UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

AN ANALYSIS OF SHORT-TERM RISK IN POWER SYSTEM PLANNING

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

Andrew P. Douglas
Norman, Oklahoma
1997

UMI Number: 9733894

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

AN ANALYSIS OF SHORT-TERM RISK IN POWER SYSTEM PLANNING

A Dissertation APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

To My Family:

William Edmund, Gayle Frances and Michael Adam

TABLE OF CONTENTS

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF TABLES

LIST OF ILLUSTRATIONS

ABSTRACT

The goal of this research is to formulate and present a methodology that evaluates short-term risk in power system planning. Specifically, this research shows how to determine the risk of short-term planning in the presence of electrical load forecast and fuel price uncertainty, both of which have a large impact on the outcome of power system production cost planning. The uncertainty in the load is described by Bayesian forecasting and fuel price uncertainty is modeled by conditional triangular probability distributions.

Classical decision analysis forms the backbone of the methodology presented herein. Throughout this dissertation, sampling theory, load forecasting theory and general engineering are applied with the aim of transforming the short-term power system planning problem into a suitable structure for decision analysis. Probabilistic sampling is used to discretize the load and fuel prices; then an electrical power production simulation model results in a unit commitment strategy and a cost of each plan. A best, i.e., minimum cost, plan can be selected and the expected cost of uncertainty can be estimated.

The results presented in this dissertation are as follows: The impact of weather forecasts on Bayesian load forecasting as a function of forecast lead time is shown in Chapter 2. Risk in the presence of load forecast uncertainty alone and risk in the presence of load forecast uncertainty together with fuel price uncertainty are shown in Chapters 4 and 5, respectively. The expected cost of uncertainty, in these chapters, is given as a function of lead time in $/MWh.

CHAPTER 1

INTRODUCTION

§ 1.1 Introduction

The goal of this chapter is to introduce to the reader the general subject matter of this dissertation, and to describe the motivation for this research topic. To begin, the classical purpose of the electric utility is presented; this discussion is followed by a synopsis of the changing electrical energy market, and a description of where an electricity supplier takes its new place.

Next, utility system planning is discussed. Within this section, the classical planning process is presented, as well as, planning in a competitive environment. Long-term planning is briefly described, because the concepts of classical long-term planning will be used to form the framework of short-term planning in the future. Following this, risk in short-term planning is discussed, and then an illustrative example shows an application of the purposed research as an effective planning tool.

To conclude the chapter, an outline of the steps taken to formulate the risk analysis model is given, and a literature review cites a list of works on which this research is based.

§ 1.2 The Electric Energy Market

In the past, the basic mission of the electric utility has been to provide its customers with their desired demand, i.e., to deliver electrical energy to the utility's end users, while minimizing operational costs, satisfying the system load and maximizing reliability. This mission is often called "The obligation to serve," and has well described the electrical utility industry in the United States. Presently, the industry is undergoing changes to its fundamental structure. This industry, which in the past has operated as a government regulated monopoly, is moving towards a free enterprise energy market, where the traditional boundaries of a well-defined service area or franchise no longer exist. These types of changes, not only impact the structure of an electric utility, but affect business practices and strategic planning, as well.

In a competitive market, the obligation to serve philosophy of system operation is outdated and can be replaced by the notion of profit maximization. Past fundamental operational concepts, such as the obligation to serve all end users' demand requirements at a minimum cost, change to a utility's opportunity to serve a customer's need for a specified length of time, while maximizing profit. In the future, an electric utility will be challenged to provide the customer with incentives, such as low energy prices and high power-quality, in order to sell its product. In addition, energy rates, which in the past were decided in rate hearings by public commissions, will be set during contract negotiations or sales agreements.

Inherent to the change of the electric utility market is a restructuring of the industry itself. The electric utility industry will most likely be split into four independent entities: generation, transmission, distribution and marketing. In the current industry infrastructure one company owns the first three sectors, while the last, power marketing, has not surfaced as a primary function. In the future, it will be possible for competitors to join the supply sector without having to build transmission or distribution systems. Utility companies will have the option to sell distribution and transmission sectors, or contract for maintenance and upkeep. With the proposed future free enterprise energy market, the possibilities are too numerous to entertain in full. However, the importance of new price structuring cannot be overstated.

In the future, the energy customer will likely pay a local distribution fee with a fixed overhead charge for Operation and Maintenance (O&M) and an energy charge for real time usage. For transmission services, the customer may pay an O&M fee, a demand charge and a distance charge. To the generation company, a power marketer or a distribution company may pay a negotiated price per MWh, but most importantly, the energy customer will be given the opportunity to share in the risk of generation scheduling and capacity allocation.

These are only a few possibilities of the direction the electric utility industry can take. However, there is at least one certainty; as the electric utility industry changes, so will the operational planning concerns and methods. This research attempts to anticipate these future planning changes by including risk in the operational planning model.

§ 1.3 Power System Operational Planning

The power system planning process will undergo changes as the electric utility industry transforms into a competitive market. Power system operational planning can be summarized as the development of operational strategies, which dictate the actions of the electric utility over a finite period, while attempting to achieve the goal of maximizing profit. In general, operational planning is grouped into short and long-term study horizons. Short-term planning takes place over the course of a day to a few months, and the long-term planning study horizon ranges from several months to several years. As this research applies a method used in classical long-term planning to the short-term planning problem of the future, the following section briefly describes long-term power system planning and the particular long-term planning model this research exploits.

§ 1.3.1 Long-Term Planning

In general, long-term system planning involves decisions regarding capacity expansion and system extension. Fuel budgeting is also included in the long-term planning process. A crucial decision in the electric utility industry is whether or not to purchase power from neighboring suppliers or to build new production facilities. Since construction lead times can run from five to ten years on large power plants, the load growth uncertainty can have a huge impact on the financial outcome of the project. Both under and over-sizing a plant can be financially devastating. In the past, capacity expansion decisions were made based on two criterion: one was the minimum percentage capacity margin; the other was the loss of load probability of one day in five years [1]. If production lead time could be shortened, the uncertainty of the load growth forecast could be reduced.

Decision theory provides the underlying framework for the solution process of such studies. This type of risk analysis involves making long-term energy and demand growth forecasts, developing a decision structure that is suitable to the problem and finding the expected cost of uncertainty due to load growth variation. See Reference [1] for a detailed treatment of the subject. Similar studies can be performed for short-term planning, where the unit commitment decision replaces the decision to build a new power plant. In the following section, the application of decision analysis and the aforementioned long-term risk model to short-term power system planning studies will be discussed.

§ 1.3.2 Short-Term Planning

In short-term planning, the goal is to allocate an appropriate mix of generation to satisfy the projected demand for the upcoming week. Note that the study horizon can be reduced to a matter of hours or extended to several weeks. Both load and fuel price forecasts play an important role in the short-term planning process. Even small variations in either can result in large operating cost swings.

For short study horizons, when weather forecast lead time permits, temperature dependent load forecasts can be used. Looking into the future, the load forecast variance increases as a function of time due, in part, to the weather forecast uncertainty. As the electrical energy industry changes into a free enterprise market, the requirement of quoting prices will become increasingly important. The cost of uncertainty that is due to the load forecast will impact prices quoted from one hour to one month ahead of time. Of course, with shorter lead times, the risk of quoting prices decreases. This type of planning problem can be approached in a manner similar to the long-term planning case described previously. Decision theory will play a large role in the solution to the risk problem. In addition, it is possible to consider more than one uncertain planning variable. This research will also detail the inclusion of fuel price uncertainty into the risk model.

§ 1.4 Risk in Short-Term Power System Planning

Risk analysis can be an useful tool in the power system planning process. In the past, operational scheduling was performed for the purpose of minimizing costs while complying with the obligation to serve. In the future when rates will be decided during contract negotiations, the electric utility must have at their disposal a set of tools to analyze and quote prices intelligently. When taking pricing issues into account, there are many new challenges the electric utility will have to overcome in order to maintain economic stability. It is in the area of pricing that the research herein provides the greatest contribution to the power system planning field.

In the past, when rates were decided in public hearings, utilities based their price structures on the expected production cost plus a fixed rate of return. Moreover, if units were committed during real time operation that were not necessary, or if a utility had to buy power from a neighboring system to cover unserved energy, the extra costs were passed on to the end user in the form of a cost adjustment. In the future, however, rates will be negotiated in advance, and prices will probably be binding for the length of the contract. The most important point here is that these contracts will more than likely not contain cost adjustment clauses. As a result, the electric utility may have a larger incentive to operate their system efficiently, and the value of accurate forecasts will increase. If the load forecast or fuel price estimate is in error, and these are the bases for future energy rates, then the supplier may have to cover the extra cost, and not the customer. Future energy rates should be quoted with uncertainties in mind, but at the same time, be competitive with respect to other suppliers in the industry. The following section describes one example of how risk analysis can be applied to the short-term power system planning process.

§ 1.5 Descriptive Example

A manager of a power company has the task of putting together a bid for a contract to provide energy to a local industrial customer. The customer is a tomato cannery, which receives its fruit from local and out-of-state growers. The desired length of the contract is for the main production season, which equates to the tomato harvest. Typically, the harvest lasts from early May to mid-September. The load factor of the cannery during this season is near unity. The load is comprised mostly of induction motors and lighting. The variation in the plant's demand is only due to shift changes, production line changes and the transitions from day to night. To the unwary manager, the cannery sounds like a great customer. How can you beat a unity load factor? There are, however, hidden problems.

The first problem is that the substation, which is owned by the plant, is extremely sensitive to voltage fluctuations. Any sudden variations in the voltage will trip the breakers. This may compromise system security, because the cannery's demand represents a large enough contribution to the local demand that grid stability is a concern. Also, the plant production supervisors tend to ignore recommendations of incremental start-up schemes, which are designed to gradually bring the plant up to full operating capacity. Another major concern is that the weather has been very erratic this year. It is expected that the consistently warm summer nights, which are required for tomatoes to ripen, will not regularly occur.

While operation in the past was not disrupted due to planned harvests and dependable weather, processing this year's harvest will be a challenge. There may be weeks without any fruit and some weeks with too much. This issue concerns the tomato company as well; they run their plant at full capacity so extra tomatoes mean wasted tomatoes. Fortunately, the plant has a few options. The over-ripe tomatoes can be run through the sauce and ketchup lines; however, the solid pack production will suffer. The possibility of irregular operation also concerns the utility manager.

In order to serve the plant's demand, an additional medium size coal-fired unit must be brought on line. With minimum up and downtime running several days, it is still better to keep the unit on line, even when the cannery experiences downtime. However, the system itself would not be optimally scheduled. How does this manager decide on a pricing scheme, and what type a structure should be used?

The extra coal-fired unit lightens the burden to some of the expensive gas peaking units; as a result, the system will run at a higher economic efficiency, when the cannery is included. The question is, to whom is the greatest portion of risk assigned. Should the electric company cover the cost of the risk, because their profit margin will be higher while serving the tomato plant, or should the tomato cannery be held accountable because of the uncertainties they impose on the system?

Fortunately our manager is equipped with a rare understanding of the price and risk elasticities of the customer. Moreover, the tomato plant supervisors will know one to two weeks ahead of time if a weekly harvest will not arrive as expected. With all of this information, the manager can reasonably negotiate as to what portion of the risk each participant should pay, as well as, determine how far ahead to quote prices. One possibility is to quote monthly prices at a higher rate due to the uncertainty in the sporadic harvests and potential lulls in activity. This tends to keeps the "spot" price of electricity constant, but the premium the cannery pays is due to uncertainty in the load forecasts and tomato harvests, which are not accurate for a month.

Another plan is to provide weekly prices because the harvest and the weather will be known well enough in advance to project the activity of the following week. If no problems occur this may lead to lower prices over time, but this situation does increase the chances of volatile energy prices. All of this information can be used with the expected cost of uncertainty to develop a price strategy suitable for the electric utility and the tomato cannery. The bottom line is that uncertainty implies risk, and the cost of this risk should not be left out of pricing considerations. This discussion is only one of many applications that can utilize risk analysis to aid in system planning decisions.

§ 1.6 Risk and the Unit Commitment Decision

As stated before, evaluating the risk in short-term planning requires the use of concepts from decision analysis, power system planning and sampling theory. The planning process is modeled as a classical decision analysis problem whose goal is to evaluate the expected cost of uncertainty. The first step is to identify the decision variable; next, the load and the fuel prices must be discretized to represent the possible states of nature; lastly, probabilities must be assigned to the states of nature.

For the purposes of this research the decision variable is the short-term unit commitment decision. The states of nature are high, medium and low values of the electrical load and fuel prices. If the future were known, the optimum strategy could always be implemented, however, as planning occurs in advance, the utility runs the risk of selecting a unit commitment strategy that will not minimize the production cost over the chosen study horizon. That is, another unit commitment may exist that would result in a lower production cost. The difference between the minimum expected cost in the presence of uncertainty and the expected cost if the future were known is said to be the risk associated with that unit commitment decision. The uncertainties that exist within the planning process are many; this research considers two very important parameters: the electrical load and fuel prices.

The electric load can be thought of as both a deterministic and a random time series. The system load, in the deterministic sense, is nothing more than the aggregation of each end user's demand, plus system losses. However, this type of description is not adequate for planning purposes. In general, the number of end users are large, and a good percentage of them cannot reliably predict their future energy consumption.

As a chronological profile of the electrical load is needed to determine a unit commitment strategy, the time series must be provided in advance of when the operational plan is to be implemented; a load forecast profile is one type of load representation that is sufficient for power system planning. Moreover, the load can be characterized as a stochastic time series. Since a great deal of effort has been devoted to estimating random processes, forecast methods are numerous. However, in order to take advantage of decision analysis, the states of nature are better modeled as discrete events. If the electrical load could be described by a probability density function, stratified sampling theory could then be applied to separate the load into individual events. Probabilities of the states of nature are also required for decision analysis, and are generated during the stratification process.

To represent the electrical load as a distribution, a forecast method should be selected based on this criterion. Bayesian forecasting is one such method; it provides both a forecast mean and variance. If the electric load could be considered a normal random process, then the forecast mean and variance, from Bayesian estimation, completely describe the normal probability distribution. At this point, the stratified sampling process can be applied. The above description forms the backbone of the purposed research, and is detailed in the following chapters.

Fuel prices also have a considerable effect on the planning process. A triangular distribution will be used to represent the fuel prices, and the same stratified sampling process, as with the load, will be used to discretize fuel prices into individual events. Because fuel prices are affected by the electrical load, additional efforts will be required to accurately model this relationship. When more than one planning parameter imposes uncertainty on the planning process and interdependencies exist, conditional densities should be used in the stratification process. Also, when multiple inputs are used, the number of simulations required to evaluate the risk can grow very large. Reductions in simulation sizes should be made if possible. A method proposed to account for this situation is derived from the concepts of statistical experimental design, and is presented in this research. The next section gives a brief outline of the major subjects covered by this research and where they can be found.

§ 1.7 Dissertation Outline

The organization of the rest of this dissertation is outlined below. A literature survey concludes this introductory chapter. Chapter 2 is devoted to Bayesian electrical load forecasting. Chapter 3 covers the necessary prerequisites from sampling theory and other modeling concerns. Chapter 4 focuses on the formulation of the solution of the expected cost of load uncertainty. Chapter 5 covers the combined problem of representing fuel price and load forecast uncertainty. Case studies are performed and included in Chapters 2, 4, and 5 from utility derived system data and National Weather Service temperature forecasts. The results are summarized and evaluated in Chapter 6. Chapter 6 also serves as a conclusion that provides the reader with a summary of the research results, as well as possible directions to further these research efforts.

§ 1.8 Literature Survey

In the following section, an outline of the references used by this research will be presented. The survey is split into three sections, each describing a different field whose concepts are drawn from within this research. The subjects described below are decision theory sampling theory, and load forecasting.

§ 1.8.1 Decision Theory

Some of the first works on decision analysis, following that of Von Neumann and Morgenstern who proved the Max-Min Principle for two-person zero-sum games in Reference [2], were authored by Raiffa and Luce [3], Howard [4,5] and North [6]. Raiffa and Luce concentrate on game theory; however, a short treatment of the decision analysis problem is included. Raiffa and Luce relate the decision problem to the two-person zero-sum game, where the decision maker plays against nature. Howard explains the basic structure and procedures of setting up the practical decision problem in a qualitative manner in Reference [4], and he takes a mathematical approach to the decision analysis problem in Reference [5]. North provides an excellent tutorial on the basic mathematical decision problem in Reference [6]. Both References [5] and [6] tie the Bayesian notion of prior probabilities to the conditional expectation of a probability density function, an important result this research will exploit. Raiffa develops the notion of the expected cost of uncertainty in Reference [7]. Esser et al. and Anders list many examples of the application of decision analysis to power system planning in References [8] and [9]. Many of the examples involve building or siting decisions; however, Esser et al. suggest that decision analysis techniques can be utilized to evaluate operational decisions, such as the long-term trade-offs of capacity expansion. References [10] and [11] make this claim as well. The proposed research will take these works in a different direction and explore the effects of uncertainty in short-term operational planning.

§ 1.8.2 Electrical Load Forecasting

Over the past few decades, many procedures have been used for short-term electrical load forecasting. The most common approaches include time series methods [12-15], regression techniques [16-19] and artificial neural network algorithms (ANN) [20-23]. ANN represents the most popular and current technology. One significant characteristic of ANN is that the method captures non-linear relationships between input and output variables. References [20] and [21] are some of the first applications of ANN to electrical load forecasting. References [22] and [23] are current sources that report mean absolute errors of 0.02 and better.

In regression analysis, the load is represented as a linear combination of exogenous inputs, which include weather and load information. References [16] and [17] are some of the first applications of multivariate regression to electrical load forecasting. References [18] and [19] comprise a set of current works on file.

The time series approach is the most widely documented methodology. Autoregressive moving average (ARMA) models are the most common time series approach that has been applied to load forecasting. The autoregressive part of ARMA is an all-pole recursive filter, while the moving average part is an all-zero recursive filter. Transfer function models may be used to incorporate explanatory variables. References [12] and [13] are some of the first works using ARMA models for load forecasting. References [14] and [15] are some of the current works in the fields.

What is required is a forecast method that is simple to implement, and that provides a forecast distribution. It is in this area where Kalman filters and Bayesian estimation excel. The Kalman filter provides a forecast distribution, including forecast variance, which is useful in risk analysis. The load signal is modeled as the output of a state space dynamic system; the model is adaptive in that the Kalman gain changes with the forecast error. Some of the first applications of Kalman filtering theory to electrical load forecasts are cited in Reference [24] and [25]. Current works in the field are covered in References [26] and [27]. A disadvantage of Kalman filters is the necessity of knowing both the process and the observation noise.

Parallel in time to the development of Kalman filters was the development of Bayesian estimation. When normal distributed noise is considered, the algorithms are identical. However, additional features were developed in Bayesian estimation theory which compensate for the noise identification problems prevalent in Kalman filtering theory. The use of discount factors in place of the state evolution noise and the addition of a variance learning algorithm for the observation noise makes Bayesian estimation an attractive forecasting model that provides a forecast distribution. The first works on Bayesian analysis are cited in References [28-30]. Current works are covered in References [31-33]. The concept of sampling with unknown parameters is well known; References [31,32] show its application to Bayesian estimation. Reference [28,33] incorporate variance learning into the fundamental Kalman filter/Bayesian estimation forecast algorithm.

This research will exploit these features of Bayesian estimation in order to further the study of lead time dependent risk analysis in power system planning.

§ 1.8.3 Sampling Theory

Fundamental to decision theory is the ability to properly structure a real-life situation into a decision tree. Since both the electrical load and future fuel prices are described by continuous probability density functions, some means of forming discrete events is necessary. Stratified sampling is employed to provide representations of discrete events for a given probability distribution. Works on general sampling theory include References [34-36]. These works, however, concentrate on discrete sampling of finite populations. References [37-40] provide excellent coverage of sampling with infinite populations described by probability density functions.

This concludes the literature survey. In the following pages, these three topics, load forecasting, stratified sampling and decision analysis, will be drawn on as this research details the development of a methodology that evaluates the risk as a function of lead time due to load forecast and fuel price uncertainty in short-term power system planning.

CHAPTER 2

SHORT-TERM BAYESIAN ELECTRICAL LOAD FORECASTING

§ 2.1 Introduction

This chapter presents a short-term load forecast methodology that is suitable for power system operational planning studies. Bayesian estimation is use to predict multiple step ahead peak forecasts using peak and average temperature forecasts as explanatory variables. Herein, the forecast model is developed and illustrated in a case study with utility derived system data. Special attention is given to the practical issue of forecasting the electrical load with imperfect weather information.

Electrical load forecasting is an integral step in the short-term power system operational planning process. Accurate hourly forecast profiles, with study horizons of up to one week, are necessary for economic dispatch, energy transactions and short-term generation scheduling. As the electric energy supply sector changes to a competitive industry, the need to operate at maximum economic efficiency has increased. Electrical load forecast accuracy greatly affects the utility's ability to achieve this goal.

Coupled with each load forecast algorithm is an inherent model error, i.e., the fundamental errors due to the model not completely representing all load characteristics of the service area. This model error is determined empirically, and usually with perfect weather information. However, to evaluate a load forecast method based on model errors alone overlooks an important factor that enters into practical applications. Most short-term load forecast algorithms use weather information to refine the load estimate; in a real time setting, weather information generally takes the form of temperature forecasts. Errors present in the temperature forecast can have a huge impact on the accuracy of the final load forecast and also the planning process.

Consider a utility that has made a one week load forecast based on imperfect weather information, and has chosen to implement the optimum unit commitment decision based on that load forecast. If the load forecast is in error, enough such that a different unit commitment strategy would have been otherwise selected, the utility will be forced to reschedule its units, if possible. This type of planning is a common procedure in real time power systems operation. However, increasing the accuracy of the load forecast will result in minimizing the number of times generation rescheduling is necessary. To fully appreciate this aspect of short-term planning, the usefulness of a load forecast algorithm should be analyzed when using temperature forecasts.

The focus of this chapter is to report on the application of Bayesian estimation to electrical load forecasting, and to present comparative results with both perfect temperature information and temperature forecasts supplied by reliable sources. Bayesian estimation is a recursive algorithm that is used in conjunction with a dynamic linear model (DLM) to represent the behavior of the electrical load. Bayesian forecasting makes use of *a priori* and *a posteriori* distributions, along with sampled data, to project estimates of the future electrical load. Bayesian forecasting has the ability to make multiple step ahead forecasts, and the DLM easily models the trend and periodic characteristics of the electrical load. In addition, explanatory or causal variables, such as weather information, can be introduced into the model to further reduce the forecast error. Also, the Bayesian forecasting method yields a forecast distribution. The desire for the availability of forecast distributions has increased as the electric utility industry changes to a competitive market, where planning under uncertainty is emphasized.

The structure of the rest of this chapter is as follows. Section 2.2 details this chapter's concise problem statement. Section 2.3 describes the Bayesian forecasting paradigm, and its application to electrical load forecasting. Section 2.4 discusses weather behavior and presents graphical results of both temperature forecast biases and standard deviations as a function of forecast lead time. Section 2.5 takes the form of an illustrative case study that shows the impact that temperature forecast errors can have on electrical load forecasts. Section 2.6 concludes the chapter with a summary and evaluation of the results.

Load forecasting methods for short-term planning fall into several different categories, such as regression analysis, artificial neural network methods, time series analysis, and state space approaches. Autoregressive moving average (ARMA) models have received a great deal of attention in the literature [12-15]. ARMA models fall into the time series category, but can be implemented in state space formulation. Linear regression models are also common [16-19]. These models minimize the squared error over the data set to determine weighting coefficients of explanatory variables. Artificial Neural Network (ANN) models are currently the most popular and reported method of load forecasting [20-23]. ANN models are unique in that they capture the non-linear relationship between the load and weather information. Kalman filtering is a state space method found in the literature [24-27]. Kalman filtering uses the state space description to recursively update forecast results. However, a difficult aspect of Kalman filtering is the selection of the process and the observation noise. In the following pages, the Bayesian estimation method of forecasting will be applied to the electrical load forecasting problem [32,33]. The updating equations of Bayesian estimation are similar to those of Kalman filtering; however, several model enhancements allow for easier implementation.

§ 2.2 Problem Statement

Given an acceptable set of historical load data, historical temperature data and temperature forecast data provided by the National Weather Service, the goal of this chapter is to present a multiple step ahead load forecast methodology, adequate for use in system planning studies. Results are presented, which provide further insight into the impact of temperature forecast errors on electrical load forecasting. The load forecast takes the form of an hourly load profile that is suitable for input to a chronological production cost simulation program, which determines the unit commitment strategy. Bayesian forecasting and dynamic linear modeling comprise the main prediction algorithm. This model is structured to predict daily peaks from one to several days in advance. Typical daily load profiles are then scaled to the forecasted peaks to provide a weekly load profile suitable for chronological production cost simulators. Note that this is different from real time forecasts used for generation dispatch, where the study horizon is generally one hour to one day. This chapter emphasizes the theoretical development of the Bayesian forecasting method and details its application to short-term electrical load forecasting. In addition, an error analysis is provided with actual and forecasted weather information used as exogenous inputs to the forecast model.

