P_{ti}^t - \sum_{i=1}^{n_h} P_{hi}^t - \sum_{i=1}^{n_n} P_{ni}^t) \\
+ \; & \textstyle\sum_{s=1}^{4} \mu^s (E_s - \sum_{t=0.25(s-1)N+1}^{0.25sN} P_{hi}^t) \\
= \; & \textstyle\sum_{i=1}^{n_t} \sum_{t=1}^{N} (F(P_{ti}^t, U_i^t) + S_i^t - \lambda^t P_{ti}^t) \\
+ \; & \textstyle\sum_{i=1}^{n_t} \sum_{t=1}^{N} \lambda^t (P_D^t - \sum_{i=1}^{n_h} P_{hi}^t - \sum_{i=1}^{n_n} P_{ni}^t) \\
+ \; & \textstyle\sum_{s=1}^{4} \mu^s (E_s - \sum_{t=0.25(s-1)N+1}^{0.25sN} P_{hi}^t)
\end{aligned}$$

We desire to minimize the Lagrange function under the constraints. When the power balance constraints and the hydro energy constraints are met, the added items in the Lagrange function will not take effect. If we always choose λ^t and μ^s such that

$$\begin{aligned}
\lambda^t (P_D^t - \sum_{i=1}^{n_t} P_{ti}^t - \sum_{i=1}^{n_h} P_{hi}^t - \sum_{i=1}^{n_n} P_{ni}^t) \geq 0 \\
\mu^s (E_s - \sum_{t=0.25(s-1)N+1}^{0.25sN} P_{hi}^t) \geq 0
\end{aligned} \tag{D.8}$$

the added items become penalty items. Therefore, minimizing L with respect to P_{ti}^t, U_{ti}^t,P_{hi}^t first while fixing λ^t, μ^s, then maximizing L with respect to λ^t, μ^s while fixing P_{ti}^t, U_{ti}^t,P_{hi}^t will drive us to the optimized solution. This is the basic idea of Lagrange relaxation method. We can also explain the method in an economic way. Suppose at every hour there is a power market. λ^t is the power energy marginal price. If λ^t is low, sellers are unwilling to produce power, the generation of commited units is not enough to compensate load, then we should raise the price λ^t. On the other hand, if λ^t is high, sellers are willing to produce power, the generation of commited units is more than load, then we should reduce the price. The same principle is also applied to hydro energy

constraints. μ^s is just the price of hydro power in season s. A buyer will buy as much hydro power as possible whenever the hydro power price μ^s is cheaper than the thermal power price λ^t; and will not buy any hydro power whenever $\mu^s > \lambda^t$; when the two prices are equal, any possible division between thermal power and hydro power is acceptable. Therefore, when μ^s is small, hydro energy will be used up too quickly, then we should increase μ^s; when μ^s is large, extra water will be left, then we should decrease μ^s. The point that we coincidentally use up all water just after we arrange the last hour's unit commitment corresponds to the solution. Once we set all the prices, every unit can use them to determine unit commitment to minimize cost by dynamic programming, and the buyer can also determine how much hydro power can be purchased. The thermal unit commitment calculation can be fulfilled independently for each unit by dynamic programming. This kind of decomposition is also shown in the second expression in equation (8). Once the individual optimization results meet all constraints, we get the global optimization solution.

In reality, we are not always lucky to get a converged solution. Because unit commitment is a typical nonlinear and discontinuous problem, there is normally no guarantee for convergence. In order to understand this, we can define the primal and dual problems as follows.

Primal problem When we fix the unit commitment pattern and search for the economic dispatch solution, we are solving the primal problem by minimizing Lagrange function. The minimization result can be denoted by J. When the generation capacity is not enough to provide load, we apply a large penalty.

Dual problem When we fix λ^t and μ^s and search for unit commitment pattern decomposedly by dynamic programming, we get the value of the dual problem. It can be denoted by q.

From the definition the primal solution J is always greater than the dual solution

q. We can define the duality gap as $\frac{J-q}{q}$. As we mentioned, there is no guarantee for the duality gap to be converged to zero. As long as the duality gap is small enough, we should stop and make final adjustments to make the solution feasible.

There are several points we should specify.

- We can start from a primal feasible solution, or a dual feasible solution. The case that all prices are zeros is a convenient dual feasible solution. The case that economic dispatch assuming all units are on is calculated for each hour, with zero output units canceled from the unit commitment is a convenient primal feasible solution. We think that starting from primal feasible solution is better, because at least we have a feasible solution.

