2007

Effects of fuel cost uncertainty on optimal energy flows in U.S.

Yan Wang
Iowa State University

Effects of fuel cost uncertainty on optimal energy flows in U.S.

by

Yan Wang

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Industrial Engineering

Program of Study Committee:
Sarah M. Ryan, Major Professor
Sigurdur Olafsson
William Q. Meeker

Iowa State University

Ames, Iowa

2007

UMI Number: 1443136

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

DEDICATION

I would like to dedicate this thesis to my parents Jinjiang Wang and Xiaoxia Shi, whose love made me confident of myself and helped me overcome all the difficulties. It is also dedicated to my grandparents Pinxiang Shi and Qiuxiang Wu, who brought me up and offered me unconditional love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, my major professor Dr. Sarah Ryan for her guidance, patience and support throughout this research and the writing of this thesis. Her insights are illuminating and her enthusiasm inspired me to pursue a high level of reseach work. I appreciate my committee members for their efforts and contributions to this work: Dr. Sigurdur Olafsson and Dr. William Q. Meeker. I would also like to thank Dr. Rahul Marathe for his help with writing the thesis in Latex and thank my colleagues Suphalat Chittamvanich and Somchan Vuthipadadon for their attentions kindly paid to my graduate study.

ABSTRACT

The research is motivated by the need for economic efficiency and risk management in the national electric system. Stochastic costs of natural gas are introduced in a generalized network flow model of the integrated power energy system to explore the effects of uncertain fuel costs on the optimal energy flows in U.S. The fuel costs are modeled as discretely distributed random variables and a rolling two-stage approach is applied to solve the stochastic recourse problem. All the data are derived from publicly available information for the year 2002. The natural gas price forecasts by the Energy Information Administration are adapted to generate scenarios that are considered in the stochastic problem. Compared to the expected value solution from the deterministic model, the recourse problem solution obtained from the stochastic model has higher total cost, lower natural gas consumption and less subregional power trade but a flow mix which is closer to the 2002 real data. Surprisingly, increasing the uncertainty level of the scenarios leads to a recourse problem solution with slightly lower total cost but this effect may be distributed to the inaccuracy of the forecasts. The comparison demonstrates the stochastic model's capability of forecasting energy flows. The stochastic model assists decision makers to better understand how the uncertain fuel costs would affect future flows within the national electric energy system.

CHAPTER 1 INTRODUCTION

1.1 Motivation

Our life style would be unthinkable without the use of electric power. The growing utilization of electric energy is a decisive prerequisite for a rapid development of industry and agriculture. To meet the demand of electrical energy which increases by 4% to 7% per year in industrial countries, considerable amounts of primary energy carriers such as coal, petroleum or natural gas must be provided for power generation. Power plants together with the production and transmission of fuels compose a network with complex structure and many uncertain factors such as fuel price are involved in the system. As a reduction of the resources of primary energy carriers takes place all over the world and the fuel prices increase continuously, there has been a great concern about both technical and economic efficiency of the production of electrical energy. Given its inherent nonlinearities and uncertainties, remarkable efforts have been made to achieve a concise and comprehensive understanding of the large electric power network and to find more economic and more reliable ways to assemble and operate it.

Quelhas (35) constructed a decision model to account for the interdependencies across time and space in the U.S. bulk energy transportation system. It included subsystems for major fossil fuels and electricity. This is a generalized minimum cost network flow model which is constituted by coal and natural gas supply and storage, electricity generation and the energy flows among them. As fuel inventories may be carried over from one period to another, the model was extended to a dynamic domain with multiple periods. After validation with year 2002 data, an overall optimization was performed at the national level and the result provided insights into ways to increase the economic efficiency of the national energy system such as better utilization of low cost generators and increase in electric power trade. While the model offers related decision makers a comprehensive analysis on the national energy system, its formulation assumes that all information is known with certainty in advance. A question is raised by researchers: can we use this model to understand the effect of uncertainty on energy movements?

The reason to propose this question is quite natural given the uncertainty involved in the energy system. Multiple factors such as severe weather, equipment failures and international political events affect fuel prices, electric supply/demand and energy transportation. Some of the uncertain elements may cause a high cost to satisfy energy demands and some even lead to serious consequences, for example, large-scale disruption of energy supply. In 2005, hurricanes Katrina and Rita hit the Gulf of Mexico area. The catastrophic event not only interrupted the local electric and coal supplies but also damaged the natural gas production and transportation facilities, which caused significant nationwide impacts. The huge potential effects caused by the great uncertainty associated with the energy system motivate us to include uncertainty in the forecast elements within the model and study their effects using stochastic programming.

Stochastic programming has been applied to numerous energy models to address the problem of uncertain prices and demand. However, most of the research in the literature is limited to regional models or a single energy resource because of the spatial complexity and the interdependencies among various resources. Therefore, it would be interesting and meaningful to research the bulk energy transportation model with stochastic programming and address solutions to provide practical guidance for the U.S. power generation and transmission systems.

1.2 Objective

The first problem we need to research is which factors to be modeled as stochastic elements and how to measure the uncertainty mathematically. The selection criterion would be importance of the uncertain factors. For example, we might not want to investigate what would happen if the coal transportation capacity on a certain route varies because the situation rarely happens and has little effect on the whole system. However, if the price of natural gas from the Gulf of Mexico fluctuates, power generation cost all over the country would be influenced because electricity generated from natural gas is also used to satisfy peak demand and the Gulf of Mexico is a major supply of the fuel. Besides importance, we should also consider the tractability of the corresponding model. If it is difficult to capture the distribution of a selected variable, then it seems that we could not define the uncertain elements mathematically in the model and could say nothing of a solution.

After the stochastic model is constructed, we should collect necessary data and draw solutions from the model with available methodologies that are used to address stochastic programming models. Since the solution is a prediction of energy movements, it will be compared to the actual flows for judgment. Also, comparison between stochastic technique and deterministic approach weighs whether the additional computation cost for stochastic programming is worthwhile or not.

In summary, the objectives are to:

- Describe the important uncertain elements in a proper way and build the stochastic bulk energy transportation model;
- Implement the model with appropriate data and solve for optimal flows;
- Judge the value of the stochastic model by comparing the solution to both actual flows and the result from the deterministic model.
- State the lessons indicated by the stochastic model and how the model performs in projection of energy movements.

1.3 Thesis Organization

Chapter 2 reviews the relevant literature about integrated energy systems and stochastic programming. The formulation of the stochastic energy model is presented in Chapter 3 and a small numeric example in this chapter illustrates the modeling methodology. Chapter 4 provides a detailed description of the model structure, data collection and the complete procedure for obtaining the solution of the optimization problem. Visualized results of both stochastic and deterministic models are presented and compared in Chapter 5. Concluding remarks and directions for future work follow in Chapter 6.

CHAPTER 2 LITERATURE REVIEW

2.1 The National Energy System

Due to the limited data availability and the complex interaction between subsystems, most energy models built in the literature are narrowed into contract/utility/region level and focus on one aspect of the whole system. Petroleum product, electric power and fuel supply and transmission systems are therefore investigated separately regardless that they are highly interconnected. However, since 1974, the Federal Energy Administration (FEA) and its predecessor, the Energy Information Administration (EIA), have developed a series of three computer-based, medium term energy modeling systems to analyze domestic energy-economy markets and the relationship among electric energy and all kinds of fuels.

The Project Independence Evaluation System (22) was the first of the three systems and was employed by the FEA prior to 1982. It was initiated in 1974 to provide a framework for the developing a national energy policy through quantitative analysis and projections of the energy system. PIES considered several objectives including fuel price sensitivity, fuel competition (the possibility of the substitution of one energy source for another), technology restriction or improvement, resource limitations, economic impact, regional variations and other external effects on the energy system. Given the large volume of information and highly interdependent nature, a modular system was employed to permit the integration of subsystems, expansion of major components and the introduction of new elements. Figure 2.1 depicts the framework of PIES where the supply, demand, and equilibrium balancing components are combined with models of the economy, assessments of non-energy resource availability, and report writers that evaluate energy solutions in terms of the environmental, economic or resource impacts. PIES successfully analyzed the U.S. national energy system with an organization of engineering, econometric, and optimization models and improved the decision-making process for the complicated large-scale, time dependent system.

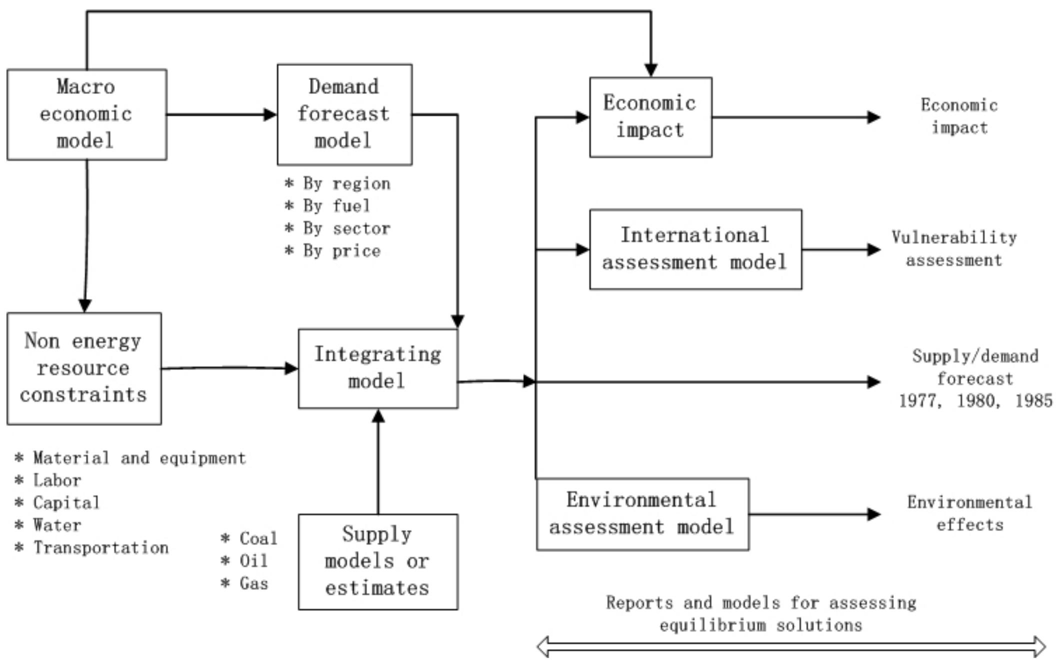

Figure 2.1 PIES Structure (22)

In 1982, PIES was updated to the Intermediate Future Forecasting System (IFFS) (29) which was used by EIA through 1993. While keeping the major objectives the same as the PIES, IFFS made a significant modification to the structure of model design, as shown in Figure 2.2. PIES built sub-models according to functions such as supply, demand and other constraints, keeping corresponding information about all the fuels in the same block. However, with the period of comprehensive energy legislation ending in the late 1970s, energy issues became more fuel specific, which motivated a model structured by fuels rather than functions. A simple integrating routine coordinates across the fuels and steps from submodel to submodel in order to capture the interaction among fuels. The new structure decomposes the model into manageable units which adopt diverse methodologies and are developed by individual groups with detailed knowledge of certain fuels. Compared to the PIES in which the person responsible for the integrating methodology becomes unreasonably overburdened by the developmental runs needed to test changes in submodels, IFFS is partitioned by fuel to avoid the complex task of integration and to balance the workload among the staff in charge of submodels.

