Market-Based Demand Response Integration in Super-Smart Grids in the Presence of Variable Renewable Generation

by

Sahand Behboodi
B.Sc., Amirkabir University of Technology, 2009
M.Sc., Power and Water University of Technology, 2012

A Dissertation Submitted in Partial Fulfilment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Mechanical Engineering

Market-Based Demand Response Integration in Super-Smart Grids in the Presence of Variable Renewable Generation

by

Sahand Behboodi
B.Sc., Amirkabir University of Technology, 2009
M.Sc., Power and Water University of Technology, 2012

Supervisory Committee

Dr. C. Crawford, Supervisor
(Department of Mechanical Engineering)

Dr. N. Djilali, Departmental Member
(Department of Mechanical Engineering)

Dr. P. Agathoklis, External Member
(Department of Electrical and Computer Engineering)

Supervisory Committee

Dr. C. Crawford, Supervisor
(Department of Mechanical Engineering)

Dr. N. Djilali, Departmental Member
(Department of Mechanical Engineering)

Dr. P. Agathoklis, External Member
(Department of Electrical and Computer Engineering)

ABSTRACT

Variable generator output levels from renewable energies is an important technical obstacle to the transition from fossil fuels to renewable resources. *Super grids* and *smart grids* are among the most effective solutions to mitigate generation variability. In a super grid, electric utilities within an interconnected system can share generation and reserve units so that they can produce electricity at a lower overall cost. Smart grids, in particular *demand response* programs, enable flexible loads such as plug-in electric vehicles and HVAC systems to consume electricity preferntially in a grid-friendly way that assists the grid operator to maintain the power balance. These solutions, in conjunction with energy storage systems, can facilitate renewable integration.

This study aims to provide an understanding of the achievable benefits from integrating demand response into wholesale and retail electricity markets, in particular in the presence of significant amounts of variable generation. Among the options for control methods for demand response, market-based approaches provide a relatively efficient use of load flexibility, without restricting consumers' autonomy or invading their privacy. In this regard, a model of demand response integration into bulk electric grids is presented to study the interaction between variable renewables and demand response in the double auction environment, on an hourly basis. The cost benefit

analysis shows that there exists an upper limit of renewable integration, and that additional solutions such as super grids and/or energy storage systems are required to go beyond this threshold.

The idea of operating an interconnection in an unified (centralized) manner is also explored. The traditional approach to the unit commitment problem is to determine the dispatch schedule of generation units to minimize the operation cost. However, in the presence of price-sensitive loads (market-based demand response), the maximization of *economic surplus* is a preferred objective to the minimization of cost. Accordingly, a surplus-maximizing hour-ahead scheduling problem is formulated, and is then tested on a system that represents a 20-area reduced model of the North America Western Interconnection for the planning year 2024. The simulation results show that the proposed scheduling method reduces the total operational costs substantially, taking advantage of renewable generation diversity.

The value of demand response is more pronounced when ancillary services (e.g. real-time power balancing and voltage/frequency regulation) are also included along with basic temporal load shifting. Relating to this, a smart charging strategy for plug-in electric vehicles is developed that enables them to participate in a 5-minute retail electricity market. The cost reduction associated with implementation of this charging strategy is compared to uncontrolled charging. In addition, an optimal operation method for thermostatically controlled loads is developed that reduces energy costs and prevents grid congestion, while maintaining the room temperature in the comfort range set by the consumer. The proposed model also includes loads in the energy imbalance market.

The simulation results show that market-based demand response can contribute to a significant cost saving at the sub-hourly level (e.g. HVAC optimal operation), but not at the super-hourly level. Therefore, we conclude that demand response programs and super grids are complementary approaches to overcoming renewable generation variation across a range of temporal and spatial scales.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

First and utmost, I would like to express my deepest gratitude to my supervisor, Dr. Curran Crawford for his immeasurable help and support throughout my study and research at the University of Victoria. I will always be grateful to him for his mentorship. Prof. Ned Djilali has also been an invaluable source of guidance and advice. His insightful discussions, suggestions and corrections were crucial in enhancing the quality of this work.

I am greatly indebted to my good friend David Chassin, whose encouragement and guidance enabled me to develop an understanding of the subject. He showed me how to think and question everything, that led to amazing years of research for me. I would also like to thank Camille Israel, my wonderful girlfriend, for helping me with the grammar.

A special appreciation is due to so many of my friends and colleagues at the Institute for Integrated Energy Systems (IESVic), for helping to foster a collaborative environment of research and learning. Thank you all for your support, enthusiasm, and encouragement.

The financial support of the Natural Resource Canada (NRCan) and the University of Victoria is gratefully acknowledged.

To my beloved parents

Chapter 1

Introduction

1.1 Background and Motivation

Inherent generation variability[1] is an important technical barrier to the transition from fossil fuels to renewable energies such as wind and solar power [2, 3]. Generation variation will break the balance of supply and demand, bringing risk to the entire electric system. In a typical system, the ramp up/down capability of base load power plants (nuclear, coal-fired and combined cycle) is not sufficient to mitigate renewable variability. Also, operating additional reserves to back up variable generation resources is often too costly. The high overall cost of renewable energies often limits the transition, despite great socio-economic benefits of these clean resources.

Super grids [4, 5] and smart grids, in particular demand response programs [6, 7], are among the effective solutions to overcome renewable generation intermittency. A super grid is an interconnected system, often at the continent scale, that ties together a number of control areas so that they can share generation and reserve units [8, 9] e.g., the European super grid [10]. Using system interties, control areas can accommodate generation and load fluctuations at a lower overall cost [11, 12]. A demand response program motivates changes in electricity use by customers through changes in the price of electricity over time, or through incentive payments at times of high market prices or when grid reliability is jeopardized [13, 14].

In an interconnection, control areas most often exchange electricity based on long-term bilateral contracts. If control areas set the import/export flows 24 hours ahead of operation, and then reset them one hour ahead according to the real-time sys-

[1] Variability is the extent to which a power source may exhibit undesired or uncontrolled changes in output [1].

tem condition, costs associated with the exchange would decrease and/or revenue would increase, because the hour-ahead forecast is more accurate than the day-ahead forecast. A centralized approach to the resource allocation process (import/export optimal schedule through system tielines) would also reduce the combined operation cost [15, 16]. Upgrading monitoring and controlling devices would allow the implementation of such improved operation methods, facilitating renewable integration.

Modernizing the electric system would accelerate the transition to renewable energies. Recent advances in information technology enables smart grids to effectively control distributed resources (generation, load and storage) that can potentially result in lowering operational costs and increasing grid reliability [17, 18]. Flexible resources such as plug-in electric vehicles [19, 20] and HVAC loads [21, 22] can provide ancillary services (e.g. energy balancing [23, 24] and frequency and voltage regulation [25, 26]) to the power grid. Demand response behaves very much like fast-acting generators when it is enabled with the appropriate automation technology [27, 28]. For this reason, demand response is sometimes referred to as virtual power generation [29, 30].

This study aims to provide a better understanding of the benefits of introducing super grid and demand response solutions to electric systems with a significant amount of intermittent renewables. As stated before, these solutions along with energy storage systems can facilitate renewable integration to a great extent. This dissertation investigates the concept of super grids in the presence of variable generation and demand response at the wholesale market level in Chapters 2–4, and explores load control methods at the retail market level in Chapters 5 and 6.

To explain the use of markets to determine optimal resource allocation (the lowest operational cost), it is important to understant the impact of energy market deregulation. There are two kinds of electricity markets: regulated and deregulated. In regulated markets, the utility sets the prices for electricity supply (typically overseen by an energy regulator, such as the BCUC overseeing BC Hydro), along with the associated transportation and distribution costs. Utilities are granted a monopoly in exchange for foregoing the ability to set prices. Consumers therefore have no choice when it comes to their electricity provider. In deregulated markets, electricity is a commodity capable of being bought, sold, and traded at current and future times [31]. Producers compete to sell electricity to consumers, which in theory leads to lower overall prices to consumers by giving them the opportunity to search for the best deal. Accordingly, deregulated markets set the price of electricity in accordance to the supply-demand balance, which also theoretically gives the most economically-efficient allocation of resources. The detailed market structures, in terms of rules, temporal breakdown and ancillary markets vary widely around the world, and can lead to more or less efficient market implementation in practice. Our objective is to treat electric loads in the same way as generation units by including them in electricity markets so that they can compete with generators and also with each other [32, 33]. We develop a market platform in which demand resources can participate. There are other approaches to use load flexibility (e.g. direct load control methods). Although their implementation could be simpler and even more effective than market-based approaches, they are a step backward from achieving full market deregulation.

Related to the super grid modeling, we will first develop a model of a wholesale electricity market to investigate the economical amount of intermittent renewables. Second, we will examine the idea of a super grid, and redefine the objective function of the unit commitment problem in the presence of demand response. Third, we will explore solutions to this problem for an interconnected system consisting of a number of electricity markets on an hourly basis. Related to the load management modeling, implementing operation control methods can help the grid operator maintain real-time energy balance at a lower cost. Both generation and load deviate from their predicted hourly values in real time, which causes a mismatch between supply and demand. In this regard, fourth, we will propose a load management strategy to charge electric vehicles in a grid-friendly way using an agent-based modeling approach. Fifth, we will develop a method of operating thermostatically controlled loads based on the transactive control paradigm, in order to reduce energy costs and prevent grid congestion.

In summary, we will analyze the idea of operating an interconnection in a centralized manner that dispatches resources taking into consideration the real-time condition of the electric system. We will then explore the idea of including electric vehicles and HVAC loads in retail energy imbalance markets.

1.2 Dissertation Outline

This dissertation consists of an introduction in Chapter 1, five research articles presented individually in Chapters 2-6, and a conclusion in Chapter 7. The first and third articles were published in the Elsevier Applied Energy Journal, the second article was presented at the CSME 2016 International Congress, the fourth article was presented at the HICSS 2016 International Conference, and the fifth article has been submitted to Elsevier Applied Energy Journal. Each paper includes its own abstract, introduction, methodology, simulations, discussions and conclusion. Chapters 2-6 are outlined as follows:

In Chapter 2, we investigate the optimal integration level of a variable resource in a typical power grid in conjunction with demand response. In this regard, we consider a wholesale electricity market in which both generators and loads participate, on an hourly basis. The demand curve includes inflexible loads that are unresponsive to price changes and flexible loads that are responsive (price-sensitive). As shown in Figure 1.1, a sigmoid (logistic) function is suggested to represent the collective response of flexible loads to price changes. In addition, an asymptotic (hockey stick shape) function is used to represent the supply curve, consisting of a flat-price segment for must-take generation units and a variant-price segment for dispatchable units. We then discuss the impact of renewable intermittency and demand flexibility on the uncertainty cost acting jointly in a double auction. We also analyze sensitivity of the optimal amount to capital cost, carbon tax and load flexibility. The simulation results suggests that additional tools such as super grids and energy storage systems are required to increase the renewable penetration beyond a certain level.

In Chapter 3, we explore the concept of super grids. Load fluctuations and generation intermittency are not strongly correlated with each other over a large interconnected system. Accordingly, the combined interconnection power fluctuations are smaller than the sum of the variations in individual control areas. Therefore, with an unified manner of operation, it is possible to mitigate the intermittency of renewable generation. A simulation is performed to evaluate the effectiveness of the proposed idea on a system that loosely represents the North America Western Interconnection. The Western Interconnection, also known as Western Electricity Coordinating Council (WECC), stretches from Western Canada South to Baja California in Mexico, reaching eastward over the Rockies to the Great Plains. A hypothetical wholesale market is assigned to British Columbia and Alberta together, and another hypothetical market to the rest of the system consolidated. With price-sensitive loads included, the market is cleared to maximize the economic surplus rather than minimize the operation costs. Figure 1.2 shows a group of wholesale markets that exchange electricity to maximize the interconnection surplus. The impact of optimal inter-area electricity transfer (through the Canada-US tieline) on the economic surplus is assessed.

In Chapter 4, we formulate an interconnection-wide optimal flow scheduling method,

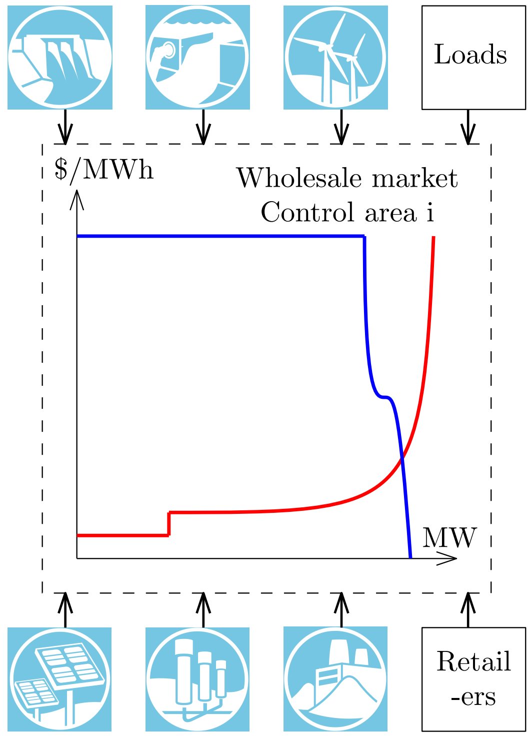


Fig. 1.1: A wholesale electricity market including different types of generation units as well as load aggregators and retailers.

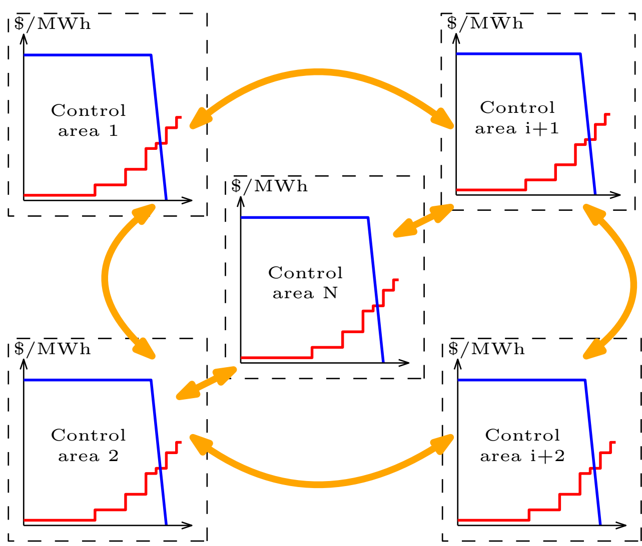


Fig. 1.2: Surplus-maximizing inter-area transfer flows within an interconnection.

and then test it on a 20-area reduced model of the WECC system. We assign a hypothetical double auction market to each area, considering the characteristics of the electric system in that area. The interconnection model captures the geographic distribution of resources as well as intertie constraints, taken from the WECC 2024 Common Case [34]. The proposed scheduling method simultaneously clears these markets to maximize the global economic surplus. However, due to technical and political barriers, it is practically impossible to operate the system in such an optimal manner. The central scheduler sets transfer flows to maximizes the global surplus. The proposed work is aimed at scenario studies on a large scale, but without going into the arcane details of bilateral contracts. This model is therefore good for looking at constrained optimal operation of the system, which places an upper bound on the achievable economic benefits of generation sharing in WECC.

All control areas are required to deliver the hourly scheduled imports/exports regardless of local real-time supply and demand fluctuations. Control areas must therefore deal with these possible energy mismatches using their local generation assets, which can be very costly. In Chapter 5 and Chapter 6, we explore the use of flexible load resources to assist grid operators in maintaining energy balance. We suggest using retail electricity markets to indirectly control distributed energy resources. Figure 1.3 illustrates a retail electricity market in which generation units and flexible loads can participate (double auction). In this regard, in Chapter 5, we propose a new charging strategy for electric vehicles to improve inter-temporal coordination between charging events and low cost periods in a real-time retail energy market. The difference between the elapsed time required for charging and the time that the vehicle is plugged allows for charging flexibility that allows consumers to take advantage of inexpensive renewable generation normally only available at particular hours of the day.

In Chapter 6, a market-based (indirect and centralized) demand response program is presented for thermostatically controlled loads under the transactive control paradigm. The role of demand response is to facilitate an accurate alignment between ON times and the most beneficial periods. We propose a bidding strategy that quantifies the load's willingness-to-pay (bid) price, taking into account both the indoor temperature state and the grid's real-time conditions. Simulation results indicate that implementation of this method of operation reduces energy costs in both heating and cooling modes, while maintaining the room temperature in the comfort range set by the consumer.

Figure 1.4 illustrates the relation between the models and market structure hierarchy presented in this work. This hierarchy can be thought of two ways. The first would be a proposal for a real-world market structure and operating mechanism. This

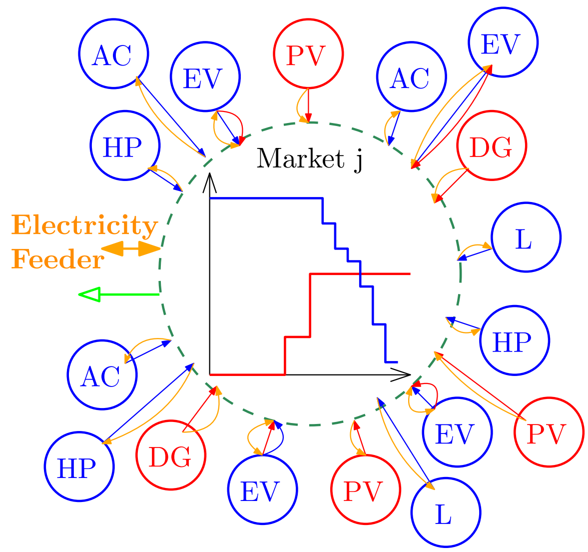


Fig. 1.3: A retail electricity market including distributed generators, solar panels, electric vehicles, heat pumps and air conditioners.

is not the purpose of this work, so that the second way of viewing the hierarchy is as a proxy, market optimization tool to determine upper bound efficiencies on real-world markets with a myriad of market structures. In summary, we will first present a model of an energy-only wholesale market in Chapter 2 (blue zone). This double auction market belongs to a control area that can potentially exchange electricity with other control areas through system tielines within an interconnection, in a dynamic manner. Under each control area, there are several retail markets, generation units and load centers. The generation units exist both directly participating in the wholesale market (large generators, e.g. coal plants or wind farms), as well as embedded in the retail market (distributed generation, such as building-mounted PV and micro CHP). The power grid operator dispatches the large generation units participating directly in the wholesale market based on the hourly schedule set at the wholesale market. Second, we develop a model of super grids to determine optimal inter-area change in Chapters 3 and 4 (green zone). The super grid model consists of a number of control areas (each with a wholesale market) that coordinate the generation an hour ahead in order to increase the global economic surplus. Third, we propose a model of 5-minute retail markets that includes electric vehicles and HVAC loads in Chapters 5 and 6 (gray zone).