- Dynamic programming is very time consuming when one year is considered. However, Lagrange relaxation method has to carry on a lot of dynamic programming calculations. Therefore, it is unrealistic to use Lagrange relaxation method to calculate unit commitment for one year.

- There is no guarantee that when the dual solution is stopped, it will be a feasible solution. Further adjustment may lose optimization. Therefore, even if we use Lagrange relaxation method, we can not guarantee to get the best solution.

- The best solution, although an economic one, may be a stressed case when the transmission system is considered. If we account for security risk, it may not be the best.

- When the number of generators increases, we have more choices to maneuver, therefore, the duality gap can be smaller. When the time length increases, we have to meet more constraints, and the duality gap tends to be larger.

Above all, the Lagrange relaxation method is not an appropriate method for one year unit commitment calculation because it is too time consuming. However, some basic

ideas such as marginal price and duality gap can be borrowed to calculate and check our priority list method.

Piecewise Linear Model for a Thermal Unit

For a thermal unit, fuel consumption cost curve is normally represented by a quadratic or a cubic curve. It is accurate but not convenient. Especially under the deregulated power market, it is inconvenient for a generation agent to tell a customer " my electricity price is a quadratic curve with coefficients a,b,c", instead, he will say "my electricity price is \$30 per MWh if you buy 5 MW or less, and \$40 per MWh if you buy 5 to 10 MW". Therefore, the piecewise linear model is more suitable. We will show that the piecewise linear model is also more convenient for calculation.

Consider Figure D.1. First we ignore the start up cost, and only consider about the operation cost. When we calculate economic dispatch, we always have a system incremental cost, which is the current price for the load increment. If a unit has marginal cost below this incremental cost, the unit should increase its power; otherwise, if a unit has marginal cost beyond the incremental cost, the unit should decrease its power. Now consider the incremental cost as it increases from 0 to $+\infty$. We denote the system incremental cost λ'. When $\lambda' <$ AB's price, the optimal output for the unit is 0. When λ'=AB's price, the optimal output can be any number between 0 and P_B. When AB's price$< \lambda' <$BC's price, the optimal output should be kept at P_B. When $\lambda'=$ BC's price, the optimal output can be any value between P_B and P_C, and so on. Therefore, if we only choose λ' to be among prices of all piecewise linear segments, we can still obtain all possible output power. If we have 4 linear segments for each unit and 24 units in operation, we only have to investigate 96 possible λ' values. Compared with the infinite number of possible λ' values when smooth curves are employed, the piecewise linear model has great advantage, and it can also give an optimal range instead of just one

point. This is very important in hydro-thermal coordination. It makes the calculation more robust to the accuracy of various shadow prices.

Now consider whether we should turn a unit on or off. We still ignore startup cost. We know λ^t is the market price at time t, and we tune it to optimize unit commitment. From equation D.1 if for a unit $F(P_{ti}^t, U_i^t) - \lambda^t P_{ti}^t \leq 0$, the unit should be on at time t, otherwise it should be off. Now consider λ^t increases from 0 to $+\infty$ in Figure D.1. When $\lambda^t <$ slope of OD, the unit should be off. When slope of OD$< \lambda^t <$slope of DE, the unit should be operated at state D. When $\lambda^t =$slope of DE, the unit can be operated at any point between D and E. When $\lambda^t >$slope of DE, the unit should be operated at state E. There is also another case, as shown in Figure D.2. In Figure D.2, when $\lambda^t \leq$slope of OE, the unit should be off; when $\lambda^t >$slope of OE, the unit should be operated at point E. Above all, the possible unit states are very limited and discontinuous in unit commitment calculation, as opposed to the volatility and continuity in economic dispatch calculation. This leads to the problem that the duality gap can not be eliminated in principle. When we ignore start up cost, we can also discretize λ^t's values. For Figure D.1's case, we can add the slope of DE to the price list; for Figure D.2's case, we can add the slope of OE to the price list. If we include the critical prices for all units and order them in sequence, we can choose λ^t only from this set. However, when startup cost is considered, we have to deal with dynamic programming, then the choice of λ^t should not be limited to such a price set. This is also a drawback of accurate Lagrange relaxation method.