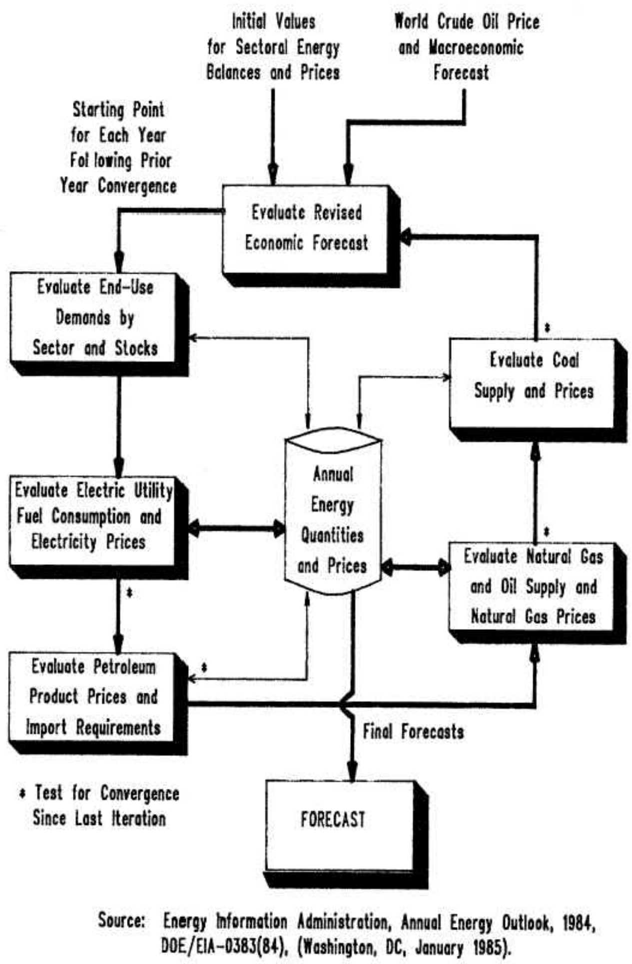

Figure 2.2 IFFS calculation flow (29)

In 1993, the IFFS was replaced by the National Energy Modeling System (NEMS) (14), which again had a new system structure. As depicted in Figure 2.3, NEMS takes advantages from both PIES and IFFS. There are two levels of subsystems. The first level is composed by function components of Supply, Conversion and Demand. Within the function blocks of Supply and Conversion, submodels are built for individual fuels, while Demand is partitioned according to end-users. Associated with advanced modeling and optimization techniques and the latest computing machines, the NEMS combines and processes more energy information than its predecessors and therefore is more capable with projections. In addition to the baseline forecast Annual Energy Outlook, NEMS generates one-time analytical reports and papers to analyze the effects of environmental impacts, existing government regulations and alternative energy policies. The system is used to test different assumptions about energy markets, to evaluate the potential impacts of new and advanced energy production, conversion and consumption technologies. It has been used for special analysis at the request of the White House, U.S. Congress, other offices of the Department of Energy who specify the scenarios and assumptions, which means the analysis produced by NEMS has an important effect on how the U.S. government regulates the energy markets. However, it is not open-source and not available for researchers and utilities to plan with.

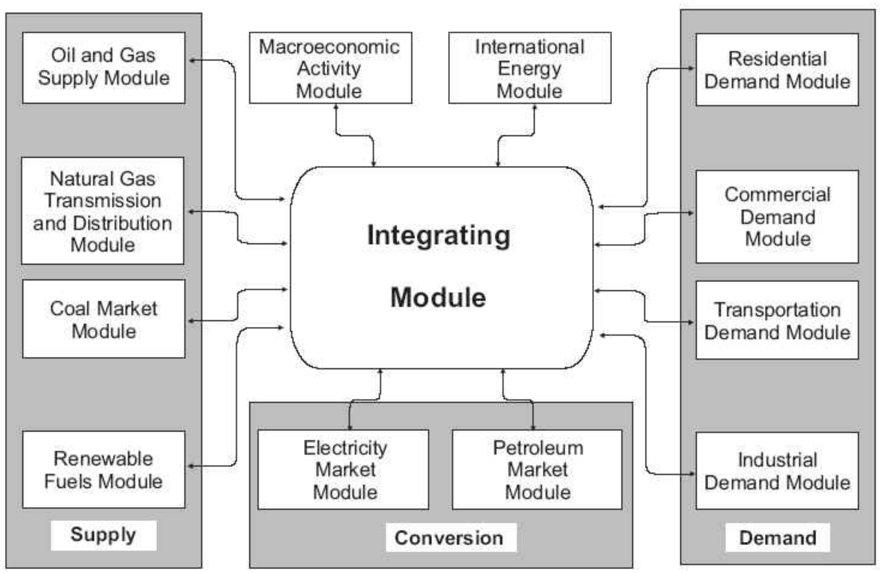

Figure 2.3 National Energy Modeling System

Quelhas developed a generalized network flow model for the U.S. electric energy system to explore economic efficiency of the energy flows from fuel suppliers to electric load centers (36). Within this decision model, fuel production, transportation, storage, electricity generation and transmission are represented by nodes and arcs included in the generalized network which is a three-level system: Coal, natural gas and electricity are partitioned into corresponding levels and connected by energy movements among different levels. All the data in this model are derived from various public available sources, such as the websites of the Energy Information Administration and the Canadian National Energy Board. The model was validated by comparing its output to the actual data published by EIA for 2002 (37). With the objective of cost minimization at the national level, the model is constrained by electricity generation/demand, fuel supply/demand and transmission capacities. It can be solved efficiently by network optimization codes and is expected to enable both public and private decision makers having limited available data and other resources to better understand the complex dynamics of interdependencies of primary fuels and electricity networks and carry out comprehensive analysis of a wide range of issues related to the energy sector.

2.2 Stochastic Programming

Whereas deterministic optimization problems are formulated with known parameters, real world problems almost invariably include some unknown parameters. Randomness in problem data poses a serious challenge for solving many linear programming problems, through which the solutions obtained are optimal for the specific problem but may not be optimal for the situation that actually occurs. Stochastic programming (SP) is a framework for modeling optimization problems that involve uncertainty. This field is currently developing rapidly with contributions from many disciplines including operations research, mathematics, and probability. Conversely, it is being applied in a wide variety of subjects ranging from agriculture to financial planning and from industrial engineering to computer networks.

The fundamental idea behind stochastic programming is the concept of recourse, which introduced by Dantzig (11) and Beale (6) independently. Recourse is the ability to take corrective action after a random event has taken place. The most widely applied and studied stochastic programming with recourse models are two-stage linear programs. Here the decision maker takes some action in the first stage, after which a random event occurs, affecting the outcome of the first-stage decision. A recourse decision can then be made in the second stage that compensates for any bad effects that might have been experienced as a result of the first-stage decision. The optimal policy from such a model is a single first-stage policy and a collection of recourse decisions defining which second-stage action should be taken in response to each random outcome. One natural generalization of the two-stage model is to extend it to many stages, each of which consists of a decision followed by a set of observations of the uncertain parameters that are gradually revealed over time. Mulvey and Vladimirou (30) specified stochastic programming to networks by dividing nodes and arcs into separate sets corresponding to the stage to which they belong. They also develop a specific decomposition method for solving multistage stochastic networks.

An alternative type of stochastic programming approach is so-called Chance-constrained stochastic programming, which was first introduced by Charnes and Cooper (10). It does not require that decisions be feasible for every outcome of the random parameters. It tries to find a decision which ensures that a set of constraints will hold with a certain probability. An application might be a delivery service that experiences random demands, and wishes to find the cheapest way to deliver its packages with a high probability.

While stochastic programming is usually characterized by a probability distribution on the parameters, robust optimization, which is a further development of chance-constrained SP, can tackle the problems where the parameters are only known within certain bounds. The goal is to find a solution which is feasible and acceptably close to optimal for all such data. Research with main contributions to the foundation of Robust Optimization includes Ben-Tal and Nemirovski (4) and Kouvelis and Yu (25). Bertsimas and Sim (5) presented a robust optimization approach which set up a parameter to control the level of robustness against conservatism. This method provides a solution satisfying a high proportion (which depends on the parameter set) of the constraints even for the worst situation.

2.3 Stochastic Programming Models in Energy

Stochastic programming models are widely used in the area of optimal allocation of energy and its related resources, where demand and prices are always unpredictable (41). Those models in power systems planning are usually divided according to the planning horizon. Long term planning models deal with 15-20 year large investments such as building thermal units and constructing hydro reservoirs and turbines. This kind of model helps us find the optimal investment to meet the uncertain future demand. Regularly, several possible future load duration curves are put forward and a straightforward recourse model is developed to address the recourse solutions for different scenarios. Murphy et al. (28) carried out a deterministic investment analysis using a new load duration curve aggregated from predicated curves and obtained the same solution as if the recourse problem is solved. Sherali et al. (39) (40) emphasized peak load pricing and discuss Murphy's model in greater detail. Gardner and Rogers (19) investigated a multi-stage problem where load duration curves are revealed over time and investments are made stage by stage. While all the demand must be satisfied in traditional monopoly-based production planning, Qiu and Girgis (33) look at the problem from a different perspective by allowing and pricing outages, which takes into account that something even worse than the worst scenario modeled could occur with the consequence of shortage.

Medium-term power planning has a 1-3 year horizon and usually deals with reservoir management, where the true cost and risk brought by the uncertain aspects of using the water are underestimated by deterministic solutions and the performance of stochastic optimization models is proved to be significant. Short term planning typically deals with problems with horizons of one week or shorter, such as unit commitment and economic dispatch.

All of the above models were developed for regulated markets. The transition of electricity markets from the old regulated regime to the deregulated system motivated the development of hybrid stochastic models where there is both a demand constraint and a wholesale market, where the producer can choose to serve the local load by his own production capacity or by buying capacity. Some stochastic programming models serve the needs of utility planners and policy makers in that they can generate derive scenarios for market prices of electricity. Important papers include Fragniere and Haurie (17), Botnen et al. (8) and Hindsberger (21).

As the electricity markets are developing into regional commodity markets, the use of standardized financial contracts such as forward contracts increases. The contract price represents the current market value of future delivery of the electricity. Hence, valuation of future production is needed in stochastic programming models in energy. These models are based on describing the uncertainty in the form of scenarios of the spot price of the commodity. Since basing the scenarios on forecasts of spot prices will not give a valuation that is consistent with the market, the stochastic programming models are in a position to value the decision flexibility using a price of risk that is consistent with the market.

Energy bidding is viewed as a short term optimization problem in which the market participant offers to buy or sell capacity to the market in the form of price-quantity pairs for given time intervals. Determining optimal bids to send to the market operator becomes a nontrivial task that can be supported by stochastic programming models. Nowak et al. (32) study this problem and present an integrated stochastic unit commitment and bidding model. Neame et al. (31) and Anderson and Philpott (1) also developed stochastic models to explore optimal energy bidding prices.

Operations scheduling in deregulated markets is divided into two categories. In the first set of problems, generation utilities are not large enough to influence electricity prices by changing the amount of generation capacity offered to the market. Scott and Read (38) investigated the other class of models in which the operators do have market power on energy price. A major limitation in these analyses is that buying and selling of contracts is in reality determined simultaneously with production.

Financial instruments such as trading in the forward market are used to reduce risk in energy market. However, since fixing income in the future does not automatically mean reduced risks, researchers made great efforts on stochastic models that manage the risk of energy trading. Mo et al. (27) and Fleten et al. (18) suggest that production scheduling and contract risk management should be integrated in order to maximize expected profit at some acceptable level of risk. However, other researchers claim that the benefits of a decoupled set of models will probably outweigh the small theoretical gain from integrating production planning and trading. All in all, the deregulated markets have not found their final forms and there are a lot more topics we can discuss and research with the tool of stochastic programming.

CHAPTER 3 MODEL FORMULATION AND ILLUSTRATIVE EXAMPLE

3.1 Deterministic Model and Its Notations

The researched national electric energy system is aggregated at a regional level which is based on the topology of the electrical grid, availability of aggregated data, and operating constraints. It is an adequate simplification of the physical and institutional complexity of the electric power industry given that data are generally available at this level (36). The whole system is modeled as a generalized minimum cost flow network. The nodes represent coal mines, natural gas wells, natural gas storage facilities and electricity transmission centers. The flows between these nodes include fuel transmission/storage and electricity transmission/subregional trade. The arc coefficients denote the efficiency of energy movement or the transferring rate from fuel to electric energy. The mathematical formulation of this model is as formula 3.1 (36). Table 3.1 shows notations used in the formula.

$$
\begin{aligned}
&Min \quad z = \textstyle\sum_{t \in T} \sum_{(i,j) \in A} c_{ij}(t) e_{ij}(t) \\
s.t. \quad &\textstyle\sum_{(j,k) \in A} e_{jk}(t) - \sum_{(i,j) \in A} r_{ij}(t) e_{ij}(t) = b_j(t) \quad \forall j \in N, \forall t \in T \\
&e_{ij,min} \leq e_{ij}(t) \leq e_{ij,max} \quad \forall (i,j) \in A, \forall t \in T
\end{aligned}
\tag{3.1}
$$

3.2 Solve Stochastic Problem via Deterministic Equivalent

Solving stochastic network flows involves both the description of uncertain elements and the methodology chosen to deal with uncertainty, which are interdependent and could not be fixed separately. On one hand, the mathematical assumption in the chosen method should

Table 3.1 Notations in deterministic model

t	The t^{th} time period.
$e_{ij}(t)$	Energy flowing from node i to node j during time t.
$b_j(t)$	Supply (if positive) or demand (if negative) at node j during time t.
$e_{ij,max}$	Upper bound on the energy flowing from node i to node j.
$e_{ij,min}$	Lower bound on the energy flowing from node i to node j.
$c_{ij}(t)$	Per unit cost of the energy flowing from node i to node j during time t.
$r_{ij}t$	Efficiency parameter associated with the arc connecting node i to node j during time t.
A	Set of arcs, $\{(i,j)\}$.
N	Set of nodes, $\{j\}$.
T	Set of time periods, $\{t\}$.

be appropriate for the description of uncertainty in this model. On the other hand, while stating the uncertain elements mathematically, we should also consider whether it is possible to estimate them from data available.