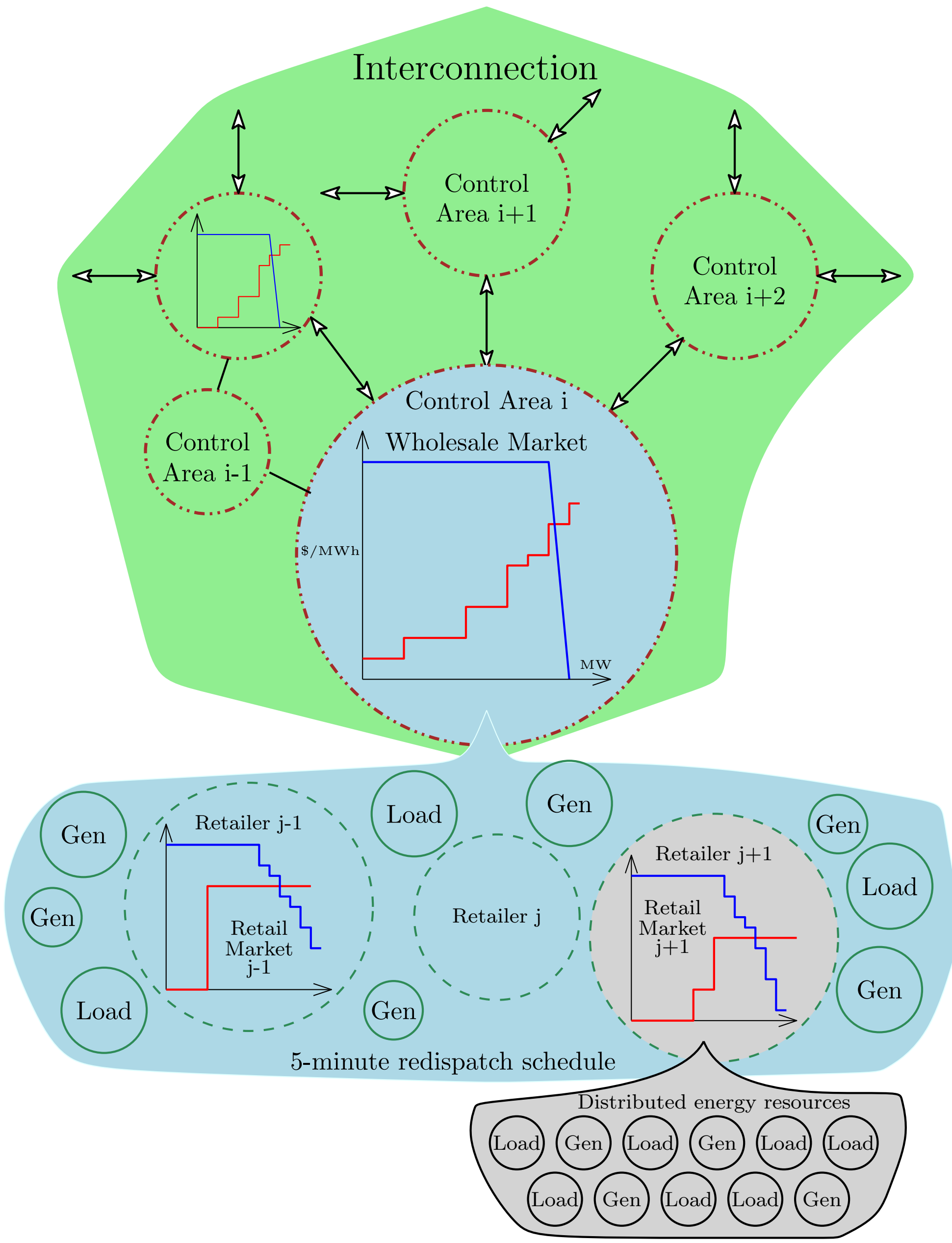


Fig. 1.4: Proposed model structure.

1.3 Research Contributions

The contributions arising from this work are listed below:

In Chapter 2,

(I) **Demand response integration into wholesale markets:** Including flexible loads in the market makes it possible to assess the interaction between must-take renewables and demand response, and also to explore the impacts on the uncertainty cost of intermittent renewables and demand response acting jointly.

(II) **Joint demand response/renewables portfolio study for a typical power grid:** Developing a methodology for portfolio studies is crucial as conducting such a study is an essential first step to development of renewable energies.

In Chapter 3,

(III) **A new resource allocation approach to incorporate market-driven demand response into the unit commitment problem:** With demand response included in a market, the settlement process is such that it maximizes the economic surplus rather than minimizing the operational costs. Accordingly, the objective function of the optimal scheduling problem (at the interconnection level) is redefined.

(IV) **The super grid concept:** We introduce the idea of operating an interconnected system in a centralized manner in terms of increasing the economic surplus, and then define new performance parameters to evaluate system interties.

In Chapter 4,

(V) **A model of the interchange export/import scheduling problem for the interconnection-wide surplus maximization objective:** We formulate an optimization problem to determine the unconstrained and constrained optimal inter-area power flows. With this model, we can then explore the impact of the electricity import/export on the economic surplus.

(VI) **The optimal inter-area transfer schedule for the Western Interconnection:** We present a reduced model of the WECC system consisting of 20 consolidated areas, each with a hypothetical wholesale market for the planning year 2024. The proposed surplus-maximizing scheduling approach is applied on this interconnection model, and simulation results are analyzed in detail.

In Chapter 5,

(VII) **Plug-in electric vehicle participation in a 5-minute retail market:** A load management scheme is developed to charge electric vehicles in a grid-friendly way, in the presence of an appreciable amount of rooftop solar PV panels.

(VIII) **The idea of smart charging:** The cost reduction associated with unidirectional charging (V1G) and bidirectional charging (V2G) scenarios is compared with the uncoordinated charging (V0G) scenario.

In Chapter 6,

(IX) **An agent-based model of a new operation method for thermostatically controlled loads considering temperature comfort range:** A bidding strategy for HVAC loads is developed that quantifies load flexibility, considering real-time grid conditions based on the transactive control paradigm. Then these loads are included in an energy imbalance market in conjunction with PV panels.

(X) **The collective behavior of HVAC loads:** We investigate load aggregator behavior in response to price changes in both heating and cooling modes.

Chapter 2

Renewable Resources Portfolio Optimization in the Presence of Demand Response

This paper was accepted to Applied Energy journal in October 2016:

Sahand Behboodi, David P Chassin, Curran Crawford and Ned Djilali. *Renewable Resources Portfolio Optimization in the Presence of Demand Response.* Applied Energy. 2016 Jan 15;162:139-48. Available online at: http://www.sciencedirect.com/science/article/pii/S030626191501301X

Sahand Behboodi has done the major part of developing the methodology, coding the simulation, and writing the text. David Chassin has helped Sahand to establish an understanding of the energy markets and the demand curve shape. David has also written the introduction section, and edited the entire manuscript.

This chapter proposes a model for demand response integration in wholesale electricity markets. We also present a cost model of integrating intermittent renewables and demand response that can be used to assess the optimal level of variable generation in an electric system.

Abstract

Demand response is viewed as a practical and relatively low-cost solution to increasing penetration of intermittent renewable generation in bulk electric power systems. This paper examines the question of what is the optimal installed capacity allocation of renewable resources in conjunction with demand response. We introduce an integrated model for total annual system cost that can be used to determine a cost-minimizing allocation of renewable asset investments. The model includes production, uncertainty, emission, capacity expansion and mothballing costs, as well as wind variability and demand elasticity to determine the hourly cost of electricity delivery. The model is applied to a 2024 planning case for British Columbia, Canada. Results show that cost is minimized at about 30% wind generation. We find that the optimal amount of renewable resource is as sensitive to installation cost as it is to a carbon tax. But we find the inter-hourly demand response magnitude is much less helpful in promoting additional renewables than intra-hourly demand elasticity.

Keywords

Demand response, Renewable integration, Power market, Portfolio optimization

Nomenclature

C	Annual cost, in \$/y.
c	Hourly cost, in \$/h.
G	Annual generation, in MWh/y.
g	Hourly generation, in MWh/h.
p	Price, in \$/MWh.
Q	Quantity, in MW.
q	Hourly demand, in MWh/h.
t	Time, in hours.
v	Normalized hourly wind production, per unit of installed wind capacity.
Greek symbols	
α	Magnitude of the variable cost component of supply curve, in \$/MWh.
β	Curvature of the supply curve, in a non-dimensional unit.
γ	Base price of the first dispatchable generation, \$/MWh.

ϵ	Emission factor of a resource, in tCO_2/MWh.
κ	Curvature of the demand curve, in a non-dimensional unit.
τ	Time-substitution delay of inter-hourly demand response, in hours.
ω	Fractional resource allocation, per unit of installed capacity.
Subscripts	
BG	Base load generation
CT	Carbon tax
D	Demand response
E	Emission
IG	Intermediate load generation
M	Market
P	Production
PG	Peak load generation
R	Reserve
SV	Scarcity value
U	Uncertainty
W	Wind

2.1 Introduction

According to the Energy Information Agency (EIA) International Energy Outlook developing economies have seen a steady growth in renewable energy resources in recent years. Wind and solar resources in particular show the strongest growth with EIA projecting that more than three quarters of all new additions in 2015 will be renewable [35]. The advantages of renewable energy are manifest and in the absence of viable alternatives to reducing greenhouse gas emissions, they are expected to remain the electricity generation resource of choice for new additions for many years to come. Unfortunately, all is not well where renewable electricity generating resources are concerned. Significant economic and operational considerations impose practical limits on the total amount of renewables that can be deployed in bulk electric power systems. Land use considerations, power system reliability, and electricity market design are among the many issues that contribute to constraints on the total deployment of renewables, particular those that rely on intermittent prime-movers, like wind and solar energy. Hydro-electric generation has long been employed as a significant renewable source of electricity. But climate change may jeopardize the magnitude and certainty with which the existing asset base can meet demand [36, 37], while lack of productive new dam siting options, population displacement, habitat destruction and fish stock degradation limit the growth of new assets. Wind power has seen rapid growth in recent years, but the need for reliable resources limits the penetration of wind generation unless additional intermittency mitigation measures are considered [2]. Solar resources are also becoming increasingly available but have intermittency challenges similar to those of wind. In addition residential rooftop solar resources are challenging the classical utility revenue model [38] and are known to cause voltage control issues in distribution systems in response to cloud transients and the diurnal cycle [39, 3]. There are also early signs that the wholesale market designs are not well suited to high penetration of renewables and the specter of revenue adequacy problems has been raised [40, 41]. Finally, the reliable, robust control and optimal operation of an increasingly complex bulk electricity system has become a very real concern [42].

The traditional utility approach to renewable intermittency is to allocate additional firm resources to replace all potentially non-firm renewables resources. These firm resources are generally provided by fast-responding fossil-fuelled thermal plants and hydro (where available) power generation as well. The need for fast-ramping resources discourages the dispatch of high-efficiency fossil and nuclear generation assets while promoting low-efficiency fossil for regulation and reserve services. The early state of development of many wholesale regulation markets precludes consideration of market-based remedies at this time, although arguably one should consider renewables before committing to any particular market design.

Demand response is generally regarded as a lower-cost alternative to fast-response generation reserves that reduces the dispatch of expensive generation resources [6, 18, 43, 44], although the response speed, magnitude and duration are important considerations [45]. The effect of demand response on the daily generation schedule is known [46] and sometimes demand response is even presented as a virtual power plant [29]. But load control strategies for demand response can be challenging to deploy [7] in part due to competing local and global objectives [47, 48] and in part due of the complexity of the load control modeling and design problem itself [27]. Numerical modeling of resource adequacy for large-scale planning problems is difficult to implement [49] and demand response models typically do not capture the salient features of load necessary to make optimal resource allocation decisions. This is particularly true when considering the interaction of renewable intermittency and demand response capabilities [50].

Effective and widely used strategies for optimizing the scheduling and operation of bulk-system resources have used markets to solve the cost-minimizing resource-allocation problem since they were proposed in the early 1980s [51]. Market-based control strategies were later adapted to building control systems [52], generalized to feeder-scale operations [53], then utility-scale operations [54], and most recently proposed for ancillary services [55]. Integrated demand dispatch mechanisms allow consideration of the combined economic impact of both intermittent generation under traditional wholesale markets and so-called "transactive" retail-side demand response dispatch system. It seems therefore possible to define global cost functions that incorporate the essential characteristics of both intermittent generation and demand response.

In recent years many have contributed relevant and very detailed models [12, 56, 57, 58] addressing the individual aspects discussed above. Wang *et al.* [59] reviewed prototyped real-time electricity markets, focusing on their market architectures and incentive policies for integrating distributed energy resources and demand response. Kwag and Kim [60] introduced a new concept of virtual generation resources, according to which marginal costs are calculated in the same manner as conventional generation marginal costs using demand response information: magnitude, duration, frequency and marginal cost. Sreedharan *et al.* [61] determined the avoided cost of demand response in a restructured market with renewables in California. Dallinger *et al.* [62] showed that a demand response program based on smart charging of electric vehicles can facilitate the integration of intermittent resources in California and Germany. Mahmoudi *et al.* [63] proposed a new wind offering strategy in which a wind power producer employs demand response to cope with power production uncertainty and market violations. To this end, the wind power producer sets contracts with a demand response aggregator. Rajeev and Ashok [64] proposed a dynamic load shifting program using real-time data in a cloud computing framework to enable the effective capacity utilisation of renewable resources. Heydarian-Forushani *et al.* [65] investigated the impacts of different electricity markets on the optimal behavior of a demand response aggregator in a renewable-based power system. Fripp [66] introduces Switch, a new open-source optimization model for long-term planning of power systems with large shares of renewable energy. Santoro *et al.* [67] used a stochastic approach based on Monte Carlo simulation technique to simulate the impacts of demand response in power systems with integrated renewable resources over one year period. They showed the optimization of demand response and renewable production reduces locational marginal prices.

Electricity consumers behavior, and in particular their response to price fluctuations is challenging to characterize and model, and researchers have modeled the behavior by using a linear demand curves to represent price responsiveness [68], invoking new methods to calculate a demand reserve offer function [69], or assuming consumers use day-ahead prices to shift daily energy consumption from hours when the price is expected to be high to hours when the price is expected to be low while maintaining total net energy consumption [70]. The results from the Olympic Peninsula demonstration project [53] and American Electrical Power gridSMART project [54] showed that the demand curves of thermostatic loads are generally sigmoid with asymptotes at the unresponsive quantity and the maximum load.

For the most part, these contributions do not collectively answer the larger question of how to determine the optimal installed capacity allocation of renewable resources when demand response is considered simultaneously. This paper introduces a model for total annual system cost that integrates renewable resource intermittency and demand response impacts in a global cost function that can be used to determine the optimal allocation of new asset investments. The new contributions of this work are: (i) formulation of the uncertainty cost of intermittent renewable resources and demand response acting jointly; (ii) an economic model of demand response interacting with renewables in markets; (iii) separation of the impact of intra-hour (short-term) demand response from inter-hour (mid to long-term) demand response; and (iv) a joint demand response/renewables portfolio study for British Columbia.

In Section 2.2 the model is described in detail, and in Section 2.3 we propose a resource portfolio optimization formulation that addresses the question of how much renewable and demand response is necessary to minimize annual cost in any given system. In Section 2.4 the model is applied to a system loosely based on the power grid of British Columbia, Canada and sensitivity analyses of the results are presented in Section 2.5.

2.2 System Description

In this section we describe the system models employed to solve the general annual-cost minimizing resource allocation problem. The model includes three categories of elements: (i) the resource models, (ii) the temporal models, and (iii) the economic models.

2.2.1 Resource Model

System resources are modeled with five classes of generation and two classes of load:

Base: Baseload generation includes all generators that are presumed to be always running when available. Baseload generation usually has a very low marginal cost but is not expected to respond to intra-hour load changes or intermittency in other generation assets.

Intermediate: Intermediate generation includes all the main energy production assets that are used to follow the normal diurnal fluctuations in demand. Intermediate generation is usually also relatively low marginal cost but is expected to have at least some ability to change output in response to intra-hour imbalances.

Peak: Peak generation usually includes only low efficiency energy production assets that are used to meet peak load events that happen infrequently. These assets are typically low capital-cost assets with high marginal costs of production, but they are expected to have excellent ability to change output quickly in response imbalances.

Intermittent: Intermittent generation generally has high first cost, but effectively zero marginal production cost. The main feature of intermittent resources is that they are essentially non-dispatchable because their production capacity is subject to uncontrollable fluctuations in the prime mover, e.g., wind, solar, or wave. As a result not only intermittent resources cannot provide any useful load following capability, but they may also contribute to increasing imbalances due to forecasting uncertainty.

Reserve: Reserve generation is usually comprised of peak generation units that are effectively never used and only held in reserve in the event of a system contingency. Because many of these non-spinning reserve units typically are not dispatched, they effectively do not generate revenue directly from production. Instead they are a cost which is typically recovered through scarcity rent on the other assets in a vertically integrated systems, or by participating in reserve markets, when they exist.

Unresponsive: Unresponsive loads include the vast majority of load in most systems. Unresponsive load generally has five components: (i) the base load, which is always present; (ii) the seasonal load, which varies according to the time of year; (iii) the long term weather component, which changes with weather; (iv) the diurnal component, which fluctuates with the daily solar cycle; and (v) the short term weather and human activity disturbances, which fluctuate on a subdaily and often subhourly basis.

Responsive: Responsive loads are all the loads that can respond to signals of various kinds, including direct and indirect (e.g. price-based) load control signals. Responsive load is generally divided into three categories: (i) curtailable load, where energy use is reduced and not replaced later, e.g., by industrial load curtailment; (ii) deferrable capacity or inter-hourly demand response, where peak demand is cut and energy use is replaced in subsequent hours, e.g., by direct load control; and (iii) fast-acting ramp response or intra-hourly demand response, where load is shifted momentarily and typically replaced within one hour, e.g., by real-time price signals.

Elasticity represents the response of consumers to dynamic pricing. The price elasticity of demand is the fractional change in demand to a given fractional change in price:

$$\eta = \frac{p}{q}\,\frac{dq}{dp} \tag{2.1}$$

where η, p and q respectively are elasticity, price and demand. Numerous studies in recent decades have examined the elasticity of demand under various tariffs. However, few of those studies [71, 72] address real-time price tariffs. In their survey of 15 demand response studies, Faruqui and Sergici [73] identified the likely range of inter-hourly elasticity of substitution as between -0.07 and -0.21. For the purposes of this paper, we use the suggested average value of -0.14. It should be noted that the long-term demand elasticity is taken out from consumers' behaviour before estimating the short-term elasticity. Sensitivity analysis on the elasticity is formed to gauge the impact of this estimate on the results.

2.2.2 Temporal Model

All system assets are typically scheduled for operation on an hourly basis in a day-ahead electricity market. Fast-acting controllable assets can be dispatched subhourly, with some assets responding at a five-minute time-scale (e.g. demand response), some at the 4 second control time-scale, and some at a subsecond electromechanical dynamic response and sub-cycle relay/protection control time scale. However, in general these are only considered insofar as they may reduce the backup reserve requirement and they do not affect the energy component of the hourly dispatch schedule.

The potential magnitude of demand response resources must be considered in terms of the bandwidths over which they can operate [74]. In general demand response that addresses intermittency is based on load resources that respond only within a time no greater than about a few hours and no less than a few minutes, the upper limit arising from limits on the customer's willingness to forgo or defer consumption, and the lower limit arising from the time update rate of the load control signal or load control lockout. For example, building thermostat-based demand response is relatively fast and essentially subhourly, whereas electric vehicle charging demand response is relativity slow and primarily super-hourly. The magnitude of the intermittency within that frequency band is the only intermittency that demand response can mitigate and therefore the only intermittency that we can consider being cancelled in the total resource pool [75].

The production cost for energy is determined hourly based on the variation in load for each hour of the year. In this study, demand response with inter-hour capability is assumed to not be significant beyond 4 hours. The proposed model dispatches price-sensitive load by comparing the real-time price and the average price over the next 4 hours. In the case of subhourly response, we assume that all fast dynamics have mean zero contribution to the hourly energy demand, but they do have a non-zero variance contribution to the power imbalance. For intermittent generation the cost of mitigating this variance is included in the cost function through the variability of the wind production. With fast-acting demand response the magnitude of its contribution is assumed to be always less than the subhourly intermittency of wind but effectively mitigates intermittent generation. Because the marginal cost of fast-acting demand response and the marginal cost of intermittent generation are both zero, they are simply cancelled and the total intermittent wind subhourly impact on the system is reduced by the amount of demand response available. The first cost of fast-acting demand response is assumed to be in addition to the deferrable load control infrastructure cost. Where automated metering infrastructure is already in place, this additional first cost should in principle be nearly zero.

2.2.3 Economic Model

The global cost function consists of variable and fixed costs. The variable costs include the following components:

Production: This cost includes the hourly cost of producing energy from the resources that were dispatched. In principle this should include subhourly cost of production as a result of redispatch to follow load and mitigate forecast deviations in intermittent resources, but we assume that this cost has zero mean over the hour.