In our program, we combine the economic dispatch incremental cost set and λ^t set together, calculate $\min F(P_{ti}, U_i) - \lambda P_{ti}$ for each λ in the set and each unit, and the optimal generation range $[P_{imin}, P_{imax}]$. Then we record all these data in a table (We call it Fee Record Table). Since the same type of units have the same $\min F(P_{ti}, U_i) - \lambda P_{ti}$ and $[P_{imin}, P_{imax}]$ values, we only have to assign one record line for each unit type instead of each unit. When we choose or adjust λ, we even do not have to know the

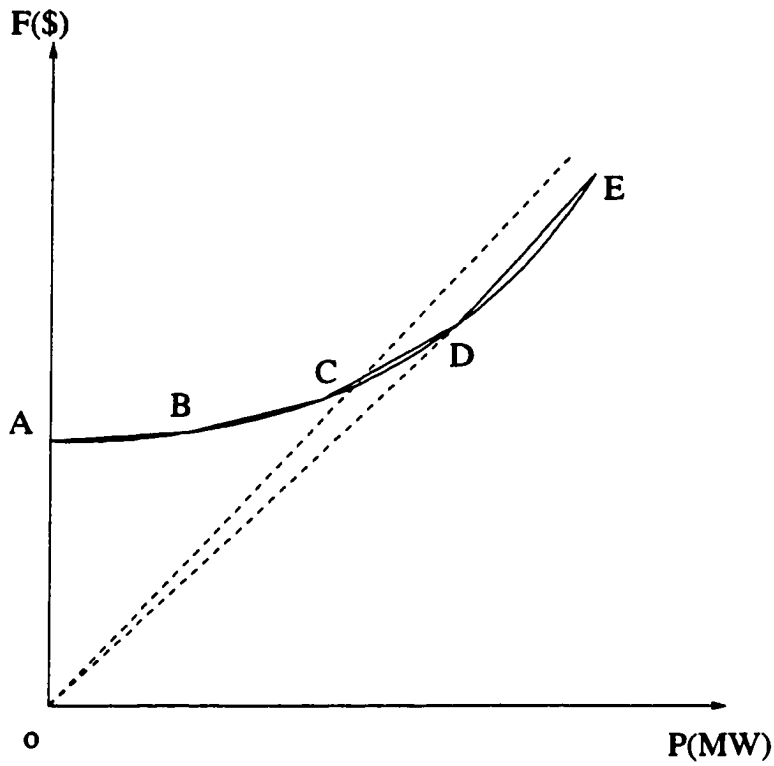

Figure D.1 The fuel consumption curve of a thermal unit-case I

real value of λ, as long as we know its index in the list. Then we can use the table to know if the generator should be on or off, what is the economic generation range if the system incremental cost is given. This technique saves a large amount of calculation time. However, the accurate Lagrange relaxation method will inevitably lose such an advantage.

Our Heuristic Method

The steps of our heuristic method is shown in pseudo code as follows.

1. Form the λ list and the Fee Record Table.
2. For s=1:4
   (1) pick up one μ^s from the λ list.
   (2) for $t = (s-1)N/4 + 1 : sN/4$

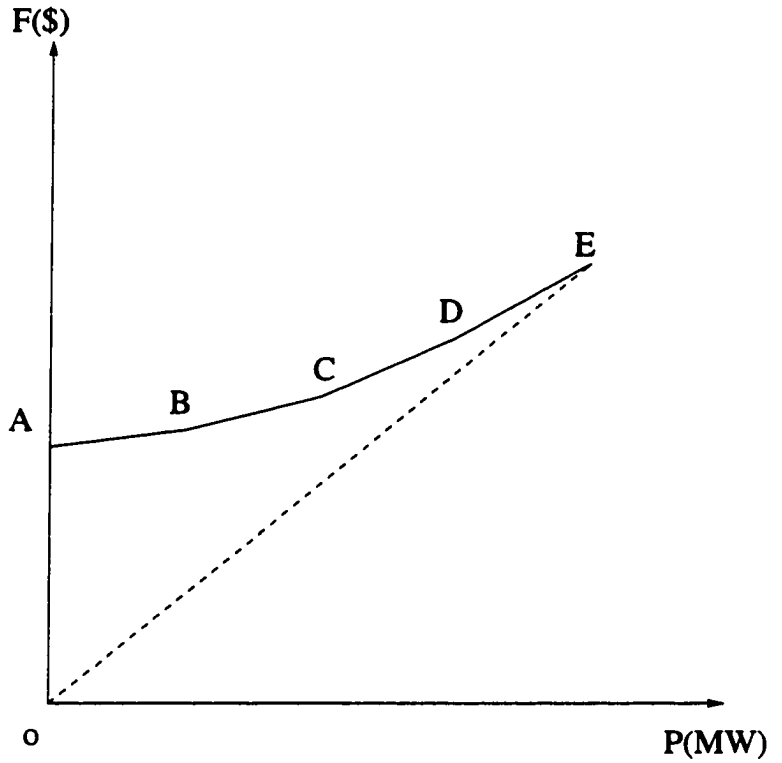

Figure D.2 The fuel consumption curve of a thermal unit–case II

a. pick up one λ^t from the λ list.

b. sum the total generation range including the hydro power.

c. if the range cover the load, record the hydro power range, go to the next t; else change λ^t, go to b.