As stated in chapter 1, we study only uncertain fuel cost and demand in this research and it is reasonable to formulate them as discrete random variables taking a finite number of realizations, which describe how the price and demand can fluctuate. The assumption of discrete distributions for the uncertain elements is common in most stochastic programming approaches, which enables us to solve the problem with the famous two-stage approach (30).

To simplify the explanation of two-stage approach, we suppress the notation t and only use the ordinary term ij to differentiate variables and parameters in different periods.

We model the cost per unit flow on a fuel acquisition arc as a random variable with K possible values:

$$Pr\{c_{ij} = c_{ij}(1)\} = p_{c_{ij}}(1), Pr\{c_{ij} = c_{ij}(2)\} = p_{c_{ij}}(2), \ldots, Pr\{c_{ij} = c_{ij}(K)\} = p_{c_{ij}}(K).$$

Similarly, the electricity load is modeled on a demand node as a random variable with L possible values:

$$Pr\{b_j = b_j(1)\} = p_{b_j}(1), Pr\{b_j = b_j(2)\} = p_{b_j}(2), \ldots, Pr\{b_j = b_j(L)\} = p_{b_j}(L).$$

Assume that there are m random cost variables and n random demand variables in the model; for the assumption of discrete distribution, we can define a scenario $s \in S$ for each combination of values:

$$\pi_s = Pr\{c_{(ij)_1} = c_{(ij)_1}(k_1), \ldots, c_{(ij)_m} = c_{(ij)_m}(k_m), b_{j_1} = b_{j_1}(l_1), \ldots, b_{j_n} = b_{j_1}(l_n)\}.$$

We also assume all the random variables are independent. Thus,

$$\pi_s = p_{c_{(ij)_1}}(k_1) \ldots p_{c_{(ij)m}}(k_m) p_{b_{j_1}}(l_1) \ldots p_{b_{j_n}}(l_n).$$

In the two-stage approach, all the arcs are divided into three sets (30). The flows on sets of first-stage arcs are decided in the first stage, then the values of all uncertain quantities are revealed and the flows on the second-stage arcs are set. Generally, the sets A of arcs and N of nodes are partitioned into disjoint subsets as follows:

- A_1= {arcs representing first-stage decisions for which all associated parameters are deterministic} ;
- A_1'= {arcs representing first-stage decisions that have associated stochastic costs, capacities or multipliers} ;
- A_2= {arcs corresponding to second-stage decisions} ;
- N_1= {nodes with all the incoming arcs in A_1 } ;
- $N_2 = N \backslash N_1$.

We also distinguish between first-stage and second-stage flows as:

- $X = \{x_{ij} | (i, j) \in A_1 \bigcup A_1'\}$;
- $Y = \{y_{ij} | (i, j) \in A_2\}$, so that $\{e_{ij}\} = X \bigcup Y$.

Finally, for node i, denote the set Δ of incident out-arcs and in-arcs as, respectively,

- $\Delta_i^+ \equiv \{(i, j) \in A\}$ and $\Delta_i^- \equiv \{j, i) \in A\}$.

Each scenario subproblem is a generalized network with the fixed topology of the given one realization of the uncertain costs and demands. The subproblem for scenario $s \in S$ is stated as formula 3.2, in which $x_{ij}(s)$ is the flow on a first stage arc in scenario s and $y_{ij}(s)$ is the flow on a second stage arc.

$$Min \quad f_s(x(s), y(s)) = \sum_{(i,j)\in A_1} c_{ij}x_{ij}(s) + \sum_{(i,j)\in A_1'} c_{ij}(s)x_{ij}(s) + \sum_{(i,j)\in A_2} c_{ij}(s)y_{ij}(s)$$

$$Subject \quad to \quad \sum_{(i,j)\in \Delta_i^+} x_{ij}(s) - \sum_{(j,i)\in \Delta_i^-} r_{ji}x_{ji}(s) = b_i \quad \forall i \in N$$

$$\sum_{(i,j)\in \{\Delta_i^+ \bigcap A_1\}} x_{ij}(s) - \sum_{(j,i)\in \{\Delta_i^- \bigcap A_1\}} r_{ji}x_{ji}(s) + \sum_{(i,j)\in \{\Delta_i^+ \bigcap A_1'\}} x_{ij}(s) - \sum_{(j,i)\in \{\Delta_i^- \bigcap A_1'\}} r_{ji}x_{ji}(s)$$

$$+ \sum_{(i,j)\in \{\Delta_i^+ \bigcap A_2\}} y_{ij}(s) - \sum_{(j,i)\in \{\Delta_i^- \bigcap A_2\}} r_{ji}y_{ji}(s) = b_i(s) \quad \forall i \in N_2 \tag{3.2}$$

$$l_{ij}^x \leq x_{ij}(s) \leq u_{ij}^x \quad \forall (i,j) \in \{A_1 \bigcup A_1'\}$$

$$l_{ij}^y \leq y_{ij}(s) \leq u_{ij}^y \quad \forall (i,j) \in A_2$$

If we know which scenario would occur, we could solve just one subproblem. However, to jointly consider all the possibilities in the solution procedure, the values of the first-stage decisions are assumed to be invariant and thus we have $x(s) = x(s') = z \quad \forall s, s' \in S; s \neq s'$.Therefore, the overall program could be stated as the deterministic equivalent problem in formula 3.3, where X is substituted with Z and π_s is the probability of scenario s.

$$Min \quad \sum_{s\in S} \pi_s f_s(z, y(s)) = \sum_{(i,j)\in A_1} c_{ij}z_{ij} + \sum_{s\in S} \pi_s[\sum_{(i,j)\in A_1'} c_{ij}(s)z_{ij} + \sum_{(i,j)\in A_2} c_{ij}(s)y_{ij}(s)]$$

$$Subject \quad to \quad \sum_{(i,j\in \Delta_i^+)} z_{ij} - \sum_{(j,i)\in \Delta_i^-} r_{ji}z_{ji} = b_i \quad \forall i \in N$$

$$\sum_{(i,j)\in \{\Delta_i^+ \bigcap A_1\}} z_{ij} - \sum_{(j,i)\in \{\Delta_i^- \bigcap A_1\}} r_{ji}z_{ji} + \sum_{(i,j)\in \{\Delta_i^+ \bigcap A_1'\}} z_{ij} - \sum_{(j,i)\in \{\Delta_i^- \bigcap A_1'\}} r_{ji}z_{ji} \tag{3.3}$$

$$+ \sum_{(i,j)\in \{\Delta_i^+ \bigcap A_2\}} y_{ij}(s) - \sum_{(j,i)\in \{\Delta_i^- \bigcap A_2\}} r_{ji}y_{ji}(s) = b_i(s) \quad \forall i \in N_2, \forall s \in S$$

$$l_{ij}^z \leq z_{ij} \leq u_{ij}^z \quad \forall (i,j) \in \{A_1 \bigcup A_1'\}$$

$$l_{ij}^y \leq y_{ij}(s) \leq u_{ij}^y \quad \forall (i,j) \in A_2, \forall s \in S$$

Say $|Z| = n_1$ and $|y(s)| = n_2$. There are m_1 constraints not related to the second stage arcs Y and m_2 constraints related to Y. The total number of different scenarios is $|S|$. Hence, the size of this deterministic equivalent formulation is $n_1 + |S|n_2$ variables and $m_1 + |S|m_2$ constraints. Solving 3.3, we get a feasible solution $(z, y(s))$ for each scenario s. The significance is that no matter which scenario would be realized, the flows on the arcs in set Z are deterministic and not affected by uncertain factors. Moreover, because the objective is the expected value of objective functions for each scenario, all the scenarios are considered jointly. Although $f_s(z, y(s))$ may not be as low as the optimal objective functions for each single scenario, it is relatively good for most of the scenarios, especially when it is uncertain which one will eventually become truth.

3.3 Numeric Example and Solutions

We apply the two-stage approach to a single period integrated energy system with 2 coal suppliers, 2 natural gas suppliers, 5 generation plants and 2 transmission centers (34). The network is shown in Figure 3.1, in which $A_1 = \{x2, x3, x4, x5, x6, x7\}$, $A_1' = \{v2, v3, v4\}$, $A_2 = \{x1, x8, v1, v5, exp, imp\}$, $N_1 = \{Coal1, Coal2, 2, 3, 4\}$, $N_2 = \{NG1, NG2, 1, 5, North, South\}$. With this case, we also explain and compare different sets of solutions.

Table 3.2 Scenarios of single-period numeric example

Scenario	NG Price ($ / mcf)	Northern Demand (MWh)	Probability
(d-, p-)	2	25000	0.4*0.6833=0.2733
(d+, p-)	2	37000	0.4*0.3167=0.1267
(d-, p+)	3	25000	0.6*0.6833=0.4100
(d+, p+)	3	37000	0.6*0.3167=0.1900
(d, p)	2.6	28800	—-

The southern NG cost (associated with arc $(NG2, 5)$) and northern demand are stochastic with the four possible scenarios shown in Table 3.2. It is assumed that the uncertain cost and demand are independent. Because a coal contract lasts for a long time and natural gas contract tends to be much shorter as a result of floating prices, coal arcs are put in the first-stage set while natural gas and transmission flows are decided in the second stage. Solve the recourse

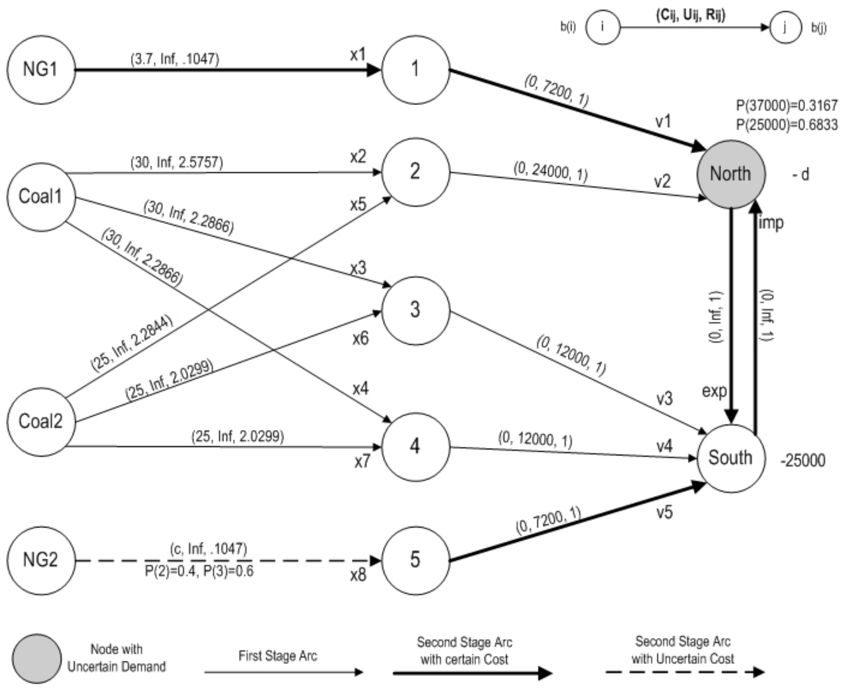

Figure 3.1 Numeric example: single-period electric energy system

problem by the deterministic equivalent 3.3 and the solution is called the ***Recourse Problem solution (RP)***, illustrated in Figure 3.2. Figure 3.3 shows the optimal flows Z on the first stage arcs and values of the recourse flows for every scenario on the second-stage arcs. As we can find in Figure 3.3, when demand is low in the North, RP uses more southern natural gas when the price is at its lower level, otherwise more northern natural gas is consumed. And Nouth exports only when its demand is low and South NG price is high.

