HEURISTIC AND EXACT APPROACHES FOR VARIANTS OF THE UNIT COMMITMENT PROBLEM

by

Jennifer Van Dinter

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Mineral Economics).

Golden, Colorado
Date 1/13/2012

Signed: Jennifer Van Dinter
Jennifer Van Dinter

Signed: Alexandra M. Newman
Dr. Alexandra M. Newman
Thesis Advisor

Golden, Colorado
Date 1/13/2012

Signed: R. G. Eggert
Dr. Roderick Eggert
Professor and Director,
Division of Economics and Business

ABSTRACT

We present three variants of the unit commitment model which determine generator schedules, associated production and storage quantities, and spinning reserve requirements. Our optimization models minimize start-up, fuel, maintenance, and shortage costs, and in certain cases, shut-down and emissions costs. Constraint sets balance load, control the operation of power generators, renewable generators and storage devices, and track reserve requirements. We capture cost functions with piecewise-linear constructs.

In Variant 1, we incorporate concave quadratic emissions curves into our model, as presented above, which results in a mixed-integer, nonlinear problem. We strengthen the formulation via cut addition. We describe an underestimation approach to obtain an initial feasible solution to our model. Finally, we constitute a Benders' master problem from the scheduling variables and a subset of those variables associated with the nonlinear constructs; the subproblem contains the storage and reserve requirement quantities, and power from generators with linear emissions curves. We demonstrate that our strengthening techniques and Benders' Decomposition approach solve our version of the unit commitment model more quickly than standard global optimization algorithms. We present numerical results based on a subset of the Colorado power system that provide insights regarding storage, renewable generators, and emissions.

Variant 2 relies on our mixed-integer linear variant of the unit commitment model, including the flexibility to model either concave or convex piecewise-linear heat input curves. We use a data set representing Xcel Energy's Colorado territory with this variant. We implement a *sliding time window heuristic* to estimate the annual cost savings Xcel might capture by reducing the amount of spinning reserves it carries on the system. We also evaluate the expected greenhouse gas emissions for

the annual horizon.

Finally, in Variant 3, we introduce a dual-time-index binary variable into Variant 2, replacing the single-time-index binary variables. This new variable prevents generators from contributing generation to the grid until they have reached minimum capacity. We also introduce methods to increase tractability of the formulation. We compare operating costs over a 48-hour horizon using this variant and Variant 2 to determine how precluding generation during ramping changes the total operating cost for the Colorado system.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF FIGURES viii

LIST OF TABLES ix

ACKNOWLEDGMENT xi

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 THE UNIT COMMITMENT MODEL WITH CONCAVE EMISSIONS COSTS: A HYBRID BENDERS' DECOMPOSITION WITH NONCONVEX MASTER PROBLEMS 4

2.1 Literature Review 6
2.1.1 Classical Unit Commitment Model and Bender's Decomposition 6
2.1.2 Incorporation of Renewable Generation and Storage Devices . 8
2.2 Model 9
2.2.1 Generator Dispatch Optimization Model Formulation 12
2.2.2 Detailed Model Description 19
2.3 Proposed Solution Techniques 22
2.3.1 Simple Cuts 22
2.3.2 Cuts Based on Auxiliary Problem Solves 25
2.3.3 Underestimators 26
2.3.4 Preprocessing 27
2.3.5 Benders' Decomposition 28
2.3.6 Benders' Cuts 34
2.3.7 Summary: A Hybrid Benders' Decomposition Algorithm . . . 35
2.4 Computational Results: Solution Time Comparisons and Test System Results 37
2.4.1 Solution Time Comparisons 39
2.4.2 Test System Results 42
2.5 Conclusion 43

CHAPTER 3 THE COST OF PROVIDING SPINNING RESERVES 45

3.1 Literature Review 46
3.1.1 Typical Unit Commitment Model 47
3.1.2 Incorporation of Storage Devices 48

3.1.3 Heuristics for Creating Long-Term Generation Schedules . . . 48
3.1.4 Demand Response and Spinning Reserves 50
3.2 Model 52
3.2.1 Generator Dispatch Optimization Model Formulation 54
3.2.2 Detailed Model Description 61
3.3 Proposed Solution Techniques 65
3.3.1 Variable and Constraint Elimination 66
3.3.2 Cuts 67
3.3.3 Modeling Language Implementation 67
3.3.4 Sliding Time Window Heuristic 68
3.4 Computational Results 71
3.4.1 Data 71
3.4.2 Numerical Results for a Coal-Dominated Data Set 73
3.4.3 Numerical Results for a Gas-Dominated Data Set 79
3.5 Conclusion 80

CHAPTER 4 COMPARISON OF SINGLE- AND DUAL-TIME-INDEX VARIABLES USING THE UNIT COMMITMENT MODEL 82

4.1 Literature Review 84
4.2 Formulations 85
4.2.1 Single-Time-Index Binary Decision Variables and Generation Logic 86
4.2.2 Dual-Time-Index Binary Decision Variables and Generation Logic 87
4.3 Solution Techniques 87
4.3.1 Variable and Constraint Elimination 87
4.3.2 Constraint Sparsification 89
4.4 Comparison of Formulations and Numerical Results 91
4.4.1 Data 91
4.4.2 Numerical Results 92
4.4.3 Sparsification Impact on Dual-Time-Index Formulation 93
4.4.4 Comparison of Single- and Dual-Time Index Formulations . . 93
4.4.5 Forcing Shut Down in the Single-Time-Index Formulations . . 95
4.5 Conclusions 96

CHAPTER 5 CONCLUSIONS 98

5.1 Research Contributions 98
5.2 Possible Research Extensions 99

REFERENCES CITED 100

APPENDIX CONTENTS OF CD-ROM 105

List of GAMS and AMPL data files 105

CD-ROM OF DATA. . . . . Pocket

LIST OF FIGURES

Figure 2.1	Forbidden and allowable dispatch schedules	10
Figure 2.2	Sample quadratic NOx curve	11
Figure 2.3	Storage values over a 24-hour time period for each of the four seasons. Note that the storage device is not usable until time period 2.	42
Figure 2.4	Renewable energy used over a 24-hour horizon for each of the four seasons.	43
Figure 3.1	For conventional generators, we prohibit instances of overlapping start-up and shut-down times. For example, if a conventional generator operates between t_0 and t', it may not also start up in v_0, which occurs before t_0, and shut down in v', which occurs after t'. Other prohibited instances of (v_0, v') include starting up in v_0 and shutting down in v' when: (i) $v_0 < t_0$ and $t_0 < v' < t'$; (ii) $t_0 < v_0 < v' < t'$; and (iii) $t_0 < v_0 < t' < v'$. A conventional generator may start up in v_0 if either $v_0 < v' < t_0$ or $t' < v_0 < v'$. In other words, the instance (v_0, v') could either precede or follow (t_0, t').	63
Figure 3.2	Generator in *steady state* at the beginning of the current window	70
Figure 3.3	Generator in *minimum down time* at the beginning of the current window	71
Figure 3.4	Allocation of steam turbine (ST) to gas turbines (GTs) to derive combined cycle unit capacities	73
Figure 3.5	Total costs for five spinning reserve scenarios.	75
Figure 3.6	CO_2 emissions for five spinning reserve scenarios.	76
Figure 3.7	NO_x and SO_x emissions for five spinning reserve scenarios.	77
Figure 4.1	Histogram of sparsified and dense formulations.	90

LIST OF TABLES

Table 2.1 Numerical results for the tightened formulation: Comparison of the monolith and the one- and two-segment Benders' approximations root relaxations 40

Table 2.2 Numerical results for Benders: Comparison of the monolith solution time and quality against that of using the one- and two-segment Benders' approximations. We state the gap relative to the optimal objective function value, which the two-segment Benders' procedure provides. 41

Table 3.1 Summary of data: Generator statistics 72

Table 3.2 Comparison of annual costs of providing varying levels of spinning reserves 74

Table 3.3 Comparison of annual emissions resulting from varying levels of spinning reserves 76

Table 3.4 Estimation of annual costs of providing varying levels of spinning reserves: LP relaxation 77

Table 3.5 Comparison of monthly costs of providing varying levels of spinning reserves 78

Table 3.6 Comparison of monthly emissions resulting from varying levels of spinning reserves 78

Table 3.7 Summary of data: Generator statistics 79

Table 3.8 Comparison of annual costs of providing varying levels of spinning reserves: Gas-dominated data set 80

Table 3.9 Comparison of annual emissions resulting from varying levels of spinning reserves: Gas-dominated data set 81

Table 4.1 Comparison dual-time-index problem statistics, 83 generators and 48-hour horizon 90

Table 4.2 Summary of data: Generator statistics 92

Table 4.3 Dual-time-index formulation with and without sparsification . 93

Table 4.4 Comparison of single- and dual-time-index problem statistics, 48-hour horizon 94

Table 4.5 Comparison of single- and dual-time-index objective function values, 48-hour horizon 95

Table 4.6 Comparison of single- and dual-time-index objective function values and forcing shut down, 48-hour horizon 96

ACKNOWLEDGMENT

I wish to thank my advisor, Dr. Alexandra Newman and my research collaborators, Dr. Steffen Rebennack at the Colorado School of Mines, Dr. Josef Kallrath at the University of Florida and Dr. Paul Denholm at the National Renewable Energy Laboratory for their guidance and perseverance with respect to unit commitment models and Benders' Decomposition. Additionally, I would like to thank my dissertation committee members: Dr. Robert Braun, Dr. Paul Denholm, Dr. Daniel Kaffine, Dr. Joy Godesiabois, and Dr. Michael Walls for their suggestions to improve this work. Finally, I wish to thank my husband, Dave Lefaive, and my father, Bernie Van Dinter, for their unfailing love, support, and encouragement over the last four years.

CHAPTER 1

INTRODUCTION

Regional Transmission Organizations (RTOs) and Independent System Operators (ISOs) use unit commitment models to determine the optimal generator dispatch and transmission of electricity based on anticipated consumer demand. These unit commitment models consider a series of complex operating decisions and identify the lowest operating cost schedule. Dispatching generators is challenging because the power grid requires balance between supply of electricity from utilities and demand from customers. The power grid cannot simply discard excess electricity nor can it fail to provide customers with power, barring a catastrophic event. Should a failure occur, adequate back-up capacity must be immediately available to fill the void.

Specifically, our three variants of the dispatch model determine when generators start up, operate and shut down based on start-up costs, maintenance costs, shortage costs, and a generator's fuel consumption. The models include fuel consumption values determined using piecewise-linear approximations of heat input curves. This type of curve relates the level of generation with the level fuel usage required for a specific generator. Emissions curves estimate the amount of pollutants emitted based on the amount of generation from each generator. Each generator exhibits a variety of operating characteristics in addition to its heat input. A generator cannot provide electricity to the power grid until it reaches a minimum steady-state operating level and each generator cannot provide more electricity than its nameplate capacity. Additionally, ramp-up and ramp-down rates restrict changes in a generator's operating level over a given time horizon.

Historically, RTOs and ISOs did not solve their mixed integer unit commitment models using traditional branch-and-bound techniques. Instead, Lagrangian Relaxation and heuristic techniques provided generator schedules; if a heuristic method is employed, the quality of the solution may not be known. Today, many of these RTOs and ISOs use mixed integer programming (MIP) branch-and-bound solution methods to produce their day-ahead schedules. Solving the problems exactly, or at least to a small optimality gap, may provide significant cost savings. The Southwest Power Pool reports an estimated annual cost savings of $103 million in 2009 and the PJM (Pennsylvania, New Jersey, Maryland) Interconnection expected savings of $100 million in 2006 (O'Neill and Krall, 2011). Thus, improving tractability of the unit commitment problem and/or more accurately modeling the power system could yield additional cost savings.

Our research contributions include both theoretical and policy work. From the theoretical perspective, we develop a new solution methodology for intractable mixed-integer, nonlinear (MINLP) unit commitment problems and we reformulate the unit commitment problem to more accurately capture generator behavior during the ramp-up and ramp-down phases. Our policy work includes assessing the operating cost of supplying spinning reserves.

In Chapter 2, we present a hybrid Benders' Decomposition algorithm for solving a mixed-integer, nonlinear (MINLP) unit commitment problem, specifically, a unit commitment problem with concave quadratic emissions curves for certain generators. Because few commercial algorithms can solve MINLPs, we use a small data set containing 15 generators. Our hybrid algorithm provides information from piecewise-linear underestimators, mathematical cuts, and information from the auxiliary problems to the traditional Benders' Decomposition algorithm in an effort to expedite convergence. We compare our results to those obtained using one of the industry standard MINLP solvers, BARON (Tawarmalani and Sahinidis, 2004). A version of this chapter has been revised and resubmitted to *Annals of Operations Research.*

Chapter 3 addresses the costs associated with the provision of spinning reserves

(SR). In the case of an unexpected outage, the unit commitment model continues to meet demand by relying on spinning reserves to instantaneously satisfy a demand shortfall. The provision of these reserves also has the potential to increase the operating costs of the power system. We use a *sliding time window heuristic* to approximate the cost of SR. We solve each 48-hour horizon of a mixed-integer linear problem (MILP) variant of the unit commitment model to (near) optimality and fix the solution for the first 24 hours. We then move our horizon forward by one day and solve using the fixed variable information. We evaluate the cost of SR using a data set which characterizes Xcel Energy's generation system for Colorado in 2005. Expected emissions of greenhouse gases are determined a posteriori.

We introduce an alternative formulation of the MILP unit commitment model in Chapter 4. We define a new, dual-time-index binary variable to be used in lieu of the single-time-index variables used by Baldick (1995), Takriti *et al.* (2000) and Sioshansi *et al.* (2008). This new variable provides greater detail regarding the ramping status of a power generator and prevents it from producing electricity before reaching minimum capacity. This additional model fidelity may be of particular importance in sub-hourly dispatch scheduling.

Given the complexity of the unit commitment problem, there are many factors we omit from the unit commitment model variants we present in this dissertation. Specifically, we ignore aspects related to transmission, electricity consumption during start up of fossil-fuel generators, and allowances on SO_2 emissions governed under the US Environmental Protection Agency's Acid Rain Program. Historically, these permits traded for approximately \$100 to \$300 per ton of SO_2 (Burtraw and Szambelan, 2009).

Chapter 5 concludes with a discussion of additional areas of research on the unit commitment problem. We also present a summary of the relevant results and research contributions of this dissertation.

CHAPTER 2

THE UNIT COMMITMENT MODEL WITH CONCAVE EMISSIONS COSTS: A HYBRID BENDERS' DECOMPOSITION WITH NONCONVEX MASTER PROBLEMS

Power systems deliver electricity to consumers at all times of the day and night. A series of complex decisions made by utility operators ensures that customer demand for electricity is satisfied. The classical unit commitment model, or generator dispatch model, determines when generators start up, operate and shut down to minimize start-up costs and generator fuel consumption based on a day-ahead scheduling procedure. Generator dispatch decisions must consider the power system requirements to balance the supply of electricity from utilities and the demand from customers. Each generator exhibits a variety of operating characteristics. For example, each generator requires a certain amount of fuel input to produce a given level of electricity; while many models assume a linear relationship, we model heat input curves as piecewise linear. A generator cannot provide electricity to the power system until it reaches a minimum steady-state operating level, and each generator cannot provide more electricity than its nominal capacity. Additionally, ramp-up and ramp-down rates restrict changes in a generator's operating level over a given time segment.

Regional Transmission Organizations (RTOs) and Independent System Operators (ISOs) employ unit commitment models, such as the one described above, to determine the lowest cost generator commitment and dispatch schedule. Often, these organizations optimize their unit commitment models over large regions including hundreds, if not thousands of generators. Thus, tractability must be preserved at all costs.

To help mitigate events such as generator failures, our unit commitment model attempts to ensure the reliability of the power supply by providing spinning reserves. Specifically, we consider secondary, or contingency, spinning reserves (SR), which are intended to deliver reserve power in order to bring back the frequency and the interchange programs to their target values (Rebours and Kirschen, 2005). We determine which power generators and/or the storage device provide spinning reserves in our program. However, due to the deterministic nature of our problem, we do not include spinning reserves in the hourly demand constraint, because we do not consider generator failures nor do we dispatch spinning reserves to the power system.

We add renewable energy sources, e.g., solar and wind energy. We also consider the use of a electricity storage device, which offers a repository for excess electricity. Drawing electricity from these devices may help to mitigate a power shortage. However, storing, and later retrieving, electricity results in significant efficiency losses. Storage devices also possess capacity restrictions and other device-specific constraints. We consider horizons of up to 24 hours, consistent with day-ahead scheduling.

Our research contributions lie in the following: (i) we formulate a detailed unit commitment model with storage, renewable sources, piecewise-linear heat input curves, and nonlinear emissions (NO_x) curves, in addition to other details included in classical models; (ii) we present an efficient formulation for this problem in the way in which we use mathematical constructs such as cuts; (iii) we suggest a tailored Benders' Decomposition formulation whose implementation allows us to solve problem instances to near-optimality; and (iv) we present insights gained from a test system using our mixed integer nonlinear program.

The remainder of the chapter is organized as follows: In Section 2.1, we provide a literature review. We present our version of the unit commitment model in Section 2.2. In Section 2.3, we describe several improvements applicable to both the monolith and to our Benders' Decomposition procedure. Section 2.4 provides numerical results contrasting a naive implementation of the monolith with our suggested improvements, including Benders' Decomposition. We also demonstrate insights gained from our problem instances. Section 2.5 concludes the chapter.

2.1 Literature Review

In this section, we review the relevant literature. First, we provide a review of Benders' Decomposition applied to unit commitment models. We then summarize the adoption of renewable energy sources and storage devices into the unit commitment problem.

2.1.1 Classical Unit Commitment Model and Bender's Decomposition

Mixed integer linear programming-based (MILP) attempts to solve the classical deterministic unit commitment model began in the late 1970s (Dillon *et al.*, 1978). Lahdelma and Ruuth (1994) propose a Benders' Decomposition algorithm tailored to MILP formulations of the unit commitment problem where two special type of constraints are present: (i) bounding of continuous variables via integer variables and (ii) logical constraints on ramp up and shut down. Their work is theoretical in nature.

The seminal work of Takriti *et al.* (2000) incorporates stochastic elements (e.g., spot market, fuel price) into the classical unit commitment model to minimize operating costs. Takriti's MILP, which includes up to 729 scenarios, relies upon Lagrangian Relaxation and Benders' Decomposition to identify an optimal solution. Moya (2005) uses Benders' Decomposition to decompose an optimal economic dispatch problem with spinning reserves and load shedding with transmission. The author does not consider renewables, storage, or piecewise-linear constructs. The master problem is solved by dynamic simulation, passing on the dispatch and reserve decisions to the linear subproblems. Computational results for a single-period problem with fewer than 10 generators are reported for a Northern Chilean power system. Lu and Shahidehpour (2005) pose a detailed model considering transmission, emissions and fuel blending. The authors suggest a Benders' Decomposition method to solve instances containing a few generators and 24 hours. Wu and Shahidehpour (2010) derive special cuts for a Benders' master problem for an MILP formulation of a network-constrained unit commitment problem. They compute strong cuts by increasing the density of the support of the Benders' cuts. Extensive computational tests are performed on different power systems with up to 5663 busses.

Several researchers have applied generalized Benders' Decomposition to the unit commitment model. This technique is applicable when the subproblem is not linear, i.e., it is either nonlinear, possesses integer variables, or both. Generalized Benders' can ensure finite convergence of the algorithm to a global optimum if implemented correctly. Duran and Grossmann (1986) present a mixed-integer nonlinear programming (MINLP) algorithm based on an outer-approximation solution technique. In this instance, the master problem is an MILP and the subproblem is nonlinear program (NLP). Interestingly, the authors note that the outer approximation algorithm provides a tighter lower bound for their minimization problem than the lower bound from the generalized Benders' Decomposition algorithm. Ma and Shahidehpour (1999) use generalized Benders' to solve a unit commitment problem with transmission security and voltage constraints for one day using a 36 unit system within the IEEE-118 bus network. They also rely upon Lagrangian Relaxation to solve the difficult master problems.

Wang *et al.* (2008) use a security-constrained MINLP unit commitment problem to evaluate the impact from uncertain wind power. Separating the MILP master problem, i.e., the unit commitment problem, from the remainder of problem allows the authors to avoid the concerns regarding global optimality of nonlinear solutions. Once the master problem provides a dispatch schedule, feasibility of the solution must be verified to ensure adequate system flexibility to manage wind scenarios. Rahimi *et al.* (2009) also present a generalized Benders' Decomposition approach for the unit commitment model. Their master problem is an MILP problem incorporating spinning reserve constraints and minimum up and down time constraints while minimizing production, shut-down and start-up costs. The subproblems are continuous quadratic economic dispatch problems to allocate power generation. Niknam *et al.* (2009) consider transmission and emissions, but omit storage and renewable devices. They too suggest a generalized Benders' Decomposition approach in which the subproblem is nonlinear due to cost functions and ramp rate constraints. The authors solve 100-generator 24-time period instances and compare costs.

2.1.2 Incorporation of Renewable Generation and Storage Devices

Gollmer *et al.* (2000) incorporate storage into their unit commitment model, which lacks some of the operational details of our model yet contains extensions such as nonlinear start-up costs and stochastic demand. They propose a Lagrangian procedure to solve instances of their problem containing 100 generators and 24 hourly time periods. Bouffard and Galiana (2008) consider the effect of incorporating wind generation on emissions from a power system. The authors develop a model in which demand is met for all wind forecast scenarios, avoiding the need to shed load.

Exarchakos *et al.* (2009) consider five different storage devices. Using a stochastic model, the authors maximize profit, including capital costs. The model determines when to use demand-side management (i.e., storing or discharging various storage devices) over either a one-day or three-day period to smooth the peaks and valleys common in a load curve. We currently model only pumped hydroelectric storage.

See Padhy (2004) for a survey of the literature containing 150 references on the unit commitment model; however, the survey emphasizes metaheuristics over finitely exact solution approaches. A variety of applied work regarding the impact of renewable energy and storage devices also exists; see Sioshansi *et al.* (2009), Tuohy *et al.* (2008) and Tuohy *et al.* (2009).

2.2 Model

Our unit commitment model minimizes total annual power system costs which include: fixed start-up and shut-down costs, operation and maintenance costs, variable fuel costs, emissions taxes, and the costs of not meeting demand. Utilities currently dispatch generators on an hourly or quarter-hourly basis throughout the year. Generally, utilities wish to run their "base load" generators (typically, coal and nuclear plants) continuously and meet variations in demand with "intermediate load" and "peaker" natural gas combustion turbine plants. Renewable generators reduce fuel use from marginally profitable plants during each time interval. Due to each generator's unique ramping characteristics, the decision to dispatch a given generator often happens well in advance of an expected change in load.

We incorporate renewable energy sources into the classical unit commitment model. Unlike conventional generators, which have a static maximum capacity, renewable generators are characterized by time-varying and non-dispatchable output. We prevent renewable generators from contributing to operating reserves in our model due to their stochastic nature.

For conventional generators, we prohibit instances of overlapping start-up and shut-down times, as shown in Figure 2.1. For example, if a conventional generator operates between t_0 and t', it may not also start up in v_0, which occurs before t_0, and shut down in v', which occurs after t'. Other prohibited instances of (v_0, v') include starting up in v_0 and shutting down in v' when: (i) $v_0 < t_0$ and $t_0 < v' < t'$; (ii) $t_0 < v_0 < v' < t'$; and (iii) $t_0 < v_0 < t' < v'$. A conventional generator may start up in v_0 if either $v_0 < v' < t_0$ or $t' < v_0 < v'$. In other words, the instance (v_0, v') could either precede or follow (t_0, t').

Consistent with the classical unit commitment model, our mathematical programming model also contains the following types of constraints: load balance (i.e., meeting demand requirements) and spinning reserve requirements based on the ca-

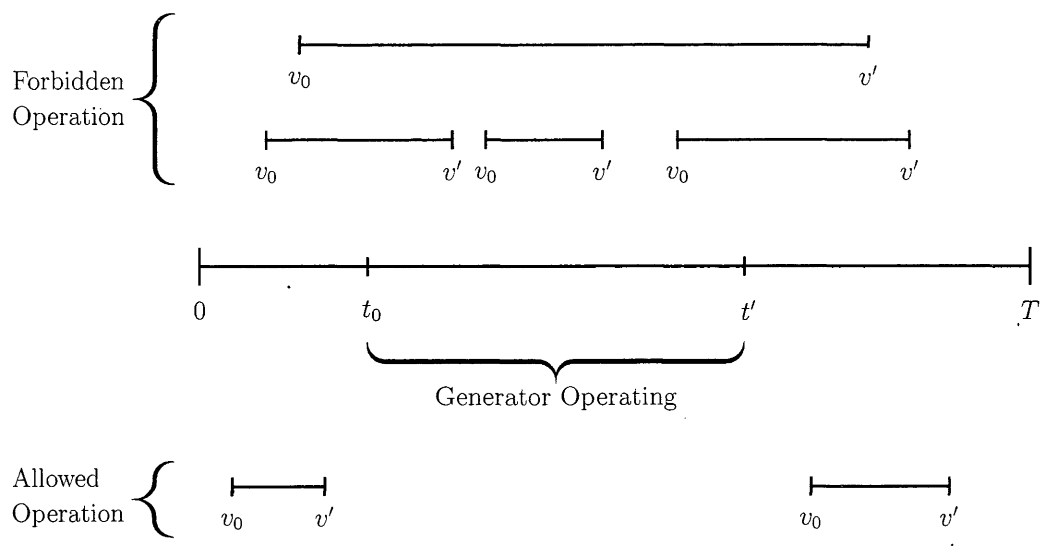

Figure 2.1. Forbidden and allowable dispatch schedules

pacity of the largest generator operating. The model also contains generator-specific constraints including: ramp rate restrictions, minimum and maximum capacity limitations, and minimum up and down times.

