Design and Valuation of Demand Response Mechanisms and Instruments for Integrating Renewable Generation Resources in a Smart Grid Environment

Final Project Report

Power Systems Engineering Research Center

Empowering Minds to Engineer
the Future Electric Energy System

Design and Valuation of Demand Response Mechanisms and Instruments for Integrating Renewable Generation Resources in a Smart Grid Environment

Final Project Report

Project Team

Shi-Jie Deng, Project Leader
Georgia Institute of Technology

Shmuel Oren
University of California at Berkeley

George Gross
University of Illinois at Urbana-Champaign

PSERC Publication 12-27

October 2012

For information about this project, contact

Shi-Jie Deng, Ph.D.
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332
Tel: 404-894-6519
Fax: 404-894-2301
Email: deng@isye.gatech.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Acknowledgements

This is the final report for the Power Systems Engineering Research Center (PSERC) research project titled "Design and Valuation of Demand Response Mechanisms and Instruments for Integrating Renewable Generation Resources in a Smart Grid Environment" (project M-23). We express our appreciation for the support provided by PSERC's industry members and by the National Science Foundation under the Industry / University Cooperative Research Center program.

We thank all PSERC members for their technical advice on the project, especially Jim Price (California ISO), Jianzhong Tong (PJM), Feng Zhao (ISO New England), Mark Westendorf (Midwest ISO), Mark Sanford (GE Energy), who are our industry advisors.

Executive Summary

As the penetration of renewable generation resources in the power grid deepens, the intermittent nature of renewable generation poses significant challenges to the reliable operation of power grids by system operators. In particular, imposing higher requirements on reserves can accommodate the increasing integration of renewable resources. The increased reserve requirement can be met using the so-called demand response resources (DRRs) that play an increasingly important role in maintaining the supply-demand balance. Appropriate DRRs create demand-side flexibility, such as shifting, reducing, increasing, or curtailing the electricity loads, instead of dealing with the increasing uncertainty in the supply-side. In this project, we investigate the design of appropriate demand response (DR) mechanisms, which create additional controllable resources to attain supply-demand equilibrium for systems with intermittent renewable generation supply.

The project is separated into three parts. Part I addresses short-term wind speed/power forecasting and modeling by means of wavelet transform techniques. Based upon the forecasts, this part also analyzes the flexibility of thermostatically controlled loads (TCLs) as a source to create demand response for absorbing the high variability of renewable generation supply. Part II develops a simulation method that explicitly represents the various sources of uncertainty and the time-dependent nature of demand response resource utilization. Part III focuses on a contract that couples the operations of renewable energy resources with deferrable loads that can shift a fixed amount of energy demand over a given time window. An overview of the work accomplished in each part is summarized as follows:

Part I: Modeling and Forecasting Wind Speed via Wavelet Transform and Controlling Thermostats to Mitigate the Variability of Renewable Resource Outputs (Georgia Tech)

In wind speed/power forecasting part, we investigate a short-term wind speed forecasting method based on maximal overlap discrete wavelet transform (MODWT). To compensate time-frequency information that is neglected in most traditional methods, we combine wavelet analysis with existing statistical models, primarily time series models and regression models. We used a newly proposed variant of wavelet transform, namely MODWT, which can both extract real time frequency information and eliminate disadvantages of traditional Discrete Wavelet Transform (DWT) by means of a highly redundant non-orthogonal transform.

To be specific, when combined with time series models, we apply the MODWT-based Multi-Resolution Analysis (MRA), which utilizes the smooth and detail components of the original wind speed series reconstructed on each single independent level and mitigates the effects of boundary conditions. Our empirical analysis in this report demonstrates that MRA-based model has an impressive forecasting performance compared to other traditional statistical models. On the other hand, when combined with regression models, we utilize MODWT coefficients, which serve as feature variables for regression approaches. We select Bayesian Additive Regression Tree (BART) model as

our regression approach, which is a boosting-based nonparametric model, designed for estimating the expectation of an unknown nonlinear function given high-dimensional input variables generated by a set of low-dimensional "weak learners". Forecasting results based on NREL wind speed data show that the training fit of using history MODWT coefficients as primary input variables outperforms that of using history wind speeds. This demonstrates the promising potential of this approach in improving the performance of the existing wind speed forecasting models.

We also examine the flexibility of thermostatically controlled loads (TCLs) as a way to manage demand for absorbing the high variability of renewable energy resources. We propose an incentive-based control mechanism design, which allows the load serving entities (LSEs) to differentiate customers by offering various rebates on thermostat set-point adjustments. We apply a game-theoretic model to address the interactive behavior between LSEs and end users in contracting the control of thermostats such that the controlled loads follow the renewable energy supplies. Several control groups are formed as a result of customers' subscription to different rebate offerings. LSEs come up with a control law for each group to minimize the costs of following the output level of concerned renewable resources, such as a wind farm. We formulate the control problem as a linear-quadratic (LQ) tracking problem with inequality constraints. The state-space representation is derived from the transfer function, which relates the TCLs of a homogeneous group of customers and their thermostat set-points (see Kundu et al. (2011)). We apply the Model Predictive Control (MPC) technique to handle the problem due to some of its nice features. First, MPC is easily applicable to a LQ problem with inequality constraints. Second, MPC must be coupled with on-line states and parameters update. In our case, the reference signal, e.g., the wind power output is updated according to our short-term forecasting method. As accurate long-term wind power forecasting is difficult to obtain, MPC is ideal here to take advantage of the accuracy of the short-term prediction, instead of obtaining a controller off-line based on an inaccurate long-term forecast. Using an example with two control groups, we show that our design yields a piece-wise linear controller for both groups, which is straightforward to implement.

Part II: Simulation of Impacts of Demand Response Resources on Power System Variable Effects (University of Illinois at Urbana-Champaign)

The DRRs now actively participate in the electricity markets as buyers of electricity and sellers of load curtailment services by reducing their loads during certain hours. For these hours, the DRRs compete to provide the load curtailment services directly against the sales of the supply-side resources. The role of DRRs has become increasingly important in ensuring that the supply-demand balance is efficiently attained. The demand profile modifications resulting from the DRR curtailments affect the market outcomes and impact the generation and transmission resource utilization. In fact, DRRs impact and, in turn, are impacted by the planning activities on the generation and transmission side and the regulatory/legislative developments. Consequently, appropriate tools for quantifying the impacts of DRRs on market performance, generation and transmission resource utilization and other variable effects are required. One particularly important need is that of a simulation tool and we focus on addressing this need. We present in this report the development of a comprehensive methodology that provides the basis for such a

simulation tool. We illustrate its application to the quantification of various variable effects of large-scale power systems incorporating DRRs.

The principal challenge in the development of our simulation methodology is to integrate the effective representation of the time-dependent transmission-constrained markets and that of the supply and demand resources so as to construct a practical approach that can be computationally tractable when implemented for simulating large-scale systems over longer-term periods. The key goal is to deploy simulation methodology to quantify the impacts of DRRs on market performance, generation dispatch, transmission usage, environment and other system variable effects. We are interested in the study of the variable effects in planning and policy analysis studies. We construct the proposed simulation approach by marrying the concepts of probabilistic simulation with snapshot-based analysis techniques. In this way, we effectively integrate the representation of the supply- and demand-side resources, electricity markets, transmission grid operations and constraints, and various sources of uncertainty in the simulation. We discuss the implementation aspects of the proposed approach to ensure computational tractability so as to allow its application to the simulation of large-scale systems over longer-term periods. We present simulation results from representative case studies to illustrate the application of the proposed approach to quantification of the variable effects of a large-scale test system. The reported results demonstrate the economic, environmental and reliability benefits the integration of DRRs can provide to the power system.

Part III: Stochastic Modeling of Multi-area Wind Power Production and Evaluation of Direct Coupling between Deferrable Load and Intermittent Renewable Resources (University of California, Berkeley)

We have developed a multi-area wind production model that can be used as part of a stochastic unit commitment with transmission constraints. In order to assess the impact of wind power production on power system operations over an entire year, it is necessary to account for the non-stationary (seasonal and diurnal) patterns of wind power production. In our model we capture both the geographical diversity and the seasonal and diurnal patterns of wind power production while accounting for the temporal and spatial correlations. The model is calibrated to NREL wind speed data providing hourly time series of wind speed at different locations over a year period. Our model is designed to capture the statistical properties of the data set in a stochastic model that enables us to simulate multiple wind scenarios with the same spatial and temporal statistical properties as the original data set while accurately reproducing the marginal distribution of wind power production at each location of the network. The model is applied to a detailed dataset of the California wind power resources corresponding to the 2012 and 2020 Renewable Portfolio Standards, to produce counterfactual forecasts of wind power production under alternative assumptions regarding renewables penetration and demand response strategies.

In particular we focused on direct coupling of wind power resources with deferrable loads such as PHEV charging, HVAC, water pumping etc. which require a certain amount of energy over a time interval but the specific load can be shifted so as to follow intermittent supply resources. One of our research objectives was to assess the merit of

such a strategy as compared to a market based approach where responsive demand is mobilized through real time pricing, and to an idealized central dispatch of deferrable load. For that, we have developed a simulation platform that emulates two settlement markets via a stochastic unit commitment model which determines reserves requirements endogenously based on supply uncertainties and system contingencies. The model was applied to a case study for California which incorporates the wind model described above for five major wind locations and employs a reduced 225 bus representation of the WECC generation and transmission system. The case study demonstrates the effectiveness of stochastic unit commitment in reducing reserves cost and the potential benefits of the direct coupling paradigm as a means to operationally hedge the supply uncertainty and variability imputed by intermittent resources.

Project Publications:

A. Papavasiliou, S. S. Oren, "Integration of Contracted Renewable Energy and Spot Market Supply to Serve Flexible Loads", 18th World Congress of the International Federation of Automatic Control, August 28 – September 2, 2011, Milano, Italy.

A. Papavasiliou and S. S. Oren, "Stochastic Modeling of Multi-Area Wind Power Production", To be presented at PMAPS 2012, Istanbul, Turkey, June 10-14, 2012.

A. Kowli, M. Negrete-Pincetic and G. Gross, "A Successful Application of the Smart Grid: Demand Response Resources". Proceedings of the IEEE PES General Meeting, Minneapolis, MN, July 2010.

A. Kowli and G. Gross, "Simulation-Based Planning Approach for Systems with Time-Dependent Resources". Proceedings of the 2010 IREP Symposium - Bulk Power System Dynamics and Control - VIII, Buzios, Rio de Janeiro, Brazil, August 1-6, 2010.

Y.B. Lu, S-J. Deng, and X. Huo, "Electricity Price Spike Prediction via Boosting Trees and Wavelet Analysis". Technical Report, Georgia Institute of Technology, Jan 2012.

L. Xu and S-J. Deng, "An Incentive-based Demand Response Contract Design for Thermostatically Controlled Loads", working paper, Georgia Institute of Technology, 2012.

Student Theses:

A. S. Kowli, "Assessment of Variable Effects of Systems With Demand Response Resources," Masters thesis, submitted to meet the degree requirements in the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign 2010.

A. Papavasiliou, "Coupling Renewable Energy Supply with Deferrable Demand", PhD Dissertation, University of California at Berkeley.

L. Xu, "Financial and Computational Models in Electricity Markets", PhD Dissertation (in progress), Georgia Institute of Technology.

Part I

Modeling and Forecasting Wind Speed via Wavelet Transform and Controlling Thermostats to Mitigate the Variability of Renewable Resource Outputs

Shi-Jie Deng
Li Xu, Xinyu Min – Graduate Students
Georgia Institute of Technology

For information about Part I, contact

Shi-Jie Deng, Ph.D.
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332
Tel: 404-894-6519
Fax: 404-894-2301
Email: deng@isye.gatech.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Table of Contents

Table of Contents (continued)

List of Figures

List of Tables

1. Introduction

1.1 Background

Power electric industry is a major source of carbon emissions. In order to reduce air pollution worldwide, legislations are pushing forward to increase renewable energy integration. In the United States, most states have their own targets of renewable energy integration, known as renewable portfolio standard (RPS). For example, California sets a target of having 33% renewable generation capacity in its total capacity by 2020. However, renewable energy resources, such as wind and solar, are costly to be integrated into the current power grid. They possess some different features from the conventional coal and gas generation resources. On one hand, they are highly variable and intermittent. You don't know when wind blows and even during the same hours on two different dates, the wind speed patterns may behave dramatically different. The increasing uncertainty in the supply side requires fast responsive generators as additional reserves, when you are short of renewable energy supply. On the other hand, renewable energy has limited dispatchability. For example, when there is excessive wind, you can't ramp up or ramp down the wind generators as conventional ones to adjust the outputs. One possible way is to curtail some wind turbines and this is a kind of wasting cheap and clean energy.

Due to the high uncertainty in renewable energy supply, for example, wind power, system operators (SO) and load serving entities (LSEs) are transferring from managing the supply to demand side management. Accurate forecasts of wind speed as well as wind power at several minutes or hours ahead help grid operators in various aspects, which include maintaining the system reliability and power quality, and optimally dispatching conventional generations. Based upon wind power forecasts, LSEs may also design and implement appropriate demand response programs to mitigate the intermittency in wind power supplies.

1.2 Overview of the Problem

Demand response contracts are such tools designed to create demand flexibility (see Deng and Xu (2009) for a detailed survey of demand response mechanisms). In addition to traditional demand response programs offered by system operators, researchers are looking for other sources of flexible loads as demand responses. Among sources of energy consumption, Department of Energy in 2007 showed that thermostatically controlled loads (TCLs) including cooling, air conditioning, and refrigerating, account for 60%-70% of U.S. household energy consumption. In the smart grid environment, LSEs are able to direct control of remote thermostatic devices via programmable thermostats. Having agreements with end users, LSEs may exercise control by adjusting the temperature set-point instead of directly curtailing the power usage. In this way, the load patterns can be quickly shifted, usually in a minute scale to follow the renewable energy outputs, for example, the wind energy. The implementation of such kind of control of

TCLs requires two steps. First, we need an accurate wind speed/power forecasting method in a short-term scale to obtain a reference-tracking signal corresponding to the wind energy production. Second, LSEs design an incentive-based energy rate plan for end-users to self-select which results in controllable TCLs for tracking the forecasted wind energy output.

1.2.1 Wind Speed Modeling and Forecasting

Wind energy is a highly important source of renewal energy, the integration of which into the large-scale electricity system has become one of the most popular research problems in the power systems field. To eliminate the fossil CO_2 emissions and to decrease the exploitation of finite resources, energy experts are anticipating a surge in the ratio of this non-polluting and sustainable energy source to the total electricity consumption.

However, without accurate wind speed forecasting, a large-scale utilization of wind energy could get current grid operators and researchers into serious troubles, both technically and economically. The primary reason of difficulty in the integration of wind energy into the current power system is that the intermittent feature and high variability of wind pose many challenges in the operations of a power grid. Since wind power generation in wind farms fluctuates with local wind speed, which is largely dependent on local weather or meteorological conditions, a significant number of additional fossil fuel driven plants as balancing power are to be built to compensate the variation in wind energy, which in turn weakens the environmental benefits and raises operation cost of wind energy.

Fortunately, accurate forecasts of wind speed as well as wind power by several minutes, hours or even days ahead help grid operators in various aspects, which include maintaining the system reliability and power quality, optimally dispatching conventional generations, and implementing appropriate demand response programs to follow wind power supplies as well. Also, the amount of balancing power could be sharply reduced and wind power efficiency would be enhanced dramatically provided such a prediction is reliable enough. For these reasons, it is of great importance to develop an accurate wind speed forecasting method to facilitate a high penetration level of wind power resources.

Previously, many researchers have proposed a number of wind speed forecasting methods with limited accuracy. Most of these methods could be classified as physical models or statistical approaches. Physical models mainly concentrate meteorological information and equations of motion of the atmosphere. These models either carry out forecasting by operational fluid dynamical simulations or implement diagnostic methodology by utilizing parameterizations of boundary layer. One of the most important physical models is Numerical Weather Predictions (NWP) model, provided by national meteorological forecasting agencies. There are also some other variant of physical models based on spatial information. On the other hand, statistical approaches stress on how to utilize merely historical wind speed data to predict future wind speed in a short-term scale. Typical statistical approaches are time series analysis and machine learning techniques, including ARX, Artificial Neural Networks (ANN), Fuzzy Logic, etc. Though it is hard

to tell which model is the best or which class of models is better, it has been generally believed that physical models have advantages in long-term forecasting while statistical models perform better in short-term prediction.

In this project, we apply novel wavelet transform techniques to wind speed-modeling methods in purpose of achieving accurate forecasts of wind speed in a short-term scale. To be specific, we propose a short-term wind speed forecasting method based on maximal overlap discrete wavelet transform (MODWT). We combine wavelet analysis with existing statistical models, primarily time series models. The intuitive reason underlying this combination is that most of time series methods require stationarity or other properties which wind speed series do not satisfy, while wavelet transform could decompose these series into sub series, which are suitable for time series modeling and forecasting, and then reconstruct them to generate a complete forecasting series. However, the ordinary discrete wavelet transform (DWT), which is mostly used in previous models, is sensitive to the different ways of separating wind speed series. This indicates that wavelets coefficients critically depend on the starting point of the time series, which means that if a different starting time point is chosen or a little shift is added to the wind speed series, the coefficients may change significantly for the same time point. To overcome this shortcoming, we apply MODWT, which serves as a modified version of DWT to eliminate the effect of starting time by means of a highly redundant non-orthogonal transform.

However, although those MODWT coefficients exhibit desirable properties of time invariance compared to that of ordinary Discrete Wavelet Transform (DWT), they sometimes incur practical problems due to boundary concerns. So, we take wavelet Multi-Resolution Analysis (MRA) into consideration. The primary reason lies in the fact that the MODWT-based MRA -- the smooth and detail components of the original wind speed series reconstructed on each single independent level -- mitigates the effects of boundary condition and thus is more suitable for time series modeling. In our experiment, the combination of time series models and MRA generate impressive forecasting results that outperform other statistical models, which shows that such a combination is suitable.

Moreover, in order to capture more complexity of wind speed series and even to incorporate more related features, we investigate another innovative wind speed forecasting approach combining Bayesian Additive Regression Trees (BART) and MODWT. BART is a boosting-based nonparametric Bayesian regression model, aimed to estimate the expectation of an unknown nonlinear function given high-dimensional input variables by a sum of low-dimensional “weak learners”. The high-dimensional input variables should be chosen appropriately; If not, the results might not be able to reflect the authentic inherent dynamics of wind speed series. Because the MODWT coefficients of wind speed essentially are the weighted averages at each scale and the weighted differences between each scale, which describe the overall features of a wind speed series, it is justifiable to use MODWT coefficients as primary input variables. Our experiment shows acceptable forecasting results for this hybrid model.

1.2.2 Demand Response Contract Design for TCLs

Modeling of TCLs is well documented in the literatures. Ihara and Schweppe (1981) start to model the dynamics of a single thermostat in the population from a physical view. Malhame and Chong (1985) describe the dynamics of aggregated TCLs by using Fokker-Planck equation. Ucak and Dokuyucu (2010) use Monte Carlo simulation to study TCLs as a direct control method. Lu et al. (2005) discuss the modeling of uncertainties in aggregated TCLs using a state queuing model. Recently, more researchers have been looking at TCLs as demand responses to track an exogenous signal. Callaway (2009) propose a minimum variance controller of manipulating thermostat set-points of a homogeneous group to track a wind power generation signal. Kundu et al. (2011) derive a transfer function relating the aggregate response of a homogeneous group of TCLs to disturbances that are applied uniformly to the thermostat set-points, and they present a linear quadratic regulator for tracking purpose. Both works assume the whole path of the tracking signal is known. But usually the error in a long-term forecast of wind power is very large, even in hours ahead forecast. Using forecasted values, as a whole trajectory to get a close loop control signal will produce large control errors in the real time operations. Koch et al. (2011) extend the idea to a heterogeneous group by using Markov chain, and apply model predictive control techniques, which only require one-step ahead forecast of wind power as a model input. In addition, the above papers assume the manipulating of set-points does not cause any discomfort of customers ensuring the room temperature stays inside the set-point dead band. However, during some extreme events, e.g. wind stops blowing or wind production is extremely low, their mechanisms do not offer sufficient amount of flexible loads for control, which prevent reducing the demand further to match the wind profile. It is crucial to maintain the system reliability in such situations, so that costly additional reserves have to come in places.

In this part, we assume LSEs design contracts, which offer incentives to end users to encourage their participation in demand responses. Customers are willing to bear some discomfort by allowing the temperature to be outside of the dead band in exchange for some rebates. In this way, LSEs gain access to more flexible loads for control to absorb the variability of renewable energy outputs. By offering various rebates on thermostat set-point adjustments, LSEs differentiate customers into several control groups. We apply a game-theoretic model to address the interactive behavior between LSEs and end users in contracting the control of thermostats such that the controlled loads follow the renewable energy supplies. Customers have a chance to decide whether they opt-in or opt-out these TCL programs based on the discomfort levels (represented by set-point adjustment limits) and the potential bill savings (rebate levels). Several control groups are formed as a result of customers' subscription to different rebate offerings. LSEs come up with a control law for each group to minimize the costs of following the output level of concerned renewable resources, such as a wind farm. We formulate the control problem as a linear-quadratic (LQ) tracking problem with inequality constraints, and apply model predictive control (MPC) technique to handle the problem. MPC is attractive here due to, first, it is easily applicable to a LQ problem with inequality constraints. Second, as an input of MPC, the reference signal is updated according to the short-term 10 minutes forecasting at each step.

1.3 Report Organization

The first part of this report investigates modeling and forecasting wind speed via wavelet transform and controlling thermostats to mitigate the variability of renewable resource outputs. The rest of the part is organized as follows: in Section 2, we address the issue of short-term wind speed forecasting via a novel application of wavelet transform. In Section 3, we address the contract design for TCLs to serve as demand response resources. We provide a game-theoretic framework studying LSEs' and customers' interactive behaviors, and a solution of LSEs' control problem by MPC. A case study with analysis of the sensitivity of contract parameters to tracking errors and total costs is also provided. We conclude in Section 5.

2. Modeling and Forecasting Wind Speed via Wavelet Transform

2.1 Motivation

Previous research has shown that statistical wind forecasting models have relatively high performance compared to physical models on short-term wind speed prediction, whose prediction window varies from 10 min to 5 hours, while physical wind forecasting models dominate on long-term wind speed prediction domain, whose prediction interval is usually around one day or 48 hours. Since wind speed prediction results serve as an input for our demand response mechanisms, and our demand response mechanisms require wind speed prediction window to be from 10 min to 30 min, it is natural to apply statistical models rather than physical model to wind speed prediction to achieve higher accuracy and better performance.

However, current statistical models for wind speed forecasting mainly focus on wind speed interactive patterns and inherit relationship on time domain, while neglect the rich information of wind speed on frequency domain or time-frequency domain. The reason lies in that traditional frequency analysis tools, like Fourier transform, cannot properly extract the frequency information from highly variable wind speed series with frequency changing with time.

As a time-frequency analysis tool, wavelet transform, on the other hand, could overcome this shortcoming efficiently and achieve Multi-Resolution Analysis (MRA). So, combining wavelet transform with traditional statistical wind speed forecasting models would definitely help us penetrate deeper into the essential relationship of wind speed time series on time-frequency domain and further improve forecasting performance.

Moreover, rather than using traditional Discrete Wavelet Transform (DWT), we apply Maximal Overlapped Discrete Wavelet Transform (MODWT) to the combination mentioned above. The coefficients of traditional DWT have great dependency on the starting time point of the time series, which might cause undesirable impact on time series analysis. But, as a novel wavelet transform variant, MODWT implements a highly redundant non-orthonormal transform, and successfully eliminates such impacts. Thus, MODWT serves as a suitable and practical mathematic tool in time-frequency analysis and further wind speed forecasting.

Although those MODWT coefficients exhibit desirable properties of time invariance compared to that of DWT, they sometimes have boundary problems, which prevents their application in time series. On the other hand, the MODWT-based MRA decomposes the original series into different scales, and then reconstructs them on each scale into the smooth and detail components, namely MRA series. Such smooth and detail components will not have any boundary problems and thus is suitable for time series modeling. That is the reason why we utilize MRA series rather than MODWT coefficients to combine with time series model.

Finally, when considering the combination of MODWT with BART model, we mainly focus on which kind of series best extracts the time-frequency information and the internal dynamics in the original wind speed time series. Since MODWT coefficients of wind speed essentially are the weighted averages at each scale and the weighted differences between each scale, these coefficients successfully describe the overall features of a wind speed series. So, it is reasonable to use MODWT coefficients as primary input variables and this turns out to be an efficient way to combine wavelet transform with regression models.

2.2 Wind Speed Forecasting via Combination of MRA and Time Series Models

2.2.1 DWT and MODWT

The definition of Discrete Wavelet Transform (DWT) is as below:

A real-valued function $\psi(t)$ is defined as a wavelet, if

$$\int_{-\infty}^{\infty} \psi(t)dt = 0$$

$$\int_{-\infty}^{\infty} \psi^2(t)dt = 1$$

The family of DWT basis for this wavelet are:

$$\psi_{j,n}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - n),\ j, n \in Z$$

The coefficients of DWT are:

$$c_{j,n} = \ < f(t), \psi_{j,n}(t) > \ = \int_{-\infty}^{+\infty} f(t)\psi_{j,n}^{*}(t)dt\,,$$

where $f(t)$ is the corresponding signal.

The Inverse of DWT is:

$$f(t) = \sum_{j,n} c_{j,n}\psi_{j,n}(t)$$

To improve the speed of calculation of DWT coefficients, Mallat proposed an algorithm (Mallat Algorithm), which has complexity as low as $n\log(n)$:

A series h_l is defined as wavelet filter, and a series g_l is defined as scale filter, if

$$\sum_{l=0}^{L-1} h_l = 0, \sum_{l=0}^{L-1} h_l h_{l+2n} = \begin{cases} 1, & \text{if } n = 0 \\ 0, & \text{o/w} \end{cases}$$

$$g_l = (-1)^{l+1} h_{L-1-l}$$

where L is the total length of the time series.

First we filter the original time series using both filters to get the scale coefficients a_1 and the wavelet coefficients d_1 of the first level; Then we filter the previous scale coefficients using the same filters to get the new scale and the wavelet coefficients of the second level; Continue going on until we get all level's information.

One step of Mallat process is showed in the figure below:

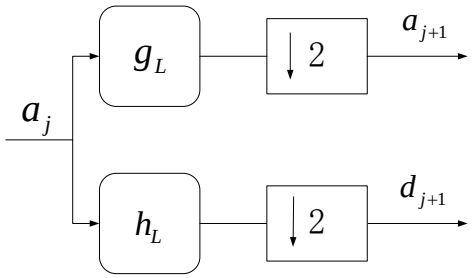


Figure 2.1 Mallat Algorithm for DWT

Although Mallat Algorithm is extremely practical and has led wavelet transform to industry, there exist some drawbacks in it. One of the most important drawbacks is that it can only deal with data whose length is a multiple of 2. When applying Mallat Algorithm to real time series, researchers have to cut the time series to certain multiples of 2 to make the algorithm work. Another drawback is that when you shift the time series along the time line while change nothing else, the corresponding wavelet coefficients are changed.

Percival and Walden proposed MODWT to overcome this shortcoming by changing the two filters a little bit, as below:

$$\tilde{h}_l = h_l / \sqrt{2}$$

$$\tilde{g}_l = g_l / \sqrt{2}$$

and they no longer downsample the output of the filters.

Percival & Walden claim that their MODWT is not orthogonal, redundant but more useful than Mallat DWT, because MODWT is shift invariant and can be applied to all data length. Its filter process is easier, as below:

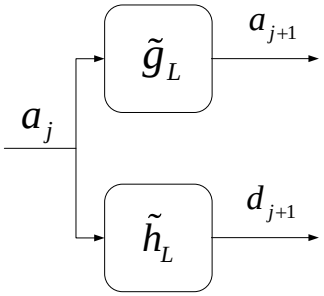


Figure 2.2 MODWT Filter Process

So we could utilize MODWT to decompose wind speed data, reconstruct it on single branch and then do forecasting.

2.2.2 Holt-Winters Model

Holt-Winters Model is a popular time series analysis approach in recent years. Its reliability, robustness and performance come from its capability to accurately handle the relationship between the level and the trend of a time series.

The additive Holt-Winters prediction function (for time series with period length p) is

$$\hat{y}(t+h) = a(t) + h \cdot b(t) + s(t+p+1+(h-1) \bmod p)$$

where

$$a(t) = \alpha[y(t) - s(t-p)] + (1-\alpha)[a(t-1) + b(t-1)]$$

$$b(t) = \beta[a(t) - a(t-1)] + (1-\beta)b(t-1)$$

$$s(t) = \gamma[y(t) - a(t)] + (1-\gamma)s(t-p)$$

It is obvious to see that the key ideas of Holt-Winters lie in its three factors: level a, trend b and season s.