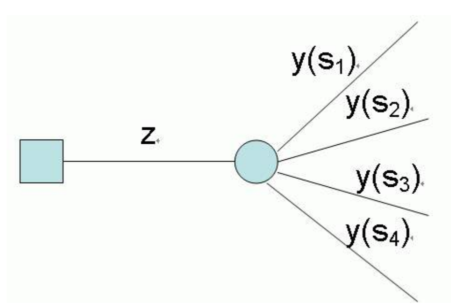

Figure 3.2 Recourse Problem Solution

We have so far embarked on formulating and solving stochastic programming models with-

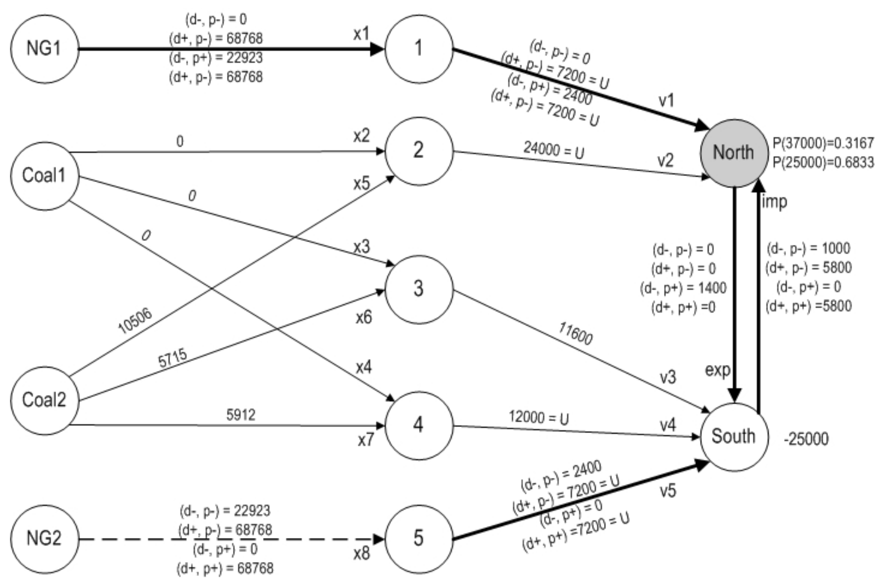

Figure 3.3 RP solution of the single-period numeric example

out much concern about whether or not it is worthwhile to do so. Hence two special terms are introduced to analyze how the decisions are affected by including uncertainty and using stochastic programming (24).

First we compare the optimal objective value of the stochastic model to the ***wait-and-see solution (WS)*** which is calculated by finding the expected value of the optimal solutions for each scenario and illustrated in figure 3.4. In the WS solution, the decision maker knows which scenario will occur before making the first-stage decisions. In a cost-minimization problem, $WS \leq RP$ and the difference between them is called the ***expected value of perfect information (EVPI)***, since it shows how much one could expect to gain if one were told what would happen before making decisions. Another interpretation is that the difference is what one would be willing to pay for that information. A large EVPI shows that randomness plays an important role in the problem, but it does not necessarily show that a deterministic model cannot function well. However, a small EVPI means that randomness plays a minor role in the model (24). Figure 3.5 shows wait-and-see solutions. Note that some first-stage flows are invariant over all scenarios, but others such as those for (Coal2, 3) and (3, South) depend on the uncertain price and demand. In this example, we have ***RP =1256824; WS =1249200;***

EVPI = RP − WS = 7624 which is 0.6% of the WS objective function.

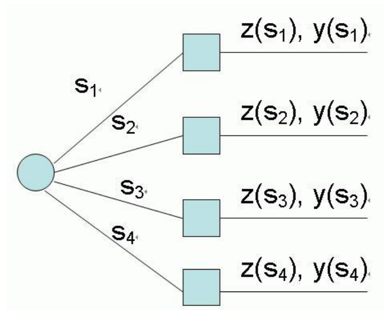

Figure 3.4 Wait-and-See solution

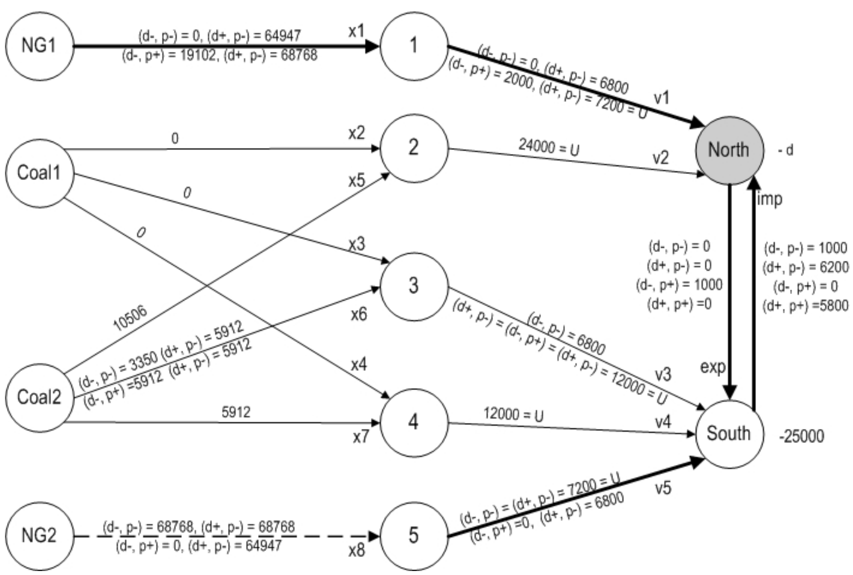

Figure 3.5 Wait-and-See solution of the single-period numeric example

The other term concerned is called the value of the stochastic solution ***(VSS)*** which is the difference of RP and EEV *(expected value of the expected value solution)*. A common approach to decision-making in an uncertain environment is to solve a deterministic problem using the expected value of each random variable. EEV is the optimal cost in the stochastic problem with the first-stage variables fixed at the values obtained from solving the deterministic problem when substituting expected values for the random variables. The second stage variables for each scenario describe the optimal recourse in that scenario given the fixed first-stage values. Because the expected value solution *(EV)* is feasible for the stochastic problem, $EEV > RP$.

The value of the stochastic solution is what one would pay for the additional solution effort to solve the recourse problem. And in this example we get ***RP =1256824; EEV = 1257079; VSS = EEV − RP = 255***. Figure 3.6 shows the expected value solution.

Averaged flows in EV, RP and WS are compared in figure 3.7. Overall, RP is closer to WS than EV and has a lower total cost. EV differs from RP most in decreased use of southern natural gas, which also leads to a difference in greater use of NG from NG1. In some scenarios, RP uses the southern natural gas because when the price is as low as $2, it is cheaper to use natural gas than the coal from Coal2. However, when EV decides the first-stage variables, it takes the expected price of $2.6 which is higher than coal and ignores the possibility that the southern NG price is lower. Hence, EV solution uses as much coal as possible. Apparently, RP is better because it keeps features of each scenario, while EV loses them in the cause of averaging.

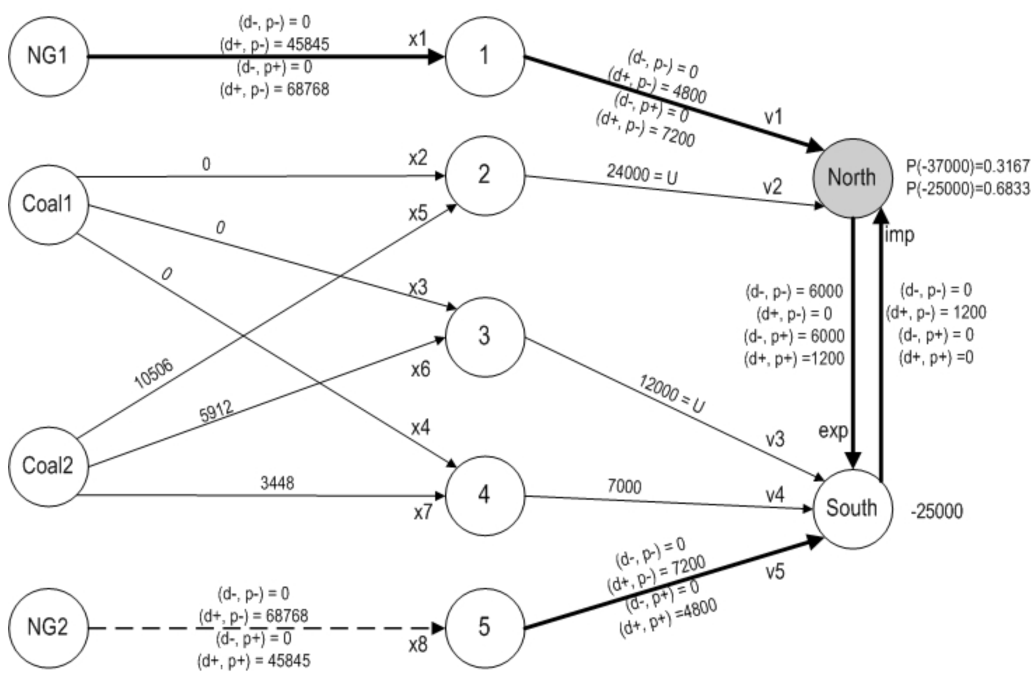

Figure 3.6 EV solution

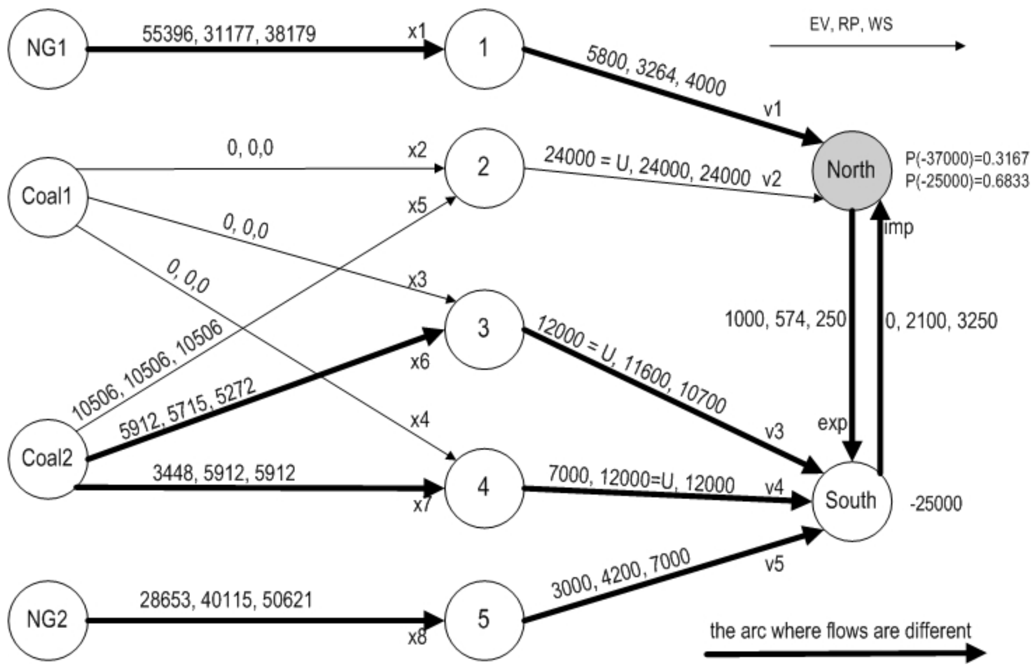

Figure 3.7 EV vs. RP vs. WS

CHAPTER 4 IMPLEMENTATION

4.1 Model Validation

In the model of the US national electric energy system, 11 coal supply nodes and 14 natural gas supply nodes are aggregated at a regional level regardless of real mine or well locations. The 17 nodes of electric transmission centers represent the NERC demand regions among which electricity is traded and transferred. For each demand region, energy generation plants are aggregated to one node if they use the same fuel type and prime mover. There are 6 different types of plants and totally 102 generation nodes in the system. Flows between the nodes represent the transportation of fuel and electric energy. The whole system is modeled as a generalized minimum cost flow network. With year 2002 data, monthly natural gas and electricity nodes and yearly coal nodes, there are totally 1290 nodes, 3480 arcs in this model. Demand, capacities and flows represent monthly (natural gas and electricity) or yearly (coal) totals. (35)

Quelhas (37) verified the model formulated in chapter 3 by comparing results from the model to actual data. As shown in Table 4.1, the first column is actual coal and NG deliveries in year 2002 and the other two columns are total flows calculated from the model. In case A, optimized coal and NG flows are solved by fixing generation and demand according to the actual data at each electricity transmission center, while Case B is solved only with fixed demand of electric power. The small difference between Case A and the actual data validates the model in the terms of the values of arc efficiencies and capacities. Comparing Case A to Case B, greater economic efficiency would be achieved if more coal is bought and more electricity is traded between sub regions.