Additionally, our model incorporates storage. In most systems, supply and demand must match perfectly, requiring ongoing adjustments to the dispatch schedule. Most other energy-related industries, including oil and natural gas production, have flexible operations that can, to some extent, store their products in inventory. Storage provides the flexibility of allowing supply to exceed demand, which means that a system supplying energy may produce at its maximum efficiency without adjusting generation level to meet demand exactly. In addition, storage allows production to fall below demand, and, in such cases, the power stored supplements production. We incorporate constraints relating to the storage device as follows: an inventory balance constraint ensures that the electricity in storage from one period to the next reflects inflows and outflows in the prior period as well as the prior period's initial storage amount. We also incorporate the storage efficiency loss in the inventory balance constraint. Finally, we restrict the overall capacity of the storage device, and flow into and out of the device based on the required minimum and maximum flow rates for each flow direction.

We incorporate detailed representations of conventional generator characteristics. Classical unit commitment models often assume linear relationships between fuel use and generation level. However, generators often become less efficient and require more fuel at higher levels of operation, reflecting a nonlinear relationship (Gollmer *et al.*, 2000). Therefore, to include more accurate representations of these conventional characteristics, we use piecewise-linear heat input curves for the generators. We also model emissions curves as nonlinear constructs (see Figure 2.2). Specifically, we approximate the NO_x emissions curves from three generators with a constant and quadratic term (with a negative coefficient) for three generators in our data set, including two combined cycle units. That is, several natural gas generators possess concave emissions curves. We assume that the other generators exhibit linear (constant rate) emissions.

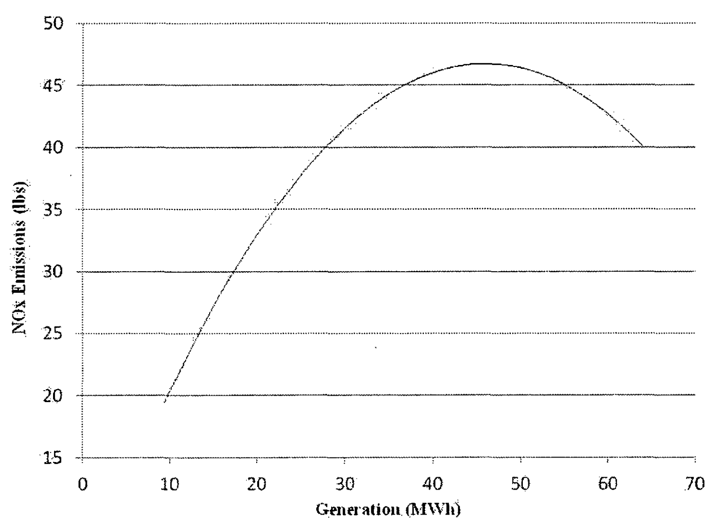

Figure 2.2. Sample quadratic NOx curve

The model we present includes many important aspects of generator commitment and dispatch, however, we do not capture all features which impact the generation schedule. Specifically, we omit transmission and distribution, similar to other unit commitment models including Gollmer *et al.* (2000) and Sioshansi (2010a). Omitting transmission constraints, we may obtain a solution with an undesirable geographic distribution of generation. Additionally, we consider NO_x emissions but not CO_2 or SO_x in this variant of the unit commitment model. Given that the US Environmental Protection Agency's Acid Rain Program limits the emission of SO_2, our generator dispatch decisions may be skewed toward coal generation.

2.2.1 Generator Dispatch Optimization Model Formulation

We now present our monolith model, ($\mathrm{P_{UCP}}$). We use lower case Roman letters to represent parameters, upper case Roman letters in regular font to represent variables, and upper case Roman letters in script font to denote sets. We use Greek letters to denote dual variable values on relevant constraints, i.e., those constraints with non-zero right-hand-side parameters, either in their own right or after binary variables are fixed at given values (see §2.3.5). Superscripts provide descriptions for various parameters while subscripts on parameters and variables represent indices, i.e., elements of a set.

Sets:

- $\mathcal{G}$: set of generators
- $\mathcal{R} \subseteq \mathcal{G}$: set of renewable generators
- $\mathcal{C} \subseteq \mathcal{G}$: set of conventional generators
- $\mathcal{I}$: set of intervals i for the piecewise-linear approximation of the heat input curves
- $\mathcal{T}$: time horizon over which the system is optimized

Data:

Prices

- p_g^f: price of fuel for generator g [$/MMBtu]
- $p_g^{\bar{s}}$: price of start up for generator g [$]
- $p_g^{\underline{s}}$: price of shut down for generator g [$]
- p_g^{om}: operation and maintenance costs for generator g [$/MWh]
- $\hat{p}$: elastic penalty for not meeting demand [$/MWh]
- p^{nox}: tax on NO_x emissions from power generators [$/lb]

Supplies and Demands

- $\underline{c}_g$: minimum hourly generation of conventional generator g [MWh]
- $\bar{c}_g$: maximum hourly generation of conventional generator g [MWh]
- c_g^s: amount of hourly spinning generation of conventional generator g [MWh]
- $\bar{c}_{gt}^r$: maximum hourly generation of renewable generator g in hour t [MWh]
- d_t: demand at hour t [MWh]

Generator Rates and Times

- r_g^{r+}: ramp-up rate of generator g [MWh/hr]
- r_g^{r-}: ramp-down rate of generator g [MWh/hr]
- r_g^{t+}: ramp-up time required for generator g [hr]
- r_g^{t-}: ramp-down time required for generator g [hr]
- $\underline{t}_g^u$: minimum up time of generator g before shut down [hr]

- $\underline{t}_g^d$: minimum down time of generator g before start up [hr]

Storage Characteristics

- s^l: percentage efficiency loss from storage
- $\underline{s}^e$: minimum quantity of energy allowed in storage [MWh]
- $\bar{s}^e$: maximum quantity of energy allowed in storage [MWh]
- $\underline{s}^{f+}$: minimum flow into storage, i.e., pumping [MWh/hr]
- $\bar{s}^{f+}$: maximum flow into storage, i.e., pumping [MWh/hr]
- $\underline{s}^{f-}$: minimum flow from storage, i.e., generating [MWh/hr]
- $\bar{s}^{f-}$: maximum flow from storage, i.e., generating [MWh/hr]

Slopes and Intercepts for Convex Piecewise Approximation

- m_{gi}: slope of the linear approximation of the heat input curve for generator g in interval i [MMBtu/MWh]
- b_g: intercept of the first segment of the linear approximation of the heat input curve for generator g [MMBtu]
- k_{gi}: length of interval i on the linear approximation of the heat input curve for generator g [MWh]

Coefficients for Quadratic Approximation

- a_g^q: coefficient of the quadratic term of the approximation of the NO_x emissions curve for generator g [lb/MWh^2]
- a_g^l: coefficient of the linear term of the approximation of the NO_x emissions curve for generator g [lb/MWh]

- a_g^c: constant of the approximation of the NO_x emissions curve for generator g [lb]

Continuous Decision Variables:

Generation

- X_{gt}: electricity production from generator g at time t [MWh]
- X_{gt}^{sp}: spinning reserves provided by generator g at time t [MWh]
- X_{gt}^{sr}: spinning reserves required for each generator g at time t [MWh]
- X_t^{st}: total spinning reserves required at time t [MWh]
- F_{gt}: fuel used by generator g at time t [MMBtu]
- $\hat{D}_t$: unmet demand at time t [MWh]

Storage

- S_t: amount of energy in storage at the beginning of time t; $S_1 = 0$ [MWh]
- S_t^+: energy inflow to storage during time t [MWh]
- S_t^-: energy outflow from storage during time t [MWh]
- S_t^{s-}: spinning reserves from storage during time t [MWh]

Piecewise-Linear Approximation

- L_{git}: hourly capacity at which generator g is operating on the heat input curve at time t in interval i [MWh]

Binary Decision Variables:

Generation

- $Y_{gt} = \begin{cases} 1 & \text{if generator } g \text{ starts up at time } t \\ 0 & \text{otherwise} \end{cases}$

- $\tilde{Y}_{gt} = \begin{cases} 1 & \text{if generator } g \text{ operates in steady state at time } t \\ 0 & \text{otherwise} \end{cases}$

- $\hat{Y}_{gt} = \begin{cases} 1 & \text{if generator } g \text{ shuts down at time } t \\ 0 & \text{otherwise} \end{cases}$

Storage

- $W_t^+ = \begin{cases} 1 & \text{if there is inflow into the storage device at time } t \\ 0 & \text{otherwise} \end{cases}$

- $W_t^- = \begin{cases} 1 & \text{if there is outflow from the storage device at time } t \\ 0 & \text{otherwise} \end{cases}$

Formulation:

$$
\begin{aligned}
(\mathrm{P_{UCP}}) \quad \min \quad & \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p_g^f F_{gt} + \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p_g^{om} X_{gt} + \sum_{t\in\mathcal{T}} \hat{p}\hat{D}_t \\
& + \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p^{nox}\left(a_g^q X_{gt}^2 + a_g^l X_{gt} + a_g^c \tilde{Y}_{gt}\right) \\
& + \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p_g^{\bar{s}} Y_{gt} + \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p_g^{\underline{s}} \hat{Y}_{gt}
\end{aligned}
$$

subject to

(see §2.2.2)

$$\sum_{g\in\mathcal{G}} X_{gt} + S_t^- - S_t^+ + \hat{D}_t = d_t \quad \forall\, t \in \mathcal{T} \qquad (\lambda_{1t}) \tag{2.1a}$$

$$X_t^{st} \geq X_{gt}^{sr} \qquad \forall g \in \mathcal{C}, t \in \mathcal{T} \tag{2.1b}$$

$$\sum_{g\in\mathcal{C}} X_{gt}^{sp} + S_t^{s-} \geq X_t^{st} \qquad \forall t \in \mathcal{T} \tag{2.1c}$$

$$X_{gt} - X_{g,t-1} + X_{gt}^{sp} \leq r_g^{r+} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T},\ \ t > 1 \qquad (\hat{\lambda}_{1gt}) \tag{2.1d}$$

$$X_{g,t-1} - X_{gt} \leq r_g^{r-} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T},\ \ t > 1 \qquad (\hat{\lambda}_{2gt}) \tag{2.1e}$$

$$X_{gt} \leq \bar{c}^r_{gt} \quad \forall\, g \in \mathcal{R}, t \in \mathcal{T} \qquad (\hat{\lambda}_{3gt}) \qquad (2.1f)$$

$$S_t = S_{t-1} + (1 - s^l)S^+_{t-1} - S^-_{t-1} \qquad \forall\, t \in \mathcal{T}, t > 1 \qquad (2.1g)$$

$$S^{s-}_t \leq S^+_t \qquad \forall t \in \mathcal{T} \qquad (2.1h)$$

$$S^-_t - S^+_t + S^{s-}_t \leq S_t \qquad \forall t \in \mathcal{T} \qquad (2.1i)$$

$$S^-_t + S^{s-}_t \leq \bar{s}^{f-} \qquad \forall\, t \in \mathcal{T} \qquad (\lambda_{2t}) \qquad (2.1j)$$

$$S_t \geq \underline{s}^e \qquad \forall\, t \in \mathcal{T} \qquad (\lambda_{3t}) \qquad (2.1k)$$

$$S_t \leq \bar{s}^e \qquad \forall\, t \in \mathcal{T} \qquad (\lambda_{4t}) \qquad (2.1l)$$

(see §2.2.2)

$$Y_{gt} \geq \tilde{Y}_{gt} - \tilde{Y}_{g,t-1} \quad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \qquad (2.2a)$$

$$\hat{Y}_{gt} \geq \tilde{Y}_{g,t-1} - \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C}, t \in \mathcal{T},\ t > 1 \qquad (2.2b)$$

$$\sum_{t'=t-\underline{t}^u_g - r^{t+}_g - r^{t-}_g}^{t} Y_{gt'} \leq \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C},\ t \in \mathcal{T},\ t > 1 \qquad (2.2c)$$

$$\sum_{t'=t-\underline{t}^d_g}^{t} \hat{Y}_{gt'} \leq 1 - \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C},\ t \in \mathcal{T} \qquad (2.2d)$$

(see §2.2.2)

$$X_{gt} + X^{sp}_{gt} \leq \bar{c}_g \tilde{Y}_{gt} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \qquad (\mu_{1gt}) \qquad (2.3a)$$

$$X^{sr}_{gt} = \bar{c}_g \tilde{Y}_{gt} \qquad \forall g \in \mathcal{C}, t \in \mathcal{T} \qquad (\mu_{2gt}) \qquad (2.3b)$$

$$X^{sp}_{gt} \leq c^s_g \tilde{Y}_{gt} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \qquad (\mu_{3gt}) \qquad (2.3c)$$

$$X_{gt} \geq \underline{c}_g \tilde{Y}_{gt} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \qquad (\mu_{4gt}) \qquad (2.3d)$$

(see §2.2.2)

$$W_t^+ + W_t^- \leq 1 \qquad \forall t \in \mathcal{T} \tag{2.4}$$

(see §2.2.2)

$$S_t^+ \leq \bar{s}^{f+} W_t^+ \qquad \forall\, t \in \mathcal{T} \qquad (\nu_{1t}) \tag{2.5a}$$

$$S_t^+ \geq \underline{s}^{f+} W_t^+ \qquad \forall\, t \in \mathcal{T} \qquad (\nu_{2t}) \tag{2.5b}$$

$$S_t^- \leq \bar{s}^{f-} W_t^- \qquad \forall\, t \in \mathcal{T} \qquad (\nu_{3t}) \tag{2.5c}$$

$$S_t^- \geq \underline{s}^{f-} W_t^- \qquad \forall\, t \in \mathcal{T} \qquad (\nu_{4t}) \tag{2.5d}$$

(see §2.2.2)

$$X_{gt} = \sum_{i \in I} L_{git} \qquad \forall\, g \in \mathcal{C},\ t \in \mathcal{T} \tag{2.6a}$$

$$L_{git} \leq k_{gi} \qquad \forall\, g \in \mathcal{C},\ \ i \in \mathcal{I},\ t \in \mathcal{T} \qquad (\hat{\theta}_{git}) \tag{2.6b}$$

(see §2.2.2)

$$F_{gt} = \sum_{i \in \mathcal{I}} m_{gi} L_{git} + b_g \tilde{Y}_{gt} \qquad \forall\, g \in \mathcal{C},\ t \in \mathcal{T} \qquad (\theta_{gt}) \tag{2.7}$$

(see §2.2.2)

$$X_{gt}, X_{g't}^{sp}, X_{g't}^{sr}, X_t^{st}, F_{g't}, L_{g'it}, \hat{D}_t, S_t, S_t^+, S_t^-, S_t^{s-} \geq 0$$

$$\forall\ \ g \in \mathcal{G},\ g' \in \mathcal{C},\ t \in \mathcal{T},\ i \in \mathcal{I}, \tag{2.8a}$$

$$Y_{g't}, \tilde{Y}_{g't}, \hat{Y}_{g't}, W_t^+, W_t^- \ \text{binary} \qquad \forall\, g' \in \mathcal{C},\ t \in \mathcal{T} \tag{2.8b}$$

2.2.2 Detailed Model Description

We next give a detailed description of the model formulation. The objective function seeks to minimize fuel costs, operating and maintenance costs, penalties associated with unmet demand, taxes on emissions of NO_x from power generators, and start-up and shut-down costs. We model the NO_x emissions using a quadratic approximation, where the quadratic term is either negative or zero, and the linear term can assume any value. (Therefore, not all generators have associated with their emissions a quadratic function, but if they do, the quadratic assumes a concave form.) Additionally, we penalize any failure to meet demand. The following subsections detail the constraints. We categorize the constraints not only by their general sense, but also by the types of variables they contain. The latter categorization is helpful for our Benders' Decomposition solution procedure (see §2.3.5).

Power Generation and Storage Constraint (2.1a) ensures that load is met using generation from conventional and renewable sources and a storage device. If generators and the storage device cannot provide sufficient generation, the elastic variable allows the demand constraint to be violated (and the objective function incurs a penalty in this case). We implement a portion of a linear minimum-maximum structure with constraint (2.1b) to identify the largest operating unit in each time period. Constraint (2.1c) ensures that sufficient spinning reserves are provided by operating generators and/or storage devices. The largest unit dictates the amount of spinning reserves that must be available in each time period. Constraints (2.1d) and (2.1e) restrict the rate of change in generation level between any two successive time periods to be less than or equal to the ramp-up or ramp-down rate, respectively. The sum of the change in generation level from one period to the next and any committed spinning reserves cannot exceed the ramp-up rate. This ensures that if a generator must provide reserves in addition to generation, the incremental ramping caused by the reserves does not force a generator to exceed its ramp-up rate. Constraint (2.1f) limits the amount of generation provided by renewable generators to a time-dependent hourly maximum output. Constraint (2.1g) ensures inventory balance with respect to the storage device, including losses incurred by storing power, for each period. Constraint (2.1h) limits the amount of spinning reserves provided by the storage device in any given period to be less than or equal to the amount of electricity flowing into the device. We note that the provision of spinning reserves only during inflow to the storage device is a simplifying assumption, necessary due to the complexity of the model. Constraint (2.1i) limits the amount of generation and reserves provided by the storage device to the amount of power already stored. Constraints (2.1j) limit the amount of outflow from the storage device to the maximum allowable. Constraints (2.1k) and (2.1l) limit the minimum and maximum hourly capacity of the storage device, respectively.

Generation Logic Constraint (2.2a) ensures that the generator has started up in at least the same period as steady-state operation occurs. Constraint (2.2b) prohibits any steady-state operation after a generator shuts down. Constraint (2.2c) enforces the minimum up-time requirement for each generator. Constraint (2.2d) enforces the minimum down-time requirement for each generator.

Generation Related to Spinning Reserve Logic Constraint (2.3a) limits the amount of energy and spinning reserves provided by an operating generator to be no more than the unit's nominal hourly capacity if it is operating. Constraint (2.3b) determines the amount of spinning reserves from each generator in each time period as a function of the generator's hourly capacity and whether or not the generator is operating. This, in conjunction with constraint (2.1b), forces spinning reserves to be greater than or equal to the maximum hourly capacity of the largest unit operating for each time period. Constraint (2.3c) limits the amount of spinning reserves a generator provides in any period based on its hourly spinning reserve capacity and its operating status. Constraint (2.3d) requires a generator to be operating at or above its minimum hourly capacity while in steady-state operation.

Storage Logic with Binary Variables Constraint (2.4) requires that flow move in at most one direction relative to the storage device, i.e., only inflow (pumping) or outflow (generation) may occur in each time period.

Storage Logic with Mixed Variables Constraints (2.5a) and (2.5b) provide upper and lower bounds on the inflow into the storage device (e.g., when pumping water into the storage pond), respectively. Constraint (2.5c) limits the amount of generation provided by the storage device to be less than or equal to the maximum outflow from the storage device. Constraint (2.5d) defines the minimum flow from the storage device per period.

Convex Piecewise Approximations with Continuous Variables The following set of constraints implements the piecewise-linear, convex approximation with respect to the continuous variables only. Constraint (2.6a) links the amount of electricity production to one or more segments of the curve. Constraint (2.6b) defines the operating level intervals for each segment. No binary variables are necessary as the piecewise-linear approximations are convex.

Convex Piecewise Approximation with Mixed Variables The following constraints link the continuous variables in §2.2.2 with the binary variables in §2.2.2. Constraint (2.7) determines the generator's fuel use based on the generation level, the independent variable, and the slope and intercept of the line segment.

Variable Type The final constraints, (2.8a) and (2.8b), restrict the variables to be nonnegative or binary, as appropriate.

2.3 Proposed Solution Techniques

Instances of our unit commitment model contain 15 generators, and up to 24 hourly time periods (corresponding to day-ahead planning). The corresponding problem sizes tax state-of-the-art commercial optimizers. Therefore, we suggest six techniques to make the model tractable: (i) addition of simple cuts; (ii) addition of cuts based on solutions to auxiliary problems; (iii) generation of underestimators; (iv) preprocessing; (v) implementation of a Benders' Decomposition algorithm, including (vi) cuts derived from Benders' solves. We detail each of these techniques here, and conclude this section with a summary of how we use our proposed techniques.

2.3.1 Simple Cuts

In order to tighten the formulation ($\mathrm{P_{UCP}}$), we add cuts. The main purpose of these cuts is to strengthen the master problem of the Benders' Algorithm (see §2.3.5) by providing insights into the Benders' subproblem structure. Saharidis *et al.* (2011) propose a similar idea for fixed-charge network problems. We can derive many cuts simply from the structure of formulation. We detail these in this subsection. In the following subsection, we describe cuts for which we must solve an auxiliary problem to derive.

- For every generator and time period, the number of times a generator has started must be greater than or equal to the number of times that the generator has shut down:

$$\sum_{t' \leq t} \dot{Y}_{gt'} \geq \sum_{t' \leq t} \hat{Y}_{gt'} \quad \forall g \in \mathcal{C}, t \in \mathcal{T} \tag{2.9}$$

- Because we assume that the storage device is initially empty, it is impossible to withdraw energy from storage in the first time period:

$$W_1^- = 0 \tag{2.10}$$

- It is impossible to withdraw energy from storage if no energy is currently stored.

$$\sum_{t'=1}^{t-1} W_{t'}^+ \geq W_t^- \quad \forall t \leq T \tag{2.11}$$

- In order to prevent an excessive number of generators from being turned on in a given time period, we limit the sum of the minimum capacities over the horizon and the maximum capacities of the renewable generators (a must-take source) to be less than or equal to the sum of the total demand and the total minimum energy withdrawn from storage.

  Letting $\bar{x}_t^{st}$ be the total amount of spinning reserves required at hour t, as determined by constraints (2.1b) and (2.3b) in ($\mathrm{P_{UCP}}$), and letting t_0 and t', respectively, be the beginning and ending time periods in the horizon under consideration (as in §2.2), we have:

$$\sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} \underline{c}_g \tilde{Y}_{gt} + \sum_{g \in \mathcal{R}} \sum_{t \in \mathcal{T}} \bar{c}_{gt}^r \leq \sum_{t \in \mathcal{T}} d_t + \bar{x}_t^{st}(t' - t_0) + \sum_{t \in \mathcal{T}} \underline{s}^{f+} W_t^+ \tag{2.12}$$

- Based on the capacity of the storage device and the minimum withdrawal rate, for any $\lfloor \frac{\bar{s}^e}{\underline{s}^{f-}} \rfloor = \alpha$ time periods, it is only possible to withdraw energy from the storage device in $\alpha - 1$ time periods.

$$\sum_{t'=t}^{t+\alpha} W_{t'}^{-} \leq \alpha - 1 \quad \forall t \in \mathcal{T} \tag{2.13}$$

- For each time period, storage must not exceed its upper bound. In the worst case, the inflow (less the loss factor) is at a minimum, and the outflow is at a maximum. Note that the loss factor is not applied to inflows in the current period.

$$\sum_{t'=1}^{t-1} (\underline{s}^{f+}(1 - s^{l})W_{t'}^{+} - \bar{s}^{f-}W_{t'}^{-}) + \underline{s}^{f+}W_{t}^{+} - \bar{s}^{f-}W_{t}^{-} \leq \bar{s}^{e} \quad \forall t \leq T \tag{2.14}$$

- For each time period, storage must not fall below its lower bound. The left hand side is computed as in (2.14), but using minimum outflow and maximum inflow.

$$\sum_{t'=1}^{t-1} (\bar{s}^{f+}(1 - s^{l})W_{t'}^{+} - \underline{s}^{f-}W_{t'}^{-}) + \bar{s}^{f+}W_{t}^{+} - \underline{s}^{f-}W_{t}^{-} \geq \underline{s}^{e} \quad \forall t \leq T \tag{2.15}$$

- To ensure that there exists adequate generation, we require that the sum of the maximum capacity of the conventional generators, the available renewable generation, and the minimum flow into storage must meet or exceed the sum of the demand and the spinning reserve requirement, less the maximum generation that can be withdrawn in any time period from storage.

$$\sum_{g \in \mathcal{C}} \bar{c}_g \tilde{y}_{gt} + \sum_{g \in \mathcal{R}} \bar{c}^r_{gt} + \underline{s}^{f-} W^+_t \geq d_t + \bar{x}^{st}_t - \bar{s}^{f-} W^-_t \quad \forall t \in \mathcal{T} \tag{2.16}$$

2.3.2 Cuts Based on Auxiliary Problem Solves

- We can bound the number of times that we can put energy into or withdraw energy from storage by solving an auxiliary model consisting of maximizing either $\sum_t W^+_t$ or $\sum_t W^-_t$, respectively, subject to all relevant constraints, and bounding the result from above.

