2012

Electricity System Expansion Studies to Consider Uncertainties and Interactions in Restructured Markets

Shan Jin
Iowa State University

Electricity system expansion studies to consider uncertainties and interactions in restructured markets

by

Shan Jin

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Industrial Engineering

Program of Study Committee:
Sarah M. Ryan, Major Professor
Lizhi Wang
Jo Min
James McCalley
Leigh Tesfatsion

Iowa State University

Ames, Iowa

2012

DEDICATIONS

To my beloved grandparents, parents, husband, and daughter. Thank you all for your love and care, enlightenment and inspiration throughout my life. Thank you for always believing in me, encouraging me to pursue a better of myself, and supporting me at any time.

To all my dearest and life-time friends who shared every single moment of the happiness and sadness, always supported and encouraged me and my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to take this great opportunity to express my thanks to those who helped me with various aspects of both my academic life of conducting the research and the writing of this thesis and personal life of giving me constant care, support and encouragement for the last five years.

First and foremost, I sincerely thank Dr. Sarah Ryan, as my major professor and as my lifetime mentor, for her guidance, patience and support throughout my graduation study. Her rich insights, rigorous scholarship, and words of encouragement have inspired me to always pursue a higher academic level; her personal integrity, selfless spirit and unlimited support have always nourished my heart and guided me to become a better person. Without her constant encouragement and unlimited support for both my academic and personal lives, I would not have been able to complete this research and the writing of this thesis. I would like to thank my committee members for their efforts and contributions to this work: Dr. Lizhi Wang, Dr. Jo Min, Dr. James McCalley and Dr. Leigh Tesfatsion, from whom I gained enormous knowledge and insightful feedbacks. I am also greatly thankful for my co-authors, Jean-Paul Watson at Sandia National Laboratory and David Woodruff at University of California, Davis, for the computational experiments with Python, and investigation of the solution accuracy and robustness. I would also like to thank my mentor, Audun Botterud, and director, Guenter Conzelmann at Center for Energy, Environmental, and Economic Systems Analysis (CEEESA), Decision and Information Sciences (DIS) Division of Argonne National Laboratory during my internships for consistently encouraging me and supporting me on the research study, and opening my eye to the updated research topics.

I greatly acknowledge the Electric Power Research Center (EPRC) of Iowa State University, Power Systems Engineering Research Center (PSERC) and Center for Energy, Environmental, and Economic Systems Analysis (CEEESA), Decision and Information Sciences (DIS) Division of Argonne National Laboratory (ANL) for the sponsorship of my research.

ABSTRACT

This dissertation concerns power system expansion planning under different market mechanisms. The thesis follows a three paper format, in which each paper emphasizes a different perspective. The first paper investigates the impact of market uncertainties on a long term centralized generation expansion planning problem. The problem is modeled as a two-stage stochastic program with uncertain fuel prices and demands, which are represented as probabilistic scenario paths in a multi-period tree. Two measurements, expected cost (EC) and Conditional Value-at-Risk (CVaR), are used to minimize, respectively, the total expected cost among scenarios and the risk of incurring high costs in unfavorable scenarios. We sample paths from the scenario tree to reduce the problem scale and determine the sufficient number of scenarios by computing confidence intervals on the objective values. The second paper studies an integrated electricity supply system including generation, transmission and fuel transportation with a restructured wholesale electricity market. This integrated system expansion problem is modeled as a bi-level program in which a centralized system expansion decision is made in the upper level and the operational decisions of multiple market participants are made in the lower level. The difficulty of solving a bi-level programming problem to global optimality is discussed and three problem relaxations obtained by reformulation are explored. The third paper solves a more realistic market-based generation and transmission expansion problem. It focuses on interactions among a centralized transmission expansion decision and decentralized generation expansion decisions. It allows each generator to make its own strategic investment and operational decisions both in

response to a transmission expansion decision and in anticipation of a market price settled by an Independent System Operator (ISO) market clearing problem. The model poses a complicated tri-level structure including an equilibrium problem with equilibrium constraints (EPEC) sub-problem. A hybrid iterative algorithm is proposed to solve the problem efficiently and reliably.

CHAPTER 1 INTRODUCTION

1.1 Background

Electrical energy has become a more and more essential part of people's lives as well as a key concern of the whole world. The increasing use of electricity has rapidly improved our human society and standard of living. However, with the global economy more reliant on sustainable development of energy, a series of problems, such as energy shortage, electricity blackout and global warming are gaining attention. All these issues contribute to the importance of maintaining a reliable and efficient electricity energy supply system.

The complexity of decision making in power system expansion planning problems arises from the diversity of technologies available for generating power, important reliability constraints of sufficient energy supply, and wholesale market restructuring. Both investment planning and operational scheduling must be considered over multiple decades because of the scale of capital investment and long lives of generation and transmission assets. The problem is complicated by the multiplicity of organizations involved in the whole integrated electricity supply system including fuel transportation, generation, transmission and distribution facilities. And it also concerns the facilities siting, environmental impact and reliability to avoid potential electricity blackout [1, 2].

Power system expansion planning is also very complicated to formulate due to the large number of uncertainties involved. Load growth has always been a significant uncertainty in generation expansion planning. It can usually be estimated by forecasting climate, population expansion or movement, and technology development. 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 [3]. Natural gas price is usually considered to be a very important uncertainty in the generation expansion planning problem. It accounts for an increasing share of total generation because it is cleaner than coal and the units can be started and stopped quickly to compensate for fluctuations in renewable energy. Increasingly, gas units have become the marginal ones so their costs have a significant impact on wholesale electricity price. Recently, a newly available resources, shale gas, has become more attractive and the annual growth in production has averaged 17% from 2006 to 2011 [4]. Due to its low price and relatively low carbon emissions, it has now become a promising resource and its use is expected to grow rapidly in the future.

The traditional capacity expansion planning problem takes a centralized perspective due to the industry's previous vertically integrated structure of generation, transmission and distribution. The factors considered in the centralized planning decisions include the system load balance, reserve requirement, investment budget, and capacity limit constraints.

However, in the 1990s, power system deregulation in United States introduced privatization and competition in the electricity market and is expected to lower electricity prices. The previous centralized integrated system was then decomposed into decentralized generator companies (GENCO), transmission owners (TRANSCO), distribution companies (DISCO), load serving entities (LSE) and individual consumers [5]. To enhance the reliability of the power grid, the Independent System Operator (ISO) was established to monitor the grid, coordinate generation and transmission, settle the power trades among power sellers and buyers, and conduct the regional transmission expansion planning and resource adequacy studies. The deregulated markets introduced competition among multiple private participants in the electricity market, to which modeling from a centralized point of view no longer applies. New models such as game theory, equilibrium program and multi-level programs are proposed to characterize the strategic behaviors and interactions among multiple players.

To provide a reliable and efficient electricity supply network, we must consider not only generation expansion to make sure that we have sufficient energy to meet future loads, but also an entire integrated electricity supply system including transmission, fuel transportation, market settlement by ISO, and LSE. Limited fuel transportation capacity results in energy shortage, higher production cost, and less generation expansion; while the layout and expansion of the transmission network greatly affects the system efficiency in terms of transmission congestion, and wind/load/reserve curtailment, The ISO is significantly important to maintain the reliable and efficient operations. LSEs buy the bulk power from the electricity wholesale market and provide services to the end-use consumers.

1.2 Problem Statement

This dissertation addresses electricity system expansion planning problems including generation expansion, transmission expansion and fuel transportation expansion. The traditional expansion problem is a cost minimization problem for a vertically-integrated entity in a regulated environment, while market based planning in a deregulated competitive market environment is more complicated because interactions from multiple market participants must be taken into account [6]. This dissertation covers both of these two major mechanisms and examines the problem from different perspectives.

The dissertation begins with a traditional generation expansion planning problem that assumes a central regulator who aims for a system-wide cost minimization decision on how many units of what type of power plants to build in which year and how much electricity is generated by each type of the power plants. In the long term planning horizon, both the investment and operational decisions are largely affected by future uncertainties. The study addresses the following issues:

1) The methodologies to model and integrate the multi-period evolution of multiple uncertain factors in a long term generation expansion planning (GEP) problem;
2) The methodologies to deal with the large scale of uncertainty in a GEP problem to achieve a tradeoff between computational complexity and solution accuracy;
3) The risk-based cost minimization model's impact on the optimal generation mix compared to the expected cost minimization model.

Besides consideration of investment and operational cost minimization, in restructured markets, the profit of expansion decision must also be taken into account to justify the expansion decision. The profit return received by an investor is determined by an electricity market price settlement. The independent system operator (ISO) matches the electricity supply bid and demand offer and settles the LMPs to maximize total market surplus of both buyers and sellers. Moreover, while investing more generation capacity, the transmission adequacy should be guaranteed to deliver the power, the fuel transportation should be sufficient to supply the generation, and the system net surplus should be increasing. Therefore, research on an integrated electricity supply system is intended to address the following:

1) The methodologies to address a mixed integer bi-level program with multiple followers that result from modeling an system net surplus maximization problem in which the expansion decisions for generation, transmission and fuel transportation are made from a centralized point of view in anticipation of the followers' operational decisions in a competitive electricity market;
2) The impact of the anticipation of the strategic GENCOs' operational decisions on the capacity expansion decisions; the system surplus, decomposed as buyer surplus, seller surplus and transmission rent; and the electricity prices.

Finally, due to the great impact of transmission expansion on individual generation expansion and operational decisions in the electricity markets, a more realistic model is investigated to account for the interaction among transmission and generation expansion with a competitive market. This work considers:

1) The development of a market based transmission and generation model that can capture the strategic behaviors of GENCOs making both generation expansion and operational decision reacting to the transmission decisions and anticipating market settlement results by an ISO;
2) The methodology to approach a solution of an equilibrium problem with equilibrium constraints (EPEC) comprised of bi-level games among GENCOs' capacity expansion problems;

3) The methodology to solve a tri-level mixed integer problem with an EPEC sub-problem resulting from a market based transmission and generation model;
4) The transmission expansion's impact on strategic GENCOs market power, system net surplus, transmission congestion, and electricity prices.

1.3 Thesis Structure

The dissertation consists of three papers. The first paper, published in *Energy Systems* [7], is reproduced in Chapter 2. The second paper, published by *IEEE Transactions on Power Systems* [8], is presented in Chapter 3 with an added appendix on incorporating carbon emission regulations. The third paper, submitted to *IEEE Transactions on Power Systems*, includes two parts with part I in Chapter 4, and part II in Chapter 5. A general conclusion in Chapter 6 summarizes the thesis.

References

[1] S.M. Ryan, J.D. McCalley, and D.L. Woodruff, "Long term resource planning for electric power systems under uncertainty", Technical Report, Mar. 2010.

[2] J. McCalley, "Introduction to electric systems expansion planning", http://home.eng.iastate.edu/~jdm/ee590/PlanningIntro.pdf, Sep. 2008.

[3] "Annual energy outlook 2009", Technical Report, Energy Information Administration, Department of Energy, Mar. 2009.

[4] "Annual energy outlook 2011", Technical Report, Energy Information Administration, Department of Energy, Apr. 2011.

[5] P.L. Joskow, "Markets for power in the United States: An interim assessment", Technical Report, Aug. 2005.

[6] R. Green, "Competition in generation: the economic foundations", *Proceedings of the IEEE,* Vol. 88, No. 2, pp.128–139, Feb. 2000.

[7] S. Jin, S.M. Ryan, J. Watson and D.L. Woodruff, "Modeling and Solving a Large-Scale Generation Expansion Planning Problem Under Uncertainty", Energy Systems, Vol. 2, No. 3-4, pp. 209-242, 2011

[8] S. Jin, S.M. Ryan, "Capacity Expansion in the Integrated Supply Network for an Electricity Market", IEEE Transactions on Power Systems, Vol. 26, No. 4, pp. 2275-2284, 2011

CHAPTER 2 MODELING AND SOLVING A LARGE-SCALE GENERATION EXPANSION PROBLEM UNDER UNCERTAINTY

Published in *Energy Systems*

Shan Jin, Sarah M. Ryan, Jean-Paul Watson and David L. Woodruff

Abstract

We formulate a generation expansion planning problem to determine the type and quantity of power plants to be constructed over each year of an extended planning horizon, considering uncertainty regarding future demand and fuel prices. Our model is expressed as a two-stage stochastic mixed-integer program, which we use to compute solutions independently minimizing the expected cost and the Conditional Value-at-Risk; i.e., the risk of significantly larger-than-expected operational costs. We introduce stochastic process models to capture demand and fuel price uncertainty, which are in turn used to generate trees that accurately represent the uncertainty space. Using a realistic problem instance based on the Midwest US, we explore two fundamental, unexplored issues that arise when solving any stochastic generation expansion model. First, we introduce and discuss the use of an algorithm for computing confidence intervals on obtained solution costs, to account for the fact that a finite sample of scenarios was used to obtain a particular solution. Second, we analyze the nature of solutions obtained under different parameterizations of this method, to assess whether the recommended solutions themselves are invariant to changes in costs. The issues are critical for decision makers who seek truly robust recommendations for generation expansion planning.

2.1 Introduction

Generation expansion planning is the problem of determining the type, quantity, and timing of power plant construction to build in order to meet increasing demand for electricity over an extended time horizon. Over the last two decades, the magnitudes and types of uncertainties confronting system planners have increased with the growth of policies to encourage generation from renewable sources, the possibility of regulation to control carbon emissions, and volatility in the prices of fossil fuels, particularly natural gas. Consequently, explicit consideration of uncertainty is required to mitigate risk in generation expansion planning models.

Extended, long-term time horizons are an additional integral component in generation expansion models, for the following reasons [47, 61]:

- Initial capital investment is expensive and the lifetime of a power plant normally ranges from 25 to 60 years. Therefore, a long term perspective is necessary to accurately evaluate alternative build schedules.
- Multiple organizations must be involved in the planning process, as the addition of new power plants typically imposes additional capacity requirements on transmission and distribution facilities. Consequently, organizations must coordinate their activities, which occur over extended time horizons.
- Long lead times may be required to obtain regulatory approval for plant construction, acquire land on which to build plants, negotiate fuel procurement contracts, and build up the required infrastructure. These considerations also imply that there is limited

maneuverability for investment decisions as new information becomes known or underlying conditions change.

Generation planning decisions must also account for operational impacts. Because the demand for electricity varies with diurnal, weekly, and seasonal patterns, different combinations of generating units are most cost-effective at different times, depending on fuel prices, availability of intermittent energy sources, and equipment outages.

Uncertainty regarding future operational conditions arises from several sources. Load (demand) growth has always been a significant source of uncertainty in generation expansion planning. Historically, it is estimated through combinations of climate forecasts, population expansion or movement models, projected economic conditions, and technology evolution. The world-wide annualized growth rate in electricity demand increased from 2% in 1990 to 4% in 2007 [25], and is projected to grow at an annual rate of 2% until 2030 [35]. Growth in electricity demand in the US gradually slowed from 9% per year in the 1950s to less than 2.5% per year in the 1990s. Recently, from 2000 to 2007, the average US growth rate in demand dropped to 1.1% per year. The slowdown in growth is projected to continue until the year 2030 [7]. In contrast, China—currently one of the fastest-growing economies—has experienced an average 14% annualized growth rate in electricity demand over the past five years [25].

The introduction of new generation technologies is also becoming important, as environmentally friendly renewable energy is receiving increased public support. The US government is considering greatly increasing the percentage of wind energy, to 20% of total electricity generation by 2030 [7], compared to 9% in 2008 [26]. Some state mandates specify a proportion of capacity rather than generation. For example, Iowa plans to increase the Renewable Portfolio Standard (RPS) to 20% of generation capacity by 2020, compared with 7% in 2007. Most of this increase will be obtained from wind energy due to the abundance of wind resources in the Midwest US. However, integration of wind generation into the electrical grid is problematic due to output uncertainty, caused by weather (wind speed) dependence [12, 13, 15, 21, 40, 42]. Hence, instead of a “capacity factor” (an average output over a year), the concept of a “capacity credit” is introduced as a measure of generation potential. The capacity credit captures the output of wind generation in the worst case; i.e., the minimal amount of power that the grid can rely upon at any given point in time. This quantity can be estimated using various methods [50–52, 67]. In addition to wind generation, clean coal [27], solar, new nuclear, and bio-based technologies should also be considered in generation expansion planning.

Environmental concerns, including emission penalties/constraints and other sources of regulatory uncertainty, also have a large potential influence on the cost effectiveness of investments in different types of power plants [39]. In particular, potential policies to limit or reduce greenhouse gas emissions would have a significant impact on generation planning. For the past several decades, tax incentives have yielded increased growth in renewable generation sources. The renewable electricity production tax credit (PTC) [28] was established as an incentive to promote renewable energy alternatives, and has significantly affected the growth of wind generation sources over the past 10 years [7]. It is likely that the PTC program will be extended in the near future.

Prices and availabilities of fuels, particularly for coal and natural gas, contribute additional sources of uncertainty in generation expansion planning. The price of coal is considered generally stable, with an average annualized growth rate of 2%. Natural gas prices fluctuate in a more unpredictable way [7], mainly depending on economic and technology development growth rates. The proportion of electricity generated by natural gas in the US in 2008 was approximately 21% [29]. Because natural gas is typically the most expensive fuel, power plants using natural gas are considered peak load generation units. Due to its large degree of uncertainty, consideration of natural gas price is critical in generation expansion planning.

Generation expansion planning involves two primary costs: investment costs and generation costs. Investment costs are dictated by decisions specifying how many units of each power plant type to build in each year of the planning horizon. Operational costs depend on the quantity of electricity generated by each plant in each year of the planning horizon, and fuel costs associated with such generation. To mitigate costs, investment decisions must take into account future uncertainties, which in turn affect operational costs. At the same time, investment decisions must satisfy additional requirements, including satisfaction of electricity demand, power generation reliability, energy resource limitations, financial budgets, maximum carbon emissions, and the minimum required electricity generation proportion for renewable energy.

In this paper, we formulate a model for the long term generation expansion problem, with the goal of deciding how many units of each type of power plant to construct, for each year in the planning horizon. The optimization objective in our model is to minimize the sum of initial investment costs and subsequent generation costs, while taking into account future uncertainty in both electricity demand and natural gas price. From a modeling perspective, we focus on support for the following issues:

– Identifying an appropriate model for the evolution of demand and fuel price uncertainty over time.
– Constructing a set of scenarios that adequately represent the evolution of uncertainty.
– Developing a test problem instance, with realistic parameter values, for the US Midwest.
– Specifying a reasonable level of risk aversion in view of the trade-off between reducing expected cost and reducing risk.

We formulate our model as a two-stage stochastic mixed-integer program, considering both minimization of the expected cost and the Conditional Value-at-Risk (CVaR). Given the computational difficulty of our model, we investigate methods for quantifying confidence intervals associated with the computed costs, using a limited number of scenarios to capture the space of future outcomes. Finally, we investigate the similarity of solutions obtained with different sets of scenarios, in order to assess the practical impact of scenario reduction on the actual decisions recommended by the model. These two computational issues are largely ignored in the literature on stochastic optimization and energy planning, yet are crucial when presenting potential solutions to decision-makers.

The remainder of the paper is organized as follows. We begin in Section 2.2 with a brief review of state-of-the-art methodologies for solving the generation expansion planning problem. In Section 2.3, we present a two-stage stochastic programming model for our generation expansion planning problem, accounting for both expected cost and Conditional

Value-at-Risk. In Section 2.4, we further discuss how to realize the computational implementation, including modeling assumptions, fitting of random variables' continuous time distributions, and generation of discrete scenarios for the case study. Section 2.5 details a case study based on the Midwest US electric power system. A procedure for computing confidence intervals on solution costs is introduced in Section 2.6; experimental results regarding both confidence intervals and solution similarity across different sets of scenarios are presented in Section 2.7. Finally, we conclude in Section 2.8.

2.2 Background

Both general capacity expansion planning problems and power-specific generation expansion problems have been studied for decades, yielding a range of different optimization models and algorithmic techniques for solving these models. We now briefly survey this literature, in order to place our research in the broader context.

Stochastic programming has been used frequently to address uncertainties in general capacity expansion problems [1, 4, 14, 55]. Robust optimization has also been studied in the context of general capacity expansion, to reduce cost variance over the range of possible scenarios [45, 54, 66]. Ahmed, King and Parija [3] describe a multistage stochastic programming model for capacity expansion, introduce a reformulation technique to reduce computational difficulty, and analyze various heuristic methods for solving large problem instances. Ahmed and Sahinidis [2] introduce a fast approximation scheme based on linear programming to solve a multi-stage stochastic integer programming model of a capacity expansion planning problem. State-of-the-art optimization methods under uncertainty are reviewed in [62], and include stochastic programming, robust optimization, probabilistic (chance-constrained) programming, fuzzy programming, and stochastic dynamic programming.

Independent of the specific model formulation, capacity expansion planning problems can pose significant computational challenges, due to the number of scenarios used to model the uncertainty, the number of decision stages in the planning horizon, the scale of the system under consideration, and the presence of integer decision variables. Thus, significant research has been devoted to the development of decomposition techniques to solve these problems more efficiently, and heuristics for obtaining high-quality approximate solutions in tractable run-times.

Laurent [41] summarizes different methodologies for scenario discretization. Scenario tree is constructed to model and evaluate the risk of the stochastic uncertainty of electricity portfolios [17]. Several techniques for constructing multi-stage scenario tree are presented in [16]. A scenario construction algorithm successively reducing the tree structure by bundling similar scenarios is introduced in [22]. Høyland and Wallace propose a generalized method applied to both single-stage scenarios and multi-stage scenarios [24]. The latter method is applied in this paper to generate the scenarios for the multi-year case study.

The state-of-the-art in generation (as well as transmission) expansion planning, in addition to an overview of optimization under uncertainty, is described in [61]. Here we mention only a few highlights. A collection of stochastic programs is discussed in [23]; one of the applications involves electricity generation capacity expansion with uncertainty arising from different modes of demand. A game-theoretic model to solve the generation expansion problem in a competitive environment is described in [9], motivated by the desire to analyze differences in the solutions relative to a centralized expansion plan. Multi-objective optimization has been applied to the generation expansion problem, in order to balance minimization of cost, environmental impact, imported fuel, and fuel price risk [49, 68]. Dynamic programming has also been used to solve the generation expansion problem [8, 39, 53].

Parallel genetic algorithms have been introduced to solve a deterministic power generation expansion problem [20]; Firmo and Legey [19] also use a genetic algorithm to yield approximate solutions to a related problem. A comparison of metaheuristic techniques for solving the generation expansion planning problem is described in [38].

In the electric power industry, some commercial packages for generation expansion planning are available, including EGEAS [18], ProMod [30], and Plexos [31, 32]. Most of these packages are based on deterministic models, although Plexos also offers support for two-stage stochastic programming. These packages 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.

We chose to formulate a two-stage stochastic programming model to represent our generation expansion planning problem for three main reasons. First, the decisions can be segmented naturally into discrete investment decisions that must be adopted before uncertain quantities are realized and continuous operational variables that can include recourse to demand and cost realizations. Second, historical data are available for fitting models for the evolution of the uncertain variables. Third, the risk of unacceptably high cost can be controlled in a tractable way by including linear constraints to compute Conditional Value-at-Risk.

2.3 A Two-stage Stochastic Mixed-integer Program for Generation Expansion Planning

We now describe a two-stage mixed-integer stochastic programming model to represent our generation expansion planning problem. We begin in Section 2.3.1 with a discussion concerning high-level assumptions underlying our optimization model. Section 2.3.2 introduces notation for model sets and associated indices and parameters, while Section 2.3.3 details the decision variables. Model constraints are described in Section 2.3.4. Finally, expected cost and Conditional Value-at-Risk minimization objectives are described respectively in Sections 2.3.5 and 2.3.6.

2.3.1 Modeling Assumptions

We assume that the optimization objective is to minimize some function of the combined investment and generation costs for pre-existing and newly constructed power plants, over the entire planning horizon. Additionally, because power outages can be both costly and disruptive, we impose monetary penalties for unmet demand. For example, outages might result in direct economic impact due to damage incurred by the electricity infrastructure, loss of data or breakdown of an assembly line, loss of life due to hospital service disruptions, and failure of public services. Finally, the model constraints enforce the following requirements, which are essential in generation expansion models: (1) because electricity currently cannot be stored economically, we require the generated electricity to meet the demand in each sub-period; (2) the load for each type of generator must be less than its planned capacity; and (3) the total number plants of each type built over the planning horizon must be less than a predefined maximum imposed due to budget, resource, or regulatory limitations.

2.3.2 Notation: Sets, Indices, and Parameters

Our generation expansion optimization model is expressed in terms of the following sets, described in conjunction with the corresponding index notation:

- $g \in G$: Types of generators.
- $y \in Y$: Years in planning horizon.
- $t \in T$: Load duration curve sub-periods.
- T_y: Set of sub-periods t in year y.
- Y_t : Year y to which sub-period t belongs.
- $\omega \in \Omega$: Scenario paths representing parameter uncertainties.

Model parameters common to all scenarios are given as follows:

- c_g: Cost per MW capacity to build a generator of type g, discounted to the beginning of the construction period. Units are $/MW.
- m_g^{max}: Maximum output capacity of installed generators of type g. Units are MW.
- h_t: Number of hours in sub-period t.

- n_g^{max}: Maximum output rating of generators of type g per hour. Units are MW.
- u_g^{max}: Maximum number of generators of type g that can be constructed over the planning horizon.
- u_g: Existing number of generators of type g at the beginning of the planning horizon.
- p_u: Penalty cost for unserved energy. Units are $/MWh.
- r: Annual interest rate, for cost discounting purposes.

As discussed in Section 2.4, uncertainty regarding future demand and fuel prices is captured by discrete scenarios. The following parameters are defined for each scenario$\omega \in \Omega$:

- $l_{gt\omega}$: Generation cost per MW hour for generators of type g in sub-period t, for scenarioω. Units are $/MWh.
- $d_{t\omega}$: Demand per hour in sub-period t for scenario ω. Units are MW.
- π_ω: Probability that scenario ω is realized; $\sum_{\omega \in \Omega} \pi_\omega = 1$

2.3.3 Decision Variables

Decision variables in our optimization model are partitioned into investment and operations categories, as follows:

- $U_{gy} \in Z^+$: (Investment) Number of generators of type g to be built in year y.
- $L_{gt\omega} \geq 0$: (Operations) The power generated by generators of type g per hour in sub-period t for scenario scenario ω. Units are MW.
- $E_{t\omega} \geq 0$: (Operations) The unserved load per hour in sub-period t for scenario scenarioω. Units are MW.

This partitioning of decision variables corresponds to that of a two-stage stochastic (mixed-integer) program [64]. The first-stage variables correspond to investment decisions over the planning horizon, and are determined prior to resolution of any uncertainty. The second-stage variables correspond to operational decisions, and are scenario-dependent because their evaluation is delayed until it is clear which specific scenario has been realized. In contrast to a multi-stage stochastic programming formulation [3], there is no recourse associated with the investment plan; all decisions are made up-front, and cannot be modified as uncertainty about the future is resolved.

A deterministic mixed-integer programming 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. The parameters in such a formulation can be taken as the planner's best guess of the outcomes of uncertain quantities, or computed as the expectation over the range of anticipated future outcomes.

2.3.4 Constraints

As indicated in Section 2.3.2, we impose limits on the total number of units for each generator type built over the planning horizon. This requirement is expressed as follows, constraining the cumulative number of units built during the planning horizon:

$$\sum_y U_{gy} \leq u_g^{max} \;\; \forall g \in G \qquad (2\text{-}1)$$

For each scenario, we impose energy balance constraints to enforce equality between the demand and the sum of electricity generated and unmet demand:

$$\sum_g L_{gt\omega} + E_{t\omega} = d_{t\omega} \; \forall t \in T, \omega \in \Omega \qquad (2\text{-}2)$$

Finally, we bound the output for each type of generator by the aggregate output rating of both existing and newly constructed units through each year in the planning horizon:

$$L_{gt\omega} \leq n_g^{max}\left(u_g + \sum_{y \leq Y_t} U_{gy}\right) \forall g \in G, t \in T, \omega \in \Omega \tag{2-3}$$

The numbers of decision variables and constraints in the deterministic (single scenario) mixed-integer program are then respectively |T | + |T ||G| + |G||Y| and |T | + |T ||G| + |G|. For the two-stage mixed-integer stochastic program, the number of decision variables and constraints are |T ||Ω| + |T ||Ω||G| + |G||Y| and |T ||Ω| + |T ||Ω||G| + |G|, respectively. Of the decision variables, |G||Y| are constrained to take integer values.

2.3.5 Minimization of Expected Cost

The most widespread optimization objective in two-stage stochastic programming is minimization of the expected cost; i.e., the sum of the first stage cost and the expected second stage costs. We denote our generation expansion optimization model with the expected cost minimization objective, subject to the constraints defined in Section 2.3.4, as GEP-EC. Formally, this objective is defined as follows:

$$\min_{U_{gy}, L_{gt\omega}, E_{t\omega}} \frac{\sum_{y \in Y} \sum_{g \in G} c_g m_g^{max} U_{gy}}{(1+r)^{y-1}} + \sum_{\omega \in \Omega} \pi_\omega \xi_\omega \tag{2-4}$$

where the per-scenario operational costs ξ_ω are defined as:

$$\xi_\omega = \frac{\sum_{y \in Y} \sum_{t \in T_y} \left(\sum_{g \in G} h_t l_{gt\omega} L_{gt\omega} + p_u h_t E_{t\omega}\right)}{(1+r)^{y-1}} \ \forall \omega \in \Omega \tag{2-5}$$

This cost formulation neglects operational costs beyond the study period, which could introduce end-of-study distortions in the first-stage decisions. To avoid those distortions, estimates of the discounted remaining operational costs could be added in the final period for each scenario. We did not attempt this in the case study because, in practice, the expansion planning problem is typically solved repeatedly over a rolling horizon and only the investment decisions for early years, which are less prone to distortion, are implemented.

The formulation *GEP-EC* is known as the extensive form of the stochastic program, in which the variables and constraints for all scenarios are explicitly represented in a single, large mathematical program. The extensive form can in principle be solved directly via commercial solvers, which is the approach we take in generating the results described in Section 2.7, because we want to make comparisons between provably optimal solutions. However, the difficulty of these problems can be considerable, especially in the presence of discrete decision variables such as the number of generators built. Consequently, researchers have developed decomposition techniques to accelerate the solution times for large-scale stochastic programs [59, 65].

2.3.6 Minimization of Conditional Value-at-Risk

The standard two-stage stochastic programming model does not take into account the potentially significant risk that the cost of one or even many scenarios far exceeds the expected cost. Various metrics have been introduced to formally quantify risk, including worst-case cost, cost variance, and the cost of a specific quantile (Value-at-Risk or VaR). Alternatively, one can focus on the cost expectation of the most costly l fraction of scenarios; i.e., the tail-conditional expectation. An easily-computed representation of the tail-conditional expectation has recently attracted significant attention in the risk analysis community. This metric is known as Conditional Value-at-Risk, or CVaR [57, 58]. CVaR is parameterized by l, $0 \leq l \leq 1$, which represents the fraction of high-cost scenarios that are to be considered by the metric. CVaR has a number of mathematically appealing properties (specifically, relative to VaR), and is particularly useful in optimization contexts because it can be expressed and minimized as a simple variant (as we discuss below) of a two-stage stochastic program [63].

We denote the generation expansion optimization model with CVaR minimization as the objective, subject to the constraints defined in Section 2.3.4, as GEP-CVAR. Formally, the CVaR optimization objective is defined as follows:

$$\min_{U_{gy}, L_{gt\omega}, E_{t\omega}, \eta, \delta_\omega} \eta + \frac{\sum_\omega \pi_\omega \delta_\omega}{l} \tag{2-6}$$

where η is an additional first-stage decision variable representing the 1-l-quantile of total investment and operational cost over all scenarios and the δ_ω denote additional second-stage, per-scenario variables. For scenario ω, δ_ω is the maximum of 0 and the excess cost beyond η if the excess is positive. To compute CVaR, it is also necessary to impose per-scenario constraints as follows:

$$\delta_\omega \geq \frac{\sum_{y \in Y} \sum_{g \in G} (c_g m_g^{max} U_{gy})}{(1+r)^{y-1}} + \xi_\omega - \eta \;\; \forall \omega \in \Omega \tag{2-7}$$

$$\delta_\omega \geq 0 \;\; \forall \omega \in \Omega \tag{2-8}$$

The quantity ξ_ω in (2-7) represents the discounted operational cost incurred under scenario ω, as defined in (2-5). The upper tail of the cost distribution represents high operational costs caused by high fuel prices and penalties from failing to satisfy large demands.

In practice, CVaR solutions are often viewed as excessively costly, so CVaR is often combined with expected-cost minimization in a weighted multi-objective scheme. For further discussions on the computation of CVaR, we refer to [63].

2.4 Scenario Tree Generation

We base our case study and test problem instance on real data we collected from the Energy Information Administration (EIA), the Midwest Independent System Operator (MISO), and the Joint Coordinated System Planning Report 2008 (JCSP) [37]. 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 that monitors the transmission system and provides safe and cost-efficient delivery of electric power across the Midwest US and one province, Manitoba, in Canada. JCSP is a joint organization in the Midwest and Northeast regions of the US formally initiated in November 2007. Both economic and reliability studies have been conducted by the JCSP 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. We now consider stochastic process models for these quantities, and propose a methodology to construct a sample-path tree that represents these processes, providing input to our GEP-EC and GEP-CVAR optimization models where paths through the tree are used as scenarios.

2.4.1 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. Because 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.

2.4.2 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 μy and variance $\sigma^2 y$ [60].

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 widely used for modeling financial markets [33]. It has the statistical property that $w(y) = \log\frac{X(y+1)}{X(y)}$ 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 $G(y)$.

2.4.3 Verification of Geometric Brownian Motion

To test whether both the annual electricity demand and the natural gas price can be represented as GBM, we obtained hourly demand data from years 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 years 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 μ_D = 0.0072, σ_D = 0.0094, μ_G = 0.037 and σ_G = 0.082. The related JMP outputs are shown in Figs. 2-1 and 2-2. 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 years 1991–2007 and years 1970–2006. Since the Shapiro-Wilk test statistic 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 statistic 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 log-normal 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 $w_D(y)$, and between $w_G(y+1)$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 Figs. 2-3 and 2-4. The R^2 for the log ratios of demand is 0.208272 and the R^2 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 zero correlation.

Another way to verify the independence is through the autocorrelation test with different lags of the time series model in JMP. The null hypothesis is that there is no autocorrelation. For the time series for historical demand in Fig. 2-5, the p-value 0.0742 with lag = 1. For the time series for historical natural gas price in Fig. 2-6, the p-value is 0.2086 with lag 1. Both of the p-values fail to reject the null hypothesis, which indicates there is no autocorrelation for the time series data with lag 1. Therefore, the assumption that both $D(y)$ and $G(y)$ follow GBM processes is supported by these tests [46].

