

2009

Long Term Power Generation Planning Under Uncertainty

Shan Jin

Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

 Part of the [Industrial Engineering Commons](#)

Recommended Citation

Jin, Shan, "Long Term Power Generation Planning Under Uncertainty" (2009). *Graduate Theses and Dissertations*. 10803.
<https://lib.dr.iastate.edu/etd/10803>

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Long term power generation planning under uncertainty

by

Shan Jin

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

Mervyn G. Marasinghe

Jo Min

Iowa State University

Ames, Iowa

2009

Copyright © Shan Jin, 2009. All rights reserved.

TABLE OF CONTENT

LIST OF FIGURES	v
LIST OF TABLES	vii
ABSTRACT	viii
CHAPTER 1 OVERVIEW	1
1.1 Introduction	1
1.2 Motivation	2
1.3 Problem Statement	5
1.4 Thesis Structure	6
CHAPTER 2. REVIEW OF LITERATURE	8
2.1 Methodologies for Power Expansion Planning Problem	8
2.2 Methodologies for Improving Computational Efficiency	9
2.3 Scenario Generation Methodologies	10
2.4 Commercial Packages	10
CHAPTER 3. METHODS AND PROCEDURES	12
3.1 Model Assumptions and Notations	12
3.2 Two-Stage Stochastic Optimization	14
3.3 Robust Optimization	16
CHAPTER 4. SCENARIO TREE GENERATION	20
4.1 Stochastic Process	20
4.1.1 Geometric Brownian Motion	21

4.1.2 Verification of Geometric Brownian Motion	22
4.1.3 Statistical Properties of Random Variables	26
4.2 Scenario Generation Method	29
4.2.1 Scenario Generation for a Single Year	29
4.2.2 Evolution over Future Years	31
CHAPTER 5. CASE STUDY	34
5.1 Scenario Reduction	34
5.2 Assumptions	35
5.2.1 Demand	35
5.2.2 Annual Interest Rate for Discounting	37
5.2.3 Build Cost for Generators	38
5.2.4 Generation Cost for Generators	40
5.2.5 Capacity for Generators	42
5.2.6 Maximum Units to Build for Generators	43
5.2.7 Scenario Probabilities	43
5.2.8 Lead Time for Generators	43
5.2.9 Penalty Costs	43
5.3 Experiment Results	44
5.4 Sensitivity to Robustness Parameter	49
5.5 Production Tax Credit	55
5.6 Sensitivity to Scenario Sampling	60
5.7 Deterministic Solutions	62

CHAPTER 6. SUMMARY AND FUTURE RESEARCH	66
6.1 Summary	66
6.2 Future Research	67
6.2.1 Assumptions and Constraints	67
6.2.2 Uncertainties	68
6.2.3 Methodologies	69
ACKNOWLEDGEMENTS	70
BIBLIOGRAPHY	71

LIST OF FIGURES

Figure 1. Log ratios of annual demand in Midwest region from year 1991-2006	23
Figure 2. Log ratios of annual natural gas price in Midwest region from year 1970-2006	23
Figure 3. Correlation of annual demand in Midwest region from year 1991-2006	24
Figure 4. Correlation of annual natural gas price in Midwest region from year 1970-2006	24
Figure 5. Correlation of total annual demand and average annual natural gas price in Midwest region from year 1991-2006	29
Figure 6. A scenario tree for multi-year horizon	32
Figure 7. Load duration curve in year 2008 (3 sub-periods)	36
Figure 8. Scenario incremental demand	37
Figure 9. Generation cost for BaseLoad, IGCC, wind and nuclear power plants	41
Figure 10. Scenario generation cost for CC and CT power plants	42
Figure 11. Investment cost, generation cost, USE cost for both stochastic and robust	46
Figure 12. Cumulative expansion planning decision in “Stochastic” solution	47
Figure 13. Cumulative expansion planning decision in “Robust 0” solution	48
Figure 14. Standard deviation vs. expected cost for stochastic and robust optimization	50
Figure 15. The cost over scenarios for five different robust solutions	51
Figure 16. Cumulative expansion planning decision in “Robust 7” solution	52
Figure 17. Cumulative expansion planning decisions in “Robust 8” solution	52
Figure 18. Cumulative expansion planning decisions in “Robust 9” solution	53
Figure 19. Generation cost for BaseLoad, IGCC, wind and nuclear power plants (PTC)	56

Figure 20. Standard deviation vs. expected cost for stochastic and robust optimization (PTC)	57
Figure 21. The cost over scenarios for five different robust solutions (PTC)	57
Figure 22. Cumulative expansion planning decision in “Robust 1” solution (PTC)	58
Figure 23. Cumulative expansion planning decision in “Robust 7” solution (PTC)	58
Figure 24. Cumulative expansion planning decision in “Robust 8” solution (PTC)	59
Figure 25. Cumulative expansion planning decision in “Robust 9” solution (PTC)	59
Figure 26. Standard deviation vs. expected cost for stochastic and robust optimization (sampling 2)	60
Figure 27. The cost over scenarios for three different robust solutions (sampling 2)	61
Figure 28. Cumulative expansion planning decision in “Robust 9” solution (sampling 2)	61
Figure 29. Cumulative expansion planning decision in “Sce 1” solution	63
Figure 30. Cumulative expansion planning decision in “Sce 2” solution	63
Figure 31. Cumulative expansion planning decision in “Sce 3” solution	64
Figure 32. Cumulative expansion planning decision in “Sce 4” solution	64
Figure 33. Cumulative expansion planning decision in “Sce 5” solution	65

LIST OF TABLES

Table 1. Time series autocorrelation with lag = 1 for demand	25
Table 2. Time series autocorrelation with lag = 1 for natural gas price	25
Table 3. Statistical properties of random variables over years	28
Table 4. Demand data for 3 sub-periods in year 2008	36
Table 5. Overnight cost for generators	39
Table 6. Capital expenditure profile for generators	39
Table 7. Generation cost related parameters for the generators in the first year	40
Table 8. Unit transformation for the natural gas price	40
Table 9. Installed capacity and generator rating for generators	42
Table 10. Max units to build for generators	43
Table 11. Experiment result for a 10-year case study	44
Table 12. The penalty cost p_v for 10 different robust optimization models	49
Table 13. Regret over scenarios for “Robust 7”, “Robust 8”, “Robust 9”, “Robust 10”	55

ABSTRACT

Generation expansion planning concerns investment and operation decisions for different types of power plants over a multi-decade horizon under various uncertainties. The goal of this research is to improve decision-making under various long term uncertainties and assure a robust generation expansion plan with low cost and risk over all possible future scenarios. In a multi-year numerical case study, we present a procedure to deal with the long term uncertainties by first modeling them as a multidimensional stochastic process and then generating a scenario tree accordingly. Two-stage stochastic programming is applied to minimize the total expected cost, and robust optimization is further applied to reduce the cost variance. Results of experiments on a realistic case study are compared. An efficient frontier of the planning solutions that illustrates the tradeoff between the cost and risk is further shown and analyzed.

CHAPTER 1 OVERVIEW

1.1 Introduction

Nowadays, energy has become a more and more essential element of people's lives, as well as a key concern of the whole world. The development of inexpensive, fossil-fuel energy instigated a new era of industrial revolution and the increasing use of that energy has rapidly improved our human society and standard of living. However, with the global economy more reliant on the sustainable development of energy, a series of problems, such as energy shortage, electricity shortage and global warming are gaining attention.

In order to deal with some of the problems, the concept of clean energy and sustainable living has gained more popularity and been widely accepted. More and more renewable power plants have been built to help ease the energy crisis, improve the environmental condition, and, at the same time, meet the increasing demand instead of the traditional fossil-fueled power plants.

All of these reasons contribute to the importance of building up a reliable and efficient electricity energy supply system for all the consumers by the decision makers of the power generation plants.

Usually, the decision making for the power generation expansion planning involves a long time horizon, from 10 to 20 years. The reason for the long term is as follows [1]:

- The initial capital investment is expensive and the lifetime of a power generation plant normally ranges from 25 to 60 years.

- Multiple organizations are required to get involved in the decision making since the power plant needs to be integrated into the whole electricity system and thus certainly will have impact on the other organizations, by construction of electric generation, transmission, or distribution facilities.
- A certain amount of land is necessary for the power generation stations, the transmission and distribution circuits, etc.
- Environmental impacts which include carbon emission of the fossil-fueled plants, wastes from the nuclear plants, audible noise, visual aesthetics and some other ecological impact must be considered.
- The energy cost must be considered based on the forecast of the future fuel prices. Also, the forecast of the electricity demand must be estimated to determine the appropriate installed capacity. Either unmet demand or surplus energy will lead to a loss due to adverse consequences of shortage.
- Reliability is also a key issue to keep in mind to avoid potential electricity interruption or insufficiency.

1.2 Motivation

Long term generation expansion planning is a power plant investment decision making problem [2, 3]. It is challenging to model and solve due to multiple objective functions [4], complexity arising from power plants of different technologies and important reliability constraints of sufficient energy supply. Both investment planning and operation scheduling must be considered over multiple decades.

Long term generation expansion planning is also very complicated to formulate because of its large number of uncertainties [5-7].

Load growth has always been a significant uncertainty in generation expansion planning. It can usually be estimated by climate forecast, population expansion or movement, economic conditions and technology development. For the long term, the world growth rate increased from average of 2% in year 1990 to 4% in year 2007 [8] and it is projected to grow with average 2% until year 2030 [9]. Growth in electricity demand in the U.S. has generally slowed down from 9% per year in the 1950s to less than 2.5% per year in the 1990s. Recently from 2000 to 2007, the average growth rate was down to 1.1% per year. And it was projected that the slowdown will still continue for the next 23 years until year 2030 [10]. As one of the fast-growing economies currently, China has experienced an electric demand growth rate as high as average 14% for the past 5 years [8].

The growth of new generation technologies has also gradually become more important because environmentally friendly renewable energy is receiving more public support currently. The US government is considering greatly enlarging the percentage of wind energy to 20% of the generation by year 2030 [11], compared to the 9% in year 2008 [12]. In Iowa, plans are to increase the Renewable Portfolio Standard (RPS), the proportion of renewable net generation over the total [13-15], to 8% of generation capacity by year 2010 and 20% by year 2020, compared with 7% in year 2007. Most of this increase will come from wind energy due to the abundance of wind resources in the Midwest. However, integration of wind generation into the power system involves more uncertainties due to its weather dependence [2, 6, 7, 16-18]. Hence, instead of “capacity factor”, an average output

over a year, “capacity credit” is introduced as a measure of generation potential. It measures the wind generation in the worst case that the power generation system can actually count on all the time. It can be estimated by different methods [3, 19-21]. Besides wind generation, clean coal [22], new nuclear and bio-based technologies are also to be taken into account in the expansion planning.

Other related environmental concerns including emission penalty and/or constraints and other regulatory uncertainty will also have a large influence on the investment decision of different types of power plants [23]. We need to take into account the potential policy to limit or reduce the greenhouse gas emissions, which would have a significant impact on the power plant planning. For the past decades, tax incentives have increased the growth of renewable generation. The renewable electricity production tax credit (PTC) [24] was established as an incentive to promote different kinds of renewable energy beyond wind generation and has had a great impact on the growth of wind generation for the past 10 years [10]. It is likely that the PTC program will be extended over a future longer term.

Prices and availability of fuels, particularly coal and natural gas, contribute additional uncertainty. Generally speaking, coal price can be considered to be more stable with an average yearly growth rate of 2%, while natural gas price fluctuates in a more unpredictable way [10], mainly depending on the economic growth rate and the technology development rate. The proportion of electricity generated by natural gas in United States for the year 2008 is around 21% [25]. Since natural gas has generally the highest fuel price, the power plants it fuels are considered as peak load generation units, and the generation cost in the future is

highly subject to the uncertainty of the natural gas price. Hence, natural gas price is usually considered to be a very important uncertainty in the generation expansion planning problem.

In some cases, the transmission capacity and congestion need to be considered as well, because an insufficient transmission network will not enable the system to meet the electricity demand by efficiently allocating the generation production.

In a generation expansion planning problem, two major costs, investment cost and generation cost, are involved, respectively depending on the investment decisions on how many units of what type of power plants to build in which year and the operational decisions on how much electricity is generated by what type of the power plants. While making these decisions, we have to take into account the future uncertainties since they could have a significant impact on both the investment decision and the generation decision and their corresponding costs. At the same time, the investment decisions should be able to satisfy some other additional requirements, such as electricity demand, power generation reliability, energy resource limitation, financial budget, maximum carbon emission, or the minimum required electricity generation proportion for the renewable energy.

1.3 Problem Statement

In this thesis, we consider a long term power generation expansion planning problem of determining how many units of what type of power plants to build in which year to minimize both the initial investment cost and the generation cost in later years, while taking account of the future uncertainties represented by different future scenarios. Besides, we also

consider the robustness of the expansion planning decisions so that the generation cost incurred in the future will not vary too much among the future scenarios.

For solving it, two-stage stochastic programming is applied to minimize the total expected cost over scenarios, and robust optimization is further considered for minimizing both the expected cost and the cost variance among scenarios.

In addition to the optimization models, we also address the following problems:

- The appropriate way to model the future uncertainties over years
- The appropriate way to generate future scenarios for a long term horizon
- Reduction of the number of scenarios that must be considered
- Model implementation with appropriate data for the Midwest region
- Comparison of the experiment results for two-stage stochastic programming and robust optimization
- Trade-off between the expected cost and cost variance achieved by robust optimization.

1.4 Thesis Structure

In Chapter 2, a comprehensive literature review on the state-of-art methodologies to solve the expansion planning problem is introduced. In Chapter 3, models of a two-stage stochastic programming and a robust optimization method are given, as well as the model assumptions and notations. In Chapter 4, we further discuss how to realize the computational implementation, including model assumptions, fitting of uncertain variables' continuous time distributions, discrete scenario generation methods and scenario reduction for a multi-year

case study. In Chapter 5, a multi-year case study based on the Midwest electric power system is conducted, and furthermore a sensitivity analysis of the penalty coefficient for the cost variance in the robust optimization model is studied. In Chapter 6, a comprehensive summary of the thesis is made and future research regarding the assumptions, uncertainties, constraints and methodologies is further discussed.

CHAPTER 2. REVIEW OF LITERATURE

2.1 Methodologies for Power Expansion Planning Problem

It has been decades since the generation expansion planning problem arose. Different optimization techniques have been applied to study the problem concentrating on different aspects of it.

A collection of stochastic programming problems is discussed in [26] and one of the applications is electrical capacity expansion problems with the uncertainty concerning the different modes of demand. The use of stochastic programming was also studied to address the uncertainties in [13, 27, 28]. A review published in 1997 of emerging techniques on generation expansion planning included many optimization techniques until that time including: expert system, fuzzy logic, neural networks, analytic hierarchy process, network flow, decomposition method, simulated annealing and genetic algorithm [29]. For the robustness concern of the planning decision, robust optimization was also studied in order to reduce the cost variance over different future scenarios [30-32]. Besides, a game-theoretic model was applied to solve the problem in a competitive environment to learn the different results from the centralized expansion planning [33]. A multi-objective technique [5, 34] can also be applied to the power generation expansion problem to minimize cost, environmental impact, imported fuel and fuel price risks. The same model was also applied in [35]. Different criteria are suggested in [36] to help make a preferable planning solution. Dynamic programming [23, 37, 38] can also be applied to the problem. State-of-the-art optimization

methods under uncertainty [39] were also reviewed in 2004 including stochastic programming, robust stochastic programming, probabilistic (chance-constraint) programming, fuzzy programming, and stochastic dynamic programming.

2.2 Methodologies for Improving Computational Efficiency

Generally speaking, the computational size of the long-term expansion planning problem sometimes can be huge. And since integer decision variables are involved, it can be very computationally difficult to solve. Thus, many studies propose alternative heuristics or other techniques to solve the problem more efficiently.

Ahmed et al. provided a multi-stage stochastic programming method, recommended a reformulation technique, and applied different heuristic methods to solve the problem in a much more efficient way [40]. A parallel genetic algorithm was proposed to solve the deterministic power generation expansion planning problem with computational benefit [41]. A genetic algorithm was also used to reduce the problem complexity in [42]. Comparison among number of meta-heuristic techniques for solving the generation expansion planning problem was studied [43]. In 2003, Ahmed developed a fast linear-programming-based, approximation scheme that efficiently solves a multi-stage stochastic integer program arising from a capacity expansion planning problem [44]. Computational effort for solving by two-stage stochastic programming was also studied by using Benders decomposition and parallel algorithm [32].