  Specifically, we can impose the following bounds:

$$\sum_{t \in \mathcal{T}} W^+_t \leq \text{maximum number of times energy can be put into storage} \tag{2.17}$$

$$\sum_{t \in \mathcal{T}} W^-_t \leq \text{maximum number of times energy can be withdrawn from storage} \tag{2.18}$$

- In order to determine the minimum up time for all generators in the system, we solve ($\mathrm{P_{UCP}}$) with the following modification: we replace the objective function as given in §2.2.1 with the expression $\sum_{gt} \tilde{Y}_{gt}$. If we minimize this sum, then we obtain a minimum amount of up time, which we use on the right-hand-side of (2.19); if we maximize this sum, then we obtain a maximum amount of up time, which we use on the right-hand-side of (2.20).

$$\sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} \tilde{Y}_{gt} \geq \text{minimum up time} \tag{2.19}$$

$$\sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} \tilde{Y}_{gt} \leq \text{maximum up time} \tag{2.20}$$

Because deriving the cuts in this section requires only the solution of (small) mixed integer linear programs, we can obtain the relevant right hand sides in a matter of seconds.

2.3.3 Underestimators

The relaxation of nonlinear terms using piecewise linear functions is a standard technique to enable the solution of global optimization problems. Our minimization problem possesses concave (i.e., nonconvex) objective function terms. By substituting piecewise linear functions for these nonconvex cost functions, we obtain an MILP problem. We seek piecewise linear functions which possess the same value as or a lower value than each point in the domain of the original function. Such underestimator functions, when replacing nonconvex terms and when employed together with the remainder of the original (convex) problem, yield an optimal objective function value that provides a valid lower bound on the optimal objective function value of the original MINLP.

For our problem instances, piecewise linear functions with one or two segments produce good approximations of the concave cost functions. A one-segment underestimator is a linear function underestimating the concave cost function. A two-segment underestimator introduces a breakpoint in the middle of the support interval of the concave function. Thus, the two segment model yields a piecewise linear function employing two linear functions which are joined at the breakpoint, and underestimates the concave cost function.

2.3.4 Preprocessing

In our preprocessing phase, we employ the following ideas:

Underestimators: Because our problem is an MINLP in which the only nonlinearities consist of concave objective function terms, replacing these terms with underestimators transforms the problem into an MILP which has the same feasible region as the original MINLP. The relationship between the MINLP and the MILP is based on the following three characteristics: (i) any solution of the MILP, when evaluated in the objective function of the MINLP, yields an upper bound on the objective function of the MINLP; (ii) the optimal objective function value of the MILP yields a lower bound on the MINLP (see §2.3.3); and (iii) for any feasible solution of the MILP, the variables contained in the master problem can be used to generate a valid Benders' (optimality) cut by solving a Benders' subproblem, an LP.

Nonlinear programs: The binary portion of the solution obtained from the underestimator model can be fixed in the original model. This yields an NLP with continuous variables associated with the concave generators. The optimal objective function value produces an upper bound on the optimal solution of the monolith which is at least as strong as that obtained from the underestimator model.

Fixed generation: We exploit the best upper bound obtained at any point in the preprocessing as follows: For each generator, solve two MILP models (incorporating the underestimators for the concave cost terms), where exactly one generator is either turned off for the entire horizon or turned on at least once. For the case in which the generator is turned off, if the optimal objective function value is worse than the best upper bound, then this particular generator must be on in any optimal monolith solution, and we can add a constraint forcing the generator to be turned on at least once during the horizon. Conversely, when forcing a generator to be turned on at least once, if the resulting optimal objective function value is worse than the best upper bound, then this particular generator cannot be used in any optimal solution of the monolith. Thus, we can add a constraint forcing this particular generated to be turned off, e.g., we can remove the generator from the model.

The algorithm terminates in preprocessing when the lower and upper bounds are within a specified tolerance.

2.3.5 Benders' Decomposition

Benders' Decomposition divides the monolith into two parts, a master problem and a subproblem. The master problem determines the values for the continuous generation and fuel use variables for the generators with the concave emissions profiles, the binary variables which control the start-up and shut-down times and steady-state operation of the conventional generators, and the logic governing the intake and outflow of the storage device. The subproblem determines the values of the remaining variables, generally, the storage of electricity, the spinning reserves, the generation and fuel use of generators with linear emissions profiles, and the unmet demand. We begin this subsection with a description of both the master problem and the subproblem. We then briefly discuss the implementation of our algorithm using these two problems. A more detailed explanation of the algorithm appears in §2.3.7.

Let us define the following notation in conjunction with the Benders' Decomposition procedure:

- k: current iteration number
- η: master problem Benders' cut variable
- $\mathcal{Q} \subseteq \mathcal{G}$: power generators with concave quadratic approximation NO_x emissions curves
- $\mathcal{L}$: the set of extreme point vectors l at the k^{th} Benders' iteration (associated with the primal problem for the case in which the primal subproblem is feasible).

Note that we use a superscript index here to differentiate the iteration number from the indices associated with the monolith.

- λ^l_{wt}, $w = 1, \ldots, 4$; $\hat{\lambda}^l_{wgt}$, $w = 1, \ldots, 3$ (see §2.2.2)
- μ^l_{wgt}, $w = 1, \ldots, 4$ (see §2.2.2)
- ν^l_{wt}, $w = 1, \ldots, 4$ (see §2.2.2)
- θ^l_{gt}; $\hat{\theta}^l_{git}$ (see §2.2.2 and §2.2.2)

- $\mathcal{M}$: the set of extreme ray vectors m at the k^{th} Benders' iteration (associated with the dual problem for the case in which the primal subproblem is infeasible); these are vectors and are subscripted (and decorated) in the same way as the duals given in the monolith, ($\mathrm{P_{UCP}}$), but are also underlined. (Again, we use a superscript index here to differentiate the iteration number from the indices associated with the monolith.)

  - $\underline{\lambda}^m_{wt}$, $w = 1, \ldots, 4$; $\hat{\underline{\lambda}}^m_{wgt}$, $w = 1, \ldots, 3$ (see §2.2.2)
  - $\underline{\mu}^m_{wgt}$, $w = 1, \ldots, 4$ (see §2.2.2)
  - $\underline{\nu}^m_{wt}$, $w = 1, \ldots, 4$ (see §2.2.2)
  - $\underline{\theta}^m_{gt}$; $\hat{\underline{\theta}}^m_{git}$ (see §2.2.2 and §2.2.2)

Note that at any given iteration k, there exists either a set of extreme point vectors l or a set of extreme ray vectors m, but not both.

The Unit Commitment Master Problem ($\mathrm{MP_{UCP}}$) of the Benders' Decomposition scheme is defined as:

$$(\mathrm{MP_{UCP}}) \quad \min \ \eta + \sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} p^{\bar{s}}_g Y_{gt} + \sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} p^{\underline{s}}_g \hat{Y}_{gt} + \sum_{g \in \mathcal{Q}} \sum_{t \in \mathcal{T}} p^f_g F_{gt}$$
$$+ \sum_{g \in \mathcal{Q}} \sum_{t \in \mathcal{T}} p^{om}_g X_{gt} + \sum_{g \in \mathcal{Q}} \sum_{t \in \mathcal{T}} p^{nox} \left(a^q_g X^2_{gt} + a^l_g X_{gt} + a^c_g \tilde{Y}_{gt} \right) \tag{2.21}$$

subject to

$$\eta \geq \sum_{t\in\mathcal{T}} \lambda_{1t}^{l} \left(d_t - \sum_{g\in\mathcal{Q}} X_{gt} \right) + \sum_{t\in\mathcal{T}} \lambda_{2t}^{l} \bar{s}^{f-} - \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} \hat{\lambda}_{1gt}^{l} r_g^{r+}$$
$$- \sum_{g\in\mathcal{Q}, t\in\mathcal{T}} \hat{\lambda}_{1gt}^{l} \left(r_g^{r+} - X_{gt} + X_{g,t-1} \right) + \sum_{g\in\mathcal{Q}, t\in\mathcal{T}} \hat{\lambda}_{2gt}^{l} r_g^{r-}$$
$$+ \sum_{g\in\mathcal{R}, t\in\mathcal{T}} \hat{\lambda}_{3gt}^{l} \bar{c}_{gt}^{r} - \sum_{t\in\mathcal{T}} \lambda_{3t}^{l} \underline{s}^{e} + \sum_{t\in\mathcal{T}} \lambda_{4t}^{l} \bar{s}^{e} \tag{2.22a}$$
$$+ \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} \mu_{1gt}^{l} \left(\bar{c}_g \tilde{Y}_{gt} \right) + \sum_{g\in\mathcal{Q}, t\in\mathcal{T}} \mu_{1gt}^{l} \left(\bar{c}_g \tilde{Y}_{gt} - X_{gt} \right)$$
$$+ \sum_{g\in\mathcal{C}, t\in\mathcal{T}} \mu_{2gt}^{l} \bar{c}_g \tilde{Y}_{gt} + \sum_{g\in\mathcal{C}, t\in\mathcal{T}} \mu_{3gt}^{l} c_g^{s} \tilde{Y}_{gt} - \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} \mu_{4gt}^{l} \underline{c}_g \tilde{Y}_{gt} \tag{2.22b}$$
$$+ \sum_{t\in\mathcal{T}} \nu_{1t}^{l} \bar{s}^{f+} W_t^{+} - \sum_{t\in\mathcal{T}} \nu_{2t}^{l} \underline{s}^{f+} W_t^{+}$$
$$+ \sum_{t\in\mathcal{T}} \nu_{3t}^{l} \bar{s}^{f-} W_t^{-} - \sum_{t\in\mathcal{T}} \nu_{4t}^{l} \underline{s}^{f-} W_t^{-} \tag{2.22c}$$
$$+ \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, i\in\mathcal{I}, t\in\mathcal{T}} \hat{\theta}_{git}^{l} k_{gi} + \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} \theta_{gt}^{l} b_g \tilde{Y}_{gt} \qquad \forall\, l \in \mathcal{L} \tag{2.22d}$$

$$0 \geq \sum_{t\in\mathcal{T}} \underline{\lambda}_{1t}^{m} \left(d_t - \sum_{g\in\mathcal{Q}} X_{gt} \right) + \sum_{t\in\mathcal{T}} \underline{\lambda}_{2t}^{m} \bar{s}^{f-} - \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} \hat{\underline{\lambda}}_{1gt}^{m} r_g^{r+}$$
$$- \sum_{g\in\mathcal{Q}, t\in\mathcal{T}} \hat{\underline{\lambda}}_{1gt}^{m} \left(r_g^{r+} - X_{gt} + X_{g,t-1} \right)$$
$$+ \sum_{g\in\mathcal{Q}, t\in\mathcal{T}} \hat{\underline{\lambda}}_{2gt}^{m} r_g^{r-} + \sum_{g\in\mathcal{R}, t\in\mathcal{T}} \hat{\underline{\lambda}}_{3gt}^{m} \bar{c}_{gt}^{r} - \sum_{t\in\mathcal{T}} \underline{\lambda}_{3t}^{m} \underline{s}^{e} + \sum_{t\in\mathcal{T}} \underline{\lambda}_{4t}^{m} \bar{s}^{e}$$
$$+ \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} \underline{\mu}_{1gt}^{m} \left(\bar{c}_g \tilde{Y}_{gt} \right) + \sum_{g\in\mathcal{Q}, t\in\mathcal{T}} \underline{\mu}_{1gt}^{m} \left(\bar{c}_g \tilde{Y}_{gt} - X_{gt} \right)$$
$$+ \sum_{g\in\mathcal{C}, t\in\mathcal{T}} \underline{\mu}_{2gt}^{m} \bar{c}_g \tilde{Y}_{gt} + \sum_{g\in\mathcal{C}, t\in\mathcal{T}} \underline{\mu}_{3gt}^{m} c_g^{s} \tilde{Y}_{gt} - \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} \underline{\mu}_{4gt}^{m} \underline{c}_g \tilde{Y}_{gt}$$
$$+ \sum_{t\in\mathcal{T}} \underline{\nu}_{1t}^{m} \bar{s}^{f+} W_t^{+} - \sum_{t\in\mathcal{T}} \underline{\nu}_{2t}^{m} \underline{s}^{f+} W_t^{+}$$
$$+ \sum_{t\in\mathcal{T}} \underline{\nu}_{3t}^{m} \bar{s}^{f-} W_t^{-} - \sum_{t\in\mathcal{T}} \underline{\nu}_{4t}^{m} \underline{s}^{f-} W_t^{-}$$
$$+ \sum_{g\in\mathcal{C}\setminus\mathcal{Q},\ i\in\mathcal{I},\ t\in\mathcal{T}} \hat{\underline{\theta}}_{git}^{m} k_{gi} + \sum_{g\in\mathcal{C}\setminus\mathcal{Q},\ t\in\mathcal{T}} \underline{\theta}_{gt}^{m} b_g \tilde{Y}_{gt} \qquad \forall m \in \mathcal{M} \tag{2.23}$$

$$\text{constraints in } \S 2.2.2 : (2.1e) \text{ for } g \in \mathcal{Q},\ t \in \mathcal{T} \tag{2.24}$$

$$\text{constraints in } \S 2.2.2 : (2.2a) - (2.2d) \text{ for } g \in \mathcal{C},\ t \in \mathcal{T} \tag{2.25}$$

$$\text{constraints in } \S 2.2.2 : (2.3d) \text{ for } g \in \mathcal{Q},\ t \in \mathcal{T} \tag{2.26}$$

$$\text{constraints in } \S 2.2.2 : (2.4) \text{ for } t \in \mathcal{T} \tag{2.27}$$

$$\text{constraints in } \S 2.2.2 : (2.6a) - (2.6b) \text{ for } g \in \mathcal{Q},\ i \in \mathcal{I},\ t \in \mathcal{T} \tag{2.28}$$

$$\text{constraints in } \S 2.2.2 : (2.7) \text{ for } g \in \mathcal{Q},\ t \in \mathcal{T} \tag{2.29}$$

$$Y_{g't}, \tilde{Y}_{g't}, \hat{Y}_{g't}, W_t^+, W_t^- \text{ binary } \forall\, g' \in \mathcal{C};\ t \in \mathcal{T}; \tag{2.30}$$

$$X_{gt}, F_{gt}, L_{git} \geq 0,\ \forall g \in \mathcal{Q}, i \in \mathcal{I}, t \in \mathcal{T}; \eta \text{ free}, \tag{2.31}$$

where the first set of constraints constitutes the *objective function cuts* ((2.22a)-(2.22d)) and the second set, (2.23), the *feasibility cuts.* The first set of terms in the first constraint, (2.22a), represents the product of the right-hand-side constants, if nonzero, the associated master problem variables, X_{gt}, $g \in \mathcal{Q}$, of the constraints in §2.2.2, and the corresponding dual variable values; the second set of terms in the first constraint, (2.22b), represents the product of the right-hand-side constants, if nonzero, the associated master problem variables X_{gt}, $g \in \mathcal{Q}$ and $\tilde{Y}_{gt}$, $g \in \mathcal{G}$, of the constraints in (§2.2.2), and the corresponding dual variable values; the third set of terms in the first constraint, (2.22c), represents the product of the right hand side containing the master problem variables W_t^+ and W_t^- and their associated coefficients (§2.2.2), and the corresponding dual variable values; the fourth set of terms in the first constraint, (2.22d), represents the product of the right hand side containing the master problem variables, $\tilde{Y}_{gt}$, and their associated coefficients (§2.2.2), and the corresponding dual variable values. Constraints (2.23), the feasibility cuts, possess the same right-hand-side terms as the optimality cuts, except the dual multipliers are extreme rays instead.

For given continuous X_{gt}, F_{gt}, L_{git} ($g \in \mathcal{Q}, i \in \mathcal{I}, t \in \mathcal{T}$) and binary Y_{gt}, $\tilde{Y}_{gt}$, $\hat{Y}_{gt}$, W_t^+, and W_t^- ($g \in \mathcal{G}, t \in \mathcal{T}$) variables, the (primal) Benders' subproblems are derived by fixing these variables of problem ($\mathrm{P}_{\mathrm{UCP}}$) to the values x^k_{gt}, f^k_{gt}, l^k_{git}, y^k_{gt}, $\tilde{y}^k_{gt}$, $\hat{y}^k_{gt}$, w^{k+}_t, and w^{k-}_t, resulting in:

$$
\begin{aligned}
&(\mathrm{SP}_{\mathrm{UCP}})(x^k_{gt}, f^k_{gt}, l^k_{git}, y^k_{gt}, \tilde{y}^k_{gt}, \hat{y}^k_{gt}, w^{k+}_t, w^{k-}_t):\\
&\quad \min \sum_{g\in\mathcal{C}\setminus\mathcal{Q}} \sum_{t\in\mathcal{T}} p^f_g F_{gt} + \sum_{g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}} p^{om}_g X_{gt} + \sum_{t\in\mathcal{T}} \hat{p}\hat{D}_t\\
&\quad \text{subject to}\\
&\qquad \text{constraints in §2.2.2}: (2.1a), (2.1d), (2.1e) \text{ for } g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}; (2.1b), (2.1c),\\
&\qquad\qquad (2.1g)-(2.1l) \text{ for } g\in\mathcal{C}, t\in\mathcal{T}; (2.1f) \text{ for } g\in\mathcal{R}, t\in\mathcal{T}\\
&\qquad \text{constraints in §2.2.2 with variables fixed to } x^k_{gt}, \tilde{y}^k_{gt}:\\
&\qquad\qquad (2.3a),\ (2.3b),\ (2.3c) \text{ for } g\in\mathcal{C}, t\in\mathcal{T};\ (2.3d) \text{ for } g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}\\
&\qquad \text{constraints in §2.2.2 with variables fixed to } w^{k+}_t, w^{k-}_t: (2.5a)-(2.5d) \text{ for } t\in\mathcal{T}\\
&\qquad \text{constraints in §2.2.2}: (2.6a)-(2.6b) \text{ for } g\in\mathcal{C}\setminus\mathcal{Q}, t\in\mathcal{T}\\
&\qquad \text{constraints in §2.2.2 with variables fixed to } \tilde{y}^k_{gt}: (2.7) \text{ for } g\in\mathcal{C}\setminus\mathcal{Q}, i\in\mathcal{I}, t\in\mathcal{T}\\
&\qquad X_{g't}, X^{sp}_{gt}, X^{sr}_{gt}, X^{st}_t, F_{\hat{g}t}, L_{\hat{g}it}, \hat{D}_t, S_t, S^+_t, S^-_t, S^{s-}_t \geq 0\\
&\qquad\quad \forall\, g\in\mathcal{C},\ \hat{g}\in\mathcal{C}\setminus\mathcal{Q},\ g'\in\{\mathcal{C}\setminus\mathcal{Q}\}\cup\mathcal{R},\ t\in\mathcal{T},\ i\in\mathcal{I}
\end{aligned}
$$

Note that constraints in the monolith are not included in the Benders' subproblem because they contain only binary variables and/or those continuous variables relegated to the master problem and, as such, are already satisfied by the solution of the Benders' master problem.

Given ($\mathrm{MP}_{\mathrm{UCP}}$) and ($\mathrm{SP}_{\mathrm{UCP}}$), we initially select values for our binary variables and those variables associated with generation from the units with concave emissions from solving the underestimator problem (see §2.3.3). We obtain values for the continuous variables (except those associated with generation from the units with concave emissions) and collect the dual variable values (or, in the case of an infeasible subproblem, the unbounded rays) from the dual subproblem. We use this information to construct an optimality cut or feasibility cuts, which we add to the master problem. We then resolve the master problem to obtain new values for the binary variables, which we pass to the subproblem as parameters. We repeat the process until the objective function values from the two problems converge. The Benders' Decomposition procedure for ($\mathrm{P_{UCP}}$) follows:

---

1. **Initialize**: Let $k = 0$; initialize the sets $\mathcal{L}$ and $\mathcal{M}$. GoTo 2.

2. **Solve** ($\mathrm{MP_{UCP}}$): Obtain the solution, $\{x^k_{g't},\ f_{g't},\ l_{g'it},\ y^k_{gt},\ \tilde{y}^k_{gt},\ \hat{y}^k_{gt},\ w^{k+}_t,\ w^{k-}_t\}$, $g' \in \mathcal{Q},\ g \in \mathcal{G},\ t \in \mathcal{T},\ i \in \mathcal{I}$ for the master problem variables, and real number η^k, i.e., the value of the cut at the k^{th} iteration. GoTo 3.

3. **Check for optimality if the subproblem in the previous iteration was feasible**: If

$$
\begin{aligned}
\eta^k \geq\ & \sum_{g \in \mathcal{C}, t \in \mathcal{T}} p^f_g F_{gt} + \sum_{g \in \mathcal{C}, t \in \mathcal{T}} p^{om}_g X_{gt} + \sum_{t \in \mathcal{T}} \hat{p} \hat{D}_t \\
& + \sum_{g \in \mathcal{C}, t \in \mathcal{T}} p^{nox} \left(a^q_g X^2_{gt} + a^l_g X_{gt} + a^c_g \tilde{Y}_{gt} \right) \\
& + \sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} p^{\bar{s}}_g Y_{gt} + \sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} p^{\underline{s}}_g \hat{Y}_{gt} - \epsilon;
\end{aligned}
$$

   ϵ a small tolerance, then stop; otherwise GoTo 4.

4. **Solve** ($\mathrm{SP_{UCP}}$)$(x^k_{g't}, f_{g't}, l_{g'it}, y^k_{gt}, \tilde{y}^k_{gt}, \hat{y}^k_{gt}, w^{k+}_t, w^{k-}_t)$ $(g' \in \mathcal{Q},\ g \in \mathcal{G},\ t \in \mathcal{T},\ i \in \mathcal{I})$ and derive cuts. There are two cases:

   (a) The subproblem is feasible (and bounded - it is bounded if the monolith is bounded). Obtain the dual multipliers from the subproblem at the k^{th} iteration from the constraints on the primal subproblem. Add an *objective function cut*, i.e., a constraint of the form (2.22a)-(2.22d), to ($\mathrm{MP_{UCP}}$) by increasing set $\mathcal{L}$ correspondingly. GoTo 5.

(b) The subproblem is (primal) infeasible, i.e., dual unbounded. Take the extreme ray from the dual subproblem. Add a *feasibility cut*, i.e., a constraint of the form (2.23), to (MP_{UCP}) by increasing set $\mathcal{M}$ correspondingly.

5. **Increase iteration counter** k: $k \leftarrow k + 1$. GoTo 2.

2.3.6 Benders' Cuts

We employ four types of cuts in some form of the Benders' master problem:

- **In the MINLP Master Problem**: The objective function value of (MP_{UCP}) must be greater than or equal to the minimum of the current lower bound from the mixed integer nonlinear program, typically obtained from the underestimators.

- **In the RMINLP Master Problem**: The objective function value of (MP_{UCP}) must be greater than or equal to the product of a scale factor (typically, 75%), and the minimum of the current lower bound from the relaxed mixed integer nonlinear program and that obtained from the underestimators.

  We apply this 75% reduction to the right-hand-side because we are using MINOS, a local solver, to solve the relaxed mixed integer nonlinear programs. Often, the local solver cannot find a good local minimum. But because the termination criterion for the relaxed solves before starting the mixed integer nonlinear programming solves is that the gap between the most recent master problem objective function value and subproblem value are close, were we to supply the master problem with a good lower bound in the first iteration, no Benders' iterations would occur. This might appear to be an advantage in the short run, but, in the long run, we are unable to exploit solving the relaxed problems to eliminate some of the solution space. By reducing the right hand side by an arbitrary factor, we generally force Benders' iterations in the relaxed problem, thereby reducing the solution space with these easier solves.

- **In the RMINLP and MINLP Master Problems**: The objective function value of (MP_{UCP}) must be less than or equal to the tightest upper bound obtained from the subproblem.

- **In the RMINLP and MINLP Master Problems**: We use CPLEX's solution pool to initially collect multiple underestimator problem solutions. We take each master problem solution, fix it in the subproblem, and obtain a set of duals so that we can add as many optimality cuts to the problem as underestimator problem solutions we generate.

2.3.7 Summary: A Hybrid Benders' Decomposition Algorithm

We implement our algorithm using the techniques in §2.3.1-§2.3.6 as follows:

---

1. **Generate the Model:**

   Strengthen the original model formulation (see §2.2.1) with the cuts established in §2.3.1 and §2.3.2. GoTo 2.