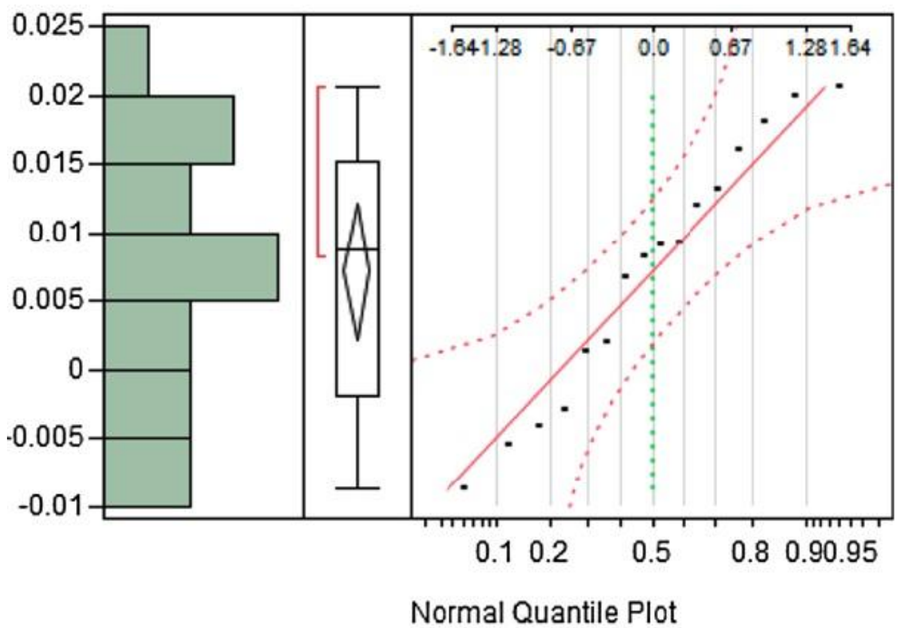

Figure 2- 1 Log Ratios of Annual Demand in the Midwest US from Year 1991-2006

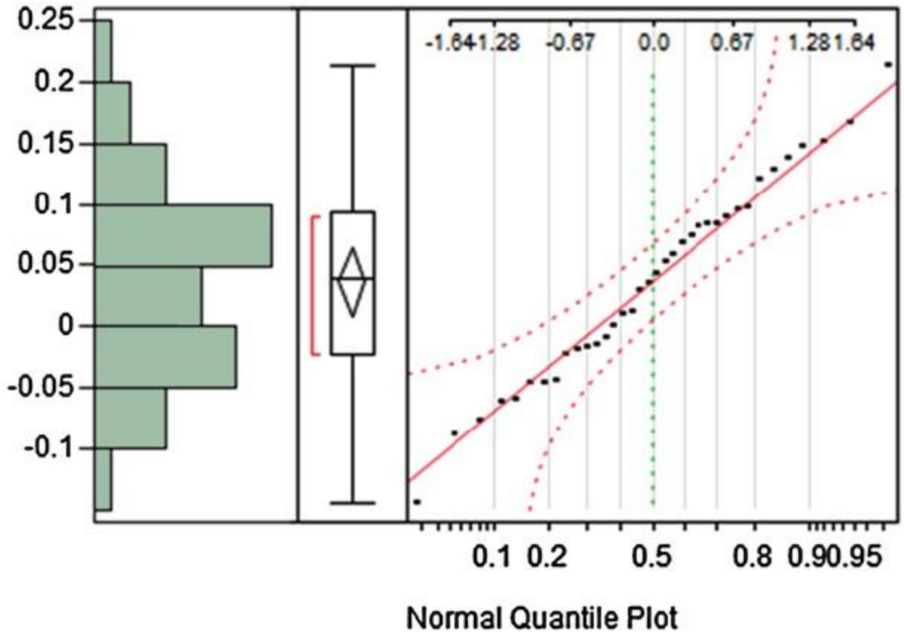

Figure 2- 2 Log Ratios of Annual Average Natural Gas Prices in the Midwest US from Year 1970-2006

2.4.4 Statistical Properties of Random Variables

Because $w(y) = \log\left(\frac{X(y+1)}{X(y)}\right)$ is normally distributed with mean μ_x and standard deviation σ_x, the ratio $\frac{X(y+1)}{X(y)}$ exhibits the log-normal distribution with mean μ_x and standard deviation σ_x. Consequently, we can derive the following statistical properties of the GBM using the following formulas for the log-normal distribution [36]:

$$E\left(\frac{X(y+1)}{X(y)}\right) = e^{\mu_X + \frac{\sigma_X^2}{2}} \quad (2\text{-}9)$$

$$Var\left(\frac{X(y+1)}{X(y)}\right) = \left(e^{\sigma_X^2} - 1\right)e^{2\mu_X + \sigma_X^2} \quad (2\text{-}10)$$

$$sk\left(\frac{X(y+1)}{X(y)}\right) = \left(e^{\sigma_X^2} + 2\right)\sqrt{e^{\sigma_X^2} - 1} \quad (2\text{-}11)$$

Let $x(y)$ denote the actual value of x in year y. Assume that the initial year of the expansion planning is year 0, and that there is no uncertainty in $x(0)$. Given (2-9)–(2-11) and the condition that $X(0) = x(0)$, we then derive conditional formulas for the evaluation of $X(y)$ as follows:

$$E(X(y+1)|X(u), 0 \le u \le y) = E\left(e^{Z(y+1)}|Z(u), 0 \le u \le y\right) = E\left(e^{Z(y)+Z(y+1)-Z(y)}|Z(u), 0 \le u \le y\right) = e^{Z(y)}E\left(e^{Z(y+1)-Z(y)}|Z(u), 0 \le u \le y\right) = x(y)E\left(\frac{X(y+1)}{X(y)}\right) = x(y)e^{\mu_X+\frac{\sigma_X^2}{2}} \quad (2\text{-}12)$$

$$Var(X(y+1)|X(u), 0 \le u \le y) = Var\left(e^{Z(y+1)}|Z(u), 0 \le u \le y\right) = Var\left(e^{Z(y)+Z(y+1)-Z(y)}|Z(u), 0 \le u \le y\right) = \left(e^{Z(y)}\right)^2 Var(e^{Z(y+1)-Z(y)}|Z(u), 0 \le u \le y) = x(y)^2 Var\left(\frac{X(y+1)}{X(y)}\right) = x(y)^2\left(e^{\sigma_X^2}-1\right)e^{2\mu_X+\sigma_X^2} \quad (2\text{-}13)$$

$$sk(X(y+1)|X(u), 0 \le u \le y) = sk\left(e^{Z(y+1)}|Z(u), 0 \le u \le y\right) = sk\left(e^{Z(y)+Z(y+1)-Z(y)}|Z(u), 0 \le u \le y\right) = \frac{E\left(e^{Z(y)}\left(e^{Z(y+1)-Z(y)}-E\left(e^{Z(y+1)-Z(y)}\right)\right)\right)^3}{\left(Var\left(e^{Z(y)}e^{Z(y+1)-Z(y)}\right)\right)^{\frac{3}{2}}} = \frac{E\left(e^{Z(y+1)-Z(y)}-E\left(e^{Z(y+1)-Z(y)}\right)\right)^3}{\left(Var\left(e^{Z(y+1)-Z(y)}\right)\right)^{\frac{3}{2}}} = sk\left(\frac{X(y+1)}{X(y)}\right) = \left(e^{\sigma_X^2}+2\right)\sqrt{e^{\sigma_X^2}-1} \quad (2\text{-}14)$$

From (2-12)–(2-14) 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, depending only on σX. The results of applying (2-12)–(2-14) to the annual demand and annual natural gas price in the Midwest region are summarized in Table 2-1. The correlation value between the two random variables in each year was also obtained by JMP as shown in Fig. 2-7. In general, the annual natural gas price and electricity demand both have increasing trends.

The R^2 value for the linear regression model of the annual demand versus the annual natural gas price is 0.75002, with a p-value <0.0001. Thus we reject the null hypothesis of 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.

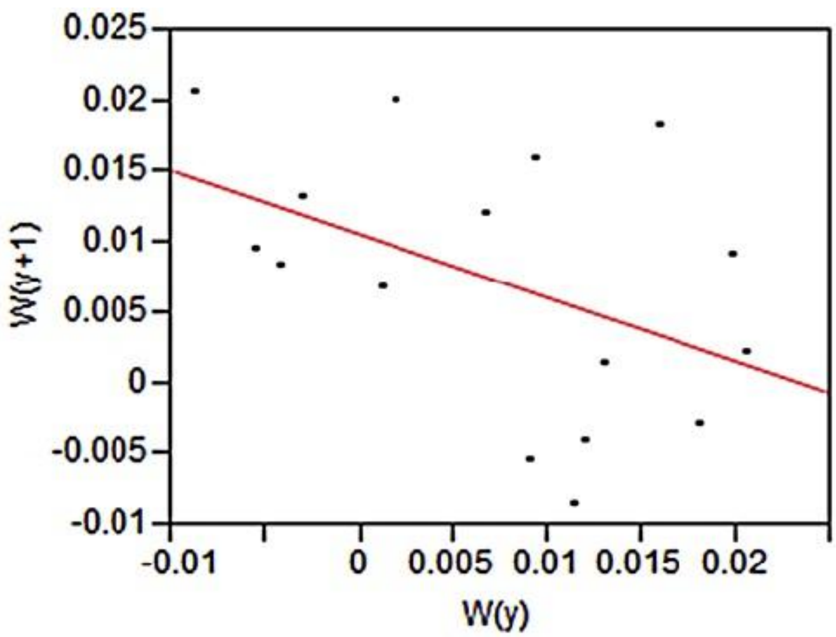

Figure 2- 3 Scatter Plot of Annual Demand Quantities over Successive Years in the Midwest US from Year 1991-2006

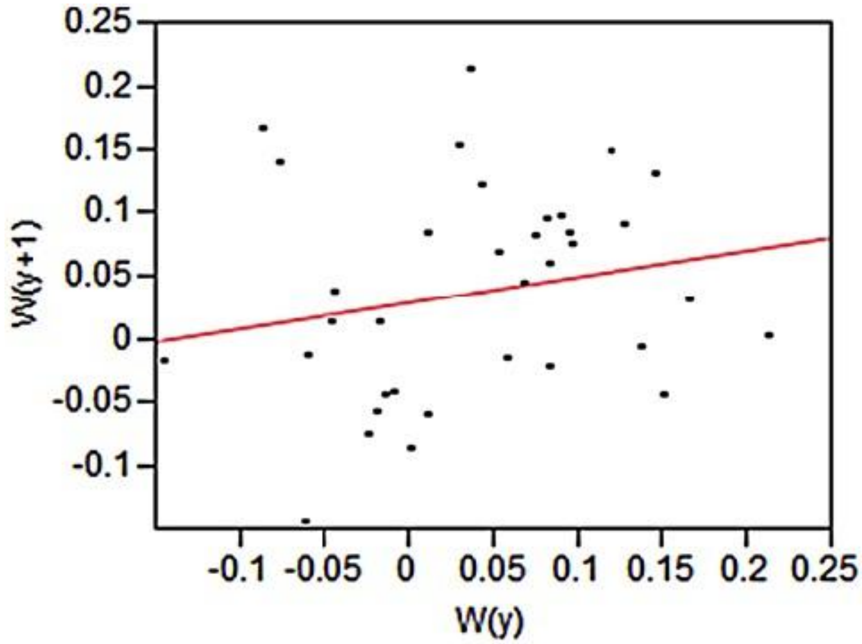

Figure 2- 4 Scatter Plot of Annual Average Natural Gas Prices over Successive Years in the Midwest US from Year 1970-2006

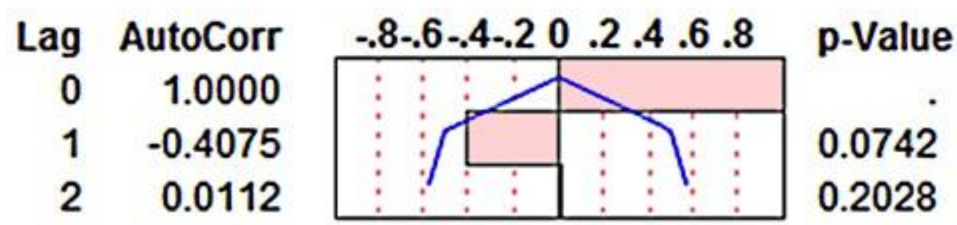

Figure 2- 5 Time Series Autocorrelation with Lag=1 for Demand in the Midwest US from Year 1991-2006

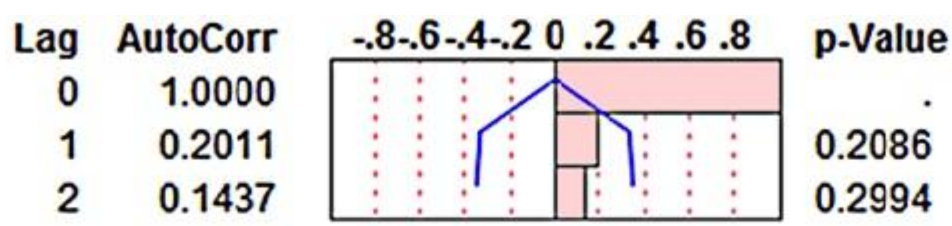

Figure 2- 6 Time Series Autocorrelation with Lag=1 for Demand in the Midwest US from Year 1970-2006

Table 2- 1 Statistical Properties of Random Variables Representing Natural Gas Price and Demand in the Mid-west US over Planning Horizon

Random variables	Statistical property	First year (y = 1)	After first year (y > 1)
Demand	Mean	1.00727 *d(0)*	1.00727 *d(y-1)*
(Billion MWhs)	Std. dev.	0.009469 *d(0)*	0.009469 *d(y-1)*
D(y)	Skewness	0.028	0.028
Natural gas price	Mean	1.041188 *g(0)*	1.041188 *g(y-1)*
($/thousand cubic feet)	Std. dev.	0.085521 *g(0)*	0.085521 *g(y-1)*
G(y)	Skewness	0.25	0.25
Gas price and demand	Correlation	0.866	0.866

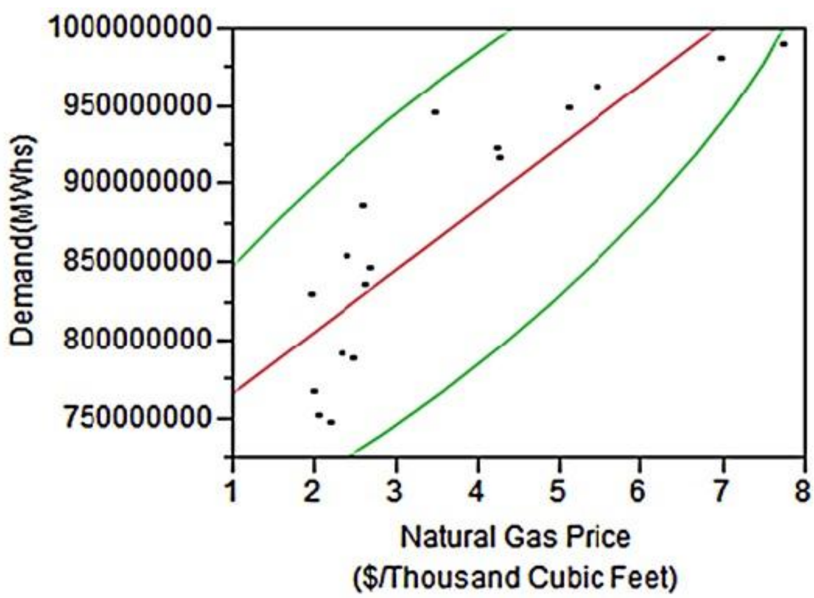

Figure 2- 7 Correlation between Total Annual Demand and Average Annual Natural Gas Price in the Midwest US from Year 1991-2006

2.4.5 Scenario Generation Method

Given distributional characterizations of uncertainty in future demand and natural gas price, we now consider the issue of generating a tree from those distributions. The paths through this tree will form the scenarios used as input to our GEP-EC and GEP-CVAR optimization models. We use (2-12)–(2-14) to specify statistics for one-step ahead samples. Let X and q respectively denote the set of representative values and the associated probability vector, and let S denote a set of labels for the elements in X. Next, let $x_s \in X$ denote a single pair of representative values for demand and gas price with label $s \in S$ and probability q_s. Finally, let $f_i(X, q)$ denote the value of the ith statistical measure of interest computed for X and q. For example, if $f_i(X, q)$ represents the sample mean vector $\bar{X}$, then $f_i(X, q)$ is computed as $\sum_{s \in S} q_s x_s$.

To generate a sample of paths through a tree that accurately represents the uncertainty space of future demand and natural gas price, we use a procedure introduced by Høyland and Wallace [24]. The foundation of this procedure is the following optimization model, in which the objective is to match as closely as possible statistical properties of the original continuous random variables and those of a set of discrete values that we will treat as if they were samples from those random variables:

$$\min_{X,q} \sum_{i \in P} \phi_i \left(f_i(X, q) - P_{VAL_i}\right)^2, \tag{2-15}$$

$$\sum_{s \in S} q_s = 1, \tag{2-16}$$

$$q_s \geq 0 \quad \forall s \in S \tag{2-17}$$

In this formulation, P (indexed by i) denotes a set of statistical measure of interest, and P_{VAL_i} denotes the value of the ith statistical measure quantified in the context of the original continuous random variables. The squared sum of differences between P_{VAL_i} and the corresponding sampled quantity $f_i(X, q)$ are then minimized. The weights ϕ_i provide a mechanism to specify the relative importance to modelers of the various measures $i \in P$. In our case study, we define P to include the mean and variances of both random variables (with ϕ_i = 2), the skewness of both random variables (with ϕ_i = 1), and the correlation between the random variables (with ϕ_i = 1). The constraints (2-16) and (2-17) ensure the q_s are interpretable as probabilities.

Høyland and Wallace discuss the issue of how to select an appropriate number of scenarios |*S*| for a given optimization problem. To avoid both underspecification and overspecification, the chosen number of statistical measures |*P*| should be similar to the number of decision variables in (2-15). In our case study, we have a two-dimensional scenario variable to represent demand and natural gas price, and the probability needs to be decided. The number of decision variables is |*S*|(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. We use |*S*| = 3 because |*S*|(2 + 1) −1 ≈ 7. Hence, the number of branches from each node in the tree representing the stochastic process is determined to be 3.

The scenario generation problem is a nonlinear mathematical program with a nonlinear objective function and linear constraints. We solve this problem using the nonlinear solver Tomlab/SNOPT, available from Matlab, which iterates to a locally optimal solution from a specified starting point. We use multiple starting points to heuristically identify a posited global optimum. The initial points for the 3 value vectors, X and associated scenario probabilities, q are assumed to be $X_1 = (\mu_D - \sigma_D, \mu_G - \sigma_G)$, $q_1 = 0.333$, $X_2 = (\mu_D, \mu_G)$, $q_2 = 0.334$, $X_3 = (\mu_D + \sigma_D, \mu_G + \sigma_G)$, $q_1 = 0.333$. The minimum possible objective value is expected to equal zero, if the specifications are consistent. However, because (2-15) is generally not a convex optimization problem, the final solution may be only locally optimal even with different initial conditions. If the derived statistical properties are still close to the specification, a locally optimal solution is also acceptable. However, if severe inaccuracy occurs, it must be resolved either resetting the weight coefficient ϕ_i or increasing the number of initial starting points.

Once the 3 representative value pairs for 2009 are generated based on $d(0)$ and $g(0)$, we generate the 2010 values similarly. Conditional statistical properties are first specified based on the 3 generated 2009 value pairs by applying (2-12)–(2-14). Then, another 3 discrete pairs are generated using (2-15)–(2-17).

The complete stochastic process tree can be recursively constructed through the end of the planning horizon. A fragment of the tree for our case study is shown in Fig. 2-8. Each column in the tree represents a single year, and each tree node represents one possible outcome for that year. For each node, the number at the top of the corresponding block represents the product of the conditional probabilities for that specific scenario path up to that node (i.e., the absolute probability of occurrence; for nodes in the final period, these are the path probabilities π_ω). The numbers in the parenthesis at the bottom of each block indicate the scenario-specific values for both demand and natural gas price. In the initial year 2008,

$d(0) = 0.57$ and $g(0) = 9.37$. The units for demand and natural gas price are respectively billion MWhs and $/thousand cubic feet.

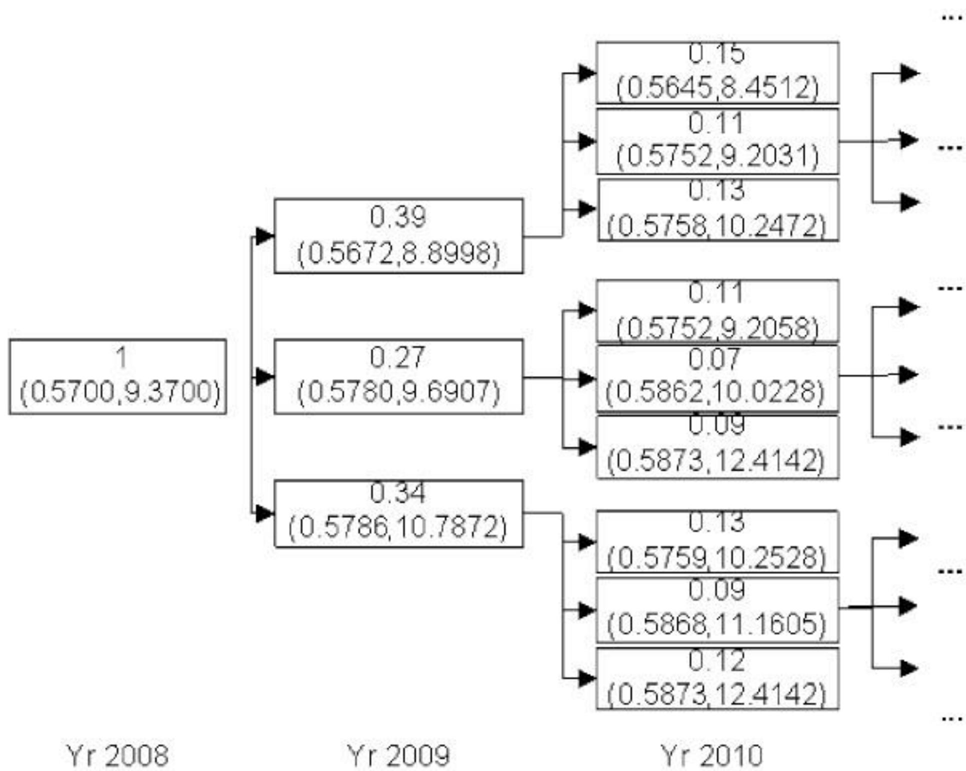

Figure 2- 8 A Fragment of the Stochastic Process Tree for a Multi-year Planning Horizon, Showing the Subpath Probability, Total Demand for the Year, and Annual Average Natural Gas Price

2.5 A Realistic Problem Instance

We now describe in detail a realistic instance of our generation expansion planning problem, derived from system data associated with the Midwest US. In particular, we describe the sources for all fixed and uncertain parameters, primarily drawn from EIA, MISO, and JCSP sources. The computational studies described subsequently in Section. 2.7 are based on this instance.

2.5.1 Demand

We divide each year in the planning horizon into three sub-periods corresponding to demand blocks: peak, medium, and low. Hourly demands during the year are ordered in a decreasing sequence, forming a load duration curve (LDC). The three demand blocks are then formed by imposing break points at the top quarter, middle half, and bottom quarter of the load curve. An example is provided in Fig. 2-9, which shows the hourly load for 2008 in the Midwest US; data are obtained from the real-time market report of MISO [34]. The demand for each sub-period, which actually represents nonconsecutive hours in the year, is simply taken as the corresponding area under the LDC. By considering only three demand blocks, we reduce the problem size and also retain information about the chronological demand variability.

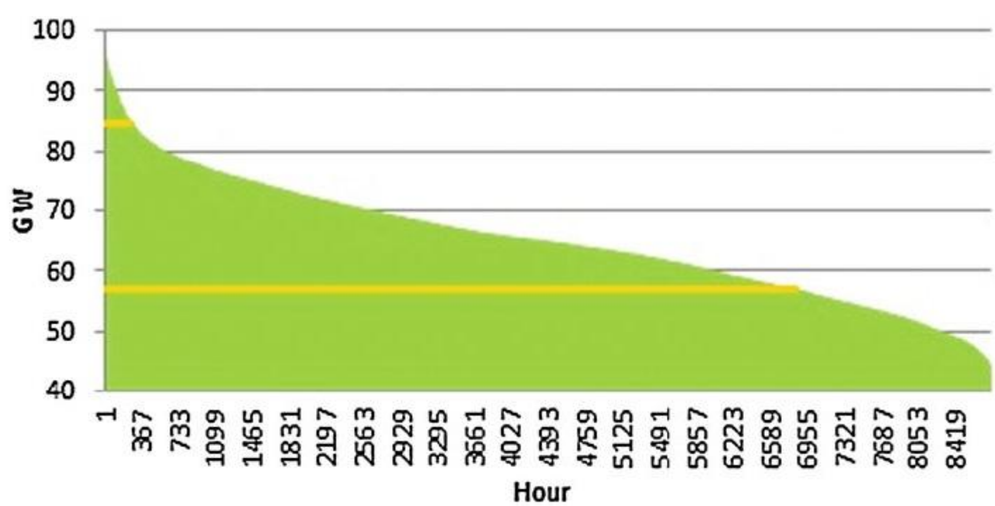

Figure 2- 9 Load Duration Curve in Year 2008 (3 Demand Blocks)

Demand data for the 3 demand blocks corresponding to these 2008 data (converted to a non-leap-year basis) is summarized in Table 2-2. The multiplier λ_t for demand block t is the ratio of average hourly demand in that block to the overall average demand for the year.

Because future LDCs are unknown, we assume they are formed via incremental additions to the reference year 2008 (ignoring leap years) and the blocks are proportional to those in the reference year 2008. That is, given a realization of total energy demand $d(y)$ in year y under the stochastic process path ω, for $t \in T(y)$ the average hourly demand $d_{t\omega}$ in MW is $\max\left[\frac{\lambda_t d(y)}{8760E+9} - d(0), 0\right]$. Because the stochastic demand process is incremental over the reference year, we set the initial numbers of generating units u_g to zero accordingly.

Table 2- 2 Baseline Sub-period Hourly Demand

t	Block	Hours (h)	Load (MWh)	λ_t
1	Peak	271	2.427E+7	1.38
2	Base	6556	4.437E+8	1.04
3	Shoulder	1933	1.006E+8	0.80
	Total	8760	5.685E+8	

2.5.2 Generator Data

We assume six different candidate generator types are available for new construction in all time periods of the planning horizon: G = { BaseLoad, CC, CT, nuclear, wind, IGCC}. Here, CC and CT respectively denote combined cycle and combustion turbine power plants, both of which are fueled by natural gas. Integrated gasification combined cycle (IGCC) power plants are fueled by coal.

With the exception of wind farms, the calculation of generator build cost is based on the capital expenditure profile suggested by the JCSP [37]. Table 2-3 shows the fraction of overnight investment costs actually expended in each year to build each type of generator. For wind farms, because the construction time is estimated at two years [37], the capital expenditure is simply split evenly. To obtain c_g, the total cost to build a single generator of type g, we sum the present value in each construction year using the discount rate r. For example, to obtain c_g for a CC plant, we first multiply the overnight build cost by the capital expenditure percentage for each year, yielding \$1.833(0.25) million for the first year, \$1.833(0.5) million for the second year, and \$1.833(0.25) million for the third year. These values are then discounted by r to the first year, and summed to form c_g. Table 2-4 reports both the overnight investment cost and the final calculated build cost cg for each generator type g. We assume that the newly built generators are able to generate electricity beginning in the year that the expansion decision is made; i.e., the lead time for building and installing a generator is ignored.

Table 2- 3 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	0.3			0.01		0.3
5	0.3			0.01		0.3
6	0.1			0.02		
7				0.03		
8				0.2		
9				0.3		
10				0.3		
11				0.1		

Table 2- 4 Overnight Cost for Generators

Technology	Overnight cost ($/MW)	Build cost c_g($/MW)
Baseload	1.833E+6	1.446E+6
CC	0.857E+6	0.795E+6
CT	0.597E+6	0.575E+6
Nuclear	2.928E+6	1.613E+6
Wind	1.713E+6	1.650E+6
IGCC	2.118E+6	1.671E+6

Table 2- 5 Generation Cost Related Parameters for the Generators in the First Year

Technology	Fuel price ($/Mbtu)	Heat rate (Btu/KWh)	Efficiency	Variable OM cost ($/MWh)
Baseload	3.37	8844	0.4	4.70
CC	9.11	7196	0.56	2.11
CT	9.11	10842	0.4	3.66
Nuclear	0.93E-3	10400	0.45	0.51
Wind	0	N/A	N/A	5.00
IGCC	3.37	8613	0.48	2.98

Power generation costs are divided into two components: (1) fuel costs and (2) variable operation and maintenance (OM) costs. All the related parameters for calculating the generation cost for year 2008 are shown in Table 2-5 from JCSP [37]. From Table 2-5, we can easily calculate the generation costs. For $t \in T_1$ and all stochastic process paths ω, $l_{gt\omega}$ is found by converting the fuel price to $/MWh using the heat rates and efficiencies, then adding the variable OM cost. In later years we made the escalation assumptions of 2% annual growth rate in the fuel prices (coal, nuclear) and 3% annual growth rate in the variable OM cost, as suggested by JCSP [37]. For the CC and CT technologies, the fuel price in $/MBtu is obtained by dividing the natural gas price in the corresponding nodes of the stochastic process tree by 1.028 MBtu per thousand cubic feet.

The installed capacity and generator ratings are based on the JCSP [37] and the generator ratings are calculated by their installed capacity multiplied by the forced outage rate (FOR), also from the JCSP [37]. 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 relevant assumptions are shown in Table 2-6. For the maximum units to build constraint over the whole planning horizon, we used the assumptions shown in Table 2-7.

Table 2- 6 Installed Capacity and Generator Rating for Generators

Type	Baseload	CC	CT	Nuclear	Wind	IGCC
Generator 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

Table 2- 7 Max Units to Build for Generators

Type	Baseload	CC	CT	Nuclear	Wind	IGCC
Max units built, u_g^{max}	4	10	10	1	45	4

2.5.3 Other

We assume an annualized interest rate of r = 0.08, based on data reported by the JCSP [37]. This rate is used to discount all future expenditures (capital investments and operational costs) to the reference year 2008. The penalty, p_u, for unserved load is 100,000 $/MWh.

2.6 Scenario Sampling and Cost Confidence Intervals

In this section, we describe procedures for computing statistical bounds on the optimal objective function values for both the GEP-EC and GEP-CVAR optimization models. In the following, we use the term "cost" loosely to denote either the expected value objective (2.4) or the CVaR objective (2.6).We begin in Section 2.6.1 by discussing the motivation behind such procedures. The specific procedure we employ is detailed in Section 2.6.2. We conclude in Section 2.6.3 with a discussion of the key practical issues surrounding the use of such bounding procedures in the context of our optimization models.

2.6.1 Motivation

In any stochastic optimization model in which uncertainty is treated via Monte Carlo sampling (in contrast to analytically), practical questions arise involving the nature of the sampling strategy employed. From the standpoint of solution validation, a key question is: Did we use a sufficient number of samples to obtain a solution with the required level of accuracy? From a computational standpoint, a related question of significant practical importance (particularly in the case of stochastic mixed-integer programs) is: What is the smallest number of samples with which we can obtain a solution possessing the required level of accuracy? The remainder of this sub-section is devoted to discussing computational procedures to answer these two questions.

Let Ω^+ denote the (infinite) space of all possible scenarios to our generation expansion planning problem, defined using the stochastic processes introduced in Section 2.4. Further, let Ω denote a finite set of scenarios generated from Ω^+ via sampling. We assume that the process used to generate Ω is sufficiently representative of the underlying continuous probability space; i.e., that our sample of 19,683 scenarios accurately reflects the statistics of the reference stochastic process. In practice, directly solving either GEP-EC and GEP-CVAR using Ω is computationally infeasible, requiring at a minimum several weeks of run-time on a modern high-performance workstation. Thus, we focus on the analysis of sub-samples $\tilde{\mathcal{E}} \subset \Omega$.

In our generation expansion planning models, the first-stage decisions U_{gy} represent the primary variables of interest, while the second-stage variables $L_{gt\omega}$ and $E_{t\omega}$ are derived variables used to quantify the operational impacts given particular U_{gy}. Further, given fixed values for the U_{gy} variables, the values for the $L_{gt\omega}$ and $E_{t\omega}$ variables can be immediately determined by solving the resulting linear program for each scenario independently. Thus, in order to assess the accuracy of a particular solution, we consider our confidence in the total cost as a function of the first-stage U_{gy} variables.

Let $F(v; \tilde{\mathcal{E}})$ denote the minimal total cost of either GEP-EC (2.4) or GEP-CVAR (2.6) given a set of scenarios $\tilde{\mathcal{E}}$, subject to the constraint that the first-stage decision variables U are fixed at v if v is feasible. If v is not feasible, then $F\left(v; \tilde{\mathcal{E}}\right) = \infty$. Finally, let z^* denote the minimal total cost for one of our optimization models if it could be computed for the entire set of scenarios Ω^+ (i.e., $z^* = \min_y F(y; \Omega^+)$).

2.6.2 A Multiple Replication Procedure

Once we obtain what we believe to be a good first-stage solution $\widehat{U}$, we wish to compute a bound on the gap between the total cost of $\widehat{U}$ and z^*. One way to perform such validation is to sample additional groups of scenarios (those not used to compute $\widehat{U}$), optimize the corresponding extensive forms, and use differences in the resulting total costs relative to $\widehat{U}$ to compute statistical bounds.

Mak, Morton and Wood [44] propose a method for computing a confidence interval for the total cost that could be achieved if it were possible to use extremely large sample sizes; i.e., sample sizes so large that the solution to the full stochastic program using all scenarios from Ω^{+} is perfectly approximated. This method is known as the Multiple Replication Procedure (MRP). Consider a finite universe of N available scenarios, from which $\hat{n} < N$ are drawn at random and without replacement. The extensive form of a stochastic program is then solved using the sample of $\hat{n}$ scenarios, resulting in a "reference" or baseline solution $\widehat{U}$. The remaining $N - \hat{n}$ scenarios are then partitioned into equal-sized groups, each containing n scenarios. To enforce the equality constraint, small numbers of scenarios may be discarded. The MRP takes as input a parameter $0 < \alpha < 1$ specifying a $1 - \alpha$ confidence level and generates n_g sets of "validation" scenario groups $\tilde{\mathcal{E}}^i, i = 1, \ldots, n_g$, such that each set has equal probability.

A high-level description of the MRP is as follows (loops are combined in the actual implementation; we consider the common random number streams variant described in [44]):

1. For each validation group $\tilde{\mathcal{E}}^i, i = 1, \ldots, n_g$, compute a gap statistic as follows:

$$G^i = F\left(\widehat{U}, \tilde{\mathcal{E}}^i\right) - \min_U F\left(U, \tilde{\mathcal{E}}^i\right)$$

2. Compute the average gap statistic $\bar{G}$ and the sample standard deviation s_G.

Each computation of G_i in Step 1 involves solving two optimization problems. First, the computation of $F(\widehat{U}, \tilde{\mathcal{E}}^i)$ requires solution of a stochastic program extensive form given fixed first stage decisions $\widehat{U}$; in the case of GEP-EC and GEP-CVAR, this simply requires the solution of a linear program due to the fixing of all first-stage integer variables. Second, the computation of $\min_U F(U, \tilde{\mathcal{E}}^i)$ requires the solution of a stochastic mixed-integer program extensive form. Thus, depending on the values of N, n_g and $\hat{n}$, the MRP can be computationally expensive.

Given the gap statistics computed above, the approximate $(1-\alpha)$-level confidence interval for the optimality gap on the total cost of baseline solution $\widehat{U}$ is then given as:

$$[0, \bar{G} + \frac{t_{n_g-1,\alpha} s_G}{\sqrt{n_g}}]$$

where $t_{n_g-1,\alpha}$, is the α tail value for a t -distribution with $n_g - 1$ degrees of freedom.

2.6.3 Research Questions

Several fundamental computational questions arise when applying the MRP in practice. These questions affect both the practical utility of the MRP and the interpretation of results obtained using the procedure. The experimental methodology described subsequently in Section 2.7.3 is explicitly designed to yield data to provide at least preliminary answers such questions, as outlined below.

Foremost among these questions is the selection of parameter values for the procedure, specifically for $\hat{n}$ and n_g. There is presently little empirical guidance to suggest specific values for MRP parameters. At one extreme, there is strong motivation to use a large (relative to N) $\hat{n}$ value, in order to obtain an accurate reference solution. On the other hand, we want to leave a sufficient number of scenarios for purposes of validation; too few scenarios will likely lead to either small n_g or n, which may in turn lead to unnecessarily conservative confidence intervals. Further, given $N - \hat{n}$ samples for validation, there is the question of how to select n_g. One could argue that small values of n_g yield more accurate solutions. However, the use of small n_g also leads to conservative confidence intervals, and it is therefore necessary to balance both the number and size of the validation groups $\tilde{\mathcal{E}}^i$. Fundamentally, as we show next in Section 2.7, different values of the MRP parameters do lead to different results, thus motivating the need to explore trends in the MRP parameter space.