2.2.3 Model Description

Our first proposed model is Wind Speed Forecasting via combination of MRA and Holt-Winters Models. Here, our model consists of two separate parts: Wavelet Transform and Time Series Analysis.

For Wavelet Transform Part, we firstly do MODWT decomposition to the historical wind speed data, and get the MODWT coefficients. Then, we do single branch reconstruction to the MODWT coefficients, and get the MRA.

For Time Series Analysis Part, after calculating each scale's MRA of the historical wind speed time series data, we input these MRA on each scale into the Holt-Winters model

and output the prediction on each scale. Finally, we synthesize all scales of wind speed to get the final wind speed prediction.

The process is shown in Figure 2.3. Note that sometimes the first level MRA detail looks like pure noise. If so, it is natural to remove these details and set them to zero.

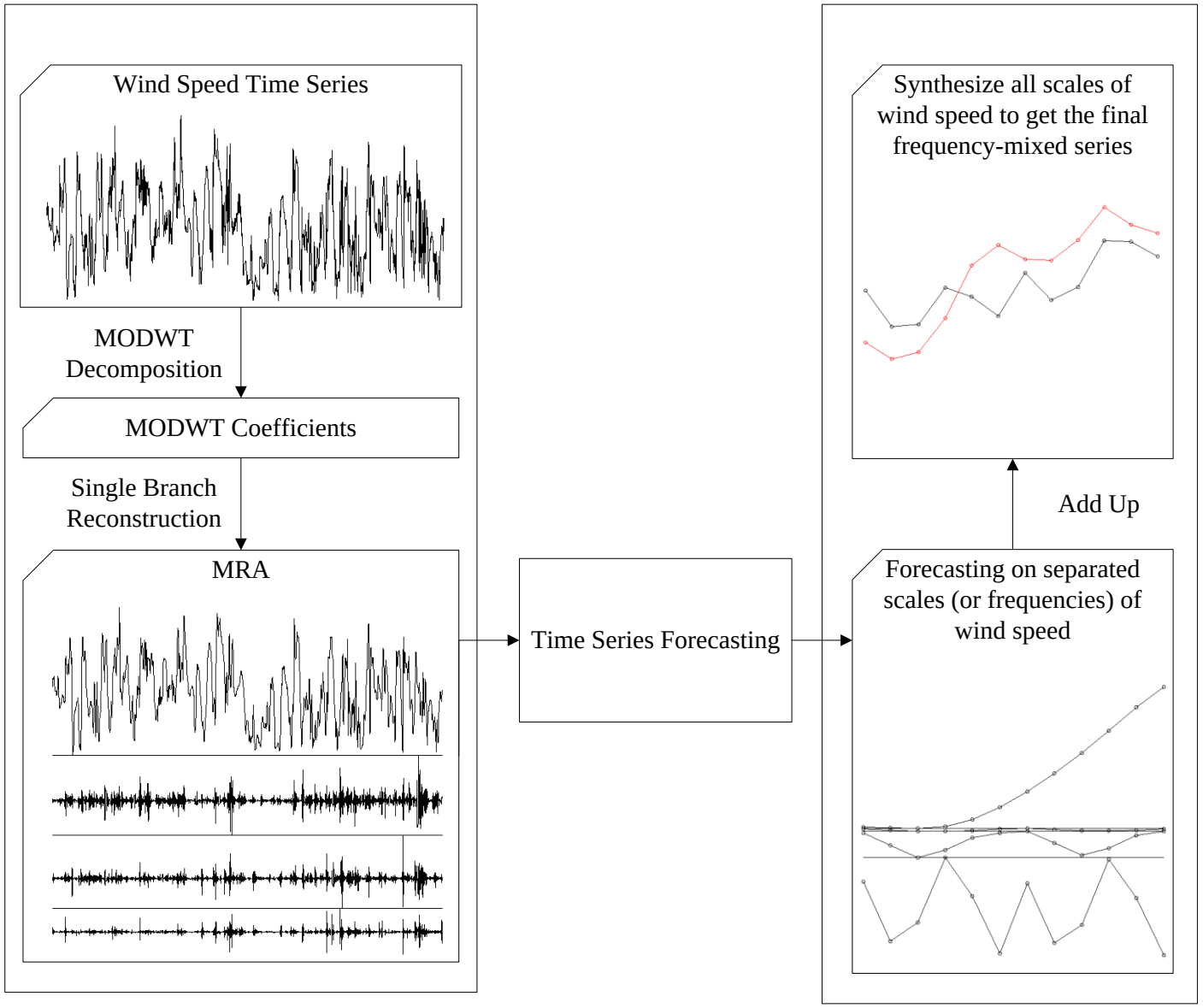


Figure 2.3 Combination of MRA and Holt-Winters

2.3 Wind Speed Forecasting via Combination of MODWT and BART

2.3.1 BART Model

Bayesian Additive Regression Trees (BART) model is a nonparametric Bayesian regression approach based on boosting algorithms and iterative Bayesian backfitting MCMC algorithm, proposed by Chipman (2008).

The essential idea of BART is to estimate an unknown function f satisfying

$$Y = f(W) + \varepsilon, \varepsilon \sim N(0, \sigma^2)$$

where Y is the wind speed to predict, W is a vector of selected features extracted from historical data and ε is the noise.

To achieve such estimation, BART approximates above function as a sum of m trees:

$$f(W) \approx \sum_{j=1}^{m} g_j(x)$$

where g_j denote one small regression tree.

Each small regression tree serves as a low dimension weak learner, whose performance is only slightly better than randomness. These weak learners add together to construct a powerful learner, which can incorporate complex and nonlinear interactive relationship between feature variables. Such a property of BART model will introduce significant flexibility to the modeling of wind speed and will achieve great forecasting performance if feature variables are carefully selected.

Feature selection is one of the most important components of BART model. On one hand, bad feature variables will mislead the learners and generate unrealistic interactive relationship of feature variables, which results in unsatisfactory training or testing performance. On the other hand, good feature variables which contain rich information on the essential structure of the problem will induce the model into the right track and generate the right tree structure, which in turn determines a good training and testing results.

Besides containing rich information, good feature variables also should not have too high dimensions; Otherwise, BART model might not be able to handle so much flexibility. Even we can improve flexibility by increasing the number of trees, this might cause undesirable over-fit problems.

The training procedure for BART mainly consists of setting prior for the regression trees and extracting information from posterior by MCMC, which is based on Metropolis-Hasting (MH) algorithm and Gibbs sampling. More training detail could be found in [10].

2.3.2 Model Description

Our second proposed model is Wind Speed Forecasting via combination of MODWT and BART Model. Here, this model also consists of two separate parts: Wavelet Transform and BART.

The wavelet transform part accomplishes the primary task of feature selection. As feature variable selection is one of the most important components for BART model, we must think of variables that keep intensive information on both time domain and frequency domain as well as that remain relatively low dimension.

MODWT wavelet coefficients can satisfy the above requirement, because each MODWT wavelet coefficient point represents the information around the corresponding time point in time-frequency domain. So we decide to use MODWT wavelet coefficients as feature variables for BART model.

We firstly split wind speed data into training set and testing set, and we use training set as historical data and use testing set to predict wind speed. We do MODWT decomposition to the training set, and select a proper number of MODWT coefficients as feature variables to feed into the BART model. After training, BART model generates a number of fully-grown regression trees. Then, we test it by inputting the MODWT coefficients of historical part of testing data into these fully-grown regression trees. Finally, we sum up all of the output from each leaf on these trees, and the result is our prediction.

The process of this approach is shown in Figure 2.4. Note that feature variables could be any kind of variables. Besides MODWT coefficients, other factors, like temperature, humidity, and pressure, which can significantly affect the future wind speed, could be easily incorporated in this BART and MODWT model. Thus, this model is really expandable and flexible to use.

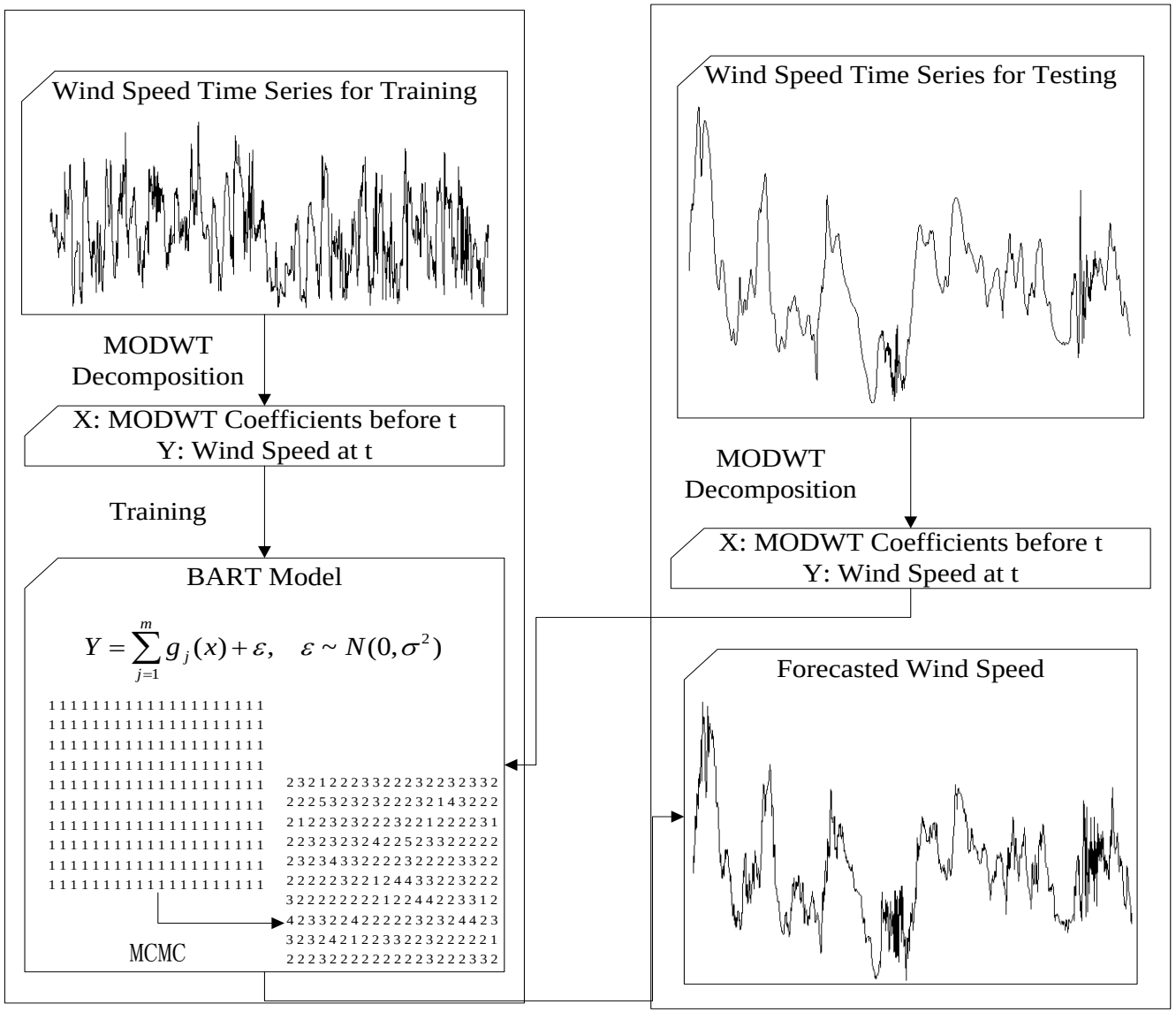


Figure 2.4 Combination of MODWT and BART

2.4 Case Study

In this section, we will testify the proposed two models using wind speed data. Moreover, we will compare these models with traditional statistical wind speed forecasting models and illustrate the reason behind their different performance.

2.4.1 Data Description

Our data are from Western Wind Dataset from National Renewable Energy Laboratory (NREL). Due to high wind speed and many wind farms in Texas, we collect 100m wind speed data with 10min interval during January 2006 at Texas. Node TX14 (-104.458,562.009) is randomly selected as the location.

To better plot the prediction errors and compare the difference of participated models, we extend our original prediction window from 10 min interval to 30 min interval, which means our prediction interval is now three time points ahead. By each time point, we will be able to use its historical wind speed data 3 time points before, but be not allowed to use any information in the future.

2.4.2 Results for Wind Speed Forecasting via MRA and Holt-Winters

We select 3 days' forecasting result to present: January 4^{th}, 10^{th}, and 17^{th}.

Since we perform Holt-Winters modeling on both smooth and detail scale, so we call it MRA-HWHW in short.

During these three days, our MODWT wavelet filter is the simplest wavelet—Haar function. The reason for choosing Haar function is to minimize the calculation time. We can use other wavelet functions, but the result will be similar, because MODWT is not sensitive to different wavelet filters, according to Percival and Walden.

The time interval that is being predicted is nearly the same, from 0:00 a.m. to 12:00 p.m..

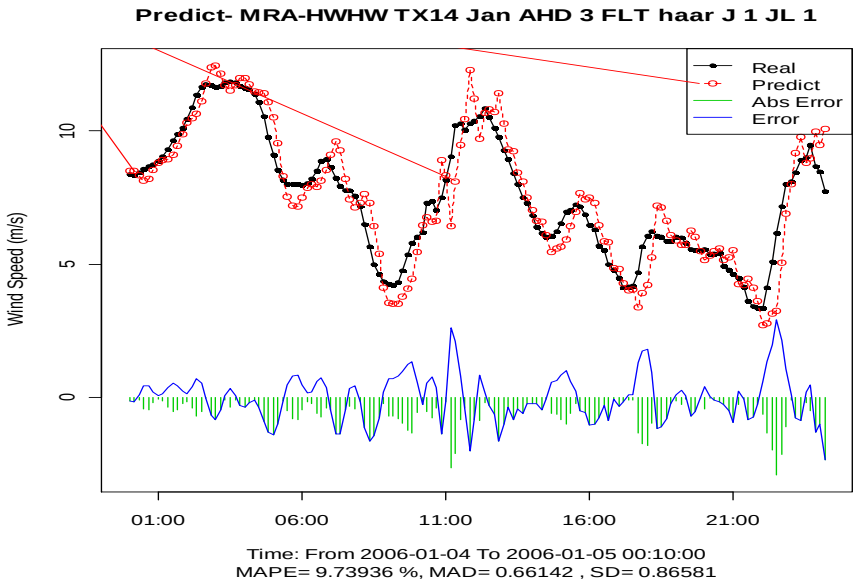


Figure 2.5 Jan 4th MRA-HWHW Forecasting Results

For Jan 4th, the MAPE is 9.74%, MAD is 0.66 while SD is 0.87. The prediction accuracy is acceptable.

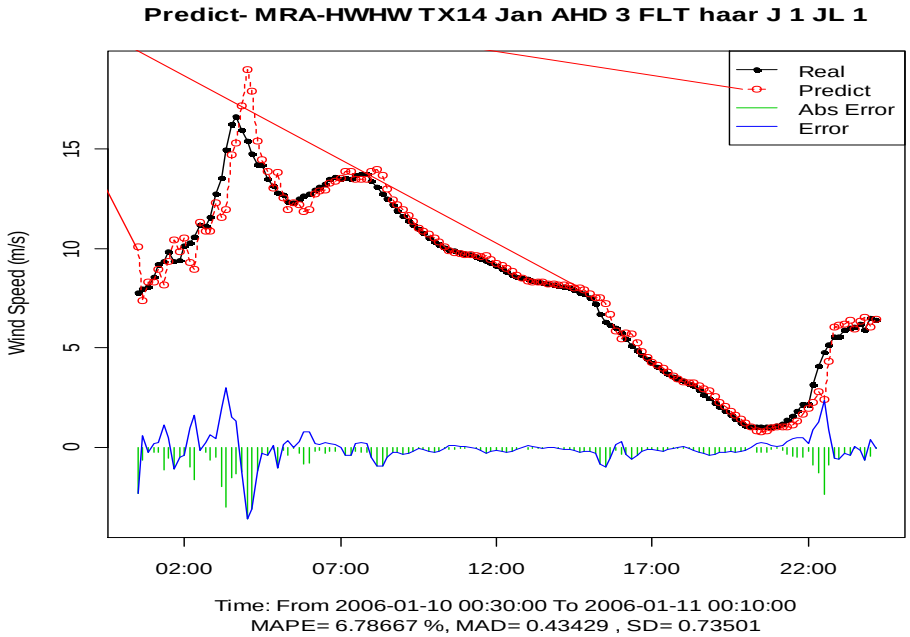


Figure 2.6 Jan 10th MRA-HWHW Forecasting Results

For Jan 10th, the MAPE is 6.79%, MAD is 0.43, and SD is 0.74. We can observe that the wind speed series on Jan 10th is pretty smooth from 7:00am to 20:00pm. So, the MAPE is quite low.

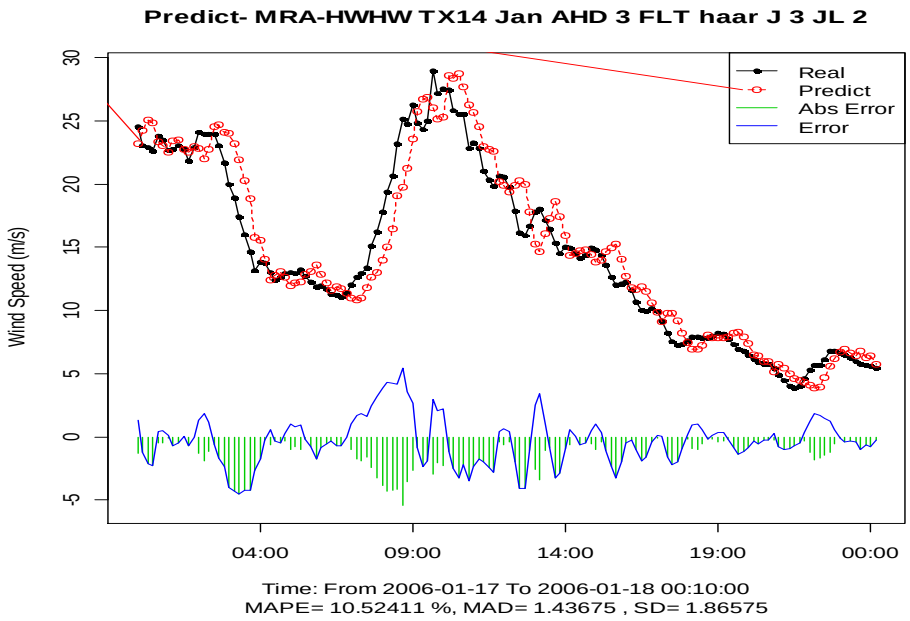


Figure 2.7 Jan 17th MRA-HWHW Forecasting Results

For Jan 17th, the MAPE is 10.52%, MAD is 1.44, and SD is 1.87. We can see that the major error happens on 9:00 a.m. where wind speed suffers a sharp change. After that change, the predicted wind speed follows the real wind speed smoothly later in the day.

Table 1 shows the comparison result on these three days in January between MRA-HWHW and other traditional statistical wind speed forecasting models. It is obvious to see that MRA-HWHW outperformed other models on nearly all measurement (MAPE, MAD and SD). The only slight failure for MRA-HWHW is on Jan 17th, when it did not perform well as compared to ARIMA.

Table 2.1 Comparison between Different Approaches

DATE	ERROR	PERSIST	HOLT-WINTERS	ARIMA	MRA-HWHW
Jan 4th	MAPE	11.90%	10.77%	10.16%	9.74%
	MAD	0.81	0.72	0.69	0.66
	SD	1.03	0.99	0.93	0.87
Jan 10th	MAPE	11.43%	7.62%	9.65%	6.79%
	MAD	0.68	0.48	0.57	0.43
	SD	0.91	0.83	0.81	0.74
Jan 17th	MAPE	10.73%	10.89%	10.31%	10.52%
	MAD	1.44	1.50	1.44	1.44
	SD	1.89	1.90	1.85	1.87

Thus, the performance of MRA-HWHW is satisfactory compared to other traditional statistical wind speed forecasting models.

2.4.3 The Strength of combination of MRA and Holt-Winters

In this section, we will illustrate why MRA-HWHW outperforms other benchmark models: MRA-HWHW has relatively good performance over periods when the wind speed has very high volatility. MRA separates wind time series into several different sub-series under different frequencies, which means it is easier for de-noising and it is more convenient for Holt-Winters to handle the relatively simple sub-series.

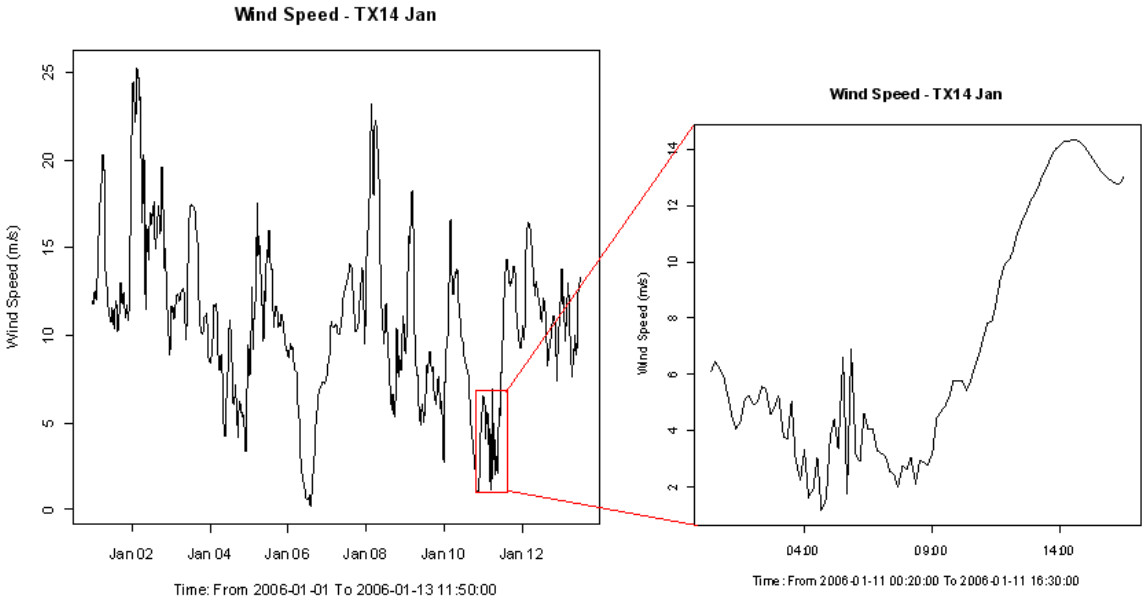


Figure 2.8 High Volatility on Jan 11th

Figure 2.8 shows that there is an abnormal period of wind speed with significantly high volatility in Jan 11th. We plot the prediction results of different models as below:

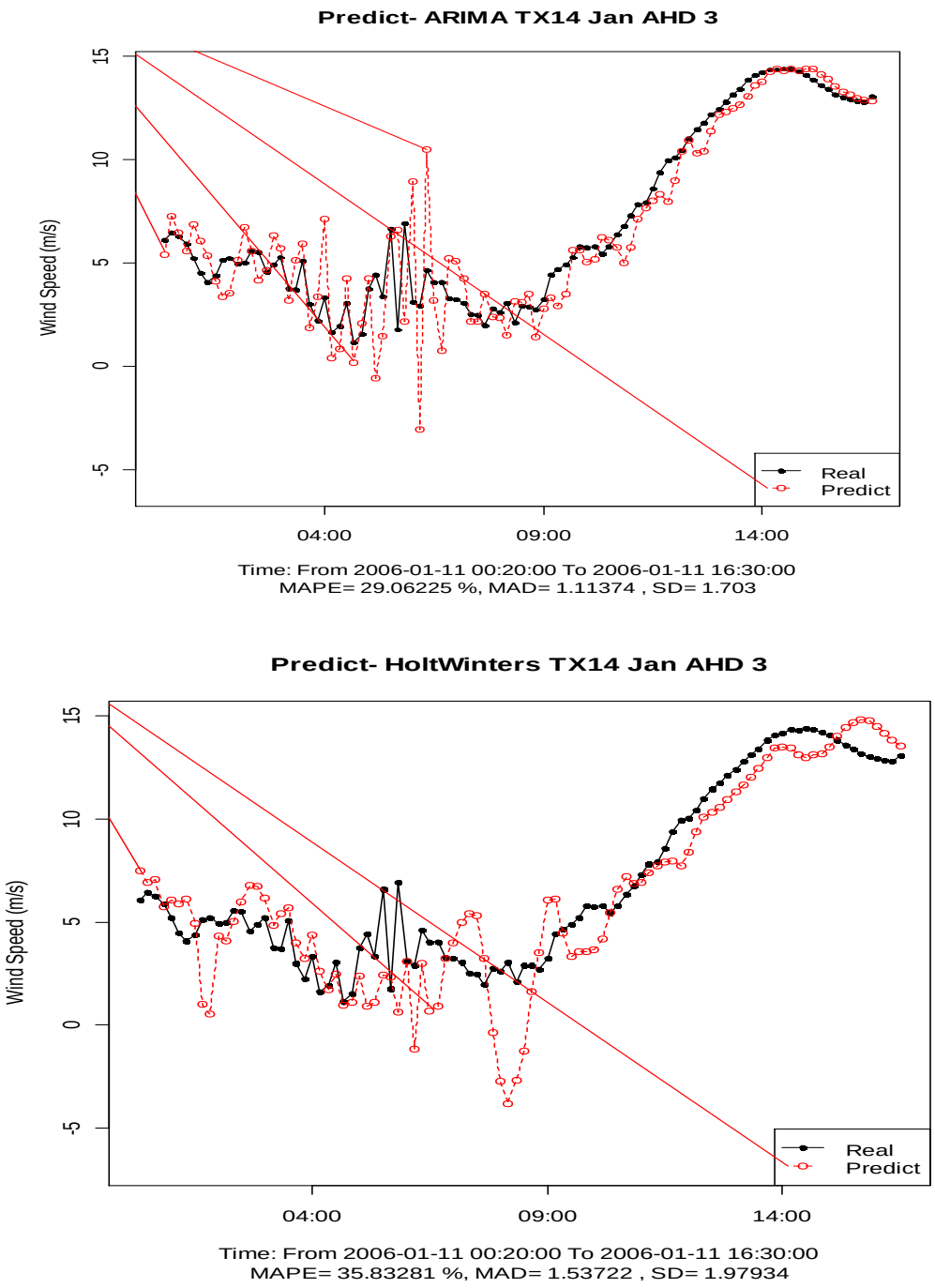


Figure 2.9 Comparison of Different Models on Highly Volatile Wind Speed

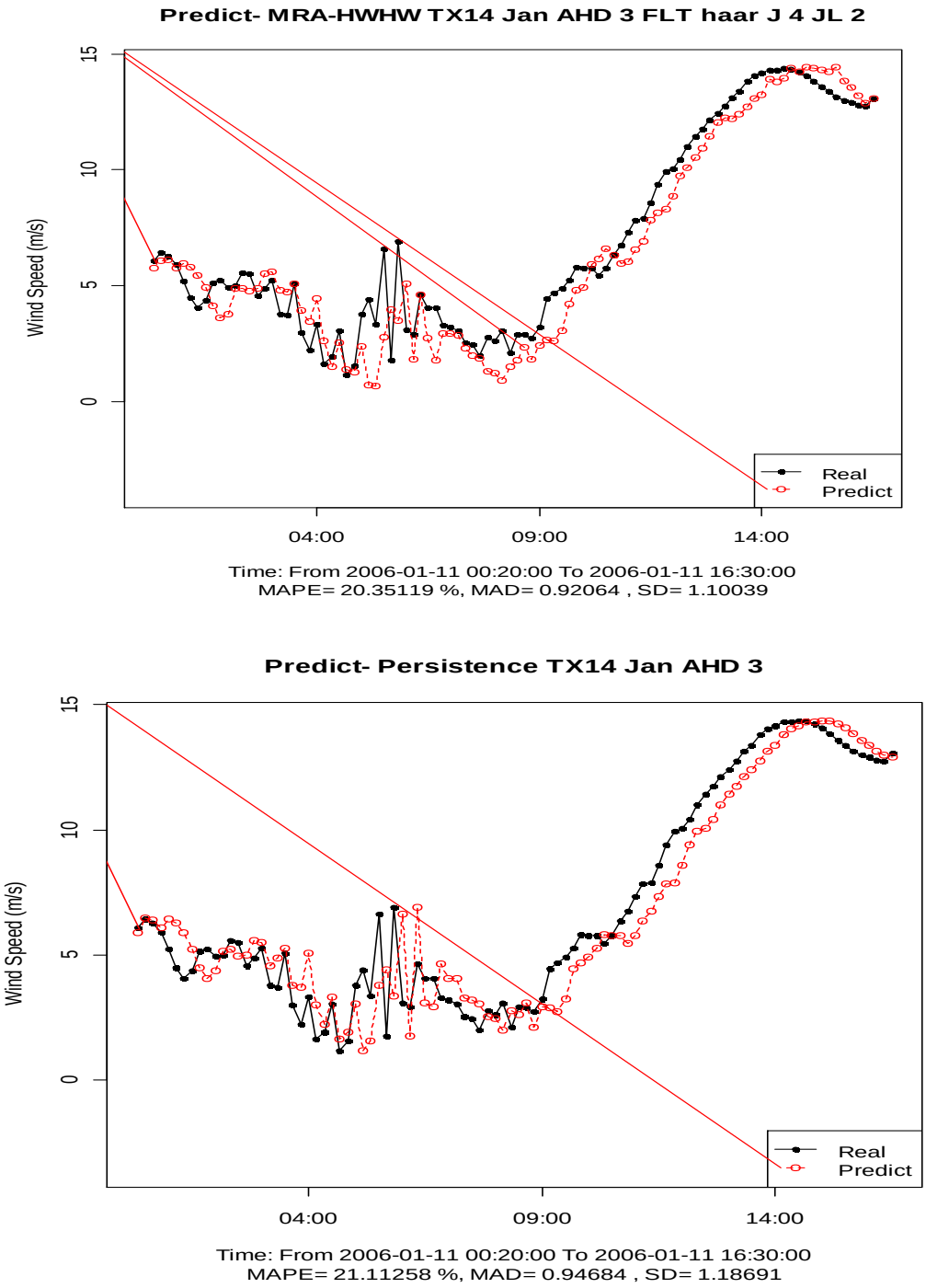


Figure 2.9 Comparison of Different Models on Highly Volatile Wind Speed (continued)

From Figure 2.9, we can find that MRA-HWHW has the lowest MAPE and remain relatively smooth during the period with high volatility. However, Holt-Winters and ARIMA tend to go up and down quickly with the variable wind speed, which results in a very bad performance.