§ 2.3 Methodology

Bayesian estimation is a recursive prediction algorithm that uses an *a priori* distribution and sampled data to update an *a posteriori* distribution for the next time increment. A dynamic linear model (DLM) is used in conjunction with this updating scheme to describe the output signal characteristics. Trends and periodic behavior are easily represented by the dynamic linear model in terms of the state space description. The DLM also has the capability to reflect the impact of explanatory variables on the forecast. These exogenous inputs in the case of electrical load forecasting take the form of up-to-date weather information and forecasts. The following development is taken from [32,33]. For a thorough treatment of the theoretical aspects of Bayesian estimation and dynamic linear models, see [32]; for practical examples, see [33]. For practical suggestions regarding software implementation, see [25].

§ 2.3.1 Bayesian Estimation

A stochastic state space description is used as the dynamic linear model (DLM). Since the DLM is to be used as an estimator only, by selecting the evolution and output matrices properly, the system will maintain observability. The state space system is given as follows:

$$\begin{aligned} \theta_t &= G\theta_{t-1} + \omega_t \quad \text{where} \quad \omega_t \sim N(0, W_t) \\ Y_t &= F_t'\theta_t + \nu_t \quad \text{where} \quad \nu_t \sim N(0, V_t) \end{aligned} \tag{2.1}$$

θ_t is the time dependent (k×1) state vector, G is the (k×k) state evolution matrix and W_t is the (k×k) covariance matrix of the zero mean process noise, ω_t. Y_t is the scalar output, F_t is the time dependent (k×1) regression vector and V_t is the scalar variance of the zero mean observation noise, ν_t. The normal noise processes, ω_t and ν_t, are both mutually and temporally independent. G describes the behavior of the state vector, θ_t, over time, which can be represented by trend and cyclic components. Although G can be time varying, in this case G is fixed. Note that the above state space system is given in general terms. Section 2.3.3 covers its application to electrical load forecasting. F_t

contains exogenous regression variables. Bayesian estimation relies on the recursive use of Bayes' rule to generate a forecast. That is, prior distributions are updated to posterior distributions via Bayes' rule and these distributions are scaled based on forecast errors. To the above multivariate normal state space description, Bayes' rule is applied. With some mathematical manipulation, a set of updating equations can be formulated. Equation (2.2) shows the *a posteriori* distribution for time t-1 and the *a priori* distribution for time t.

$$
\begin{aligned}
(\theta_{t-1} | D_{t-1}) &\sim T_{n_{t-1}}(m_{t-1}, C_{t-1}) \\
(\theta_{t-1} | D_{t}) &\sim T_{n_{t-1}}(a_{t}, R_{t}) \\
a_t &= Gm_{t-1} \\
R_t &= \delta^{-1} \cdot GC_{t-1}G' \\
S_t &= d_{t-1} \cdot n_{t-1}^{-1}
\end{aligned}
\tag{2.2}
$$

D_{t-1} is the complete state knowledge at time t-1. C_{t-1} is the (k×k) *a posteriori* covariance matrix and m_{t-1} is the (k×1) *a posteriori* state mean vector. D_t is the state knowledge at time t. The updated *a priori* mean and covariance of time t are a_t and R_t, respectively, where n_{t-1} is the degrees of freedom of the student-t distribution and S_t is an estimate of the scalar output variance at time t. d_{t-1} and n_{t-1} are also parameters of a Gamma probability distribution used in the derivation of Equation (2.2). The forecast distribution is shown below.

$$
\begin{aligned}
(Y_t | D_{t-1}) &\sim T_{n_{t-1}}(f_t, Q_t) \\
f_t &= F_t' a_t \\
Q_t &= F_t R_t F_t' + S_t
\end{aligned}
\tag{2.3}
$$

Where f_t and Q_t are the scalar forecast distribution mean and variance. The one step ahead point forecast is generally taken to be the forecast distribution mean. The updated *a posteriori* distribution is give by:

$$
\begin{aligned}
(\theta_t | D_t) &\sim N(m_t, C_t) \\
m_t &= a_t + A_t e_t \\
C_t &= U_t R_t U_t' + A_t S_t A_t' \\
U_t &= (I - A_t F_t') \\
e_t &= Y_t - f_t \\
A_t &= R_t F_t Q_t^{-1} \\
n_t &= n_{t-1} + 1 \\
d_t &= d_{t-1} + e^2 S_t Q_t^{-1}
\end{aligned}
\tag{2.4}
$$

where m_t and C_t denote the updated *a posteriori* mean and covariance, respectively. A_t is the adaptive factor and e_t is the forecast error. The adaptive factor, A_t, and the covariance matrix, C_t, are asymptotic in nature and converge over time. In our practical application, W_t, the process noise covariance, is replaced by a scalar discount factor, and a variance learning scheme is used in lieu of V_t, the observation noise variance. The updating equation for C_t is somewhat different than what is cited in the Bayesian literature. After a number of iterations, the covariance matrix, C_t, becomes non-symmetric, when implemented on a digital computer. Instead, the Joseph stabilized equation is used. This symmetric version guarantees the positive semi-definiteness of C_t in the presence of round-off and truncation error, and is often used in software implementation. See Reference [25] for further details.

At times, it is desirable to make multiple step ahead forecasts. In this case, the k^{th} step ahead forecast distribution is given as follows:

$$\begin{aligned}(Y_{t+k} \mid D_t) &\sim T_{n_{t-1}}(f_t(k), Q_t(k)) \\ f_t(k) &= F_t' a_t(k) \\ Q_t(k) &= F_t R_t(k) F_t' + S_t \\ a_t(k) &= G m_t(k-1) \\ R_t(k) &= \delta^{-1} \cdot G R_t(k-1) G'\end{aligned} \tag{2.5}$$

§ 2.3.2 Model Identification

Until now, only the general system has been discussed, and no mention has been made regarding the selection of G, F, d and δ. G is the state evolution matrix with dimensions (k×k). G can be time varying, but for purposes of this research, G is time invariant. The selection of G depends on which fundamental signal components (trend, periodic behavior, causal relationships) are to be modeled. G is a block diagonal matrix where each block corresponds to a signal component form. A portion of the regression vector, F, is devoted to each component block in G. A thorough treatment of this subject is given in [32].

Once the model has been identified based on the output signal characteristics, a word or two on the selection of initial conditions is warranted. Due to the asymptotic nature of both m and C, if enough lead time data exist, almost any reasonable guess will be fine, provided m and C are not zero. For example, in the case of forecasting the daily peak load explained later in this chapter, if m is selected "arbitrarily" and C is large, twenty days of lead time data is sufficient for m and C, to converge. It is certainly possible to select m and C with better accuracy, but this requires much experience and has little impact on the forecast results after the transient period. However, the selection of both d and δ has an effect on the forecast outcome. d should be chosen close to the actual output variance, because in practice, the convergence of d is slow. The discount factor, δ, is basically a tuning parameter to refine the accuracy of the forecast. There are many different forecast error criteria, which can be used to evaluate the accuracy while adjusting δ. The mean absolute error, MAE, or the mean squared error, MSE, are reasonable methods to try. The bulk of the previous discussion is taken from [32] and [33], which provide excellent theoretical development as well as practical examples of Bayesian forecasting.

§ 2.3.3 Electrical Load Forecasting

This section covers the application of Bayesian estimation to electrical load forecasting. The first decision to make is whether to use an hourly model or a daily peak model. There are certain difficulties that arise in the practical implementation of hourly models. In order to use an hourly model, an hourly temperature profile is required, if the causal relationship between temperature and the electrical load is to be exploited. In usual practice, weather forecast services only provide peak, valley, and average temperature predictions for the following day and the rest of the week. Even if a typical temperature profile is scaled to match the peak, valley and mean temperature predictions, forecasting an entire week would require a multiple step ahead forecast of up to 168 hours; this situation is not reasonable as the forecast variance can grow very large for this many iterations without updating. In addition, the model order is larger than 170th; even with the Joseph stabilized algorithm, numerical instability is still a concern. Another approach is to use a daily peak model where the peak load is forecasted, and then a typical load profile is linearly scaled to generate the required load profile. Historical load data are generally available to find typical load profiles, and these profiles tend to be good representations, when drastic weather changes are not present. Furthermore, the periodic representation is only seventh order for the weekly load cycle, and the model order including regression variables is 15^{th}. In fact, it is this model that has been implementation herein.

The model uses as regression variables, the previous day mean, peak, and hour 24 temperature and load, as well as, temperature forecasts of the next days peak and average temperature. A full harmonic DLM description is used to represent the loads weekly cyclic behavior with a period of seven. A first order trend component is also included. For this system, F is shown as follows:

$$F' = \left[1 \; L_{AVG(t-1)} \; L_{PK(t-1)} \; L_{24(t-1)} \; T_{AVG(t-1)} \; T_{PK(t-1)} \; L_{24(t-1)} \; T_{AVG(t)} \; T_{PK(t)} \; 1 \; 0 \; 1 \; 0 \; 1 \; 0\right] \quad (2.6)$$

$L_{AVG(t-1)}$, $L_{PK(t-1)}$ and $L_{24(t-1)}$ are the previous day average, peak and hour 24 loads, respectively. $T_{AVG(t-1)}$, $T_{PK(t-1)}$ and $T_{24(t-1)}$ are the previous day average, peak and hour 24 temperatures, respectively. $T_{AVG(t)}$ and $T_{PK(t)}$ are the forecasted daily average and peak temperatures, respectively. Note that t in Equation (2.6) is in days.

As to initial conditions, m can be selected "arbitrarily" and C as $1\times10^{7}I_{15\cdot15}$. The discount factor which gave rise to the lowest MSE was found to be 0.93. This indicated that past information remains fairly important to the forecast results. Although a block discounting method can be used whereby each component block has its own discount factor, in practical experience with load forecasting, it has been found that block discounting had little impact on the forecast results. The model tended to yield a forecast standard deviation of 200 and so d_0 was selected as $(200)^2$, or 40000.

Now that all of the initial conditions have been specified for the peak load forecast model, a scheme must be used to linearly scale a typical day to generate a load forecast profile. This process can be reduced to the solution of a 2×2 linear set of equations. The known quantities at the time of the forecast are the load estimate at hour 24 of the previous day and the typical load at hour 24 of the previous day. We also have the peak load of the typical day and the forecasted peak load, all of which can be used to solve for the parameters a, b in the following system of equations [19]:

$$\begin{aligned} \hat{l}(p) &= a + b l_t(p) \\ \hat{l}(0) &= a + b l_t(0) \end{aligned} \tag{2.7}$$

$\hat{l}(p)$ is the peak forecast, $l_t(p)$ is the typical day peak load, $\hat{l}(0)$ is the load estimate at hour 24 of the previous day and $l_t(0)$ is the typical load at hour 24 of the previous day. When solving for a and b we have:

$$\begin{aligned} a &= \frac{l_t(p)\hat{l}(0) - l_t(0)\hat{l}(p)}{l_t(p) - l_t(0)} \\ b &= \frac{\hat{l}(p) - \hat{l}(0)}{l_t(p) - l_t(0)} \end{aligned} \tag{2.8}$$

Once a and b are found, the typical day can be scaled to fit the peak load forecast.

$$\hat{l}(i) = a + b l_t(i) \quad \forall\, i = 1,\ldots,24 \tag{2.9}$$

This process guarantees that the typical day passes through the peak load forecast and the load estimate at hour 24 of the previous day. Note that the typical load at hour 24 of the previous day is required; therefore, a typical week must be generated. If a forecast variance profile is desired for this research, Equation (2.9) can be rewritten as:

$$\hat{l}(i) = \left(\frac{l_t(p) - l_t(i)}{l_t(p) - l_t(0)} \right) \hat{l}(0) + \left(\frac{l_t(i) - l_t(0)}{l_t(p) - l_t(0)} \right) \hat{l}(p) \tag{2.10}$$

The variance of $\hat{l}(i)$ is:

$$\begin{aligned} Var[\hat{l}(j)] &= \left(\frac{l_t(p) - l_t(j)}{l_t(p) - l_t(0)} \right)^2 Var[\hat{l}(0)] \\ &+ \left(\frac{l_t(j) - l_t(0)}{l_t(p) - l_t(0)} \right)^2 Var[\hat{l}(p)] \\ &+ 2 \frac{[l_t(p) - l_t(j)][l_t(j) - l_t(0)]}{[l_t(p) - l_t(0)]^2} Cov[\hat{l}(0), \hat{l}(p)] \end{aligned} \tag{2.11}$$

In the case that forecast is a one day ahead forecast, $\hat{l}(0)$ is not an estimate, but is a known value with zero variance. When multiple step ahead load forecasts are desired, note that by using the estimate of the previous day hour 24 load in the scaling equation, continuity of the load shape is preserved. This is very important when the load forecasts are intended to be used in a chronological production cost simulation. This completes the development of Bayesian estimation and electrical load forecasting. The following section will discuss temperature forecast errors.

§ 2.4 Weather Forecast Accuracy

Since the goal of this research is to capture the effects of weather forecast uncertainty in short-term load forecasting, it is important to note that a sizable portion of the load forecast error is due to a lack of accuracy in the weather forecast. Figures 2.1 - 2.4 are plots of the seasonal weather forecast standard deviations as a function of forecast lead time, from several forecast sources within the service area of the illustrative case study. Due to legal ramifications, we are not at liberty to cite these sources by name, with the exception of the National Weather Service signified by the dotted line in all figures. It is interesting to note that all seasons, with the exception of winter have similar variance characteristics. Notice the dip present after day three in Figures 2.2 - 2.4. Note also that the summer season has the smallest standard deviation. In addition to the standard deviation profile, the temperature forecast biases are shown in Figures 2.5 - 2.8. The bias is the estimate subtracted from the true value. The temperature bias plots show that all seasons have a negative bias (temperature forecast is high), with the exception of spring.