At a higher level of abstraction, there is the question of stability of MRP results: Given a fixed set of scenarios, how variable are results obtained with different partitions of those scenarios among the baseline $\hat{n}$ group and the n_g validation groups? This variability has not been addressed previously in the literature, and as we show in Section 2.7, its presence affects both the aggregation and interpretation of results from the MRP procedure.

Next, there is the up-front question of how to select a proper value of N. In our case study, we conjectured that only a fraction of the 19,683 scenarios were ultimately required to obtain reasonably small confidence intervals. This conjecture was based on a series of experiments described below in Section 2.7.2, in which we sub-sample an increasing number N of scenarios, solve the resulting stochastic program extensive form, and observe the stability in the total cost as a function of N. We observe stability by approximately N = 1,000, so we conjectured that at most this number of scenarios would be required to obtain reasonably tight confidence intervals on optimal solution cost. This conjecture is substantiated by the experiments discussed in Section 2.7. However, we observe that in general, selection of N is more art than science, requiring a careful balance between the value of N, the computational time associated with the MRP procedure, and the desired tightness of the obtained confidence intervals.

Finally, in addition to consideration of cost confidence intervals, there is the question of solution similarity: How different are the solutions obtained under different MRP parameterizations or replications? Answers to this question directly affect the decision-maker, as it is possible that even if observed solution costs are disparate, the underlying solutions may not be—implying that the existence of cost variability may have little impact on any final investment decisions.

2.7 Experimental Results

We now report and analyze the results obtained by executing the MRP on our two optimization models. In Section 2.7.1, we briefly discuss the computational tools used to model and solve our GEP-EC and GEP-CVAR formulations. Our experimental methodology involving the MRP is then detailed in Section 2.7.3. We present results for GEP-EC and GEP-CVAR in Sections 2.7.4 and 2.7.5, respectively. Finally, we conclude with a discussion regarding the structural similarity of solutions obtained with different parameterizations of the MRP in Section 2.7.6.

2.7.1 Implementation

We modeled the GEP-EC and GEP-CVAR two-stage stochastic programs using the PySP software package [56, 70]. PySP is a Python-based, open-source tool co-developed and maintained by Sandia National Laboratories, the University of California Davis, Texas A&M, and others. It is part of IBM's COIN-OR open-source initiative for operations research software [10]. The data corresponding to the case study instance described in Sections 2.4 and 2.5 was encoded in PySP's native data format, similar to that used by the commercial modeling tool AMPL [5]. Both the PySP model and data files can be obtained by contacting the authors.

Stochastic programs in PySP are specified in a scenario-oriented manner. The base deterministic (i.e., single-scenario) optimization model must be specified first. Then, users specify the data defining the scenario tree for the problem instance under consideration, including the stages, tree nodes, branching probabilities, and variable-to-stage assignments. Finally, data specifying parameter values for each scenario is supplied. PySP then uses this data to construct an internal representation of the corresponding stochastic program, by creating instances for each scenario in the scenario tree and constructing constraints to enforce variable non-anticipativity at each composite node. A comprehensive description of the use and implementation of PySP can be found in [70].

We leverage integrated algorithms within PySP to execute all MRP trials described below. Specifically, PySP provides a generic MRP implementation, and functionality to generate and solve extensive forms of stochastic programs. PySP leverages commercial solvers to obtain solutions to the mixed-integer extensive forms, including CPLEX 12.2 [11], which was used for all experiments. All experiments were executed on a modern 2.93 GHz Intel Xeon 8-core workstation (each core is hyper-threaded), with 96 GB of RAM, running Linux.

2.7.2 Cost Stability and Extensive Form Run-times

In order to conduct experiments concerning how to best allocate N scenarios between those used to compute the best possible solution and those used to compute a confidence interval, we first conduct experiments to determine the largest computationally feasible *N* and to verify that the optimal solution values obtained using the full *N* scenarios are reasonably stable. The optimal solution cost of both GEP-EC and GEP-CVAR will stabilize toward an asymptotic value as $\hat{n} \to \infty$. It is an empirical question as to how fast this convergence occurs for finite sample sizes. However, knowledge of the empirical convergence rate in the low sample count regime suggests a heuristic for selecting the parameter *N* associated with tests of the MRP procedure.

Table 2- 8 The Stability of Optimal Solution Costs for the GEP-EC and GEP-CVAR Optimization Models. Columns Report the Solution Cost (USD) and Run-time (wall clock seconds), as a Function of the Number of Scenarios *N*

N	Optimal solution cost		Run-time	
	GEP-EC	GEP-CVAR	GEP-EC	GEP-CVAR
10	1.91175E+10	2.18873E+10	17.49	13.19
100	1.81437E+10	2.49147E+10	57.93	145.9
250	1.72728E+10	2.45407E+10	986.45	1622.11
500	1.74896E+10	2.38776E+10	5176.26	1547.48
1000	1.77731E+10	2.36481E+10	52442.93	8402.54

In Table 2-8, we show the optimal solution cost and the run-time required to solve the corresponding stochastic program extensive form, for a range of *N*. For GEP-CVAR, we use l= 0.05. The scenarios in each sample were drawn randomly and without replacement from the full scenario tree (containing all 19,683 scenarios); different random seeds were used for each value of N, so any overlap in the selected scenarios is by chance. Run-time is reported in terms of wall clock time, while costs are measured in USD. We observe that CPLEX 12.2 is multi-thread capable, and routinely uses 16 threads during execution of these problems on our compute hardware. Thus, execution times on less powerful platforms will be significantly longer.

Analyzing the results in Table 2-8, we observe significant growth in run-time as a function of N. This is consistent with the observed empirical difficulty of stochastic mixed-integer programs reported in the literature, particularly at this scale. Further, the memory requirements are non-trivial, exceeding 8 GB of RAM for the larger runs. Given the large run-times at $N = 1000$, it is clear that solution of either model given our complete scenario tree is intractable via the extensive form approach, and would likely require significant effort even leveraging decomposition-based approaches—which in turn would require significant implementation effort and tuning. There is no consistent trend in the run-time differences between the GEP-EC and GEP-CVAR runs. The large difference at $N = 1000$ is an artifact of the particular sample chosen; mixed-integer solvers are known to exhibit significant variability in run times as problem data is changed.

The optimal solution costs for both GEP-EC and GEP-CVAR vary significantly for $N \leq 100$, but start to stabilize once $N \geq 250$. When $N = 1000$, we further observe (not reported) significant stability across replications of the experiment. These two trends lead us to heuristically select $N = 1000$ for all replications of the MRP procedure reported subsequently.

2.7.3 Experimental Methodology

To experiment with ways to allocate the scenarios in a fixed sample between finding a reference solution and computing its confidence interval, we draw a random sample of N = 1000 scenarios from the full scenario tree described in Section 2.4. We then execute a full factorial experiment for both the GEP-EC and GEP-CVAR optimization models, executing the MRP for each of the following combinations of parameters:

– $\hat{n} \in \{70, 140, 280, 420, 560\}$

– $n_g \in \{2, 5, 10, 20, 40\}$

These values were selected to obtain results over the spectrum of MRP parameterizations. As discussed in Section 2.6.3, there is little empirical evidence to guide selection of $\hat{n}$ and n_g.

For each combination of $\hat{n}$ and n_g we execute five MRP trials, varying the random seed used to partition the set of N scenarios into the set $\widehat{\Omega}$ of $\hat{n}$ scenarios, which is used to compute the reference solution, and the n_g validation groups, which are used to compute a confidence interval. To reduce variance, we use identical sets of random seeds across trials involving different $\hat{n}$ and n_g. For each trial, we record the confidence interval width on the optimal solution cost $f(\widehat{U}) \equiv f(\widehat{U}, \widehat{\Omega})$ for level α = 0.05. The particular value of α does not affect the qualitative nature of the results presented below (for example, moving from α = 0.05 to α = 0.01 roughly doubles the confidence interval width in the worst case), and was selected arbitrarily. For all GEP-CVAR runs, we use l = 0.05.

2.7.4 Expected Cost Minimization

The results obtained from executing the MRP on our GEP-EC optimization model are shown in Table 2-9. For each MRP parameterization, the table shows results for both a single, arbitrary MRP trial and aggregated statistics taken over all five MRP trials. Units for entries associated with $f(\widehat{U})$ and the 95% confidence interval (CI) are given in billions USD.

Table 2- 9 Results of Applying the Multiple Replication Procedure (MRP) to the GEP-EC optimization model. The MRP Input Parameters are Reported in the Columns Labels $\hat{n}$, n_g, and n. Outputs are Reported for the MRP as well as Summary Statistics taken over Five MRP Trials to Verify the Stability of the Procedure.

$\hat{n}$	n_g	n	MRP results		Verification: five MRP trials	
			$f(\widehat{U})$	95% CI	$f(\widehat{U})$ std. dev.	95% CI std. dev.
70	2	465	17.362	0.021	0.42	0.22
70	5	186	17.362	0.055	0.42	0.19
70	10	93	17.362	0.051	0.42	0.20
70	20	46	17.362	0.124	0.42	0.19
70	40	23	17.362	0.290	0.42	0.19
140	2	430	17.214	0.026	0.34	0.05
140	5	172	17.214	0.070	0.34	0.04
140	10	86	17.214	0.074	0.34	0.03
140	20	43	17.214	0.101	0.34	0.02
140	40	21	17.214	0.241	0.34	0.03
280	2	360	17.151	0.053	0.39	0.03
280	5	144	17.151	0.074	0.39	0.03
280	10	72	17.151	0.101	0.39	0.03
280	20	36	17.151	0.241	0.39	0.02
280	40	18	17.151	0.452	0.39	0.03
420	2	290	17.377	0.017	0.24	0.04
420	5	116	17.377	0.039	0.24	0.05
420	10	58	17.377	0.110	0.24	0.03
420	20	29	17.377	0.224	0.24	0.03
420	40	14	17.377	0.522	0.24	0.02
560	2	220	17.377	0.020	0.14	0.04

Table 2- 9 (continued)

560	5	88	17.377	0.063	0.14	0.04
560	10	44	17.377	0.103	0.14	0.05
560	20	22	17.377	0.198	0.14	0.07
560	40	11	17.377	0.571	0.14	0.07

We begin by considering the results obtained for one replication of the MRP, whose outputs are reported in the columns labeled "$f(\widehat{U})$"and "95% CI". Recall that we use (for a given MRP trial) a fixed random seed across experiments involving different combinations of $\hat{n}$ and n_g. This seed is used to randomize the list of scenarios, which are then sequentially partitioned into the baseline set containing $\hat{n}$ scenarios and the n_g validation groups. Consequently, the value of $f(\widehat{U})$ is identical in all trials in which only n_g is varied. We first observe that the cost $f(\widehat{U})$ of the baseline solution $\widehat{U}$ is remarkably stable as $\hat{n}$ is varied. This is consistent with the experiments reported in Section 2.7.2. Further, we would expect (and indeed observe) more stability in Table 2-9 because scenarios are accumulated as $\hat{n}$ is increased—due to the use of a common random number seed across the individual MRP trials, and the sequential partitioning of the scenarios.

Next, we analyze the widths of the 95% confidence interval on the cost $f(\widehat{U})$ as $\hat{n}$ and n_g are varied, considering a single MRP trial. The largest CI width obtained is approximately 571 million USD, representing less than 3.5% of corresponding total cost of 17.377 billion USD. These results strongly suggest that accurate estimates of the optimal cost to our GEP-EC optimization model can be obtained using a remarkably small number of scenarios—especially relative to the full scenario tree. Unexpectedly, we observe that CI widths for a fixed n̂ are generally increasing in n_g; the tightest confidence intervals appear at n_g = 2, despite the trivial number of degrees of freedom in the resulting t test. This behavior is partially explained by the fact that for even modest n_g, the number of samples in each group is too small to observe stability in $f(\widehat{U})$. However, this explanation fails to account for the fact that the pattern holds even when the baseline solution is obtained with $\hat{n}$ = 70. Fundamentally, the increase in the number of degrees of freedom in the t distribution is not offset by the stability in the total costs obtained from the validation scenario groups. Recall that the CI width is proportional to both the standard deviation of the sample optimality gap and the value of the t -statistic, and is inversely proportional to the square root of the number of scenario groups. The CI width monotonically increases as a function of n_g for the GEP-EC optimization model (this behavior is not universal, but depends on both the specific optimization problem under consideration and the particulars of the MRP parameterization) because the sample optimality gap variance overcomes the benefit of increased $\sqrt{n_g}$ and the number of degrees of freedom in the t distribution. Given a fixed n_g, CI widths are reasonably consistent across different $\hat{n}$. Remarkably, the lowest CI widths obtained represent less than 0.5% of $f(\widehat{U})$.

Finally, we consider the variability over multiple trials of MRP results, given fixed $\hat{n}$ and n_g. Such variability is generated by varying the random seed used to order the scenarios prior to partitioning them into the baseline and validation groups. The standard deviations of $f(\widehat{U})$ reported in Table 2-9 indicate that the stability of the MRP increases—as expected—with growth in $\hat{n}$; taking a larger proportion of the N = 1000 scenarios in Ω will necessarily decrease variability. However, even with modest n̂, the standard deviations are comparatively small, indicating less than 5% deviation from the single-trial MRP result. The variability in CI widths obtained by the MRP also rapidly decreases as n̂ is increased, roughly leveling off once $\hat{n} \geq 140$. Fundamentally, the five-trial MRP results indicate that the most accurate results are obtained using a large number of scenarios to form the baseline solution $\hat{U}$, yielding minimal negligible variability in both $f(\hat{U})$ and CI width.

Overall, the results presented in Table 2-9 illustrate that very high-quality estimates of the minimal total cost for the GEP-EC optimization model can be obtained by using surprisingly small numbers of scenarios. Specifically, the computed CI widths range from less than 0.2% to roughly 3.5% of the baseline solution cost $f(\hat{U})$. Given the nature of long-term planning models—particularly the modeling assumptions employed and the range of uncertainties not explicitly considered—such tight confidence intervals strongly suggest that further efforts should be focused on improving the optimization model and uncertainty characterization, rather than using larger scenario samples in the existing GEP-EC model.

Table 2- 10 Results of Applying the Multiple Replication Procedure (MRP) to the GEP-CVAR Optimization Model. The MRP Input Parameters are Reported in the Columns Labels $\hat{n}$, n_g, and n. Outputs are Reported for the MRP as well as Summary Statistics taken over Five MRP Trials to Verify the Stability of the Procedure.

$\hat{n}$	n_g	n	MRP		Verification: five MRP trials	
			$f(\hat{U})$	95% CI	$f(\hat{U})$ std. dev.	95% CI std. dev.
70	2	465	23.341	0.265	0.70	1.97
70	5	186	23.341	0.365	0.70	1.78
70	10	93	23.341	0.509	0.70	1.93
70	20	46	23.341	0.784	0.70	1.93
70	40	23	23.341	1.260	0.70	1.98

Table 2- 10 (continued)

140	2	430	23.575	0.113	0.55	1.36
140	5	172	23.575	0.199	0.55	1.19
140	10	86	23.575	0.349	0.55	1.15
140	20	43	23.575	0.635	0.55	1.12
140	40	21	23.575	1.240	0.55	1.26
280	2	360	23.534	0.338	0.50	0.25
280	5	144	23.534	0.454	0.50	0.22
280	10	72	23.534	0.557	0.50	0.17
280	20	36	23.534	0.975	0.50	0.19
280	40	18	23.534	1.588	0.50	0.21
420	2	290	23.630	0.487	0.38	0.29
420	5	116	23.630	0.461	0.38	0.28
420	10	58	23.630	0.755	0.38	0.18
420	20	29	23.630	1.113	0.38	0.23
420	40	14	23.630	1.817	0.38	0.18
560	2	220	23.907	0.150	0.13	0.26
560	5	88	23.907	0.282	0.13	0.15
560	10	44	23.907	0.548	0.13	0.16
560	20	22	23.907	0.825	0.13	0.18
560	40	11	23.907	1.551	0.13	0.19

2.7.5 Conditional Value-at-Risk Minimization

The results obtained for executing the MRP on our GEP-CVAR optimization model are shown in Table 2-10. Considering the results for an individual MRP trial, we again observe stability in $f(\widehat{U})$---despite the more sensitive nature of the optimization metric. However, the associated confidence intervals are substantially wider than those observed for the GEP-EC model. On average, the solution costs are approximately 40% larger than the GEP-EC results, given fixed $\hat{n}$ and n_g. While the growth in cost is necessary, the magnitude of growth is highly problem-dependent. Relative to other studies involving planning models in unrelated domains, e.g., see [69], the increase is modest. We observe trends in MRP parameter-result correlations that are similar in most cases to those observed for GEP-EC. Specifically, CI widths are monotonically increasing functions of n_g given a fixed $\hat{n}$.

The largest difference between the single-trial GEP-EC and GEP-CVAR MRP results are the computed confidence interval widths—which range from approximately 1% of the base cost to over 9% of the base solution cost. However, even with a limited number of scenarios, parameterizations of the MRP with modest $\hat{n}$ and small n_g indicate that the optimality gaps associated with the baseline solutions are relatively small; i.e., within a few percent.

Next, we consider variability of MRP trials given fixed $\hat{n}$ and n_g. The variability in both $f(\widehat{U})$ and the confidence interval widths are modestly higher than those observed for GEP-EC. The increase can be attributed to the sensitivity of CVaR; there are comparatively few high-cost scenarios, and the results are as a consequence sensitive to the distribution of those high-cost scenarios among the baseline and validation scenario groups. Overall, the increased variability in MRP results indicates that GEP-CVAR solution costs should be interpreted more carefully than those for GEP-EC; i.e., that there is a significant risk of deviation from the computed CVaR cost metric.

2.7.6 The Structural Similarity of Solutions

Confidence intervals on optimal solution costs are important pieces of information for decision-makers, but they only capture one dimension of “stability” in a solution. A complementary dimension of solution stability considers the degree to which the solutions themselves differ, independent of any variability that may occur in the objection function.

In order to make statements regarding comparisons of many solutions, a metric in the space of solutions is needed. As we shall see, a metric that takes into account covariances will require some degree of dimension reduction in the metric space. Suppose our solutions are represented by column vectors of length *p*.

The Euclidean similarity or “distance” metric for two solutions *x* and *y* can be written as $d_I(x,y) = \sqrt{(x-y)^T I^{-1}(x-y)}$, where *I* is the *p* by *p* identity matrix. The Euclidean metric is very unsatisfying for quantifying the difference between solutions to the GEP-EC and GEP-CVAR optimization models, because it ignores (for example) the fact that a difference of 2 CT generating units in a pair of solutions is conceptually (to domain experts) much less significant than a difference of 2 nuclear generating units. This situation could be remedied if we somehow knew an appropriate variance for each generator type. With this information, we could form a $p \times p$ matrix *S* with the standard deviations on the diagonal, and instead use the distance metric $d_S(x,y) = \sqrt{(x-y)^T S^{-1}(x-y)}$.

A more subtle issue is the consideration of correlations that would be expected between generator selections. To address this concern, we consider the Mahalanobis distance [43] metric, given by:

$$d_{\Sigma}(x,y) = \sqrt{(x-y)^T \Sigma^{-1}(x-y)}$$

where $\sum$ is a $p \times p$ positive definite matrix. In particular, if $\sum$ is an appropriate covariance matrix, then we have a distance that is scaled for both variance and covariance. The Mahalanobis metric is widely used to characterize distances between vectors and groups of vectors (e.g., see [6, 48].) There are some theoretical properties of the Mahalanobis metric that can provide quantitative insights under specific conditions. In particular, if x has a multi-variate normal distribution $(\mu, \sum)$ then $d_{\sum}^2(x, \mu)$ obeys a χ^2 distribution with p degrees of freedom.

When comparing solutions to generation expansion planning models, the investment variables are generally the drivers of structural differences of interest to decision makers. To reduce the problem dimension, we consider the total number of each type of generator over time prescribed by a particular solution U^*; i.e., we will examine the vector whose components are formed from the components $x_g = \sum_y U_{gy}^*$, for all $g \in G$.

Table 2-11 reports statistics for the Mahalanobis distances for both the GEP-EC and GEP-CVAR optimization models. The statistics are computed using the five $\widehat{U}$ reference solutions obtained across replications of the MRP procedure, for the range of $\hat{n}$. The aggregation procedure is then as follows. For each $\hat{n}$ and each optimization objective (GEP-EC and GEP-CVAR), we form a "population" using the $\widehat{U}$ solutions associated with the other $\hat{n}$ values. Considering a fixed $\hat{n}$, twenty replicates are then available for use in estimating the mean and covariance. Each solution is represented by investments in each generator type aggregated over time, as described previously. In our particular study, there is no variance in the number of CC and BaseLoad generator types, so they are ignored, leaving vectors with p = 4. For reference, the 95th percentile of a χ^2 with four degrees of freedom is 9.49.

Table 2- 11 The Average Pair-wise Mahalanobis Distances for the Generation Expansion Planning Investment Decisions Aggregated over Time, for Both the GEP-EC and GEP-CVAR Optimization Objectives

$\hat{n}$	Objective Function	
	EC	CVAR
70	17.3	12.1
140	3.7	7.9
280	5.8	4.5
420	3.0	2.5
560	5.3	5.4

The solutions generated for $\hat{n}$ =70 differ significantly from the solutions for other $\hat{n}$ for both optimization objectives, which is not unexpected given the small number of samples used to form the reference solutions $\widehat{U}$. However, the average pair-wise solution distance rapidly drops as $\hat{n}$ grows, albeit somewhat more slowly for GEP-CVAR than for GEP-EC (specifically, the drop at $\hat{n}$ =140 is not as punctuated in GEP-CVAR as for GEP-EC. However, in both cases the differences are very small once $\hat{n} \geq 280$.

A key question is: Are the solutions obtained for GEP-CVAR statistically and qualitatively different than those obtained for GEP-EC. To get a quantitative measure of the distance we take as our population the solutions for GEP-EC and drop those for $\hat{n}$= 70 since they seem to be different. That is, we take as our population the solutions from MRP replicates for GEP-EC with $\hat{n} \geq 140$, so we have 20 vectors of aggregated investment decisions that can be used to estimate a mean and covariance for the sample. The smallest Mahalanobis distance to the aggregated investment decisions for any of the GEP-CVAR solutions was 108 and the average distance to those solutions was 281. Hence, we have strong evidence based on the metric that the GEP-CVAR solutions are structurally much different from the GEP-EC solutions.

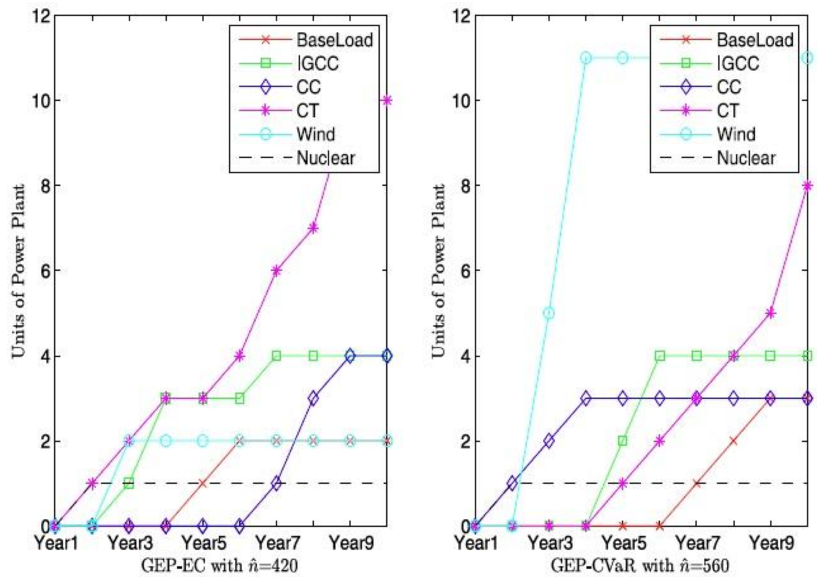

Figure 2- 10 The GEP-EC Solution with $\hat{n}$ = 420 and the GEP-CVAR Solution with $\hat{n}$ = 560

To illustrate the differences between solutions obtained using the GEP-EC and GEP-CVAR optimization models, we make use of solutions with tight confidence intervals. The lowest CI widths obtained for GEP-EC represent 0.12% of the cost estimate with $\hat{n}$ = 420, n_g = 2 and n = 290. The total cost is 17.337 billion USD and the 10-year optimal expansion plan is shown in left hand side of Fig. 2-10. For GEP-CVAR, the lowest CI widths obtained represent 0.36% of the point estimate with $\hat{n}$ = 560, n_g = 2 and n = 220. The CVaR value is 23.907 billion USD and the 10-year optimal expansion plan is shown in right hand side of Fig. 2-10. The GEP-CVAR planning decision is a bit different from the GEP-EC plan shown in the left side of Fig. 2-10. It greatly increases the wind power capacity, which has almost no operational costs, and decreases the number of gas-fired CT plants. The latter occurs because the GEP-CVAR solution avoids the risk of high gas prices in the future. The total cost of the planning decision with more CT units is subject to the natural gas price volatility, as well as the high electricity demand. Besides, GEP-CVAR should be more costly than GEP-EC since it is more likely to suggest building sufficient generation capacity to ensure meeting high demand in a small percentage of extremely high-cost scenarios.

Overall, our analysis of solution stability reinforces the general conclusions obtained with the MRP procedure. In the case of the GEP-EC optimization model, the $\widehat{U}$ reference solutions obtained using a small number of scenarios are stable in terms of both cost and solution structure. In the case of GEP-CVAR, the solution stability results indicate that even with larger-than-desirable confidence intervals on solution cost, the underlying structural characteristics of the resulting solutions are not significantly different. In other words, the cost uncertainty does not appear to translate into significant differences in the recommendations provided to a potential decision maker.

2.8 Conclusions

We have introduced a novel formulation of the generation expansion planning problem, with the goal of determining the types and quantities of available generator types to build during each year of a long-term planning horizon. We formulate the problem as a two-stage mixed-integer stochastic program, considering minimization of both expected cost and the Conditional Value-at-Risk. The optimization models incorporate two sources of uncertainty regarding the future: natural gas fuel prices and demand for electricity. We propose a stochastic process model describing the evolution of these parameters, and use this model to construct a scenario tree for input to the stochastic program. As a case study, we introduce a planning problem based on the US Midwest generation infrastructure.

Direct solution of a large-scale, mixed-integer stochastic generation expansion planning problem is computationally prohibitive. Consequently, the full scenario tree must be down-sampled for purposes of computational tractability. This downsampling leads to a variety of computational issues, which must be addressed to accurately represent the approximated solution to a decision maker. Specifically, it is necessary to quantify the stability of the approximate solution—error is necessarily present due to the approximation of a stochastic process by a scenario tree small enough to facilitate tractable solution of the corresponding stochastic program. To address the issue of solution cost stability, we apply the Multiple Replication Procedure of Mak, Morton, and Wood [44] to compute confidence intervals on the optimality gap. Our results indicate that the optimality gaps obtained when minimizing expected cost are very small, while those obtained when minimizing Conditional Value-at-Risk are large enough to be of concern. Independent of cost, the solutions obtained under different samplings of scenarios are structurally very similar for both optimization metrics. This suggests that the presence of even moderate optimality gaps for solution cost have little impact on the final solution recommended to a decision-maker. Our results indicate that limited-scale sampling of a very large scenario tree is sufficient in our particular problem to yield high-accuracy solutions.

Acknowledgements

Support for this work at Iowa State University was provided by its Electric Power Research Center. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94-AL85000. The Sandia and University of California Davis authors were funded in part by the Department of Energy's Office of Science.

References

1. Ahmed, S.: Introduction to stochastic integer programming. http://www.stoprog.org/SPIntro/intro2sip.html (2009)

2. Ahmed, S., Sahinidis, N.: An approximation scheme for stochastic integer programs arising in capacity expansion. Oper. Res. 51(3), 461–471 (2003)

3. Ahmed, S., King, A., Parija, G.: A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. J. Glob. Optim. 26, 3–24 (2003)

4. Albornoz, V., Benario, P., Rojas,M.: A two-stage stochastic integer programming model for a thermal power system expansion. Int. Trans. Oper. Res. 11, 243–257 (2004)

5. AMPL: A modeling language for mathematical programming. http://www.ampl.com (2010)

6. Anderson, T.: An Introduction to Multivariate Statistical Analysis, 3rd edn. Wiley, New York (2003)

7. Annual Energy Outlook 2009: Tech. rep., Energy Information Administration, Department of Energy (2009)

8. Booth, R.: Optimal generation planning considering uncertainty. IEEE Trans. Power Appar. Syst. PAS-91, 70–77 (1972)

9. Chuang, A., Wu, F., Varaiya, P.: A game-theoretic model for generation expansion planning: problem formulation and numerical comparisons. IEEE Trans. Power Syst. 16(4), 885–891 (2001)

10. COIN-OR: COmputational INfrastructure for Operations Research. http://www.coin-or.org (2010)

11. CPLEX: http://www.cplex.com (2010)

12. DeMeo, E., Grant,W., Milligan,M., Schuerger,M.:Wind plant generation. IEEE Power EnergyMag. 3(6), 38–46 (2005)

13. Denny, E., O'Malley, M.: Wind generation power system operation, and emissions reduction. IEEE Trans. Power Syst. 21(1), 341–347 (2006)

14. Dentcheva, D., Romisch, W.: Optimal power generation under uncertainty via stochastic programming. Stoch. Program. Methods Tech. Appl., pp. 22–56 (1998)

15. Doherty, R., Outhred, H., O'Malley, M.: Establishing the role that wind generation may have in future generation portfolios. IEEE Trans. Power Syst. 21(3), 1415–1422 (2006)

16. Dupacova, J., Consigli, G.,Wallace, S.: Scenarios for multistage stochastic programs. Ann. Oper. Res. 100, 25–53 (2000)

17. Eichhorn, A., Heitsch, H., Romisch, W.: Stochastic optimization of electricity portfolios: scenario tree modeling and risk management. In: Rebennack, S., Pardalos, P., Pereira, M., Iliadis, N. (eds.) Handbook of Power Systems, vol. II, pp. 405–432. Springer, Berlin (2010)

18. Electric Generation Expansion Analysis System (EGEAS): Tech. rep., Electric Power Research Institute (2009)

19. Firmo, H., Legey, L.: Generation expansion planning: an iterative genetic algorithm approach. IEEE Trans. Power Syst. 17(3), 901–906 (2002)

20. Fukuyama, Y., Chiang, H.: A parallel genetic algorithm for generation expansion planning. IEEE Trans. Power Syst. 11(2), 955–961 (1996)

21. Garcia-Gonzalez, J., de la Muela, R., Santos, L., Gonzalez, A.: Stochastic joint optimization of wind generation and pumped-storage units in an electricity market. IEEE Trans. Power Syst. 23(2), 460–468 (2008)

22. Growe-Kruska, N., Heitsch, H., Romisch, W.: Scenario reduction and scenario tree construction for power management problems. In: IEEE Bologna Power Tech. Conference, Bologna, Italy (2003)

23. Holmes, D.: A collection of stochastic programming problems. Tech. rep. (1994)

24. Høyland, K., Wallace, S.: Generating scenario tree for multistage decision problems. Manag. Sci. 47(2), 295–307 (2001)

25. http://www.eia.doe.gov/international (2009)

26. http://www.eia.doe.gov/cneaf/electricity/epm/table1_1.html (2009)

27. http://en.wikipedia.org/wiki/Clean_coal#Clean_coal_technology (2009)

28. http://en.wikipedia.org/wiki/Renewable_energy_development#Production_tax_credits (2009)

29. http://www.eia.doe.gov/cneaf/electricity/epm/table1_1.html (2009)

30. http://www.gesys.com/pro-mod.htm (2009)

31. http://www.plexos.info/wiki (2009)

32. http://www.plexossolutions.com (2009)

33. http://en.wikipedia.org/wiki/Geometric_Brownian_motion (2009)

34. http://www.midwestmarket.org (2009)

35. International Energy Outlook 2009: Tech. rep., Energy Information Administration, Department of Energy (2009)

36. Johnson, N., Kotz, S.: Continuous Univariate Distributions, 2nd edn. Wiley Series in Probability and Mathematical Statistics, vol. 1. Wiley, New York (1994). Chap. 3: Lognormal distribution

37. Joint Coordinated System Planning Report 2008: Tech. rep. (2009)

38. Kanna, S., Slochanal, S., Padhy, N.: Application and comparison of metaheuristic techniques to generation expansion planning problem. IEEE Trans. Power Syst. 20(1), 466–475 (2005)

39. Karaki, S., Chaaban, F., Al-Nakhl, N., Tarhini, K.: Power generation expansion planning with environmental consideration for Lebanon. Int. J. Electr. Power Energy Syst. 24, 611–619 (2002)

40. Karki, R., Billinton, R.: Cost-effective wind energy utilization for reliable power supply. IEEE Trans. Energy Convers. 19(2), 435–440 (2004)

41. Laurent, A.: A scenario generation algorithm for multistage stochastic programming: application for asset allocation models with derivatives. Ph.D. thesis, University of Lugano, Lugano, Switzerland

42. Lund, H.: Large-scale integration of wind power into different energy systems. Energy 30(13), 2402–2412 (2005)

43. Mahalanobis, P.: On the generalised distance in statistics. In: Proceedings of the National Institute of Sciences of India, vol. 2, pp. 49–55 (1936)

44. Mak, W., Morton, D., Wood, R.: Monte Carlo bounding techniques for determining solution quality in stochastic programs. Oper. Res. Lett. 24, 47–56 (1999)

45. Malcolm, S., Zenios, S.: Robust optimization for power systems capacity expansion under uncertainty. J. Oper. Res. Soc. 45(9), 1040–1049 (1994)

46. Marathe, R., Ryan, S.: On the validity of the geometric Brownian motion assumption. Eng. Econ. 50(2), 159–192 (2005)

47. McCalley, J.: Introduction to electric systems expansion planning. http://home.eng.iastate.edu/~jdm/ee590/PlanningIntro.pdf (2008)

48. McLachlan, G.: Discriminant Analysis and Statistical Pattern Recognition. Wiley Interscience, New York (1992)

49. Meza, J., Yildirim, M., Masud, A.: A model for the multiperiod multiobjective power generation expansion problem. IEEE Trans. Power Syst. 22(2) (2007)

50. Milligan, M.: Measuring wind plant capacity value. Tech. rep., National Renewable Energy Laboratory, Colorado (1996)

51. Milligan,M.: Variance estimates of wind plant capacity credit. Tech. rep., National Renewable Energy Laboratory, Colorado (1996)

52. Milligan, M.: Modeling utility-scale wind power plants. Part 2: Capacity credit.Wind Energy 3, 106–206 (2000)

53. Mo, B., Hegge, J., Wangensteen, I.: Stochastic generation expansion planning by means of stochastic dynamic programming. IEEE Trans. Power Syst. 6(2), 662–668 (1991)

54. Mulvey, J., Vanderbei, R., Zenios, S.: Robust optimization of large-scale systems. Oper. Res. 43(2), 264–281 (1995)

55. Paulun, T., Haubrich, H.J.: Long-term and expansion planning for electrical networks considering uncertainties. In: Rebennack, S., Pardalos, P., Pereira, M., Iliadis, N. (eds.) Handbook of Power Systems, vol. I, pp. 391–408. Springer, Berlin (2010)

56. PySP: PySP: Python-based stochastic programming. https://software.sandia.gov/trac/coopr/wiki/PySP (2011)

57. Rockafellar, R.: Coherent approaches to risk in optimization under uncertainty. In: Tutorials in Operation Research. INFORMS Annual Meeting (2007)

58. Rockafellar, R., Uryasev, S.: Optimization of Conditional Value-at-Risk. J. Risk 2, 21–41 (2000)

59. Rockafellar, R.T., Wets, R.J.B.: Scenarios and policy aggregation in optimization under uncertainty. Math. Oper. Res. 16(1), 119–147 (1991)

60. Ross, S.: Brownian motion and stationary processes. In: Introduction to Probability Models, 9th edn. Elsevier, Amsterdam (2007). Chapter 10

61. Ryan, S., McCalley, J., Woodruff, D.: Long term resource planning for electric power systems under uncertainty. Technical report, Iowa State University (2011)

62. Sahinidis, N.: Optimization under uncertainty: state-of-the-art and opportunities. Comput. Chem. Eng 28, 971–983 (2004)

63. Schultz, R., Tiedemann, S.: Conditional Value-at-Risk in stochastic programs with mixed-integer recourse. Math. Program. 105(2–3), 365–386 (2005)

64. Shapiro, A., Dentcheva, D., Ruszczynski, A.: Lectures on Stochastic Programming: Modeling and Theory. Society for Industrial and Applied Mathematics, Philadelphia (2009)

65. Slyke, R.M.V.,Wets, R.J.: L-shaped linear programs with applications to optimal control and stochastic programming. SIAM J. Appl. Math. 17, 638–663 (1969)

66. Takriti, S., Ahmed, S.: On robust optimization of two-stage systems. Math. Program., Ser. A 99, 106–126 (2004)

67. Voorspools, K., D'haeseleer,W.: An analytical formula for the capacity credit of wind power. Renew. Energy 31, 45–54 (2006)

68. Voropai, N., Ivanova, E.: Multi-criteria decision analysis techniques in electric power system expansion planning. Int. J. Electr. Power Energy Syst. 24, 71–78 (2002)

69. Watson, J., Murray, R., Hart, W.: Formulation and optimization of robust sensor placement problems for drinking water contamination warning systems. J. Infrastruct. Syst. 15(4), 330–339 (2009)

70. Watson, J.P., Woodruff, D., Hart, W.: Modeling and solving stochastic programs in Python. Math. Program. Comput. (2011) (to appear)

CHAPTER 3 CAPACITY EXPANSION IN THE INTEGRATED SUPPLY NETWORK FOR AN ELECTRICITY MARKET

Published in *IEEE Transactions on Power Systems*

Shan Jin, Sarah M. Ryan

Abstract

Constraints in fuel supply, electricity generation and transmission interact to affect the welfare of strategic generators and price-sensitive consumers. We consider a mixed integer bilevel programming model in which the leader makes capacity expansion decisions in the fuel transportation, generation, and transmission infrastructure of the electricity supply network to maximize social welfare less investment cost. Based on the leader's expansion decisions, the multiple followers including the fuel suppliers, ISO and generation companies simultaneously optimize their respective objectives of cost, social welfare, and profit. The bilevel program is formulated as a mathematical program with complementarity constraints. The computational challenge posed by the discrete character of transmission expansions has been managed by multiple model reformulations. A lower bound provided by a nonlinear programming reformulation increases the efficiency of solving a binary variable reformulation to global optimality. A single-level optimization relaxation serves as a competitive benchmark to assess the effect of generator strategic operational behavior on the optimal capacity configuration.