2.3 Scenario Generation Methodologies

The long-term generation expansion planning problem is also a multi-period problem. When we consider the uncertainties for a multi-period problem, techniques for scenarios generation and reduction, and construction of a scenario tree are required.

Laurent summarized different methodologies for scenario discretization [29]. Several techniques for constructing multi-stage scenario tree were presented in [45]. A scenario construction algorithm successively reducing the tree structure by bundling similar scenarios was introduced in [46]. Hoyland and Wallace proposed a generalized method applied for both single-stage scenario and multi-stage scenarios [47]. Their method will be applied to generate the scenarios for the multi-year case study.

2.4 Commercial Packages

In the electric power industry, some commercial packages are also available such as EGEAS [48], ProMod [49], and Plexos [50, 51], most of which are based on deterministic models. They are also widely used in practice to approximate a stochastic programming model to address the future uncertainties by solving the different deterministic models based on one of the specific generated future scenarios at each time. Robust optimization is approximated in an ad hoc way by identifying common elements of the optimal plans found for different futures.

In this thesis, we propose a new procedure to model the multi-stage generation expansion planning problem in two-stage stochastic programming and robust optimization by

combining them with a multi-year scenario tree generation method representing the future uncertainties. We first introduce continuous time random variables to model the future uncertainties over the years, and then verify their stochastic process as a geometric Brownian motion. Based on the statistical specifications of the geometric Brownian motion, a methodology is further applied to generate the discrete scenarios for each year until a scenario tree has been constructed recursively. At last, naïve sampling is used to reduce the number of scenarios in order to improve the computational efficiency of the optimization models.

CHAPTER 3. METHODS AND PROCEDURES

3.1 Model Assumptions and Notations

In the following sections, we introduce three different formulations: two-stage stochastic optimization and its special case, deterministic optimization; and robust optimization.

Regarding the objective function of all three models, both investment and generation cost of the power plants are minimized. In addition, we take into account penalties for unmet demand, since serious power outage is always costly and disruptive. It might result in the direct economic damage due to the destruction of the electricity infrastructure, loss of data or breakdown of an assembly line, the loss of a life of a patient who is in the middle of a surgery in the hospital, failure of public services and regional confusion. The constraints considered in these models are essential for this type of problem: because electricity cannot be stored economically, we require the energy to meet the demand in each sub-period, the load of each type of generator to be less than its planned capacity, and the number of newly built plants to be less than the maximum limit because of the limitation of either budget or other resources.

For the stochastic and robust models, multiple uncertainties are incorporated by bringing in scenario decision variables and scenario parameters with different values over scenarios. The notation of decision variables, scenario decision variables, parameters, and scenario parameters are as follows:

- Indices

g : Type of generator

y : Year

t : Load duration curve sub-period

T_y : Set of sub-periods t in year y

Y_t : The year to which sub-period t belongs

s : Scenario

- First Stage Decision Variables

$U_{g,y}$: Units of generators of type g to be built in year y (integer)

- Second Stage Decision Variables

$L_{g,t,s}$: Load generated by generators of type g in sub-period t under scenario s ,
MWh

$E_{t,s}$: Unserved energy (USE) in sub-period t under scenario s , MWh

- Parameters

c_g : Total cost to build a generator of type g , discounted to beginning of
construction period, \$/MW

m_g^{\max} : Installed capacity of generators of type g , MW

n_g^{\max} : Maximum generation rating of generators of type g over a year, MW

u_g^{\max} : Maximum units built for generators of type g for the whole planning
horizon

u_g : Existing units of generators of type g at the beginning of the planning horizon

p_u : Penalty cost for unmet energy, \$/MWh

p_v : Penalty coefficient of cost variance over scenarios

r : Annual interest rate for cost discounting

- Scenario Parameters

$l_{g,t,s}$: Generation cost for generators of type g in sub-period t under scenario s ,

\$/MWh

$d_{t,s}$: Demand in sub-period t under scenario s , MWh

π_s : Probability that scenarios s occurs

3.2 Two-Stage Stochastic Optimization

The two-stage stochastic optimization model formulates future uncertainties as different discrete scenarios. It is assumed that the investment decisions must be made at the beginning of the planning horizon before any future uncertainties are revealed, and once it has been decided, it remains the same over decades no matter what future scenario occurs. The operational decisions can be made afterwards, depending on both the future scenario and the previous investment decisions. Hence, it is essential to analyze the future uncertainties and make robust expansion planning decisions in the first place to ensure a total cost as low as possible under any of the future conditions.

Each scenario is determined by its own parameters and addressed by its decision variables. There are two types of decision variables in this formulation. The investment

decision variables, $U_{g,y}$, are also referred to as the first stage decision variables since they have to be decided at the beginning of each year before the outcomes of any future uncertainties are revealed. Once the decision is made, it has to carry on over the years, and will no longer be changed at all. On the other hand, operational decision variables, $L_{g,t,s}$ and $E_{t,s}$, are scenario dependent, which are also referred to as the second-stage decision variables, since their decision can be delayed until after the realization of some certain scenario described by the scenario parameters, $l_{g,t,s}$ and $d_{t,s}$.

A probability value π_s , aggregating to 1 over all scenarios, is assigned to each scenario. The objective is to minimize both the investment cost and the expected generation cost over scenarios, and the constraints required must be satisfied for every scenario.

The two-stage stochastic formulation is as follows:

- Objective function: Minimize present value of the investment cost and the expectation of the sum of the load cost and penalty cost for unmet demand

$$\min \sum_y \left(\frac{\sum_g (c_g m_g^{\max} U_{g,y}) + \sum_s \pi_s \left(\sum_{t \in T_y} \left(\sum_g (l_{g,t,s} L_{g,t,s}) + p_u E_{t,s} \right) \right)}{(1+r)^y} \right) \quad (1)$$

- Energy constraints: The generation and unserved energy should equal demand in each sub-period t for each scenario

$$\sum_g L_{g,t,s} + E_{t,s} = d_{t,s} \quad \forall t, s \quad (2)$$

- Maximum generation constraints: Load generation of each type of generator g should be less than or equal to its aggregate rating of both existing units and newly built units so far in sub-period t for each scenario s

$$L_{g,t,s} \leq n_g^{\max}(u_g + \sum_{y \leq Y_t} U_{g,y}) \quad \forall g,t,s \quad (3)$$

- Maximum units to build constraints: The number of newly built units of each type of generator g in each year should be less or equal to its maximum limit

$$\sum_y U_{g,y} \leq u_g^{\max} \quad \forall g \quad (4)$$

A deterministic formulation can be seen as a special case of the two-stage stochastic formulation when there is a single scenario that occurs with probability 1. This might represent the planner's best guess of the outcomes of uncertain quantities.

If G is the number of generator type, S the number of scenarios, Y the number of years in the planning horizon, and T the total number of the sub-periods, then both the deterministic model and the two-stage stochastic programming model are mixed integer programming problems, with $T+T \cdot G+G$ constraints and $T+T \cdot G+G \cdot Y$ decision variables, and $T \cdot S+T \cdot S \cdot G+G$ constraints and $T \cdot S+T \cdot S \cdot G+G \cdot Y$ decision variables, respectively. Of the decision variables, $G \cdot Y$ are constrained to take integer values. For this thesis, they are both solved by Tomlab/CPLEX in Matlab.

3.3 Robust Optimization

The two-stage stochastic model deals with uncertainty by minimizing expected cost. However, it does not take into account the risk that the cost of a particular scenario far exceeds its expected value. Risk can be measured mathematically in different ways. It can be

assessed by a possible bad scenario, a worst-case analysis, expectation, standard deviation, a specified probability quantile, a value-at-risk or a conditional value-at-risk [52]. In this thesis, we measure the risk by the cost variance over scenarios, which reflects a preference that the cost among scenarios does not differ too much. The robustness of the planning decision implies that overall cost will be more likely to stay stable over all the possible scenarios.

The robust formulation not only minimizes the expected cost over scenarios, but also generates a smaller cost variance among scenarios to ensure less difference resulting from scenarios. We include the variance in the robust formulation by taking it as an additional component of the objective function with a penalty coefficient p_v . The objective of the robust formulation is as follows:

- Objective function: Minimize the expectation of both investment and operation cost, the penalty for unmet energy, and the penalty for the cost variance over scenarios

$$\min \sum_s \pi_s \xi_s + p_v \sum_s \pi_s \left(\xi_s - \sum_{s'} \pi_{s'} \xi_{s'} \right)^2, \text{ where} \quad (5)$$

$$\xi_s = \sum_y \left(\frac{\sum_g (c_g m_g^{\max} U_{g,y}) + \sum_{t \in T_y} \left(\sum_g (l_{g,t,s} L_{g,t,s}) + p_u E_{t,s} \right)}{(1+r)^y} \right)$$

The quantity ξ_s in equation (5) represents the discounted investment cost and generation cost incurred under the scenario s . The constraints are the same as in the two-stage stochastic formulation.

Because of the variance term, the robust optimization model is a quadratic mixed integer programming problem. Assume column vector $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_s, \mathbf{X}_{inv}]'$, with \mathbf{X}_s representing the scenario decision variables under scenario s , and \mathbf{X}_{inv} representing the investment decision variables, and let $\mathbf{C}^s = [C_1^s, \dots, C_s^s, C_{inv}^s]'$ represent the corresponding multipliers in $\xi_s = \mathbf{C}^{s'} \mathbf{X}$ with C_s^s the scenario parameters under scenario s , C_{inv}^s the investment cost parameters, and $C_1^s = \dots = C_{s-1}^s = C_{s+1}^s = \dots = C_s^s = 0$. The variance can be then transformed to the following format:

$$\begin{aligned}
\sum_s \pi_s \left(\xi_s - \sum_{s'} \pi_{s'} \xi_{s'} \right)^2 &= \sum_s \pi_s \left(\xi_s^2 - 2\xi_s \sum_{s'} \pi_{s'} \xi_{s'} + \left(\sum_{s'} \pi_{s'} \xi_{s'} \right) \left(\sum_{s'} \pi_{s'} \xi_{s'} \right) \right) \\
&= \sum_s \pi_s \xi_s^2 - 2 \sum_s \pi_s \xi_s \sum_{s'} \pi_{s'} \xi_{s'} + \sum_s \pi_s \left(\sum_{s'} \pi_{s'} \xi_{s'} \right) \left(\sum_{s'} \pi_{s'} \xi_{s'} \right) \\
&= \sum_s \pi_s \xi_s^2 - 2 \sum_s \pi_s \xi_s \sum_{s'} \pi_{s'} \xi_{s'} + \left(\sum_{s'} \pi_{s'} \xi_{s'} \right) \left(\sum_{s'} \pi_{s'} \xi_{s'} \right) \\
&= \sum_s \pi_s \xi_s^2 - \sum_s \sum_{s'} \pi_s \pi_{s'} \xi_s \xi_{s'} \\
&= \sum_s \pi_s \mathbf{C}^{s'} \mathbf{X} \mathbf{C}^{s'} \mathbf{X} - \sum_s \sum_{s'} \pi_s \pi_{s'} \mathbf{C}^{s'} \mathbf{X} \mathbf{C}^{s''} \mathbf{X} \\
&= \sum_s \mathbf{X}' \pi_s \mathbf{C}^s \mathbf{C}^{s'} \mathbf{X} - \sum_s \sum_{s'} \mathbf{X}' \pi_s \pi_{s'} \mathbf{C}^s \mathbf{C}^{s''} \mathbf{X} \\
&= \mathbf{X}' \left(\sum_s \pi_s \mathbf{C}^s \mathbf{C}^{s'} \right) \mathbf{X} - \mathbf{X}' \left(\sum_s \sum_{s'} \pi_s \pi_{s'} \mathbf{C}^s \mathbf{C}^{s''} \right) \mathbf{X} \\
&= \mathbf{X}' \left[\left(\sum_s \pi_s \mathbf{C}^s \mathbf{C}^{s'} \right) - \left(\sum_s \sum_{s'} \pi_s \pi_{s'} \mathbf{C}^s \mathbf{C}^{s''} \right) \right] \mathbf{X}
\end{aligned} \tag{6}$$

Denote $F = \sum_s \pi_s \mathbf{C}^s \mathbf{C}^{s'} - \sum_s \sum_{s'} \pi_s \pi_{s'} \mathbf{C}^s \mathbf{C}^{s''}$ as the quadratic matrix in the robust optimization model. Since the variance $\sum_s \pi_s \left(\xi_s - \sum_{s'} \pi_{s'} \xi_{s'} \right)^2$ is nonnegative, we can conclude that for any decision variables \mathbf{X} , $\mathbf{X}' F \mathbf{X} \geq 0$ from equation (6). Based on this derivation, it follows that the quadratic matrix F is a positive semi-definite matrix, which ensures a global minimum solution in this case.

The robust optimization model has the same size as the two-stage stochastic program with $T \cdot S + T \cdot S \cdot G + G$ constraints and $T \cdot S + T \cdot S \cdot G + G \cdot Y$ variables, $G \cdot Y$ of which are integer. This quadratic mixed integer programming problem can also be solved by the CPLEX of TOMLAB/CPLEX in Matlab. However it takes a much longer time to solve than the linear mixed integer programming problem, as the size of problem increases.

CHAPTER 4. SCENARIO TREE GENERATION

For implementation, we collected the real data of year 2008 from the Energy Information Administration (EIA), Midwest Independent System Operator (MISO) and Joint Coordinated System Planning Report 2008 (JCSP) [53]. EIA is an independent statistical agency providing data, analysis and future projection within the U.S. Department of Energy (DOE). MISO is an independent system operator and the regional transmission organization which monitors the transmission system and provides safe and cost-economic delivery of electric power across Midwest United States and one state, Manitoba, in Canada. JCSP is a joint organization in the Midwest and Northeast regions of America formally initiated in November 2007. Both economic and reliability studies have been conducted to develop a conceptual overlay to accommodate the potential 20% wind energy mandate in the future years. Year 2008 is considered as the reference year in our case study, since all the assumptions made for the later years are based on the 2008 data.

The uncertainties considered in the case study are both electricity demand and natural gas price.

4.1 Stochastic Process

In order to model the future uncertainties over multiple years, demand and natural gas price, respectively represented by $D(y)$ and $G(y)$, are considered as continuous time random variables. We need to fit a model for their evolution over time.

Since both the demand and natural gas price are usually modeled with an annual growth rate relative to the previous year, which is equivalent to geometric growth over time, and these annual growth rates in different years are taken to be mutually independent, we need to find an appropriate stochastic process which best satisfies these characteristics to model the uncertainties.

4.1.1 Geometric Brownian Motion

A continuous time stochastic process $Z(y)$ is a Brownian motion with drift coefficient μ and variance parameter σ^2 if $Z(0)=0$, $Z(y)$ has stationary and independent increments, and $Z(y)$ is normally distributed with mean μt and variance $t\sigma^2$ [54].

If $Z(y)$ is a Brownian motion with drift coefficient μ and variance parameter σ^2 , then the stochastic process $X(y) = e^{Z(y)}$ is defined as a Geometric Brownian motion (GBM), which is mostly applicable for modeling the financial market [55]. It has the statistical property that $w(y) = \log\left(\frac{X(y+1)}{X(y)}\right)$ is normally distributed with mean μ_x and standard deviation σ_x . In addition, the log ratios $w(y)$ are mutually independent.

Considering that the continuous time random variables, annual electricity demand and natural gas price, also possess the similar characteristic, with an annual geometric growth rate uncorrelated in different years, GBM might be a reasonable assumption for the random variables $D(y)$ and $G(y)$.

4.1.2 Verification of Geometric Brownian Motion

To test that both the annual electricity demand and the natural gas price can be represented as GBM, we obtained hourly demand data from year 1991 to 2007 for the Midwest region from the MISO website, and calculated the average annual natural gas price data from EIA by state in Midwest region from year 1970 to 2006, weighted by their consumption.