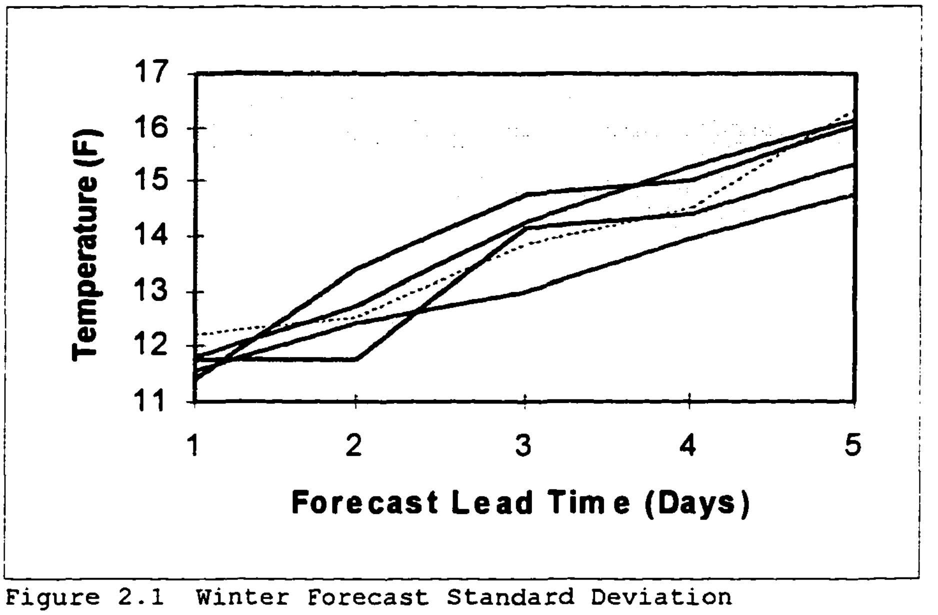

Figure 2.1 Winter Forecast Standard Deviation

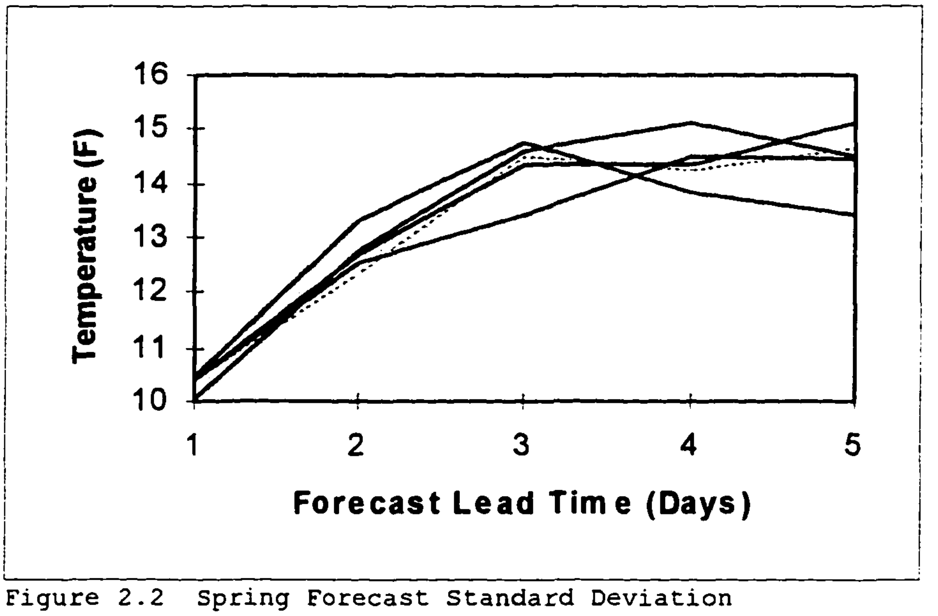

Figure 2.2 Spring Forecast Standard Deviation

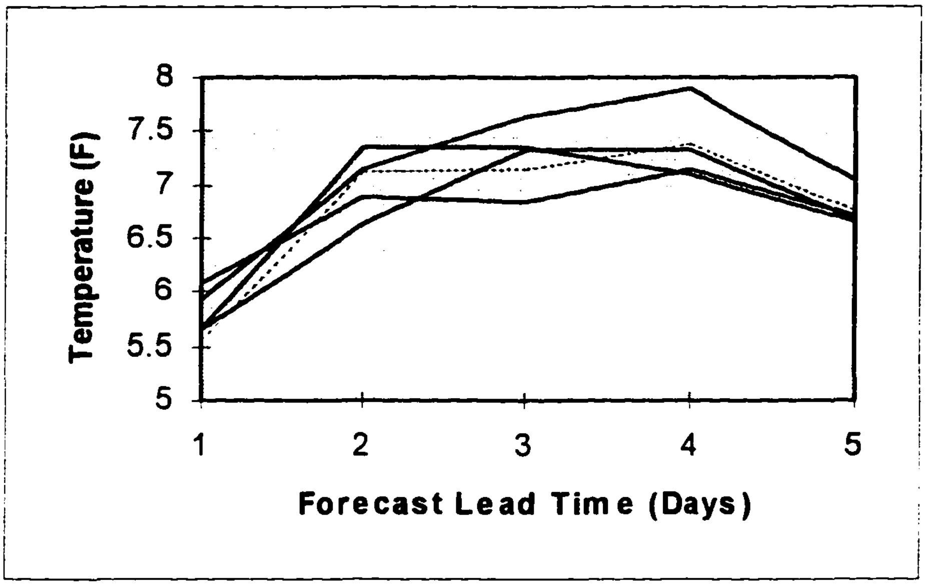

Figure 2.3 Summer Forecast Standard Deviation

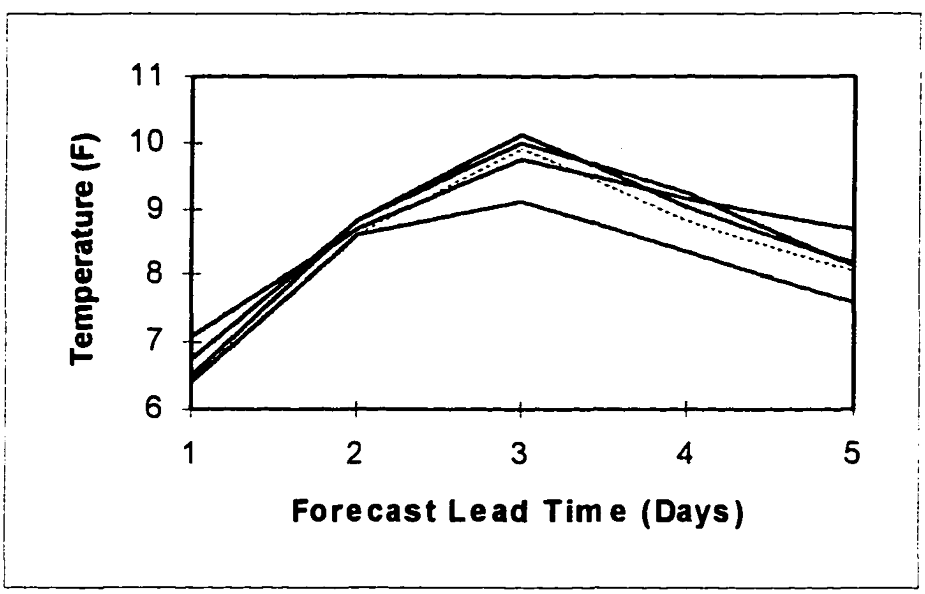

Figure 2.4 Autumn Forecast Standard Deviation

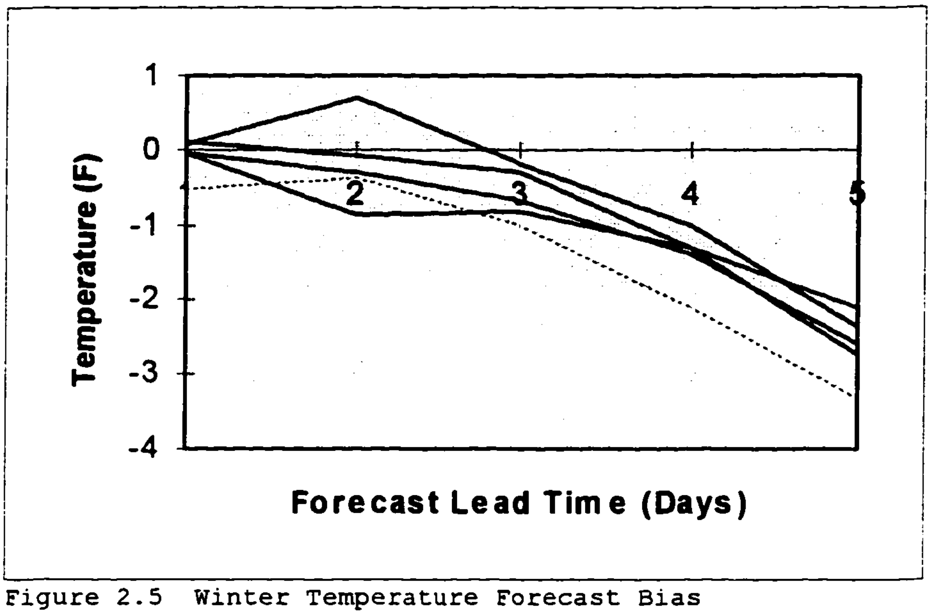

Figure 2.5 Winter Temperature Forecast Bias

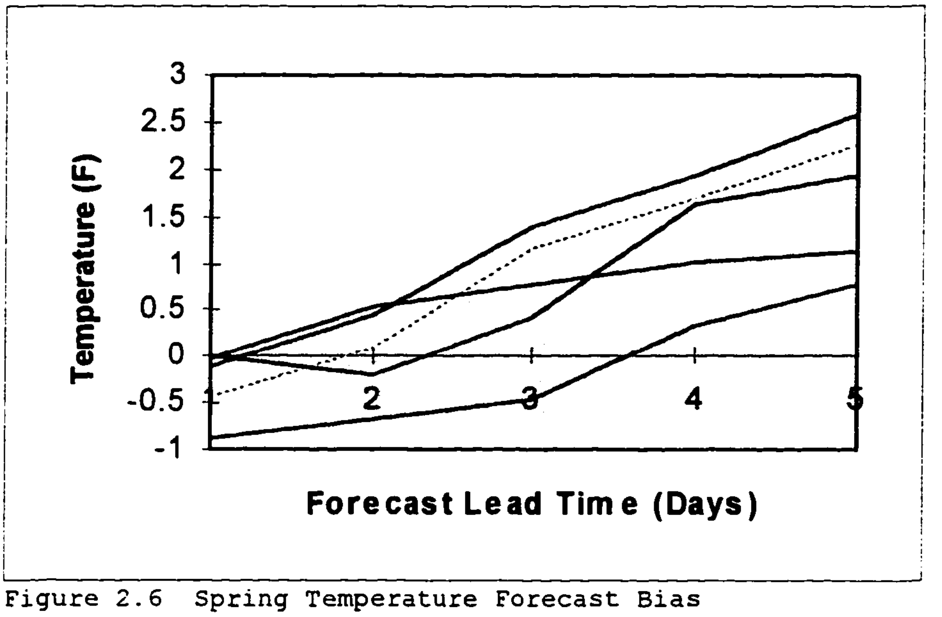

Figure 2.6 Spring Temperature Forecast Bias

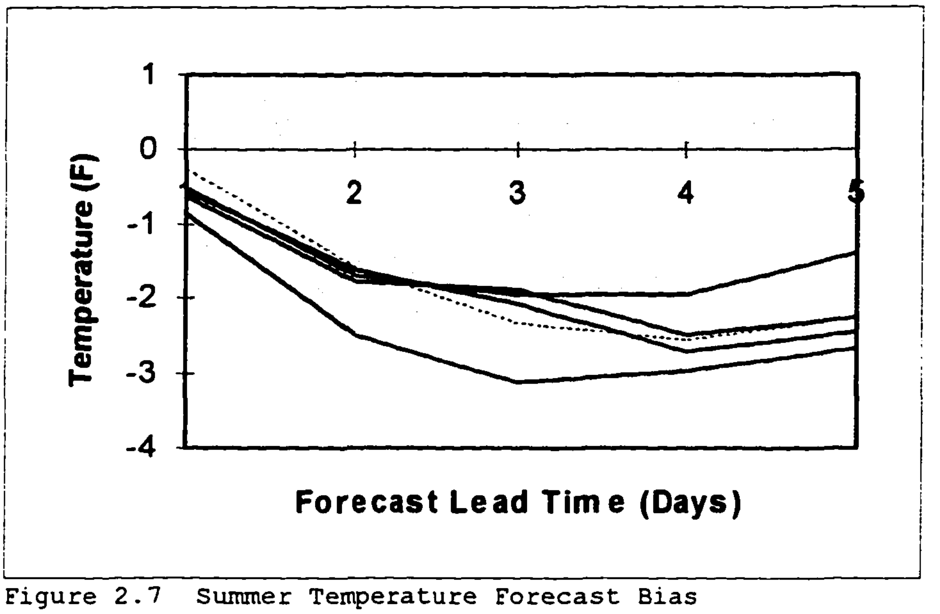

Figure 2.7 Summer Temperature Forecast Bias

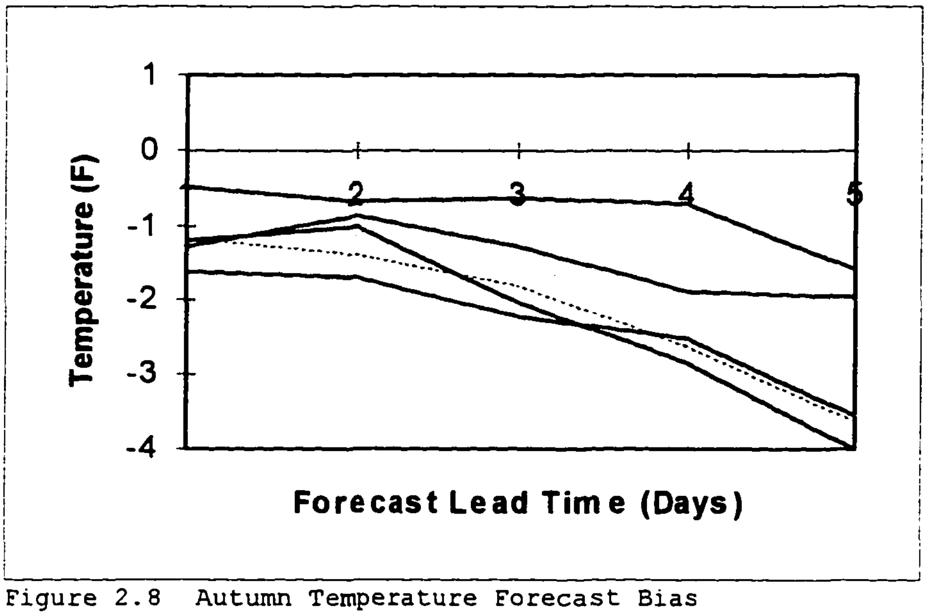

Figure 2.8 Autumn Temperature Forecast Bias

The plot in Figure 2.1 (winter) appears as expected. The temperature forecast standard deviation increases with lead time. However, Figures 2.2-2.4 (spring, summer, autumn) indicate that the standard deviation either levels out or decreases past day 2 or 3. These results are somewhat counterintuitive. An F test was performed at a 5% significance level to check if this behavior can be attributed to chance alone. For Figure 2.2 (spring), there is enough evidence to support the null hypothesis of equal variances from day three and beyond. For Figures 2.3 and 2.4 (summer, autumn), again there is enough evidence to support the null hypothesis of equal variances from day two and beyond.

§ 2.5 Illustrative Case Study

The case study preformed uses data from a medium size electric utility whose annual peak is around 5,000 MW and whose annual energy is roughly 23,000 GWH. Unit data for this system is detailed in Appendix 2.

In order to further understand the effects of forecast model error and the error present in the forecast due to weather forecasts, the following relationship is helpful [41].

$$E[(y - E[y|x])^2] = E[y^2] - E[E^2[y|x]] \qquad (2.12)$$

When put in terms of load forecasting, we have the following:

$$E[e^2] = E[(e - E[e|t=T])^2] + E[E^2[e|t]] \qquad (2.13)$$

Where e is the load forecast error and t denotes the temperature forecast random variable. Note in the above equation that $E[(e - E[e|t=T])^2]$ corresponds to both the forecast error variance due to inaccurate modeling and to variance that is inherent in the random model. Such is the case when actual temperatures are used in the forecast scheme. $E[E^2[e|t]]$ represents the portion of the load forecasting error that is due to weather forecast inaccuracies. That is, the errors that are present when temperature forecasts are used in place of actual weather information.

Below in Figures 2.9 - 2.12 are plots which show the peak normalized errors as a function of lead time for load forecasts made with forecasted temperatures and load forecasts made with actual temperatures. In each plot, the dashed line corresponds to load forecasts made with the National Weather Services' temperature forecasts, and the line marked with the diamonds are load forecasts made with actual weather data.

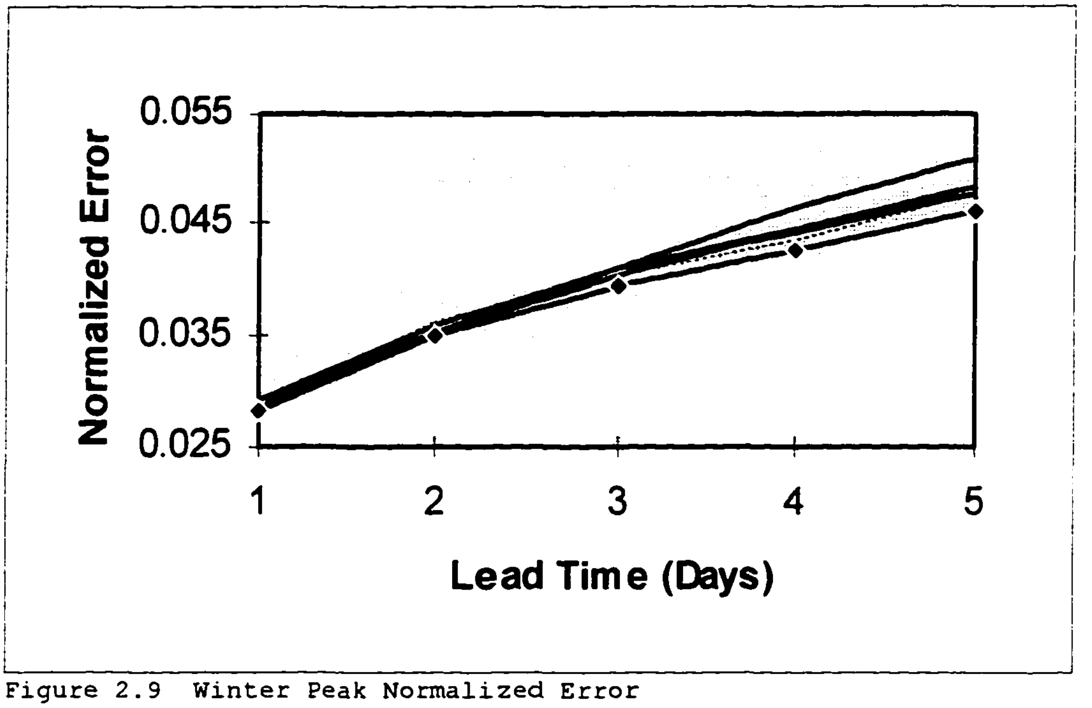

Figure 2.9 Winter Peak Normalized Error
◊ ~ Forecasts w/ Actual Temperatures

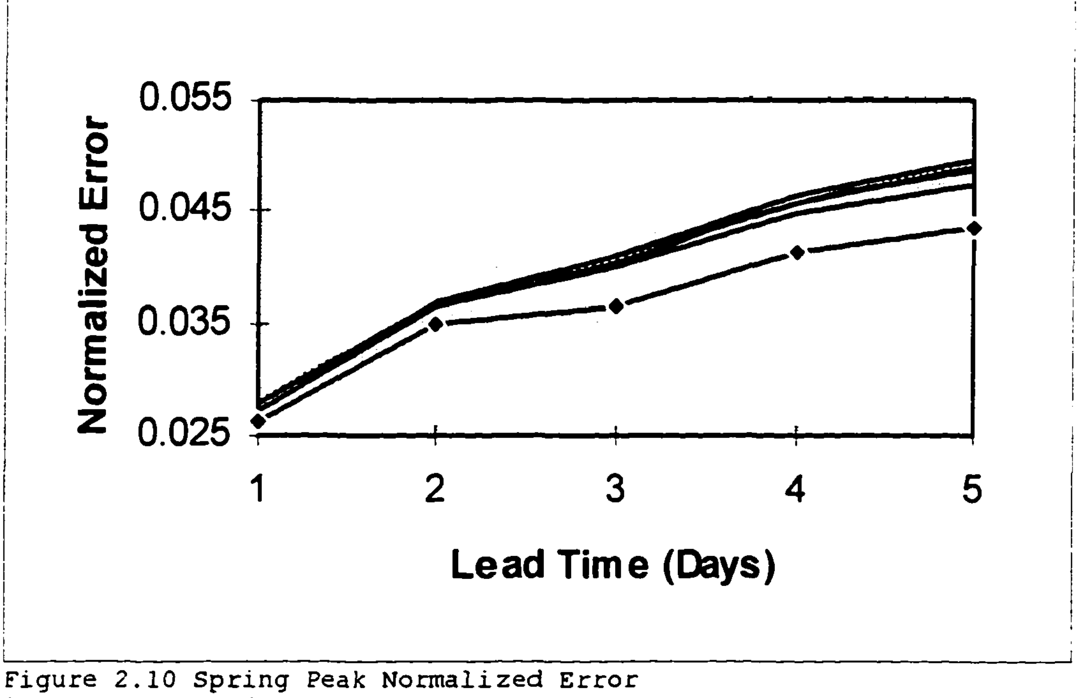

Figure 2.10 Spring Peak Normalized Error
◊ ~ Forecasts w/ Actual Temperatures

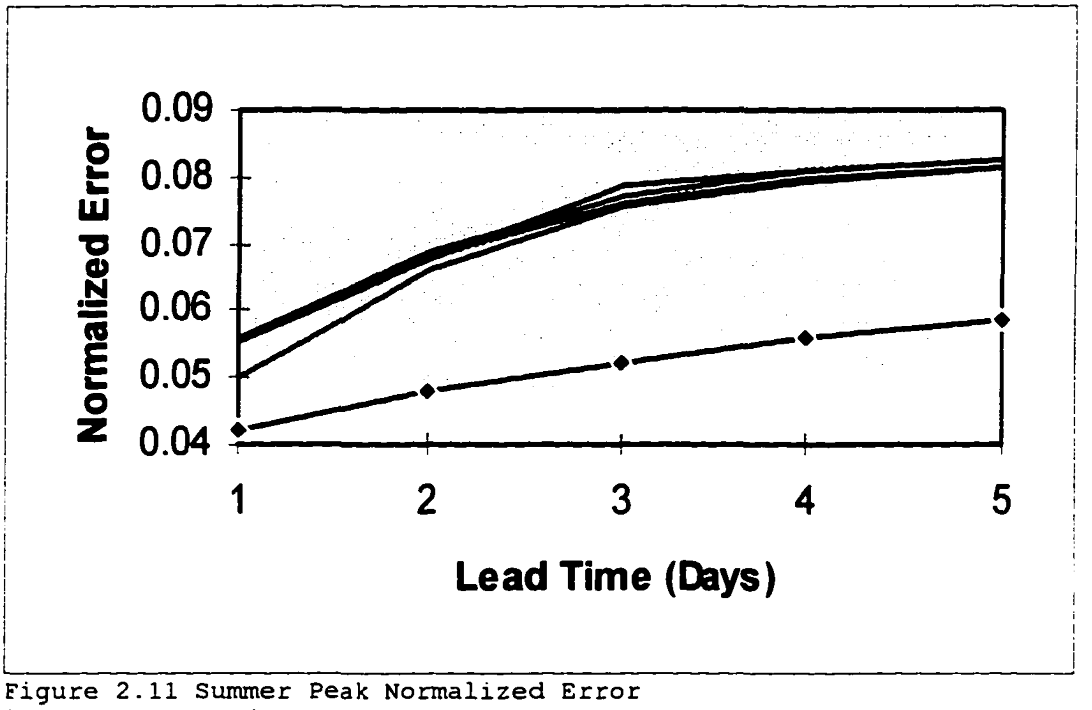

Figure 2.11 Summer Peak Normalized Error
◊ ~ Forecasts w/ Actual Temperatures

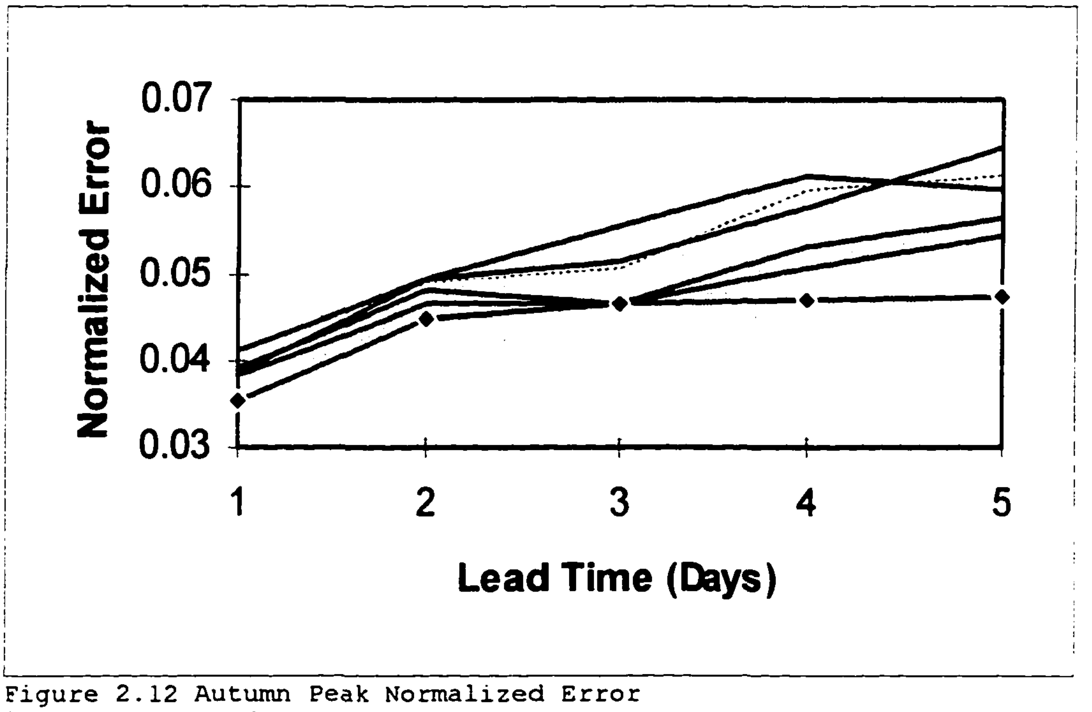

Figure 2.12 Autumn Peak Normalized Error
◊ ~ Forecasts w/ Actual Temperatures

§ 2.6 Conclusion

From the graphs above, one can see that the impact of temperature forecast errors differ with the annual seasons. While winter and spring have similar traits, summer is dramatically different. From Figure 2.11, the summer season, we see that the difference between load forecasting with actual temperatures verses forecasting with predicted temperatures can vary from thirty to fifty percent. It is also important to note that the summer standard deviations are the smallest of any of the seasons. The reason for this is due to the fact that the summer load is extremely weather sensitive. On the other hand, From the winter plot we see an example of a season which is temperature non-sensitive.

A technique using Bayesian forecasting to estimate multiple day ahead load forecast distribution profiles has been presented. With data from a medium sized utility and temperature forecasts from the National Weather Service, results are given as a function of forecast lead time and prove to be reasonable. In the following pages, the forecast mean and variance profiles will be used in conjunction with sampling theory to describe the states of nature of the electrical load to be used in the decision analysis risk model.

CHAPTER 3

MODELING CONSIDERATIONS

§ 3.1 Introduction

When developing a decision tree structure that describes a practical problem, four tasks must be accomplished. The decision variables need to be identified, states of nature must be defined and probabilities assigned to the possible outcomes. Also, the terminal node cost functionals have to be evaluated. Since the goal of this research is to evaluate risk due to uncertainty in short-term planning, the natural decision variable is the short-term unit commitment decision.

§ 3.2 Stochastic Unit Commitment vs. Risk

The short-term unit commitment problem has a finite number of solutions, one of which will be optimal for a given load and set of fuel prices. Unfortunately, the solution set is too large to enumerate exhaustively. A great deal of research has been done in this area to find a near optimal solution in a timely manner. Dynamic Programing, LaGrange Relaxation and Sequential Bidding are among the most popular solution methods. Each method requires a load forecast profile and fuel price estimates to determine the generation schedule. However, if one seeks the true optimal solution, but does not fix the load, and instead considers the load to be a random process, finding a solution that minimizes the expected production cost is a problem that has yet to be solved. This is known as the stochastic unit commitment problem.

This problem can be modeled using decision analysis. Consider a decision tree with finitely many courses of action, e.g., unit commitment strategies, and infinitely many states of nature corresponding to the electrical load. Note that the finite number of unit commitment strategies is still very large. The goal of this problem is to determine the unit commitment strategy that minimizes the expected cost, over all possible load profiles and fuel prices. It is important to remember that the emphasis lies in determining the unit commitment strategy; the evaluation of the production cost is secondary, even though the production cost is used as a minimization criterion. This research does not attempt to solve the stochastic unit commitment problem.

To evaluate the risk due to load forecast and fuel price uncertainty, instead of infinitely many states of nature, only three will be considered. These states are denoted high, medium and low for both the load and fuel prices. The set of admissible strategies are those unit commitment decisions that correspond to the load and fuel prices set to high, medium and low. It is not proposed that the solution strategies to the risk decision tree model are the same as the stochastic unit commitment problem. The unit commitment strategies corresponding to the high, medium and low states that have been discussed previously are not the only solutions. Others are certain to exist and one may be optimal.

Since the expected production cost corresponding to the solution of the risk model will be greater than or equal to that of the stochastic unit commitment problem, the risk model forms an upper bound, which is fairly tight in most practical circumstances. By reducing the states of nature, a feasible solution is attainable. However, it cannot be understated that, for the purposes of this research, the expected cost of the risk model is emphasized, the derived unit commitment strategy is secondary. The rest of the chapter will cover the development of a method to classify the load and fuel prices into states of nature and to assign probabilities to the possible outcomes.

§ 3.3 Sampling Theory

From sampling theory, to reduce study costs, a small number of samples from well chosen sub-populations can sufficiently represent an entire population. The sub-populations are called strata, and each stratum corresponds to a certain part of the total population that exhibits a similar trait. It is sometimes convenient, for the purposes of theoretical investigation, to represent the domain of a probability distribution function by a set of mutually exclusive strata. In this case, the population can assume infinitely many values. Furthermore, each stratum can be represented by its conditional expected value.

The electrical load for a given hour, which can be considered normal by the central limit theorem [42-44], and fuel prices, which are considered to follow a triangular distribution, will be represented as probability density functions so as to utilize sampling theory to synthesize the states of nature of the decision problem. For each distribution, the high and low strata correspond to the upper and lower p% of the area under the respective density function. The medium stratum corresponds to the (100-2p)% of the area remaining in the center.

An assumption that is made in this research is that the load is well behaved and maintains a similarly shaped hourly profile. It is felt that the set of load profiles which lie within certain bounds can be adequately described by the profile of hourly expected values. In Figure 3.1, the load is assumed to be similar in shape to Series D or C, which lies between the bounds of Series A and B. During different seasons of the year, the shape of the load profile may change, but these changes do not occur drastically. However, if the load profile is sporadic and is shaped like Series E, this type of analysis and characterization would not be appropriate.

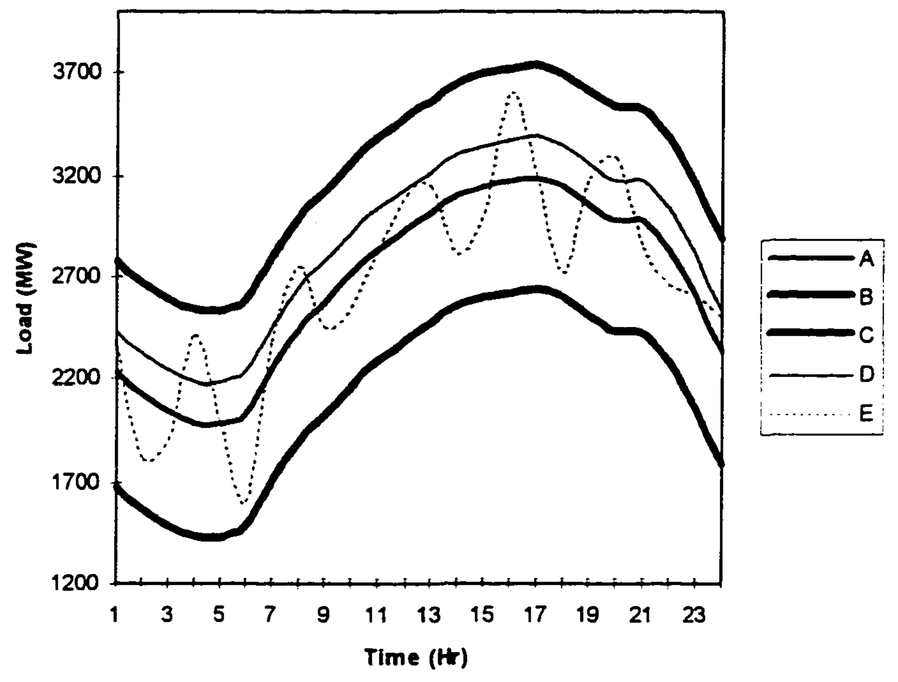

Figure 3.1 Load Profiles: Standard and Irregular

When using stratified sampling, two important decisions must be made. First is the number of strata to be used, and second is choosing strata boundaries. Selecting the number of strata is a matter of cost. In this case, the cost is computer simulation time. With the need for enumerating many combinations of varying load profiles, fuel prices and unit commitment strategies, if k is the number of possible actions, and m is the number of states of nature, then $k \times m$ simulations will be required to complete the study. Theoretically speaking, one could choose twenty partitions, but for only a small increase in accuracy a huge increase in computer simulation time is required.

Stratum boundary selection is also an important part of the sampling process. Reference [34] has suggested that as a first cut, the strata should be partitioned so that each stratum has an equal number of members. Since our population is a continuous density function, choosing the partition boundaries such that the area under the density function within each stratum is equal would do. However, Reference [40] states that a better choice is to select the strata boundaries such that the conditional variances of the strata are equal. When the population is the electrical load, which is described by the normal distribution, a solution can be approximated by using the standard normal distribution and the numerical bisection method. From Reference [40], the solution to the following yields the strata boundaries.

$$E[(z - \overline{z})^2 \mid z \in H_i] = Constant \quad \forall\, i = 1,2, \ldots, p \qquad (3.1)$$

Where p denotes the number of strata and H_i is the i^{th} partition. When p is large this is a difficult problem to solve. However, when p is three, a solution can be found by exploiting the symmetrical properties of the standard normal density function. When $p = 3$, we have the following partitions:

$$\{ (-\infty,-0.848);\ (-0.848,0.848);\ (0.848,\infty) \}$$

The probabilities for these partitions are approximately:

$$\{ 0.2,\ 0.6,\ 0.2 \}$$

Once the strata and partition boundaries have been selected, the conditional expected value of each stratum serves as the representation of an event. Below the expected value of z in (a,b) is shown.

$$E[z \mid z \in (a,b)] = \frac{\int_a^b z f(z)\,dz}{\int_a^b f(z)\,dz} \tag{3.2}$$

Where $f(z)$ is the standard normal probability density function as shown below.

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-z^2/2\right\} dz \tag{3.3}$$

When $a = 0.848$ and $b = \infty$ in Equation (3.2), the expected value of the high stratum is 1.389. Figure 3.2 shows the standard normal density function partitioned with equal conditional variances.