(3) if the hydro energy range cover the given energy E_s, our unit commitment for season s is ok, otherwise adjust the hydro power price μ^s, go to (2).

3. Check for the minimum gap constraint, change unit state from 1 (cold reserve) to 2 (hot reserve) whenever there is a violation.

4. For the obtained unit commitment pattern, apply economic dispatch. Hot reserve units also participate in economic dispatch.

5. If a thermal unit has power output, set it to be at state 3; if it has 0 output but it is at state 3, set it at 2.

6. If there are more consecutive 2s than minimum gap, set them to 1 (cold reserve). This is because in the IEEE RTS'96 system, when the duration is longer than minimum gap, stay in state 1 (cold reserve) is always more economic than stay in state 2 (hot reserve).
7. Arrange the power proportionally among generators.
8. Check for P_{gmin} constraints and make appropriate adjustments.

Steps 7 and 8 need to be explained in details. Just before step 7 is run, our program can guarantee the following two things.

- In every hour t, $P^t_{gmin} \leq P^t_D \leq P^t_{gmax}$, where P^t_{gmin} and P^t_{gmax} are the total minimum optimal generation and total maximum generation respectively.
- In every season s, $E^s_{hmin} \leq E_s \leq E^s_{hmax}$, where E^s_{hmin} and E^s_{hmax} are total minimum hydro energy and total maximum hydro energy of season s respectively.

We want to arrange power among units proportionally. That is to say:

1. If for time t, the minimum economic hydro power output is P^t_{hmin} for one hydro unit, and the maximum is P^t_{hmax}, then the actual hydro power output for the unit can be chosen as

$$P^t_h = P^t_{hmin} + \gamma^s_h (P^t_{hmax} - P^t_{hmin}) \tag{D.9}$$

where

$$\gamma^s_h = (E_s - E^s_{hmin})/(E^s_{hmax} - E^s_{hmin}) \tag{D.10}$$

Such a choice can guarantee us to use up all hydro energy, and every operational hydro unit outputs the same power since they are of the same type.

2. After subtracting the hydro power both from P^t_{gmin}, P^t_{gmax} and from P^t_D, the new values should satisfy $P^t_{gmin} \leq P^t_D \leq P^t_{gmax}$. Therefore, if we choose

$$\gamma^t = (P^t_D - P^t_{gmin})/(P^t_{gmax} - P^t_{gmin}) \tag{D.11}$$

we can arrange every unit which is in operation to output power

$$P_{ti}^{t} = P_{timin}^{t} + \gamma^{t}(P_{timax}^{t} - P_{timin}^{t}) \tag{D.12}$$

Such a choice can guarantee us to meet the load demand precisely, and the same type units output the same power if they are in operation.

Step 8 is used when the minimum unit output constraints are taken into account. These constraints are not listed in our previous mathematical model. However, if we choose 5% to be the minimum thermal unit output ratio, which is reasonable, we find that we have several violations in our solution. This can be adjusted in the following way. If the output of a thermal unit is less than 5%, we sum the output of all units with the same type together, if the output is more than 5% of one unit's capacity, we just concentrate those power on one unit and transfer the other units to state 2 or 1 according to minimum gap constraint; if the output is less than 5%, we ignore the power and transfer all such units to state 2 or 1 according to minimum gap constraint. Such an adjustment may lead to guaranteed transmission loss ratio less than 6%, but this is not a serious problem.

Duality Gap Calculation

After these adjustments, we can calculate the total fuel cost including start up cost of the whole system for one year. If we use \$3.0/MBtu, \$1.5/MBtu, and \$0.65/MBtu as the price of oil, coal and nuclear fuel, which are typical, then we obtain the system total fuel cost for one year is $\$\ 2.54 \times 10^{8}$. Now we want to know how good it is. We know this is just the value of the primal problem J. If we fix the current λ^{t} and use the dynamic programming to optimize unit commitment, we are calculating the dual problem's value q. If there is no unit commitment changed, then we have already found the optimal solution. Unfortunately this is not the case for this system. We found $q = 1.78 \times 10^{8}$,

and the duality gap $(J-q)/q$ is 0.43. As far as one year long is considered, the duality gap is quite small. And it can be easy to check that the dual solution is not a feasible solution, some constraints are violated. Therefore q is the lower bound of the optimal solution, it is the ideal goal that we can never attain. The optimal solution should be between \$ $2.54 \sim 1.78 \times 10^8$. If someone is not satisfied with our heuristic result of J, our result is at least a good start for further calculation by Lagrange relaxation method.