The annual data were first transformed to logarithm format by computing $w_D(y) = \log\left(\frac{D(y+1)}{D(y)}\right)$ and $w_G(y) = \log\left(\frac{G(y+1)}{G(y)}\right)$, and then statistical software JMP was used to fit a normal distribution to the data. By performing a goodness of fit test on each data series, we found that both $w_D(y)$ and $w_G(y)$ are consistent with observations from normal distributions, $N(\mu_D, \sigma_D)$ and $N(\mu_G, \sigma_G)$, respectively with $\mu_D = 0.0072$, $\sigma_D = 0.0094$, $\mu_G = 0.037$ and $\sigma_G = 0.082$. The related JMP outputs are shown in Figures 1 and 2. They show the histogram, moment and normal probability plot of the log ratios of the demand and natural gas price in the Midwest region respectively from year 1991-2007 and year 1970-2006.

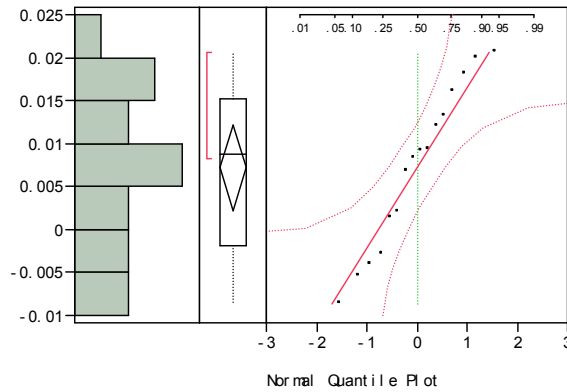


Figure 1. Log ratios of annual demand in Midwest region from year 1991-2006

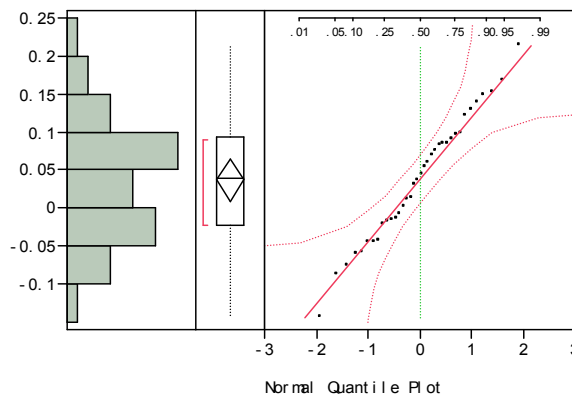


Figure 2. Log ratios of annual natural gas price in Midwest region from year 1970-2006

Since the Shapiro-Wilk test statistics for log ratios of demand is 0.951568 and p-value is 0.5149, it fails to reject the null hypothesis that the data is from the normal distribution. Similarly, since the Shapiro-Wilk test statistics for log ratios of natural gas price is 0.985879 and p-value is 0.9237, it fails to reject the null hypothesis that the data is from the normal distribution as well. Thus, we conclude that the lognormal distribution is a reasonable representation for each data set.

Besides the test of normal distribution, we also test the correlation between $w_D(y+1)$ and $w_D(y)$, $w_G(y+1)$ and $w_G(y)$ for each y , and furthermore confirm the independence of successive values of both $w_D(y)$ and $w_G(y)$. The related JMP outputs are shown in Figures 3 and 4.

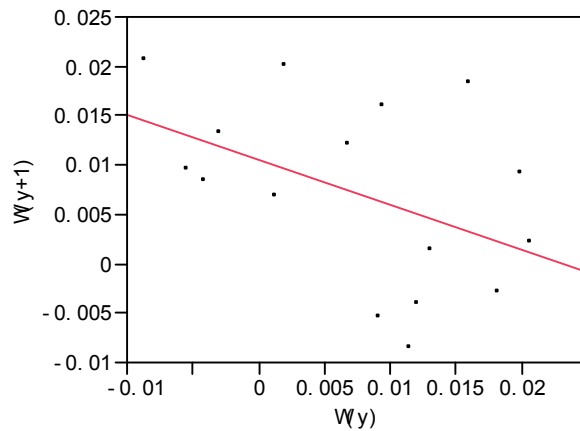


Figure 3. Correlation of annual demand in Midwest region from year 1991-2006

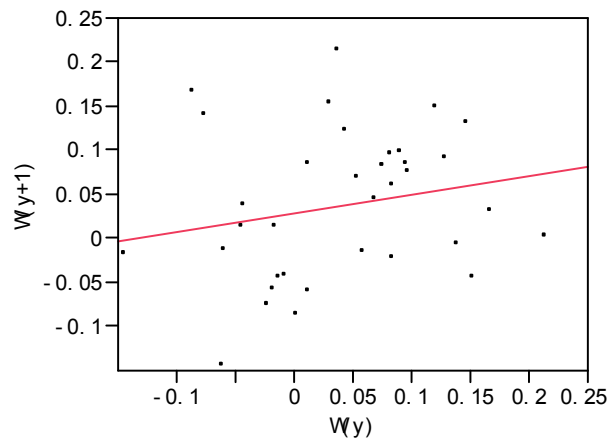


Figure 4. Correlation of annual natural gas price in Midwest region from year 1970-2006

The R-Square for the log ratios of demand is 0.208272 and the R-Square for the log ratios of natural gas price is 0.041814, and the p-values are respectively 0.0756 and 0.2387, thus we fail to reject the null hypothesis of the zero correlation.

Therefore, the assumption that both $D(y)$ and $G(y)$ are GBM has been verified [56].

Another way to verify the independence is through the autocorrelation test with different lags of the time series model in JMP.

Table 1. Time series autocorrelation with lag = 1 for demand





Lag	AutoCorr	Plot Autocorr	p-Value
0	1.0000		.
1	-0.4075		0.0742

Table 2. Time series autocorrelation with lag = 1 for natural gas price

Lag	AutoCorr	Plot Autocorr	p-Value
0	1.0000		.
1	0.2011		0.2086

The null hypothesis is that there is no autocorrelation. For the time series result for historical demand in Table 1, the p-value 0.0742 with lag =1. For the time series result for historical natural gas price in Table 2, the p-value is 0.2086 with lag = 1. Both of the p-values fail to reject the null hypothesis, which indicates there are no autocorrelation for the time series data with lag = 1.

4.1.3 Statistical Properties of Random Variables

Since $w(y) = \log\left(\frac{X(y+1)}{X(y)}\right)$ is normally distributed with mean μ_x and standard deviation σ_x , the ratio $\frac{X(y+1)}{X(y)}$ satisfies the lognormal distribution with mean μ_x and standard deviation σ_x , based on which we can further derive the following statistical properties of the GBM by the following formulas for the lognormal distribution [57]:

$$E\left(\frac{X(y+1)}{X(y)}\right) = e^{\mu_x + \frac{\sigma_x^2}{2}} \quad (7)$$

$$\text{Var}\left(\frac{X(y+1)}{X(y)}\right) = (e^{\sigma_x^2} - 1)e^{2\mu_x + \sigma_x^2} \quad (8)$$

$$\text{sk}\left(\frac{X(y+1)}{X(y)}\right) = (e^{\sigma_x^2} + 2)\sqrt{e^{\sigma_x^2} - 1} \quad (9)$$

Denote by $x(y)$ the actual value in year y . Assume that the initial year of the expansion planning is year 0 and there is no uncertainty in year 0 with known $x(0)$, based on which we can continue to calculate the conditional mean, standard deviation and skewness of $X(1)$ for the next year.

Given equations (7), (8), (9) and $X(0) = x(0)$, we derive the conditional formulas as follows in equation (10), (11), (12):

$$\begin{aligned}
E(X(y+1)|X(u), 0 \leq u \leq y) &= E(e^{Z(y+1)} | Z(u), 0 \leq u \leq y) \\
&= E(e^{Z(y)+Z(y+1)-Z(y)} | Z(u), 0 \leq u \leq y) \\
&= e^{Z(y)} E(e^{Z(y+1)-Z(y)} | Z(u), 0 \leq u \leq y) \\
&= X(y) E\left(\frac{X(y+1)}{X(y)}\right) = X(y) e^{\mu_x + \frac{\sigma_x^2}{2}}
\end{aligned} \tag{10}$$

$$\begin{aligned}
Var(X(y+1)|X(u), 0 \leq u \leq y) &= Var(e^{Z(y+1)} | Z(u), 0 \leq u \leq y) \\
&= Var(e^{Z(y)+Z(y+1)-Z(y)} | Z(u), 0 \leq u \leq y) \\
&= (e^{Z(y)})^2 Var(e^{Z(y+1)-Z(y)} | Z(u), 0 \leq u \leq y) \\
&= X(y)^2 Var\left(\frac{X(y+1)}{X(y)}\right) \\
&= X(y)^2 (e^{\sigma_x^2} - 1) e^{2\mu_x + \sigma_x^2}
\end{aligned} \tag{11}$$

$$\begin{aligned}
sk(X(y+1)|X(u), 0 \leq u \leq y) &= sk(e^{Z(y+1)} | Z(u), 0 \leq u \leq y) \\
&= sk(e^{Z(y)+Z(y+1)-Z(y)} | Z(u), 0 \leq u \leq y) \\
&= \frac{E\left(e^{Z(y)} \left(e^{Z(y+1)-Z(y)} - E(e^{Z(y+1)-Z(y)})\right)\right)^3}{\left(Var\left(e^{Z(y)} \left(e^{Z(y+1)-Z(y)}\right)\right)\right)^{\frac{3}{2}}} \\
&= \frac{E\left(e^{Z(y+1)-Z(y)} - E(e^{Z(y+1)-Z(y)})\right)^3}{\left(Var\left(e^{Z(y+1)-Z(y)}\right)\right)^{\frac{3}{2}}} \\
&= sk\left(\frac{X(y+1)}{X(y)}\right) = (e^{\sigma_x^2} + 2) \sqrt{e^{\sigma_x^2} - 1}
\end{aligned} \tag{12}$$

From the equations (10), (11) and (12) for the conditional statistical properties, the conditional expectation and variance in later years both depend on the values for the previous

year. However, the skewness is independent over the years, and thus remains the same, only depending on σ_x .

Apply (10), (11) and (12) to the annual demand and annual natural gas price in the Midwest region and the derived results are summarized in Table 3.

Table 3. Statistical properties of random variables over years

Random Variables	Statistical property	First Year ($y = 1$)	After First Year ($y > 1$)
Demand (Billion MWhs), $D(y)$	Mean	$1.00727 d(0)$	$1.00727 d(y-1)$
	Standard deviation	$0.009469 d(0)$	$0.009469 d(y-1)$
	Skewness	0.028	0.028
Natural Gas Price (\$/Thousand Cubic Feet), $G(y)$	Mean	$1.041188 g(0)$	$1.041188 g(y-1)$
	Standard deviation	$0.085521 g(0)$	$0.085521 g(y-1)$
	Skewness	0.25	0.25

The correlation value between the two random variables in each year was also obtained by JMP as shown in Figure 5. In general, the annual natural gas price and electricity demand both have increasing trends over the year.

The R-Square for the linear fitness between the annual demand and annual natural gas price is 0.75002, and the p-value is $<.0001^*$, thus we reject the null hypothesis of the zero correlation.

A correlation of 0.866 was indicated by the JMP outputs. Hence, there is a strong positive correlation between the total annual electricity demand and average annual natural gas price over the years.

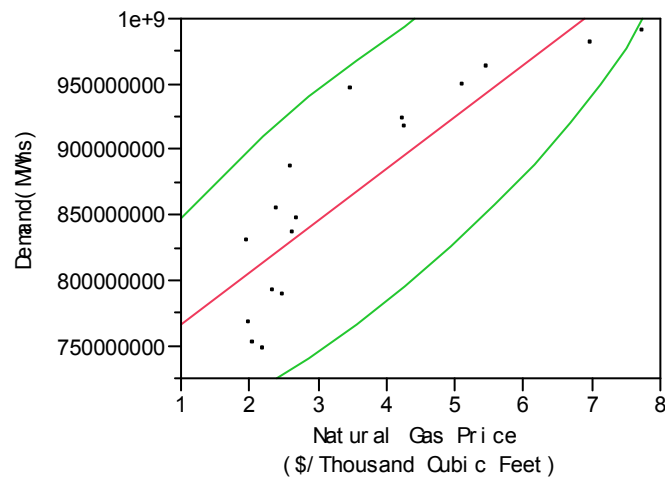


Figure 5. Correlation of total annual demand and average annual natural gas price in Midwest region from year 1991-2006

4.2 Scenario Generation Method

4.2.1 Scenario Generation for a Single Year

Once the distribution assumption for the uncertain variables has been made, we can further generate discrete scenarios to well represent the random variables. We construct

scenario vector \mathbf{X}_s with the same dimension as the number of uncertain variables, and scenario probability π_s , to approximate their statistical specifications $f_i(\mathbf{X}_s, \pi_s)$ to the statistical specification P_{VAL_i} of the original continuous space. For instance, if $f_i(\mathbf{X}_s, \pi_s)$ represents mean \bar{X} , then $f_i(\mathbf{X}_s, \pi_s)$ can be calculated as $\sum_s \pi_s \mathbf{X}_s$.

The optimization model for the scenario generating method [47] is shown in (13).

$$\begin{aligned} \min_{\mathbf{X}_s, \pi_s} \sum_{i \in P} \omega_i (f_i(\mathbf{X}_s, \pi_s) - P_{VAL_i})^2 \\ \sum \pi_s = 1 \\ \pi_s \geq 0 \end{aligned} \quad (13)$$

In equation (13), P is a set of statistical properties, the ones that we already know from the original distribution, i stands for one specification from set P , and P_{VAL_i} is the value of the i th statistical specification. The square norm of distance from the original P_{VAL_i} and the generated scenarios $f_i(\mathbf{X}_s, \pi_s)$ are measured and minimized. A set of weights ω_i can be manually specified depending on personal preference. In our case study, we use weights of 2 for the means and variances of both random variables, and 1 for the skewnesses and correlation. The constraints simply indicate that the scenario probabilities add up to one and are non-negative.

Hoyland and Wallace [47] also discuss the appropriate number of scenarios for the optimization problem. To avoid both underspecifications and overspecifications, the number of statistical specifications should be close to the number of decision variables. In our case study, we have a two-dimensional scenario variable to represent both demand and natural gas price, and one scenario probability needs to be decided. The number of decision variables is

$S \cdot (2+1) - 1$, since all the scenario probabilities adding to 1 eliminates one degree of freedom. Regarding the number of the statistical specifications, there are 7, including the mean, standard deviation, and skewness of each of the two random variables, as well as their correlation. Because the minimal S that leads to 7 statistical specifications is 3, the number of scenarios is determined to be 3 at a time.

This scenario generation problem is a nonlinear programming problem with nonlinear objective function and linear constraints. It can be solved by the nonlinear solver Tomlab/SNOPT in Matlab, which requires single or multiple starting points for the iteration leading to the optimal solution. The initial points for the 3 scenario vectors and scenario probabilities are assumed to be $X_1 = (\mu_D - \sigma_D, \mu_G - \sigma_G)'$, $\pi_1 = 0.333$, $X_2 = (\mu_D, \mu_G)'$, $\pi_2 = 0.333$, $X_3 = (\mu_D + \sigma_D, \mu_G + \sigma_G)'$ and $\pi_3 = 0.333$. The minimum possible objective value is expected to equal zero, if the specifications are consistent. However, since (13) is generally not a convex optimization problem, the final solution might end up with a local optimal solution, which has a nonzero objective value. If the derived statistical properties are still close to the specification, the local solution is also acceptable. But if severe inaccuracy occurs, we might need to resolve it by either resetting the weight coefficient ω_i or increasing the number of initial starting points for a benefit to the statistical specifications.

4.2.2 Evolution over Future Years

Once the 3 scenarios for year 2009 are generated based on the known information $d(0)$ and $g(0)$, we generate the scenarios for year 2010 similarly. Conditional statistical properties are first specified based on the 3 generated scenario outcomes of year 2009 by

applying equations (10), (11) and (12). Then another 3 discrete scenarios are generated from equation (13).

The final scenario tree can be recursively constructed accordingly until the end of the planning horizon. A fragment of the scenario tree for our 10-year case study is shown in Figure 6.