Since the fuel prices are described by the triangular distribution, to be consistent, the partitions will be selected using the (20%,60%,20%) area allotment. These distributions are unconditional distributions, and as said before, the expected value of each stratum represents a possible state of nature for both the load profile and fuel prices. The probabilities being (.2,.6,.2) for the independent states of nature corresponding to low, medium and high.

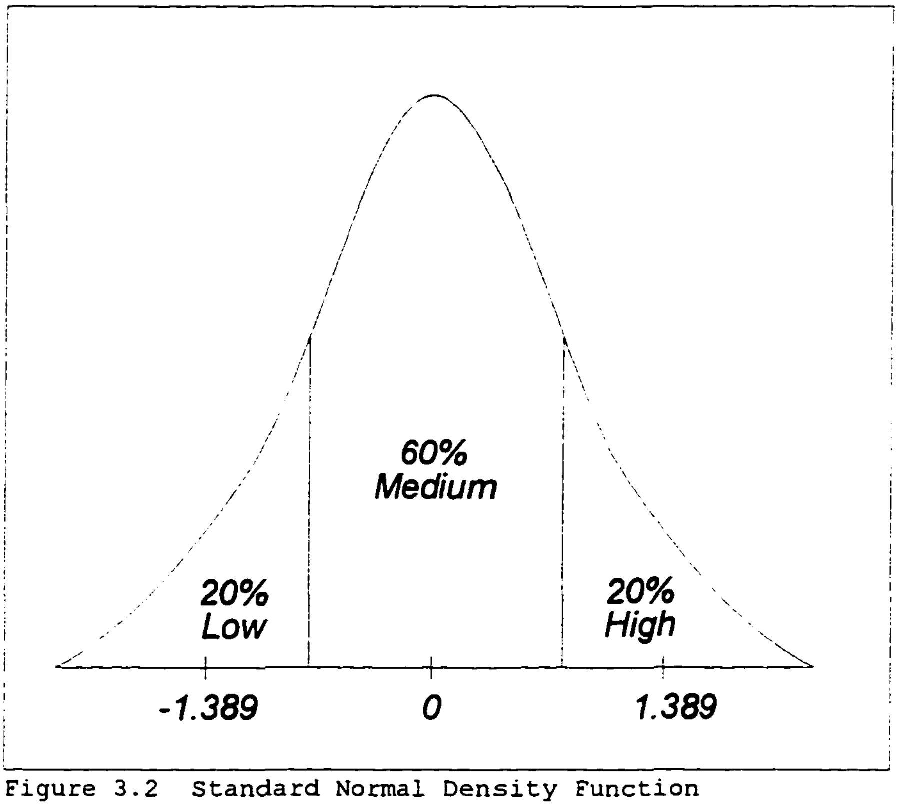

Figure 3.2 Standard Normal Density Function

However, interdependencies exist between fuel prices and varying loads, and between different fuels. Using conditional densities is a better representation, when evaluating risk with multiple dependent uncertain input variables.

§ 3.4 Conditional Distributions

When considering only load forecast uncertainty, we do not account for dependencies in the input variables. However, if it is desired to evaluate the risk due to both load forecast and fuel price uncertainty, the interdependencies that exist between the load and the fuel prices should not be overlooked. In the following pages, a method for considering these relationships is presented.

The easiest way to account for these dependencies is to specify the conditional densities. However, these conditional density functions are rarely known; herein a method is devised to approximate the generation of these conditional densities with a minimum number of requirements. Only the unconditional densities and the correlation coefficients for each input variable will be necessary.

Consider the two following relationships, which are valid for two normal random variables, *x1* and *x2* [45]:

$$\begin{aligned} \mu_{x1|x2} &= \mu_{x1} + \rho_{12}\left(\frac{\sigma_{x1}}{\sigma_{x2}}\right)(x2 - \mu_{x2}) \\ \sigma^2_{x1|x2} &= \sigma^2_{x1}(1 - \rho^2_{12}) \end{aligned} \qquad (3.4)$$

Where $\mu_{x1|x2}$ and $\sigma^2_{x1|x2}$ are the conditional mean and variance of *x1* given *x2* = *X2*; μ_{x1}, μ_{x2}, σ^2_{x1} and σ^2_{x2} are the unconditional means and variances of *x1* and *x2*; and ρ_{12} is the correlation coefficient between *x1* and *x2*. If each the fuel price densities were normal, then the above equations would be sufficient to formulate any conditional density. However, the fuel price density functions are assumed to be triangular. Nonetheless, the above relationships can be used as a guide to transform unconditional triangular distributions into approximate triangular conditional distributions. Note what occurs to an unconditional normal distribution when Equation (3.4) is applied. The unconditional mean is shifted on the abscissa by a factor involving the unconditional variances, the correlation coefficient, etc. The unconditional variance is then scaled by the correlation coefficient. The same will be done to the unconditional triangular distribution.

Consider the following:

load ~ $N(\mu_L, \sigma^2_L)$

fuel ~ $Tri(f_L, f_M, f_H)$

$HL = \mu_L + 1.389\sigma_L$ -- High Load

where f_L, f_M and f_H are the low, most likely, and high values of the triangular fuel price distribution. Since the triangular distribution is described by its coordinates, shifting the distribution can be accomplished by adding the same shifting factor, as described before, to each low, most likely and high value of the unconditional distribution. The conditional expected value is shown as follows:

$$\mu_{f|HL} = \mu_f + \rho_{f,L}\left(\frac{\sigma_f}{\sigma_L}\right)(HL - \mu_L) \tag{3.5}$$

Due to the definition of HL, above, Equation (3.5) can be simplified.

$$\begin{aligned}
\mu_{f|HL} &= \mu_f + \rho_{f,L}\left(\frac{\sigma_f}{\sigma_L}\right)(\mu_L + 1.389\sigma_L - \mu_L) \\
&= \mu_f + 1.389\sigma_L\rho_{f,L}\left(\frac{\sigma_f}{\sigma_L}\right) \\
&= \mu_f + 1.389\sigma_f\rho_{f,L}
\end{aligned} \tag{3.6}$$

From Equation (3.6), it can be seen that shifting the unconditional fuel price mean by $1.389\rho_{f,L}\sigma_f$ produces the desired conditional mean. The same shifting effect can be obtained by adding this factor to each of the triangular distribution's coordinate values. The shifted distribution is as follows:

$$\begin{gathered}
fuel_{SHIFTED} \sim Tri[f_L + C, f_M + C, f_H + C] \\
C = 1.389\rho_{f,L}\sigma_f
\end{gathered} \tag{3.7}$$

The conditional variance is found by the following:

$$\sigma^2_{f|HL} = \sigma^2_f(1 - \rho^2_{f,L}) \tag{3.8}$$

The low and high coordinates of the shifted unconditional triangular distribution must be changed to reflect the new conditional variance. If the unconditional triangular density is not symmetric about the most likely value, efforts must be made to maintain the same skewed shape during the transformation. This can be accomplished by solving for the conditional coordinates, while ensuring that the ratio of the areas of the triangles of either side of the most likely value remain constant. This added constraint is reduced to the following:

$$r = \frac{f_M - f_L}{f_H - f_M} = \frac{f_{M|HL} - f_{L|HL}}{f_{H|HL} - f_{M|HL}} \tag{3.9}$$

where $fl_{L|HL}$, $fl_{M|HL}$ and $fl_{H|HL}$ are the conditional coordinates. The scaled triangular distribution is the following:

$$f|HL \sim Tri[f_{L|HL}, f_{M|HL}, f_{H|HL}]$$

$$\sim Tri\left[f_{M|HL} - r\sqrt{\frac{18\sigma^2_{f|HL}}{1+r+r^2}}\ ,\ f_{M|HL}\ ,\ f_{M|HL} + \sqrt{\frac{18\sigma^2_{f|HL}}{1+r+r^2}}\right] \tag{3.10}$$

where $f_{M|HL} = f_M + 1.389\rho_{f,L}\sigma_f$ and $\sigma^2_{f|HL}$ is the conditional variance. See Appendix 1 for the proof of the above relationship. The stratified sampling technique described earlier in the chapter can then be applied to the conditional distributions, derived from the above transformation. Figure 3.3 shows a graphical representation of the above method.

The next chapter, Chapter 4, will discuss the risk due to load forecast uncertainty alone. Chapter 4 applies the techniques described earlier in this chapter to formulate the short-term planning problem in terms of a suitable decision analysis problem that is feasibly analyzed. In Chapter 5, the method for considering the conditional relationships between the load and the fuel prices will be used when fuel price uncertainty is added to the risk problem. One more modeling concern is the number of simulations required to evaluate the multiple fuel risk model. As the number of fuels increase, the required number of simulations grow exponentially. A method for reducing the model order is presented in this dissertation, but is deferred until Chapter 5.

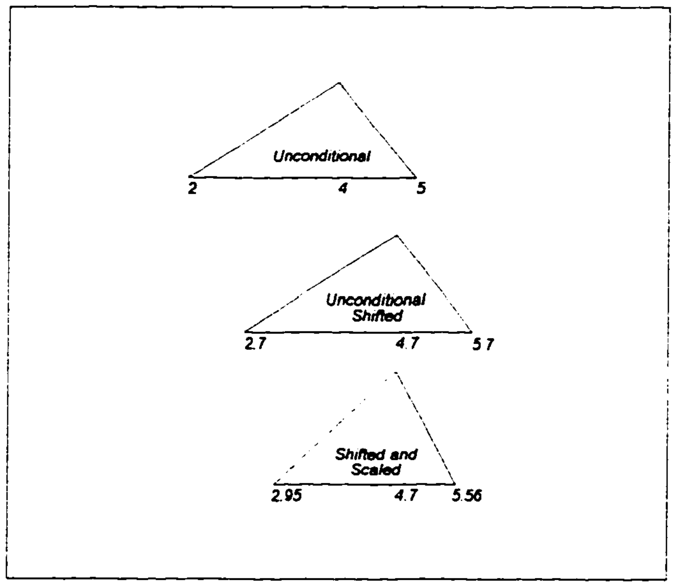

Figure 3.3 Triangular Transformation

CHAPTER 4

RISK DUE TO LOAD UNCERTAINTY

§ 4.1 Introduction

This chapter presents a methodology to analyze the risk of short-term power system operational planning in the presence of electrical load forecast uncertainty. As this methodology requires an estimate of the load forecast variance, the Bayesian load forecaster, presented in Chapter 2, is used in the practical implementation. The results are expressed as a function of forecast lead time from one to five days into the future, in terms of $/MWh. The risk due to load forecast uncertainty is based on the forecast variance, and is found by determining the expected cost of perfect information. The risk evaluation method is illustrated in a case study with utility derived system data and temperature forecast data from the National Weather Service.

As the energy market moves from a government regulated monopoly to a competitive free enterprise industry, the analysis of operational risk will become increasingly important. In the future, energy prices will most likely be determined contractually, instead of the current practice of public rate hearings. As a result, the contract lead time and contract price become important parameters in the short-term power system operational planning process.

Uncertainties impose additional risk in short-term planning, because of the unit commitment decision's large influence. Electrical load forecasts and spot fuel price estimates are two such uncertain quantities. For example, if the load forecast for the next week is in error, the electric utility runs the risk of committing either too much or too little capacity. In either case, the electric utility can incur additional costs. Furthermore, as the load forecast lead time increases, so does the forecast uncertainty. A similar situation exists when a significant proportion of the electrical utility's natural gas is supplied by the daily spot market. In the future free enterprise system, where prices may be quoted with a lead time of one hour to one month, new tools to analyze the risk due to load forecast and fuel price uncertainty will prove helpful in contract negotiation and energy price structuring. In this chapter the effects of load forecast uncertainty on short-term planning are explored; the effects of fuel price uncertainty are deferred until the next chapter. The impact of this planning variable is represented by the operational risk. Electric utilities employ various types of load forecasting models. The choice of forecast method generally depends on the study horizon lead time. When the study horizon lead time is short, conditional forecasts, which make use of weather forecast information to refine the load estimate, can be used to perform daily and weekly planning tasks. However, when the study horizon lead time is significant, unconditional forecasts, which only have the benefit of historical load or climatological data, may be the only feasible estimate. This is normally the case when the study horizon lead time is large enough to question the validity of the weather forecasts. Five days lead time is the limit for reliable temperature forecasts in the service area of our case study.

The aim of this chapter is to present a general procedure for evaluating the operational risk that is independent of the load forecast algorithm. The risk methodology demonstrated in this chapter is expressed as a function of forecast lead time. The load forecast variance, which differs from method to method, is an underlying quantity in the risk calculation, but can be generated from any load forecast source that estimates the forecast variance. Unfortunately, only a few load forecasting techniques yield a forecast distribution that includes a variance estimate. Bayesian estimation is one such method, and is the technique used in this chapter's illustrative case study.

Statistical decision theory and probabilistic sampling theory are used to evaluate the effects of load forecast uncertainty on short-term planning. Decision analysis has proven valuable, when evaluating system extension options [8-11]. In addition, the technique has been applied to capacity expansion studies with respect to annual load growth [1]. In this chapter, we take these concepts in a new direction and evaluate the risk of system planning in the short-term. From decision analysis, the evaluation of risk is straight forward. The Expected Cost Of Uncertainty (ECOU), or otherwise called the expected value of perfect information, serves very well as a risk indicator. The main task is to formulate and transform the short-term planning problem into a suitable structure for analysis using decision theory. Hence, decision variables must be defined and identified, states of nature have to be chosen and probabilities corresponding to the states of nature must be assigned.

For the purposes of this research, we define the decision variable to be the short-term unit commitment decision. The states of nature correspond to high, medium and low electrical load classes. In particular, efforts are made herein to describe the electrical load as discrete events derived from probability densities. The representations of the load is then assigned a likelihood of occurrence. This problem has a two-part solution. First, sampling theory is used to form discrete load categories via conditional expected values [34-40]. Secondly, the load forecast distribution profiles are scaled with the conditional expectations obtained from the stratification process. The load profiles are then used as input for a chronological production cost simulator to determine the unit commitment strategy and the production cost. Probabilities are available from the conditional expectations.

The rest of this chapter is structured as follows: Section 4.2 provides a concise problem statement. Section 4.3 describes the risk calculation methodology. This section contains four subsections. Subsection 4.3.1 covers electrical load forecasting. In this chapter, we use Bayesian forecasting to predict the conditional forecast distribution. Those aspects of probabilistic stratified sampling necessary to classify and categorize the electrical load are covered in Subsection 4.3.2. Subsection 4.3.3 details attributes of production cost simulation important to risk calculations. The fundamentals of decision analysis used to calculate short-term risk are described in Subsection 4.3.4. Section 4.4 is a practical implementation of the presented methodology, which provides an illustrative case study. Utility derived system data from a medium size electric company are used in Section 4.4, as well as, temperature forecast data from the National Weather Service. Section 4.5 concludes the chapter.

§ 4.2 Problem Statement

Given an acceptable set of historical load data, historical temperature data and short lead time temperature forecasts; the goal of this chapter is to determine the daily expected value of perfect information due to electrical load forecast uncertainty as a function of lead time. The study horizon for this research is limited to five days into the future. The results will provide a measure of uncertainty for each day of the study horizon. From the historical data and temperature forecasts, one through five day ahead load forecast distributions are generated, which, by means of sampling theory, are transformed into three individual events, denoted as high, medium and low load profiles.

Using decision analysis, where the decision variable is the unit commitment decision, we find the expected value of both nature's decision tree and the clairvoyant's decision tree. Nature's tree represents decision making under uncertainty and the clairvoyant's tree emulates decision making, when the future is known with 100% certainty. The difference in expected cost between the two scenarios yields the expected cost of load forecast uncertainty [7]. By using forecast distributions with varying study horizons, the expected cost of uncertainty can be expressed as a function of lead time, which can be a useful tool in short-term power system planning.

§ 4.2.1 Assumptions

Fundamental to this analysis are several points. First is that some method of determining the optimal unit commitment strategy and short-term production cost must be available. *Scheduler*, a resource allocation and scheduling tool developed by Power Costs, Inc. of Norman, Oklahoma, is used in this research. Although the determination of the expected cost of uncertainty is independent of the forecast method, the electrical load forecast must take the form of a time series profile suitable for use with a chronological production cost simulator.

Second is that the solution of the decision problem in terms of the strategy, which corresponds to the smallest expected cost, should not be confused with the solution to the stochastic unit commitment problem. Although both the stochastic unit commitment problem and the risk problem can be modeled using the same decision tree, the emphasis of our research rests on the derived expected cost while the solution to the stochastic unit commitment problem emphasizes the strategy that yields the minimum expected cost. Allowing the load to take on uniformly high, medium and low values is an approximation, and not a true representation of all possible load scenarios. The strategy corresponding to the solution of the decision problem will not, in general, correspond to the solution of the stochastic unit commitment problem. While the two strategies could be the same, it is not true in every case.

§ 4.3 Methodology

Probabilistic sampling theory will be used later to represent the load as a set of individual events. Necessary for that process is a load forecast distribution profile, which includes both mean and standard deviation. The following section will discuss, briefly, a method to generate distribution profiles.

§ 4.3.1 Forecast Distributions

Conditional load forecasts are those electrical load forecasts in which weather information can be used to refine the prediction. Many load forecast procedures are in use today. Among them are: artificial neural network algorithms, time series analysis, least squares methods, and state space approaches [26]. Some forecast techniques give only a point estimate and not a forecast distribution as required by this research. However, Bayesian estimation, which is in the state space category, does provide the necessary forecast distribution.

Bayesian estimation is a recursive prediction algorithm that uses an *a priori* distribution and sampled data to update to an *a posteriori* distribution for the next time increment. A dynamic linear model (DLM) is used in conjunction with this updating scheme to describe the output signal characteristics. Trends and periodic behavior are easily represented by the dynamic linear model in terms of the state space description. The DLM also has the capability to reflect the impact of explanatory variables on the forecast. These exogenous inputs, in the case of load forecasting, take the form of current weather forecast information. For a thorough treatment of the theoretical aspects of Bayesian estimation and dynamic linear models, see Chapter 2 and Reference [30-33].

Since the goal of this research is to capture the effects of load forecast uncertainty in short-term planning, it is important to note that a sizable portion of the load forecast uncertainty is due to a lack of accuracy in the weather forecast. For our particular case, temperature forecast errors can double the load forecast errors. Figures 2.9 - 2.12 in Chapter 2 are plots of the seasonal weather forecast standard deviations as a function of forecast lead time from several forecast sources within the service area used in the illustrative case study.

§ 4.3.2 Probabilistic Stratified Sampling

In the following section, aspects of probabilistic stratified sampling will be discussed and used to classify and categorize the electrical load. An important assumption this research makes is that the load at any given hour can be considered a normal random variable [42-44]. This is a reasonable assumption because the system load is actually an aggregation of all end users connected to the grid. If each customer is considered a random variable, then their summed loads will approximate a normal distribution by the Central Limit Theorem [44].

From the concepts of sampling theory, a distribution can be partitioned into a finite number of strata, where each stratum mean is a representation of all members in the stratum [34]. By choosing the strata wisely, sampling can be reduced to save effort. Since there are infinitely many values the electrical load can assume for a given hour, the goal is to represent the electrical load, described as a normal distribution, by the conditional expected value of well chosen partitions within the distribution. In so doing, not only does this discretize the load into a finite number of events, but also assigns probabilities to each event. Since forecast distribution profiles will be available, the process can be implemented for each hour of the load forecast profile. See Chapter 3 for further details.

However, in Chapter 3 the selection of strata and partition boundaries are shown only for the standard normal distribution. In fact, this process must be repeated for each hour of the forecast distribution profile. Since the load is normal, there are two short cuts that reduce the calculation workload. Any normal distribution can be represented in scaled terms of the standard normal distribution. Therefore, the conditional expectation of the standard normal curve, stratified as previously discussed, need to be found only once. The load profiles are found by scaling the hourly and standard deviation profiles. Furthermore, the standard normal distribution is symmetric about the ordinate so the conditional expectation of the medium stratum is zero and the low conditional expectation is opposite in sign from that of the high partition. Hence, only one conditional expectation need be found. By using the equation below, high, medium and low load profiles can be generated with the load forecast and standard deviation profiles.

$$
\begin{aligned}
l_L(j,k) &= \hat{l}(j,k) - 1.389\hat{\sigma}(j,k) \\
l_M(j,k) &= \hat{l}(j,k) \\
l_H(j,k) &= \hat{l}(j,k) + 1.389\hat{\sigma}(j,k)
\end{aligned}
\tag{4.1}
$$

Where l(j,k) is the appropriate load of j^{th} day and the k^{th} hour. Note that the above equations are independent of the forecast method.

§ 4.3.3 Production Cost Simulators

To effectively plan future courses of action, an essential tool for the system planner is the production cost simulator; a computer simulation package that determines unit commitment strategies and the minimum expected production cost, given load profiles, fuel price estimates and other system characteristics. There are two predominant methods of simulating power system production and activity. The Beleriaux-Booth method provides an analytical solution, where system demands and unit reliability characteristics are represented by probability distribution functions [46,47]. Since this method is based on the load duration curve, temporal information is lost. The chronological production cost simulator is the other method of evaluating power system planning options and strategies. This method requires the use of chronological load profiles as inputs and performs statistically based Monte Carlo simulations, which consider unit outage characteristics.

This research uses a chronological production cost simulator, because temporal information is important. Since the study horizon for this research is limited to short-term, where obtaining the unit commitment strategy is the main goal, unit outage characteristics will be ignored. *Scheduler*, a product of Power Costs, Inc. of Norman, Oklahoma is the chronological production cost simulator used in this research, because it provides all of the necessary tools of analysis. Not only does the program have the ability to simulate both fixed and flexible unit commitment strategies, but *Scheduler* also has the unique characteristic of using a dynamically linked library (DLL) as the main optimization engine, separate from the graphical user interface. This allows the user to run automated multiple simulations by calling the DLL from an external program, which is an essential requirement of this type of research.

Once the load profiles have been categorized and the high, medium and low load profiles have been generated, the profiles are used as input load files for a short-term deterministic production cost simulation, in order to determine associated cost and unit commitment strategy. It is important to note that the production cost simulator program must have the ability to evaluate production costs using both fixed and optimizing unit commitment strategies. Moreover, it is helpful, but not necessary, to have macro programing features in order to perform automated multiple simulation runs. The reasons for these requirements will become apparent in the next section.

§ 4.3.4 Expected Cost of Perfect Information

The goal of this chapter is to find the expected cost of uncertainty due to electrical load forecast variation. This is easily accomplished using the concepts of decision analysis. In its basic form, decision analysis models the available options in the form of a decision tree with decision nodes representing events that are under the control of the decision maker and chance nodes representing uncertain events outside the control of the decision maker. Assessed probabilities for each state of nature must be assigned, as well, as values for each outcome represented by terminal nodes [8]. See [7] for a thorough treatment of the subject.

Figure 4.1, on the following page, is the decision tree that corresponds to the standard short-term power system operational planning problem with electrical load forecast uncertainty. That is, a short-term unit commitment strategy is developed, deterministically, using one to five day load forecast profiles. The decision variable is the unit commitment decision and is denoted by the square decision node in the figure. HUC, MUC and LUC are different unit commitment strategies, where HUC is the unit commitment strategy of the high load scenario. MUC and LUC are defined similarly for medium and low load situations. The unit commitment decision is made prior to the week in question, and for the purposes of comparison, rescheduling is not considered.

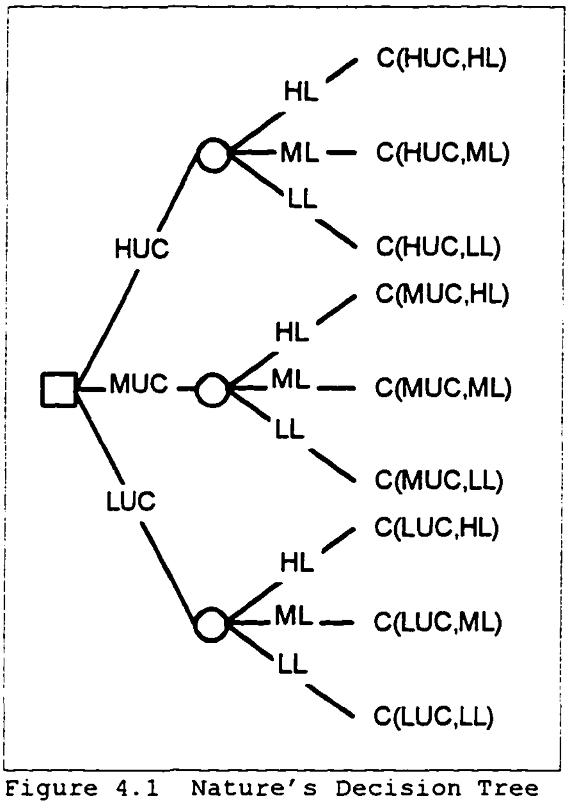

Figure 4.1 Nature's Decision Tree

Tied to each circular chance node are the states of nature. For our case we have high, medium and low load representations denoted by HL, ML and LL in the figure. Assigned to each state of nature are probabilities P(HL), P(ML) and P(LL) (Not shown in the figure). The terminal nodes of the tree, outboard right of the figure, correspond to cost functions. For example, C(HUC,ML) is the weekly operational cost if the high unit commitment strategy was selected and the load turned out to be medium. Each of the nine possible outcomes are defined in this fashion.