3.1 Introduction

We consider an integrated supply network for an electricity market including fuel transportation, electricity transmission and individual generation companies. The decision makers in each level of the supply network have distinct objectives to optimize. The fuel suppliers, who deliver the fuel to the power generation companies, want to minimize their fuel transportation cost that includes both the fuel cost and fuel delivery cost. The independent system operator (ISO) who settles the locational marginal prices (LMPs) and dispatches the electricity through the transmission network, aims to maximize the total welfare of both the sellers and the buyers of electricity in a wholesale market. The individual generation companies, who buy the fuel, generate electricity and sell it at the LMPs, wish to maximize their profits. All of these decision makers optimize simultaneously in the electricity market subject to capacity constraints.

The overall welfare of market participants could be increased by capacity expansions to relieve constraints. Expansions at different levels and locations in the supply network could increase the availability of low-cost fuel, enable higher utilization of efficient generation resources, and level out the LMPs. These decisions are the responsibility of the facility owners, who naturally determine capacity investments to achieve their own objectives. However, such decentralized expansion decisions may not be optimal for the whole integrated electricity supply system. In this paper, we examine the decisions that a leader would make on behalf of the overall system, to maximize the total welfare less the total investment cost. The results of optimizing from this global perspective reveal the interactions among constraints at different levels and identify bottlenecks in the integrated supply network. They could be used as a target in the development of consistent incentives or regulations to encourage the lower level players to make individual decisions that more closely approximate the global optimum for the overall system.

Considering both the investment and operational levels of decision-making, we propose a mixed integer bilevel programming model. The upper level leader makes decisions to expand the capacity of the integrated supply system. Once the capacity expansion decisions are made, lower level decision makers including the fuel suppliers, the ISO and the generation companies simultaneously make their optimal operating decisions to realize their respective objectives. We set up a simple two-period model in which capacity investments are made in the first period based on equivalent hourly costs, and the system is operated in the second period, which represents a typical hour in a future scenario. The model could be elaborated to account for sequential expansions over time and/or multiple future scenarios of demand, fuel costs, and other uncertain conditions.

Given the capacity decisions in the integrated supply network, the simultaneous optimization of the lower level sub-problems, which mutually interact, leads to an equilibrium problem. It has been thoroughly studied and solved as a mixed complementarity problem (MCP) [1] and validated by comparison with a computational agent simulation [2]. In the restructured wholesale electricity market, the prices submitted by the generator companies are determined according to their marginal costs, which depend mostly on the fuel costs [3]. Here we assume that the fuel suppliers, represented for simplicity by a single fictional fuel dispatcher, decide the quantities of the fuel shipped over various routes to minimize the cost of satisfying demands of all of the generator companies. The ISO manages the electricity wholesale market where the buyer (inverse) demand functions are linear functions of the LMPs, and makes the electricity dispatch decision. The individual generation companies, considered as Cournot competitors, pay attention to the price differentials among the electricity nodes and determine their electricity quantities to sell under a type of bounded rationality [4].

In general, mixed integer bilevel programming (MIBLP) problems are hard to solve. Moore and Bard [5] presented the challenge of the MIBLP problem and proposed a series of heuristics to efficiently find a good feasible solution. DeNegre et al. [6] further proposed a branch and cut method to improve the branch and bound algorithm in [5]. Colson et al. [7] reviewed methodologies and applications of bilevel programming problems and described their connections with mathematical programs with equilibrium constraints (MPECs). The bilevel programming technique was introduced to model the integrated system of multiple participants usually with different objective functions. Many applications involving restructured electricity markets have formulated the ISO market clearing problem as the lower level. Hu and Ralph [8] modeled a game among the consumers and generators submitting the bids to the ISO in the upper level. There are also formulations considering the bidding strategy of the generation companies in the upper level [9, 10]. Generation and transmission expansion decisions by individual participants can also be modeled as the upper level with the lower level representing the market outcomes [11, 12].

In our model, the integer variables appear only in the upper level. Therefore, the Karush–Kuhn–Tucker (KKT) optimality conditions can still be applied to the lower level problems. Upon applying the KKT conditions to all the sub-problems, the bilevel problem becomes a mathematical program with complementarity constraints (MPCC). The MPCC is a special case of MPEC, which has attracted great attention for the past decade as more and more engineering and economic applications involve equilibrium modeling. Ferris and Pang [13] gave a comprehensive summary of the engineering and economic applications of MPCC and the available solution algorithms. Its solution requires an equivalent reformulation of complementarity constraints and global convergence the solution can be guaranteed only under certain conditions. Hu et al. [14] presented a methodology to find the global optimal solution of a linear program with linear complementarity constraints by reformulating the constraints with binary variables.

The contributions of this paper are fourfold: (1) We investigated a fuel transportation, generation and transmission expansion problem of an integrated electricity supply system in which an equilibrium in a restructured market is reached by the fuel suppliers, generator companies and ISO solving simultaneous and interdependent optimization problems. Instead of letting each market player make his own expansion decision, we find the optimal expansion decision for the whole integrated system from the global perspective. (2) We incorporated discrete transmission expansion decisions by using binary variables in the direct current optimal power flow constraints. (3) The problem is formulated as a bilevel programming problem. The challenge posed by the discrete decision variables makes it difficult to achieve global optimality. We provided three problem reformulations to bound the objective and find a global optimum. (4) The model and solution procedure are illustrated by a small case study that shows how the global expansion decision affects the LMPs of each node in the transmission grid, the buyers' surplus, the sellers' surplus and the transmission rents.

3.2 Model Formulation

Table 3- 1 Sets of Nodes and Arcs

Set	Description	Indices
N	Electricity nodes	*i, j*
F	Fuel supply nodes	*g*
L	Transmission lines	*ij*
A	Fuel supply lines	*gj*
N_g	Set of electricity nodes supplied by fuel supply node *g*	*j*
F_j	Set of fuel supply nodes supplying the electricity node *j*	*g*

Table 3- 2 Decision Variables

Decision Variable		Description
Upper Level	***nU***	Fuel transport capacities after expansion
	nV	Generation capacities after expansion
	z	Binary decision variables for new transmission lines
	q	Demand satisfied at electricity nodes
Lower Level	***x***	Quantities of fuel delivered
	$\boldsymbol{\theta}$	Voltage angles at electricity nodes
	f	Electricity flows on transmission lines

Table 3- 2 (continued)

y	Generation amounts at electricity nodes
η	(Scalar) price at the reference electricity node

Table 3- 3 Parameters

Parameter	Description
a	Intercepts of electricity demand prices as linear functions of quantities
b	Slopes of electricity demand prices as linear functions of quantities
c	Costs per MWh-equivalent of fuel transported
fc	Investment cost for fuel transportation network expansion
gc	Investment costs for generation expansion
tc	Investment costs for transmission line expansion
θ^{max}	Maximum values for voltage angles
θ^{min}	Minimum values for voltage angles
ϕ	Nodal electricity price premia at electricity nodes
π	Nodal prices for fuel delivered to generators
V	Generation capacities at electricity nodes
U	Capacities of fuel supply arcs
K	Capacities of transmission lines
W	Quantities of fuel available at fuel supply nodes
B	Susceptances of transmission lines

We formulate a bilevel capacity expansion problem of an integrated electricity supply network, where we also optimize the sub-problems of fuel suppliers, generators and the ISO in the lower level.

Table 3-1 indicates the sets of nodes and arcs of the integrated electricity supply network. Tables 3-2 and 3-3 respectively give the notation for both decision variables and parameters of the model. All fuel quantities are expressed in MWh equivalents. The variable η is a scalar and all the other variables and parameters are vectors. All fuel quantities are expressed in MWh equivalents. Appendix 3.A shows how to incorporate heat rates and efficiencies of converting fuels to electricity.

3.2.1 Mixed Integer Bilevel Program (MIBLP)

The objective function of the upper level is to maximize the social welfare including the buyers' surplus, power producers' surplus and transmission rents, less the investment costs of expansions in fuel transportation, transmission network and power generation capacities. We expand capacity of the existing assets for both the fuel network and generation. For the transmission network, we assume the capacity expansions are realized by building new lines selected from a set of candidates. All investment costs are linear.

$$\max_{nU,nV,z,q,x} \sum_{j}\left(\frac{1}{2}b_j q_j^{\,2} + a_j q_j\right) - \sum_{gj\in A} c_{gj}x_{gj} - \sum_{gj\in A} fc_{gj}\left(nU_{gj} - U_{gj}\right) - \sum_{j\in N} gc_j\left(nV_j - V_j\right) - \sum_{ij\in L} tc_{ij}K_{ij}z_{ij} \tag{3-1}$$

The electricity demand at each electricity node is defined as a linear function of the electricity nodal price. The intercept and slope of the inverse demand function are respectively a_j and b_j with $b_j < 0$ at electricity node j. Fixed (inelastic) demands can also be incorporated in the lower level model [15].

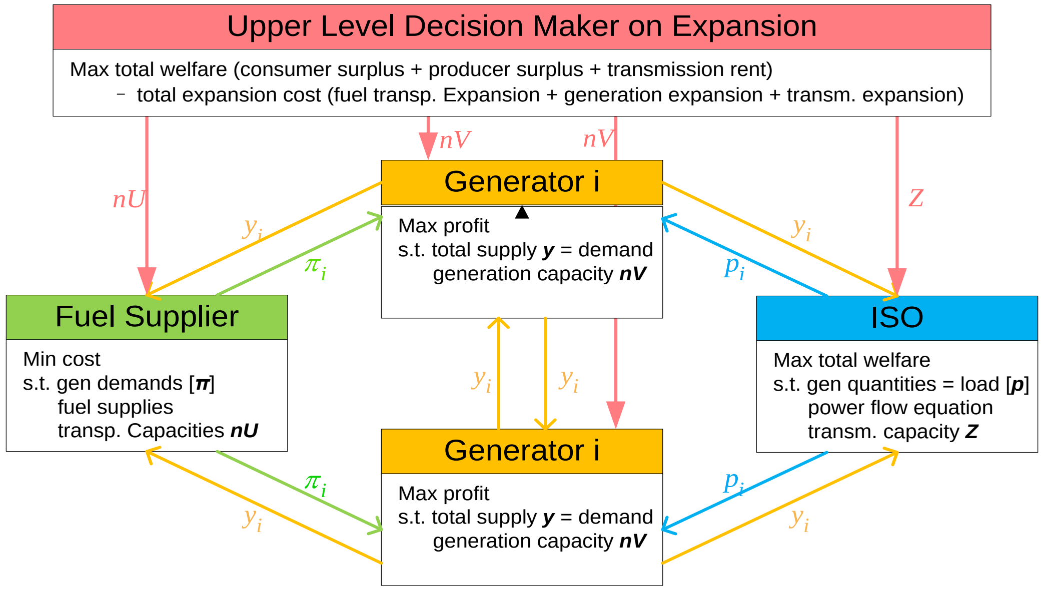

Figure 3- 1 The Bilevel Program with the Interacting Lower Level Optimization Problems

For the lower level optimization problems, we consider the three major market participants of the electricity supply network consisting of fuel suppliers, ISO and generators, who all optimize under the same electricity market conditions. The fuel dispatcher minimizes the fuel transportation cost delivered to the electricity producers; the ISO maximizes the social welfare of participants in the wholesale electricity market; and the generators maximize their profits from selling the electricity.

However, these three different optimization problems interact with each other as shown in Fig. 3-1. The dual variables π in eqn. (3-4) of the fuel dispatcher's problem are the marginal cost parameters in the generators' problems. The dual variables ϕ in eqn. (3-16) of the ISO's problem are the electricity nodal price premium parameters in the generator problems. The electricity quantities determined by the generators are the fuel demand parameters in the fuel dispatcher's problem and the electricity supply quantities for the ISO's problem.

Given the generation amount y_j required at each electricity node *j*, the fuel dispatcher's optimization problem is:

Fuel Dispatcher's decision problem

$$\min_{x \geq 0} \quad \sum_{gj \in A} c_{gj} x_{gj} \tag{3-2}$$

$$s.t. \; -\sum_{j \in N_g} x_{gj} \geq -W_g, \quad \forall g \in F \quad [\omega_g \geq 0] \tag{3-3}$$

$$\sum_{g \in F_j} x_{gj} = y_j, \quad \forall j \in N \quad [\pi_j] \tag{3-4}$$

$$-x_{gj} \geq -nU_{gj}, \quad \forall gj \in A \quad [\rho_{gj} \geq 0] \tag{3-5}$$

$$x_{gj} \geq 0 \tag{3-6}$$

The fuel dispatcher aims to minimize the transportation cost subject to the constraints of fuel supply capacity (3-3), electricity demand (3-4) and the fuel transportation arc capacity (3-5).

The market operator, ISO, seeks to maximize the social welfare based on the direct current optimal power flow (DCOPF) model with the full-structured form [16, 17]. Given the generation amounts y_j at the electricity nodes, the ISO's decision problem is:

ISO's decision problem

$$\max_{q,\theta,f} \quad \sum_j \left(\frac{1}{2} b_j q_j^{\,2} + a_j q_j \right) \tag{3-7}$$

$$s.t.\ \ q_j + \sum_i f_{ji} - \sum_i f_{ij} = y_j,\ \ \forall j \in N \quad [p_j] \tag{3-8}$$

$$\theta_j \le \theta_j^{\max},\ \ \forall j \in N \quad [\alpha_j^+ \ge 0] \tag{3-9}$$

$$-\theta_j \le -\theta_j^{\min},\ \ \forall j \in N \quad [\alpha_j^- \ge 0] \tag{3-10}$$

$$f_{ij} - B_{ij}(\theta_i - \theta_j) - \left(1 - z_{ij}\right) M_{ij} \le 0,\ \ \forall ij \in L \ \ [\gamma_{ij}^+ \ge 0] \tag{3-11}$$

$$B_{ij}(\theta_i - \theta_j) - f_{ij} - \left(1 - z_{ij}\right) M_{ij} \le 0,\ \ \forall ij \in L \ \ [\gamma_{ij}^- \ge 0] \tag{3-12}$$

$$f_{ij} \le z_{ij} K_{ij},\ \ \forall ij \in L \quad [\lambda_{ij}^+ \ge 0] \tag{3-13}$$

$$-f_{ij} \le z_{ij} K_{ij},\ \ \forall ij \in L \quad [\lambda_{ij}^- \ge 0] \tag{3-14}$$

The ISO's decision problem, based on the full structure DCOPF model of [17], is equivalent to the reduced structure DCOPF model in [1].

Condition (3-8) represents the flow balance at each electricity node. Constraints (3-9) and (3-10) give the bounds on each voltage angle. Equations (3-11) and (3-12) incorporate the physical characteristics of the transmission grid so that the (linearized) power flow equations will be always satisfied. The maximum capacity of each transmission line is enforced by equations (3-13) and (3-14).

Instead of the standard power flow equation $B_{ij}(\theta_i - \theta_j) - f_{ij} = 0$, binary decision variables z and big value M are used in (3-11) and (3-12) to represent discrete investment decisions on new transmission lines in a manner similar to a transmission-switching model [16, 18]. The variable z_{ij} indicates the existence of the transmission line *ij*. If the candidate transmission line has been added, then z_{ij} equals 1, the value of M_{ij} does not matter at all, and the two inequalities are equivalent to the traditional power flow equation. On the other hand, if z_{ij} is 0, then the value of M_{ij} matters. The value of M_{ij} should be large enough to impose no additional constraint on $B_{ij}(\theta_i - \theta_j) - f_{ij}$. We fixed all the z variables corresponding to the lines which exist prior to the expansion decision to be 1. For the candidate transmission lines, their corresponding z variables can be either 1 or 0, to represent building those lines or not.

Regarding the parameters *M*, excessively large values might cause numerical difficulties when solving the problem. One of the assumptions of the DCOPF model is that the voltage angle difference of any transmission line is quite small [17]. Here we adopt the assumption in [16] with upper and lower bounds on θ of ± 0.6. Because the electricity flow f is also bounded by *K*, the quantity $\left| B_{ij}(\theta_i - \theta_j) - f_{ij} \right|$ is bounded by $1.2B_{ij} + K_{ij}$, which therefore represents a sufficiently large value of M_{ij}.

We assume that the electricity wholesale market takes the form of oligopolistic competition. The multiple generators are modeled as Cournot competitors in the electricity wholesale market. Each of them determines its electricity quantity to sell. Besides Cournot model, there are also many other approaches available to model the generator's competition in the electricity market [19]. A more realistic approach is the supply function equilibrium game that allows each firm to submit a bid function with different quantity offered given different market price. However, it suffers from computational inefficiency and multiplicity of equilibrium due to its non-convexity. Another popular approach is Bertrand game in which the producer makes the decision on selling price instead of quantity. Therefore it is more likely to give a similar result as in a competitive electricity market where the prices are all set to the marginal cost. This market behavior does not agree with the oligopolistic competition that we assumed. The Cournot model is not as realistic as the supply function equilibrium model. However, it is much easier to solve and in a long run the market behavior is close to Cournot result [19, 20]. Since the total amount of electricity generated affects the electricity market prices, here we assume that each generator also determines the LMP η at the reference electricity node [4].

The LMP at the reference electricity node is

$$\eta = p_{ref} \tag{3-15}$$

and the price premium at each node is then defined as

$$\phi_j = p_j - \eta \tag{3-16}$$

where the LMPs p are dual variables of the market clearing constraint (3-8).

Given the fuel price π_i and price premium ϕ_i at its node, which are derived from the dual variables of the fuel supplier's decision problem and ISO's decision problem, respectively, the generator's decision problem is:

Generator's decision problem

$$\max_{y_i \geq 0, \eta} \quad (\eta + \phi_i - \pi_i) y_i \tag{3-17}$$

$$s.t.\ y_i - \eta \sum_j \frac{1}{b_j} = \sum_j \frac{\phi_j - a_j}{b_j} - \sum_{j \neq i} y_j, \quad \forall i \in N\ [\beta_i] \tag{3-18}$$

$$y_i \leq nV_i, \quad \forall i \in N \quad [\mu_i \geq 0] \tag{3-19}$$

$$y_i \geq 0 \tag{3-20}$$

Equation (3-18) represents $\sum_i q_i = \sum_i y_i$, the balance of total demand and total generation amount in terms of the residual demand seen by generator i. Constraint (3-19) indicates the maximum generation capacity. The generator's problem can also be extended to take carbon emission regulations into account, as described in Appendix 3.B.

If we explore only the lower level optimization problems by simply ignoring the upper level objective function, the lower level is equivalent to the problem studied by Ryan, et al. [1]. The only difference is the equivalent modification of the ISO's decision problem to incorporate the transmission expansion decision. We verified our equivalent new model with fixed capacities by comparing its numerical results with those in [2]. Because the existence of Nash equilibrium has been proved in [1], it also holds for our lower level problems. To explore the potential multiplicity of equilibria, we solved both the maximization and minimization problems for multiple different objective functions in the numerical instance of Section 3.4 with investment variables fixed, and they all returned with the same equilibrium solution, which suggests that the equilibrium is unique in that instance.

3.2.2 Mathematical Program with Complementarity Constraints (MPCC)

The MPCC problem is to optimize an objective function subject to complementarity constraints that can be expressed with the standardized format $0 \le x \perp f(x) \ge 0$.

The mixed integer bilevel programming problem presented in Section 3.2.1 has integer decision variables only in the upper level. Thus, the lower level optimization problem can be reformulated in terms of complementarity constraints by applying the KKT conditions to each player's optimization problem. This transforms the original bilevel program into an equivalent MPCC with a mixed integer quadratic objective function. The objective function is given by equation (3-1) and the full set of constraints is:

$$0 \le x_{gj} \perp c_{gj} + \omega_g - \pi_j + \rho_{gj} \ge 0, \ \forall gj \in A \tag{3-21}$$

$$\sum_{g \in F_j} x_{gj} = y_j, \ \forall j \in N \tag{3-22}$$

$$0 \le \omega_g \perp W_g - \sum_{j \in N_g} x_{gj} \ge 0, \ \forall g \in F \tag{3-23}$$

$$0 \le \rho_{gj} \perp nU_{gj} - x_{gj} \ge 0, \ \forall gj \in A \tag{3-24}$$

$$a_j + b_j q_j - p_j = 0, \ \forall j \in N \tag{3-25}$$

$$\alpha_j^- - \alpha_j^+ + \sum_{i,ji \in L} B_{ji}\left(\gamma_{ji}^+ - \gamma_{ji}^-\right) - \sum_{i,ij \in L} B_{ij}\left(\gamma_{ij}^+ - \gamma_{ij}^+\right) = 0, \forall j \in N \tag{3-26}$$

$$p_j - p_i - \gamma_{ij}^+ + \gamma_{ij}^- - \lambda_{ij}^+ + \lambda_{ij}^- = 0, \ \forall ij \in L \tag{3-27}$$

$$q_j + \sum_{ji \in L} f_{ji} - \sum_{ij \in L} f_{ij} = y_j, \ \forall j \in N \tag{3-28}$$

$$0 \le \theta_j^{\max} - \theta_j \perp \alpha_j^+ \ge 0, \ \forall j \in N \tag{3-29}$$

$$0 \le -\theta_j^{\min} + \theta_j \perp \alpha_j^- \ge 0, \ \forall j \in N \tag{3-30}$$

$$0 \le B_{ij}\left(\theta_i - \theta_j\right) - f_{ij} + \left(1 - z_{ij}\right) M_{ij} \perp \gamma_{ij}^+ \ge 0, \ \forall ij \in L \tag{3-31}$$

$$0 \le -B_{ij}\left(\theta_i - \theta_j\right) + f_{ij} + \left(1 - z_{ij}\right) M_{ij} \perp \gamma_{ij}^- \ge 0, \ \forall ij \in L \tag{3-32}$$

$$0 \le z_{ij} K_{ij} - f_{ij} \perp \lambda_{ij}^+ \ge 0, \ \forall ij \in L \tag{3-33}$$

$$0 \le z_{ij} K_{ij} + f_{ij} \perp \lambda_j^- \ge 0, \ \forall ij \in L \tag{3-34}$$

$$0 \le y_j \perp -\eta - \phi_j + \pi_j + \beta_j + \mu_j \ge 0, \ \forall j \in N \tag{3-35}$$

$$y_j + \sum_{i \in N} \frac{1}{b_i} \beta_j = 0,\ \forall j \in N \tag{3-36}$$

$$0 \le \mu_j \perp nV_j - y_j \ge 0,\ \forall j \in N \tag{3-37}$$

$$nU_{gj} \ge U_{gj},\ nV_j \ge V_j,\ z_{ij} \in \{0,1\} \tag{3-38}$$

Constraints (3-21) – (3-24) are the KKT conditions for the fuel dispatcher's problem (3-2) – (3-6), while (3-25) – (3-34) are the KKT conditions for the ISO's decision problem (3-7) – (3-14), and (3-35) – (3-38) are the accumulated KKT conditions for all generation companies' problems (3-17) – (3-20).

Due to the nonconvexity of the feasible region, the MPCC problem is difficult to solve. In the next section, we outline three reformulations and describe how they help to identify and evaluate the global optimal solution.

3.3 Reformulation and Solution

3.3.1 Nonlinear Programming Reformulation (MPCC-NLP)

To solve the MPCC, its complementarity constraints must be reformulated. One method is to transform the complementarity constraints into nonlinear functions. Consider a generic complementarity constraint as:

$$0 \le r \perp s \ge 0 \tag{39}$$

The product reformulation replaces it with constraints that r and s are nonnegative and $r's = 0$ [21]. The complementarity constraint can also be expressed in terms of a nonlinear complementarity problem (NCP) function $\Phi(r,s)$ that satisfies $\Phi(r,s)=0$ if and only if $r's=0$ and $r,s \ge 0$. An example of $\Phi(r,s)$ is the Fischer-Burmeister function

$\sqrt{r^2+s^2}-r-s$ [22, 23]. Both of these methods maintain the reformulated complementarity conditions as constraints. A third method is to penalize positive values of the reformulated nonlinear function $r's$ in the objective function.

These three methods are available as options in the NLPEC solver in the General Algebraic Modeling System (GAMS) [24]. For the numerical instance in Section 3.4, no optimal solutions were found with either the Fischer-Burmeister function or the penalty reformulations. Local optimality was achieved by the product reformulation, which converted the MPCC to a mixed integer nonlinear programming problem. However, global optimality of the solution is not guaranteed. We can conclude only that the solution is feasible but not necessarily globally optimal for the MIBLP problem. Its objective value is therefore a lower bound on the optimal value.

To identify the global optimum of the MIBLP problem, we further explored two additional methods to solve the problem described in sections 3.3.2 and 3.3.3 respectively.

3.3.2 Single-Level Mixed Integer Quadratic Program (1-level MIQP)

The MIBLP model assumes that a central decision maker anticipates the lower level decision makers' reactions to his investment decision on the capacity expansion of all the facilities involved and makes the optimal decision to maximize the total benefit. All of the lower level decision makers will respectively make their optimal operational decisions, given the leader's decision.

In the 1-level MIQP relaxation, we assume that there is only one centralized decision maker in the market making all investment and operational decisions to optimize the benefit of the whole system, while satisfying all the physical constraints from each part of the integrated network. In this case, the generator companies are no longer able to make their own strategic decision to maximize their profit but simply accept the market optimal decisions.

This relaxed problem can be derived by removing all of the objective functions of the lower level optimization problems. The objective function of relaxed MIQP problem is equation (3-1), and the constraints are equations (3-3) – (3-6), (3-8) – (3-14) and (3-18) – (3-20). Because the ISO and fuel dispatch objectives are already included in (3-1), only the generator strategic capacity is removed. Since the problem is a relaxation of the original problem, its optimal solution provides an upper bound for the MIBLP problem, and therefore can be used to bound the optimality gap once a feasible solution is provided.

The difference of the optimal objective values derived from the MPCC-NLP and the 1-level MIQP gives a range in which the global optimal objective value must lie. If the gap between them is small enough, the global solution can be well approximated by the solution of MPCC-NLP problem.

3.3.3 Binary Variables Reformulated Mathematical Program with Complementarity Constraints (MPCC-BIN)

To solve the problem more efficiently and, more importantly, to obtain the global optimal solution, we converted MPCC problem into an equivalent mixed integer quadratic program by introducing a set of binary variables σ and the large parameters M [14]. For instance, the reformulation of equation (3-39) is:

$$0 \le r \le M\sigma \quad (3\text{-}40)$$

$$0 \le s \le M\left(1-\sigma\right) \quad (3\text{-}41)$$

The complementarity constraints have been converted into the mixed integer linear constraints, and the whole problem becomes an MIQP problem. The big value M and the binary variable σ ensure that either r or s must be zero.

These inequalities ensure that either r or s must be zero. The complementarity constraints have been converted into the mixed integer linear constraints, and the whole problem becomes an MIQP problem.

The constraints (3-22), (3-25) – (3-28), (3-36) and (3-38) of MPCC problem remain the same and the constraints (3-21), (3-23) – (3-24), (3-29) – (3-35) and (3-37) are reformulated by the binary variables. A large number of binary variables are introduced in the MPCC-BIN reformulation.

3.4 Numerical Results

We studied a six node transmission network with four fuel suppliers illustrated in Fig. 3-2 [2].

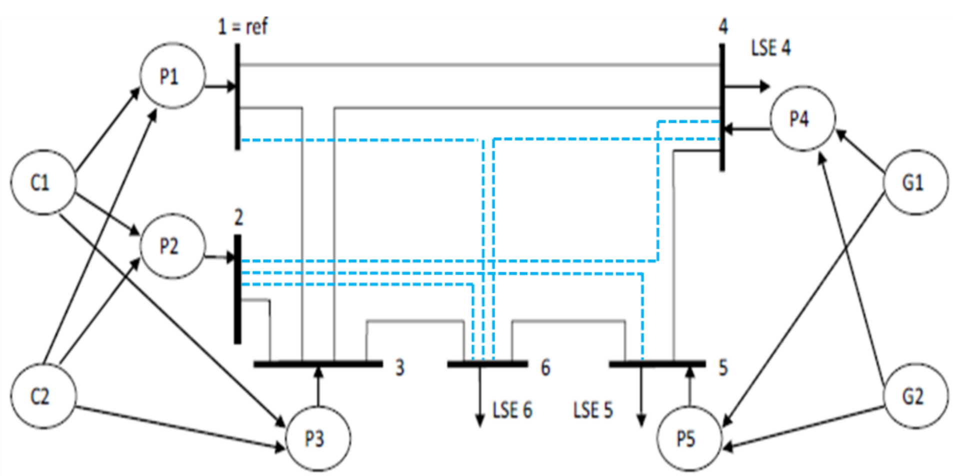

Figure 3- 2 Integrated Electricity Supply Network including the Fuel Suppliers, Transmission Grid and Generators

Table 3- 4 Investment Costs of Fuel Delivery, Generation and Transmission Capacity

Investment Cost	Coal	Natural Gas
fc ($/MWh)	1.5	4
gc ($/MWh)	10	6
tc ($/MWh)	4	

Table 3- 5 Fuel Capacity and Transportation Cost

Fuel Suppliers		Electricity Nodes				
		1	2	3	4	5
$U_{g,j}$ (MWh equivalent)	C1	200	30	30	0	0
	C2	800	500	500	0	0
	C1	0	0	0	30	60
	C2	0	0	0	200	800
c_{gj} ($/MWh)	C1	65	73	70	-	-
	C2	80	75	72	-	-
	G1	-	-	-	120	115
	G2	-	-	-	125	122

The generators P1, P2 and P3 are coal-fired plants, supplied by coal suppliers C1 and C2, and the generators P4 and P5 are natural gas-fired plants supplied by two natural gas suppliers as in G1 and G2. The LSEs 4, 5 and 6 represent the electricity loads. The solid lines are the existing transmission lines and the dashed lines are the candidate transmission lines for possible expansion.

Table 3- 6 Generation Capacity and Parameters for Inverse Demand Function

Electricity Nodes	V_j (MW)	b_j ($/MWh/MWh)	a_j ($/MWh)
1	600	-Inf	0
2	400	-Inf	0
3	400	-Inf	0
4	1000	-0.08	250
5	600	-0.08	300
6	N/A	-0.08	350

The numerical results are based on a single hour. All of the expansion costs are also estimated on an hourly basis. The generation and transmission capital costs are derived from Joint Coordinated System Planning report [25]. All of the other parameters in Table 3-6 and 3-7 are based on [2]. Tables 3-4, 3-5, 3-6 and 3-7 give the parameters of the case study.

Table 3- 7 Transmission Capacity, Susceptance of the Network and Initial Status of the Transmission Lines

Transmission Line	Transmission Capacity $K_{i,j}$ (MW)	Susceptance B_{ij} $\left(\Omega^{-1}\right)$	z_{ij}^0
(1, 3)	400	156.25	1
(1, 4)	240	33.67	1
(2, 3)	1000	Inf	1
(3, 4)	150	32.89	1
(3, 6)	250	35.59	1

Table 3- 7 (continued)

(4, 5)	240	33.67	1
(5, 6)	350	92.59	1
(2, 6)	400	32.89	Can
(2, 5)	400	35.59	Can
(1, 6)	400	32.89	Can
(1, 2)	400	156.25	Can
(2, 4)	400	32.89	Can
(4, 6)	400	33.67	Can

Table 3- 8 Numerical Results of Original Equilibirum MPCC-NLP, 1-level MIQP and MPCC-BIN problems

Decision Variables	Problems			
	Original Equilibrium	MPCC-NLP	1-Level MIQP	MPCC-BIN
Objective Value	286191	559895	575362	564275
Social Welfare	286191	598973	618312	600173
Consumer Surplus	*77798*	*242068*	*333332*	*243554*
Generator Surplus	*119318*	*160840*	*52715*	*148060*
Transmission Rent	*89075*	*196065*	*232265*	*208559*
Fuel Expan. Cost	0	12174	14598	12173
Gen. Expan. Cost	0	17305	20353	15725

Table 3- 8 (continued)

Trans. Expan. Cost		0	9600	8000	8000
	(C1,1)	0	1018	755	755
	(C1,2)	0	18	67	125
Fuel Expansion	(C1,3)	0	131	360	312
	(G1,4)	0	722	860	713
	(G1,5)	0	1883	2346	1883
	1	0	628	365	365
Gen. Expansion	4	0	0	0	0
	5	0	1353	1816	1353
	(2,6)	0	1	1	1
	(2,5)	0	1	1	1
	(1,6)	0	1	1	1
Trans. Expansion	(1,2)	0	1	0	0
	(2,4)	0	1	1	1
	(4,6)	0	1	1	1
	(C1,1)	200	1218	955	955
	(C1,2)	30	48	97	155
Fuel Delivered	(C1,3)	30	161	390	342
	(C2,1)	157	0	0	0
	(C2,2)	30	0	0	0

Table 3- 8 (continued)

	(C2,3)	142	0	10	0
	(G1,4)	130	852	990	843
	(G1,5)	60	1943	2406	1943
	(G2,4)	790	0	10	0
	(G2,5)	540	0	10	0
Amount generated by Generator	1	357	1218	955	955
	2	60	48	97	155
	3	172	161	400	342
	4	920	852	1000	843
	5	600	1943	2416	1943
Generation Consumed	4	1256	1341	1501	1344
	5	386	1665	2138	1665
	6	468	1218	1229	1229
Electricity Price	1	90	97	73	90
	2	77	76	75	79
	3	77	76	75	79
	4	150	143	130	142
	5	269	167	129	167
	6	313	253	252	252

Table 3- 8 (continued)

	(1,2)	0	344	0	0
	(1,3)	168	344	400	400
	(1,4)	190	228	240	240
	(1,6)	0	303	315	315
	(2,3)	60	-105	-400	-342
	(2,4)	0	150	150	150
Electricity Flow	(2,5)	0	115	115	115
	(2,6)	0	231	231	231
	(3,4)	150	150	150	150
	(3,6)	250	250	250	250
	(4,5)	4	-44	-44	-44
	(4,6)	0	83	83	83
	(5,6)	218	350	350	350

We implemented all the problem formulations: MPCC-NLP, 1-level MIQP, and MPCC-BIN, via the modeling language of GAMS and called its inner solvers to solve the problems. The original equilibrium represents the solution to the lower level problem only and is found by the PATH solver [26]. The MPCC-NLP is solved by the DICOPT solver [27] which cannot guarantee global optimality. The 1-level MIQP and MPCC-BIN problems are both solved by the CPLEX solver [29] to global optimality. The numerical results are indicated in Table 3-8. The "Original Equilibrium" solution in Table 3-8 gives the MPCC-

NLP results without any expansion. The MPCC-BIN is solved with each value of *M* equal to 10000. The methodology to derive appropriate values for *M* is described in Appendix 3.C.

