Multistage Stochastic Programming with Parametric Cost Function Approximations

Raymond Theodore Perkins III

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Operations Research and Financial Engineering
Adviser: Warren B. Powell

June 2018

Abstract

A widely used heuristic for solving stochastic optimization problems is to use a deterministic rolling horizon procedure which has been modified to handle uncertainty (e.g. buffer stocks, schedule slack). This approach has been criticized for its use of a deterministic approximation of a stochastic problem, which is the major motivation for stochastic programming. This dissertation recasts this debate by identifying both deterministic and stochastic approaches as policies for solving a stochastic base model, which may be a simulator or the real world. Stochastic lookahead models (stochastic programming) require a range of approximations to keep the problem tractable. By contrast, so-called deterministic models are actually parametrically modified cost function approximations which use parametric adjustments to the objective function and/or the constraints. These parameters are then optimized in a stochastic base model which does not require making any of the types of simplifications required by stochastic programming. This dissertation formalizes this strategy, describes a gradient-based stochastic search strategy to optimize policies, and presents a series of energy related numerical experiments to illustrate the efficacy of this approach.

Acknowledgements

I would like to express my sincere gratitude to my advisor Prof. Warren Powell for his guidance and mentorship during my doctoral studies. Warren introduced me to the amazing world of stochastic optimization and instilled in me a passion for the subject.

I would like to thank Prof. William Massey, Prof. Yongpei Guan, and Prof. Mengdi Wang for serving on my dissertation committee and their valuable feedback.

I would like to thank all of the members of CASTLE lab for their support. Thank you to Juliana, Lina, Dionysius, Tsvetan, Haitham, Stephan, Kobby, Donghun, Weidong, Joe, Yingfei, and Daniel. I also want to thank my collaborators at Air Liquide, Ian and Ajit.

I am grateful for the community of friends I have made over the past couple of years. Thank you to Brenda, Mrs. Blessing, Mr. James, Koushiki, Ezelle, Craig, Jamal, Sama, Colin, Leslie, Solomon, Dr. Horne, and Kobby. In particular, I want to thank Kobby Aboagye for being an amazing friend, big brother, and inspiration throughout this journey. I could not have done this without him.

Most importantly, I want to thank my family for their support. In particular, I want to thank my parents for their love and sacrifice over the years. I want to thank my best friend and sister, Ebony, for being the best sibling in the world. I want to thank my partner, Oby, for her endless love, patience, and support. Finally, I want to thank my grandparents: Bernice, Fanny, Theodore, and Raymond for their sacrifices and faith. This dissertation is dedicated to them.

To Bernice, Fanny, Theodore, and Raymond

Contents

Chapter 1

Introduction

There has been a long history in industry of using deterministic optimization models to make decisions that are then implemented in a stochastic setting. Energy companies use deterministic forecasts of wind, solar and loads to plan energy generation (Wallace and Fleten (2003)); airlines use deterministic estimates of flight times to schedule aircraft and crews (Lan et al. (2006)); and retailers use deterministic estimates of demands and travel times to plan inventories (Harrison and Van Mieghem (1999)). These models have been widely criticized in the research community for not accounting for uncertainty, which often motivates the use of large-scale stochastic programming models which explicitly model uncertainty in future outcomes (Mulvey et al. (1995) and Birge and Louveaux (2011a)). These large-scale stochastic programs have been applied to unit commitment (Jin et al. (2011)), hydroelectric planning (Carpentier et al. (2015)), and transportation (Lium et al. (2009)). These models use large scenario trees to approximate potential future events, but result in very large-scale optimization models that can be quite hard to solve in practice.

In this thesis, we make the case that these previous approaches ignore the true problem that is being solved, which is always stochastic. The so-called "deterministic models" used in industry are almost always parametrically modified deterministic approximations, where the modifications are designed to handle uncertainty. Both the "deterministic models" and the "stochastic models" (formulated using the framework of stochastic programming) are examples of lookahead policies to solve a stochastic optimization problem. The stochastic optimization problem is to find the best policy which is typically tested using a simulator, but may be field tested in an online environment (the real world).

We characterize these modified deterministic models as *parametric cost function approximations* which puts them into the same category as other parameterized policies that are well known in the research community working on policy search (Ng and Jordan (2000a), Peshkin et al. (2000a), Hu et al. (2007a), Deisenroth (2011), and Mannor et al. (2003)). A parallel community has evolved under the name simulation-optimization (see the recent edited volume Fu (2015)), where powerful tools have been developed based on the idea of taking derivatives of simulations (see the extensive literature on derivatives of simulations covered in Glasserman (1991), Ho (1992), Kusher, Harold; Ying (2003), Cao (2009)); a nice tutorial is given in Chau and Fu, Michael C, Huashuai Qu (2014). Much of this literature focuses on derivatives of discrete event simulations with respect to static parameters such as a buffer stock. Our strategy also exhibits static parameters, but in the form of a parameterized modifications of constraints in a policy that involves solving a linear program. This use of modified linear programs is new to the policy search literature, where "policies" are typically parametric models such as linear models ("affine policies"), structured nonlinear models (such as (s,S) policies for inventories) or neural networks.

1.1 Model Notation

Sequential, stochastic decision problems require a richer notation than standard linear programs and deterministic problems. For the sake of notational consistency, we follow the canonical model in Powell (2011) which breaks dynamic programs into five dimensions:

- The **state variable**, S_t, is all the information at time t that is necessary and sufficient to model the system from time t onward.

- A **decision**, x_t, is an n-dimensional vector that must satisfy $x_t \in \mathcal{X}_t$, which is typically a set of linear constraints. Decisions are determined by a decision function (policy) which we denote by $X_t^\pi(S_t)$, where π carries the information that determines the structure and parameters that define the function.

- The **exogenous information**, W_t, describes the information that first becomes known at time t. We let $\omega \in \Omega$ be a sample path of $W_1, \ldots, W_T$. Let $\mathcal{F}$ be the sigma algebra on Ω, and let $\mathcal{P}$ be a probability measure on $(\Omega, \mathcal{F})$, giving us a probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Next let $\mathcal{F}_t = \sigma(W_1, \ldots, W_t)$ be the sigma-algebra generated by $W_1, \ldots, W_t$, where $(\mathcal{F}_t)_{t=1}^T$ forms a filtration. The information W_t may depend on the state S_t and/or the action x_t, which means it depends on the policy. If this is the case, we write our probability space as $(\Omega^\pi, \mathcal{F}^\pi, \mathcal{P}^\pi)$, with the associated expectation operator $\mathbb{E}^\pi$.

- The **transition function**, $S^M(\cdot)$, describes how each state variable evolves over time, which we designate using

$$S_{t+1} = S^M(S_t, x_t, W_{t+1}). \tag{1.1}$$

- The **objective function** is used to evaluate the effectiveness of a policy or sequence of decisions. It minimizes the expected sum of the costs $C(S_t, x_t)$ in each time period t over a finite horizon, where we seek to find the policy that solves

$$\min_{\pi \in \Pi} \; \mathbb{E}^\pi \left[\sum_{t=0}^{T} C(S_t, X_t^\pi(S_t)) \,\middle|\, S_0 \right], \tag{1.2}$$

where $S_{t+1} = S^M(S_t, X_t^\pi(S_t), W_{t+1})$. We use $\mathbb{E}^\pi(\cdot)$ since the exogenous variables in the model may be affected by the decisions generated by our policy. Therefore we express the expectation as dependent on the policy. Since stochastic problems incorporate uncertainty in the model a variety of risk measures can be used in replacement of expectation. Equation (1.2), along with the transition function and the exogenous information process, is called the *base model*.

This canonical model can be used to model virtually *any* sequential, stochastic decision problem as long as we are using expectations instead of risk measures. We use this setting to put different policies onto a standard footing for comparison. In the next section we describe the major classes of policies that we can draw from to solve the problem. We use this framework to review the literature.

We state this canonical model because it sets up our modeling framework, which is fundamentally different than the standard paradigm of stochastic programming (for multistage problems). However, it sets the foundation for searching over policies which is fundamental to our approach.

1.2 Classes of Policies

There are two fundamental strategies for identifying policies. The first is policy search, where we search over different classes of functions $f \in \mathcal{F}$ and different parameters $\theta \in \Theta^f$ in each class (see Robbins and Monro (1951a) and Spall et al. (2003)). Policy search is written as

$$\min_{\pi=(f,\theta^f)\in(\mathcal{F},\Theta^f)} \mathbb{E}\left\{ \sum_{t=0}^{T} C(S_t, X_t^\pi(S_t|\theta^f)) \middle| S_0 \right\}. \tag{1.3}$$

Policies that can be identified using policy search come in two classes:

Policy function approximations (PFAs) These include linear or nonlinear models, neural networks, and locally parametric functions. For example a linear model, also known as an affine policy, might be written

$$X^{PFA}(S_t|\theta) = \theta_0 + \theta_1\phi_1(S_t) + \theta_2\phi_2(S_2) + \ldots$$

PFAs (using any of a wide range of approximation strategies) have been widely studied in the computer science literature under the umbrella of policy search, most commonly using parametric functions. A few examples of parametric policies are the Boltzmann exploration policy (Sutton et al. (1999)), linear decision rules (see Bertsimas and Goyal (2012), Hadjiyiannis et al. (2011), and Bertsimas et al. (2011)), and neural networks (Bengio (2009) and Levine and Abbeel (2014)). See Ng and Jordan (2000a), Peshkin et al. (2000a), Hu et al. (2007a), Deisenroth (2011), and Mannor et al. (2003) for a sample.

Cost function approximations (CFAs) Here we use parametrically modified costs and constraints that are then minimized. These are written

$$X^{CFA}(S_t|\theta) = \underset{x_t \in \mathcal{X}^\pi(\theta)}{\operatorname{argmin}} \bar{C}^\pi(S_t, x_t|\theta).$$

CFAs are widely used in industry for complex problems such as scheduling energy generation or planning supply chains, but they have not been studied formally in the research literature.

In special cases, PFAs and CFAs may produce optimal policies, although generally we are looking for the best within a class.

The second strategy is to construct policies based on lookahead models, where we capture the value of the downstream impact of a decision x_t made while in state S_t.

An optimal policy can be written

$$X_t^*(S_t) = \operatorname*{argmin}_{x_t \in \mathcal{X}_t} \left(C(S_t, x_t) + \mathbb{E}\left\{ \min_{\pi \in \Pi} \mathbb{E}\left\{ \sum_{t'=t+1}^{T} C(S_{t'}, X^\pi(S_{t'})) \big| S_{t+1} \right\} \big| S_t, x_t \right\} \right). \quad (1.4)$$

Equation (1.4) is basically Bellman's equation, but it is computable only for very special instances (Puterman (2014)). For example, to model a decision tree the policies π in (1.4) would be a lookup table expressing the action to be taken out of every decision node.

There are two core strategies for approximating the lookahead portion in (1.4):

Value function approximations (VFAs) Here we approximate the lookahead portion using a value function. Standard practice is to write the value function $V_t(S_t)$ around the (pre-decision) state S_t as

$$V_{t+1}(S_{t+1}) = \min_{\pi \in \Pi} \mathbb{E}\left\{ \sum_{t'=t+1}^{T} C(S_t, X^\pi(S_t)) \big| S_{t+1} \right\}.$$

Since we typically cannot compute $V_{t+1}(S_{t+1})$ exactly, we replace it with a value function approximation $\overline{V}_{t+1}(S_{t+1})$, in which case we would write our policy as

$$X_t^{VFA}(S_t) = \operatorname*{argmax}_{x_t \in \mathcal{X}_t} \big(C(S_t, x_t) + \mathbb{E}\{\overline{V}_{t+1}(S_{t+1}) | S_t\} \big).$$

Often it is easier to use the post-decision state S_t^x (the state immediately after a decision has been made) which captures the entire lookahead term in equation (1.4). This allows us to write our policy without the imbedded expectation

$$X_t^{VFA}(S_t) = \operatorname*{argmax}_{x_t \in \mathcal{X}_t} \big(C(S_t, x_t) + \overline{V}_t^x(S_t^x) \big).$$

Eliminating the expectation opens the door to solving problems where x_t is high-dimensional (but only if $\overline{V}_t^x(S_t^x)$ is concave). Value function approximations have been widely studied under the umbrellas of approximate dynamic programming (see Powell (2011), and Bertsekas (2010)) and reinforcement learning (Sutton and Barto, 1998). Specialized methods have evolved for handling convex problems that arise in multistage linear programming such as stochastic dual dynamic programming (SDDP) (see Pereira and Pinto (1991), Shapiro et al. (2009), Shapiro (2011), and Philpott and Guan (2008)) or piecewise linear, separable value functions (Powell et al. (2004), Topaloglu and Powell (2006)).

Direct lookahead approximations (DLAs) When the lookahead problem cannot be reasonably approximated by a value function, it is often necessary to turn to a direct lookahead approximation, where we replace the model with an approximation for the purpose of approximating the future. In this case our policy (1.4) can be written

$$X_t^{DLA}(S_t) = \operatorname*{argmin}_{x_t \in \mathcal{X}_t} \left(C(S_t, x_t) + \tilde{\mathbb{E}} \left\{ \min_{\tilde{\pi} \in \tilde{\Pi}} \tilde{\mathbb{E}} \left\{ \sum_{t'=t+1}^{t+H} \tilde{C}(\tilde{S}_{tt'}, \tilde{X}^{\tilde{\pi}}(\tilde{S}_{tt'})) | \tilde{S}_{t,t+1} \right\} | S_t, x_t \right\} \right). \tag{1.5}$$

Here, all variables (states and decisions) in the lookahead model are indicated with tilde's, and are indexed by t (the time at which the lookahead model is instantiated) and t' (the time period within the lookahead horizon). Lookahead models are typically characterized by five types of approximations: 1) the horizon, 2) the staging of information and decisions (multistage problems may be approximated by two-stages), 3) the outcome space (we may use a deterministic lookahead or a sampled stochastic), 4) discretization (of states, actions, and time periods), and 5) holding some information static that varies in the base model (a common assumption is to hold a forecast constant within the lookahead model). Lookahead models can take a variety of forms: deterministic lookahead models, also referred to as rolling horizon procedures (Sethi and Sorger, 1991) or model predictive control (Camacho and Alba, 2013), decision trees (which can be approximated using Monte Carlo tree search) for discrete actions, or stochastic programming models using scenario trees (see Birge and Louveaux (2011b) and Donohue and Birge (2006)).

Policy search, whether we are using PFAs or CFAs, requires tuning parameters in our base objective function (1.2). By contrast, policies based on lookahead approximations depend on developing the best approximation of the future that can be handled computationally, although these still need to be evaluated using (1.2).

1.3 Thesis Outline

The main theme of this thesis is introducing and formalizing the class of decision making polices known as parametric cost function approximations (CFA) which use deterministic optimization problems that have been parametrically modified to account for uncertainty. This thesis is organized as follows. In chapter 2, we introduce and develop the idea of parameterized cost function approximations as a tool for solving important classes of stochastic optimization problems. Then we show the approach is computationally comparable to solving deterministic approximations, with the exception that the parametric modifications have to be optimized, typically in a simulator that avoids the many approximations made in stochastic lookahead models. We derive the policy gradients for parameterized right-hand sides using the properties of the underlying linear program and introduce a gradient-based policy search algorithm for determining parameter values. Finally, we illustrate different styles of parametric approximations using the context of a nonstationary energy inventory problem, and quantify the benefits relative to a basic deterministic lookahead without adjustments.

In chapter 3, we apply the CFA to the difficult problem of making lagged commitments while managing a portfolio of energy resources including steam and gas turbine generators. This particular problem requires making commitments several hours and days in the advance. This chapter provides a proper base model of a stochastic, lagged resource allocation problem in the context of energy portfolio management. Additionally, we introduce a family of parameterizations of a deterministic lookahead model designed to produce robust policies that respond in a realistic way to the level of uncertainty in forecasts. This chapter also provides a set of sufficient conditions to prove the convergence of our data-driven gradient based search algorithm. Finally, we demonstrate empirically that our method produces high quality solutions relative to unmodified deterministic lookahead policies on a library of lagged energy portfolio problems.

In chapter 4, we apply the CFA to the complex problem of managing a network of industrial gas production plants, a hydrogen storage cavern, a diverse set of customers, and access to electricity and natural gas commodity markets. We formally describe the problem and propose a parametrically modified operating policy. We then presents a series of experiments to demonstrate the use of our model, the efficacy of our gradient-based search algorithm, and analyze the performance of solutions under varying operating environments.

1.4 Publications and Presentation

The material in this thesis has been presented at various conferences: INFORMS Annual Meeting, Federal Energy Regulatory Commission (FERC)'s Technical Conference on Increasing Real-Time and Day-Ahead Market Efficiency through Improved Software, International Conference of Stochastic Programming, and INFORMS Optimization Society Conference.

Chapter 2

Parametric Cost Function Approximations

2.1 Introduction

We consider the idea, used for years in industry, that an effective way to solve complex stochastic optimization problems is to shift the modeling of the stochastics from a lookahead approximation, where even deterministic lookahead models can be hard to solve, to the stochastic base model, which typically is implemented as a simulator but which might also be the real world. Tuning a model in a stochastic simulator makes it possible to handle arbitrarily complex dynamics, avoiding the many approximations (such as two-stage models, exogenous information that is independent of decisions) that are standard in stochastic programming.

The parametric cost function approximation (CFA) is conceptually rooted in this idea. This class of policies opens up a fundamentally new approach for providing practical tools for solving high-dimensional, stochastic programming problems. It provides an alternative to classical stochastic programming with its focus on optimizing a stochastic lookahead model which requires a variety of approximations to make it computationally tractable. The parametric CFA makes it possible to incorporate problem structure, such as the recognition that robust solutions can be achieved using standard methods such as schedule slack and/or buffer stocks. Furthermore, the parametric CFA makes it possible to incorporate problem structure for handling uncertainty. Some examples include:

- Air freight companies plan for equipment problems by maintaining spare aircraft at different locations around the country.
- Hospitals can handle uncertainty in blood donations and the demand for blood by maintaining supplies of O-minus blood, which can be used by anyone.
- Grid operators handle uncertainty in generator failures, as well as uncertainty in energy from wind and solar, by requiring generating reserves.

Central to this approach is the ability to manage uncertainty by recognizing effective strategies for responding to unexpected events. We argue that this structure is apparent in many settings, especially in complex resource allocation problems. Parametric cost function approximations make it possible to exploit these structural properties. For example, it may be obvious that the way to handle uncertainty when planning energy generators in a unit commitment problem is to require extra reserves at all times of the day. A stochastic programming model encourages this behavior, but the requirement for a manageable number of scenarios will produce the required reserve only when one of the scenarios requires it. Imposing a reserve constraint (which is a kind of cost function approximation) allows us to impose this requirement at all times of the day, and to tune this requirement under very realistic conditions. At a minimum, we offer that our approach represents an interesting, and very practical, alternative to stochastic programming.

Designing a parametric cost function approximation closely parallels the design of any parametric statistical model, which is part art (creating the model) and part science (fitting the model). To illustrate the process of designing a parametric cost function approximation, we use the setting of a time-dependent stochastic inventory planning problem that arises in the context of energy storage, but could represent any inventory planning setting. We assume we have access to rolling forecasts where forecast errors are based on careful modeling of actual and predicted values for energy loads, generation from renewable sources, and prices. The combination of the time-dependent nature and the availability of rolling forecasts which are updated each time period make this problem a natural setting for lookahead models, where the challenge is how to handle uncertainty. We have selected this problem since it is relatively small, simplifying the extensive computational work. However, our methodology is scalable to any problem setting which is currently being solved using a deterministic model.

This chapter makes the following contributions. 1) We introduce and develop the idea of parameterized cost function approximations as a tool for solving important classes of stochastic optimization problems, shifting the focus from solving complex, stochastic lookahead models to optimizing a stochastic base model. This approach is computationally comparable to solving deterministic approximations, with the exception that the parametric modifications have to be optimized, typically in a simulator that avoids the many approximations made in stochastic lookahead models. 2) We derive the policy gradients for parameterized right-hand sides using the properties of the underlying linear program. 3) We illustrate different styles of parametric approximations using the context of a nonstationary energy inventory problem, and quantify the benefits over a basic deterministic lookahead without adjustments.

The chapter is organized as follows. Section 2.2 introduces the basic concept and alternative structures of the CFA. Section 2.3 describes the derivation of the gradient of the base model with respect to the policy parameters. Section 2.3 describes gradient-based stochastic search strategy to optimize our parameterized policies. Section 2.4 formally describes a time-dependent stochastic energy storage problem. Finally, section 2.5 presents a series of numerical results.

2.2 The Parametric Cost Function Approximation

We extend the concept of policy search to include parameterized optimization problems. The parametric Cost Function Approximation (CFA) draws on the structural simplicity of deterministic lookahead models and myopic policies, but allows more flexibility by adding tunable parameters. This puts this methodology in the same class as parametric policy function approximations widely used in the policy search literature, with the only difference that our parameterized functions are inside an optimization problem, making them more useful for high dimensional problems.

Basic Idea

Since the idea of a parametric cost function approximation is new, we begin by outlining the general strategy, and then demonstrate how to apply it for our energy storage problem. We propose using parameterized optimization problems such as

$$X_t^{\pi}(S_t|\theta) = \underset{x_t \in \mathcal{X}^{\pi}(\theta)}{\operatorname{argmin}} \left\{ C(S_t, x_t) + \sum_{f \in \mathcal{F}} \theta_f^c \phi_f(S_t, x_t) : A_t x_t = \bar{b}_t^{\pi}(\theta^b) \right\} \tag{2.1}$$

as a type of parameterized policy. Here the index π signifies the structure of the modified set of constraints, θ^c is the vector of cost function parameters, θ^b is the vector of constraint parameters, and ϕ_f are the basis functions corresponding to features $f \in \mathcal{F}$.

Parametric terms can be added to the cost function or constraints of a myopic or deterministic lookahead model. In the following example, parameters have been added as an error correction term to the objective function as well as to the model constraints, giving us

$$X_t^\pi(S_t|\theta) = \underset{x_t \in \mathcal{X}_t}{\operatorname{argmin}} \left(C(S_t, x_t) + \sum_f \theta_f^c \phi_f(S_t, x_t) \right) \tag{2.2}$$

subject to

$$A_t x_t = b_t + D\theta^b,$$

where D is a scaling matrix. We emphasize that the cost correction term should not be confused as a value function approximation, because we make no attempt to approximate the downstream value of being in a state.

Whether the parameterizations are in the objective function, or in the constraints, the specification of a parametric CFA parallels the specification of any statistical model (or policy). The structure of the model is the "art" that draws on the knowledge and insights of the modeler. Finding the best CFA, which involves finding the best θ, is the science which draws on the power of classical search algorithms.

A hybrid Lookahead-CFA policy

There are many problems that naturally lend themselves to a lookahead policy (for example, to incorporate a forecast or to produce a plan over time), but where there is interest in making the policy more robust than a pure deterministic lookahead using point forecasts. For this important class (which is the problem we face), we can create a hybrid policy where a deterministic lookahead has parametric modifications that have to be tuned using policy search. When parameters are applied to the constraints it is possible to incorporate easily recognizable problem structure. For example, a supply chain management problem can handle uncertainty through buffer stocks, while an airline scheduling model might handle stochastic delays using schedule slack. A grid operator planning energy generation in the future might schedule reserve capacity to account for uncertainty in forecasts of demand, as well as energy from wind and solar. As with all policy search procedures, there is no guarantee that the resulting policy will be optimal unless the parameterized space of policies includes the optimal policy. However, we can find the optimal policy within the parameterized class, which may reflect operational limitations. We note that while parametric cost function approximations are widely used in industry, optimizing within the parametric class is not done.

Structure of the cost function approximation

Parametric terms can be appended to existing constraints, and new parameterized constraints can be added to the existing model. Often the problem setting will influence how policy constraints should be parameterized. Consider the energy storage problem where a manager must satisfy the power demand of a building. The manager has a stochastic supply of renewable energy at no cost, an unlimited supply from the power grid at a stochastic price, and access to a local rechargeable storage device. Every period the manager must determine what combination of energy sources to use to satisfy the power demand, how much energy to store, and how much to sell back to the grid. Given the manager has access to point forecasts of future exogenous information he or she can use the following lookahead policy to determine how to allocate their energy,

$$X_t^{\text{D-LA}}(S_t) = \underset{x_t,(\tilde{x}_{tt'},t'=t+1,\ldots,t+H)}{\operatorname{argmin}} \left(C(S_t, x_t) + \left[\sum_{t'=t+1}^{t+H} \tilde{c}_{tt'} \tilde{x}_{tt'} \right] \right) \tag{2.3}$$

where $\tilde{S}_{t,t'+1} = \tilde{S}^M(\tilde{S}_{tt'}, \tilde{x}_{tt'}, \widetilde{W}_{t,t'+1})$ and H is the size of the lookahead horizon. Here, all variables (states and decisions) in the lookahead model are indicated with tilde's, and are indexed by t (the time at which the lookahead model is instantiated) and t' (the time period within the lookahead horizon). It is important to note that if the contribution function, transition function, and constraints of $X^\pi(\cdot)$ are linear, this policy can be expressed as the following linear program

$$X_t^{\text{D-LA}}(S_t) = \underset{x_t, \tilde{x}_{tt'}, t'=t+1,t+H}{\text{argmin}} c_t x_t + \tilde{c}_t \tilde{x}_t : A_t x_t \leq b_t,\ \tilde{A}_t x_t \leq \tilde{b}_t,\ \ x_t \geq 0\}, \tag{2.4}$$

where $\tilde{c}_t = \{\tilde{c}_{tt'} : t' = t+1, ..., t+H\}$, $\tilde{A}_t = \{\tilde{A}_{tt'} : t' = t+1, ..., t+H\}$, and $\tilde{b}_t = \{\tilde{b}_{tt'} : t' = t+1, ..., t+H\}$. There are different ways to parameterize the previous policy, but since all the uncertainty in our problem is restricted to the right hand side constraints, we will only parameterize the vector b_t. Once parameterized our policy becomes

$$X_t^{\text{LA-CFA}}(S_t|\theta) = \underset{x_t, \tilde{x}_{tt'}, t'=t+1,t+H}{\text{argmin}} c_t x_t + \tilde{c}_t \tilde{x}_t : A_t x_t \leq b_t(\theta),\ \tilde{A}_t x_t \leq \tilde{b}_t(\theta),\ x_t \geq 0\} \tag{2.5}$$

where θ is a vector of tunable parameters. Parametric modifications can be designed specifically to capture the structure of a particular policy. The idea to use buffers and inventory constraints to manage storage is intuitive and easily incorporated into a deterministic lookahead. In the previous energy storage problem a lower buffer guarantees the decision maker will always have access to some stored energy. Conversely, an upper threshold will make sure some storage space remains in the battery in order to capitalize on unexpected gusts of wind (for example). For the energy storage problem, we represent the amount of energy stored in the battery as the variable R_t and the approximated future amount of energy in storage at time t' given the information available at time t as $R_{tt'}$. Thus,

$$\theta^L \leq R_{tt'} \leq \theta^U \text{ for } t' > t. \tag{2.6}$$

Although it can greatly increase the parameter space, the upper and lower bounds can can also depend on $(t'-t)$, as in

$$\theta^L_{t'-t} \leq R_{tt'} \leq \theta^U_{t'-t} \text{ for } t' > t. \tag{2.7}$$

The resulting modified deterministic problem is no harder to solve than the original deterministic problem (where $\theta^L = 0$ and $\theta^U = R^{max}$). We now have to use policy search techniques to optimize θ. Below we suggest different ways of parameterizing the right hand side adjustment.

Policy parameterizations come in a variety of forms. A simple form is a lookup table indexed by time as in equation (2.7). Although it may be simple, a lookup table model for θ means that the dimensionality increases with the horizon which can complicate the policy search process.

In the energy storage example, $f^E_{tt'}$ represents the forecast of the amount of renewable energy available at time t' given the information available at time t. A policy maker may use the parametrization $\theta_{t'-t} \cdot f^E_{tt'}$ to intentionally overestimate or underestimate the amount of future renewable energy. The policy maker may set $\theta_{t'-t} \leq 1$ to make the policy more robust and avoid the risk of running out of energy.

This type of parameterization is not limited to just modifying the point forecast of exogenous information. If the modeler has sufficient information such as the cumulative distribution function of $E_{t'}$, $F^E_{t'}(\cdot)$, he or she may even exchange the point forecast $f^E_{tt'}$ with the quantile function

$$Q^E_{t'}(\theta_{t'-t}) = \inf\left\{w \in \mathbb{R} : \theta_{t'-t} \leq F^{E_{tt'}}(w)\right\}. \tag{2.8}$$

In this case $\theta_{t'-t}$ is still a parameter of the policy and determines how aggressively or passively the policy stores energy. The lookup table in time parameterization is best if the relationship between parameters in different periods is unknown.

Instead of having an adjustment $\theta_\tau = \theta_{t'-t}$ for each time $t+\tau$ in the future, we can use instead a parametric function of τ, which reduces the number of parameters that we have to estimate. For example, we might use the parametric adjustment given by

$$\theta^L \cdot e^{\alpha\tau} \leq R_{tt'} \leq \theta^U \cdot e^{\beta\tau} \text{ for } t' > t \text{ and } \alpha, \beta \in \mathbb{R}. \tag{2.9}$$

These parametric functions of time can also be used to directly modify the forecasts in the lookahead model. For example, in the energy storage example, the policy maker may use the parameterization $f^E_{tt'} \cdot \theta_1 e^{\theta_2 \cdot (t'-t)}$ to replace the forecasted amount of future renewable energy, $f^E_{tt'}$.

2.3 The CFA gradient algorithm

Policy Function Approximations (PFAs) and Cost Function Approximations (CFAs) are structurally different in the sense that one uses analytic functions and the other uses parameterized optimization problems, respectively, to make decisions. However, they are both subclasses of the same general class of parameterized policies, $X^\pi_t : \mathcal{S}_t \times \Theta \to \mathcal{X}_t$, and their optimal parameterization, θ^*, can be found by solving

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmin}} \; F(\theta), \tag{2.10}$$

where

$$F(\theta) = \mathbb{E}\left[\sum_{t=0}^{T} C(S_t, X^\pi_t(S_t|\theta)) \,\Big|\, S_0\right] \tag{2.11}$$

and $S_{t+1} = S^M(S_t, X^\pi_t(S_t|\theta), W_{t+1})$. If $F(\cdot)$ is well defined, finite valued, convex, and continuous at every θ in the nonempty, closed, bounded, and convex set $\Theta \subset \mathbb{R}^n$, then an optimal $\theta^* \in \Theta$ exists. We can use a stochastic gradient algorithm to search for θ^*, (see Robbins and Monro (1951b)) given W_t is a stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and there exists a stochastic subgradient, $g^n \in \partial_\theta F(\theta^{n-1})$ that satisfies the following assumptions:

A1) $\mathbb{E}\left[g^{n+1} \cdot (\bar{\theta}^n - \theta^*)\middle|\mathcal{F}^n\right] \geq 0,$

A2) $|g^n| \leq B_g,$

A3) For any θ where $|\theta - \theta^*| > \delta, \delta > 0$, there exists $\epsilon > 0$ such that $\mathbb{E}[g^{n+1}|\mathcal{F}^n] > \epsilon$.

This is true regardless of whether the policy $X_t^\pi(\cdot|\theta)$ is a CFA or PFA. There are several ways to generate stochastic subgradients that satisfy the previous conditions. If the cumulative reward of a single sample path, $\bar{F}(\cdot,\omega)$, is convex and differentiable for every $\omega \in \Omega$ and θ is an interior point of Θ, then the gradient, $\nabla_\theta \bar{F}$, of

$$\bar{F}(\theta,\omega) = \sum_{t=0}^{T} C\left(S_t(\omega), X_t^\pi(S_t(\omega)|\theta)\middle|\theta\right) \tag{2.12}$$

where $S_{t+1}(\omega) = S^M(S_t(\omega), X_t^\pi(S_t(\omega)), W_{t+1}(\omega))$, satisfies the conditions required for the previously mentioned iterative algorithm (see Strassen (1964)). This subgradient can be calculated recursively and is described in the following proposition.