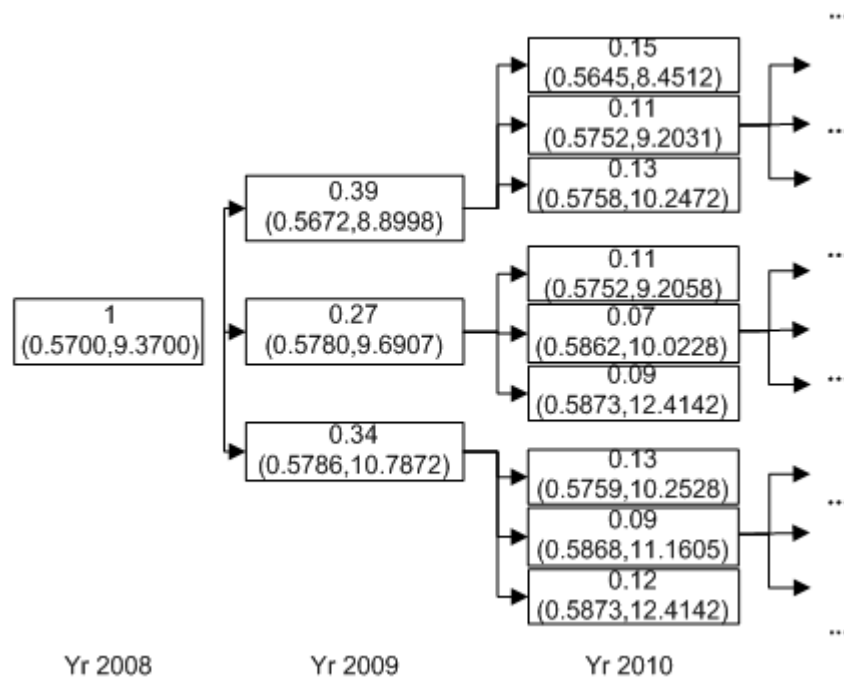


Figure 6. A scenario tree for multi-year horizon

Each column in the scenario tree represents one single year, and each tree node represents one possible outcome for a year. For each node, the number on the top stands for the product of the probabilities for that specific scenario path up to that tree node. The numbers in the parenthesis on the bottom are the scenario values for both demand and natural gas price. The initial year in this case is year 2008, with the known value $d(0)=0.57$ and

$g(0)=9.37$. The units for demand and natural gas price are respectively billion MWhs and \$/thousand cubic feet.

In year 2008, demand is 0.57 and natural gas price is 9.37. Given the node information for 2008, statistical specifications for year 2009 can be calculated based on Table 3, and three possible scenarios can be generated for year 2009 by solving the optimization problem (13): with probability 0.3911, the demand is 0.5672 and natural gas price is 8.8998; with probability 0.2689, the demand is 0.5780 and the natural gas price is 9.6907; with probability 0.3400, the demand is 0.5786 and the natural gas price is 10.7872. For year 2010, in the same way, based on each of the possible scenarios generated for year 2009, statistical specifications are calculated for year 2010 and 3 scenarios are generated. We recursively update the statistical specification for each year based on one of the scenarios in previous year until the end of the planning horizon, year 2017.

CHAPTER 5. CASE STUDY

In this chapter, a multi-year case study is formulated. Parameter and scenario assumptions are made and experiment results for the deterministic model, two-stage stochastic programming model and robust optimization model are illustrated and compared. We also examine the tradeoff between cost and risk in robust optimization model by varying the penalty cost p_v for the cost variance.

5.1 Scenario Reduction

In Figure 6, each tree node has three branches. Since the number of scenarios increases exponentially, at the final year it will be tremendously huge. For instance, if we have a 10 year horizon, then the total number of scenarios in year 10 will be $3^{(10-1)} = 19,683$. We define one of the scenarios at the end of horizon with all their parent nodes back to the initial year as one scenario path. We also refer to a scenario path as a “scenario” in the remainder of this thesis for the multi-year horizon case study.

In order to reduce the computational complexity, we need some scenario reduction technique. To select the scenarios for the case study, we used the naïve sampling to randomly select a small number of scenarios we need for the case study and rescaled the scenario probabilities to make them add up to 1.

For the case study in this Chapter, we selected 5 scenarios for the illustration simplicity. We assume the planning horizon 10 years, and the 6 different types of power plant. Hence, the deterministic model is a 216×270 mixed integer programming problem,

two-stage stochastic programming is a 1056×1110 mixed integer programming problem and robust optimization is a 1056×1110 quadratic mixed integer programming with the 1110×1110 quadratic matrix F .

5.2 Assumptions

For the multi-year case study, we made the following parameter and scenario assumptions. Most of the data come from the EIA, MISO and JCSP.

We assume six different generators, BaseLoad, CC, CT, nuclear, wind and IGCC, as the candidate generators to invest for the future expansion planning. CC, representing a combined cycle power plant, and CT, representing a combustion turbine power plant, are both fueled by natural gas. IGCC, an integrated gasification combined cycle power plant, is fueled by coal.

5.2.1 Demand

For each year, we assume 3 sub-periods, the peak, medium and low. We order hourly demand from the highest to the lowest as one load duration curve (LDC), and respectively separate them into three sub-periods by the top quarter, middle half and bottom quarter of the load as shown in Figure 7. The accumulated load for each sub-period stands for the demand in it. By considering only three sub-periods, we reduce the problem size and also retain the chronological demand variability. Figure 7 is an example based on the Midwest hourly load in year 2008 from the real-time market report of MISO [58].

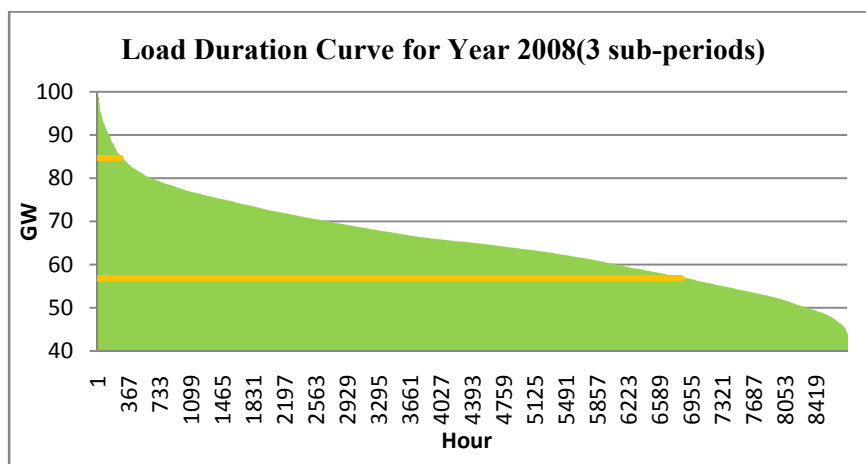


Figure 7. Load duration curve in year 2008 (3 sub-periods)

Thus the demand data for 3 sub-periods in year 2008 can be summarized in Table 4. Since the LDCs for the later years are unknown, we assume them to be the same as the one in year 2008. Once the assumption of the net generation has been made, we can further scale its demand load for each sub-period according to Table 4.

Table 4. Demand data for 3 sub-periods in year 2008

Hours(h)	Demand(MWh)
271	0.02431*10E9
6574	0.44637*10E9
1938	0.10097*10E9

For the demands to satisfy by the newly built power plants in our case studies, we also assume them to be the incremental demands from the reference year 2008. For example, if we have total generation 570 million MWhs in year 2008, and 578 million MWhs is

predicted in year 2009, then the demand for year 2009 will be 8 million MWhs. But if we have a demand decrease, then the incremental demand will be simply assumed to be zero. And we assume the u_g to be zero accordingly.

In the case study, the demand is considered to be an uncertain variable. The 5 different scenario paths for it are shown in Figure 8. Since we assume the demand as the incremental demand compared with the initial year, the demand data for year 2008 are all zeros.

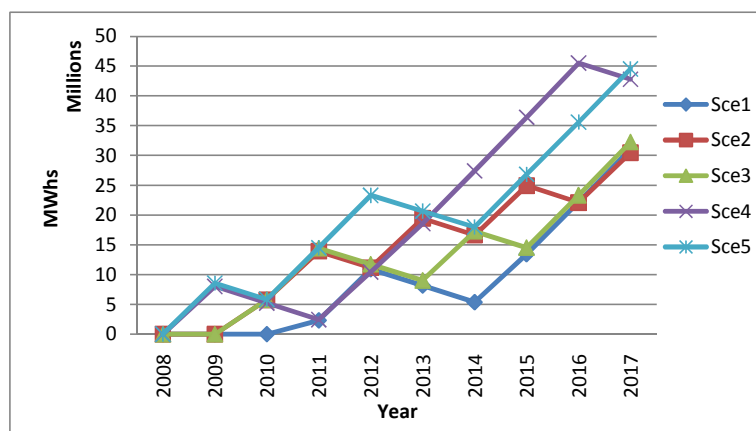


Figure 8. Scenario incremental demand

5.2.2 Annual Interest Rate for Discounting

We assume the annual interest rate $r = 0.08$ based on JCSP [53]. This rate is used to discount the future expenditures to the present value in year 2008, which represents the reference year, as well as the initial year.

5.2.3 Build Cost for Generators

The calculation of the build cost c_g for the generators except the wind farm is based on the capital expenditure profile suggested in JCSP [53], shown in Table 4. Since the construction time for a wind farm is 2 years from JCSP [53], 50% of the capital expenditure for each year is assumed. To obtain c_g , we sum up the present value for each year by using the discount rate r . Table 5 indicates both the overnight investment cost and the final calculated build cost c_g . Table 6 shows all the percentages of the overnight investment cost actually spent in each year to build the generators.

For instance, if we want to calculate the build cost for CC power plant, we first simply multiply the overnight build cost by the capital expenditure percentage for each year. Then we will get for the first year $1.833 \cdot 10^6 \cdot 0.25$, the second year $1.833 \cdot 10^6 \cdot 0.5$ and the third year $1.833 \cdot 10^6 \cdot 0.25$. We discounted them by r to the present value in the first year and summed them up in equation (14).

$$\begin{aligned}
 & 1.833 \cdot 10^6 \cdot 0.25 + \frac{1.833 \cdot 10^6 \cdot 0.5}{(1+0.08)} + \frac{1.833 \cdot 10^6 \cdot 0.25}{(1+0.08)^2} \\
 & = 214250 + 396759.3 + 183684.8 \\
 & = 794694.1
 \end{aligned} \tag{14}$$

Thus the build cost for CC is \$794,694.10/MW.

Table 5. Overnight cost for generators

	<i>Overnight Building Cost (\$/MW)</i>	<i>Build cost c_g (\$/MW)</i>
BaseLoad	1.833*10E6	1.446*10E6
CC	0.857*10E6	0.795*10E6
CT	0.597*10E6	0.575*10E6
Nuclear	2.928*10E6	1.613*10E6
Wind	1.713*10E6	1.650*10E6
IGCC	2.118*10E6	1.671*10E6

Table 6. Capital expenditure profile for generators

Year	BaseLoad	CC	CT	Nuclear	Wind	IGCC
1	0.02	0.25	0.5	0.01	0.5	0.02
2	0.03	0.5	0.5	0.01	0.5	0.03
3	0.25	0.25		0.01		0.25
4	0.30			0.01		0.3
5	0.30			0.01		0.3
6	0.1			0.02		0.1
7				0.03		
8				0.2		
9				0.3		
10				0.3		
11				0.1		

5.2.4 Generation Cost for Generators

Generation cost involves two parts: the variable operation and maintenance (O&M) cost and the fuel cost. All the related parameters for calculating the generation cost for year 2008 are shown in Tables 7 and 8 from JCSP [53]. From Table 7, we can easily calculate the generation cost for each generator.

Table 7. Generation cost related parameters for the generators in the first year

	<i>Fuel Price(\$/Mbtu)</i>	<i>Heat Rate(Btu/kwh)</i>	<i>Efficiency</i>	<i>Variable O&M(\$/MWh)</i>
BaseLoad	3.37151	8844	0.4	4.7
CC	$G(0)/1.028$	7196	0.56	2.11
CT	$G(0)/1.028$	10842	0.4	3.66
Nuclear	0.00093	10400	0.45	0.51
Wind	0	N/A	N/A	5
IGCC	3.37151	8613	0.48	2.98

Since CC and CT plants are fueled by natural gas, their generation costs are random variables depending on scenarios. In order to transform the natural gas price, \$/thousand cubic feet, into the formal format of the energy price, we made the change based on Table 8.

Table 8. Unit transformation for the natural gas price

	<i>\$/thousand cubic feet</i>	<i>Btu/thousand cubic feet</i>	<i>\$/Mbtu</i>
Natural Gas Price	$G(0)$	$1.028*10E6$	$G(0)/1.028$

As for the later years, we made the escalation assumptions that 2% annual growth rate was applied to the fuel price (coal, nuclear, wind) and 3% annual growth rate was applied to the variable O&M cost, suggested by JCSP [53]. For the CC and CT power plant, since the fuel price is a continuous random variable, we can simply replace the $G(0)$ in Tables 7 and 8 by $G(y)$, and calculate the fuel cost.

The generation cost for the units not fueled by natural gas over the years are in Figure 9. And the generation cost for the CC and CT plants under 5 different scenarios are in Figure 10.

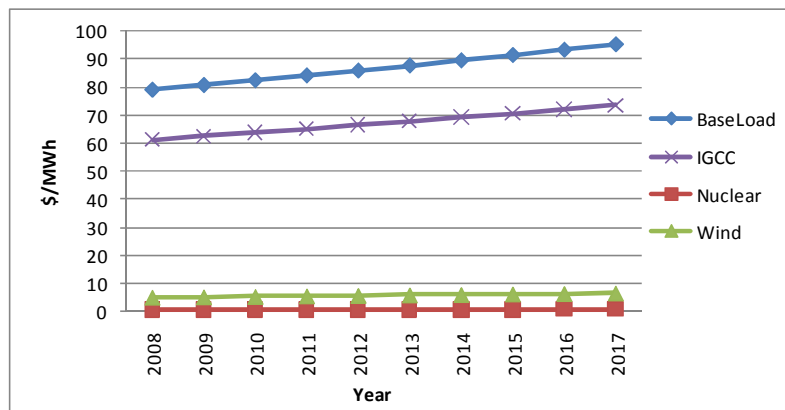


Figure 9. Generation cost for BaseLoad, IGCC, wind and nuclear power plants

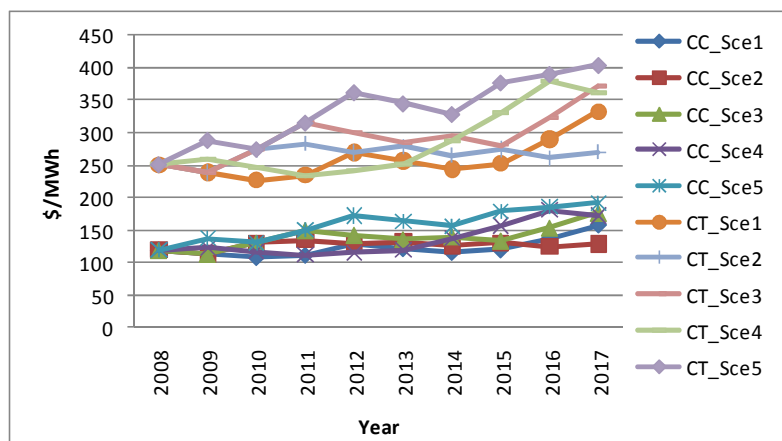


Figure 10. Scenario generation cost for CC and CT power plants

5.2.5 Capacity for Generators

The installed capacity and generator ratings are based on the JCSP [53] and the generator ratings are calculated by their installed capacity multiplied by the forced outage rate (FOR), also from the JCSP [53]. The installed capacity is for calculating the investment cost of the power generation expansion, and the rating is considered as a maximum capacity for the electricity generation in the future daily operation. The assumptions for them are shown in Table 9.

Table 9. Installed capacity and generator rating for generators

Type	Baseload	CC	CT	Nuclear	Wind	IGCC
Generators, g	1	2	3	4	5	6
Install Capacity(MW), m_g^{\max}	1200	400	400	1200	500	600
Generator Rating(MW), n_g^{\max}	1130	390	380	1180	175	560

5.2.6 Maximum Units to Build for Generators

For the maximum units to build constraint over the whole planning horizon, we used the following assumption in Table 10.

Table 10. Max units to build for generators

<i>Type</i>	<i>Baseload</i>	<i>CC</i>	<i>CT</i>	<i>Nuclear</i>	<i>Wind</i>	<i>IGCC</i>
Max Units Built, u_g^{\max}	4	10	10	1	45	4

5.2.7 Scenario Probabilities

A sample of scenarios was taken using a random number generator so that the probability, s , selected was proportional to its probability π_s . The scenario probabilities for the 5 randomly selected scenarios are respectively 6.89491×10^{-5} , 4.05860×10^{-5} , 8.86573×10^{-5} , 5.00286×10^{-5} and 5.30519×10^{-5} . We rescale them to 0.2289, 0.1347, 0.2943, 0.1660 and 0.1761 so that they add up to 1.