The BARON NLP solver could find a global optimum for MPCC-NLP if all mathematical expressions have finite lower and upper bounds [28]. We did not pursue this avenue because finding upper bounds is equivalent to identifying large enough values of *M* for the binary reformulation and finding lower bounds requires additional effort.

Let the optimal objective values found by solving MPCC-NLP, 1-level MIQP and MPCC-BIN be ζ_1, ζ_2 and ζ_3 and the global optimum of the MIBLP problem be ζ_{opt}. The ptimality gap of the MPCC solution is $\left|\zeta_1-\zeta_{opt}\right|$. Since the 1-level MIQP problem is a relaxation, ζ_2 is an upper bound of ζ_{opt}. Therefore, the optimality gap is bounded by $\left|\zeta_1-\zeta_2\right|$, which indicates how far the obtained optimal solution might be from the global optimal solution. The optimality gap by percentage can also be defined as $\left|\zeta_1-\zeta_{opt}\right|/\zeta_{opt}\times 100\%$, bounded from above by $\left|\zeta_1-\zeta_2\right|/\zeta_2\times 100\%$. In our numerical study, the optimality gap $\left|\zeta_1-\zeta_2\right|$ is 15467 and the bound on the percentage optimality gap is 2.76%, which implies that the feasible solution solved by MPCC-NLP is within 2.76% of the global optimum.

The MPCC-BIN reformulation is also equivalent to the MIBLP problem. Moreover, it is also a maximization problem with a concave quadratic objective function. Therefore, the global optimal solution is guaranteed. The CPLEX solver verifies convexity by checking that the Hessian matrix of both objective function and the constraints is positive semi-definite. This allows inclusion of a lower bound, but not an upper bound, constraint on the objective function value. We compared the computational time of the problem with and without the lower bound obtained from MPCC-NLP. The computation time in seconds for solving the original equilibrium, MPCC-NLP, and 1-level MIQP are, respectively, 0.484 and 0.125. It takes 0.953 s to solve MPCC-BIN without any bound, and 0.64 s with the lower bound provided by MPCC-NLP. That is, the bound improves the computational efficiency by nearly 33%.

The optimal solutions suggest that the decision maker should build the candidate transmission lines (2, 6), (2, 5), (1, 6), (2, 4) and (4, 6) to achieve the global optimum. It is obvious in Table 3-4 that the coal-fired generators are much cheaper than the natural gas-fired generators. Thus the electricity is more likely to flow from the left to right in Figure 3-2, especially given that all of the loads are located on nodes 4, 5 and 6. Before making the expansion decision, transmission congestion exists on lines (3, 4) and (3, 6). Without the accessibility of the cheaper electricity, the LMPs on LSE nodes 4, 5 and 6 are much higher, which also suggests that more transmission lines are required to help deliver the electricity from left to right. The expansion made on the candidate transmission lines increases the transmission capacity to deliver more electricity from the coal-fired generators to the loads, and thus will certainly help to balance the electricity prices of the network.

As for the coal generators, the cheapest fuel source is C1. Therefore the decisions have been made to expand the arcs (C1, 1), (C1, 2) and (C1, 3). Similar decisions have been made to expand the natural gas transportation from the relatively cheaper source G1. Even though the natural gas costs are about twice as high as the coal costs, it makes the gas transport expansion decision due to the transmission capacity limits. The congestion makes it impossible to use all of the electricity available from the coal-fired plants.

The decision on the generation expansion matches the expansion on the fuel transportation and the transmission grid.

The expansions help generate and deliver more cheap electricity to satisfy the demand and thus improve the balance in the electricity prices on the nodes. It also leads to an increase of sellers' and buyers' surplus. Although the price differences among the nodes has been decreased, the transmission congestion still exists and the increasing number of transmission lines results in even more transmission rents in total. All of the effects achieve an increase in overall welfare of the integrated electricity supply system.

Also from Table 3-8, by comparing the results of both MIQP and MPCC-BIN problems, we are able to see how the strategic decisions made by generation companies affect the performance of the electricity market, due to the fact that MIQP problem is a relaxation of the MIBLP problem obtained by only eliminating the strategic behavior of the generators. Without generator strategic operational behavior, more fuel supply and generation facilities are expanded so that the electricity prices are lower, which results in a large increase in buyer surplus and decrease in generator surplus.

3.5 Conclusions

In this paper, we investigated a capacity expansion bilevel programming problem. In the lower level, we take into account an integrated electricity supply system including the fuel transportation, generation and transmission, as well as the interactions among them in a restructured electricity market, where the consumer demand is modeled as a linear function of the electricity price. Capacity expansion decisions are made by an upper level decision maker from a global point of view.

In the absence of strategic operational decisions by generators, the total social welfare increases. Electricity buyers are better off while the generators are worse off. Fuel and generation facilities are expanded more, which leads to lower electricity prices. In our numerical study, generator strategic operational decisions reduce the welfare less investment cost by 2%.

We used two reformulations of the MPCC to efficiently identify a global optimum. The NLP reformulation takes less time to solve the problem but its solution is not guaranteed to be globally optimal. It provides a feasible solution and lower bound on the optimal value. On the other hand, the binary formulation takes more time to solve the problem, but it is able to identify the global optimal solution. Including the lower bound derived from MPCC-NLP significantly improves its computation time. However, it takes effort to find an appropriate *M* value as a tight bound for the mathematical expressions in the complementarity constraints to implement the MPCC-BIN reformulation. Small values of *M* could eliminate the optimal solution but excessively large ones increase the computation time. The relaxed 1-level MIQP provides upper bounds on the optimal value and optimality gap. It can be solved easily, but will not necessarily provide a feasible solution for the original problem.

A six bus case study is provided to illustrate the three methodologies and give the combined expansion results in fuel transportation, generation and transmission. We also analyze the effect of the global optimal expansion decision on the integrated electricity supply system.

For future research, the investment costs for the fuel transportation, generation and transmission subsystems can be further extended to nonlinear cost functions that incorporate economies of scale. A dynamic decision making process can also be represented to optimize investment decisions in multiple periods over a long term horizon. The major uncertainties including natural gas cost and electricity demand can also be taken into account in the model. To do so, we can first generate different future scenarios for the uncertainties and then incorporate them into the model as a stochastic MPCC problem. Furthermore, we can also compare the results of the model in our paper with the ones from a more realistic point of view in which every asset owner makes his own capacity expansion decisions. This comparison will show how much the optimal decisions identified from a global point of view could benefit the integrated electricity supply system and provide possible targets for policy design.

Appendix 3.A Elaboration of Fuel Dispatcher's Decision Problem

Two fuel resources, coal and natural gas, which normally have different units \$/ton and \$/thousand cubic feet, are considered in Section IV. To make the units of different fuel types match and most importantly compatible with the unit of energy MWh, we converted the units of fuel into MWh equivalents and set their cost parameters $\boldsymbol{c}$, capacity limit parameters $\boldsymbol{W}$ and $\boldsymbol{nU}$, and investment costs $\boldsymbol{fc}$ accordingly. It is also possible to directly model the fuel dispatcher problem with original units of fuel resources by incorporating the heat rate $\boldsymbol{H}$ and efficiency ε conversion into the model. The fuel cost $\boldsymbol{c}$ can also further be decomposed into two parts: fuel cost ***cl*** and the delivery cost ***del***. In this case, the objective function Eq. (3-1) is changed to:

$$\max_{nU,nV,z,q} \sum_j \left(\frac{1}{2} b_j q_j^{\ 2} + a_j q_j \right) - \sum_{gj \in A} \left(cl_{gj} + del_{gj} \right) x_{gj}$$
$$- \sum_{gj \in A} fc_{gj} \left(nU_{gj} - U_{gj} \right) - \sum_{j \in N} gc_j \left(nV_j - V_j \right) - \sum_{ij \in L} tc_{ij} K_{ij} z_{ij}$$

And the fuel dispatcher's decision problem is revised as:

$$\min_{x \ge 0} \sum_{gj \in A} \left(cl_{gj} + del_{gj} \right) x_{gj}$$

$$s.t. \ - \sum_{j \in N_g} x_{gj} \ge -W_g, \quad \forall g \in F \quad [\omega_g \ge 0]$$

$$\sum_{g \in F_j} x_{gj} e_g = \frac{y_j \varepsilon_j}{H_j}, \quad \forall j \in N \qquad [\pi_j]$$

$$-x_{gj} \ge -nU_{gj}, \ \forall gj \in A \quad [\rho_{gj} \ge 0]$$

$$x_{gj} \ge 0$$

Since the units of ***x*** are \$/ton and \$/thousand cubic feet respectively, we have parameter ***e*** (BTU/ton or BTU/thousand cubic feet) to convert both of them into \$/BTU.

The revision affects the units of π, which also represent the marginal costs in the objective function of generator's problem. It can be expressed as:

$$\max_{y_i \ge 0, \eta} \left(\eta + \phi_i - \frac{H_j}{\varepsilon_j} \pi_i \right) y_i$$

Appendix 3.B Incorporation of Carbon Emission Regulations

One of the ways to consider carbon emission concerns is simply to adopt a carbon emission cost p_{co2} with the unit $/ton. The objective function of the generator's problem is restated as:

$$\max_{y_i \geq 0, \eta} \quad (\eta + \phi_i - \pi_i) y_i - p_{co2} E_i y_i$$

The parameter $\boldsymbol{E}$ is the tons of carbon emission emitted per MWh of energy depending on different generation technologies.

Another way to incorporate the carbon emission is in a cap-and-trade system. Each of the generation companies is given a certain number of carbon emission allowances $\boldsymbol{N}$. Generation companies are allowed to trade the allowances as long as the total carbon emissions of the system are within the limit [30]. The objective function of the generator's problem is then changed to:

$$\max_{y_i \geq 0, \eta} \quad (\eta + \phi_i - \pi_i) y_i - p_{co2} (E_i y_i - N_i)$$

If the carbon emission $\boldsymbol{E_i y_i}$ exceeds its allowance, the generation company must buy allowances from others for the extra emissions. Otherwise, the generation company can make a profit by selling its unused allowances. In addition, a market clearing equilibrium constraints must be added to the upper level optimization problem [30]:

$$0 \leq p_{co2} \perp \left(\sum_i E_i y_i - \sum_i N_i \right) \leq 0 \qquad (2\text{-}42)$$

If the total carbon emission is less than the total amount of the allowance, the carbon allowance trade is free. Otherwise, there is price p_{co2}>0 for buying each ton of the carbon allowance.

Appendix 3.C Setting the Values for M in the Binary Reformulation

One way to set the *M* value is to roughly estimate the biggest possible values for all the *r* and *s* in the equilibrium constraints, which is also equivalent to estimate the upper bounds of the dual and primal variables.

For the fuel dispatcher's decision problem, Eq (2-21), (2-23) and (2-24) are the relevant equilibrium constraints. The primal variables $\boldsymbol{x}$ is bounded by $\boldsymbol{nU}$, the new fuel transportation capacity after expansion. The expansion will not be infinite due to the limited electricity demand. The dual variables ω, $\boldsymbol{\pi}$ and $\boldsymbol{\rho}$ represent how much the objective function changes if the right-hand side (RHS) changes by 1 unit. For Eq. (2-3) and (2-5), the largest possible change in objective function happens when an extra unit of cheapest fuel resource becomes available and substitutes one unit of the most expensive fuel resource, which is estimated as $\overline{c_{gj}} - \underline{c_{gj}} = 60$ in this specific instance. Likewise, we can also obtain the upper bound $\overline{c_{gj}}$ =125 for $\boldsymbol{\pi}$ which represents the marginal fuel cost.

For the generator's decision problem, we estimated the bound of the variables in the same manner so that $\boldsymbol{y}$ is bounded by $\boldsymbol{nV}$, and $\boldsymbol{\mu_g}$ and $\boldsymbol{\beta_j}$ are both bounded by $\overline{a_j} - \underline{\pi_j}$ =285.

For ISO's decision problem with equilibrium constraints (2-29) – (2-34), $\boldsymbol{f}$, $\boldsymbol{\theta}$ and $\boldsymbol{q}$ are bounded by $\boldsymbol{K}$, $\boldsymbol{\theta_{max}}$ and $\boldsymbol{nV}$, respectively. The upper bound for $\boldsymbol{p}$ is $\boldsymbol{a}$, the intercepts of the inverse demand functions. The dual variables of the voltage angle constraints, $\boldsymbol{\alpha}^+$ and $\boldsymbol{\alpha}^-$, will not affect the objective function because the voltages are always within the bounds. The dual variables $\boldsymbol{\lambda}^+$ and $\boldsymbol{\lambda}^-$ indicate how much the welfare changes if the capacity limit changes by one unit. Since it is quite difficult to estimate the impact of a one unit flow change on $\boldsymbol{q}$, we can roughly evaluate the largest possible change in $\boldsymbol{q}$ and the welfare accordingly. Likewise, we obtain the bounds for γ^+ and γ^-.

According to the rough estimating result of the bounds, letting each value of *M* be 10,000 will be large enough.

Another way to give the *M* an appropriate guess is to take advantage of the MPCC-NLP solution. It provides a general idea of neighborhoods for the optimal values of the variables, which can be used to estimate a tighter and more realistic *M*. Based on our MPCC-NLP solution, we estimate that 5000 should be large enough for *M*.

In our case study of MPCC-BIN, we tried out different values for *M*, ranging from 2000 to 100000. The results indicate that all of these values are valid for *M* because they all result in the same optimal solution. Better computational performance could be achieved by setting different *M* values for each of the mathematical expressions in the complementarity constraints.

References

[1] S.M. Ryan, A. Downward, A.B. Philpott and G. Zakeri, “Welfare effects of expansions in equilibrium models of an electricity market with fuel network”, *IEEE Transactions on Power Systems*, Vol. 25, No. 3, pp. 1337-1349, 2010

[2] S.M. Ryan, “Market outcomes in a congested electricity system with fuel supply network”, *Proceedings of IEEE Power Engineering Society General Meeting*, Calgary, Jul 2009

[3] Electric Power Annual Report, Energy Information Administration, Jan 2010, http://www.eia.doe.gov/cneaf/electricity/epa/epa_sum.html, viewed in Aug 2010

[4] J. Yao, I. Adler and S.S. Oren, "Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network", *Operations Research*, Vol. 56, No. 1, pp. 34-47, 2008

[5] J.T. Moore and J.F. Bard, "The mixed integer bilevel programming problem", *Operations Research*, Vol. 38, No. 5, pp. 911-921, 1990

[6] S.T. DeNegre and T.K. Ralphs, "A branch-and-cut algorithm for integer bilevel linear programs", *Operations Research and Cyber-Infrastructure*, Vol. 47, No. 2, pp. 65-78, 2009

[7] B. Colson, P. Marcotte and G. Savard, "An overview of bilevel optimization", *Annals of Operations Research*, Vol. 153, No. 1, pp. 235-256, 2007

[8] X. Hu, D. Ralph, "Using EPECs to model bilevel games in a restructured electricity market with locational prices", *Operations Research*, Vol. 55, No. 5, pp. 809-827, 2007

[9] S. Soleymani, A.M. Ranjbar, A.R. Shirani, "New approach to bidding strategies of generating companies in day ahead energy market", *Energy Conversion and Management*, Vol. 49, No. 6, pp. 1493-1499, 2008

[10] T. Li and M. Shahidehpour, "Strategic bidding of transmission-constrained GENCOs with incomplete information", *IEEE Transaction on Power Systems*, Vol. 20, No. 1, pp. 437-447, 2005

[11] R. Garcia-Bertrand, D. Kirschen, A.J. Conejo, “Optimal investments in generation capacity under uncertainty”, *Proceedings of 16th Power Systems Computation Conference,* Glasgow, United Kingdom, Jul 2008

[12] L.P. Garces, A.J. Conejo, R. Garcia-Bertrand and R. Romero, “A bilevel approach to transmission expansion planning within a market environment”, *IEEE Transactions on Power Systems*, Vol. 24, No. 3, pp. 1513-1522, 2009

[13] M.C. Ferris and J.S. Pang, “Engineering and economic applications of complementarity problems”, *SIAM Review*, Vol. 39, No. 4, pp. 669-713, 1997

[14] J. Hu, J.E. Mitchell, J.S. Pang, K.P. Bennett and G. Kunapuli, “On the global solution of linear programs with linear complementarity constraints”, *SIAM Journal on Optimization*, Vol. 19, No. 1, pp. 445-471, 2008

[15] S.M. Ryan, “Demand price sensitivity and market power in a congested fuel and electricity network”, *Proceedings of IEEE Power Engineering Society General Meeting,* Minneapolis, Jul 2010

[16] K.W. Hedman, R.P. O’Neil, E.B. Fisher and S.S. Oren, “Optimal transmission switching with contingency analysis”, *IEEE Transactions on Power System*, Vol. 24, No. 3, pp. 1577-1586, 2009

[17] H. Liu, L. Tesfatsion, and A.A. Chowdhury, “Derivation of locational marginal prices for restructured wholesale power markets”, *Journal of Energy Markets,* Vol. 2, No. 1, pp. 3-27, 2009

[18] E.B. Fisher, R.P. O'Neill, M.C. Ferris, “Optimal transmission switching”, *IEEE Transactions on Power Systems*, Vol. 23, No. 3, pp. 1346-1355, 2008

[19] C.J. Day, B.F. Hobbs, J. Pang, "Oligopolistic competition in power networks*:* a conjectured supply function approach*", IEEE Transaction on Power System*, Vol. 17, No. 3, pp. 597-607, 2002

[20] Z. Younes, M. Ilic, "Generation strategies for gaming transmission constraints*:* will the deregulated electric power market be an oligopoly?", *Decision Support Systems*, Vol. 24, No. 3-4, pp. 207-222, 1999

[21] M.C. Ferris, S.P. Dirkse, A. Meeraus, "Mathematical programs with equilibirum constraints: automatic reformulation and solution via constrained optimization", T.J. Kehoe et al., eds. *Frontiers in Applied General Equilibrium Modeling*, Cambridge University Press, pp. 67-94, 2005

[22] A. Fischer, "A special Newton-type optimization method", *Optimization*, Vol. 24, pp. 269-284, 1992

[23] A. Fischer, "An NCP-function and its use for the solution of complementarity problems", in: D. Du, L. Qi and R. Womersley, eds., *Recent Advances in Nonsmooth Optimization*, pp. 88–105, 1995

[24] NLPEC solver, http://www.gams.com/dd/docs/solvers/nlpec.pdf, viewed in May 2010

[25] JCSP'08 Report Volume 1 – Economic Assessment, http://www.jcspstudy.org/, viewed in Feb 2010

[26] PATH solver, http://www.gams.com/dd/docs/solvers/path.pdf, viewed in May 2010

[27] DICOPT solver, http://www.gams.com/dd/docs/solvers/dicopt.pdf, viewed in May 2010

[28] BARON solver, http://www.gams.com/dd/docs/solvers/baron.pdf, viewed in Oct 2010

[29] CPLEX solver, http://www.gams.com/dd/docs/solvers/cplex.pdf, viewed in Sep 2010

[30] Y. Chen, B.F. Hobbs, “An oligopolistic power market model with tradable NOx permits”, *IEEE Transactions on Power Systems*, Vol. 20, No. 1, 2005

CHAPTER 4 A TRI-LEVEL MODEL WITH AN EPEC SUB-PROBLEM FOR CENTRALIZED TRANSMISSION AND DECENTRALIZED GENERATION EXPANSION PLANNING FOR AN ELECTRICITY MARKET: PART I

Submitted to *IEEE Transactions on Power Systems*

Shan Jin and Sarah M. Ryan

Abstract

We develop a tri-level model of transmission and generation expansion planning in a deregulated power market environment. Due to long planning/construction lead times and concerns for network reliability, transmission expansion is considered in the top level as a centralized decision. In the second level, multiple decentralized GENCOs make their own capacity expansion decisions while anticipating a wholesale electricity market equilibrium in the third level. The collection of bi-level games in the lower two levels forms an equilibrium problem with equilibrium constraints (EPEC) that can be approached by either the diagonalization method (DM) or a complementarity problem (CP) reformulation. We propose a hybrid iterative solution algorithm that combines a CP reformulation of the tri-level problem and DM solutions of the EPEC sub-problem.

4.1 Introduction

The traditional capacity expansion planning problem takes a centralized perspective compatible with the industry's previous vertically integrated structure of generation, transmission and distribution. In the 1990s, the power systems in United States experienced a transition from regulation to deregulation that is expected to lead to an efficient market and lower electricity prices. It has been a long and difficult transition process. Increasingly across the U.S., the electricity market is composed by separate generation companies (GENCOs), transmission owners (TRANSCOs), distribution companies (DISCOs) and load serving entities (LSEs) [1].The Independent System Operator (ISO) has been developed to monitor the grid, ensure reliability and settle the electricity market for a region. The restructuring process introduced a market environment to which the traditional centralized planning methodologies no longer apply. New models are needed to study market interactions created by different competitors, the strategic decisions of market players and their decisions' impact on wholesale electricity trade. The ISOs and regional reliability councils, who conduct transmission planning studies and reliability assessment, will be interested in how GENCOs' strategic expansion decisions react to the transmission planning decisions, and the performance of the electricity market in response to both the transmission and generation expansions.

To provide a reliable electricity supply network, we must not only consider generation expansion to make sure that we have sufficient energy to meet future loads, but also take into account the entire integrated wholesale electricity supply system including transmission and market clearing by the ISO. All these decisions have great impact on market behavior. Transmission congestion due to insufficient transmission capacity can cause spikes in the locational marginal prices (LMPs) or even load curtailment in extreme cases. The ISO is important to maintain reliable and efficient grid operations. LSEs, who are the buyers in the wholesale market, play important roles in distributing the electricity to retail customers. In restructured markets, expansion decisions may be justified by potential profit increases rather than cost reductions. The profit return received by an investor is determined by an electricity market price settlement. The system operator (ISO) matches the electricity supply bids and demand offers and settles the LMPs to maximize total market surplus of both buyers and sellers. Typically this is done on an hourly basis in a day ahead market and every 5 minutes in the real-time market. Moreover, while investing in more generation capacity, the transmission adequacy should be guaranteed so that the newly installed power can be transported to where the demand is located. Due to deregulation of electricity market, each GENCO can take its strategic expansion and bidding decisions to the market.

We propose a method to solve a market-based generation and transmission expansion problem. In our model, each GENCO anticipates prices settled by an ISO market clearing problem when making its own investment and operational decisions. At the same time, the GENCOs' decisions are also made in response to transmission planning decisions because sufficient transmission capacity is essential for GENCOs to reap additional profits from delivering energy from expanded capacity to the load locations. GENCOs will hesitate to expand if a high level of grid congestion is likely to result in future generation curtailment. On the other hand, too much transmission capacity does not favor expansion either, due to the low electricity prices which provide no incentive for investment. Therefore, the transmission expansion planning decision must be considered in a market based generation expansion planning problem. Although in a deregulated market, transmission lines are owned by individual TRANSCOs, the overall transmission expansion planning decision remains centralized to guarantee reliability of the transmission grid. Therefore, our model includes a centralized transmission planning decision by the ISO, who is mainly in charge of the reliability of the regional market. The ISO conducts a resource adequacy study, anticipates the expansion and dispatch decisions by multiple GENCOs, and decides where to expand the grid. Thus, our market-based model reflects all of these market integrations among GENCOs and the ISO. It captures both dependence of the GENCOs' expansion decisions on the ISO's centralized transmission planning decision and their anticipations of wholesale electricity market settlement after expansion.

We formulate the generation and transmission expansion planning problem as a mixed integer tri-level program, where the discrete centralized transmission planning decisions occur in the first level, multi-GENCOs' generation expansion decisions constitute the second level, and an electricity market equilibrium problem forms the third level. Modeling transmission planning in the top level is consistent with a principle that transmission planning should proactively influence generation investment [2]. The lower level interactions are based on our previous model [3], including strategic behavior by the GENCOs. Because the tri-level structure with a sub-problem of bi-level games poses solution difficulties, algorithms will first be proposed to solve the collection of bi-level games. This collection can be reformulated as an equilibrium program with equilibrium constraints (EPEC), to which two currently available methodologies discussed in [4] can be applied. We propose a hybrid iterative algorithm to solve the entire tri-level programming problem by exploiting the advantages of both EPEC solution methods. In part II of this paper in Chapter 5, case studies of 6, 30, and 118 bus test systems are presented to illustrate how the algorithm works to optimize the transmission expansion plan in anticipation of generation expansion decisions and market equilibria.

The contributions of this paper are fourfold: 1) We propose a novel formulation of centralized transmission and decentralized generation expansion planning as an integrated tri-level optimization problem with a sub-problem of bi-level games. 2) The solution challenges posed by the problem's multi-level and bi-level games structure are addressed by first reformulating the non-convex sub-problem as an EPEC and solving it by the diagonalization method (DM) as multiple mathematical programs with equilibrium constraints (MPECs). Since the concavity of each maximization MPEC is not guaranteed, we also propose a way to verify the solution as a local (approximate) Nash equilibrium (NE) point. 3) We apply a complementarity problem (CP) reformulation to the entire tri-level programming problem to search for promising transmission expansion plans. 4) We develop a novel hybrid iterative algorithm that can successfully solve the entire tri-level expansion planning model.

In Section 4.2, a thorough literature review is given. The model is presented in Section 4.3. Sections 4.4 and 4.5, respectively, illustrate the algorithms to solve an NE game of bi-level games, and a tri-level programming problem with a bi-level games sub-problem. Section 4.6 concludes the paper.

4.1 Literature Review

In a restructured market, generation expansion decisions are no longer taken from a centralized perspective. Instead, each individual GENCO makes its own decision. Because the decision made by a strategic player is always subject to a market settlement result, many studies of restructured electricity markets formulate an ISO market clearing problem as a lower level problem of the decision maker. Generation or transmission expansion decisions by an individual participant are modeled in the upper level with the lower level representing the market outcomes [5][6]. A two-tier, multi-period, multi-GENCO equilibrium capacity expansion model was proposed in [7]. A capacity expansion problem of strategic multi-GENCO bi-level games was presented and a co-evolutionary algorithm was applied to search for the NE solution in [8]. A GENCO's expansion decision with uncertainties of other GENCOs' bidding prices in electricity market was studied [9]. There are also formulations considering the bidding strategies of the generation companies in the upper level [10][11]. A multi-GENCO bi-level problem has been studied subject to a market clearing problem in the lower level and was solved as an EPEC [12]. A review of traditional and market based transmission expansion planning methodologies was summarized in [13].

In the EPEC sub-problem of our tri-level model, each strategic GENCO optimizes its capacity expansion planning decision in the upper level in anticipation of an electricity market equilibrium problem in the lower level, which results in a bi-level program (BLP). BLP models are widely applied in recent papers for a deregulated electricity market to model individual GENCOs' capacity expansion decisions and/or bidding strategies while anticipating the market settlement results. Algorithms are also proposed to solve the BLP. Colson et al. [14] reviewed methodologies and applications of bi-level programs and described their connections with mathematical programs with equilibrium constraints (MPECs). DeNegre et al. [15] further proposed a branch and cut method to improve the branch and bound algorithm in [16]. Ruiz et al. [17] and Kazempour et al. [18] exploited duality theory to equivalently reformulate a bi-level program that combines GENCOs' decisions and electricity market clearing into a mixed integer linear program that can be solved to optimality. Gabriel [19] proposed a technique to linearize a bilinear element in the objective function of a bi-level program and improve computational reliability. Wogrin et al. [20] investigated a stochastic bi-level multi-year generation expansion problem subject to a market clearing problem with conjectured-price-response in the lower level and solved it as a stochastic MPEC. Kazempour et al. [21] also investigated a stochastic bi-level generation investment problem and solved it by Benders decomposition.

The technique of solving a bi-level program by replacing the lower level problems with their equivalent Karush-Kuhn-Tucker (KKT) optimality conditions can be applied to our model. Upon applying the KKT conditions to all the problems at the lower level for every market participant, the bi-level problem becomes a mathematical program with complementarity constraints (MPCC). MPCC is a special case of MPEC, which has attracted great attention for the past decade as more and more engineering and economic applications involve equilibrium modeling. Ferris and Pang [22] gave a comprehensive summary of the engineering and economic applications of MPCC and the available solution algorithms. Its solution requires an equivalent reformulation of complementarity constraints and global convergence of the solution can be guaranteed only under certain conditions. Hu et al. [23] presented a methodology to find the global optimal solution of a linear program with linear complementarity constraints by reformulating the constraints with binary variables.

In a restructured competitive power market in which multiple strategic players make their decisions simultaneously, game theory is widely applied to model and investigate the competitive market behaviors. A single level Cournot game of multiple GENCOs making both capacity expansion and operational decision was studied, and an equilibrium solution was iteratively solved by a DM algorithm [24]. Three models of solving a single level Cournot capacity game under different economic schemes were presented in [25]. Given the parameter assumptions on demand and two types of candidate units, the existence and uniqueness of the Cournot equilibrium solutions were also discussed and proved.

Table 4- 1 Comparison with Different Models Proposed in Previous Relevant Literature Review

	[26]	[27]	[28]	[29][30]	[31]	Our Model
Transm. Expan.	Cent.; Exist./new line expan.; Max. net surplus	Decent.; New line expan.; Max. net profit	Cent.; Exist./new line expan.; Multi-criteria	Cent.; Exist. line augment.; Min. oper. and invest. cost	Cent; Exist./new line expan.; Min. oper. and invest. cost	Cent; New line expan.; Max. net surplus
Gen. Expan.	Decent.; Continuous	Decent.; Binary	Decent.; Cont.	Decent; Binary	Decent.; Continuous	Decent.; Continuous
Multi-Period Expan.	No	Yes	Yes	No	No	No
ISO's Market Problem	Max. surplus	Min. system cost, min. loss of energy prob.	Max. surplus	Min. oper. cost	Min oper. cost	Max. surplus
GENCO's Operational Problem	Strategic (Cournot)	Competitve	Strategic (pair of price and quantity)	Strategic (pair of price and quantity)	Competitive	Strategic (Cournot)
Operational Uncertainty	Yes	Yes	No	No	Yes	No
Solution Method	Optimization of Bi-level Games	Simulation of an Iterative Procedure	Search-based and Agent-based Method	Genetic Algorithm	Linearization and MILP Reformulation	Iterative algorithm with Optimization of Bi-level Games

When more dependencies among planning decisions and interactions among market players are taken into account, the planning model takes on a more complicated structure involving multi-level or bi-level games. Sauma and Oren [26] studied a multi-GENCO equilibrium expansion planning model with anticipation of an ISO market clearing problem, and evaluated the transmission expansion's effect on the social welfare of the system by considering different transmission expansion plans. For various candidate transmission expansion decisions, the bi-level games were solved by an iterative DM algorithm. Hu and Ralph [4] investigated a sufficient condition for existence of pure-strategy Nash equilibrium of bi-level games, discussed the concepts of local Nash and Nash stationary equilibrium, and proposed a DM of iteratively solving each single bi-level optimization problem and a CP reformulation to solve the bi-level games. Roh et al. [27] developed an iterative process to solve a generation and transmission planning problem by simulating the interactions among GENCOs, TRANSCOs and ISO with consideration of uncertainty, profit from the market clearing decision, and transmission reliability. Motamedi et al. [28] proposed a transmission expansion framework to take into account the expansion reaction from decentralized GENCOs and also integrated an operational optimization in restructured electricity market. The problem was formulated as a four level model and it was approached by agent-based system and search-based techniques. Hesamzadeh et al. [29][30] studied a new framework of transmission augmentation planning problem with strategic generation expansion and operational decision and solved a tri-level program by a genetic algorithm. Pozo et al. [31] studied a three-level generation and transmission model, and converted it into single level mixed integer linear programming problem. Table 4-1 compares the tri-level model that we propose with the multi-level generation and transmission expansion models investigated in the previous papers. Our formulation is similar to the one in [26] but we include the transmission plan as a decision variable in the optimization problem rather than a parameter. Our tri-level model has a similar structure to that investigated in [31]. However, we consider price-responsive demand functions and strategic interactions among the generators at the operational level. The objective of the system operator, to maximize the total net surplus, cannot be reduced to minimizing cost. The problem structure is also similar to [29][30] but we consider expansion as new transmission lines rather than augmentation of the existing circuits, price-responsive demand functions, Cournot competition among GENCOs in the operational level, and surplus maximization as the objective function for the system operator.

4.2 Model and Formulation

4.2.1 Model Assumption

1) For simplicity, we formulate a static model with a single hour of operation and no uncertainty. Thus, the third (operational) level represents a typical hour in a single future scenario for market conditions. The model can be extended to incorporate multiple periods and probabilistic scenarios at the expense of increased computational time.

2) For the transmission expansion, we only consider building new lines and do not consider expanding the capacity of the existing lines. However, the model can be easily extended to include line expansion without changing the structure of the problem.

3) The transmission and generation expansion costs in the top two levels are both modeled as linear and are discounted to form equivalent hourly costs.
4) A price-responsive linear demand function is modeled for each LSE.
5) Each generator at each bus is owned by a single GENCO. As a Cournot competitor, each GENCO makes his own decision on the generation quantity to sell in the electricity market under a type of bounded rationality [32]. A quadratic generation cost function that will not be affected by the capacity expansion is assumed.
6) We assume the market equilibrium in the third level is simultaneously determined by Cournot competition among the GENCOs and an ISO market clearing problem. Its equivalent linear complementarity problem (LCP) reformulation generates a unique NE solution due to the concave objective functions and convex feasible regions [26]. Although the Cournot model simplifies the actual market structure, its broad market outcomes have been validated against an agent-based simulation of bid and offer matching [33]. In contrast, the multiplicity of solutions [10][12] to supply function equilibrium formulations may obscure the effects of upper-level capacity expansion decisions.

4.2.2 Model Formulation

The problem is formulated as a tri-level model with the ISO's discrete transmission expansion decisions on the first level, multi-GENCOs' separate generation expansion decisions on the second level, and the multiple market players' operational decisions in the third level. By extending the bi-level model in [3], we decentralize the generation expansion decisions by separate GENCOs where, in the lower level, the GENCOs and ISO simultaneously optimize their own operational benefits. The GENCO decides its generation level, and the ISO allocates energy to the LSEs to maximize the total system surplus. Different from the lower level model in [3], we do not consider a fuel supply problem to account for the fuel availability and fuel transportation capacity. Instead, for a simplified version, we include a fuel capacity constraint in each GENCO's operational decision problem to represent any combination of fuel supply and transportation capacity constraints. Since the generation decisions are influenced by the transmission grid, it is assumed that the ISO makes a centralized transmission expansion decision in anticipation of the expansion decisions made by its followers, the GENCOS.