Thus, the combination of MRA and Holt-Winters could effectively eliminate the impact of short-term high volatility on forecasting performance, and outperforms other traditional statistical models on that aspect.

2.4.4 Results for Wind Speed Forecasting via MODWT and BART

Our training set is wind speed data from Jan 1st to Jan 8th. Our testing set is wind speed data of Jan 9th.

The parameter for the BART model is: The tree number is 200. We use 6 historical MODWT coefficients as feature variables.

Figure 2.10 shows the training result from Jan 1st to Jan 8th. We can find that the MAPE is as low as around 4%.

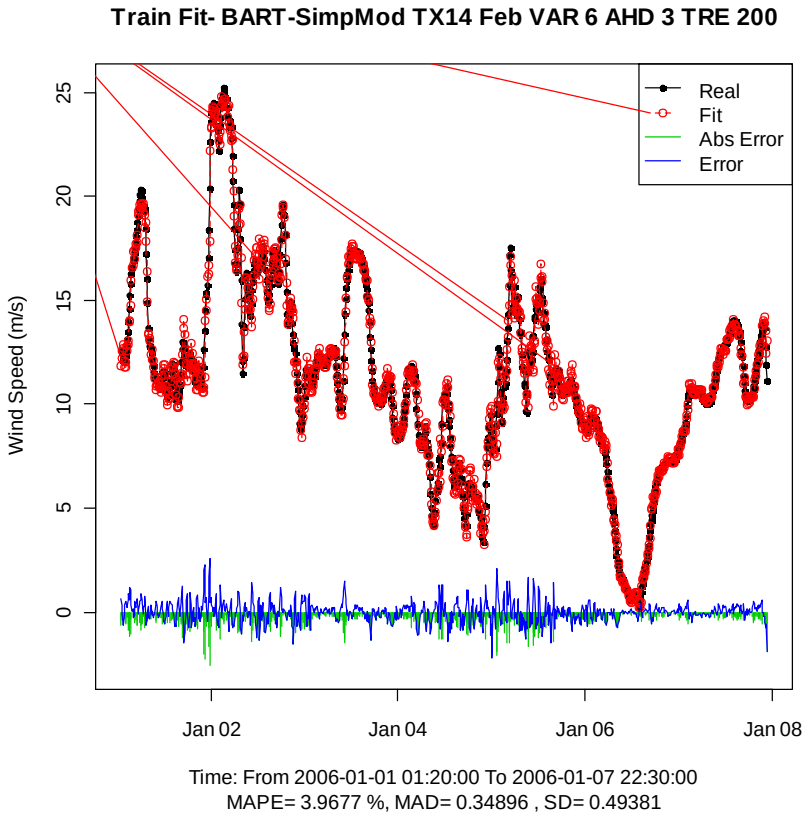


Figure 2.10 Training Result for BART and MODWT

Figure 2.11 shows the testing result, and we can see that the MAPE is a little bit higher than that of MRA-HWHW on Jan 10th, but it is still acceptable.

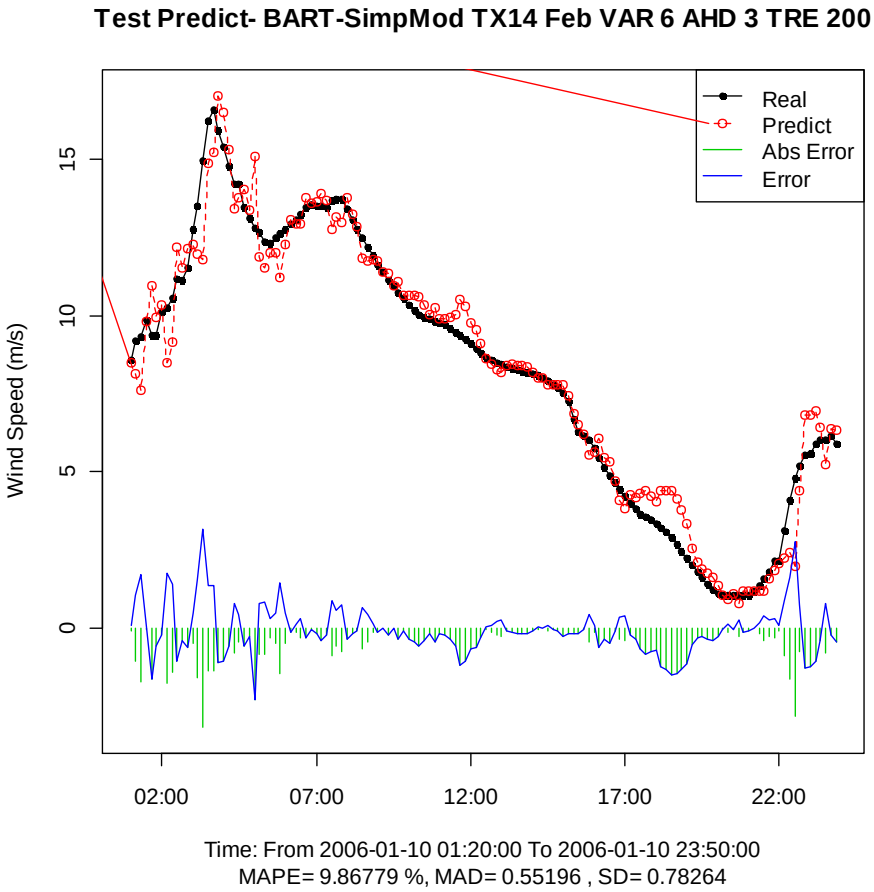


Figure 2.11 Testing Result for BART and MODWT

Note that here we do not use any information on Jan 9th, which indicates that such regression models based on training and testing might have advantages over traditional time series data if some large pieces of data are missing or not available. Also, more related information like temperature, humidity or pressure could be incorporated into this model easily. Thus, this model shows great extensibility on wind speed forecasting.

3. Incentive-Based Demand Response Contract Design for TCLs

3.1 Contract Structure

LSEs offer a pricing menu $\mathbf{p(u;a)} = \bar{\mathbf{p}} - \mathbf{au}$, where $\bar{\mathbf{p}}_{n' 1} = \bar{p} \times \mathbf{1}_{n' 1}$ with $\bar{p}$ the regular fixed retail rate charged, $\mathbf{a}_{n' n} = diag(a_1, a_2, \cdots a_n)$ is the discount matrix specifying discount parameters a_i for each group, and $\mathbf{u}_{n' 1}$ is the controller (the change of temperature set point). There is an upper bound $\mathbf{u}_0$ and a lower bound $-\mathbf{u}_0$ for the controller, specifying the limits of control for each group. n is the number of groups under control. Under cooling environment, LSEs need to raise the temperature set-point in order to reduce the load, i.e., they exercise control $\mathbf{u} > 0$, then the customers will enjoy a discounted rate $\mathbf{p(u;a)}$, otherwise, customers pay the regular rate $\bar{p}$. In this way, LSEs are able to differentiate customers by offering them different products to choose from. The table in Figure 3.1 gives an illustration of a pricing menu for two groups.

The customers observe the control limits $\mathbf{u}_0$ and the estimated savings $\mathbf{s}$ from each pricing scheme offered, and decide which contract to subscribe or not to subscribe depending on their personal preferences or utility functions $\mathbf{U(u, s)}$. Then given a design parameter set $\mathbf{(a, u)}$, there is a best response of customers' subscription $\mathbf{N = f(s; a, u)}$. With customers' subscription, LSEs derive an optimal control strategy for each group to minimize its control cost, which provides an estimated savings for each group $\mathbf{s = g(N; a, u)}$. Then we need to find an equilibrium such that $\mathbf{(N^*, s^*) = (f, g)(N^*, s^*; a, u)}$, meaning the customers and the LSEs will not deviate from this point for a given $\mathbf{(a, u)}$. Therefore, given parameters $\mathbf{(u_0, a)}$, a control strategy can be found under this partial equilibrium.

$\mathbf{(u_0, a)}$ are key contract parameters. LSEs can modify these parameters to get a different subscription from customers, and achieve their targets such as desired total costs, maximum tracking errors, etc. For example, higher discounts for all groups will increase the whole subscription, and as a result, will also increase the loads available to release. Higher discounts on a single group will increase the subscription of that group and change the whole structure of control groups.

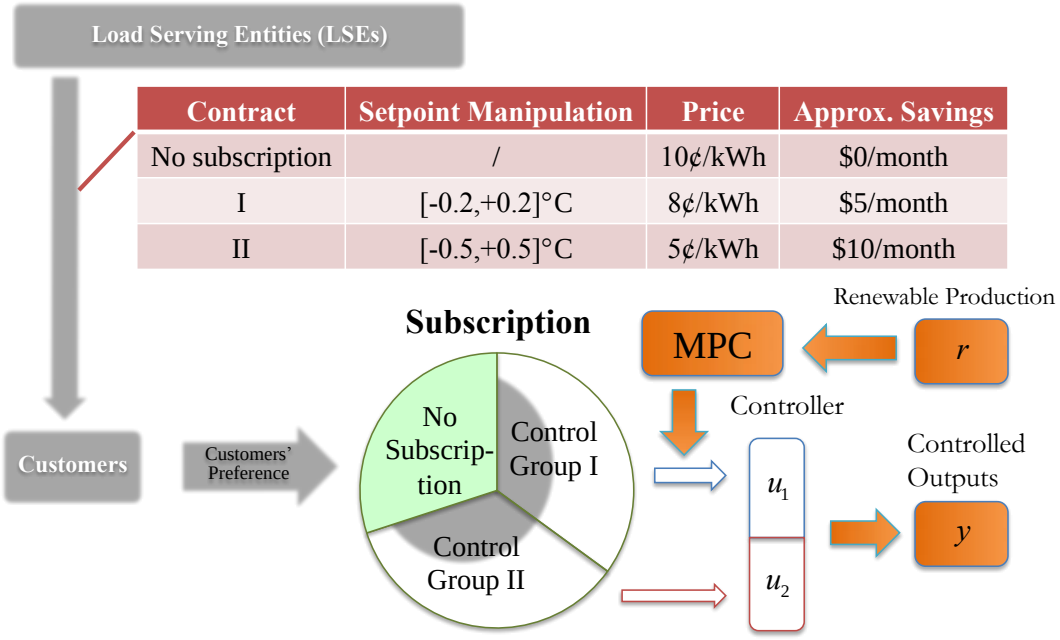


Figure 3.1 An Overview of the Model

3.2 Model Description

In this section, we describe the state-space form of modeling TCLs. If we were given design parameters $\mathbf{(u_0, a)}$ and customers' subscription **N**, how can we develop a control strategy? We formulate this as a linear quadratic tracking problem with inequality constraints and solve it by Model Predictive Control (MPC). We also study customers' preferences and investigate the interactive behavior between LSEs and customers. Finally, we propose an algorithm to find an equilibrium $\mathbf{(N^*, s^*)}$given $\mathbf{(u_0, a)}$.

3.2.1 State-Space Representation

Kundu et al. (2011) proposed a state-space model for homogeneous TCLs as follows:

$$\dot{x} = Ax + Bu$$

$$y = Cx + Du$$

where the model output y is the change of total power outputs from the steady-state value and the input u is the shift in thermostats set points. They derive an analytical treatment of homogeneous TCLs from the transfer function, which can be easily extended to handle n groups.

For simplicity, we consider 2 groups. Let N_1, N_2 respectively denote the number of customers sub- scribing to contract 1 and 2. We assume each group is homogeneous, and the dynamics of TCLs for each group follows the above model. The total power outputs: $y = y_1 + y_2$. The state-space model for TCLs with two inputs u_1, u_2 and one output y can be written as:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ y &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \end{aligned}$$

$$\text{where } \mathbf{x} = \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix}_{4\times 1}, \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_{2\times 1}, \mathbf{A} = \begin{bmatrix} \mathbf{A_1} & 0 \\ 0 & \mathbf{A_2} \end{bmatrix}_{4\times 4}, \mathbf{B} = \begin{bmatrix} \mathbf{B_1} & 0 \\ 0 & \mathbf{B_2} \end{bmatrix}_{4\times 2},$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C_1} & \mathbf{C_2} \end{bmatrix}_{1\times 4}, \mathbf{D} = \begin{bmatrix} d_1 & d_2 \end{bmatrix}_{1\times 2}.$$

$$\mathbf{A_i} = \begin{bmatrix} -2\sigma & -\omega \\ \frac{\sigma^2+\omega^2}{\omega} & 0 \end{bmatrix}, \mathbf{B_i} = \begin{bmatrix} \omega_i \Delta_i \\ 0 \end{bmatrix}, \mathbf{C_i} = \begin{bmatrix} -1 & 0 \end{bmatrix}. \text{ where}$$

$$\Delta_i = \frac{5\sqrt{15}C(\theta_a - \theta_+)(PR - \theta_a + \theta_+)}{\eta(P^2R^2 + 3PR(\theta_a - \theta_+) - 3(\theta_a - \theta_+)^2)^{3/2}} \cdot \frac{(3PR - \theta_a + \theta_+)N_i}{(T_{c0} + T_{h0})},$$

$$\omega = \frac{2\sqrt{15}C(\theta_a - \theta_+)(PR - \theta_a + \theta_+)}{CR\delta\sqrt{P^2R^2 + 3PR(\theta_a - \theta_+) - 3(\theta_a - \theta_+)^2}},$$

$$d_i = \frac{N_i}{\eta R}.$$

In the above equation, θ_a is ambient temperature, C is thermal capacitance, R is thermal resistance, P is TCL's rated power, θ_+ is temperature upper bound, d is dead-band width, η is electrical efficiency, s is damping coefficient, T_{c0} is total cooling time, and T_{h0} is total heating time.

3.2.2 Model Predictive Control for TCLs

Given design parameters **(a, u)** and customers' subscription **N**, we would like to minimize the following cost:

$$\begin{aligned} J(t,x) &= \min_{\mathbf{u}} \frac{1}{2}\int_0^T (y_t - r_t)^T Q(y_t - r_t) + \mathbf{u^T R u} dt \\ & \quad -\mathbf{u_0} \le \mathbf{u} \le \mathbf{u_0}, \end{aligned} \tag{3.1}$$

where r_t is a reference wind power output at time *t*, which we would like to follow as close as possible, $Q = 1$, $\mathbf{R} = \begin{bmatrix} r_w a_1 & 0 \\ 0 & r_w a_2 \end{bmatrix}$, and r_w is the weight of the importance of control cost.

In the traditional control theory, we need to know the whole trajectory of r_t in [0,T]. However, forecasting the wind power production long term ahead, say one day or even several hours, is quite hard and not accurate, because of the high variability of wind. An inaccurate forecast as a reference may cause the controlled outputs to deviate from the desired trajectory in the long run. We apply the Model Predictive Control (MPC) technique to handle the problem due to some of its nice features. MPC handles the problem by discretizing the continuous system and does on-line optimization across a time window with updated states and inputs at each step. At time *k*, it finds N_c number of controllers to minimize the cost across N_p steps. The reference signal during N_p steps is assumed to be constant at r_k. Thus, only the states estimates and wind power forecasts at the step, which are fairly accurate, are needed to perform the optimization. In addition, MPC is easily applicable to a LQ problem with inequality constraints.

We discrete the system:

$$\begin{aligned} \mathbf{x_{k+1}} &= \mathbf{A x_k} + \mathbf{B u_k} \\ y_k &= \mathbf{C x_k} + \mathbf{D u_k}, \end{aligned}$$

and the augmented system is:

$$\mathbf{Y}_{N_p \times 1} = \mathbf{F}_{N_p \times 4} \mathbf{x}_{k_{4 \times 1}} + \mathbf{G}_{N_p \times 2N_c} \mathbf{U}_{2N_c \times 1}.$$

At time *k*, we solve:

$$\begin{aligned} J_k &= \min_{\mathbf{U}} (\mathbf{S} - \mathbf{Y})^{\mathbf{T}} (\mathbf{S} - \mathbf{Y}) + \mathbf{U}^{\mathbf{T}} \bar{\mathbf{R}} \mathbf{U} \\ & \mathbf{M U} \leq \mathbf{m_0}, \end{aligned}$$

where

$$\mathbf{S} = r_k \mathbf{1}_{N_p \times 1}, \bar{\mathbf{R}} = \mathrm{diag}(a_1, a_2, a_1, a_2, \cdots, a_1, a_2) r_w, \mathbf{M} = \begin{bmatrix} I_{2N_c \times 2N_c} \\ -I_{2N_c \times 2N_c} \end{bmatrix}, \mathbf{m_0} = \mathbf{u_0} \mathbf{1}_{N_c \times 1}.$$

Without constraints, we have $\mathbf{U} = (\mathbf{G^T G} + \bar{\mathbf{R}})^{-1} \mathbf{G^T} (\mathbf{S} - \bar{\mathbf{F}} \mathbf{x_k}) = -\mathbf{K_x} \bar{\mathbf{x}}_\mathbf{k} + \mathbf{K_r r_k}$. Here, we

use Active Set method to solve the constrained quadratic programming.

3.2.3 Customers' Preference and Participation

Given design parameters **(a, u)**, the control law in the previous section gives the best response of the estimated savings **s** to the customers' subscription **N**. We define a utility function of customers so that we can find the best response of **N** to the estimated savings **s**.

Customers face a choice between $\mathbf{u}_0$ representing the discomfort level and **s** cash savings of utility bill. We use the following exponential utility function to model the customers' choice:

$$U(u_0, s) = 1 - \frac{1}{2}(e^{-\frac{\alpha}{u_0}} + e^{-\frac{s}{b}})$$

where α is a uniform random number describing customers' different levels of risk aversion, and b is the regular utility bill expenses. This is actually an equally weighted average of two exponential utility functions, one for the "good" $1/u_0$, and the other for the "good" s/b. Figure 3.2 plots a typical utility function with α = 1, s = 0. We find that as the limit of set-point adjustment u_0 increases, the utility U decreases slowly at the beginning (not much discomfort with small adjustment in set-point), and decreases fast above a certain level (very uncomfortable above this level). We also compute the marginal utility for s/b and $1/u_0$:

$$MU_{s/b} = \frac{1}{2b}e^{-\frac{s}{b}}, MU_{1/u_0} = -\frac{\alpha}{2u_0^2}e^{-\frac{\alpha}{u_0}}.$$

$MU_{s/b} > 0$ indicates that s/b is a "good", and has diminishing marginal utility as usual. $MU_{1/u_0} < 0$ indicates that u_0 is a "bad", and its marginal utility decreases first and then increases. This is due to the change of marginal utility becomes small when the temperature is adjusted above a certain high level. The marginal rate of substitution is calculated as well:

$$MRS = -\frac{MU_{1/u_0}}{MU_{s/b}} = \frac{b\alpha}{u_0^2}e^{-\frac{\alpha}{u_0}+\frac{s}{b}}.$$

It is decreasing when u_0 increases, meaning that with larger limit of set-point adjustment, you need to get larger rebate in return to maintain the utility. Based on the utility function and given the estimated savings **s**, customers choose the contract, which gives them the highest utility. The total number of customers subscribing to each contract **N** is the best response to the estimated savings **s**.

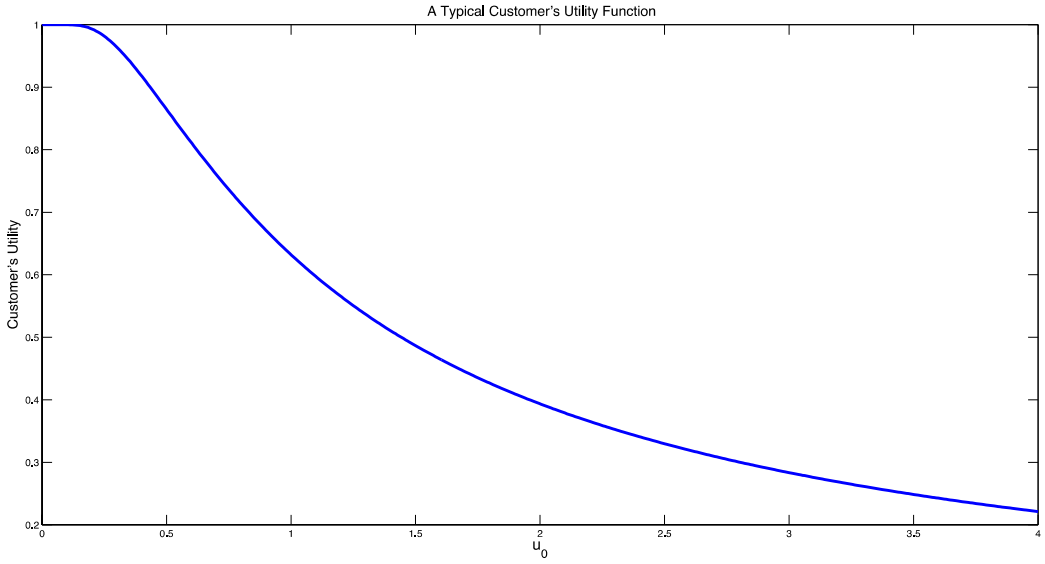


Figure 3.2 A Typical Utility Function

3.2.4 Equilibrium and Contract Design

Given design parameters $\mathbf{(a, u)}$, customers' subscription **N** best respond to the estimated saving **s** is determined by customers' utility function. Knowing customers' subscription **N**, LSEs find a control law by solving the control problem (3.1). Then the estimated savings s equals normal expenses - (controlled power output discount price):

$$\mathbf{s} = \int_0^T \mathbf{y_t p_t} dt.$$

Given an initial guess of the estimated savings $\mathbf{s_0}$, at each iteration *i*, we find $\mathbf{N}_i$ responding to $\mathbf{s}_{i-1}$, and find $\mathbf{s}_i$ responding to $\mathbf{N}_i$ until these numbers do not change. This gives an equilibrium point $\mathbf{(N^*, s^*)}$. Algorithm 1 describes the detail steps of finding such equilibrium.

Based on the above discussion, given design parameters $\mathbf{(a, u)}$, we can find customers' subscription **N** and an optimal control signal **u** from MPC to estimate savings **s**. We can adjust these design parameters $\mathbf{(a, u)}$ to achieve some desired targets, say maximum errors from wind power profile, or total control costs, etc.

Algorithm 1 Iteration Algorithm of Subscription Equilibrium

1. Choose intitial estimated savings $\mathbf{s_0}$.

2. Loop $i = 1$ to $NumIteration$. **Input:** $\mathbf{s_i}$.

3. $\mathbf{N} = (N_j)$, $j = 1, \cdots, n$.

4. Customer find j maxmizing

$$U(\mathbf{u_0}, \mathbf{s_i}) = 1 - \frac{1}{2}(e^{-\frac{\alpha}{\mathbf{u_0}}} + e^{-\frac{\mathbf{s_i}}{b}}).$$

5. N_j++.

6. Obtain best response $\mathbf{N}$ to $\mathbf{s}_i$.

7. Solve MPC to get $\mathbf{y_t}$.

8. End Loop. **Output:**

$$\mathbf{s_{i+1}} = \int_0^T \mathbf{y_t p_t} dt.$$

9. Loop ends when $\mathbf{N}$ do not change.

3.3 A Case Study

In this section, we perform a case study to illustrate how customers subscribe to given TCL contracts, and how LSEs design a control law to track a wind power output signal. At a certain location, there is a population of N = 10,000 potential customers receiving the contract offer. We assume a homogeneous group of customers with the following thermostat parameters in Table 3.1. The LSE offers two pricing menus to separate customers into two control groups. We use NREL 10 minutes forecasting wind power data on January 7, 2006 in Western U.S. as our reference signal. This grid point contains 10 wind turbines with capacity of 30MW. Time horizon of tracking is T = 24 (hours), and step size Δt = 1/6 (10 minutes), matching the wind power forecasts. The contract parameters are shown in Table 3.2. Customers who do not subscribe to any of these contracts pay a fixed regular rate at 10¢/kWh, while participators of group I receive 1.5¢/kWh rebate for every degree (◦C) adjustment and those of group II receive 2¢/kWh.

Table 3.1 Thermostat Parameters

Ambient temp.	Thermal capacitance	Thermal resistance	TCL's rated power
θ_a	C	R	P
32°C	10kWh/°C	2°C/kW	14kW
Dead-band width	Temp. set-point	Electrical efficiency	Damping coefficient
δ	θ_0	η	σ
1°C	21.5°C	2.5	0.002 hr^{-1}

Table 3.2 Contract Parameters

Regular rate	Marginal rebate I	Marginal rebate II	Temp. change limit I	Temp. change limit II
$\bar{p}$	a_1	a_2	u_{10}	u_{20}
10¢/kWh	1.5¢/kWh°C	2¢/kWh°C	0.5°C	1.5°C

Customers choose which contract to subscribe based on their own preferences. We set $a \in$ [1.2,6.7] to represent the different risk aversion in the utility functions, and average bill expenses b = 480¢ in the steady state. We have an initial guess of savings s_{01} = 15¢, s_{02} = 50¢, and do 15 iterations. In each iteration, we obtain a subscription N_1 and N_2, in which each customer maximizes the utility corresponding to the savings. Then we implement MPC to solve for a control signal and the estimated savings under this control for use in the next iteration. In MPC, prediction step N_p = 7, control step N_c = 7, and weight r_w = 105. Figure 3.3 shows that the numbers of customers subscribing to contract I and contract II converge to N_1 = 3063 (30.6%), and N_2 = 5648 (56.5%) respectively. And there are N_0 = 1289 (12.9%) customers who do not participate. With such information of subscription as an equilibrium point, we generate the optimal output trajectory and the control signal in Figure 3.4.