Uncertainty: This is actually defined as producer surplus [76]. But because this cost arises primarily from the requirement to maintain dispatchable resources with non-zero marginal costs to mitigate for the uncertainty in non-dispatchable resources with zero marginal (as well as variability in the unresponsive load) we choose to call this *the cost of uncertainty* due to the intermittency of lower or zero cost resources. As we will see below, this definition has the significant advantage of allowing us to easily relate the magnitude of the resource uncertainty to the cost impact of that uncertainty as the allocation of that resource changes. For example, the uncertainty cost of a small allocation of wind is counterintuitively much higher than it is for a large allocation of wind simply because as we add more wind, the resources being used to mitigate its intermittency are dispatched from lower down the supply curve. This effect is independent of and in addition to smoothing effects [77] that results from geographic resource diversity.

Emission: This cost is considered by the introduction of a carbon tax at the point of CO_2 emission.

The fixed costs include the following components:

Wind: Increasing wind allocation requires an investment in the installation of new units, which is represented by a levelized cost of energy on what would otherwise be a zero production cost. However, strict application of market-pricing regards this as a sunk cost once the unit is installed, which is why intermittent wind, solar and wave units are effectively zero-marginal cost relative to fossil units. Thus the first cost of new units is captured separately in the model in order to avoid having to account for these in the production cost.

Demand response: There is very little recent data on the first cost of demand response installation. Borenstein provided a quote from Comverge in 2002 where the estimate was $1000 per customer [78]. If we assume that each customer can provide about 10 kW of controllable load on peak, the cost of controllable demand response capacity is around $50,000/MW.

Reserve: As the allocation of wind is increased, a proportion of non-spinning reserve is not required but continues to incur costs.

2.3 Problem Formulation

In this section, the cost minimization problem is stated in the standard form, and its components are quantified considering renewable intermittency and demand response effects.

2.3.1 Hourly Cost

To derive the annual cost function we begin with the hourly costs, which will then be integrated over a year. The hourly cost includes the market-based energy cost, and the intermittency and the demand response impacts discussed above.

Market Cost

We use a mathematical formulation for the market cost based on an asymptotic supply curve for production cost, which is combined with a cost arising from the producer surplus, which we refer to as the uncertainty cost when intermittent resources are considered. Consider the supply curve illustrated in Figure 2.1a where producers bid their marginal costs (which is zero for wind) and are paid the clearing price, which is the marginal cost of the last unit dispatched. At any given time, the region enclosed by the market clearing quantity q_M and price p_M is the market cost $c_M = p_M\, q_M$. So the region under the supply curve is the total production cost which is the sum of the individual production costs of the each generating unit $c_P = \sum_{n=1}^{N} p_n \ g_n$ and the region above the curve but below the clearing price is total producer surplus associated with each unit $c_U = \sum_{n=1}^{N} (p_M - p_n) \ g_n$. The supply curve can be expressed as an asymptotic function of the quantity:

$$p = \alpha \ \left(1 - \frac{q}{Q_{cap}}\right)^{-\beta} + \gamma \tag{2.2}$$

where α and β determine the magnitude of the scarcity rent and the curvature of the supply curve respectively, and γ is the minimum bidding price of the first dispatchable unit. The system capacity Q_{cap} is the maximum observed demand Q_{max} with the supply requirement reserve factor ω_R, so $Q_{cap} = Q_{max} \ (1 + \omega_R)$. The cost function for any particular time t is then:

$$c_{M[t]} = \alpha \ q_{M[t]} \ \left(1 - \frac{q_{M[t]}}{Q_{cap}}\right)^{-\beta} + \gamma \ q_{M[t]} \tag{2.3}$$

We can express the energy production cost as:

$$c_{P[t]} = \int_0^{q_{M[t]}} p \ dq = \frac{\alpha \ Q_{cap}}{\beta - 1} \ \left\{\left(1 - \frac{q_{M[t]}}{Q_{cap}}\right)^{-\beta+1} - 1\right\} + \gamma \ q_{M[t]} \tag{2.4}$$

Then the uncertainty cost $c_U = c_M - c_P$ at any particular time is:

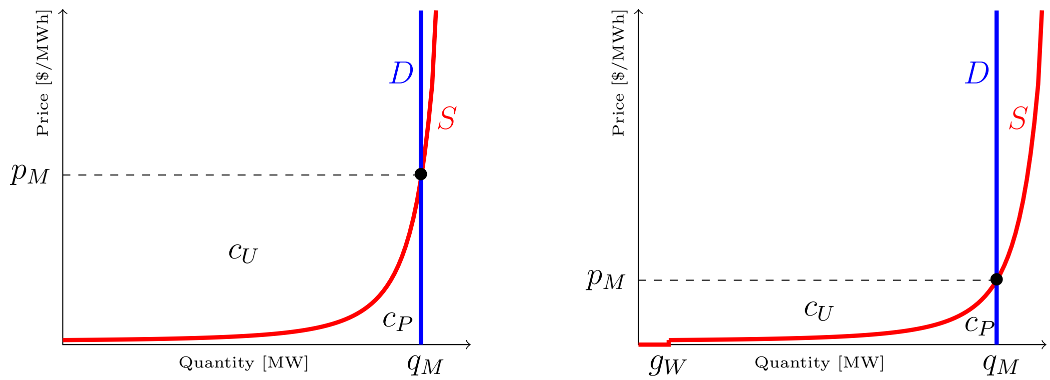
$$c_{U[t]} = \int_0^{p_{M[t]}} q \ dp = \frac{\alpha \ Q_{cap}}{\beta - 1} \ \left\{1 - \left(1 - \frac{q_{M[t]}}{Q_{cap}}\right)^{-\beta} \ \left(1 - \frac{q_{M[t]}}{Q_{cap}} \ \beta\right)\right\} \tag{2.5}$$

Wind Intermittency Effect

When renewable resources are active, they are dispatched below the baseload resources in the supply merit order, and therefore they shift supply curve accordingly. We assume that renewable production cost is zero because the marginal cost of all wind is zero. But the producer surplus can be large, as shown in Figure 2.1b. Deducting the wind generation g_W from the demand, the clearing price is:

$$p_M = \alpha \left(1 - \frac{q_{M[t]} - g_{W[t]}}{Q_{cap}}\right)^{-\beta} + \gamma \tag{2.6}$$

where wind generation is $g_{W[t]} = v_{W[t]} \ \omega_W \ Q_{max}$, and v_W is the normalized wind



(a) Supply curve without wind **(b)** Supply curve with wind

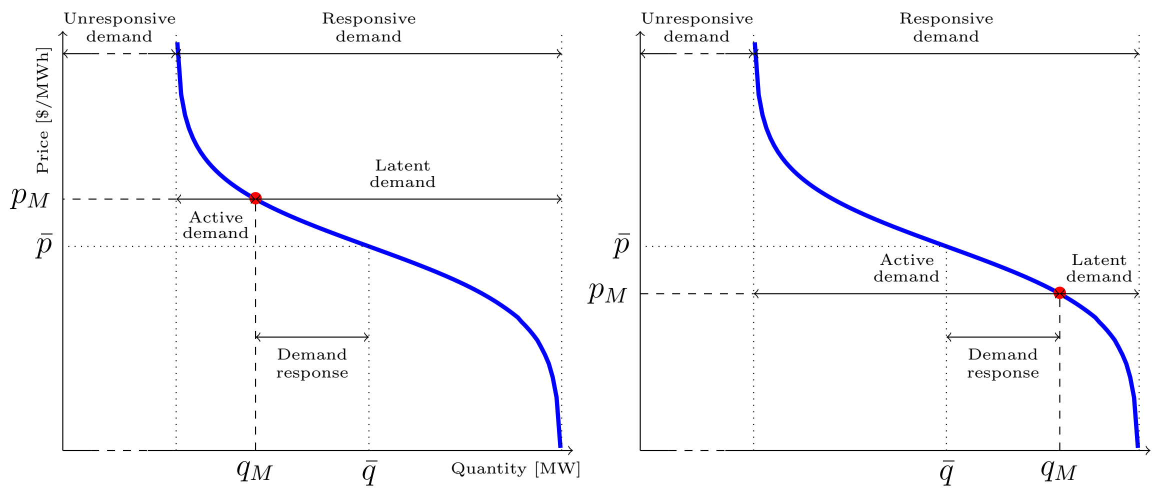
Fig. 2.1: Single auction electricity market: demand curve (blue) and supply curve (red).

generation pattern taken from historical data.

We make an important observation about the cost of uncertainty: if the supply curve is flat, the cost of uncertainty is zero even though there may be very high variability associated with the lowest cost resource. In other words, if the lowest cost resource is highly uncertain, but can be replaced by other similarly low-cost resources, then the cost of uncertainty may be in fact near zero. Of course, this condition is predicated on the notion that excess resources are "curtailed", which may not always the case with today's wind resources. But this possibility suggests that any attempt to optimize a resource portfolio where unlimited highly uncertain resources are permitted will necessarily result in an optimal allocation where a large amount of low cost/high uncertainty resources are acquired and only the uncertain resources are used.

Demand Response Effect

Being sensitive to electricity price, customers change their demand in response to price fluctuations. With demand response included, the total quantity consumed is given as the summation of price unresponsive and responsive demands. The particular form of the sigmoid function is not readily deduced from the field data, but one can presume that it arises from the discrete choice statistics of the consumers based on the random utility model [79]. According to this model, comfort governs the outcome with the highest utility going to the customers with the highest demand for comfort. The net benefit to each customer depends on an unobservable characteristic a and an observable one b, such that the utility of choosing x is $a + bx + \delta$ where



(a) Demand curtailment

(b) Demand recovery

Fig. 2.2: Behavior of price-responsive demand.

δ is a random independent error. The action corresponding to that choice is taken when $a + bx + \delta > 0$. The probability of taking the action is then proportional to $(1 + e^{-(a+bx)})^{-1}$. The behavior of the demand response model under curtailment and recovery is illustrated in Figure 2.2a and Figure 2.2b respectively.

The proposed model determines the active responsive load considering the mean of expected prices $\bar{p}$ within the next τ hours. We use this model to express all demand curves from automated demand response agents such as HVAC thermostats and electric vehicle chargers as taking the form:

$$q_{D[t]} = \frac{2\ \omega_D\ \bar{q}_{[t]}}{1 + e^{\kappa\left(\frac{p_M}{\bar{p}} - 1\right)}} \tag{2.7}$$

where ω_D is demand response allocation and κ is the demand response function curvature. It should be noted that q is the total load when the demand is completely blind to the price. The responsive demand q_D changes to clear market at quantity of $q_M = \bar{q}\ (1 - \omega_D) + q_D$ and its associated price p_M. We consider only demand response that is capable to shift the load for more than 1 hour and treats all subhourly demand response as mean-zero magnitude. The average elasticity of demand $\bar{\eta}$ is then given over a τ hours time window, and defined based on the instantaneous elasticity. Combining this definition with the equation for the demand response we find: $\kappa = -2\ \bar{\eta}$. Figure 2.3 illustrates the interaction between the proposed demand response model (under curtailment) and supply model in a double auction market. The horizontal

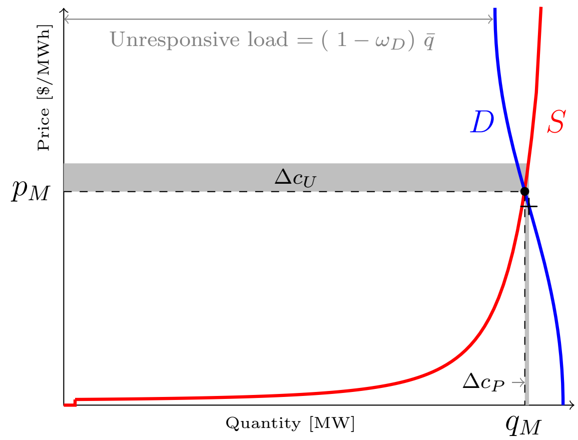


Fig. 2.3: Double auction electricity market.

and vertical shades respectively show the reduction of uncertainty and production costs as a result of demand response implementation.

2.3.2 Annualized Effective Electricity Price

Our objective is to minimize the total annual cost, which is a function of the renewable resource and demand response penetration levels. The total annual cost, consisting of annual production, uncertainty, emission, wind capacity expansion, supply reserve and demand response (e.g. labor and hardware) installation costs, is computed for different combinations of design variables ω_W and ω_D. We express results in annualized effective electricity price p_{eff}, which is a more easily understood criterion, and by definition is the total annual cost divided by the annual demand:

$$p_{eff(\omega_W,\omega_D)} = \frac{C_P + C_U + C_E + C_W + C_R + C_D}{\sum_{t=1}^{8760} q_{M[t]}} \tag{2.8}$$

where $C_P = \sum_{t=1}^{8760} c_{P[t]}$ and $C_U = \sum_{t=1}^{8760} c_{U[t]}$ are the annual production and uncertainty costs. These costs are obtained across the entire year with a hourly time resolution. The emission cost C_E is:

$$C_{E(\omega_W,\omega_D)} = F_{CT}\ (\epsilon_{BG}\ G_{BG} + \epsilon_{IG}\ G_{IG} + \epsilon_{PG}\ G_{PG}) \tag{2.9}$$

where ϵ and G are the carbon intensity and annual generation of base, intermediate and peak load generation units; F_{CT} is the carbon tax. The cost of adding a new wind unit is assumed quadratic due to market scarcity for large magnitude resource additions. The wind installation cost C_W therefore is computed as:

$$C_{W(\omega_W)} = F_{SV}\ (\omega_W\ Q_{max})^2 + F_W\ \omega_W\ Q_{max} \tag{2.10}$$

where F_{SV} and F_W are scarcity value and average wind initial cost. The cost of unused supply reserves C_R is:

$$C_{R(\omega_W,\omega_D)} = F_R\ (Q_{cap} - (1+\omega_R)\ [q_{[t]} - g_{W[t]}]_{max}) \tag{2.11}$$

where F_R is the average cost for unused reserve capacity. Finally, the cost of demand response C_D infrastructure is assumed as:

$$C_{D(\omega_D)} = F_D\ \omega_D\ Q_{max} \tag{2.12}$$

where F_D is the estimated cost for demand response.

2.4 Application

We apply the proposed model to a hypothetical electric system based on the planning model for the province of British Columbia, Canada used by the Western Electricity Coordinating Council for the year 2024 [34]. The hourly load forecast and wind generation profiles of the province are taken from a 10 year-ahead planning case. British Columbia's power system is not deregulated, so we use a hypothetical energy market with characteristics of a deregulated market, as shown in Table 2.1. To estimate the emission costs, the baseload generation type is assumed to be a zero emission resource (e.g. hydro), and intermediate and peak units are combined and simple cycles respectively. Figure 2.4 shows the supply curve with a supply reserve requirement of 14%.

Table 2.1: Market characteristics

Variable	Base	Intermediate	Peak	Reserve	Unit
Capacity	5300	3500	3500	1700	MW
Emission factor [80, 81]	5	450	670	670	tCO_2e/MWh
Minimum bid	15	25	65	1006	\$/MWh

The first cost of wind is determined by averaging the direct capital cost of 111 potential onshore wind site in British Columbia [82]. The assumed costs and demand

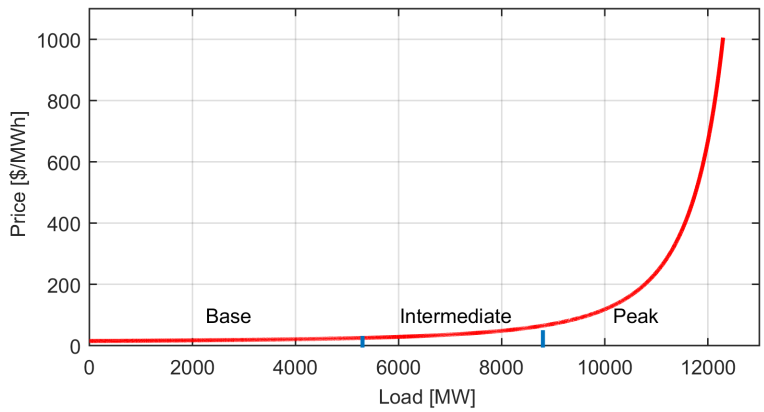


Fig. 2.4: Supply curve.

response parameters are tabulated in Table 2.2. We estimate the scarcity value based on the wind turbine price trends in the US over the past decade [83].

Table 2.2: Assumed values

Variable	Symbol	Value	Unit	Source
Carbon emission tax	F_{CT}	30	$/tCO$_2$	[84]
Demand response cost	F_D	50000	$/MW	[78]
Interest rate	i	3	%	
Elasticity	η	-14	%	[73]
Peak load	Q_{max}	12300	MW	Table 2.1
Scarcity value	F_{SV}	7	$/MW^2	
Time-substitution	τ	4	h	
Unnecessary supply reserve cost	F_R	100000	$/MW-year	
Wind installation cost	F_W	3210000	$/MW	[82]

The cost model is applied across a range of wind penetration levels. Figure 2.5 is a plot of objective function, allowing identification of an optimal level. The penetration level of 100% is the case where the wind capacity equals to the maximum demand Q_{max}. With demand response considered, the optimal wind capacity is slightly less than 3860 MW, or 31.2% of the system load on peak. In other words, the reduction in the combined annual production and uncertainty costs is greater than the wind installation capital cost up to 3860 MW. By May 2015, British Columbia had 4 wind farms currently supplying power to the grid with a nameplate capacity of 487 MW, and another 4 wind farms in development with a nameplate capacity of 230 MW [85].

The scenarios presented in Figure 2.6 and Figure 2.7 are as follows:

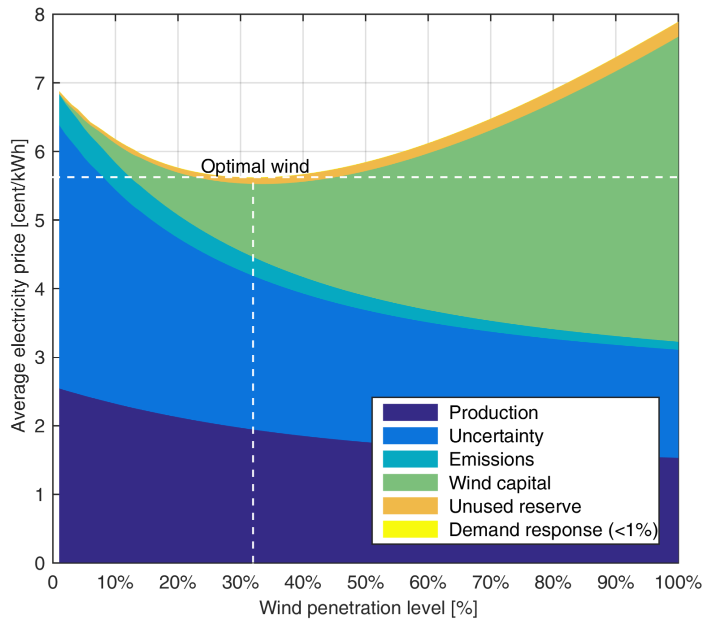


Fig. 2.5: Objective function for various possible renewable capacity allocations with active demand response

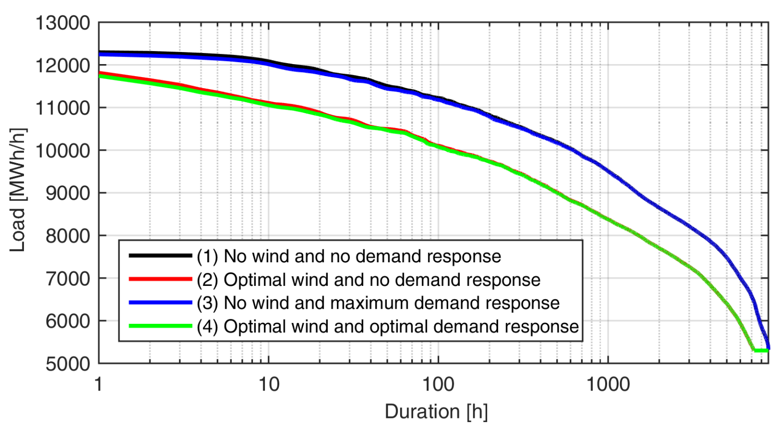


Fig. 2.6: Load duration curve

(1) No wind and no demand response;

(2) Optimal wind and no demand response;

(3) No wind and maximum demand response; and

(4) Optimal wind and optimal demand response.