Table 4- 2 Sets of Nodes and Arcs

Set	Description	Indices
N	Electricity nodes	*i, j, k*
L	Transmission lines	*ij*
N_{gen}	Set of electricity nodes where a GENCO is located	*i, j, k*

Table 4- 3 Decision Variable Vectors

Primal Decision Variable		Description
First Level	$\boldsymbol{z}$	Vector of binary decision variables for transmission lines, with elements for existing lines set fixed to 1
Second Level	$\boldsymbol{V^{new}}$	Generation capacities after expansion, MW
Third Level	$\boldsymbol{q}$	Demand satisfied at electricity nodes, MW
	$\boldsymbol{\theta}$	Voltage angles at electricity nodes

Table 4- 3 (continued)

$\boldsymbol{f}$	Electricity flows on transmission lines, MW
$\boldsymbol{y}$	Generation amounts at electricity nodes, MW
η	(Scalar) price at the reference electricity node, \$/MWh
$\boldsymbol{\phi}$	Nodal electricity price premia at electricity nodes, \$/MWh

Table 4- 4 Parameter Vectors

Param.	Description
$\boldsymbol{a}$	Intercepts of electricity demand prices as linear functions of quantities, \$/MWh
$\boldsymbol{b}$	Slopes of electricity demand prices as linear functions of quantities, \$/MWh/MWh
$\boldsymbol{c}$	Linear coefficient of the generation cost function, \$/MWh
$\boldsymbol{e}$	Quadratic coefficient of the generation cost function, \$/MWh/MWh
$\boldsymbol{c^{gexp}}$	Invest. costs for generation expansion discounted on an hourly basis, \$/MW
$\boldsymbol{c^{trexp}}$	Invest. costs for transmission line expansion discounted on an hourly basis, \$/MW
$\boldsymbol{\theta^{max}}$	Maximum values for voltage angles
$\boldsymbol{\theta^{min}}$	Minimum values for voltage angles
$\boldsymbol{V}$	Generation capacities at electricity nodes, MW
$\boldsymbol{U}$	Fuel availability, MW
$\boldsymbol{K}$	Capacities of transmission lines, MW
$\boldsymbol{B}$	Susceptances of transmission lines, Ω^{-1}

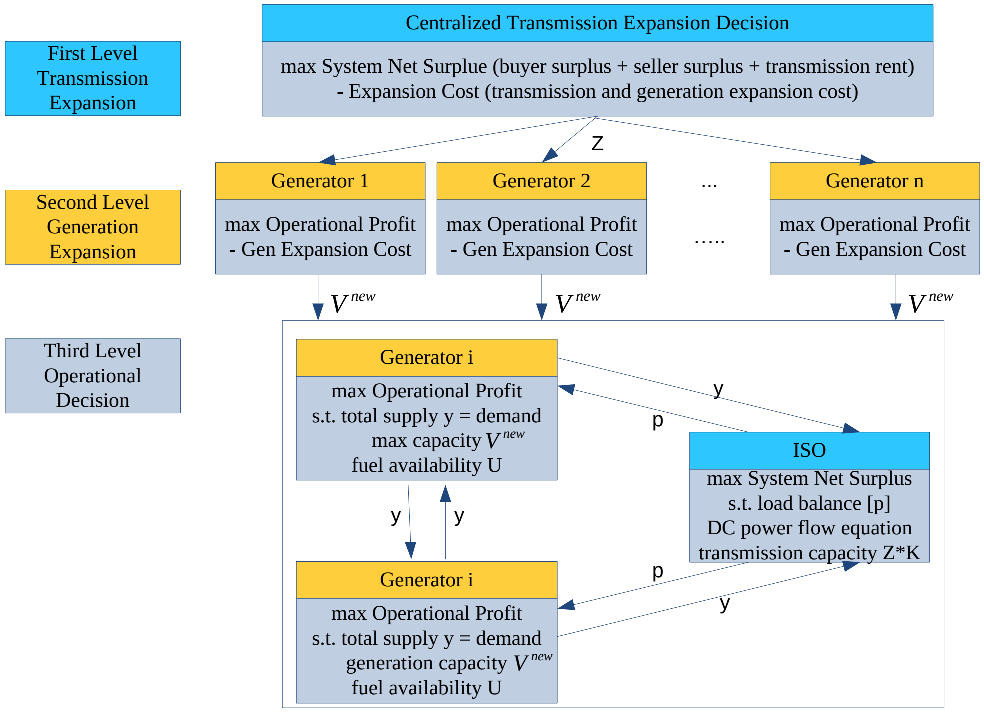

Figure 4- 1 A Tri-level Integrated Generation and Transmission Expansion Planning Model

The model is illustrated in Figure 4-1, and the sets, indices, primal decision variables and parameters are listed in Table 4-2, Table 4-3, and Table 4-4, respectively. A full mathematical formulation of the mixed integer tri-level nonlinear programming model is proposed as below, where the variables in the brackets to the right of constraints are their corresponding dual multipliers.

- First Level: From a system point of view, the ISO collects the information about future loads and resources and makes a centralized decision, z, on transmission expansion to maximize system net surplus, equivalent to system total surplus less generation and transmission expansion cost:

$$\max_{z} \sum_{j\in N}\left(\frac{1}{2}b_j q_j^2 + a_j q_j\right) - \sum_{j\in N_{gen}}\left(c_j y_j + e_j y_j^2\right) - \sum_{ij\in L} c_{ij}^{trexp} K_{ij} z_{ij}$$
$$- \sum_{j\in N_{gen}} c_j^{gexp}\left(V_j^{new} - V_j\right) \quad (4\text{-}1)$$

- Second Level: Each GENCO k makes its own expansion decision V_k^{new}, given the other GENCOs' decisions and in anticipation of market clearing results. Each GENCO k maximizes its net operating profit from selling the power in the electricity market less the expansion cost:

$$max_{V_k^{new}} (p_k - c_k - e_k y_k) y_k - c_k^{gexp} (V_k^{new} - V_k) \quad (4\text{-}2)$$

$$\text{s.t.} \qquad V_k^{new} - V_k \geq 0 \ [\mu nVV_k \geq 0] \quad (4\text{-}3)$$

- Third Level: The ISO optimizes both the sellers' and buyers' surplus, and transmission rent, the total of which is given by $\sum_{j\in N}\left(\frac{1}{2}b_j q_j^2 + a_j q_j\right) - \sum_{j\in N_{gen}}\left(c_j y_j + e_j y_j^2\right)$. Because $\sum_{j\in N_{gen}}\left(c_j y_j + e_j y_j^2\right)$ remains constant in this optimization problem, it is equivalent to maximize $\sum_{j\in N}\left(\frac{1}{2}b_j q_j^2 + a_j q_j\right)$. Constraint (4-5) gives the load balance on each electricity nodes. Equations (4-6) and (4-7) give the bounds on voltage angles. Equations (4-8) and (4-9) are the linearized power flow equations. The thermal transmission limits are enforced by constraints (4-10) and (4-11). M is a big value so that when z is 1, $f_{ij} = B_{ij}\left(\theta_i - \theta_j\right)$; otherwise, the constraints (4-8) and (4-9) are relaxed. Here, with a direct current optimal power flow approximation, we assume the voltage angle ranges are within ±0.6, so that $M_{ij} = 1.2B_{ij}$ is big enough.

$$\max_{q,\theta,f} \sum_j \left(\frac{1}{2}b_j q_j^2 + a_j q_j\right) \quad (4\text{-}4)$$

$$s.t. q_j + \sum_{ji} f_{ji} - \sum_{ij} f_{ij} = y_j, \ \forall j \in N \ \left[p_j\right] \quad (4\text{-}5)$$

$$\theta_i \le \theta_j^{max},\ \forall j \in N\ \left[\alpha_j^+ \ge 0\right] \tag{4-6}$$

$$-\theta_j \le -\theta_j^{min},\ \forall j \in N\ \left[\alpha_j^- \ge 0\right] \tag{4-7}$$

$$f_{ij} - B_{ij}\left(\theta_i - \theta_j\right) \le \left(1 - z_{ij}\right)M_{ij},\ \forall ij \in L\ \left[\gamma_{ij}^+ \ge 0\right] \tag{4-8}$$

$$-f_{ij} + B_{ij}\left(\theta_i - \theta_j\right) \le \left(1 - z_{ij}\right)M_{ij},\ \forall ij \in L\ \left[\gamma_{ij}^- \ge 0\right] \tag{4-9}$$

$$f_{ij} \le z_{ij}K_{ij},\ \forall ij \in L\ \ [\lambda_{ij}^+ \ge 0] \tag{4-10}$$

$$-f_{ij} \le z_{ij}K_{ij},\ \forall ij \in L\ \ [\lambda_{ij}^- \ge 0] \tag{4-11}$$

$$q_j \ge 0,\ \forall j \in N\ \ [\delta_j \ge 0] \tag{4-12}$$

Simultaneously, each GENCO i maximizes its operational profit with anticipation of their decisions' effect on the reference price, η, and determines its quantity to sell [32]. Equation (4-14) implies balance of total demand and generation, (4-15) indicates that the generation level must not exceed its capacity, and (4-16) imposes a constraint on fuel availability, which also implies an upper bound on the generation level y and expanded capacity V_i^{new}, that restricts the feasible region and makes the problem computationally easier to solve.

$$max_{y_i,\eta}\ (\eta + \phi_i - c_i - e_i y_i)\, y_i \tag{4-13}$$

$$s.t.\, y_i + \eta \textstyle\sum_j \frac{1}{b_j} = \sum_j \frac{\phi_j - a_j}{b_j} - \sum_{j \ne i} y_j\ \ [\beta_i] \tag{4-14}$$

$$y_i \le V_i^{new}\ \ [\mu_i \ge 0] \tag{4-15}$$

$$y_i \le U_i\ \ [\rho_i \ge 0] \tag{4-16}$$

$$y_i \ge 0\ \ \ [\zeta_i \ge 0] \tag{4-17}$$

The reference node LMP and price premia at non-reference nodes that appear in (4-13) and (4-14) are respectively defined in equations (4-18) and (4-19), which link the dual variables, p_j , in the ISO's problem and the reference price, η , in the GENCO's problem.

$$\eta = p_{ref} \tag{4-18}$$

$$\phi_j = p_j - \eta, \ \forall j \in N \tag{4-19}$$

Working from the bottom to the top, the sub-problem in the third level is an equilibrium problem formed by combining equations (4-13) – (4-17) for each GENCO with (4-4) – (4-12), (4-18), and (4-19); the bi-level sub-problem in the lower two levels consists of the objective function (4-2) subject to constraint (4-3) and the decision variables (q, etc,), with corresponding dual variables, being optimal in the lower level equilibrium problem (4-4) – (4-19); and the entire tri-level program includes the objective function (4-1) subject to the decision variables V^{new} being optimal in the bi-level sub-problem (4-2) – (4-19).

4.3 Algorithm to Solve an EPEC Sub-problem

To approach the optimization result of the tri-level expansion problem with electricity market, we first study its bi-level sub-problem, equations (4-2) – (4-19), without considering the ISO's centralized transmission expansion decisions in the first level. It involves generation expansion decisions from multiple GENCOs, where each of their separate optimization problems is a bi-level problem with multiple followers including the other GENCOs and the ISO. Each GENCO's bi-level problem in equations (4-2) – (4-19), can be reformulated as an MPEC by replacing the lower level optimization problems in equations (4-4) – (4-17) with their equivalent first-order optimality conditions in equations (4-20) – (4-

35) [4]. Each perpendicular constraint, e.g. $\theta^{max} - \theta_j \perp \alpha_j^+$, can be further converted to an equivalent nonlinear reformulation, e.g., $(\theta^{max} - \theta_j)\alpha_j^+ = 0$. Therefore, for each equilibrium constraint, there are three corresponding dual variables: two for the equations or inequalities and one for the nonlinear reformulation of the perpendicular relationship, e.g., in equation (4-24), $\mu\theta_{kj}^+ \geq 0$ and $\mu\alpha_{kj}^+ \geq 0$ are dual variables for equations $\theta^{max} - \theta_j \geq 0$ and $\alpha_j^+ \geq 0$, respectively while $\mu\theta\alpha_{kj}^+$ is the dual variable for the equation $(\theta^{max} - \theta_j)\alpha_j^+ = 0$. In each GENCO k's problem, the other GENCOs' capacity expansion decisions are considered as fixed parameters. The subscript of dual variables starting with k indicates the specific sets of dual variables for GENCO k. Also the dual variables for (4-18) and (4-19) are, respectively, $\mu\eta_k$ and $\mu\phi_{kj}$.

$$b_j q_j + a_j - p_j + \delta_j = 0,\ \forall j \in N\ \ [\mu dq_{kj}] \tag{4-20}$$

$$-\alpha_j^+ + \alpha_j^- - \textstyle\sum_{i,} \gamma_{ij}^+ B_{ij} - \sum_i \gamma_{ji}^+ B_{ji} + \sum_i \gamma_{ij}^- B_{ij} - \sum_i \gamma_{ji}^- B_{ji} = 0,\ \forall j \in N\ \ [\mu d\theta_{kj}] \tag{4-21}$$

$$p_j - p_i - \gamma_{ij}^+ + \gamma_{ij}^- - \lambda_{ij}^+ + \lambda_{ij}^- = 0,\ \forall ij \in L\ \ [\mu df_{kij}] \tag{4-22}$$

$$q_j + \textstyle\sum_{ji} f_{ji} - \sum_{ij} f_{ij} = y_j,\ \forall j \in N\ \ [\mu qy_{kj}] \tag{4-23}$$

$$\theta^{max} - \theta_j \geq 0 \perp \alpha_j^+ \geq 0, \forall j \in N\ [\mu\theta_{kj}^+ \geq 0, \mu\alpha_{kj}^+ \geq 0, \mu\theta\alpha_{kj}^+] \tag{4-24}$$

$$\theta_j - \theta^{min} \geq 0 \perp \alpha_j^- \geq 0, \forall j \in N\ [\mu\theta_{kj}^- \geq 0, \mu\alpha_{kj}^- \geq 0, \mu\theta\alpha_{kj}^-] \tag{4-25}$$

$$-f_{ji} + B_{ji}(\theta_j - \theta_i) + (1 - z_{ji})B_{ji}1.2 \geq 0 \perp \gamma_{ij}^+ \geq 0, \forall ij \in L\ \ [\mu f_{kij}^+ \geq 0, \mu\gamma_{kij}^+ \geq 0, \mu f\gamma_{kij}^+] \tag{4-26}$$

$$f_{ji} - B_{ji}(\theta_j - \theta_i) + (1 - z_{ji})B_{ji}1.2 \geq 0 \perp \gamma_{ij}^- \geq 0, \forall ij \in L\ \ [\mu f_{kij}^- \geq 0, \mu\gamma_{kij}^- \geq 0, \mu f\gamma_{kij}^-] \tag{4-27}$$

$$z_{ji}K_{ji} - f_{ji} \geq 0 \perp \lambda_{ij}^+ \geq 0, \forall ij \in L\ [\mu K_{kij}^+ \geq 0, \mu\lambda_{kij}^+ \geq 0, \mu K\lambda_{kij}^+] \tag{4-28}$$

$$z_{ji}K_{ji} + f_{ji} \geq 0 \perp \lambda_{ij}^- \geq 0, \forall ij \in L\ [\mu K_{kij}^+ \geq 0, \mu\lambda_{kij}^+ \geq 0, \mu K\lambda_{kij}^+] \tag{4-29}$$

$$q_j \geq 0 \perp \delta_j \geq 0, \forall j \in N\ \ [\mu q_{kj} \geq 0, \mu\delta_{kj} \geq 0, \mu q\delta_{kj}] \tag{4-30}$$

$$\eta + \phi_j - c_j - 2e_j y_j - \beta_j - \mu_j - \rho_j + \zeta_j = 0, \forall j \in N_{gen} \left[\mu d y_{kj}\right] \qquad (4\text{-}31)$$

$$y_j + \beta_j \sum_i \frac{1}{b_i} = 0,\ \forall j \in N_{gen} \ \left[\mu d\eta_{kj}\right] \qquad (4\text{-}32)$$

$$V_j^{new} - y_j \geq 0 \perp \mu_j \geq 0, \forall j \in N \left[\mu n V_{kj} \geq 0, \mu\mu_{kj} \geq 0, \mu n V\mu_{kj}\right] \qquad (4\text{-}33)$$

$$U_j - y_j \geq 0 \perp \rho_j \geq 0, \forall j \in N \left[\mu U_{kj} \geq 0, \mu\rho_{kj} \geq 0, \mu U\rho_{kj} \geq 0\right] \qquad (4\text{-}34)$$

$$y_j \geq 0 \perp \zeta_j \geq 0,\ \forall j \in N \ \left[\mu y_{kj} \geq 0, \mu\zeta_{kj} \geq 0, \mu y\zeta_{kj}\right] \qquad (4\text{-}35)$$

4.3.1 Diagonalization Method (DM)

One way to find an equilibrium solution, if one exists, is to iteratively solve each GENCO's problem by fixing the other GENCOs' expansion decisions to their current optimal solutions, which is called DM in [4]. In other words, the optimal solution determined by each GENCO should be identical to the value that the other GENCOs assume as a model parameter of their own optimization models. However, the existence of a pure Nash equilibrium (NE) strategy is not guaranteed for the EPEC sub-problem, and the GENCOs' expansion decisions, V_k^{new} , can oscillate, usually among two or more different values within a small range, generally by 1-3% and at most 5% from our computational experience. Therefore, we define a maximum number of iteration cycles and an approximate NE solution as the average of the subsequential limiting solutions, which will be further illustrated in Part II of this paper in Chapter 5.

Table 4- 5 Solving the EPEC Sub-problem by DM Algorithm

DM Algorithm

Input parameters $V_{2,..,|N_{gen}|}^{newequ}$;

Let ConvergenceFlag = 0, Cycle=0;

Table 4- 5 (continued)

While (ConvergenceFlag = 0 and Cycle ≤MaxCycle)

Let ConvergenceFlag = 1;

For GENCO k = 1 to $|N_{gen}|$

Cycle = Cycle+1;

Solve GENCO k's MPEC problem, Equations (4-2), (4-3), (4-18), (4-19), (4-20)-(4-35), with optimal solution V_k^{new*};

If $\left|V_k^{new*} - V_k^{newequ}\right| \geq \varepsilon$

ConvergenceFlag = 0;

End If;

Let $V_k^{newequ} = nV_k^{new*}$;

End For

End While

Output $V_{1,..,|N_{gen}|}^{new} = V_{1,..,|N_{gen}|}^{newequ}$;

The DM algorithm is illustrated in Table 4-5. First, we solve the MPEC for GENCO 1 by initializing the values of other GENCOs' expansion decisions $V_{2,..|N_{gen}|}^{newequ}$ as model parameters. Once the optimal solution V_1^{new*} for GENCO 1 is obtained, it is considered as a model input V_1^{newequ} to the next MPEC problem of GENCO 2 while $V_{3,..,|N_{gen}|}^{newequ}$ remain the same for the new problem. The iteration continues until the MPEC problem of the final GENCO is solved. Because the decision V_k^{new} can only be changed when the MPEC of

GENCO k is solved, V_k^{new} is updated once during a round of n iterations (one for each GENCO). If, after a round of iterations, the changes in value of each GENCO's decision are all within the predefined threshold ε, we conclude that an equilibrium has been identified and stop the iteration process. Otherwise, we continue the iteration until the predetermined limit is reached. The MPEC problem can be solved in GAMS by using the solver NLPEC [34], which first reformulates our MPEC into a nonlinear program (NLP), and then calls the NLP solver CONOPT [35] to solve the problem.

4.3.2 Complementarity Problem (CP) Reformulation

Another way to find a NE solution of the equilibrium bi-level sub-problem is through CP reformulation [4]. Combining the MPEC problem for each GENCO results in an equilibrium program with equilibrium constraints (EPEC). The CP reformulation combines the KKT conditions of each MPEC to reformulate the EPEC into a mixed complementarity problem (MCP). Given the lack of convexity of each MPEC, optimality to the CP reformulation is only a necessary condition for a solution to be an equilibrium of the original bi-level games, but not a sufficient condition. Specifically, the solution found by CP reformulation is a stationary point of the original EPEC problem. The derivation of the CP reformulation can be found in Appendix 4.A.

4.4 A Hybrid Iterative Algorithm to Solve the Tri-level Programming Problem

Upon CP reformulation of equations (4-2) – (4-19), the tri-level problem (4-1) – (4-19), can be converted into a single level optimization problem with a set of nonlinear, linear and complementarity constraints (4-38) – (4-95) as shown in Appendix 4.A. The reformulated problem, consisting of the equations (4-1) and (4-38) – (4-95), is a generalized MPEC with mixed integer, linear, nonlinear and equilibrium constraints, and can be solved by the NLPEC solver in GAMS. The NLPEC solver provides several different reformulation methodologies to approach MPEC problems [34]. The type of MPEC reformulation we found the most successful and reliable is to penalize violation of the equilibrium constraints in the objective function by first converting each equilibrium constraint, $0 \leq x \perp y \geq 0$, to its equivalent constraint set: $x \geq 0, y \geq 0, xy = 0$, and then including a term, $1/\mu^{pen}xy$, in the objective function. As the reciprocal penalty parameter μ^{pen} iteratively decreases to zero, the penalty applied to $xy \neq 0$ increases until solutions eventually approximate $xy = 0$ [34].

The two currently available algorithms to solve bi-level games as an EPEC are DM and CP reformulation. Based on our computational studies reported in Part II in Chapter 5, given a certain transmission expansion planning decision, the performance of DM is quite stable in successfully identifying the (approximate) Nash Equilibrium (NE) of the bi-level games; while the CP reformulation, since it is not an equivalent reformulation of the original bi-level game, can only find a stationary point and provide a bound for the original problem. However, the benefit of CP reformulation is to be able to solve the entire tri-level problem as a single level problem that includes the transmission expansion decisions z_{ij}. With all these considerations in mind, we propose a hybrid iterative algorithm that takes advantage of both methods. It first solves a master problem, transformed from the CP reformulation, to propose a transmission expansion decision $z^{master-n}$ in the *n*-th iteration. Given that transmission expansion plan, it employs DM to find an (approximate) NE point of the game of bi-level games. The iterative solution procedure is illustrated in Figure 4-2, and a detail explanation of the hybrid algorithm is in Appendix 4.B.

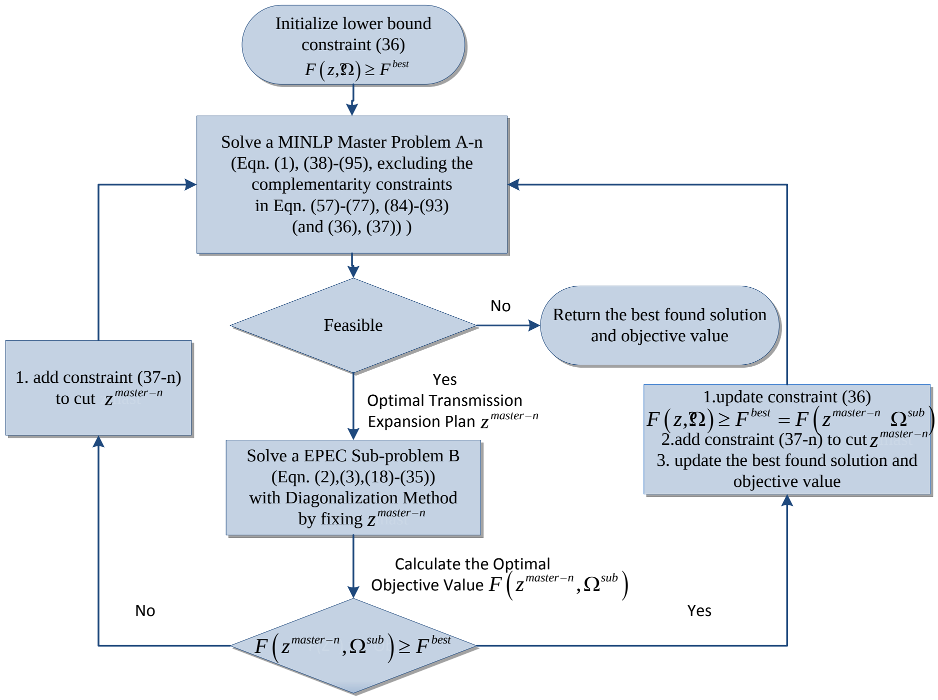

Figure 4- 2 A Hybrid Iterative Algorithm to Solve a Tri-level Problem with an EPEC Sub-problem

The steps of the algorithm are detailed as follows:

- Step 1: Set n to 0 and initialize the best found system net surplus, F^{best}, to 0, and let the objective value lower bound constraint (4-36) of MINLP master problem greater than F^{best}. Go to Step 2.

$$F(z, \Omega) \geq F^{best},$$

where $\Omega = \{V_j^{new}, q, y\}$, $F(z,\Omega) = \sum_j \left(\frac{1}{2} b_j q_j^2 + a_j q_j\right) - \sum_{j \in N_{gen}} (c_j y_j + e_j y_j^2) - \sum_{ij \in L} c_{ij}^{trexp} K_{ij} z_{ij} - \sum_{j \in N_{gen}} c_j^{gexp} (V_j^{new} - V_j)$ (4-36)

- Step 2: Solve a MINLP master problem A-n. If the MINLP master problem A-n can be successfully solved to an optimal solution, we fix the transmission expansion decision

$z^{master-n}$ and continue to solve the EPEC sub-problem B with DM. Go to Step 3. Otherwise, the algorithm is terminated and best solution found so far is returned as the final solution.

- Step 3: With optimal solution $\Omega^{sub} = \{V_j^{new\,sub}, q^{sub}, y^{sub}\}$ found by DM, the system net surplus $F(z^{master-n}, \Omega^{sub})$ is calculated. If $F(z^{master-n}, \Omega^{sub}) \geq F^{best}$, go to Step 4. Otherwise, go to Step 5.
- Step 4: Update constraint (4-36) with $F^{best} = F(z^{master-n}, \Omega^{sub})$, and add a constraint (4-37-*n*) to cut $z^{master-n}$ point and update the best found solution and its objective value, $F(z^{master-n}, \Omega^{sub})$. Let $n = n + 1$, and go to Step 2.

$$\sum_{ij, z_{ij}^{master-n}=1} (1 - z_{ij}) + \sum_{ij, z_{ij}^{master-n}=0} (z_{ij}) \geq 1 \qquad \text{(4-37-}n\text{)}$$

- Step 5: Add a constraint (4-37-*n*) to cut z^{master}. Let $n = n + 1$, and go to Step 2.

4.5 Conclusions

In this paper, we consider an integrated transmission and generation expansion planning problem for a restructured electricity market environment. We propose a novel tri-level programming model, where a centralized transmission expansion planning decision in the top level is made in anticipation of the multi-GENCOs' responses in terms of their generation expansion decisions in the middle level, while each GENCO also makes its capacity expansion decision by anticipating the electricity market equilibrium results achieved by all the GENCOs making their generation decisions, and an ISO's market clearing problem in the bottom level.

The tri-level programming model includes an EPEC problem, which can be approached by either the diagonalization method (DM) or a complementarity problem (CP) reformulation. To solve the tri-level optimization problem, a hybrid iterative algorithm is proposed by taking advantage of the strengths of both DM algorithm and CP reformulation. The benefit of applying CP reformulation is its capability to transform the tri-level model into a single level MINLP problem, which can be solved by the DICOPT solver, and identify a promising transmission planning decision in each iteration. On the other hand, given a preselected transmission expansion decision, the DM algorithm works more reliably and efficiently to find the corresponding (approximate) Nash Equilibrium (NE) point for the generation expansion bi-level games.

The problem we consider in this paper is a static model that considers only a single hour in a future year, which will always result in the generation levels y being equal to the new capacity levels V^{new}. In the future, we can extend the model and the algorithm to take into account multiple periods and uncertainties.

Part II of the paper in Chapter 5 will continue with numerical results.

Appendix 4.A CP Reformulation for Multiple GENCO's EPEC Sub-problem

The CP reformulation for the EPEC sub-problem, equations (4-38) – (4-95), are derived as following, based on [4]. L represents the Lagrange function of the second level problem, equations (4-2) – (4-19), with nonlinear reformulation, equations (4-20) – (4-35), replacing the third level problem, equations (4-4) – (4-19).

4.A.1 Partial Derivatives of Lagrange Function

There are all together 19 sets of equality constraints with dimension $|N_{gen}| + 7|N_{gen}||N| + 5|N_{gen}||L| + 6|N_{gen}||N_{gen}|$. They have the same number of unrestricted variables to match them, respectively μdq_{kj}, $\mu d\theta_{kj}$, μdf_{kij}, μdy_{kj}, $\mu d\eta_{kj}$, $\mu U\rho_{kj}$, μqy_{kj}, $\mu\theta\alpha_{kj}^{+}$, $\mu\theta\alpha_{kj}^{-}$, $\mu f\gamma_{ij}^{+}$, $\mu f\gamma_{ij}^{-}$, $\mu K\lambda_{ij}^{+}$, $\mu K\lambda_{ij}^{-}$, $\mu q\delta_{kj}$, μyy_{kj}, $\mu nV\mu_{kj}$, $\mu y\zeta_{kj}$, $\mu\eta_k$, $\mu\phi_{kj}$.

$$L_{nV_k} = -c_k^{gexp} + \mu nVV_k + \mu nV_{k,k} + \mu nV\mu_{kk}\mu_k = 0, \forall k \in N_{gen}, k \in N_{gen} \quad (4\text{-}38)$$

$$L_{q_j} = \mu dq_{kj}b_j + \mu qy_{kj} + \mu q_{kj} + \mu q\delta_{kj}\delta_j = 0, \ \forall k \in N_{gen}, j \in N \quad (4\text{-}39)$$

$$L_{\theta_j} = -\mu\theta_{kj}^{+} - \mu\theta\alpha_{kj}^{+}\alpha_j^{+} + \mu\theta_{kj}^{-} + \mu\theta\alpha_{kj}^{-}\alpha_j^{-} - \sum_i(\mu f_{kij}^{+}B_{ij} + \mu f\gamma_{kij}^{+}\gamma_{ij}^{+}B_{ij} - \mu f_{kij}^{-}B_{ij} - \mu f\gamma_{kij}^{-}\gamma_{ij}^{-}B_{ij}) + \sum_i(\mu f_{kji}^{+}B_{ji} + \mu f\gamma_{kji}^{+}\gamma_{ji}^{+}B_{ji} - \mu f_{kji}^{-}B_{ji} - \mu f\gamma_{kji}^{-}\gamma_{ji}^{-}B_{ji}) = 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}40)$$

$$L_{f_{ij}} = -\mu qy_{kj} + \mu qy_{ki} - \mu f_{kij}^{+} - \mu f\gamma_{kij}^{+}\gamma_{ij}^{+} + \mu f_{kij}^{-} + \mu f\gamma_{kij}^{-}\gamma_{ji}^{-} - \mu K_{kij}^{+} - \mu K\lambda_{kij}^{+}\lambda_{ij}^{+} + \mu K_{kij}^{-} + \mu K\lambda_{kij}^{-}\lambda_{ij}^{-} = 0, \forall k \in N_{gen}, ij \in L \quad (4\text{-}41)$$

$$L_{y_j} = p_{j=k} - c_{j=k} - 2e_{j=k}y_{j=k} - 2e_j\mu dy_{kj} + \mu d\eta_{kj} + \sum_j \mu yy_j - \mu nV_{kj} - \mu nV\mu_{kj}\mu_j + \mu y_{kj} + \mu y\zeta_{kj}\zeta_j - \mu qy_{kj} = 0, \forall k \in N_{gen}, j \in N_{gen} \quad (4\text{-}42)$$

$$L_{\eta} = \sum_j \mu dy_{kj} + \mu\eta_k + \sum_j \mu\phi_{kj} = 0, \forall k \in N_{gen} \quad (4\text{-}43)$$

$$L_{\phi_j} = \mu dy_{kj} + \mu\phi_{kj} = 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}44)$$

$$L_{p_j} = y_{j=k} - \mu dq_{kj} + \sum_i \mu df_{kij} - \sum_i \mu df_{kji} - \mu\phi_{kj} = 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}45)$$

$$L_{\alpha_j^{+}} = -\mu d\theta_{kj} + \mu\alpha_{kj}^{+} + \mu\theta\alpha_{kj}^{+}(\theta^{max} - \theta_j) = 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}46)$$

$$L_{\alpha_j^{-}} = \mu d\theta_{kj} + \mu\alpha_{kj}^{-} + \mu\theta\alpha_{kj}^{-}(\theta_j - \theta^{min}) = 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}47)$$

$$L_{\gamma_{ij}^{+}} = -\mu d\theta_{kj}B_{ij} + \mu d\theta_{ki}B_{ij} - \mu df_{kij} + \mu\gamma_{kij}^{+} + \mu f\gamma_{kij}^{+}[-f_{ij} + B_{ij}(\theta_i - \theta_j) +$$

$$(1 - z_{ij})M_{ij}] = 0, \forall k \in N_{gen}, ij \in L \quad (4\text{-}48)$$

$$L_{\gamma_{ij}^{-}} = \mu d\theta_{kj}B_{ij} - \mu d\theta_{ki}B_{ij} + \mu df_{kij} + \mu\gamma_{kij}^{-} + \mu f\gamma_{kij}^{+}[f_{ij} - B_{ij}(\theta_i - \theta_j) +$$

$$(1 - z_{ij})M_{ij}] = 0, \forall k \in N_{gen}, ij \in L \quad (4\text{-}49)$$

$$L_{\lambda_{ij}^{+}} = -\mu df_{kij} + \mu\lambda_{kij}^{+} + \mu K\lambda_{kij}^{+}(z_{ij}K_{ij} - f_{ij}) = 0, \forall k \in N_{gen}, ij \in L \quad (4\text{-}50)$$

$$L_{\lambda_{ij}^{-}} = \mu df_{kij} + \mu\lambda_{kij}^{-} + \mu K\lambda_{kij}^{-}(z_{ij}K_{ij} + f_{ij}) = 0, \forall k \in N_{gen}, ij \in L \quad (4\text{-}51)$$

$$L_{\delta_j} = \mu dq_{kj} + \mu\delta_{kj} + \mu q\delta_{kj}q_j = 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}52)$$

$$L_{\beta_j} = -\mu dy_{kj} + \mu d\eta_{kj}\sum_j \frac{1}{b_j} = 0, \forall k \in N_{gen}, j \in N_{gen} \quad (4\text{-}53)$$

$$L_{\mu_j} = -\mu dy_{kj} + \mu\mu_{kj} + \mu nV\mu_{kj}(V_j^{new} - y_j) = 0, \forall k \in N_{gen}, j \in N_{gen} \quad (4\text{-}54)$$

$$L_{\rho_j} = -\mu dy_{kj} + \mu\rho_{kj} + \mu U\rho_{kj}(U_j - y_j) = 0, \forall k \in N_{gen}, j \in N_{gen} \quad (4\text{-}55)$$

$$L_{\zeta_j} = \mu dy_{kj} + \mu\zeta_{kj} + \mu y\zeta_{kj}y_j = 0, \forall k \in N_{gen}, j \in N_{gen} \quad (4\text{-}56)$$

4.A.2 KKT Conditions Derived from the Optimization Problems in the Second Level

There are all together 21 sets of inequality constraints with dimension of $|\mathrm{N_{gen}}| + 6|\mathrm{N_{gen}}||\mathrm{N}| + 8|\mathrm{N_{gen}}||\mathrm{L}| + 6|\mathrm{N_{gen}}||\mathrm{N_{gen}}|$. They have the same number of positive variables to match them, showed as the dual variables in the constraints.