Proposition 1. *Assume $\bar{F}(\cdot,\omega)$ is convex for every $\omega \in \Omega$, θ is an interior point of Θ, and $F(\cdot)$ is finite valued in the neighborhood of θ, then*

$$\nabla_\theta F(\theta) = \mathbb{E}[\nabla_\theta \bar{F}(\theta,\omega)]$$

where

$$\nabla_\theta \bar{F} = \left(\frac{\partial C_0}{\partial X_0}\cdot\frac{\partial X_0}{\partial \theta}\right) + \sum_{t'=1}^{T}\left[\left(\frac{\partial C_{t'}}{\partial S_{t'}}\cdot\frac{\partial S_{t'}}{\partial \theta}\right) + \left(\frac{\partial C_{t'}}{\partial X_{t'}(S_t|\theta)}\cdot\left(\frac{\partial X_{t'}(S_t|\theta)}{\partial S_{t'}}\cdot\frac{\partial S_{t'}}{\partial \theta} + \frac{\partial X_{t'}(S_t|\theta)}{\partial \theta}\right)\right)\right], \tag{2.13}$$

and

$$\frac{\partial S_{t'}}{\partial \theta} = \frac{\partial S_{t'}}{\partial S_{t'-1}} \cdot \frac{\partial S_{t'-1}}{\partial \theta} + \frac{\partial S_{t'}}{\partial X_{t'-1}(S_{t-1}|\theta)} \cdot \left[\frac{\partial X_{t;-1}(S_{t-1}|\theta)}{\partial S_{t'-1}} \cdot \frac{\partial S_{t'-1}}{\partial \theta} + \frac{\partial X_{t'-1}(S_{t-1}|\theta)}{\partial \theta}\right].$$

Proof. If $\bar{F}(\cdot, \omega)$ is convex for every $\omega \in \Omega$, θ is an interior point of Θ, and $F(\cdot)$ is finite valued in the neighborhood of θ, then by theorem 7.47 of Shapiro et al. (2009)

$$\nabla_\theta \mathbb{E} F(\theta, W) = \mathbb{E} \nabla_\theta F(\theta, W).$$

Applying the chain rule, we find

$$\begin{aligned}
\nabla_\theta \bar{F} &= \frac{\partial}{\partial \theta}\left[C_0(S_0, X_0(S_0|\theta)) + \sum_{t'=1}^{T} C(S_{t'}, X_{t'}(S_{t'}|\theta))\right] \\
&= \frac{\partial}{\partial \theta} C_0(S_0, X_0(S_0|\theta)) + \frac{\partial}{\partial \theta}\left[\sum_{t'=1}^{T} C(S_{t'}, X_{t'}(S_{t'}|\theta))\right] \\
&= \left(\frac{\partial C_0}{\partial X_0} \cdot \frac{\partial X_0}{\partial \theta}\right) + \left[\sum_{t'=1}^{T} \frac{\partial}{\partial \theta} C(S_{t'}, X_{t'}(S_{t'}|\theta))\right] \\
&= \left(\frac{\partial C_0}{\partial X_0} \cdot \frac{\partial X_0}{\partial \theta}\right) + \sum_{t'=1}^{T} \left[\left(\frac{\partial C_{t'}}{\partial S_{t'}} \cdot \frac{\partial S_{t'}}{\partial \theta}\right) + \left(\frac{\partial C_{t'}}{\partial X_{t'}(S_{t'}|\theta)} \cdot \frac{\partial X_{t'}(S_{t'}|\theta)}{\partial \theta}\right)\right] \\
&= \left(\frac{\partial C_0}{\partial X_0} \cdot \frac{\partial X_0}{\partial \theta}\right) + \sum_{t'=1}^{T} \left[\left(\frac{\partial C_{t'}}{\partial S_{t'}} \cdot \frac{\partial S_{t'}}{\partial \theta}\right) + \left(\frac{\partial C_{t'}}{\partial X_{t'}(S_{t'}|\theta)} \cdot \left(\frac{\partial X_{t'}(S_{t'}|\theta)}{\partial S_{t'}} \cdot \frac{\partial S_{t'}}{\partial \theta} + \frac{\partial X_{t'}(S_{t'}|\theta)}{\partial \theta}\right)\right)\right],
\end{aligned}$$

where

$$\frac{\partial S_{t'}}{\partial \theta} = \frac{\partial S_{t'}}{\partial S_{t'-1}} \cdot \frac{\partial S_{t'-1}}{\partial \theta} + \frac{\partial S_{t'}}{\partial X_{t'-1}(S_{t-1}|\theta)} \cdot \left[\frac{\partial X_{t;-1}(S_{t-1}|\theta)}{\partial S_{t'-1}} \cdot \frac{\partial S_{t'-1}}{\partial \theta} + \frac{\partial X_{t'-1}(S_{t-1}|\theta)}{\partial \theta}\right].$$

■

Computing the gradient of linear cumulative reward

If the objective function in equation (3.37) is a linear function of the decisions, x_t, the parametric CFA policy, $X^\pi_t(S_t|\theta)$, which determines the decision, x_t, can be written as the following linear program

$$X_t^\pi(S_t|\theta) = \underset{x_t, (\tilde{x}_{tt'}), t'=t+1,\dots,t+H}{\operatorname{argmin}} \quad c_t x_t + \sum_{t'=t+1}^{t+H} \tilde{c}_{tt'} \tilde{x}_{tt'},$$

where $\tilde{A}_t \tilde{x}_t \leq \tilde{b}_t(\theta, S_t)$ and $\tilde{x}_t^T = [\tilde{x}_{t,t}^T, ..., \tilde{x}_{t,T}^T]$. The state variable, S_t, includes the point estimates, $(\widetilde{W}_{tt'})_{t'=t+1,\dots,t+H}$, that are used to approximate exogenous information. If this policy is written as a linear program where the state and approximated exogenous information is only in the right hand side constraints, $\tilde{b}_t(\theta, S_t)$, then a subgradient can be calculated recursively and is described in the following proposition:

Proposition 1. *Given $\bar{F}(\theta, \omega)$ is concave in θ for every $\omega \in \Omega$, θ is an interior point of Θ, and the contribution function $C(x)$ is a linear function of x, the transition function $S_t = S^M(S_{t-1}, x_{t-1}, W_t)$ is linear in S_{t-1} and x_{t-1}, $F(\cdot)$ is finite valued in the neighborhood of θ, and the policy, $X_t^\pi(S_t|\theta)$ is defined as*

$$X_t^\pi(S_t|\theta) = \underset{x_t, (\tilde{x}_{tt'}), t'=t+1,\dots,t+H}{\operatorname{argmin}} \quad c_t x_t + \sum_{t'=t+1}^{t+H} \tilde{c}_{tt'} \tilde{x}_{tt'} \tag{2.14}$$

where $\tilde{A}_t \tilde{x}_t \leq \tilde{b}_t(\theta, S_t)$, B_t is the basis matrix corresponding to the basic variables for the optimal solution of (3.40), and $\tilde{x}_t^T = [\tilde{x}_{tt}, \dots, \tilde{x}_{tT}]$. Then

$$\nabla_\theta F(\theta) = \mathbb{E}[\nabla_\theta \bar{F}(\theta, \omega)]$$

where

$$\nabla_\theta \bar{F}(\theta, \omega) = \sum_{t=1}^{T} \left(\nabla_\theta \tilde{b}_t(\theta, S_t) + \nabla_{S_t} \tilde{b}_t(\theta, S_t) \cdot \nabla_\theta S_t \right)^T \cdot \left(B_t^{-1} \right)^T \cdot c_t, \tag{2.15}$$

$$\nabla_\theta S_t = \nabla_{S_{t-1}} S^M(S_{t-1}, x_{t-1}, W_t) \cdot \nabla_\theta S_{t-1} + \nabla_{x_{t-1}} S^M(S_{t-1}, x_{t-1}, W_t) \cdot \nabla_\theta x_{t-1}. \tag{2.16}$$

Proof. If $\bar{F}(\cdot,\omega)$ is convex for every $\omega \in \Omega$, θ is an interior point of Θ, and $F(\cdot)$ is finite valued in the neighborhood of θ then we define our policy as equation (3.40). If the contribution function $C(x)$ is a linear function of x, the transition function, $S_t = S^M(S_{t-1}, x_{t-1}, W_t)$, is linear, and the policy, $X_t^\pi(S_t|\theta)$ is defined as

$$X_t^\pi(S_t|\theta) = \underset{x_t}{\operatorname{argmax}} \left(c_t x_t + \max_{\tilde{x}_{tt'}, t'=t+1,\ldots,t+H} \sum_{t'=t+1}^{t+H} \tilde{c}_{tt'} \tilde{x}_{tt'} \right)$$

where $\tilde{A}_t \tilde{x}_t \leq \tilde{b}_t(S_t, \theta)$ and $\tilde{x}_t^T = [\tilde{x}_{tt}, \ldots, \tilde{x}_{t,T}]$. Then

$$\begin{aligned}
\nabla_\theta \bar{F}(\theta,\omega) &= \nabla_\theta \left[\sum_{t=0}^{T} c_t^T x_t(S_t(\omega)|\theta) \right] \\
&= \sum_{t=0}^{T} \left[\nabla_\theta \left(c_t^T x_t(S_t(\omega)|\theta) \right) \right] \\
&= \sum_{t=0}^{T} \left[\nabla_\theta x_t(S_t(\omega)|\theta)^T \right] \cdot c_t \\
&= \sum_{t=0}^{T} \nabla_\theta \left[B_t^{-1} \cdot \tilde{b}_t(\theta, S_t(\omega)) \right]^T \cdot c_t \\
&= \sum_{t=0}^{T} \nabla_\theta \left[\tilde{b}_t(\theta, S_t(\omega)) \cdot (B_t^{-1})^T \right] \cdot c_t \\
&= \sum_{t=0}^{T} \left(\nabla_\theta \tilde{b}_t(\theta, S_t(\omega))^T \right) \cdot \left(B_t^{-1} \right)^T \cdot c_t \\
&= \sum_{t=0}^{T} \left(\nabla_\theta \tilde{b}_t(\theta, S_t(\omega)) + \nabla_{S_t(\omega)} \tilde{b}_t(\theta, S_t(\omega)) \cdot \nabla_\theta S_t(\omega) \right)^T \cdot \left(B_t^{-1} \right)^T \cdot c_t.
\end{aligned} \tag{2.17}$$

■

The algorithm

The ability to calculate an estimator of $\nabla_\theta F(\theta)$ allows us to use stochastic approximation techniques to determine the optimal parameters, θ, of the CFA policy, $X_t^\pi(\cdot|\theta)$. Below is the iterative algorithm we use to tune our CFA policies

Algorithm 1 CFA Gradient Algorithm

1: Initialize θ^0, N, and k:

2: **for** n = 1, 2, 3, ..., N **do**

3: Generate a trajectory ω^n where

$$S_{t+1}^n(\omega^n) = S^M(S_t^n(\omega^n), X_t^\pi(S_t^n(\omega^n)|\theta^{n-1}), W_{t+1}(\omega^n))$$

4: Compute the gradient estimator using equation (2.13)

5: Update policy parameters, θ

$$\theta^n = \theta^{n-1} + \alpha_{n-1}\nabla_\theta \bar{F}(\theta^{n-1}, \omega^n)|_{\theta=\theta^{n-1}} \tag{2.18}$$

where the stepsizes α_n satisfy conditions

B1) $\alpha_n > 0$, a.s.

B2) $\sum_{n=0}^\infty \alpha_n = \infty$, a.s.

B3) $\mathbb{E}\left[\sum_{n=0}^\infty (\alpha_n)^2\right] < \infty$.

If $F(\cdot)$ is continuous and finite valued in the neighborhood of every θ, in the nonempty, closed, bounded, and convex set $\Theta \subset \mathbb{R}^n$ such that $\bar{F}(\cdot, \omega)$ is convex for every $\omega \in \Omega$ where θ is an interior point of Θ, then

$$\lim_{n\to\infty} \theta^n \longrightarrow \theta^* \text{ a.s.}$$

Although any stepsize rule that satisfies the previous conditions will guarantee asymptotic convergence, we prefer parameterized rules that can be tuned for quicker convergence rates. Therefore, we limit our evaluation of the algorithm to how well it does within N iterations. The CFA Algorithm can be described as a policy, $\theta^{\pi}(S^n)$, with a state variable, $S^n = \theta^n$ plus any parameters needed to compute the stepsize policy, and where π describes the structure of the stepsize rule. If $\theta^{\pi,n}$ is the estimate of θ using stepsize rule π after n iterations, then our goal is to find the rule that produces the best performance (in expectation) after we have exhausted our budget of N iterations. Thus, we wish to solve

$$\min_{\pi} \mathbb{E}\bar{F}(\theta^{\pi,N}, W). \tag{2.19}$$

We wish to find the best stepsize rule that maximizes terminal value within N iterations. For our numerical example we use the adaptive gradient algorithm, ADAGRAD, as our step size rule (Duchi et al. (2011a)). ADAGRAD modifies the individual step size for the updated parameter, θ, based on previously observed gradients using

$$\alpha_n = \frac{\eta}{\sqrt{G_t + \epsilon}} \tag{2.20}$$

where η is a scalar learning rate, $G \in \mathbb{R}^{d \times d}$ is a diagonal matrix where each diagonal element is the sum of the squares of the gradients with respect to θ up to the current iteration n, while ϵ is a smoothing term that avoids division by zero. For our simulations we set $\eta = 0.1$.

2.4 An Energy Storage Application

We use the setting of an energy storage application to show how we can use a parametric CFA to produce robust policies using rolling forecasts of varying quality. Our problem is designed to test the capabilities of the algorithm, rather than representing an accurate model of a specific energy storage application.

In our setting a smart grid manager must satisfy a recurring power demand with a stochastic supply of renewable energy, limited supply of energy from the main power grid at a stochastic price, and access to a local rechargeable storage devices. This system is graphically represented in Figure 2.1.

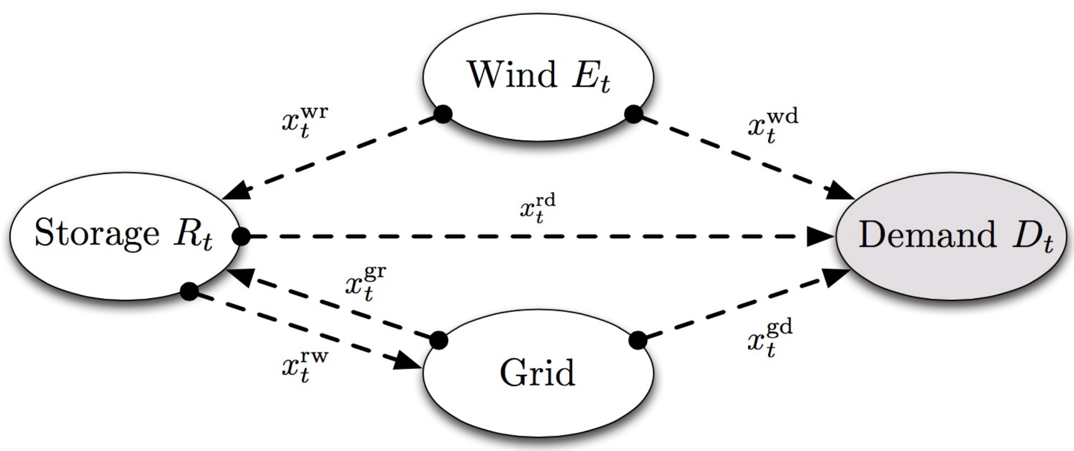


Figure 2.1: Energy system schematics

Every hour the manager must determine what combination of energy sources to use to satisfy the power demand, how much energy to store, and how much to sell back to the grid. The state variable at time t, S_t, includes the level of energy in storage, R_t, the amount of energy available from wind, E_t, the spot price of electricity, P_t, the demand D_t, and the energy available from the grid G_t at time t. The state of the system can be represented by the following five dimensional vector,

$$S_t = (R_t, E_t, P_t, D_t, G_t) \tag{2.21}$$

where $R_t \in [0, R_{\max}]$ is the level of energy in storage at time t. The demand, D_t, has a deterministic seasonal structure

$$D_t = \lfloor \max\{0, 100 - 50 \sin\left(\frac{5\pi t}{T}\right)\} \rfloor. \tag{2.22}$$

At the beginning of every period t the manager must combine energy from the following sources to satisfy the demand, D_t:

1. Energy currently in storage (represented by a decision x_t^{rd});
2. Newly available wind energy (represented by a decision x_t^{wd});
3. And energy from the grid (represented by a decision x_t^{gd}).

Additionally, the manager must decide how much renewable energy to store, x_t^{wr}, how much energy to sell to the grid at price P_t, x_t^{rg}, and how much energy to buy from the grid and store, x_t^{gr}. The manager's decision is defined as the following vector

$$x_t = (x_t^{wd}, x_t^{gd}, x^{rd}, x_t^{wr}, x_t^{gr}, x_t^{rg})^T \geq 0 \tag{2.23}$$

given the following constraints:

$$\begin{aligned}
x_t^{wd} + \beta^d x_t^{rd} + x_t^{gd} &\leq D_t, & (2.24)\\
x_t^{gd} + x_t^{gr} &\leq G_t, & (2.25)\\
x_t^{rd} + x_t^{rg} &\leq R_t, & (2.26)\\
x_t^{wr} + x_t^{gr} &\leq R_{\max} - R_t, & (2.27)\\
x_t^{wr} + x_t^{wd} &\leq E_t, & (2.28)\\
x_t^{wr} + x_t^{gr} &\leq \gamma^c, & (2.29)\\
x_t^{rd} + x_t^{rg} &\leq \gamma^d & (2.30)
\end{aligned}$$

where γ^c and γ^d are the maximum amount of energy that can be charged or discharged from the storage device. Typically, γ^c and γ^d are the same.

The transition function, $S^M(\cdot)$, explicitly describes the relationship between the state of the model at time t and $t+1$,

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

where $W_{t+1} = (E_{t+1}, P_{t+1}, D_{t+1})$ is the exogenous information revealed at $t+1$. In our numerical experiments, we assumed that W_{t+1} was independent of S_t, but the CFA algorithm can work with any sample path provided by an exogenous source. The relationship of storage levels between periods is defined as

$$R_{t+1} = R_t - x_t^{rd} + \beta^c x_t^{wr} + \beta^c x_t^{gr} - x_t^{rg} \tag{2.31}$$

where $\beta^c \in (0,1)$ and $\beta^d \in (0,1)$, are the charge and discharge efficiencies. For a given state S_t and decision x_t, we define:

$$C(S_t, x_t) = P_t \cdot (x_t^{wd} + \beta^d x^{rd} + x^{gd} + \beta^d x_t^{rg} - x_t^{gr} - x_t^{gd}) - C^{\text{penalty}} \cdot \big(D_t - x_t^{wd} - \beta^d x^{rd} - x^{gd}\big) \tag{2.32}$$

where C^{penalty} is the penalty of not satisfying demand and $C(S_t, x_t)$ is the profit realized at t given the current state is S_t and the decision is x_t. The objective is to find the policy π that solves

$$\max_{\pi \in \Pi} \ \mathbb{E}^{\pi} \left[\sum_{t=0}^{T} C(S_t, X_t^{\pi}(S_t)) \,\middle|\, S_0 \right] \tag{2.33}$$

subject to (2.23) - (2.31) for $t \in [0, T]$.

Renewable energy model

Our model below is designed in part to create complex nonstationary behaviors to test the ability of our policy to exploit forecasts while managing uncertainty. We use a hidden Markov model (Durante et al. (2017)) to create a very realistic model of the stochastic process describing the generation of renewable energy and make the amount of energy available from the grid a function of time. This model generates forecast errors based on an underlying crossing time distribution, the consecutive periods of time for which the observed energy produced is above or below the forecast.

These errors are modeled using a two-level Markov model with two state variables that evolve on different time scales. The primary state variable, which contains all the pertinent information to approximate the current period's error distribution, evolves at every discrete point of time. The secondary state variable, also known as the crossing state of the system, contains the sign of the error and the duration of how long the sample path has been above or below the forecast. Unlike the primary state variable, this secondary state variable is only updated when forecast errors change signs. Forecast errors are then generated using a distribution selected by a second level Markov model conditioned on the crossing state of the system.

A sample path of renewable energy and its respective forecast can be viewed in figure 2.2. Arrows have been added to identify crossing times. This is an example of a complex stochastic process that causes problems for stochastic lookahead models. For example, it is very common when using the stochastic dual decomposition procedure (SDDP) to assume interstage independence, which means that W_t and W_{t+1} are independent, which is simply not the case in practice (Shapiro et al. (2013) and Dupačová and Sladký (2002)). However, capturing this dynamic in a stochastic lookahead model is quite difficult.

Our CFA methodology, however, can easily handle these more complex stochastic models since we only need to be able to simulate the process in the base model. We

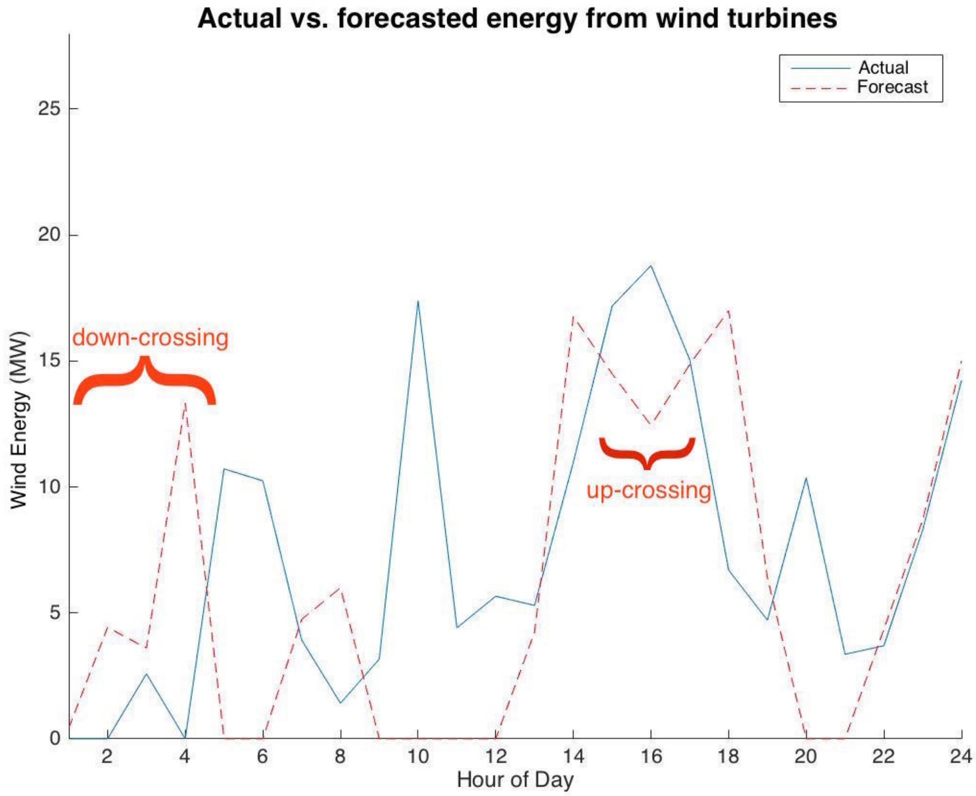


Figure 2.2: Sample path of renewable energy (E_t)

manipulate the quality of the renewable energy forecast by multiplying the forecast errors by the forecast quality, σ_f. This allows us to modify the quality of our forecast without modifying the observed stochastic process (P_t).

The amount of energy available from the main grid at t, G_t is defined as:

$$G_t = \min\left\{ \max\left\{90 - 50\sin\left(\frac{5\pi t}{2T}\right), G_{\min}\right\} G_{\max}\right\} \tag{2.34}$$

where $G_{\min}$ is the minimum energy always accessible from the grid, $G_{\max}$ is the maximum energy every accessible.

Spot price model

The spot price (P_t) of electricity at time t is a sinusoidal stochastic function defined as:

$$P_t = \min \left\{ \max \left\{ \frac{P_{\max} + P_{\min}}{2} - (P_{\max} - P_{\min}) \cdot \sin \left(\frac{5\pi t}{2T} \right) + \epsilon_t, P_{\min} \right\} P_{\max} \right\} \tag{2.35}$$

where $\epsilon \sim \mathcal{N}(\mu_p, \sigma_p)$, $P_{\min}$ is the minimum price allowed, $P_{\max}$ is the maximum price allowed, μ_p is expected value of the change in price, and σ_p is the standard deviation of the change in price. Since spot prices occasionally go below zero $P_{\min}$ may have a negative value. This is also the price at which energy can be purchased and sold to and from the grid. Sample paths of the stochastic process S_t are displayed in figure 2.3.

Since the price process, P_t, is stochastic, forecasts of P_t must be generated for both the deterministic lookahead and the CFA. In our model, we create forecasts by using correlated perturbations of the observed prices. This allows us to control the quality of the forecast without modifying the observed prices P_t. In our simulation we begin by creating a set of observed prices $P_1, \ldots, P_T$, which we treat as coming from history. We then create a series of forecasts where $F^P_{tt'} = \mathbb{E}_t[P_{t'}]$ given the information available at time t. The process $F^P_{tt'}$ satisfies the following conditions:

1. The spot price, P_t, is defined as

$$P_t = F^P_{t,t} \ \forall \ t \in [1, T] \tag{2.36}$$

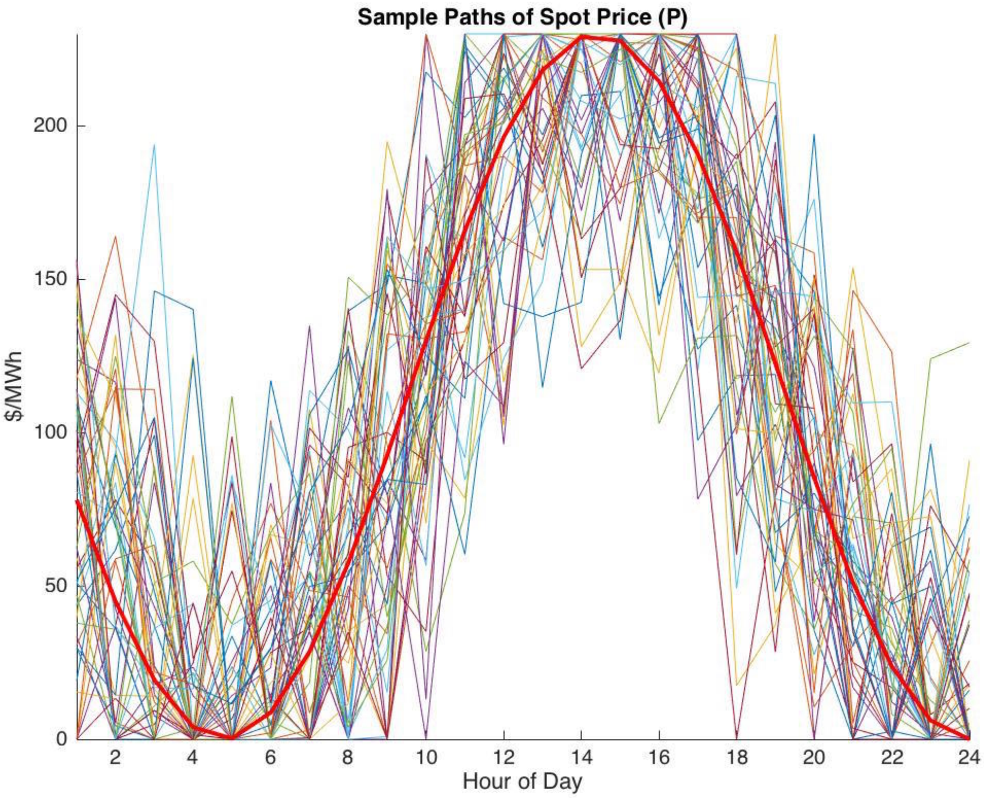


Figure 2.3: Sample paths of spot prices (P_t)

2. The stochastic process, $P_{tt'}$, is calculated from,

$$F^P_{t-1,t'} = \min\left\{ \max\left\{\rho_t, P_{\min}\right\}, P_{\max}\right\} \quad t' \geq t \tag{2.37}$$

where $\rho_t \sim \mathcal{N}(P_{tt'}, \sigma_f^2)$. We can directly control the quality of the forecast by varying σ_f, where $\sigma_f = 0$ means the forecast is perfect, while increasing σ_f degrades the quality of the forecast. Figure 2.4 compares the forecasted price path at time $t = 5$ to the observed price path for $\sigma_f = 5$.

Policy Parameterizations

If the contribution function, transition function and constraints are linear, a deterministic lookahead policy can be constructed as a linear program if point forecasts

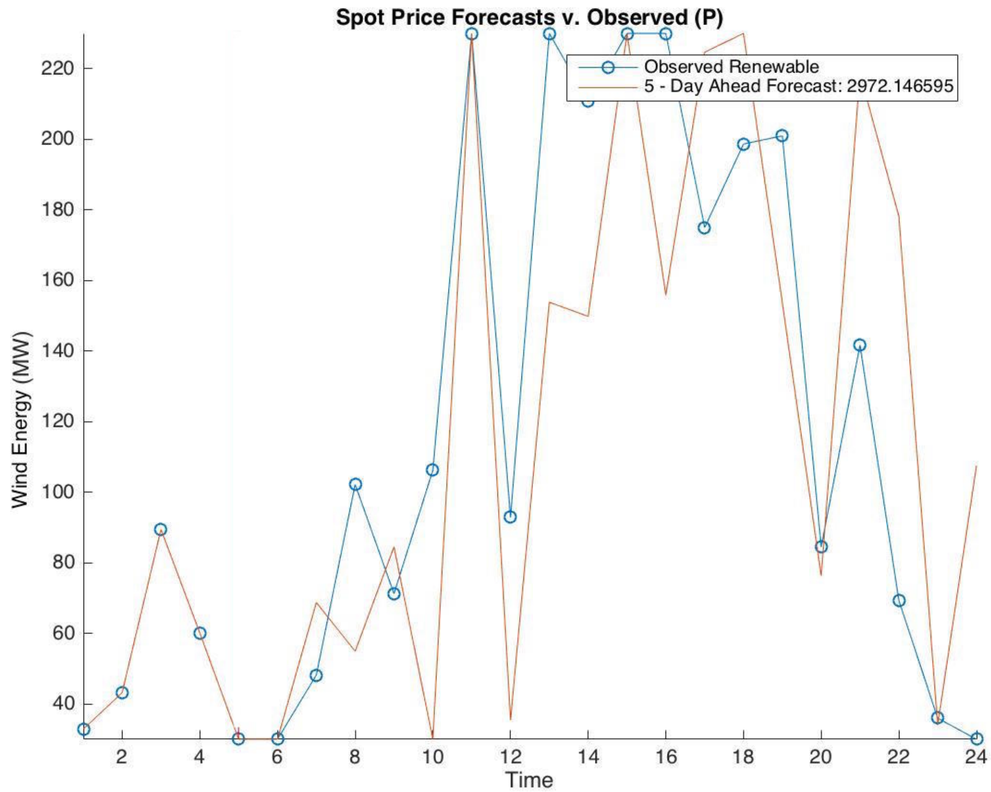


Figure 2.4: Spot price forecast ($P_{tt'}$) v. Observed price (P_t)

of exogenous information are provided. For our deterministic lookahead we optimize equation (2.3) subject to constraints (2.24) - (2.30).

for $t' \in [t+1, t+H]$. We call this deterministic lookahead policy the benchmark policy, and use it to estimate the degree to which the parameterized policies are able to improve the results in the presence of uncertainty.