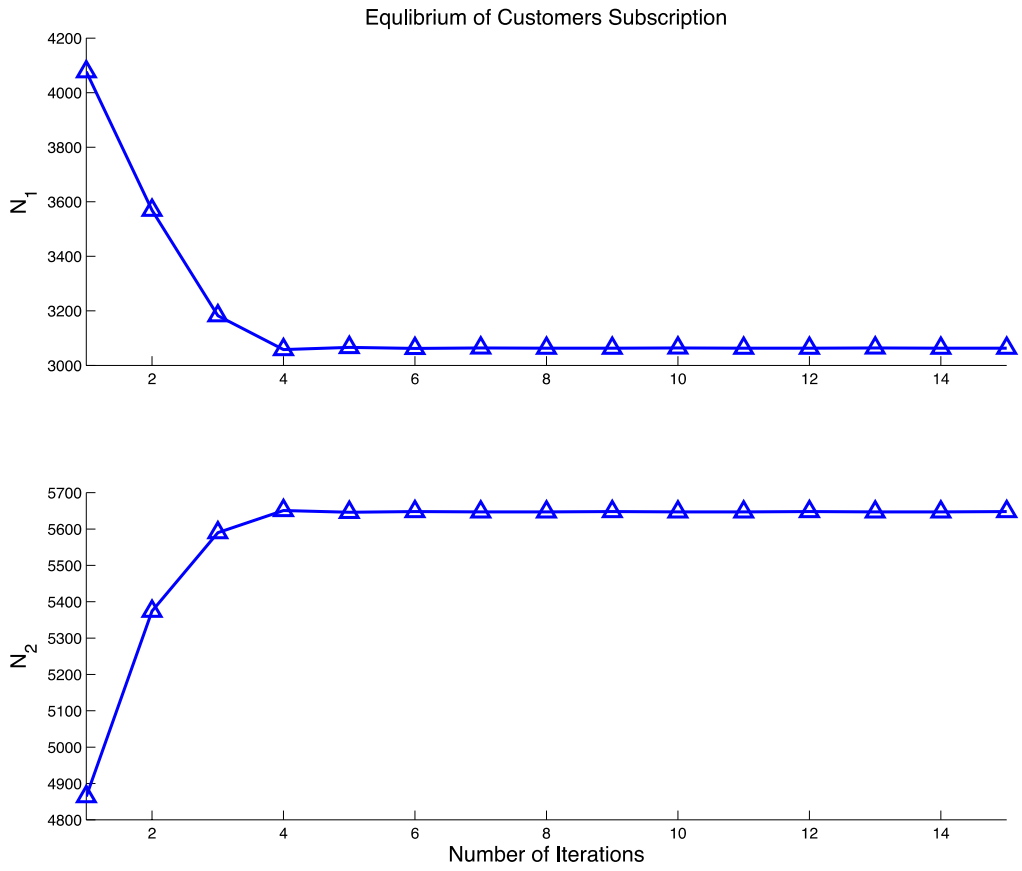


Figure 3.3 Convergence of Customers' Participation

In Figure 3.4, the first plot shows TCLs verses tracking wind power signal, the second plot shows the tracking errors, and the third one and forth one shows the control signal of adjusting the set-point of group I and II respectively. We find that large errors occur during wind shortage. For example, during 400th to 600th min, both groups almost reach their set-point change upper limits preventing them to lower the power consumption. When there is wind power production surplus, for example, during 800th to 1,000th min, temperatures are set lower to increase the demand to match the wind power profile. Figure 3.5 plots the power output for group I and II respectively under the control law. The output tends to offset each other to achieve the desired trajectory of wind power signal. We list several performance metrics in Table 3.3.

Table 3.3 Tracking Performance

Max Err. (MW)	Percent Err	Min NRMSE (MW)	Max NRMSE (MW)	s_1 (¢/day)	s_2 (¢/day)	Total Cost ($/day)
0.33	2.13%	0.0046	0.0173	11.23	58.00	3678.11

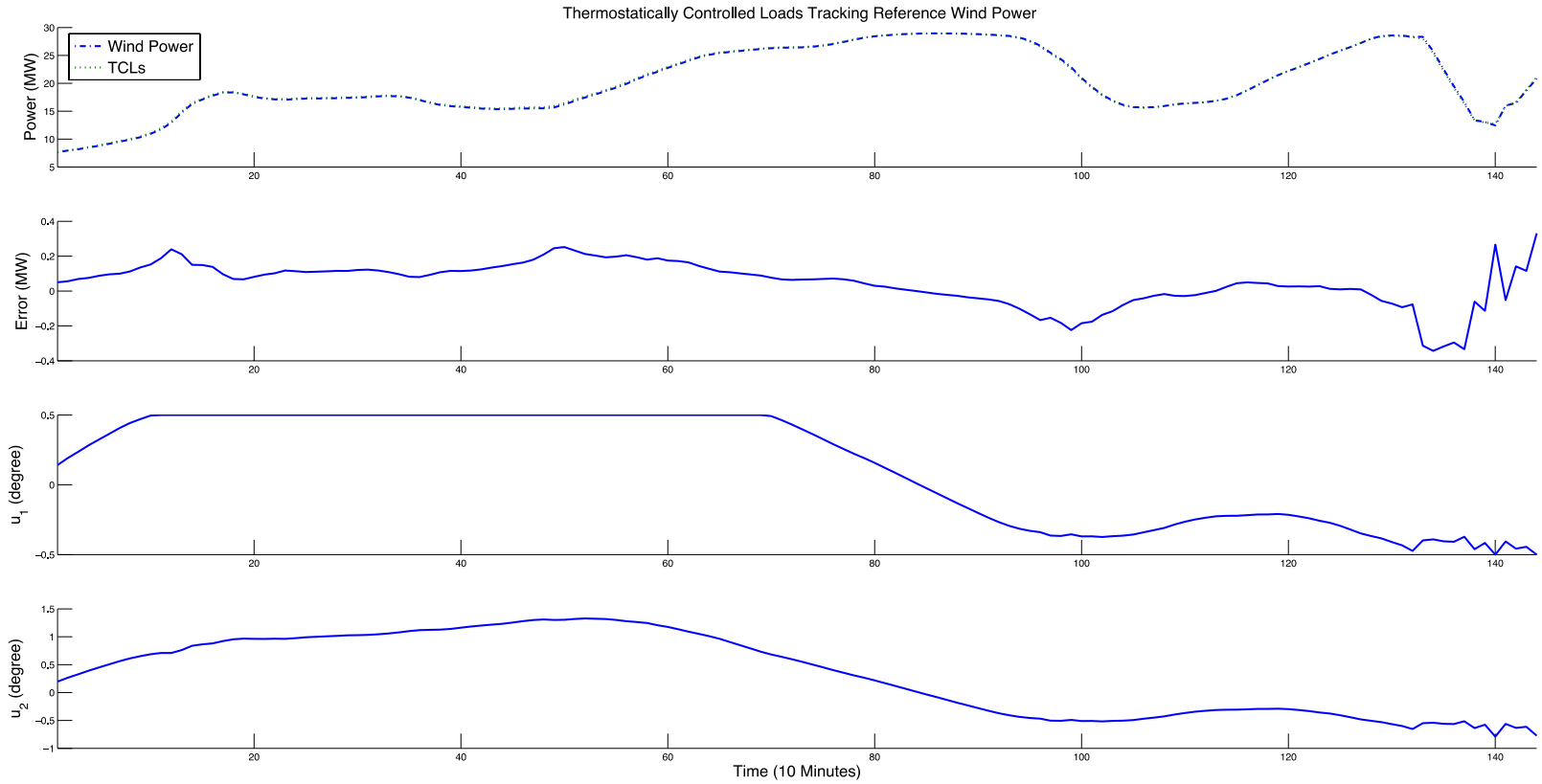


Figure 3.4 Controlled Load at Equilibrium vs. Wind Power

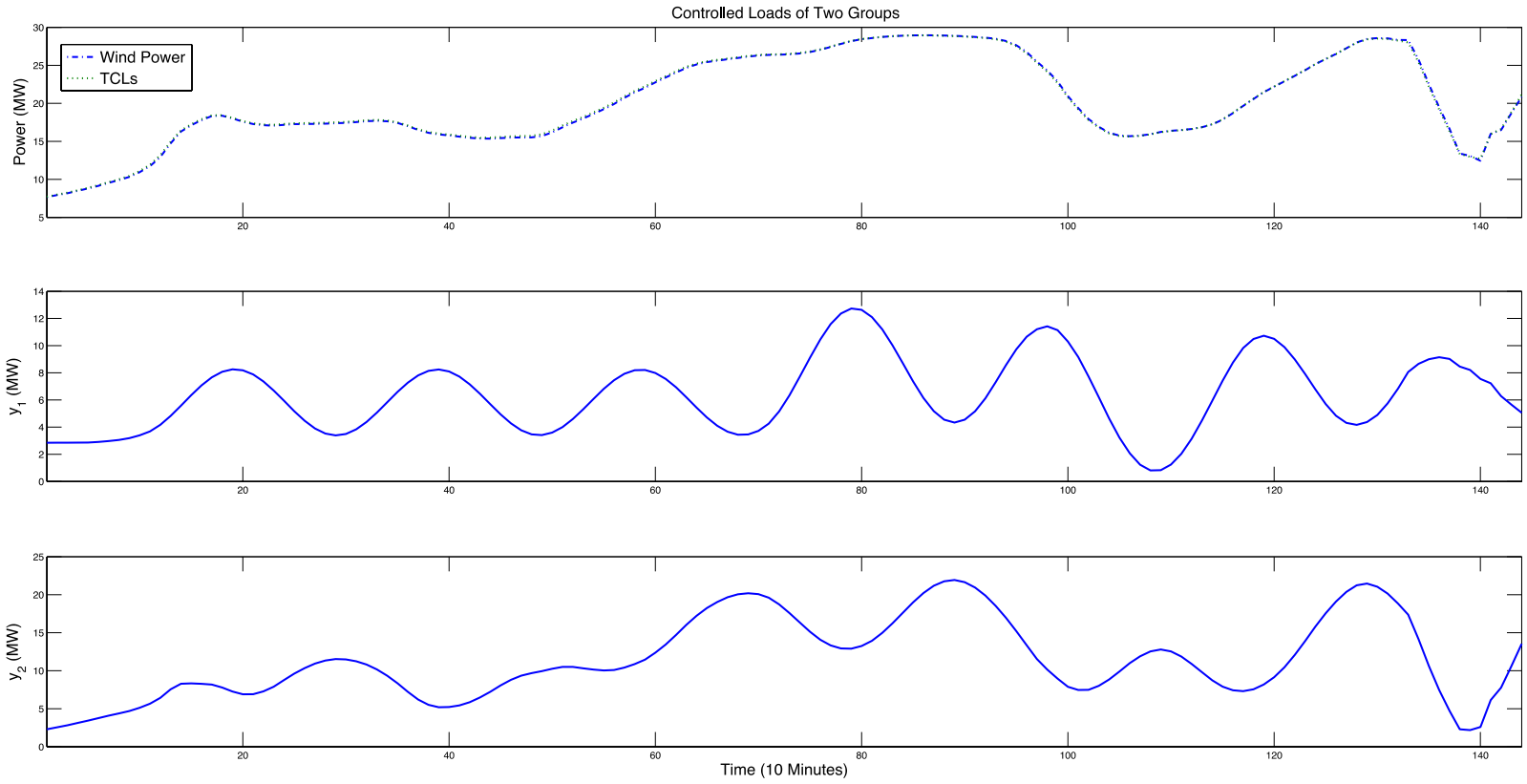


Figure 3.5 Controlled Loads of Two Groups

Finally, we investigate how design parameters (a, u_0) in the contracts affect the desired targets. We define two measurements, the total cost and the maximum error that the output deviates from the referred wind power. We set marginal rebate of two groups be the same, i.e., $a_1 = a_2 = a$. Fix set-point adjustment limit u_0, we vary a in [1.2,2] ¢/(kWh ^{o}C). Top chart in Figure 3.6 shows the customers' subscription versus the marginal rebates. We find that as marginal rebate a increases, the total number of participating customers increases (sum of group I and II). Customers who did not participate at low rebate level will switch to either group I or group II depending on their personal utility function to benefit from the increasing savings. Bottom plot in Figure 3.6 shows the maximum error and total cost versus the marginal rebates. As marginal rebate increases, the total costs increase definitely, and the maximum error decreases eventually.

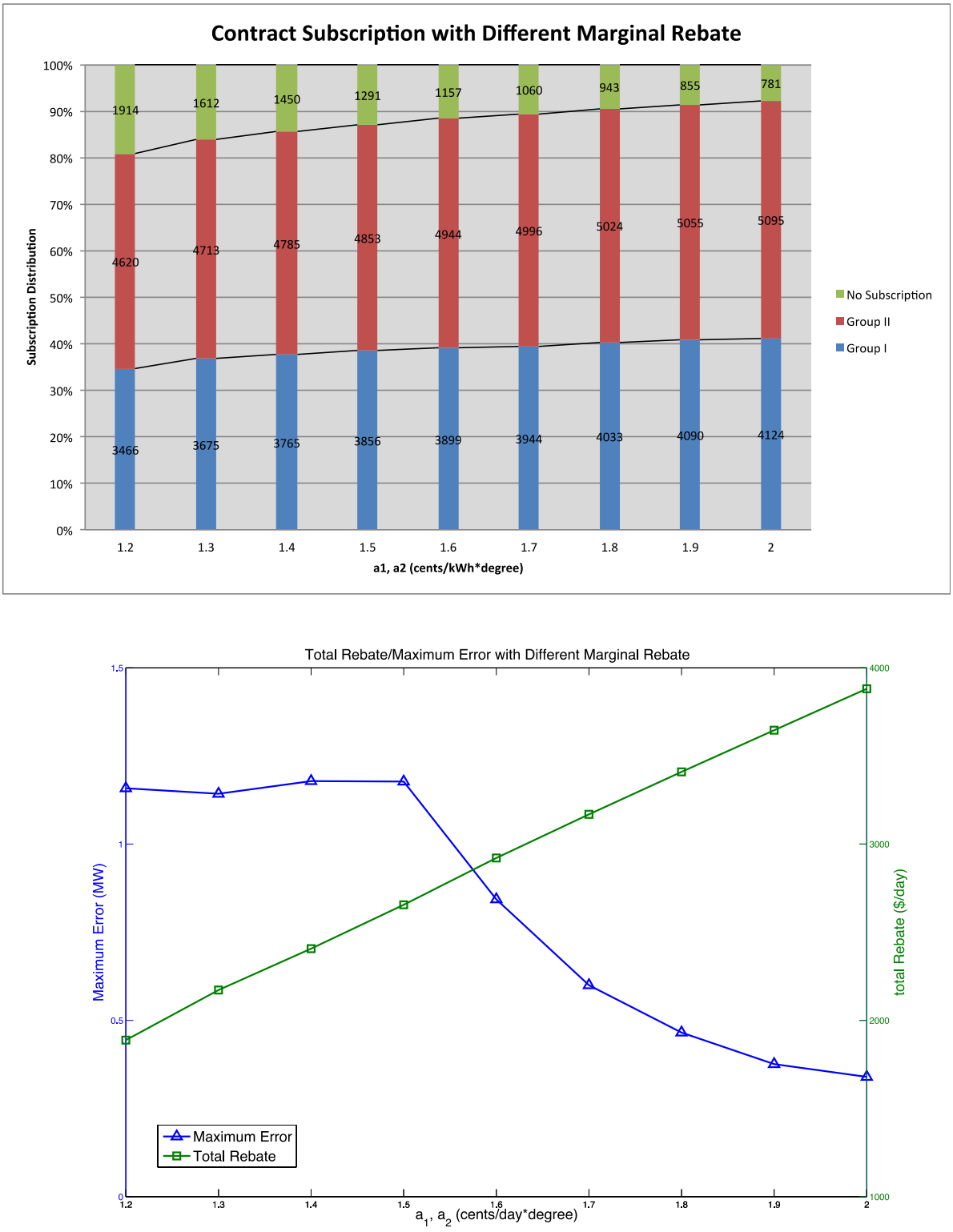
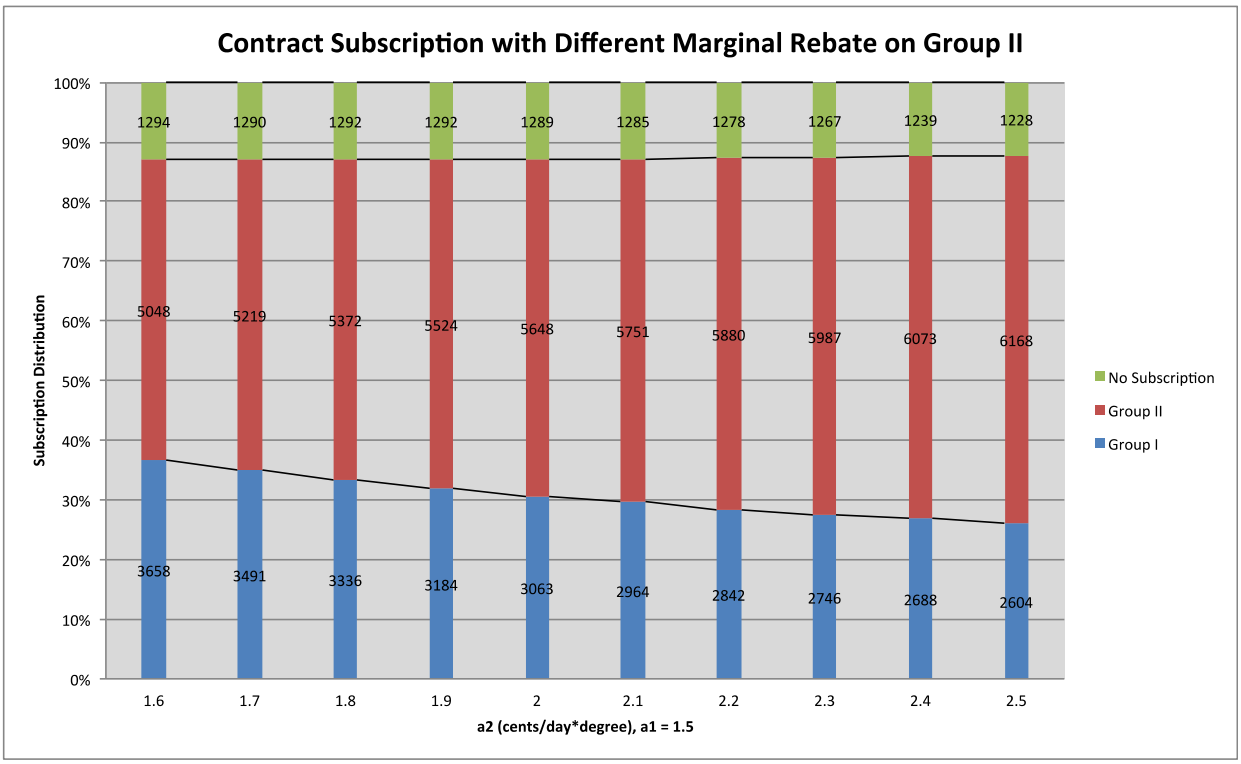


Figure 3.6 Customers' Participation, Maximum Error and Total Cost vs. Marginal Rebate

In order to attract more customers to contract II so that LSEs are more flexible in load control, we fix the marginal rebate of group I, a_1 at 1.5¢/(kWh ^{o}C), but vary a_2 - the marginal rebate of group II. With set-point adjustment limit u_0 fixed, we vary a_2 in [1.6,2.5] ¢/(kWh ^{o}C). Top chart in Figure 3.7 shows the customers subscription versus the marginal rebate of group II. We find that as marginal rebate of group II a_2 increases, customers switch from group I to group II, and those do not subscribe stay almost unchanged. Bottom plot in Figure 3.7 shows the maximum error and total cost versus the marginal rebate of group II. As the marginal rebate of group II increases, the total costs increase definitely, and the maximum error decreases. There is a tradeoff of LSEs to set the marginal rebate level of group II. Their decisions depend on the performance metric used.



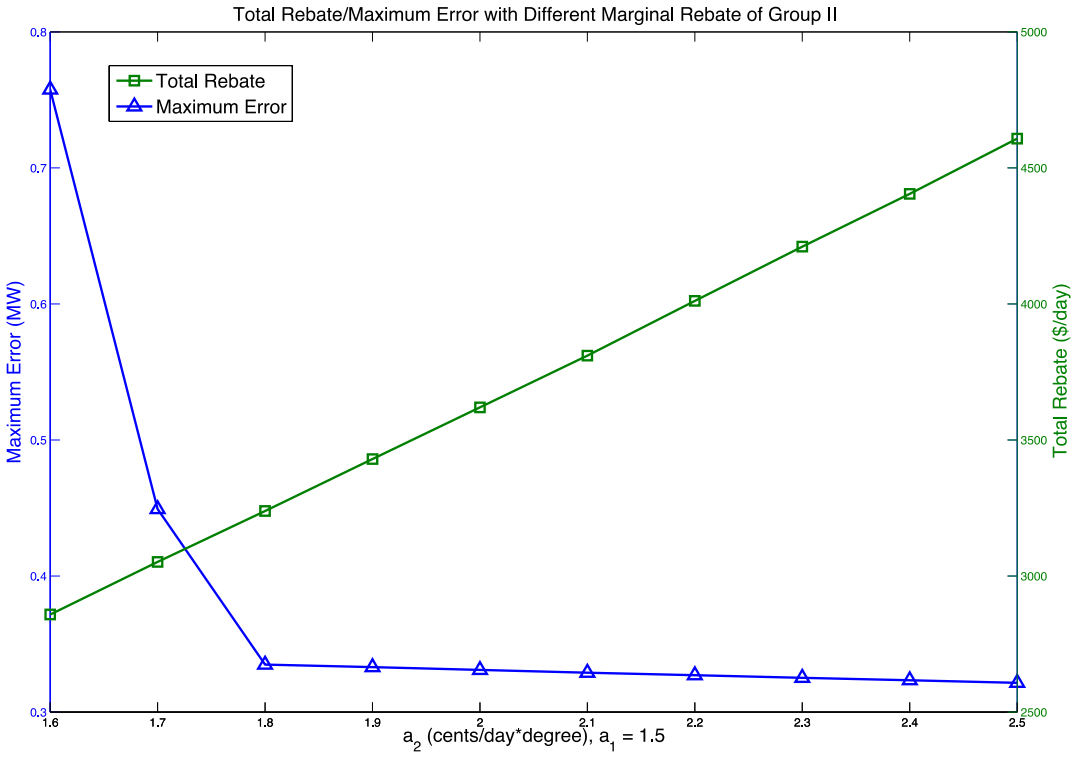


Figure 3.7 Customers' Participation, Maximum Error and Total Cost vs. Marginal Rebate of Group II

We also investigate how changes in set-point adjustment limits affect the tracking performance. Fix the marginal rebate at a_1 = 1.5¢/(kWh ^{o}C) and a_2 = 1.8¢/(kWh ^{o}C), we vary the set-point limits of group I, u_{10} in [0.4, 0.6] degree and group II, u_{20} in [1.5, 1.9] degree. Top plot in Figure 3.8 shows the maximum error and total cost versus limits of group I. We find that as adjustment limit u_{10} increases, the maximum error (total cost) decreases (increases) at the beginning but increases (decreases) afterward. This is it is possible to get more savings from group I, encouraging them to switch from contract II to contract I. Then less loads are available for control resulting large tracking errors. Bottom plot in Figure 3.8 shows the maximum error and total cost versus limits of group II. We find that as the adjustment limit u_{20} increases, the maximum error decreases but the total cost increases. LSEs should carefully choose contract parameters based on their own evaluation metrics.

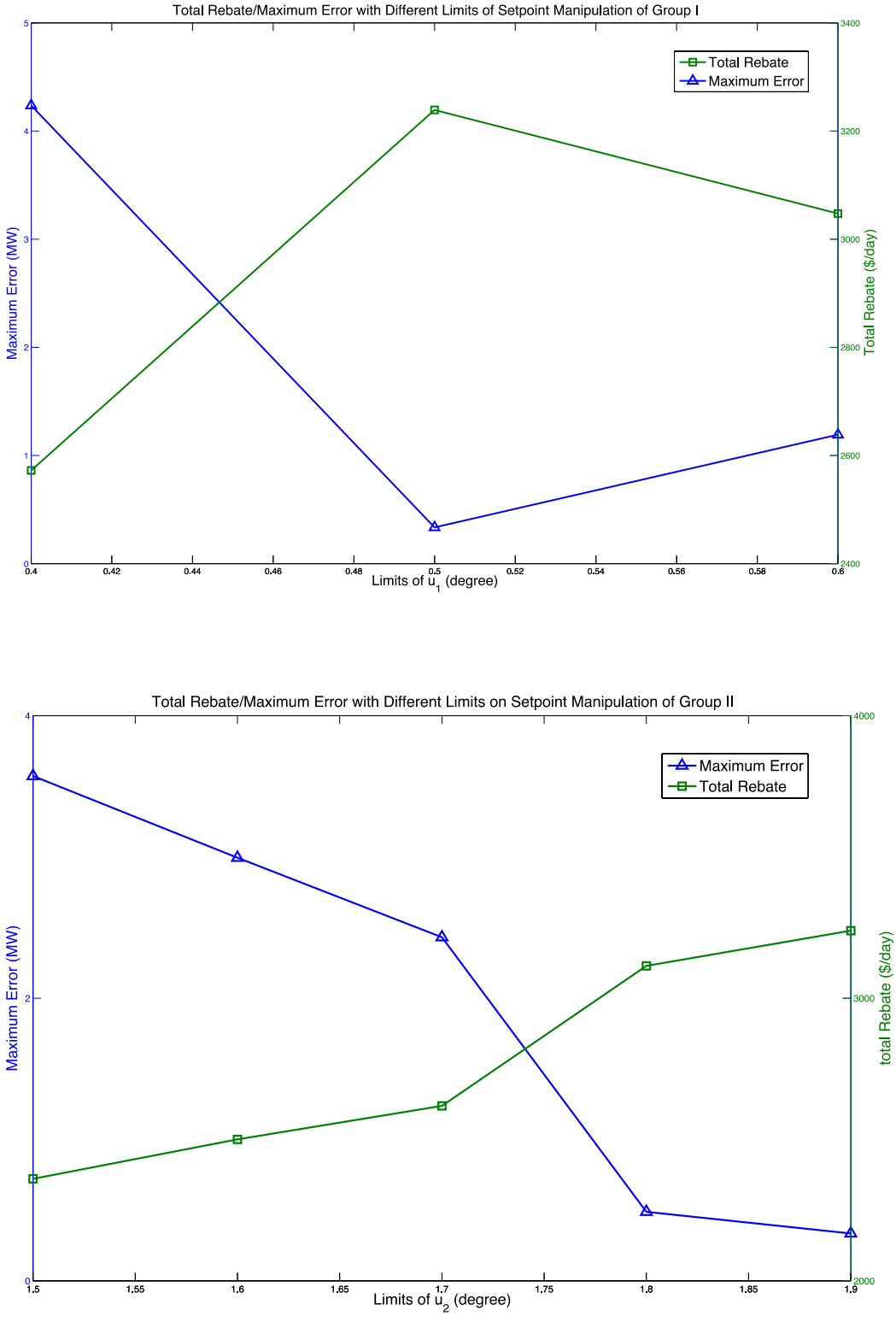


Figure 3.8 Maximum Error and Total Cost vs. Set-point Adjustment Limits

4. Conclusions

In this part, we propose two novel wind speed-forecasting approaches via wavelet transform to predict short-term wind speed, which serves as an input for our design of demand response mechanism. The first statistical forecasting model applies Multi-Resolution Analysis to the original wind speed time series by decomposing the series into different frequencies, which is more suitable for Holt-Winters modeling. The second wind speed forecasting model extracts the time-frequency information, namely MODWT coefficients, which turns out to be appropriate feature variables for BART model. Through boosting and MCMC algorithm, the prior sum of trees in BART will be trained into fully-grown decision trees, which reflect the expectation of future wind speed based on the complex and nonlinear interaction of feature variables. The performance of both models is acceptable. The first model has higher prediction accuracy while the second model has better extensibility.

We also propose an incentive-based demand response contract design for TCLs to absorb the intermittency in renewable energy production. LSEs design contracts, which offer incentives to end users to encourage their participation. The contract offers different levels of set-points adjustment limits and potential savings from rebates, so customers are separated into different control groups. Customers observe the contract offers, and decide which contract to subscribe or do not subscribe depending on their utility functions which measure a tradeoff between bearing discomfort and getting savings on bills. LSEs then observe the population subscription, and design an optimal control law such that the controlled load follows the renewable energy output. We apply MPC to solve for the optimal control signal due to the nice feature of MPC. It handles inequality constraints in the linear quadratic control, and only requires one-step further forecasting of wind power signal. We apply a game-theoretic model to address the interactive behavior between LSEs and end users in contracting the control of thermostats. Both LSEs and customers find their best responsive actions to the other party, and then we use iteration to find an equilibrium, which LSEs and customers do not want to deviate from.

This mechanism greatly increases the flexibility of LSEs to gain access to customers' thermostatically controlled loads. They change the control group structure by simply changing the contract parameters in order to get more loads released. In the case study, we show that TCLs are able to track a given wind power output closely with properly chosen parameters. LSEs can vary the contract parameters to achieve desired targets, e.g., minimizing the maximum tracking error or the total cost of implementing these contracts. We conclude that TCL contract design, as a type of demand response, is a very flexible and effective way to help LSEs mitigate the variability and intermittency of renewable energy.

References

[1] Callaway, D. S. (2009). Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. Energy Conversion & Management 50(9), 1389–1400.

[2] Deng, S.-J. and L. Xu (2009). Mean-risk efficient portfolio analysis of demand response and supply resources. Energy 34, 1523–1529.

[3] Ihara, S. and F. Schweppe (1981). Physically based modeling of cold load pickup. IEEE Trans Power App Syst 100 (9), 4142 – 4150.

[4] Koch, S., J. Mathieu, and D. Callaway (2011). Modeling and control of aggregated heterogeneous thermostatically controlled loads for ancillary services. In Proceedings of 17th Power Systems Computation Conference, Stockholm, Sweden.

[5] Kundu, S., N. Sinitsyn, S. Backhaus, and I. Hiskens (2011, Jan). Modeling and control of thermostatically controlled loads. arXiv:1101.2157v1.