$$V_k^{new} - V_k \geq 0 \perp \mu nVV_k \geq 0, \forall k \in N_{gen} \quad (4\text{-}57)$$

$$\theta^{max} - \theta_j \geq 0 \perp \mu\theta_{kj}^{+} \geq 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}58)$$

$$\theta_j - \theta^{min} \geq 0 \perp \mu\theta_{kj}^{-} \geq 0, \forall k \in N_{gen}, j \in N \quad (4\text{-}59)$$

$$-f_{ij} + B_{ij}(\theta_i - \theta_j) + (1 - z_{ij})M \geq 0 \perp \mu f_{kij}^{+} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-60}$$

$$f_{ij} - B_{ij}(\theta_i - \theta_j) + (1 - z_{ij})M \geq 0 \perp \mu f_{kij}^{-} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-61}$$

$$z_{ij}K_{ij} - f_{ij} \geq 0 \perp \mu K_{kij}^{+} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-62}$$

$$z_{ij}K_{ij} + f_{ij} \geq 0 \perp \mu K_{kij}^{-} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-63}$$

$$q_j \geq 0 \perp \mu q_{kj} \geq 0, \forall k \in N_{gen}, j \in N \tag{4-64}$$

$$\alpha_j^{+} \geq 0 \perp \mu \alpha_{kj}^{+} \geq 0, \forall k \in N_{gen}, j \in N \tag{4-65}$$

$$\alpha_j^{-} \geq 0 \perp \mu \alpha_{kj}^{-} \geq 0, \forall k \in N_{gen}, j \in N \tag{4-66}$$

$$\gamma_{ij}^{+} \geq 0 \perp \mu \gamma_{kij}^{+} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-67}$$

$$\gamma_{ij}^{-} \geq 0 \perp \mu \gamma_{kij}^{-} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-68}$$

$$\lambda_{ij}^{+} \geq 0 \perp \mu \lambda_{kij}^{+} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-69}$$

$$\lambda_{ij}^{-} \geq 0 \perp \mu \lambda_{kij}^{-} \geq 0, \forall k \in N_{gen}, ij \in L \tag{4-70}$$

$$\delta_j \geq 0 \perp \mu \delta_{kj} \geq 0, \forall k \in N_{gen}, j \in N \tag{4-71}$$

$$V_j^{new} - y_j \geq 0 \perp \mu n V_{kj} \geq 0, \forall k \in N_{gen}, j \in N_{gen} \tag{4-72}$$

$$U_j - y_j \geq 0 \perp \mu U_{kj} \geq 0, \forall k \in N_{gen}, j \in N_{gen} \tag{4-73}$$

$$y_j \geq 0 \perp \mu y_{kj} \geq 0, \forall k \in N_{gen}, j \in N_{gen} \tag{4-74}$$

$$\mu_j \geq 0 \perp \mu \mu_{kj} \geq 0, \forall k \in N_{gen}, j \in N_{gen} \tag{4-75}$$

$$\rho_j \geq 0 \perp \mu \rho_{kj} \geq 0, \forall k \in N_{gen}, j \in N_{gen} \tag{4-76}$$

$$\zeta_j \geq 0 \perp \mu \zeta_{kj} \geq 0, \forall k \in N_{gen}, j \in N_{gen} \tag{4-77}$$

4.A.3 Equivalent KKT Conditions Derived from the Optimization Problems in the Third Level

There are all together 18 sets of constraints, among which there are 8 sets of equality constraints with dimension of $1 + 2|\mathrm{N_{gen}}| + 4|\mathrm{N}| + |\mathrm{L}|$ and 10 sets of inequality constraints with dimension of $3|\mathrm{N_{gen}}| + 3|\mathrm{N}| + 4|\mathrm{L}|$. The equality constraints have the same amount of unrestricted variables to match them, respectively q_j, θ_j, f_{ij}, y_j, η, p_j, β_j, ϕ_j, while the inequality constraints have the same number of positive variables to match them, showed as the dual variables in the constraints.

$$b_j q_j + a_j - p_j + \delta_j = 0, \forall j \in N \tag{4-78}$$

$$-\alpha_j^+ + \alpha_j^- - \sum_{i,ij\in L}\left(B_{ij}\gamma_{ij}^+ - B_{ij}\gamma_{ij}^-\right) + \sum_{i,ji\in L}\left(B_{ji}\gamma_{ji}^+ - B_{ji}\gamma_{ji}^-\right) = 0, \forall j \in N \tag{4-79}$$

$$p_j - p_i - \gamma_{ij}^+ + \gamma_{ij}^- - \lambda_{ij}^+ + \lambda_{ij}^- = 0, \forall ij \in L \tag{4-80}$$

$$\eta + \phi_j - c_j - 2e_j y_j - \beta_j - \mu_j - \rho_j + \zeta_j = 0, \forall j \in N_{gen} \tag{4-81}$$

$$y_j + \beta_j \sum_j \frac{1}{b_j} = 0, \forall j \in N_{gen} \tag{4-82}$$

$$q_j + \sum_i f_{ji} - \sum_i f_{ij} = y_j, j \in N \tag{4-83}$$

$$\theta^{max} - \theta_j \geq 0 \perp \alpha_j^+ \geq 0, \forall j \in N \tag{4-84}$$

$$\theta_j - \theta^{min} \geq 0 \perp \alpha_j^- \geq 0, \forall j \in N \tag{4-85}$$

$$-f_{ij} + B_{ij}\left(\theta_i - \theta_j\right) + \left(1 - z_{ij}\right)M_{ij} \geq 0 \perp \gamma_{ij}^+ \geq 0, \forall ij \in L \tag{4-86}$$

$$f_{ij} - B_{ij}\left(\theta_i - \theta_j\right) + \left(1 - z_{ij}\right)M_{ij} \geq 0 \perp \gamma_{ij}^- \geq 0, \forall ij \in L \tag{4-87}$$

$$z_{ij}K_{ij} - f_{ij} \geq 0 \perp \lambda_{ij}^+ \geq 0, \forall ij \in L \tag{4-88}$$

$$z_{ij}K_{ij} + f_{ij} \geq 0 \perp \lambda_{ij}^- \geq 0, \forall ij \in L \tag{4-89}$$

$$q_j \geq 0 \perp \delta_j \geq 0, \forall j \in N \tag{4-90}$$

$$V_j^{new} - y_j \geq 0 \perp \mu_j \geq 0, j \in N_{gen} \tag{4-91}$$

$$U_j - y_j \geq 0 \perp \rho_j \geq 0, j \in N_{gen} \tag{4-92}$$

$$y_j \geq 0 \perp \zeta_j \geq 0, j \in N_{gen} \tag{4-93}$$

$$\eta = p_{ref} \tag{4-94}$$

$$\phi_j = p_j - \eta, \forall j \in N \tag{4-95}$$

Appendix 4.B Details of the Hybrid Algorithm Solving the Tri-level Problem

To improve the computational efficiency and algorithm stability, the complementarity constraints are excluded from the master problem. Therefore the initial master problem A-*0* includes the objective function (4-1), and all the nonlinear, and linear constraints (4-38) – (4-95) converted from the CP reformulation except for the complementarity constraints (4-57) – (4-77) and (4-84) – (4-93). The master problem is updated after each iteration and defined as master problem A-*n* after the *n*th iteration. For $n \geq 1$, the problem A-*n* includes constraint (4-36) to update the best lower bound for the objective function, and constraints (4-37-*m*), $m = 1, .., n$, to duplicating a previously proposed $\mathrm{z}^{\mathrm{master-n}}$. Therefore, problem A-*n* is a mixed integer nonlinear program (MINLP), which can be solved by the DICOPT solver [36] in GAMS. The sub-problem as bi-level games is equivalently reformulated as an EPEC sub-problem B including equations (4-2), (4-3) and (4-18) – (4-35) for all the GENCOs.

Here, we define the original tri-level problem C, (4-1) – (4-19); and the reformulated single level problem D, (4-1) and (4-38) – (4-95), based on CP reformulation. The EPEC sub-problem B shares the same feasible region as the original tri-level problem C. The reformulated single level problem D solves for the stationary points, including the optimal solution, for the original problem C. The master problem A-*0* is a further relaxation of problem D by removing all the complementarity constraints, and thus the optimal solution for the original problem C is also included in the feasible region of master problem A-*0*. However the feasible solution for problem A-*0* is not necessarily feasible for the original tri-level problem C, or the EPEC sub-problem B.

The purpose of solving the MINLP master problem A-*n* at each major iteration in Step 2 is to search for a possibly better transmission plan $z^{master-n}$ with $F(z^{master-n}, \Omega^{master-n}) \geq F^{best}$. However, since the feasible solution for master problem A-*0* is not necessarily feasible to the original tri-level problem C, the $\Omega^{master-n}$ solved might not be feasible for the original problem due to the relaxation of the complementarity constraints in (4-57) – (4-77) and (4-84) – (4-93) from the CP reformulation.

To examine whether the $z^{master-n}$ indeed results in a higher net surplus than F^{best}, an EPEC sub-problem B, which has the same feasible region as the original tri-level problem C, is solved in Step 3 given $z^{master-n}$. If the $F\left(z^{master-n}, \Omega^{sub}\right)$ is better than F^{best}, $z^{master-n}$ is confirmed to be better, so we updated constraint (4-36) by letting F^{best} equal to $F\left(z^{master-n}, \Omega^{sub}\right)$. Meanwhile, no matter whether $F\left(z^{master-n}, \Omega^{sub}\right)$ is better than F^{best} or not, we add constraint (4-37-n) to prevent the $z^{master-n}$ solution from being found in the MINLP master problem again. With the iteratively reduced feasible region of the master problem constrained by equations (4-36) and (4-37-n), it will ultimately arrive a point where there is no longer any feasible solution z that can generate a higher surplus than F^{best} in the master problem. Since the optimal solution of the original tri-level problem C must be a feasible solution of the initial master problem A-*0* transformed from the CP reformulation, having relaxed the complementarity constraints, there is no need to continue the search. Thus the algorithm terminates with the best z found so far and its corresponding F^{best}.

References

[1] P.L. Joskow, "Markets for power in the United States: An interim assessment," Energy Journal, vol. 27, no. 1, pp. 1-36, 2006.

[2] M. Awad, K.E. Casey, A.S. Geevarghese, J.C. Miller, A.F. Rahimi, A.Y. Sheffrin, M. Zhang, E. Toolson, G. Drayton, B.F. Hobbs, and F.A. Wolak, "Economic Assessment of Transmission Upgrades: Application of the California ISO Approach," in Restructured Electric Power Systems: Analysis of Electricity Markets with Equilibrium Models, Power Engineering Series, Ed.: J. Wiley & Sons/IEEE Press, July 2010, ch. 7, pp. 241-270.

[3] S. Jin and S.M.Ryan, "Capacity expansion in the integrated supply network for an electricity market," IEEE Transactions on Power Systems, vol. 26, no. 4, pp. 2275–2284, 2011.

[4] X. Hu and D. Ralph, "Using EPECs to model bilevel games in restructured electricity markets with locational prices," Operations Research, vol. 55, no. 5, pp. 809–827, 2007.

[5] R. Garcia-Bertrand, D. Kirschen and A.J. Conejo, "Optimal investments in generation capacity under uncertainty," in Proceedings of 16th Power Systems Computation Conference, 2008.

[6] L.P. Garces, A.J. Conejo, R. Garcia-Bertrand and R. Romero, "A bilevel approach totransmission expansion planning within a market environment," IEEE Transactionson Power Systems, vol. 24, no. 3, pp. 1513–1522, 2009.

[7] V. Nanduri, T.K. Das and P. Rocha, "Generation capacity epxansion in energy markets using a two-level game theoretic model," IEEE Transactions on Power Systems, vol. 24, no. 3, pp. 1165–1172, 2009.

[8] J. Wang, M. Shahidehpour, Z. Li and A. Botterud, "Strategic generation capacity expansion planning with incomplete information," IEEE Transactions on Power Systems, vol. 24, no. 2, pp. 1002–1010, 2009.

[9] J. Su and F.F. Wu, "Evaluation of generation expansion investment under competitive market environment," in Proceedings of IEEE Power and Energy Society General Meeting, 2005, pp. 2136–2140.

[10] T. Li and M. Shahidehpour, "Strategic bidding of transmission-constrained GENCOs with incomplete information," IEEE Transactions on Power Systems, vol. 20, no. 1, pp. 437–447, 2005.

[11] S. Soleymani, A.M. Ranjbar and A.R. Shirani, "New approach to bidding strategies of generating companies in day ahead energy market," Energy Conversion and Management, vol. 49, no. 6, pp. 1493–1499, 2008.

[12] C. Ruiz, A.J. Conejo and Y. Smeers, "Equilibria in an oligopolistic electricity pool with stepwise offer curves," IEEE Transmission on Power Systems, vol. 27, no. 2, pp. 752-761, 2012.

[13] F.F. Wu, F.L. Zheng and F.S. Wen, "Transmission investment and expansion planning in a restructured electricity market," Energy, vol. 31, no. 6-7, pp. 954–966, 2006.

[14] B. Colson, P. Marcotte and G. Savard, "An overview of bilevel optimization," Annals of Operations Research, vol. 153, no. 1, pp. 235–256, 2007.

[15] S.T. DeNegre and T.K. Ralphs, "A branch-and-cut algorithm for integer bilevel linear programs," Operations Research and Cyber-Infrastructure, vol. 47, no. 2, pp. 65–78, 2009.

[16] J.T. Moore and J.F. Bard, "The mixed integer bilevel programming problem," Operations Research, vol. 38, no. 5, pp. 911–921, 1990.

[17] C. Ruiz and A.J. Conejo, "Pool strategy of a producer with endogenous formation of locational marginal prices," IEEE Transactions on Power Systems, vol. 24, no. 4, pp. 1855–1866, 2009.

[18] S.J. Kazempour, A.J. Conejo and C. Ruiz, "Strategic generation investment using a complementarity approach," IEEE Transactions on Power Systems, vol. 26, no. 2, pp. 940–948, 2011.

[19] S.A. Gabriel and F.U. Leuthold, "Solving discretely-constrained MPEC problems with applications in electric power markets," Energy Economics, vol. 32, no. 1, pp. 3–14, 2010.

[20] S. Wogrin, E. Centeno and J. Barquin, "Generation capacity expansion in liberalized electricity markets: a stochastic MPEC approach," IEEE Transactions on Power Systems, vol. 26, no. 4, pp. 2526 - 2532, 2011.

[21] S.J. Kazempour and A.J. Conejo, "Strategic generation investment under uncertainty via Benders decomposition," IEEE Transactions on Power Systems, vol. 27, no. 1, pp. 424-432, 2012.

[22] M.C. Ferris and J.S. Pang, "Engineering and economic applications of complementarity problems," SIAM Review, vol. 39, no. 4, pp. 669–713, 1997.

[23] J. Hu, J.E. Mitchell, J.S. Pang, K.P. Bennett and G. Kunapuli, "On the global solution of linear programs with linear complementarity constraints," SIAM Journalon Optimization, vol. 19, no. 1, pp. 445–471, 2008.

[24] A.S. Chuang, F. Wu and P. Varaiya, "A game-theoretic model for generation expansionplanning: Problem formulation and numerical comparisons," IEEE Transactionons on Power Systems, vol. 16, no. 4, pp. 885-891, 2001.

[25] F.H. Murphy and Y. Smeers, "Generation capacity expansion in imperfectly competitive restructured electricity markets," Operations Research, vol. 53, no. 4, pp. 646–661, 2005.

[26] E.E. Sauma and S.S. Oren, "Proactive planning and valuation of transmission investments in restructured electricity markets," Journal of Regulatory Economics, vol. 30, no. 3, pp. 261–290, 2006.

[27] J.H. Roh, M. Shahidehpour and L. Wu, "Market-based generation and transmission planning with uncertainties," IEEE Transactions on Power Systems, vol. 24, no. 3, pp. 1587–1598, 2009.

[28] A. Motamedi, H. Zareipour, M.O. Buygi and W.D. Rosehart, "A transmission planning framework considering future generation expansions in electricity markets," IEEE Transactions on Power Systems, vol. 25, no. 4, pp. 1987–1995, 2010.

[29] M.R. Hesamzadeh, D.R. Biggar, N. Hosseinzadeh and P.J. Wolfs, "Transmission augmentation with mathematical modeling of market power and strategic generation

expansion - part I," IEEE Transactions on Power Systems, vol. 26, no. 4, pp. 2040–2048, 2011.

[30] M.R. Hesamzadeh, D.R. Biggar, N. Hosseinzadeh and P.J. Wolfs, "Transmission augmentation with mathematical modeling of market power and strategic generation expansion - part II," IEEE Transactions on Power Systems, vol. 26, no. 4, pp. 2049–2057, 2011.

[31] D. Pozo, E.E. Sauma and J. Contreras, "A three-level static MILP model for generation and transmission expansion planning," IEEE Transactions on Power Systems, forthcoming.

[32] J. Yao, I. Adler and S.S. Oren, "Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network," Operations Research, vol. 56, no. 1, pp. 34-47, 2008.

[33] S.M. Ryan, "Market outcomes in a congested electricity system with fuel supply network," in Proceedings of the IEEE Power Engineering Society General Meeting, Calgary, 2009.

[34] NLPEC Solver Manual. [Online]. http://www.gams.com/dd/docs/solvers/nlpec.pdf

[35] CONOPT Solver Manual. [Online]. http://www.gams.com/dd/docs/solvers/conopt.pdf

[36] DICOPT Solver Manual. [Online]. http://www.gams.com/dd/docs/solvers/dicopt.pdf

CHAPTER 5 A TRI-LEVEL MODEL WITH AN EPEC SUB-PROBLEM FOR CENTRALIZED TRANSMISSION AND DECENTRALIZED GENERATION EXPANSION PLANNING FOR AN ELECTRICITY MARKET: PART II

Submitted to *IEEE Transactions on Power Systems*

Shan Jin and Sarah M. Ryan

Abstract

We study a tri-level integrated transmission and generation expansion planning problem in a deregulated power market environment. The collection of bi-level sub-problems in the lower two levels is an equilibrium problem with equilibrium constraints (EPEC) that can be approached by either the diagonalization method (DM) or a complementarity problem (CP) reformulation. This paper is a continuation of its Part I in Chapter 4, in which a hybrid iterative algorithm is proposed to solve the tri-level problem by iteratively applying the CP reformulation of the tri-level problem to propose solutions and evaluating them in the EPEC sub-problem by DM. It focuses on the numerical results obtained by the hybrid algorithm for a 6 bus system, a modified IEEE 30 bus system and a IEEE 118 bus system. In the numerical instances, the (approximate) Nash equilibrium point for the sub-problem can be verified by examining local concavity.

5.1 Introduction

This paper is a continuation of its Part I in Chapter 4. There, we formulate a generation and transmission expansion planning problem as a mixed integer tri-level programming problem, with the centralized transmission planning decision in the first level, multi-GENCOs' generation expansion decisions in the second level, and an electricity market equilibrium problem in the third level.

A bi-level centralized generation, transmission and fuel transportation expansion problem was formulated in [1], in which a centralized expansion decision is made in the top level and an operational game among GENCOs, ISO and a fuel supplier was modeled in the lower level. In Part I of this paper, we extend the previous model to a tri-level expansion model with consideration of the strategic expansion decision made by each GENCO in the second level and a centralized transmission expansion decision in the first level. Different from [1], we do not account for fuel network expansion and dispatch decisions in this paper. Instead we add a fuel supply capacity constraint into each GENCO's operational problem to represent a potential fuel limit. In the third level, each GENCO makes its strategic operational decision while the ISO clears the market.

In a restructured market, both the expansion planning and the operational decisions are no longer decided from a centralized perspective. Instead, each individual GENCO makes its own expansion decision in anticipation of market clearing by an integrated system operator (ISO), and the GENCOs submit their supply function bids in the day-ahead market. Therefore the generation expansion problem can be modeled as bi-level games among the GENCOs. In our model, instead of modeling the GENCO's bidding pair of quantity and price, we assume that each GENCO makes its decision on generation quantity under a type of bounded rationality [2].

Part I in Chapter 4 includes a full literature review of the modeling aspects of our formulation. Here, we focus on previous related numerical results. Nanduri et al. [3] modeled a two-tier matrix game for a multi-period, multi-GENCO capacity expansion model with an investment game in the upper level and a supply function game in the lower level with consideration of a transmission network. They proposed an algorithm to solve the matrix game to its optimal Nash equilibrium (NE) point, and applied it to a 5 bus system. Wang et al. [4] investigated bi-level games for a multi-GENCO capacity expansion planning problem in which GENCOs make their capacity and bidding decisions in the upper level and ISO clears the market in the lower level, proposed a co-evolutionary algorithm with pattern search, and applied it to an 8 bus network system, to identify the NE solution of the competition. Li et al. [5] and Soleymani et al. [6] modeled bi-level games with GENCOs' bidding decisions in the upper level and ISO clearing the market in the lower level, for which two iterative methods were illustrated in [5] for an 8 bus system, and a search based algorithm was applied to a 6 bus system in [6]. Ruiz et al. [7] studied a multi-GENCO bi-level bidding problem subject to a market clearing problem in the lower level, and solved it as an EPEC problem, which can be reformulated as a mixed integer linear programming problem. A case study of an IEEE Reliability Test System (RTS) [8] was presented.

When the GENCOs, modeled as Cournot competitors, make their expansion decisions in anticipation of the market clearing results, their decisions are also affected by the transmission capacity. Sauma and Oren et al. [9] modeled a multi-GENCO capacity expansion problem for a restructured electricity market, given various transmission expansion plans, as bi-level games and evaluated the social welfare of the system. An iterative algorithm to solve the bi-level games was illustrated on a 30 bus system. Roh et al. [10] simulated the interactions among GENCOs, TRANSCOs and ISO, and applied an iterative algorithm to a 6 bus system to solve a generation and transmission planning problem. The algorithm first solved resource planning problems of each GENCO and TRANSCO to maximize its profit with forecasted locational marginal price (LMP) and flowgate marginal price (FMP). Within each iteration, an ISO reliability check problem evaluated the system reliability in terms of loss of energy probability and provided capacity signals to the resource planning problem; while an ISO total social cost minimizing problem updated the LMP and FMP and provided price signals to the resource planning problems. Motamedi et al. [11] proposed a framework to consider decentralized GENCOs' reactions to the transmission expansion decision and anticipations of clearing prices from a restructured electricity market, formulated it as a four level model approached by agent-based system and search-based techniques, and applied it to a 5 bus system. Hesamzadeh et al. [12] solved a tri-level transmission augmentation planning problem with strategic generation expansion and operational decisions by a hybrid bi-level /island parallel genetic algorithm, tested on an IEEE 14-bus system. The first level minimizes the social cost including the transmission augmentation cost and the system operational cost. The bottom two levels are bi-level games, in which GENCOs, on the top level, maximize their profits by determining a price and quantity bid pair and expansion level, with anticipation of a social cost minimization problem based on a security constrained economic dispatch model in the bottom level. Pozo et al. [13]

studied a tri-level generation and transmission model, in which the investment and operational cost is minimized in the first level, the GENCOs maximize profit from expansion in the second level, and a market equilibrium problem with perfect competition among the GENCOs forms the third level. The GENCOs' MPEC problems were combined into a mixed integer linear programming problem by linearization of the nonlinear components in the objective functions and mixed integer reformulation of the equilibrium constraints. The model was tested on a 34-bus realistic power system in Chile. The problem we study is most similar to [9] but we treat the transmission plan as a decision variable in the optimization problem rather than a parameter. Our tri-level model also has a similar structure to the model investigated in [13]. However, we consider price-responsive linear demand functions and strategic operational decisions by the GENCOs. The problem structure is also similar to [12] but we apply mathematical programming to approach the solution instead of a heuristic. We test the solution accuracy and scalability on a 6 bus system, a modified IEEE 30-bus system, and the IEEE 118 bus system.

The model in our paper has a complicated tri-level structure with an equilibrium bi-level sub-problem and is difficult to solve. Algorithms are first proposed to solve the equilibrium bi-level sub-problem. Because bi-level games can be reformulated into an equilibrium problem with equilibrium constraints (EPEC), two currently available methodologies, diagonalization method (DM) and complementarity problem reformulation (CP), discussed in [15], are applied in Part I to help reformulate and solve the EPEC sub-problem and the tri-level problem. Further, we propose a hybrid iterative algorithm in Part I of the paper in Chapter 4 to solve the entire tri-level-programming problem by taking advantage of both these methods.

In this paper, three case studies are presented to illustrate how the algorithm proposed in Part I in Chapter 4 works to find the best transmission expansion plan, which can generate the largest net surplus in the system, in anticipation of generation expansion, production and market clearing decisions.

In Section 5.2, the numerical results are presented. Section 5.3 concludes the paper.

5.2 Numerical Results

For illustration, the hybrid algorithm has been applied to a small 6 bus system, a modified IEEE 30 bus test system, and the IEEE 118 bus sytem. In the 6 bus system, all the transmission planning options can be enumerated so that we are able to validate the global optimality of the solution found by the hybrid algorithm. The 30 bus system tests the scalability of the method and allows comparison with previous results in [9]. In the 118 bus system, global optimality among a restricted, realistic, set of transmission expansion options is verified. All the computational results are computed in GAMS23.4, and run on a 3.4GHz Intel Pentium 4 processor with 4 GB of RAM and 64 bit windows 7 system.

5.2.1 6 Bus System

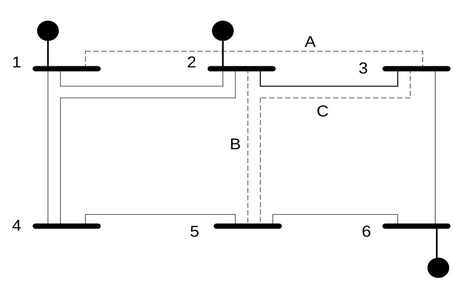

Figure 5- 1 A 6 Bus Test System with Three Candidate Lines

Table 5- 1 Model Parameters for Bus Nodes

j	V_j (MW)	c_j^{gexp} ($/MW)	c_j ($/MWh)	e_j ($/MWh/MWh)	U_j (MW)	a_j ($/MWh)	b_j ($/MWh/MWh)
1	80	10	20	0.0625	1800	100	-1
2	50	10	20	0.0625	1800	100	-1
3	n/a	n/a	n/a	n/a	0	120	-1
4	n/a	n/a	n/a	n/a	0	120	-1
5	n/a	n/a	n/a	n/a	0	120	-1
6	20	6	40	0.2500	1200	100	-1

Table 5- 2 Model Parameters for Transmission Lines

(i,j)	K_{ij} (MW)	$B_{ij}(\Omega^{-1})$	c_{ij}^{trexp}($/MW)	Line Status
(1,2)	200	5.9	n/a	Existing
(1,3) A	100	30	4	Candidate
(1,4)	50	3.9	n/a	Existing
(2,3)	100	27	n/a	Existing
(2,4)	100	5.1	n/a	Existing
(2,5) B	100	30	4	Candidate
(3,5) C	100	30	4	Candidate
(3,6)	100	55.5	n/a	Existing
(4,5)	50	27.0	n/a	Existing
(5,6)	100	5.1	n/a	Existing

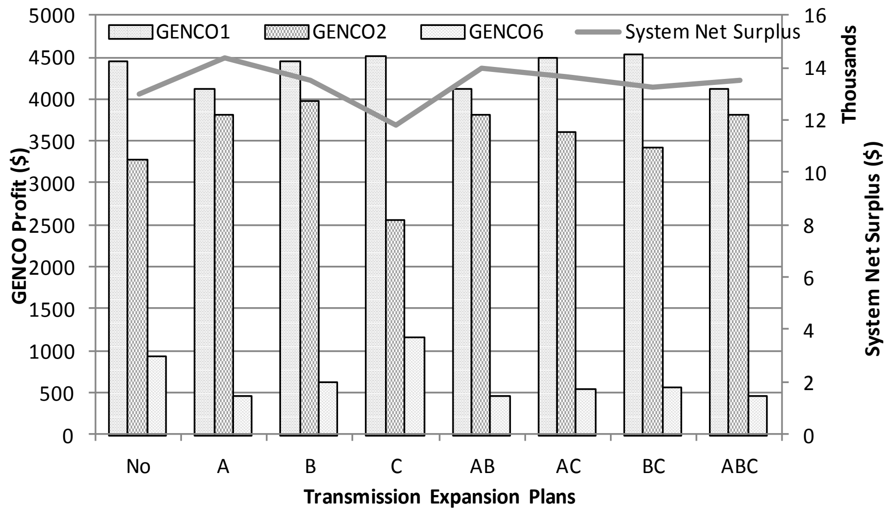

Figure 5- 2 Net Surplus and GENCOs' Net Profits with Different Transmission Expansion Plans for A 6 Bus System

For a demonstration case, we present a 6 bus network with three GENCOs on Buses 1, 2 and 6, and three candidate transmission lines shown in Figure 5-1, where solid lines represent the existing transmission lines and dotted lines represent the candidate lines.

All the model parameters are presented in Table 5-1 and Table 5-2. The three candidate lines are called lines A, B and C. The initial values V_j^{newequ} for the DM algorithm are set to equal their current values V_j.

Each of the eight feasible transmission expansion solutions can be evaluated in the equilibrium bi-level sub-problem by DM. Figure 5-2 compares the system net surplus and the profit for each GENCO among all the solutions, and it indicates that building transmission line A only is the global optimal solution according to system net surplus.

The iterative results obtained by the hybrid algorithm are presented in Table 5-3. In the first iteration, the master problem identifies building line A only as a promising initial transmission planning decision. Assuming this line is built, DM identifies an (approximate) NE for the multi-GENCO expansion decisions. The lower bound for the upper-level objective value and best found solution are updated, and a cut constraint is added to the master problem to eliminate the current solution of building line A only. The master problem becomes infeasible at its second iteration, which implies that all the other transmission plans other than building line A only cannot produce a net surplus in the master problem higher than 14348.46. Therefore, the algorithm terminates with the best found solution of building line A only, which is the global optimal solution of the original tri-level programming problem, as shown in Figure 5-2. The hybrid algorithm takes only two major iterations and one DM evaluation to find the optimal solution within 104.41 seconds. In comparison, enumerating all solutions and evaluating each by DM requires a total of 542.82 seconds.

Table 5-4 summarizes the detailed results obtained with different transmission expansion plans, where we can draw the same conclusion from the total net surplus, that the global optimal solution is to build line A only. From Table 5-4, when there is no transmission expansion, the system experiences congestion in line (2, 3). When one transmission line, A, is built, the congestion is relieved and the electricity price decreases. The GENCOs have less market power to drive a high market price by expanding and generating less. Instead, the GENCOs maximize their profit by making the expansions to sell more power. Therefore, the buyers receive more electricity with lower prices, which results in higher buyer surplus. Compared with plan "None", the buyer surplus and seller surplus both increase. Because the increase in the system surplus is sufficient to cover the cost of building transmission line A and extra generation expansion cost, plan "A" is much more favored than plan "None".

Unlike in plan "A", in plans "B" and "C", the network congestion has not been eliminated. Although building only transmission line B leads to a slight decrease in electricity price, and, thus, an increase in both buyer surplus and seller surplus, the overall system net surplus is not as high as in plan "A". In plan "C", the system congestion becomes even worse, which leads to higher electricity prices, and the buyer surplus and seller surplus both decrease. Plans "AB" and "ABC" result in the same generation expansion level, and quantity consumed as plan "A", but at a higher transmission expansion cost. Therefore it is obvious that plan "A" is preferred. Plans "AC" and "BC" generally help to relieve the congestion and increase the system efficiency with a higher system net surplus. However they have higher transmission investment cost and lower increase in system surplus than plan "A".

We observe that the best transmission expansion plan can not only increase the total net surplus but also guide the market participants to achieve a win-win situation in which total buyer and seller surpluses can be increased by 22% and 7%, respectively. The total net surplus increase comes mostly from the increasing total buyer surplus, which is driven by the increasing generation capacity expansions and the lower electricity prices.

Table 5- 3 Iterative Results of the Hybrid Algorithm to Solve a 6 Bus Case Study

Major Iter.	**MINLP Master Problem A with CP Reformulation**			**EPEC Sub-problem B with DM**	**Adding Constraints**	
	Status	z^{master}	**Net Surplus** $F(z^{master}, \Omega^{master})$	**Net Surplus** $F(z^{master}, \Omega^{sub})$	**Lower Bound** F^{best}	**Cut Point** z^{master}
1	Feasible	A	15244.07	14348.46	14348.46	A
2	Infeasible					

Table 5- 4 Detail Results with Different Transmission Expansion Plans

		No	A	B	C	AB	AC	BC	ABC
Total Surplus		13502	15908	14690	12561	15908	15509	14921	15908
Total Buyer Surplus		3929	6347	4808	3394	6347	5871	5130	6347
Total Seller Surplus		9202	9561	9861	8596	9561	9638	9791	9561
Total Transmission Rent		371	0	21	571	0	0	0	0
Total Generation Investment Cost		535	1159	800	362	1159	1043	870	1159
Total Transmission Investment Cost		0	400	400	400	800	800	800	1200
Total Net Surplus		12967	14348	13490	11799	13948	13666	13251	13548
Gen. Expan. Level, V_j^{new} = Gen. Level, y_j	1	99.98	120.36	105.93	98.20	120.36	126.39	117.48	120.36
	2	74.63	120.36	100.09	54.38	120.36	101.51	93.93	120.36
	6	34.86	28.66	26.58	42.68	28.66	30.66	29.34	28.66
Quantity Consumed, q_j	1	27.23	34.90	28.88	26.04	34.90	33.09	30.13	34.90
	2	28.17	34.90	28.90	28.62	34.90	33.09	30.13	34.90
	3	41.39	54.90	48.54	38.98	54.90	53.09	50.13	54.90
	4	45.81	54.90	48.87	42.15	54.90	53.09	50.13	54.90
	5	45.16	54.90	48.86	40.37	54.90	53.09	50.13	54.90
	6	21.71	34.90	28.56	19.10	34.90	33.09	30.13	34.90

Table 5- 4 (continued)

Electricity Price, p_j	1	72.77	65.10	71.12	73.96	65.10	66.91	69.88	65.10
	2	71.83	65.10	71.10	71.38	65.10	66.91	69.88	65.10
	3	78.61	65.10	71.46	81.02	65.10	66.91	69.88	65.10
	4	74.19	65.10	71.13	77.85	65.10	66.91	69.88	65.10
	5	74.84	65.10	71.14	79.63	65.10	66.91	69.88	65.10
	6	78.29	65.10	71.44	80.90	65.10	66.91	69.88	65.10
Flow, f_{ij}	(1,2)	31.10	0.66	41.88	37.39	9.10	3.15	48.27	6.89
	(1,3)	0.00	51.87	0.00	0.00	58.41	71.54	0.00	63.41
	(1,4)	41.64	32.93	35.17	34.77	17.95	18.61	39.09	15.17
	(2,3)	**50.00**	43.64	**50.00**	**50.00**	10.91	49.96	49.28	25.55
	(2,4)	27.57	42.48	9.79	13.15	15.60	21.61	9.39	13.88
	(2,5)	0.00	0.00	53.28	0.00	68.06	0.00	53.40	52.92
	(3,5)	0.00	0.00	0.00	28.22	0.00	57.26	-1.35	24.53
	(3,6)	8.61	40.62	1.46	-17.20	14.42	11.14	0.51	9.53
	(4,5)	23.40	20.51	-3.90	5.77	-21.35	12.88	-1.65	-25.85
	(5,6)	-21.75	-34.38	0.52	-6.38	-8.18	-8.71	0.28	-3.29

5.2.2 Modified IEEE 30 Bus Test System

The modified IEEE 30 Bus Test System includes six generators on nodes 1, 2, 13, 22, 23 and 27, thirty-nine transmission lines, and ten candidate transmission lines. Based on the 30 bus case study in [9], the model parameters are set up as shown in Appendix 5.A.