In duality calculation, the most difficult and time consuming part is the dynamic programming calculation. How to take the minimum gap constraint into consideration is the key to design the dynamic programming. Here we use a technique we called "state expansion" technique. In our technique, a thermal unit has more than 3 possible states when it is not in maintanance, state 1 is expanded to only 1 consecutive state that is 1, two consecutive states that are 1s,$\cdots$, g consecutive states that are 1s, where g denotes the minimum gap. Therefore, in our calculation a unit may have as many as $g+2$ possbile states at a time. At time t, for every possible state, we can find the most economic path, therefore, we only have to record $g+2$ optimal paths and update them hour by hour till we meet the end of the year. However, when the maintanance is met, we change all the paths to the most optimal path, and begin new dynamic programming after the maintanance is done.

Our software is written in matlab and runs on HP 166MHz station. It spends about half an hour to calculate unit commitment, while it spends 4 to 5 hours to calculate duality gap. This is one reason why we no longer use the Lagrange relaxation method.

Power Flow Calculations and Final adjustments

After determining the unit commitment, we find bus 23 always has units in operation and the generation capacity is always large. Therefore, it is convenient for us to choose it to be the swing bus. We then carry out power flow calculation for each hour. All power

flows are converged and only line 7–8 sometimes exceeds the continuous thermal overload rate. When the 3 units connected with bus 7 are on, we reduce the power output of the 3 units to mitigate the thermal overload of line 7–8. The power flow calculations still converge, and the thermal overload cases are confined to 6 hours, all are the thermal overload of line 7–8. The most severe case has the equivalent $S = \sqrt{P^2 + Q^2}/V_* = 176.6$ MVA, exceeds the continuous rate 175 MVA. Here P and Q are in Unit MVA, V is line voltage's per unit value. All these cases occur when there is no unit at bus 7 in operation. On the other hand, such a minor violation is tolerable especially for a short time. Therefore, we have already arranged system states at every hour for one year. Further risk analysis can proceed based on our calculation result.

BIBLIOGRAPHY

[1] IEEE PICA 97 Tutorial Committee, *Open Access Impact on Energy Scheduling Pricing and Control*, Columbus, Ohio: IEEE Power Industry Computer Applications, 1997.

[2] Richard P. O'Neill and Ben Hobbs and David Mead and Michael Rothkopf, "Short-term electric auction markets: ISOs, information, the inconvenience of non-convexity, and inappropriate behavior", *Proceedings of Bulk Power System Dynamics and Control IV-Restructuring*, pp. 3-28, Santorni, Greece, August 1998.

[3] Gerald B. Sheble, "Alternative implementations of auction mechanisms", *Proceedings of Bulk Power System Dynamics and Control IV-Restructuring*, pp. 69-80, Santorni, Greece, August 1998.

[4] Junji Kubokawa, Hiroshi Sasaki, Kenji Okada, Hiroshi Asano, Ryuichi Yokoyama, "Evaluation of Reserve Supply Cost in the Japanese Electricity Market", *Proceedings of Bulk Power System Dynamic and Control IV-Restructuring*, pp. 341-350, Santorni, Greece, August 1998.

[5] Jung-Uk Lim, Kwang Y. Lee, Young-Moon Park, Hee-Seog Koh, "An analytic approach to cost allocation for multiple wheeling transactions", *Proceedings of Bulk Power System Dynamic and Control IV-Restructuring*, pp. 719-724, Santorni, Greece, August 1998.

[6] Paul Haase, "Charting power system security", *EPRI journal*, pp. 27-31, September/October 1998.

[7] E. Henley and H. Kumamoto, *Probabilistic Risk Assessment*, New York: IEEE Press, 1991.

[8] J. Grandell, *Aspects of Risk Theory*, Berlin: Springer-Verlag, 1991.

[9] Emmett J. Vaughan, *Risk Management*, New York: John Wiley & Sons, 1997.

[10] IEEE, *The New IEEE Standard Dictionary of Electrical and Electronic Terms*, 5th edition, Piscataway: IEEE Press, 1993.

[11] Agustin A. Irizarry-Rivera, *Risk Management for Dynamic Security Constrained Electric Power Systems*, Ph.D. dissertation, Ames, Iowa: Iowa State University, 1997.

[12] Roy Billinton, *Reliability Evaluation of Power Systems*, London: Plenum Press, 1996.

[13] CIGRE Task Force 38.03.12, *Power System Security Assessment: A Position Paper*, Paris: CIGRE, 1997.