[6] Lu, N., D. Chassin, and S. Widergren (2005). Modeling uncertainties in aggregated thermostatically controlled loads using a state queuing model. IEEE Transactions on Power Systems 20 (2), 725–733.

[7] Malhame, R. and C. Chong (1985). Electric-load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system. IEEE Trans Automat Contr, 30, 854–860.

[8] Papavasiliou, A. and S. S. Oren (2011). Large-scale integration of deferrable demand and renewable energy sources. submitted to IEEE Transactions on Power Systems Special section on Electricity Market Operations.

[9] Ucak, C. and G. Dokuyucu (2010). Investigation of thermostat-set control as a new direct load control method. Department of Electrical Engineering, Electrical & Electronics Faculty, Istanbul Technical University.

[10] Chipman, HA, George, EI, and McCulloch, RE (2010). BART: Bayesian Additive Regression Trees. Annals of Applied Statistics, 4, 266-298.

[11] Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., & Draxl, C. (2011). The State-Of-The-Art in Short-Term Prediction of Wind Power: A Literature Overview, 2nd edition, ANEMOS.plus.

[12] Landberg, L., Giebel, G., Nielsen, H. A., Nielsen, T. and Madsen, H. (2003). Short-term Prediction—An Overview. Wind Energy, 6: 273–280. doi: 10.1002/we.96.

[13] Pinson, P., Nielsen, H. Aa., Møller, J. K., Madsen, H. and Kariniotakis, G. N. (2007), Non-parametric probabilistic forecasts of wind power: required properties

and evaluation. Wind Energ., 10: 497–516. doi: 10.1002/we.230.

[14] J.P.S. Catalão, H.M.I. Pousinho, V.M.F. Mendes (2011). Short-term wind power forecasting in Portugal by neural networks and wavelet transform. Renewable Energy, Volume 36, Issue 4, April 2011, Pages 1245-1251, ISSN 0960-1481, 10.1016/j.renene.

[15] M.C Alexiadis, P.S Dokopoulos, H.S Sahsamanoglou, I.M Manousaridis (1998). Short-term forecasting of wind speed and related electrical power, Solar Energy, Volume 63, Issue 1, July 1998, Pages 61-68, ISSN 0038-092X, 10.1016/S0038-092X(98)00032-2.

[16] Donald B. Percival, Andrew T. Walden (2006). Wavelet Methods for Time Series Analysis. Volume 4 of Cambridge Series In Statistical And Probabilistic Mathematics, Cambridge University Press, 2006.

Part II

Simulation of Impacts of Demand Response Resources on Power System Variable Effects

George Gross
Anupama Kowli
University of Illinois at Urbana-Champaign

For information about Part II, contact:

George Gross
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
1406 W. Green Street
Urbana, IL 61801
Email: gross@illinois.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Table of Contents

List of Figures

List of Tables

1. Introduction

The push towards sustainability, increasing electricity prices, technological advances and policy initiatives at the federal and state level drive the efforts to harness active participation of the consumers in the North American electricity markets [1]. Such efforts have resulted in the implementation several demand-side activities which, in turn, have created many new players in the electricity industry. In particular, there is a new class of consumers, called *demand response resources* (*DRR*s), whose role has become increasingly important in ensuring that the supply-demand balance is efficiently attained [2], [3]. In addition to purchasing electricity in the markets, the *DRR*s can sell load curtailment services in those markets by reducing their loads during certain hours. For these hours, the *DRR*s compete to provide the load curtailment services directly against the supply-side resources that sell their generation outputs and so provide the *IGO* with additional degrees of freedom in maintaining economic and reliable power system operations. The *time-dependent DRR* deployments impact the loading on the system, thereby affect the generation and transmission resource utilization, as well as the market outcomes.

The potential load reduction from the *DRR*s in the year 2008 was estimated to be close to 41 *GW*, which represented approximately 5.8 *%* of 2008 U.S. peak load [4]. This capacity is significant in terms of preventing the need for the use of some existing and the possible construction of new peaking units as well as deferring the need for new transmission capacity [5], [6]. The implementation of the smart grid with a deeper penetration of renewable supply resources is expected to entail a broader deployment of *DRR*s [7]. Indeed, the projections of the peak load reductions potential in 2019 range between 38 GW and 188 GW across the country, which roughly constitutes about 4 to 20 *%* of the nation-wide peak demand [8]. As the penetration of *DRR*s deepens, their impacts on market outcomes and system operations become more pronounced and propagate to generation and transmission planning and, eventually, entail regulatory/legislative initiatives. Consequently, there is a need for appropriate tools for quantifying the impacts of *DRR*s on market performance, generation dispatch, transmission usage, emissions and other system variable effects. A particularly critical requirement is that of a simulation tool to emulate the behavior of a power system with integrated *DRR*s; we address this requirement. We present, in this report, the development of a simulation approach that provides the basis of such a tool. We illustrate the application of the proposed approach to various planning and analysis studies.

While the impacts of *DRR*s on the system variable effects over short- and medium-term periods have been studied to some extent [9]–[15], our focus is on the longer-term study periods. We are interested in assessing the impacts of *DRR*s on the variable effects that are, typically, evaluated in planning and policy analysis studies. Consequently, explicit representation of the various sources of uncertainty which impact system and market operations along with the representation, with the appropriate level of detail, of the supply- and demand-side resources, the transmission grid, the market clearing operations,

the market structure and operating policies, is the key requirement of the simulation approach.

We develop a systematic, computationally efficient approach that allows the representation of the time-dependent nature of the *DRR* deployments and the transmission-constrained market clearing operations, as well as that of the uncertainty inherent in power systems and the policies in effect. Our approach is designed to be applicable to longer-term analysis and to capture the uncertainty in future developments. The time-dependent nature of *DRR*s leads to the use of the snapshot-based market performance analysis as the basic building block of the proposed approach. Given the longer-term nature of the studies, we make use of probabilistic simulation concepts and the associated uniformity assumptions in the construction of the approach. In this way, we capture the effects of uncertainty on the system and market operations. In the implementation of the proposed approach, we pay careful attention to reducing the computation burden. For this purpose, we devise effective means to specify the simulation periods and also the deployment of the *Latin hypercube sampling* scheme to provide an efficient way to approximate the probability distributions of the market outcomes. Thus, we bring about computational tractability in the quantification of the system variable effects; this is particularly significant for applications to large-scale systems over longer-term periods.

The proposed approach provides, for the first time, a useful mechanism to assess the impacts of effective *DRR* utilization on resource investment decisions, transmission planning and system reliability. The approach has a wide range of applications – from justifying investments in *DRR*s to investigating dynamic pricing schemes, from formulating policies for the more widespread use of demand-side resources to devising effective strategies for their utilization and from study of alternative market designs to quantifying environmental benefits of the smart grid implementations. The proposed approach is useful for *IGO*s, *ESP*s, generation and transmission asset investors, regulators and policy makers to make better informed decisions.

The report contains five additional sections. We describe in section 2 the modeling details and in section 3 the proposed approach. In section 4, we discuss the implementation steps to bring about computational tractability in the practical applications of our approach. We report representative application studies in section 5 in which we demonstrate the capability of the proposed approach to quantitatively assess the range of *DRR* impacts on a large test system. We conclude with a summary and directions for future work in section 6.

2. Models and Metrics

We devote this section to the description of the models and metrics used in our approach. At the outset, we state the assumptions introduce into our modeling. We assume that throughout each simulation period the resource mix, the transmission grid, the market structure, the operating policies and the seasonality effects remain unchanged so that the loads and the resources exhibit uniform characteristics. We assume that a forecast of the

aggregate system load, consisting of both fixed and price-sensitive load components, is specified for the simulation period and that each component is known. To simplify the discussion, we consider the supply system to consist of only controllable, i.e., dispatchable, units. We assume that the impacts of the unit commitment decisions over the simulation period are incorporated into the specified loading list of the supply resources (*SR*s). We also assume that the maximum load curtailment capacities offered by the *DRR*s and their associated load recovery impacts are *a priori* specified for a typical week day (weekend day) and that such diurnal week day (weekend day) *DRR* behavior is assumed valid for each week day (weekend day) of the simulation period. Also, we assume that there is no uncertainty in the availability of the *DRR*s.

We consider the central pool market structure [16] that is widely adopted in many jurisdictions. For the purposes of this study, we assume the behavior of each market participant is unaffected by that of the other participants and we ignore any strategic behavior on the part of either the buyers or the sellers. For market clearing purposes, we consider a lossless transmission network and assume that the *DC* power flow conditions hold at all times [17]. In the event of insufficient supply due to resource inadequacy or forced outages of generators and/or transmission congestion, the *IGO* uses a regulatorily specified price cap λ^{max} to clear the market [16]. Whenever such an event occurs, the electricity is priced at λ^{max}.

We consider the day-ahead electricity markets (*DAM*s) and explicitly represent their impacts. We find it convenient, therefore, to adopt an hour as the smallest indecomposable unit of time and we view the system to be in steady-state in each hour in the simulation period. The resolution chosen does not allow the representation of any phenomenon of duration shorter than an hour and, consequently, such phenomena are entirely ignored in the simulation. We denote by H the index set of the hours $\{1, 2, \ldots, H\}$ in the simulation period, where H is the number of hours. We denote by $\mathrm{H}^{c} \subset \mathrm{H}$ the subset of hours in which *DRR*s are allowed to submit load curtailment offers. And, we denote by $\mathrm{H}^{r} \subset \mathrm{H}$, the subset of hours in which *DRR*s recover the load curtailments are collected in the set H^{c}. We impose the restriction $\mathrm{H}^{r} \bigcap \mathrm{H}^{c} = \emptyset$ so that the beneficial effects of load curtailments are not attenuated due to load recovery actions [9], [10].

We review the *IGO*'s market clearing problem for a specified hour $h \in \mathrm{H}$ such that $h \notin \mathrm{H}^{c}$ and $h \notin \mathrm{H}^{r}$. We use $\mathrm{S}(\mathrm{B})$ to denote the collection of supply-side sellers (demand-side buyers). Each seller (buyer) submits its price and quantity offer (bid), indicating the willingness to sell to (buy from) the *IGO* in the hour h. The *IGO* uses this information to clear the hour h market. The market clearing depends on the physical network [18] and, therefore, we require a representation of the transmission grid for the assessment of market performance.

We consider the transmission grid to consist of $(N+1)$ buses with J transmission lines. We use the set $\mathrm{N} = \{0, 1, \ldots, N\}$ to denote the index set of the buses in the network, with

the bus *0* denoting the slack bus. We denote by $J = \{1, 2, \ldots, J\}$ the index set of the lines and the transformers which connect the buses of the network. We associate with each line/transformer $j \in J$ the real power flow limit f_j^{max}. We denote by $\underline{\underline{A}}$ the J x N reduced branch-to-node incidence matrix, by $\underline{\underline{B}}_d$ the J x J diagonal branch susceptance matrix, by $\underline{\underline{B}}$ the N x N reduced nodal susceptance matrix and by $\underline{b}_0$ the N x 1 column vector of the augmented susceptance matrix corresponding to the slack node. We use $\underline{\underline{A}}$, $\underline{\underline{B}}_d$, $\underline{\underline{B}}$ and $\underline{b}_0$ in the expressions for the *DC* power flow equations and transmission constraints [18] to describe the key characteristics of the transmission system for the hour h system snapshot.

The market clearing explicitly considers the feasibility of the transactions cleared in the market with respect to the transmission constraints and, determines at each node $n \in \mathrm{N}$, the real power supply (demand) $p_n^g\left(p_n^d\right)$. We define the vectors $\underline{p}^g \equiv \left[p_1^g, p_2^g, \ldots, p_N^g\right]^T$ and $\underline{p}^d \equiv \left[p_1^d, p_2^d, \ldots, p_N^d\right]^T$. The *IGO* market clearing for hour h entails the solution of an *OPF* with the objective to maximize the auction surplus[1] S. We use $\beta_n^b(\cdot)$ to represent the aggregated benefits of the buyers at node *n* and $\gamma_n^s(\cdot)$ to represent the aggregated costs incurred by the *IGO* for the offers of the *SR*s at node n. Then, we state the *OPF* as

$$\max \ \mathrm{S} = \sum_{n \in \mathrm{N}} \left\{\beta_n^b\left(p_n^d\right) - \gamma_n^s\left(p_n^g\right)\right\} \qquad (1a)$$

$$\textit{subject to} \quad p_0^g - p_0^d = \underline{b}_0^T \underline{\theta} \ \leftrightarrow \lambda_0 \qquad (1b)$$

$$\underline{p}^g - \underline{p}^d = \underline{\underline{B}}\,\underline{\theta} \ \leftrightarrow \underline{\lambda} \qquad (1c)$$

$$\underline{\underline{B}}_d \underline{\underline{A}}\,\underline{\theta} \le \underline{f^{\max}} \ \leftrightarrow \underline{\rho}\,, \qquad (1d)$$

with the hour h notation suppressed. We use $\underline{\theta} \equiv \left[\theta_1, \theta_2, \ldots, \theta_N\right]^T$ and $f^{\max} \equiv \left[f_1^{\max}, f_2^{\max}, \ldots, f_J^{\max}\right]^T$ to denote the vectors of the bus voltage phase angles and the real power line flow limits, respectively. We use λ_0, $\underline{\lambda}$ and $\underline{\rho}$ to denote the dual variables associated with the constraints in (1b)-(1d). The constraints and the variables in (1) correspond to the hour h system snapshot.

We denote the *OPF* problem in (1) by $\mathrm{M}(\mathrm{S},\mathrm{B})|_h$. We consider both fixed as well as price-sensitive demand bids in this *OPF* problem which is solved to clear the hour h market. The fixed demand bid, where the buyer exhibits an unlimited willingness to pay for electricity, can be viewed as a special case of the price-sensitive demand bids for which a specified quantity is submitted without price information.

[1] We use the terminology auction surplus instead of social welfare because the offer (bid) data for the sellers (buyers) need not necessarily reflect the true costs (benefits) of the players.

The market performance for the hour h is quantified from the market clearing given by the solution $\mathrm{M}(\mathrm{S},\mathrm{B})|_h$.When $\mathrm{M}(\mathrm{S},\mathrm{B})|_h$ is feasible, the optimal values of its decision and dual variables are used to determine the market outcomes. We use $\left[p_n^g\right]^*|_h$ and $\left[p_n^d\right]^*|_h$ to denote the total supply and demand, respectively, cleared at the node *n*. The total load cleared in hour h is computed as

$$\ell^{\mathrm{S}}|_h = \sum_{n\in\mathrm{N}} \left[p_n^d\right]^*\Big|_h , \tag{2}$$

which clearly equals the total generation of the supply-side resources. The optimum values $\left[\lambda_n\right]^*\Big|_h$ of the dual variables associated with the nodal power balance constraints provide the locational marginal prices (*LMP*s) at each node $n\in\mathrm{N}$. Each *MWh* is sold (bought) at node *n* at the price $\left[\lambda_n\right]^*\Big|_h$. Then, the total hourly supply-side payments are

$$w^{\mathrm{S}}\Big|_h = \sum_{n\in\mathrm{N}} \left[\lambda_n\right]^*\Big|_h \cdot \left[p_n^g\right]^*\Big|_h , \tag{3}$$

and the total hourly demand-side payments are

$$w^{\mathrm{B}}\Big|_h = \sum_{n\in\mathrm{N}} \left[\lambda_n\right]^*\Big|_h \cdot \left[p_n^d\right]^*\Big|_h . \tag{4}$$

The hourly congestion rents are given by the difference between $w^{\mathrm{B}}\Big|_h$ and $w^{\mathrm{S}}\Big|_h$:

$$\kappa\Big|_h = \sum_{n\in\mathrm{N}} \left[\lambda_n\right]^*\Big|_h \cdot \left(\left[p_n^d\right]^*\Big|_h - \left[p_n^g\right]^*\Big|_h \right). \tag{5}$$

In the event that $\mathrm{M}(\mathrm{S},\mathrm{B})|_h$has no feasible solution because the fixed demand requirements cannot be satisfied due to shortfall in total generation capacity, some fraction of the fixed load at a subset of the nodes cannot be supplied, leading to a *loss of load* event. The *IGO* determines the unserved load $u_n\Big|_h$ at each node *n* in the subset. Then, the total unserved demand due to loss of load in hour h is given by

$$\mathscr{U}\Big|_h = \sum_{n\in\mathrm{N}} u_n\Big|_h . \tag{6}$$

Clearly, whenever $\mathrm{M}(\mathrm{S},\mathrm{B})\Big|_h$ is feasible, $\mathrm{U}\Big|_h = 0$. When $\mathscr{U}\Big|_h > 0$, the *LMP*(s) at the node(s) with $u_n\Big|_h > 0$ are set to the price cap λ^{max} and, then, the payments and congestion rents are evaluated. The metrics in (2)-(6) provide the measures of the market performance for an arbitrary hour h without *DRR* load curtailments or load recovery.

We need to make certain modifications to the market clearing problem in (1) to incorporate curtailment and recovery effects. We incorporate the impacts of *DRR*s as market players in each hour $h^c \in \mathrm{H}^c$ by extending the *DAM* representation and solving the *IGO*'s modified market clearing problem for the curtailment hours. By definition, a *DRR* is a buyer $b \in \mathrm{B}$ whose load curtailment cannot exceed his load in any hour $h^c \in \mathrm{H}^c$. We partition B into two non-overlapping subsets – the collection of buyers with demand response capability $\hat{\mathrm{B}}$, and its complement $\bar{\mathrm{B}}$, the collection of *pure* buyers that provide no load curtailment. Each *DRR* $\hat{b} \in \hat{\mathrm{B}}$ is both a buyer purchasing electricity and a seller selling load curtailment. We represent these actions of *DRR* $\hat{b} \in \hat{\mathrm{B}}$ in terms of its demand bid and its curtailment offer. The curtailment offer of *DRR* $\hat{b}$ consists of the maximum load it can curtail and the price at which the curtailment is provided. Whenever the hour h^c curtailment offer of a *DRR* is accepted, he becomes a *net* buyer and purchases only the net *MWh* for the remaining load above the load curtailed in that hour. Otherwise, he acts as pure buyer and purchases the *MWh* cleared in the market for its load without any curtailment.

The hour h^c market clearing determines the *DRR* load curtailment $\hat{p}^c_n$ at each node $n \in \mathrm{N}$ and the aggregated costs incurred $\gamma^{\hat{b}}_n(\cdot)$ by the *IGO* for the accepted *DRR* offers at node *n*. Note that the term $\left(p^d_n - \hat{p}^c_n\right)$ represents the net load cleared at that node. So, we use the vector $\hat{p}^c \equiv \left[\hat{p}^c_1, \hat{p}^c_2, \ldots, \hat{p}^c_N\right]^T$ and the term $\hat{p}^c_0$ in the equations (1b) and (1c) that represent the nodal power balance constraints and the cost term $\gamma^{\hat{b}}_n(\cdot)$ in (1a) to state the market clearing problem for the hour $h^c \in H^c$, again with the hour h^c notation suppressed. We denote the hour h^c *OPF* by $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\Big|_{h^c}$, where we use the fact that $\mathrm{B} = \hat{\mathrm{B}} \cup \bar{\mathrm{B}}$ explicitly represents the *DRR* players. We use the solution of $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\Big|_{h^c}$ to find the market clearing quantities for the hour h^c. Although each *DRR* provides a load curtailment service, it is compensated on a *$/MWh* basis as if it provided the energy saved by its curtailment. The *IGO* recovers the compensation payments for the *DRR*s by collecting an additional curtailment service charge υ *$/MWh* from all the buyers whose demand bids get cleared in the hour h^c market. Therefore, a buyer at the node *n* pays $\left([\lambda_n]^*\big|_{h^c} + v\Big|_{h^c}\right)$ *$/MWh* for his electricity purchases in the hour h^c. The value of $v\Big|_{h^c}$ depends on the specific compensation scheme used. Assuming that the *DRR*s are compensated at the corresponding *LMP*s, we compute the total compensation $w^{\hat{\mathrm{B}}}\Big|_{h^c}$ paid to the *DRR*s whose bids are accepted in the hour h^c as

$$w^{\hat{\mathrm{B}}}\big|_{h^c} = \sum_{n \in N} [\lambda]^*_n\big|_{h^c} \cdot \left[\hat{p}^c_n\right]^*\Big|_{h^c}, \qquad (7)$$

so that the "uplift" for the *DRR* curtailments incurred by each buyer is

$$\nu\big|_{h^c} = \frac{w^{\hat{\mathrm{B}}}\big|_{h^c}}{\ell^{\mathrm{S}}\big|_{h^c}} \quad . \tag{8}$$

We use the nodal market clearing quantities and the uplift charge to evaluate the metrics (2)-(6) for each hour $h^c \in \mathrm{H}^c$.

Since the load curtailment by the *DRR*s whose offers are accepted in the hour $h^c \in \mathrm{H}^c$ may induce a deferred load in some subsequent hours in which the load recovery occurs [9], [10], the simulation needs to appropriately represent the *load recovery effects*. We use curtailment recovery factor (*CRF*) $\mathrm{X}^{\hat{b}}_{h^c,h^r}$ to specify the fraction of *DRR* $\hat{b}$'s load curtailment in hour $h^c \in \mathrm{H}^c$ that is recovered in the hour $h^r \in \mathrm{H}^r$. The *CRF* is used to compute the recovery load p^r_n at each affected node $n \in N$ [4]. Clearly, the term $\left(p^d_n + \hat{p}^r_n\right)$ represents the total load at the node n with the load recovery effects taken into account for hour h^r. Then, we can use the vector $\underline{\hat{p}}^r \equiv \left[\hat{p}^r_1, \hat{p}^r_2, \ldots, \hat{p}^r_N\right]^T$ and p^r_0 to formulate the market clearing problem $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\big|_{h^r}$ for the hour h^r. We use the solutions of $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\big|_{h^r}$ to evaluate the market performance for each hour $h^r \in \mathrm{H}^r$ by using the suitably modified expressions of (2)-(6).

The hourly market clearing problem and the metrics defined using its solution serve as the basic building block in the quantification of the impacts of *DRR*s. We aggregate the variable effects for all the H hours in set H to quantify the variable effects of the entire simulation period. However, whenever a simulation involves future time periods, we need to explicitly consider the various sources of uncertainty which impact the operations of the markets and the system. We apply well-known probabilistic simulation notions to capture the uncertainty impacts on the system variable effects [17], [19].

We take into account the uncertainty due to the variability of the loads, the availability of the *SR*s and the clearing of the transmission-constrained markets. Under the set of adopted assumptions, we view the load over the simulation period as a random variable (*r.v.*) $\underset{\sim}{L}$ and use the forecasted hourly load data $\left\{\ell\big|_h : \forall h \in \mathrm{H}\right\}$ to estimate its cumulative distribution function (*c.d.f.*). Similarly, we use the discrete *r.v.* $\underset{\sim}{A}^i$ to represent the multi-state capacities for loading the *SR* $i \in \mathrm{T}$, where I represents the set of *SR*s in the system. We denote by $\alpha^i\big|_h$ the realization of $\underset{\sim}{A}^i$ for the hour h.

Now, the offers and the bids of the players for a particular hour h in the simulation period depend on the realizations $\ell\big|_h$ and $\alpha^i\big|_h$'s of the load and the available generation capacities, respectively. The uncertain load and available capacities result in uncertain market clearing outcomes, which we represent as *r.v.*s. We denote the cleared load that is

met by the supply resources by $\underset{\sim}{L}^{S}$, the generation output of *SR* i by $\underset{\sim}{P}^{i}$, the node n *LMP* by $\underset{\sim}{\Lambda}_{n}$, the supply-(demand-)side payments by $\underset{\sim}{W}^{S}$ $\left(\underset{\sim}{W}^{B}\right)$, the congestion rents by $\underset{\sim}{K}$ and the unserved demand by $\underset{\sim}{U}$. We use the outcomes of the hourly markets operated in the simulation period to approximate the *c.d.f.*s of each of these *r.v.*s.

We quantify the variable effects of the system by evaluating the expected values of the market outcome *r.v.s*. The expected supply-side payments for the simulation period are denoted by

$$\mathrm{W}^{S} = H \cdot E\left\{\underset{\sim}{W}^{S}\right\} \quad . \tag{9}$$

We use analogous expressions to evaluate the expected demand-side payments W^{B} and the expected congestion rents K. We compute the expected energy supplied by *SR* $i \in \mathrm{I}$ for the simulation period using

$$\mathrm{E}^{i} = H \cdot E\left\{\underset{\sim}{P}^{i}\right\} \quad . \tag{10}$$

We compute the contribution of the *SR* i to the expected emissions of the system over the simulation period using E^{i}. We incorporate the impacts of *DRR* curtailments, market clearing and transmission constraints in the computation of the reliability metrics by using the unserved demand *r.v.* $\underset{\sim}{U}$ to compute the loss of load probability *LOLP*,

$$LOLP = P\left\{\underset{\sim}{U} > 0\right\} \tag{11}$$

and expected unserved energy $\mathscr{U}$,

$$\mathscr{U} = H \cdot E\left\{\underset{\sim}{U} \middle| \underset{\sim}{U} > 0\right\} \cdot LOLP \, . \tag{12}$$

We use these economic measures and the reliability metrics to evaluate the complete set of variable effects of interest.

We construct the proposed simulation approach making use of the hourly snapshot-based market clearing and the probabilistic models described in this section. We describe the approach in the next section.

3. Proposed Approach

In this section, we describe the proposed approach to emulate, over longer-term periods, the operations of the power system and electricity markets with *DRR*s and to quantify the system variable effects. An important requirement is the explicit consideration of the time-varying nature of *DRR* deployments and the representation, with appropriate level of detail, of the loads and resources, transmission grid, the market clearing operations, the

market structure and various sources of uncertainty. We discuss the approach for a single simulation period for which the assumptions in section II are satisfied.

Conceptually, the market clearing for each hour h in the simulation period is performed for the particular realizations of $\underset{\sim}{L}$ and $\underset{\sim}{A}^i$, the demand and supply *r.v.*s. The realization ℓ of $\underset{\sim}{L}$ manifests itself in terms of the sum of the fixed price load and price-sensitive bids submitted by the buyers. The realization α^i of $\underset{\sim}{A}^i$ represents the maximum capacity each unit $i \in \mathrm{I}$ offers in the market. In the absence of strategic behavior, the capacity α^i is offered to the market. In the event of a forced outage, $\alpha^i = 0$ and so unit i cannot contribute towards meeting the load in hour h. Indeed, the load sample ℓ and the sampled available capacities α^i are used to construct demand and supply curves, respectively, for the hour h market. Since the solution of $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\Big|_h$ is used to compute the realization of each market outcome *r.v.* – $\underset{\sim}{L}^{\mathrm{S}}$, $\underset{\sim}{P}^i$, $\underset{\sim}{\Lambda}_n$, W^{S}, $\underset{\sim}{W}^{\mathrm{B}}$, $\underset{\sim}{\mathrm{K}}$ and $\underset{\sim}{\mathrm{U}}$ – corresponding to the realized values of $\underset{\sim}{L}$ and $\underset{\sim}{A}^i$, we may view the *OPF* $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\Big|_h$ as mapping the realizations ℓ and α^i – the *input* – into the realizations ℓ^{S}, p^i, λ_n, w^{S}, w^{B}, K and U of the corresponding market outcome *r.v.*s – the *output* – for each hour $h \in \mathrm{H}$. The resulting collection of the market outcome realizations constitutes the sample space that we use to approximate the *c.d.f.*s of the market outcome *r.v.*s. In this way, we construct approximations for each *c.d.f.* of interest and then proceed with the evaluation of the metrics in (9)-(12).

Clearly, the solution of $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\Big|_h$ for the hour $h \in \mathrm{H}$ depends on whether hour $h \in \mathrm{H}^c$ or $h \in \mathrm{H}^r$ or $h \notin \mathrm{H}^c \cup \mathrm{H}^r$ and on the load curtailment capacity offered/associated load recovery effects. The time-dependent nature of the *DRR*s and that of the market clearing require their representation in the simulation. This requirement precludes the direct application of the conventional probabilistic simulation framework [17], [19], in which time is abstracted out. As the load $\underset{\sim}{L}$ lacks temporal information, a modification of the probabilistic simulation is necessary to incorporate the impacts of *DRR* deployment. The modification makes use of the forecasted hourly loads and the *DRR* usage schedule, be it for a typical weekday and a typical weekend day. We make use of the uniform load characteristics exhibited over the entire simulation period, i.e., in every hour of that period, to incorporate the effects of the time-varying *DRR*s. In fact, the probability distribution of the load *r.v.* $\underset{\sim}{L}$ holds for every hour in the period so that the approximate *c.d.f.* of $\underset{\sim}{L}$ is constructed from the collection of sampled load values for each hour $h \in \mathrm{H}$. Therefore, we can partition the collection of the hourly load samples $\left\{\ell\big|_h : h \in \mathrm{H}\right\}$ into 24 week day hour subsets and 24 weekend day hour subsets, and use these subsets to approximate the conditional *c.d.f.* of $\underset{\sim}{L}$ conditioned on each such hour. Indeed, we can use load samples from the partitioned subsets to construct corresponding subsets of market outcome realizations obtained via the *OPF* $\mathrm{M}\left(\mathrm{S}, \hat{\mathrm{B}} \cup \bar{\mathrm{B}}\right)\Big|_h$ mapping and approximate the conditional *c.d.f.*s of the market outcome *r.v.*s conditioned on each

such hour. The *c.d.f.*s of the market outcome *r.v.*s are then approximated as the probability weighted average of the conditional *c.d.f.*s for the 24 week day and 24 weekend hours. The details on the hourly conditioning and *c.d.f.* evaluation are presented in [12, pp. 53-67]. We construct the proposed approach through the systematic application of the hourly conditioning scheme.