2. **Preprocessing:**

   (1) Use the preprocessing steps as described in §2.3.4 to obtain an initial lower bound (LB) and an initial upper bound (UB) on the optimal objective function value of the monolith. Collect ϑ (e.g., 10) solutions from the underestimator model using the CPLEX solution pool within a pool gap of ψ (e.g., 10^{-5}). Obtain ϑ Benders' optimality cuts by solving ϑ LP subproblems.

(2) If $(UB - LB)/LB \leq \varepsilon$ (e.g., 0.001), then stop; otherwise GoTo 3.

3. **Benders' Applied to the Relaxed Master Problem**:

(1) Solve the RMINLP master problem using a local solver (e.g., Minos) to obtain an objective function value z_{RMINLP}.

(2) Generate a feasibility cut or an optimality cut by solving the Benders' subproblem (an LP) using the solution obtained in (1) yielding objective function value z^*_{SUB}.

(3) If $(z^*_{\text{SUB}} - z_{\text{RMINLP}})/z_{\text{RMINLP}} \leq \varepsilon$, then GoTo 4; otherwise Return To 3.

4. **Benders' Applied to the Master Problem**:

(1) Solve the MINLP master problem using a local solver (e.g., AlphaECP).

(2) Pass the solution obtained in (1) to a global solver (e.g., BARON), and solve the MINLP master problem to a gap of ω (e.g., 0.0005). Let the lower bound of the MINLP master problem be z_{MINLP}.

(3) If z_{MINLP} is larger then current lower bound, then update LB.

(4) Generate a feasibility or an optimality cut by solving the Benders' subproblem (an LP) using the solution obtained in (2).

(5) If z^*_{SUB} is lower than the current upper bound, then update UB.

(6) If $(UB - LB)/LB \leq \varepsilon$, then stop; otherwise Return To 4.

2.4 Computational Results: Solution Time Comparisons and Test System Results

We model our hybrid algorithm in GAMS (Brooke *et al.* (1988), Bussieck and Meeraus (2004)), Version 23.7.3, and employ the following solvers: Minos, Version 5.51 (Murtagh *et al.* (2011)), BARON 9.3.1 (Tawarmalani and Sahinidis (2004)), CPLEX 12.3 (IBM (2011)), and AlphaECP 2.04.02 (Westerlund and Lastusilta (2011)). We run our computational tests on a Sun X5675 server with Intel Xeon processors and 48 GB RAM.

The National Renewable Energy Laboratory (NREL) provided the generator and load data for the Colorado power system from the GridView database, and also provided all available continuous monitoring emission system (CEMS) data for the Colorado generators (US Environmental Protection Agency, 2005). We include a subset of these generators in operation during the year 2005 in our analysis. Specifically, we include 15 generators of which 14 are gas (two of these comprise a combined cycle unit), and one is a coal generator. The minimum capacity for a single power unit is 2.6 MW and the maximum capacity for a single unit is 178.3 MW. There is no minimum capacity for the renewable generators; the winter, spring, summer, and fall peak levels for a single renewable generator are 23.5 MW, 23.7 MW, 11.0 MW, and 19.7 MW, respectively. The total installed capacity for the gas units is 522.5 MW, while that of the coal unit is 178.3 MW. The available hourly renewable generation ranges between 43 MW and 81.6 MW. We scale several parameters, including demand, available renewable capacity, and the size and flow rates of the storage device to coordinate with our subset of generators. We assume the storage device can hold 50 MWh of energy and has a capacity rating of 12 MWh. We modify certain parameters which NREL deemed to be unreasonable, e.g., minimum down times longer than 48 hours are reduced to 24 hours. We also use the CEMS data to develop our own piecewise-linear heat rate curves and nonlinear NO_X emissions curves.

For the combined cycle units in the system, we identify the gas turbines and the steam turbines. In order to model these units accurately, we include each gas turbine in the data set as well as a combined gas and steam turbine. To develop the combined gas and steam turbine's operating characteristics, we allocate the steam turbine's operating characteristics and costs across the gas turbines based on each gas turbine's relative maximum capacities. For example, assuming a combined cycle unit with three gas turbines, each with a maximum capacity of 150 MW, and a steam turbine with a maximum capacity of 300 MW, we increase each gas turbine's maximum capacity by 100 MW to 250 MW. For the combined gas and steam units, we use the steam turbine's maximum ramp-up, minimum up and ramp-down times, which are the more conservative of the two.

We differentiate instances by altering the time horizon and season; specifically, we solve 12-, 18-, and 24-hour models for each of four seasons (winter, spring, summer, and fall). Our 15-generator, 12-hour problem instances contain 3,826 constraints, 2,185 variables (including 804 binaries), and 16,468 nonzeros in the matrix of left-hand-side coefficients. The corresponding 18-hour problem instances contain 5,757 constraints, 3,277 variables, of which 1,206 are binary, and 28,207 nonzeros. The 24-hour instances possess 7,690 constraints, 4,369 variables, including 1,608 binaries, and 41,822 nonzeros.

We run computational tests with two goals: (i) we demonstrate the strength of our solution techniques described in §2.3, and (ii) we show how our model provides insight into the operation of a real-world power system. Specifically, regarding (i), we show how each of the procedures, as appropriate, outlined in §2.3 affects convergence and solution time of our Benders' procedure relative to solving the monolith. Then, regarding (ii), we measure the extent to which the storage device is used and the contribution of renewables to the system. We assess the emissions differences when measured with constant, versus quadratic, curves.

2.4.1 Solution Time Comparisons

While the cuts described in §2.3.1 and §2.3.2 do not yield tighter lower bounds, they greatly reduce solution time for the Benders' procedure. For example, the 12-hour winter season instance, with the cuts, requires 172 iterations to converge; without the cuts, the same instance requires 189 iterations for convergence. The 12-hour fall season instance requires 177 Benders' iterations (including 137 relaxed MINLP solves) with the cuts; without the cuts, the procedure runs for more than 3,400 iterations without improving the objective for more than 2,200 iterations and without ever commencing MINLP solves.

In the 12-hour winter season model, we generate six feasibility cuts. Similarly, for our 12-hour fall season model, only 23 cuts are feasibility cuts. The incorporation of an elasticized demand constraint and valid inequalities significantly reduces the likelihood of generating solutions from the Benders' master problem which result in infeasible subproblems. If we chose not to include these features in our model, we would consider implementing an approach such as the maximum feasible subproblem, presented in Saharidis and Ierapetritou (2010), to try to increase the chance of generating solutions which are feasible for the Benders' subproblem.

We solve each of twelve instances using the formulation given in §2.2 with the cuts described in the first two subsections of §2.3. We employ the one- and two-segment underestimators separately, and compare the root relaxation value against that of the mixed integer nonlinear program (see Table 2.1). We provide the monolith the same benefit with respect to cuts, with the exception of fixed generation (see §2.3.4) which requires bounds to determine which generators should be fixed on or off.

The percent difference between the root relaxation of the mixed integer nonlinear program with the nonlinear curves approximated with one- and two-segment underestimators is about the same and ranges between 0% and 3% relative to the optimal solution. That is, these underestimators yield tight lower bounds. On the other hand, the percent difference between the root relaxation of the mixed integer nonlinear program as given in (P_{UCP}) and the optimal solution ranges between 25% and over 100%. In other words, these nonlinear programming bounds are substantially weaker than those obtained from their piecewise linear counterparts.

Table 2.1. Numerical results for the tightened formulation: Comparison of the monolith and the one- and two-segment Benders' approximations root relaxations

	Percent difference between root node and optimal solution		
	Two-segment underestimator	**One-segment underestimator**	**Monolith**
Winter-12	3.16	2.28	104.89
Spring-12	0.69	0.59	53.18
Summer-12	1.52	1.39	24.98
Fall-12	1.60	0.40	48.29
Winter-18	2.63	2.04	87.28
Spring-18	2.12	2.04	56.55
Summer-18	1.55	1.43	36.69
Fall-18	2.08	1.07	55.26
Winter-24	2.13	1.28	50.44
Spring-24	1.55	1.46	52.29
Summer-24	1.94	1.94	27.86
Fall-24	1.47	1.47	43.63

The Benders' Decomposition procedure (see Table 2.2) follows the algorithm described in §2.3.7. To expedite solutions for 24-hour time horizon model instances, we implement cut aging, i.e., we do not consider Benders' master problem cuts that have not been active for 120 iterations. We initially solve relaxed mixed integer nonlinear programming problems (RMINLPs), i.e., nonlinear programming problems, corresponding to the master problem until the objectives from the RMINLP master problem and the subproblem are within 0.1% for the 12-hour models or within 0.15% for the 18- and 24-hour models. At that point, we solve mixed integer nonlinear programming problems representing the master problem for the remainder of the run. Within each iteration, we solve the mixed integer nonlinear programming problems to within 0.05%. For each MINLP master problem, we provide an initial solution, if any is found within 500 seconds, from the local solver AlphaECP.

The piecewise-linear methodologies perform well relative to the implementation of the mixed integer nonlinear monolith, both in terms of solution quality, if applicable, and in terms of solution time. In particular, the two-segment Benders' method significantly outperforms implementation of the monolith, and is also substantially superior to the one-segment Benders' method in seven of the twelve instances. In the other five instances, the solutions lie close to the end points of the NO_x curves, allowing a one-segment underestimator to adequately approximate the curve. However, the performance gains from fewer binaries in the one-segment Benders' model versus the two-segment model are small enough that the two-segment Benders would still be a safe implementation choice.

Table 2.2. Numerical results for Benders: Comparison of the monolith solution time and quality against that of using the one- and two-segment Benders' approximations. We state the gap relative to the optimal objective function value, which the two-segment Benders' procedure provides.

Horizon length [h] and season	**Two-seg. Benders**		**One-seg. Benders**		**Monolith**	
	Obj.	**Solution time**	**Gap%**	**Solution time**	**Gap%**	**Solution time**
Winter-12	189,355	7	0.09	12,505	0	761
Spring-12	256,971	7	0.02	8	0	482
Summer-12	418,142	13	0.01	10	0	1235
Fall-12	265,659	7	0.10	1,447	0	121
Winter-18	311,546	42	0.34	♭	14.33	♭
Spring-18	397,084	79	0.15	♭	7.55	♭
Summer-18	497,987	37	0.17	♭	2.61	20,933♯
Fall-18	392,357	17	0.09	14	9.78	♭
Winter-24	486,413	635	0.38	♭	15.71	♭
Spring-24	482,222	183	0.19	♭	8.92	♭
Summer-24	632,311	107	0.02	63	7.33	♭
Fall-24	513,736	85	0.14	33	N/A	♭

♭: Timed out after 10 hours
♯: Out of memory
Solution times in seconds

2.4.2 Test System Results

We show in Figure 2.3 the extent to which the storage device is used over a 24-hour time horizon and each season for which we obtain solutions. The storage device is used extensively in the representative fall day and to a much lesser extent in the corresponding spring and summer days. In all cases but the winter day, storage capacity falls to zero during some hour(s) in the middle of the day.

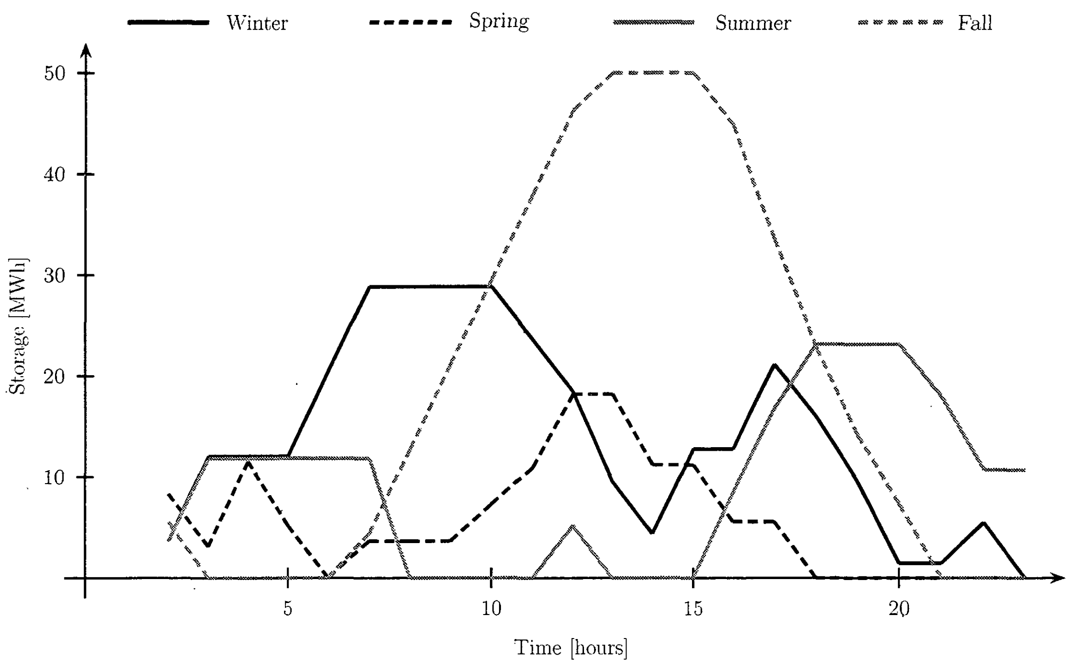

Figure 2.3. Storage values over a 24-hour time period for each of the four seasons. Note that the storage device is not usable until time period 2.

We present in Figure 2.4 the renewable energy scheduled for the a 24-hour time horizon and the four seasons. We note that no curtailment of renewable energy occurs in these four windows, and even during the spring peak, all renewable energy was consumed.

We compare the NO_x emissions predicted using the concave quadratic curves, i.e., for the generators in set $\mathcal{Q}$, to constant emissions rates. Using the four 24-hour horizon models, we find that a constant NO_x emissions rate (rather than a quadratic

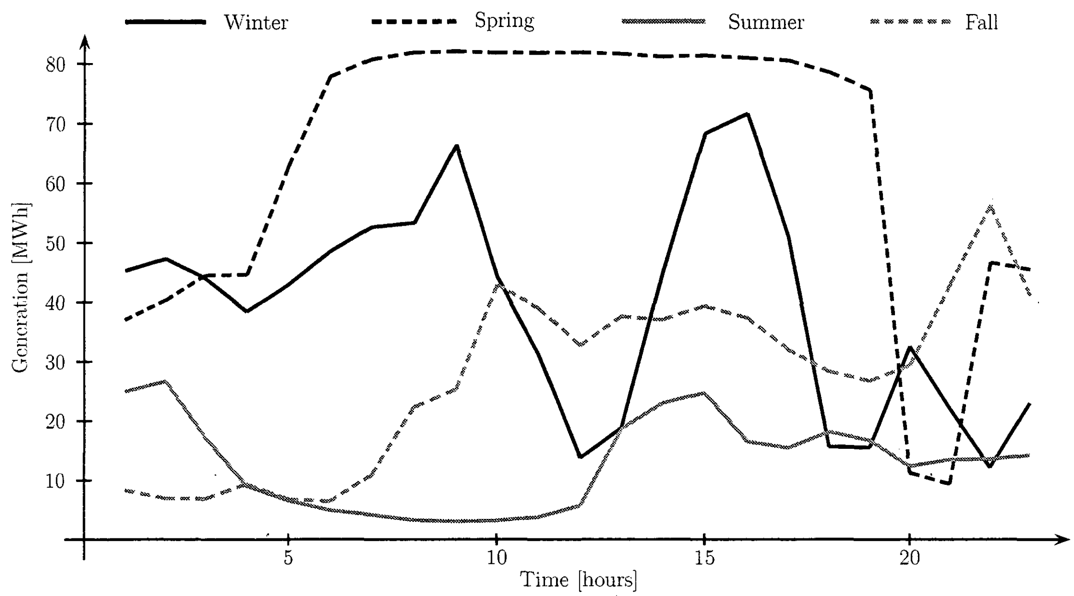

Figure 2.4. Renewable energy used over a 24-hour horizon for each of the four seasons.

curve) results in an average overestimation in emissions on the order of 1.9%. For all four models considered, the discrepancy in emissions between the constant rate and quadratic approximations ranges from -0.7% to 4.7%.

2.5 Conclusion

We present a mixed-integer nonlinear program that minimizes fixed and variable costs of operation, emissions costs, and penalty costs for failure to meet demand. We incorporate not only standard constraints regarding the operation of the generators, but also details such as renewable generators, storage devices, piecewise linear heat input curves, and nonlinear emissions curves. Because the model instances we test exhaust the capabilities of state-of-the-art mixed integer nonlinear solvers, we implement several enhancements to increase model tractability and computationally demonstrate their usefulness.

Our enhancements allow us to quickly solve otherwise intractable mixed-integer nonlinear programs to near optimality, or in many cases, to optimality. Our results demonstrate that the hybrid algorithm with two-segment piecewise linear underestimators converges in seconds compared to some monolithic models which do not achieve the required gap after 10 hours. Thus, this hybrid Benders' Decomposition algorithm offers the opportunity for RTOs and ISOs to introduce nonlinear constructs into their unit commitment models while maintaining tractability.

CHAPTER 3

THE COST OF PROVIDING SPINNING RESERVES

Utilities must generate enough electricity to meet customer demand. This presents a challenge, as utilities must anticipate quickly changing consumer needs, while ensuring that power plants operate within specifications. The unit commitment model provides utilities with a lowest cost schedule to meet their anticipated customer demand without violating the operational limits of their power plants. The typical unit commitment model also ensures adequate spinning, or contingency, reserves are always available in case of an unexpected generator failure. Spinning reserves (SR) are present as excess online, or spinning, capacity which allows a power system to respond to uncertain failures almost instantaneously. Local regulatory agencies, such as an Independent System Operator (ISO), usually establish the required level of SR. In the absence of SR, some customers would experience curtailment of power for a short period of time.

The unit commitment model typically contains three distinct single-time-index binary variables which indicate whether a generator starts up, shuts down or operates in steady state in a given time period. Furthermore, the model also incorporates continuous variables which describe how the generators in steady-state operation meet demand and spinning reserve requirements. We provide a unit commitment model containing these features, as well as renewable energy sources and a storage device. Specifically, we include wind turbines and a pumped hydro storage device in our system. Our storage device not only acts as a repository for excess electricity, but it also has the capability of providing spinning reserves. As with any storage device, significant efficiency losses occur. The storage device has capacity and flow rate limits which must be observed. The addition of renewables and storage in our unit commitment model results in a more accurate representation of Xcel's power system in Colorado.

Our research contributions lie in the following: (i) we formulate a unit commitment model which includes features such as storage, renewable sources, and piecewise-linear convex or concave heat input curves, in addition to other details included in published models; (ii) we implement a *sliding time window heuristic* to obtain feasible generator dispatch schedules of reasonable quality for up to five different SR requirements for an annual planning horizon; (iii) we evaluate the short-term (monthly) cost savings achieved by reducing the required level of spinning reserves for a system with storage; (iv) we estimate annual cost savings which might be achieved by reducing SR; and (v) we present expected emissions of CO_2, NO_x and SO_2 for these annual schedules, which are determined a posteriori.

The remainder of the chapter is organized as follows: In Section 3.1, we provide a literature review. In Section 3.2, we present our version of the unit commitment model. In Section 3.3, we describe opportunities to tighten the formulation and introduce our *sliding time window heuristic.* In Section 3.4, we contrast operating costs with and without the spinning reserve requirements enforced. Section 3.5 concludes the chapter.

3.1 Literature Review

In this section, we provide a review of relevant literature. We first discuss typical unit commitment models used today, including their size and techniques for solving such problems. Additionally, we discuss how storage devices have been used in other published models. We then review heuristic methods for solving unit commitment models for longer horizons, e.g., a year. Finally, we summarize research into demand response and the costs associated with providing SR.

3.1.1 Typical Unit Commitment Model

The first attempts to solve mixed integer linear programming (MILP)-based deterministic unit commitment problems date back to the late 1970s and consider fewer than 20 generators dispatched over a 24 hour horizon (Dillon *et al.*, 1978). Baldick (1995) presents a unit commitment problem with typical generator constraints, e.g., capacity, ramp rates, and up- and down-time constraints, as well as line flow and voltage constraints. In order to solve a small problem with 10 generators over a 24 hour horizon, the author relies upon Lagrangian Relaxation. A comparison of using Lagrangian Relaxation to solving the monolith outright with branch-and-bound appears in Li and Shahidehpour (2005).

Since these early attempts to solve the unit commitment problem, the model has been applied to many different research questions. Sioshansi *et al.* (2008) compare welfare transfers between consumers and producers based on solutions obtained using Lagrangian Relaxation versus solving the monolith directly. The authors study the ISO New England, which consists of 276 generators. The authors incorporate a time horizon of 24 hours, constant marginal costs, and omit transmission in order to develop a tractable MILP. The results of this practical application dwarf most published results, which focus on test systems.

In practice, branch-and-bound solution techniques applied to MILP unit commitment models are still considered a "new" methodology. O'Neill and Krall (2011) survey six Regional Transmission Organizations (RTOs) and/or ISOs regarding their use of day-ahead unit commitment models. Based on the responses, the majority of the organizations uses or is currently evaluating MILP-based models. Of the organizations that responded, the look-ahead time frame may be as long as four hours with quarter-hour model fidelity. The only organization to report an optimization gap for their unit commitment model optimization, the California ISO (CAISO), recently increased the quality of their solution from 0.5% to 0.05%. However, it is unclear from the report if these organizations develop their own unit commitment models or purchase commercial software from vendors such as Ventyx (Ventyx, 2011), Power Optimisation (Power Optimisation, 2011), Plexos (Energy Exemplar, 2011), and Nexant (Nexant, 2011). These software packages offer a variety of features including generator operating constraints, pumped storage device constraints, and transmission (Power Optimisation, 2011). Some software also considers more sophisticated features such as buying and selling power to and from the grid (Nexant, 2011) and decision analysis tools (Ventyx, 2011).

3.1.2 Incorporation of Storage Devices

Several authors have incorporated aspects of renewable generation and/or storage devices into their unit commitment models to assess the economic implications of these devices, including Gollmer *et al.* (2000). Sioshansi *et al.* (2009) evaluate the ability of storage devices to decrease arbitrage opportunities in electricity markets from 2002 to 2007 by smoothing, or shifting, load. The model consists of a linear program (LP) which is solved in successive two-week intervals. The model characterizes a generic storage device, constraining its operation based on its efficiency, storage capacity and ability to generate power.

Additionally, Denholm and Sioshansi (Forthcoming) propose two MILPs to measure the benefits and costs of co-locating wind farms and compressed air energy storage (CAES) devices. The authors consider the transmission infrastructure required to transport the energy from a wind farm in a rural area to an urban center. See Sioshansi (2010b) for a comprehensive summary of relevant literature regarding storage devices and their economic implications on power systems.

3.1.3 Heuristics for Creating Long-Term Generation Schedules

To better understand the impact of storage and renewable generation on power systems, heuristics are often employed to provide an estimate of the benefits of these technologies. Tuohy *et al.* (2007) determine how operating costs change as reserve requirements fall due to the benefit of better wind forecasts. The authors compare the results they obtain from their own unit commitment and generator dispatch model to an industry standard "rounded relaxation" heuristic solution obtained using the Plexos software (Energy Exemplar, 2011). The authors use a rolling horizon method, which varies from one hour to 24 hours, to quantify these operating costs.

Sioshansi and Short (2009) model the Electricity Reliability Council of Texas (ERCOT) system with high wind penetration using a combination of two models: a day-ahead unit commitment model and a real-time dispatch model. The day-ahead model, with hourly fidelity, is solved repeatedly to develop a commitment schedule for 165 days. The real-time dispatch model then uses 15-minute time steps to dispatch the selected generators to meet demand. Sioshansi (2010a) models the ERCOT system, including wind generation and a storage device, using a more sophisticated, 48-hour unit commitment model. Both generator commitment and dispatch are determined for the two-day horizon. The generators' states at the end of the horizon are passed to the next 48-hour unit commitment model to ensure a feasible schedule. An hourly dispatch model then determines the generation level for each unit. Given that the system consists of 375 generators, commitment decisions are made in three-hour increments to improve model tractability.

Sioshansi and Miller (Forthcoming) construct a unit commitment model which incorporates the use of plug-in hybrid electric vehicles. The authors solve the model repeatedly, overlapping 48-hour horizons in order to assess the impacts of plug-in hybrid electric vehicles on emissions from power generation in Texas. The window has a 24-hour overlap; thus, the generators' states at the end of the first 24 hours become fixed for the next window, and the process repeats for one year. This model focuses on power generators and plug-in hybrid electric vehicles only. It does not consider renewable energy sources or storage devices. Sioshansi *et al.* (2010) perform a similar analysis on the impact of plug-in hybrid electric vehicles for the Ohio region.