Table 4.1 Total flows comparison: 2002 actual data and the model

Result	Actual	Case A	Case B
Coal deliveries (million tons)	976	953	1,054
Natural gas deliveries (million mcf)	5,398	5,125	3,615
Electricity net trade (thousand GWh)	N/A	205	306

4.2 Rolling Two-stage Procedure

The solution in Case B is optimal for the whole system given that all data on costs, capacities, supplies, demands and efficiencies are known with certainty beforehand. However, it is impossible to achieve this good solution in reality because in January 2002, decision makers did not know what the exact price of natural gas would be in October 2002, for instance. Instead, they had to base their decisions on forecasts of future costs. As discussed in chapter 3, the stochastic problem with uncertain fuel cost can be solved by a two-stage approach. However, there is a problem blocking us from simply applying the method to this case. The two-stage approach only works when we have no more than 2 periods because it is required that all uncertain elements are revealed at the beginning of the second stage. In our 2002 model, we have 12 periods (months) and the natural gas price for March would not be revealed at the beginning of February. To address this problem, we applied the two-stage approach repeatedly in a rolling procedure.

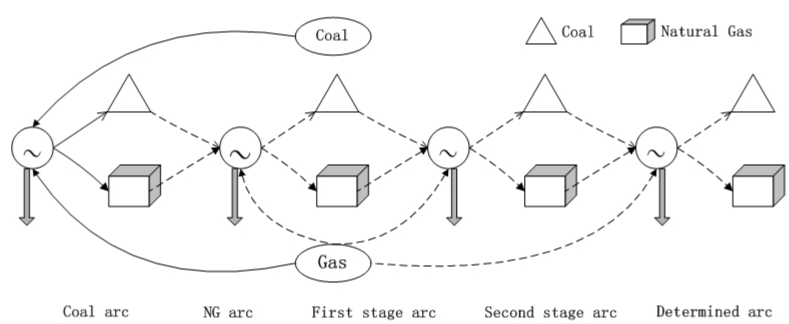

Figure 4.1 Rolling two-stage approach: the first period

This rolling procedure is illustrated by Figures 4.1 to 4.5 for a 4-period problem. We assume a single coal purchase is made in period 1 but natural gas is purchased each period.

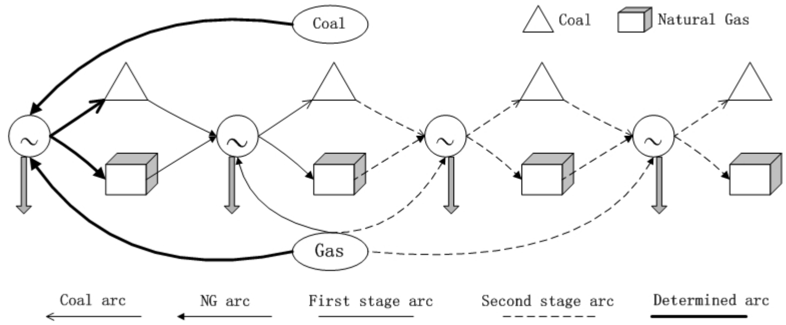

Figure 4.2 Rolling two-stage approach: the second period

All demands are assumed known. At the very beginning, period 1 is the first stage, and all the remaining periods are the second stage. One set of fuel cost forecasts is used to generate scenarios. After the problem has been solved with the 2-stage approach, we keep the first stage decisions. After one time step, we remove the period 1 decision variables and roll to period 2. This time, period 2 becomes the first stage, and period 3 and 4 are the second stage. With new information coming in, the fuel cost forecast can be adjusted. A new set of forecasts is used to generate scenarios. Then flows on period 2 are decided and we roll to the next period, adopting the updated price forecast. When it rolls to the last period, there is no uncertain cost anymore, so we just solve this small size deterministic problem. This procedure allows a simulation of the actual decision process in time steps as the decision-maker applies the two-stage approach repeatedly with updated information about uncertain elements.

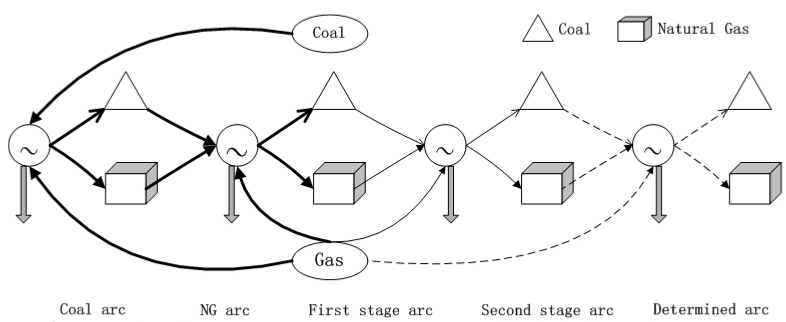

Figure 4.3 Rolling two-stage approach: the third period

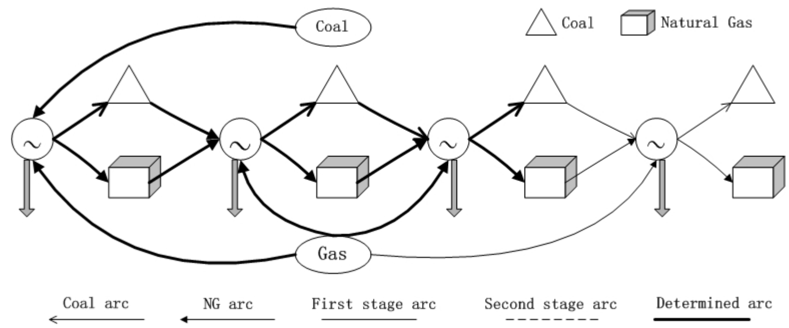

Figure 4.4 Rolling two-stage approach: the forth period

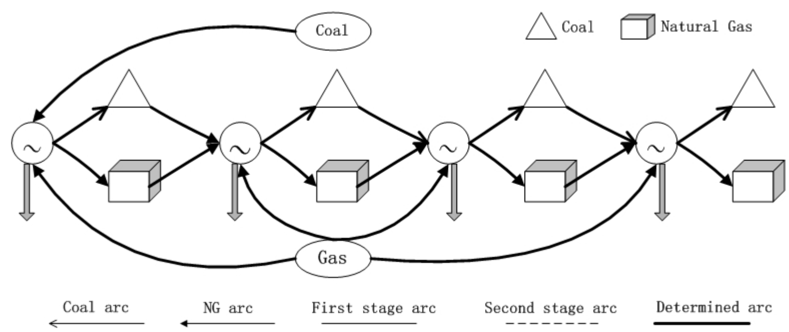

Figure 4.5 Rolling two-stage approach: the end

4.3 Price Forecast

To simulate the rolling procedure, we still need to generate scenarios with predicted fuel costs. The long term fuel cost graph in Figure 4.6 is from EIA Annual Energy Review. The coal price is quite flat so it is treated as fixed. Natural gas price is much more variable and therefore modeled as an uncertain cost in the stochastic model.

EIA provides a monthly updated Short Term Energy Outlook, which *"industry participants and energy analysts regularly adopt as a 'best estimate' of future energy outcomes"* (7). We use the 2002 data to generate scenarios. Figure 4.7 (15) was released in January 2002 with estimated NG prices for the whole year. Figure 4.8 (16), released in January 2003, has actual 2002 NG prices. The price estimate is in the rectangle in the first graph but the actual price in the second graph is out of the rectangle which indicates inaccuracy in price forecast. So even though the outlook from EIA is one of the most convincing data sets based on which utilities and others conducted resource planning and modeling studies, there still exists much

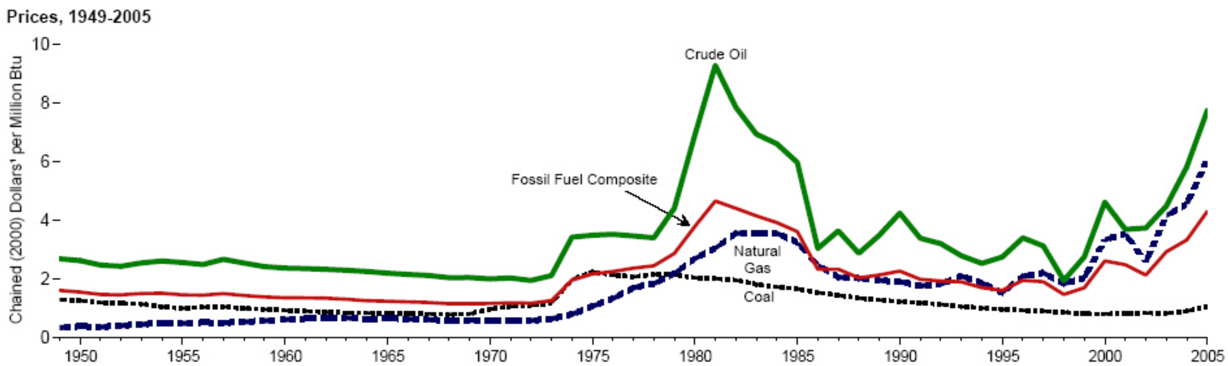

Figure 4.6 Long term fossil fuel cost trends

inaccuracy and uncertainty.

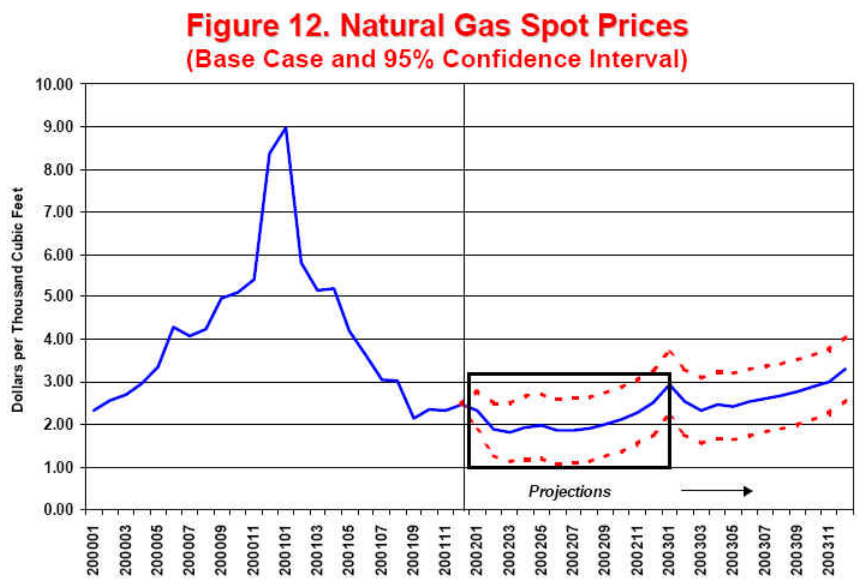

Figure 4.7 EIA short-term natural gas price outlook, Jan. 2002

Based on the EIA data, uncertain NG cost is modeled as a discrete random variable. There are 3 possible values for each period. The mean is set to equal the "base case" represented by the solid line. The low value is the lower confidence limit shown in Figure 4.7 and the high value is the upper confidence limit. Both extreme values have the same probability $p\{c_t = LCL = \widehat{x}_t\} = p\{c_t = UCL = \widehat{x}_t\} = p_t$, so that $p\{c_t = \widehat{x}_t\} = 1 - 2p_t$. The variance of the random variable $Var(c_t) = 2p_t(CIW_t)^2$ depends on both p and the width of the confidence