- **Capacity Constraints**: This parameterization limits the amount of energy in storage and guarantees there is capacity to purchase inexpensive energy. An upper bound constraint is easily created by multiplying the capacity of the storage device, $R_{\max}$ by the parameter $\theta_{t'-t}$. This changes (2.27) to

$$x^{wr}_{tt'} + x^{gr}_{tt'} \leq R_{\max} \cdot \theta^U_{t'-t} - F^R_{tt'} \tag{2.38}$$

where $\theta_{t'} \in [0,1]$ and $t' \in [t, t+H]$. Parameterized lower constraints are incorporated into the policy by creating the additional linear constraints

$$-x_t^{rd} - x_t^{rg} + R_t \geq R_{\max} \cdot \theta^L_{t'-t} \tag{2.39}$$

where $\theta^L_{t'} \in [0,1]$ and $t' \in [t+1, t+H]$.

- **Lookup table forecast parameterization** - Overestimating or underestimating forecasts of renewable energy influences how aggressively a policy will store energy. We modify the forecast of renewable energy for each period of the lookahead model with a unique parameter θ_τ. This parameterization is a lookup table representation because there is a different θ for each lookahead period, $\tau = 0, 1, 2, ...$ This changes (2.28) to

$$x^{wr}_{tt'} + x^{wd}_{tt'} \leq F^{E_{tt'}} \cdot \theta_{t'-t}. \tag{2.40}$$

  where $t' \in [t+1, t+H]$ and $\tau = t' - t$. If $\theta_\tau < 1$ the policy will be more robust and decrease the risk of running out of energy. Conversely, if $\theta_\tau > 1$ the policy will be more aggressive and less adamant about maintaining large energy reserves.

- **Constant forecast parameterization** - Instead of using a unique parameter for every period, this parameterization uses a single scalar to modify the forecast amount of renewable energy for the entire horizon. The policy constraints (2.28) are changed to

$$x^{wr}_{tt'} + x^{wd}_{tt'} \leq F^{E_{tt'}} \cdot \theta. \tag{2.41}$$

- **Exponential Function** - Instead of calculating a set of parameters for every period within the lookahead model we make our parameterization a function of time and two parameters. The policy constraints (2.28) are then changed to

$$x_{tt'}^{wr} + x_{tt'}^{wd} \leq F^{E_{tt'}} \cdot \theta_1 \cdot e^{\theta_2 \cdot (t'-t)}. \tag{2.42}$$

2.5 Numerical Results

To demonstrate the capability of the CFA and Algorithm 1, we test parameterizations, (2.38)-(2.42), of the deterministic lookahead policy defined by equation (2.3) on variations of the previously described energy storage problem. We provide the benchmark policy and parameterized policies the same forecasts of exogenous information. Our goal is to show that parameterizing the benchmark policy and using the CFA gradient algorithm to determine parameter values can improve the performance of the benchmark policy. We say a parameterization, $\pi(\theta)$, outperforms the nonparametric benchmark policy if it has positive *policy improvement*, $\Delta F^{\pi}(\theta)$. We define the policy improvement, $\Delta F^{\pi}(\theta)$, of parameterization $\pi(\theta)$ as

$$\Delta F^{\pi}(\theta) = \frac{F^{\pi(\theta)} - F^{\text{D-LA}}}{|F^{\text{D-LA}}|} \tag{2.43}$$

where $F^{\pi}(\theta)$ is the average profit generated by parametrization $\pi(\theta)$ and $F^{\text{D-LA}}$ is the average profit generated by the unparameterized deterministic lookahead policy described by equation (2.3).

One of the most prominent advantages of the CFA is its ability to handle uncertainty without restrictions on the structure of the dynamics. By varying the forecast quality, σ_f, of the energy storage problem we demonstrate the abilities of the CFA and the CFA Gradient Algorithm to detect different levels of uncertainty and adapt accordingly. Table 2.1 presents the performance of each parameterization over varying forecast qualities.

	$\sigma_f = 20$	$\sigma_f = 25$	$\sigma_f = 30$	$\sigma_f = 35$
Constant	13%	13%	16%	17%
Lookup	20%	22%	26%	25%
Expo	14%	22%	26%	26%
Capacity Constraint	0.00%	0.00%	0.00%	0.00%

Table 2.1: This table displays the percentage improvement obtained by parameterized policies relative to the deterministic benchmark for varying forecast qualities, σ_f. These values are calculated using 500 simulations.

As the uncertainty and forecast error increases the performance of the benchmark policy deteriorates and the average profit generated decreases since it is unable to deal with uncertainty. The average profit of the parameterizations also deteriorate as forecast errors increase, but does so at a slower rate than the benchmark policy. Although the added noise to the forecast makes the problem more difficult, the parameterized policy is able to adapt and perform better than the standard deterministic lookahead policy. This explains the positive relationship between the the Constant, Lookup table, and Exponential parameterizations improvements and forecast quality.

As the forecast uncertainty increases, the parameterized policies adapt by discounting the forecast to limit the risk of paying penalties for not satisfying demand. This phenomenon can be seen in figure 2.5 which shows the relationship between profits and θ for the constant parameterization and different forecast qualities.

Factoring the Forecast

All policies prefer to underestimate the future renewable levels by setting $\theta_\tau < 1$ for all $\tau \in [0, H]$. Figures 2.6 and 2.7 show how θ_τ for the parameterizations described by (2.40) - (2.42) behave as functions of τ.

Notice how θ_τ decreases for each subsequent lookahead period for the lookup table and exponential adjustment functions. This is a consequence of the diminishing marginal improvement for each additional period in the lookahead model. As seen in figure 2.5, as the forecast error, σ_f, increases θ_τ decreases. This implies the algorithm

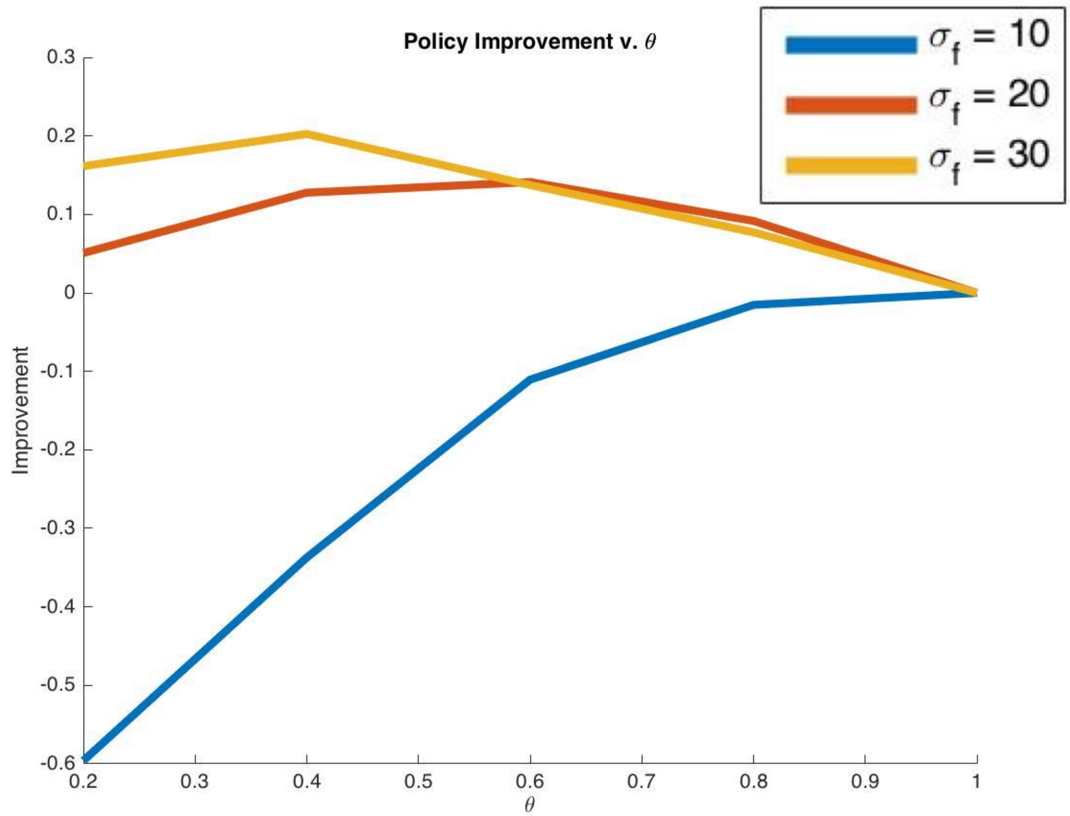
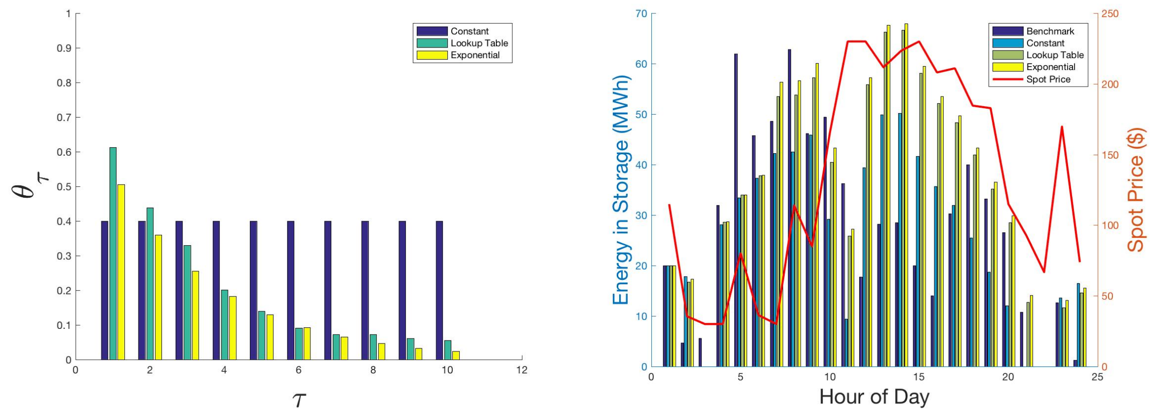


Figure 2.5: Policy Improvement over deterministic benchmark v. θ for constant parameterizations

recognizes that as the forecast error increases the forecast is less reliable. The policy determines that it is better to just expect no renewable energy than to depend on the forecast.

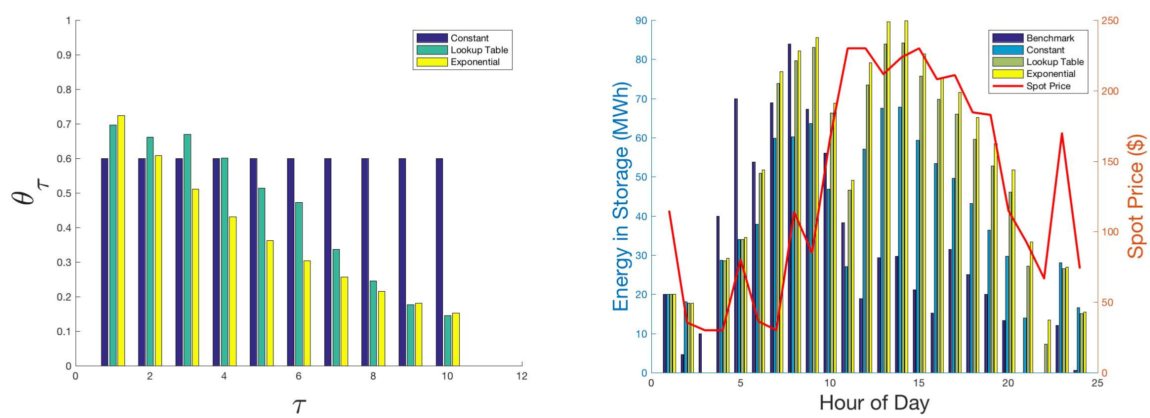
Capacity Constraints

The capacity constraint parameterization, described by equations (2.38) and (2.39), was the only parameterization that did not generate positive improvement in the provided problem settings. Setting an upper limit on the storage in the lookahead model decreases the amount of energy placed into storage during the current state. This maintains lower storage levels than the benchmark policy and limits the purchased energy from the grid for storage. However, this also limits the ability of the



(a) θ_τ for $\sigma_f = 35$ and $\gamma = R_{\max}$ (b) R_t for $\sigma_f = 35$ and $\gamma = .R_{\max}$

Figure 2.6: Figure 2.6a compares the θ_τ values for the *Constant*, *Lookup Table*, and *Exponential* parameterizations when $\sigma_f = 35$ and $\gamma = \cdot R_{\max}$. Figure 2.6b compares the storage levels, R_t of the different parameterizations over $t \in [1, 24]$ for the same conditions.



(a) θ_τ for $\sigma_f = 20$ and $\gamma = R_{\max}$ (b) R_t for $\sigma_f = 20$ and $\gamma = R_{\max}$

Figure 2.7: Figure 2.7a compares the θ_τ values for the *Constant*, *Lookup Table*, and *Exponential* parameterizations when $\sigma_f = 30$ and $\gamma = R_{\max}$. Figure 2.7b compares the storage levels, R_t of the different parameterizations over $t \in [1, 24]$ for the same conditions.

parameterized policy to sell excess energy to the grid for profit. This can be seen in figure 2.8.

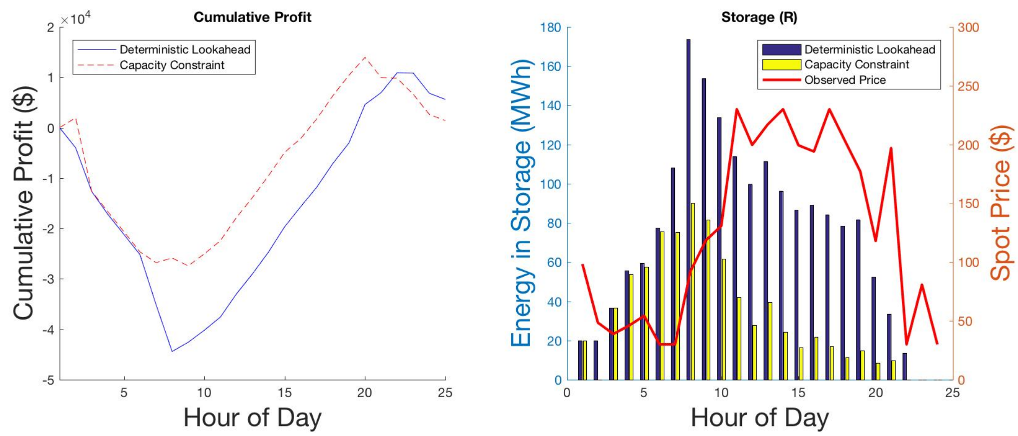


Figure 2.8: Capacity constraint parameterization sample path where $\theta = [.3, 0]$

Notice how the cumulative profit of the parameterized policy is greater than that of the benchmark policy until $t = 20$ in figure 2.8. The parameterized policy achieves this by maintaining lower storage associated costs. However, as the simulation approaches $t = T$ the benchmark policy begins to sell off excess storage. Since the storage for the parameterized policy is constantly lower than the benchmark it misses the additional returns. Setting a lower limit has the reverse effect on storage. This can be seen in figure 2.9. By requiring a certain amount of energy in the storage device in the

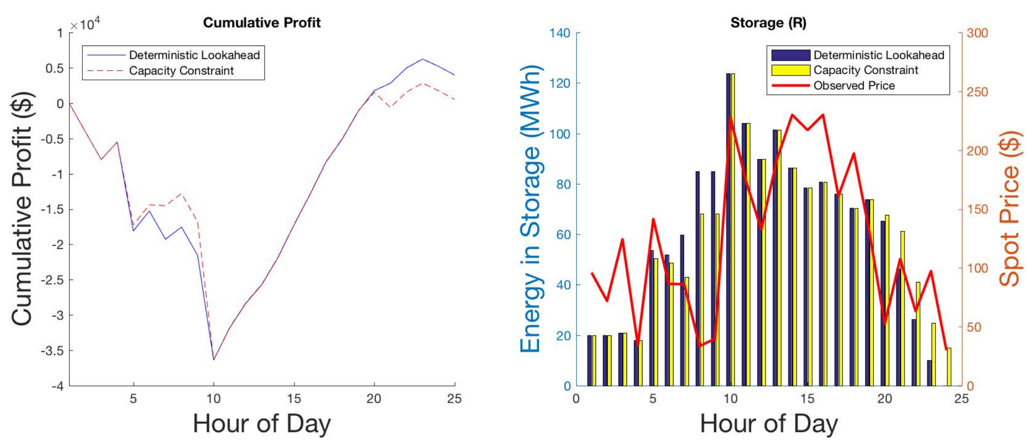


Figure 2.9: Capacity constraint parameterization sample path where $\theta = [1, .05]$

lookahead policy, the policy is unable to sell as much excess energy to the grid during the current period as the benchmark policy. This limits the ability of the policy to generate revenue. The CFA Algorithm seemed to recognize these problems and did not limit the capacity constraints in the lookahead model as much. Although it could not improve the lookahead policy by modifying the capacity constraints, it still identified the optimal $\theta^* = [1, 0]$ for the parameterization form.

Finally, we applied the methodology for a problem with perfect forecasts. As we would expect (but there are never guarantees), the CFA gradient algorithm finds the optimal policy by setting $\theta_\tau \approx 1 \forall \tau = 1, \ldots, H$. This is shown in figure 2.10.

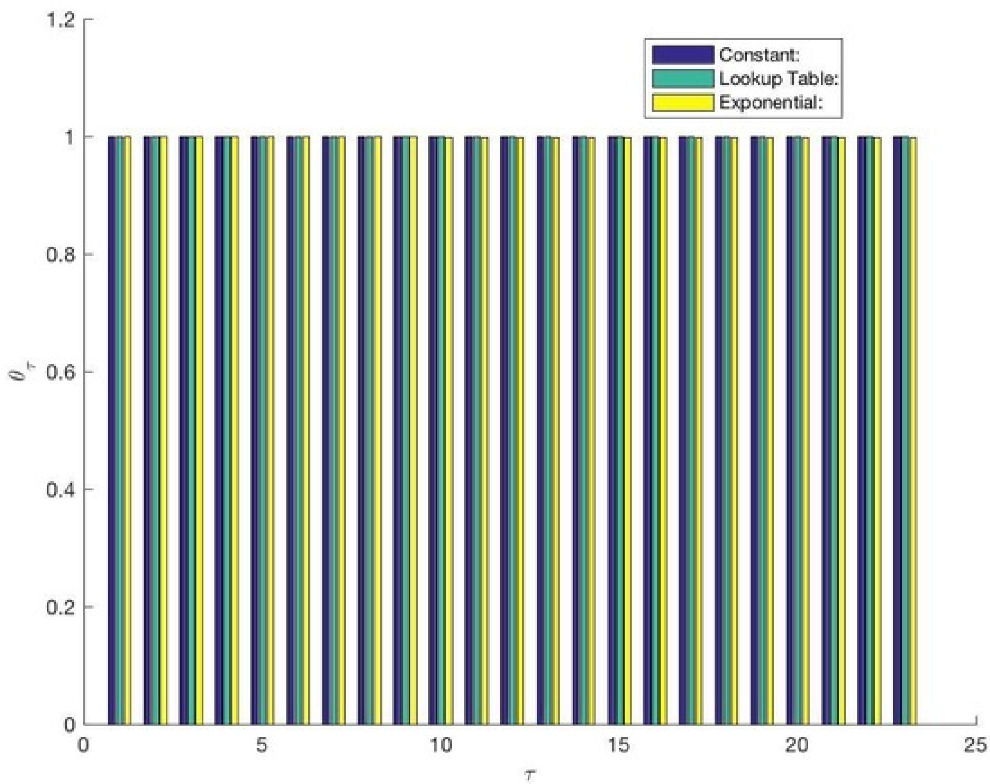


Figure 2.10: θ_τ values for the Constant, Lookup, and Exponential parameterized policies given a perfect forecast after 500 simulations.

We note in closing that our parametric CFA policy is stationary, in that the parameters θ are not time-dependent. As a result, we only have to apply the gradient

CFA algorithm once in an offline setting, as long as characteristics such as the quality of the forecast remain the same.

2.6 Conclusion

This work builds upon a long history of using deterministic optimization models to solve sequential stochastic problems. Unlike other deterministic methods, our class of methods, *parametric cost function approximations*, parametrically modify deterministic approximations to account for problem uncertainty. Our particular use of modified linear programs and the CFA Gradient Algorithm represent a fundamentally new approach to solving stochastic programming problems. Our method allows us to exploit the structural properties of the problem while capturing the complex dynamics of the full base model, rather than accepting the approximations required in a stochastic lookahead model. We have demonstrated this class of policies in the context of a complex, time-dependent energy storage problem with forecasts. For our numerical work we selected an energy storage problem that is relatively small to simplify the extensive computational work. However, our methodology is scalable to any problem setting which is currently being solved using a deterministic model.

An important feature of our approach is that it can handle complex dynamics, as long as we are able to compute the derivatives in equations (2.13). For example, we were able to handle the complex hidden semi-Markov model used to represent renewable energy described in section 6.1. Our methodology would not be affected if this were replaced with any other time series model, or even an observed sample path from history (for which there is no model).

The parametric cost function approximation represents an alternative to stochastic lookahead models that represent the foundation of stochastic programming. Parametric CFAs require some intuition into how uncertainty might affect the optimal solution.

We would argue that this requirement parallels the design of any parametric statistical model, and hence enjoys a long history. We believe there are many problems where practitioners have a good sense of how uncertainty affects the solution. However, further research will be required to determine how well this methodology performs compared to classical stochastic programming models based on scenario trees.

Chapter 3

Managing Energy Portfolios using Parametric Cost Function Approximations

3.1 Introduction

In a deregulated electricity market power producers are exposed to risk from volatile fuel and electricity prices as well as time-varying electricity loads. Faced with the potential of expensive production costs, electricity retailers can reduce their own costs by making advance commitments, but this requires making decisions using more uncertain forecasts. This uncertainty can be mitigated by committing resources (e.g. gas turbines) with much smaller notification times, but at higher cost. For example, at time t (which might be noon today), a retailer may need to plan how much steam generation to bring online at time t' tomorrow (say, 2pm). This decision requires knowing how much energy might be available from a wind farm using a day-ahead forecast that is quite poor. However, the retailer will get a better forecast at 1pm tomorrow, which can inform decisions for using gas turbines which we do not have to commit until 1pm tomorrow. The higher cost of waiting and using a short-notification time generator may be offset by reducing the possibility of committing excess resources in advance.

Lagged decisions can be handled in a deterministic lookahead model, as is currently done in all unit commitment models used by industry. Such models carry an inherent bias toward making early commitments to take advantage of lower costs. However, this standard approximation does not properly account for differences in forecast errors for different lead times. The challenge is properly modeling the uncertainties in forecasts within a planning horizon, and the stochastic decisions that result in response to the arrival of new information.

Lagged resource planning problems arise in many settings, including supply chain management, transportation and logistics, but for this paper our particular focus is in energy systems. There are a number of papers modeling energy management problems as stochastic programs (see for example Takriti et al. (1996), Wallace and Fleten (2003), Philpott et al. (2000), Shapiro et al. (2013)). There are several papers that focus on problems related to lagged resource planning problems in the energy management community. Conejo et al. (2008) uses a large scale mixed-integer linear program to manage an electricity producer's portfolio of long-term electricity forward contracts using alternative risk measures such as Conditional Value-at-Risk (CVaR). Singh et al. (2009) simulates a large-scale multi-stage, stochastic, mixed-integer linear program to optimize the capacity expansion plan of an electricity distribution network in New Zealand. Takriti et al. (2000) presents a stochastic model for the unit commitment that incorporates power trading. These models use large scenario trees to approximate potential future events, but result in very large-scale optimization models that can be quite hard to solve in practice.

Every paper on stochastic unit commitment (and the vast majority of papers in stochastic programming) uses a standard two-stage modeling approximation, where the stochastic future (which might span one or two days) is realized all at once, after which all decisions after the initial decision are made using perfect information. This modeling approximation ignores the evolution of information and the changing quality in forecasts. For example, a day-ahead forecast of wind at 3pm will be much less accurate than the hour-ahead forecast. The model has to trade off potentially lower costs of energy committed in the day-ahead market against the potentially higher costs of decisions made with shorter notification times.

A popular modeling strategy for multistage linear programs is the use of stochastic dual dynamic programming (SDDP), which has been widely used in hydroelectric planning (see Pereira and Pinto (1991), Shapiro et al. (2013), and Philpott and Guan (2008)). SDDP falls under the broad umbrella of approximate dynamic programming, which approximates the future through value function approximations using methods such as Benders cuts or statistical models (Powell, 2011). VFA-based policies are not able to explicitly model forecasts in the state variable. As a result, forecasts have to be treated as latent variables.

By contrast, lookahead models such as scenario trees handle the forecast directly in the formation of the scenario tree, but standard practice (for computational reasons) is to use a two-stage approximation, where decisions in the future are allowed to see the entire future. The two-stage approximation ignores the characteristic of lagged problems which requires that we recognize at time t that a decision may be made at time $t' > t$ to be implemented at time $t'' > t'$ using the forecast that will be available at time t'.

In this chapter, we transition to the difficult problem of making lagged commitments while managing a portfolio of energy resources including steam and gas turbine generators. This setting requires making a decision now to make a commitment several hours and days in the future, which still captures our ability to make shorter term commitments, typically at higher prices but with more accurate forecasts.

Our method is be able to make the tradeoff between forecast reliability and cost. For example, imagine that we have a system that can make very reliable short-term forecasts. Now imagine that these short-term contracts are the lowest cost option. If this is the case, then an effective model and algorithm should be able to recognize that we should never schedule fossil fuel generators very early in advance, but this would not be the case if we used any current stochastic or deterministic lookahead model. We note that in our work in energy, it is not uniformly the case that forecasts become more accurate with shorter notification times, and costs do not increase monotonically as the notification time decreases.

This chapter makes the following contributions. 1) We provide, apparently for the first time, a proper base model of a stochastic, lagged resource allocation problem in the context of energy portfolio management. 2) We introduce a family of parameterizations of a deterministic lookahead model designed to produce robust policies that respond in a realistic way to the level of uncertainty in forecasts. 3) We demonstrate empirically that our method produces high quality solutions relative to unmodified deterministic lookahead policies on a library of lagged energy portfolio problems.

Our presentation is organized as follows. In section 3.2 we discuss the inherent bias of deterministic lookahead models in lagged problems. The modeling framework for the lagged energy portfolio problem is given in section 3.3. We then provide an overview of the different classes of policies and alternative designs for parametric CFAs in section 3.4. Section 3.6 presents a series of numerical results.

3.2 Bias of deterministic models

Although deterministic lookahead models are commonly used to handled lagged decision problems, these models exhibit clear biases and limitations. Consider the case when prices increase and forecast quality improves the later decisions are made. If costs decline when we make commitments farther in the future, deterministic models will have a natural bias of placing orders early because they are unable to account for the issue of forecast accuracy, and our ability to respond to better forecasts. This bias can be corrected by strategically parameterizing a deterministic model. To illustrate this idea we consider an extension of the classical newsvendor problem (Arrow et al., 1952). In the classical problem a newsvendor sells daily newspapers at a static price, p. Every day the vendor must decide how many newspapers to purchase, x, for a given cost, c, before knowing how many customers will demand newspapers, $\hat{D}$. The objective of the vendor is to determine how many newspapers to purchase in order to maximize expected profits. Formally, this problem is written as

$$\max_{x} \quad p \cdot \mathbb{E}\left[\max(x, \hat{D})\right] - cx.$$

This is a classical single-period problem where a decision x is made, then exogenous information is observed, $\hat{D}$, and the final outcome is measured. However, if we allow the vendor multiple opportunities to purchase inventory the problem becomes a lagged resource allocation problem. In the two-stage newsvendor problem the vendor can order inventory one day in advance, x_1, at the marginal cost c_1 or place orders once $\hat{D}$ is known, x_2, at the marginal cost c_2 where $c_1 < c_2$. In this case, the demand $\hat{D}$ is an unobserved random variable when x_1 is made, but known when x_2 is determined. If the cost of ordering inventory at the latest opportunity, c_2, is less than the selling price, p, then the vendor will satisfy the order of every customer and $x_2 = (D - x_1)^+$.

Formally, we describe the problem as

$$\max_{x_1} \mathbb{E}\left[p \min(\hat{D}, x_1 + x_2) - c_1 x_1 - c_2 x_2)\right].$$

Given the vendor has a day-ahead forecast of the demand, f_1^D, we can use the following deterministic lookahead policy to determine how much inventory to order using

$$x_1^* = \underset{x_1}{\operatorname{argmax}} \left[p f_1^D - c_1 x_1 - c_2 x_2 \right] \tag{3.1}$$

subject to

$$\begin{aligned} x_1 + x_2 &= f_1^D, \\ x &\geq 0. \end{aligned} \tag{3.2}$$

If the forecast f_1^D is perfect, then the policy (3.1) is optimal and the problem is trivial. Since this generally not the case, the vendor must consider the potential of ordering too much inventory in advance or waiting and ordering the right amount at a higher cost. This lookahead model will always order the forecasted demand, f_1^D, without considering the risk of ordering too much or too little. However, we can mitigate the risk of ordering too much by adding the following parameterized constraint,

$$x_1 \leq f_1^D \cdot \theta, \tag{3.3}$$

to the policy in equation (3.1) where $\theta \geq 0$. When $\theta = 1$, the parameterized policy defined in equation (3.3) is identical to the unmodified policy described in equation (3.1). Depending on the value of θ the decisions of the parameterized policy can deviate from unmodified lookahead model such that the policy may purchase more or less inventory than is forecasted.

Consider the case where $c_1 = \$5$, $c_2 = \$7$, $p = \$10$, the forecast of demand is a noisy observation such that $f_1^D \sim \mathcal{N}(D, \sigma^f)$, and the realization is $\hat{D} = 10$. Figure 3.1 illustrates the relationship between the policy parameter, θ, and the expected cumulative profit for the parametric policy for different σ^f. Figure 3.1 shows that as

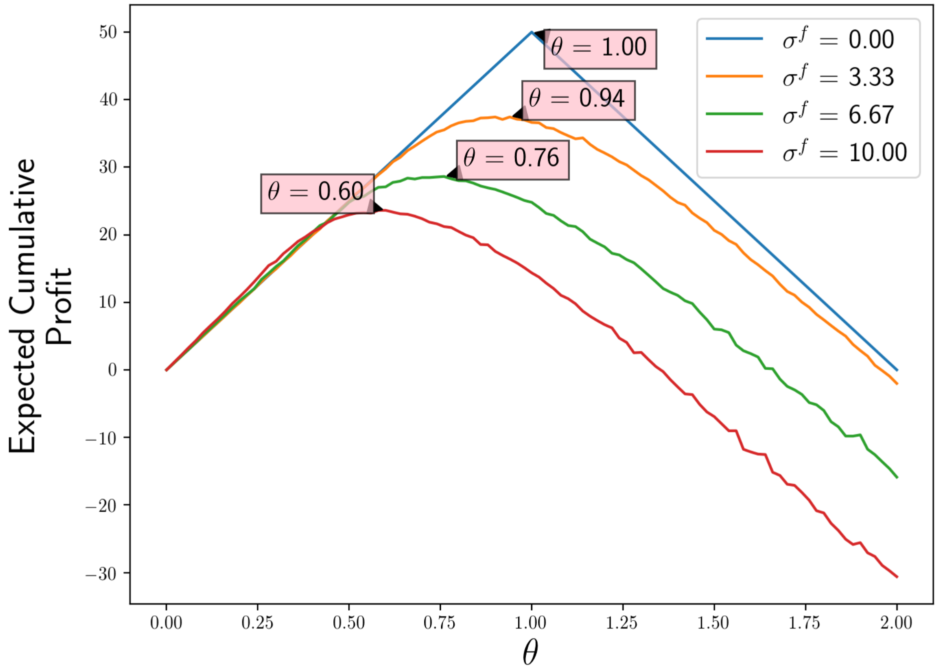


Figure 3.1: This figure compares the expected cumulative profits for parametric policy defined by equation (3.3) for different parameter values, θ, and forecast qualities σ^f where $f_1^D \sim \mathcal{N}(\mathbb{E}[\hat{D}], \sigma^f)$. The empirical expected values were calculated using 10^6 independent observations.

the amount of noise in the forecast, σ^f, increases the vendor can increase expected profits by discounting the forecast, or setting $\theta < 1$. Figure 3.2 shows the optimal distribution of orders for each of the tested forecast qualities, σ^f. As the forecast noise increases the optimal policy is to order less inventory in advance and to order the expensive inventory later. The parameterized policy exploits the fact that it is less expensive to wait and order inventory than to have a surplus.