Weber *et al.* (2009) present the National Renewable Energy Laboratory's (NREL) Wind Integration in Liberalized Markets (WILMAR) model, an MILP. The model also contains the WILMAR Scenario Tree Tool to incorporate uncertainty in wind generation and load. The WILMAR code possesses a rolling, or sliding, time horizon to allow forecasts of wind and load to be updated as better information becomes available. The initial three hours of the horizon are treated deterministically; time horizons for stochastic optimization vary between 12 and 36 hours. Tuohy *et al.* (2008) and Tuohy *et al.* (2009) evaluate the WILMAR stochastic unit commitment problem after introducing uncertain wind generation into the system. An MILP version of the WILMAR model provides the schedule and economic dispatch of the generators. The three-stage WILMAR Scenario Tree Tool characterizes the stochastic wind and load parameters. The authors solve the problem repeatedly, moving forward in time in three-hour increments until they obtain a schedule for the entire year. In Tuohy and O'Malley (2009), the authors use the WILMAR model with the Scenario Tree Tool to determine the benefits of pumped storage devices on systems with high levels of wind generation. A sliding time window heuristic, rolling forward every three hours, determines the commitment and dispatch schedule, taking into consideration the availability of the pumped storage device.

3.1.4 Demand Response and Spinning Reserves

The unit commitment model typically provides information about spinning reserves as well as generator dispatch. Recent emphasis on incorporating demand response into the power system increases the complexity of the traditional unit commitment model. Each industry has its own definition of "demand response." In economics, demand response may refer to how consumers alter their behavior based on electricity rates, such as Californians did in response to the 2000-2001 energy crisis (Reiss and White, 2003). In the context of our work, demand response refers to load shedding, particularly of customers who have expressed a willingness to have their demand curtailed for short periods of time.

Martins *et al.* (1996) develop a multiple-objective linear program to minimize a generator expansion budget, environmental costs and environmental impact from SO_2 and NO_x emissions while meeting load over a multi-year horizon. The authors consider load shedding as an alternative to SR when developing the capital budget. Because such considerations are the primary focus of their work, the annual time horizon of the budget optimization model differs from a typical hourly dispatch model. Luo *et al.* (1998) also develop an MILP to minimize the net cost of shedding load from the perspective of a large industrial customer. This load-side strategy allows the customer to reduce or increase its power consumption using a "control event" algorithm. The authors also consider the customer's production targets when determining load shedding and restoration.

More recently, Bouffard *et al.* (2005a) and Bouffard *et al.* (2005b) expand the electrical engineering definition of demand response in evaluating power systems by reducing operating reserves. Reduction of operating reserves could result in load shedding in the event of a generator or line failure. The authors also quantify the amount of excess reserves within a power system using a stochastic model of a simple three-generator, three-line and three-bus system operating over a four-hour time horizon. Binary variables with a single time index represent whether the generator operates in a given time period. The model maximizes social welfare in light of possible failures of both generators and lines. While load shedding results in a penalty to the objective function, the model does not explicitly limit total shed load. The authors also exclude non-spinning reserves, i.e., reserves offered by non-operational generators, owing to intractability of larger-scale models.

Ortega-Vazquez *et al.* (2006) present a stochastic unit commitment model which provides a time-varying optimal level of SR, noting that traditional models generally overcompensate for potential outages. The authors incorporate the cost of shedding load into the objective function and maximize social welfare. Additionally, this research incorporates a piecewise-linear approximation of expected energy not supplied (EENS). The piecewise approximation has limitations, particularly due to convexity of the EENS frontier. However, since the approximation tends to overestimate the EENS, the unit commitment model solution may provide more SR than necessary.

Madaeni and Sioshansi (Forthcoming) compare cost savings of two models: a stochastic unit commitment model incorporating a three-stage wind availability scenario tree and a unit commitment model with demand response and real-time pricing. The models predict cost savings for ERCOT, which has significant wind generation. The results indicate that a unit commitment model with real-time pricing offers the potential for greater cost savings than a stochastic unit commitment model considering uncertain wind demand. If implemented together, the total costs savings from both real-time pricing and stochastic wind forecasts are less than the additive savings from the two independent features.

3.2 Model

The unit commitment model we present minimizes total operating costs for a power system including: fixed start-up costs, fuel and maintenance costs, and the costs of not meeting demand. Our system includes "base load" coal units as well as "intermediate" combined cycle units and "peaker" natural gas units. We assume that utilities dispatch these generators on an hourly basis throughout the year. Because we consider a system with a high penetration of coal generation, we anticipate a solution in which the base load units operate steadily throughout the year, with intermediate and peaker units providing supplemental generation during high demand hours.

Our model contains typical constraints found in unit commitment models, both for the power system as a whole and for the individual generators. Load must be balanced, i.e., we must meet demand, and we have to provide adequate spinning reserves, which must be equal to the maximum capacity of the largest generator operating at a given time. For the generators, we limit their operation based on ramp rate, minimum and maximum capacity, and minimum up and down times. Unit commitment models typically use a linear approximation to describe the relationship between fuel use and generation level. However, some generators operate more efficiently at levels below their maximum capacity, resulting in a nonlinear relationship between generation and fuel consumption (Gollmer *et al.*, 2000). Therefore, to include more accurate representations of these conventional characteristics, we use piecewise-linear approximations of each generator's heat input curve. Each conventional generator possesses a unique heat input curve, which we approximate with between one and four linear intervals depending on the convexity or concavity of the curve.

We specify a slope for each segment for all of the piecewise-linear approximations, regardless of convexity. For the convex approximations, we also define the length, or amount of generation, modeled in each interval, as well as the first segment's intercept. The concave piecewise-linear approximations are more complex, because we must specify the intercept for each segment, define the lower and upper bounds on each interval, and ensure that a generator operates in at most one segment at a time.

We embellish the typical unit commitment model by including generation from wind energy sources and energy storage using a pumped hydro device. Renewable generators operate very differently than conventional generators because their maximum output in a given time period varies as a function of the amount of wind available at any time. Due to this variability, renewable generators are unsuitable for contributing to spinning reserves. We include a pumped hydro storage device which is capable of supplying power when needed, typically during peak periods, and storing power when there is an excess due to low demand. The storage device may also provide spinning reserves, if necessary. Limitations on the device's operation include: minimum and maximum capacity of the device, limitations on the flow into and out of the device, as well as inventory balance. All power stored in the pumped hydro device is subject to efficiency losses.

Although we consider many aspects related to generator commitment and dispatch, our model cannot include all features which impact the generation schedule. We do not consider transmission. Thus, our solution could concentrate generation in a region without adequate transmission capacity, making the generation schedule infeasible upon implementation. We also do not incorporate the impacts of the SO_2 cap-and-trade system in place under the US Environmental Protection Agency's Acid Rain Program, which may make coal generators appear more cost effective in our analysis. We also assume constant coal and natural gas prices throughout the year. Finally, we assume that load and the maximum available renewable generation are known with certainty because we model our system deterministically.

3.2.1 Generator Dispatch Optimization Model Formulation

We now present our model, ($\mathrm{P_{UCP}}$). We use lower case Roman letters to represent parameters, upper case Roman letters in regular font to represent variables, and upper case Roman letters in script font to denote sets. Superscripts provide descriptions for various parameters while subscripts on parameters and variables represent indices, i.e., elements of a set. A mixed-integer nonlinear variant of this model appears in Van Dinter *et al.* (2011b).

Sets:

- $\mathcal{G}$: set of generators
- $\mathcal{R} \subseteq \mathcal{G}$: set of renewable generators
- $\mathcal{C} \subseteq \mathcal{G}$: set of conventional generators
- $\hat{\mathcal{C}} \subseteq \mathcal{G}$: set of conventional generators with concave piecewise-linear heat input curves

- $\check{\mathcal{C}} \subseteq \mathcal{G}$: set of conventional generators with convex piecewise-linear heat input curves
- $\mathcal{I}$: set of intervals i for the piecewise-linear approximation of the heat input curves
- $\mathcal{T}$: time horizon over which the system is optimized

Data:

Prices

- p^f_g: price of fuel for conventional generator g [\$/MMBtu]
- $p^{\bar{s}}_g$: price of start up for conventional generator g [\$]
- p^{om}_g: operation and maintenance costs for conventional generator g [\$/MWh]
- $\hat{p}$: elastic penalty for not meeting demand [\$/MWh]

Supplies and Demands

- $\underline{c}_g$: minimum hourly generation of conventional generator g [MWh]
- $\bar{c}_g$: maximum hourly generation of conventional generator g [MWh]
- c^s_g: amount of hourly spinning generation of conventional generator g [MWh]
- $\bar{c}^r_{gt}$: maximum hourly generation of renewable generator g in hour t [MWh]
- d_t: demand at hour t [MWh]

Generator Rates and Times

- r^{r+}_g: ramp-up rate of generator g [MWh/hr]
- r^{r-}_g: ramp-down rate of generator g [MWh/hr]
- r^{t+}_g: ramp-up time required for generator g [hr]

- r_g^{t-}: ramp-down time required for generator g [hr]
- $\underline{t}_g^u$: minimum up time of generator g before shut down [hr]
- $\underline{t}_g^d$: minimum down time of generator g before start up [hr]

Storage Characteristics

- s^l: percentage efficiency loss from storage
- $\underline{s}^e$: minimum quantity of energy allowed in storage [MWh]
- $\bar{s}^e$: maximum quantity of energy allowed in storage [MWh]
- $\underline{s}^{f+}$: minimum flow into storage, i.e., pumping [MWh/hr]
- $\bar{s}^{f+}$: maximum flow into storage, i.e., pumping [MWh/hr]
- $\underline{s}^{f-}$: minimum flow from storage, i.e., generating [MWh/hr]
- $\bar{s}^{f-}$: maximum flow from storage, i.e., generating [MWh/hr]

Slopes and Intercepts for Convex Piecewise Approximation

- m_{gi}: slope of the linear approximation of the heat input curve for generator g in interval i [MMBtu/MWh]
- $\hat{b}_{gi}$: intercept of interval i of the concave linear approximation of the heat input curve for generator g [MMBtu]
- $\check{b}_g$: intercept of the first segment of the convex linear approximation of the heat input curve for generator g [MMBtu]
- $\hat{k}_{gi}$: length of interval i on the concave linear approximation of the heat input curve for generator g [MWh]
- $\check{k}_{gi}$: breakpoint of interval i on the convex linear approximation of the heat input curve for generator g [MWh]

Continuous Decision Variables:

Generation

- X_{gt}: electricity production from generator g at time t [MWh]
- X_{gt}^{sp}: spinning reserves provided by generator g at time t [MWh]
- X_{gt}^{sr}: spinning reserves required for each generator g at time t [MWh]
- X_{t}^{st}: total spinning reserves required at time t [MWh]
- F_{gt}: fuel used by generator g at time t [MMBtu]
- $\hat{D}_t$: unmet demand at time t [MWh]

Storage

- S_t: amount of energy in storage at the beginning of time t; $S_1 = 0$ [MWh]
- S_t^{+}: energy inflow to storage by the end of time t [MWh]
- S_t^{-}: energy outflow from storage by the end of time t [MWh]
- S_t^{s-}: spinning reserves from storage by the end of time t [MWh]

Piecewise-Linear Approximation

- L_{git}: hourly capacity at which generator g is operating on the heat input curve at time t in interval i [MWh]

Binary Decision Variables:

Generation

- $Y_{gt} = \begin{cases} 1 & \text{if generator } g \text{ starts up at time } t \\ 0 & \text{otherwise} \end{cases}$
- $\tilde{Y}_{gt} = \begin{cases} 1 & \text{if generator } g \text{ operates in steady state at time } t \\ 0 & \text{otherwise} \end{cases}$

- $\hat{Y}_{gt} = \begin{cases} 1 & \text{if generator } g \text{ shuts down at time } t \\ 0 & \text{otherwise} \end{cases}$

Storage

- $W_t^+ = \begin{cases} 1 & \text{if there is inflow into the storage device at time } t \\ 0 & \text{otherwise} \end{cases}$

- $W_t^- = \begin{cases} 1 & \text{if there is outflow from the storage device at time } t \\ 0 & \text{otherwise} \end{cases}$

Piecewise-Linear Approximation

- $Z_{git} = \begin{cases} 1 & \text{if generator } g \text{ operates in interval } i \text{ of the heat input curve at time } t \\ 0 & \text{otherwise} \end{cases}$

Formulation:

$$(\mathrm{P_{UCP}}) \quad \min \quad \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p_g^f F_{gt} + \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p_g^{om} X_{gt} + \sum_{t\in\mathcal{T}} \hat{p}\hat{D}_t + \sum_{g\in\mathcal{C}}\sum_{t\in\mathcal{T}} p_g^{\bar{s}} Y_{gt}$$

subject to

(see §3.2.2)

$$\sum_{g\in\mathcal{G}} X_{gt} + S_t^- - S_t^+ + \hat{D}_t = d_t \quad \forall\, t \in \mathcal{T} \tag{3.1a}$$

$$X_t^{st} \geq X_{gt}^{sr} \qquad \forall g \in \mathcal{C}, t \in \mathcal{T} \tag{3.1b}$$

$$\sum_{g\in\mathcal{C}} X_{gt}^{sp} + S_t^{s-} \geq X_t^{st} \qquad \forall t \in \mathcal{T} \tag{3.1c}$$

$$X_{gt} - X_{g,t-1} + X_{gt}^{sp} \leq r_g^{r+} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T},\ \ t > 1 \tag{3.1d}$$

$$X_{g,t-1} - X_{gt} \leq r_g^{r-} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T},\ \ t > 1 \tag{3.1e}$$

$$X_{gt} \leq \bar{c}_{gt}^r \quad \forall\, g \in \mathcal{R}, t \in \mathcal{T} \tag{3.1f}$$

$$S_t = S_{t-1} + (1 - s^l)S_{t-1}^+ - S_{t-1}^- \qquad \forall\, t \in \mathcal{T}, t > 1 \tag{3.1g}$$

$$S_t^{s-} \leq S_t^+ + \left(\bar{s}^{f-} W_t^- - S_t^-\right) \qquad \forall t \in \mathcal{T} \tag{3.1h}$$

$$S_t^- - S_t^+ + S_t^{s-} \leq S_t \qquad \forall t \in \mathcal{T} \tag{3.1i}$$

$$S_t^- + S_t^{s-} \leq \bar{s}^{f-} \qquad \forall\, t \in \mathcal{T} \tag{3.1j}$$

$$S_t \geq \underline{s}^e \qquad \forall\, t \in \mathcal{T} \tag{3.1k}$$

$$S_t \leq \bar{s}^e \qquad \forall\, t \in \mathcal{T} \tag{3.1l}$$

(see §3.2.2)

$$Y_{gt} \geq \tilde{Y}_{gt} - \tilde{Y}_{g,t-1} \quad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \tag{3.2a}$$

$$\hat{Y}_{gt} \geq \tilde{Y}_{g,t-1} - \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C}, t \in \mathcal{T},\ t > 1 \tag{3.2b}$$

$$\sum_{t'=t-\underline{t}_g^u - r_g^{t+} - r_g^{t-}}^{t} Y_{gt'} \leq \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C},\ t \in \mathcal{T},\ t > 1 \tag{3.2c}$$

$$\sum_{t'=t-\underline{t}_g^d}^{t} \hat{Y}_{gt'} \leq 1 - \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C},\ t \in \mathcal{T} \tag{3.2d}$$

(see §3.2.2)

$$X_{gt} + X_{gt}^{sp} \leq \bar{c}_g \tilde{Y}_{gt} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \tag{3.3a}$$

$$X_{gt}^{sr} = \bar{c}_g \tilde{Y}_{gt} \qquad \forall g \in \mathcal{C}, t \in \mathcal{T} \tag{3.3b}$$

$$X_{gt}^{sp} \leq c_g^s \tilde{Y}_{gt} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \tag{3.3c}$$

$$X_{gt} \geq \underline{c}_g \tilde{Y}_{gt} \qquad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \tag{3.3d}$$

(see §3.2.2)

$$W_t^+ + W_t^- \leq 1 \qquad \forall t \in \mathcal{T} \tag{3.4}$$

(see §3.2.2)

$$S_t^+ \leq \bar{s}^{f+} W_t^+ \qquad \forall\, t \in \mathcal{T} \tag{3.5a}$$

$$S_t^+ \geq \underline{s}^{f+} W_t^+ \qquad \forall\, t \in \mathcal{T} \tag{3.5b}$$

$$S_t^- \leq \bar{s}^{f-} W_t^- \qquad \forall\, t \in \mathcal{T} \tag{3.5c}$$

$$S_t^- \geq \underline{s}^{f-} W_t^- \qquad \forall\, t \in \mathcal{T} \tag{3.5d}$$

(see §3.2.2)

$$X_{gt} = \sum_{i \in I} L_{git} \qquad \forall\, g \in \check{\mathcal{C}},\ t \in \mathcal{T} \tag{3.6a}$$

$$L_{git} \leq \check{k}_{gi} \qquad \forall\, g \in \check{\mathcal{C}},\ i \in \mathcal{I},\ t \in \mathcal{T} \tag{3.6b}$$

(see §3.2.2)

$$F_{gt} = \sum_{i \in \mathcal{I}} m_{gi} L_{git} + \check{b}_g \tilde{Y}_{gt} \qquad \forall\, g \in \check{\mathcal{C}},\ t \in \mathcal{T} \tag{3.7}$$

(see §3.2.2)

$$X_{gt} = \sum_{i \in I} L_{git} \qquad \forall\, g \in \hat{\mathcal{C}},\ t \in \mathcal{T} \tag{3.8}$$

(see §3.2.2)

$$\sum_{i \in \mathcal{I}} Z_{git} \leq 1 \qquad \forall\, g \in \hat{\mathcal{C}},\ t \in \mathcal{T} \tag{3.9}$$

(see §3.2.2)

$$F_{gt} = \sum_{i \in \mathcal{I}} (m_{gi} L_{git} + \hat{b}_{gi} Z_{git}) \qquad \forall\, g \in \hat{\mathcal{C}},\ t \in \mathcal{T} \tag{3.10a}$$

$$\hat{k}_{g,i-1} Z_{git} \leq L_{git} \qquad \forall\, g \in \hat{\mathcal{C}},\ \ i \in \mathcal{I},\ i > 1, t \in \mathcal{T} \tag{3.10b}$$

$$L_{git} \leq \hat{k}_{gi} Z_{git} \qquad \forall\, g \in \hat{\mathcal{C}},\ \ i \in \mathcal{I},\ t \in \mathcal{T} \tag{3.10c}$$

(see §3.2.2)

$$X_{gt}, X^{sp}_{g't}, X^{sr}_{g't}, X^{st}_{t}, F_{g't}, L_{g'it}, \hat{D}_t, S_t, S^{+}_{t}, S^{-}_{t}, S^{s-}_{t} \geq 0$$

$$\forall\ \ g \in \mathcal{G},\ g' \in \mathcal{C},\ t \in \mathcal{T},\ i \in \mathcal{I}, \tag{3.11a}$$

$$Y_{gt}, \tilde{Y}_{gt}, \hat{Y}_{gt}, W^{+}_{t}, W^{-}_{t}, Z_{g'it}\ \text{ binary} \qquad \forall\, g \in \mathcal{C},\ g' \in \hat{\mathcal{C}},\ i \in \mathcal{I},\ t \in \mathcal{T} \tag{3.11b}$$

3.2.2 Detailed Model Description

We next give a detailed description of the model formulation. The objective function seeks to minimize fuel costs, operating and maintenance costs, penalties associated with unmet demand, and start-up costs. The following subsections detail the constraints. We categorize the constraints not only by their general sense, but also by the types of variables they contain.

Power Generation and Storage Constraint (3.1a) ensures that load is met using generation from conventional and renewable sources and a storage device. If generators and the storage device cannot provide sufficient generation, the elastic variable allows the demand constraint to be violated (and the objective function incurs a penalty in this case). We implement a portion of a linear minimum-maximum structure with constraint (3.1b) to identify the largest operating unit in each time period. Constraint (3.1c) ensures that sufficient spinning reserves are provided by operating generators and/or storage devices. The largest unit dictates the amount of spinning reserves that must be available in each time period. Constraints (3.1d) and (3.1e) restrict the rate of change in generation level between any two successive time periods to be less than or equal to the ramp-up or ramp-down rate, respectively. The sum of the change in generation level from one period to the next and any generator's committed spinning reserves cannot exceed that unit's ramp-up rate. This ensures that if a generator must provide reserves in addition to generation, the incremental ramping caused by the reserves does not force a generator to exceed its ramp-up rate. Constraint (3.1f) limits the amount of generation provided by renewable generators to a time-dependent hourly maximum output. Constraint (3.1g) ensures inventory balance with respect to the storage device, including losses incurred by storing power, for each period. Constraint (3.1h) limits the amount of spinning reserves provided by the storage device in any given period to be less than or equal to either (i) the amount of electricity flowing into the device if there is flow in during a given period or (ii) the difference between the maximum flow out of the device and the actual flow out if there is flow out during a given period. Constraint (3.1i) limits the amount of generation and reserves provided by the storage device to the amount of power already stored. Constraints (3.1j) limit the amount of outflow from the storage device to the maximum allowable. Constraints (3.1k) and (3.1l) limit the minimum and maximum hourly capacity of the storage device, respectively.

Generation Logic Constraint (3.2a) ensures that the generator has started up in at least the same period as steady-state operation occurs. Constraint (3.2b) prohibits any steady-state operation after a generator shuts down. See Figure 3.1 for a graphical representation of these constraints. Additionally, constraint (3.2c) enforces the minimum up-time requirement for each generator. Constraint (3.2d) enforces the minimum down-time requirement for each generator.

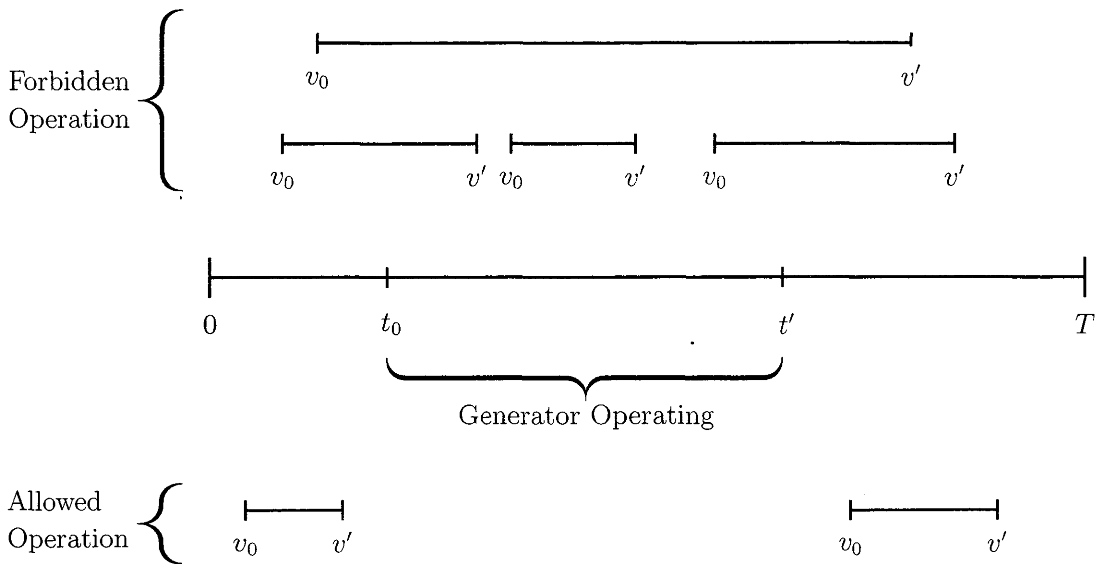

Figure 3.1. For conventional generators, we prohibit instances of overlapping start-up and shut-down times. For example, if a conventional generator operates between t_0 and t', it may not also start up in v_0, which occurs before t_0, and shut down in v', which occurs after t'. Other prohibited instances of (v_0, v') include starting up in v_0 and shutting down in v' when: (i) $v_0 < t_0$ and $t_0 < v' < t'$; (ii) $t_0 < v_0 < v' < t'$; and (iii) $t_0 < v_0 < t' < v'$. A conventional generator may start up in v_0 if either $v_0 < v' < t_0$ or $t' < v_0 < v'$. In other words, the instance (v_0, v') could either precede or follow (t_0, t').

Generation Related to Spinning Reserve Logic Constraint (3.3a) limits the amount of energy and spinning reserves provided by an operating generator to be no more than the unit's nominal hourly capacity if it is operating. Constraint (3.3b) determines the amount of spinning reserves from each generator in each time period as a function of the generator's hourly capacity and whether or not the generator is operating. This, in conjunction with constraint (3.1b), forces spinning reserves to be greater than or equal to the maximum hourly capacity of the largest unit operating for each time period. Constraint (3.3c) limits the amount of spinning reserves a generator provides in any period based on its hourly spinning reserve capacity and its operating status. Constraint (3.3d) requires a generator to be operating at or above its minimum hourly capacity while in steady-state operation.

Storage Logic with Binary Variables Constraint (3.4) requires that flow move in at most one direction relative to the storage device, i.e., only inflow (pumping) or outflow (generation) may occur in each time period.