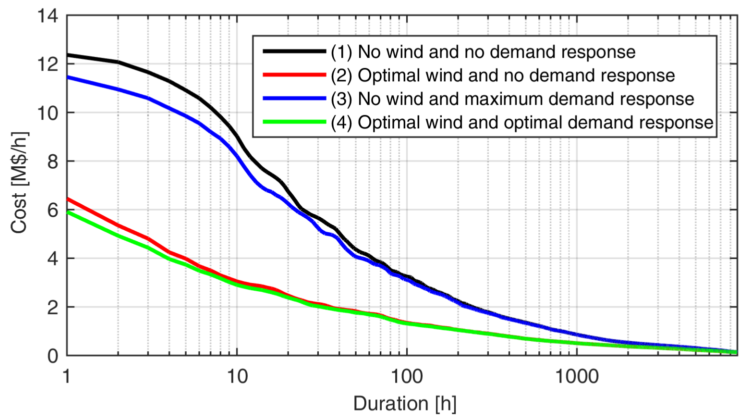


Fig. 2.7: Cost duration curve.

The fraction of maximum responsive demand to total demand was established at 10% and the optimization determined that this is the corner solution for the optimal demand response. A logarithmic scale on the duration axis is used to emphasize the results during peak hours where they have the greatest impact on overall costs. Table 2.3 compares annual generation of all scenarios.

Table 2.3: Annual generation by type

	Scenario				
Resource	**(1)**	**(2)**	**(3)**	**(4)**	**Unit**
Intermittent	0	8500	0	8515	GWh
Base	46555	46555	46555	46555	GWh
Intermediate	19845	12635	19862	12636	GWh
Peak	1753	464	1729	453	GWh

Figure 2.8 illustrates the impact of the demand response fraction on the annual production and uncertainty costs and also on its installation cost. The saving impact of demand response on the uncertainty cost is much greater than its impact on production cost.

Figure 2.9 shows the behaviour of demand response pricing over the study year. This illustrates the degree to which demand response is reacting when hourly prices are different from expected price. The negative points (red) are hours during which the market price is higher than the average of the next 4 hours; therefore, the responsive demand is postponed.

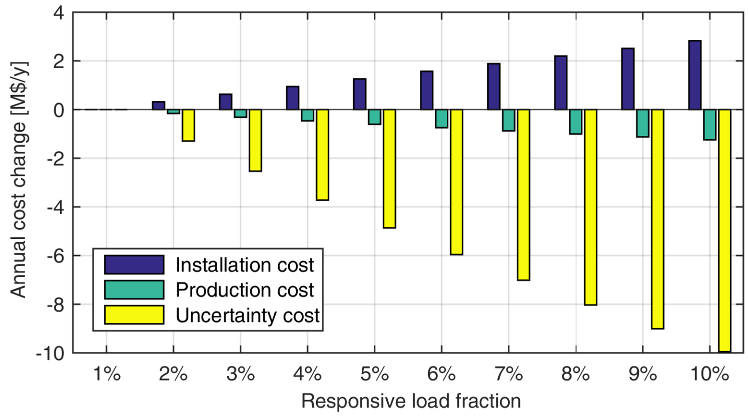


Fig. 2.8: Demand response cost and benefits.

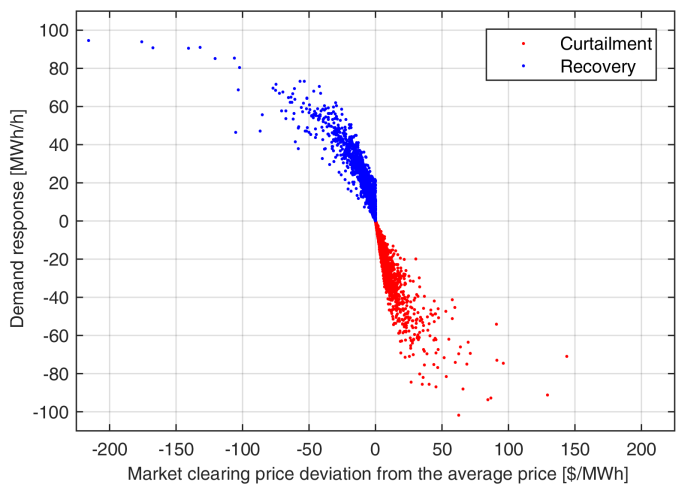
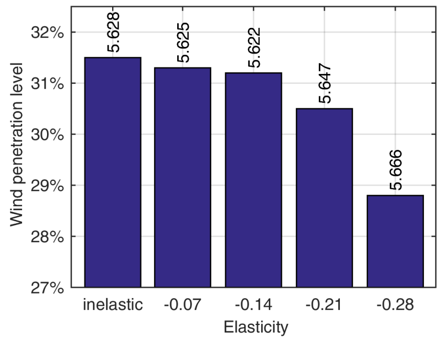


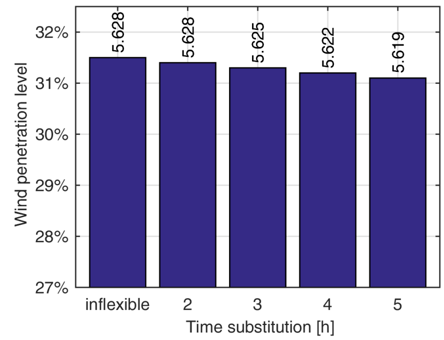
Fig. 2.9: Demand response behaviour.

2.5 Discussion and Sensitivity Analysis

At the hourly scale of energy markets, only a slight benefit from demand response can be observed. The sensitivity of the optimal wind allocation to elasticity and time-substitution of demand response are shown in Figure 2.10. This result shows that strategies to increase load shifting horizon and demand elasticity have no significant impact on the effective electricity price for the optimal wind case. This suggests that reasoning based on the inter-hour forward energy prices does not offer a significant benefit when compared to accounting for only the intra-hour price fluctuations. This emphasizes the importance of analyzing thermostatic (intra-hourly) demand response using short-term fluctuations in prices, separately from storage-based (inter-hourly)



(a) Price elasticity of demand



(b) Time substitution

Fig. 2.10: Demand response impact on optimal wind allocation and effective electricity price (¢/kWh).

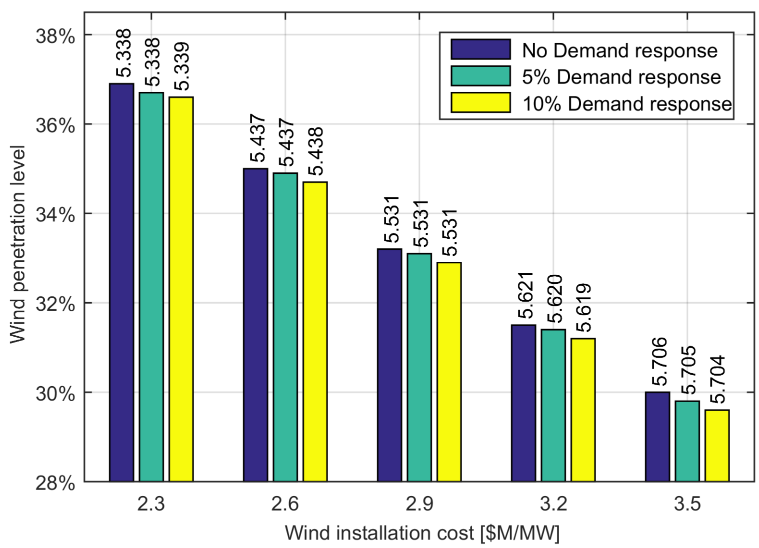


Fig. 2.11: Wind installation cost impact on optimal wind allocation and effective electricity price (¢/kWh).

demand response using slow price fluctuations.

The nominal wind installation cost assumed for this study is \$3.21 M/MW. However wind turbine costs are expected to decrease over time. Figure 2.11 shows the sensitivity on the wind installation cost. For a 30% decrease in wind capacity cost, we observe a 5.5% increase in wind capacity and a corresponding 0.28 ¢/kWh decrease in electricity price.

A carbon tax is widely regarded as one of the most effective tools regulators have to encourage power producers to invest on clean energy resources. Figure 2.12 shows

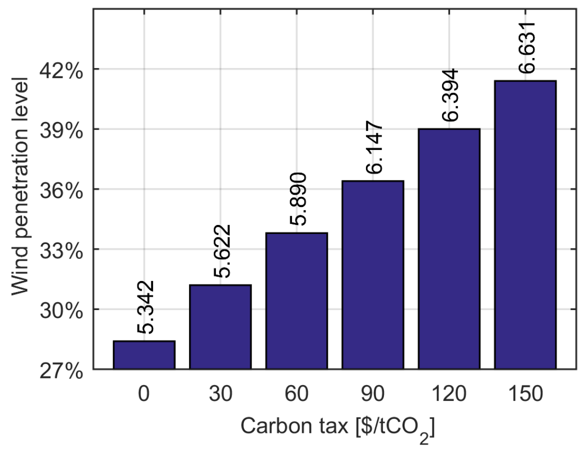


Fig. 2.12: Carbon emission tax impact on optimal wind allocation and effective electricity price (in ¢/kWh).

the sensitivity to the level of carbon tax and suggests the optimal wind penetration changes more than 10% for a range of reasonable carbon taxes expected in the foreseeable future.

Since we consider a fixed generation schedule for baseload units, we must assume excess wind generation is curtailed rather than redispatching baseload units. However if the baseload generation can go to part load, renewable curtailment is reduced. Figure 2.13 shows the sensitivity on the part load factor range. From this analysis, part load does not have an appreciable impact on the optimal allocation intermittent resources. The cost impact of curtailing wind rather than redispatching baseload is insignificant because of the low cost during off peak load hours when this is expected to occur.

2.6 Conclusions

In this paper we introduce a simple cost model of renewable integration and demand response that can be used to determine the optimal mix of generation and demand response resources. We use numerical methods to obtain the optimal mixtures of renewable generation and demand response resources given a fixed portfolio of conventional generation assets, wind patterns and energy use. The model incorporates production, uncertainty, emission costs, as well as capacity expansion and mothballing costs, and considers wind variability and demand response impacts to determine the hourly price of electricity delivery. Supply is divided into intermittent, base, interme-

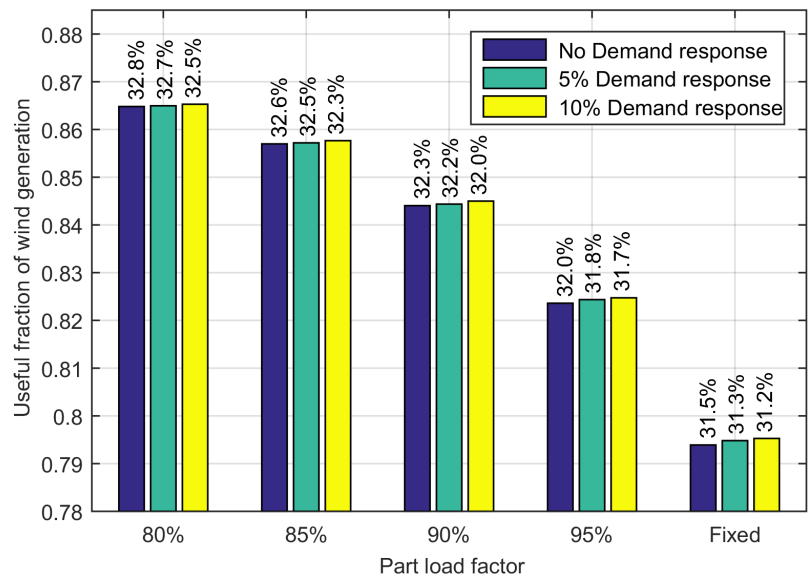


Fig. 2.13: Part load factor impact on useful wind generation fraction and optimal wind allocation.

diate, peak and reserves, while load is divided into unresponsive and responsive. The load model includes inter-hourly impact of fast (e.g. building thermostat) and slow (e.g. electric vehicles) responses to price variations. The temporal model includes time-substitution in demand up to 4 hours.

The model is tested by optimizing the 2024 planning case for British Columbia at the hourly level. We find that cost is minimized with about 31.5% renewable generation. The cost reduction relative to the current level is about 15%. The optimal renewable mix decreases to 31.2% when 10% demand response is considered with a very small cost impact. We find that demand response does not have a significant impact on cost at the hourly level, which suggests that future work must include sub-hourly load behavior to properly consider its full potential. The results also suggest that the optimal level of renewable resources is not very sensitive to demand elasticity, but it is highly sensitive to a carbon tax and renewable resource first cost.

Chapter 3

Optimal Inter-Area Transfer in the Presence of Demand Response and Renewable Electricity Generation

This paper was presented at CSME international congress on June 27th 2016:

Sahand Behboodi, David P Chassin, Curran Crawford, Ned Djilali. *Optimal Inter-Area Transfer in the Presence of Demand Response and Renewable Electricity Generation.* In proceedings of Canadian Society for Mechanical Engineering (CSME) International Congress 2016.

Sahand Behboodi has done the major part of developing the theory, performing the simulation, and preparing the paper. David Chassin has helped Sahand to establish an understanding of the double auction theory and the economic surplus calculation. David has also revised the manuscript.

This chapter discusses the impact of incorporating demand response in the unit commitment problem. It also provides an insight to the super grid concept, and assesses the impact of electricity imports/exports on the economic surplus.

Abstract

This paper describes the impact of demand response and intermittent renewable resources integration on electricity generation and inter-change scheduling. A surplus

maximizing method is proposed and tested on a hypothetical system of two consolidated areas that loosely represent the North America's Western Interconnection (separated by the Canada-United States border). An hourly electricity market is assigned to each area, and the power exchange that achieves the maximum surplus is obtained for the planning year of 2024. The solution is then modified to account for the existing path transfer capacity, 3150 MW North-South and 3000 MW South-North. The path economic utilization factor, the ratio of the surplus increase to the maximum surplus increase, is 31%. The economic power transfer distribution factor, the metric used to quantify the sensitivity of the power flow with respect to the price difference, is 67 MW/($/MWh). In addition, the optimal schedule is sought for cases when path transfer capacity is expanded by 1000 MW and 2000 MW. The same study is performed for year 2030, assuming the additional wind electricity meets the demand growth. Results show the economic utilization factor increases to 32%, and the economic power transfer distribution factor increases to 84 MW/($/MWh).

Keywords

Demand response; variable generation; scheduling; surplus; electricity market; tie-line transfer capacity.

3.1 Introduction

Growing participation of renewable resources in the overall generation fleet has increased pressure on generators responsible for ensuring reliability to provide resources that mitigate renewable intermittency, without increasing overall GHG emissions. Demand response is widely regarded as a potentially significant class of zero-carbon reliability resources that can displace carbon-intensive reliability resources, such as natural gas combustion turbines and/or energy constrained reliability resources such as hydro-electric generators. The US Department of Energy has adopted a definition of demand response that is now widely recognized for its inclusiveness [14]: load variations in response to change in both financial incentives and/or reliability signals over time. The interaction between demand response and renewable resources in electricity markets is a well-studied topic [18, 46, 59, 60, 67, 86]. Given an inelastic demand, the market finds the generation and inter-change schedules that minimizes the operational costs; however, in the presence of demand response, the market finds the schedule that maximizes the total surplus [87]. Total surplus is defined as consumer monetary value minus producer cost [76]. We present a methodology to investigate the impact of demand response implementation on generation and inter-change scheduling. The proposed method is tested on a hypothetical system consisting of two areas that loosely represent the North America's Western Interconnection, known as the Western Electricity Coordinating Council (WECC), including 14 Western States, two Canadian Provinces and Baja California in Mexico. Area 1 is the consolidation of British Columbia and Alberta power grids, and Area 2 is the rest of the system. In interconnected systems, transmission tie-lines enable balancing authorities to exchange electricity and share operating reserves. In the WECC system, balancing authorities exchange electricity to reduce their operational costs, although the interconnection is not operated according to a true optimal schedule, because of local regulations. As the penetration of variable generation resources increases, balancing authorities can collaborate more effectively to benefit from the geographical diversity of renewable resources in the interconnection, which requires an enhanced transmission system. Previous work [11] has shown potential savings in production cost due to consolidation of balancing authorities, with 8% wind and 3% solar energy penetration, ranges from 2.4% to 3.2% of the total yearly production cost, considering transmission congestion; the full copper-sheet consolidation of WECC shows an additional 1.4% improvement. A recent study [4] showed that the deployment of wind and solar power can reduce CO_2 emissions in the US by up to 80% relative to 1990 levels, without an increase in electricity price, by moving away from a regionally divided electricity sector to a national system enabled by high-voltage direct-current transmission lines. The inter-area exchange schedule is usually updated on an hourly basis; however, intra-hour scheduling should perhaps be used, since it has substantial cost benefits, particularly for cases with high penetrations of variable generation. A 10-minute exchange schedule has been shown to reduce the WECC production cost by 4% and 6% for intermittent renewable penetrations of 11% and 33%, respectively [88]. This paper examines the impact of demand response integration and transmission expansion on generation and inter-change schedules in the presence of variable generation resources, for WECC year 2024 and 2030 case studies. In Section 3.2, a detailed description on the proposed scheduling method is provided. In Section 3.3, optimal flow is determined for given assumptions, and the surplus increase is studied.

3.2 Methodology

3.2.1 Market Basics

The demand curve consists of a flat segment, including price-unresponsive loads, and a sloping segment including responsive loads. An approximation for the demand function is [89]:

$$Q_L(P) = \frac{Q_R}{1 + e^{2\ \eta}(1 - \frac{P}{\bar{P}})} + Q_U, \tag{3.1}$$

where Q_R is the maximum responsive demand, Q_U is the unresponsive load, is the short-term price elasticity of demand, P is the dispatch price, and $\bar{P}$ is the average price. At equilibrium the responsive portion of demand has a symmetric shape, with an inflection point at $(\frac{Q_R}{2} + Q_U, \bar{P})$. The supply curve consists of a flat segment including must-run (usually zero-marginal cost) units and another sloping segment including non-zero bid units, which together form the *hockey stick* shape supply curve. An approximation for the supply function is [89]:

$$P(Q_G, Q_N) = c_1(1 - \frac{Q_G - Q_N}{Q_{max}})^{-c_2} + c_3, \tag{3.2}$$

where c_1, c_2 and c_3 are supply curve constants determined by the generation mixture, Q_{max} is the control area's maximum possible load with reserve requirements, and Q_N is non-dispatchable (must-run) generation. Figure 3.1 shows the clearing condition. The intersection of these curves (equilibrium point) is the stand-alone clearing condition (Q_S, P_S). The blue and red hatched areas are the consumer and producer surpluses respectively.

Figure 3.2 illustrates the impact of electricity export on the surplus. As the exports increase, the local consumption (Q_L) and consumer surplus decrease, while local production (Q_G) and producer surplus increase, as a result of the new higher clearing price (P_C). The total surplus increase is the solid red area. Similarly, for the importing jurisdiction, the production and producer surpluses decrease, while the local consumption, consumer and total surpluses increase. The producer surplus is obtained as:

$$\frac{P_C - c_3}{c_2 - 1}\ c_2\ Q_G - \frac{P_c - c_2\ c_3}{c_2 - 1}\ Q_N - \frac{P_c - c_1 - c_3}{c_2 - 1}\ Q_{max}. \tag{3.3}$$

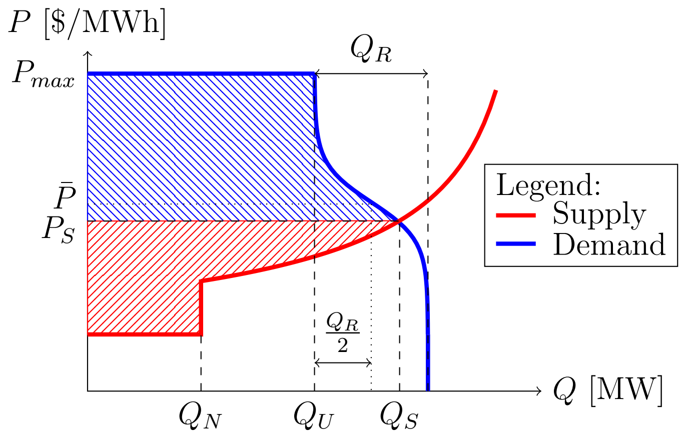


Fig. 3.1: Market clearing process.