For the purposes of this research, the cost functionals are determined by a chronological production cost simulator run under the appropriate set of circumstances. The decision tree can be evaluated by a method called averaging out and folding back [7]. The quantity that is determined from the tree's evaluation is the expected value. The decision tree in the above figure is sometimes called nature's decision tree. The calculation of the expected value of nature's tree, denoted EV(N), is shown as follows:

$$EV(N) = \min_{uc}\left\{\begin{array}{l} P(HL)\cdot C(HUC,HL) + P(ML)\cdot C(HUC,ML) + P(LL)\cdot C(HUC,LL), \\ P(HL)\cdot C(MUC,HL) + P(ML)\cdot C(MUC,ML) + P(LL)\cdot C(MUC,LL), \\ P(HL)\cdot C(LUC,HL) + P(ML)\cdot C(LUC,ML) + P(LL)\cdot C(LUC,LL) \end{array}\right\} \tag{4.2}$$

Another decision tree configuration is called the clairvoyant's tree, and is shown on the following page. This is the case where the decision maker has perfect knowledge of the future. Notice in the figure that only one terminal node is tied to each decision node. Presumably, the decision maker will chose to implement the optimal unit commitment strategy if the future is known with complete certainty.

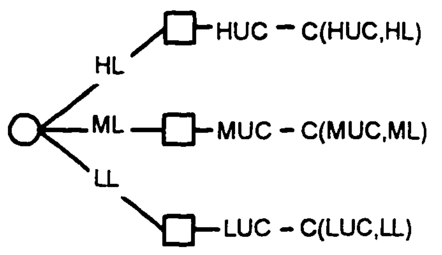

Figure 4.2 Clairvoyant's Decision Tree

The expected value of the clairvoyant's tree, denoted EV(C), is shown in Equation (4.3).

$$EV(C) = P(HL)\cdot C(HUC,HL) + P(ML)\cdot C(MUC,ML) + P(LL)\cdot C(LUC,LL) \tag{4.3}$$

The difference between EV(N) and EV(C) is called the expected cost of uncertainty, denoted ECOU, and is shown in Equation (4.4) [7].

$$ECOU = EV(N) - EV(C) \tag{4.4}$$

The ECOU is the most the decision maker should pay for load forecasting services, whether it be in engineering costs or to a weather forecast service. The ECOU is also the risk due to load forecast uncertainty. In the decision analysis sense, the ECOU due to load forecast inaccuracies is a good indicator of the danger present in planning under longer forecast lead times.

There are two possible ways to find the daily ECOU as a function of forecast lead time. One method is the perform one day unit commitments for each day of the study horizon and find the corresponding ECOU of each day. A problem with this method is that the one day unit commitment does not capture startup and shutdown times necessary for plant operation. These factors are an essential part of realistic planning and should not be overlooked. Instead, we first find the ECOU for k days by considering the k day unit commitment strategy starting from the present to k days into the future.

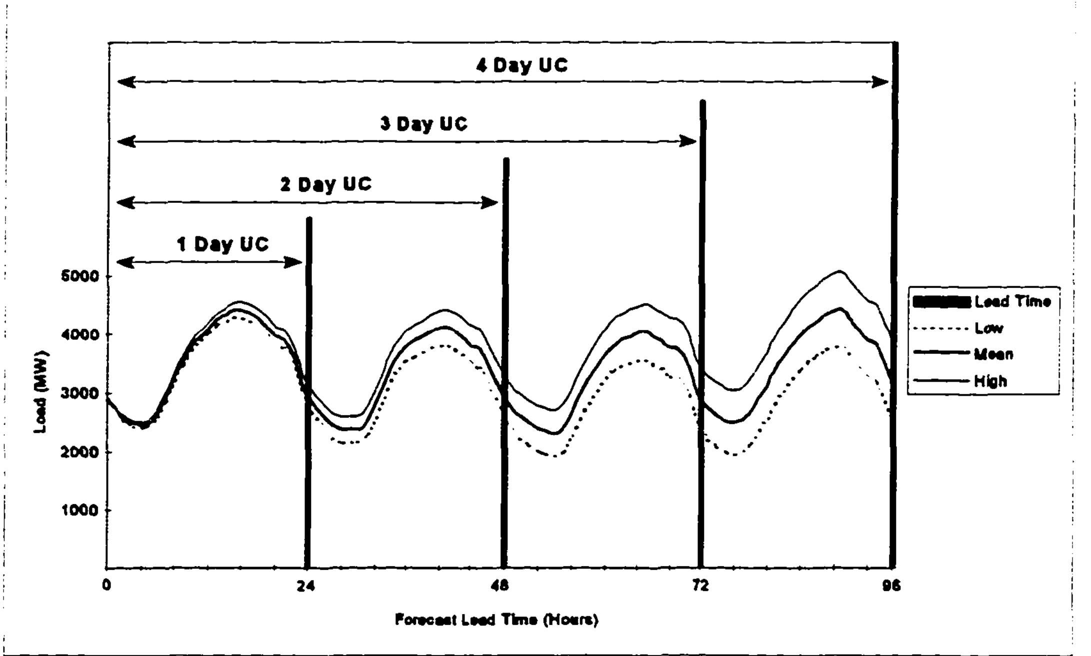

Figure 4.3 Period ECOU as a Function of Lead Time

In this way, startup and shutdown times are reflected in the ECOU, and a more realistic representation of the passage of time is presented. See Figure 4.3 for a graphical view. We define the ECOU for one day, say day k, to be the difference between the period ECOU(k) and the period ECOU($k-1$), where ECOU(j) denotes the expected cost of uncertainty when the decision variable is the unit commitment decision with a study horizon ranging from the present to j days into the future. Note that ECOU(0) = 0.

Consider the following: A power supplier has the responsibility for quoting daily energy prices for the next week. The guideline that the power supplier uses for setting the daily energy price for the k^{th} day into the future is defined to be the aggregation of the expected production cost, the daily ECOU of day k and a thirty percent profit margin. Given the daily expected costs which are based on a load forecast and obtained by performing a short-term deterministic unit commitment simulation, the price for the first day is the expected cost plus ECOU(1) plus thirty percent. To price day two, however, we cannot simply add the expected cost and ECOU(2), because ECOU(2) contains the cost of uncertainty of day one as well. Instead, we need only add in the cost due to uncertainty not recovered in the price of day one; that is, ECOU(2) - ECOU(1). Therefore, the price of day two becomes the expected cost of day two, plus ECOU(2) - ECOU(1) plus a profit margin of thirty percent; this method is defined similarly for k days into the future. For our simplified pricing task we have the following:

$$P(k) = (1+PM)(C(k) + ECOU(k) - ECOU(k-1)) \qquad (4.5)$$

Where $P(k)$ is the price of day k, PM is the profit margin, $C(k)$ is the expected production cost for day k and ECOU(k) is the expected cost of uncertainty whose decision variable is the k day unit commitment.

It is important to note that the high, medium and low unit commitments that have been discussed previously are not the only solutions. Others are certain to exist, and one may be optimal. The selection of the true optimal unit commitment decision in the presence of infinitely many loads is better known as the stochastic unit commitment problem. This research does not attempt to solve this problem. Since we may not have identified the true optimal unit commitment, the ECOU, found from the suggested methodology, may be slightly higher. This implies that Equation (4.4) represents an upper bound of the true ECOU. Practical experience has shown that the optimal ECOU is roughly fifteen percent lower. However, this accounts for less than one percent of the production cost for the period in question.

§ 4.4 Illustrative Case Study

The case study performed used data from a medium size electric utility whose annual peak is around 5,000 MW and whose annual energy is roughly 23,000 GWH. See Appendix 2 for a detailed list of generator data. Figures 4.4 - 4.7, at the end of this chapter, show plots of the expected cost of uncertainty due to load forecast uncertainty as a function of lead time, in $/MWh.

§ 4.5 Conclusion

Based on the above results we can see that the period expected cost of uncertainty (The ECOU from the present to *k* days into the future) exhibits reasonable behavior. That is, the period ECOU increases with lead time. However, the daily ECOU does not. Such is the case when the period ECOU increases by a smaller amount with the passage of time. Is this reasonable? From the plots of the temperature forecast variance, the reader should note that the daily cost of uncertainty tracks the temperature forecast standard deviation. Winter is the best seasonal example. A spike in the daily ECOU occurs on day three, similar to the plot of the temperature standard deviation.

The above analysis considers risk as a function of lead time due to electrical load forecast uncertainty. However, the load forecast in not the only uncertain variable in short-term power system planning that has a significant impact on the unit commitment strategy. Spot market fuel prices also affect the short-term planning process. In the next chapter, fuel price uncertainty is included in the risk analysis. Chapter 5 also considers the effects of conditional relationships between input variables, such as the interdependencies between load and fuel prices, and presents a methodology to reduce the model order by approximation to decrease computer simulation time.

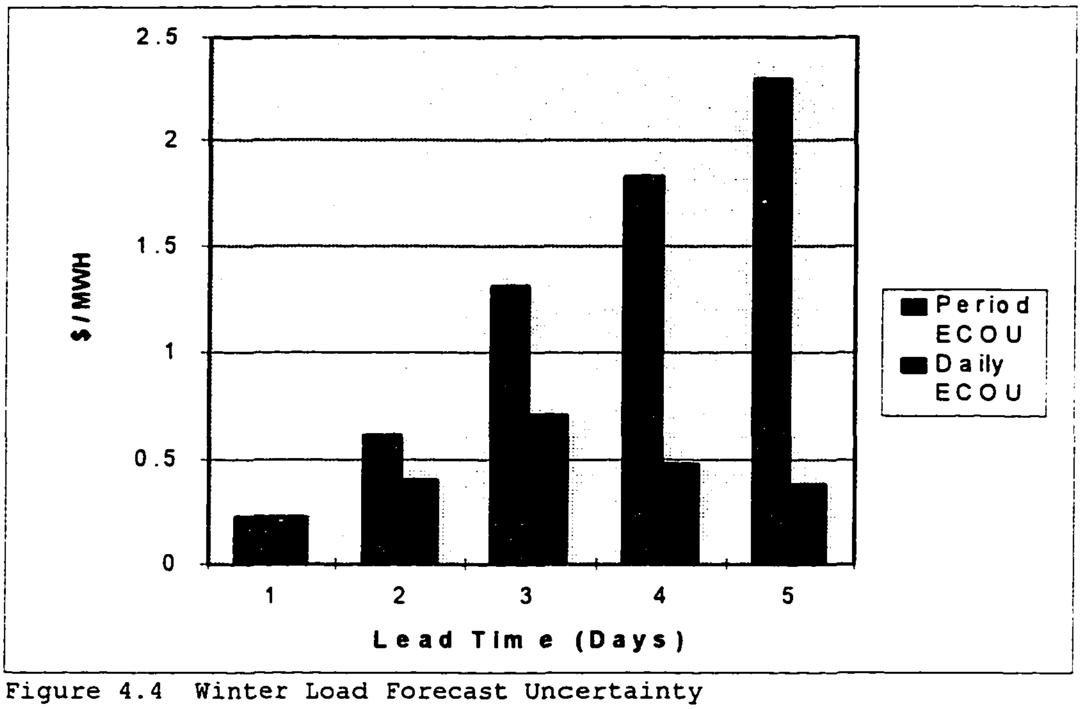

Figure 4.4 Winter Load Forecast Uncertainty

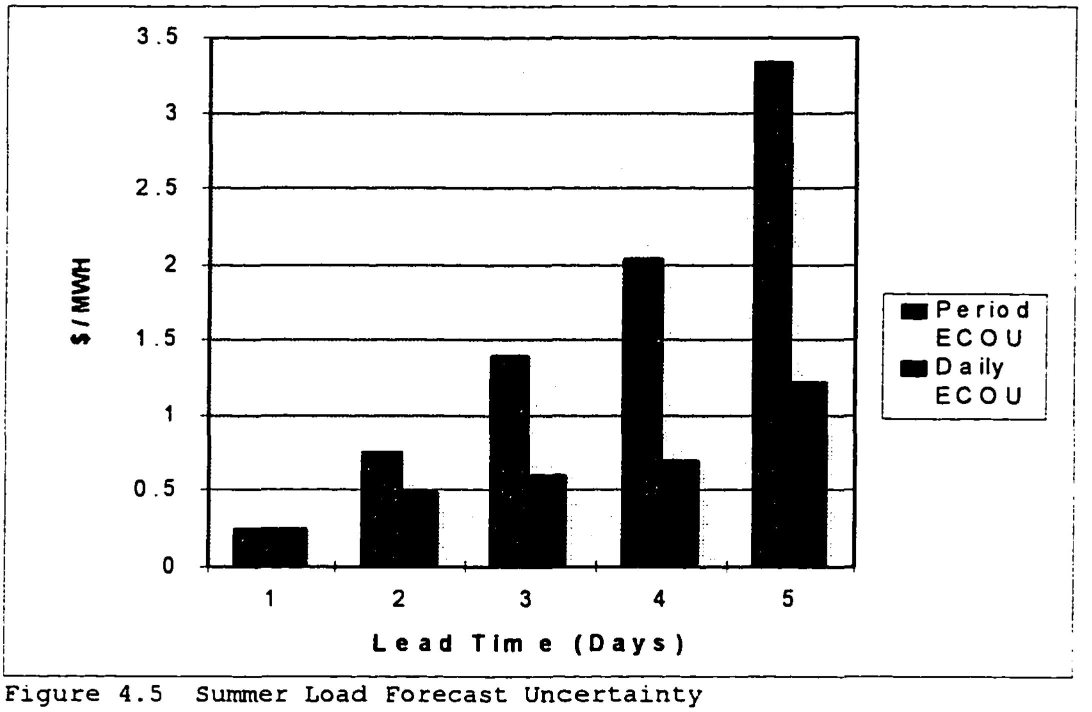

Figure 4.5 Summer Load Forecast Uncertainty

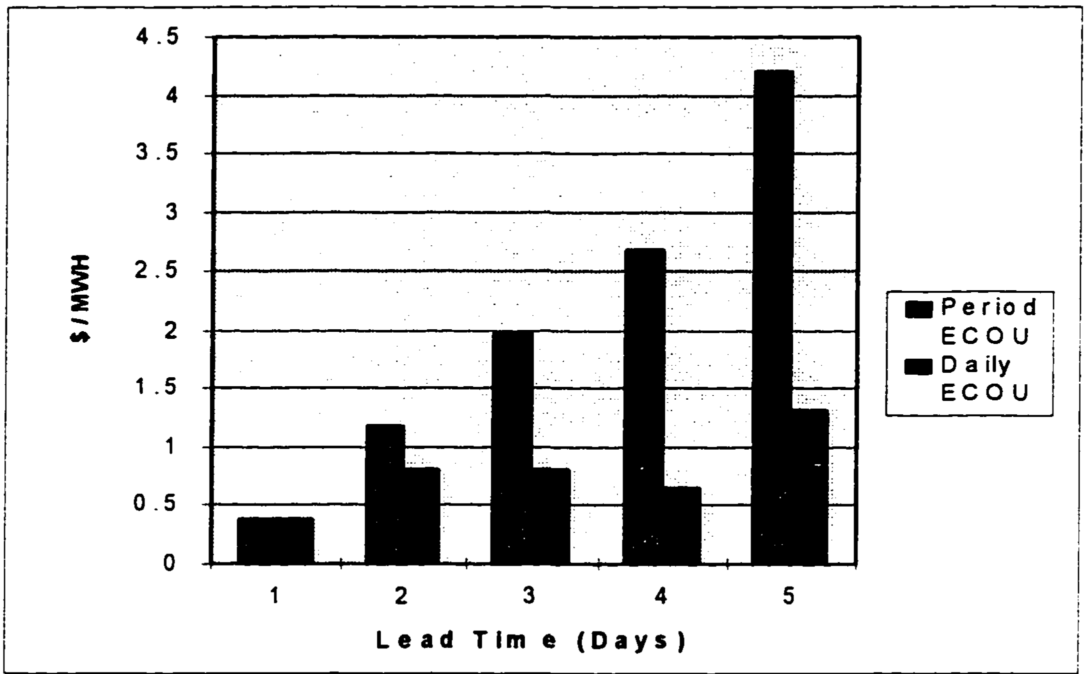

Figure 4.6 Spring Load Forecast Uncertainty

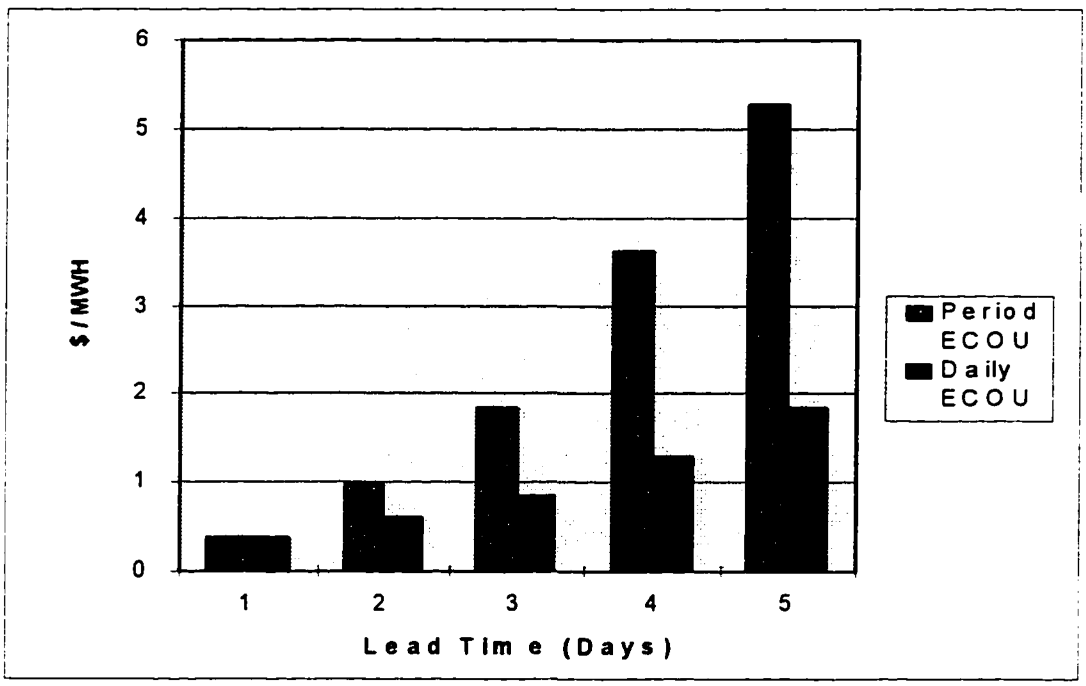

Figure 4.7 Autumn Load Forecast Uncertainty

CHAPTER 5

RISK DUE TO LOAD AND FUEL PRICE UNCERTAINTY

§ 5.1 Introduction

This chapter presents a methodology to analyze the risk of short-term power system operational planning in the presence of electrical load forecast and fuel price uncertainty. As the methodology will require an estimate of the load forecast variance, the Bayesian load forecaster from Chapter 2 will be used. In addition, a decision tree structure, similar to that of Chapter 4, will be the starting point for including fuel price uncertainty in the risk model. Due to the interdependencies that exist between the load and fuel prices, the results form Chapter 3 will be used to account for these relationships. The illustrative case study's results are expressed as a function of forecast lead time from one to five days into the future, in terms of \$/MWh. The risk evaluation method that is implemented in this chapter's illustrative case is taken from the utility derived data set and National Weather Service temperature forecast data that have been used throughout this dissertation.

As shown in the previous chapter, load forecast uncertainty can have a large impact on the expected operational cost, in some cases the ECOU was almost 3 $/MWh. However, the electrical load is not the only variable in the short-term power system planning process that has a degree of uncertainty linked to it. Fuel prices also tend to fluctuate throughout the year, especially spot market natural gas. As a result, fuel price uncertainties should not be overlooked, when evaluating risk in short-term planning. As will be shown in the following pages, the risk due to fuel price uncertainty is significant, as well.

However, the dimensionality of the decision tree cannot just be increased without further analysis. When evaluating risk models with multiple input variables, where dependencies may exist, conditional densities should be used in place of unconditional densities. Furthermore, the model order quickly grows very large as the number of input variables increase. Exhaustive simulations become impractical, when three or more random inputs are considered. It is common for electric utilities to have five or more fuel suppliers; because of this, additional efforts are made herein to reduce the model order by using a reasonable approximation.

The rest of this chapter will be structured as follows: Section 5.2 provides a concise problem statement. Section 5.3 uses results from Chapter 3 to represent dependent inputs. Section 5.4 details a method to reduce the over all model order and the number of simulations required to evaluate the risk model with multiple inputs. Section 5.5 covers the formulation of the new decision tree that is to be evaluated by the techniques described in Chapter 4. Section 5.6 is the illustrative case study which make use of the standard data set found in the Appendix 2.

§ 5.2 Problem Statement

Given an acceptable set of historical load data, historical temperature data, fuel price densities, fuel price and load correlations, and short lead time temperature forecasts; the goal of this chapter is to determine the daily expected cost of uncertainty, due to load forecast and fuel price uncertainty, as a function of lead time. The study horizon for the research is limited to five days into the future. The results will provide a measure of uncertainty for each day of the study horizon.

From the historical data and temperature forecasts, one through five day ahead load forecast distributions are generated, which by means of sampling theory, are transformed into three individual events, denoted high, medium and low. These are the same load profiles used in the previous chapter. Using the unconditional fuel price densities and a set of correlation coefficients, conditional densities for the fuel price distributions are created. These densities dictate the formation of individual fuel price events using sampling theory, as well.

Using decision analysis, where the decision variable is the unit commitment decision, the expected value of both nature's decision tree and the clairvoyant's decision tree are found. Nature's tree represents decision making under uncertainty, where both the load and fuel prices are unknown to the decision maker. The clairvoyant's tree represents decision making with complete knowledge of the future. The difference in expected costs between the two scenarios yields the expected cost of load forecast and fuel price uncertainty [7]. By using forecast distributions with varying study horizons, the expected cost of uncertainty can be expressed as a function of lead time, which can be a useful tool in short-term power system planning.

§ 5.2.1 Assumptions

Fundamental to this analysis are two points. First is that the assumptions made in the previous chapter are also valid here. That is, the load is assumed to be well behaved, and the stochastic unit commitment problem is not considered. Secondly, fuel prices are given as unconditional triangular probability density functions, and these density functions stay constant over the entire study horizon, which is different from the load forecast distribution profiles.

§ 5.3 Dependent Input Variables

Using the formulae describing the conditional normal density function as a guideline, found in Chapter 3, a method for approximating conditional triangular densities was developed, to account for interdependencies that may exist between input variables. The triangular fuel price distributions will be adjusted using Equations (3.9) and (3.10) in Chapter 3. This implies that the unconditional triangular distributions must be known, as well as, the correlation coefficients. Note that if a correlation coefficient is zero, then the conditional density defaults to the unconditional density. However, a zero correlation coefficient, in general, does not necessarily imply independence. When multiple fuels are considered the matrix form of Equation (3.4), in Chapter 3, is the most efficient method of generating the conditional means and variances [44]. Consider the following:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \quad X_1 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad X_2 = \begin{bmatrix} x_{k+1} \\ x_{k+2} \\ \vdots \\ x_m \end{bmatrix} \quad \mu_X = \begin{bmatrix} \mu_{x1} \\ \mu_{x2} \end{bmatrix} \quad \Sigma_X = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{5.1}$$

Let X be a m-dimensional multivariate normal random vector and let X_1 and X_2 be two sub random vectors of X, such that the augmentation of X_1 and X_2 is X. μ_{x1} and μ_{x2} denote the unconditional means of X_1 and X_2. Σ_{11}, Σ_{22}, Σ_{12} and Σ_{21} are the unconditional covariance matrices of X_1 and X_2. Note that μ_{x1} is $(k \times 1)$ and Σ_{11} is $(k \times k)$. The following can be shown [44]:

$$
\begin{aligned}
\mu_{X_1|X_2} &= \mu_{X_1} + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_{X_2}) \\
\Sigma_{X_1|X_2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}
\end{aligned}
\tag{5.2}
$$

Where $\mu_{x1|x2}$ and $\Sigma_{x1|x2}$ are the conditional mean and covariance of *X1* given X2. To use Equation (5.2), set $k = 1$. To find the i^{th} conditional fuel distribution, the vector X can be restructured such that the i^{th} fuel in question is the first element. The covariance matrices are formed using the correlation coefficients. The most difficult aspect of this method is the matrix inverse in Equation (5.2). However, because the covariance matrix is usually positive definite, an inverse should exist. As to its calculation, if the numerical ranges of the elements are great, numerical instability may occur. In that case, a linear transformation can be applied to bring as many elements of the covariance matrix to a similar order as possible, and thus limiting the round-off and truncation error.