Storage Logic with Mixed Variables Constraints (3.5a) and (3.5b) provide upper and lower bounds, respectively, on the inflow into the storage device (e.g., when pumping water into the storage pond). Constraint (3.5c) limits the amount of generation provided by the storage device to be less than or equal to the maximum outflow from the storage device. Constraint (3.5d) defines the minimum flow from the storage device per period.

Convex Piecewise Approximations with Continuous Variables The following set of constraints implements the piecewise-linear, convex approximation with respect to the continuous variables only. Constraint (3.6a) links the amount of generation to one or more specific segments of the curve. Constraint (3.6b) defines the operating level intervals for each segment. No binary segment indicator variables are necessary for these convex piecewise-linear approximations.

Convex Piecewise Approximation with Mixed Variables The following constraints link the convex continuous variables in §3.2.2 with the convex binary variables in §3.2.2. Constraint (3.7) determines the generator's fuel use based on the generation level, the independent variable, and the slope and intercept of the line segment.

Concave Piecewise Approximations with Continuous Variables The following set of constraints implements the concave piecewise-linear approximation with respect to the continuous variables only. Constraint (3.8) links the amount of generation to a specific segment of the curve.

Concave Piecewise Approximation with Binary Variables Constraint (3.9) limits the generator to being in at most one segment of the concave heat input curve in any time period.

Concave Piecewise Approximation with Mixed Variables The following constraints link the concave continuous variables in §3.2.2 with the concave binary variables in §3.2.2. Constraint (3.10a) determines the generator's fuel use based on the generation level, the independent variable, and the slope and intercept of the line segment. Constraints (3.10b) and (3.10c) define the operating level intervals for each segment using a binary variable.

Variable Type The final constraints, (3.11a) and (3.11b), restrict the variables to be nonnegative or binary, as appropriate.

3.3 Proposed Solution Techniques

While utilities wish to generate yearly generator dispatch schedules, annual unit commitment models with hourly or sub-hourly dispatch schedules may be intractable. For example, a variant of the unit commitment is NP-hard (Tseng, 1996). Our empirical results in §3.4 concur with this finding. To improve solution time and increase the size of models we can solve, we eliminate unnecessary variables and constraints and implement a cut related to the concave piecewise-linear heat input curves. We also identify ways to reduce computational overhead through modeling language usage. Finally, we must employ a *sliding time window heuristic* to obtain an annual schedule. This heuristic technique allows us to decompose the annual model into smaller horizons which we later concatenate to yield an annual schedule.

3.3.1 Variable and Constraint Elimination

While we prefer to incorporate the additional detail of piecewise-linear heat input curves, their presence, particularly for concave piecewise-linear curves, further complicates the unit commitment model because these curves require binary variables. Thus, we choose to not only define binary variables for the generators in $\hat{\mathcal{C}}$, but also to use power sets to generate binaries corresponding to the specific number of intervals each generator with a concave heat input curve possesses. We use the following notation:

- $\mathcal{J}$: set of breakpoints j for piecewise-linear approximations of heat input curves
- $\mathcal{I} \subseteq \mathcal{J}$: set of intervals i (consisting of the total number of breakpoints j) for piecewise-linear approximations of heat input curves
- $\mathcal{A}(\mathcal{J})$: set of generators with $|\mathcal{J}|$ breakpoints, i.e., breakpoints $1, 2, \ldots, |\mathcal{J}|$.

Thus, while set $\mathcal{I}$ may contain elements 1 through 4, if generator g's concave heat input curve only has two segments, we only define Z_{g1t} and Z_{g2t}.

To expedite model generation, we define the following parameter-dependent sets, i.e., sets whose elements depend on data, to eliminate unnecessary constraints:

- $\bar{\mathcal{T}}$: Intervals of time during which a generator is in a specific operating state, e.g., either in steady-state operation (including ramp up, steady state and ramp down) or shut down:
  - $\bar{\mathcal{T}}_A \subseteq \mathcal{T}$: An interval of the time horizon in which conventional generators operate in steady-state, i.e., not ramping up or down. For example, for a generator operating at hour t in the interval $[t_0, t']$:

    $$t \geq t_0 + r_g^{t+},\ t \leq t' - r_g^{t-}, t' \geq t_0 + r_g^{t+} + \underline{t}_g^u + r_g^{t-}$$

  - $\bar{\mathcal{T}}_B \subseteq \mathcal{T}$: An interval of the time horizon that defines the longest possible duration of steady-state operation:

$$1 + \min_{g \in G} \{r_g^{t+}\} \leq t \leq |\mathcal{T}| - \min_{g \in G} \{r_g^{t-}\}$$

Then, we employ these sets to eliminate irrelevant constraints and streamline summations in the formulation, e.g.:

$$X_t^{st} \geq X_{gt}^{sr} \qquad \forall g \in \mathcal{C}, t \in \bar{\mathcal{T}}_A$$

$$\sum_{g \in G} X_{gt} + S_t^- - S_t^+ + \hat{D}_t = d_t \qquad \forall\, t \in \bar{\mathcal{T}}_B$$

3.3.2 Cuts

We know that for each operating generator with a concave heat input curve, it must operate in one interval on its heat input curve. Conversely, for each non-operating generator, it cannot be operating anywhere on its heat input curve. From this information, we derive the following cut:

$$\sum_{i \in \mathcal{I} \ni i \leq |\mathcal{J}|-1} Z_{git} = \tilde{Y}_{gt} \quad \forall j \in \mathcal{J}, g \in \mathcal{A}(\mathcal{J}) \cap \hat{\mathcal{C}}, t \in \mathcal{T}.$$

3.3.3 Modeling Language Implementation

We use a modeling language (AMPL, 2009) both to write our problem instances and to implement the *sliding time window heuristic* algorithm. Therefore, we must utilize memory-saving constructs to the extent possible. Specifically, we: (i) define subsets over time periods a priori, rather than computing the elements in the set on each occurrence, e.g., subsets $\mathcal{T}_\mathcal{A}$ and $\mathcal{T}_\mathcal{B}$ (see §3.3.1); (ii) eliminate some constraint qualifiers by including the qualification in the set, e.g., defining sets of feasible pairs of t_0 and t' for each generator; (iii) and streamline expressions containing multiple terms with similar summations.

3.3.4 Sliding Time Window Heuristic

A sliding time window heuristic, also known as a *fix-and-relax* technique (Pochet and Wolsey, 2006), provides us with a feasible annual schedule for the system by decomposing the annual horizon into two-day intervals. This heuristic method is based on the idea of fixing some binary variable values while allowing others to remain unfixed in a given model instance. For each iteration of this heuristic, we divide our model into (i) an interval over which the model is solved, i.e., 48 hours; and (ii) a shorter window in which the variables are fixed at their optimal values (with respect to the iteration), i.e., 24 hours. Once we fix values for the binary variables in the first time window, e.g., the first 24 hours, we move forward in time and optimize over the second interval, i.e., we perform the second iteration. This iteration provides us with the binary variable values for the second time window, which we then fix. We terminate the procedure after having solved the model for the last interval in the horizon. Feasibility of the generator schedule is guaranteed because constraint (3.1a) permits a demand shortfall, and we have installed capacity in excess of the sum of the peak demand and spinning reserves.

Let $\underline{\tau}$ be the first time period of any given time window, and $\bar{\tau}$ be the last time period in the window. We introduce the following subsets of the set of conventional generators, $\mathcal{C}$, within the set of generators, $\mathcal{G}$:

- $\mathcal{C}'$: generators in steady state at the beginning of current window, regardless of up time requirements
- $\tilde{\mathcal{C}}$: generators which are in steady state but have not completed their minimum up time requirement at the beginning of the current window
- $\bar{\mathcal{C}}$: generators which have not completed their minimum down time at the beginning of the current window

We introduce the following additional parameters with respect to a current time window

- $\tilde{h}_g$, $g \in \mathcal{C}$: hours of steady-state operation which must be completed at the beginning of the current window
- $\bar{h}_g, g \in \mathcal{C}$: hours of minimum down time which must be completed at the beginning of the current window

We modify the objective to avoid incurring start-up costs from generators already in operation:

$$(\mathrm{P_{UCP}}) \quad \min \quad \sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} p_g^f F_{gt} + \sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T}} p_g^{om} X_{gt} + \sum_{t \in \mathcal{T}} \hat{p}\hat{D}_t + \sum_{g \in \mathcal{C}} \sum_{t \in \mathcal{T} \ni t > \underline{\mathcal{I}}} p_g^{\bar{s}} Y_{gt} + \sum_{g \in \mathcal{C} \setminus \mathcal{C}'} p_g^{\bar{s}} Y_{g\underline{\mathcal{I}}} \tag{3.12}$$

We add the following constraints:

1. If a generator has not met its minimum up time requirement, it must continue to operate for at least the remaining required minimum up time:

$$\tilde{Y}_{gt} = 1 \quad \forall g \in \tilde{\mathcal{C}}, t \in \mathcal{T} \ni t < \underline{\mathcal{I}} + \tilde{h}_g \tag{3.13}$$

We depict this new constraint in Figure 3.2. We identify our starting time as t_0, which occurs in the previous window. In this example, the generator has not completed its total ramp up, minimum up time and ramp down prior to the end of the horizon. The end of the horizon is indicated by the dotted vertical line. Thus, we must enforce constraint (3.13) to ensure that the generator completes the ramping and up time requirements in the current window. After this point, any time t' represents an acceptable shut down time.

2. If a generator shuts down prior to the beginning of the current window, but has not completed its required minimum down time, it must continue to be shut down for the remaining minimum down time:

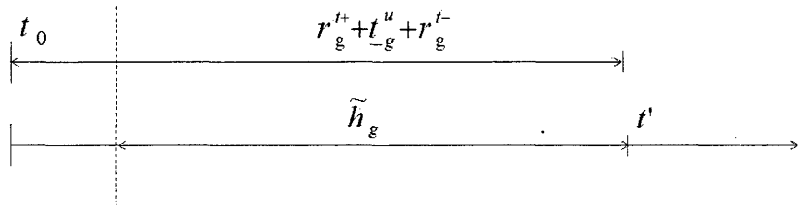

Figure 3.2. Generator in *steady state* at the beginning of the current window

$$Y_{gt} = 0 \quad \forall g \in \bar{\mathcal{C}}, t \in \mathcal{T} \ni t \leq \bar{\tau} + \bar{h}_g \tag{3.14}$$

In Figure 3.3, the generator has only completed a portion of its minimum downtime requirement, $\underline{t}_g^d$, in the previous horizon. We denote the end of the prior horizon and the beginning of the current horizon with a dotted vertical line. In the current horizon, $\hat{h}_g$ of down time must still be enforced. The first acceptable start-up time for this generator occurs at t_0.

In addition to the new constraints, we modify two existing constraints to maintain feasibility from one window to another. We fix the continuous variables associated with the power generation and quantity in storage in the period prior to the current window to ensure that ramp rates for power units are not violated and that the storage values are consistent with storage levels in the current interval. In our implementation, each interval contains 48 time periods, and each window contains 24 time periods.

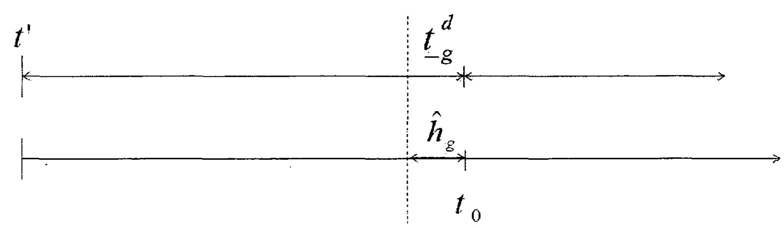

Figure 3.3. Generator in *minimum down time* at the beginning of the current window

3.4 Computational Results

Our data represents Xcel Energy's Colorado power system. We present results of estimated annual costs for five levels of spinning reserves, ranging from 0% to 100% of the mandated amount, using our *sliding time window heuristic* for this data set. Additionally, we compute expected emissions based on the schedules we obtain.

3.4.1 Data

We implement our *sliding time window heuristic* using a portion of the Grid View database provided by the National Renewable Energy Laboratory (NREL). Specifically, our data set contains generator and load information representing the Colorado power system in 2005. We use fuel consumption and emissions data from the US Environmental Protection Agency's continuous monitoring emission system (CEMS) data (US Environmental Protection Agency, 2005) to develop our own piecewise-linear heat input curves. We include all power generators in operation during the year 2005 in our analysis. Specifically, we include 83 power generators: 59 of these are gas-fired plants and 24 of these are coal-fired plants. We also include eight renewable generators. See Table 3.1 for a summary of generator and storage device capacities. The total installed capacity for the gas units is 3892.9 MW, while the total for the coal units is 4422.4 MW. The governing authority has specified that the Colorado power system must provide 177 MWh of spinning reserve, regardless of which generators operate. Thus, our implementation varies slightly from the formulation presented in §3.2.1 in that we do not use the minimum-maximum construct, which appears in constraints (3.1b) and (3.3b), to determine the largest unit operating in a given hour.

We modify certain parameters which NREL deems to be unreasonable. In particular, for any generator with a combined ramp-up, minimum up time and ramp-down time exceeding 24 hours, we reduce the minimum up time such that the combined time totals 24 hours. Similarly, we reduce any minimum down times that exceed one day to 24 hours.

Table 3.1. Summary of data: Generator statistics

Characteristic	**Minimum**	**Maximum**
Power Generator with Smallest Minimum Capacity	2.6 MW	10 MW
Power Generator with Largest Maximum Capacity	149.46 MW	498.2 MW
Aggregate Hourly Available Renewable Capacity	39 MW	1020 MW
Pumped Hydro Storage Device Capacity	0 MWh	975 MWh

We also model combined cycle units with significant detail. Using the GridView data, we identify the type of component: gas turbine or steam turbine. Each gas turbine may operate in single cycle mode according to the parameters specified in the data set. We also allow these gas turbines to have the option of operating in combined cycle mode. We generate the combined cycle mode data by evenly allocating the steam turbine's characteristics and costs across all of the gas turbines in the combined cycle unit. For a three-gas-turbine (150 MW capacity each), one-steam-turbine system (300 MW capacity), we increase each gas turbine's maximum capacity by 100 MW to

250 MW when it operates in combined cycle mode. Additionally, we use the steam turbine's maximum ramp-up, minimum up and ramp-down times for combined cycle operation, which are more conservative than the corresponding values for the single cycle mode data.

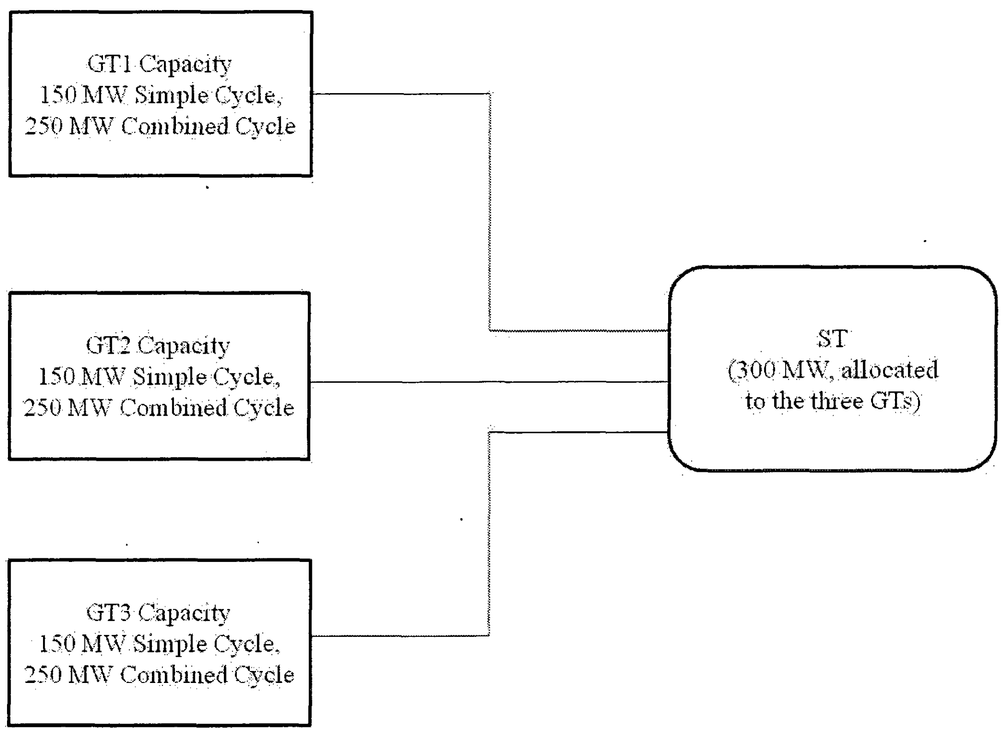

Figure 3.4. Allocation of steam turbine (ST) to gas turbines (GTs) to derive combined cycle unit capacities

3.4.2 Numerical Results for a Coal-Dominated Data Set

We use AMPL version 20090327 (AMPL, 2009) and CPLEX version 12.1 (IBM, 2010) to generate our annual dispatch schedule and annual operating cost estimates. We determine emissions after identifying the optimal dispatch schedule for the 48-hour horizon, i.e., we do not penalize emissions in the optimization model. Each 48-hour instance contains approximately 32,000 variables, 13,000 of them binary, and 40,000 constraints. We attempt to solve each horizon to optimality. We enforce a

1,000 second time limit to ensure that if an optimal solution cannot be identified quickly, we maintain a reasonable solution time for the heuristic. Of the five annual scenarios solved, at most 39 instances failed to converge within the 1,000 seconds, and the largest gap remaining after the time limit is 0.07%. We stress that while we solve each window to (near) optimality, the annual costs and emissions we describe in this section are only an estimate of the actual results that could be obtained with this system. For each horizon solved, the model contains only data corresponding to that single 48-hour horizon. Thus, an optimal schedule for a single window could result in a poor schedule for a subsequent window.

In Table 3.2, we present the annual start-up, operating and maintenance, and fuel costs obtained for our Colorado system considering five levels of spinning reserves ranging from 0% to 100% of the mandated amount. Comparing the 0% SR and 100% SR cases, significantly higher start-up costs, on the order of 5.5%, occur when the mandated spinning reserves must be provided. Enforcing the spinning reserve requirement also results in approximately 0.7% higher maintenance costs. Fuel costs vary by only 0.1% between the two extreme cases of 0% and 100% SR. The results show that total operating costs decrease by approximately 0.6% as the spinning reserve requirement declines from the mandated 100% case to 0% case. We also include Figure 3.5, which graphically demonstrates the cost savings presented in Table 3.2.

Table 3.2. Comparison of annual costs of providing varying levels of spinning reserves

	0% SR	**25% SR**	**50% SR**	**75% SR**	**100% SR**
Start-up	68,285,316	69,134,256	69,580,486	71,076,528	72,049,564
O&M	68,295,400	68,440,600	68,437,300	68,663,900	68,762,800
Fuel	681,855,000	681,970,000	682,187,000	682,273,000	682,611,000
Total	818,435,716	819,544,856	820,204,786	822,013,428	823,423,364

All costs in US dollars
O&M refers to operating and maintenance

We compute the amount of CO_2, NO_x and SO_x emitted based on the annual

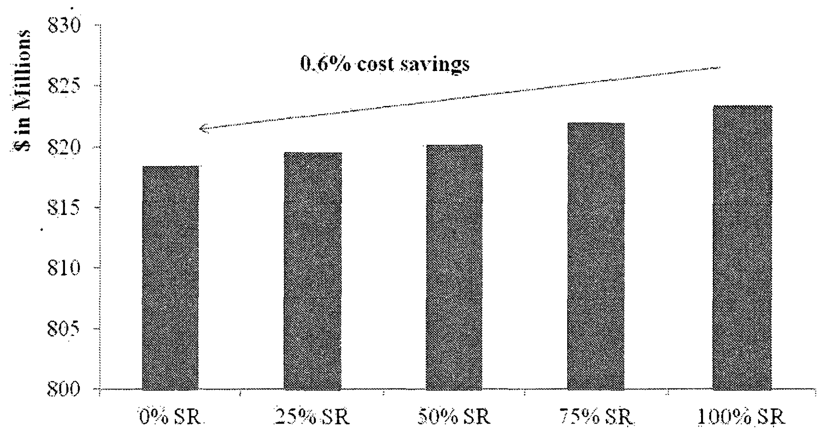

Figure 3.5. Total costs for five spinning reserve scenarios.

schedules obtained for each level of spinning reserves. The results in Table 3.3 indicate that all three types of greenhouse gas emissions increase as the spinning reserve requirement decreases. For CO_2 emissions, this result is intuitive because the majority of our generators has convex piecewise-linear heat input curves. Thus, the higher the generation level, the more fuel consumed and the more CO_2 emitted.

We also see the mix of fuel shift away from gas units as we change the level of spinning reserves, which explains the increase in NO_x and SO_2. The mix of generation shifts from 94.6% coal and 5.4% gas, when 177 MWh of SR are provided, to 94.9% coal and 5.1% gas, when the reserve requirement is eliminated. Figure 3.6 shows the increase in CO_2 emissions due to a decrease in the spinning reserve requirement. Similarly, Figure 3.7 shows the increase in NO_x and SO_x emissions for the five spinning reserve scenarios.

Current hardware and software can solve the linear programming (LP) relaxation of the annual model (see Table 3.4). We use these results to evaluate the quality of

Table 3.3. Comparison of annual emissions resulting from varying levels of spinning reserves

	0% SR	**25% SR**	**50% SR**	**75% SR**	**100% SR**
CO_2	30,019,450	29,994,900	29,996,200	29,960,000	29,943,100
NO_x	103,045,000	102,957,000	102,982,000	102,779,000	102,618,000
SO_x	94,673,600	94,578,000	94,580,700	94,426,400	94,300,200

CO_2 emissions in tons, NO_x and SO_x in pounds

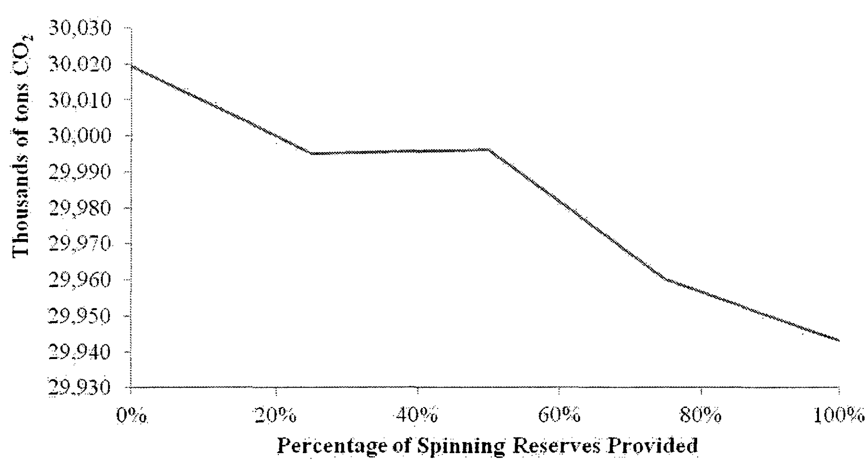

Figure 3.6. CO_2 emissions for five spinning reserve scenarios.

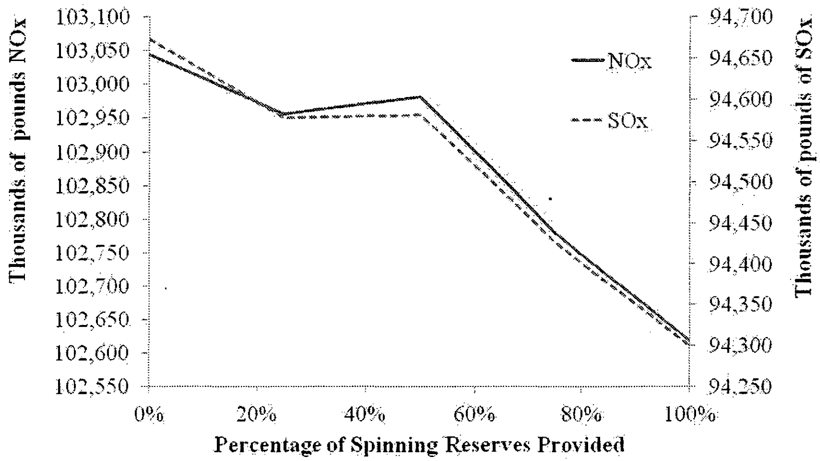

Figure 3.7. NO_x and SO_x emissions for five spinning reserve scenarios.

the solutions obtained from our *sliding time window heuristic*. The lower bounds (LP relaxation objective function values) for both the 0% SR and 100% SR cases are approximately 9% below the annual costs for the same cases obtained using our *sliding time window heuristic*. Based on these results, the solutions obtained by our hueristic are within 9% of optimality.