The consumer surplus for the elastic demands is:

$$(P_{max} - P_C)(Q_U + Q_R) - \frac{Q_R\ \bar{P}}{2\ \eta} \ln \frac{1 + e^{2\ \eta}(1 - \frac{P_C}{\bar{P}})}{1 + e^{2\ \eta}(1 - \frac{P_{max}}{\bar{P}})}, \tag{3.4}$$

and for the inelastic demands is:

$$(P_{max} - P_C)\ Q_L. \tag{3.5}$$

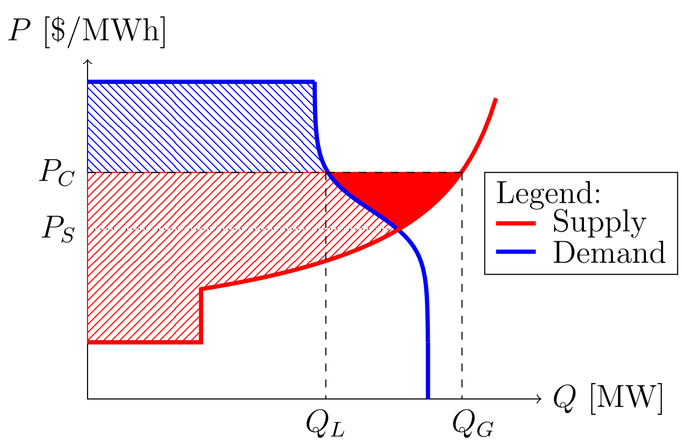


Fig. 3.2: The impact of electricity export on surplus.

3.2.2 Problem Description

The purpose of interconnection scheduler is to establish inter-area flows such that the system (global) surplus increases, while considering system constraints. In every area the surplus includes the stand-alone surplus, which is invariant with respect to inter-area flows, and the additional inter-change surplus, which is variable. The scheduler maximizes the summation of additional surplus over the entire interconnected system's areas:

$$\max_{f} \sum_{m=1}^{M} interchange\ surplus_m, \tag{3.6}$$

where f are the tie-line flows. The maximum surplus condition is achieved when the clearing price is uniform in the entire interconnection. However, there may be no flow solution that satisfies the ideal exports/imports within tie-line flow constraints. In this case, we seek the tie-line flow that increases the surplus as much as possible given the constraints. The ratio of the additional surplus to the maximum additional surplus is a metric that defines the transmission system performance, which we refer to as the economic utilization factor (EUF). Accordingly, the path EUF is zero for stand-alone (no exchange) condition, and is 100% for the *copper-sheet* (unconstrained flow) condition. Determination of the optimal flow is straightforward for a two-area system, because only one path exists. In hours that the optimal flow is beyond the path limit, the constraint is active, thus the actual flow should be truncated to the path transfer capacity, and the quantities and prices should be updated accordingly. Another performance metric, the economic power transfer distribution factor ($EPTDF$), is defined to quantify the sensitivity of flow with respect to price difference:

$$EPTDF = \frac{1}{8760} \sum_{m=1}^{M} \left(\frac{f_{1\to 2}}{P_{s2} - P_{s1}}\right)_t, \tag{3.7}$$

where P_{s1} and P_{s2} are stand-alone prices at hour t in Area 1 and Area 2, respectively. The proposed method is summarized in the following steps. First, the stand-alone (the minimum global surplus), copper-sheet (the maximum global surplus) and constrained-flow solutions are obtained, assuming the load is completely inelastic, for given hourly demand and must-run generation profiles. Notice that the non-flat segment of the supply curve is assumed fixed for the whole study year. The opportunity price, defined as the next hour price for any hour, is found. Second, the demand curve is stated based on the logistic function given in Equation 3.1, with an inflection point at the mean value of the hour price and the opportunity price, and a demand elasticity of –1. Third, the market clearing process is repeated with the elastic demand, and stand-alone demand and price are obtained. Next, the inter-change schedule that equalizes the prices is determined, and the annual maximum surplus increase is computed. At the end, the exchange is adjusted according to the path transfer capacity, and the actual surplus increase, the path EUF and $EPTDF$ are computed.

3.2.3 Model Inputs

An important assumption here is that each area has a uniform price, with no binding internal constraints that would result in different price zones within the area. This is not always a valid assumption for peak hours, but is easily remedied by increasing the number of areas defined in the model. The model inputs are taken from the WECC 2024 common case [34]. Figure 3.3 shows the forecasted hourly load duration curve for the WECC 2024 model. We aggregate the intermittent renewable electricity (wind, solar and run-of-river hydro) in each area, to find the must-run generation. These resources have zero-marginal cost. Table 3.1 shows the capacity and capacity factor of wind electricity in Area 1 (Canada) and Area 2 (United States), taken from [34].

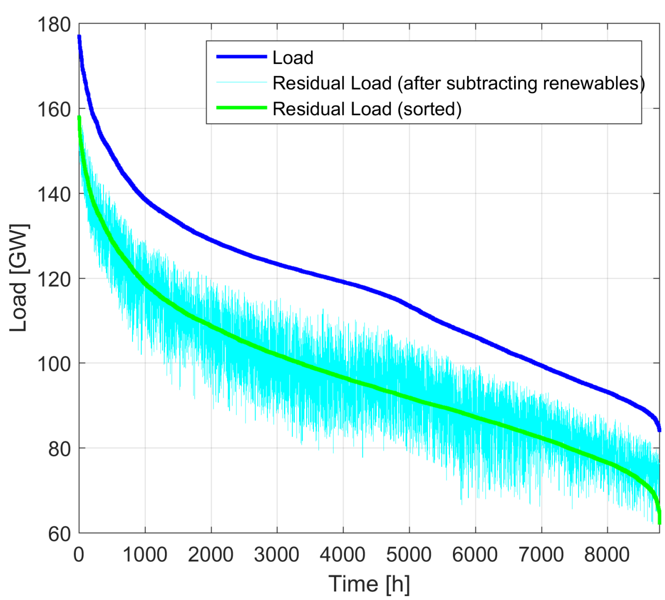


Fig. 3.3: WECC load duration curve.

Supply curve constants for a typical curve are taken from [90] as $c_1 = 4\$/MWh$, $c_2 = 2.6$ and $c_3 = 11\$/MWh$, and the market price cap is 1000 \$/MWh. Regarding

Table 3.1: Wind characteristics

Type	Area 1	Area 2
Capacity (MW)	3234	25918
Capacity factor (%)	33.26	28.53
Correlation coefficient	0.27	

the demand curve, we assume that the responsive portion of demand will be 10% if the cleared price is the same as the average price. In extremes, the magnitude of responsive demand will be: twice as the average price case if the price is zero; and zero if the price is at its cap.

3.3 Results and Discussion

In this section, the simulation results for year 2024 and 2030 are provided. Demand response implementation moderates the price, as is clearly shown in Figure 3.4 and Figure 3.5.

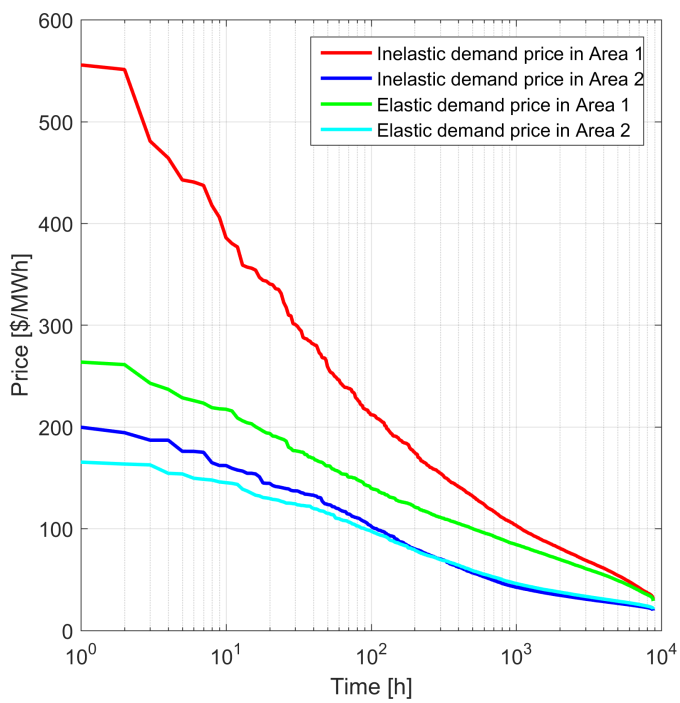


Fig. 3.4: Price duration curve (stand-alone) year 2024.

Figure 3.6 and Figure 3.7 show both unconstrained and constrained flow versus the difference between stand-alone prices under the inelastic and elastic loads. The

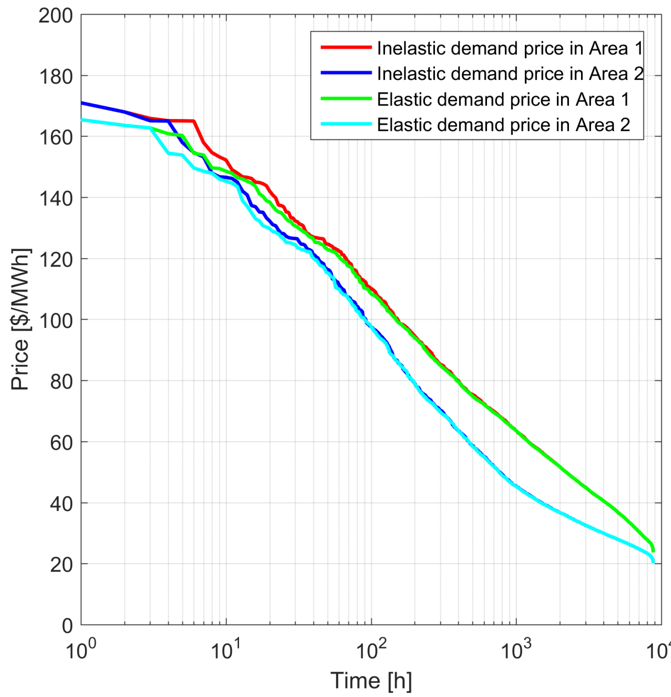


Fig. 3.5: Price duration curve (with transfer constraint) year 2024.

positive direction of path flow is from Area 1 to Area 2. The price difference is the stand-alone price in Area 2 minus the stand-alone price in Area 1. The unconstrained-flow $EPTDF$ is 80 MW/($/MWh), and constrained-flow is 40 MW/($/MWh) for the inelastic load. Similarly, the unconstrained-flow $EPTDF$ is 151 MW/($/MWh) and constrained-flow is 67 MW/($/MWh) for the elastic load. The price difference causes a relatively greater demand in the expensive side of the path, and a smaller demand in the cheap side of the path; therefore, increasing the path flow.

Figure 3.8 illustrates the path utilization duration curves. As expected, the magnitude of flow under the elastic demand case is greater than with the inelastic demand. The flow is truncated according to the path transfer capacity expansions.

The maximum annual surplus increase under the inelastic and the elastic demands are $703 M and $635 M, respectively. The annual surplus increase for the existing path transfer capacity under the inelastic and the elastic demands are $298 M and $156 M. The annual surplus increase (ASI), EUF and $EPTDF$ are summarized in Table 3.2. The path capacity expansion increases both EUF and $EPTDF$. A cost-benefit analysis would be required to determine the optimal transfer capacity of the path.

The same analysis is performed for the year 2030, assuming a case where wind

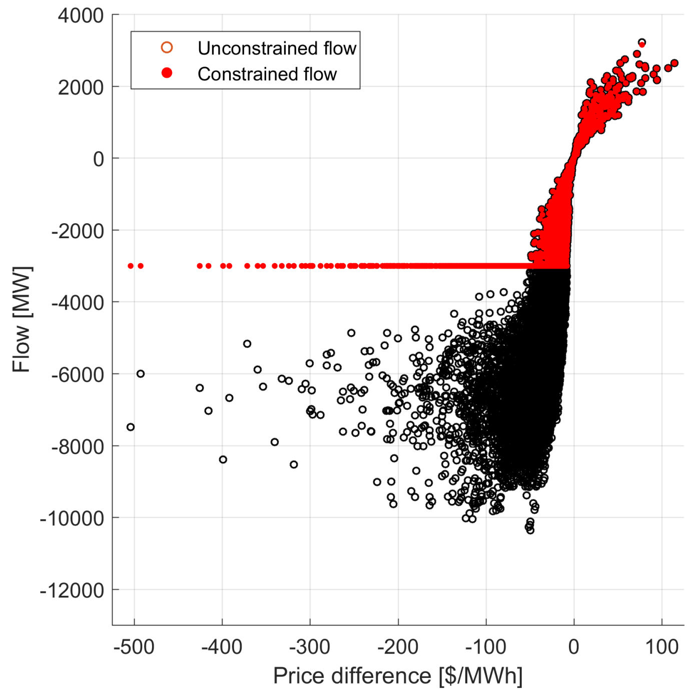


Fig. 3.6: Flow sensitivity to the price difference (inelastic load) year 2024.

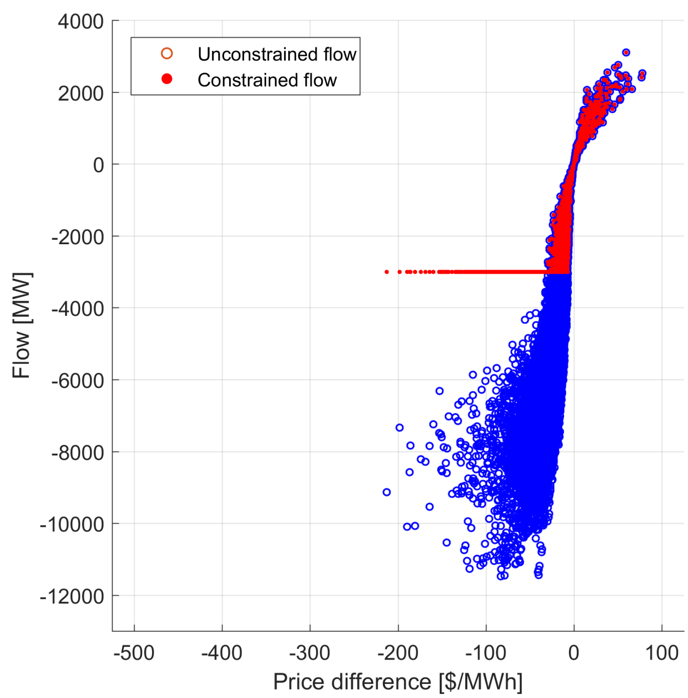


Fig. 3.7: Flow sensitivity to the price difference (elastic load) year 2024.

capacity increases such that the annual generation from additional wind is equal to the demand growth (0.7% per year [4]). Figure 3.9 and Figure 3.10 show the price

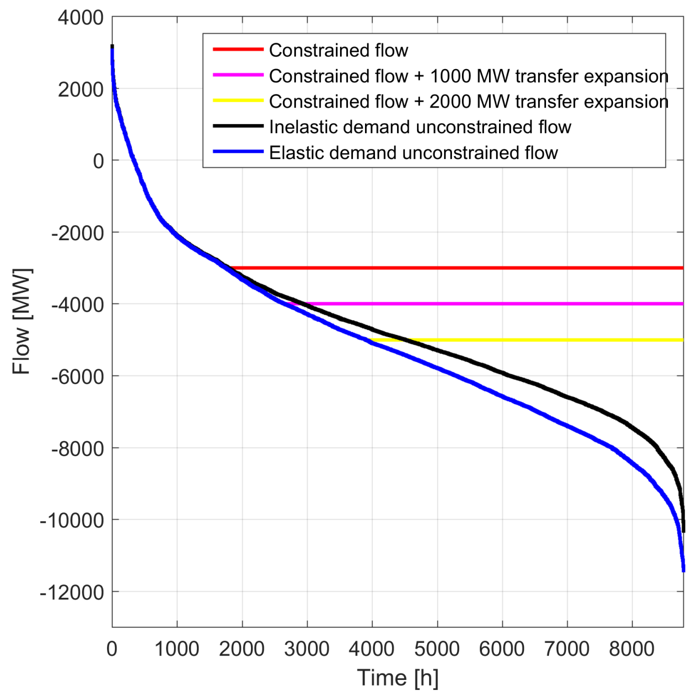


Fig. 3.8: Path utilization duration curve for year 2024.

Table 3.2: Results for year 2024

	Inelastic Demand			**Elastic Demand**		
Scenario	**ASI [M$]**	**EUF [%]**	**EPTDF [MW/($/MWh)]**	**ASI [M$]**	**EUF [%]**	**EPTDF [MW/($/MWh)]**
Stand-alone	0	0	0	0	0	0
Existing transfer capacity	298	42	40	196	31	67
1000 MW transfer expansion	430	61	52	302	47	88
2000 MW transfer expansion	545	78	63	406	64	107
Copper-sheet	703	100	88	635	100	151

duration curves under inelastic and elastic demands. In comparison to the year 2024, the market clearing price is lower. The reason is wind, which is a zero-marginal cost resource, is a greater portion of the generation mixture, and influences the market price more significantly. This means that as wind penetration level increases, the producer surplus decreases, and maybe some supply units cannot recover their investment costs.

Figure 3.11 illustrates the path utilization duration curves. In comparison with

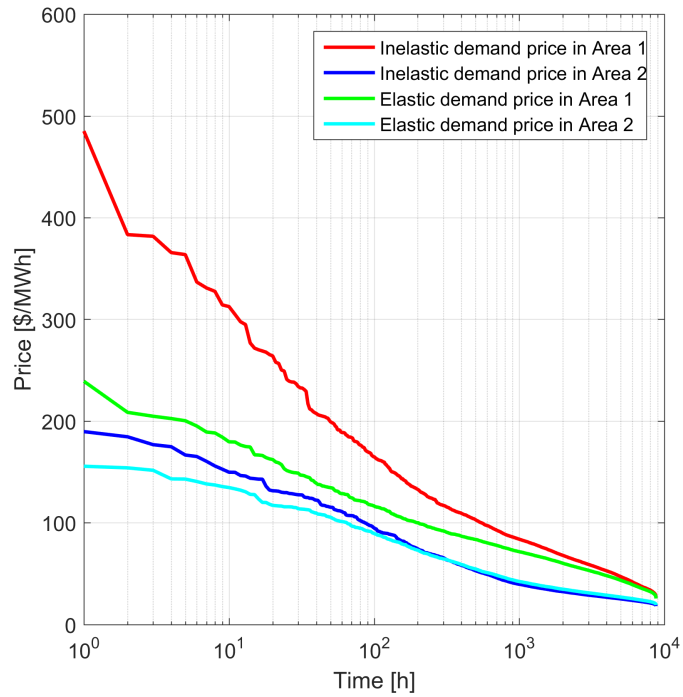


Fig. 3.9: Price duration curve (stand-alone) year 2030.