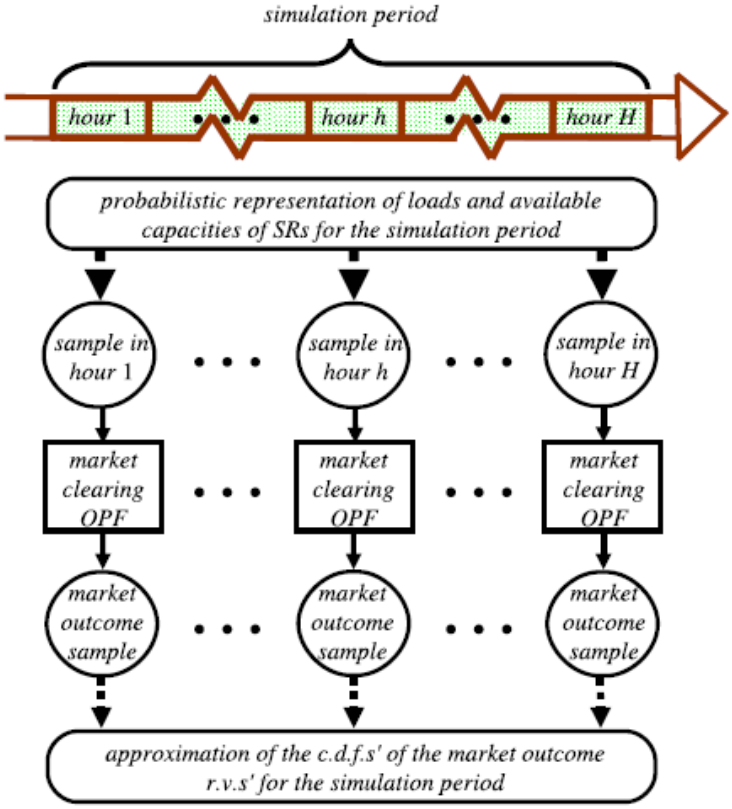


Fig. 1. Conceptual illustration of the scheme to construct the *c.d.f.s* of the market outcome *r.v.s*

We implement the proposed approach to quantify the variable effects of a power system with *DRR*s. The approach effectively captures the interactions between the *DRR* deployments, generation dispatch, transmission usage and congestion, market structure and market clearing outcomes under the specific policies in effect. The explicit representation of the various sources of uncertainty provides a realistic emulation of the hour-by-hour operation of the actual system over the simulation period. We do need to manage effectively the computing burden associated with the simulation of a large-scale system over a longer-term period. We describe in the next section the steps that need to be taken to ensure computational tractability in implementing the approach.

4. Implementational Aspects

In this section, we discuss the implementational aspects of the proposed approach by focusing on the steps taken to ensure its computational tractability and make practical its application to the simulation of large-scale systems over longer-term periods. For a study spanning multiple years, we first partition the *study period* into T non-overlapping *simulation periods* to capture seasonal effects, changes in the resource mix and the transmission grid, maintenance schedules as well as the introduction of new policies. Any change entails the specification of a new simulation period and the introduction of the assumption that the change persists over that entire simulation period. As such, the simulation periods may be of unequal duration. We denote the index set of the non-overlapping simulation periods by $\mathrm{T} \subset \{1, \ldots, T\}$. We apply the proposed approach to each of the T simulation periods and aggregate the variable effects for these simulation periods, taking into account the time value of money where appropriate, to compute the variable effects for the study period.

To make the computing tasks tractable, we choose judiciously the simulation periods for a study and use numerically efficient computation of the hourly snapshots in each simulation period. Typically, we specify the simulation periods to be 168-hour weeks. Then, the set T consists of $52 \cdot y$ indices, where y is the number of years in the study period. Now, we can take advantage of the fact that several weeks in a year have similar load patterns because of the seasonal nature of electricity demand. So, we can use representative weeks to reduce the number of simulations below $52 \cdot y$. However, since the set of *SR*s on scheduled maintenance may differ across the weeks with similar load patterns, the number of representative weeks increases. The index set of the subset of the selected representative weeks is $\mathrm{T}' \subset \mathcal{T}$. We associate with each representative week $t' \in \mathrm{T}'$ a corresponding number, $\psi_{t'}$, of weeks it represents and apply the proposed approach to simulate each representative week $t' \in \mathrm{T}'$. The resulting assessments for that week are weighted by $\psi_{t'}$ to compute its contribution to the study period variable effects. In this way, we simulate representative weeks in the study period to effectively reduce the computations required.

We can further implement numerically efficient schemes for the simulation of each representative week. The approximation of the market outcome *r.v.*s *c.d.f.*s for a representative week entails extensive computations to take into account the many possible input sample realizations corresponding to the realized values of the demand and supply *r.v.*s. Consequently, we need an effective sampling technique to produce input sample realizations which are representative of all the possible realizations of the demand and supply *r.v.*s.

Latin hypercube sampling or *LHS* [20] is a variance reduction technique that uses a stratified sampling approach to produce multiple samples realizations for a set of *r.v.*s with known distributions. To obtain M sample realizations of a *r.v.*, we partition the domain of the *c.d.f.* of the *r.v.* into equal probability M intervals. Then, M realizations of the *r.v.* are chosen, one from each such interval. When sampling across multi-variate

distributions of a set of demand and supply *r.v.s*, we perform *LHS* on each component *r.v.* and obtain M sample realizations of each *r.v.* We then construct M input samples by pairing the sampled values of each component *r.v.* The pairing scheme depends on the correlations and dependence between the component *r.v.*s [21]. The M sample realizations of the demand and the supply *r.v.*s result in a computationally efficient approximation of the *c.d.f.*s of the market outcome *r.v.*s and their expected values [22]. Unlike random sampling, the stratified sampling approach of the *LHS* ensures that the entire distribution of the demand and supply *r.v.*s – including the extreme regions of the *c.d.f.*s – is sampled, providing useful means for evaluating the reliability metrics *LOLP* and U [22]. To choose an appropriate sample size M, we use a frequently-employed rule of thumb and require that the sample size be larger than the number of *r.v.*s being sampled by a factor η, where $\eta \in [1.5, 20]$. Our extensive testing has borne out that this rule of thumb is useful for the large-scale systems we tested.

The judicious selection of representative simulation periods and the use of *LHS* technique to build statistically representative sample realizations for the approximation of the *c.d.f.*s of the market outcome *r.v.*s bring about computational tractability in the implementation of the proposed approach. This implementation strategy for the proposed approach is particularly effective in the study of large-scale systems over longer-term periods. We illustrate, in the next section, the application of our approach with the implementational scheme to answer a wide variety of *what-if* questions for planning and policy analysis.

5. Application Studies

We illustrate the application of the proposed simulation approach to a set of studies on a large-scale test system. The goal of these studies is to investigate the ramifications of the integration of *DRR*s into the power system. We investigate the impacts of varying levels of *DRR* penetration and load recovery effects and analyze the economic, environmental and reliability benefits associated with the effective utilization of *DRR*s. The results discussed here are representative of those in our extensive tests of the proposed approach. These results demonstrate the capabilities of the proposed approach to study various power system planning and policy analysis issues.

The test system we use for the simulations consists of 241 buses and 555 transmission lines and is representative of largescale *ISO* networks. We explicitly consider the constraints imposed by the real power transfer capabilities of the transmission grid. There are loads connected at 130 buses in the network. We use the historical load shapes of the Midwest *ISO* system for the year 2006 [24]. The aggregate average hourly load for the system is 70 *GW*, with the annual peak load of 117 *GW* in the summer. There are generators connected at 152 buses in the network. The supply-side resource mix capacity composition is given in Table I. The supply system consists of 766 generators. Each generation unit has a pre-specified maintenance schedule. We use the emissions factors in [25] to estimate the expected CO_2 emissions from the supply-side generation.

TABLE I
The Capacity Composition of the Supply-Side Resource Mix in *GW*

coal	CCGT	peakers	others	total
70	21	24	20	135

We designate some of the loads in the system as *DRR*s. We use the total *DRR* capacity – expressed as a fraction of the annual system peak load – as the penetration level parameter and study the impacts of varying *DRR* penetration levels. The impacts of the load recovery consequences for the *DRR* curtailments are studied by varying the associated *CRF*s $\chi^{\hat{b}}_{h^c,h^r}$'s. In the simulations reported here, we consider each buyer as submitting fixed demand bids in the hourly *DAM*s. We limit the hours *DRR*s can offer load curtailments in the *DAM*s from the hour ending at 9:00 to the hour ending at 18:00 for each weekday. The load recovery hours are from the hour ending at 23:00 to the hour ending at 6:00.

We limit our analysis to a single year to get insights into the nature of results obtained. The seasonality effects, the load patterns and the maintenance scheduling requirements allow us to reduce the 52 weeks in the year to 15 representative weeks. Each representative week appears at least once. In the *LHS* deployment, we select sample sizes between 500 and 1000. We use the test system simulation without the *DRR*s, which we denote as D_0, to provide the variable effects for the reference case of our studies. We assess the impacts of the *DRR*s with respect to the variable effects in D_0 . These are summarized in Table II.

TABLE II
Variable Effects of Case D_0 for the Simulated Year

min, avg and max values of the cleared load (*GW*)		average values of the economic metrics	
base load	50.806	electricity payments	*$* 4.453 M
average load	69.910	congestion rents	*$* 67,752
peak load	117.658	CO_2 emissions	11,823 tons

We discuss first the implementation of *DRR*s into the test system with a total capacity of 5.6 *GW*, which is approximately 5 *%* of the peak load. We examine two specific cases: $D_{05,00}$ without recovery of the *DRR* curtailments and $D_{05,70}$ with 70 *%* recovery

TABLE III
Minimum, Average And Maximum Values for the Hourly Cleared Loads in *GW*

case	base load	average load	peak load
$D_{05,00}$	50.806	69.830	112.720
$D_{05,70}$	52.242	69.575	112.720

of the curtailments offered and accepted in the *DAM*s. We summarize the load-related metrics for each of these cases in Table III from which we can evaluate the effect of *DRR*s on reducing the loads in the peak load hours of the study period. The reduction in the load is observed in approximately 25 *%* of the hours in the year. The load reduction results in a visible downward shift in the load duration curve (*LDC*), as shown in Fig. 2. Note the *LDC* shown here is constructed using the hourly cleared loads.

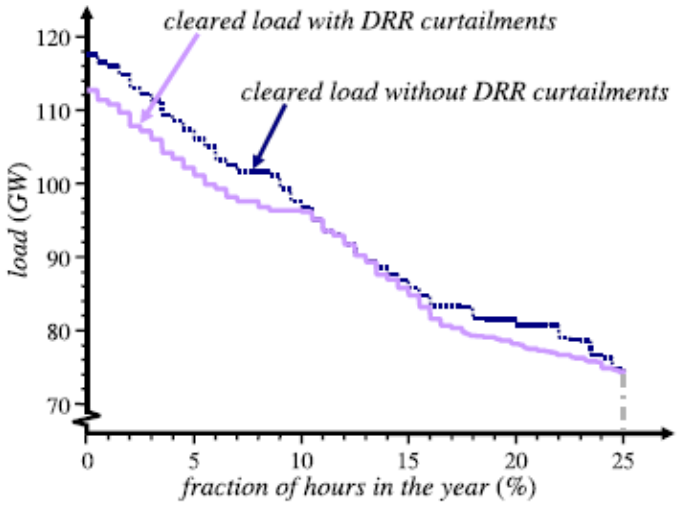


Fig. 2. The impacts of *DRR* curtailments on the cleared loads for the peak hours of the year as seen from the high load portion of the annual *LDC*

The reduction in the annual peak load from 117.658 to 112.720 *GW* increases the capacity margin of the system by 5 *%*; from 14.74 to 19.77 *%*. When the load recovery effects are considered in case $D_{05,00}$, the increases in the cleared loads during the off-peak hours lead to an upward shift in the *LDC*, including a larger base load value, as shown in Fig. 3.

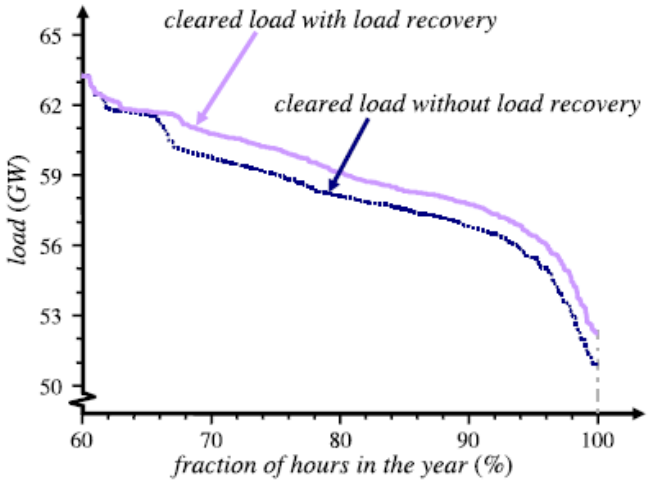


Fig. 3. The impacts of *DRR* curtailments recovery during the off-peak hours of the year as seen from the low load portion of the annual LDC

The *DRR*-provided peak shaving results in a number of important economic and environmental benefits as measured by the electricity payments, congestion rents and CO_2 emissions. We tabulate the average values of these metrics for the $D_{05,00}$ and $D_{05,70}$ cases in Table IV.

TABLE IV
Average Values for the Hourly Metrics

case	electricity payments (M$)	congestion rents ($)	CO_2 emissions (tons)
$D_{05,00}$	4.069	44,350	11,574
$D_{05,70}$	4.150	45,515	11,738

The *DRR* reduced system loads during the peak hours decrease the *LMP*s in those hours, which in turn, reduce the annual electricity payments of all the buyers by nearly as much as 11%. We note that the congestion rents in these two *DRR* cases are significantly lower than the reference case – by as much as 35% – indicating that the *DRR* deployments can drastically impact network flows and provide transmission congestion relief. Moreover, the decrease in the energy consumption implies that the generation for the system is also reduced, and hence we have a decrease in the CO_2 emissions.

To assess the impacts of *DRR* penetration and load recovery effects, we simulate cases with varying capacity of *DRR*s and different values of the *CRF*s $\chi^{\hat{b}}_{h^c,h^r}$'s. We illustrate in Fig. 4 that as *DRR* penetration deepens, the capacity margin increases and the aggregate

annual congestion rents decrease. Since the *DRR* provided peak load reduction is not impacted

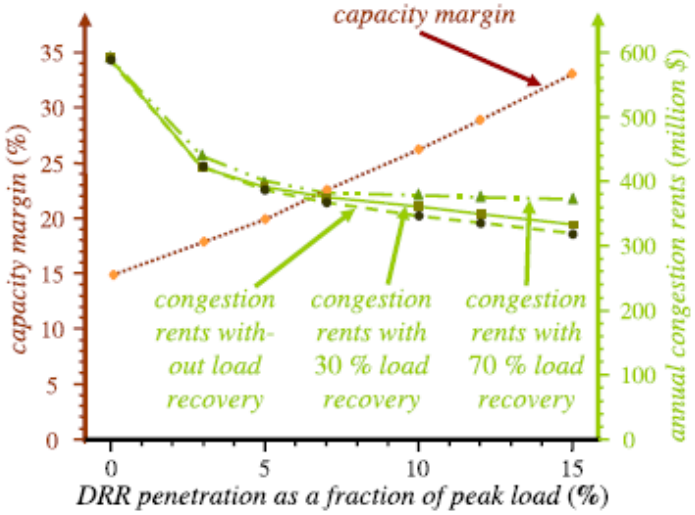


Fig. 4. The impacts of deeper *DRR* penetration and Increased load recovery effects on the system capacity margin and congestion rents

by load recovery, the capacity margin depends on the level of *DRR* penetration. The total congestion rents collected by the *IGO* during the year are impacted by both the amount of load curtailed and the amount of load recovered by the *DRR*s. As the load recovery effects become more pronounced, the reduction in the congestion rents with deeper *DRR* penetration decreases markedly. Similar plots are available to illustrate the nature of the total annual electricity payments. While the energy savings decrease as more load is recovered, the total electricity payments in *DRR* cases with different penetration levels and load recovery effects remain lower than those in the reference case D_0 . Shifting the load from peak to off-peak hours due to *DRR* deployment results in the utilization of more economic generation resources. Therefore, not withstanding the load recovery effects, a deeper *DRR* penetration results in economic benefits in terms of reduced electricity payments and congestion rents, representing the more efficient utilization of the generation and transmission resources.

We also investigate the ability of *DRR*s to defer the need for additional generation capacity. We compare the D_0 and $D_{05,00}$ cases under a 3 % load growth scenario and no additions to/retirements from the supply-side resource mix. We compute the hourly cleared loads in these cases and compare them to the forecasted loads. We note a capacity shortage in 2 % of the hours of the year for the reference case D_0 and resulting loss of load at some nodes. However, no shortage of capacity arises in the $D_{05,00}$ case due to the reduced loads during the peak hours. The illustration in Fig. 5

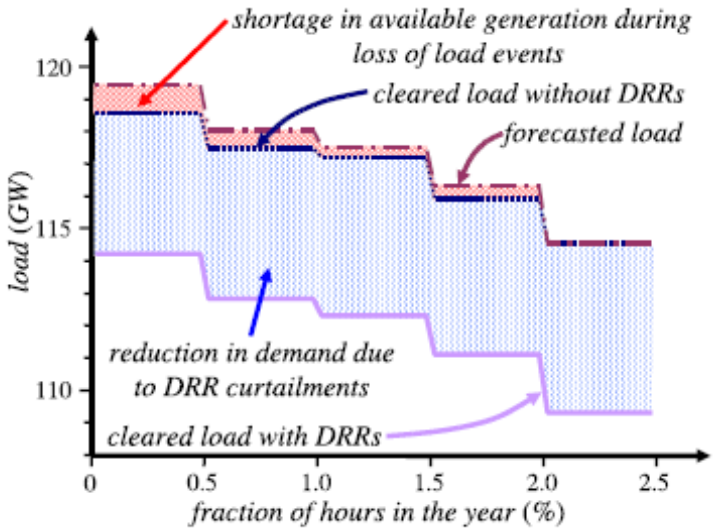


Fig. 5. Reduction in loss of load events due to *DRR* deployment

clearly indicates the capacity shortage in the peak load hours in D_0. Such shortage effects may be overcome by adding approximately 1 *GW* of peaking capacity. Clearly, the effective use of *DRR*s is able to defer the need for such additional capacity. Therefore, *DRR* curtailments can impact the need for additional generation capacity.

We also investigate the impacts of *DRR* deployments on the need for transmission reinforcement. We compute the aggregate annual congestion rents for the cases without *DRR*s and with 5 *% DRR* capacity for four different configurations of the transmission grid: the existing grid, the existing grid with line τ_a upgraded, the existing grid with lines τ_a and τ_b upgraded, and, the existing grid with lines τ_a, τ_b and τc upgraded. We plot the congestion rents in Fig. 6 for four different grid configurations for the two cases. For simplicity, we use *CRF*s $\chi^{\hat{b}}_{h^c,h^r} = 0$ in the 5 *% DRR* case but we obtain similar plots for other values of $\chi^{\hat{b}}_{h^c,h^r}$. We note from Fig. 6 that as more transmission lines are upgraded,

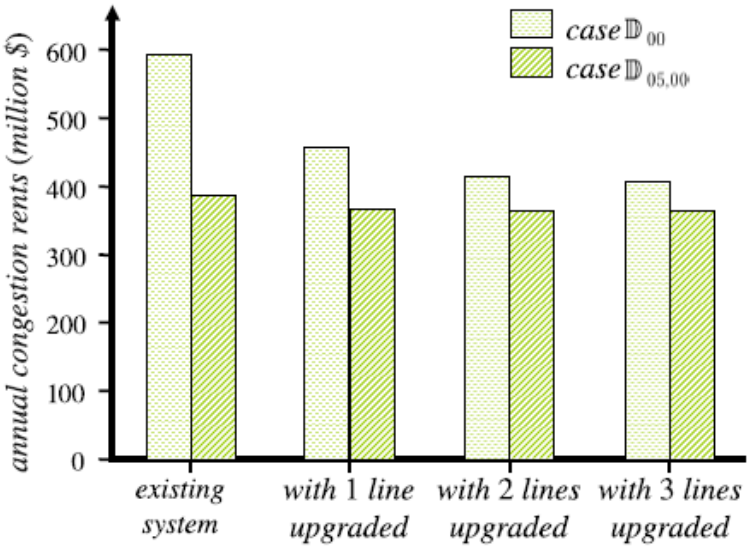


Fig. 6. Congestion rents for the system with transmission line upgrades in the $\mathbb{D}_0$ and $\mathbb{D}_{05,00}$

the transfer capability of the system improves and we observe lower congestion rents for both cases – with and without *DRR*s. A significant finding of this exercise is that the lowest congestion rents are *$*406 million for the system with 3 line upgrades for the reference case D_0 , which are higher than the congestion rents of *$*387 million on the existing transmission system for the *DRR* case $D_{05,00}$. We observe similar results when load recovery effects are considered. We conclude, consequently, that effective utilization of the *DRR*s can lead to more reduction in the congestion rents than undertaking the capital intensive projects such as transmission line upgrades. Indeed, the integration of *DRR*s into the power system may defer the need for additional transmission and not only for new supply-side resources.

In each of the many cases studied, we have quantified the beneficial impacts of *DRR*s in reducing the system electricity payments, the transmission congestion and the loss of load probability. These benefits vary with the load recovery effects and the penetration of *DRR*s. Such quantification, as given in the representative results discussed, demonstrates the effectiveness of the proposed approach for use in a broad range of applications.

6. Concluding Remarks

In this report, we presented a proposed a simulation approach for evaluating the variable effects of power systems with integrated *DRR*s. The ability to quantify the impacts of *DRR*s on the economics of electricity supply, environmental effects of electricity production and reliability of the system makes the approach very useful in regulatory filing studies, long-term planning and policy analysis. We demonstrate the capability of the proposed approach to answer a broad range of *what-if* questions for a realistic-sized power system. We present results from extensive simulation studies which quantify

impacts of *DRR*s on the system variable effects for different *DRR* penetration levels and varying load recovery characteristics. Our investigations provide practical insights into the significant role played by the *DRR*s in the efficient utilization of generation and transmission resources and in increased competition in the electricity markets to bring about lower prices to consumers.

As the penetration of renewable resources – such as wind and solar – deepens, the simulation approach proposed here needs to be extended to include their impacts on the system and market operations. The modeling must carefully represent the lack of control capability and the intermittency effects associated with the renewable generation along with the uncertainty associated with prediction of the wind/solar output patterns. In addition, energy storage units – be they utility-scale storage devices or large aggregations of plug-in hybrid electric vehicles – whose flexibility can be effectively harnessed in the management of the intermittent renewable resources need to be explicitly represented. Future work will focus on extending the proposed approach to incorporate renewables, storage and other time-dependent resources.

References

[1] R. Masiello, "Demand response: The other side of the curve [guest editorial]," *Power and Energy Magazine, IEEE*, vol. 8, no. 3, pp. 18 –18, may-june 2010.

[2] L. Ruff, "Economic principles of demand response in electricity," Tech. Rep., October 2002, prepared for Edison Electric Institute, Washington, D.C.

[3] U. S. Department of Energy, "Benefits of demand response in electricity markets and recommendations for achieving them," February 2006. [Online]. Available: *{*http://www.oe.energy.gov/DocumentsandMedia/congress 1252d.pdf*}*

[4] FERC Staff, "Assessment of demand response and advanced metering," Federal Energy Regulatory Commission, Tech. Rep., December 2008. [Online]. Available: http://www.ferc.gov/legal/staff-reports 12-08-demand-response.pdf

[5] A. Faruqui, R. Hledik, S. Newell, and H. Pfeifenberger, "The power of 5 percent," *The Electricity Journal*, vol. 20, no. 8, pp. 68–77, October 2007.

[6] A. Kowli, "Assessment of variable effects of systems with demand response resources," Master's thesis, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 2009. [Online]. Available: http://energy.ece.uiuc.edu/gross/ papers/Dissertations/Kowli.pdf

[7] K. Hamilton and N. Gulhar, "Taking demand response to the next level," *Power and Energy Magazine, IEEE*, vol. 8, no. 3, pp. 60 –65, may-june 2010.

[8] “A national assessment of demand response potential,” Federal Energy Regulatory Commission, Tech. Rep., June 2009. [Online]. Available: http://www.ferc.gov/legal/staff-reports/06-09-demand-response.pdf

[9] G. Strbac, E. Farmer, and B. Cory, “Framework for the incorporation of demand-side in a competitive electricity market,” *IEE Proceedings - Generation, Transmission and Distribution*, vol. 143, no. 3, pp. 232–237, May 1996.

[10] A. Borghetti, G. Gross, and C. A. Nucci, “Auctions with explicit demand-side bidding in competitive electricity markets,” in *The Next Generation of Electric Power Unit Commitment Models*, B. F. Hobbs, M. H. Rothkopf, R. P. O’Neill, and H. po Chao, Eds. Norwell, MA: Kluwer Academic Publishers, 2001, vol. 36, pp. 53–74.

[11] C.-L. Su and D. Kirschen, “Quantifying the effect of demand response on electricity markets,” *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1199–1207, August 2009.

[12] R. Walawalkar, S. Blumsack, J. Apt, and S. Fernands, “Analyzing PJM’s economic demand response program,” in *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, July 2008, pp. 1–9.

[13] E. Bompard, R. Napoli, and B. Wan, “The effect of the programs for demand response incentives in competitive electricity markets,” *European Transactions on Electrical Power*, vol. 19, no. 1, pp. 127–139, January 2009.

[14] V. Stanojevic, V. Silva, D. Pudjianto, G. Strbac, P. Lang, and D. Macleman, “Application of storage and demand side management to optimize existing network capacity,” in *Electricity Distribution, 2009 20th International Conference and Exhibition on*, June 2009, pp. 1–4.

[15] R. Earle, E. P. Kahn, and E. Macan, “Measuring the capacity impacts of demand response,” *The Electricity Journal*, vol. 22, no. 6, pp. 47 – 58, July 2009.

[16] S. Stoft, *Power System Economics: Designing Markets for Electricity*. New York, NY: Wiley-IEEE Press, 2002.

[17] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation and Control*, 2nd ed. New York, NY: John Wiley and Sons, Inc., 1996.

[18] M. Liu and G. Gross, “Framework for the design and analysis of congestion revenue rights,” *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 243–251, Feb. 2004.

[19] R. Sullivan, *Power System Planning*. New York, NY: McGraw Hill International Book Company, 1977.

[20] M. McKay, R. Beckman, and W. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 1, pp. 239–245, 1979.

[21] M. Stein, "Large sample properties of simulations using latin hypercube sampling," *Technometrics*, vol. 29, no. 2, pp. 143–151, 1987.

[22] J. Linderoth, A. Shapiro, and S. Wright, "The empirical behavior of sampling methods for stochastic programming," *Annals of Operations Research*, vol. 142, pp. 219 – 245, 2006.

[23] P. Jirutitijaroen and C. Singh, "Comparison of simulation methods for power system reliability indexes and their distributions," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 486–493, May 2008.