This problem can be further generalized so the vendor may have T opportunities to purchase inventory before customer arrives. In this case, demand $\hat{D}_T$, is only revealed at time T, but the forecasts of the demand, $(f_{t,T}^D)_t^T$ evolves over time according to

$$f_{t+1,T}^D = f_{t,T}^D + \epsilon_t,$$

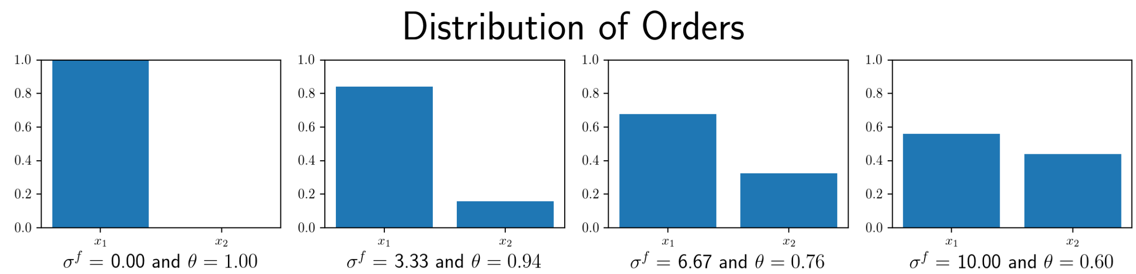


Figure 3.2: This figure presents the optimal distribution of ordering decisions for the parametric policy defined by equation (3.3) for different forecast qualities σ^f. The empirical expected values were calculated using 10^6 independent observations.

where $\epsilon_t \sim \mathcal{N}(0, \sigma^f)$ and $f^D_{T,T} = \hat{D}_T$. The decisions are determined using a policy, $X^\pi_t(S_t)$, which uses the $\mathcal{F}_t$-measurable state variable $S_t = (f^D_t, R_t)$ to determine how much inventory to order at time t. To keep track of the aggregate amount of inventory already ordered, we introduce the sequence of variables, $(R_t)^T_{t=1}$, where

$$R_{t+1} = R_t + x_t.$$

The T-period problem is formally described as

$$\max_\pi \mathbb{E}\left[p \min\left(\hat{D}_T, \sum_{t=1}^{T} X^\pi_t(S_t)\right) - \sum_{t=1}^{T} c_t X^\pi_t(S_t)\right], \tag{3.4}$$

where the objective is to find the policy, π, that maximizes profit. We can solve this problem using the following parameterized lookahead policy to determine how much inventory to purchase for every period $1 \leq t \leq T$.

$$X^\pi_t(S_t) = \underset{\tilde{x}_{t,t}}{\text{argmax}} \left[p \cdot f^D_{t,T} - \sum_{t'=t}^{T} c_{t'}, \tilde{x}_{t,t'}\right] \tag{3.5}$$

$$\text{subject to}$$
$$\sum_{t'=t}^{T} \tilde{x}_{t,t'} \leq f^D_{t,T} - R_t,$$

$$\tilde{x}_{t,t} \leq \theta_{t,T} f^D_{t,T},$$
$$\tilde{x} \geq 0,$$

where the decision $x_{tt'}$ represents the estimated amount of inventory to purchase at time t' given information available at time t and $\theta_{t,T} = 1+\beta(T-t)$. By modifying the policy parameter, β, we can control the timing of when the vendor orders inventory. Figure 3.3 compares the expected cumulative profits for the parametric policy defined in equation (3.5) when $T = 5$ and

$$c_t = 5 + (t-1)\frac{5}{T-1}.$$

With this function, costs increase as we order later, which encourages ordering earlier.

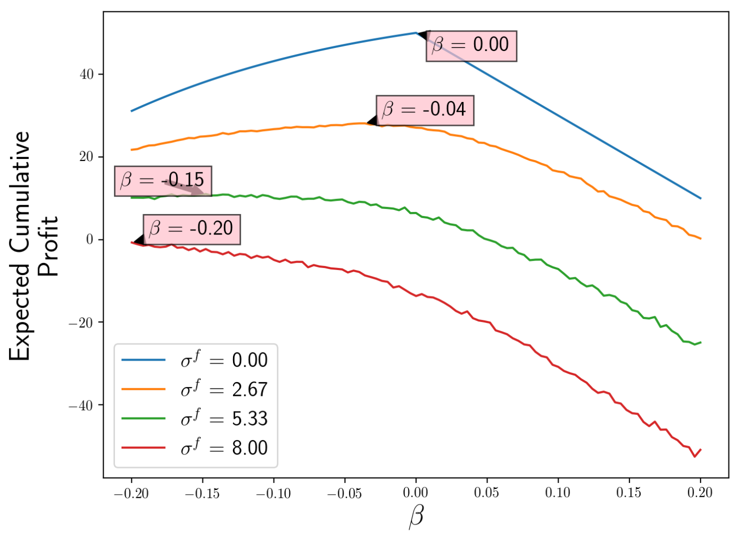


Figure 3.3: This figure compares the expected cumulative profits for parametric policy defined by equation (3.5) for different parameter values, θ, and forecast qualities σ^f. The empirical expected values were calculated using 10^6 independent observations and $T = 5$.

Figure 3.3 illustrates the same behavior as seen in the two period newsvendor problem. As the amount of noise in the forecast increases, the optimal policy is to further discount the forecast of demand and wait to purchase inventory. This behavior can clearly be seen in figure 3.4.

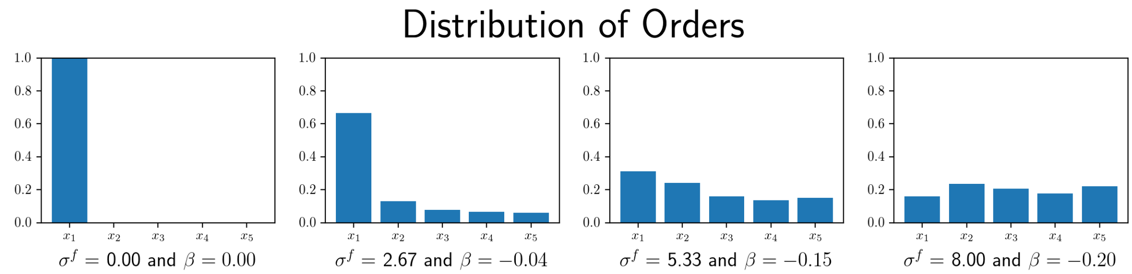


Figure 3.4: This figure presents the optimal distribution of ordering decisions for the parametric policy defined by equation (3.5) for different forecast qualities σ^f. The empirical expected values were calculated using 10^6 independent observations.

Although the multiperiod newsvendor problem is much simpler than the lagged energy commitment problem we present in this chapter, it demonstrates why a parameterized cost function approximation can outperform unmodified deterministic polices in a lagged setting. When the cost of ordering energy increases over time, it is better to order early, but this requires using less accurate forecasts which requires balancing cost against forecast accuracy. Below, we study a more complex lagged problem where a range of nested decisions $x_{tt'}$ can be made for many time periods t', where energy is committed anywhere from time $t = 1$ to $t = t'$ for $t' = 1, ..., T$. In addition, we assume excess energy in storage at time t' can be held forward to time $t' + 1$, which means the problem does not decompose by time t'.

3.3 Lagged energy portfolio model

An electricity retailer must satisfy a stochastic hourly load over a finite horizon, T, while managing a portfolio of energy resources. The portfolio contains dispatchable (typically steam and gas turbine) generators, $\mathcal{G}$, intermittent renewable energy sources, $\mathcal{E}$, and a single storage device. Each source of energy is characterized by a notification time that reflects startup times, planning processes, and internal operational practices. We use the notation $x_t = (x_{tt'})_{t' \geq t}$ as a vector of energy commitments made at time t to be delivered over a horizon $t \leq t' \leq T$. Any decision $x_{tt'}$ is treated as a commitment that cannot be changed. We note different types of decisions made be made at different time periods (e.g. steam decisions may be made at noon each day, while decisions about gas turbines may be made hourly). These notification times can span minutes to 24 hours or more. Renewables cannot be controlled (that is, they are not "dispatchable"), but are described by an evolving set of forecasts, which parallel the lagged plans. Any unsatisfied load must be met using electricity purchased from the electricity spot market at a stochastic price. The retailer may store electricity from the renewable sources or spot market to use at a later date. The retailer's objective is to maximize expected cumulative returns.

We follow the framework described in chapter 1 to present our lagged energy portfolio model.

State variable

At time t the state variable S_t contains the following:

> $(f^L_{tt'})_{t'=t,\ldots,t+H}$ = forecasts of future loads at time t' given information available at time t,
>
> $((f_{tt'e})_{t'=t,\ldots,t+H})_{e\in\mathcal{E}}$ = forecasts of electricity available from renewable sources, $e \in \mathcal{E}$ at time t' given information available at time t,
>
> $(f^P_{tt'})_{t'=t,\ldots,t+H}$ = forecasts of the spot price for electricity at time t' given information available at time t,
>
> $\left(P^{\text{gen}}_t(\tau_g)\right)_{g\in\mathcal{G}}$ = prices of purchasing electricity from the set of dispatachable generators, $\mathcal{G}$, at time t. Specifically, $P^{\text{gen}}_t(\tau_g)$ is the price of ordering from generator $g \in \mathcal{G}$ at time t to be delivered at $t + \tau_g$, where τ_g is the unique notification time of the generator $g \in \mathcal{G}$.

$(R^{\text{gen}}_{tt'})_{t'=t,\ldots,t+H}$ = aggregate amount of energy from the generators from the set of dispatachable generators, $\mathcal{G}$, already committed for time t',

R^{stor}_t = the amount of energy in storage at time t,

$(P_{te})_{e\in\mathcal{E}}$ = the marginal prices to use electricity generated by renewable sources, $e \in \mathcal{E}$.

Formally, we define the state variable as

$$S_t = \Big(R^{\text{stor}}_t, (R^{\text{gen}}_{tt'})_{t'=t,\ldots,t+H}, (f^L_{tt'})_{t'=t,\ldots,t+H}, \\ ((f_{tt'e})_{t'=t,\ldots,t+H})_{e\in\mathcal{E}}, (f^P_{tt'})_{t'=t,\ldots,t+H}, \left(P^{\text{gen}}_t(\tau_g)\right)_{g\in\mathcal{G}} \Big), \tag{3.6}$$

where $R^{\text{stor}}_t \in [0, R^{\max}]$. Forecasts, $f^X_{tt'}$, are estimates of some underlying random variable, $X_{t'}$, at time t', given what we know at time t. We denote the actual observation of the underlying variable at time t as $X_t = f^X_{tt}$. We note that our handling of forecasts is important in that they are modeled explicitly in the state variable, and not as a latent variable. This is the reason (as we show below) that our parametric fit of the CFA does not have to be reoptimized when the forecasts are changed, giving us a stationary policy even in the presence of a highly time dependent problem.

The cost of purchasing electricity in advance from an energy generator is a function of the notification time of generator g, τ_g. This function is known as the *lagged price function*, $\lambda(\cdot) : \mathbb{N} \to \mathbb{R}$. By varying the shape of the lagged price function we can create very different problem settings. We discuss this function in greater detail and provide examples in section **??** with the numerical experiments.

Decision variables

At time t the electricity retailer must determine the flow of electricity between the following sources:

$(x^{\text{ren-load}}_{t,e})_{e\in\mathcal{E}}$ = the renewable sources, $e \in \mathcal{E}$, to load,

$(x^{\text{gen-load}}_{tt'g})_{t'=t+1,...,t+H}$ = the energy committed at time t from generator g to come online at time t',

$x^{\text{spot-load}}_t$ = the spot market to load,

$(x^{\text{ren-stor}}_{t,e})_{e\in\mathcal{E}}$ = the renewable sources, $e \in \mathcal{E}$ to storage,

$x^{\text{spot-stor}}_t$ = the electricity spot market to storage,

$x^{\text{stor-spot}}_t$ = the storage device to the spot market,

$x^{\text{stor-load}}_t$ = the storage device to the load.

Formally, we define the decision, x_t, at time t as

$$x_t = \left((x^{\text{ren-load}}_{t,e})_{e\in\mathcal{E}}, (x^{\text{gen-load}}_{tt'g})_{t'=t+1,...,t+H}, x^{\text{spot-load}}_t, (x^{\text{ren-stor}}_{t,e})_{e\in\mathcal{E}}, x^{\text{spot-stor}}_t, x^{\text{stor-spot}}_t, x^{\text{stor-load}}_t \right) \tag{3.7}$$

where $x_t \geq 0$. The decisions relating to electricity generated by renewable sources, $e \in \mathcal{E}$ are constrained by the amount of available renewable energy,

$$0 \leq x^{\text{ren-load}}_{t,e} + x^{\text{ren-stor}}_{t,e} \leq f^e_{tt} \;\; \forall \;\; t = 0, ..., T \text{ and } e \in \mathcal{E}. \tag{3.8}$$

The variables, $\gamma^c \in [0, 1]$ and $\gamma^d \in [0, 1]$ are the charge and discharge rates of the electricity storage device. They decide how quickly energy can be deposited and withdrawn from storage. The following constraint describes how quickly electricity can be deposited in storage,

$$0 \leq x^{\text{spot-stor}}_t + \sum_{e\in\mathcal{E}} x^{\text{ren-stor}}_{t,e} \leq \min\left(\gamma^c R^{\max}, R^{\max} - R^{\text{stor}}_t\right). \tag{3.9}$$

The following constraint describes how quickly electricity can be removed from storage,

$$0 \leq x_t^{\text{stor-load}} + x_t^{\text{stor-spot}} \leq \min\left(\gamma^d R^{\max}, R_t^{\text{stor}}\right). \tag{3.10}$$

The load at t, f_{tt}^L, must always be satisfied. We use the following constraints to enforce this requirement,

$$f_{tt}^L \leq \sum_{e \in \mathcal{E}} x_{t,e}^{\text{ren-load}} + R_t + x_t^{\text{spot-load}} + x_t^{\text{stor-load}} \quad \forall \quad t = 0, ..., T. \tag{3.11}$$

There is also a limit on how much electricity can be ordered from generator $g \in \mathcal{G}$ every period. We define this constraint as

$$0 \leq x_{tt'g}^{\text{gen-load}} \leq R_g^{\max} \quad \forall \quad t = 0, ..., T \;\text{ and } \forall g \in \mathcal{G}. \tag{3.12}$$

Decisions are determined by a decision function (policy) which we denote by $X_t^\pi(S_t)$, where π carries the information that determines the structure and parameters that define the function. When we wish to make the dependence on the parameters explicit, we write the policy as $X_t^\pi(S_t|\theta)$.

Exogenous Information

The exogenous information, W_t, at time t contains the changes in forecasts of future loads $((\hat{f}_{tt'}^L)_{t'=t,...,t+H})$, available electricity from renewable sources $\left(\left((\hat{f}_{tt'e}^{\text{ren}})_{t'=t,...,t+H}\right)_{e \in \mathcal{E}}\right)$, and the spot price of electricity $\left((\hat{f}_{tt'}^P)_{t'=t,...,t+H}\right)$. We formally define the exogenous information at time t as

$$W_t = \left(\left(\hat{f}_{tt'}^L\right)_{t'=t,...,t+H}, \left((\hat{f}_{tt'e}^{\text{ren}})_{t'=t,...,t+H}\right)_{e \in \mathcal{E}}, \left(\hat{f}_{tt'}^P\right)_{t'=t,...,t+H} \right). \tag{3.13}$$

Updated forecasts are dependent on past forecasts such that forecasts improve over time. We define the forecasts as noisy approximations of the underlying stochastic variable $X_{t \leq t' \leq H}$ where $\hat{f}^X_{t+1,t'} \sim \mathcal{N}(0, (\sigma^f)^2)$ and

$$f^X_{t+1,t'} = f^X_{t,t'} + \hat{f}^X_{t+1,t'}. \tag{3.14}$$

The stochastic error terms $(\hat{f}^P_{tt'})_{t'=t,\dots,t+H}$ and $\left((\hat{f}^{\text{ren}}_{tt'e})_{t'=t,\dots,t+H}\right)_{e \in \mathcal{E}}$ are generated using the univariate crossing state hidden semi-Markov model (HSMM) from Durante et al. (2017) to capture the distribution of wind speed and electricity spot prices exceeding and falling below a forecast series, termed crossing times. These models include hidden states which captures for how long wind speed is above or below a forecast as well as distributions of the change in wind speed conditioned on both the prior wind speed and the hidden state. In our simulations the periodic load, (f^L_{tt}), has the stochastic seasonal structure

$$f^L_{tt} = 1200 + 200 \sin\left(\frac{5\pi t}{2T}\right) + \epsilon_L, \tag{3.15}$$

where $\epsilon_L \sim \mathbb{N}(0, 50)$.

Transition function

The transition function $S^M(\cdot)$ explicitly describes the relationship between the state of the model at time t and $t+1$,

$$S_{t+1} = S^M(S_t, x_t, W_{t+1}) \tag{3.16}$$

where $W_{t+1} = \left((\hat{f}^L_{tt'})_{t'=t,\dots,t+H}, \left((\hat{f}^{\text{ren}}_{tt'e})_{t'=t,\dots,t+H}\right)_{e \in \mathcal{E}}, (\hat{f}^P_{tt'})_{t'=t,\dots,t+H}\right)$ is the exogenous information revealed at time $t+1$. Every period t the system must update all forecasts and electricity scheduled for future periods. We formally describe the transition function for this system as

$$S^M(S_t, x_t, W_{t+1}) = \begin{cases} R^{\text{gen}}_{t+1,t'} &= R^{\text{gen}}_{tt'} + x^{\text{gen-load}}_{tt'g} \;\; \forall \;\; t' = t+1, ..., t+H \\ R_{t+1} &= R^{\text{stor}}_t + x^{\text{spot-stor}}_t + x^{\text{wind-stor}}_t - x^{\text{stor-load}}_t - x^{\text{stor-spot}}_t \\ f^L_{t+1,t'} &= f^L_{t,t'} + \hat{f}^L_{t+1,t'} \\ f_{t+1,t',e} &= f_{t,t',e} + \hat{f}_{t+1,t',e} \\ f^P_{t+1,t'} &= f^P_{t,t'} + \hat{f}^P_{t+1,t'} \end{cases} \tag{3.17}$$

where the transition function is required to satisfy the constraints (3.8)-(3.12).

Objective function

For a given state, S_t, and decision, x_t, at time t we define the contribution function, $C_t(S_t, x_t)$ as

$$\begin{aligned} C_t(S_t, x_t) = f^P_{tt} \cdot (f^L_{tt} - x^{\text{spot-load}}_t + x^{\text{stor-spot}}_t - x^{\text{spot-stor}}_t - R_{tt}) \\ - \sum_{e \in \mathcal{E}} (x^{\text{ren-load}}_{t,e} \cdot P_{te}) - \sum_{g \in \mathcal{G}} (x^{\text{gen-load}}_{t,t+\tau_g,g} \cdot f^{\text{gen}}_{tt}(\tau_g)). \end{aligned} \tag{3.18}$$

The objective is to find the policy, $X^\pi(S_t)$, that solves

$$\max_{\pi \in \Pi} \; \mathbb{E} \left[\sum_{t=0}^{T} C_t(S_t, X^\pi(S_t)) \middle| S_0 \right] \tag{3.19}$$

where the state evolves according to the transition function in (3.17).

3.4 Policies

In this chapter, we propose a hybrid based on an approximate lookahead that is parameterized to provide better performance. We then use policy search to tune the parameterization. Below we begin by presenting a deterministic lookahead, and then present a parameterization designed to accommodate the errors introduced by the deterministic approximation.

Deterministic lookahead

We are particularly interested in deterministic lookahead models where exogenous information, $(W_{t'})_{t'=t}^{t+H}$, is replaced with an estimate of $W_{t'}$ made at time t, $f^W_{tt'}$. To make a distinction between the base model (which is the problem we are trying to solve) and the lookahead model, we use the same notation as in the base model, but we introduce tilde's on all the variables. Each variable carries a triple time index, $\tilde{x}_{t,t',t''}$, where t is the current point in time in the base model (equation (3.19)), which determines what we know. We may be planning a decision that might be made at time $t' \geq t$ (in the lookahead model) to be implemented at time t'' (also in the lookahead model). In addition, the decision made at time t' still has to consider the uncertainty at time t''. Given this notation, the deterministic lookahead model is defined as

$$X_t^{DLA}(S_t) = \underset{(x_{t,t''})\,\forall\, t''=t,\ldots,T}{\operatorname{argmax}} \left\{ C_t\left(S_t, (x_{tt''})_{t''=t,\ldots,T}\right) + \sum_{t'=t+1}^{T} \tilde{C}_{tt'}\left(\tilde{S}_{tt'}, (\tilde{x}_{tt't''})_{t''=t,\ldots,T}\right) \right\} \tag{3.20}$$

$\tilde{S}_{t,t'+1} = \tilde{S}^M(\tilde{S}_{tt'}, (\tilde{x}_{ttt''})_{t''=t,\ldots,T}, \tilde{f}^W_{t,t'+1})$. If the transition and contribution functions are linear functions we can represent (3.20) as the following linear program

$$X_t^{\text{DLA}}(S_t) = \underset{(x_{tt''})\,\forall\, t''=t,\ldots,T}{\operatorname{argmax}} \left\{ \sum_{t''=t}^{T} (c_{tt''} x_{tt''}) + \max_{(x_{tt't''})\,\forall\, t'>t} \left(\sum_{t'=t}^{T} \sum_{t''=t'}^{T} (\tilde{c}_{t't''} \tilde{x}_{tt't''}) \right) \right\} \tag{3.21}$$

where we model transitions in the lookahead model using $S_{t,t'+1} = S^M(S_{tt'}, \tilde{x}_t, \tilde{W}_{t,t'+1})$ and $\tilde{x}_t = (\tilde{x}_{ttt''})_{t''=t,\ldots,T}$.

For this problem we define a deterministic lookahead model that is constructed by substituting all of the exogenous information in (3.19) with point forecasts. We adapt the linear contribution function (3.18) to use the triple index notation,

$$\tilde{C}_{tt'}\left(\tilde{S}_{tt'},\tilde{x}_t\right) = \tilde{f}^P_{tt'}\cdot(\tilde{f}^L_{tt'} - \tilde{x}^{\text{spot}}_{tt't'} + \tilde{x}^{\text{stor-spot}}_{tt't'} - \tilde{x}^{\text{spot-stor}}_{tt't'}) \\ - \sum_{e\in\mathcal{E}}(x^{\text{ren-load}}_{tt't',e}\cdot P_{t'e}) - \sum_{g\in\mathcal{G}}(x^{\text{gen-load}}_{t,t',t'+\tau_g,g}\cdot P^{\text{gen}}_t(\tau_g)), \tag{3.22}$$

where $\tilde{x}_t = (\tilde{x}_{tt't''})_{t''\in[t',t'+H]}$. We formally define $\tilde{S}_{tt'}$ as

$$\tilde{S}_{tt'} = \left((\tilde{f}^L_{tt'})_{t'=t,\ldots,t+H}, (\tilde{f}^E_{tt'})_{t'=t,\ldots,t+H}, (\tilde{f}^P_{tt'})_{t'=t,\ldots,t+H}, (\tilde{R}^{\text{gen}}_{tt't''})_{t''=t',\ldots,t+H}\right), \tag{3.23}$$

where $\tilde{R}^{\text{gen}}_{tt't''}$ represents the amount of committed energy from generators at time t' in the lookahead model to be used at t'' in the lookahead model generated at time t (in the base model). We define $\tilde{R}^{\text{stor}}_{tt'}$ as the estimate of electricity in storage at time t' given the information available at time t. Given this notation we define the deterministic policy, $X^{\text{D-LA}}_t$, as

$$X^{\text{D-LA}}_t = \underset{\{x_{tt'}\ t'=t,\ldots,t+H\}}{\operatorname{argmax}} \left\{C_t\left(S_t,(x_{tt'})_{t'\in[t,t+H]}\right) + \sum_{t'=t+1}^{t+H}\tilde{C}_{tt'}\left(\tilde{S}_{tt'},(\tilde{x}_{tt't''})_{t''\in[t,t+H]}\right)\right\} \tag{3.24}$$

subject to

$$\tilde{R}^{\text{gen}}_{t,t'+1,t''} = \tilde{R}^{\text{gen}}_{t,t',t''} + \sum_{g\in\mathcal{G}}\tilde{x}^{\text{gen-load}}_{tt't''g}\ \forall\ t''=t'+1,...,t+H, \tag{3.25}$$

$$\tilde{R}^{\text{stor}}_{t,t'+1} = \tilde{R}^{\text{stor}}_{tt'} + \tilde{x}^{\text{spot-stor}}_{tt't'} + \sum_{e\in\mathcal{E}}\tilde{x}^{\text{ren-stor}}_{tt't'e} - \tilde{x}^{\text{stor-load}}_{tt't'} - \tilde{x}^{\text{stor-spot}}_{tt't'}\quad t'\in[t,t+H], \tag{3.26}$$

$$0 \le \tilde{x}^{\text{gen-load}}_{tt't''g} \le R^{\max}_g\ \forall\ t\le t'\le t''\le t+H \text{ and } \forall g\in\mathcal{G}, \tag{3.27}$$

$$\tilde{x}^{\text{ren-load}}_{tt't'e} + \tilde{x}^{\text{ren-stor}}_{tt'te} \le \tilde{f}^E_{tt'e}\ \ \forall\ \ t'\in[t,t+H] \text{ and } \forall\ e\in\mathcal{E}, \tag{3.28}$$

$$0 \leq \tilde{x}^{\text{spot-stor}}_{tt't} + \sum_{e \in \mathcal{E}} \tilde{x}^{\text{ren-stor}}_{tt'te} \leq \min(\gamma^c R^{\max}, R^{\max} - R^{\text{stor}}_{tt'}) \quad t' \in [t, t+H], \tag{3.29}$$

$$0 \leq \tilde{x}^{\text{stor-load}}_{tt't} + \tilde{x}^{\text{stor-spot}}_{tt't} \leq \min(\gamma^d R^{\max}, R^{\text{stor}}_{tt'}) \quad t' \in [t, t+H], \tag{3.30}$$

$$\tilde{f}^L_{tt} \leq \sum_{e \in \mathcal{E}} \tilde{x}^{\text{ren-load}}_{tt't'e} + \tilde{R}^{\text{gen}}_{tt't'} + \tilde{x}^{\text{spot-load}}_{tt't'} + \tilde{x}^{\text{stor-load}}_{tt't'} \quad t' \in [t, t+H]. \tag{3.31}$$

Equations (3.25) are transition constraints for energy generated by energy generators. Constraint (3.26) is a transition constraint for the energy level of the storage device. The constraint (3.27) limits how much energy can be ordered from generator g during a single period. The constraint (3.28) limits the policy from using more intermittent energy sources than forecasted. Constraints (3.29) and (3.30) determine how much energy can be deposited and withdrawn from the storage device. The constraint (3.31) guarantees that the load is always satisfied.

Parametric Cost Function Approximation

The obvious weakness of the deterministic lookahead policy is that it does not account for the uncertainties in the forecasts. This means we are counting on them being correct, and ignoring problems when they are too high (an issue for the prices or loads) or too low (such as the energy from renewables). To overcome this weakness, we introduce the idea of using a parametric cost function approximation (or CFA), where the deterministic lookahead is modified in a way that we hope will better accommodate uncertainty. While we could modify either the cost function or the constraints, in this paper we are going to use constraint modifications, which are then optimized using our stochastic base model in equation (2.33). We then train the parametrically modified policy using a data-driven approach that makes it possible to handle arbitrary richness in the underlying stochastic process.

The benchmark policy, defined by equations (3.24)-(3.31), is easily converted into a parametric cost function approximation by modifying existing constraints and/or appending new parametric constraints. The selection of these parametric constraints parallel the selection of any statistical model. The structure of the model is an *art* that draws on the knowledge and imagination of the modeler. Once the structure of the model has been determined, finding the best parameter values, is the science which draws on the power of classical search algorithms. For the previously described lagged energy storage problem we propose two parametric constraint modifications. We modify the constraints controlling how much energy can be ordered from the energy generators, known as *order constraints*, to mitigate the effects of uncertainty in spot prices, renewables and loads. Specifically, we modify the order constraint (3.27), which limits how much energy from energy generator g can be ordered in advance, with the parameter $\theta^{\text{gen}} = (\theta^{\text{gen}}_1, ..., \theta^{\text{gen}}_H) \in [0,1]^H$ such that

$$0 \leq \tilde{x}^{\text{gen-load}}_{ttt'g} \leq \theta^{\text{gen}}_{t'-t} R^{\max}_g. \tag{3.32}$$

Overestimating or underestimating forecasts of renewable energy influences how aggressively a policy will store or order energy for the future. We parameterize the forecasts of renewable energy sources $e \in \mathcal{E}$ in constraint (3.28) with the parameter, $\theta^{\text{ren}}_e = (\theta^{\text{ren}}_{1,e}, ..., \theta^{\text{ren}}_{H,e}) \in [0,\infty)^H$ such that

$$\tilde{x}^{\text{ren-load}}_{tt't'e} + \tilde{x}^{\text{ren-stor}}_{tt'te} \leq \theta^{\text{ren}}_{t'-t,e} \tilde{f}^E_{tt'e} \quad \forall \;\; t' \in [t, t+H] \text{ and } \forall \;\; e \in \mathcal{E}, \tag{3.33}$$

Constraints (3.27) and (3.28) can be parameterized multiple ways, below are a few:

- **Lookup table lagged price function parameterization** - This parameterization uses a unique parameter, $\theta^{\text{gen}}_{t'-t}$ and/or $\theta^{\text{ren}}_{t'-t}$, to modify each period in the lookahead model.

- **Linear lagged price function parameterization** - Searching through an H-dimensional parameter space is computationally expensive. If we make $\theta^{\text{gen}}_{t'-t}$ and/ or $\theta^{\text{ren}}_{t'-t}$ a function of the *time until implementation*, $t'-t$, we can significantly reduce the number of tunable policy parameters and policy search time. The linear parameterization defines the parameters $\theta_{t'-t}$ as a linear function of the parameter $\theta = (\theta_1, \theta_2)$ such that

$$\theta_{t'-t} = \theta_1 + \theta_2 \cdot (t' - t). \tag{3.34}$$

- **Exponential lagged price function parameterization** - This parameterization defines $\theta_{t'-t}$ as an exponential function of the two parameters $\theta = (\theta_1, \theta_2)$ such that

$$\theta_{t'-t} = \theta_1 e^{\theta_2 \cdot ('t-t)}. \tag{3.35}$$

  This exponential curve dramatically reduces the number of parameters to be fitted.

3.5 The Algorithm

Parametric cost function approximations are members of the general class of parameterized models, $X^\pi_t : \mathcal{S}_t \times \Theta \to \mathcal{X}_t$. The optimal parameterization, θ^*, of policy, X^π_t, can be found by solving

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmax}} \; F(\theta), \tag{3.36}$$

where

$$F(\theta) = \mathbb{E}\left[\sum_{t=0}^{T} C_t(S_t, X_t^\pi(S_t|\theta)) \mid S_0\right] \tag{3.37}$$

and where $S_{t+1} = S^M(S_t, X_t^\pi(S_t|\theta), W_{t+1}^n)$. This problem can be solved using an iterative stochastic gradient algorithm such as the CFA Gradient algorithm, described in Algorithm 1, (see Spall et al. (2003)).

The gradient

There are multiple ways to approximate $\nabla_\theta F$ for Algorithm 1 (See Spall et al. (2003)). Given the following assumptions we can calculate an estimator of $\nabla_\theta F$ using a single sample path ω:

A1) The cumulative reward of a single sample path, $\bar{F}(\cdot, \omega)$, is uniformly bounded for every $\omega \in \Omega$.

A2) $\bar{F}(\theta, \omega)$, is differentiable for every $\theta \in \Theta$.