5.2.8 Lead Time for Generators

In our case study, assuming that the newly built generators are able to generate electricity ever since the first year that we made the expansion decision, i.e., the lead time for building and installing a generator is ignored.

5.2.9 Penalty Costs

The penalty for USE p_u is 100,000 \$/MWh and the penalty for cost variance p_v is 1.

5.3 Experiment Results

The experiment results for the 10-year case study are in Table 11.

Table 11. Experiment result for a 10-year case study

<i>Method</i>		<i>Total Expected Cost</i>					<i>Robustness</i>	
		<i>(Billions of \$)</i>					<i>(Billions of \$)</i>	
		<i>Scce 1</i>	<i>Scce 2</i>	<i>Scce 3</i>	<i>Scce 4</i>	<i>Scce 5</i>	<i>Expected</i>	<i>Standard Deviation</i>
Deterministic	Scce 1	5.8	1234.2	893.0	3507.8	2983.9	1538.2	1297.9
	Scce 2	7.8	7.7	11.3	1598.8	1340.8	507.8	695.2
	Scce 3	6.4	32.2	7.2	1988.7	1195.6	548.7	779.5
	Scce 4	8.6	96.8	98.5	12.4	568.7	146.2	199.3
	Scce 5	9.2	9.8	9.5	62.9	11.9	18.8	19.7
	Expected	6.8	104.8	91.0	770.2	511.7	260.4	284.1
Stochastic		7.8	9.2	8.6	13.6	12.8	10.1	2.3
Robust0		10.3	11.7	11.0	14.1	13.7	11.9	1.5

We have six deterministic models, a two-stage stochastic programming model and a robust optimization model (with solution called “Robust 0” because $p_v = 10^0$). The deterministic models have the fixed parameters. For the first five scenario models, the expansion planning decisions are determined by only considering one of the scenarios at a time while ignoring the other four. However, although it is a deterministic model, the “Expected” method considers the future uncertainties by assuming its parameters based on the expected value of all the scenario values.

The different models give different optimal expansion planning decisions, based on which, we can further optimize the actual “Total Cost” under each of the future scenarios. For instance, cost for the scenario 1 model under scenario 1 is found by solving (1) - (4) using $s = 1$ with $\pi_s = 1$. Denote the expansion plan as $U^1 = [U_{g,y}^1]$. The cost of this solution if scenario s occurs is

$$z_1^1 = \sum_y \left(\frac{\sum_g (c_g \cdot m_g^{\max} \cdot U_{g,y}^1)}{(1+r)^y} \right) + \min \sum_y \left(\frac{\sum_{t \in T_y} \left(\sum_g (l_{g,t,1} L_{g,t,1}) + p_u E_{t,1} \right)}{(1+r)^y} \right)$$

$$\text{s.t.} \quad \sum_g L_{g,t,1} + E_{t,1} = d_{t,1} \quad \forall t \quad (15)$$

$$L_{g,t,1} \leq n_g^{\max} \left(u_g + \sum_{y \leq Y_t} U_{g,y} \right) \quad \forall g, t$$

$$\sum_y U_{g,y} \leq u_g^{\max} \quad \forall g$$

Since each of the scenarios happens with some specific probability, we calculate the “Expected” $\bar{z}^1 = \sum_s \pi_s \cdot z_s^1$ to determine whether an expansion planning decision is cost-efficient or not. A standard deviation $\sqrt{\sum_s \pi_s \cdot (z_s^1 - \bar{z}^1)^2}$ among the scenarios has also been calculated to indicate the robustness of the expansion planning decision.

In general, based on Table 11, the “Stochastic” and “Robust 0” solutions perform much better than the deterministic models on both expected cost and robustness of the solution. More specifically, “Stochastic” gives the smallest expected total cost, and “Robust 0” gives the least standard deviation of the total cost amid scenarios.

In Figure 11, we made cost comparison amid scenarios for both “Stochastic” and “Robust 0”. It shows the three cost components contributing to the total cost: investment cost (INV), generation cost (GEN), and penalty cost for unmet demand (USE), which are zeros

for all of the scenarios in this case study. And the total costs are shown on the top at each column.

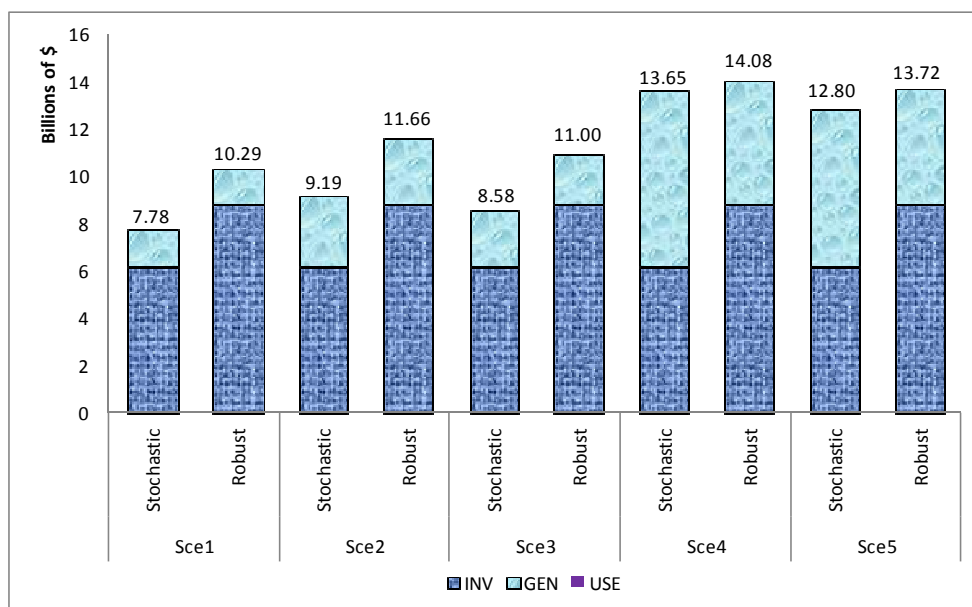


Figure 11. Investment cost, generation cost, USE cost for both stochastic and robust

Regarding the investment cost, it remains the same under all the scenarios for both the “Stochastic” and “Robust 0” solutions since once the expansion planning decision is made, it will not change for any future scenario.

As we can see from the Figure 11, the “Robust 0” solution spends around 1/3 more on the generation expansion investment than the “Stochastic” to ensure a more robust investment decision, which can easily adjust to the future scenarios: if it’s a scenario of high demand, more generation expansion investment will help meet the demand to avoid the penalty cost for USE; if it’s a scenario of either extremely low or high natural gas price, it will help save generation cost as well by altering the generation preference for different type

of power plants. In Figure 11, it also indicates less generation cost of the “Robust 0” compared with the “Stochastic”. However, although the “Robust 0” solution might save the generation cost, or sometimes the USE cost, its total costs for the scenarios, in this case study, are all higher than the “Stochastic”.

The expansion planning decisions for “Stochastic” and “Robust 0” are shown in Figures 12 and 13.

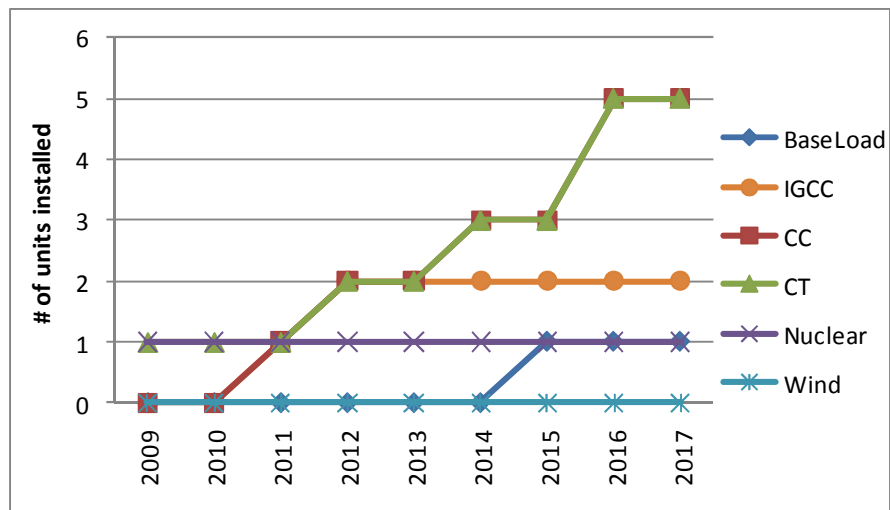


Figure 12. Cumulative expansion planning decision in “Stochastic” solution

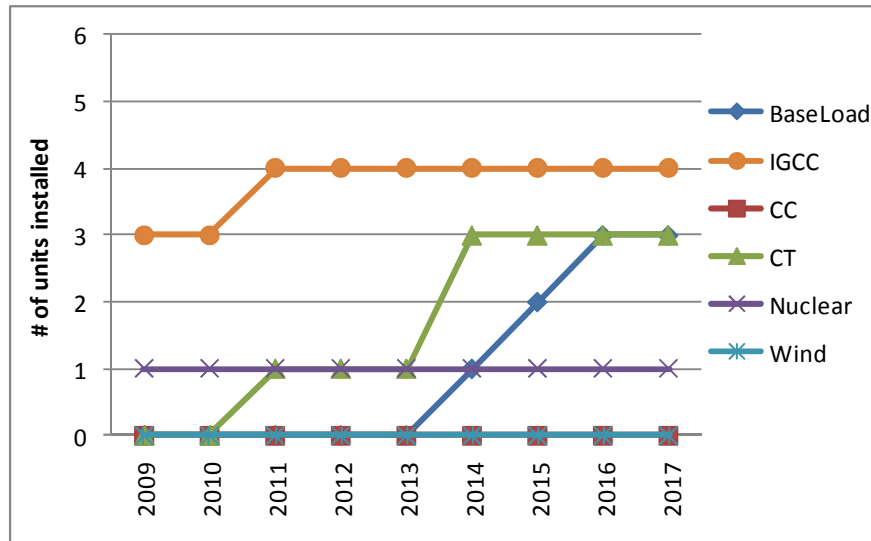


Figure 13. Cumulative expansion planning decision in “Robust 0” solution

In the “Stochastic” model, the tradeoff is between the investment cost, and the generation cost and USE cost. It’s also a tradeoff of the total cost among scenarios. “Stochastic” is looking for the minimum total expected cost. In this case study, the tradeoff is specifically the potential large amount of generation cost for both scenario 4 and 5 with high natural gas price and high demand, and the cost saved by expansion investment on CC and CT plants.

On the other hand, for the “Robust 0” plan, the tradeoff is between the total expected cost and the cost variance amid scenarios. With $p_v=1$ as the weight of the cost variance in the model, in this case study the “Robust 0” solution builds more coal-fueled plants such as BaseLoad and IGCC, but much fewer natural gas-fueled plants, such as CC and CT. Since both the building cost and the generation cost of CC plants are much more expensive than the CT’s, the “Robust 0” solution doesn’t consider building the CC plants at all.

Both the “Stochastic” and “Robust 0” solutions do not make investment decision in wind farm due to its relatively more expensive building cost and much lower capacity credits. In addition, because the production tax credits for the renewable energy were not included in this case study, it makes the wind farm less attractive than it actually is in the real world nowadays. However, since the PTC has been allowed to lapse in the past, it might not be something that the planners assume for the expansion planning decision.

5.4 Sensitivity to Robustness Parameter

In this section, we will continue to discuss the penalty parameter p_v for the robust optimization. The weight p_v plays a very significant role in the tradeoff between the total expected cost and the cost variance. We don’t want it to overemphasize the importance of the cost variance relative to the expected cost to the extent that it even overlooks our major concern, the total expected cost over scenarios.

A series of “Robust” models is studied with different penalty cost p_v . The value for the p_v , and the experiment results are summarized in Table 12 and Figure 14.

Table 12. The penalty cost p_v for 10 different robust optimization models

	1	2	3	4	5	6	7	8	9	10
p_v	10E-1	10E-2	10E-3	10E-4	10E-5	10E-6	10E-7	10E-8	10E-9	10E-10

In Figure 14, we compared the experiment results of the 10 additional robust models with the stochastic and robust models studied in section 5.3. From experiments 1 through 10,

p_v keeps decreasing. Hence, generally speaking, the total expected cost will decrease, and the standard deviation will increase, since we are putting more weight on the total expected cost while we are concerned less about the cost variance.

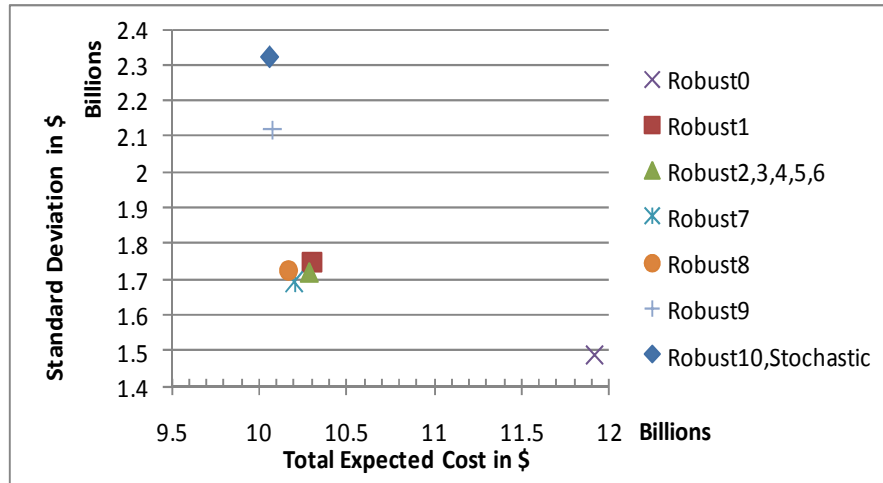


Figure 14. Standard deviation vs. expected cost for stochastic and robust optimization

In Figure 14, we can see that the “Robust 2, 3, 4, 5 and 6” models have the same results. And when p_v decreases to 10^{-10} , the “Robust 10” model has the same solution as the stochastic model.

As we discussed before, for the robust model, there is a tradeoff between minimizing the expected cost and minimizing the cost variance amid scenarios. In Figure 14, “Robust 2, 3, 4, 5, 6” is dominated by “Robust 7”, and “Robust 1” is dominated by “Robust 8”. Thus, they will not be considered as the best planning decisions. As for the rest of the solutions, further comparison must be made based on their total cost for each of the scenarios, shown in Figure 15.

In Figure 15, it is very easy to conclude that the “Robust 0” solution is the only one dominated by another solution, thus it can’t be the best available planning decision. For the other four, the decision can be made based on the planner’s preference, or some other assumptions.

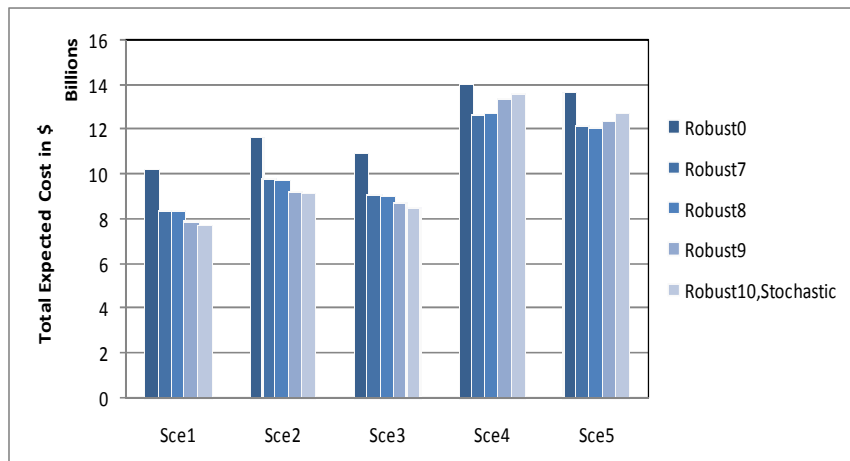


Figure 15. The cost over scenarios for five different robust solutions

In Figures 16-18, the expansion planning decisions for the “Robust 7”, “Robust 8” and “Robust 9” solutions are shown.