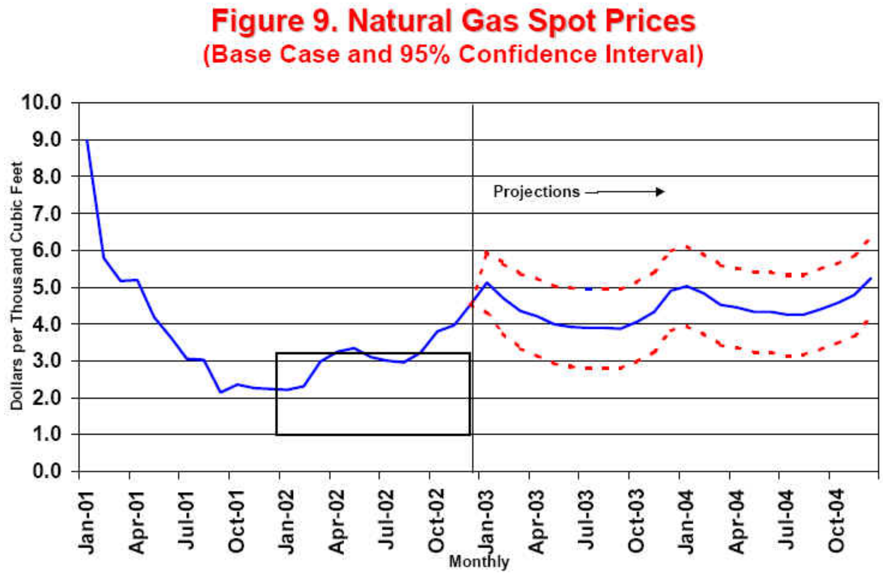

Figure 4.8 EIA short-term natural gas price outlook, Jan. 2003

interval. It is reasonable to set a larger value of p for more remote periods because we are more uncertain about the forecast. Case 1 is the base case we will investigate in next chapter. The "narrow" confidence interval is as wide as shown in the figure. In case 2, both p and CI are enlarged to study the effect of increasing uncertainty. The variance of the cost distribution in case 2 is 8 times that in case 1.

- Case 1: p_2=0.05, p_3=0.125, p_4=0.2, narrow CIW;
- Case 2: p_2=0.1, p_3=0.25, p_4=0.4, wide CIW = 2*narrow CIW.

Note that, whereas EIA predicts national average NG price, we use regional prices in the model. Given the fact that regional prices are generated by multiplying the national price by the regional factors (35), it is assumed that predictions of the regional prices have the same trend and are calculated by multiplying national price estimate by the factors. Since NG imports from Canada play a very important role in the U.S. national NG consumption, it is necessary to take those NG prices as uncertain elements, too. To generate the forecast for the price of natural gas imported from Canada, we first find the average gap between the actual NG prices in Canada and in U.S.A and then add the difference to the U.S. national NG price forecast.

4.4 Aggregation

The problem with the two-stage deterministic equivalent is that it enlarges the problem size. As mentioned in chapter 3, the deterministic model has 1290 nodes and 3480 arcs. Being consistent with the assumption that natural gas price has 3 possible values in each month from February to December, the biggest problem in the rolling procedure has $223 + (97 \times 11) \times 3^{11} = 189,016,072$ nodes and $521 + (29 \times 11) \times 3^{11} = 524,178,494$ arcs, which is absolutely non-trivial to solve on a regular PC without any decomposition.

To reduce the problem size, we aggregated the monthly model into monthly-quarterly (first three months and the rest three quarters) and quarterly models by adding up capacities and demands and averaging fuel costs and compared the flows in deterministic problem, as shown in Table 4.2. Since the maximum difference of flows between the monthly model and the quarterly model is only 3%, it is reasonable to aggregate the model into quarter level to achieve computational tractability. After aggregation, the problem size is reduced by several orders of magnitude. The largest deterministic equivalent has $157 + 296 \times 3^3 = 8149$ nodes and $521 + 807 \times 3^3 = 22310$ arcs, which could be solved by a 4G memory PC in less than 1 second.

Table 4.2 Total flows comparison: monthly, month-quarterly and quarterly

Result	**M**	**MQ**	**Q**	**(MQ-M)/M**	**(Q-M)/M**
Coal deliveries (million tons)	1,053	1,057	1,058	0.386%	0.497%
Natural gas deliveries (million Mcf)	3,615	3,560	3,608	-1.522%	-0.198%
Electricity generation from coal (million GWh)	2,117	2,121	2,121	0.197%	0.231%
Electricity generation from NG (million GWh)	414	410	409	-0.997%	-1.232%
Net trade (million GWh)	381	383	371	0.533%	-2.698%
Total costs (billion $)	96.896	96.850	97.102	-0.048%	0.212%

4.5 NG Consumptions Other Than Electricity

In the optimal solution of the deterministic model, more than 60 million dollars of the approximately $ 100 million total cost is spent on the NG consumptions other than electricity, which significantly affects the decisions on electric flows. However, we can not simply disregard this part of NG consumptions because flows used to meet those demands as well as the gas for electricity generation are constrained by total supply and pipeline capacity. Originally, the demands for other NG consumptions were assigned to the NG transmission nodes. To optimize the cost of meeting electric demand only while preserving these constraints, we modified the NG subsystem. One way to modify the network is to add a new set of nodes with those demands to the original problem where the arcs connecting the new nodes to the NG supply nodes are associated with zero cost. In this way, we eliminate the huge cost brought by other NG consumptions and keep their capacity impacts. The modification to the network is illustrated in Figure 4.9. If $C_4 = C_1$, then the new graph is the same as the old one. If $C_4 = 0$, then the cost for other NG consumptions is removed.

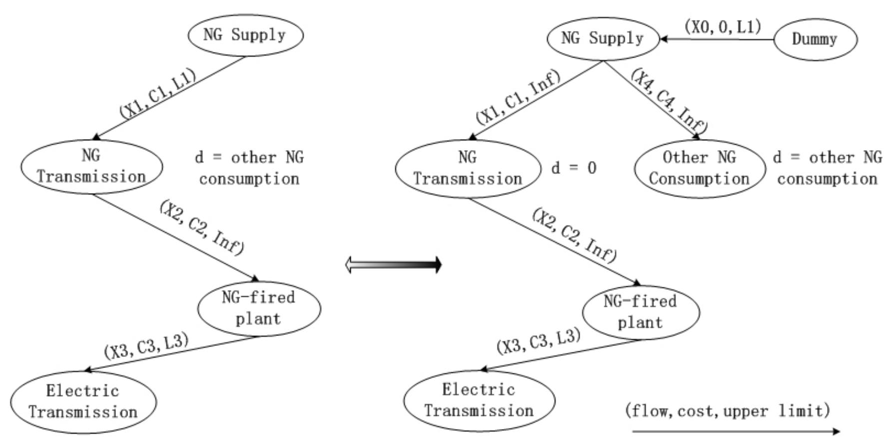

Figure 4.9 Addition node for NG consumptions other than electricity

Since the network is complex, it is non-trivial to modify it as illustrated in Figure 4.8. For example, the two types of natural gas flows should be combined with respect to capacity but separated according to costs. So we solve a relaxed network problem which has the same structure as the original network and check whether the optimal solution is feasible for the modified one, which indicates it is also the optimal solution for the modified network. The idea is shown is Figure 4.10. When the optimal solution to the relaxed problem is feasible for the modified network, the strategy avoids modification of the network and saves considerable time. However, if it is not feasible, we still have to modify the network and solve the modified problem. Here we are lucky that the solution obtained in the quarterly aggregation is actually feasible. However, for the monthly model without the aggregation, it is not. For further investigation, the modification will be implemented in the future.

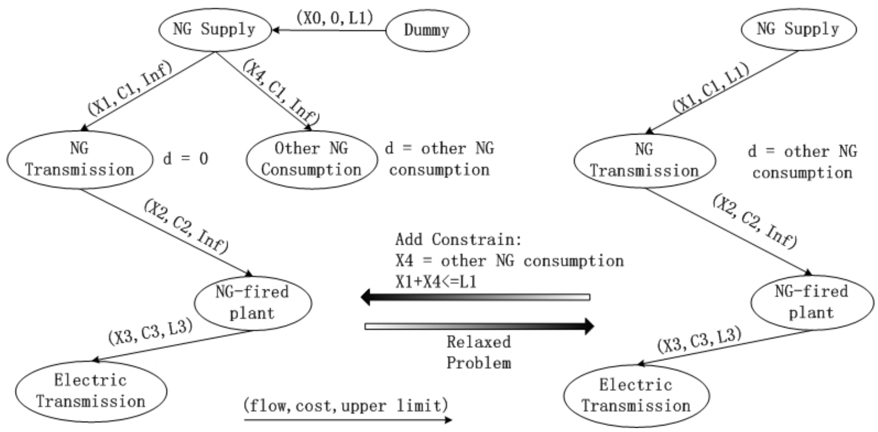

Figure 4.10 Solving the problem by relaxation

CHAPTER 5 RESULTS

5.1 Stochastic Model vs. Deterministic Model

The model formulated in chapter 3 and implemented as described in chapter 4 is solved by three different approaches which lead to three sets of solutions. The Wait and See (WS) solution is obtained from solving a deterministic problem with the actual fuel price. The expected value (EV) solution is also from a deterministic problem but solved by replacing the actual price with the mean value of its forecast. The Recourse Problem (RP) solution is obtained by solving the stochastic problem through the rolling 2-stage procedure.

We first compare the total flows in each solution in Table 5.1. The total cost refers to actual cost for each solution but not necessarily the objective function value of that solution. For consistency, the EV and RP solutions are evaluated using the actual (WS) costs. When uncertainty is considered in the RP solution, coal deliveries decrease and NG deliveries increase; especially, imports from Canada are more than doubled compared to EV. As a result, electricity generated from coal-fired plants is reduced and more electricity is generated from natural gas. In the stochastic case, net trade within sub regions decreases by 11%. One reason to explain the reduction of trade is that when people are not sure of future price which determines how much benefits can be earned from trading, they tend to avoid it and save the transportation cost.

Compared to RP, the EV solution is closer to the WS, the optimal solution with perfect information. However, RP is closer to the 2002 actual data than both EV and WS, as shown in Table 5.2. We can not compare the total costs because actual flows are not available at this level of detail. The comparison indicates that while EV and WS rely more on coal, RP has a similar trend as the actual data to use more natural gas. This interesting result comes

Table 5.1 Total flows comparison

Result	WS	EV	RP	(RP-EV)/EV
Coal deliveries (million tons)	1,072	1,071	1,018	-5.01%
Canada Natural gas deliveries (million Mcf)	119	210	467	122.76%
Domestic Natural gas deliveries (million Mcf)	3,719	3,651	4,544	24.45%
Total Natural gas deliveries (million Mcf)	3,839	3,861	5,011	29.79%
Electricity generation from coal (thousand GWh)	2,121	2,121	1,997	-5.81%
Electricity generation from NG (thousand GWh)	410	410	533	29.88%
Net trade (thousand GWh)	350	346	309	-10.88%
Total costs (billion $)	35,694	35,996	38,405	6.69%

from the greater realism of the stochastic model: we modeled the uncertain factors that people making decisions faced in reality. And therefore, the stochastic model can be utilized as a tool to investigate and predict how the whole system would react in the real world.

Table 5.2 EV and RP compared to 2002 actual data

Result	Actual	EV	RP
Coal deliveries (million tons)	976	1,071	1,018
Natural gas deliveries (million Mcf)	5,398	3,861	5,011

Besides total flows, it is also beneficial to look at sub regional flows. Figure 5.1, natural gas flows from supply areas to power plants, shows that EV and RP make different decisions on how much to buy at each natural gas supply area. The randomization of natural gas cost not only changes the total flows but also has an inevitable impact on the amount of natural gas purchased from each supply area.

Natural gas storage levels in EV and RP are compared in figure 5.2 with the dashed line showing forecasted price trend. When the uncertain factor is introduced, the system stores more natural gas as for future uncertainty. And the storage level in RP is more consistent with the price outlook than that in EV. Figure 5.3 shows net trade amount at each electricity transmission center. At most places, exports or imports decline because of future price uncertainty, which corresponds to the decrease of total power trade in the total flows comparison.