[14] CIGRE Study Committee 38, *Power System Reliability Analysis Application Guide*, Paris: CIGRE, 1987.

[15] T. Dy Liacco, "The adaptive reliability control system", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, pp. 517-531, May 1967.

[16] T. Dy Liacco, "Real-time computer control of power systems", *Proceedings of IEEE*, vol. 62, pp. 874-891, July 1974.

[17] T. Dy Liacco, "System security: the computer's role", *IEEE Spectrum*, vol. 15, pp. 43-50, May 1967.

[18] North American Electric Reliability Council, *NERC Planning Standards*, Princeton, New Jersey: NERC, September 1997.

[19] B. Porretta, D. L. Kiguel, G. A. Hamoud, E. G. Neudorf, "A comprehensive approach for adequacy and security evaluation of bulk power systems", *IEEE Transaction on Power Systems*, vol. 6, no. 2, pp. 433-441, August 1991.

[20] TRELSS User Group (TUG) Meeting & Training Workshop, *Transmission Reliability Evaluation*, EPRI Report, September 1998.

[21] CIGRE Task Force 13 of Advisory Group 38.03, "Sequential probabilistic methods for power system operation and planning", *CIGRE Proceedings*, Paris, pp. 69-99, August 1998.

[22] James D. McCalley and Vijay Vittal, *Risk Based Security Assessment*, EPRI Final Report, Project W08604-01, 1998.

[23] G.E.P. Box, G.M. Jenkins, D. W. Bacon, "Models for Forecasting Seasonal and Nonseasonal Time Series", in *Spectral Analysis of Time Series*, pp. 146-151, New York: Wiley, 1967.

[24] G.E.P. Box and G. M. Jenkins, "Some Recent Advances in Forecasting and Control", *Applied Statistic Society*, vol. B24, pp. 1321-1328, 1968.

[25] G.E.P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, revised ed., London: Holden-Day, 1976.

[26] G.E.P. Box, G. M. Jenkins, G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed., Englewood Cliffs, New Jersey: Prentice Hall, 1994.

[27] J.W. Cooley and J.W. Tukey, "An algorithm for the machine calculation of complex fourier series", *Math. Computation*, vol. 19, pp. 297-301, 1965.

[28] R. E. Kalman, "New methods and results in linear filtering and prediction theory", *ASME J. Basic Engineering*, vol. 83, pp. 35-45, 1960.

[29] Timothy Masters, *Neural, Novel & Hybrid Algorithms for Time Series Prediction*, New York: John Wiley & Sons Inc., 1995.

[30] D. Brillinger, *New Directions in Time Series Analysis*, Berlin: Springer Verlag, 1992.

[31] A. Piras, A. Germond, B. Buchenel, K. Imhof, Y. Jaccard, "Heterogeneous artificial neural network for short term electrical load forecasting", *IEEE Trans. on Power Systems*, vol. 11, no. 1, pp. 397-402, February 1996.

[32] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, A. Azeem, J. Farah, C. Drake, "Practical experiences with an adaptive neural network short-term load forecasting system", *IEEE Trans. on Power Systems*, vol. 10, no. 1, pp. 254-265, February 1995.

[33] Alireza Khotanzad, Rey-Chue Hwang, Alireza Abaye, "An Adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities", *IEEE Trans. on Power Systems*, vol. 10, no. 3, pp. 1716-1722, August 1995.

[34] Alex D. Papalexopoulos, Shangyou Hao, Tie-Mao Peng, "Implementation of a neural network based load forecasting model for the EMS", *IEEE Trans. on Power Systems*, vol. 9, no. 4, pp. 1956-1962, November 1994.

[35] T. S. Dillon, "Short term load forecasting using an adaptive neural network", *Electrical Power & Energy Systems*, vol. 13, no. 4, pp. 186-192, August 1996.

[36] Shin-Tzo Chen, David C. Yu, Alireza R. Moghaddamjo, "Weather sensitive short term load forecasting using nonfully connected artificial neural network", *IEEE Trans. on Power Systems*, vol. 7, no. 3, pp. 1098-1105, August 1992.

[37] Hiroyuki Mori, Hidenori Kobayashi, "Optimal fuzzy inference for short-term load forecasting", *IEEE Trans. on Power Systems*, vol. 11, no. 1, pp. 390-396, February 1996.

[38] Hiroyuki Mori, Shouichi Urano, "Short-term load forecasting with chaos time series analysis", Proceedings of the International Conference on Intelligent Systems Applications to Power Systems, pp. 133-137, Janurary 1996.