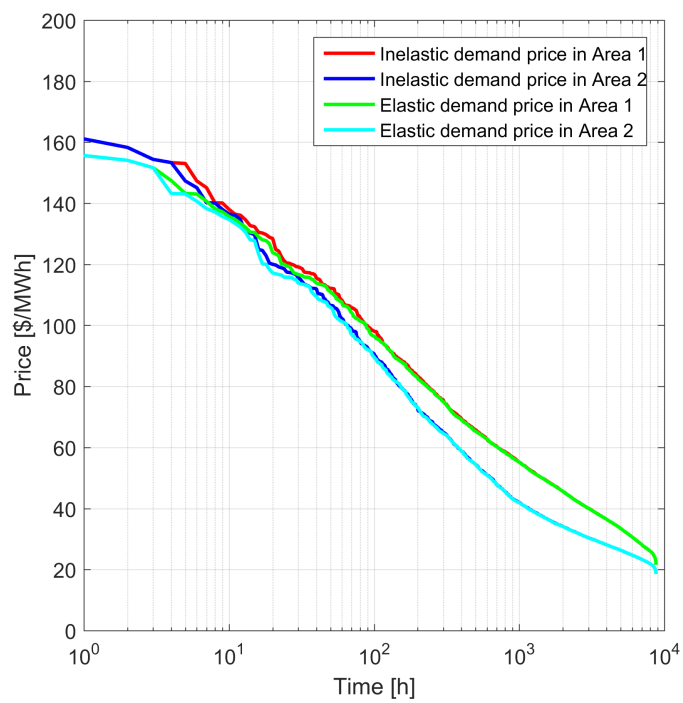


Fig. 3.10: Price duration curve (with transfer constraint) year 2030.

year 2024, the magnitude of the optimal flow is greater, because the demand and the wind generation are greater, and a greater flow is required to balance market prices.

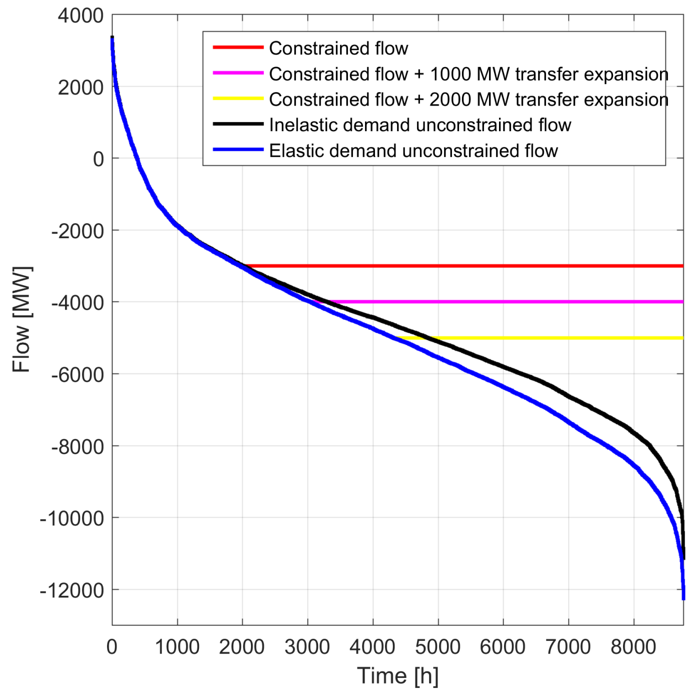


Fig. 3.11: Path utilization duration curve for year 2030.

Simulation results for 2030 are listed in Table 3.3. In comparison to 2024, the EUF is almost the same, but the $EPTDF$ increases, because the flow is greater and the price difference is smaller in 2030 relatively.

Table 3.3: Results for year 2030

	Inelastic Demand			**Elastic Demand**		
Scenario	**ASI** [M$]	**EUF** [%]	**EPTDF** [MW/($/MWh)]	**ASI** [M$]	**EUF** [%]	**EPTDF** [MW/($/MWh)]
Stand-alone	0	0	0	0	0	0
Existing transfer capacity	232	42	54	160	32	84
1000 MW transfer expansion	333	60	69	243	48	109
2000 MW transfer expansion	419	76	83	323	64	132
Copper-sheet	549	100	106	502	100	185

3.4 Conclusion

This paper assesses the impact of demand response integration and transmission capacity expansion on the electricity generation and inter-change scheduling. A surplus maximizing scheduling method is proposed and tested on a hypothetical system consisting of two consolidated areas that loosely represents the Western Interconnection. For given 2024 forecast data with hourly resolution, the stand-alone, copper-sheet (unconstrained-flow) and constrained-flow schedules are determined and compared. The additional annual surplus is \$196 M under the elastic demand, for the existing path transfer capacity, which is only one-third of the maximum possible surplus increase. The path economic utilization factor and the economic power transfer distribution factor shows the sensitivity of the flow to the price difference is found 67 MW/(\$/MWh). The results for the year 2030, where additional wind electricity supplies the demand growth alone, show that a greater penetration of wind causes the path to be utilized relatively more. Future work is to develop a multi-area model for the WECC interconnection, and solve the constraint flow problem.

Chapter 4

Interconnection-wide Hour-ahead Scheduling in the Presence of Intermittent Renewables and Demand Response: a Surplus Maximizing Approach

This paper was accepted to Applied Energy Journal in December 2016:

Sahand Behboodi, David P Chassin, Ned Djilali and Curran Crawford. *Interconnection-wide Hour-ahead Scheduling in the Presence of Intermittent Renewables and Demand Response: a Surplus Maximizing Approach* Applied Energy 189 (2017): 336-351. Available online at: http://www.sciencedirect.com/science/article/pii/S0306261916318165

Sahand Behboodi has done the major part of defining the problem, formulating the objective function, performing the optimization, and writing the paper. David Chassin has been a help to Sahand to code the model in Matlab. David has also modified the manuscript.

This chapter addresses the topic of resource allocation in an interconnected system in presence of significant amount of intermittent renewables and demand response. A new approach to determine the optimal inter-area transfer schedule is presented that maximizes the system's economic surplus. This method is demonstrated on a reduced model of the North America Western Interconnection.

Abstract

This paper describes a new approach for solving the multi-area electricity resource allocation problem when considering both intermittent renewables and demand response. The method determines the hourly inter-area export/import set that maximizes the interconnection (global) surplus satisfying transmission, generation and load constraints. The optimal inter-area transfer set effectively makes the electricity price uniform over the interconnection apart from constrained areas, which overall increases the consumer surplus more than it decreases the producer surplus. The method is computationally efficient and suitable for use in simulations that depend on optimal scheduling models. The method is demonstrated on a system that represents North America Western Interconnection for the planning year of 2024. Simulation results indicate that effective use of interties reduces the system operation cost substantially. Excluding demand response, both the unconstrained and the constrained scheduling solutions decrease the global production cost (and equivalently increase the global economic surplus) by \$12.30B and \$10.67B per year, respectively, when compared to the standalone case in which each control area relies only on its local supply resources. This cost saving is equal to 25% and 22% of the annual production cost. Including 5% demand response, the constrained solution decreases the annual production cost by \$10.70B, while increases the annual surplus by \$9.32B in comparison to the standalone case.

Keywords

Demand response, energy market, renewable intermittency, resource allocation, western interconnection

Highlights

1. A new approach for electricity resource allocation that includes price-elastic loads

2. A new model of interconnection-scale scheduling that maximizes economic surplus

3. A demonstration of the scheduling method on the North America Western Interconnection

Nomenclature

A	Market state solution matrix.
b	Market condition vector.
d	Demand curve slope, in \$/(MWh.MW).
e	Net export, in MW.
f	Transfer flow, in MW.
p	Price, in \$/MWh.
p_{max}	Must-serve load price, in \$/MWh.
p_{min}	Must-take generation price, in \$/MWh.
q	Quantity, in MW.
s	Supply curve slope, in \$/(MWh.MW).
x	Market state vector.
Y	Connectivity matrix.
α	Degree of demand inelasticity.
Δ	Difference operator.
ω	Combined demand and supply slope, in \$/(MWh.MW).
Ω	Diagonal matrix of combined slopes.
Subscripts	
0	Standalone
c	Clearing
d	Demand
p	Price
q	Quantity
r	Responsive
s	Supply
u	Unresponsive
w	Must-take

4.1 Introduction

Most jurisdictions in North America have adopted renewable energy portfolio policies as part of efforts to reduce greenhouse gas emissions. The inherent intermittency of renewables is the main challenge to the large-scale integration of these clean resources at high penetration levels. The traditional utility approach to generation variability is to operate reserve units, which are usually more costly and may increase emissions. Demand response is a zero-emission and potentially lower-cost alternative to the use of generation reserves. It also benefits the flexible load through payment for their services, and benefits all consumers through lowered electricity costs. The US Department of Energy has adopted a definition of demand response that is now widely recognized for its inclusiveness [14]: "load variations in response to changes in both financial incentives and/or reliability signals over time".

The idea of including demand response in electricity markets is discussed in a large body of recent works. The impact of demand response integration on peak energy consumption, energy price and emissions under load uncertainty is analyzed in [91]. A model of demand response participation in real-time markets to minimize the operation cost considering the load elasticity is formulated in [92]. The interaction between renewable intermittency and demand response in the market environment is investigated in [50].

Load fluctuations and renewable generation intermittency are generally not strongly correlated with each other over a large interconnected system that includes multiple balancing authorities [9]. As a result, the combined interconnection power fluctuations are smaller than the sum of the variations in individual balancing authorities. Neighbouring jurisdictions can take advantage of the geographical diversity of renewable resources within the system, and cooperate more effectively to mitigate the intermittency of renewable power generation. This cooperation, which is beneficial from both reliability and economic viewpoints, requires an enhanced transmission system, sometimes referred to as a "supergrid" [8]. A recent study of consolidation of balancing authorities in the US [4] showed that if planners moved away from a regionally divided electricity system to a national system using high-voltage direct-current transmission lines then the deployment of wind and solar power could reduce CO_2 emissions by up to 80% relative to 1990 levels, without an increase in electricity price.

Resource scheduling using locational marginal price (LMP) has been the foundation of modern electricity system operations since the early 1980s when it was first introduced [93]. The basic LMP solution was subsequently extended to perform security constrained economic dispatch (SCED) to satisfy operational constraints. This family of solutions has been deployed very successfully by transmission system operators [94]. However, the LMP formulation considers load to be essentially inelastic. Approaches to compensating demand response that allow consideration of price sensitive loads have been examined [95]. For policy-makers seeking to study the widespread development of renewable resources and the impact of demand response in system operation, the preferred LMP/SCED solution to the resource scheduling problem presents a significant barrier to adoption because the system models are typically constructed in a manner that assumes: (i) the system operation is dominated by supply resources with significant and relatively consistent fixed and variable cost components throughout an interconnected system, and (ii) demand is essentially inelastic and predictable. Solutions to the demand response problem include those proposed by the US Federal Energy Regulatory Commission [96]. Unfortunately, renewable resources such as wind and solar do not fulfill assumption (i), and short-term redispatchable demand does not conform well to assumption (ii).

A deep understanding of the interconnection-scale impact of demand response integration is difficult to achieve in the absence of accurate resource allocation models that properly consider the system-wide impact of demand response on locational energy price calculations and generation resource allocation. This is even more important for the case of large interconnected systems where mixed pricing mechanisms are extant, such as in the Western Electricity Coordinating Council (WECC). In the WECC some regions have fully developed energy markets and others do not, and multiple balancing authorities operate and interact through a myriad of bilateral contracts and other financial arrangements including some as obscure as the Columbia River Treaty [97]. In an effort to address these barriers and to study optimal operation of large-scale interconnections, we are motivated to find a more flexible and general model of the resource scheduling problem based on energy pricing. In the absence of price sensitive loads, the problem of unit commitment is to determine the hourly generation schedule in a way that minimizes the operational costs, which equivalently maximizes the economic surplus (social welfare) [76, 98]. Therefore solving the traditional LMP problem is sufficient. However, when a significant amount of price sensitive loads is present, minimizing cost is no longer a satisfactory objective, and maximizing surplus is preferred [87], as described in Section 4.2. Surplus maximization for the unit commitment problem has been already formulated at the balancing authority level [18, 46, 99]. In the present work, we are interested in analyzing this problem at the interconnection level. More precisely, we seek a set of inter-area power transfers that maximizes the global surplus, which is defined as the sum of consumer and producer surpluses over all balancing authorities in an interconnection.

The optimal operation of the interconnection helps utilities produce electricity with a lower cost, integrate more intermittent renewables, and defer or cancel costly investments in grid infrastructure. Previous work [11] has shown the potential annual savings in production cost due to consolidation of balancing authorities ranges from 2.4% to 3.2%, considering transmission congestions. The full coppersheet consolidation of the WECC system provides an additional 1.4% improvement. However this study did not consider the impact of demand response on system resource allocation. Load management assists the WECC system operators in dealing with uncertainty in demand and intermittent resource output [100].

We consider one important reference to be the inelastic demand scenario in which the WECC system as a whole is operated in the most economically efficient manner. This scenario is unlikely because of various jurisdictional regulations, but it does provide an upper bound for the achievable system-wide economic benefits considering the transmission constraints. Thus, demand response with different penetration levels across the system, and potential surplus increases are evaluated with respect to this best achievable performance absent demand response. It should be noted that the outcome of this work is a model to study system-wide scenarios for planning rather than proposing an operational tool to be used in WECC.

Energy scheduling and reserve scheduling can be performed simultaneously to achieve a more economically-efficient use of both supply and demand resources, particularly when reserve sharing is anticipated to be significant [101, 102]. We exclude reserve scheduling in this formulation/simulation for simplicity. But we recognize that including it is an essential future work.

The new contributions of this work are: (i) a new resource allocation method that incorporates the market-driven demand response into the unit commitment problem; (ii) a new model of the interchange export/import scheduling problem for the global (interconnection-wide) surplus maximization objective; and (iii) a demonstration of the proposed model on the Western Interconnection for the planning year 2024. In Section 4.2 an overview on the double auction market theory is provided, and in Section 4.3 the scheduling problem considering demand price responsiveness is formulated. In Section 4.4, a reduced model of the WECC system is presented, and in Section 4.5, simulation results are discussed.

4.2 Electricity Markets

The ultimate purpose of the proposed model of resource allocation is to determine the hourly schedule for the most economically efficient mix of supply and demand resources that does not violate the transfer constraints on system interties within an interconnection. This hourly schedule is essential to establishing the intra-hourly control reference for both supply and demand resource dispatch, which is needed to evaluate the performance of short-term demand response control strategies. To obtain this schedule, an hour-by-hour virtual double auction wholesale market is assigned to each of the system's control areas in which both supply and demand resources participate. Although the magnitude of the responsive load is relatively small, it is still worth considering its impact on the optimal schedule, particularly in the presence of substantial intermittent renewable resources. One key advantage of using the proposed model is that it facilitates modeling of inter-temporal effects such as demand response by load shifting and recovery after load curtailment.

4.2.1 Market clearing process

We begin by considering how prices and quantities are computed in an area. In general, both supply and demand curves are composed of a series of linear segments (sell or buy bidding quantities/prices). For simplicity, we assume supply curves only include a flat segment for non-dispatchable generators (e.g. wind and nuclear power) and a variant segment for dispatchable generators. Similarly, we assume demand curves consist of a flat segment for must-serve loads and a variant segment for deferrable loads (e.g. electric vehicles). The supply and demand curves are given as:

$$p_s = \begin{cases} s\ (q_s - q_w) + p_{min} & q_s > q_w \\ p_{min} & \text{otherwise,} \end{cases} \tag{4.1a}$$

and

$$p_d = \begin{cases} d\ (q_d - q_u) + p_{max} & q_d > q_u \\ p_{max} & \text{otherwise,} \end{cases} \tag{4.1b}$$

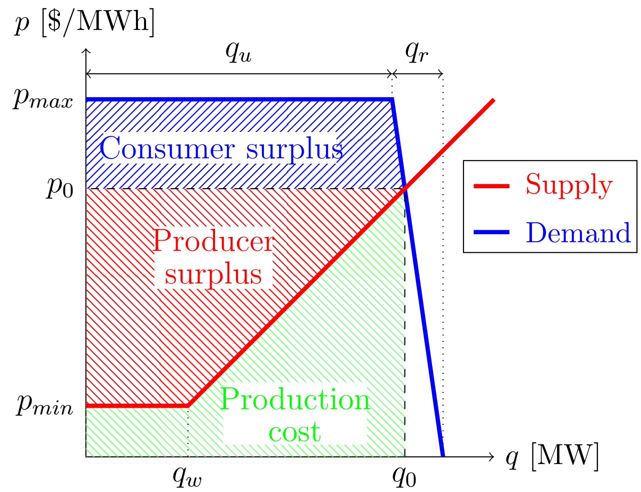


Fig. 4.1: A double auction electricity market.

respectively, where p_s is the supply price, $s > 0$ is the slope of the supply curve, q_s is the supply quantity, q_w is must-take (non-dispatchable) generation, p_{min} is the must-take price, p_d is the demand price, $d < 0$ is the slope of the demand curve, q_d is the demand quantity, q_u is must-serve load, and p_{max} is the must-serve price. The slope of an inelastic demand curve is infinite ($d \to -\infty$). Supply and demand curves are illustrated in Figure 4.1.

The supply and demand prices, p_s and p_d respectively, of energy as well as the supply and demand quantities, q_s and q_d respectively, of power must satisfy the linear system:

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{4.2a}$$

where the matrix $\mathbf{A}$ represents the area supply and demand, the vector $\mathbf{x}$ represents the prices and quantities for supply and demand, and the vector $\mathbf{b}$ represents the solution condition. In the case of the simple linear supply and demand system above, we have:

$$\mathbf{A} = \begin{bmatrix} -d & 1 & 0 & 0 \\ 0 & 0 & -s & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} q_d \\ p_d \\ q_s \\ p_s \end{bmatrix}, \text{ and } \quad \mathbf{b} = \begin{bmatrix} p_{max} - d\ q_u \\ p_{min} - s\ q_w \\ \Delta q \\ \Delta p \end{bmatrix}, \tag{4.2b}$$

where Δq is the export (or import if $\Delta q < 0$) quantity and Δp is the price subsidy (or tax if $\Delta p < 0$). Note that if both s and d are infinite, this formulation is not appropriate.

The solution price and quantity at which supply equals demand in a standalone area (both Δq and Δp are zero) is found by solving the *standalone area problem*:

$$\mathbf{x}_0 = \mathbf{A}^{-1}\mathbf{b} = \begin{bmatrix} q_0 & p_0 & q_0 & p_0 \end{bmatrix}^{\mathrm{T}}. \tag{4.2c}$$

The supply and demand quantities for a given price p can be found by solving the *quantity problem*:

$$\mathbf{x}_p = \mathbf{A}_p{}^{-1}\mathbf{b}_p = \begin{bmatrix} q_d & p & q_s & p \end{bmatrix}^{\mathrm{T}}, \tag{4.3a}$$

where

$$\mathbf{A}_p = \begin{bmatrix} -d & 1 & 0 & 0 \\ 0 & 0 & -s & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \text{ and } \mathbf{b}_p = \begin{bmatrix} p_{max} - d\ q_u \\ p_{min} - s\ q_w \\ p \\ p \end{bmatrix}. \tag{4.3b}$$

Similarly, the supply and demand prices for a given quantity q can be found by solving the *price problem*:

$$\mathbf{x}_{\mathbf{q}} = \mathbf{A}_q{}^{-1}\mathbf{b}_q = \begin{bmatrix} q & p_d & q & p_s \end{bmatrix}^{\mathrm{T}}, \tag{4.4a}$$

where

$$\mathbf{A}_q = \begin{bmatrix} -d & 1 & 0 & 0 \\ 0 & 0 & -s & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and } \mathbf{b}_q = \begin{bmatrix} p_{max} - d\ q_u \\ p_{min} - s\ q_w \\ q \\ q \end{bmatrix}. \tag{4.4b}$$

Another important problem required to determine the multi-area schedule is finding the price at which the difference between supply and demand quantities (export) is Δq. The clearing price, supply and demand quantities can be found by solving the *export problem* (or *import problem* for $\Delta q < 0$):

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}_{\Delta q} = \begin{bmatrix} q_d & p & q_s & p \end{bmatrix}^{\mathrm{T}}, \tag{4.5a}$$

where

$$\mathbf{b}_{\Delta q} = \begin{bmatrix} p_{max} - d\ q_u \\ p_{min} - s\ q_w \\ \Delta q \\ 0 \end{bmatrix}. \tag{4.5b}$$

In addition, if there is a difference between supply and demand prices such as a renewable subsidy or a carbon tax, the clearing quantity, supply and demand prices can be found by solving the *subsidy problem* (or *tax problem* for $\Delta p < 0$):

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}_{\Delta p} = \begin{bmatrix} q & p_d & q & p_s \end{bmatrix}^{\mathrm{T}}, \tag{4.6a}$$

where

$$\mathbf{b}_{\Delta p} = \begin{bmatrix} p_{max} - d\ q_u \\ p_{min} - s\ q_w \\ 0 \\ \Delta p \end{bmatrix}. \tag{4.6b}$$

4.2.2 Surplus calculation

The double auction market finds the quantity and price that maximize the total economic surplus, which is geometrically the area between the supply and demand curves, as shown in Figure 4.1. By definition, the total surplus is the summation of the producer surplus (the red hatched area) and the consumer surplus (the blue hatched area) [76]. The green hatched area is the electricity production cost.