If the cumulative reward function, $F(\cdot)$, satisfies conditions (A1) and (A2) then an estimator of $\nabla_\theta F(\cdot)$ is computed as

$$\nabla_\theta \bar{F} = \left(\frac{\partial C_0}{\partial X_0} \cdot \frac{\partial X_0}{\partial \theta}\right) + \sum_{t'=1}^{T}\left[\left(\frac{\partial C_{t'}}{\partial S_{t'}} \cdot \frac{\partial S_{t'}}{\partial \theta}\right) + \left(\frac{\partial C_{t'}}{\partial X_{t'}(S_t|\theta)} \cdot \left(\frac{\partial X_{t'}(S_t|\theta)}{\partial S_{t'}} \cdot \frac{\partial S_{t'}}{\partial \theta} + \frac{\partial X_{t'}(S_t|\theta)}{\partial \theta}\right)\right)\right], \tag{3.38}$$

where

$$\frac{\partial S_{t'}}{\partial \theta} = \frac{\partial S_{t'}}{\partial S_{t'-1}} \cdot \frac{\partial S_{t'-1}}{\partial \theta} + \frac{\partial S_{t'}}{\partial X_{t'-1}(S_{t-1}|\theta)} \cdot \left[\frac{\partial X_{t;-1}(S_{t-1}|\theta)}{\partial S_{t'-1}} \cdot \frac{\partial S_{t'-1}}{\partial \theta} + \frac{\partial X_{t'-1}(S_{t-1}|\theta)}{\partial \theta}\right]. \tag{3.39}$$

If the policy is a linear program such that

$$X_t^\pi(S_t|\theta) = \underset{x_t, (\tilde{x}_{tt'}), t'=t+1,\ldots,t+H}{\operatorname{argmin}} \quad c_t x_t + \sum_{t'=t+1}^{t+H} \tilde{c}_{tt'} \tilde{x}_{tt'} \tag{3.40}$$

where $\tilde{A}_t\tilde{x}_t \leq \tilde{b}_t(\theta, S_t)$, B_t is the basis matrix corresponding to the basic variables for the optimal solution of (3.40), and $\tilde{x}_t^T = [\tilde{x}_{tt}, \ldots, \tilde{x}_{tT}]$. Then an estimator of $\nabla_\theta F(\cdot)$ can be computed as

$$\nabla_\theta \bar{F}(\theta, \omega) = \sum_{t=1}^{T} \left(\nabla_\theta \tilde{b}_t(\theta, S_t) + \nabla_{S_t} \tilde{b}_t(\theta, S_t) \cdot \nabla_\theta S_t \right)^T \cdot \left(B_t^{-1} \right)^T \cdot c_t, \tag{3.41}$$

where

$$\nabla_\theta S_t = \nabla_{S_{t-1}} S^M(S_{t-1}, x_{t-1}, W_t) \cdot \nabla_\theta S_{t-1} + \nabla_{x_{t-1}} S^M(S_{t-1}, x_{t-1}, W_t) \cdot \nabla_\theta x_{t-1}. \tag{3.42}$$

The stepsize

Given an estimator and an appropriately selected stepsizes $(\alpha_n)_{n=0}^{\infty}$, Algorithm 1 will converge to a stationary point of $F(\cdot)$. Specifically, the sequence of (possibly stochastic) stepsizes, $(\alpha_n)_{n=0}^{\infty}$, must satisfy the following:

B1) $\alpha_n > 0$, a.s.

B2) $\sum_{n=0}^{\infty} \alpha_n = \infty$, a.s.

B3) $\mathbb{E}\left[\sum_{n=0}^{\infty} (\alpha_n)^2\right] < \infty$.

If $F(\cdot)$ is convex in Θ, Algorithm 1 will be a global minimum, but if not, it could be a local minimum or even a saddle point. This is true asympototically, but in practice we prefer parameterized rules that can be tuned for quicker convergence rates. Therefore, we limit our evaluation of the algorithm to how well it does within N iterations. We evaluated several stepsize rules and found the Adaptive Gradient Algorithm (AdaGrad) to be the best. The main benefit of AdaGrad is that it features self-scaling (similar to the gain matrix of the Kalman filter), which eliminates the need to manually tune the learning rate for each dimension of θ. Instead, we only have to tune a single scalar η. Additionally, AdaGrad theoretically has tighter regret bounds than standard stepsize algorithms (see Duchi et al. (2011b)).

AdaGrad modifies the individual step size for the updated parameter, θ, based on previously observed gradients using

$$\alpha_n = \frac{\eta}{\sqrt{G_t + \epsilon}} \tag{3.43}$$

where η is a scalar learning rate, $G \in \mathbb{R}^{d \times d}$ is a diagonal matrix where each diagonal element is the sum of the squares of the gradients with respect to θ up to the current iteration n, while ϵ is a smoothing term that avoids division by zero. For our simulations we set $\eta = 0.1$ and $\epsilon = 10^{-8}$.

Stochastic Updating

It is important to emphasize that the stochastic gradient equations (3.38)-(3.41) are data-driven, which means they can use any stochastic process. We do not require a formal stochastic model, and nor do we impose any restrictions on the structure of the stochastic process. The data may exhibit arbitrary inter-temporal correlations. In addition, the noise in the process may decrease or increase in time, which is something that happens in practice. For example, we might have a process where the single-period lookahead forecast is perfect. If the cost of waiting until the last minute is low, the method will learn to wait until the last minute to take advantage of low costs along with a high quality forecast.

3.6 Experimental Testing

In this section, we compare the deterministic lookahead model, defined by equation (3.24), to several parametric CFA policies tuned by Algorithm 1, using Monte Carlo simulation on different variations of the lagged energy storage problem, defined by equations (3.17) - (3.19). We test the following parameterizations of the deterministic lookahead model, defined by equation (3.24):

- **Constraint (Lookup)** - This parameterization uses the vector, $\theta^{\text{gen}} = (\theta_1^{\text{gen}}, ..., \theta_H^{\text{gen}}) \in [0,1]^H$, where θ_h^{gen} modifies the order constraint (3.27) for each period of the lookahead model. This parameterization directly limits how much energy can be ordered in advance.

- **Forecast (Lookup)** - This parameterization uses the vector, $\theta_e^{\text{ren}} = (\theta_{1,e}^{\text{ren}}, ..., \theta_{H,e}^{\text{ren}}) \in [0,\infty)^H$, where $\theta_{h,e}^{\text{ren}}$ factors the forecasts of renewable energy sources $e \in \mathcal{E}$ in constraint (3.28) for each period of the lookahead model.

- **Both (Lookup)** - This parameterization uses unique parameters, $\theta_{t'-t}^{\text{gen}}$ and $\theta_{t'-t}^{\text{ren}}$, to modify both the order constraints and forecasts of renewable energy sources for each period in the lookahead model.

We test these policies on the following lagged price function

$$\lambda^{\text{linear}}(\tau) \quad = \quad 35 - \tau\frac{15}{12}. \tag{3.44}$$

This lagged price function is designed to demonstrate the ability of our model to handle the case when the cost of purchasing electricity in advance increases and forecast errors decrease the later an order is placed. We note that forecast errors do not always decrease as forecasting horizons shorten and that these modeling assumptions regarding lagged prices are not universal. Forecasts of renewable energy sources commonly use persistence forecasting over short horizons, and meteorological models for longer horizons. We make these assumptions for demonstrative purposes. The CFA is not limited to these assumptions because the parametric CFA makes no assumptions about the structure of the stochastic errors in forecasts, or in the behavior of the advance costs.

Figure 3.5 presents the individual performance of each parameterization relative to the unmodified benchmark for a variety of forecast qualities, σ^f. When the forecast

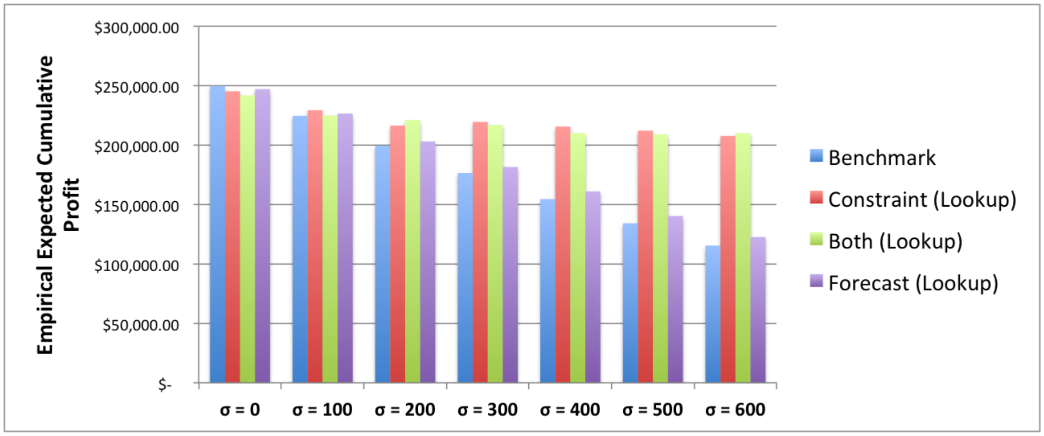
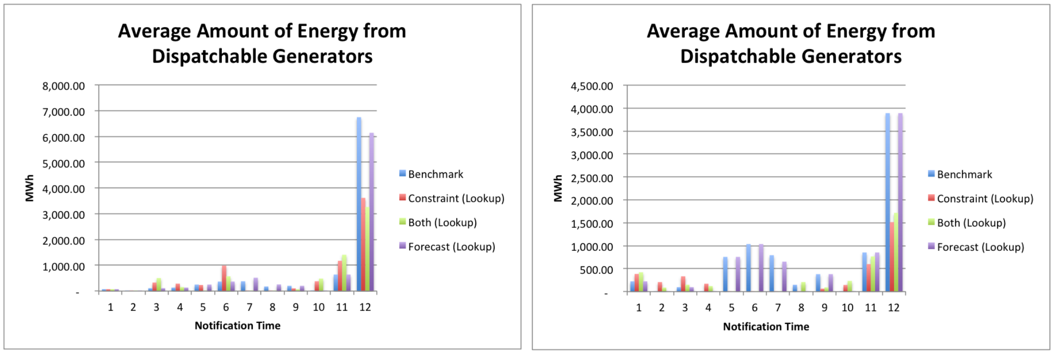


Figure 3.5: This table displays the empirical expected cumulative profit for each parameterized policy and the benchmark over varying forecast qualities, σ^f, and the lagged price function described by (3.44).

is perfect, $\sigma^f = 0$, the deterministic benchmark policy is the optimal policy. Because a policy with a perfect forecast in this scenario will never order a surplus, the order constraints for all the periods of the lookahead model, except for the first two periods, are non-binding. As long as the order constraints that modify the last two periods of the lookahead model are 1.0, the parameters that modify the order constraints of the other periods can be any value greater than or equal to 0 and the policy will still be optimal. The policies that parameterize just the forecasts of future renewable energy sources, described by equation (3.33), are also identical to the deterministic benchmark given a perfect forecast and all the parameters are equal to one.

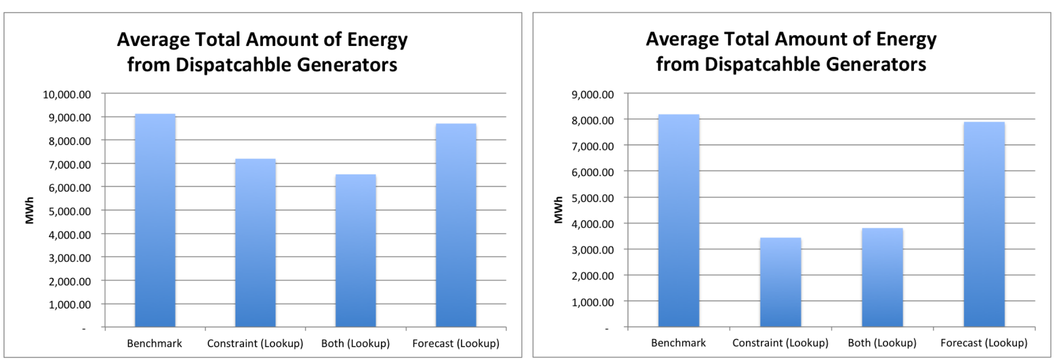
As the amount of uncertainty in the forecast, σ^f, increases, the unmodified lookahead model is no longer optimal. Figure 3.6 presents the distribution energy commitments for the different policies when $\sigma^f = 300$ and $\sigma^f = 600$. These results demonstrate the inability of the unmodified lookahead model to balance the benefit of ordering inexpensive energy against the risk of using poor forecasts. Specifically, we notice the parameterized policies avoid the bias to make early energy commitments and have more dispersed distributions of energy commitments than the unmodified benchmark policy. This is the same behavior observed in the multi-period newsvendor example in section 3.2. By waiting later to purchase energy these policies avoid



(a) $\sigma^f = 200$ (b) $\sigma^f = 600$

Figure 3.6: This table displays the distribution of energy commitments for all of the tested policies for varying degrees of forecast quality and when the lagged price function is described by equation (3.44). The graphs show that the benchmark deterministic lookahead orders farther into the future than the parameterized lookup policies.

having a surplus of energy they cannot use. In addition to ordering energy later, the parameterized policies also rely more heavily on other sources of energy than just the dispatchable energy generators, $\mathcal{G}$. Figure 3.7 shows how as σ^f increases the parameterized policies order less energy from the dispatchable generators. Instead, these policies rely more heavily on electricity from the spot market.



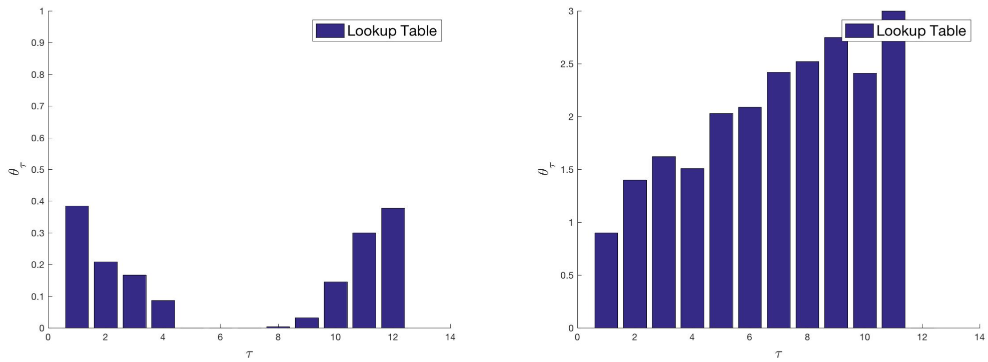
(a) $\sigma^f = 200$ (b) $\sigma^f = 600$

Figure 3.7: This table displays the average total amount of energy ordered from dispatchable generators for all of the tested policies for varying degrees of forecast quality and when the lagged price function is described by equation (3.44).

Figure 3.8a shows that the policies with parameterized order constraints, described by equations (3.32) - (3.35), avoid surplus by reducing order constraints in the lookahead model such that $\theta^{\text{gen}}_{t'-t} < 1 \; \forall \; t' \in [t, t+H]$. It is interesting to note the parabolic shape of the parameters that modify the order constraints, $(\theta^{\text{gen}}_{t'-t})^{t+H}_{t'=t}$. By gradually decreasing the order constraints of the last four periods of the lookahead model these policies are able to take advantage of the low prices associated with long notification times and avoid the risk of ordering too much. These policies also eliminate the ability to order energy four to eight periods in advance because it is less expensive, on average, to purchase electricity from the spot market, at $25.78 per MWh. Another interesting observation is that as σ^f increases these policies do not experience a significant decrease in expected cumulative profit because the parameters θ^{gen} do not modify a forecast. Hence, the optimal order constraint parameterization, θ^{gen}, seems to be the same for all $\sigma^f > 200$.

Figure 3.8b shows the policy with parameterized forecasts of renewable energy supply are modified such that $\theta^{\text{ren}}_{t'-t} > 1 \quad \forall \quad t' \in [2, t+H]$. Consequently, the parameterized policy orders less energy in advance from energy generators because it expects to satisfy a larger portion of its future load with renewable energy sources in the future. It is important to note these are not frequently binding constraints. Thus, these constraints do not influence policy decisions nearly as much as the order constraints. Therefore, they are not as effective at reducing the bias to make early energy commitments.

These experiments also show that the inventory parameterization in equation (3.32) is much more effective that the parameterization of the forecasts in equation (3.33). Finding the best parameterization is comparable to finding the best model specification in a statistical model. Our algorithms can find the best parameters for a given model specification, but at the moment we need intuition, as is the case in



(a) Order constraint parameter values (b) Forecast parameter values

Figure 3.8: Parameter values of parameterized policies when the lagged price function is the trivial lagged price function described by equation (3.44) and $\sigma^f = 60$.

designing statistical models, to design the structure of the parameterization. This is clearly an area for further research.

3.7 Conclusion

In this chapter, we describe a finite-horizon lagged resource problem that is subject to stochastic prices and renewable sources, with the purpose of demonstrating the capability of parametrically modified deterministic lookahead models, known as Cost Function Approximations (CFA). These parameterized deterministic policies allows us to exploit the structural properties of sequential decision problems while also capturing their complex dynamics. Consequently, these models are capable of handling much greater richness than are typically captured in stochastic lookahead models (Birge and Louveaux (2011a)). We demonstrate the efficacy of this approach to lagged resource allocation problems with the multi-period newsvendor problem and a significantly more complex energy portfolio management problem. Our numerical results illustrate the abilities of the CFA to handle uncertainty without restrictions on the structure of the dynamics. Specifically, we show the ability of the CFA to balance the tradeoff between forecast reliability and cost for a library of lagged energy portfolio problems.

Chapter 4

An Optimization Model for Natural Gas Supply Portfolios of an Industrial Gas Producer

4.1 Introduction

Industrial gas companies have to manage the volatility of electricity and natural gas prices in the process of converting air into products such as purified oxygen and nitrogen. These sources of uncertainty are especially important because industrial gas production processes require large amounts of energy. In addition to energy related risks, producers deal with many other sources of uncertainty including customer demands and transmission outages.

In this chapter we propose a detailed, dynamic model of a network of industrial gas production plants, a hydrogen storage cavern, a diverse set of customers, and electricity and natural gas commodity markets. We pay special attention to the planning of lagged decisions in the management of monthly and daily natural gas deliveries and contracts. This problem requires anticipating decisions that might be made in the future. While approximate dynamic programming is a powerful tool for a wide range of problems, lagged problems remain a problem class that has resisted this solution approach.

We demonstrate how our model can be used to test and identify hydrogen storage policies for minimizing the probability of a production shortage. We also demonstrate the robustness of these strategies on a library of problems. As with any industry, suppliers must control their costs, which means managing both average costs and the risks of volatility in costs. Because of their high dependence on natural gas and electricity, industrial gas suppliers are specifically interested in maximizing daily cashflows and minimizing the dependence of those cashflows on energy commodity prices. We demonstrate how this model can be used to test and identify natural gas nomination strategies to minimize the dependence of the daily profit of the supplier and energy commodity market. We also demonstrate the robustness of these strategies for different degrees of natural gas spot market volatility.

This chapter is organized as follows. Section 4.2 provides a brief literature review of natural gas supply, energy portfolio management, and energy related stochastic optimization. Section 4.3 presents a dynamic model of an industrial gas producer with access to a hydrogen storage cavern, a natural gas hub, an electricity grid, and a diverse set of customers. In this section we formally describe the problem by defining the five elements of a stochastic optimization problem according to Powell (2011). We also make a special effort to distinguish between the stochastic base model which can span from several months to several years, and the daily and monthly operating policies. Section 4.4 describes a parametric operating policy and a policy search algorithm to tune it. Section 4.5 presents a series of experiments to demonstrate the use of our model and analyze the performance of solutions under varying operating conditions. Finally, section 4.6 concludes the chapter.

4.2 Literature Review

The related literature encompasses multiple areas of research including natural gas supply, energy portfolio management, inventory management, and the wide range of energy related stochastic optimization problems (see Pilipovic (1998), Silver et al. (1998), and Powell and Meisel (2015)). We consider not only literature that focuses specifically on industrial gas producers, but the broader collection of similar optimization problems as well. Given its analogous problem structure, we pay particular attention to the problem of managing short term natural gas contracts for combined cycle power plants operating in deregulated electricity markets. Chen and Baldick (2007) proposes a utility-maximization-based policy to optimize a short-term natural gas supply portfolio for a natural gas fired power plant. The approach considers the financial risks associated with energy commodities and adjusts the natural gas supply of the electric utility company to satisfy a designated risk preference. Jirutitijaroen et al. (2013) uses a two-stage stochastic program to manage a portfolio of short term natural gas contracts for a power plant with a stochastic recurring customer demand. Takriti et al. (2000) presents a stochastic model for a unit commitment problem with extremely volatile electricity spot prices that incorporates power trading with fuel constraints.

There exists a broader area of research on the valuing and optimal trading of natural gas storage contracts (see Lai et al. (2010), Secomandi (2010), Lai et al. (2011), and Löhndorf and Wozabal (2017)). Secomandi (2010) shows the optimal trading policy for a risk-neutral commodities merchant is a two stage base stock policy. Though this paper uses natural gas as an example commodity, the paper focuses on commodity storage in general. This community tends to consider only the perspective of commodity merchants with no intention of using the commodity outside of trading. Lai et al. (2011) develops a heuristic model to value real options to store liquefied natural gas (LNG) that incorporates shipping, price evolution, inventory control, and sales in the wholesale natural gas market. Secomandi et al. (2010) studies model selection methods for natural gas storage real option price models and hedging strategies given modeling errors. All of the previous literature ignores the combined problem of securing natural gas through financial markets for the purpose of satisfying an exogenous stochastic demand.

Stochastic programming applied to energy related problems is a widely studied area of research (see Jirutitijaroen et al. (2013), Carpentier et al. (2015), and Singh et al. (2009)). These approaches use large scenario trees to approximate potential future events, but result in very large-scale optimization models that can be quite hard to solve in practice. In this chapter we forgo the scenario tree approach and use a modified deterministic model to account for problem uncertainty. Perkins and Powell (2017) formalizes the idea of modeling and tuning parametrically modified deterministic optimization models known as parametric cost function approximations. The authors argue for an approach that shifts the modeling of stochastics from an approximate of a lookahead model to the stochastic base model, which is typically implemented in a simulator (but might also be the real world). Tuning a parametric model in a stochastic simulator makes it possible to handle arbitrarily complex dynamics. This additionally allows us to avoid the many approximations (such as two-stage models, exogenous information that is independent of decisions) that are standard in stochastic programming. Perkins and Powell (2017) uses gradient based methods to determine policy parameters. Their method builds upon the Robbins-Monro algorithm (Robbins and Monro, 1951a) and a rich literature of stochastic gradient algorithms and their applications to Markov decision processes (Spall et al., 2003) and policy search (see Peshkin et al. (2000b), Ng and Jordan (2000b), Hu et al. (2007b), and Deisenroth (2011)).

4.3 The stochastic base model

The base model steps forward in daily increments, optimizing over a year long horizon. It determines how the supplier manages its collection of industrial gas producing plants to satisfy the recurring daily stochastic demand of customers. The supplier has multiple individual customers for each type of industrial gas. We denote the set of industrial gases as $g \in \mathcal{G}$ where $\mathcal{G} = \{H, N, O, \text{steam}\}$ and use the notation $\mathcal{J}^g$ to represent the set of customers demanding gas g. Each of these customers has a unique contract with the supplier which determines their marginal costs as a function of either daily or monthly natural gas prices. To satisfy the demand of customers the supplier has access to three sets of production plants including Steam Methane Reformers (SMRs), air separation units (ASUs), and cogeneration units which we denote as $\mathcal{U}^{\text{SMR}}$, $\mathcal{U}^{\text{ASU}}$, and $\mathcal{U}^{\text{cogen}}$ respectively. Hydrogen customers, $\mathcal{J}^H$, may be supplied by any SMR unit in $\mathcal{U}^{\text{SMR}}$. Either oxygen or nitrogen customers can be supplied by any ASU in $\mathcal{U}^{\text{ASU}}$. However, each steam consumer can only be supplied by a single cogeneration unit. Thus, the set of indices denoting steam customers and the set of cogeneration units are equivalent, $\mathcal{J}^{\text{steam}} = \mathcal{U}^{\text{cogen}}$.

In addition to satisfying the daily demand of customers, the industrial gas supplier must purchase enough natural gas and electricity to fuel their production processes. The supplier must decide once a month how much natural gas to purchase through forward contracts to be delivered in uniform daily increments. Once the supplier has purchased the natural gas they must decide whether to have it delivered via an interruptible or uninterruptible pipeline. The interruptible pipeline carries the risk of random delivery interruptions which will force the industrial gas supplier to purchase natural gas from the spot market at a potentially higher price. The uninterruptible pipeline does not have the risk of nondelivery, but is more expensive to use. The supplier may use a combination of both forms of delivery. Figure 4.1 visually describes the dynamics of this problem.

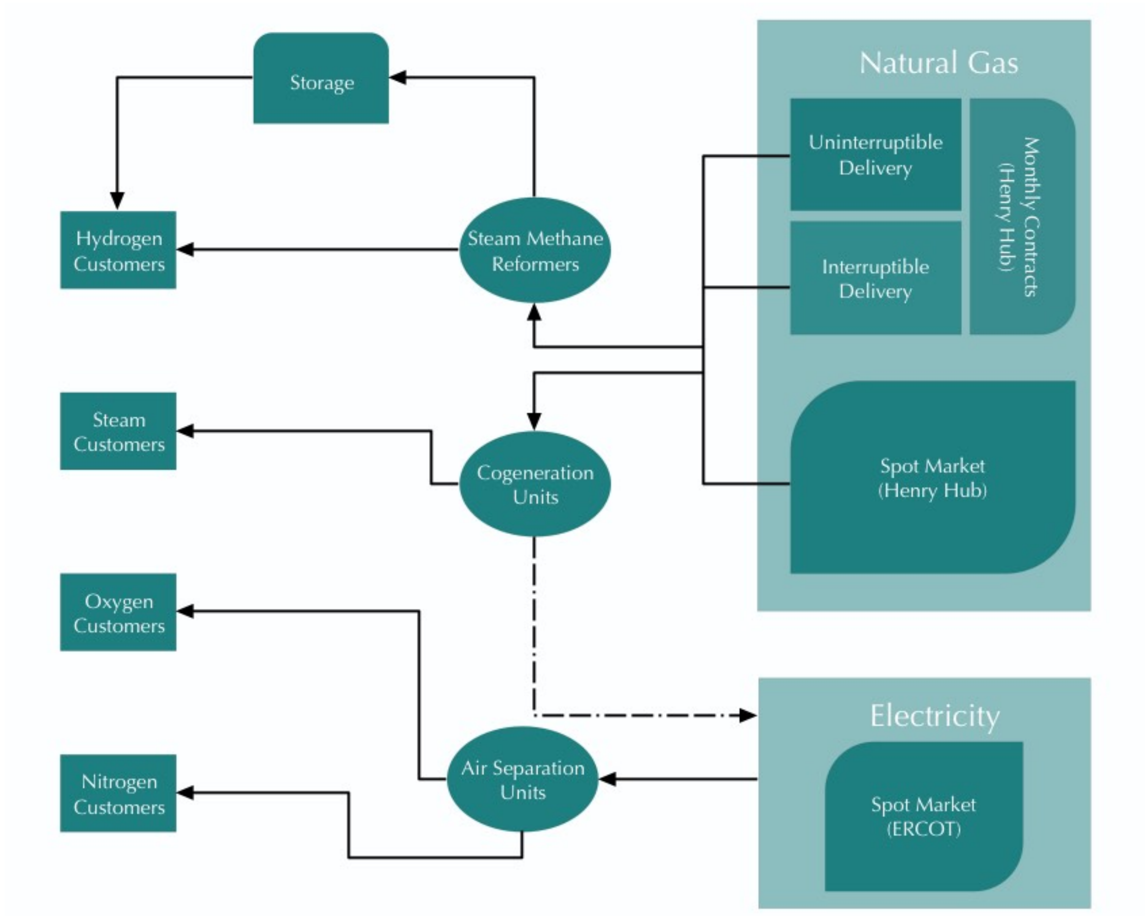


Figure 4.1: This diagram represents the industrial gas supplier's problem of maximizing expected profits while always satisfying their customer's random daily demands. The supplier has access to multiple plants to satisfy these demands. The air separation units run entirely off of electricity purchased from ERCOT's spot market. The steam methane reformers and cogeneration units require natural gas instead. Natural gas can be purchased in the day ahead market or through monthly contracts. The cogeneration units may also be used to produce electricity that is sold back to ERCOT.

We define the following notation:

Indices: $(\mathcal{G}, (\mathcal{J}_g)_{g\in\mathcal{G}}, (\mathcal{U}_g)_{g\in\mathcal{G}}, t, \mathcal{M}_m)$

$\mathcal{G}$ = The set of industrial gases, $\mathcal{G} = \{H, N, O, \text{steam}\}$, where $g \in \mathcal{G}$ is the generic index for the set,

$\mathcal{J}^g$ = The set of customers, $\mathcal{J}^g$, that demand industrial gas g, where $j \in \mathcal{J}^g$ is the generic index,

$\mathcal{U}^{\text{SMR}}$ = The set of steam-methane reformers (SMRs) producing hydrogen for customers, $\mathcal{J}^H$,

$\mathcal{U}^{\text{ASU}}$ = The set of air separation units (ASUs) producing oxygen and nitrogen for customers, $\mathcal{J}^O \cup \mathcal{J}^N$,

$\mathcal{U}^{\text{cogen}}$ = The set of cogeneration units producing steam for customers, $\mathcal{J}^{\text{steam}}$,

t = The day for which we are solving the problem. We denote the day of the month and the month of year that t represents with the functions $d(t)$ and $m(t)$, respectively,

$\mathcal{M}_m$ = The set of all days that fall within the month m. Formally, we define this set as $\mathcal{M}_m = \{t : m(t) = m\}$,

$|\mathcal{M}_m|$ = The number of days in the month m.

Policy parameter: θ^L

θ^L = This parameter guarantees there a minimum amount of hydrogen in storage to satisfy customer demand in case there is an SMR outage.

Static State Variables:

$(\gamma_i^{\textbf{SMR}}, \gamma_i^{\textbf{ASU}}, \gamma_i^{\textbf{cogen}}, \beta^{\textbf{ele}}, \phi_i^{\textbf{SMR}}, \phi_i^{\textbf{ASU}}, R^{H,\max})$

γ_i^{SMR} = The conversion efficiency of natural gas to hydrogen of SMR $i \in \mathcal{U}^{\text{SMR}}$,

γ_i^{ASU} = The conversion efficiency of electricity to oxygen and nitrogen of ASU $i \in \mathcal{U}^{\text{ASU}}$,

γ_i^{cogen} = The conversion efficiency of natural gas to steam of cogeneration unit $i \in \mathcal{U}^{\text{cogen}}$,

β^{ele} = The conversion efficiency from steam to electricity,

ϕ_i^{SMR} = The production capacity of SMR $i \in \mathcal{U}^{\text{SMR}}$,

ϕ_i^{ASU} = The production capacity of ASU $i \in \mathcal{U}^{\text{ASU}}$,

$R^{H,\max}$ = The hydrogen cavern total capacity.