[24] Midwest ISO staff, archived data in the Market Reports Section on the Midwest ISO website, July 2008. [Online]. Available: http://www.midwestiso.org/publish/Folder/7be606 10b7aacd66e-7da30a48324a?rev=6

[25] S. Meyers, C. Marnay, K. Schumacher, and J. Sathaye, "Estimating carbon emissions avoided by electricity generation and efficiency projects: A standardized method," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-46063, July 2000. [Online]. Available: http://eetd.lbl.gov/ea/ems/reports/46063.pdf

Part III

Stochastic Modeling of Multi-Area Wind Power Production and Evaluation of Direct Coupling between Deferrable Load and Intermittent Renewable Resources

Shmuel S. Oren
Anthony Papavasiliou – Post-Doctoral Researcher
University of California at Berkeley

For information about this project, contact

Shmuel S. Oren, Ph.D.
Department of Industrial Engineering and Operations Research
University of California at Berkeley
Berkeley, CA 94720
Tel: 510-642-1836
Fax: 510-642-1403
Email: oren@ieor.berkeley.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Table of Contents

List of Figures

1. Introduction

1.1 Multi-Area Wind Production

The large-scale integration of renewable power supply in power systems has recently motivated researchers to consider stochastic unit commitment policies for committing reserves in order to guarantee the reliable operation of the grid. Such studies include the work of Ruiz et al. (Ruiz, 2009b), Wang et al. (Wang, 2008), Constantinescu et al. (Constantinescu, 2011), Tuohy et al. (Tuohy, 2009), Morales et al. (Morales, 2009b), Bouffard et al. (Bouffard, 2008) and Papavasiliou et al. (Papavasiliou, 2010). Stochastic unit commitment models explicitly account for uncertainty in the formulation of the unit commitment problem and therefore have the potential to outperform ad-hoc deterministic reserve rules that are used in practice. The formulation of the stochastic unit commitment problem requires explicit modeling of the uncertain parameters in the unit commitment problem in terms of a few appropriately weighted representative scenarios.

Uncertainty in power system operations can be categorized between discrete and continuous disturbances. Discrete disturbances refer to the failure of equipment such as generators and transmission lines. Continuous disturbances include parameters of the unit commitment problem that vary smoothly such as electricity demand and renewable power production.

Transmission constraints strongly affect the optimal rule for allocating reserves in each area of the network. In order to account for transmission constraints, operators often use ad-hoc import constraints for determining locational reserve requirements. Import constraints can be categorized between `bubble' constraints and inter-tie constraints. `Bubble' constraints limit the total amount of power that can flow into a load pocket in order to ensure that the unit commitment schedule reserves sufficient transfer capability on the lines in order to protect against the possibility of generation capacity failure within the load pocket. On the other hand, inter-tie constraints limit the amount of power that can flow over inter-ties in order to protect the system against the failure of major corridors that bring significant amounts of power from outside the system. Both types of constraints are formulated on an ad-hoc basis, in the sense that there is no formal methodology for determining the set of lines belonging to an import constraint and the limit on the amount of power that can flow on the import set. The complexity of committing reserves in the presence of transmission constraints has been demonstrated by various authors, including Arroyo and Galiana (Arroyo, 2005), Galiana et al. (Galiana, 2005) and Bouffard et al. (Bouffard, 2005). Beyond their influence on reserve requirements, transmission constraints also affect the cost of operating the system. This is due both to the fact that transmission constraints reduce the flexibility of dispatching conventional generators in the system, and also due to the fact that they result in the waste of renewable energy supply.

The inclusion of transmission constraints in the unit commitment model necessitates the development of a multi-area wind production model. Moreover, in order to assess the impact of wind power production on power system operations over an entire year, it is

necessary to account for the non-stationary (seasonal and diurnal) patterns of wind power production. This paper presents a multi-area stochastic wind production model that captures the seasonal and diurnal patterns of wind power production, accounts for the temporal and spatial correlations of the original data set and accurately reproduces the marginal distribution of wind power production at each location of the network. Moreover, the proposed model is applied to a detailed dataset of the California wind power resources corresponding to the 2012 and 2020 Renewable Portfolio Standards.

1.2 Integrating Renewable Supply with Deferrable Demand

The proliferation of renewable energy sources in the United States is taking place at an unprecedented pace. The federal government is coordinating these efforts, with state regulations further advancing renewable energy integration targets. The American Clean Energy and Security Act (2009) set a target of sourcing 20% of US electricity consumption from renewable energy by 2020 and also set various goals for limiting reliance on nonrenewable resources.

In 2002, the state of California enacted the Participating Intermittent Resources Program, which facilitates the integration of renewable energy sources. The California Renewable Portfolio Standard requires 20\% of energy supply in the state to be sourced from renewable sources by 2012.

In the California ISO renewable integration report, (CAISO, 2007) voiced concerns about the impacts of large scale renewable energy integration on the capacity requirements and ramping requirements of load following and regulation services.

These increased reserve requirements represent a significant barrier for the large-scale integration of renewable power. An alternative to the costly investment in backup generation capacity is to exploit the flexibility of electricity demand. In close analogy to the policy coordination that is taking place for renewable energy integration, demand-side management is also being coordinated both at the federal level and in individual states. The American Clean Energy and Security Act (2009) has allocated $3.4 billion in order to spawn the development and deployment of the necessary technology to enable active management of electricity demand.

Anticipating the importance of demand-side flexibility, the California electricity market rules have been adapted in order to accommodate the participation of demand resources through the recent Market Redesign and Technology Upgrade of 2007. Two of the major state utilities, San Diego Gas and Electric and Pacific Gas and Electric, are deploying smart metering throughout their respective service areas.

The most efficient approach for exploiting demand-side flexibility would be to establish real-time pricing at the retail level, a possibility which is discussed by (Borenstein, 2002). However, there is strong political opposition to this approach as it exposes retail consumers to the volatility of electricity prices. In an alternative approach, which is described by (Kirby, 1999) and (Kirby, 2003), flexible loads can participate in the ancillary services market. An aggregator could bid on behalf of a population of loads for

providing capacity services to the system operator and would coordinate load consumption either through prices or direct control. The technical feasibility of demand side aggregation for the provision of spinning reserve has been studied in practice by (Eto, 2007).

However, it is necessary to define market products that correspond to the types of ancillary services that loads can actually offer, which raises the need for reform in existing electricity markets.

In this paper we analyze a direct contractual agreement between deferrable loads and renewable generators. In the proposed contract, loads request a certain quantity of energy within a deadline. An aggregator is then responsible for serving these requests within their deadlines by relying primarily on renewable resources and, to a limited extent, on the real-time market.

The proposed contract closely matches dynamic scheduling, as described in (Hirst, 1997), whereby demand and supply resources from different control areas coordinate their schedules in order to produce zero net output to the remaining system. Dynamic scheduling is currently implemented in the Electric Reliability Council of Texas.

The specification of demand flexibility in terms of requests for a fixed amount of energy within a fixed time horizon is a natural description for a wide variety of flexible energy needs, such as EV charging, water pumping, and various residential consumptions. There is also a natural complementarity between coupling renewable resources with deferrable requests. Due to the fact that renewable power supply is more predictable over a certain time horizon than in any given moment in time, it is easier to fulfill requests that extend over a time window. In addition, the contract relieves the system operator from the obligation of procuring reserves for protecting against intermittent renewable supply since renewable resources appear "behind the meter" . The significant capital savings that stem from avoided investment in backup reserves can be shared with deferrable loads in order to incent their flexible behavior. Although the coupled system may resort to the spot market to a limited extent, the coupling contract effectively transfers the risk of renewable power variability from the system operator to deferrable loads.

A disadvantage of the proposed approach is that the bilateral commitment between loads and renewable generators results in trading inefficiencies. Coupling also reduces the effect geographical smoothing in renewable energy supply. Moreover, the contract requires direct load control by the aggregator, which might be undesirable to consumers.

For coupling contracts to work, renewables must be exposed to some of the risk they impose on the system. This can be achieved by forcing renewable suppliers to bid in the day-ahead market with a penalty for deviation, by removing feed-in tariffs, or by exposing renewable suppliers to the risk of spilling excess wind.

1.3 Report Organization

The first part of this report investigates a stochastic model for multi-area wind production that is used for planning reserves in transmission-constrained systems with large amounts of integrated renewable power supply. In section 2.1 we present the methodology for calibrating the multi-area wind production model and for simulating the process. In Section 2.2 we present a case study of the California power system. In section 2.3 we summarize the conclusions of our study. In Section 3 we present a contract for integrating renewable energy supply and electricity spot markets for serving deferrable electric loads in order to mitigate renewable energy intermittency. The model is formulated in section 3.1. In section 3.2 we present simulation results. We conclude in Section 4.

2. Stochastic Modeling of Multi-Area Wind Power Production

2.1 Methodology

The nonlinear dependence of wind power production on wind speed raises challenges in the statistical modeling of wind power production. It is therefore common in the wind power modeling literature to model wind speed and use a static power curve to calculate the corresponding wind power production.

The task of modeling wind speed consists of removing seasonal and daily patterns from wind speed data, fitting the resulting process to a parametric or non-parametric distribution, and fitting an appropriate time series model to the underlying noise in order to capture the strong temporal correlation of wind speed time series. Early work on wind power modeling was performed by Brown et al. (Brown, 1984). The authors list various parametric distributions for fitting wind speed data, such as the Weibull, inverse Gaussian and exponential distribution. The authors use an exponential function to transform their data to an approximately Gaussian data set. They remove hourly means and estimate the order of an appropriate autoregressive model and they use the Yule-Walker equations (Box-Jenkins, 1976) to estimate the parameters of the autoregressive model. Torres et al. (Torres, 2005) follow the same methodology as Brown et al. (Brown, 1984). The authors use autoregressive moving average models and find that these more general models provide a more satisfactory fit.

Transmission constraints have recently prompted researchers to develop multi-area wind production models. Moreover, diurnal and seasonal patterns of wind power production need to be accounted for in order to assess the impact of wind integration on power system operations over an entire year. In recent work, Morales et al. (Morales, 2010) develop a multi-area wind speed model by using a noise vector that drives a vector autoregressive process. In order to simplify the calibration of the model, the authors assume a diagonal matrix of autoregressive coefficients, which implies that spatial correlations among wind speed in various locations are captured fully by the underlying noise vector. The calibration and simulation model that we use extends the approach of Morales et al. (Morales, 2010) in order to account for seasonal and diurnal wind speed patterns.

2.1.1 Calibration

Given a multi-area data set y_{kt}, where k indexes location and t indexes time, the first step of the calibration procedure is to remove diurnal and seasonal patterns. We normalize the data by subtracting the hourly mean and dividing by the hourly standard deviation in order to obtain a stationary data set y_{kt}^{S} for each location. Systematic patterns can be monthly, seasonal, or may even vary between weekdays and weekends, as is the case for load data. In each case, the appropriate portion of the data set should be chosen for estimating the mean and variance. In the present study, the data is partitioned by month.

We next filter the data set in order to obtain an approximately Gaussian stationary data set y_{kt}^{GS}. Brown et al. (Brown, 1984), Torres et al. (Torres, 2005) and Morales et al. (Morales, 2010) use this approach for transforming Weibull-distributed wind speed data to Gaussian data, and Callaway (Callaway, 2010) uses a non-parametric transformation. In the single-area wind integration study of Papavasiliou et al. (Papavasiliou, 2010), the authors find that the inverse Gaussian distribution provides a satisfactory fit for the data set. For the multi-area wind integration study presented in this paper, no single parametric distribution provides a close fit for the observed data in all locations, therefore we fit an empirical distribution $\hat{F}_k$to the data of each location k.

The resulting time series y_{kt}^{GS} can be modeled by an autoregressive model:

$$y_{k,t+1}^{GS} = \sum_{j=0}^{p} \hat{\phi}_{kj} y_{k,t-j}^{GS} + \hat{\omega}_{kt}, \qquad (1)$$

where $\hat{\omega}_{kt}$ is the estimated noise and $\hat{\varphi}_{kj}$ are the estimated coefficients of the autoregressive model.

The calibration process is summarized in the following steps:

Step (a). Remove systematic seasonal and diurnal effects:

$$y_{kt}^{S} = \frac{y_{kt} - \hat{\mu}_{kmt}}{\hat{\sigma}_{kmt}}, \qquad (2)$$

where y_{kt} is the data, y_{kt}^{S} is the transformed stationary data, and $\hat{\mu}_{kmt}$ and $\hat{\sigma}_{kmt}$ are the sample mean and standard deviation respectively for location k, epoch (e.g. month or season) m and hour t.

Step (b). Transform the data in order to obtain a Gaussian stationary data set:

$$y_{kt}^{GS} = N^{-1}(\hat{F}_k(y_{kt}^{S})), \qquad (3)$$

where y_{kt}^{GS} is the transformed stationary data that follows a Gaussian distribution, $N^{-1}$is the inverse of the cumulative distribution function of the normal distribution and $\hat{F}_k$ is the cumulative function of the (parametric or non-parametric) fit for the data in location k.

Step (c). Use the Yule-Walker equations (Box-Jenkins, 1976) to estimate the autoregressive parameters and covariance matrix of the residual noise obtained from Eq. (1).

2.1.2 Simulation

In order to simulate multi-area wind power production, we assume that the process is driven by an autoregressive `noise' vector. For K locations and p periods of lag the model is:

$$Y_{k,t+1} = \sum_{j=0}^{p} \phi_{kj} Y_{k,t-j} + \omega_{kt}, \qquad (4)$$

where $\Phi = (\phi_{kj})$ is the matrix of autoregressive parameters and ω_{kt} are independent, identically distributed, multivariate Gaussian random variables with mean 0 and covariance matrix Σ. The simulation of the multi-area process can then be summarized in the following steps:

Step (a). Generate autoregressive noise of order p by using the estimated autoregressive parameters and variance.

$$Y^{GS}_{k,t+1} = \sum_{j=0}^{p} \hat{\phi}_{kj} Y^{GS}_{k,t-j} + \omega_{kt}, \qquad (5)$$

where $\omega_{kt} = (\hat{L}\omega)_k$, ω are independent standard normal random vectors with K entries, $\hat{L}$ is the Cholesky factorization of $\hat{\Sigma}$ and Y^{GS}_{kt} is the Gaussian stationary autoregressive process for location k.

Step (b). Transform the resulting process such that it obeys the non-Gaussian distribution of the original stationary data:

$$Y^{S}_{kt} = \hat{F}_k^{-1}(N(Y^{GS}_{kt})) \qquad (6)$$

where Y_{kt}^{S} is the stationary, non-Gaussian process, N is the cumulative distribution function of the normal distribution and $\hat{F}_k^{-1}$ is the inverse of the cumulative function of the data for each location.

Step (c). Transform Y_{kt}^{S} by its seasonal and hourly mean and variance

$$Y_{kt} = \hat{\sigma}_{kmt} Y_{kt}^{S} + \hat{\mu}_{kmt}, \qquad (7)$$

where Y_{kt} is the resulting process that is non-stationary and distributed according to the original data for each location.

Step (d). Use an approximation of the aggregate power curve for each location to simulate wind power production:

$$P_{kt} = \hat{P}_k(Y_{kt}), \qquad (8)$$

where P_{kt} is the simulated wind power production process for each location.

2.2 Results

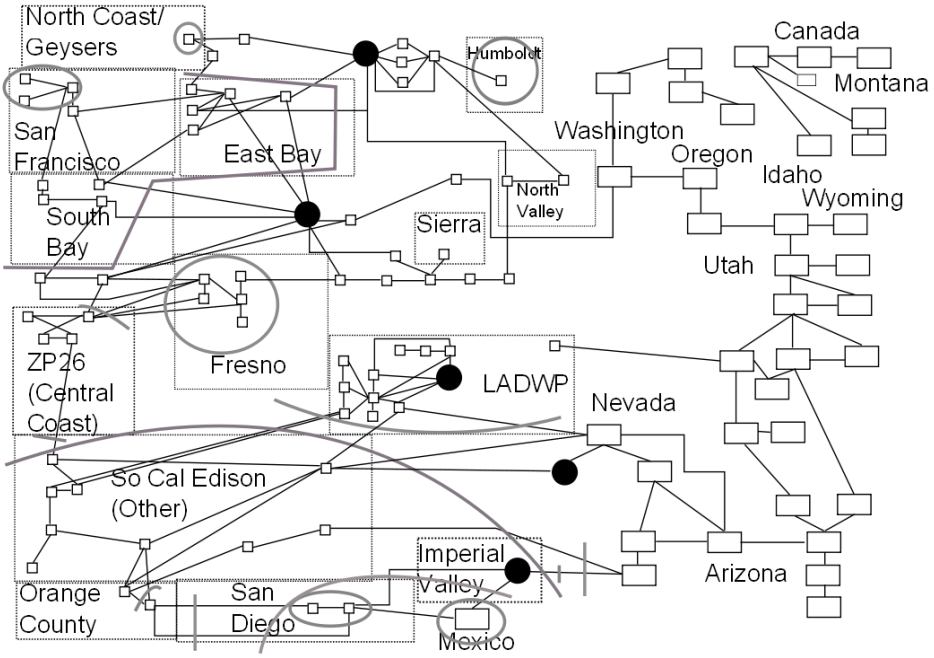


Figure 1: A schematic of the WECC model

County	Existing	Moderate	Deep
Altamont	954	954	1,086
Clark	-	-	1,500
Imperial	-	-	2,075
Solano	348	848	1,149
Tehachapi	1,346	4,886	8,333
Total	2,766	6,688	14,143

Figure 2: Current and projected capacity of wind power installations

We use the multi-area wind production model to study the impacts of large-scale renewable energy integration in the California power system. We use a model of the California ISO with imports from the Western Electricity Coordinating Council (WECC) that is also used by Yu et al. (Yu, 2010) and is described in detail by Papavasiliou et al. (Papavasiliou, 2010).

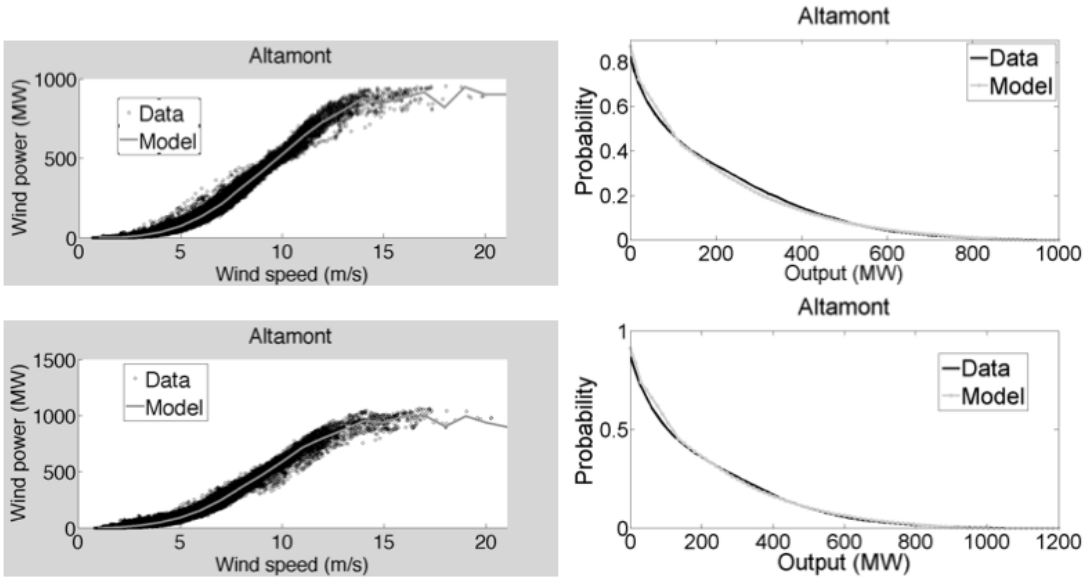


Figure 3: In reading order: power curves (left) and complementary cumulative probability distribution of wind power output (right) for Altamont for the moderate (up) and deep (down) integration study

2.2.1 Data

We use wind speed and wind power production data from the 2006 data set of the National Renewable Energy Laboratory (NREL) Western Wind and Solar Integration

Study (WWSIS), described by Potter et al. (Potter, 2008). We study two wind integration cases. The first represents a moderate energy integration level for wind power corresponding to the 2012 integration target of California, and the second case represents a deep integration level corresponding to the 2020 integration target. Ex post we have estimated that the moderate integration case corresponds to approximately 7% wind energy penetration, while the deep integration case corresponds to approximately 14% wind energy penetration. In the subsequent analysis we will refer to these cases as moderate and deep integration respectively.

In order to collect data for each case, we examined the interconnection queue of the California ISO until 2020, and placed individual wind generators in our model by matching the geographical locations of planned wind power installations with the corresponding wind park data in the WWSIS data set. In Fig. 2 we present the location of existing wind generation capacity, as well as capacity for the moderate and deep integration cases.

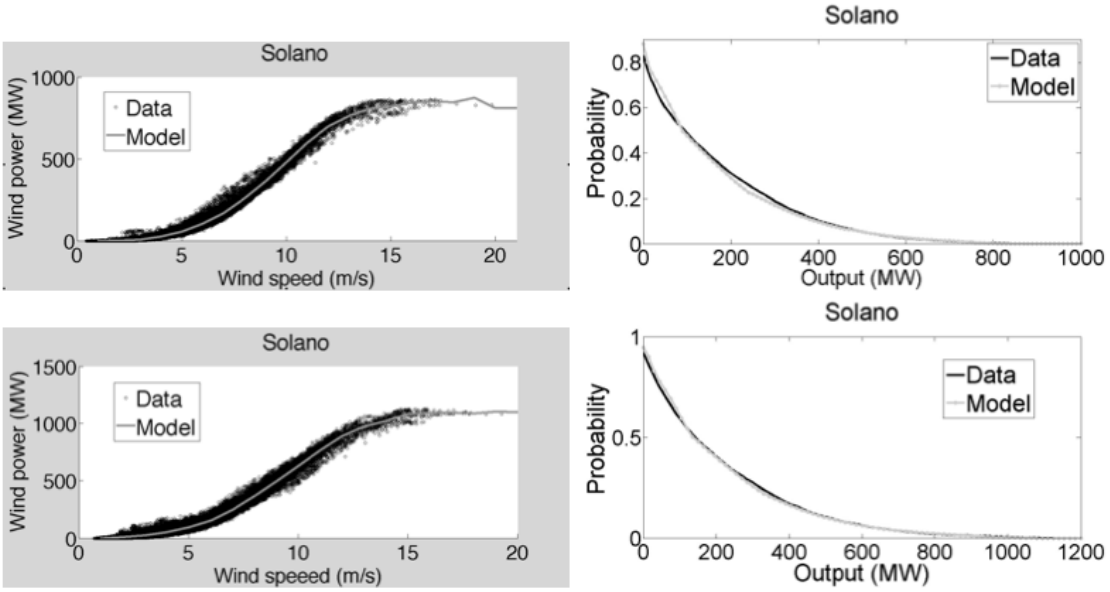


Figure 4: In reading order: power curves (left) and complementary cumulative probability distribution of wind output (right) for Solano for the moderate (up) and deep (down) integration study.

In Fig. 1 we present a schematic diagram of the WECC model. The dashed boxes represent load and generation pockets. The thick solid lines represent the import constraints discussed in the introduction. Each thick solid line intersects a set of transmission lines over which the total amount of power cannot exceed a certain limit.

The wind generators of Fig. 2 are located in the five buses that are depicted as solid black circles. In order of appearance from top to bottom, these wind sites are Solano, Altamont, Tehachapi, Clark and Imperial.

2.2.2 Results

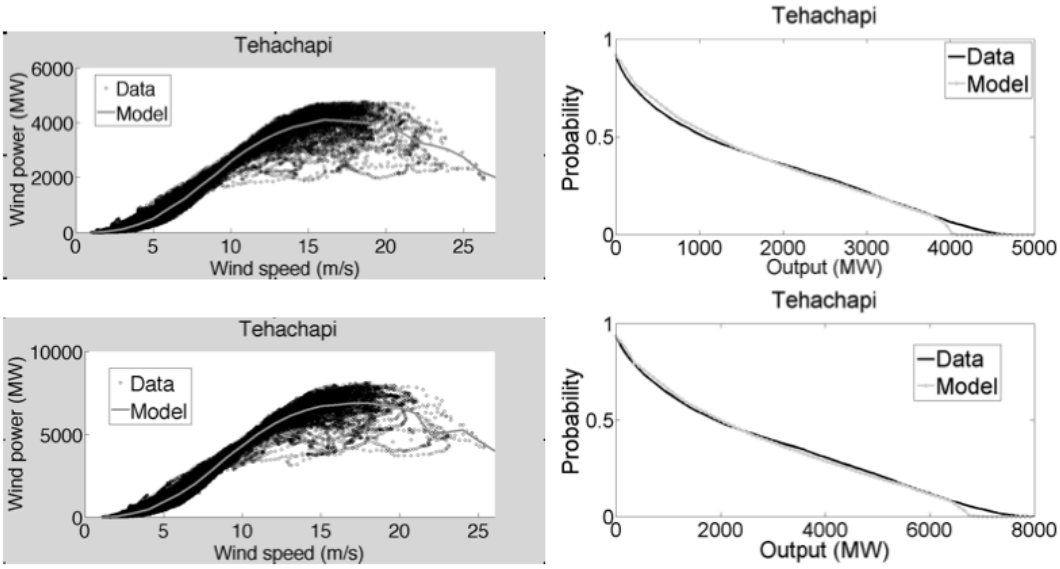


Figure 5: In reading order: power curves (left) and complementary cumulative probability distribution of wind output (right) for Tehachapi for the moderate (up) and deep (down) integration study.

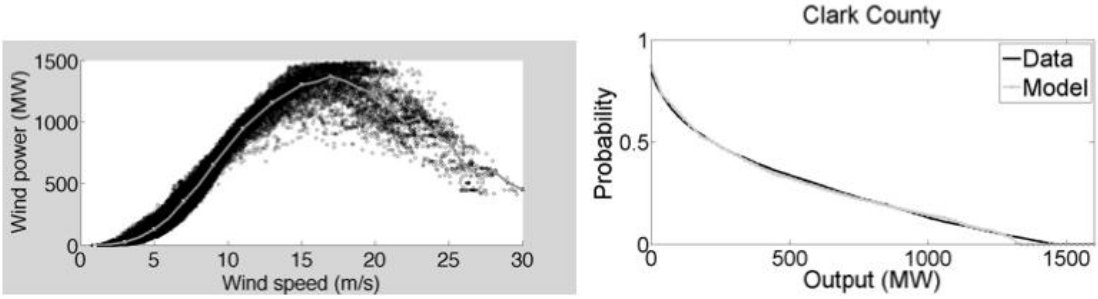


Figure 6: In reading order: power curves (left) and complementary cumulative probability distribution of wind output (right) for Clark County for the deep integration study.

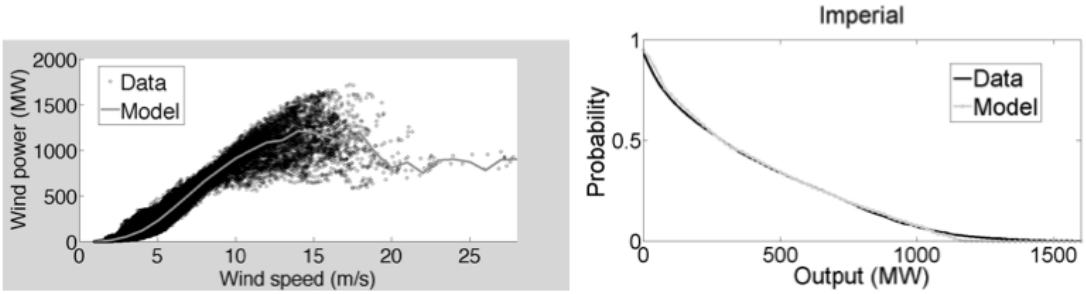


Figure 7: In reading order: power curves (left) and complementary cumulative probability distribution of wind output (right) for Imperial Valley for the deep integration study.

In Fig. 3 we present the approximate power curve and the fit of the complementary cumulative probability distribution of wind output to the data for the Altamont area for both integration studies. The corresponding results for Solano and Tehachapi are presented in Figs. 4, 5 respectively.

In Figs. 6, 7 we present results for the deep integration case for Clark County and Imperial Valley respectively.

As the figures indicate, the primary source of discrepancy in the model is the approximate power curve. Note that the complementary cumulative probability distribution deviates from the data only for high wind output levels for the Tehachapi area, and to a lesser extent for the Solano area. From the power curve of the Tehachapi area we note that the scatter plot of wind speed to wind power exhibits a significant spread. This is due to the fact that Tehachapi covers a wide geographic area with wind parks located in most regions of the area. As a result, the power curve cannot reproduce the high-power results observed in the data. %The reason is that Tehachapi covers a wide geographic area, making average wind speed an inaccurate proxy for the actual wind speed in each sub-region of the Tehachapi area, while at the same time most of the wind capacity of the state is installed in this area which means that most of the locations of this region had wind generators installed and therefore what wind existed there mattered.

In order to alleviate this problem, we experimented with further partitioning the Tehachapi area in smaller regions. However, this introduced greater inaccuracy to the model due to the higher dimensions of the correlation matrix Σ. As a result, we chose to model five areas as the best compromise between capturing locational dependencies and retrieving marginal wind speed distributions at each location. Similarly, the Solano area exhibits a noticeable spread in the scatter diagram between wind speed and wind power production.

The aforementioned drawback is acceptable in the context of unit commitment studies of wind integration. Wind power variability affects reserve requirements due to the fact that wind power production often reaches a near-zero level for extended periods of time. We note from the complementary cumulative probability distribution of wind output that the behavior of wind power production is accurately depicted at low wind production levels.