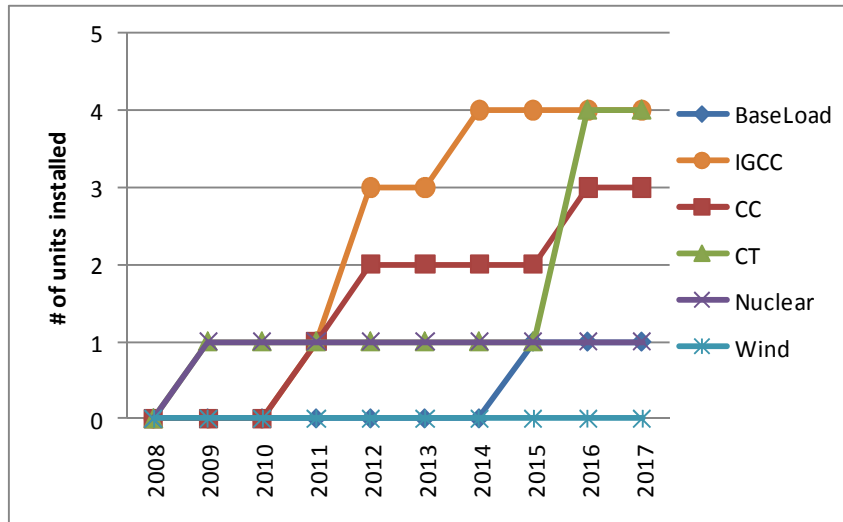


Figure 16. Cumulative expansion planning decision in "Robust 7" solution

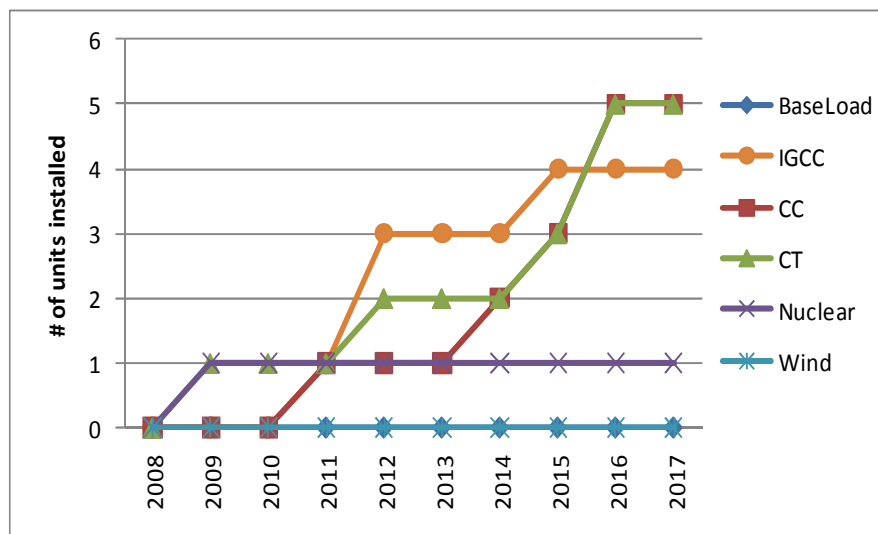


Figure 17. Cumulative expansion planning decisions in "Robust 8" solution

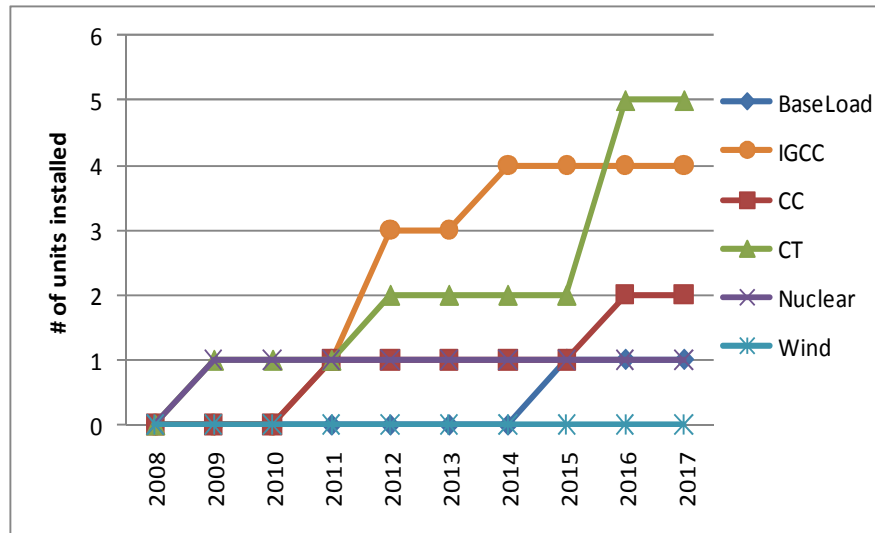


Figure 18. Cumulative expansion planning decisions in “Robust 9” solution

Since “Robust 0” has $p_v = 1$, “Stochastic” has $p_v = 0$ and “Robust 7, 8, 9” respectively have $p_v = 10^{-7}$, $p_v = 10^{-8}$ and $p_v = 10^{-9}$, “Robust 7, 8, 9” represent the efficient solutions between the “Robust 0” and “Stochastic”.

In Figures 16-18, BaseLoad and IGCC are coal-fired, CC and CT are natural gas-fired, and Nuclear and Wind are considered as “green” energy. It appears throughout “Robust 0”, “Robust 7”, “Robust 8”, “Robust 9” and “Stochastic” that green energy sources are not used much, coal reduces cost variance and natural gas reduces expected cost. As p_v decreases, the derived optimal solutions build more CC and CT plants with the lowest investment costs, though there is a potential risk of the high price of the future natural gas. On the other hand, the optimal solutions reduce the expansion units for BaseLoad plant, which has a relatively low and stable fuel price but a more expensive initial investment cost.

In addition, the expansion decisions on CC and CT tend to become more attractive for the last few years of the whole planning horizon. This is due to the different investment cost and generation cost of the different types of power plants. The model incorporates the tradeoff between the initial capital cost and generation cost in the later years. At the beginning of the planning horizon, although the initial investment cost for the coal-fired plants and nuclear plant are much more expensive than the ones for CC and CT, the much lower generation cost will help save more in the future years till the end of the horizon. When it comes to the last few years of the planning horizon, when the generation cost saved by the end of the horizon is no longer able to cover the expensive capital cost, CC and CT with the lowest investment cost becomes more economic. In this case, the model ignores the end effect of the 10-year planning horizon.

By varying p_v , we made different assumptions on the importance of total expected cost and cost variance. One of these expansion planning decisions can be further selected based on the planner's preference, as long as the two following criteria are satisfied:

- The solution is not dominated by any of the other solutions regarding both the total expected cost and cost variance amid scenarios
- The solution is not dominated by any of the other solutions regarding the total cost of all the scenarios.

For the final four candidate planning decisions, the final optimal one can be selected based on the planner's preference or some preferable criteria. For instance, if we want to select the solution with the minimal maximum regret over all the scenarios, we can calculate

their regret under each scenario, get the maximum and select the one with the minimum value of the maximum regret.

In Table 13, a regret table is shown.

Table 13. Regret over scenarios for “Robust 7”, “Robust 8”, “Robust 9”, “Robust 10”

	Regret (Billions of \$)					Maximum Regret (Billions of \$)
	Sce1	Sce2	Sce3	Sce4	Sce5	
Robust7	0.6372	0.6063	0.5683	0	0.0233	0.6372
Robust8	0.5782	0.5473	0.5093	0.0329	0	0.5782
Robust9	0.1904	0.1196	0.1693	0.6934	0.2617	0.6934
Robust10	0	0	0	0.8915	0.6195	0.8915

From Table 13, among the four candidate robust solutions, “Robust 10” is the optimal for scenarios 1-3, “Robust 7” is optimal for scenario 4, and “Robust 8” is optimal for scenario 5. The maximum regret in the last column indicates that “Robust 8” has the minimum maximum regret among the four with 5.782×10^8 . Hence, based on this criterion, “Robust 8” is selected as our final expansion planning decision.

5.5 Production Tax Credit

For the past few years, the U.S. government has already initiated a production tax credit program to provide a tax incentive promoting the renewable energy. For the electricity

generated by wind energy, \$0.021/KWh credits will be given. This might significantly alter our expansion planning decisions concluded before.

If we take account of the \$0.021/KWh credit for the electricity generation by the wind power, the assumption for the generation costs will be changed as shown in Figure 19.

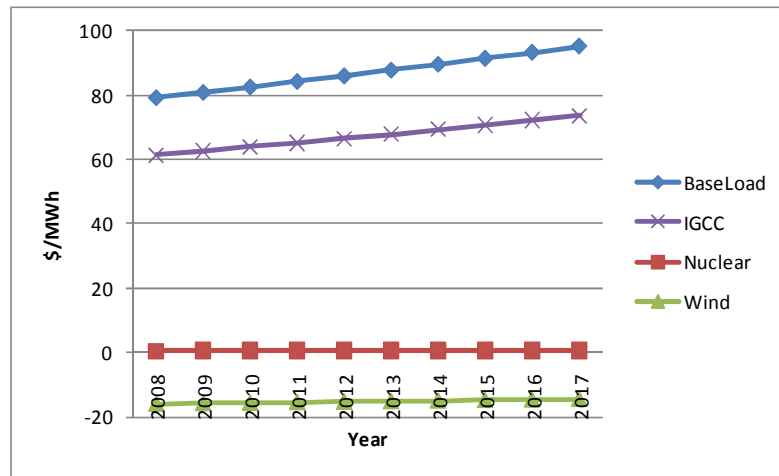


Figure 19. Generation cost for BaseLoad, IGCC, wind and nuclear power plants (PTC)

The generation cost for BaseLoad, IGCC and nuclear remain the same, while the generation cost for the wind power drops down to be even negative.

In the same way as in the section 5.4, a series of robust solutions is shown in Figure 20.

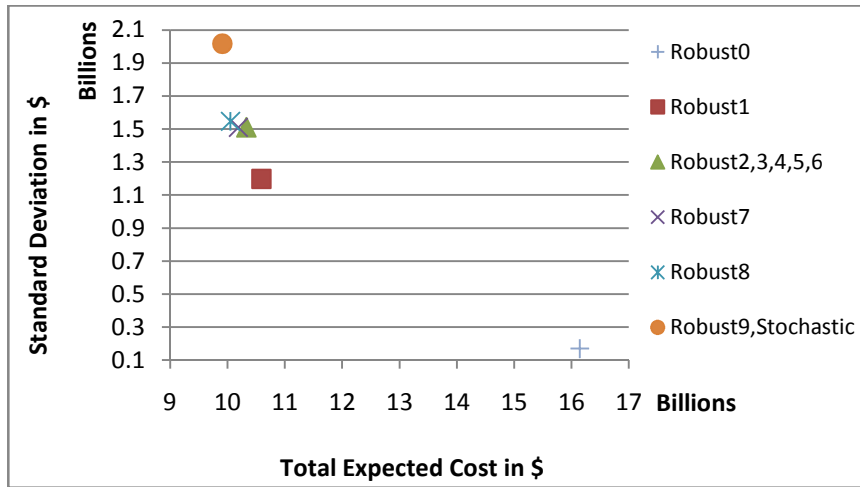


Figure 20. Standard deviation vs. expected cost for stochastic and robust optimization (PTC)

“Robust 2, 3, 4, 5, 6” is dominated by “Robust 7”. The costs over scenarios for all the solutions on the efficient frontier are shown in Figure 21.

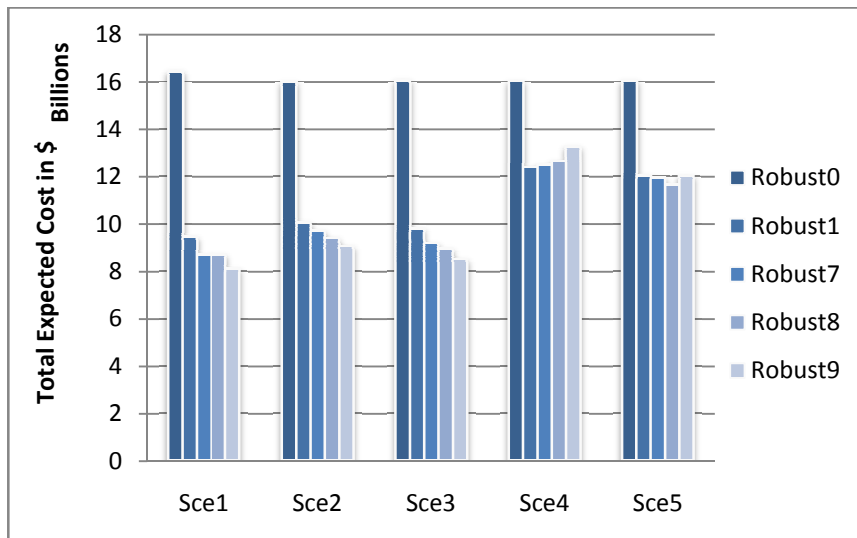


Figure 21. The cost over scenarios for five different robust solutions (PTC)

Except “Robust 0”, the other four solutions are non-dominated. The planner can further select one of them based on the preference. The cumulative expansion planning decisions for them are respectively shown in Figures 22-25.

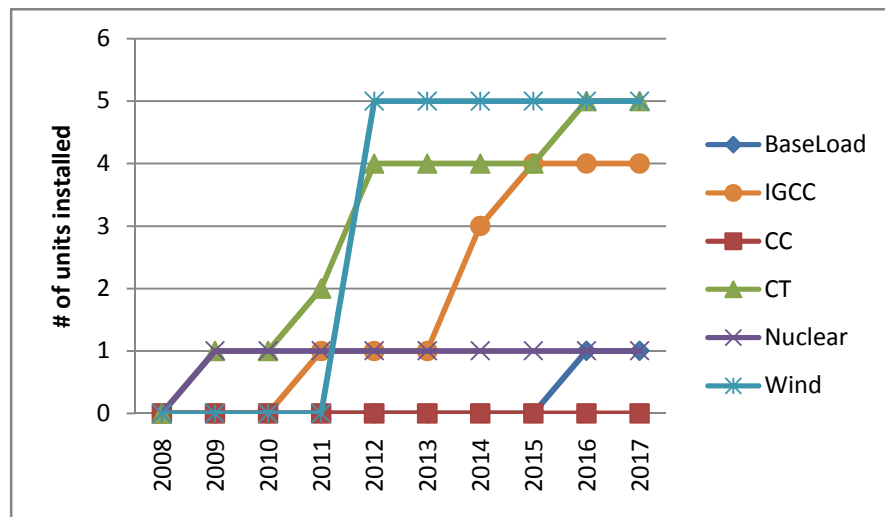


Figure 22. Cumulative expansion planning decision in “Robust 1” solution (PTC)

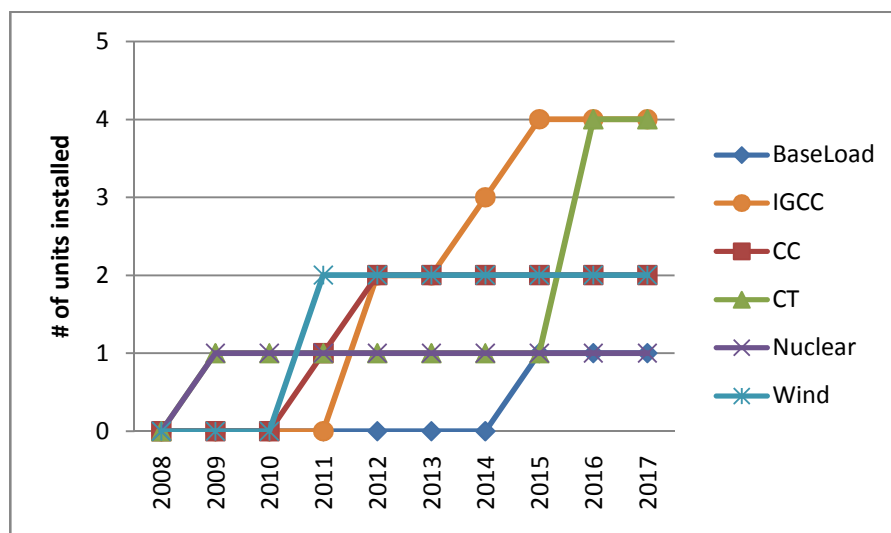


Figure 23. Cumulative expansion planning decision in “Robust 7” solution (PTC)