Dynamic State Variable:

$$S_t = \left(R_t^{NG-\mathbf{int}}, R_t^{NG-\mathbf{uni}}, c_t^{\mathbf{int}}, c_t^{\mathbf{int}}, P_t^{\mathbf{NG\text{-}spot}}, P_t^{\mathbf{NG\text{-}fut}}, D_{t,j}, \mathbb{1}_t^{\mathbf{int}}, \left(\mathbb{1}_{t,i}^{\mathbf{SMR}}\right)_{i\in\mathcal{U}^H}, P_{t,j}^g, E_t\right)$$

$R_t^{NG-\text{int}}$ = The amount natural gas purchased through forward contracts at the end of the bid week, when $d(t) = 1$ and $m(t) = m$, to be delivered in uniform daily increments throughout the month, $\mathcal{M}_m$, via the interruptible pipeline,

$R_t^{NG-\text{uni}}$ = The amount natural gas purchased through forward contracts at the end of the bid week, when $d(t) = 1$ and $m(t) = m$, to be delivered in uniform daily increments throughout the month, $\mathcal{M}_m$, via the uninterruptible pipeline,

$P_t^{NG-\text{month}}$ = The marginal cost of purchasing natural gas through forward contracts at the end of the bid week, when $d(t) = 1$ and $m(t) = m$, to be delivered in uniform daily increments over the succeeding month, $\mathcal{M}_m$,

c_t^{int} = The marginal cost of purchasing and delivering natural gas in uniform daily increments over the succeeding month, $\mathcal{M}_m$, via an interruptible pipeline. This value includes the monthly index price, a premium between $(\$-.005, \$.005)$ per MMBTU of natural gas, and transportation costs. For our numerical example we use the definition

$$c_t^{\text{int}} = P_t^{NG-\text{month}} + .05$$

where interruptible transportation costs are \$.05/MMBTU and the premium is \$0,

c_t^{uni} = The marginal cost of purchasing and delivering natural gas in uniform daily increments over the succeeding month, $\mathcal{M}_m$, via an uninterruptible pipeline. For our numerical example we use the definition

$$c_t^{\text{uni}} = P_t^{NG-\text{month}} + .1$$

where interruptible transportation costs are \$.1/MMBTU and the premium is \$0,

$P_t^{\text{NG-fut}}$ = The listed futures contract price of natural gas at time t, to be delivered during the next month, $m(t)+1$,

$P_t^{\text{NG-spot}}$ = The spot price of natural gas at time t,

$D_{t,j}$ = The demand of customer $j \in \mathcal{J}^g$ at time t,

$f_{t,j}^D$ = The forecasted demand of customer $j \in \mathcal{J}^g$ at time t. These forecasts are based on the demand of customer, j, from the previous year,

$\mathbb{1}_t^{\text{int}}$ = The status of the interruptible natural gas pipeline at time t where

$$\mathbb{1}_t^{\text{int}} = \begin{cases} 0 & \text{the pipeline is interrupted} \\ 1 & \text{the pipeline is not interrupted} \end{cases}.$$

R_t^H = The amount of hydrogen stored in the salt cavern at time t. The storage capacity of the cavern is limited by $R_t^H \in [0, R^{H,\max}]$,

$P^g_{t,j}$ = The marginal revenue from selling a unit of gas g to customer $j \in \mathcal{J}^g$ at time t. Formally, we define this as

$$P^g_{t,j} = P_{0,j} \times \left([1 - \alpha_j] + \alpha_j \left(\frac{\beta_{tj}}{NG_0} \right) \right), \tag{4.2}$$

where the parameters NG_0, $P_{0,j}$, and α_j are unique for each customer $j \in \mathcal{J}^g$. The parameter NG_0 is the spot price of natural gas when the contract was signed. Parameter α_j determines how much energy prices influence prices for customer j. The parameter $\beta_{tj} = P_t^{\text{NG-spot}}$ if the prices customer j pays are indexed on daily natural gas prices. Conversely, $\beta_{tj} = P_t^{\text{NG-month}}$ if those prices are indexed on monthly natural prices,

$\mathbb{1}^{\text{SMR}}_{t,i}$ = The outage status of SMR unit $i \in \mathcal{U}^{\text{SMR}}$ where

$$\mathbb{1}^{\text{SMR}}_{t,i} = \begin{cases} 0 & \text{unit } i \text{ is down} \\ 1 & \text{unit } i \text{ is running} \end{cases} .$$

E_t = The spot price of electricity at time t,

Γ^{uni} = The daily maximum amount of natural gas that can be ordered from the uninterruptible pipeline,

μ^{SMR} = The daily minimum percentage of total production capacity, ϕ_i^{SMR}, that must be produced by every running SMR.

Decision variables:

$$\left(x_t^{NG-\mathbf{int}}, x_t^{NG-\mathbf{uni}}, \left((x_{t,i,j})_{j \in \mathcal{J}^g \,\&\, i \in \mathcal{U}^g} \right)_{g \in \mathcal{G}}, x_t^{sell}, x_t^{buy}, x_t^{NG-\mathbf{day}}, \quad x_t^{\mathbf{res}}, x_t^{\mathbf{stor}}, x_{t,i}^{\mathbf{stor\text{-}in}} \right)$$

Monthly decisions - These decisions are made only once a month, when $\{t : d(t) = 1\}$.

$x_t^{NG-\text{int}}$ = The fixed daily amount of natural gas ordered from the interruptible pipeline at the beginning of month m when $d(t) = 1$ and $m(t) = m$, to be delivered in uniform daily increments over the succeeding month, $\mathcal{M}_m$,

$x_t^{NG-\text{uni}}$ = The fixed daily amount of natural gas ordered from the uninterruptible pipeline at the beginning of month m when $d(t) = 1$ and $m(t) = m$, to be delivered in uniform daily increments over the succeeding month, $\mathcal{M}_m$. Daily deliveries are limited by the contract agreements with natural gas pipelines such that

$$x_t^{NG-\text{uni}} < \Gamma^{\text{uni}}.$$

Daily decisions - These decisions are made throughout the month, when $\{t : d(t) \geq 1\}$.

$x_{t,i,j}$ = The amount of industrial gas g produced by plant i to be sold to customer $j \in \mathcal{J}^g$, where $i \in \mathcal{U}^{\text{SMR}}$ if $g = \text{H}$, $i \in \mathcal{U}^{\text{ASU}}$ if $g \in \{\text{N}, \text{O}\}$, or $i \in \mathcal{U}^{\text{cogen}}$ if $g = \text{steam}$,

x_t^{sell} = The amount electricity to produce from the cogeneration units sold to ERCOT on the spot market,

x_t^{buy} = The amount of electricity purchased from the spot market to power the air separation units,

$x_t^{NG-\text{day}}$ = The amount of natural gas purchased from the spot market,

x_t^{res} = The amount of unused natural gas sold back to the spot market,

$x_{t,i}^{\text{stor-in}}$ = The amount of hydrogen from SMR unit $i \in \mathcal{U}^{\text{SMR}}$ deposited into storage at time t,

$x_{t,j}^{\text{stor-out}}$ = The amount of hydrogen withdrawn from storage to satisfy demand for customer $j \in \mathcal{J}_H$ at time t.

Production decisions are limited by the production capacities of the associated production plants such that

$$\sum_{j \in \mathcal{J}_H} x_{t,i,j} \leq \phi_i^{\text{SMR}},$$

$$\sum_{j \in \mathcal{J}_O} x_{t,i,j} + \sum_{j \in \mathcal{J}_N} x_{t,i,j} \leq \phi_i^{\text{ASU}}.$$

Hydrogen production is additionally constrained by production minimums requirements, μ^{SMR}, and occasional outages, $\mathbb{1}_{t,i}^{\text{SMR}}$, such that

$$\sum_{j \in \mathcal{J}_H} x_{t,i,j} + x_{t,i}^{\text{stor-in}} \leq \mathbb{1}_{t,i}^{\text{SMR}} \phi_i^{\text{SMR}} \ \forall \ i \in \mathcal{U}^{\text{SMR}},$$

$$\sum_{j \in \mathcal{J}_H} x_{t,i,j} + x_{t,i}^{\text{stor-in}} \geq \mu^{\text{SMR}} \mathbb{1}_{t,i}^{\text{SMR}} \phi_i^{\text{SMR}} \ \forall \ i \in \mathcal{U}^{\text{SMR}}.$$

The amount of hydrogen deposited and withdrawn is limited by the total and current capacity of the salt cavern. Thus,

$$\sum_{j \in \mathcal{J}_H} x_{t,j}^{\text{stor-out}} \leq R^{H,\max} - R_t^H,$$

$$\sum_{j \in \mathcal{J}_H} x_{t,i}^{\text{stor-in}} \leq R_t^H - R^{H,\max},$$

$$x \geq 0.$$

Exogenous information process: $\left(\hat{D}_{t,j}, \hat{E}_t, \hat{P}_t^{NG-\mathbf{day}}, \hat{P}_t^{NG-\mathbf{day}}, \hat{P}_t^{NG-\mathbf{fut}}, \hat{P}_t^{NG-\mathbf{month}}\right)$

The exogenous information is a random variable that captures the stochastic updating of forecasts as well as the randomly occurring equipment failure and natural gas pipeline deliveries. Let $\omega = \{W_0, W_1, ..., W_T\}$ for $\omega \in \Omega$ be a sample path and $\mathcal{F}$ a sigma-algebra on Ω. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the filtration $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_T = \mathcal{F}$. Our exogenous information, $\{W_t, t = 1, 2, ..., T\}$, is a stochastic process adapted to the filtration $\{\mathcal{F}_t, t = 0, 1, 2, ..., T\}$. This information includes the following:

$$
\begin{aligned}
\hat{D}_{t,j} &= \text{The forecast errors for the daily demand of customer } j \in \mathcal{J} \text{ at time } t,\\
\hat{E}_t &= \text{The forecast error for the electricity spot price at time } t,\\
\hat{P}_t^{NG-\text{day}} &= \text{The forecast error for the spot price of natural gas},\\
\hat{P}_t^{NG-\text{fut}} &= \text{The forecast error for the one month ahead futures contract prices},\\
\hat{P}_t^{NG-\text{month}} &= \text{The forecast error of the monthly natural gas index prices before trans-}\\
&\quad\ \text{portation costs are included},
\end{aligned}
$$

Customer demand

The supplier uses customer demands from the current month of the previous year, $f_{t,i}^D$ where $m(t) = m$, to determine how much natural gas to purchase through forward contracts for the current month, $\mathcal{M}_m$. These forecasts are not accurate enough to predict customer orders for the individual days of the month, but they do provide an estimate of how customer orders and cancellations are distributed across the month. We model the forecasted, $f_{t,i}^D$, and observed, $D_{t,i}$ customer demands with the respective processes

$$
f_{t+1,i}^D = \bar{D}_{t+1,i} + \hat{f}_{t+1,i}^D,
$$

$$
D_{t+1,i} = \bar{D}_{t+1,i} + \hat{D}_{t+1,i},
$$

where the reference series, $\bar{D}_{t+1,i}$, is produced using a mean reverting stochastic process fitted to a moving average of historical customer demand. The stochastic error terms, $\hat{D}_{t+1,i}$ and $\hat{f}^D_{t+1,i}$ are generated using the univariate crossing state hidden semi-Markov model (HSMM) from Durante et al. (2017) to capture the distribution of customer orders exceeding and falling below the reference series, termed *crossing times*. We derive our customer demand models using hypothetical data, built in collaboration with Air Liquide. Figure 4.2 shows a simulated sample path of the daily demand of a single oxygen customer.

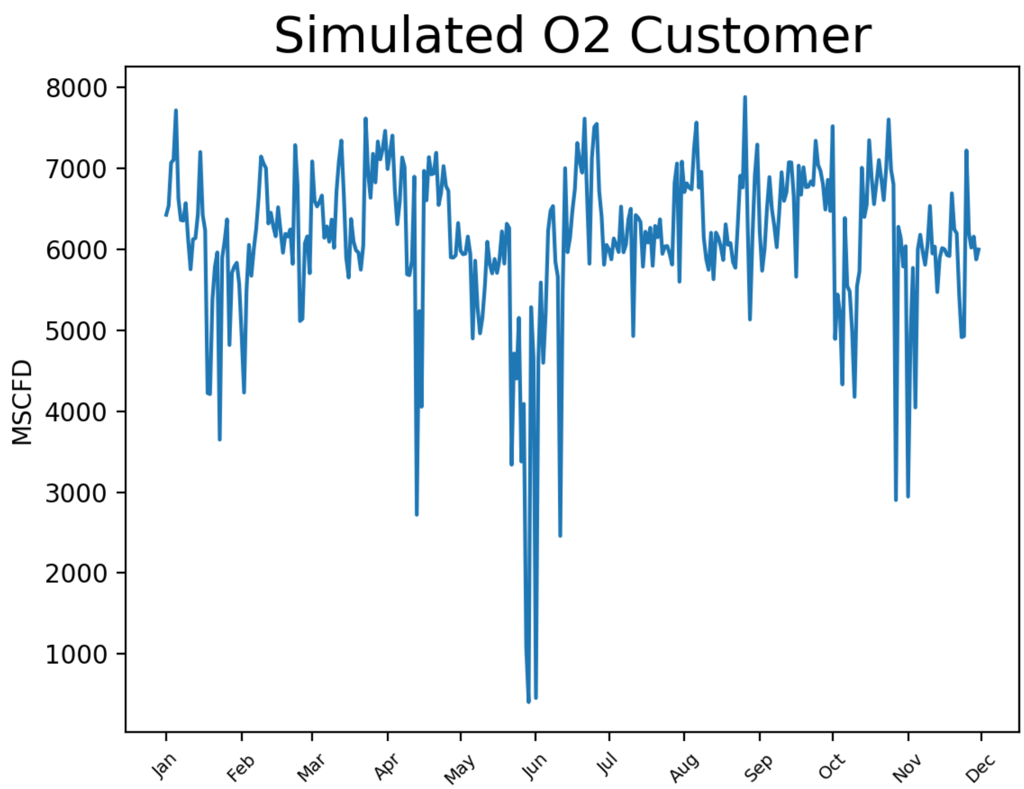


Figure 4.2: This is figure is a simulated sample path of the daily demand of a single oxygen customer.

Natural Gas Prices

We use a one-factor Schwartz model (see Schwartz (1997)) fit to historical data to simulate natural gas spot prices, $P_t^{NG-\text{spot}}$. This process models natural gas spot prices as an exponential function of an Ornstein-Uhlenbeck process given by

$$dt \log P_t^{NG-\text{spot}} = \kappa(m - \log P_t^{NG-\text{spot}})dt + \eta dW_t. \tag{4.2}$$

We then generate natural gas futures contract prices, $P_t^{NG-\text{fut}}$, such that

$$P_{t+1}^{NG-\text{fut}} = P_{t+1}^{NG-\text{spot}} + \hat{P}_{t+1}^{NG-\text{fut}},$$

where $\hat{P}_{t+1}^{NG-\text{fut}}$ is generated using a hidden semi-Markov model that captures the distribution of times the spot price falls above and below the futures price. We determine the monthly index price, $P_t^{NG-\text{month}}$, as the average futures contract price during bid week. We fit our natural gas spot and futures contract price models using 2014 - 2017 Henry Hub and 2014 - 2017 New York Mercantile Exchange (NYMEX) data, respectively.

Electricity Prices

Electricity spot prices exhibit spiky behavior and are influenced significantly by natural gas prices. To capture the complex dependence structure of electricity and natural gas prices we use the model proposed in Coulon et al. (2013) to model electricity spot prices, E_t, as exponential functions of natural gas prices, $P_t^{NG-\text{spot}}$, using the relationship

$$E_{t+1} = \alpha \exp(P_{t+1}^{NG-\text{spot}} + \beta) + \hat{E}_{t+1},$$

where instead of using a regime switching model to capture the volatile behavior of electricity spot prices we generate $\hat{E}_{t+1}$ with the crossing state model described in

Durante et al. (2017) with price distributions conditioned on different periods of the year. This allows us to imitate both the spiky and seasonal behavior of electricity spot prices. The data our models used to simulate spot prices are derived from 2014 - 2017 Electric Reliability Council of Texas (ERCOT) spot price data. Figure 4.3 shows a sample path of the simulated natural gas and electricity prices.

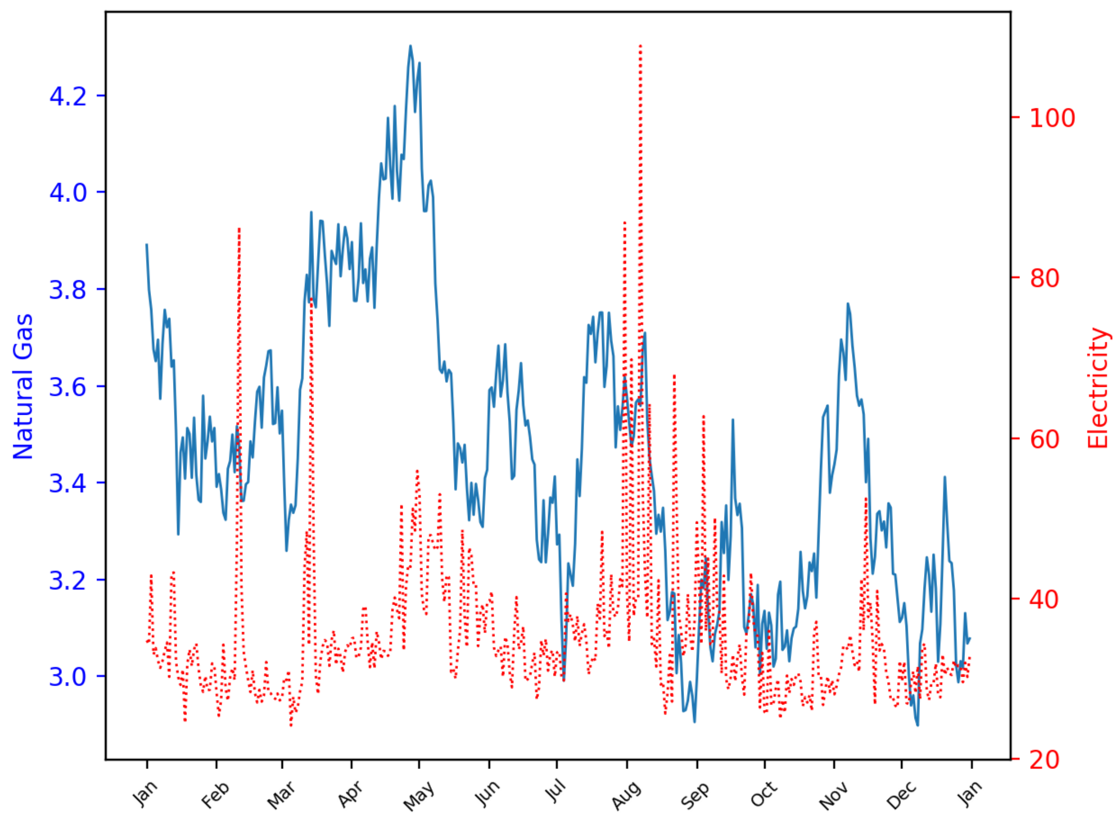


Figure 4.3: This is figure is a simulated sample path of electricity prices. Where the blue line represents the daily natural gas price and red dotted line is the daily weighted-average price price of electricity.

Transition function

The transition function, $S^M(\cdot)$, describes how each state variable evolves over time, which we designate using

$$S_{t+1} = S^M(S_t, x_t, W_{t+1}),$$

where W_t is the exogenous information. The exogenous information process W_t evolves through time according to the following equations:

$$W_{t+1} = \begin{cases} E_{t+1} & = \alpha \exp(P_{t+1}^{NG-\text{spot}} + \beta) + \hat{E}_{t+1}, \\ f_{t+1,i}^{D} & = \bar{D}_{t+1,i} + \hat{f}_{t+1,i}^{D} \quad \forall i \in (\mathcal{U}^g)_{g \in \mathcal{G}}, \\ D_{t+1,i} & = \bar{D}_{t+1,i} + \hat{D}_{t+1,i} \quad \forall i \in (\mathcal{U}^g)_{g \in \mathcal{G}}, \\ P_{t+1}^{NG-\text{fut}} & = P_{t+1}^{NG-\text{spot}} + \hat{P}_{t+1}^{NG-\text{fut}}, \\ P_{t+1}^{NG-\text{spot}} & = P_{t}^{NG-\text{spot}} + \hat{P}_{t+1}^{NG-\text{spot}}. \end{cases}$$

The amount of hydrogen in the hydrogen cavern, R_t^H, evolves according to

$$R_{t+1}^H = R_t^H + \sum_{i \in \mathcal{U}^H} x_{t,i}^{\text{stor-in}} - \sum_{j \in \mathcal{J}^H} x_{t,j}^{\text{stor-out}}.$$

And the amount of natural gas purchased through forward contracts from the uninterrupted and interrupted pipelines at the beginning of the month are set by

$$R_{t'}^{NG-\text{uni}} = \begin{cases} x_t^{NG-\text{uni}} & d(t') \geq 1 \;\&\; m(t') = m, \\ 0 & \text{otherwise}, \end{cases}$$

and

$$R_t^{NG-\text{int}} = \begin{cases} x_t^{NG-\text{int}} & d(t') \geq 1 \;\&\; m(t') = m, \\ 0 & \text{otherwise}, \end{cases}$$

where $d(t) = 1$ and $m(t) = m$.

Objective Function:

The objective of the industrial gas firm is to maximize expected cumulative profits. Given a state, S_t, and decision, x_t, at time $t = (m, d)$ the daily profit of the industrial gas supplier, $C_t(S_t, x_t)$, is defined as

$$C_t(S_t, x_t) = \begin{cases} \sum_{j \in \mathcal{J}_H} \sum_{i \in \mathcal{U}^{\text{SMR}*}} P^H_{t,j} x_{t,i,j} + \sum_{g \in \{O,N\}} \sum_{j \in \mathcal{J}_g} \sum_{i \in \mathcal{U}^{\text{ASU}}} x_{t,i,j} \left(P^g_{t,j} - \gamma^g_i E_t\right) & \\ + \sum_{j \in \mathcal{J}_{\text{steam}}} \sum_{i \in \mathcal{U}^{\text{cogen}}} x_{t,i,j} \left(P^{\text{steam}}_{t,j} + \beta^{\text{ele}}_i E_t\right) + P^{\text{NG-spot}}_t \left(.8 x^{\text{res}}_t - x^{NG-\text{day}}_t\right) & d(t) \neq 1, \\ & \\ |\mathcal{M}_m| C^{\text{int}}_t x^{NG-\text{int}}_t + |\mathcal{M}_m| C^{\text{uni}}_t x^{NG-\text{uni}}_t & d(t) = 1, \end{cases} \tag{4.3}$$

where $\mathcal{U}^{\text{SMR}*} = \{\mathcal{U}^{\text{SMR}} \cup \text{storage}\}$. The objective is find the policy, π, that solves

$$\max_{\pi \in \Pi} \ \mathbb{E}\left[\sum_{t=0}^{T} C_t(S_t, X^\pi_t(S_t)) \middle| S_0\right] \tag{4.4}$$

where $T = \sum_{m=1}^{M} |\mathcal{M}_m|$.

4.4 Operating policy

We use multiple linear programs to manage the daily and monthly operations of the network of production plants, storage, and customers. Our daily policy determines the daily flows of fuel to production plants, delivery of product to customers, and manages the inflow and outflow of hydrogen to the salt cavern. We parameterize the constraints of this linear program to determine a hydrogen reserve level for the salt cavern. The monthly policy is a deterministic lookahead model of the forecasted daily operations of the network for following month. This policy determines the amount of natural gas to purchase via future contracts. We parameterize this linear program to determine what portion of the forecasted demand for natural gas to satisfy through future contracts.

Our aggregate policy π is represented by the parameterized piecewise function

$$X_t^{\pi}(S_t) = \begin{cases} X_t^{\pi-\text{month}}(S_t|\theta) & d(t) = 1, \\ X_t^{\pi-\text{day}}(S_t|\theta) & d(t) \geq 1. \end{cases} \tag{4.5}$$

For the set of daily decisions indexed by $t \in \{t : d(t) \geq 1\}$, the policy $X_t^{\pi-\text{day}}(S_t)$ is a parameterized linear program. We provide a detailed description of the daily policy in subsection 4.4. In addition to satisfying the daily demand of customers and managing the hydrogen storage cavern the supplier must determine how much natural gas to purchase at a indexed price to be delivered in uniform daily increments over the succeeding month, $\mathcal{M}_m$. The supplier can only make this decision at the end of the bid week for each month, $\{t = (m, 1) : m = 1, ..., M\}$. Thus, when $t \in \{t = (m, d) : d = 1\}$ the policy $X_t^{\pi}(S_t)$ is a different parameterized lookahead model, which we describe in subsection 4.4. We emphasize that this policy, π is not an optimal policy, but we can obtain relatively robust behavior by tuning parameter, θ, in the base model given by equation (4.3).

The daily optimization policy $X_t^{\pi-\textbf{day}}(S_t|\theta)$

For day t where $d(t) \geq 1$, policy π is defined as the following parameterized linear program:

$$\begin{aligned} X_t^{\pi-\text{day}}(S_t|\theta) = \underset{x}{\text{argmax}} & \sum_{j \in \mathcal{J}_H} \sum_{i \in \mathcal{U}^{\text{SMR}}} P_{t,j}^H x_{t,i,j} + \sum_{g \in \{O,N\}} \sum_{j \in \mathcal{J}_g} \sum_{i \in \mathcal{U}^{\text{ASU}}} x_{t,i,j} \left(P_{t,j}^g - \gamma_i^g E_t \right) \\ & + \sum_{j \in \mathcal{J}_{\text{steam}}} \sum_{i \in \mathcal{U}^{\text{cogen}}} x_{t,i,j} \left(P_{t,j}^{\text{steam}} + \beta_i^{\text{ele}} E_t \right) + P_t^{\text{NG-spot}}(.8 x_t^{\text{res}} - x_t^{NG-\text{day}}) \\ & + \sum_{J \in \mathcal{J}_H} P_{t,j}^H x_{t,j}^{\text{stor-out}} \end{aligned} \tag{4.6}$$

This linear program is solved subject to the following constraints

$$\sum_{i \in \mathcal{U}^{\text{ASU}}} x_{t,i,j} = D_{t,j} \ \forall \ j \in (\mathcal{J}^g)_{g \in \{N,O\}}, \tag{4.7}$$

$$x_{t,j,j} = D_{t,j} \ \forall \ j \in \mathcal{J}^{\text{steam}}, \tag{4.8}$$

$$\sum_{i \in \mathcal{U}^{\text{SMR}}} x_{t,i,j} + x_{t,j}^{\text{stor-out}} = D_{t,j} \ \forall \ j \in \mathcal{J}^H, \tag{4.9}$$

$$\sum_{i \in \mathcal{U}^{\text{SMR}}} \left(\sum_{j \in \mathcal{J}_H} \gamma_i^H x_{t,i,j} + \gamma_i^H x_{t,i}^{\text{stor-in}} \right) + \sum_{j \in \mathcal{J}_{\text{steam}}} \sum_{i \in \mathcal{U}^{\text{cogen}}} \gamma_i^{\text{steam}} x_{t,i,j} = \mathbb{1}_t^{\text{int}} R^{NG-\text{int}} + R^{NG-\text{uni}} + x_t^{NG-\text{day}} - x_t^{\text{res}} \tag{4.10}$$

$$\sum_{j \in \mathcal{J}_H} x_{t,i,j} \leq \phi_i^{\text{SMR}} \quad \forall \, i \in \mathcal{U}^{\text{SMR}}, \tag{4.11}$$

$$\sum_{j \in \mathcal{J}_O} x_{t,i,j} + \sum_{j \in \mathcal{J}_N} x_{t,i,j} \leq \phi_i^{\text{ASU}} \quad \forall \, i \in \mathcal{U}^{\text{ASU}}, \tag{4.12}$$

$$\sum_{j \in \mathcal{J}_H} x_{t,i,j} + x_{t,i}^{\text{stor-in}} \leq \mathbb{1}_{t,i}^{\text{SMR}} \phi_i^{\text{SMR}} \quad \forall \, i \in \mathcal{U}^{\text{SMR}}, \tag{4.13}$$

$$\sum_{j \in \mathcal{J}_H} x_{t,i,j} + x_{t,i}^{\text{stor-in}} \geq .7 \mathbb{1}_{t,i}^{\text{SMR}} \phi_i^{\text{SMR}} \quad \forall \, i \in \mathcal{U}^{\text{SMR}}, \tag{4.14}$$

$$\sum_{j \in \mathcal{J}_H} x_{t,j}^{\text{stor-out}} \leq R^{H,\max} - R_t^H \tag{4.15}$$

$$\sum_{j \in \mathcal{J}_H} x_{t,i}^{\text{stor-in}} \leq R_t^H - \theta_{s(t)}^L R^{H,\max} \tag{4.16}$$

$$\sum_{j \in \mathcal{J}_H} x_{t,i}^{\text{stor-in}} \geq \max(\theta_{s(t)}^L R^{H,\max} - R_t^H, 0) \tag{4.17}$$

$$x \geq 0 \quad . \tag{4.18}$$

The constraints are as follows:

Equation (4.7) - Demand constraint for nitrogen and oxygen, i.e. guarantees that every unit of demand for nitrogen and oxygen is satisfied.

Equation (4.8) - Demand constraint for steam, guarantees that each steam consumer can only be supplied by a single cogeneration unit.

Equation (4.9) - Demand constraint for hydrogen, guarantees that the demand of every hydrogen customer is satisfied.

Equation (4.10) - Natural gas constraint, i.e. forces the supplier to purchase more natural gas from the spot market if the amount of natural gas needed for production exceeds the amount of natural gas purchased through forward contracts. This constraint also requires the supplier to sell unused natural gas from the interruptible and uninterruptible pipeline on the spot market.

Equations (4.11) and (4.12) - Production capacity constraints enforce the production constraint of each production plant.

Equation (4.13) - SMR unit outages, i.e. restricts hydrogen production from unit $i \in \mathcal{U}_H$ in the event of an outage.

Equation (4.14) - SMR hydrogen production minimum, i.e. the supplier cannot turn off an SMR or lower production below a ratio (e.g., 70% of maximum capacity). All of the surplus hydrogen produce is place in storage.

Equation (4.16) - Hydrogen cavern withdrawal limits ensure the amount of hydrogen in storage is always greater than some buffer, $\theta^L R^{H,\max}$ where $\theta^L \in [0,1]$.

Equation (4.15) - Hydrogen cavern deposit constraints ensure the amount of hydrogen in storage never exceeds storage capacity.

We incorporate additional constraints in the linear program to enforce the storage policy for the hydrogen cavern. These additional constraints change depending on the state of the system. Equation (4.16) describes the hydrogen cavern withdrawal policy. These constraints ensure the amount of hydrogen in storage is always greater than some buffer, $\theta^L R^{H,\max}$. This constraint is only violated if there is an SMR outage and customer demands cannot be satisfied by the remaining SMR units. When this happens equation (4.16) becomes

$$\sum_{j \in \mathcal{J}_H} x_{t,i}^{\text{stor-in}} \leq R_t^H - R^{H,\max} \qquad \text{if} \qquad \sum_{i \in \mathcal{U}^{\text{SMR}}} \mathbb{1}_{t,i}^{\text{SMR}} \phi_i^H < \sum_{j \in \mathcal{J}_H} D_j \tag{4.19}$$

This allows the supplier to use reserve hydrogen and maintain reliability in the event a customer demand exceeds production capabilities.

Because the supplier must observe multiple constraints on the flow of products and resources, we model the daily policy as a parametric linear program. The parametric modifications are designed to account for problem uncertainty. This approach allows us to work with arbitrarily complex stochastic processes without requiring us to make limiting approximations. We emphasize that solving this linear program is not an optimal policy, but we can obtain relatively robust behavior by tuning the parameter θ^L in the base model given by equation (4.3).

The monthly optimization policy: $X_t^{\pi-\mathbf{month}}(S_t|\theta)$

We now consider the monthly decision to purchase natural gas forward contracts. At the end of bid week, $t \in \{t : d(t) = 1, m(t) = 1, ..., M\}$, the supplier must determine how much natural gas to be delivered in uniform daily increments over the succeeding month. We use a parametrically modified deterministic lookahead model of the month to make this decision. To distinguish the elements of our lookahead model from the base model we use a double index notation where all variables (states and decisions) in the lookahead model are indicated with tildes ($\sim$), and are indexed by t (the time at which the lookahead model is instantiated) and t' (the time period within the lookahead horizon). Before we proceed with the monthly optimization policy, we need to introduce some new notation.