[39] A.G. Bakirtzis, J. B. Theocharis, S. J. Kiartzis, K. J. Satsios, "Short term load forecasting using fuzzy neural networks", *IEEE Trans. on Power Systems*, vol. 10, no. 3, pp. 1518-1524, August 1995.

[40] Kwang-Ho Kim, Jong-Keun Park, Kab-Ju Hwang, Sung-Hak Kim, "Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems", *IEEE Trans. on Power Systems*, vol. 10, no. 3, pp. 1534-1539, August 1995.

[41] Dipti Srinivasan, "Demand forecasting using fuzzy neural computation, with special emphasis on week end and public holiday forecasting", *IEEE Trans. on Power Systems*, vol. 10, no. 4, pp. 1897-1903, November 1995.

[42] D. Srinivasan, "Forecasting daily load curves using a hybrid fuzzy-neural approach", *IEE Proceedings on Generation, Transmission and Distribution*, vol. 141, no. 6, pp. 561-567, November 1994.

[43] E. L. Silva, M. Morozowski, L. G. S. Fonseca, G. C. Oliveira, A. C. G. Melo, J. C. O. Mello, "Transmission constrained maintenance scheduling of generating units:

a stochastic programming approach", *IEEE Trans. on Power Systems*, vol. 10, no. 2, pp. 695-701, May 1995.

[44] M. K. C. Marwali, and S. M. Shahidehpour, "Deterministic approach to generation and transmission maintenance scheduling with network constraints", *Electrical Power Systems Research*, vol. 47, no. 2, pp. 101-113, October 1998.

[45] X. Wang and J. R. McDonald, *Modern Power System Planning*, London: McGraw-Hill Book Company, 1994.

[46] R. Yokoyama and T. Niimura, "Thermal generating unit maintenance scheduling by multi-stage application of simulated annealing", *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pt. 2, vol. 2, pp. 1531-1535, October 1996.

[47] Hiroshi Sasaki, Yoshiyuki Takiuchi, Junji Kukobawa, "Application of artificial networks to maintenance scheduling covering of thermal units over several consecutive years", *Electrical Engineering in Japan*, vol. 116, no. 4, pp. 64-74, April 1996.

[48] Amril Aman, Anantram Balakrishnan, Vijay Chandru, "On-line maintenance of optimal machine schedules", *Sadhana-Academy Proceedings in Engineering Sciences*, vol. 22, pt. 2, pp. 257-279, April 1997.

[49] Kenneth L. Kiper, "Risk informed on-line maintenance at Seabrook Station", *International Conference on Nuclear Engineering*, pp. 240-241, May 1997.

[50] Allen J. Wood and Bruce F. Wollenberg, *Power Generation Operation, and Control*, New York: John Wiley & Sons, 1996.

[51] Sridhar Kondragunta, *Genetic Algorithm Unit Commitment Program*, M.S. Thesis, Ames, Iowa: Iowa State University, 1997.

[52] Fred N. Lee, "The application of commitment utilization factor to thermal unit commitment", *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 691-698, May 1991.

[53] Thomas J. Overbye, "Power flow measure for unsolvable cases", *IEEE Transactions on Power Systems*, vol. 9, no. 3, pp. 1359-1365, August 1993.

[54] Roy Billinton and Saleh Aboreshaid, "Voltage stability considerations in composite power system reliability evaluation", *IEEE Transactions on Power Systems*, vol. 13, no. 2, pp. 655-660, May 1998.

[55] A.C.G. Melo, J.C.O. Mello, S. Granville, "The effects of voltage collapse problems in the reliability evaluation of composite systems", *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 480-488, February 1997.

[56] S. Granville, J. C. o. Mello, A. C. G. Melo, "Application of interior point methods to power flow unsolvability", *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 1096-1103, May 1996.

[57] X. Wang, G. C. Ejebe, J. Tong, J. G. Waight, "Preventive/Corrective control for voltage stability using direct interior point method", *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 878-883, August 1998.

[58] P. Kessel and H. Clavitsch, "Estimating the voltage stability of a power system", *IEEE Transaction on Power Delivery*, vol. 1, no. 3, pp. 550-557, July 1986.

[59] Powertech Labs, Inc., *Release Notes for VSTAB Version 4.1 Survey*, British Columbia, Canada: Powertech Inc. Press, January 1996.

[60] The Reliability Test System Task Force of the Application of Probability Methods Subcommittee, "The IEEE Reliability Test System-1996", *IEEE PWRS WM*, pp. 326-334, 1996.

[61] Youjie Dai, James D. McCalley, Vijay Vittal, "Expansion and Enhancement of Direct Interior Point Algorithm on Power System Maximum Loadability", *IEEE Power Industry Computer Applications 99*, pp. 234-241, April 1999.