Different from [9], we assume a quadratic generation cost function that is not affected by the increasing generation capacity, and we do not consider expanding the capacity of the existing transmission lines. All the GENCOs have the same generation cost function with $c_j = 10$ and $e_j = 0.0625$. The ten candidate lines are labeled as A through J, among which the lines B, E, G, and H are the proposed new lines in [9]. The total number of all transmission expansion options totals $2^{10} = 1024$, which makes evaluation of each by DM computationally prohibitive. The network is in Figure 5-3, where solid lines represent the existing transmission lines and dotted lines represent the candidate lines.

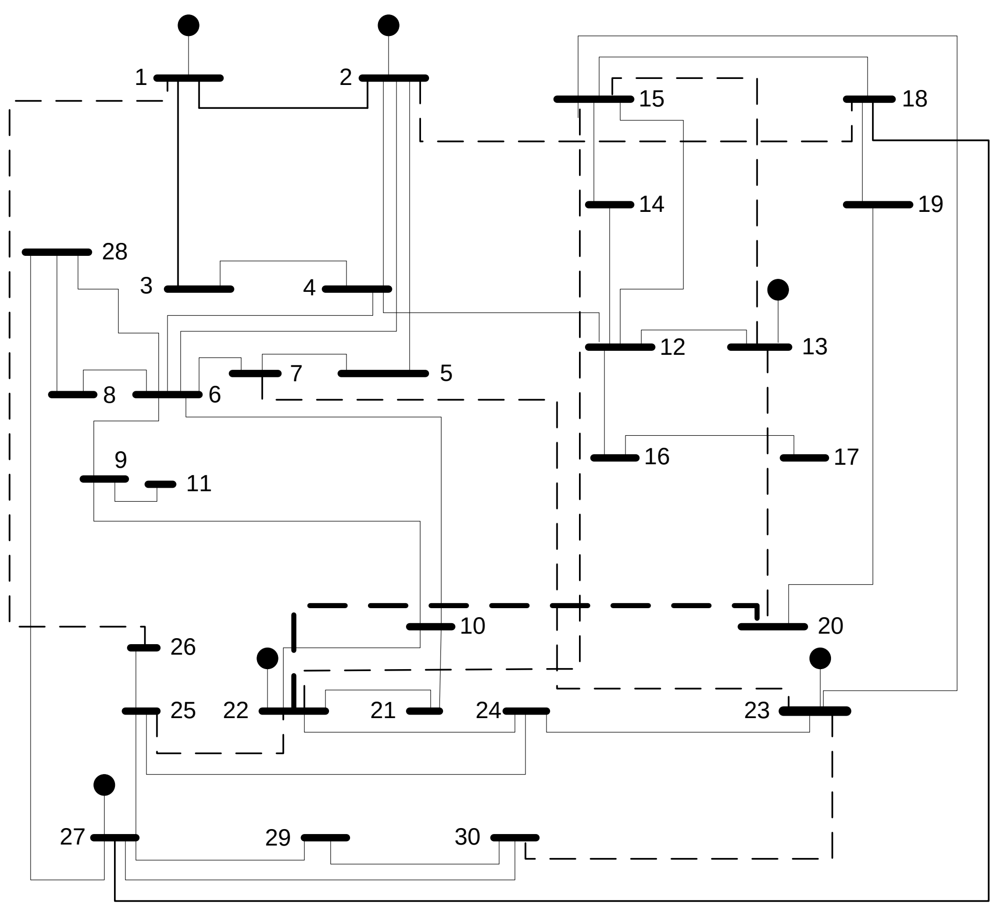

Figure 5- 3 A Modified IEEE 30 Bus Test System with Ten Candidate Lines

The larger problem size causes computational difficulty to solve the MINLP master problem at the beginning of each major iteration. Because the purpose of the master problem is to identify a promising transmission planning decision, it can be further relaxed by ignoring equations obtained from the partial derivatives of the Lagrangian with respect to the dual variables; i.e., equations (4-43) - (4-54) in Part I of this paper in Chapter 4.

The iterative results obtained by the hybrid algorithm are given in Table 5-5. In the fourth major iteration the algorithm finds the optimal solution, which is to build only candidate line H. This result also appears to be consistent with the case study results found in [9]. Except for the adjustment of the parameters due to model differences, the 30 bus case study is the same as the one in [9]. Besides the instance with ten candidate lines, we also examine a 30 bus case study with the four new transmission lines B, E, G, and H, suggested in [9], and the results also indicate building line H only. All the 16 feasible transmission expansion solutions can be evaluated in the EPEC sub-problem by DM. Figure 5-4 compares the system net surplus and the profit for each GENCO given all transmission expansion options, and it indicates that building transmission line D only is the global optimal transmission expansion decision. Although we do not have the DM solutions for all 1024 transmission expansion options to validate the best solution found by the hybrid algorithm, based on the results of the 30 bus instance with four candidate lines in Figure 5-4, it is very likely that building line H only is the global optimal solution for the tri-level expansion planning problem. The total computational time for the hybrid algorithm is 5591.97 seconds.

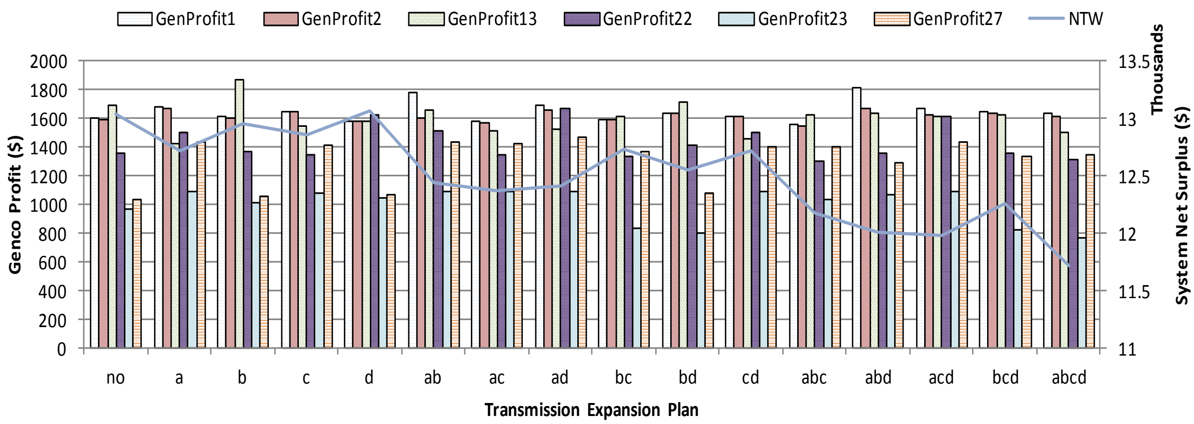

Figure 5- 4 Net Surplus and GENCOs' Net Profits with Different Transmission Expansion Plans for a modified IEEE 30 Bus Test System

Table 5- 5 Iterative Results of the Hybrid Algorithm to Solve a Modified IEEE 30 Bus Test System

Major Iter.	**MINLP Master Problem A with CP Reformulation**			**EPEC Sub-problem B with DM**	**Adding Constraints**	
	Status	z^{master}	**Net Surplus** $F(z^{master}, \Omega^{master})$	**Net Surplus** $F(z^{master}, \Omega^{sub})$	**Lower Bound** F^{best}	**Cut Point** z^{master}
1	Feasible	No	13235.34	13038.62	13038.62	No
2	Feasible	B	13057.90	12727.90	13038.62	B
3	Feasible	E	13216.10	12957.11	13038.62	E
4	**Feasible**	**H**	**13246.07**	**13066.56**	**13066.56**	**H**
5	Infeasible					

Given expansion on candidate line H, the DM results for the optimal generation capacity vector, V^{new}, are indicated in Figure 5-5. Based on much computational experience, the optimal solutions usually stabilize within 10 to 15 rounds of iterations, so we set 25 as a maximum number of iteration cycles to terminate the DM algorithm when it is impossible to find an exact NE point. Because there are six different GENCOs making their capacity decisions in each round of iteration, it results in a maximum of 150 MPEC solution iterations.

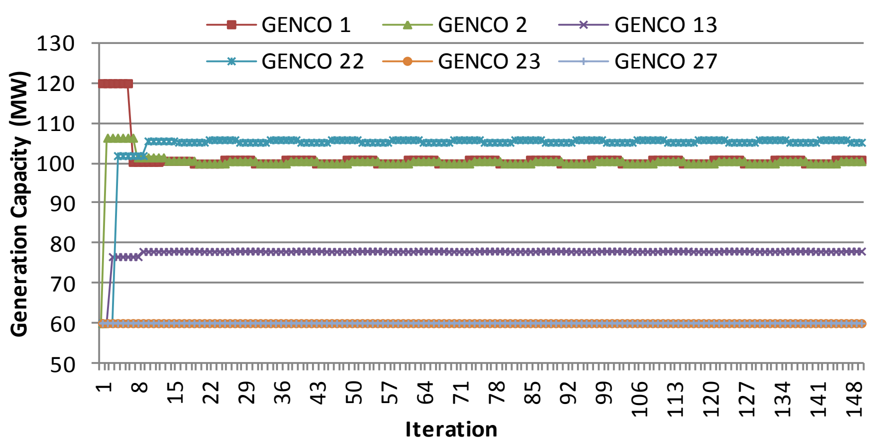

Figure 5- 5 Iteration of Expansion Capacity V^{new} with Transmission Expansion on Line H

From Figure 5-5, since the optimal solution does not converge, we infer existence of a mixed, rather than pure, Nash strategy. The GENCOs' decisions oscillate within a small range of approximately 1%: GENCO 1 slightly adjusts its decision between the values 99.94 and 100.88; GENCO 2 between the values 99.94 and 100.52; GENCO 13 between 77.94 and 78.03; GENCO 22 between 105.24 and 105.87; while both GENCOs 23 and 27 are converged to a capacity of 60 MW each. In this case, we can simply define an approximate equilibrium point by averaging the two capacity values for a generator, so that the optimal generation expansion capacities V_k^{new} , j = 1, 2, 13, 22, 23 and 27, are approximately [100.41, 100.23, 77.99, 105.56, 60, 60] and the generation levels y_j , j = 1, 2, 13, 22, 23 and 27, are the same as their new generation capacities. The net profits for each GENCO are [1586.50, 1583.39, 1581.42, 1632.46, 1050.94, 1076.19], and the total system net surplus is 13066.56.

5.2.3 Nash Equilibrium (NE) Solution Validation

The DM algorithm is applied to iteratively solve each single bi-level programming problem, reformulated as an equivalent mathematic program with equilibrium constraints (MPEC) including equations (4-2), (4-3), (4-18) – (4-35) for each specific GENCO *k*, within the EPEC sub-problem B. The (approximate) convergence point is an NE point, where no GENCO can improve its profit by changing only its own capacity expansion decisions while all the other GENCOs' decision remain fixed. To ensure the (approximate) convergence point, each GENCO's MPEC in the DM iteration should be solved to its local optimality. However the objective function (4-2) of each MPEC in Part I in Chapter 4 is nonlinear and not ensured to be concave, which implies no guarantee for the global optimality. To validate that the solution found by DM approximates an NE, we must further investigate the objective values of neighboring points.

For each single MPEC, we can reformulate the lower level problem by introducing the binary variables μz and and a big value M, and converting all the equilibrium constraints of the form $0 \le x \perp y \ge 0$ into $0 \le x, 0 \le y, x \le M\mu z, y \le M(1-\mu z)$as in our previous paper [1]. Upon this reformulation, the MPEC becomes a single level programming problem with mixed integer linear constraints and nonlinear objective function given by equation (4-

2) in Part I of Chapter 4. For GENCO 1, given the optimal solutions of the other GENCOs as the model parameters, we evaluate the net benefits for the neighboring points of the optimal solution, 100.41. Variable V_1^{new} can be fixed from its existing capacity 60 to 120 to investigate the change of the objective values in response to it. Once the variable V_1^{new} is fixed, its optimal generation variable $y_1 = V_1^{new}$ can be determined, since there is no incentive to expand beyond the actual generation level that is needed[1]. In this case, the single level nonlinear mixed integer programming (MIP) problem has been transformed to a single level linear MIP problem, which can be solved to its global optimality by CPLEX. The only variable involved in the objective function is p. In the case of making no expansion, like GENCO 23 and 27, because V_k^{new} has to be higher than the existing capacity, 60, we evaluate the net profits for the neighboring area by fixing V_k^{new} from 60 to a predetermined higher value. Figure 5-6 presents the relationship among the objective value, the net profit of GENCO 1, and its capacity decision V_1^{new}. The net profit at the top is an enlargement of the bottom one. It indicates concavity of the GENCO's objective, given in equation (4-2) of Part I in Chapter 1, as a function of V_1^{new} with the global optimal solution between 100 and 101, which is consistent with the approximate optimal point 100.41 found by the DM algorithm. The same test can be applied to each GENCO to validate the global optimality of each GENCO's MPEC problem, which further verifies that the optimal solution found by the DM algorithm is indeed the local NE point of the EPEC sub-problem B.

---

[1] Here, the model is a one time period model. However, in a stochastic EPEC, the generation level y does not necessarily equal the capacity level V^{new} in all scenarios. To validate an (approximate) NE point in the stochastic EPEC case, we can also first fix a GENCO k's capacity level V_k^{new} in a range of values. Given each fixed value V_k^{new}, we can solve a set of pairs of optimal generation level and dual price $\{y_s(V_k^{new}), p_s(V_k^{new})\}$ under each scenario s, so that we can calculate profit for GENCO k expanding at V_k^{new} in equation (4-2) in Part I of the paper. We can compare those profits for GENCO k at all the different capacity level V_k^{new} and find the V_k^{new} with the highest profit.

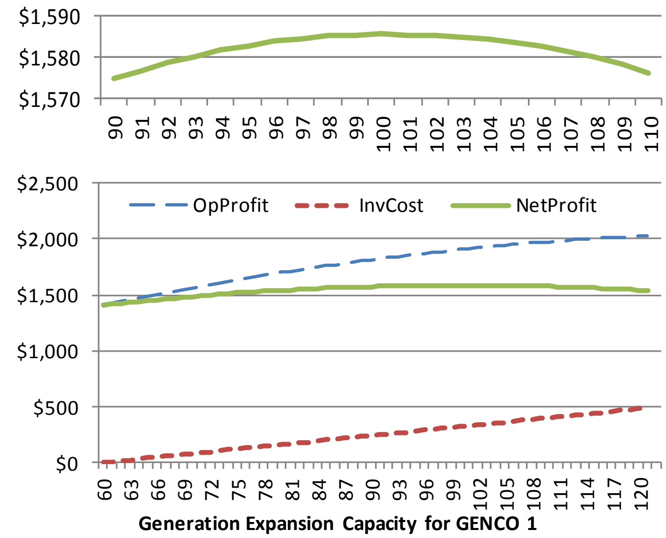

Figure 5- 6 Investment Cost, Operational Profit and Net Profit by Expansion Capacity V_1^{new}

5.2.4 IEEE 118 Bus Test System

The algorithm was also tested on a standard IEEE 118 bus system with 54 generators, 179 existing lines and 4 candidate lines. The candidate lines were selected as likely to help relieve the congestion in the existing system. All the nodes have the same linear demand functions with $a_j = 100$ and $b_j = -1$. The capacities of existing lines are assumed to be K_{ij} =50, and K_{ij} = 100 for the candidate lines. Detailed parameter assumptions for GENCOs and transmission lines are shown in Tables 5-8 and Table 5-9 in Appendix 5.B.

The algorithm identified the best solution at the first major iteration and found two more feasible, though inferior, solutions in the second and third rounds. The best solution returned is to build the transmission lines A, C and D. We observe that even after building three candidate lines, system congestion still exists.

We also obtained the global optimal solution of the 118 bus case study by enumerating all the 16 possible transmission expansion options, and verified that the best solution found by the algorithm turned out to be globally optimal in this instance.

5.3 Conclusions

In this paper, we consider an integrated market-based transmission and generation expansion planning problem in a deregulated electricity market environment. The novel tri-level programming model proposed in Part I of the paper includes an equilibrium bi-level sub-problem, also known as an EPEC, which can be solved by either a diagonalization method (DM) or a complementarity problem (CP) reformulation. To approach the tri-level optimization problem, a hybrid iterative algorithm is proposed in Part I of the paper in Chapter 4 by taking advantage of both methods.

The proposed algorithm has been tested both on three systems. In the smallest instance, where all the feasible transmission expansion solutions can be enumerated, the solution found by the hybrid algorithm has been shown to be globally optimal. The solutions of the 30 and 118 bus systems were also successfully found by the hybrid algorithm. To deal with the cases where a pure Nash equilibrium strategy does not exist, an approximate NE point has been defined. Finally, a method has been proposed to validate the (approximate) NE point found by DM algorithm.

Appendix 5.A Model Parameters for Modified IEEE 30 Bus Test System

Table 5- 6 Model Parameters for Bus Nodes of a Modified IEEE 30 Bus Test System

j	V_j(MW)	c_j^{gexp}($/MW)	b_j($/MW/MW)	a_j($/MW)
1	60	8	-1	50
2	60	8	-1	50
3	0	0	-1	60
4	0	0	-1	55
5	0	0	-1	50
6	0	0	-1	50
7	0	0	-1	60
8	0	0	-1	55
9	0	0	-1	50
10	0	0	-1	55
11	0	0	-1	50
12	0	0	-1	55
13	60	8	-1	50
14	0	0	-1	55
15	0	0	-1	55
16	0	0	-1	50
17	0	0	-1	55
18	0	0	-1	50
19	0	0	-1	55
20	0	0	-1	50

Table 5- 6 (continued)

21	0	0	-1	50
22	60	8	-1	50
23	60	8	-1	60
24	0	0	-1	55
25	0	0	-1	50
26	0	0	-1	50
27	60	8	-1	50
28	0	0	-1	50
29	0	0	-1	50
30	0	0	-1	55

Table 5- 7 Model Parameters for Transmission Lines of a Modified IEEE 30 Bus Test System

(i,j)	K_{ij} (MW)	$B_{ij}(\Omega^{-1})$	$c_{ij}^{trexp}(\$/\text{MW})$	Line Status
(1,2)	130	15	n/a	Existing
(1,3)	130	4.92	n/a	Existing
(2,4)	65	5.23	n/a	Existing
(3,4)	130	23.53	n/a	Existing
(2,5)	130	4.71	n/a	Existing
(2,6)	65	5	n/a	Existing
(4,6)	90	23.53	n/a	Existing

Table 5- 7 (continued)

(5,7)	70	7.1	n/a	Existing
(6,7)	130	10.96	n/a	Existing
(6,8)	32	23.53	n/a	Existing
(6,9)	65	4.76	n/a	Existing
(6,10)	32	1.79	n/a	Existing
(9,11)	65	4.76	n/a	Existing
(9,10)	65	9.09	n/a	Existing
(4,12)	65	3.85	n/a	Existing
(12,13)	65	7.14	n/a	Existing
(12,14)	32	3.17	n/a	Existing
(12,15)	32	5.96	n/a	Existing
(12,16)	32	4.16	n/a	Existing
(14,15)	16	2.26	n/a	Existing
(16,17)	16	4.47	n/a	Existing
(15,18)	16	3.64	n/a	Existing
(18,19)	16	6.34	n/a	Existing
(19,20)	32	12.07	n/a	Existing
(10,21)	32	12.07	n/a	Existing
(10,22)	32	5.47	n/a	Existing
(21,22)	32	40	n/a	Existing
(15,23)	16	4	n/a	Existing

Table 5- 7 (continued)

(22,24)	16	3.85	n/a	Existing
(23,24)	16	3.01	n/a	Existing
(24,25)	16	2.28	n/a	Existing
(25,26)	16	1.84	n/a	Existing
(25,27)	16	3.74	n/a	Existing
(27,28)	65	2.5	n/a	Existing
(27,29)	16	1.87	n/a	Existing
(27,30)	16	1.3	n/a	Existing
(29,30)	16	1.73	n/a	Existing
(8,28)	32	4.59	n/a	Existing
(6,28)	32	15	n/a	Existing
(1,26) A	100	23.53	4	Candidate
(2,18) B	100	23.53	4	Candidate
(7,23) C	100	23.53	4	Candidate
(13,15) D	100	23.53	4	Candidate
(13,20) E	100	23.53	4	Candidate
(15,22) F	100	23.53	4	Candidate
(18,27) G	100	23.53	4	Candidate
(20,22) H	100	23.53	4	Candidate
(22,25) I	100	23.53	4	Candidate
(23,30) J	100	23.53	4	Candidate

Appendix 5.B Model Parameters for IEEE 118 Bus Test System

Table 5- 8 Model Parameters for GENCOs of a IEEE 118 Bus Test System

j	V_j(MW)	c_j(\$/MW/MW)	e_j(\$/MW/MW)	c_j^{gexp}(\$/MW)
1	100	20	0.0625	10
4	100	20	0.0625	10
6	100	20	0.0625	10
8	100	20	0.0625	10
10	550	20	0.0625	10
12	185	20	0.0625	10
15	100	20	0.0625	10
18	100	20	0.0625	10
19	100	20	0.0625	10
24	100	20	0.0625	10
25	320	20	0.0625	10
26	414	20	0.0625	10
27	100	20	0.0625	10
31	107	20	0.0625	10
32	100	20	0.0625	10
34	100	20	0.0625	10

Table 5- 8 (continued)

36	100	20	0.0625	10
40	100	20	0.0625	10
42	100	20	0.0625	10
46	119	20	0.0625	10
49	304	20	0.0625	10
54	148	20	0.0625	10
55	100	20	0.0625	10
56	100	20	0.0625	10
59	255	20	0.0625	10
61	260	20	0.0625	10
62	100	20	0.0625	10
65	491	40	0.125	6
66	492	40	0.125	6
69	805	40	0.125	6
70	100	40	0.125	6
72	100	40	0.125	6
73	100	40	0.125	6
74	100	40	0.125	6
76	100	40	0.125	6
77	100	40	0.125	6
80	577	40	0.125	6

Table 5- 8 (continued)

85	100	40	0.125	6
87	104	40	0.125	6
89	707	40	0.125	6
90	100	40	0.125	6
91	100	40	0.125	6
92	100	40	0.125	6
99	100	40	0.125	6
100	352	40	0.125	6
103	140	40	0.125	6
104	100	40	0.125	6
105	100	40	0.125	6
107	100	40	0.125	6
110	100	40	0.125	6
111	136	40	0.125	6
112	100	40	0.125	6
113	100	40	0.125	6
116	100	40	0.125	6

Table 5- 9 Model Parameters for Transmission Lines of a IEEE 118 Bus Test System

(i,j)	$B_{ij}(\Omega^{-1})$	Line Status	(i,j)	$B_{ij}(\Omega^{-1})$	Line Status
(1.2)	9	Existing	(56.57)	9	Existing

Table 5- 9 (continued)

(1.3)	22	Existing	(56.58)	9	Existing
(2.12)	15	Existing	(56.59)	4	Existing
(3.5)	9	Existing	(59.60)	7	Existing
(3.12)	6	Existing	(59.61)	6	Existing
(4.5)	120	Existing	(59.63)	26	Existing
(4.11)	13	Existing	(60.61)	71	Existing
(5.6)	18	Existing	(60.62)	17	Existing
(5.8)	37	Existing	(61.62)	25	Existing
(5.11)	13	Existing	(61.64)	37	Existing
(6.7)	46	Existing	(62.66)	4	Existing
(7.12)	28	Existing	(62.67)	8	Existing
(8.9)	33	Existing	(63.64)	50	Existing
(8.30)	20	Existing	(64.65)	33	Existing
(9.10)	31	Existing	(65.66)	27	Existing
(11.12)	47	Existing	(65.68)	62	Existing
(11.13)	13	Existing	(66.67)	9	Existing
(12.14)	13	Existing	(68.69)	27	Existing
(12.16)	11	Existing	(68.81)	49	Existing
(12.117)	7	Existing	(68.116)	245	Existing
(13.15)	4	Existing	(69.70)	7	Existing
(14.15)	5	Existing	(69.75)	7	Existing

Table 5- 9 (continued)

(15.17)	21	Existing	(69.77)	9	Existing
(15.19)	23	Existing	(70.71)	27	Existing
(15.33)	7	Existing	(70.74)	7	Existing
(16.17)	5	Existing	(70.75)	6	Existing
(17.18)	19	Existing	(71.72)	5	Existing
(17.30)	26	Existing	(71.73)	21	Existing
(17.31)	6	Existing	(74.75)	23	Existing
(17.113)	30	Existing	(75.77)	5	Existing
(18.19)	19	Existing	(75.118)	19	Existing
(19.20)	8	Existing	(76.77)	6	Existing
(19.34)	4	Existing	(76.118)	17	Existing
(20.21)	11	Existing	(77.78)	74	Existing
(21.22)	10	Existing	(77.80)	18	Existing
(22.23)	6	Existing	(77.82)	10	Existing
(23.24)	19	Existing	(78.79)	39	Existing
(23.25)	12	Existing	(79.80)	14	Existing
(23.32)	8	Existing	(80.81)	27	Existing
(24.70)	2	Existing	(80.96)	5	Existing
(24.72)	5	Existing	(80.97)	10	Existing
(25.26)	26	Existing	(80.98)	9	Existing
(25.27)	6	Existing	(80.99)	5	Existing

Table 5- 9 (continued)

(26.30)	12	Existing	(82.83)	25	Existing
(27.28)	11	Existing	(82.96)	17	Existing
(27.32)	12	Existing	(83.84)	6	Existing
(27.115)	13	Existing	(83.85)	6	Existing
(28.29)	10	Existing	(84.85)	13	Existing
(29.31)	27	Existing	(85.86)	8	Existing
(30.38)	18	Existing	(85.88)	9	Existing
(31.32)	9	Existing	(85.89)	6	Existing
(32.113)	5	Existing	(86.87)	5	Existing
(32.114)	16	Existing	(88.89)	14	Existing
(33.37)	6	Existing	(89.90)	5	Existing
(34.36)	34	Existing	(89.92)	19	Existing
(34.37)	99	Existing	(90.91)	11	Existing
(34.43)	6	Existing	(91.92)	7	Existing
(35.36)	94	Existing	(92.93)	11	Existing
(35.37)	19	Existing	(92.94)	6	Existing
(37.38)	27	Existing	(92.100)	3	Existing
(37.39)	9	Existing	(92.102)	17	Existing
(37.40)	5	Existing	(93.94)	13	Existing
(38.65)	10	Existing	(94.95)	21	Existing

Table 5- 9 (continued)

(39.40)	15	Existing	(94.96)	11	Existing
(40.41)	19	Existing	(94.100)	16	Existing
(40.42)	5	Existing	(95.96)	17	Existing
(41.42)	7	Existing	(96.97)	11	Existing
(42.49)	3	Existing	(98.100)	5	Existing
(43.44)	4	Existing	(99.100)	12	Existing
(44.45)	10	Existing	(100.101)	8	Existing
(45.46)	7	Existing	(100.103)	17	Existing
(45.49)	5	Existing	(100.104)	5	Existing
(46.47)	7	Existing	(100.106)	4	Existing
(46.48)	5	Existing	(101.102)	9	Existing
(47.49)	15	Existing	(103.104)	6	Existing
(47.69)	3	Existing	(103.105)	6	Existing
(48.49)	18	Existing	(103.110)	5	Existing
(49.50)	12	Existing	(104.105)	25	Existing
(49.51)	6	Existing	(105.106)	17	Existing
(49.54)	3	Existing	(105.107)	5	Existing
(49.66)	10	Existing	(105.108)	13	Existing
(49.69)	3	Existing	(106.107)	5	Existing
(50.57)	7	Existing	(108.109)	31	Existing
(51.52)	15	Existing	(109.110)	12	Existing

Table 5- 9 (continued)

(51.58)	12	Existing	(110.111)	12	Existing
(52.53)	6	Existing	(110.112)	14	Existing
(53.54)	8	Existing	(114.115)	92	Existing
(54.55)	13	Existing	(20.38)A	30	Candidate
(54.56)	97	Existing	(21.69)B	30	Candidate
(54.59)	4	Existing	(61.65)C	30	Candidate
(55.56)	60	Existing	(25.89)D	30	Candidate
(55.59)	4	Existing			

References

[1] S. Jin and S.M.Ryan, "Capacity expansion in the integrated supply network for an electricity market," IEEE Transactions on Power Systems, vol. 26, no. 4, p. 2275–2284, 2011.

[2] J. Yao, I. Adler and S.S. Oren, "Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network," Operations Research, vol. 56, no. 1, pp. 34-47, 2008.

[3] V. Nanduri, T.K. Das and P. Rocha, "Generation capacity epxansion in energy markets using a two-level game theoretic model," IEEE Transactions on Power Systems, vol. 24, no. 3, p. 1165–1172, 2009.

[4] J. Wang, M. Shahidehpour, Z. Li and A. Botterud, "Strategic generation capacity expansion planning with incomplete information," IEEE Transactions on Power Systems, vol. 24, no. 2, p. 1002–1010, 2009.

[5] T. Li and M. Shahidehpour, "Strategic bidding of transmission-constrained GENCOs with incomplete information," IEEE Transactions on Power Systems, vol. 20, no. 1, p. 437–447, 2005.

[6] S. Soleymani, A.M. Ranjbar and A.R. Shirani, "New approach to bidding strategies of generating companies in day ahead energy market," Energy Conversion and Management, vol. 49, no. 6, p. 1493–1499, 2008.

[7] C. Ruiz, A.J. Conejo and Y. Smeers, "Equilibria in an oligopolistic electricity pool with stepwise offer curves," IEEE Transmission on Power Systems, vol. 27, no. 2, pp. 752-761, 2012.

[8] Reliability System Task Force, "The IEEE reliability test system-1996. A report prepared by the reliability test system task force of the application," IEEE Transactions on Power Systems, vol. 14, no. 3, p. 1010–1020, Aug. 1999.

[9] E.E. Sauma and S.S. Oren, "Proactive planning and valuation of transmission investments in restructured electricity markets," Journal of Regulatory Economics, vol. 30, no. 3, p. 261–290, 2006.

[10] J.H. Roh, M. Shahidehpour and L. Wu, "Market-based generation and transmission planning with uncertainties," IEEE Transactions on Power Systems, vol. 24, no. 3, p. 1587–1598, 2009.

[11] A. Motamedi, H. Zareipour, M.O. Buygi and W.D. Rosehart, "A transmission planning framework considering future generation expansions in electricity markets," IEEE Transactions on Power Systems, vol. 25, no. 4, p. 1987–1995, 2010.

[12] M.R. Hesamzadeh, D.R. Biggar, N. Hosseinzadeh and P.J. Wolfs, "Transmission augmentation with mathematical modeling of market power and strategic generation expansion - part II," IEEE Transactions on Power Systems, vol. 26, no. 4, p. 2049–2057, 2011.

[13] D. Pozo, E.E. Sauma and J. Contreras, "A three-level static MILP model for generation and transmission expansion planning," IEEE Transactions on Power Systems, vol. forthcoming.

[14] M.R. Hesamzadeh, D.R. Biggar, N. Hosseinzadeh and P.J. Wolfs, "Transmission augmentation with mathematical modeling of market power and strategic generation expansion - part I," IEEE Transactions on Power Systems, vol. 26, no. 4, p. 2040–2048, 2011.

[15] X. Hu and D. Ralph, "Using EPECs to model bilevel games in restructured electricity markets with locational prices," Operations Research, vol. 55, no. 5, p. 809–827, 2007.

CHAPTER 6 GENERAL CONCLUSIONS

This dissertation is a combination of three papers to provide well rounded insights addressing our research questions based on solving expansion planning problems under different market mechanisms. In a restructured electricity market there are multiple decision makers at various levels making decisions that interact and they all have to deal with uncertainty. Each decision maker must consider two types of decision, an expansion decision for the long term and an operational decision for participating in an electricity market. This is a problem with multiple periods and multiple dimensions and is too complicated to approach as one single problem. Therefore, the problem is decomposed and insight is gained by exploring it from different perspectives.

From a centralized perspective, the first paper, presented in Chapter 2, considers a long term generation expansion problem with minimization of two metrics, both the expected cost (EC) and Conditional Value-at-Risk (CVaR), by a two-stage stochastic integer program. The two integrated future uncertainties are modeled as a stochastic process and a scenario tree is constructed to represent their evolution over multiple periods by a statistical property matching technique. A Midwest ISO (MISO) based generation expansion case study is tested. To address computational complexity, scenario sampling is applied to generate a scenario subset to represent the entire uncertainty space, and the Multiple Replication Procedure (MRP) is used to compute the confidence interval on the optimality gap and evaluate the stability of the obtained approximate solution. The numerical results indicate the optimality gaps for the minimization of EC are very small compared to those for minimizing CVaR. The solutions under different scenario sampling subsets are structurally similar. The results also imply that a highly accurate solution can be obtained based on a relatively small ample of a large scenario space.

From the perspective of an integrated electricity supply system where generation expansion decisions could largely depend on the availability of other facilities, the second paper, presented in Chapter 3, adopts the system point of view in centralized capacity expansion considering the interactions of multiple decentralized participants in a competitive electricity market. It considers an integrated electricity supply system including fuel transportation, generation and transmission. The model incorporates a discrete transmission expansion decision, a competitive market with price-responsive demand, and two potential carbon policies which are presented as model extensions in an Appendix to the chapter. The difficulty of solving a bi-level programming problem to its global optimality is discussed and three problem relaxations obtained by reformulation are proposed to explore the problem and solve it to its global optimality. A 6 bus case study is presented to illustrate the global expansion decisions' effect on the market. The results indicate that the total net social welfare, defined as the welfare less investment cost, reduces by 2% with generators' strategic operational decisions, compared to the non-strategic decisions. With generators being non-strategic, the total social surplus increases with better-off electricity buyers and worse-off generators, and fuel transportation and generation capacity expansion increase, which lead to lower electricity prices. To avoid computational complexity, the model is deterministic and static. For future research, a multi-period investment and operational decisions can be considered in the model and major uncertain factors can also be taken into account by scenarios. The extended problem forms a stochastic bi-level program, which is also known as a stochastic mathematical program with complementarity constraints (MPEC) problem. Furthermore, comparison can be made between our model with a centralized capacity expansion decisions from a global point of view and the system from a more realistic point of view. The difference will tell how much we could benefit from a centralized point of view for the capacity decision making and provide insightful guidance for a potential policy design.

Finally, the third paper, presented in Chapters 4 and 5, proposes a novel hybrid algorithm to solve a more realistic market based generation and transmission expansion problem. It allows each decentralized GENCO to make its own investment and operational decisions in anticipation of a market price settled by an ISO market clearing problem and in response to a centralized transmission expansion decision by the ISO. The model poses a complicated tri-level structure including an EPEC sub-problem, in which each of its MPECs represents an individual GENCO's decision making problem. Algorithms are proposed to first solve the EPEC sub-problem and then the tri-level programming problem. The Complementarity Problem (CP) reformulation is capable to transform the entire EPEC problem into a set of constraints, so that the tri-level programs can be converted into a single-level mixed integer nonlinear program, which can be solved to identify a promising transmission plan. The Diagonalization Method (DM) is capable to solve the EPEC sub-problem reliably and efficiently given a predetermined transmission plan. A hybrid algorithm formed by by combining the two methods was first tested on a 6 bus system, in which the results demonstrate that the algorithm solved the instance to its global optimality. A modified

IEEE 30 bus test system is solved by the algorithm to test its scalability and solution of it has also been found successfully. An approximate Nash equilibrium (NE) is defined in the cases when the pure NE does not exist, and an analysis of the local concavity of the obtained (approximate) NE solution was conducted. A IEEE 118 bus test system is also solved by the algorithm and its solution has been validated as the globally optimal solution. For future research, an extended model with multi-period capacity expansion and operational decisions can be considered. The assumption of the one-period model makes the generation level always equal to the capacity level, which is not the case in reality. The algorithm's scalability on a stochastic version of the problem with consideration of the uncertainty variables can be tested in the future. Further investigation of the conditions for the existence and uniqueness of the NE solution is also needed.