Parameter: θ^{NG}

θ^{NG} = The fraction of forecasted demand for natural gas to be filled through one month forward contracts. $\theta^{NG} \geq 0$, Monthly state variable: $S_t = (c_t^{\text{int}}, c_t^{\text{uni}}, P_t^{\text{NG-spot}}, (\tilde{f}_{t,j}^D)_{j\in\mathcal{J}_g})$

c_t^{int} = The marginal cost of purchasing natural gas from the interruptible pipeline at a indexed price to be delivered in uniform daily increments over the succeeding month,

c_t^{uni} = The marginal cost of purchasing natural gas from the uninterruptible pipeline at a indexed price to be delivered in uniform daily increments over the succeeding month,

$\tilde{f}_{tt',j}^D$ = The forecasted demand of customer $j \in \mathcal{J}^g$ at time t' given information available at time t. These forecasts are based on the demand of customers from the previous year,

$|\mathcal{M}_m|$ = The number of days in the month m,

Monthly decision variables $x_t = ((\tilde{x}_{tt',ij})_{j\in\mathcal{J}^g}, x_t^{NG-\text{int}}, x_t^{NG-\text{uni}}, \tilde{x}_{tt'}^{NG})$

$\tilde{x}_{tt',ij}$ = The predicted amount of industrial gas g produced by plant i to be sold to customer $j \in \mathcal{J}^g$ during day t' of the lookahead model given information available at time t. For this lookahead model we only consider $g \in \{H, \text{steam}\}$ since they are the only products that require natural gas,

$x_t^{NG-\text{int}}$ = The fixed daily amount of natural gas ordered from the interruptible pipeline to be delivered in uniform daily increments over the succeeding month,

$x_t^{NG-\text{uni}}$ = The fixed daily amount of natural gas ordered from the uninterruptible pipeline to be delivered in uniform daily increments over the succeeding month. Daily deliveries from the uninterruptible pipeline are limited by the contract agreements such that

$$x_t^{NG-\text{uni}} < \Gamma^{\text{uni}}.$$

$\tilde{x}^{NG}_{tt'}$ = The predicted total amount of natural gas needed to satisfy the expected customer demand for day t' of month $m(t)$.

A lookahead policy (in the form of a linear program) uses point forecasts of the customer demands for the following month to determine how much natural gas to order. The linear program solved at time t, where $d(t) = 0$ and $m(t) = m$, is as follows:

$$X^\pi_t(S_t) = \underset{x}{\text{argmin}} \;\; |\mathcal{M}_m| c^{\text{int}}_t x^{NG-\text{int}}_t + |\mathcal{M}_m| c^{\text{uni}}_t x^{NG-\text{uni}}_t \tag{4.20}$$

This is solved subject to the following constraints (where $d(t) = 0$, and $t' \in \{t : d(t) \neq 0, m(t) = m\}$).

$$\tilde{x}_{tt',j} = \tilde{f}^D_{tt',j}, \;\; j \in \mathcal{J}^{\text{steam}} \quad \& \quad t' \in \mathcal{M}_m \tag{4.21}$$

$$\sum_{i \in \mathcal{U}^{\text{SMR}}} \tilde{x}_{tt',ij} = \tilde{f}^D_{tt',j}, \;\; j \in \mathcal{J}^H \quad \& \quad t' \in \mathcal{M}_m, \tag{4.22}$$

$$\sum_{j \in \mathcal{J}_g} \tilde{x}_{tt',ij} \leq \phi^g_i \;\; \forall \; i \in (\mathcal{U}^g)_{g \in \mathcal{G}} \quad \& \quad t' \in \mathcal{M}_m, \tag{4.23}$$

$$\tilde{x}^{NG}_{tt'} = \sum_{j \in \mathcal{J}_H} \sum_{i \in \mathcal{U}^{\text{SMR}}} \gamma^H_i \tilde{x}_{tt',ij} + \sum_{j \in \mathcal{J}_{\text{steam}}} \gamma^{\text{steam}}_i \tilde{x}_{tt',jj}, \; t' \in \mathcal{M}_m \tag{4.24}$$

$$x^{NG-\text{int}}_t + x^{NG-\text{uni}}_t = \frac{\theta^{NG}}{|\mathcal{M}_m|} \sum_{t' \in \mathcal{M}_m} \tilde{x}^{NG}_{tt'}, \tag{4.25}$$

$$x^{NG-\text{uni}}_t \leq \Gamma^{\text{uni}}, \tag{4.26}$$

$$x \geq 0. \tag{4.27}$$

The constraints are as follows:

Equation (4.21) - The policy must satisfy all of the forecasted demand for steam in the lookahead model.

Equation (4.22) - The policy must satisfy all of the forecasted demand for hydrogen in the lookahead model.

Equation (4.23) - Production capacity constraint enforces the production constraint of each gas producing unit.

Equations (4.24) and (4.25) ensures the total amount of natural gas ordered is equal to the average daily amount of natural gas needed over month $\mathcal{M}_m$.

Equation (4.26) - Uninterruptible constraint sets an upper limit on the amount of natural gas delivered by the uninterruptible pipeline.

The monthly policy is a linear program which we solve using Gurobi.

Stochastic Gradient Algorithm:

We next consider the problem of determining policy parameters for the daily and monthly operating policies described in subsections 4.4 and 4.4, respectively. The policy described by equation (4.5) belongs to the general class of parameterized policies, $X_t^\pi : \mathcal{S}_t \times \Theta \longrightarrow \mathcal{X}_t$. The optimal parameterization, $\theta^* \in \Theta$, of policy X_t^π can be found by solving the stochastic approximation problem

$$\theta^* = \operatorname*{argmax}_{\theta \in \Theta} \mathbb{E}\left[F(\theta, \omega) \,\middle|\, S_0\right], \tag{4.28}$$

where

$$F(\theta, \omega) = \sum_{t=0}^{T} C_t(S_t, X_t^\pi(S_t|\theta))$$

such that $S_{t+1} = S^M(S_t, X_t^\pi(S_t|\theta), W_{t+1}(\omega))$. $\mathbb{E}\left[F(\theta, \omega) \,\middle|\, S_0\right]$ is an unknown, non-convex, non-smooth function, that we cannot evaluate directly. Instead we use numerical gradients generated using our base model, and the finite difference iterative stochastic optimization algorithm, described in Algorithm 2, to solve Equation (4.28)

(Kiefer and Wolfowitz, 1952). We recognize the Kiefer Wolfowitz algorithm requires for each gradient computation, at least $p + 1$ simulations where $\theta \in \mathbb{R}^p$. If p is large this can require substantial computation. However, in our case, $p = 2$ and computationally manageable for our purposes.

Algorithm 2 Finite-Difference algorithm

Initialize $\theta^0 \in \mathbb{R}^p$ and N:

for n = 1, 2, 3, ..., N **do**

$$\theta^n = \theta^{n-1} + \alpha_{n-1} \nabla_\theta F(\theta^n, \omega^n)$$

where

$$\nabla_\theta F(\theta^n, \omega^n) \approx \begin{bmatrix} \frac{F(\theta^{n-1} + c_n \xi_1, \omega_{n,1}) - F(\theta^{n-1}, \omega_{n,0})}{c_n} \\ \vdots \\ \frac{F(\theta^{n-1} + c_n \xi_p, \omega_{n,p}) - F(\theta^{n-1}, \omega_{n,0})}{c_n} \end{bmatrix},$$

ξ_i denotes a vector with a 1 in the ith place and 0's elsewhere.

Kiefer and Wolfowitz (1952) requires the sequence $(c_n)_{n=1}^{\infty}$ and $(\alpha_n)_{n=1}^{\infty}$ are infinite sequences of positive numbers such that

$$c_n \to 0 \text{ as } n \to \infty,$$

$$\sum_{n=0}^{\infty} \alpha_n = \infty,$$

$$\sum_{n=0}^{\infty} \alpha_n c_n < \infty$$

$$\sum_{n=0}^{\infty} \alpha_n^2 c_n^{-2} < \infty.$$

For our numerical work we use the Root Mean Square Propagation algorithm, RMSProp, to generate $(\alpha_n)_{n=1}^{\infty}$ and define $c_n = n^{-1/3}$ (see Tieleman and Hinton (2012)).

RMSProp eliminates the need to tune the learning rate for each dimension of θ by dividing the learning rate by a running average of the magnitude of the recent gradients. Given the RMSProp algorithm

$$\alpha_n = \frac{\eta}{\sqrt{v_n} + \epsilon}$$

where

$$v_n = \gamma v_{n-1} + (1-\gamma)(\nabla_\theta F(\theta^n, \omega^n))^2,$$

and $\gamma \in [0,1]$ is the forgetting factor. For our simulations we set $N = 200$, $\eta = 0.1$, $\epsilon = 10^{-8}$, and $\gamma = .9$.

To ensure our policy parameters values are near-optimal, we perform a grid search over the parameters for both the daily and monthly sub-policies. We first discretize the parameter space, run a thousand simulations, and compute an empirical expected cumulative profit for the industrial gas network described in tables 4.1 and 4.2. The parameters of the customer contracts, described in table 4.1, are generated randomly. However, we constrain the parameter P_0 such that it is always greater than the product of the conversion efficiency for the associated production plant and the price of natural gas when the contract was created, NG_0. This constraint guarantees that the marginal revenue for that customer is non-negative when the current price of natural gas is less than or equal to the spot price of natural gas when the contract was formed, NG_0.

ID	Gas Type	P_0	NG_0	α_j	β_j	Penalty (per MSCF)
0	H2	2.950	4.309	0.340	month	10.00
1	H2	3.200	4.427	0.280	spot	10.00
2	H2	3.450	4.147	0.220	spot	10.00
3	H2	3.700	4.020	0.360	spot	10.00

4	H2	2.850	3.769	0.300	month	10.00
5	H2	3.100	3.526	0.240	spot	10.00
0	O2	1.140	3.686	0.280	month	10.00
1	O2	1.380	4.463	0.360	spot	10.00
2	O2	1.620	4.456	0.240	spot	10.00
3	O2	0.960	4.433	0.320	spot	10.00
4	O2	1.200	4.061	0.200	month	10.00
5	O2	1.440	3.883	0.280	spot	10.00
6	O2	1.680	3.807	0.360	spot	10.00
7	O2	1.020	4.360	0.240	spot	10.00
8	O2	1.260	4.318	0.320	month	10.00
9	O2	1.500	3.582	0.200	spot	10.00
10	O2	1.740	4.220	0.280	spot	10.00
11	O2	1.080	3.662	0.360	spot	10.00
0	N2	1.390	3.559	0.220	month	10.00
1	N2	1.780	4.198	0.240	spot	10.00
2	N2	1.270	3.591	0.260	spot	10.00
3	N2	1.660	3.555	0.280	spot	10.00
4	N2	1.150	3.501	0.300	month	10.00
5	N2	1.540	3.666	0.320	spot	10.00
6	N2	1.030	4.147	0.340	spot	10.00
7	N2	1.420	3.518	0.360	spot	10.00
8	N2	1.810	3.680	0.380	month	10.00
9	N2	1.300	4.302	0.400	spot	10.00
10	N2	1.690	3.905	0.220	spot	10.00
11	N2	1.180	3.719	0.240	spot	10.00
12	N2	1.570	3.799	0.260	month	10.00

13	N2	1.060	4.450	0.280	spot	10.00
14	N2	1.450	3.575	0.300	spot	10.00
15	N2	1.840	4.216	0.320	spot	10.00
16	N2	1.330	4.074	0.340	month	10.00
17	N2	1.720	3.655	0.360	spot	10.00
0	steam	100.400	2.836	0.360	month	10.00

Table 4.1: Customer contract details

	Capacity	**Qty.**	**Spec Power**	**Prod. Min.**
Steam Methane Reformer	35,000 MSCF	1	.45 MMBTU/MSCF	70%
Air Separation Unit	100,000 MSCF	1	0.018 MWh/MSCF	-
Cogeneration Plant	100,000 MMBTU	1	0.011 MMBTU/ MMBTU	-

Table 4.2: Production plant details

The point with the maximum expected cumulative profit from the grid search will be a near optimal parameterization. We recognize that this method may not result in the exact optimal solution since the optimization is run over a discretized grid. Figure 4.4 is the two dimensional heat maps of expected cumulative profit as functions of daily and monthly policy parameters, respectively. The expected cumulative profit ranges from black (minimum) to white(maximum).

To demonstrate the validity of Algorithm 2, we run the algorithm from different initial parameterization. If the solutions produced by the algorithm lie in the white area of the heat map in figure 4.4, regardless of the initial point then we say the algorithm performs reliability. We represent each run of the algorithm using the arrows in figure 4.4. The base and head of each arrow represents the initial point and solution of the algorithm, respectively. Our experiments imply the algorithm consistently finds a near-optimal solution. In table 4.3 we compare the solutions of our algorithm with best solutions from the grid search.

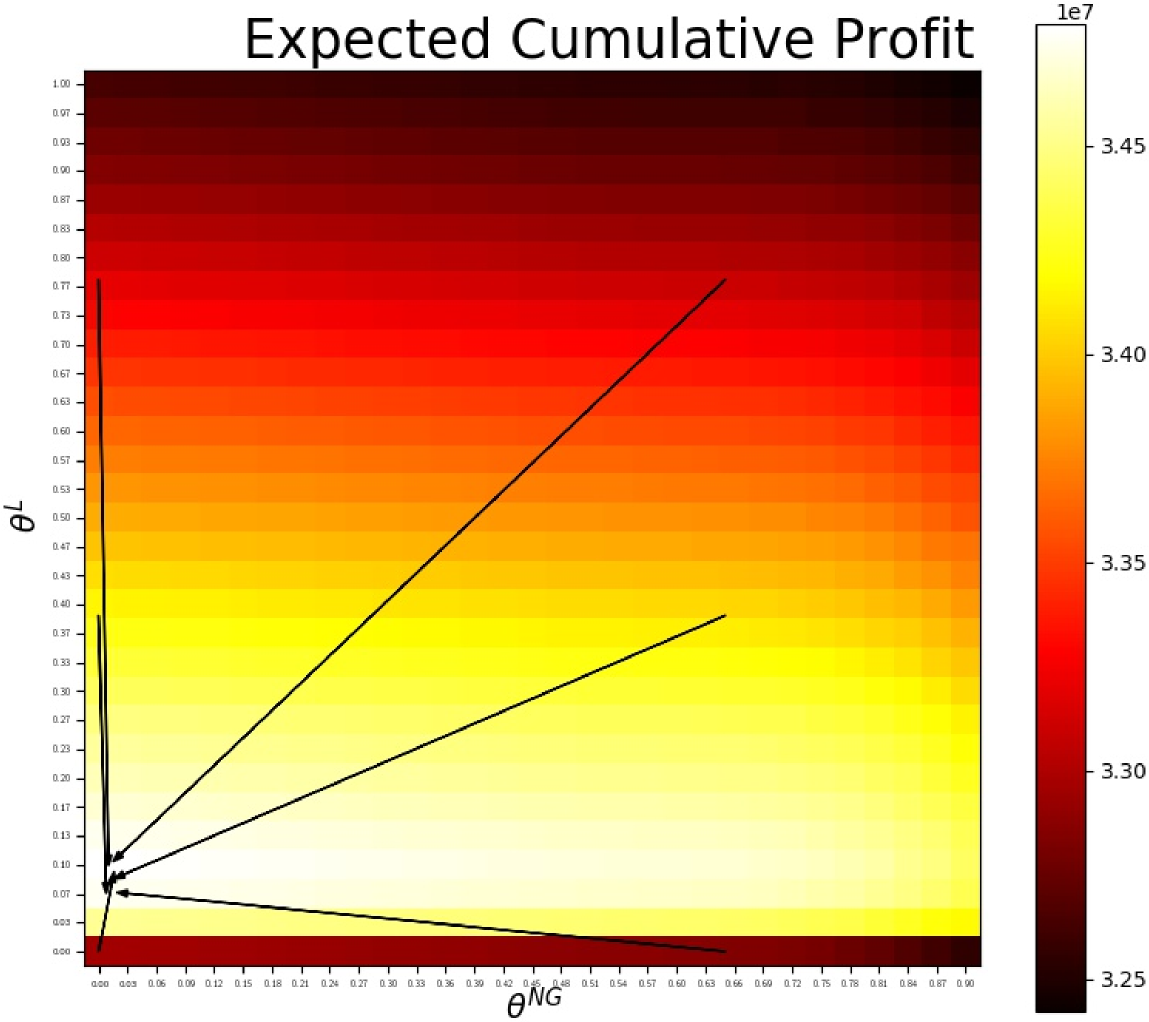


Figure 4.4: This is a heat map of the expected cumulative profit as a function of policy parameters. The expected cumulative profit ranges from black (minimum) to white(maximum). The black arrows represent different runs of Algorithm 2 for different initial points. The base and head of each arrow represents the initial point and solution of the algorithm, respectively.

The optimal poliy parameters (θ)

In this section we more closely examine the optimal policy parameter, $\theta = (\theta^L, \theta^{NG})$, given different simulation assumptions. We particularly focus on the relationship between the optimal hydrogen reserve level, θ^L, and the reliability of hydrogen producing steam methane reformers (SMR). SMRs regularly experience maintenance issues that render them unable to produce hydrogen. These outages occur randomly and typ-

Optimal Point (Algorithm 2)			Best six points (Grid Search)	
Initial $\theta = (\theta^L, \theta^{NG})$	Optimal $\theta = (\theta^L, \theta^{NG})$	Cumulative Profit (\$ $\times 10^6$)	Optimal $\theta = (\theta^L, \theta^{NG})$	Cumulative Profit (\$ $\times 10^6$)
(0.000, 0.000)	(0.079, 0.038)	36.243	(0.090, 0.000)	36.290
(0.000, 0.700)	(0.079, 0.083)	36.209	(0.090, 0.033)	36.285
(0.350, 0.000)	(0.068, 0.038)	36.223	(0.090, 0.067)	36.281
(0.350, 0.700)	(0.068, 0.078)	36.190	(0.090, 0.100)	36.276
(0.700, 0.000)	(0.093, 0.053)	36.214	(0.090, 0.133)	36.271
(0.700, 0.700)	(0.088, 0.097)	36.155	(0.090, 0.167)	36.266

Table 4.3: Search results of Algorithm 2 starting from various initial points (left) and best results from rudimentary gird search (right). These results were generated using the industrial gas network described in tables 4.1 and 4.2.

ically last for multiple days. During an outage, if there is not enough hydrogen in storage the demands of hydrogen customers will go unsatisfied and the industrial gas supplier will have to pay a penalty fee for each MSCF of hydrogen unsatisfied.

For the following set of experiments we find the set of optimal policy parameters using the stochastic gradient algorithm described by Algorithm 2 for different SMR outage durations, SMR outage frequencies, and customer penalties. We calculate these values using the industrial gas network configuration described in tables 4.1 and 4.2. Table 4.4 displays the results of these three different sets of experiments. The first section of table 4.4 shows the optimal policy parameters, θ, for varying SMR outage durations which range from zero days to a month. The second set of experiments, considers varying SMR outage frequencies from once to ten times a year. The final set of experiments considers different degrees of penalties for unsatisfied hydrogen customer demands.

As expected the optimal reserve parameter, θ^L, increases with both the duration and frequency of SMR outages. However, when SMR outages are completely eliminated θ^L does not go to zero. This occurs because the demands of hydrogen customers occasionally exceed production capacity and require hydrogen from storage. One may also expect that as the penalty for not satisfying customer demands increases so would, θ^L. However, table 4.4 (third part) shows this may not always

SMR Outage Duration (days)	**SMR Outage** Frequency (per annum)	**Penalty** (\$/MSCF)	**Opt.** $\theta = (\theta^L, \theta^{NG})$
0	1	10	(0.071, 0.608)
6	1	10	(0.148, 0.608)
12	1	10	(0.259, 0.608)
18	1	10	(0.378, 0.608)
24	1	10	(0.516, 0.140)
30	1	10	(0.647, 0.140)
4	1	10	(0.100, 0.608)
4	2	10	(0.131, 0.038)
4	4	10	(0.208, 0.001)
4	6	10	(0.267, 0.078)
4	8	10	(0.288, 0.000)
4	10	10	(0.295, 0.379)
4	1	0	(0.067, 0.108)
4	1	2	(0.100, 0.108)
4	1	4	(0.100, 0.108)
4	1	6	(0.100, 0.108)
4	1	8	(0.111, 0.057)
4	1	10	(0.111, 0.057)

Table 4.4: Optimal θ for different Steam Methane Reformers outage frequencies and durations and penalties for unsatisfied customer demand.

be true. As expected, as the penalty increases so does the optimal θ^L, however once the penalty exceeds \$10 per unit of unsatisfied customer demand the optimal value for the parameter θ^L remains constant. This can be attributed to the fact that the probability of running out of storage when the SMR only experiences one three-day outage a year is less than 5% when $\theta^L \geq .12$. Hence, increasing θ^L would only lower expected cumulative profit. This can be seen in table 4.5 which shows the relationship between θ^L and the probability of depleting the hydrogen storage.

θ^L	Probability of depleting storage	Percentage of sales missed
0.000	100.00%	1.44%
0.043	30.00%	0.12%
0.086	6.67%	0.03%
0.129	3.33%	0.01%

Figure 4.5: This table shows the relationship between θ^L and the probability of depleting the hydrogen storage.

4.5 Policy Studies

In this section we demonstrate the use of our model in the context of multiple policy questions. These include:

- Is it possible to reduce the dependence of the cash flows of the supplier and the performance of the market?
- What is the sensitivity of the parameters of the operating policy to natural gas price volatility? (e.g. for scenarios where natural gas volatility is twice its current volatility.)

All of our experiments assume the industrial gas supplier has access to a single steam-methane reformer, an air separation unit, one cogeneration plant, and a salt cavern to store hydrogen. Tables 4.1 and 4.2 contain the technical details of each production unit and the contract parameters, introduced in equation (4.2), for our simulated customers. Our problem horizon is one year, $T = 366$ days, where the first month of the problem horizon is January.

Experiment 1: Correlation of profit and fuel

The shareholders of an industrial gas supplier are interested in minimizing their exposure to external market forces. Specifically, they want to minimize the absolute correlation between the daily profit of the supplier and the overall performance of the market. For our first set of experiments, we find the optimal monthly operating policy parameter, θ^{NG}, which maximizes the expected cumulative returns of the industrial gas producer given a certain level of exposure to natural gas spot prices. We use the absolute value of the correlation of the daily profit of the industrial gas supplier and the daily natural gas spot price to measure the industrial gas supplier's exposure to natural gas prices. The correlation between natural gas spot prices and the daily profit of the supplier is negative for this particular network configuration. Our objective is to minimize the dependence of daily profit on fuel prices. Specifically, we want to minimize the absolute value of the correlation.

To determine the optimal monthly operating policy parameter we discretize the parameter space, $\theta^{NG} \in [0, 1]$, and perform a thousand simulations to approximate the expected cumulative profit, standard deviation of cumulative profits, and the correlation of daily profit and natural gas spot prices. Figure 4.6 illustrate the relationship between the monthly operating policy parameter, θ^{NG}, and the empirical expected cumulative profits of the industrial gas supplier.

Figure 4.6 shows that as the industrial supplier orders more natural gas through forward contracts they increase the correlation of their daily profits to natural gas spot prices. Since, the correlation of their daily profits to natural gas prices is negatively correlated this decreases the exposure of the supplier to energy prices. There is a cost for the reduction in exposure since the supplier pays a premium to purchase the contracts. An interesting observation from figure 4.6 is the steep decline in expected cumulative profit for $\theta^{NG} \geq .8$. This decline is the result of the transaction costs of selling unused natural gas back on the spot market.

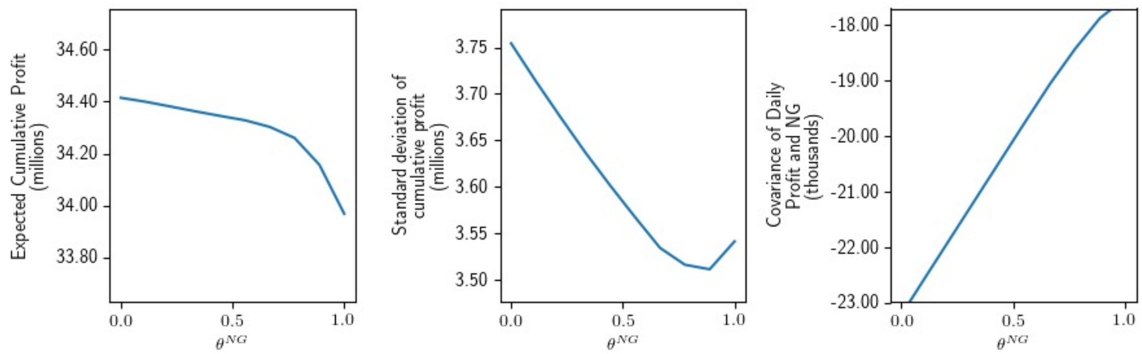


Figure 4.6: These figures illustrate the relationship between natural gas forward contracts and expected profits. The first figure shows the correlation between daily profits and the spot price of natural gas prices as a function of the amount of forward contracts purchased. The second figure shows the standard deviation of cumulative profits as a function of the amount of forward contracts purchased.

If the objective of the supplier is to just maximize expected cumulative profit, the optimal monthly policy would avoid purchasing natural gas forward contracts. This is why the monthly sub-policy parameters, θ^{NG}, found using Algorithm 2, are very small. However, if the objective is to maximize expected cumulative profits and limit exposure to natural gas prices, the supplier must determine how much profit they are willing to forfeit to reduce their exposure to fuel prices. Figure 4.7 shows the tradeoff between exposure to natural gas prices and expected cumulative profits.

The results displayed in figure 4.7 show the industrial gas supplier can reduce the absolute correlation of their daily profits to natural gas prices by nearly 22% and only sacrifice about 1% of their expected cumulative profit. If the objective of the industrial gas supplier is to simply maximize expected cumulative profits and solve equation (4.3), regardless of their exposure to natural gas prices, then Algorithm 2 is an effective and computationally tractable method. However, if the supplier is interested in maximizing expected profit given some threshold or maximizing another risk measure they must rely on gradient-free methods (see Powell and Ryzhov (2012)).

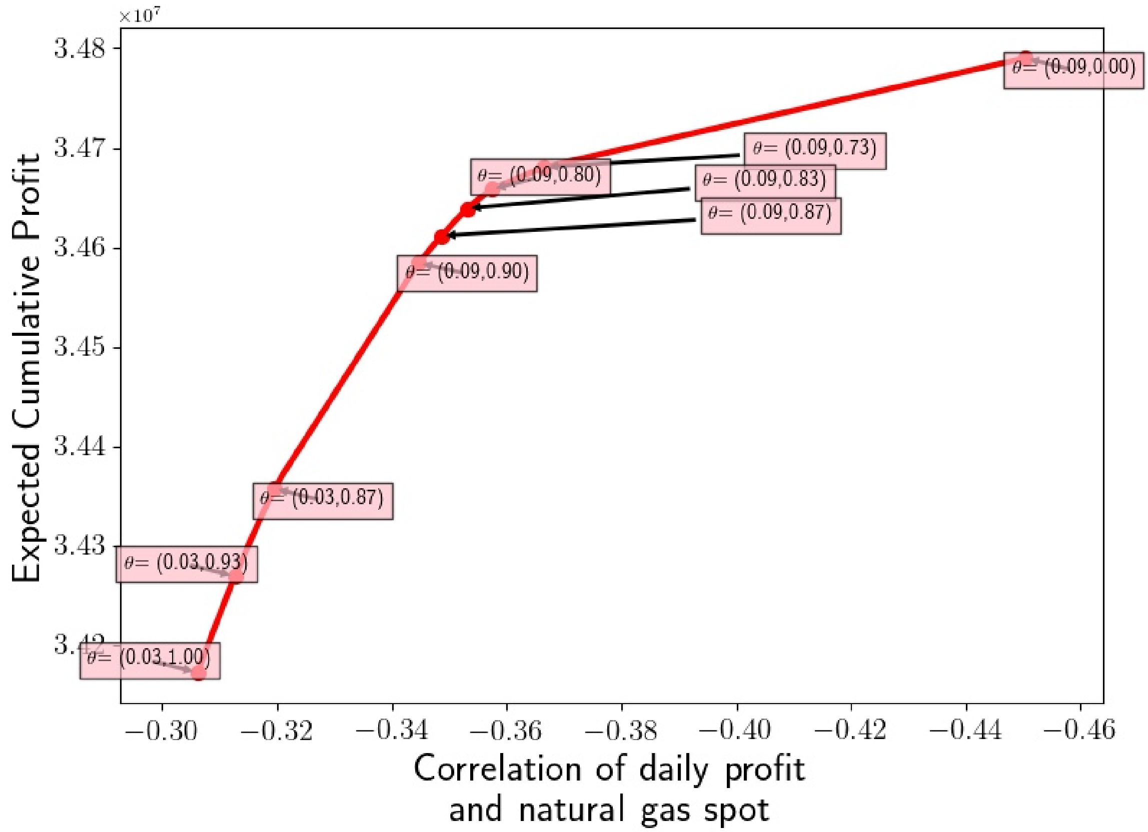


Figure 4.7: This figure illustrates the tradeoff between exposure to natural gas spot prices and expected cumulative costs. The x-axis represents the correlation between daily profits and natural gas spot prices while the y-axis represents the expected cumulative costs.

Experiment 2: Natural Gas Volatility

For our second experiment, we investigate the effect of natural gas price volatility, η, on the policy parameters, $\theta = (\theta^L, \theta^{NG})$. The parameter η is the standard deviation of the daily log-returns of natural gas spot prices. Figure 4.8 illustrates the relationship between the parameter η and the standard deviation of simulated natural gas spot prices.