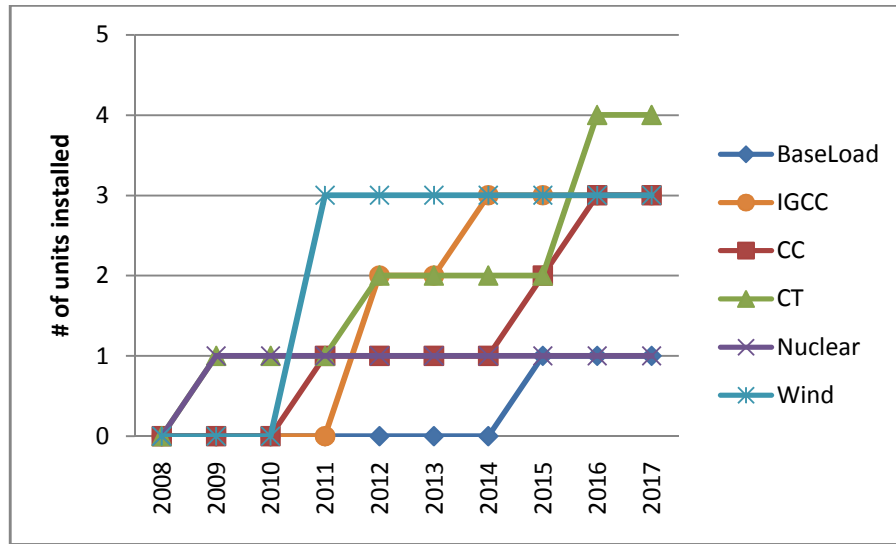


Figure 24. Cumulative expansion planning decision in “Robust 8” solution (PTC)

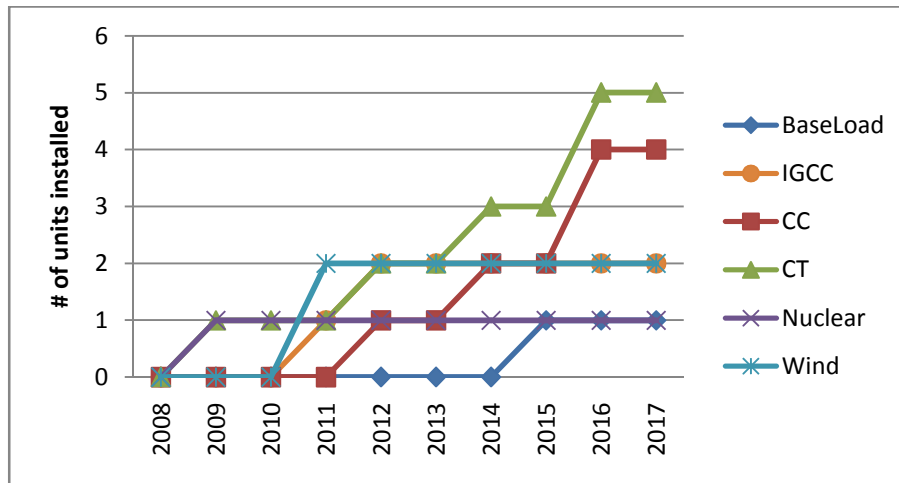


Figure 25. Cumulative expansion planning decision in “Robust 9” solution (PTC)

Compared with the expansion planning decisions without PTC, the robust solutions under the PTC program show much more emphasis on the renewable energy, specifically, the wind plants. All of the four candidate expansion planning decisions suggest making an

expansion on wind plants. As the p_v decreases, the installed units of wind plants decrease as well, the same as the coal-fired power plants, since they all have the same characteristics with the relatively expensive investment cost and low generation cost.

5.6 Sensitivity to Scenario Sampling

In this section, additional experiment results for another 10-year case study with 5 scenarios without the production tax credit are presented.

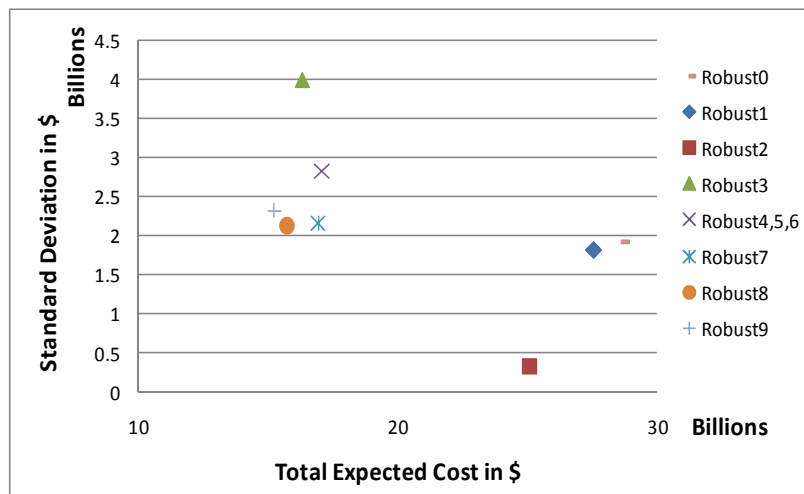


Figure 26. Standard deviation vs. expected cost for stochastic and robust optimization (sampling 2)

The tradeoff between standard deviation and expected cost for stochastic and all the robust solutions are shown in Figure 26. “Robust 2, 8, 9” are the solutions on the efficient frontier.

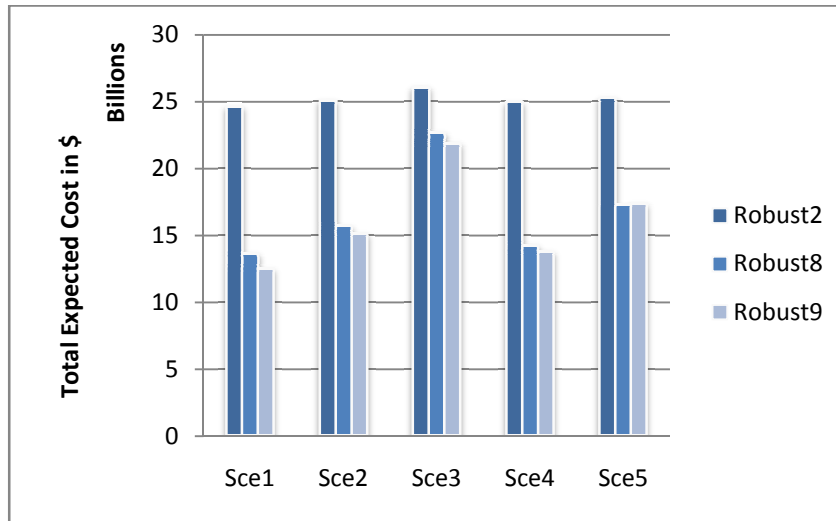


Figure 27. The cost over scenarios for three different robust solutions (sampling 2)

The costs over scenarios for three different robust solutions on the efficient frontier are compared in Figure 27. Since the “Robust 2” and “Robust 8” are dominated by “Robust 9”, we select the “Robust 9” as the best planning decision in this case. The cumulative expansion planning decisions for “Robust 9” is shown in Figure 28.

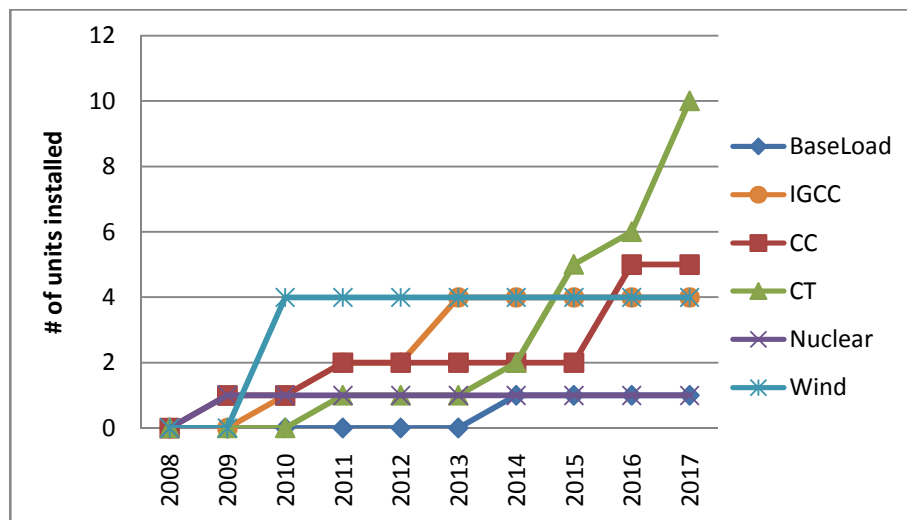


Figure 28. Cumulative expansion planning decision in “Robust 9” solution (sampling 2)

Compared with the experiment results from the previous 10-year case study with 5 scenarios, the selected planning decision in this case is very different, due to the selection of the different 5 scenarios from the previous one.

We expand more capacity in general in this case, and a few investments in wind plants are suggested as well. Since the different selection of the 5 scenarios has the great impact on the final expansion planning decisions, it suggested that 5 scenario over the 10 years is not enough for obtaining a truly robust solution.

With the number of scenarios increasing, the computational efficiency will be greatly affected. For instance, the robust optimization will be a 2106×2160 mixed quadratic integer programming problem for a 10 scenario case study. The computational efficiency also depends on the specific problem. With the same problem size, some of the robust models are solved within couple of minutes, some of them take a couple of hours, and some others are hardly solvable. More effort on the improvement of the computational efficiency is further required.

5.7 Deterministic Solutions

In this section, the deterministic solutions under each scenario of the first 10-year case study in Chapter 5 corresponding to the experiment results in Table 11 are presented.

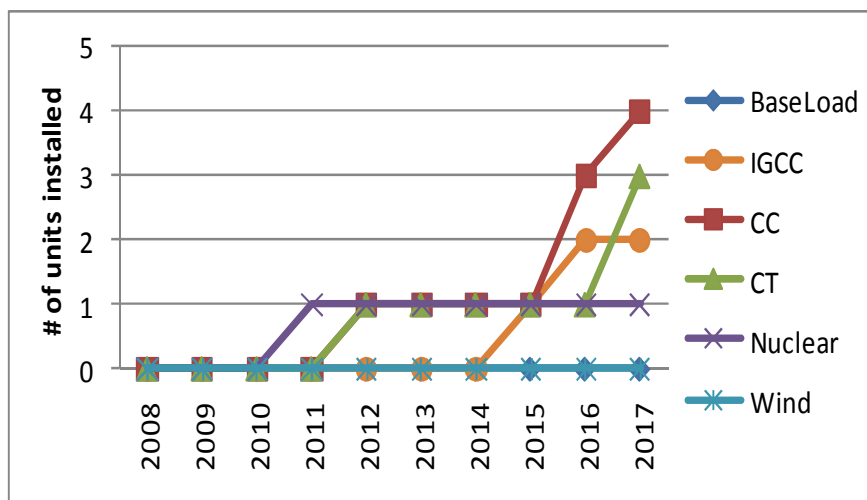


Figure 29. Cumulative expansion planning decision in “Sce 1” solution

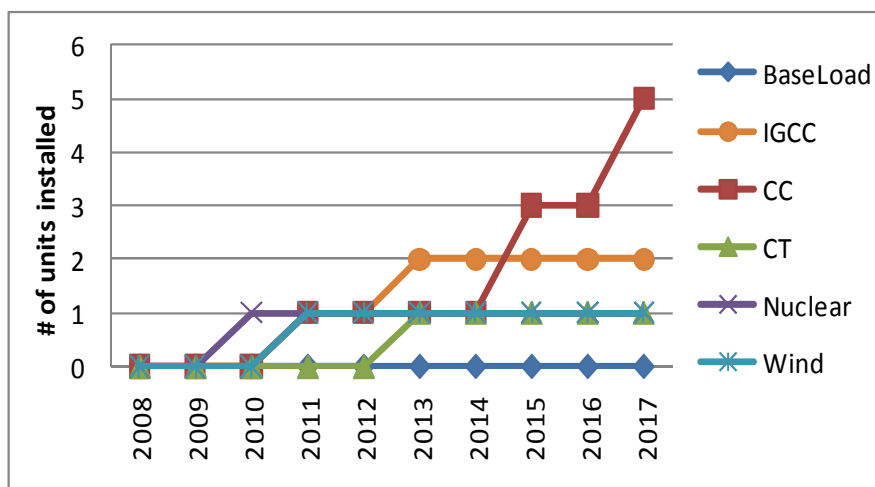


Figure 30. Cumulative expansion planning decision in “Sce 2” solution

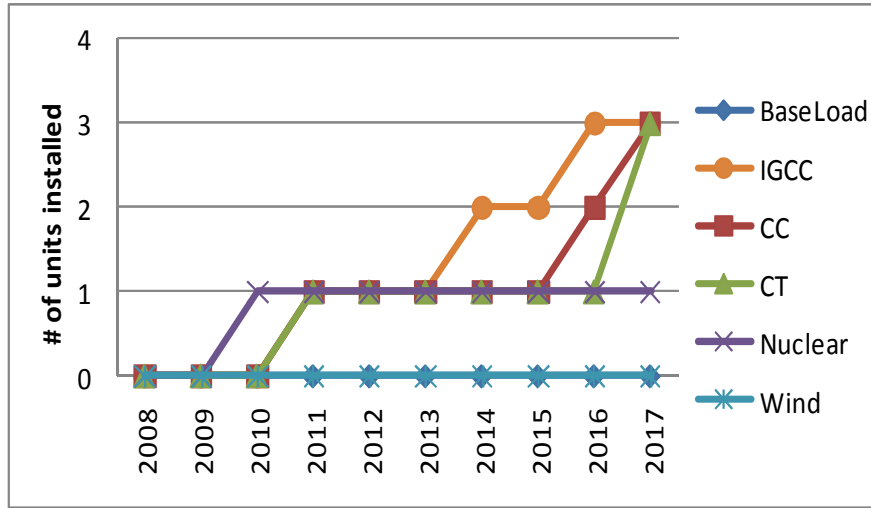


Figure 31. Cumulative expansion planning decision in "Scen 3" solution

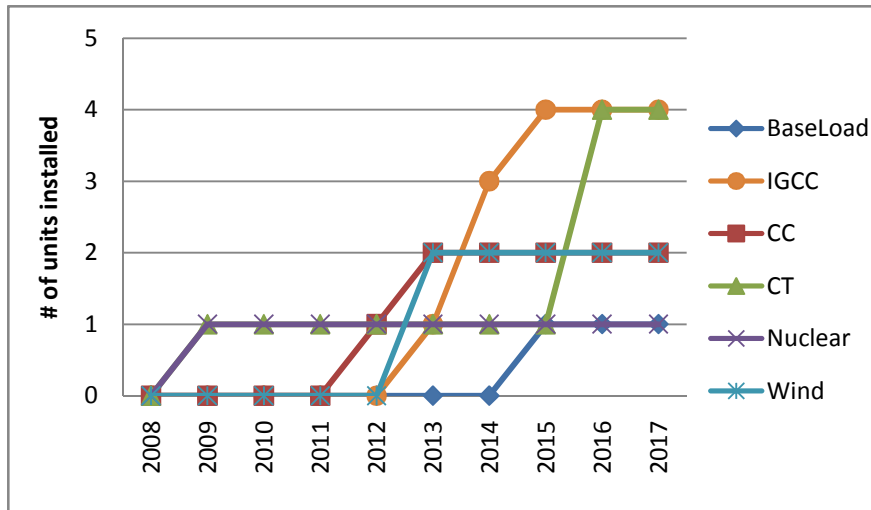


Figure 32. Cumulative expansion planning decision in "Scen 4" solution

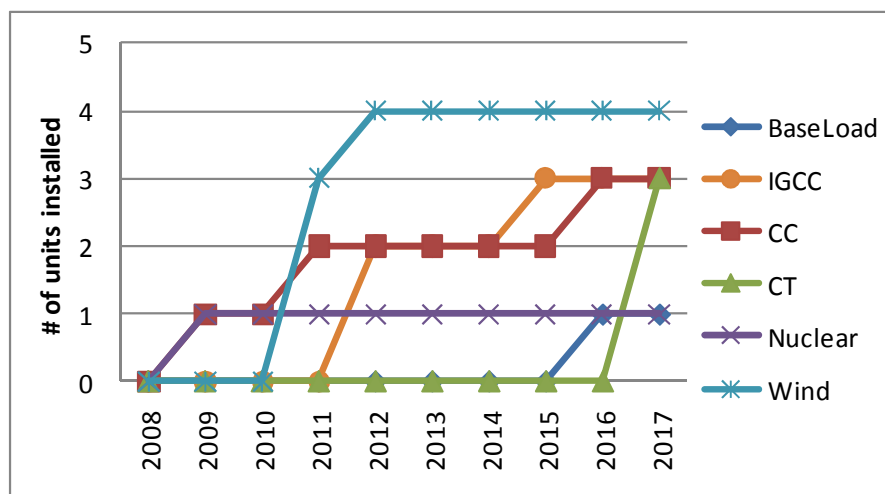


Figure 33. Cumulative expansion planning decision in “Scen 5” solution

The cumulative expansion planning decision of “Scen 1, 2, 3, 4 and 5” are respectively the optimal solutions for each of the scenarios. In the real world, without a stochastic model, the planner compares and analyzes each of the solutions based on each different scenario, and takes the planning decisions in common to come up with a robust solution.