Table 3.4. Estimation of annual costs of providing varying levels of spinning reserves: LP relaxation

	0% SR	**100% SR**
Start-up	2,638,880	3,229,320
O&M	69,720,900	70,129,700
Fuel	67,9715,000	680,409,000
Total	752,074,780	753,768,020

All costs in US dollars
O&M refers to operating and maintenance

All annual figures result from the implementation of our *sliding time window heuristic*, which is not an exact solution technique. Thus, while we demonstrate potential cost savings and emissions increases from a reduction in spinning reserves, we cannot conclusively state that these results would be achieved with the reduction of the spinning reserve requirement. To provide an exact amount of cost savings, the model would need to be solved to optimality for the entire year. However, current hardware and software cannot solve such an instance of our problem, but it can solve one month models to a gap of approximately 0.05% or better. We present the cost savings and emissions results obtained for a single month using the coal-dominated data set in Tables 3.5 and 3.6.

Table 3.5. Comparison of monthly costs of providing varying levels of spinning reserves

	0% SR	**100% SR**
Start-up	374,804	457,717
O&M	4,983,830	5,021,270
Fuel	52,925,300	53,036,700
Total	58,283,934	58,515,687
MIP Gap	0.010%	0.035%

All costs in US dollars
O&M refers to operating and maintenance

Table 3.6. Comparison of monthly emissions resulting from varying levels of spinning reserves

	0% SR	**100% SR**
CO_2	2,401,950	2,392,890
NO_x	8,263,000	8,226,330
SO_x	7,676,150	7,626,050

CO_2 emissions in tons
NO_x and SO_x in pounds

Based on these monthly results, we conclusively show that cost savings on the order of 0.35% can be achieved for our data set. We also demonstrate that greenhouse gas emissions will increase by approximately: 0.4% for CO_2, 0.5% for NO_x, and 0.7% for SO_x.

3.4.3 Numerical Results for a Gas-Dominated Data Set

In order to provide further insights regarding the operating costs and emissions resulting from spinning reserves, we create a gas-dominated system and determine an annual commitment and dispatch schedule using our *sliding time window heuristic*. We derive the data from our original data set, removing the majority of the coal-fired generators and duplicating gas generators to achieve approximately the same installed capacity for the total system. The new data set contains 75 gas-fired generators and four coal-fired units. Total installed capacity of the gas-fired units is 6,603.0 MW and total installed capacity for the coal-fired generators is 736.2 MW. A summary of the data set appears in Table 3.7. We make no other modifications to the data set used in §3.4.2. See §3.4.1 for the full description of the original data set.

Table 3.7. Summary of data: Generator statistics

Characteristic	**Minimum**	**Maximum**
Power Generator with Smallest Minimum Capacity	2.6 MW	10 MW
Power Generator with Largest Maximum Capacity	123 MW	410 MW
Aggregate Hourly Available Renewable Capacity	39 MW	1020 MW
Pumped Hydro Storage Device Capacity	0 MWh	975 MWh

We solve each 48-hour window to an optimality gap of at least 0.05%. Based on the results in Table 3.8, cost savings on the order of 0.5% may still be achievable. Similar to the coal-dominated findings, start-up cost savings are significant due to the need for online quick-response units in the 100% SR case. These results are consistent with those for the coal-dominated system.

The impact of SR on greenhouse gas emissions for the gas-dominated system is not as clear. NO_x emissions increase as spinning reserves decline from the 100% SR case to the 0% SR case, while CO_2 and SO_x emissions remain relatively constant. The relatively constant fuel mix for the 0% and 100% SR scenarios support this ambiguous result. The percentage of generation from gas units for the 0% SR case is 78.36% and the generation from gas units is 78.35% for the 100% SR case. Table 3.9 shows the estimated annual emissions for the gas-dominated data set.

3.5 Conclusion

We present a mixed-integer linear program that minimizes fixed and variable costs of operation, and penalty costs for failure to meet demand. We incorporate not only standard constraints regarding the operation of the generators, but also details such as renewable generators, storage devices, and piecewise-linear convex and/or concave heat input curves. We implement several enhancements to increase model tractability and demonstrate their usefulness via computational results performed on a real data set for the Colorado region.

For a coal-dominated power system, our results demonstrate that reducing the spinning reserve requirement likely leads to a reduction in total annual operating costs. However, annual greenhouse gas emissions may increase because the use of

Table 3.8. Comparison of annual costs of providing varying levels of spinning reserves: Gas-dominated data set

	0% SR	**25% SR**	**50% SR**	**75% SR**	**100% SR**
Start-up	330,700	331,619	334,823	335,327	337,889
O&M	182,609	182,588	182,505	182,371	182,268
Fuel	812,737	812,825	812,345	812,895	812,877
Total	1,326,046	1,327,032	1,329,673	1,330,593	1,333,034

All costs in thousands of US dollars
O&M refers to operating and maintenance

Table 3.9. Comparison of annual emissions resulting from varying levels of spinning reserves: Gas-dominated data set

	0% SR	**25% SR**	**50% SR**	**75% SR**	**100% SR**
CO_2	16,485,100	16,485,850	16,478,750	16,486,350	16,486,550
NO_x	40,727,100	40,727,900	40,734,100	40,711,900	40,720,600
SO_x	27,071,300	27,070,800	27,072,700	27,069,000	27,076,300

CO_2 emissions in tons, NO_x and SO_x in pounds

gas turbines may decline as quick-response capacity is no longer needed to provide reserves. On a monthly basis, we show that eliminating the spinning reserve requirement for Xcel's Colorado territory will lead to a cost reduction of approximately 0.4% for a single month. We also confirm that greenhouse gas emissions will increase over this horizon if spinning reserves are not provided.

Using a manually generated, gas-dominated system, we observe that cost savings may still be achieved by lowering the spinning reserve requirement. However, the change in emissions due to the decrease in spinning reserves is ambiguous. More analysis using data from an actual gas-dominated system should be performed to verify these results.

CHAPTER 4

COMPARISON OF SINGLE- AND DUAL-TIME-INDEX VARIABLES USING THE UNIT COMMITMENT MODEL

The typical unit commitment model utilizes single-time-index binary variables, e.g., Y_{gt}, which assume a value of one if we start generator g in period t, zero otherwise (Sioshansi *et al.*, 2008). Using this variable definition, three distinct types of binary variables, each representing a different operating state, i.e., start up, steady state and shut down, must be defined to approximate the operation of generators within an actual power system. The single-time index formulation in the current literature has a shortcoming; it fails to preclude generation during ramp-up and ramp-down periods. Modification of this formulation through the introduction of lagged steady-state binary variables might preclude generation during ramp periods. However, discussion of such an idea does not appear to exist.

For small gas units which ramp up to minimum capacity in fewer than 10 minutes, these ramp times may be trivial and thus, the overstatement of generation might be minimal. However, large combined cycle and coal units could require two to three hours to ramp up to minimum capacity (Harvey, 2010). Our particular data set, which characterizes Xcel Energy's Colorado territory, contains ramp times ranging from nearly instantaneous to 45 minutes. Considering the sum of the minimum up, minimum down, ramp up, and ramp down times, a large coal generator could cycle, i.e., start up and shut down, 21 times per month. Thus, a unit commitment model for a system of coal and/or combined cycle units, especially with sub-hourly fidelity, could significantly overstate the amount of available generation by using the single-time-index formulation without a lag parameter.

We address this lack of accuracy regarding generation during ramp up or ramp down using a dual-time-index formulation. We use a single binary variable to control both the start up and shut down of a generator, as well as to determine the initial ramp-up and final ramp-down periods. This variable, $Y_{gt_0t'}$, assumes a value of one if we start generator g in period t_0 and shut generator g down in period t', zero otherwise. The dual time indices allow us to define the acceptable time horizons over which a generator may operate. Specifically, the interval between t_0 and t' must be long enough to allow for ramp up, minimum up time and ramp down. A single-time-index formulation cannot incorporate all of this information into the variables; constraints must be implemented to ensure that ramp up, minimum up time and ramp down are satisfied. We will compare our dual-time-index formulation to two single-time-index formulations: (i) we will consider only the minimum up time requirement for each generator and ignore initial ramp-up and final ramp-down times and (ii) we will add the initial ramp-up and final ramp-down times to the minimum up time requirement, forcing the generator to remain on for a longer period of time than in (i). We note that the second formulation still does not preclude generation in the ramp-up and ramp-down periods.

The dual-time-index formulation, in addition to precluding generation during ramping, also enables us to determine a generator's state at any time t. If generator g starts in t_0, it must ramp up for a predetermined period of time until it reaches minimum capacity, at which time it may begin providing electricity to the grid. Similarly, if generator g shuts down in t', we can determine exactly when it began to ramp down and stopped generating. With this information, we can preclude any generation in periods of ramping and ensure that the minimum up time requirement is satisfied.

A significant drawback of the dual-time-index formulation, from an operational perspective, is that the t' subscript forces all generators that start up in the horizon to also shut down. Conversely, the single-time-index formulation, as presented in the literature and in our formulation in Chapter 3, does not require shut down of any operating generator near the end of the horizon. Thus, the dual-time-index formulation may need to start up additional generators at the end of the horizon to meet load as steady-state generators begin to ramp down. This requirement, in turn, may increase the objective function value of the dual-time-index formulation. The only way to mitigate the shut-down requirement would be to extend the time horizon by the maximum amount of ramp-down time, considering all of the generators in the system. However, increasing the horizon may unnecessarily increase problem size by adding more dual-time-index binary variables and constraints.

The remainder of the chapter is organized as follows: In Section 4.1, we provide a literature review. In Section 4.2, we present the relevant single- and dual-time-index variable definitions and the constraints which control the generator commitment decision, e.g., start up, shut down, steady-state operation, and minimum up and down time. In Section 4.3, we describe opportunities to eliminate variables from the monolith as well as opportunities to sparsify two dense constraints. In Section 4.4, we discuss numerical results contrasting operating costs using the two different formulations. Section 4.5 concludes the chapter.

4.1 Literature Review

All literature we review regarding the unit commitment model contains only single-time-index binary variables. Baldick (1995) presents a unit commitment problem which uses a single-time-index binary variable to indicate if the unit is online. Gollmer *et al.* (2000) similarly use one single-time-index binary variable to determine if a generator operates in time t. Sioshansi *et al.* (2008) use three single-time-index binary variables to indicate if the generator starts up, operates in steady state, or shuts down in period t. The authors typically do not disclose whether the minimum up time requirements include ramp-up and ramp-down times. Additional summaries of unit commitment models appear in §2.1 and §3.1.

Applications of dual-time-index variables generally appear outside of the unit commitment literature. Brown *et al.* (1991) develop a mixed-integer linear program (MILP) to determine the optimal schedule to procure and update the US Army's helicopter fleet. The authors use a dual-time-index binary indicator variable to determine if a production line is opened at the beginning of year t_0 and closed at the end of year t'. The linking of the opening and closing of a line lends itself to the dual-time-index formulation because there must be continuous use of the line while open, and the sequencing of a particular line may depend on other actions determined by the model. Brown *et al.* (2003) present an MILP which selects which space and launch systems the US Air Force Space Command should purchase, and determines when to launch the systems. The dual-time-index variables are based on predetermined pairs of starting and ending years which result from the life expectancy of each system. Castro and Grossmann (2005) use a dual-time-index variable definition for a continuous-time MILP characterizing a multistage production plant. To simplify the problem, the authors assume a limited amount of time that may elapse between the starting time, t, and the ending time t'.

4.2 Formulations

We compare a single-time-index formulation described in Chapter 3 with a dual-time-index formulation. For the dual-time-index formulation, we use a similar objective function, minimizing total system costs, the same continuous variables, and the same types of constraints as the mixed-integer linear unit commitment formulation. As necessary, we replace the single-time-index binary variables with the dual-time-index variable. A version of the dual-time-index formulation appears in Van Dinter (2010). We do not include transmission or emissions in either optimization model.

We present the relevant sets and parameters required for both of the formulations.

Sets:

- $\mathcal{C}$: set of conventional generators

- $\mathcal{T}$: time horizon over which the system is optimized

Generator-Specific Parameters:

- r_g^{r+}: ramp-up rate of generator g [MWh/hr]
- r_g^{r-}: ramp-down rate of generator g [MWh/hr]
- r_g^{t+}: ramp-up time required for generator g [hr]
- r_g^{t-}: ramp-down time required for generator g [hr]
- $\underline{t}_g^u$: minimum up time of generator g before shut down [hr]
- $\underline{t}_g^d$: minimum down time of generator g before start up [hr]

We first describe the typical formulation, i.e., the single-time-index variable method, as described in Chapter 3, and we subsequently present our alternative dual-time-index formulation.

4.2.1 Single-Time-Index Binary Decision Variables and Generation Logic

- $Y_{gt} = \begin{cases} 1 & \text{if generator } g \text{ starts up at time } t \\ 0 & \text{otherwise} \end{cases}$
- $\tilde{Y}_{gt} = \begin{cases} 1 & \text{if generator } g \text{ operates in steady state at time } t \\ 0 & \text{otherwise} \end{cases}$
- $\hat{Y}_{gt} = \begin{cases} 1 & \text{if generator } g \text{ shuts down at time } t \\ 0 & \text{otherwise} \end{cases}$

$$Y_{gt} \geq \tilde{Y}_{gt} - \tilde{Y}_{g,t-1} \quad \forall\, g \in \mathcal{C}, t \in \mathcal{T} \tag{4.1}$$

$$\hat{Y}_{gt} \geq \tilde{Y}_{g,t-1} - \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C}, t \in \mathcal{T},\ t > 1 \tag{4.2}$$

$$\textstyle\sum_{t'=t-\underline{t}_g^u-r_g^{t+}-r_g^{t-}}^{t} Y_{gt'} \leq \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C},\ t \in \mathcal{T},\ t > 1 \tag{4.3}$$

$$\textstyle\sum_{t'=t-\underline{t}_g^d}^{t} \hat{Y}_{gt'} \leq 1 - \tilde{Y}_{gt} \quad \forall\, g \in \mathcal{C},\ t \in \mathcal{T} \tag{4.4}$$

Constraint (4.1) ensures that the generator has started up in at least the same period as steady-state operation occurs. Constraint (4.2) prohibits any steady-state operation after a generator shuts down. Constraint (4.3) enforces the minimum up time requirement for each generator. Constraint (4.4) enforces the minimum down time requirement for each generator.

4.2.2 Dual-Time-Index Binary Decision Variables and Generation Logic

- $Y_{gt_0t'} = \begin{cases} 1 & \text{if generator } g \text{ starts up at time } t_0 \text{ and shuts down at time } t' \\ 0 & \text{otherwise} \end{cases}$

In addition to the dual-time-index variable definition, we must also constrain the scheduling of generators to preclude overlapping intervals of operation and ensure that the minimum up time and minimum down time requirements are satisfied. If we do not carefully define the feasible combinations of t_0 and t', we could generate $|\mathcal{C}| \times |\mathcal{T}| \times |\mathcal{T}|$ dual-time-index variables. We describe the tightening of the combinations of t_0 and t' in §4.3.1 to generate the fewest variables possible.

4.3 Solution Techniques

We employ two techniques to improve model tractability. First, we use a concise variable definition and eliminate unnecessary variables and constraints, reducing the problem size. We then use variable substitution to sparsify certain dense constraints.

4.3.1 Variable and Constraint Elimination

We undertake several steps to eliminate unnecessary instances of $Y_{gt_0t'}$. We rely on the mathematical definition of variable $Y_{gt_0t'}$ to eliminate infeasible combinations of conventional generator start-up and shut-down times. For example, consider generator g starting in hour t_0. Generator g has a ramp-up time of r_g^{t+} hours and a ramp-down time of r_g^{t-} hours. Additionally, generator g cannot shut down unless it has operated for at least $\underline{t}_g^u$ hours. Therefore, generator g cannot shut down until at least $r_g^{t+} + \underline{t}_g^u + r_g^{t-}$ hours since its start up. This information allows us to eliminate all binary variables for generator g with combinations of (t_0, t') shorter than this minimum operating horizon, implicitly implementing the minimum up time requirement. Additionally, we can incorporate the minimum down time requirement by restricting the instances of v_0 and v' to occur after the minimum down time is satisfied.

To capture this variable elimination, we define the following parameter-dependent sets, i.e., sets whose elements depend on data:

- $\bar{\mathcal{T}}$: Intervals of time during which a generator is in a specific operating state, e.g., either operating (including ramp up, steady state and ramp down) or shut down:
  - $\bar{\mathcal{T}}_1 \subseteq \mathcal{T}$: Intervals of the time horizon in which a generator g can only be in one state at one time, e.g., for a generator that operates in the interval $[t_0, t']$, it cannot operate in any of the following V_i intervals:
    - V_1: $v' \geq v_0 + r_g^{t+} + \underline{t}_g^u + r_g^{t-}, t_0 \leq v_0 < t', v' > t'$
    - V_2: $v' \geq v_0 + r_g^{t+} + \underline{t}_g^u + r_g^{t-}, v_0 < t_0, t_0 < v' \leq t'$
    - V_3: $v' \geq v_0 + r_g^{t+} + \underline{t}_g^u + r_g^{t-}, v_0 < t_0, v' > t'$
    - V_4: $v' \geq v_0 + r_g^{t+} + \underline{t}_g^u + r_g^{t-}, v_0 > t_0, v' \leq t'$
  - $\bar{\mathcal{T}}_2 \subseteq \mathcal{T}$: An interval of the time horizon in which a generator g can operate given that it has previously operated in the interval $[t_0, t']$, and has met its minimum down time requirement, e.g., for a generator operating in $[t_0, t']$ followed by $[v_0, v']$:

$$t' \geq t_0 + r_g^{t+} + \underline{t}_g^u + r_g^{t-},\ v_0 \geq t' + \underline{t}_g^d,\ t_0 \neq v_0,\ v' \geq v_0 + r_g^{t+} + \underline{t}_g^u + r_g^{t-}$$

Then, we employ these sets to eliminate irrelevant constraints and streamline summations in the formulation, e.g.:

$$Y_{gt_0t'} + \sum_{v_0 \in \bar{\mathcal{T}}_1} \sum_{v' \in \bar{\mathcal{T}}_1} Y_{gv_0v'} \leq 1 \qquad \forall\, g \in \mathcal{C},\ t_0,\ t' \in \bar{\mathcal{T}}_1 \cap \bar{\mathcal{T}}_2$$

4.3.2 Constraint Sparsification

The dual-time-index variable definition results in very dense rows. We therefore formulate these constraints as follows (Leyffer, 2010). Considering constraint 4.5:

$$Y_{gt_0t'} + \sum \sum Y_{gv_0v'} \leq 1 \quad \forall\, g \in \mathcal{C},\ t_0,\ t' \in \bar{\mathcal{T}}_1 \cap \bar{\mathcal{T}}_2$$

$$Y_{gt_0t'} + \sum_{v_0 \in \bar{\mathcal{T}}_1} \sum_{v' \in \bar{\mathcal{T}}_1} Y_{gv_0v'} \leq 1 \quad \forall\, g \in \mathcal{C},\ t_0,\ t' \in \bar{\mathcal{T}}_1 \cap \bar{\mathcal{T}}_2 \tag{4.5}$$

we define a new decision variable:

$$\tilde{Y}_{gt} = \begin{cases} 1 & \text{if generator } g \text{ starts up in time } t \\ 0 & \text{otherwise} \end{cases}$$

and replace (4.5) with the following two constraints:

$$\tilde{Y}_{gt_0} = \sum_{t' \in \bar{\mathcal{T}}_1} Y_{gt_0t'} \quad \forall\, g \in \mathcal{C},\ t_0 \in \bar{\mathcal{T}}_1 \tag{4.6}$$

$$\tilde{Y}_{gt} + \sum_{t' \in \bar{\mathcal{T}}_1} Y_{gt_0t'} \leq 1 \quad \forall\, g \in \mathcal{C},\ t \in \mathcal{T},\ t_0 \in \bar{\mathcal{T}}_1,\ t_0 \neq t \tag{4.7}$$

In Table 4.1, we summarize the reduction in problem size, particularly in the number of nonzeros, after sparsifying the dense constraints in the dual-time-index formulation. Reducing the number of nonzeros in the problem can significantly improve model tractability, as matrix computations are less onerous. Note that we need not enforce integrality on variable $\tilde{Y}_{gt}$ because (4.6) does so implicitly.

In addition to the numerical statistics, Figure 4.1 illustrates the difference in density of the A matrix, i.e., the matrix of left-hand-side constraint coefficients, with and without the sparse constraints for a similar instance of the model we describe in this chapter. Both with the denser and with the sparser constraints, many rows contain between zero and fifty non-zero coefficients, and are therefore not reported here due to the scale of the histogram. Both formulations contain approximately the same number of constraints with up to 300 non-zeros. The difference in sparsity

Table 4.1. Comparison dual-time-index problem statistics, 83 generators and 48-hour horizon

	Sparse	Dense
Constraints	101,299	155,759
Continuous Variables	31,444	17,805
Binary Variables	69,855	76,323
Nonzeros	3,806,588	56,165,500

Sparse: (4.6) and (4.7)
Dense: (4.5)

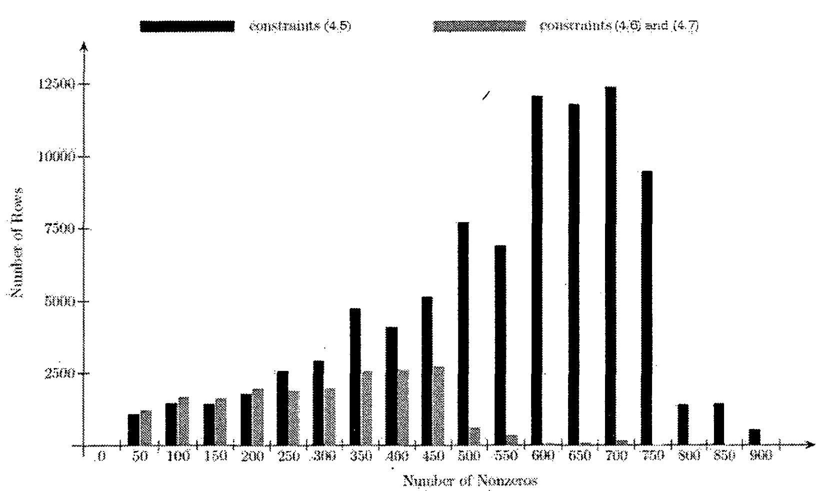

Figure 4.1. Histogram of sparsified and dense formulations.

between formulations starts to become clear for the occurrences of between 350 and 450 non-zeros per constraint. After this point, the dense formulation contains an astronomical number of constraints with nonzero entries relative to the sparse formulation. In other words, while the sparse formulation contains almost no constraints with more than 500 non-zeros, the dense formulation contains tens of thousands of such constraints.

4.4 Comparison of Formulations and Numerical Results

First, we present a summary of the data we use to obtain our numerical results. Using this data, which corresponds to Xcel's Colorado region in 2005, we compare several variants of single- and dual-time-index formulations. We primarily focus on differences in problem size, model tractability, and objective function value.

4.4.1 Data

We use generator and hourly load data provided by the National Renewable Energy Laboratory (NREL). The data, obtained from the GridView database, represents the Colorado power system in 2005. We use the US Environmental Protection Agency's continuous monitoring emission system (CEMS) data (US Environmental Protection Agency, 2005) to develop our own piecewise-linear heat input curves. We include all generators operating during the year 2005 in our data set. In particular, we include 59 gas-fired plants, 24 coal-fired plants and eight renewable generators in our data set. Table 4.2 contains a summary of generator and storage device capacities. The total installed power capacity is 8,315.3 MW. The installed capacity of coal units is 4422.4 MW, while the total for the gas units is 3892.9 MW.

We modify certain parameters based on recommendations from NREL. For example, if a generator has a combined ramp-up time, minimum up time and ramp-down time of 24 hours or more, we shorten the minimum up time requirement such that the combined time totals 24 hours. Similarly, any minimum down times longer than one day are shortened to 24 hours. We do not modify any ramp rates and we note that the longest ramp time is 45 minutes.

Table 4.2. Summary of data: Generator statistics

Characteristic	Min Capacity	Max Capacity
Power Generator with Smallest Minimum Capacity	2.6 MW	10 MW
Power Generator with Largest Maximum Capacity	149.46 MW	498.2 MW
Aggregate Hourly Available Renewable Capacity	39 MW	1020 MW
Pumped Hydro Storage Device Capacity	0 MWh	975 MWh

We incorporated detailed operation of combined cycle units in our model. We first determine the type of combined cycle component: gas turbine or steam turbine. All gas turbines may operate in single cycle mode, adhering to the parameters specified in the data set. Alternatively, these gas turbines may operate in combined cycle mode. We evenly allocate the steam turbine's operating characteristics and costs across all of the gas turbines in the combined cycle unit to determine the maximum capacity of each gas turbine operating in combined cycle mode. For a 3 $\times$ 1 system with three 150-MW gas turbines and one 300-MW steam turbine, we increase each gas turbine's maximum capacity by 100 MW to 250 MW when it operates in combined cycle mode. We use the more conservative steam turbine data for maximum ramp-up, minimum up and ramp-down times for combined cycle operation.