§ 5.4 Model Order

Another consideration for multiple input cases is the model order. For n fuels and the load forecast, the total number of simulations required to evaluate the expected cost of uncertainty is $3^{2(n+1)}$. This relationship is based on three states of nature per fuel plus three for the electrical load. The squared term is due to the need for evaluating the production cost of each course of action together with all states of nature. If an electric utility has five fuel suppliers, 3^{12} or 531441 simulations are necessary. Two methods can be sought to reduce the number of simulations.

First, the overall model order can be reduced by neglecting those fuels which either have an insignificant contribution to the total fuel cost or those fuels which remain predominantly invariant over the study horizon. The criterion for making these types of judgements is engineering experience. The second method involves a sample reduction technique, similar to the central composite design used for finding regression surfaces [35]. Consider Figure 5.1, with one fuel and the electrical load, on the following page.

By sampling only the extremes and the most likely cases, denoted by the X's in the figure, the number of possible states of nature decrease. As the courses of action correspond to the total enumerated number of states of nature, the strategies are reduced as well! For example, both the courses of action and the states of nature in the 2-d case are reduced from 9 to 5. Therefore, simulations for the above 2-d case are reduced from 81 to 25. Consider the two fuel case in Figure 5.2.

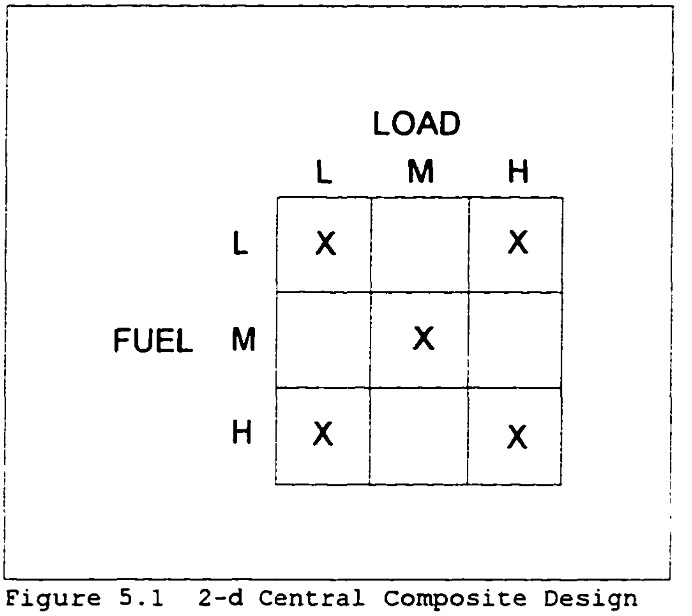

Figure 5.1 2-d Central Composite Design

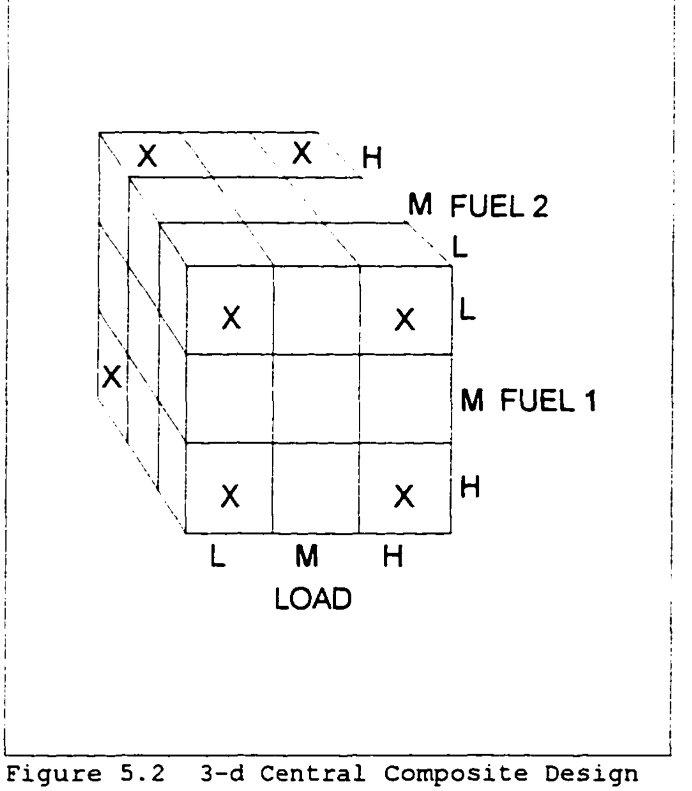

Figure 5.2 3-d Central Composite Design

Notice that state HHH (back right) and the case MMM (center) are not shown. With this design, the total number of simulations is reduced from 729 to 81, because the courses of action and states of nature have be reduced from 27 to 9. For the general reduced order sampling case with *n* fuels, $(2^{(n+1)}+1)^2$ simulations are required. If an electric utility has five fuel suppliers, the number of simulations is reduced from 3^{12} to 4225, which is a 99% reduction in simulation time. Using today's modern desktop computer, Pentium 166 MHZ with 32MB Ram, the simulation time is reduced from 45 days to 12 hours, (10 sec/run). Note that if the simulation time exceeds the study horizon, the model is probably too large.

This is a significant improvement. The probabilities of each state need to be adjusted as well. The center state, MM or MMM in the above figures, will assigned the weight of 0.6. The remaining probabilities corresponding to the extreme cases will carry a weight of $0.4/(2^{(n+1)})$, for *n* fuels, to ensure that the probabilities sum to unity. This provides a reasonable approximation. Table 5.1 is a tabulation of such an approximation. The expected production cost, found by using the reduced order approximation, is within ½% of that found using the full order model. The production cost variance was found to be within 1% of that found with the full order model. Other probability assignments are possible. For example, the conditional probabilities of each sampled state can be used. As they do not sum to one, scaling should be performed. Although, this method provided adequate results for mean estimates, within 10%, this method was in error for variance estimates by a factor of two. As a result the first probability assignment scheme was implemented.

Production Cost ($)	Expected Cost ($)	Cost Variation (^{2})	Expected Cost Reduced Order ($)	Cost Variation Reduced Order (^{2})
1277224	10218	1.301E+11	63861	8.134E+11
2289594	54950	2.190E+11		
2858262	22866	4.811E+10	142913	3.007E+11
2871311	68911	1.428E+11		
5377091	387151	3.183E+08		
5938295	142519	9.456E+09		
3101282	24810	3.905E+10	155064	2.441E+11
7605397	182530	1.264E+11		
9166252	73330	1.189E+11	458313	7.433E+11
1417202	34013	3.638E+11		
2706795	194889	4.881E+11		
3498059	83953	7.885E+10		
3096933	222979	3.528E+11		
5835502	1260468	5.951E+10	3501301	1.653E+11
6643596	478339	1.279E+11		
3410314	81848	8.667E+10		
8188963	589605	5.965E+11		
9908103	237794	5.073E+11		
1581018	12648	1.113E+11	79051	6.955E+11
3103786	74491	1.169E+11		
4160836	33287	1.058E+10	208042	6.610E+10
3386922	81286	8.881E+10		
6310095	454327	7.193E+10		
7358918	176614	1.007E+11		
3919292	31354	1.549E+10	195965	9.679E+10
8746654	209920	2.834E+11		
10687224	85498	2.313E+11	534361	1.445E+12
	5310599	**4.526E+12**	**5338871**	**4.571E+12**

Table 5.1 Model Reduction Example

§ 5.5 Decision Analysis

The calculation of the expected cost of uncertainty is no different than in Chapter 4, when only load forecast uncertainty was considered. That is, the ECOU is still the expected value of the clairvoyant's case minus the expected cost of nature's case, minimized over all possible courses of action. The tree however, is somewhat larger. Added to each of the possible load states are the states of nature and courses of action that correspond to the varying fuel prices. Figure 5.3 shows one branch of nature's tree when two fuels are considered.

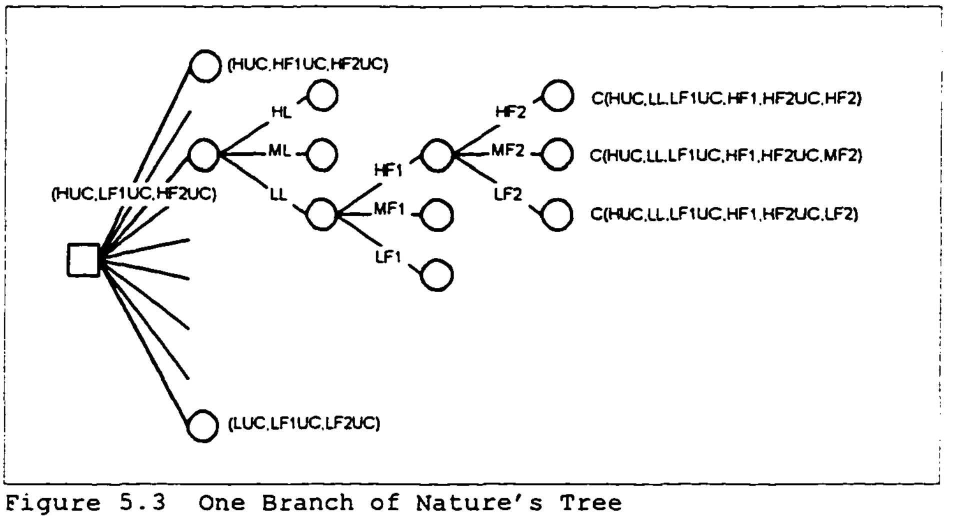

Figure 5.3 One Branch of Nature's Tree

This tree representes the case where the decision maker has uncertain knowledge of the future states of nature. C(HLUC,ML,LF1UC,MF1,HF2UC,HF2) is defined as the cost of selecting a unit commitment strategy for a high load, a low price for fuel #1 and a high price for fuel #2; given the load is medium, the price of fuel #1 is medium and the price of fuel #2 is high. The above figure is only one branch of nature's tree. The rest of the tree is drawn similarly. The expected value of nature's tree is the expected cost linked to the unit commitment strategy which yields the smallest expected cost when minimized over all admissible strategies. The idea is the same as Equation (4.2) in Chapter 4. In set notation we can write the following:

$$EV(N) = \min_{i \in \Lambda}[\bar{c}(i)] \tag{5.3}$$

where Λ is the set of all admissible courses of action and the expected production cost, denoted $\bar{c}(i)$, is defined as follows:

$$\bar{c}(i) = \sum_{j \in \Omega} p(j)\, c(i,j) \tag{5.4}$$

where Ω is the set of all viable states of nature, p(j) is the probability corresponding to the j^{th} state of nature and c(i,j) is the production cost resulting from the selection of the j^{th} course of action, when the i^{th} state of nature occurred. Note that the above formulae are valid for both the reduced order case, as well as the full order model. Only the contents of the set describing the states of nature and the courses of action change.

The clairvoyant's tree is the other decision tree, one branch of which is shown in Figure 5.4.

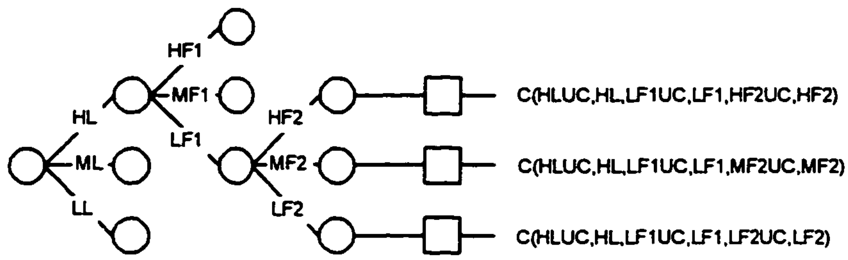

Figure 5.4 One Branch of the Clairvoyant's Tree

Again, the clairvoyant's tree represents that case where the decision maker has prefect knowledge of the future states of nature. C(HLUC,HL,LF1UC,LF1,HF2UC,HF2) is defined as the cost of selecting a unit commitment strategy for a high load, a low price for fuel #1 and a high price for fuel #2; given the load is high, the price of fuel #1 is low and the price of fuel #2 is high. This is the "optimal" case in the decision analysis sense, but not the stochastic unit commitment sense.

The expected value of the clairvoyant's tree, EV(C) as denoted in Chapter 4, is found by applying Equation 5.4 to the clairvoyant's tree. In this case however, there is only one logical course of action that corresponds to any one state of nature, and that is the unit commitment strategy that minimizes the production cost over the state of nature in question. When finding the expected cost of uncertainty for the reduced order case, only the center and the extremes are considered to be admissible courses of action; the same set also makes up the possible states of nature. The new probability assignments are used as well.

§ 5.6 Case Study

The standard system is used here in this case study, refer to Appendix 2 for more information. Assumptions made in this case study are that three of the five fuels represent so little of the overall study horizon cost (less than 2%) that they can be ignored. Only coal and spot natural gas will be considered. In addition, spinning reserve will remain constant over the study horizon, and rescheduling will not be performed. The following are the coal and gas price distributions:

Coal ~ Tri(0.600, 0.827, 1.050)
Gas ~ Tri(2.500, 2.760, 3.000)

The following are the correlation coefficients between the fuels and the load:

	Load	**Coal**	**Gas**
Load	1.0	0.8	0.7
Coal	0.8	1.0	0.5
Gas	0.7	0.5	1.0

Note that the correlation matrix should be positive definite. A sufficient condition for positive definiteness is that the matrix be symmetric, real and all eigenvalues must be positive [32]. In the above case, the eigenvalues are (0.1514,0.5078,2.4308). Table 5.2 shows the conditional fuel prices, which are generated using Equation (5.2). The states are defined such that position one is the state of the load, position two is the state of the coal price and position three is the state of the gas price.

State	Probability	Coal	Gas
LLL	0.0500	0.693	2.603
LLH	0.0500	0.699	2.770
LHL	0.0500	0.819	2.611
LHH	0.0500	0.825	2.778
HLL	0.0500	0.828	2.736
HLH	0.0500	0.834	2.899
HHL	0.0500	0.953	2.744
HHH	0.0500	0.959	2.907
MMM	0.6000	0.825	2.752

Table 5.2 Conditional Fuel Prices

For example, 0.828 is the price of coal, given the load is high and the price of gas is low. Note also that the load profiles are generated using Equation (4.1) from Chapter 4. Figures 5.5 - 5.8, in the following pages, show the period and daily ECOU, due to load forecast uncertainty and fuel price uncertainty, as a function of lead time. The method for evaluation is the same method developed in Chapter 4, Equation (4.5); however, the expected value of nature's tree and the clairvoyants tree are replace with the expressions developed in this chapter.

§ 5.7 Conclusion

Note that the plots have similar shapes to those of the previous chapter. However, the magnitude of the daily ECOU is much higher. It is easily seen that when fuel price uncertainty is included in the risk model, the expected cost of uncertainty increases dramatically. In the next chapter, the results of both Chapter 4 and 5 are compared. In addition, the following chapter will serve as a conclusion to this dissertation. A summary outline will detail the methods used to evaluate the risk due to load forecast and fuel price uncertainty, as well as, expand on recommendations for future research possibilities.

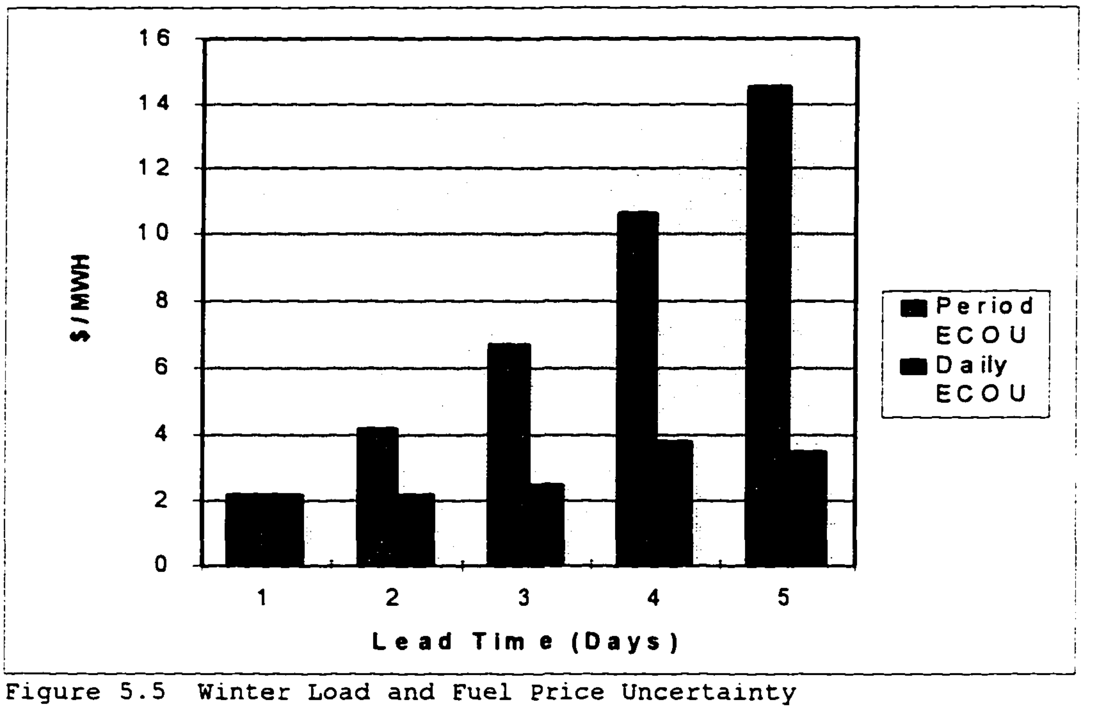

Figure 5.5 Winter Load and Fuel Price Uncertainty

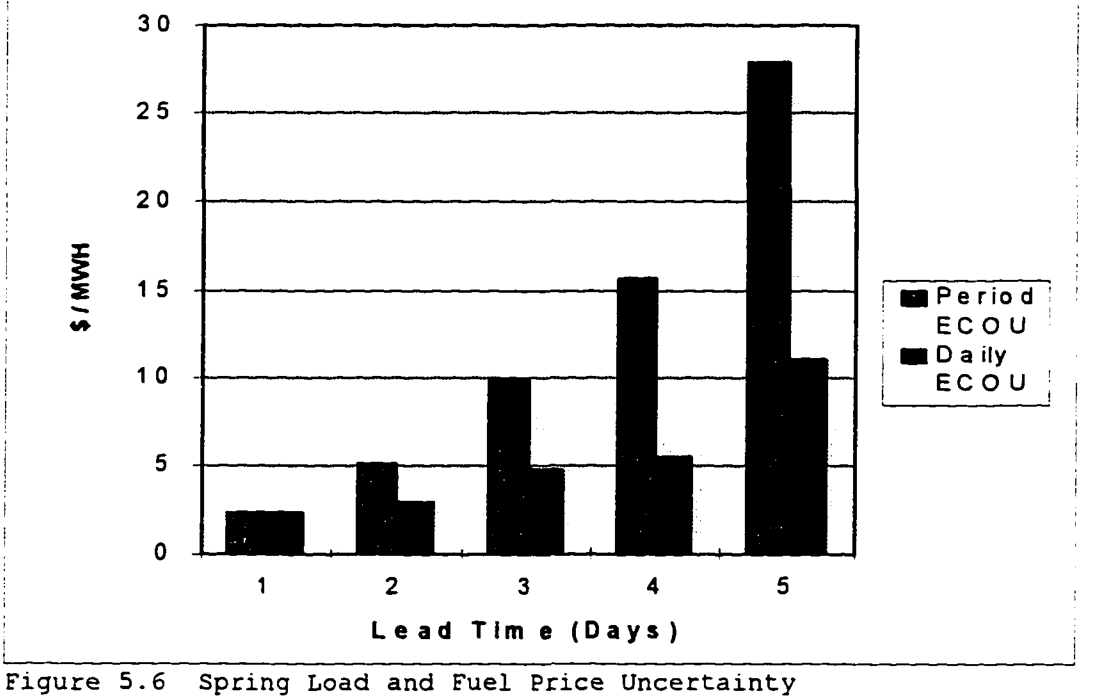

Figure 5.6 Spring Load and Fuel Price Uncertainty

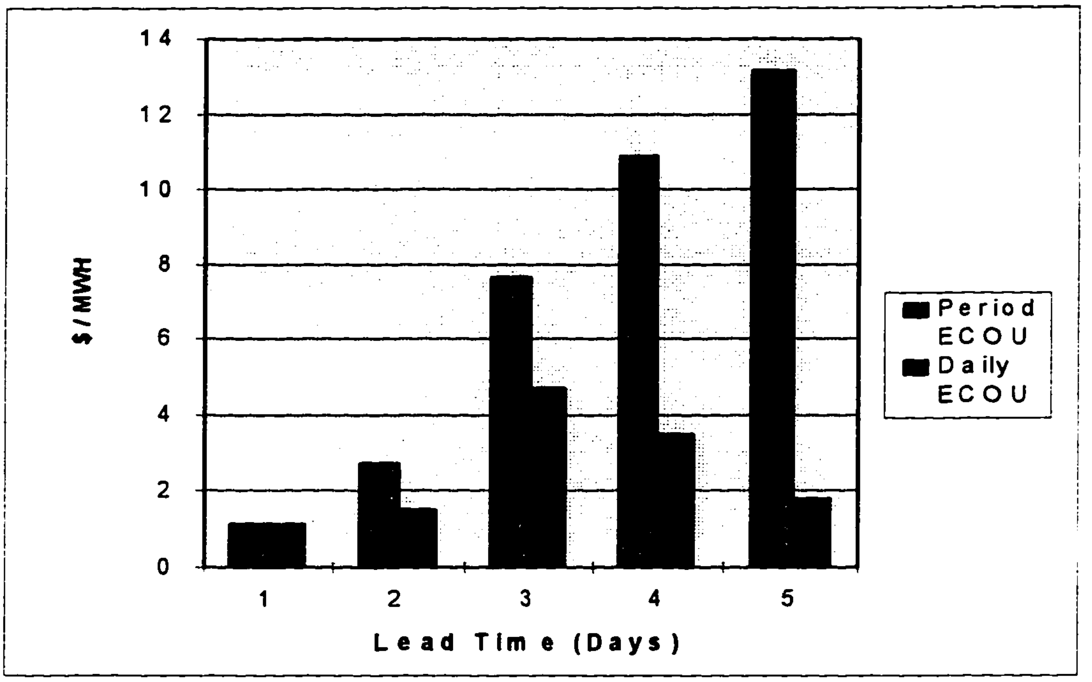

Figure 5.7 Summer Load and Fuel Price Uncertainty

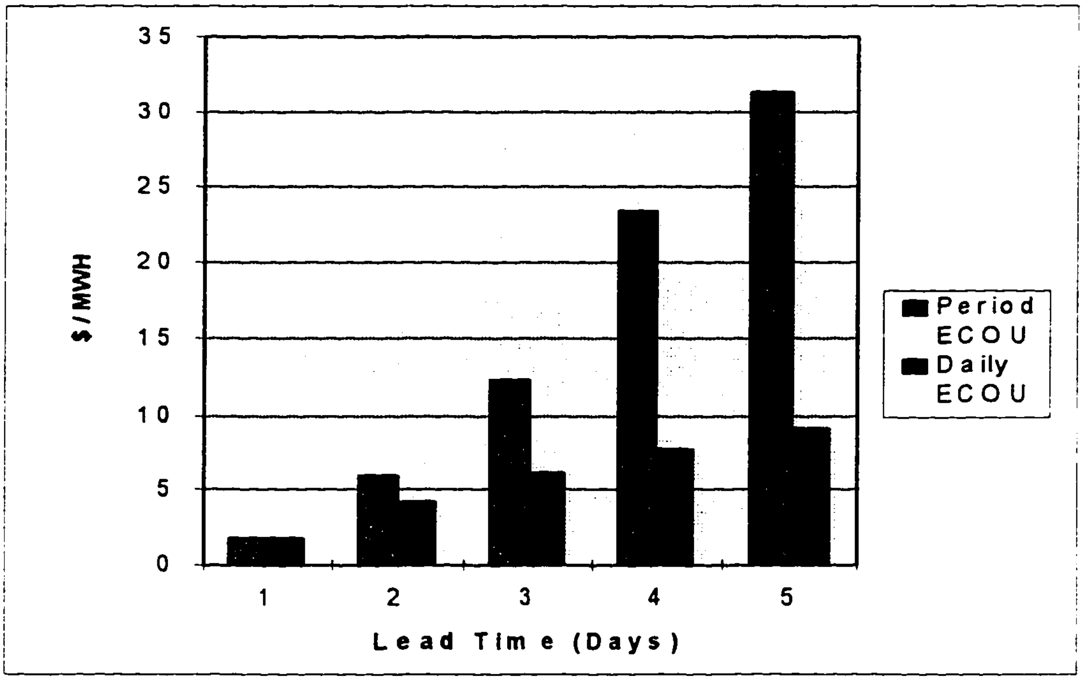

Figure 5.8 Autumn Load and Fuel Price Uncertainty

CHAPTER 6

CONCLUSION

§ 6.1 Introduction

The structure of this chapter will be as follows: first a discussion of the results from Chapters 4 and 5 is presented. Then a brief summary of the research reported herein is included. Lastly, a recommendation for further research efforts is provided.