[62] G.A.N. Mbamalu, M.E. El-Hawary, Ferial El-Hawary, "A comparison of probabilistic power flow with deterministic based solutions", *Electric Machines and Power Systems*, vol. 24, no. 5, pp. 511-520, 1996.

[63] B. Borkowska, "Probabilistic load flow", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 752-759, 1974.

[64] O. A. Klitin, "Stochastic load flows", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-94, no. 2, pp. 299-309, March/April 1975.

[65] R. N. Allan, A. M. Leite da Silva, A. A. Abu-Nasser, R. C. Burchett, "Discrete convolution in power system reliability", *IEEE Transactions on Reliability*, vol. R-30, no. 5, pp. 452-456, December 1981.

[66] R. N. Allan, "Probabilistic load flow using multilinearisations", *IEE Proceedings*, vol. 128, pt. c, no. 5, pp. 280-287, September 1981.

[67] A. M. Leite da Silva, V. L. Arienti, R. N. Allan, "Probabilistic load flow considering dependence between input nodal powers", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-103, no. 6, pp. 1524-1530, June 1984.

[68] H. R. Sirisena and E. P. M. Brown, "Representation of non-Gaussian probability distributions in stochastic load-flow studies by the method of Gaussian sum approximations", *IEE Proceedings*, vol. 130, pt. C, no. 4, pp. 165-171, July 1983.

[69] Marian Sobierajski, "Optimal stochastic load flows", *Electrical Power Systems Research*, no. 2, pp. 71-75, 1979.

[70] Roy Billinton and Wenyuan Li, *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*, London: Plenum Press, 1994.

[71] B.C. Lesieutre and J. Hockenberry, "Uncertainty analysis of power system simulations and ATC calcualtions using the probabilistic collocation method", *Proceedings of Bulk Power System Dynamics and Control IV-Restructuring*, pp. 573-584, Santorini, Greece, August 1998.

[72] A. M. Leite da Silva, S. M. P. Ribeiro, V. L. Arienti, R. N. Allan, M. B. F. Do Coutto, "Probabilistic load flow techniques applied to power system expansion planning", *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1047-1053, November 1990.

[73] Youjie Dai, James D. McCalley, Vijay Vittal, "Annual Risk assessment for system thermal overload", *Proceedings of American Power Conference*, pp. 88-93, Chicago, April 1998.

[74] Hua Wan, James D. McCalley, Vijay Vittal, "Increasing Thermal Rating by Risk Analysis", *IEEE PES Winter Meeting*, pp. 128-135, February 1998.

[75] Weihui Fu, James D. McCalley, Vijay Vittal, "Risk-based assessment of transformer thermal loading capability", *Proceedings of North American Power Symposium*, pp. 118-123, October 1998.

[76] James R. Buck, *Economic Risk Decisions in Engineering and Management*, Ames, Iowa: Iowa State University Press, 1989.

[77] Peter J. Brockwell and Richard A. Davis, *Introduction to Time Series and Forecasting*, Berlin: Springer, 1996.

[78] Kongde Zhong, *Digital Signal Processing*, Beijing: Tsing Hua University Press, 1988 (in Chinese).

[79] John G. Proakis and Dimitris G. Manolakis, *Digital Signal Processing Principles, Algorithms, and Applications*, Upper Saddle River, New Jersey: Prentice Hall, 1996.

[80] James D. Hamilton, *Time Series Analysis*, Princeton, New Jersey: Princeton University Press, 1994.

[81] L. L. Garver, "Effective load carrying capability of generating units", *IEEE Transaction on Power Apparatus and Systems*, pp. 476-483, August 1966.

[82] Ihsan Karabulut, *Edgeworth Expansion and Bootstrap Approximation for the Distribution of M-estimators of A Simple Linear Regression Parameter without Cramer's Condition*, Master Thesis, Ames, Iowa: Iowa State University Press, 1991.

[83] R. N., Allan and M. R. G. Al-Shakarchi, "Probabilistic Techniques in a.c. Load Flow Analysis", *Proceedings of IEE*, vol. 124, pt. c, no. 6, pp. 531-536, 1976.

[84] Youjie Dai, James D. McCalley, Vijay Vittal", "Stochastic load model identification and its possible applications", *Proceedings of North American Power Symposium*, pp. 244-251, October 1997.

[85] Youjie Dai, James D. McCalley, Vijay Vittal, "A heurisitc method to arrange unit commitment for one year considering hydro-thermal coordination", *Proceedings of North American Power Symposium*, pp. 382-387, October 1998.