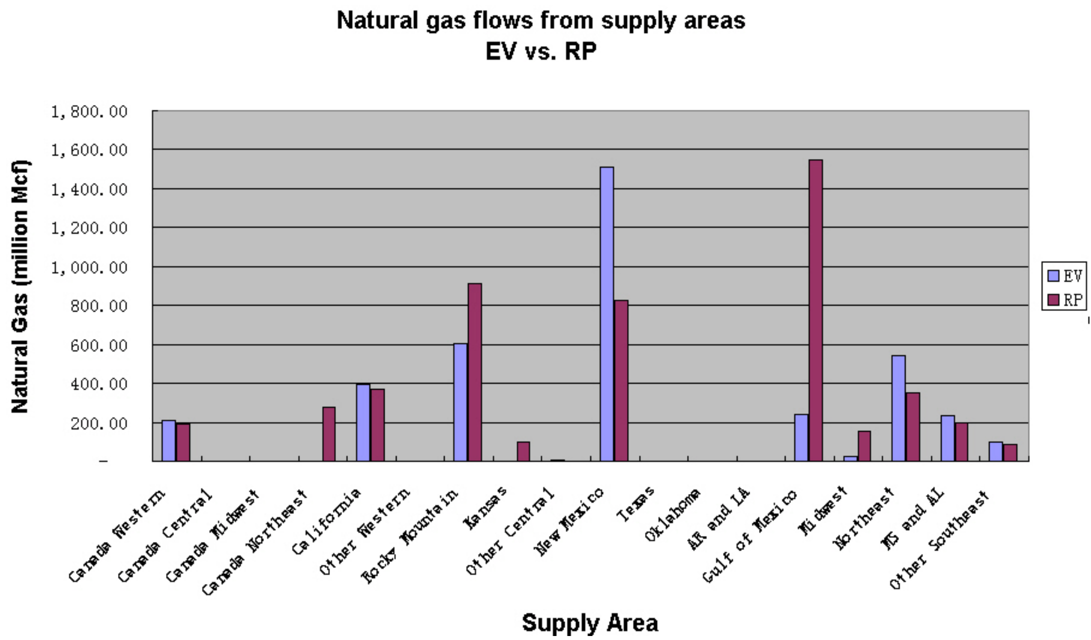

Figure 5.1 Natural gas flows from supply areas: EV vs. RP

5.2 Increased Uncertainty

As shown in the previous chapter, the actual price of natural gas does not always lie in the predicted confidence interval. The inaccuracy of prediction inspired us to study the impact of degree of the uncertainty. We increase the variance of random variables by changing both the CI width and the scenario distribution. Case 1 is the benchmark case used in previous analysis and case 2, 3 and 4 are similar cases but with larger variances. Solutions are compared in table 5.3.

- RP1: case 1 result, where $p_2 = 0.05, p_3 = 0.125, p_4 = 0.2$, narrow CI, $variance = Var(1)$;
- RP2: case 2 result, where $p_2 = 0.1, p_3 = 0.25, p_4 = 0.4$, wide CI, $Var(2) = 8 * Var(1)$;
- RP3: case 3 result, where $p_2 = 0.05, p_3 = 0.125, p_4 = 0.2$, wide CI, $Var(3) = 4 * Var(1)$;
- RP4: case 4 result, where $p_2 = 0.1, p_3 = 0.25, p_4 = 0.4$, narrow CI, $Var(4) = 2 * Var(1)$.

The comparison result is unusual. When we raise the uncertainty level by placing more weight on extreme values, we get a solution with lower cost, which contradicts the intuitive

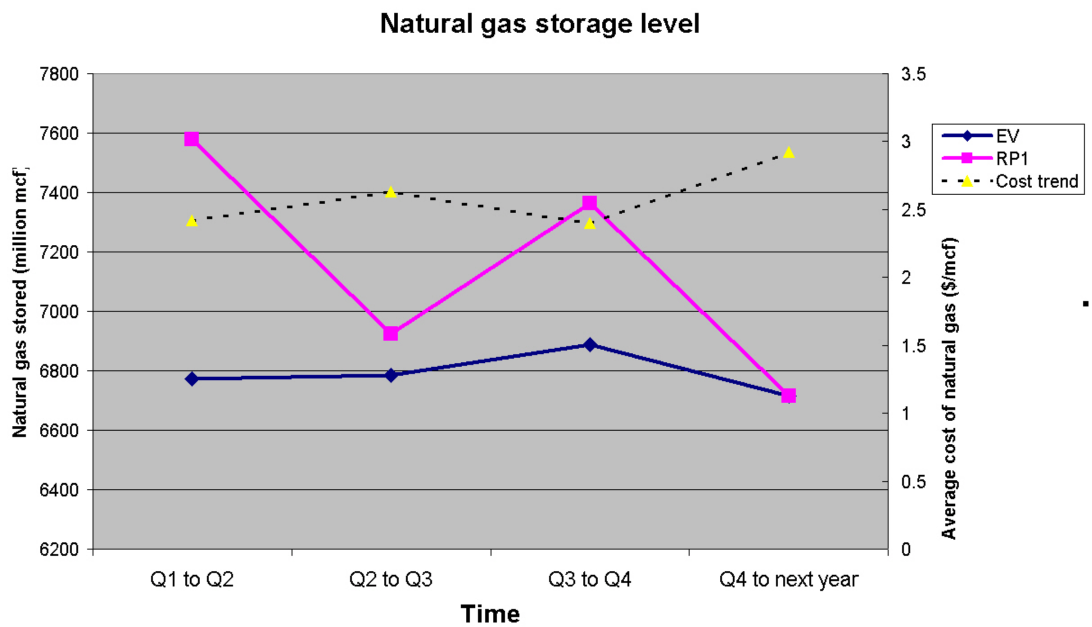

Figure 5.2 Natural gas storage level: EV vs. RP

Table 5.3 Total flows comparison: WS vs. Case 1-4

Result	WS	RP1	RP2	RP3	RP4
Coal deliveries (tons)	1,072	1,018	1,017	1,016	1,016
Canada Natural gas deliveries (Mcf)	119	467	212	467	418
Domestic Natural gas deliveries (Mcf)	3,719	4,544	4,829	4,563	4,606
Electricity generation from coal (GWh)	2,121	1,997	1,994	1,995	1,994
Electricity generation from NG (GWh)	410	533	535	535	536
Net trade (GWh)	350	309	307	311	306
Total costs (billion $)	35,694	38,405	38,318	38,438	38,395

expectation that we need to pay more for the increased uncertainty. Increasing the width of the intervals without changing probabilities has a mixed effect. Similarly for natural gas storage (Figure 5.4), less fuel is stored in Case 2 when variance of fuel price is much larger than that in Case 1. Although the result is not intuitive, it presents the relationship between accuracy of forecast and degree of uncertainty. In the 2002 data the actual price is almost totally out of the confidence interval estimated by EIA, which implies 2002 is a year when natural gas price rose much higher than people had expected. The forecast intervals of Case 2 and Case 3 contain the actual prices but those of Case 1 and Case 4 do not. In this situation, increasing uncertainty actually helps to adjust the forecast when it is not accurate enough. With a more

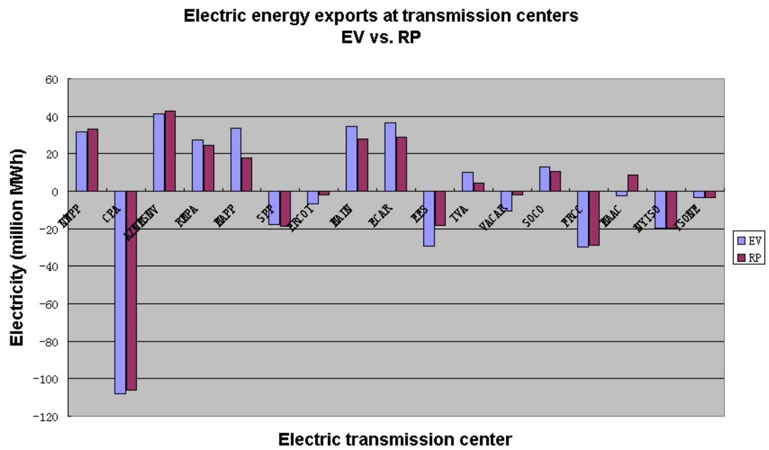

Figure 5.3 Electric energy exports at transmission centers: EV vs. RP

accurate perception of uncertainty, the decision maker chooses decisions in early periods that require less expensive adjustment in later periods.

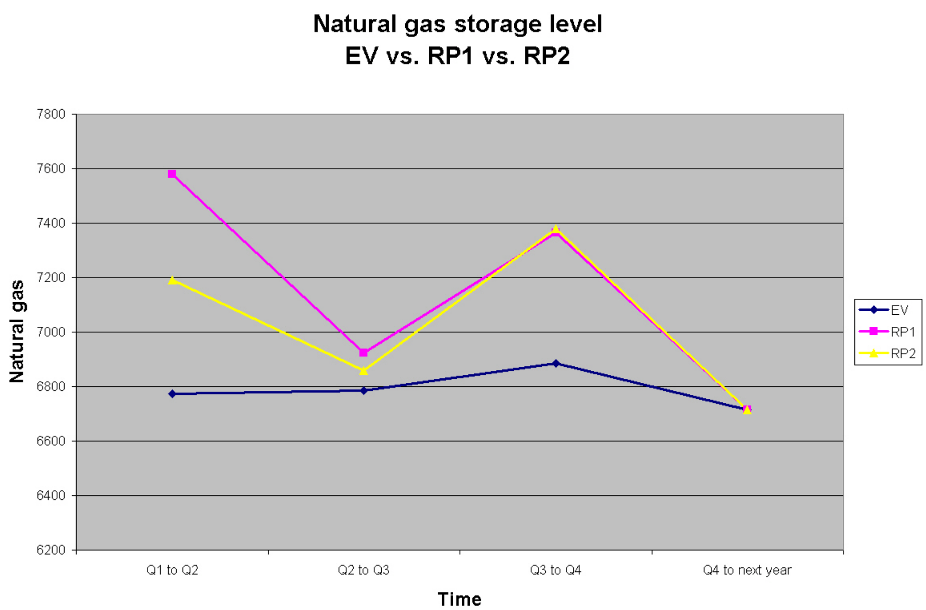

Figure 5.4 Natural gas storage levels: WS vs. RP1 vs. RP2

5.3 Summary

To conclude, since the stochastic case includes some underlying uncertain factors, the generation mix under stochastic costs is more like the actual situation than the deterministic case. Thus, the stochastic network flow model can be adopted to estimate the actual situation that happens in reality and help navigate all parts involved in the system. In a more detailed sense, while coal flows are stable with uncertain NG costs, decisions on natural gas flows vary a lot. Imports from Canada are especially sensitive to cost uncertainty. Besides, more natural gas is stored when uncertain costs are accounted for. There is less electricity trade between sub regions. Last but not least, an inaccurate forecast has a negative impact on total costs and proper control of the degree of uncertainty would offset some of its negative effect.

CHAPTER 6 CONCLUSION AND FUTURE WORK

6.1 Conclusion

To explore and forecast the U.S. electric energy system under uncertain, stochastic fuel costs are included in a model of the bulk energy transportation system (35), which is composed of coal, natural gas and electricity subsystems and validated with year 2002 data. As the uncertain elements are modeled as discretely distributed random variables, we use a two-stage recourse approach to solve the stochastic problem. A small electric network example illustrates the two-stage method and the difference between the flows in the stochastic model and those in the deterministic model.

In the implementation, the two-stage approach is applied in a rolling procedure to solve the multi-period network, in which the fuel costs are revealed period by period. The scenarios of the natural gas costs are derived from the natural gas Short Term Energy Outlook 2002 by EIA. The natural gas consumptions other than power generation are separated by transferring the demands to a new set of nodes so as to avoid excessive effects by NG uses outside the power system.