Wind production ramping also has the potential of affecting reserve requirements. Our model accounts for the inter-temporal fluctuations of wind power supply by isolating monthly and diurnal patterns and by using a time series model for wind speed.

3. Integration of Contracted Renewable Energy and Spot Market Supply to Serve Flexible Loads

3.1 Model

In this section we formulate the optimal control problem that a renewable power supplier faces under our proposed contract. The supplier has the task of serving a flexible consumer from a freely available renewable resource which is backed up by a spot market for electricity. Wind generation and the spot price of electricity are driven by two correlated mean-reverting Ornstein-Uhlenbeck processes. The aggregator procures energy from the spot market with the objective to minimize the cost of unserved energy and expenditures in the spot market. We solve the optimal control problem by using dynamic programming. In particular, we use recombinant lattices for modeling the electricity prices and wind power supply, following the methodology which is outlined in (Deng, 2003) and (Deng, 2005).

3.1.1 Price and wind models

We consider 8 day types in our analysis, which represent weekdays and weekends for each season. We use the Weibull distribution for modeling wind power generation, and the lognormal distribution for modeling prices. Both processes are driven by an underlying first-order autoregressive process which obeys the following dynamic model:

$$\begin{aligned} n_{t+1}^{s} &= n_t^{s} + \kappa^{s}(\theta^{s} - n_t^{s})\Delta t + \sigma^{s}\omega_1\sqrt{\Delta t} \\ n_{t+1}^{\lambda} &= n_t^{\lambda} + \kappa^{\lambda}(\theta^{\lambda} - n_t^{\lambda})\Delta t + \rho\sigma^{\lambda}\omega_1\sqrt{\Delta t} + \\ &\quad \sqrt{(1-\rho^2)}\sigma^{\lambda}\omega_2\sqrt{\Delta t}, \end{aligned} \tag{9}$$

where n_t^s and n_t^λ are the noise terms of the wind and price models respectively, ω_1 and ω_2 are independent standard normal random variables, Δt is the time step, θ^s and θ^λ represent the average trends of the wind and price noise respectively, the variance terms σ^s and σ^λ capture the effect of random shocks, κ^s and κ^λ model the rate at which the processes return to their mean value and ρ is a correlation coefficient which couples the evolution of the two processes.

Wind power supply is strongly influenced by seasonal and diurnal effects. (Most, 2009) discuss the need to remove these deterministic effects from the data in order to obtain the residual process which can be used for calibrating the parameters of the mean-reverting process. The seasonal and diurnal patterns of wind generation are captured by μ_t^s, the average value of hourly wind production for each day type. We assume that the ratio $\frac{s_t}{\mu_t^s}$ follows the Weibull distribution with parameters k, λ. Following Eq. 1 of (Brown, 1984), section 2.1 of (Torres, 2005) and Eq. 2 of (Morales, 2010) for transforming Weibull-

distributed data to Gaussian data, as well as the nonparametric transformation that is used in Eqs. 8, 9 of (Callaway, 2010), we use the inverse transform sampling method to obtain wind power, s_t, as a function of the underlying noise, n_t^s, and the real-time price signal λ_t as a function of n_t^λ:

$$\begin{aligned} s_t &= \mu_t^s \lambda \exp\left(k^{-1} \log\log(1 - N(n_t^s))\right) \\ \lambda_t &= \mu_t^\lambda \exp(n_t^\lambda), \end{aligned} \tag{10}$$

where s_t is the value of wind generation, $\lambda_t$is the real time electricity price, $\mu_t^s$and μ_t^λ are the average values of hourly wind production and hourly real-time electricity prices respectively, and N is the cumulative distribution function of the normal distribution.

3.1.2 Problem formulation

The dynamic optimization problem has a three-dimensional state vector, $x_t = (\lambda_t, s_t, r_t)$, where λ_t is the spot price of electricity, s_t is the amount of wind which is freely available and r_t represents the remaining quantity of demand. The residual energy r_t evolves according to $r_{t+1} = r_t - u_t \Delta t$, where u_t, our control, is the amount of power supplied to the consumer in period t. We model the two-dimensional stochastic process (λ_t, s_t) with a trinomial recombinant lattice model. In particular, the dynamics of the underlying process are assumed to obey the following:

$$n_{t+1}^{s,j} = \begin{cases} n_t^s + \sigma^s \sqrt{\frac{3}{2}} \sqrt{\Delta t}, & j = 1 \\ n_t^s, & j = 2 \\ n_t^s - \sigma^s \sqrt{\frac{3}{2}} \sqrt{\Delta t}, & j = 3 \end{cases}$$

$$n_{t+1}^{\lambda,j} = \begin{cases} n_t^\lambda + (\sqrt{3}\rho + \sqrt{1-\rho^2}) \sigma^\lambda \sqrt{\frac{\Delta t}{2}}, & j = 1 \\ n_t^\lambda - \sigma^\lambda \sqrt{1-\rho^2} \frac{2}{\sqrt{2}} \sqrt{\Delta t}, & j = 2 \\ n_t^\lambda - (\sqrt{3}\rho - \sqrt{1-\rho^2}) \sigma^\lambda \sqrt{\frac{\Delta t}{2}}, & j = 3. \end{cases} \tag{11}$$

Each state j is visited with a probability $p^j(n_t^\lambda, n_t^s)$ which depends on the current state. The transition probabilities are defined in (Deng, 2005).

The use of a recombinant lattice is motivated by computational efficiency. Note from the dynamics of Eqs. (11) that a given point in the state space in period t+1 can be visited by various points in the state space in period t. In effect, it is shown in (Deng, 2003) that the size of the state space grows quadratically in the horizon of the problem, which is to be contrasted to an exponential growth in the size of the state space when a trinomial lattice does not recombine. As a result, we are able to control the growth of the state space, and this enables us to solve the optimal control problem using the dynamic programming algorithm. (Deng, 2003) also prove that the transition probabilities and state space can be constructed such that, as Δt converges to zero, the discrete process converges in distribution to the continuous time mean reverting Ornstein-Uhlenbeck process.

In order to calibrate the wind model parameters, we used the 2006 National Renewable Energy Laboratory (NREL) database for an integration level of 14,143 MW in California. The upper frame of Fig. 8 shows the sample cumulative distribution function of the trinomial wind generation model overlaid on the sample cumulative density function of the wind dataset. The lower frame of Fig. 8 shows the cumulative distribution function of the electricity price lattice model.

The objective of the optimal control problem is to minimize the following expected cost:

$$\min_{\phi_t(x_t)} \mathbb{E}[\sum_{t=1}^{N} \lambda_t(\mu_t(x_t) - s_t)^+ \Delta t + \rho r_N], \tag{12}$$

where $\phi_t(x)$ represents the rate at which the resource is supplied and N is the number of periods. The state vector has the following initial condition: $r_1 = R$, where R is the amount of energy demand to be satisfied. The control u_t cannot exceed an upper bound on the rate of supply, $u_t \leq C_p$. In the case where power is supplied to n electric vehicle batteries, C_p equals nC, where C is be the nominal rating of an electric vehicle battery. We also limit the amount of energy that can be procured in the real-time electricity market at each period to C_m by introducing the constraint $u_t - s_t \leq C_p$ in order to transfer the risk of wind power variability from the system operator to the aggregator.

Since the coupled system may rely only up to C_m on system reserves, there is a possibility that residual demand may not be fully satisfied. Unsatisfied energy incurs a penalty ρ.

Finally, we denote as K the nameplate capacity of renewable power resource.

3.2 Results

3.2.1 Data

As we mentioned in section 3.1, we have estimated parameters for eight day types, corresponding to weekdays and weekends of each season.

The wind data used in this study is sourced from the NREL 2006 western wind database. The locations of the wind generation sites that are used for the study represent an integration target of 14,143 MW, based on the data presented in the CAISO report by (CAISO, 2007) and the 2010 California generation interconnection queue.

There is a total of 2,648 MW of wind power currently connected in the California system. The locations of the wind sites are presented in Fig. 2. Wind data was sampled from the NREL database according to the locations that are described in Fig. 2.

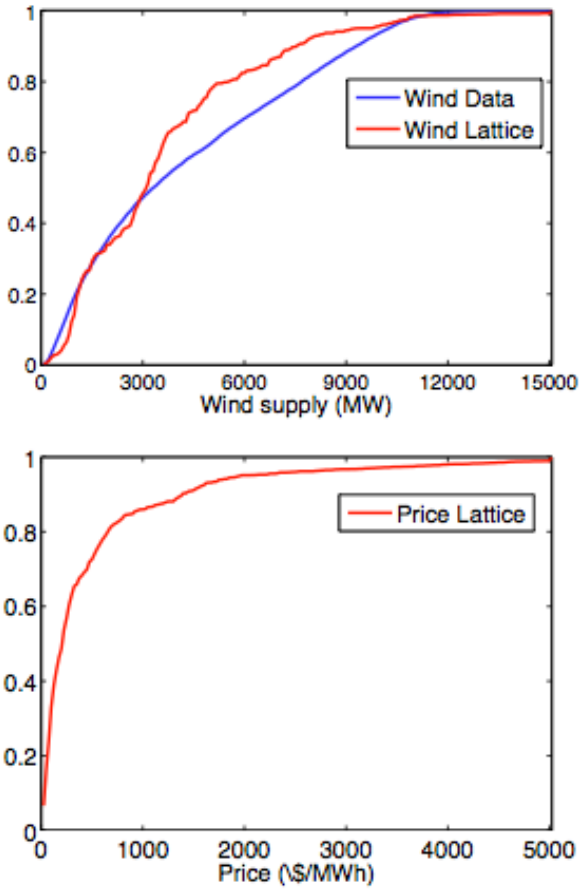


Figure 8: Probability density function of wind output (top) and real time electricity prices (bottom) for deep wind integration case.

We solve the problem for 24 hourly intervals. We consider 6 levels of charge for the control problem. The baseline level of wind integration is K = 14,143 MW, corresponding to the 33% California renewable integration target, and the baseline capacity constraint of consumers is $C_p = 15000$ MW. We also impose a limit of $C_m = 2000$ MW in spot market participation. We have chosen C_m to be a small fraction of C_p in order to test our intuition that deferrable energy requests couple well with renewable power supply over an extended time horizon. If the coupled system were required to resort primarily to the real time market in order to perform adequately, then there is little reason to consider coupling contracts as a means of utilizing demand flexibility to absorb renewable supply fluctuations. The baseline level of total energy demand is R = 80,000 MWh, which is 80 percent of the average daily wind output in the 33% wind integration target. The requests span over 24 hours. If we assume that a typical electric vehicle has a power rating of 3.6 kW and a mileage of 0.25 kWh per mile, the baseline demand model roughly represents the electricity demand of 4.167 million electric vehicles which travel 96 miles per vehicle per day. The cost of unserved energy in the baseline is $\rho = 5{,}000$ \$/MWh. This value is selected as an estimate of the average cost of not serving flexible energy requests for vehicle charging.

3.2.2 Relative performance of policies

We now compare the performance of the dynamic programming policy, the clairvoyant policy which has advance knowledge of the outcome in each realization, a naive charging policy whereby consumers are served as fast as possible, and a model predictive control policy.

The model predictive controller applies the optimal solution in step k to the optimal control problem that results if the stochastic processes were to follow their unperturbed dynamics for the remaining horizon.

The real-time market expenses of all policies are compared in Fig. 9 for the baseline scenario. The second column represents the cost of the clairvoyant policy in bold type. The other columns present the relative performance of the other three policies with respect to the clairvoyant policy. The results are then averaged according to the frequency of each day type, and the relative real-time market costs of each policy are presented in both dollar figures as well as a percentage of the cost of the clairvoyant policy in the last two rows of the table. The overall performance of all policies, including the penalty of unserved energy, is compared in Fig. 10. As in Fig. 9, the results are presented relative to the clairvoyant policy. We observe that although model predictive control performs better than the naive policy in terms of real-time market expenditures, overall it performs worse than the naive policy due to the fact that it incurs high penalties for unserved energy.

	Cost ($) Clair	Δ Cost ($) Naive	Δ Cost ($) MPC	Δ Cost ($) Dyn Prog
WinterWD	**164,420**	1,325,680	165,580	241,640
SpringWD	**85,289**	952,011	111,611	167,471
SummerWD	**2,157,500**	976,100	467,400	321,900
FallWD	**1,001,900**	1,348,500	599,000	462,000
WinterWE	**169,250**	1,453,850	177,570	265,280
SpringWE	**96,988**	1,102,712	150,912	177,532
SummerWE	**1,770,200**	1,093,100	455,700	305,600
FallWE	**1,132,200**	1,612,000	693,500	511,800
Total	**835,086**	1,197,671	345,474	303,053
improv. (%)		143.4	41.4	36.3

Figure 9: Cost of procuring electricity from the real-time market (baseline scenario).

	Cost ($) Clair	Δ Cost ($) Naive	Δ Cost ($) MPC	Δ Cost ($) Dyn Prog
Market	**835,086**	1,976,671	345,474	303,053
Unserved	**436,161**	0	1,433,482	12,500
Total	**1,271,247**	1,197,671	1,778,956	315,553
improv. (%)		94.21	139.9	24.8

Figure 10: Relative performance of policies (baseline scenario).

3.2.3 Sensitivity on load capacity (C_p)

The sensitivity of the results on the capacity of the load are presented in Fig. 11 (left). The first frame describes the relative real-time market expenses of the three policies. Beyond $C_p = 10{,}000$ MW the system is not achieving any significant gains by further increasing load capacity, since it is relatively infrequent that wind generation exceeds this level.

The second frame presents the percentage of wind that is shed on average.

The naive policy places an upper bound on the amount of wind shedding, while the clairvoyant policy places a lower bound since it relies on wind power as much as possible. The last frame presents the mix that is used for serving load for the case of the dynamic programming policy. As C_p increases, spot market procurements are replaced by

freely available wind power supply. It is notable that unserved energy is negligible even for $C_p = 5{,}000$ MW.

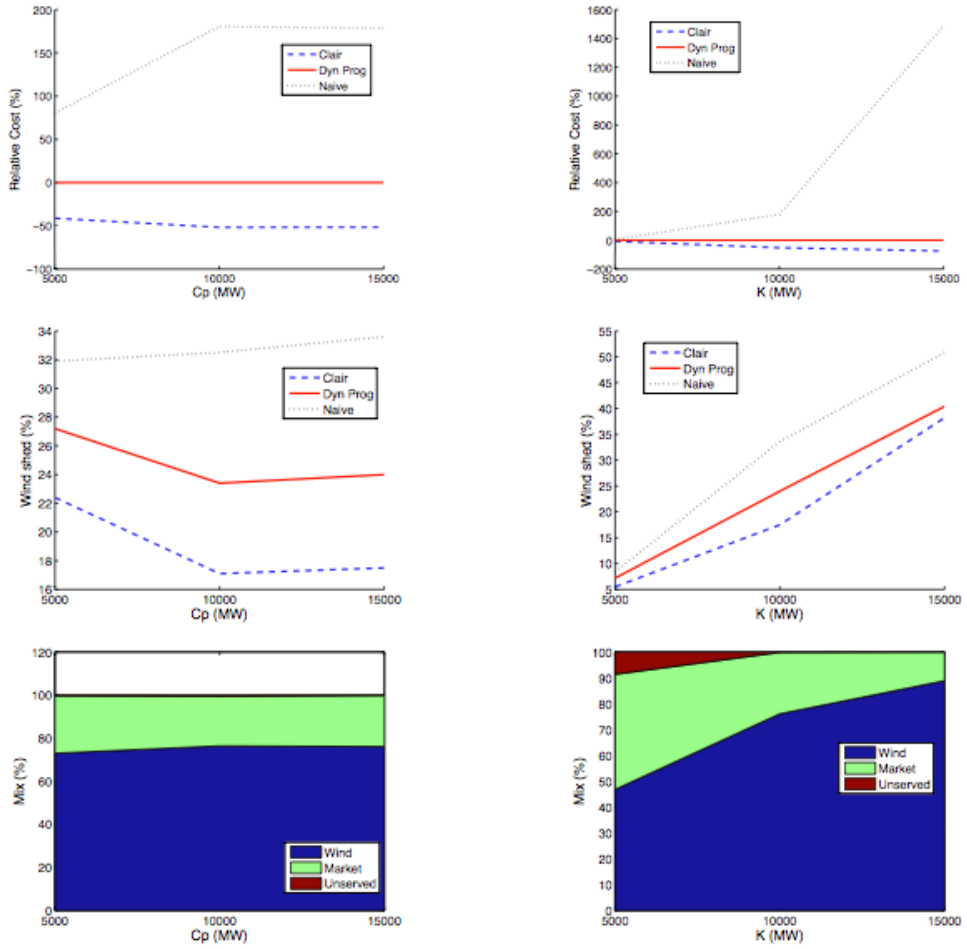


Figure 11: Sensitivity results for varying C_p (left) and K (right).

3.2.4 Sensitivity on wind power capacity (K)

In Fig. 11 (right) we present the sensitivity analysis results for varying K. We note that there is a significant amount of wind shedding for high K, with the naive policy becoming highly inefficient due to the fact that it is not taking advantage of the excess supply of wind. We also note that there is a significant amount of unserved energy for the case of low K. The first frame shows that all policies incur almost the same cost at the real-time market, because it is almost always optimal to buy as much as possible from the real-time market.

3.2.5 Sensitivity analysis on spot market participation (C_m)

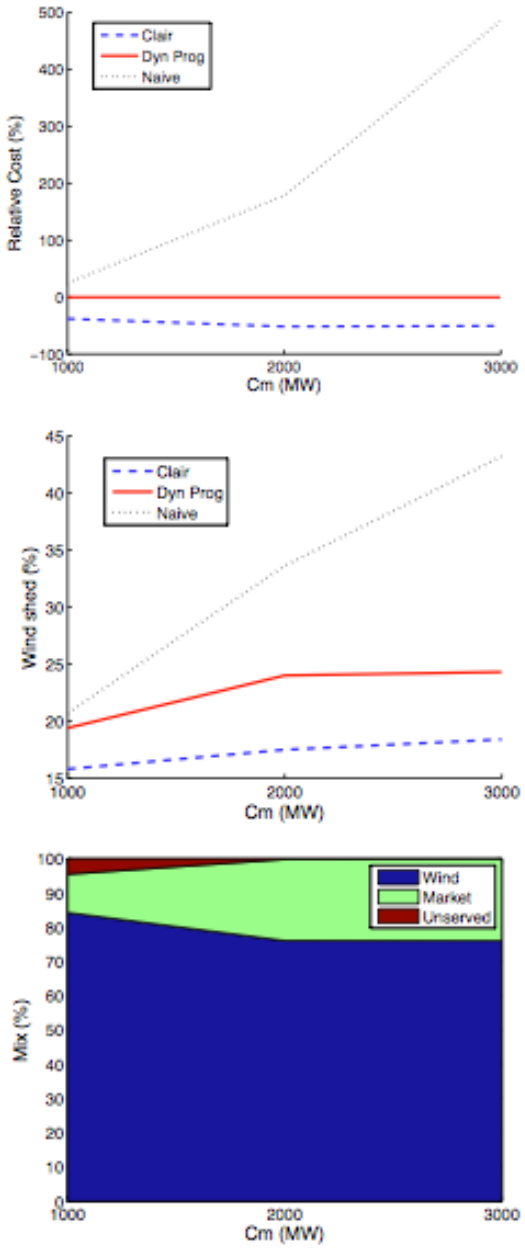


Figure 12: Sensitivity results for varying C_m.

The sensitivity analysis on the degree of spot market participation is shown in Fig. 12. From the first frame we observe that despite the fact that C_m increases, the performance of the clairvoyant does not improve relative to the performance of the dynamic programming policy. As C_m increases, the naive policy becomes remarkably inefficient due to exposure in the spot market. In the second frame we note that as C_m increases the

naive policy procures more power from the market and uses less wind. In the last frame we observe that for $C_m = 1{,}000$ MW a notable amount of load is unserved, while as C_m increases wind is replaced by market procurements and unserved energy diminishes.

3.2.6 Sensitivity on cost of unserved energy (ρ)

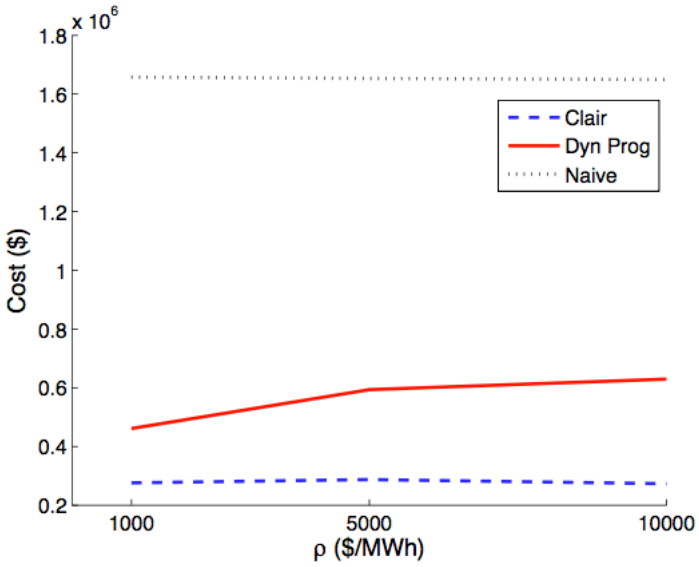


Figure 13: Sensitivity results for varying ρ.

In Fig. 13 we present the results of the sensitivity analysis on ρ. In this figure we present dollar amounts of real-time market expenses, instead of expenses relative to the dynamic programming policy. The naive policy remains insensitive to ρ, whereas the dynamic programming policy incurs greater expenses in the real-time market as ρ increases. The real-time market expenses of the clairvoyant policy are insensitive to ρ. We do not present sensitivity results on the amount of wind shedding, and the mix which is used to satisfy demand since these results are insensitive to ρ. It is noted that for the minimum price of unserved energy, $\rho = 1{,}000$ \$/MWh, the clairvoyant policy deliberately leaves load unserved for some of the day types as this is more economical than buying from the market during periods of high prices, whereas for higher values of ρ the clairvoyant policy never leaves load unserved unless the other constraints in the problem force such an outcome. Moreover, for $\rho = 1{,}000$ \$/MWh the dynamic programming policy has 348 MWh of unserved load, approximately three times greater than the amount of unserved load for $\rho = 5{,}000$ \$/MWh, which is equal to 112 MWh.

4. Conclusions

We have presented a stochastic model of multi-area wind production that can be used in stochastic unit commitment studies of renewable energy integration. We fit a time series model of wind speed and use a piecewise linear approximation of the regional power curve to simulate wind power. We account for monthly and diurnal patterns of wind speed and use a time series model for reproducing temporal correlation. We represent spatial correlation by introducing a correlation matrix in the noise that drives the vector autoregressive process. We present simulation results for two wind integration studies of the California power system that correspond to the wind integration targets of 2012 and 2020. We observe that the fitness of the model to the data depends largely on the accuracy of the piecewise linear approximation of the power curve. Increasing the number of regions improves the accuracy of the power curve approximation, at the cost of increasing the size of the correlation matrix that drives the wind speed process.

We have proposed a contract for utilizing renewable generation to mitigate the impacts of renewable power variability and unpredictability. The contract transfers the risk of not serving load from the system operator to consumers, and results in an optimal control problem in which an aggregator seeks to optimize the extent to which loads are backed up by a volatile spot market for electricity. We compare four policies for serving flexible loads, the dynamic programming policy, the clairvoyant policy, a naive policy and a model predictive control policy. For the baseline scenario we find that the model predictive control policy incurs lower real-time market costs than the naive policy, however it is not able to outperform the naive charging policy overall, due to large penalties for unserved load. We also present sensitivity results on real-time market expenditures, wind power utilization, and the optimal mix for serving energy requests with respect to various problem parameters.

References

[1] P. A. Ruiz, R. C. Philbrick, and P. W. Sauer, "Wind power day ahead uncertainty management through stochastic unit commitment policies," in Power Systems Conference and Exposition, March 2009, pp. 1–9.

[2] J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Transactions on Power Systems,* vol. 23, no. 3, pp. 1319–1327, August 2008.

[3] E. M. Constantinescu, V. M. Zavala, M. Rocklin, S. Lee, and M. Anitescu, "A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation," *IEEE Transactions on Power Systems,* vol. 26, no. 1, pp. 431–441, February 2011.

[4] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, "Unit commitment for systems with high wind penetration," *IEEE Transactions on Power Systems,* vol. 24, no. 2, pp. 592–601, May 2009.

[5] J. M. Morales, A. J. Conejo, and J. Perez-Ruiz, "Economic valuation of reserves in power systems with high penetration of wind power," *IEEE Transactions on Power Systems,* vol. 24, no. 2, pp. 900–910, May 2009.

[6] F. Bouffard and F. D. Galiana, "Stochastic security for operations planning with significant wind power generation," *IEEE Transactions on Power Systems,* vol. 23, no. 2, pp. 306–316, May 2008.

[7] A. Papavasiliou, S. S. Oren, and R. P. O'Neill, "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," *IEEE Transactions on Power Systems,* vol. 26, no. 4, pp. 2197–2206, November 2011.

[8] J. M. Arroyo and F. D. Galiana, "Energy and reserve pricing in security and network-constrained electricity markets," *IEEE Transactions on Power Systems,* vol. 20, no. 2, pp. 634–643, May 2005.

[9] F. D. Galianaa, F. Bouffard, J. M. Arroyo, and J. F. Restrepo., "Scheduling and pricing of coupled energy and primary, secondary, and tertiary reserves," *Proceedings of the IEEE,* vol. 93, no. 11, pp. 1970–1983, November 2005.

[10] F. Bouffard, F. D. Galiana, and A. J. Conejo., "Market-clearing with stochastic security," *IEEE Transactions on Power Systems,* vol. 20, no. 4, pp. 1827–1835, November 2005.

[11] B. G. Brown, R. W. Katz, and A. H. Murhpy, "Time series models to simulate and forecast wind speed and wind power," *Journal of Climate and Applied Meteorology,* vol. 23, pp. 1184–1195, 1984.

[12] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day, 1976.

[13] J. L. Torres, A. Garcia, M. D. Blas, and A. D. Francisco, "Forecast of hourly wind speed with ARMA models in Navarre (Spain)," *Solar Energy*, vol. 79, no. 1, pp. 65–77, July 2005.

[14] J. M. Morales, R. Minguez, and A. J. Conejo, "A methodology to generate statistically dependent wind speed scenarios," *Applied Energy*, vol. 87, pp. 843–855, 2010.

[15] D. Callaway, "Sequential reliability forecasting for wind energy: Temperature dependence and probability distributions," *IEEE Transactions on Energy Conversion*, vol. 25, pp. 577–585, June 2010.

[16] N.-P. Yu, C.-C. Liu, and J. Price, "Evaluation of market rules using a multi-agent system method," *IEEE Transactions on Power Systems*, vol. 25, pp. 470–479, February 2010.

[17] C. W. Potter, D. Lew, J. McCaa, S. Cheng, S. Eichelberger, and E. Grimit, "Creating the dataset for the western wind and solar integration study (U.S.A.)," in *7th International Workshop on Large Scale Integration of Wind Power and on Transmission Networks for Offshore Wind Farms*, Madrid, Spain, May 2008.

[18] CAISO, "The California ISO controlled grid generation queue as of January 8, 2010," 2010.

[19] Borenstein, S., Jaske, M., and Rosenfeld, A. (2002). *Dynamic pricing, advanced metering and demand response in electricity markets*. Technical report, University of California Energy Institute.

[20] CAISO (2010). The California ISO controlled grid generation queue as of January 8, 2010. URL http://www. caiso.com/14e9/14e9ddda1ebf0.pdf.

[21] Deng, S.J. and Oren, S.S. (2003). Incorporating operational characteristics and start-up costs in option-based valuation of power capacity. *Probability in the Engineering and Informational Sciences*, 17, 155–181.

[22] Deng, S.J. and Oren, S.S. (2005). Applications of Stochastic Programming, chapter 31, 655–667. *MPS-SIAM Series on Optimization*.

[23] Eto, J. (2007). *Demand response spinning reserve demonstration*. Technical report, Lawrence Berkeley National Laboratory.

[24] Hirst, E. and Kirby, B. (1997). Ancillary-service details: Dynamic scheduling. Technical report, Oak Ridge National Laboratory.

[25] Hirst, E. and Kirby, B. (1999). *Load as a resource in providing ancillary services*.

Technical report, Oak Ridge National Laboratory.

[26] Kirby, B.J. (2003). Spinning reserve from responsive loads. Technical report, Oak Ridge National Laboratory.

[27] Loutan, C. and Hawkins, D. (2007). *Integration of renewable resources. transmission and operating issues and recommendations for integrating renewable resources on the California ISO-controlled grid*. Technical report, California Independent System Operator.

[28] Most, D. and Keles, D. (2009). A survey of stochastic modeling approaches for liberalised electricity markets. *European Journal of Operations Research*, 207(2), 543– 556.