The idea of global surplus maximization is presented through an example. Consider two control areas with identical demand curves and slightly different supply curves, as illustrated in Figure 4.2. The must-take supply resource is 2000 MW in Area A, and 1000 MW in Area B. If these areas are not connected, their markets will be cleared under the standalone condition as shown in Table 4.1.

When we consider the unconstrained transfer problem, the generation units in Area A that did not get dispatched when the tieline was not operating are now dispatched to serve additional load in Area B. Despite Area A, the generation in Area B decreases up to the point where the price difference between two areas is zero as illustrated in Figure 4.3. Notice that the clearing price increases in Area A and decreases in Area B. As a result, demand quantity decreases in Area A and increases in Area B such that the difference between supply and demand quantities is 500 MW in Area A and -500 MW in Area B. Electricity exchange, like any other economic transaction, increases the global surplus (social welfare), as tabulated in Table 4.2. The maximum global surplus which is obtained under this condition is referred to as the *coppersheet solution.*

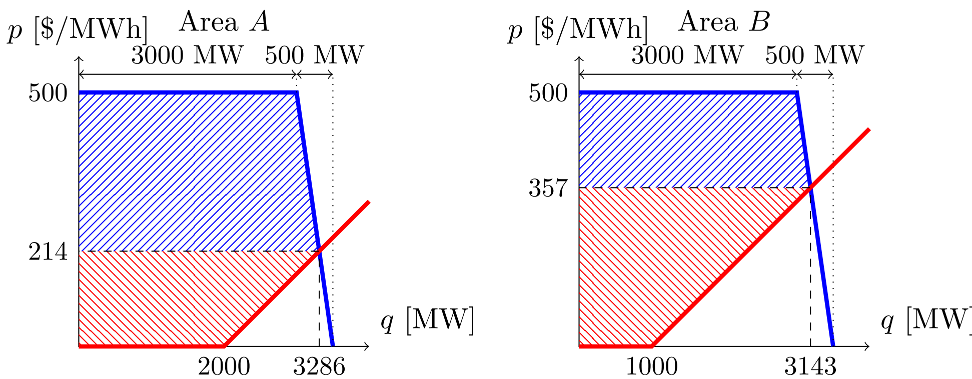


Fig. 4.2: Standalone markets.

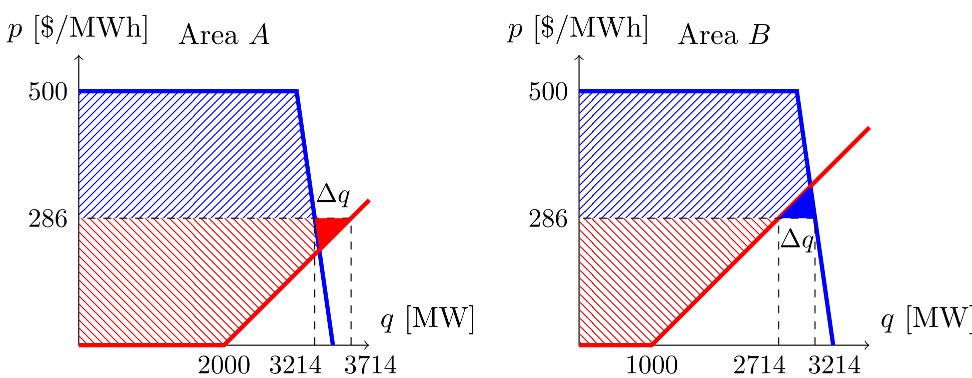


Fig. 4.3: Interconnected markets with unconstrained transfer capacity.

Now we constrain the tieline's transfer capacity to 400 MW, which is insufficient to equalize the prices, as illustrated in Figure 4.4. Table 4.3 shows that the global surplus associated with the constrained transfer solution ($2666k) is between the standalone ($2643k) and coppersheet surpluses ($2679k). A surplus-maximizing scheduler identifies the tieline flow such that it either zeros out the price difference or fully utilizes the transfer capacity. This scheduling problem becomes more complex when the number of areas and tielines is increased.

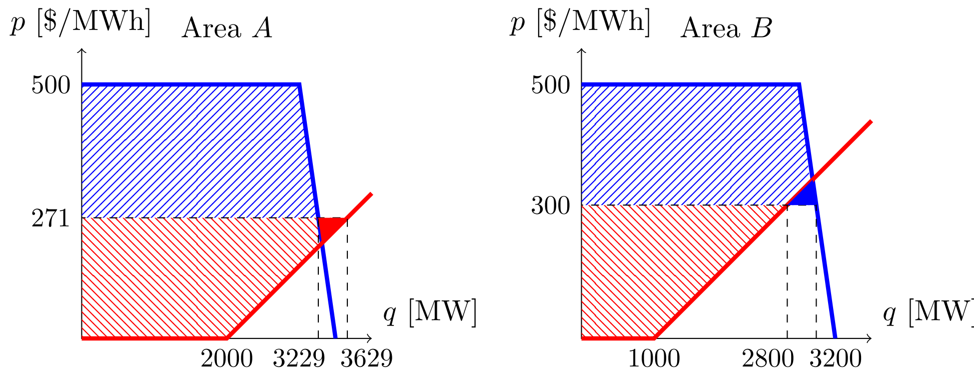


Fig. 4.4: Interconnected markets with constrained transfer capacity.

Table 4.1: Standalone schedule and surplus (zero MW exchange)

Control area	Price [$/MWh]	Demand [MW]	Supply [MW]	Consumer sur. [k$/h]	Producer sur. [k$/h]	Total sur. [k$/h]
A	214	3286	3286	898	566	1464
B	375	3143	3143	439	740	1179
Total	284	6429	6429	1337	1306	2643

Table 4.2: Coppersheet schedule and surplus (500 MW exchange)

Control area	Price [$/MWh]	Demand [MW]	Supply [MW]	Consumer sur. [k$/h]	Producer sur. [k$/h]	Total sur. [k$/h]
A	286	3214	3714	666	816	1482
B	286	3214	2714	666	531	1197
Total	286	6428	6428	1332	1347	2679

Table 4.3: Constrained transfer schedule and surplus (400 MW exchange)

Control area	Price [$/MWh]	Demand [MW]	Supply [MW]	Consumer sur. [k$/h]	Producer sur. [k$/h]	Total sur. [k$/h]
A	271	3629	3229	713	763	1476
B	300	2800	3200	620	570	1190
Total	284	6429	6429	1333	1333	2666

4.3 Inter-area Transfer Scheduling

In this section, we formulate the inter-area transfer scheduling problem, and determine the export/import schedule that maximizes the global surplus excluding and including transmission constraints, i.e., the coppersheet and constrained transfer solutions, respectively.

4.3.1 Surplus maximization

The objective function maximizes the global surplus by varying the local supply and demand dispatch quantities over the interconnection. The optimization effectively aims to maximize the sum of the areas circumscribed by the triangles bounded by the supply and demand curves and the price after export/import (see Figure 4.3 and Figure 4.4). Because the standalone surplus is invariant, it can be ignored in the formulation for surplus maximization. The optimization problem for an N-control area system is stated as:

$$\max_{q_{s_n},\ q_{d_n}} \sum_{n=1}^{N} \frac{1}{2}(p_{c_n} - p_{0_n})\ (q_{s_n} - q_{d_n}), \tag{4.7a}$$

subject to:

$$q_{u_n} \leq q_{d_n} \leq q_{u_n} + q_{r_n}. \tag{4.7b}$$

Note that supply and demand constraints are not needed because they are fully captured by the supply and demand curves. The objective function can be simplified to (see Appendix 1):

$$\min \frac{1}{2} \sum_{n=1}^{N} \omega_n\ e_n^2, \tag{4.7c}$$

where e_n is the difference between the supply and the demand quantities (net export), and $\omega_n = \frac{s_n\ d_n}{s_n - d_n}$ is the combined demand and supply slope in area n.

Negative prices are allowed in the optimization, but the surplus calculation is modified whenever the clearing price before power exchange (p_0) or after (p_c) is negative. As illustrated in Figure 4.5, the positive-price region (red area) is considered as the true surplus increase, excluding the negative-price region (yellow area). If the standalone price is negative in an area (left side market), as long as the export cannot push the clearing price beyond 0 \$/MWh, the producer surplus is zero, and the supply quantity equals the maximum demand ($q_u + q_r$) plus the export (e). Notice that the resource allocation is at the hour-ahead scheduling level not at the real-time operation level, thus a situation in which the generation exceeds the required power does not occur so that generators never pay a penalty for over-production simply because they can never be allowed to over-produce. The actual surplus appears beyond 0 \$/MWh, where the supply quantity exceeds the local must-take generation. When the clearing price is negative (right side market), the actual import will be the dif-

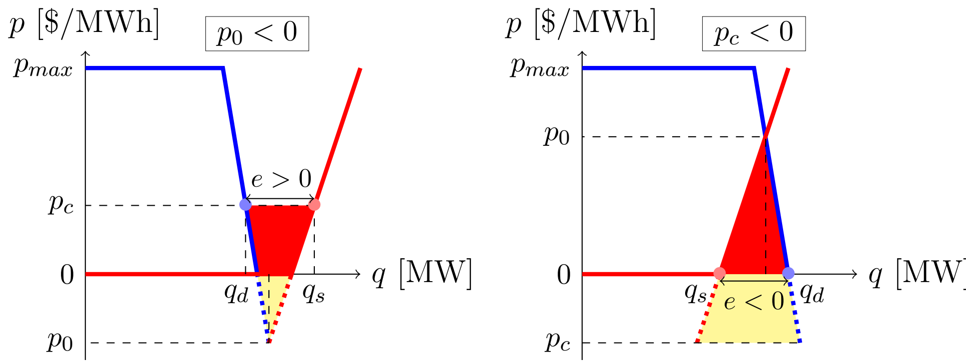


Fig. 4.5: Surplus calculation with a negative price.

ference between the maximum demand $(q_u + q_r)$ and the local must-take generation (q_s), which happens at 0 \$/MWh. Thus, the surplus increase beyond this price is virtual and will never be realized during operation. If both prices are negative, the surplus increase is zero—if an area has too much excess power generation such that after exporting to its neighboring jurisdictions the clearing price is still negative, the area's income from exporting is zero.

4.3.2 Coppersheet solution

Under the maximum global surplus, we expect the price to be the same over the entire interconnection, which gives the most economically efficient allocation of supply and demand resources. To obtain this *coppersheet price*, we construct the coppersheet supply curve by combining all supply curves, and the coppersheet demand curve by combining all demand curves. From Equation 4.2c, the intersection point determines the coppersheet quantity $\tilde{q}$ and price $\tilde{p}$:

$$\begin{bmatrix} \tilde{q} \\ \tilde{p} \\ \tilde{q} \\ \tilde{p} \end{bmatrix} = \begin{bmatrix} -\tilde{d} & 1 & 0 & 0 \\ 0 & 0 & -\tilde{s} & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{cap} - \tilde{d} \sum_{n=1}^{N} q_{u_n} \\ \tilde{s} \sum_{n=1}^{N} q_{w_n} \\ 0 \\ 0 \end{bmatrix}, \tag{4.8a}$$

where coppersheet supply and demand slopes are:

$$\tilde{s} = \left(\sum_{n=1}^{N} s_n^{-1} \right)^{-1}, \tag{4.8b}$$

and

$$\tilde{d} = \left(\sum_{n=1}^{N} d_n^{-1} \right)^{-1}, \tag{4.8c}$$

respectively. We assume the must-take and must-serve prices are the same everywhere, e.g., the minimum price of zero and the maximum price of the market cap. With the coppersheet price we obtain the ideal net exports, $\tilde{\mathbf{e}} = \begin{bmatrix} \tilde{e}_1 & \cdots & \tilde{e}_n \end{bmatrix}^{\mathrm{T}}$, from Equation 4.5a.

We can find the unconstrained transfer flow solution using the graph theory framework. The connectivity matrix for an N-area L-line interconnection is:

$$\mathbf{Y} = \begin{bmatrix} y_{(1,1)} & \cdots & y_{(1,L)} \\ \vdots & \ddots & \vdots \\ y_{(N,1)} & \cdots & y_{(N,L)} \end{bmatrix}, \tag{4.9a}$$

where the value for a line l from area n to area m is $y_{(n,l)} = +1$ when $n < m$ and $y_{(n,l)} = -1$ when $n > m$, and $y_{(n,l)} = 0$ when $n = m$ or no line connects areas n and m. Given the matrix $\mathbf{Y}$ and the ideal net export set $\tilde{\mathbf{e}}$ we can solve for the unconstrained flows on the L lines:

$$\tilde{\mathbf{f}} = \mathbf{Y}^{+}\,\tilde{\mathbf{e}} = \begin{bmatrix} \tilde{f}_1 & \cdots & \tilde{f}_L \end{bmatrix}^{\mathrm{T}}, \tag{4.9b}$$

where $\mathbf{Y}^{+}$ is the Penrose pseudo-inverse of $\mathbf{Y}$. Among the possible flow solutions that result in the ideal net exports, this solution has the smallest 2-norm [103], so arguably the transmission loss is minimum if tielines have similar loss factors.

4.3.3 Constrained-transfer solution

As described above, the maximum global surplus is achieved under a condition when the exports are those obtained from the coppersheet solution. This solution may assign a flow to a tieline over its transfer limit. In this case, we seek a flow set that does not violate transmission constraints but still results in the ideal net export values. If such a flow set does not exist, we will seek the flow set that results in a global surplus as close to the maximum surplus as possible. The objective function can be rewritten as:

$$\min_{\mathbf{e}} \mathbf{e}^{\mathrm{T}}\,\boldsymbol{\Omega}\,\mathbf{e} = \min_{\mathbf{f}} (\mathbf{Y}\,\mathbf{f})^{\mathrm{T}}\,\boldsymbol{\Omega}\,(\mathbf{Y}\,\mathbf{f}), \tag{4.10a}$$

subject to:

$$\mathbf{f_{min}} \leq \mathbf{f} \leq \mathbf{f_{max}}, \tag{4.10b}$$

where $\boldsymbol{\Omega}$ is the $N \times N$ diagonal matrix of $\begin{bmatrix}\omega_1 & \cdots & \omega_N\end{bmatrix}$. In any feasible solution, as long as there is a price gradient on a tieline its associated transfer flow equals its limit:

$$f_{n\to m} = \begin{cases} f_{n\to m_{min}} & \text{for} \quad p_m < p_n, \\ f_{n\to m_{max}} & \text{for} \quad p_m > p_n. \end{cases} \tag{4.10c}$$

The price gradient vector $\delta\mathbf{p}$ is obtained from:

$$\delta\mathbf{p} \;= -\mathbf{Y}^{\mathrm{T}}\;\mathbf{p}, \tag{4.10d}$$

where $\mathbf{p}$ is the price vector, calculated from (see Appendix 1):

$$\mathbf{p} \;= \mathbf{p}_0 \;- \boldsymbol{\Omega}\;\mathbf{e} \;=\; \mathbf{p}_0 \;- \boldsymbol{\Omega}\;\mathbf{Y}\;\mathbf{f}. \tag{4.10e}$$

We use a sequential quadratic programming (SQP) method [104] implemented by Matlab's optimization toolbox function `fmincon` [105] to solve this optimization problem and determine the optimum flow set. The objective function is convex, and also the objective and constraints are differentiable functions, thus the solution to the optimization is the global optimum [106]. Perhaps a good initial point to feed into the solver is the coppersheet solution with overloaded tielines adjusted to their capacity limits.

4.4 Western Interconnection

In this section, we present an interconnected system model that loosely represents the Western Interconnection. The WECC system extends from Canada to Mexico and includes the provinces of Alberta and British Columbia in Canada, the northern portion of Baja California in Mexico, and all or portions of 14 western US states. As of March 2016, there are 38 control areas (balancing authorities) in the WECC system [107]. California and Alberta are the only regions with full energy markets in the interconnection.

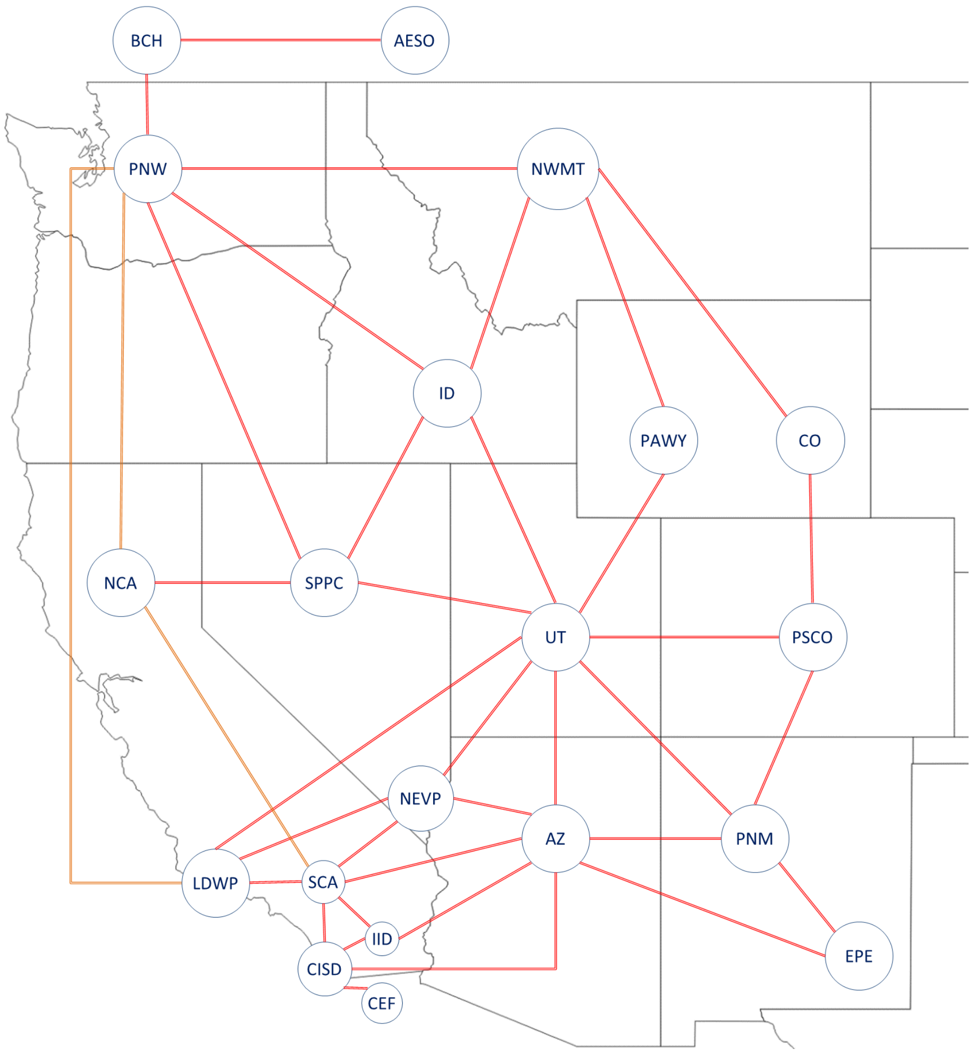


Fig. 4.6: The bubble pipeline view of the 20-consolidated area WECC model.