Compared to the recourse problem solution, the expected value solution which is obtained from the deterministic model with expected future fuel costs is closer to the optimal solution with perfect information. However, the recourse problem solution, which includes more natural gas consumption, less sub-regional trade and higher natural gas storage levels, is more like what actually happened in year 2002, which demonstrates the model's ability to project energy flows. With increased uncertainty, represented by the variance of the fuel cost forecast, the stochastic model generates a solution with lower costs than in the base case (case 1), which implies the interdependency among the accuracy of forecasts, degree of uncertainty and the value of the solution from the stochastic model.

The specific contribution of this research can be summarized as follows:

- Incorporated random fuel costs into the integrated electric energy system and solved the stochastic model via a rolling two-stage approach with 2002 data;
- Identified the stochastic model's capability of forecasting behavior by comparing it to the deterministic model and to the data from real world;
- Proposed the idea of using the stochastic model to project energy flows and provide instruction for real world decision makers.

6.2 Future Work

6.2.1 Load decomposition

One of the largest differences between the deterministic model and the stochastic one is that the latter consumes more natural gas for power generation. Later we found that the structural model may overemphasize coal by aggregating electricity demands over months and ignoring the daily/hourly variation. Because some of the electricity generated from NG-fired generation plants is usually used to satisfy the peak demand, the model might reduce the need for using natural gas in peaking units. It is worthwhile to consider disaggregating the structural model with respect to time to see the effect. Load decomposition is one of the possible approaches. The monthly electric load can be decomposed into 2 or more levels, one representing the peak hours and the other one for ordinary hours. Through load decomposition, we would make sure whether the difference in NG consumption levels is caused by including the uncertain fuel costs.

6.2.2 Monthly model

As indicated in Chapter 4, the stochastic model is aggregated from monthly model to quarterly because of computational restrictions. The great number of scenarios enlarges the problem size and makes it difficult to be solved on a regular PC, but we can find various effective methodologies in the literature to handle large-scale linear problems.

Benders decomposition (3) and other approaches derived from it are one series of schemes that decompose large size problem into small subproblems. When Benders decomposition is applied to two-stage stochastic linear problems, the first stage is formulated as the master problem providing lower bounds, and a subproblem is formed for each scenario and all the subproblems together generate upper bounds and cuts for the master problem. The lower bound and upper bound finally converge at the optimal solution. Benders decomposition keeps both the master and sub problems solvable and maintains the problem size comparable to that of the deterministic problem.

The drawback of decomposition is that it is time consuming to solve all the sub problems iteration by iteration given the large number of scenarios. Hence, sampling techniques are employed to reduce the number of sub problems. Lavenberg and Welch (26) discuss the efficiency of control variables in Monte Carlo sampling. Dantzig and Glynn (13) and Infanger (23) used importance sampling which is an improvement of Monte Carlo sampling.

In addition to decomposition and sampling, recent research on scenario reduction by Dupacova et al (12) also addresses the large-scale problem. The scenario reduction algorithm, which selects most significant scenarios with respect to perturbations of their probabilities measured in terms of a Fortet-Mourier probability metric, guarantees the degree of optimality corresponding to the number of scenarios selected. Besides this general approach, there are some heuristic for certain type of problems. Carino et al. (9) choose scenarios according to desired mean and standard deviation. Beltratti et al. (2) separate the scenario tree into extreme scenarios and the most likely ones and certain fraction of scenarios from each cluster are retained to represent the stochastic situation.

In summary, a monthly model can be solved with extra computational efforts. And the result from the disaggregated model would provide a more accurate projection of energy flows.

6.2.3 2005 data with Katrina

We have been using year 2002 data in the stochastic model. An apparent feature of this year is that the actual natural gas prices went much beyond the predicted figures. Analysis based on 2002 data could help people prepare for the unexpected rise or drop of fuel price. Similarly, study on data from alternative years would provide insights in other uncertain aspects of the energy system. In this sense, 2005 is a special year worthy of note because the unanticipated hurricanes struck the Gulf of Mexico and brought severe impacts on both local and national energy system (20). The record shows that when the natural gas production and transmission was interrupted during the catastrophic event, coal storage became a crucial source to maintain the energy supply. Therefore, a stochastic model based on 2005 data would explore the energy system in terms of sudden disruption of production and transmission and the proper fuel storage level in order to avoid huge cost brought by unforeseen events.

6.2.4 Emission constraints

While the European Union has imposed CO_2 emission regulations on power generation, the U.S. government also considers reduction of greenhouse gas emissions, but currently regulates SO_2 emission to prevent acid rain. While EIA uses the National Energy Modeling System to analyze the effects of existing and proposed government regulations, we can add emission restrictions to the stochastic model. Since the actually policy has not been decided yet, various sets of constraints, which stand for possible regulations regarding carbon emission, can be evaluated simultaneously through stochastic programming.

BIBLIOGRAPHY

[1] Anderson, E.J., A.B. Philpott, 2002. On supply function bidding in electricity markets. C. Greengard, A. ruszczynski (eds.), Decision Making under Uncertainty: Energy and Power. *Vol. 128 of IMA volumes on Mathematics and its Applications*, Springer-Verlag: 115-134.

[2] Beltratti, A., A. Consiglio and S.A. Zenios, 1999. Scenario modeling for the management of international bond portfolios. *Annals of Operations Research*, Vol. 85(1): 227-247.

[3] Benders, J.F., 1962. Partitioning procedures for solving mixed integer variables programming problems. *Numerische Mathematik*, Vol. 4: 238-252.

[4] Ben-Tal, A. and A. Nemirovski, 1998. Robust convex optimization. *Mathematics of Operations Research*, Vol. 23(4): 769-805.

[5] Bertsimas, D. and M. Sim, 2003. Robust Discrete Optimization and Network Flows. *Mathematical Programming*, Ser. B 98: 49C71.

[6] Beale, E., 1955. On minimizing a convex function subject to linear inequalities. *J. Roy. Statist. Soc.*, Ser. B(17): 173-184.

[7] Bolinger, M., R. Wiser and W. Golove, 2006. Accounting for fuel price risk when comparing renewable to gas-fired generation: the role of forward natural gas prices. *Energy Policy*, Vol. 34: 706-720.

[8] Botnen, O.J., A., Johannesen, A. Haugstad, S. Kroken, O. Froystein, June 1992. Modelling of hydropower scheduling in a national/international context. *E. Broch, D.K. Lysne (eds.), Proceedings of the 2nc international Conference on Hydropower Development (Hydropower 92)*, Lillehammer, Norway, Balkema, Rotterdam: 575-584.

[9] Carino, D.R., D.H. Myers, W.T. Ziemba, 1998. Concepts, technical issues, and uses of the Russell-Yasuda Kasai financial planning model. *Operations Research*, Vol. 46(4): 450-462.

[10] Charnes, A. and W.W. Cooper, 1959. Chance-constrained Programming. *Management Science*, Vol. 6(1):73-79.

[11] Dantig, G.B., 1955. Linear Programming under Uncertainty. *Management Science*, Vol. 1: 197-206

[12] Dupacova, J., N. Growe-Kuska and W. Romisch, 2003. Scenario reduction in stochastic programming: an approach using probability metrics. *Mathematical Programming*, Vol. 95(3): 493-511.

[13] Dantzig, G.B. and P.W. Glynn, 1990. Parallel processors for planning under uncertainty. *Annals of Operations Research*, Vol. 22(1): 1-21.

[14] Energy Information Administration. The National Energy Modeling System: An overview 2003. *DOE/EIA-0581(2003)*, Washington, DC, March 2003.

[15] Energy Information Administration. Short-term Energy Outlook. Washington, DC, Jan. 2002.

[16] Energy Information Administration. Short-term Energy Outlook. Washington, DC, Jan. 2003.

[17] Fragniere, E., A. Haurie, 1996. A stochastic programming model for energy/environment choices under uncertainty. *International Journal of Environment and Pollution*, Vol 6(4-6): 587-603.

[18] Fleten, S.-E., S.W. Wallace, W.T. Ziemba, 2002. Hedging electricity portfolios using stochastic programming. C. Greengard, A. ruszczynski (eds.), Decision Making under Uncertainty: Energy and Power. *Vol. 128 of IMA volumes on Mathematics and its Applications*, Springer-Verlag: 71-94.

[19] Gardner, D.T., J.S. Rogers, 1999. Planning electric power systems under demand uncertainty with different technology lead times. *Management Science*, Vol. 45: 1289-1306.

[20] Gil, E., 2007. Integrated network flow model for a reliability assessment of the national electric energy system. *Ph. D. thesis*, Iowa State University.

[21] Hindsberger, M., 2003. Interconnected hydro-thermal systems. *Ph.D. thesis*, Technical University of Denmark.

[22] Hogan, W.W., 1975. Energy Policy Models for Project Independence. *Computers and Operations Research*, Vol. 2: 251C271.

[23] Infanger, G., 1992. Monte Carlo (importance) sampling within a benders decomposition algorithm for stochastic linear programs. *Annals of Operations Research*, Vol. 39(1): 65-95.

[24] Kall, P. and Wallace, S.W., 1994. Stochastic Programming. *JOHN WILEY & SONS*.

[25] Kouvelis, P. and G. Yu, 1997. Robust Discrete Optimization and Its Application. *Kluwer Academic Publishers Boston*.

[26] Lavenberg, S.S. and P.D. Welch, 1981. A perspective on the use of control variables to increase the efficiency of Monte Carlo simulations. *Management Science*, Vol. 27(3): 322-335.

[27] Mo, B., A. Gjelsvik, A. Grundt, 2001. Integrated risk management of hydro power scheduling and contract management. *IEEE Transactions on Power Systems*, Vol. 16(2): 216-221.

[28] Murphy, F.H., S. Sen, A.L. Soyster, 1982. Electric utility capacity expansion planning with uncertain load forecasts. *IIE Transactions*, Vol. 14(1): 52-59.

[29] Murphy, F.H., J.J. Conti, S.H. Shaw and R. Sanders, 1988. Modeling and Forecasting Energy Markets with the Intermediate Future Forecasting System. *Operations Research*, Vol. 36(3): 406C420.

[30] Mulvey, J.M. and H. Vladimirou, 1991. Solving Multistage Stochastic Networks: An application of Scenario Aggregation. *Networks*, Vol. 21: 619-643.

[31] Neame, P.J., A.B. Philpott, G. Pritchard, 1999. Offer stack optimization for price takers in electricity markets. *Proceedings of the 34th Annual Conference of ORNZ*, University of Waikato, Hamilton, New Zealand: 3-12.

[32] Nowak, M.P. and W. Romisch, 2000. Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Annals of Operations Research*, Vol. 100, 251-272.

[33] Qiu, J., A.A. Girgis, 1993. Optimization of power-system reliability level by stochastic-programming. *Electric Power Systems Research*, Vol. 26(2): 87-95.

[34] Quelhas, A.M., E Gil. and J.D. McCalley, 2005. Nodal prices in an integrated energy system. *Int. J. Critical Infrastructures*, Vol. XX(X), to be published.

[35] Quelhas, A.M., 2006. Economics Efficiencies of the Flows from the Primary Resource Suppliers to the Electric Load Centers *Ph. D. thesis*, Iowa State University.

[36] Quelhas, A.M., E. Gil, J.D. McCalley and S.M. Ryan, 2006. A multiperiod generalized network flow model of the U.S. integrated energy system: PartI: model description. *IEEE Transactions on Power Systems*, Vol. XX(X), to be published.

[37] Quelhas, A.M., E. Gil, and J.D. McCalley, 2006. A multiperiod generalized network flow model of the U.S. integrated energy system: PartI: Simulation Results. *IEEE Transactions on Power Systems*, Vol. XX(X), to be published.

[38] Scott, T.J., E.G. Read, 1996. Modelling hydro reservoir operation in a deregulated electricity sector. *International Transactions in Operations Research*, Vol. 3(3-4), 209-221.

[39] Sherali, H.D., A.L. Soyster, F.H. Murphy and S. Sen, 1982. Linear programming based analysis of marginal cost pricing in electric utility capacity expansion. *European Journal of Operational Research*, Vol. 11: 349-360.

[40] Sherali, H.D., A.L. Soyster, F.H. Murphy and S. Sen, 1984. Intertemporal allocation of capital costs in electric utility capacity expansion planning under uncertainty. *Management Science*, Vol. 30: 1-19.

[41] Wallace, S.W. and S.-E. Fleten, 2003. Stochastic Programming Models in Energy. *Handbooks in OR & MS*, Vol 10(10): 637-677.