§ 6.2 Summary of Results

In Chapters 4 and 5, methods for evaluating the risk due to load forecast uncertainty alone and load forecast together with fuel price uncertainty were presented. In Chapter 4, the electrical load forecast was considered to be the only random process, while all other planning variables were held constant.

The results of Chapter 4 indicate what effects the load forecasting error can have on the short-term planning process. The results were given in \$/MWh as a function of forecast lead time. Table 6.1 shows the tabulated results from the illustrative case study performed in Chapter 4.

Season	Lead Time (Days)	Total ECOU ($)	Differenced ECOU ($)	Daily Energy (MWh)	Period ECOU ($/MWh)	Daily ECOU ($/MWh)
Winter	1	11700	11700	53388	0.219	0.219
	2	34633	22933	56956	0.608	0.403
	3	75817	41184	57665	1.315	0.714
	4	102310	26493	55808	1.833	0.475
	5	122797	20487	53597	2.291	0.382
Spring	1	25025	25025	65566	0.382	0.382
	2	80005	54980	67938	1.178	0.809
	3	134619	54614	67740	1.987	0.806
	4	177006	42387	66100	2.678	0.641
	5	257371	80365	61259	4.201	1.312
Summer	1	22524	22524	92814	0.243	0.243
	2	66599	44075	89213	0.747	0.494
	3	116511	49912	83756	1.391	0.596
	4	177017	60506	86718	2.041	0.698
	5	278226	101209	83366	3.337	1.214
Fall	1	21967	21967	59743	0.368	0.368
	2	57511	35544	58283	0.987	0.610
	3	104974	47463	56746	1.850	0.836
	4	161937	56963	44755	3.618	1.273
	5	248762	86825	47183	5.272	1.840

Table 6.1 Load Forecast Uncertainty

The period ECOU results are monotone increasing as is expected. The daily ECOU results, however, do not always increase with lead time. As the daily ECOU is a differenced quantity, when the period ECOU increases, but with decreasing slope, the daily ECOU will decrease with time as well. The daily ECOU can never be negative, as long as the period ECOU is increasing.

In Chapter 4 the load forecasts are the only uncertain planning variable. In this case, it is easy to see that the planning uncertainty tracks the load forecast uncertainty, which follows the weather forecast error. The following figure shows an example plot of these uncertainties. Note the ordinate scale has been left off because here the shapes are important and not the absolute magnitudes. Note that it was shown in Chapter 2 that the dip in the temperature forecast errors are not statistically significant.

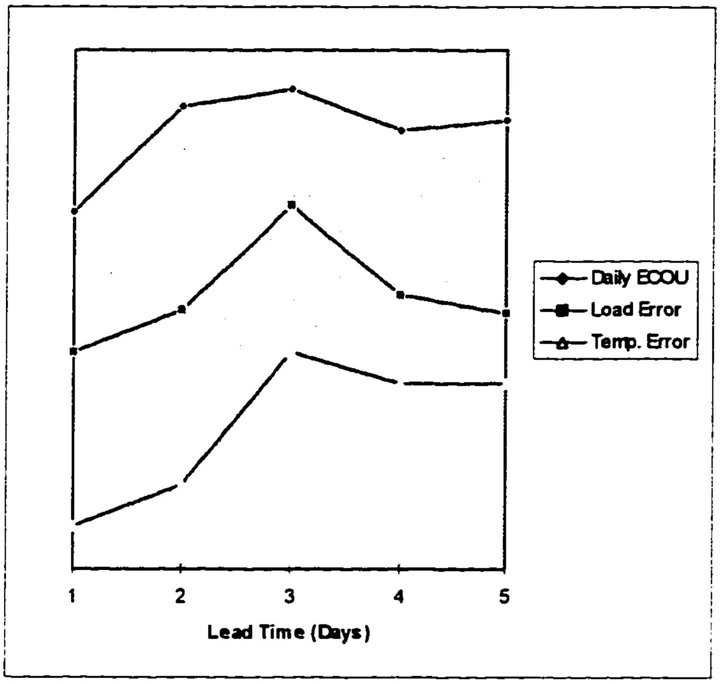

Figure 6.1 Uncertainty Shapes

Chapter 5 includes fuel price uncertainty into the risk model. The main addition that chapter 5 makes is reducing the multi-input variable risk model order and accounting for the interdependencies that may exist between the load and fuel prices. The results of Chapter 5 were given in the same format as Chapter 4. Table 6.2 shows the results of Chapter 5 in tabular form.

Season	Lead Time (Days)	Total ECOU ($)	Differenced ECOU ($)	Daily Energy (MWh)	Period ECOU ($/MWh)	Daily ECOU ($/MWh)
Winter	1	115116	115116	53388	2.156	2.156
	2	238136	123020	56956	4.181	2.160
	3	382702	144566	57665	6.637	2.507
	4	593942	211240	55808	10.643	3.785
	5	780209	186267	53597	14.557	3.475
Spring	1	152570	152570	65566	2.327	2.327
	2	350922	198352	67938	5.165	2.920
	3	672177	321255	67740	9.923	4.742
	4	1031278	359101	66100	15.602	5.433
	5	1711511	680233	61259	27.939	11.104
Summer	1	106588	106588	92814	1.148	1.148
	2	244939	138351	89213	2.746	1.551
	3	640816	395877	83756	7.651	4.727
	4	943444	302628	86718	10.879	3.490
	5	1092805	149361	83366	13.108	1.792
Fall	1	107574	107574	59743	1.801	1.801
	2	352042	244468	58283	6.040	4.194
	3	697165	345123	56746	12.286	6.082
	4	1046754	349589	44755	23.388	7.811
	5	1474895	428141	47183	31.259	9.074

Table 6.2 Load Forecast and Fuel Price Uncertainty

Notice that the fuel price uncertainty increases the over all uncertainty significantly. In a way, these results are unrealistic for the specific case listed. In general, the price of coal will not fluctuate over a weeks time, because coal is purchased in bulk, and in advance. However, this situation is realistic for a utility when two main fuels have volatile prices.

In addition, the correlation coefficients between the load and fuel prices have an impact on the model outcome as well. When studies are run for the completely dependent case and the completely independent case, the expected cost of uncertainty for these cases represent lower and upper bound for the risk, respectively, where ρ_{ij} varies from zero to one. This last section provides a summary of the research methodology and recommendations for the future research.

§ 6.3 Summary of Research and Recommendations

As a final analysis, this research presents a method for evaluating the risk due to load forecast and fuel price uncertainty. To achieve this goal, the following steps were taken. A load forecast method was developed in Chapter 2 that used Bayesian estimation to arrive at a load profile prediction for one to k days into the future. Bayesian estimation was chosen because the algorithm provides an estimation of the forecast variation. With both the load forecast profiles, standard deviation profiles and concepts from stratified sampling, three load scenarios were generated: high, medium and low. To each scenario, a probability was assigned. The results were used in conjunction with decision analysis to evaluate the ECOU, which is the expected value of decision making under uncertainty minus the expected value of decision making with perfect information. The cost functionals of the decision problem were determined by a chronological production cost simulator using the high, medium and low load scenarios as input. As the load forecasts were made from one through five days in advance, the expected cost of uncertainty was expressed as a function of lead time. The results were given in $/MWh. Fuel price uncertainty was then added to the risk model. This addition increased the dimensionality of the problem, so efforts were made to reduce the model order. Since the load and fuel prices have a dependent relationship, a method to approximate conditional triangular densities was developed as well.

The risk model, however, is not yet complete. Initial assumptions made regarding spinning reserve, the price of unserved energy and the utility's ability to reschedule generation units can be relaxed. Spinning reserve is determined based on a percentage of the annual system peak load. Suppose that in the future, spinning reserve can vary as a function of the system demand. To reflect this situation in the risk model, the spinning reserve can be evaluated on an hourly basis and included in the production cost simulation. To better mimic the competitive energy market, the price of unserved energy can be specified on an hourly basis as well. This price should follow the marginal cost of energy, as it varies over time. In real time operation, electric utilities can reschedule generation units; this constraint can be relaxed and included in the risk model also. It is in these directions that our research attention will be focused in the future.

CHAPTER 7

REFERENCES

1. Sachdeva, A., "Expected Cost of the Lead Time of Power Plant Construction," Master's Thesis, University of Oklahoma, 1991.
2. Von Neumann, J., and Morgenstern, O., *The Theory of Games and Economic Behavior, 3rd Edition,* Princeton University Press, Princeton, 1955.
3. Luce, R Duncan, and Raiffa, Howard, *Games and Decisions,* John Wiley and Sons, New York, 1957.
4. Howard, Ronald A., "Decision Analysis: Applied Decision Theory," *Proceedings of the 4th International Conference on Operational Research,* pp. 55-71, 1966.
5. Howard, Ronald A., "The Foundations of Decision Analysis," *IEEE Transactions on Systems, Science and Cybernetics,* Vol. SCC-4, No. 3, pp. 211-219, September 1968.
6. North, D. Warner, "A Tutorial Introduction to Decision Theory," *IEEE Transactions on Systems, Science and Cybernetics,* Vol. SCC-4, No. 3, pp. 200-210, September 1968.
7. Raiffa, Howard, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty.* McGraw Hill, Inc., New York, 1968.
8. Esser, W. F., Ghose, P. and Chen, K., "Decision Analysis for Electric Power System Engineering and Management," *IEEE Transactions on Power Apparatus and Systems,* Vol. PAS-96, No. 2, pp. 447-456, March/April 1977.
9. Anders, George J., *Probability Concepts in Electric Power Systems,* John Wiley and Sons, New York, 1990.
10. Douglas, John, "Road Map For Complex Decisions," *EPRI Journal,* pp. 6-13, September 1983.
11. Decision Focus, Inc., "Costs and Benefits of Over vs. Under Capacity in Electric Power System Planning," *EPRI Report EA-927,* EPRI, Palo Alto, CA, October 1978.
12. Galiana, F.D., et al., "Identification of Stochastic Electric Load Models from Physical Data," *IEEE Transactions on Automatic Control,* Vol. AC-19, No. 6, December 1974.

13. Box. G.E.P., Jenkins, G.M., and Reinsel, G.C., *Time Series Analysis: Forecasting and Control*. Prentice Hall, Englewool Cliffs, New Jersey, 1994.
14. Mbamalu, G.A.N. and El-Hawary, M.E., "Load Forecasting via Suboptimal Seasonal Auto-regressive Models and Iteratively Reweighted Least Squares Estimation," *IEEE Transactions on Power Systems*, Vol. 8, No. 1, pp. 343-348, February 1993.
15. Fan, J.Y. and McDonald J.D., "A Real-Time Implementation of Short-Term Load Forecasting for Distribution Power Systems," *IEEE Transactions on Power Systems*, Vol. 9, No. 2, pp. 988-994, May 1994.
16. Heinemann, G.T., et al., "The Relationship Between Summer Weather and Summer Loads - A Regression Analysis," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-85, No. 12, pp. 1144-1154, November 1966.
17. Draper, N.R., and Smith, H., *Applied Regression Analysis, 2nd Ed* John Wiley & Sons, New York, 1981.
18. Haida, T. and Muto, S., " Regression Based Peak Load Forecasting Using a Transformation Technique," *IEEE Transactions on Power Systems*, Vol. 9, No. 4, pp. 1788-1794, November 1994.
19. Chiang, Jia-Yo, et al., "An Implementation of a Regression Based Short-Term Load Forecasting Model for Microsoft Excel," Frontiers of Power Conference, 1996.
20. Park, D., et al., "Electric Load Forecasting Using an Artificial Neural Network," *IEEE Transactions on Power Systems*, Vol. 6, No. 2, pp. 442-449, May 1991.
21. Peng, T.M., et al., "Advancement in the Application of Neural Networks for Short-Term Load Forecasting," *IEEE Transactions on Power Systems*, Vol. 7, No. 1, pp. 250-257, February 1992.
22. Thomas, R.J., et al., "Power System Dynamic Modeling Using Artificial Neural Networks," *IEEE Transactions on Power Systems*, Vol. 9, No. 4, pp. 1868-1874, November 1994.
23. Mohammed, O., et al., "Practical Experiences with An Adaptive Neural Network Short-Term Load Forecasting System," *IEEE Transactions on Power Systems*, Vol. 10, No. 1, pp. 254-265, February 1995.
24. Sharma, B.E. and Mahalanabis A.K., "Recursive Short-Term Load-Forecasting Algorithm," *IEE Proceedings*, Vol. 121, No. 1, pp. 59-62, January 1974.
25. Gelb, A., *Applied Optimal Estimation*, MIT Press, Massachusetts, 1974.
26. Moghram, I. and Rahman, S., "Analysis and Evaluation of Five Short-Term Load Forecasting Techniques," *IEEE*

Transactions on Power Systems, Vol. 4, No. 4, pp. 1484-1491, October 1989.

27. Park, J.H., et al., "Composite Modeling for Adaptive Short-Term Load Forecasting," *IEEE Transactions on Power Systems*, Vol. 6, No. 2, pp. 450-457, May 1991.
28. DeGroot, M.H., *Optimal Statistical Decisions*. McGraw Hill, Inc., New York, 1970.
29. Box, G.E.P., Tiao, G.C., *Bayesian Inference in Statistical Analysis*. Addison-Wesley Publishing Company, New York, 1973.
30. Press, S.J., *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference, 2nd Ed.* Robert E. Krieger Publishing Company, Malabar, Florida, 1982.
31. Broemeling, L.D., *Bayesian Analysis of Linear Models*. Marcel Dekker, Inc., New York, 1985.
32. West M., and Harrison, J., *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York, 1989.
33. Pole, A., West M., and Harrison, J., *Applied Bayesian Forecasting and Time Series Analysis*. Chapman and Hall Publishing Co., New York, 1994.
34. Cochran, W.G., *Sampling Techniques*. John Wiley & Sons, New York, 1963.
35. Cochran, W.G. and Cox, G.M.,*Experimental Designs*. John Wiley & Sons, New York, 1957.
36. Johnson, N.L., and Leone, Fred C., *Statistics and Experimental Design in Engineering and the Physical Sciences: Volume II*. John Wiley & Sons, New York, 1977.
37. Dalenius, T., "The Problem of Optimum Stratification," *Skandinavisk Aktuarietidskrift*, Vol 33 #3-4, pp. 203-213, 1950.
38. Dalenius, T., and Gurney, M., "The Problem of Optimum Stratification II," *Skandinavisk Aktuarietidskrift*, Vol 34 #3-4, pp. 133-148, 1951.
39. Dalenius, T., and Hodges, J.L. Jr., "The Choice of Stratification Points," *Skandinavisk Aktuarietidskrift*, Vol 40 #3-4, pp. 198-203, 1957.
40. Hammersley, J.M., and Handscomb, D.C., *Monte Carlo Methods*. John Wiley & Sons, New York, pg. 56, 1964.
41. Papoulis, A., *Probability Random Variables, and Stochastic Processes*, 3rd Ed., McGraw-Hill, New York, 1991.
42. Breipohl, A.M. and Lee, F.N., "A Stochastic Load Model for Use in Operating Reserve Evaluation," *IEE Third International Conference Proceedings*, No. 338, pp. 123-126, 1991.
43. Breipohl, A.M., Lee, F.N., Zhai, D. and Adapa, R., "A Gauss-Markov Load Model For Application in Risk

Evaluation and Production Simulation," *IEEE-Transactions on Power Apparatus and Systems*, Vol. 7, No. 4, pp. 1493-1499, November 1992.

44. Shanmugan, K.S. and Breipohl A.M., *Random Signals: Detection, Estimation and Data Analysis*, John Wiley and Sons, New York, 1988.
45. Breipohl, A.M., *Probabilistic Systems Analysis*, John Wiley and Sons, New York, 1970.
46. Booth, R.R., "A Power System Simulation Model Based on Probability Analysis," *IEEE-Transactions on Power Apparatus and Systems*, PAS-91, 1972.
47. Chiang, Jia-Yo, et al., "Stochastic Load Variation Modeling for Estimating the Variance of Production Cost," Submitted to *IEEE-Transaction on Power Systems*, 1997.
48. Billington, Roy, *Power System Reliability Evaluation*. Gordon and Breach Science Publishers, New York, 1970.
49. Wood, Allen J., and Wollenberg, Bruce F., *Power Generation Operation & Control*. John Wiley & Sons, New York, 1984.
50. Gross, Charles A., *Power System Analysis*. John Wiley & Sons, New York, 1986.
51. Stoll, Harry G., *Least-Cost Utility Planning*. John Wiley & Sons, New York, 1989.
52. Luenberger, David G., *Optimization by Vector Space Methods*. John Wiley & Sons, New York, 1969.
53. Burden, R.L. and Faires, D.F., *Numerical Analysis*, PWS Publishing Company, Boston, Massachusetts, 1993.
54. Larson, R.J. and Marx, M.L., *An Introduction to Mathematical Statistics and its Applications, 2nd Ed.*, Prentice Hall, New Jersey, 1986.
55. Lewis, F.L., *Optimal Estimation, with an Introduction to Stochastic Control Theory*, John Wiley and Sons, New York, 1986.
56. Tong, Y.L., *The Multivariate Normal Distribution*, Springer-Verlag, New York, 1990.
57. Evans, M.,et al., *Statistical Distributions, 2nd Ed.*,John Wiley & Sons, Inc., New York, 1993.
58. Buzzi-Ferraris, G., *Scientific C++: Building Numerical Libraries the Object-Oriented Way*, Addison-Wesley Publishing Co., New York, 1994.
59. Press, H.W., *Numerical Recipes in C: The Art of Scientific Computing, 2nd Ed.*, Cambridge University Press, New York, 1992.

APPENDIX 1

PROOF OF TRIANGULAR VARIANCE SCALING

This appendix serves as the proof of the variance scaling equation used in Chapters 3 and 5. Consider the following skewed triangular distribution:

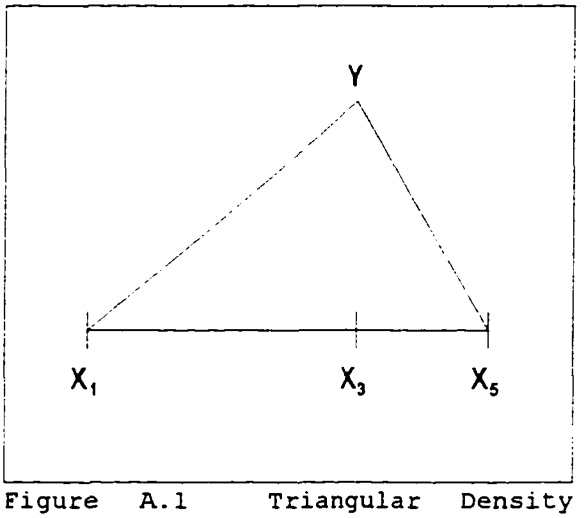

Figure A.1 Triangular Density Function

Given the center coordinate X_3, the variance σ^2, the desired skewedness r and the coordinate Y such that the area of the triangle is unity; the goal is to find X_1 and X_5. It is assumed that $0 < X_1 < X_3 < X_5$. The number r is the area of the left hand triangle divided by that of the right hand triangle, where the left hand triangle is (X_1,X_3,Y) and the right hand triangle is (X_3,X_5,Y). This ratio can be reduced to the following expression:

$$r = \frac{(X_3 - X_1)}{(X_5 - X_3)} \tag{A.1}$$

The variance of a triangular distribution is as follows[1]:

$$\sigma^2 = \frac{1}{18}\left[X_1^2 + X_3^2 + X_5^2 - \left(X_1X_3 + X_1X_5 + X_3X_5\right)\right] \tag{A.2}$$

Equation (A.2) can be rewritten as

$$\begin{aligned} 18\sigma^2 &= \left(X_1^2 - 2X_1X_3 + X_3^2\right) + \left(X_5X_3 - X_5X_1 - X_3^2 + X_3X_1\right) + \left(X_3^2 - 2X_3X_5 + X_5^2\right) \\ &= \left(X_1 - X_3\right)^2 + \left(X_1 - X_3\right)\left(X_3 - X_5\right) + \left(X_3 - X_5\right)^2 \\ &= \left(X_1 - X_3\right)^2\left(\frac{X_3 - X_5}{X_3 - X_5}\right)^2 + \left(X_1 - X_3\right)\frac{\left(X_3 - X_5\right)^2}{X_3 - X_5} + \left(X_3 - X_5\right)^2 \\ &= \left[1 + \frac{X_3 - X_1}{X_5 - X_3} + \left(\frac{X_3 - X_1}{X_5 - X_3}\right)^2\right]\left(X_5 - X_3\right)^2 \end{aligned} \tag{A.3}$$

Note that $X_3 < X_5$. Equation (A.3) is simplified to

$$18\sigma^2 = (X_5 - X_3)^2\left(1 + r + r^2\right) \tag{A.4}$$

and further simplified to

$$X_5 = X_3 + \sqrt{\frac{18\sigma^2}{1+r+r^2}} \tag{A.5}$$

Note that $0 < 1 + r + r^2$. From Equation (A.1) we have

$$X_1 = (1+r)X_3 - rX_5 \tag{A.6}$$

By using the results of (A.5) in (A.6), the following completes the proof.

$$X_1 = X_3 - r\sqrt{\frac{18\sigma^2}{1+r+r^2}} \tag{A.7}$$

Q.E.D.

---

[1] Evans, M., et al., *Statistical Distributions, 2nd Ed.*, John Wiley & Sons, Inc., New York, 1993, pg. 149.

APPENDIX 2

TEST SYSTEM DATA

The data set used for this research is taken from a medium size electric utility whose annual peak is approximately 5,000 MW and whose annual energy is roughly 23,000 GWH. Five years of hourly chronological load and temperature data is available, as well as, one year of peak and valley temperature forecasts from the National Weather Service. Listed below are the unit data.