In the following experiments, we use the same network configuration as the previous section, but vary the volatility of natural gas spot price model, η. Figures 4.9 and 4.10 compare the operating policy parameter combinations for different levels of natural gas model volatility, η. The blue dots represent the tested parameter com-

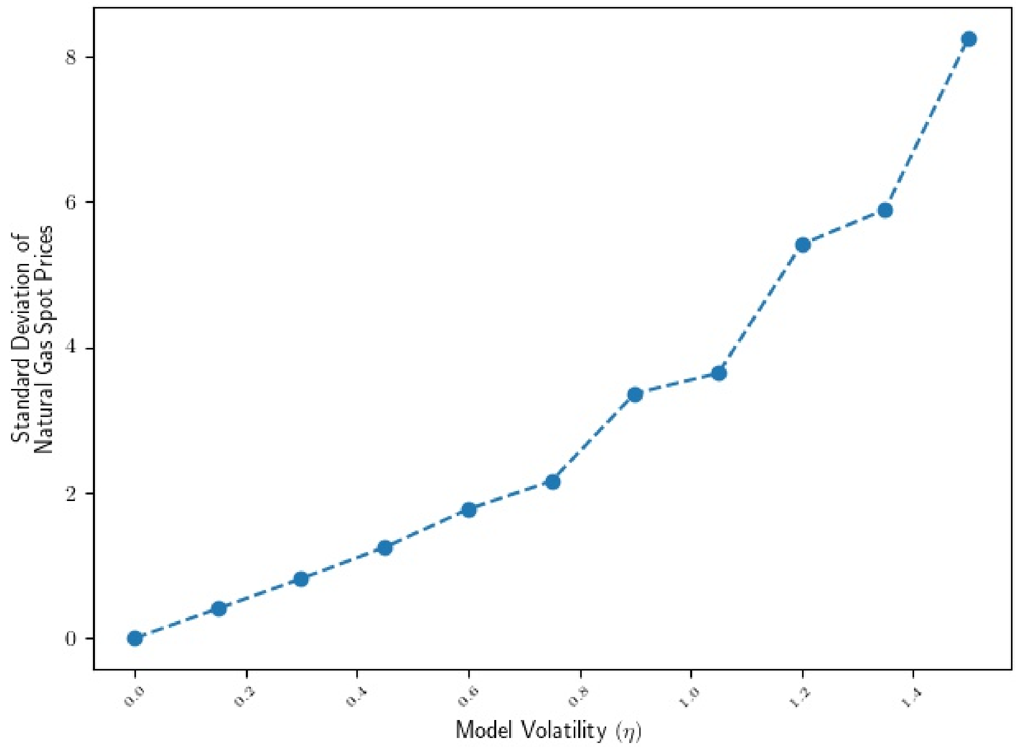
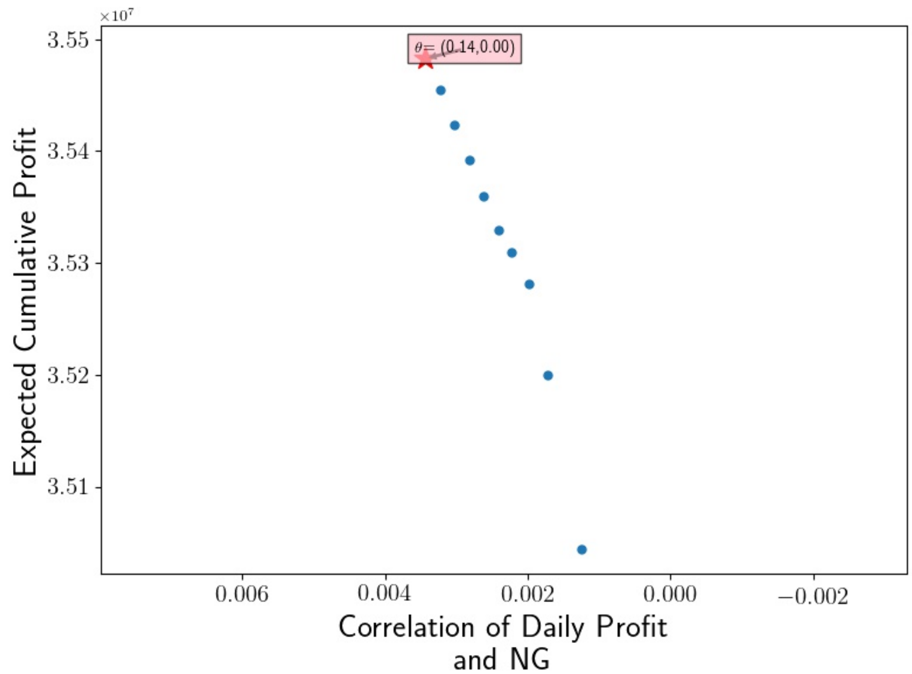
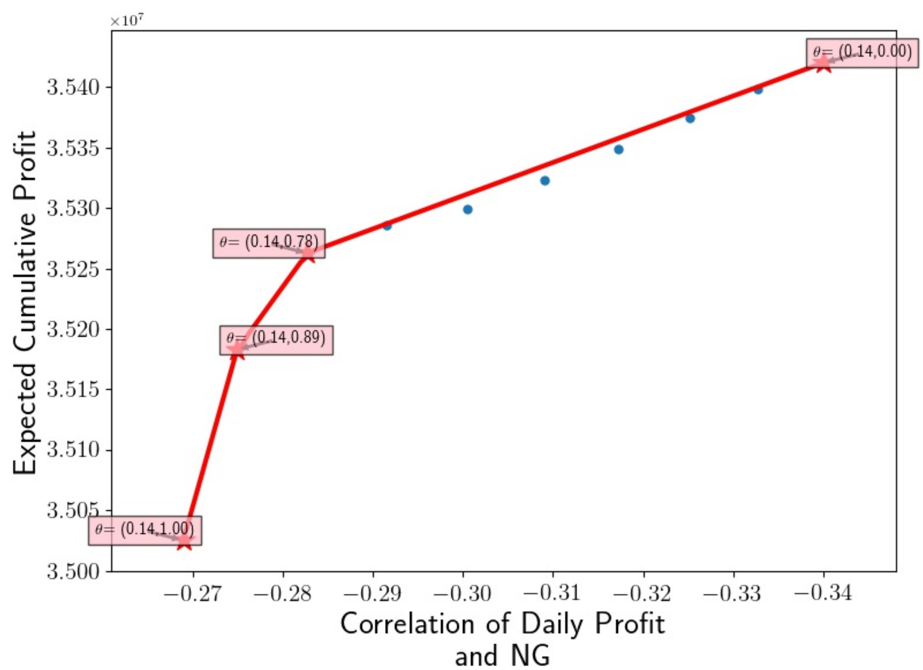


Figure 4.8: The relationship between the parameter η and the standard deviation of natural gas spot prices.

binations. The red dots are the optimal set of parameters, commonly referred to as the efficiency frontier, that offer the highest expected profit for a defined level of risk. The red line is the lower bound of the efficiency frontier given the tested policy parameters. If a parameter combination lies below the efficient frontier it is considered sub-optimal, because there exists another combination that produces a greater expected profit for a lesser or equal level of risk.

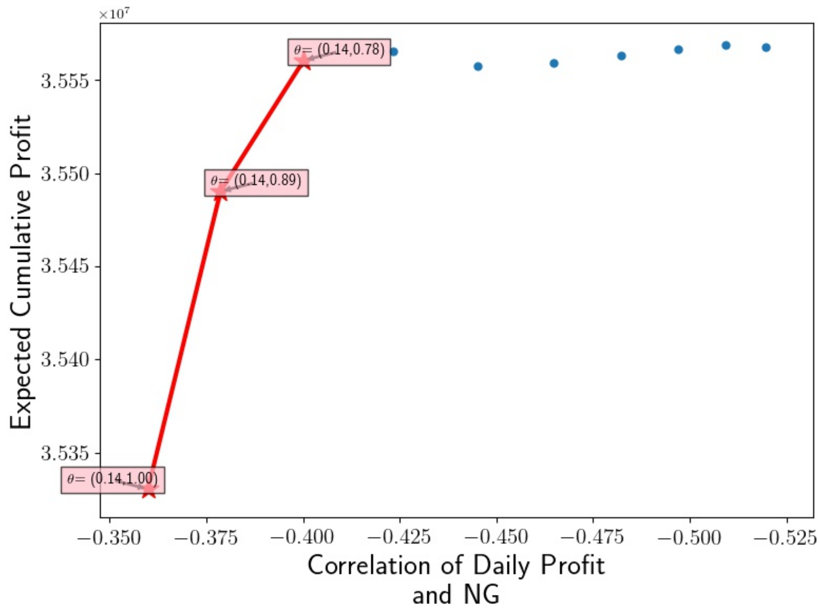


(a) $\eta = 0$

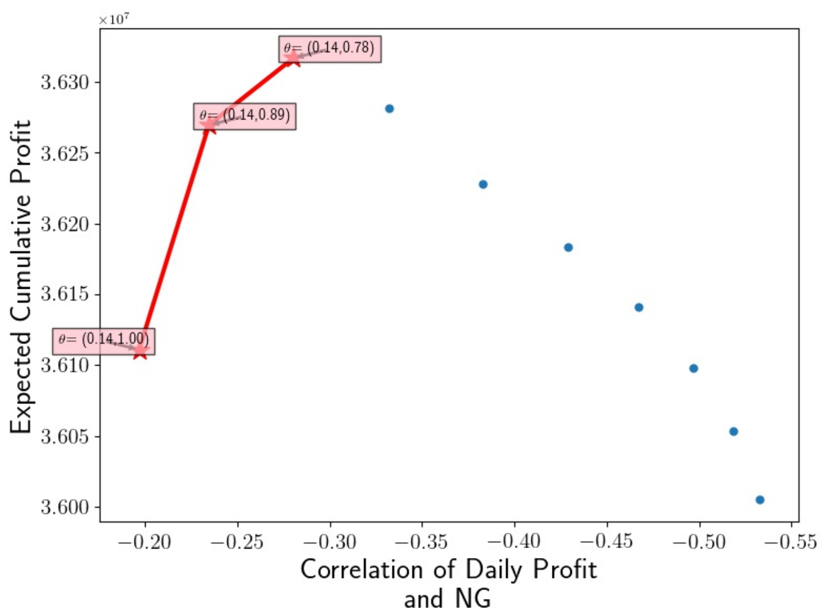


(b) $\eta = .5$

Figure 4.9: These figures display the tradeoff between exposure to natural gas spot prices and expected cumulative costs where risk is measured as the absolute correlation of daily profit and spot price of natural gas. The individual figures represent this tradeoff when $\eta = 0$ and $\eta = .5$.



(a) $\eta = 1$



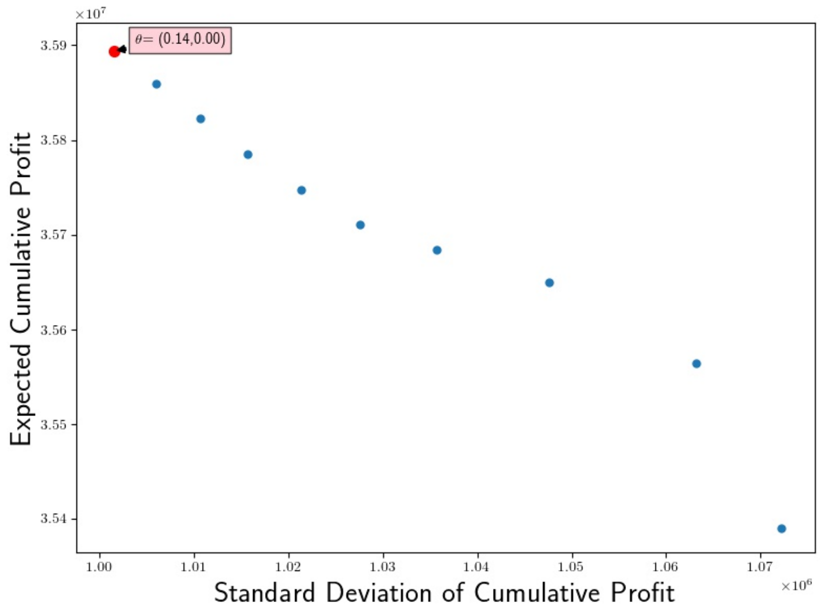
(b) $\eta = 1.5$

Figure 4.10: These figures display the tradeoff between exposure to natural gas spot prices and expected cumulative costs where risk is measured as the absolute correlation of daily profit and spot price of natural gas. The individual figures represent this tradeoff when $\eta = 1$ and $\eta = 1.5$.

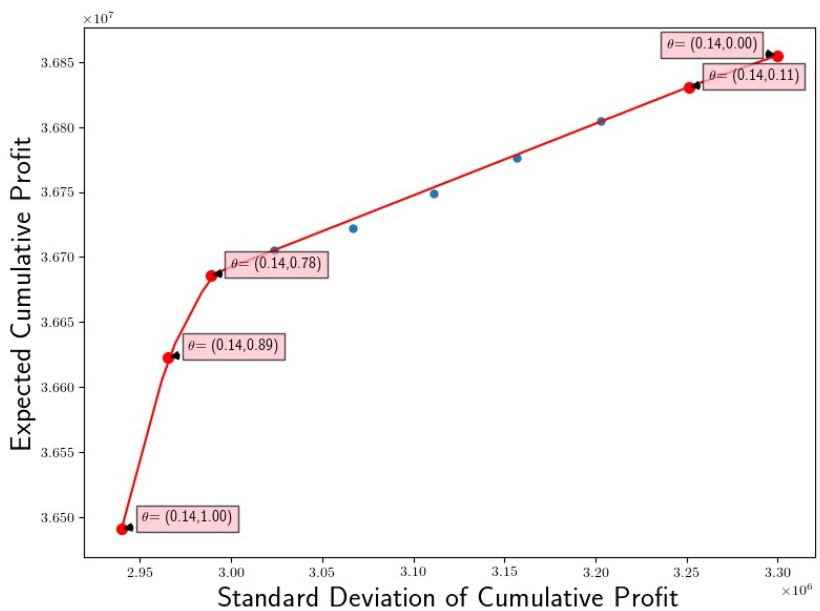
In figures 4.9 and 4.10, we can see the substantial benefit of using forward contracts. For all of the cases where $\eta > 0$, there is an interval of favorable θ^{NG} values between $.8 \leq \theta^{NG} \leq .9$. Within this range there is a substantial reduction in the absolute correlation between natural gas spot prices and daily profit, but very little decrease in expected cumulative profit. As expected, when $\eta = 0$ there is no benefit for purchasing forward contracts. In this case the premium paid for purchasing natural gas through forward contracts just lowers the expected cumulative profit. Hence, for this case the optimal $\theta^{NG} = 0$.

An interesting observation from figures 4.9 and 4.10 is how the shape of the efficiency frontier changes with the volatility of natural gas prices. As η increases the expected cumulative profit of policies where $\theta^{NG} < .78$ decrease. This demonstrates that for highly volatile periods purchasing natural gas through forward contracts both increases the expected cumulative profit and decreases the exposure of the supplier to natural gas spot prices. This is expected, since forward contracts provide a hedge against sharp increases in natural gas prices. Although the general shape of the efficiency frontier changes as the volatility of natural gas prices increases, the interval of favorable θ^{NG} values is invariant. This suggests that near optimal policy parameters, θ, for the operating policy defined in equation (4.5), should be near optimal for all cases where $\eta > 0$.

Figures 4.11 and 4.12 also compare the operating policy parameter combinations for different levels of natural gas model volatility, η. However, unlike figures 4.9 and 4.10, this figure uses the variance of the cumulative profits of the supplier as a risk measure.

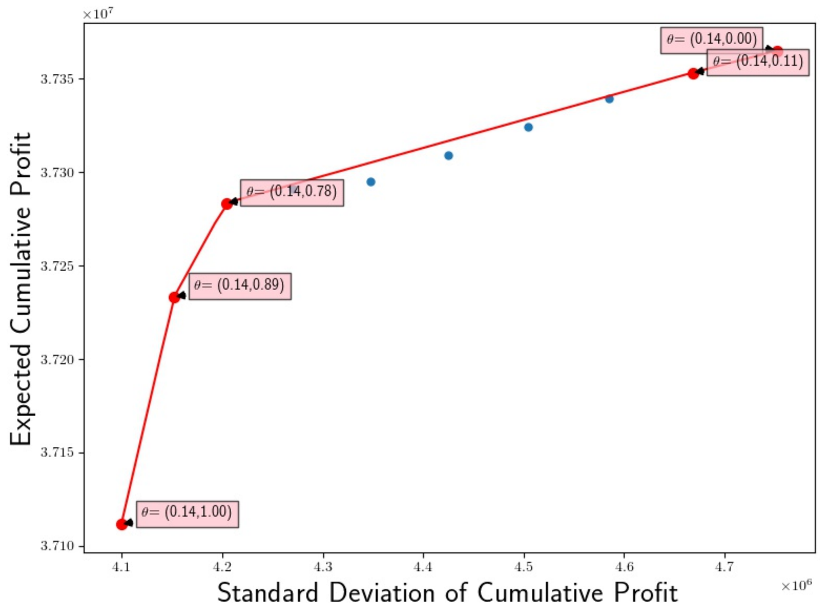


(a) $\eta = 0$

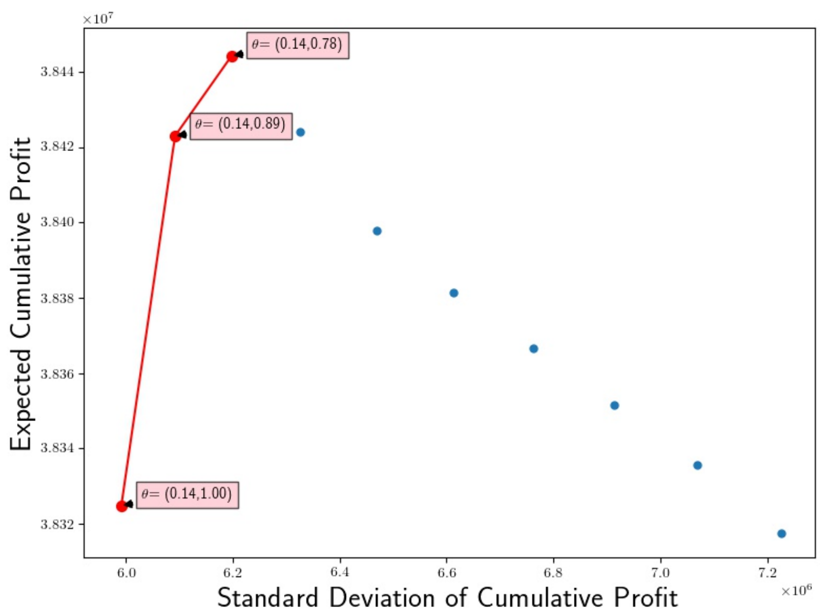


(b) $\eta = .5$

Figure 4.11: These figures display the trade off between exposure to natural gas spot prices and expected cumulative costs where risk is measured as the standard deviation of cumulative profit. The individual figures represent this tradeoff when $\eta = 0$ and $\eta = .5$.



(a) $\eta = 1$



(b) $\eta = 1.5$

Figure 4.12: These figures display the trade off between exposure to natural gas spot prices and expected cumulative costs where risk is measured as the standard deviation of cumulative profit. The individual figures represent this tradeoff when $\eta = 1$ and $\eta = 1.5$.

In figures 4.11 and 4.12, we observe very similar behavior to figures 4.9 and 4.10. When $\eta > 0$, $.8 < \theta^{NG} < .9$ is the interval of favorable θ^{NG} values. An interesting observation is that this interval is very close to the favorable interval when the correlation of natural gas prices and daily profit is the risk measure. This is expected since forward contracts reduce the variance of the daily profit of the supplier. Consequently, this reduces both the variance of cumulative profits and the covariance of the daily profit of the supplier to natural gas spot prices.

4.6 Conclusion

In this chapter, we introduce a new dynamic model that is capable of analyzing and identifying optimal natural gas procurement strategies, hydrogen storage, and customer contract negotiations. This model can be used for a wide array of experiments and analysis. We model the operations of an industrial gas producer with access to a hydrogen storage cavern, a natural gas hub, an electricity grid, and a diverse set of customers.

Our mathematical model carefully captures the distinction between a stochastic base model and an operating policy using a parametric cost function approximation. This model pays particular attention to the planning of lagged decisions in the management of monthly and daily natural gas deliveries and contracts.

We also design several experiments to better understand how to effectively manage a portfolio of short term natural gas contracts and hydrogen storage cavern. First, we illustrate the relationship between the optimal level of hydrogen reserve and the reliability of production plants. Second, we evaluate multiple strategies for minimizing the exposure of the industrial gas supplier to volatile energy prices. In addition to this, we investigate the relationship between optimal natural gas procurement strategies and market volatility. The results of our experiments show that the optimal hydrogen reserve level increases with both the duration and frequency of steam-methane reformer outages. However, the optimal reserve level is invariant once the penalty for not satisfying customer demands exceeds \$10 per unit. Our analysis of natural gas procurement strategies shows we can reduce the absolute correlation of our daily profits to natural gas prices by nearly 22% and only sacrifice about 1% of our expected cumulative profit by satisfying 78% of our forecasted natural gas demand through forward contracts. We also show the optimal natural gas procurement strategy is invariant when the volatility of natural gas spot prices is positive.

Future extensions of this work can include the addition of other risk-measure and gradient-free optimization methods (see Powell and Ryzhov (2012)).

Chapter 5

Conclusion & Future Research

This thesis introduces and formalizes a new class of decision making polices known as parametric cost function approximations (CFA) which use deterministic optimization problems that have been parametrically modified to account for uncertainty. In Chapter 2, we formally introduce the concept of parametric cost function approximations and the CFA Gradient Algorithm. We show how these parametrically modified linear programs and the CFA Gradient Algorithm allows us to exploit the structural properties of stochastic sequential problems while capturing the complex dynamics of the full base model. In this chapter we demonstrate this class of policies in the context of a complex, time-dependent energy storage problem with forecasts. In Chapter 3, we demonstrate the parametric cost function approximation in the context of the difficult problem of making lagged commitments while managing a portfolio of energy resources. This setting requires the decision maker to consider the tradeoff between forecast reliability and cost. We provide a proper base model of a stochastic, lagged resource allocation problem in the context of energy portfolio management. We also demonstrate empirically that our method produces high quality solutions relative to unmodified deterministic lookahead policies on a library of these lagged energy portfolio problems. In Chapter 4, we introduce a new dynamic model, which uses a parametric cos function approximation, that is capable of analyzing and identifying optimal natural gas procurement strategies, hydrogen storage, and customer contract negotiations. We design a several experiments to better understand how to effectively manage a portfolio of short term natural gas contracts and hydrogen storage cavern. We use this model to illustrate the relationship between the optimal level of hydrogen reserve and the reliability of production plants, evaluate multiple strategies for minimizing the exposure of the industrial gas supplier to volatile energy prices, and investigate the relationship between optimal natural gas procurement strategies and market volatility.

Though this work has introduced the concept of the Parametric Cost Function Approximation and demonstrated its potential in a multitude of settings, there exist significantly more research to be done. This new class of policies offers a new breadth of research possibilities such as identifying other appropriate problem classes and policy structures. We also recognize that gradient-based search mechanism are not always possible. Therefore developing derivative-free stochastic search methods for tuning CFAs is another potential area of future work, as well as designing methods to do adaptive search in an online setting.

Bibliography

Arrow, K. J., Harris, T., and Marschak, J. (1952). Optimal inventory policy. *Econometrica*, 20(1):133.

Bengio, Y. (2009). Continuous control with deep reinforcement learning. *Foundations and Trends® in Machine Learning*, 2(1):1–127.

Bertsekas, D. P. (2010). Dynamic Programming and Optimal Control 3rd Edition , Volume II by Chapter 6 Approximate Dynamic Programming Approximate Dynamic Programming. *Control*, II:1–200.

Bertsimas, D. and Goyal, V. (2012). On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming*, 134(2):491–531.

Bertsimas, D., Iancu, D. A., and Parrilo, P. A. (2011). A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control*, 56(12):2809–2824.

Birge, J. R. and Louveaux, F. (2011a). *Basic Properties and Theory.*

Birge, J. R. and Louveaux, F. (2011b). *Introduction to Stochastic Programming.*

Camacho, E. and Alba, C. (2013). *Model predictive control.*

Cao, X. R. (2009). Stochastic learning and optimization-A sensitivity-based approach. *Annual Reviews in Control*, 33(1):11–24.

Carpentier, P. L., Gendreau, M., and Bastin, F. (2015). Managing hydroelectric reservoirs over an extended horizon using benders decomposition with a memory loss assumption. *IEEE Transactions on Power Systems*, 30(2):563–572.

Chau, M. and Fu, Michael C, Huashuai Qu, I. O. R. (2014). Simulation Optimization: A Tutorial Overview and Recent Developments in Gradient-Based Method. *Proceedings of the 2014 Winter Simulation Conference*, pages 21–35.

Chen, H. and Baldick, R. (2007). Optimizing short-term natural gas supply portfolio for electric utility companies. *IEEE Transactions on Power Systems*, 22(1):232–239.

Conejo, A. J., Garc??a-Bertrand, R., Carri??n, M., Caballero, n., and de Andr??s, A. (2008). Optimal involvement in futures markets of a power producer. *IEEE Transactions on Power Systems*, 23(2):703–711.

Coulon, M., Powell, W. B., and Sircar, R. (2013). A model for hedging load and price risk in the Texas electricity market. *Energy Economics*, 40:976–988.

Deisenroth, M. P. (2011). A Survey on Policy Search for Robotics. *Foundations and Trends in Robotics*, 2(1):1–142.

Donohue, C. and Birge, J. (2006). The abridged nested decomposition method for multistage stochastic linear programs with relatively complete recourse. *Algorithmic Operations Research*, 1:20–30.

Duchi, J., Hazan, E., and Singer, Y. (2011a). Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research*, 12:2121–2159.

Duchi, J., Hazan, E., and Singer, Y. (2011b). Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research*, 12:2121–2159.

Dupačová, J. and Sladký, K. (2002). Comparison of multistage stochastic programs with recourse and stochastic dynamic programs with discrete time. *ZAMM Zeitschrift fur Angewandte Mathematik und Mechanik*, 82(11-12):753–765.

Durante, J. L., Nascimento, J., and Powell, W. B. (2017). Backward Approximate Dynamic Programming with Hidden Semi-Markov Stochastic Models in Energy Storage Optimization.

Fu, M. C. (2015). *Handbook of Simulation Optimization (International Series in Operations Research \& Management Science)*.

Glasserman, P. (1991). *Gradient Estimation via Perturbation Analysis*.

Hadjiyiannis, M. J., Goulart, P. J., and Kuhn, D. (2011). An efficient method to estimate the suboptimality of affine controllers. *IEEE Transactions on Automatic Control*, 56(12):2841–2853.

Harrison, J. M. and Van Mieghem, J. A. (1999). Multi-resource investment strategies: Operational hedging under demand uncertainty. *European Journal of Operational Research*, 113(1):17–29.

Ho, Y. (1992). *Discrete Event Dynamic Systems: Analyzing Complexity and Performance in the Modern World*.

Hu, J., Fu, M. C., Ramezani, V. R., and Marcus, S. I. (2007a). An Evolutionary Random Policy Search Algorithm for Solving Markov Decision Processes. *INFORMS Journal on Computing*, 19(2):161–174.

Hu, J., Fu, M. C., Ramezani, V. R., and Marcus, S. I. (2007b). An Evolutionary Random Policy Search Algorithm for Solving Markov Decision Processes. *INFORMS Journal on Computing*, 19(2):161–174.

Jin, S., Ryan, S. M., Watson, J.-P., and Woodruff, D. L. (2011). Modeling and solving a large-scale generation expansion planning problem under uncertainty. *Energy Systems*, 2(3-4):209–242.

Jirutitijaroen, P., Kim, S., Kittithreerapronchai, O., and Prina, J. (2013). An optimization model for natural gas supply portfolios of a power generation company. *Applied Energy*, 107:1–9.

Kiefer, J. and Wolfowitz, J. (1952). Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*.

Kusher, Harold; Ying, G. (2003). *Stochastic Approximmation and Recursive Algorithms and Applications*.

Lai, G., Margot, F., and Secomandi, N. (2010). An Approximate Dynamic Programming Approach to Benchmark Practice-Based Heuristics for Natural Gas Storage Valuation. *Operations Research*, 58(January 2016):564–582.

Lai, G., Wang, M. X., Kekre, S., Scheller-Wolf, A., and Secomandi, N. (2011). Valuation of storage at a liquefied natural gas terminal. *Operations Research*, 59(3):602–616.

Lan, S., Clarke, J.-P., and Barnhart, C. (2006). Planning for Robust Airline Operations: Optimizing Aircraft Routings and Flight Departure Times to Minimize Passenger Disruptions. *Transportation Science*, 40(1):15–28.

Levine, S. and Abbeel, P. (2014). Learning Neural Network Policies with Guided Policy Search under Unknown Dynamics. *Advances in Neural Information Processing Systems*, pages 1–3.

Lium, A.-G., Crainic, T. G., and Wallace, S. W. (2009). A Study of Demand Stochasticity in Service Network Design. *Transportation Science*, 43(2):144–157.

Löhndorf, N. and Wozabal, D. (2017). Indifference pricing of natural gas storage contracts.

Mannor, S., Rubinstein, R., and Gat, Y. (2003). The cross entropy method for fast policy search. *Machine Learning-International Workshop Then Conference-*, 20(2):512.

Mulvey, J. M., Vanderbei, R. J., Zenios, S. A., Apr, N. M., Mulvey, J. M., and Vanderbei, R. J. (1995). Robust Optimization of Large-Scale Systems Published. 43(2):264–281.

Ng, A. Y. and Jordan, M. (2000a). PEGASUS: A Policy Search Method for Large MDPs and POMDPs. *Conference on Uncertainty in Artificial Intelligence*, 94720:406–415.

Ng, A. Y. and Jordan, M. (2000b). PEGASUS: A Policy Search Method for Large MDPs and POMDPs. *Conference on Uncertainty in Artificial Intelligence*, 94720:406–415.

Pereira, M. V. F. and Pinto, L. M. V. G. (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52:359–375.

Perkins, R. T. and Powell, W. B. (2017). Stochastic Optimization with Parametric Cost Function Approximations.

Peshkin, L., Kim, K.-E., Meuleau, N., and Kaelbling, L. P. (2000a). Learning to Cooperate via Policy Search. *Association for Uncertainty in Artificial Intelligence*, pages 489–496.

Peshkin, L., Kim, K.-E., Meuleau, N., and Kaelbling, L. P. (2000b). Learning to Cooperate via Policy Search. *Association for Uncertainty in Artificial Intelligence*, pages 489–496.

Philpott, A. B., Craddock, M., and Waterer, H. (2000). Hydro-electric unit commitment subject to uncertain demand. *European Journal of Operational Research*, 125(2):410–424.

Philpott, A. B. and Guan, Z. (2008). On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450–455.

Pilipovic, D. (1998). *Energy risk: Valuing and managing energy derivatives*.

Powell, W., Ruszczyński, A., and Topaloglu, H. (2004). Learning Algorithms for Separable Approximations of Discrete Stochastic Optimization Problems. *Mathematics of Operations Research*, 29(4):814–836.

Powell, W. B. (2011). *Approximate Dynamic Programming: Solving the Curses of Dimensionality: Second Edition*.

Powell, W. B. and Meisel, S. (2015). Tutorial on Stochastic Optimization in Energy—Part I: Modeling and Policies. *IEEE Transactions on Power Systems*, 31(2):1–9.

Powell, W. B. and Ryzhov, I. O. (2012). *Optimal learning*.

Puterman, M. L. (2014). *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

Robbins, H. and Monro, S. (1951a). A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, 22(3):400–407.

Robbins, H. and Monro, S. (1951b). A Stochastic Approximation Method. *Ann. Math. Stat.*, 22(3):400–407.

Schwartz, E. S. (1997). The Stochastic Behaviour of Commodity Prices: Implication for Valuation and Hedging. *The Journal of Finance*, 52(3):923–973.

Secomandi, N. (2010). Optimal Commodity Trading with a Capacitated Storage Asset. *Management Science*, 56(3):449–467.

Secomandi, N., Lai, G., Scheller-wolf, A., and Seppi, D. (2010). The Effect of Model Error on the Valuation and Hedging of Natural Gas Storage. *New York*.

Sethi, S. and Sorger, G. (1991). A THEORY OF ROLLING HORIZON DECISION MAKING *. *Annals of Operations Research*, 29:387–416.

Shapiro, A. (2011). Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1):63–72.

Shapiro, A., Dentcheva, D., and Ruszczynski, A. (2009). *Lectures on stochastic programming: modeling and theory*.

Shapiro, A., Tekaya, W., Da Costa, J. P., and Soares, M. P. (2013). Risk neutral and risk averse Stochastic Dual Dynamic Programming method. *European Journal of Operational Research*, 224(2):375–391.

Silver, E. A., Pyke, D. F., and Peterson, R. (1998). Inventory Management and Production Planning and Scheduling. *Managerial and Decision Economics*, 11(January):297–315.

Singh, K. J., Philpott, a. B., and Wood, R. K. (2009). Dantzig-Wolfe Decomposition for Solving Multistage Stochastic Capacity-Planning Problems. *Operations Research*, 57(June 2014):1271–1286.

Spall, J. C., John Wiley & Sons., and Wiley InterScience (Online service) (2003). *Introduction to stochastic search and optimization : estimation, simulation, and control*.

Strassen, V. (1964). The exisitence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439.

Sutton, R. S. and Barto, A. G. (1998). Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks*, 9(5):1054.

Sutton, R. S., Mcallester, D., Singh, S., and Mansour, Y. (1999). Policy Gradient Methods for Reinforcement Learning with Function Approximation. *In Advances in Neural Information Processing Systems 12*, pages 1057–1063.

Takriti, S., Birge, J. R., and Long, E. (1996). A stochastic model for the unit commitment problem. *Power Systems, IEEE Transactions on*, 11(3):1497–1508.

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.

Tieleman, T. and Hinton, G. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4:26—-31.

Topaloglu, H. and Powell, W. B. (2006). Dynamic-Programming Approximations for Stochastic Time-Staged Integer Multicommodity-Flow Problems. *INFORMS Journal on Computing*, 18(1):31–42.

Wallace, S. W. and Fleten, S.-E. (2003). Stochastic Programming Models in Energy. *Handbooks in Operations Research and Management Science*, 10:637–677.