Based on Figures 29-33, the investment decision on nuclear plant seems to be the most robust always with a new installed unit. The second robust decision is for BaseLoad plant, with either zero or one installed unit during the whole planning horizon. For the expansion decisions on the other four types of generators, they are less robust over the scenarios. One way to make these investment decisions could be based on their expected values. However, this does not take risk into account.

CHAPTER 6. SUMMARY AND FUTURE RESEARCH

6.1 Summary

In this thesis, two optimization formulations, two-stage stochastic programming and robust optimization, are applied to the power generation expansion system to help make the planning decision on how many units of which type of generator to be build in which year.

A multi-year case study in the Midwest region of Unite States has been conducted, and two major uncertainties, the electricity demand and the natural gas price, are assumed. They are modeled as two continuous time random variables following geometric Brownian motions. We further studied the statistical properties of the random variables in order to generate a scenario tree over years for the case study and applied naïve sampling to reduce the number of scenarios for the case study.

The experiment results were analyzed and compared with the deterministic methods to indicate the benefit of both the two-stage stochastic programming and robust optimization. Besides, an experiment for the sensitivity of the robust solutions to the robustness parameter p_v has been conducted. Criteria are suggested for selecting a best robust expansion planning solution considering minimization of both total expected cost and cost variance. In addition, we also analyzed the effect that the production tax credit would have on the expansion planning decisions.

6.2 Future Research

6.2.1 Assumptions and Constraints

In the case study, we assume the existing units for all the types of the generators are zero and annual demand as the incremental demand from year to year. In the future study, a more realistic model should be implemented based on the real demand in each year instead of the annual incremental demand, and incorporating the currently existing generator units.

For the maximum units to build for each generator, a more practical assumption should be made subject to the availability of the energy, the transmission capacity or financial budget.

The lead time for constructing the different types of power plants must be taken into account, as well as the life time for both the existing power plants and newly built power plants.

To minimize the total expected cost of the power generation expansion planning, cost other than building cost and generation cost will be included, such as the annual generation fixed cost.

For the model implementation, we will further investigate the following assumptions in the case study:

- How does the assumption on the number of the sub-periods in a year representing the actually hourly demand affect the planning decision?
- How does the number of scenarios in the experiment affect the planning decision?

Besides making a more appropriate assumption for the case study, more constraints will be considered. For instance, the loss of load probability (LOLP), generally assumed to be 0.1 days per year or less in practice, will be considered as a constraints to ensure the reliability of the power generation generated by the optimization models.

6.2.2 Uncertainties

Besides the demand and natural gas price, more uncertainties should be taken into account in the future research work.

Regarding the increasing concern on the carbon emission and global warming issue, the concept of green energy and sustaining economic development has gained more popularity. In the Midwest region, due to the abundance of the wind resource, wind farms have become more and more attractive for the power expansion investment. However, to integrate the wind energy into the power system, more uncertainties will get involved. Since the wind resource largely depends on the uncertain weather, there is a potential risk to rely on the wind generation. The capacity credit of the wind generation over time will be considered as a very significant uncertainty in the future research work.

To better encourage the generation expansion investment in renewable energies, government provides the financial support known as the production tax credits, which could greatly affect the planner's investment decisions towards the renewable power plants. Both the ongoing government policy and the future potential incentive for the renewable energy are other major uncertainties involved.

Besides the promotion of the renewable energy, limitation on the carbon emission is also another global concern. A potential carbon emission cost or a cap-and-trade system will

potentially be established in the near future. These are also the uncertainties we will focus on in the future research. In some generation expansion planning situations, the transmission capacity and congestion need to be considered as well.

6.2.3 Methodologies

In this thesis, the multi-period problem is solved by the two-stage stochastic programming. In the future study, we can further consider the multi-stage stochastic programming to enable more flexibility for the investment decisions in the later years. By comparing the optimization solutions of both of them, we can study the value of the multi-stage stochastic programming

For the robust optimization model in the thesis, we use the cost variance to mathematically measure the risk of the uncertainties in the future. In the future research, we can further study a most appropriate way to measure the risk. Different measurements and their mathematical characteristics for modeling the risk are summarized in [52]. These risk measurements includes a bad scenario, a worst-case analysis, expectation, standard deviation, specified probability quantiles, a value-at-risk or a conditional value-at-risk [52].

Besides, based on the experiment result of the case studies in the thesis, it ignored the effect of the generation cost of the power plants after the planning horizon. Thus, a more realistic model able to mitigate the end-of-study effect needs to be developed.

In addition to the development of a more appropriate model, further effort is also required for improving the computational efficiency. An easy-to-solve approximation scheme to the original models can be considered to help alleviate the computational burden. We can also apply the Benders decomposition to speed up the computational performance.

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, I thank Dr. Sarah Ryan for her guidance, patience and support throughout this research and the writing of this thesis. Her insights and words of encouragement have often inspired me to always pursue a higher academic level. I would also like to thank the electric power research center (EPRC) of Iowa State University for the sponsorship on my research, as well as my committee members for their efforts and contributions to this work: Dr. Jo Min and Dr. Mervyn G. Marasinghe. I would also like to thank Dr. James D. McCalley for his help and support on this research, and Marcus M. Edvall for technical support on Tomlab/Cplex.

BIBLIOGRAPHY

- [1] J. McCalley, “Introduction to Electric Systems Expansion Planning”, EE590 Course Note, <http://home.eng.iastate.edu/~jdm/ee590/PlanningIntro.pdf>, viewed in Sep 2008.
- [2] E.A. DeMeo, W. Grant, M.R. Milligan and M.J. Schuerger, “Wind plant generation”, *IEEE Power & Energy Magazine*, Vol. 3(6), pp. 38–46, 2005.
- [3] M.R. Milligan, “Variance estimates of wind plant capacity credit”, National Renewable Energy Laboratory, Colorado, 1996.
- [4] B.G. Gorenstin, N.M. Campodonico, J.P. Costa and M.V.F. Pereira, “Power system expansion planning under uncertainty”, *IEEE Transaction on Power System*, Vol. 8(1), pp. 129–136, 1993.
- [5] N.I. Voropai and E.Y. Ivanova, “Multi-criteria decision analysis techniques in electric power system expansion planning”, *Electrical Power and Energy System*, Vol. 24, pp. 71–78, 2002.
- [6] E. Denny and M. O’Malley, “Wind generation, power system operation, and emissions reduction”, *IEEE Transaction on Power Systems*, Vol. 21(1), pp. 341–347, 2006.
- [7] H. Lund, “Large-scale integration of wind power into different energy systems”, *Energy*, Vol. 30(13), pp. 2402–2412, 2005.
- [8] <http://www.eia.doe.gov/international/>, Energy Information Administration, Department of Energy, viewed in Jun 2009.
- [9] “International Energy Outlook 2009”, Energy Information Administration, Department of Energy, Mar 2009.

- [10] “Annual Energy Outlook 2009”, Energy Information Administration, Department of Energy, Mar 2009.
- [11] R.T. Rockafellar, “Coherent approaches to risk in optimization under uncertainty”, in Tutorial in Operation Research, INFORMS Annual Meeting, 2007.
- [12] http://www.eia.doe.gov/cneaf/electricity/epm/table1_1.html,
- [13] Energy Information Administration, Department of Energy, viewed in Jun 2009.
- [14] D. Dentcheva and W. Romisch, “Optimal power generation under uncertainty via stochastic programming”, in Stochastic Programming Methods and Technical Applications (K.Marti and P.Kall, Eds.), Lecture Notes in Economics and Mathematical Systems 458, pp. 22–56, 1998.
- [15] C.J. Andrews, “Evaluating risk management strategies in resource planning”, *IEEE Transactions on Power System*, Vol. 10(1), pp. 420–426, 1995.
- [16] http://tonto.eia.doe.gov/energy_in_brief/renewable_energy.cfm, viewed in Apr 2009.
- [17] R. Doherty, H. Outhred and M. O’Malley, “Establishing the role that wind generation may have in future generation portfolios”, *IEEE Transactions on Power Systems*, Vol. 21(3), pp. 1415–1422, 2006.
- [18] R. Karki and R. Billinton, “Cost-effective wind energy utilization for reliable power supply”, *IEEE Transactions on Energy Conversion*, Vol. 19(2), pp. 435–440, 2004.
- [19] J. Garcia-Gonzalez, R.M.R. de la Muela, L.M. Santos and A.M. Gonzalez, “Stochastic joint optimization of wind generation and pumped-storage units in an electricity market”, *IEEE Transactions on Power Systems*, Vol. 23(2), pp. 460–468, 2008.

- [20] K.R. Voorspools and W.D. D'haeseleer, "An analytical formula for the capacity credit of wind power", *Renewable Energy*, Vol. 31, pp. 45–54, 2006.
- [21] M.R. Milligan, "Measuring wind plant capacity value", National Renewable Energy Laboratory, Colorado, 1996.
- [22] M.R. Milligan, "Modeling utility-scale wind power plants. Part2: Capacity credit", *Wind Energy*, Vol. 3, pp. 106–206, 2000.
- [23] http://en.wikipedia.org/wiki/Clean_coal#Clean_coal_technology, viewed in Jun 2009.
- [24] S.H. Karaki, F.B. Chaaban, N. Al-Nakhl and K.A. Tarhini, "Power generation expansion planning with environmental consideration for Lebanon", *Electrical Power and Energy Systems*, Vol. 24, pp. 611–619, 2002.
- [25] http://en.wikipedia.org/wiki/Renewable_energy_development#Production_tax_credits, viewed in Jun 2009.
- [26] http://www.eia.doe.gov/cneaf/electricity/epm/table1_1.html, viewed in Jun 2009.
- [27] D. Holmes, "A Collection of Stochastic Programming Problem", Technical Report 94-11, Apr 1994.
- [28] S. Ahmed, Introduction to stochastic integer programming, <http://www.stoprolog.org/SPIntro/intro2sip.html>, viewed in Jan 2009.
- [29] V.M. Alborno, P. Benario and M.E. Rojas, "A two-stage stochastic integer programming model for a thermal power system expansion", *International Transactions in Operational Research*, Vol.11, pp. 243-257, 2004.

- [30] A. Laurent, “A Scenario Generation Algorithm for Multistage Stochastic Programming: Application for Asset Allocation Models with Derivatives”, Ph.D. Thesis in Economics, University of Lugano, Lugano, Switzerland, Mar 2006.
- [31] S. Takriti and S. Ahmed, “On robust optimization of two-stage systems”, *Mathematical Programming Series A* 99, pp. 106–126, 2004.
- [32] S. Malcolm and S. Zenios, “Robust optimization for power systems capacity expansion under uncertainty”, *Journal of the Operational Research Society*, Vol. 45(9), pp. 1040–1049, 1994.
- [33] J.M. Mulvey, R.J. Vanderber and S.A. Zenios, “Robust Optimization of Large-Scale Systems”, *Operations Research*, Vol. 43(2), pp. 264-281, Mar-Apr 1995.
- [34] A.S. Chuang, F. Wu and P. Varaiya, “A Game-Theoretic Model for Generation Expansion Planning: Problem Formulation and Numerical Comparisons”, *IEEE Transaction on Power System*, Vol. 16(4), Nov 2001.
- [35] J.L.C. Meza, M.B. Yildirim and A.S. M.Masud, “A Model for the Multiperiod Multiobjective Power Generation Expansion Problem”, *IEEE Transactions on Power Systems*, Vol. 22(2), May 2007.
- [36] S. Pokharel and K. Ponnambalam, “Investment planning for electricity generation expansion”, *International Journal of Energy Research*, Vol. 21, pp. 185–194, 1997.
- [37] E.V. Geert, “Increased uncertainty a new challenge for power system planners”, IEE Colloquium, 1998.
- [38] R.R. Booth, “Optimal generation planning considering uncertainty”, *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-91, pp. 70–77, 1972.

- [39] B. Mo, J. Hegge and I. Wangensteen, “Stochastic generation expansion planning by means of stochastic dynamic programming”, *IEEE Transaction on Power Systems*, Vol. 6(2), pp. 662–668, 1991.
- [40] N.V. Sahinidis, “Optimization under uncertainty: State-of-the-art and opportunities”, *Computers and Chemical Engineering*, Vol. 28, pp. 971-983, 2004.
- [41] S. Ahmed, A.J. King and G. Parija, “A Multi-Stage Stochastic Integer Programming Approach for Capacity Expansion under Uncertainty”, *Journal of Global Optimization*, Vol. 26, pp. 3-24, 2003.
- [42] Y. Fukuyama and H. Chiang, “A Parallel Genetic Algorithm for Generation Expansion Planning”, *IEEE Transaction on Power System*, Vol. 11(2), pp. 955-961, May 1996.
- [43] H.T. Firmo and L.F.L. Legey, “Generation expansion planning: An iterative genetic algorithm approach”, *IEEE Transactions on Power System*, Vol. 17(3), pp. 901–906, 2002.
- [44] S. Kanna, S.M.R. Slochanal and N.P. Padhy, “Application and Comparison of Metaheuristic Techniques to Generation Expansion Planning Problem”, *IEEE Transactions on Power Systems*, Vol. 20(1), Feb 2005.
- [45] S. Ahmed and N. Sahinidis, “An Approximation Scheme for Stochastic Integer Programs Arising in Capacity Expansion”, *Operations Research*, Vol. 51(3), pp. 461-471, May-Jun 2003.
- [46] J. Dupacova, G. Consigli and S.W. Wallace, “Scenarios for multistage stochastic programs”, *Annals of Operations Research*, Vol. 100, pp. 25-53, 2000.

- [47] N. Growe-Kuska, H. Heitsch and W. Romisch, “Scenario reduction and scenario tree construction for power management problems”, IEEE Bologna Power Tech Conference, Bologna, Italy, Jun 23th-26th, 2003.
- [48] K. Hoyland and S.W. Wallace, “Generating scenario tree for multistage decision problems”, *Management Science*, Vol. 47(2), pp. 295–307, 2001.
- [49] Electric generation expansion analysis system (EGEAS), Electric Power Research Institute, <http://mydocs.epri.com/docs/public/000000000001015475.pdf>, viewed in Jan 2009.
- [50] <http://www.gesys.com/pro-mod.htm>, viewed Jan 2009.
- [51] <http://www.plexos.info/wiki/>, viewed Jan 2009.
- [52] <http://www.plexossolutions.com/>, viewed Jan 2009.
- [53] R.T. Rockafellar, “Coherent approaches to risk in optimization under uncertainty”, *Inform*s, pp. 38-61, 2007.
- [54] Jointed Coordinated System Planning Report 2008, <http://www.jcspstudy.org/>, viewed in Feb 2009.
- [55] S.M. Ross, “Chapter 10: Brownian motion and stationary processes”, *Introduction to probability models*, 9th edition, Elsevier, 2007
- [56] http://en.wikipedia.org/wiki/Geometric_Brownian_motion, viewed in Jun 2009.
- [57] R.R. Marathe and S.M. Ryan, “On the validity of the geometric Brownian motion assumption”, *The Engineering Economist*, Vol. 50(2), pp. 159-192, Apr-Jun 2005.

- [58] N.L. Johnson and S. Kotz, “Chapter 3: Lognormal Distribution”, *Continuous Univariate Distributions Volume 1*, 2nd edition, Wiley Series in Probability and Mathematical Statistics, 1994
- [59] http://www.midwestmarket.org/home/Market%20Reports?type=rt_ex&list=month, Midwest ISO, viewed in Jun 2009.