Unit #	Type	Pmax	Pmin	Ramp Rate	Min Up	Min Down	Heat Rate a0	Heat Rate a1	Heat Rate a2
1	Gas	74	22	4	24	24	8.62E+01	9.40E+00	1.90E-02
2	Gas	48	12	2	4	4	9.40E+01	1.54E+01	1.00E-03
3	Gas	178	45	5	24	24	2.30E+02	7.41E+00	6.00E-03
4	Gas	238	125	5	24	24	3.49E+02	6.94E+00	1.60E-03
5	Gas	412	100	10	24	24	5.87E+02	8.12E+00	3.00E-03
6	Gas	184	45	5	24	24	2.11E+02	8.09E+00	4.00E-03
7	Coal	500	150	10	72	72	5.72E+02	9.01E+00	6.50E-04
8	Coal	500	150	10	72	72	5.72E+02	9.01E+00	6.50E-04
9	Coal	500	150	10	72	72	5.55E+02	8.20E+00	2.00E-03
10	Gas	58	15	2	24	24	1.20E+02	7.80E+00	3.45E-02
11	Gas	57	15	2	24	24	7.58E+01	9.57E+00	3.00E-04
12	Gas	122	30	3	24	24	1.46E+02	8.15E+00	1.57E-03
13	Gas	260	60	5	24	24	3.13E+02	6.99E+00	9.70E-03
14	Gas	64	16	2	4	4	3.20E+01	1.54E+01	1.00E-03
15	Gas	530	150	10	24	24	3.53E+02	8.22E+00	1.45E-03
16	Gas	507	150	10	24	24	4.93E+02	8.22E+00	1.00E-03
17	Gas	500	150	10	24	24	7.40E+02	7.22E+00	2.70E-03
18	Gas	19	18	2	4	4	3.80E+01	1.54E+01	1.00E-03
19	Coal	505	150	10	72	72	7.01E+02	6.07E+04	5.00E-03
20	Coal	510	150	10	72	72	6.71E+02	7.16E+00	3.50E-03
21	Gas	11	10	2	4	4	2.20E+01	1.54E+01	1.00E-03
	Type	Capacity	CF %		Type	Capacity	CF %		
#22	Cogen	110	95	#23	Cogen	160	65		
#24	Cogen	160	65	#25	Cogen	64	85		

Table A.1 Unit Data

APPENDIX 3

SOURCE CODE

This appendix contains the main source code functions used to perform the illustrative case studies. Code written in C++ is formatted such that when lines are continued, they follow on the next line without indication. C++ Lines always end with a ";". Code written in BASIC has an "_" to signify the continuation of a line and a carriage return to signify the end of the line.

```
/*****************************************************/
//      Short-term Load Forecasting
//      Bayesian Estimation
//      FORECAST.CPP
/*****************************************************/
#include <time.h>
#include <math.h>
#include <ctype.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <fstream.h>
#include <iomanip.h>
#include <iostream.h>
#include "utility.hpp"

int main(void) {
   ifstream ifs("dtbase.bin", ios::in | ios::binary);
   enum simType { forecast, actual } runType;
   Dtbase Dt[6];
   FILE *matrixIn;
   double e,xa,xb,ae,pk_4cast,hr24_4cast;
   double n = 1,d = 40000,del = 0.93;
   int lt = 5,order = 15,i,j,k,l;
   Matrix[2] G(order,order,matrixIn); Matrix GT = G;
   Transpose(&GT);
   Matrix FT(1,order,matrixIn); Matrix F = FT;
```

[2] Buzzi-Ferraris, G., *Scientific C++: Building Numerical Libraries the Object-Oriented Way*, Addison-Wesley Publishing Co., New York, 1994.

```
Matrix eye(order,order,matrixIn);
Matrix Q(1,1),S0(1,1,d),S1(1,1);
Matrix m(order,1,matrixIn);
Matrix C = 100000.0*eye;
runType = forecast;
//runType = actual;
do {
   ifs.read((char *)&Dt[0],sizeof(Dt[0]));
} while(Dt[0].cal.year != 1994);
do {
   ifs.read((char *)&Dt[1],sizeof(Dt[1]));
   m = G*m;
   C = G*C*GT/del;
   FT(1,2) = Dt[0].sLd.ld[23];
   FT(1,3) = Dt[0].sLd.ld[Dt[0].tLd.pkHr];
   FT(1,4) = Dt[0].sLd.mu;
   FT(1,5) = Dt[0].aTemp.temp[23];
   FT(1,6) = Dt[0].aTemp.pk;
   FT(1,7) = Dt[0].aTemp.mu;
   if(runType == forecast) {
      FT(1,8) = Dt[1].aTemp.mu4[0];
      FT(1,9) = Dt[1].aTemp.pk4[0];
   }
   else {
      FT(1,8) = Dt[1].aTemp.mu;
      FT(1,9) = Dt[1].aTemp.pk;
   }
   F = FT; Transpose(&F);
   Q = FT*C*F+S0;
   Matrix f = FT*m;
   pk_4cast = f(1,1)+Dt[1].bLd.ld[Dt[1].tLd.pkHr];
   e = Dt[1].aLd.ld[Dt[1].tLd.pkHr]-pk_4cast;
   S1(1,1) = d/double(n);
   Matrix A = C*F/Q(1,1);
   Matrix AT = A; Transpose(&AT);
   Matrix U = (eye-A*FT);
   Matrix UT = U; Transpose(&UT);
   C = S1(1,1)*(U*C*UT+A*S0*AT)/S0(1,1);
   m = m+e*A;
   n = n*del+1;
   d = d*del+S0(1,1)*e*e/Q(1,1);
   S0 = S1;
   Dt[0] = Dt[1];
} while (Dt[0].cal.julian_date < 60);
for(l=spring;l<=winter;l++) {
   for(k=1;k<=91;k++) {
      ifs.read((char *)&Dt[1],sizeof(Dt[1]));
      m = G*m;
      C = G*C*GT/del;
      FT(1,2)=Dt[0].sLd.ld[23];
```

```
FT(1,3)=Dt[0].sLd.ld[Dt[0].tLd.pkHr];
FT(1,4)=Dt[0].sLd.mu;
FT(1,5)=Dt[0].aTemp.temp[23];
FT(1,6)=Dt[0].aTemp.pk;
FT(1,7)=Dt[0].aTemp.mu;
if(runType == forecast) {
   FT(1,8) = Dt[1].aTemp.mu4[0];
   FT(1,9) = Dt[1].aTemp.pk4[0];
}
else {
   FT(1,8) = Dt[1].aTemp.mu;
   FT(1,9) = Dt[1].aTemp.pk;
}
F = FT; Transpose(&F);
Q = FT*C*F+S0;
Matrix f = FT*m;
pk_4cast = f(1,1)+Dt[1].bLd.ld[Dt[1].tLd.pkHr];
e = Dt[1].aLd.ld[Dt[1].tLd.pkHr]-pk_4cast;
Matrix mSave = m;
Matrix fSave = f;
Matrix FTSave = FT;
Matrix CSave = C;
Matrix QSave = Q;
xa=(Dt[1].tLd.pk*Dt[0].aLd.ld[23]-
   Dt[0].tLd.ld[23]*pk_4cast)/
   (Dt[1].tLd.pk-Dt[0].tLd.ld[23]);
xb = (pk_4cast-Dt[0].aLd.ld[23])/
   (Dt[1].tLd.pk-Dt[0].tLd.ld[23]);
hr24_4cast = xa+xb*Dt[1].tLd.ld[23];
for(i=1;i<lt;i++) {
   ifs.read((char *)&Dt[i+1],sizeof(Dt[i+1]));
   m = G*m;
   C = G*C*GT/del;
   if(runType == forecast) {
      FT(1,8) = Dt[1].aTemp.mu4[i];
      FT(1,9) = Dt[1].aTemp.pk4[i];
   }
   else {
      FT(1,8) = Dt[i+1].aTemp.mu;
      FT(1,9) = Dt[i+1].aTemp.pk;
   }
   Q = FT*C*F+(i+1)*S0;
   f = FT*m;
   pk_4cast=f(1,1)+
      Dt[i+1].bLd.ld[Dt[i+1].tLd.pkHr];
   ae=Dt[i+1].aLd.ld[Dt[i+1].tLd.pkHr]-pk_4cast;
   xa = (Dt[i+1].tLd.pk*hr24_4cast-
      Dt[i].tLd.ld[23]*pk_4cast)/
      (Dt[i+1].tLd.pk-Dt[i].tLd.ld[23]);
   xb = (pk_4cast-hr24_4cast)/
```

```
            (Dt[i+1].tLd.pk-Dt[i].tLd.ld[23]);
         hr24_4cast=xa+xb*Dt[i+1].tLd.ld[23];
      }
      if(lt != 1) {
         ifs.seekg(-lt*int(sizeof(Dt[1])),ios::cur);
         ifs.read((char *)&Dt[1],sizeof(Dt[1]));
      }
      m = mSave;
      f = fSave;
      FT = FTSave;
      C = CSave;
      Q = QSave;
      S1(1,1) = d/n;
      Matrix A = C*F/Q(1,1);
      Matrix AT = A; Transpose(&AT);
      Matrix U = (eye-A*FT);
      Matrix UT = U; Transpose(&UT);
      //Joeseph Algorithm pp. 305-306 Gelb
      //Applied Optimal Estimation
      //Ensures symmetry of C and numercially stable
      C = S1(1,1)*(U*C*UT+A*S0*AT)/S0(1,1);
      m = m+e*A;n = n*del+1;
      d = d*del+S0(1,1)*e*e/Q(1,1);
      S0 = S1;
      Dt[0] = Dt[1];
   }
  }
return(0);
}
```

```
'********************************************************'
'*      Risk due to Load Forecast Uncertainty
'*
'*      ECOULOAD.XLS
'********************************************************'

Public Sub loadUncertainty(ByVal ext As Integer, _
   ByVal loadExt As String)

Dim i As Integer, j As Integer, order As Integer:
order = 3
Dim ucIndex As Integer, loadIndex As Integer, _
   fuelIndex As Integer, fuelNumberIndex As Integer
Dim returnPCI As Integer, sys_invl As String, _
   stsysm01 As String, fileExt As String
Dim xVector(4) As Single, yVector(4) As Single, _
   zVector(4) As Single, totalCount As Integer
totalCount = 0
Dim fuelProb(2) As Single, loadProb(2) As Single, _
   loadFileNames(2) As String, ucFileNames() As String
Dim testNumber As Double, lowIndex As Integer, _
   numberOfFuels As Integer
Dim expcetedCost() As Double, z() As Single, _
   costs() As Long, clair As Double
Dim studyType As String: studyType = getStudyType()
Dim currentSheet As Object
Dim goAhead As Integer, simNumber As Integer, _
   timeLeft As Double
Dim t0 As Long, timeString As String

ReDim ucFileNames(order - 1) As String
ReDim costs(order - 1, order - 1) As Long
ReDim expectedCost(order - 1) As Double
Set currentSheet = Sheets(CStr(iniSheet()))
loadFileNames(0) = "low" + loadExt
loadFileNames(1) = "mean" + loadExt
loadFileNames(2) = "high" + loadExt
testNumber = 1E+20
lowIndex = 0
sys_invl = driveLetter + _
   ":\scheduler\dbdata\sys_invl." + runExt
stsysm01 = driveLetter + _
   ":\scheduler\output\stsysm01." + runExt
loadProb(0) = 0.2: loadProb(1) = 0.6: loadProb(2) = 0.2
fuelProb(0) = 0.2: fuelProb(1) = 0.6: fuelProb(2) = 0.2
For loadIndex = 0 To 2
  ucFileNames(loadIndex) = driveLetter + _
     ":\scheduler\dbdata\ucfile." + AgetExt(loadIndex)
Next loadIndex
t0 = Timer
```

```
If putLoadFile(loadFileNames(1)) = False Then Exit Sub
If putUCTag("flexable") = False Then Exit Sub
For loadIndex = 0 To 2
   If putLoadFile(loadFileNames(loadIndex)) = _
      False Then Exit Sub
   returnPCI = callPCI(ext)
   FileCopy driveLetter + ":\scheduler\dbdata\unicom01."+_
      runExt, ucFileNames(loadIndex)
   costs(loadIndex, loadIndex) = getProdCost()
   currentSheet.Cells(2 + loadIndex + loadIndex * 3, 1) =
   costs(loadIndex, loadIndex)
   totalCount = totalCount + 1
Next loadIndex
If putUCTag("fixed") = False Then Exit Sub
For ucIndex = 0 To 2
   FileCopy ucFileNames(ucIndex), driveLetter + _
   ":\scheduler\dbdata\unicom01." + runExt
   For loadIndex = 0 To 2
      If putLoadFile(loadFileNames(loadIndex)) = _
         False Then Exit Sub
      If ucIndex <> loadIndex Then
         returnPCI = callPCI(ext)
         costs(ucIndex, loadIndex) = getProdCost()
         currentSheet.Cells(2 + 3 * ucIndex + _
            loadIndex, 1) = costs(ucIndex, loadIndex)
         totalCount = totalCount + 1
      End If
   Next loadIndex
Next ucIndex
clair = 0
For ucIndex = 0 To 2
   expectedCost(ucIndex) = 0
   For loadIndex = 0 To 2
      If loadIndex = ucIndex Then clair = clair + _
         fuelProb(loadIndex) * costs(ucIndex, loadIndex)
      expectedCost(ucIndex) = expectedCost(ucIndex) + _
         fuelProb(loadIndex) * costs(ucIndex, loadIndex)
   Next loadIndex
   If expectedCost(ucIndex) < testNumber Then _
      testNumber = expectedCost(ucIndex)
Next ucIndex
End Sub
```

```
'*******************************************************'
'*      Risk due to Load Forecast and
'*      Fuel Price Uncertainty
'*
'*      ECOUFUEL.XLS
'*******************************************************'

Public Sub fueluncertainty(ByVal ext As Integer, _
   ByRef x() As Single, ByVal loadExt As String)
'scalar integer definitions
   Dim i As Integer, j As Integer, k As Integer, _
   totalNumber As Integer, simNumber As Integer, _
      loadIndex As Integer, fuelNumberIndex As Integer, _
      goAhead As Integer, returnPCI As Integer, _
      numberOfFuels As Integer, fuelIndex1 As Integer, _
      fuelIndex2 As Integer, fuelIndex3 As Integer, _
      fuelIndex4 As Integer, fuelIndex5 As Integer, _
      fuelIndex6 As Integer, fuelIndex7 As Integer, _
      cellNumber As Integer, fuelstate(7) As Integer, _
      ucIndex As Integer
'scalar string definitions
   Dim studyType As String, sys_invl As String, _
      stsysm01 As String
'vector string definitions
   Dim loadFileNames(2) As String, states(2) As String, _
      ucFileNames() As String
'vector single definitions
   Dim xVector(4) As Single, yVector(4) As Single, _
      zVector(4) As Single, z() As Single
'scalar double definitions
   Dim timeLeft As Double, probs As Double, _
      muPrior As Double, sigmaPrior As Double, _
      sigmaPosterior As Double, r As Double, _
      newfuel As Double, testNumber As Double, _
      clair As Double
'vector double definitions
   Dim shiftValue(7) As Double, expectedCost() As Double
'scalar long definitions
   Dim t0 As Long, costs() As Long
'scalar object definitions
   Dim currentSheet As Object
'scalar matrix definitions
   Dim sigma As Matrix, sigma12 As Matrix, _
      sigma21 As Matrix, sigma22 As Matrix, _
      statevector As Matrix, result1 As Matrix, _
      result2 As Matrix, result3 As Matrix, _
      expFuel As Matrix, varfuel As Matrix, _
      triFuel As Matrix, covFuel As Matrix

loadFileNames(0) = "low" + loadExt
```

```
loadFileNames(1) = "mean" + loadExt
loadFileNames(2) = "high" + loadExt

states(0) = "L"
states(1) = "M"
states(2) = "H"

numberOfFuels = AgetNumberOfFuels()
numberOfFuels = 2

sys_invl = driveLetter + _
  ":\scheduler\dbdata\sys_invl." + runExt
stsysm01 = driveLetter + _
  ":\scheduler\output\stsysm01." + runExt
ChDrive driveLetter
Set currentSheet = Sheets(CStr(iniSheet()))
totalNumber = 2 ^ (numberOfFuels + 1) + 1
probs = 0.4 / (totalNumber - 1)
testNumber = 1E+20
clair = 0

ReDim ucFileNames(totalNumber) As String
ReDim costs(totalNumber, totalNumber) As Long
ReDim expectedCost(totalNumber) As Double

'set rows and columns for matrices
expFuel.cols = 3
expFuel.Rows = numberOfFuels
varfuel.cols = 3
varfuel.Rows = numberOfFuels
triFuel.cols = 5
triFuel.Rows = numberOfFuels
covFuel.cols = numberOfFuels
covFuel.Rows = numberOfFuels

'determines the triangular distribtuion moments
For j = 0 To numberOfFuels - 1
  For i = 0 To 4
    xVector(i) = x(j, I)
  Next i
  If (Coords(xVector, yVector, 0.1) = 0) Then Exit Sub
  For i = 0 To 4
    triFuel.vals(j + 1, i + 1) = xVector(i)
  Next I
'Tcond??? uses x and y-Vectors as inputs,
'zVector is the ouput
  TcondExp xVector, yVector, zVector
  For i = 0 To 2
    expFuel.vals(j + 1, i + 1) = Csng(zVector(i))
  Next I
```

```
   TcondVar xVector, yVector, zVector
   For i = 0 To 2
      varfuel.vals(j + 1, i + 1) = Csng(zVector(i))
   Next I
Next j

Dim rhoArray As Variant
rhoArray = Array(1,0.8,0.7,0.8,1,0.5,0.7,0.5,1)
For j = 0 To 7
   For i = 0 To 7
      corr(j, i) = CSng(rhoArray(i + j * 8))
   Next I
Next j

'make the covariance matrix
'load the diagonal first with unconditional variances
For i = 1 To numberOfFuels
   covFuel.vals(i, i) = tVar(triFuel.vals(i, 1), _
      triFuel.vals(i, 3), triFuel.vals(i, 5))
Next i

'generate off diagonals with correlation coefficients
For i = 1 To numberOfFuels
   For j = 1 To numberOfFuels
      covFuel.vals(i, j) = covFuel.vals(i, i) * _
         covFuel.vals(j, j) * corr(i, j)
   Next j
Next i

For loadIndex = 0 To totalNumber - 1
   ucFileNames(loadIndex) = driveLetter + _
      ":\scheduler\dbdata\ucfile." + AgetExt(loadIndex)
Next loadIndex
Application.ScreenUpdating = True

simNumber = -1
If putUCTag("flexable") = False Then Exit Sub
For loadIndex = 0 To 2 Step 2
   If putLoadFile(loadFileNames(loadIndex)) = _
      False Then Exit Sub
   fuelstate(0) = loadIndex
   For fuelIndex1 = 0 To 2 Step 2
      fuelstate(1) = fuelIndex1
      For fuelIndex2 = 0 To 2 Step 2
         fuelstate(2) = fuelIndex2
         If numberOfFuels = 1 Then Exit For
         If numberOfFuels = 2 Then
            simNumber = simNumber + 1
            cellNumber = 3 + simNumber + _
               simNumber * totalNumber
```

```
            currentSheet.Cells(3 + simNumber + _
               simNumber * totalNumber, 7) = _
               states(fuelstate(0))
            For k = 1 To numberOfFuels
               currentSheet.Cells(cellNumber, 7) = _
                  currentSheet.Cells(cellNumber, 7) + _
                  states(fuelstate(k))
               newfuel = condFuelPrice(k, covFuel, _
                  numberOfFuels, fuelstate, expFuel, _
                  varfuel, triFuel)
               currentSheet.Cells(cellNumber, 7 + k) = _
                  newfuel
               If (putFuelPrice(k - 1, newfuel) = 0) Then _
                  Exit Sub
            Next k
            returnPCI = callPCI(ext)
            FileCopy driveLetter + _
               ":\scheduler\dbdata\unicom01." + runExt, _
               ucFileNames(simNumber)
            costs(simNumber, simNumber) = getProdCost()
            currentSheet.Cells(cellNumber, 1) = _
               costs(simNumber, simNumber)
            Application.StatusBar = Cstr(simNumber)
            End If
      Next fuelIndex2
   Next fuelIndex1
Next loadIndex

probs = 0.6
loadIndex = 1
simNumber = simNumber + 1
cellNumber = 3 + simNumber + simNumber * totalNumber
currentSheet.Cells(cellNumber, 7) = states(1)
For k = 1 To numberOfFuels
   newfuel = expFuel.vals(k, 2)
   currentSheet.Cells(cellNumber, 7 + k) = newfuel
   currentSheet.Cells(cellNumber, 7) = _
   currentSheet.Cells(cellNumber, 7) + states(1)
   If (putFuelPrice(k - 1, newfuel) = 0) Then Exit Sub
Next k
returnPCI = callPCI(ext)
FileCopy driveLetter + ":\scheduler\dbdata\unicom01." + _
   runExt, ucFileNames(simNumber)
costs(simNumber, simNumber) = getProdCost()
currentSheet.Cells(cellNumber, 1) = costs(simNumber, _
   simNumber)
Application.StatusBar = "0," + Cstr(simNumber)


If putUCTag("fixed") = False Then Exit Sub
```

```
For ucIndex = 0 To 8
   simNumber = -1
   FileCopy ucFileNames(ucIndex), driveLetter + _
      ":\scheduler\dbdata\unicom01." + runExt
   For loadIndex = 0 To 2 Step 2
      If putLoadFile(loadFileNames(loadIndex)) = _
      False Then Exit Sub
      fuelstate(0) = loadIndex
      For fuelIndex1 = 0 To 2 Step 2
         fuelstate(1) = fuelIndex1
         For fuelIndex2 = 0 To 2 Step 2
            fuelstate(2) = fuelIndex2
            If numberOfFuels = 2 Then
               simNumber = simNumber + 1
               cellNumber = 3 + simNumber + _
                  ucIndex * totalNumber
                  If ucIndex <> (simNumber) Then
                  currentSheet.Cells(cellNumber, _
                     7) = states(fuelstate(0))
                  For k = 1 To numberOfFuels
                     currentSheet.Cells(cellNumber, _
                        7) = currentSheet.Cells(3 + _
                        simNumber, 7) + _
                        states(fuelstate(k))
                     newfuel = condFuelPrice(k, _
                        covFuel, numberOfFuels, _
                        fuelstate, expFuel, varfuel, _
                        triFuel)
                     currentSheet.Cells(cellNumber, _
                        7 + k) = newfuel
                     If (putFuelPrice(k - 1, _
                        newfuel) = 0) Then Exit Sub
                  Next k
                  FileCopy driveLetter + _
                     ":\scheduler\dbdata\unicom01." _
                        + runExt, ucFileNames(simNumber)
                  returnPCI = callPCI(ext)
                  costs(ucIndex, simNumber) = _
                     getProdCost()
                  currentSheet.Cells(cellNumber, _
                     1) =  costs(ucIndex, simNumber)
                  Application.StatusBar = _
                     Cstr(ucIndex) + "," + _
                     Cstr(simNumber)
               End If
            End If
         Next fuelIndex2
      Next fuelIndex1
   Next loadIndex
```

```
   probs = 0.6
   loadIndex = 1
   simNumber = simNumber + 1
   cellNumber = 3 + simNumber + ucIndex * totalNumber
   If ucIndex <> simNumber Then
      currentSheet.Cells(cellNumber, 7) = states(1)
      For k = 1 To numberOfFuels
         newfuel = expFuel.vals(k, 2)
         currentSheet.Cells(cellNumber, 7 + k) = newfuel
         currentSheet.Cells(cellNumber, 7) = _
            currentSheet.Cells(3 + simNumber, 7) + states(1)
         If (putFuelPrice(k - 1, newfuel) = 0) Then Exit Sub
      Next k
   FileCopy driveLetter + _
         ":\scheduler\dbdata\unicom01." + runExt, _
            ucFileNames(simNumber)
         returnPCI = callPCI(ext)
         costs(ucIndex, simNumber) = getProdCost()
         currentSheet.Cells(cellNumber, 1)=costs(ucIndex, _
            simNumber)
         Application.StatusBar = CStr(ucIndex) + "," + _
            Cstr(simNumber)
   End If
Next ucIndex

Dim ii As Integer
ii = 0
For ucIndex = 0 To 8
   expectedCost(ucIndex) = 0
   For loadIndex = 0 To 8
      If loadIndex = 8 Then: probs = 0.6
      Else: probs = 0.05: End If
      If loadIndex = ucIndex Then clair = clair + _
         probs * costs(ucIndex, loadIndex)
      expectedCost(ucIndex) = expectedCost(ucIndex) + _
         probs * costs(ucIndex, loadIndex)
   Next loadIndex
   If expectedCost(ucIndex) < testNumber Then
      testNumber = expectedCost(ucIndex)
      ii = ucIndex
   End If
Next ucIndex

End Sub
```

```
'*********************************************************'
'*      Conditional Fuel Price
'*
'*      ECOUFUEL.XLS
'*********************************************************'

Function condFuelPrice(ByVal k As Integer, _
   covFuel As Matrix, numberOfFuels As Integer, _
   fuelstate() As Integer, expFuel As Matrix, _
   varfuel As Matrix, triFuel As Matrix) As Double

Dim sigma As Matrix, sigma12 As Matrix, _
   sigma21 As Matrix, sigma22 As Matrix, _
   statevectorPrior As Matrix, statevector As Matrix, _
   result1 As Matrix, result2 As Matrix, _
   result3 As Matrix

Dim muPrior, sigmaPrior, sigmaPosterior
Dim xVector(4) As Single, yVector(4) As Single, _
   zVector(4) As Single
Dim r As Double, newfuel As Double
Dim j As Integer, shiftValue(7) As Double

'(1) get matrix conditional mean
copyMatrix covFuel, sigma
swapRC sigma, 1, k
pickMatrix sigma,sigma12,1,1,2,numberOfFuels
pickMatrix sigma,sigma21,2,numberOfFuels,1,1
pickMatrix sigma,sigma22,2,numberOfFuels,2,numberOfFuels

For j = 1 To numberOfFuels
   statevectorPrior.vals(j, 1) = expFuel.vals(j, _
      fuelstate(j) + 1) - expFuel.vals(j, 1 + 1)
Next j

statevectorPrior.Rows = numberOfFuels
statevectorPrior.cols = 1
swapR statevectorPrior, 1, k
pickMatrix statevectorPrior, statevector, 2, _
   numberOfFuels, 1, 1

copyMatrix sigma22, result1
If invMatrix(sigma22, result1)[3] <> True Then
   condFuelPrice = 0
   Exit Function
End If
```

[3] Press, H.W., *Numerical Recipes in C: The Art of Scientific Computing, 2nd Ed.*, Cambridge University Press, New York, 1992.

```
multMatrix sigma12, result1, result1
multMatrix result1, statevector, result2
muPrior = triFuel.vals(k, 3) + result2.vals(1, 1)
multMatrix result1, sigma21, result3
sigmaPrior   =   varfuel.vals(k,   fuelstate(k)   +   1)   -
result3.vals(1, 1)

'(2) get scalar conditional mean
If fuelstate(0) = 0 Then
   shiftValue(k) = -1.389 * corr(0, k) * Sqr(sigmaPrior)
ElseIf fuelstate(0) = 2 Then
   shiftValue(k) = 1.389 * corr(0, k) * Sqr(sigmaPrior)
Else
   shiftValue(k) = 0
End If

sigmaPosterior = sigmaPrior * (1 - corr(0, k) ^ 2)

'shift and scale the new distribution
r = (triFuel.vals(k, 3) - triFuel.vals(k, 1)) / _
   (triFuel.vals(k, 5) - triFuel.vals(k, 3))
xVector(1) = 0
xVector(3) = 0
xVector(2) = muPrior + shiftValue(k)
xVector(4) = xVector(2) + Sqr(18 * sigmaPosterior / _
   (1 + r + r ^ 2))
xVector(0) = (1 + r) * xVector(2) - r * xVector(4)
yVector(0) = 0
yVector(1) = 0
yVector(3) = 0
yVector(4) = 0
yVector(2) = 2 / (xVector(4) - xVector(0))
Call coords2(xVector, yVector, 0.1)
Call TcondExp(xVector, yVector, zVector)

'setfuel price
condFuelPrice = zVector(fuelstate(k))
End Function
```