4.4.2 Numerical Results

We are able to solve instances with our monolith to a reasonable gap of 0.05% or less using AMPL (AMPL, 2009) and CPLEX (IBM, 2010), by including the enhancements mentioned in §4.3.1 and §4.3.2. Our results represent a system of about 85 generators and 120 time periods.

4.4.3 Sparsification Impact on Dual-Time-Index Formulation

We solve two instances of the dual-time-index formulation to evaluate the enhancements described in the first two subsections of §4.3. Note that for all results in this section, the objective function values from the optimized models are given in millions of $US. The first row of the second and third columns of Table 4.3 reports objective function values of the integer program of the sparsified model after it reaches a 0.0% gap. The third row of Table 4.3 shows the solution time benefits of using constraints (4.6) and (4.7), rather than the originally formulated constraint (4.5). Although the smallest instance shows a small reduction in solution time from the sparse constraints, the 48-hour sparsified formulation demonstrates a significant improvement. We obtain an optimal solution using the sparse formulation in fewer than 500 seconds versus a solution within 0.05% of optimality within a time limit of 3,600 seconds using the dense formulation.

Table 4.3. Dual-time-index formulation with and without sparsification

	Sparse		Dense	
	24 hours	**48 hours**	**24 hours**	**48 hours**
Objective Function	1,841,523	4,028,698	1,841,523	4,028,801
Final MIP gap	0.00%	0.00	0.00%	0.05%
CPLEX time [sec.]	3	476	12	[b]

[b] exceeded 3,600 second time limit

Based on the results in Table 4.3, we conclude that the sparsified dual-time-index model performs better than the dense formulation. Thus, we only use the sparsified model in our subsequent analyses between the dual- and single-time-index models.

4.4.4 Comparison of Single- and Dual-Time Index Formulations

We compare the sizes of the three types of problems: the single-time-index formulation with and without incorporating the ramp-up and ramp-down times into the minimum up time requirement as well as the dual-time-index formulation. We summarize this information in Table 4.4. In general, the two single-time-index models yield similar problem sizes in terms of the number of variables and constraints, with the only difference occurring in the number of nonzeros. However, the size of the dual-time-index model is vastly different from the single-time-index models. The dual-time-index model has more than twice as many constraints as the single-time-index models and has roughly seven times as many binary variables. The number of nonzeros is significantly higher, with the sparsified model having 3.8 million nonzeros compared to roughly 170,000 nonzeros in the single-time-index models. Thus, even though we have sparsified the dual-time-index formulation, it is still significantly more dense than the single-time-index formulation.

Table 4.4. Comparison of single- and dual-time-index problem statistics, 48-hour horizon

	Case A	**Case B**	**Case C**
Constraints	40,185	40,185	101,299
Continuous Variables	18,555	18,555	31,444
Binary Variables	12,814	12,814	69,855
Nonzeros	167,871	174,660	3,806,588

Case A: Single time index, ramp times ignored
Case B: Single time index, ramp times times added to minimum up time requirement
Case C: Dual time index, dense constraints sparsified

In Table 4.5, we present the objective function values for the two single-time-index instances as well as the dual-time-index model with sparsification. We find that the dual-time-index objective function value is approximately 0.6% higher than the objective function values provided by the single-time-index models. Because all generators must shut down by the end of the horizon and generation is precluded during ramp up and ramp down, extra generators may need to start up both early and late in the horizon to avoid demand shortfalls. The start up of these additional units likely contributes to the higher objective function value.

Table 4.5. Comparison of single- and dual-time-index objective function values, 48-hour horizon

	Case A	**Case B**	**Case C**
Objective Function	4,003,433	4,004,621	4,028,698
CPLEX time [sec.]	4,772	4,210	476

Case A: Single time index, ramp times ignored
Case B: Single time index, ramp times times added to minimum up time requirement
Case C: Dual time index, dense constraints reformulated
All instances solved to optimality

Despite a denser A matrix shown in Table 4.4, the dual-time-index model solves significantly faster than the single-time-index models (see Table 4.5). We attribute our ability to solve the dual-time-index instances not just to the enhancements we describe in §4.3.1, but also to our dual time-indexed variable, which has the advantages over the traditional single time-indexed variable of: (i) enabling CPLEX to conduct its probing (fixing variable values a priori) and branching routines more efficiently, and (ii) reducing the number of necessary constraints, specifically, those linking the start-up and shut-down binary variables, which, in turn, helps to expedite solving linear programming problems at the nodes, especially when complexity necessitates the presence of other constraints, e.g., those controlling the piecewise-linear functions (Klotz, 2009).

4.4.5 Forcing Shut Down in the Single-Time-Index Formulations

One significant difference between the single- and dual-time-index formulations exists. The dual-time-index formulation implicitly assumes that all generators will shut down at the end of the horizon. The single-time-index formulation, as presented in Chapter 3, does not contain the same assumption. In order to make the models comparable with respect to this issue, we incorporate a constraint into the single-time-index formulation which forces all generators to shut down by the end of the horizon. We find that even with this additional constraint, the single-time-index formulation understates the total system costs for a 48-hour horizon (see Table 4.6) due to our inability to preclude generation during ramp up and ramp down.

Table 4.6. Comparison of single- and dual-time-index objective function values and forcing shut down, 48-hour horizon

	Case A	**Case B**	**Case C**
Objective Function	4,004,621	4,010,871	4,028,698

Case A: Single time index, ramp times times added to minimum up time requirement
Case B: Single time index, ramp times times added to minimum up time requirement and shut down forced
Case C: Dual time index, dense constraints reformulated
All instances solved to optimality

4.5 Conclusions

We introduce a dual-time-index variable definition based on the typical unit commitment model's single-time-index variable definition. The variable incorporates important details regarding the operational behavior of power generators, specifically, precluding generation during ramp-up and ramp-down times. This level of modeling precision typically does not appear in the current unit commitment literature. The dual-time-index formulation allows the user to determine the times at which a generator completes ramp up to minimum capacity or begins ramp down. We present computational results of the two formulations, using our own unit commitment models described in Van Dinter *et al.* (2011a) and Van Dinter (2010).

We note one obvious drawback of using the dual-time-index variable: all generators must shut down by the end of the horizon. While we have modified the single-time-index formulation to more closely resemble the dual-time-index formulation by implementing a constraint to force all generators to shut down at the end of the horizon, the formulations are not identical in terms of precluding generation during ramping. A single-time-index formulation with lag variables and/or parameters to preclude generation during ramping should be compared to the dual-time-index formulation to confirm our results.

CHAPTER 5

CONCLUSIONS

We conclude this dissertation with a discussion of our contributions and opportunities for additional research. Specifically, in §5.1, we provide a thorough list of our results and research contributions. In §5.2, we briefly summarize a few of the opportunities to extend the research presented in this dissertation.

5.1 Research Contributions

This dissertation addresses variants of the unit commitment problem. First, we present a hybrid Benders' Decomposition algorithm designed to solve mixed-integer, nonlinear versions of this problem. We use piecewise-linear underestimators to establish strong bounds on the problem and use information from auxiliary problems to determine a priori which generators must start or must remain off during the time horizon. To supplement the basic Benders' Decomposition algorithm, we develop a series of cuts which supplies useful information regarding the storage device and the generation required to meet demand.

We implement a *sliding time window heuristic* to quantify the operating costs of providing spinning reserves for Xcel's Colorado territory. For our particular data set, we find a potential annual cost savings of approximately 0.6%. Interestingly, we find that as reserves are eliminated from the system, the fuel mix shifts slightly away from gas and towards coal generation, resulting in higher greenhouse gas emissions.

Finally, we introduce an alternative formulation for the unit commitment problem which offers greater accuracy by prohibiting generation during ramp-up and ramp-down periods. We present preliminary results comparing our dual-time-index formulation to the single-time-index formulation used by Sioshansi *et al.* (2008) and others. Based on these results, the dual-time-index formulation solves in a shorter period of time than the single-time-index formulation. Additionally, the single-time-index formulation appears to understate operating costs slightly by allowing generators to produce electricity before they reach minimum capacity.

5.2 Possible Research Extensions

- In Chapter 2, we use piecewise-linear underestimators to provide a lower bound on the problem. Incorporating piecewise-linear overestimators would likely provide valuable information to the Benders' master problem and could facilitate convergence.

- In Chapter 3, incorporating the cuts derived in Chapter 2 may yield faster solution times, reducing the total solution time required for the *sliding time window heuristic* and allowing optimal solutions to be obtained at all iterations. Furthermore, solving single instances with longer time horizons to optimality would provide additional support for our results regarding costs and emissions.

- In Chapter 4, comparing the two formulations on different data sets, e.g., ones that include longer ramp times or have sub-hourly dispatch fidelity, may provide greater insight into the differences in solution quality and accuracy. Additionally, column generation may improve tractability, given the large number of variables in the problem relative to the number of constraints.

REFERENCES CITED

AMPL. 2009. *AMPL Version 2009.03.27.* AMPL Optimization LLC.

Baldick, R. 1995. The Generalized Unit Commitment Problem. *IEEE Transactions on Power Systems*, **10**(1), 465–475.

Bouffard, F., and Galiana, F. 2008. Stochastic Security for Operations Planning with Significant Wind Power Generation. *IEEE Transactions on Power Systems*, **23**(2), 306–316.

Bouffard, F., Galiana, F., and Conejo, A. 2005a. Market-Clearing with Stochastic Security-Part I: Formulation. *IEEE Transactions on Power Systems*, **20**(4), 1818–1826.

Bouffard, F., Galiana, F., and Conejo, A. 2005b. Market-Clearing with Stochastic Security-Part II: Case Studies. *IEEE Transactions on Power Systems*, **20**(4), 1827–1835.

Brooke, A., Kendrick, D., Meeraus, A., and Rosenthal, R. 1988. *GAMS, a User's Guide.* Scientific Press.

Brown, G., Clemence, R., Teufert, W., and Wood, R. 1991. An Optimization Model for Modernizing the Army's Helicopter Fleet. *Interfaces*, **21**(4), 39–52.

Brown, G., Dell, R., Holtz, H., and Newman, A. 2003. How US Air Force Space Command Optimizes Long-Term Investment in Space Systems. *Interfaces*, **33**(4), 1–14.

Burtraw, D., and Szambelan, S. J. 2009. U.S. Emissions Trading Markets for SO2 and NOx. *Discussion Paper*, **9**(40), 1–42.

Bussieck, M., and Meeraus, A. 2004. General Algebraic Modeling System (GAMS). *Pages 137–157 of:* Kallrath, J. (ed), *Modeling Languages in Mathematical Optimization.* Norwell, MA: Kluwer Academic Publishers.

Castro, P., and Grossmann, I. 2005. New Continuous-Time MILP Model for the Short-Term Scheduling of Multistage Batch Plants. *Industrial & Engineering Chemistry Research*, **44**(24), 9175–9190.

Denholm, P., and Sioshansi, R. Forthcoming. The Value of Energy Storage with Wind in Transmission-Constrained Electric Power Systems.

Dillon, T., Edwin, K., Kochs, H., and Taud, R. 1978. Integer Programming Approach to the Problem of Optimal Unit Commitment with Probabilistic Reserve Determination. *IEEE Transactions on Power Apparatus and Systems*, **PAS-97**(6), 2154–2166.

Duran, M., and Grossmann, I. 1986. An Outer-Approximation Algorithm for a Class of Mixed-integer Nonlinear Programs. *Mathematical Programming*, **36**(3), 307–339. 10.1007/BF02592064.

Energy Exemplar. 2011. *http://www.energyexemplar.com/*.

Exarchakos, L., Leach, M., and Exarchakos, G. 2009. Modelling Electricity Storage Systems Management Under the Influence of Demand-Side Management Programmes. *International Journal of Energy Research*, **33**(1), 62–76.

Gollmer, R., Nowak, M., Römisch, W., and Schultz, R. 2000. Unit Commitment in Power Generation – a Basic Model and Some Extensions. *Annals of Operations Research*, **96**, 167–189.

Harvey, L.D.D. 2010. *Energy and the New Reality 1: Energy Efficiency and the Demand for Energy Services*. Earthscan.

IBM. 2010. *IBM ILOG AMPL Version 12.1 User's Guide: Standard (Command-Line) Version Including CPLEX Directives*. Incline Village, NV.

IBM. 2011. *IBM ILOG GAMS Version 12.3 User's Guide: Standard (Command-line) Version Including CPLEX Directives*. Incline village, NV.

Klotz, E. 2009 (September). *Personal communication*.

Lahdelma, R., and Ruuth, S. 1994. A Modified Benders' Decomposition Technique for Solving Large Scale Unit Commitment Problems. *Pages 737–745 of:* Henry, Jacques, and Yvon, Jean-Pierre (eds), *System Modelling and Optimization*. Lecture Notes in Control and Information Sciences, vol. 197. Springer Berlin / Heidelberg. 10.1007/BFb0035523.

Leyffer, S. 2010 (March). *Personal communication*.

Li, T., and Shahidehpour, M. 2005. Price-Based Unit Commitment: A Case of Lagrangian Relaxation Versus Mixed Integer Programming. *IEEE Transactions on Power Systems*, **20**(4), 2015–2025.

Lu, B., and Shahidehpour, M. 2005. Unit Commitment with Flexible Generating Units. *IEEE Transactions on Power Systems*, **20**(2), 1022–1034.

Luo, Z., Kumar, R., Sottile, J., and Yingling, J. 1998. An MILP Formulation for Load-Side Demand Control. *Electric Machines and Power Systems*, **26**(9), 935–949.

Ma, H., and Shahidehpour, S. 1999. Unit Commitment with Transmission Security and Voltage Constraints. *IEEE Transactions on Power Systems*, **14**(2), 757–764.

Madaeni, S., and Sioshansi, R. Forthcoming. The Impacts of Stochastic Programming and Demand Response on Wind Integration. *Energy Systems*.

Martins, A., Coelho, C., and Antunes, C. 1996. A Multiple Objective Linear Programming Approach to Power Generation Planning with Demand-Side Management (DSM). *International Transactions in Operational Research*, **3**(3-4), 305–317.

Moya, O. 2005. A Spinning Reserve, Load Shedding, and Economic Dispatch Solution by Bender's Decomposition. *IEEE Transactions on Power Systems*, **20**(1), 384–388.

Murtagh, B., Saunders, M., Murray, W., and Gill, P. 2011. *Minos Version 5.51 User's Guide*. GAMS.

Nexant. 2011. *http://www.nexant.com/default.aspx*.

Niknam, T., Khodaei, A., and Fallahi, F. 2009. A New Decomposition Approach for the Thermal Unit Commitment Problem. *Applied Energy*, **86**(9), 1667–1674.

O'Neill, R., Dautel T., and Krall, E. 2011. *Recent ISO Software Enhancements and Future Software and Modeling Plans*. Washington D.C.: Federal Energy Regulatory Commission, Office of Energy Policy and Innovation.

Ortega-Vazquez, M., Kirschen, D., and Pudjianto, D. 2006. Optimising the Scheduling of Spinning Reserve Considering the Cost of Interruptions. *IEE Proceedings-Generation, Transmission and Distribution*, **153**(5), 570–575.

Padhy, N. 2004. Unit Commitment-A Bibliographical Survey. *IEEE Transactions on Power Systems*, **19**(2), 1196–1205.

Pochet, Y., and Wolsey, L. 2006. *Production Planning by Mixed Integer Programming*. Springer Verlag.

Power Optimisation. 2011. *http://www.powerop.co.uk/*.

Rahimi, S., Niknam, T., and Fallahi, F. 2009. A New Approach Based on Benders Decomposition for Unit Commitment Problem. *World Applied Sciences Journal*, **6**(12), 1665–1672.

Rebours, Y., and Kirschen, D. 2005 (September). *What Is Spinning Reserve?* Tech. rept. Release 1. University of Manchester.

Reiss, P., and White, M. 2003. Demand and Pricing in Electricity Markets: Evidence from San Diego During California's Energy Crisis. *NBER Working Paper*, September.

Saharidis, G., Boile, M., and Theofanis, S. 2011. Initialization of the Benders Master Problem Using Valid Inequalities Applied to Fixed-Charge Network Problems. *Expert Systems with Applications*, **38**(6), 6627–6636.

Saharidis, Georgios K.D., and Ierapetritou, Marianthi G. 2010. Improving Benders Decomposition Using Maximum Feasible Subsystem (MFS) Cut Generation Strategy. *Computers & Chemical Engineering*, **34**(8), 1237–1245.

Sioshansi, R. 2010a. Evaluating the Impacts of Real-Time Pricing on the Cost and Value of Wind Generation. *IEEE Transactions on Power Systems*, **25**(2), 741–748.

Sioshansi, R. 2010b. Some Policy and Research Questions Related to Energy Storage. *EUI RSCAS, Loyola de Palacio Programme on Energy Policy*, **58**.

Sioshansi, R., and Miller, J. Forthcoming. Plug-In Hybrid Electric Vehicles Can Be Clean and Economical in Dirty Power Systems. *Energy Policy.*

Sioshansi, R., and Short, W. 2009. Evaluating the Impacts of Real-Time Pricing on the Usage of Wind Generation. *IEEE Transactions on Power Systems*, **24**(2), 516–524.

Sioshansi, R., O'Neill, R., and Oren, S. 2008. Economic Consequences of Alternative Solution Methods for Centralized Unit Commitment in Day-Ahead Electricity Markets. *IEEE Transactions on Power Systems*, **23**(2), 344–352.

Sioshansi, R., Denholm, P., Jenkin, T., and Weiss, J. 2009. Estimating the Value of Electricity Storage in PJM: Arbitrage and Some Welfare Effects. *Energy Economics*, **31**(2), 269–277.

Sioshansi, R., Fagiani, R., and Marano, V. 2010. Cost and Emissions Impacts of Plug-In Hybrid Vehicles on the Ohio Power System. *Energy Policy*, **38**(11), 6703–6712.

Takriti, S., Krasenbrink, B., and Wu, L. S.-Y. 2000. Incorporating Fuel Constraints and Electricity Spot Prices into the Stochastic Unit Commitment Problem. *Operations Research*, **48**(2), 268–280.

Tawarmalani, M., and Sahinidis, N. 2004. Global Optimization of Mixed-Integer Nonlinear Programs: A Theoretical and Computational Study. *Mathematical Programming*, **99**(3), 563–591.

Tseng, C. 1996. On Power System Generation Unit Commitment Problems. *University of California, Berkeley PhD Dissertation.*

Tuohy, A., and O'Malley, M. 2009 (July). Impact of Pumped Storage on Power Systems with Increasing Wind Penetration. *Pages 1–8 of: Power Energy Society General Meeting, 2009. PES '09. IEEE.*

Tuohy, A., Denny, E., and O'Malley, M. 2007 (July). Rolling Unit Commitment for Systems with Significant Installed Wind Capacity. *Pages 1380–1385 of: 2007 IEEE Lausanne Power Tech.*

Tuohy, A., Meibom, P., Denny, E., and O'Malley, M. 2008 (May). Benefits of Stochastic Scheduling for Power Systems with Significant Installed Wind Power. *Pages 1–7 of: PMAPS '08. Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems, 2008.*

Tuohy, A., Meibom, P., Denny, E., and O'Malley, M. 2009. Unit Commitment for Systems with Significant Wind Penetration. *IEEE Transactions on Power Systems*, **24**(2), 592–601.

US Environmental Protection Agency. 2005. *Generator, Load and CEMS Data.* http://www.epa.gov/ttn/emc/cem.html.

Van Dinter, J. 2010. Predicting Cost Savings from Demand Response Using the Unit Commitment Model. *Colorado School of Mines PhD Proposal*, May.

Van Dinter, J., Denholm, P., and Newman, A. 2011a. The Cost of Providing Spinning Reserves. *Working paper.*

Van Dinter, J., Rebennack, S., Kallrath, J., Denholm, P., and Newman, A. 2011b. The Unit Commitment Model with Concave Emissions Costs: A Hybrid Benders Decomposition with Nonconvex Master Problems. *Working paper.*

Ventyx. 2011. *http://www.ventyx.com/.*

Wang, J., Shahidehpour, M., and Li, Z. 2008. Security-Constrained Unit Commitment with Volatile Wind Power Generation. *IEEE Transactions on Power Systems*, **23**(3), 1319–1327.

Weber, C., Meibom, P., Barth, R., and Brand, H. 2009. WILMAR: A Stochastic Programming Tool to Analyze the Large-Scale Integration of Wind Energy. *Pages 437–458 of:* Pardalos, P. M., Kallrath, J., Rebennack, S., and Scheidt, M. (eds), *Optimization in the Energy Industry.* Energy Systems. Springer Berlin Heidelberg. 10.1007/978-3-540-88965-6_19.

Westerlund, T., and Lastusilta, T. 2011. *AlphaECP Version 2.04.02 User's Guide.* GAMS.

Wu, L., and Shahidehpour, M. 2010. Accelerating the Benders Decomposition for Network-Constrained Unit Commitment Problems. *Energy Systems*, **1**(3), 339–376. 10.1007/s12667-010-0015-4.

APPENDIX

CONTENTS OF CD-ROM

List of GAMS and AMPL data files

In this appendix, we list all of the data files which are contained on the CD-ROM in the pocket of this dissertation. In general, we list the files containing sets first, and then list the files containing parameters. We list the parameter files according to order in which the parameters appear in the formulation in §2.2 and §3.2. If the file contains more than one parameter, the order the file is based on the initial appearance of the first parameter. These files are text files which can be opened using a basic text editor program on a Windows, Mac, or UNIX platform.

The GAMS directory contains the following files:

- sampleDataSetsConvexExtra.dat: All data sets
- sampleDataParamConvexExtranewQ4.dat: All parameter values, except demand and maximum renewable energy
- demand1.dat through demand20.dat: Scenarios of demand and maximum renewable energy data.
- underestimator-1seg.dat and underestimator-2seg.dat: Data for the one-segment and two-segment underestimators, respectively

The AMPL-Coal directory contains the following files related to the coal-dominated data set:

- PSCO_generator_sets_710slide.dat: Sets of generators

- PSCO_fuel_curves_610slide.dat: Sets of power generators determined by number of breakpoints in heat input curves
- PSCO_combinedcyclev3.dat: Combined cycle sets
- PSCO_fuel_sets_610slide.dat: Fuel type consumed by power generator
- PSCO_Psets.dat: Sets of generators in a given state at beginning of horizon
- PSCO_emissions_setsslide.dat: Set of power generators for which US Environmental Protection Agency emissions data exist
- PSCO_fuel_type_610slide.dat: Fuel price for power generators
- PSCO_su_610slide.dat: Power generator start-up costs
- PSCO_omslide.dat: Power generator operating and maintenance costs
- PSCO_min_max_cap_610slide.dat: Power generator minimum and maximum capacities
- PSCO_mxr_710.dat: Renewable generator maximum generation
- PSCO_load.dat: Hourly load data
- PSCO_rur_rdr_610slide.dat: Power generator ramp-up and ramp-down rates
- PSCO_r_s_211slide.dat: Power generator ramp-up and ramp-down times
- PSCO_ut_dt_610slide.dat: Power generator up and down times
- PSCO_storage.dat: Storage device parameters
- PSCO_b_610slide.dat: Power generator fuel consumption at minimum capacity
- PSCO_k_m_610slide.dat: Power generator slope and intercept data
- PSCO_emissions_rates_610slide.dat: Power generator emission rates

The AMPL-Gas directory contains the following files related to the gas-dominated data set:

- PSCO_generator_sets_710slide.dat: Sets of generators
- PSCO_fuel_curves_610slide.dat: Sets of power generators determined by number of breakpoints in heat input curves
- PSCO_combinedcyclev3.dat: Combined cycle sets
- PSCO_fuel_sets_610slide.dat: Fuel type consumed by power generator
- PSCO_Psets.dat: Sets of generators in a given state at beginning of horizon
- PSCO_emissions_setsslide.dat: Set of power generators for which US Environmental Protection Agency emissions data exist
- PSCO_fuel_type_610slide.dat: Fuel price for power generators
- PSCO_su_610slide.dat: Power generator start-up costs
- PSCO_omslide.dat: Power generator operating and maintenance costs
- PSCO_min_max_cap_610slide.dat: Power generator minimum and maximum capacities
- PSCO_mxr_710.dat: Renewable generator maximum generation
- PSCO_load.dat: Hourly load data
- PSCO_rur_rdr_610slide.dat: Power generator ramp-up and ramp-down rates
- PSCO_r_s_211slide.dat: Power generator ramp-up and ramp-down times
- PSCO_ut_dt_610slide.dat: Power generator up and down times
- PSCO_storage.dat: Storage device parameters

- PSCO_b_610slide.dat: Power generator fuel consumption at minimum capacity
- PSCO_k_m_610slide.dat: Power generator slope and intercept data
- PSCO_emissions_rates_610slide.dat: Power generator emission rates