4.4.1 Bulk system model

A 20-area reduced model of the WECC system, illustrated in Figure 4.6, is built by combining two or more balancing authorities into one consolidated area. The system interties (groups of transmission lines) are assumed lossless. The intertie transfer limits are given in Appendix 2, estimated from the WECC 2024 common case [34]. In this model, the internal transmission constraints within the control areas are ignored, although area demand is scaled up by a constant factor that approximately accounts for these losses. The peak load and annual energy consumption forecasts as well as the internal loss factors are provided in Table 4.4, extracted from the WECC 2024 common case [34]. Intermittent (wind, solar and run-of-river hydro), base and

Table 4.4: Demand forecast and internal loss data in 2024

Area number	Consolidated area	Peak [MW]	Energy [GWh/year]	Loss factor
1	AESO	16095	115061	1.019
2	BCH	12542	69326	1.020
3	PNW = AVA + BPA + CHPD + DOPD + GCPD + PACW + PGE + PSEI + SCL + TPWR	33384	184103	1.025
4	NWMT	1898	12163	1.023
5	PAWY	1681	11028	1.013
6	NCA = BANC + CIPB + CIPV + TIDC	30626	144848	1.043
7	SPPC	2447	15784	1.026
8	ID = IPFE + IPMV + IPTV	4157	19290	1.036
9	UT = PAID + PAUT	8443	39362	1.028
10	CO = WACM + WAUW	5867	34863	1.023
11	LDWP	7789	34129	1.027
12	NEVP	7034	30083	1.045
13	PSCO	8130	41027	1.028
14	SCA = CISC + VEA	26847	119573	1.040
15	AZ = AZPS + SRP + TEPC + WALC	23596	109534	1.026
16	CISD	5573	26702	1.037
17	IID	1342	4836	1.043
18	PNM	3136	16449	1.026
19	CEF	3255	15452	1.033
20	EPE	2391	11106	1.032

dispatchable generation capacities are given in Table 4.5 also from the WECC 2024 common case.

4.4.2 Market model

We assign a hypothetical market to every consolidated area. The supply curve consists of a flat segment for must-take generators (wind, solar, run-of-river hydro, biomass, geothermal and nuclear units), and a variant segment for dispatchable generators (coal-fired, hydro and gas-fired units). The hourly power generation of intermittent generation resources (wind, solar and run-of-river) are extracted from the WECC 2024 common case. Once constructed, there is effectively no marginal cost of producing renewable energy, thus it will produce at a price as low 0 \$/MWh. We also assign a price of 0 \$/MWh to 15% of hydro capacity, 50% of coal capacity and all of nuclear

Table 4.5: Supply data (aggregated installed capacity) in 2024

Area	Non-dispatchable (must-take) generation Intermittent [MW]	Base (invariable) [MW]	Dispatchable generation [MW]
AESO	2275.2	3710	17733
BCH	815.3	3531	17000
PNW	14432.0	5845	36675
NWMT	664.2	1227	1993
PAWY	1315.3	1648	2105
NCA	5092.7	5351	38684
SPPC	800.0	1118	3287
ID	660.3	374	2845
UT	256.5	2183	8136
CO	656.4	1807	4310
LDWP	687.0	89	8727
NEVP	75.7	426	13085
PSCO	2441.1	1616	11645
SCA	6028.1	977	28645
AZ	3220.6	8601	39353
CISD	432.8	34	8558
IID	34.4	1170	1514
PNM	840.3	910	4504
CFE	384.2	697	6600
EPE	1.3	0	3512

capacity so that they will be treated as a must-take resource because once started these generators are costly to shut down and thus exhibit a negative marginal cost of operation below these levels.

For simplicity we assume that the variant segment is linear, and the bidding prices of the cheapest and most expensive dispatchable units are the same as the must-take (zero) and must-serve (here 500 \$/MWh) prices, respectively. Note that must-take units operate continuously (except during curtailment hours) so their start-up and shut-down costs are not considered. In addition, for simplicity, the start-up and shut-down costs of dispatchable units are not currently considered in the model.

Similarly, the demand curve consists of an unresponsive segment for must-serve loads and a responsive segment for deferrable loads such as water heaters, electric vehicles, heating, ventilation and air conditioning loads. The responsive load can respond to the market price such that it is zero when price is at the market cap, and is all-in when price is zero. A simple demand model is developed to obtain the hourly responsive load function. Given that the wide variety of demand response functions available in literature can all be linearized in the neighborhood of the operating point, a linear model of demand is chosen for the multi-area resource allocation focus of this study.

The hourly demand forecast q_h in every balancing authority is available in the WECC 2024 common case. A portion of q_h is flexible and potentially responsive to the market price, and the remainder is unresponsive. For simplicity, we assume the relative magnitude of the unresponsive portion, $\alpha = \frac{q_u}{q_h}$, remains constant over the entire year. To approximate the responsive demand, we make the additional assumption that responsive load is linearly sensitive to the price fluctuation in a 4-hour future time window. If the average of inelastic-demand clearing prices over a 4-hour future time window (from a 4-hour ahead forecast) is less than the inelastic-demand clearing price at a given hour, then the cleared elastic demand will be lower than the inelastic demand. The motivation behind this behavior is driven by the opportunity to postpone a proportion of responsive load in order to take advantage of lower prices within the next 4 hours.

Consider $\bar{p}$ is the price at which the cleared quantity of the elastic and inelastic demands are equal. We assume $\bar{p}$ is equal to the average of estimated inelastic-demand clearing prices in a 4-hour future time window. If the hourly price is equal to the average price, there is no economic incentive to adjust flexible loads. Using a linear demand curve, the maximum responsive load is:

$$q_r = \frac{(1-\alpha)\ p_{cap}}{p_{cap} - \bar{p}}\ q_h. \tag{4.11}$$

Figure 4.7 shows an hour at which must-take generation, inelastic demand and the associated price are 1000 MW, 3250 MW and 321 \$/MWh, respectively. Consider the case for which the forecast shows the must-take generation will increase to 1500 MW after one hour, 2000 MW after two hours and 2500 MW after three hours, therefore the associated prices will be 250 \$/MWh, 179 \$/MWh and 107 \$/MWh, respectively if the demand remains constant. The average price of these four hours is 214 \$/MWh. Assuming 80% of the demand is inelastic, then according to the model if the hour inelastic-demand price were 214 \$/MWh, the elastic demand quantity would be the same as the inelastic demand quantity (3250 MW). Because the hourly price (321 \$/MWh) is greater than the average price, the responsive load adjusts such than the clearing price and cleared quantity are 295 ($<$ 321) \$/MWh and 3066 ($<$ 3250) MW, respectively. This demand function gives an approximately constant energy consumption model over a 4-hour window.

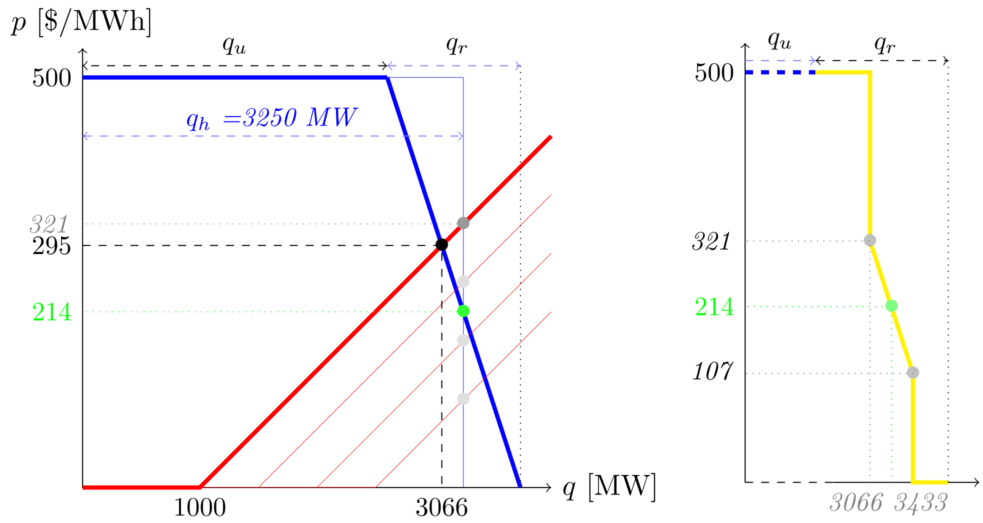


Fig. 4.7: Responsive load shape example.

Although it may seem that a relatively large fraction of the demand is responsive, the active responsive demand is typically a smaller fraction. The dispatched quantity of the responsive demand is implicitly restricted between the quantities associated with the cheapest and the most expensive prices estimated in a 4 hour future time window, not between zero and q_r. This restriction avoids the load control saturation and oscillation observed in some studies of short-term dispatchable demand response [53, 54, 108].

4.5 Results and discussion

We evaluate the proposed scheduling method on the reduced WECC baseline model for each 8784 hours of the year 2024. First, a simulation is performed assuming the demand is completely inelastic. Second, another simulation is carried out for a case in which 5% of demand is price sensitive. To set up the responsive demand functions, the 4-hour rolling average prices (in each area) are gathered from the first simulation and used in the second one. All presented figures are from the second simulation with the transfer constraint limits set to 75% of the rated transfer capacity. Finally, several simulations are performed to explore the impact of various levels of demand response, and of relaxation of the transfer constraint limit on global cost.

4.5.1 Inelastic demand

The simulation results, excluding demand response, show that the unconstrained and the constrained transfer solutions decrease the global electricity production cost (and correspondingly increase the global surplus) by \$12.30B and \$ 10.67B per year, respectively, when compared with the standalone condition case. This cost reduction is equal to 25% and 22% of the standalone production cost respectively. The \$12.30B reduction is the difference between the two extreme coppersheet and standalone conditions. In WECC and many other jurisdictions, balancing authorities exchange electricity according to long-term and short-term bilateral contracts. Therefore, without the proposed unified operation the production cost is lower than the standalone condition case. In addition, in the model, we assume hydro (up to 15% of total nominal capacity) and baseload thermal (up to 50% of total nominal capacity) power plants are treated as must-take resources (base generation) to satisfy operational constraints. Thus, the reported percentage of cost saving, which is the cost reduction divided to the total cost is over-estimated because its denominator excludes a portion of the operation cost.

Table 4.6 compares the producer cost reduction and surplus reduction in each consolidated area under unconstrained and constrained solutions. Negative cost reduction indicates an increase in production cost in that area. We note that although the magnitude of the decreased producer surplus is greater than the magnitude of the decreased production cost in many consolidated areas, the magnitude of the total surplus reduction is smaller than the magnitude of the total cost reduction.

4.5.2 Elastic demand

Coppersheet condition

Figure 4.8 illustrates the aggregated hourly demand with 5% responsive fraction, intermittent generation and base generation (both must-take) in the summer peak month of July and the winter peak month of December 2024. In addition, the net demand, which is the difference between the aggregated demand and the must-take generation is illustrated. The color of the net demand represents the associated coppersheet price. Notice that dispatchable generation is also equal to the net demand. As expected, smaller net demand is associated with lower prices.

Figure 4.9 shows the maximum global cost reduction (relative to standalone global

Table 4.6: Producer cost and surplus reduction for 100% inelastic demand for 2024 (in M$/year)

	Producer cost reduction		**Producer surplus reduction**	
	Unconstrained	**Constrained**	**Unconstrained**	**Constrained**
AESO	6642	1797	11421	2820
BCH	55	149	79	284
PNW	35	479	-552	588
NWMT	-260	-236	-1622	-1516
PAWY	-297	-246	-2513	-2132
NCA	-393	121	-817	145
SPPC	-297	-254	-1058	-933
ID	1873	1907	2630	2696
UT	453	557	763	997
CO	1244	1296	2978	3147
LDWP	2304	2422	2498	2632
NEVP	-183	-4	-214	-8
PSCO	-454	-306	-1044	-753
SCA	5697	6083	7215	7771
AZ	-4624	-4082	-12111	-10968
CISD	992	1108	1061	1187
IID	-213	-192	-1467	-1375
PNM	-324	-262	-873	-735
CFE	-521	-287	-873	-506
EPE	573	622	573	622
Total	12301	10671	6075	3965

cost) versus the standard deviation of standalone prices (for each of 8784 hours in the year 2024). Negative prices are considered zero when computing the standard deviation. The color represents the associated coppersheet price. A larger standard deviation value indicates a greater resource diversity, which usually creates a greater opportunity to reduce the production cost by exchanging power within the interconnection.

Constrained transfer condition

Figure 4.10 illustrates fluctuations of the hourly global cost reduction under both the unconstrained and the constrained transfer conditions in July and December 2024. Clearly, the cost reduction is strongly correlated with demand. The color represents the *economic utilization factor*, which is a proposed performance index for the transmission system defines as the ratio of the cost reduction associated with the

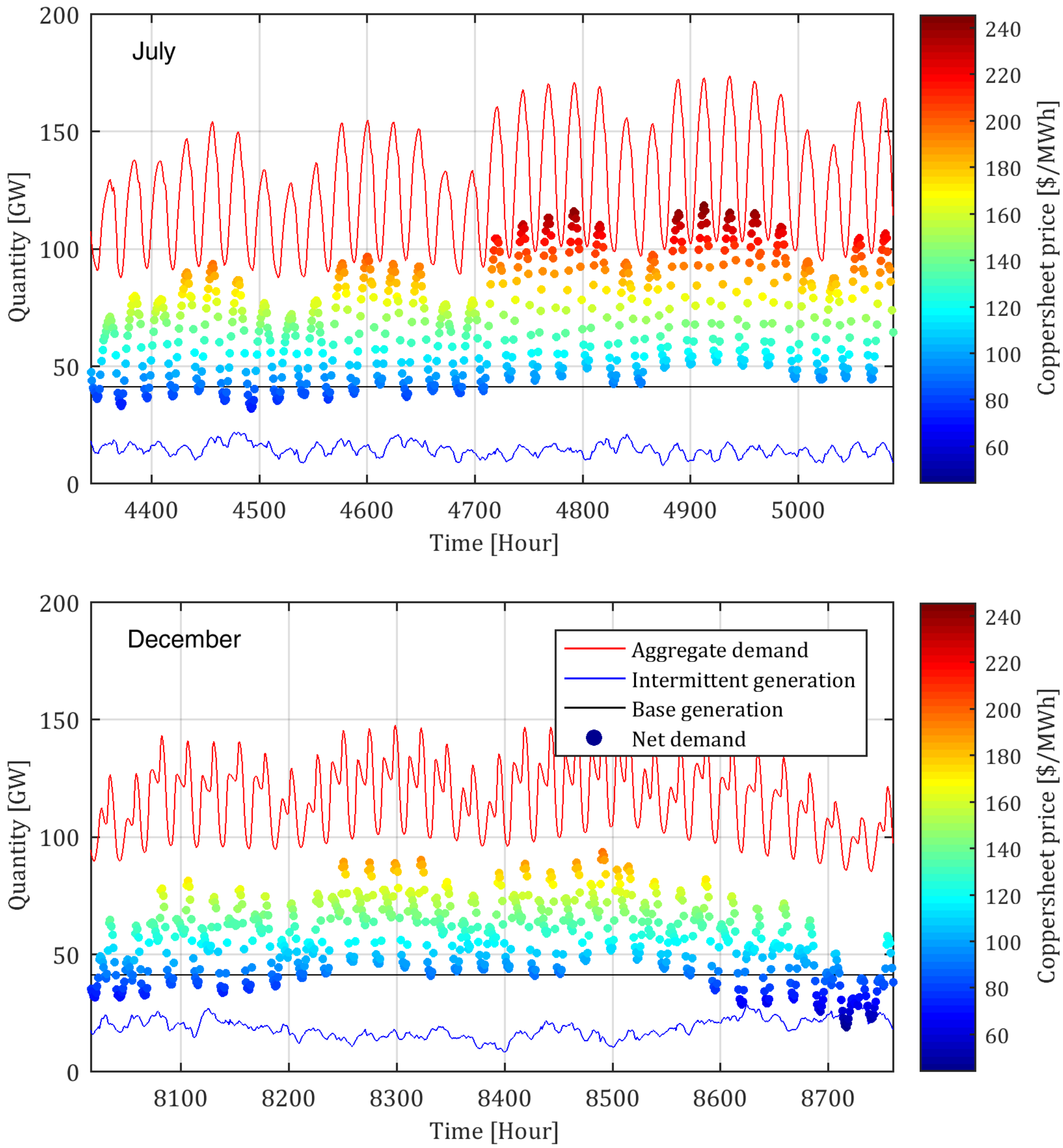


Fig. 4.8: Coppersheet demand, must-take and associated prices.

constrained transfer solution to the cost reduction associated with the unconstrained transfer solution on an hourly basis. We note that there is a greater opportunity to reduce the production cost in December than in July, despite the lower aggregate demand in December. The reason is the magnitude of net demand is more diverse over the system in winter than in summer. In other words, the relatively lower demand in winter enables some areas to feed cheap electricity to the rest of the system,

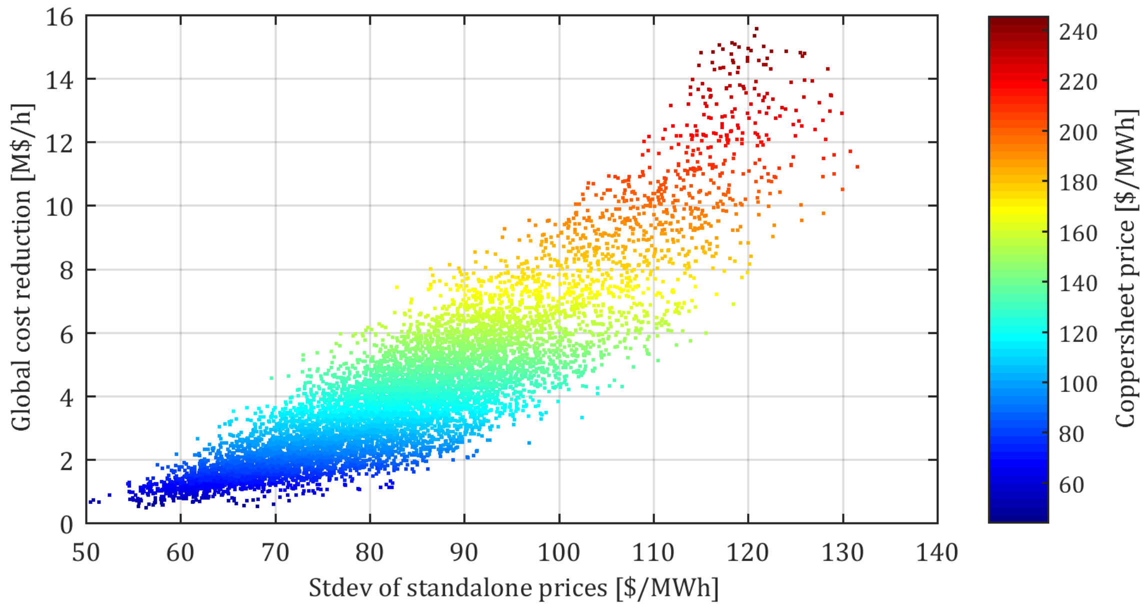


Fig. 4.9: Global cost reduction vs. standard deviation of standalone prices for each hour.

so other areas do not schedule their more expensive units. However, because of transfer constraints the coppersheet export/import schedule is not usually achievable in December, thus the economic utilization factor in December is considerably lower than in July.

Figure 4.11 shows a set of the "economic utilization factor" and the "standard deviation of clearing price over areas" pairs. The utopia point is at the top left corner, where the clearing price is uniform over the entire interconnection. As the standard deviation approaches zero, the economic utilization factor increases. The color represents the mean value of clearing prices.

Figure 4.12 shows the optimal transfer set at the system peak hour. The node color represents the price associated with the standalone condition case, which varies between 0 $/MWh and 500 $/MWh with a standard deviation of 124 $/MWh. Because of the electricity exchange, the clearing price decreases in importing areas and increases in exporting areas. In the peak hour, the clearing price in AESO decreases to 251 $/MWh, in BCH and CFE increases to 251 $/MWh and 236 $/MWh respectively, and in the rest of the system becomes 241 $/MWh. Consequently, suppliers will gather less surplus in importing and more surplus in exporting areas in comparison to the standalone condition. The magnitude of the decreased surplus in importing areas is greater than the magnitude of increased surplus in exporting areas thus sharing

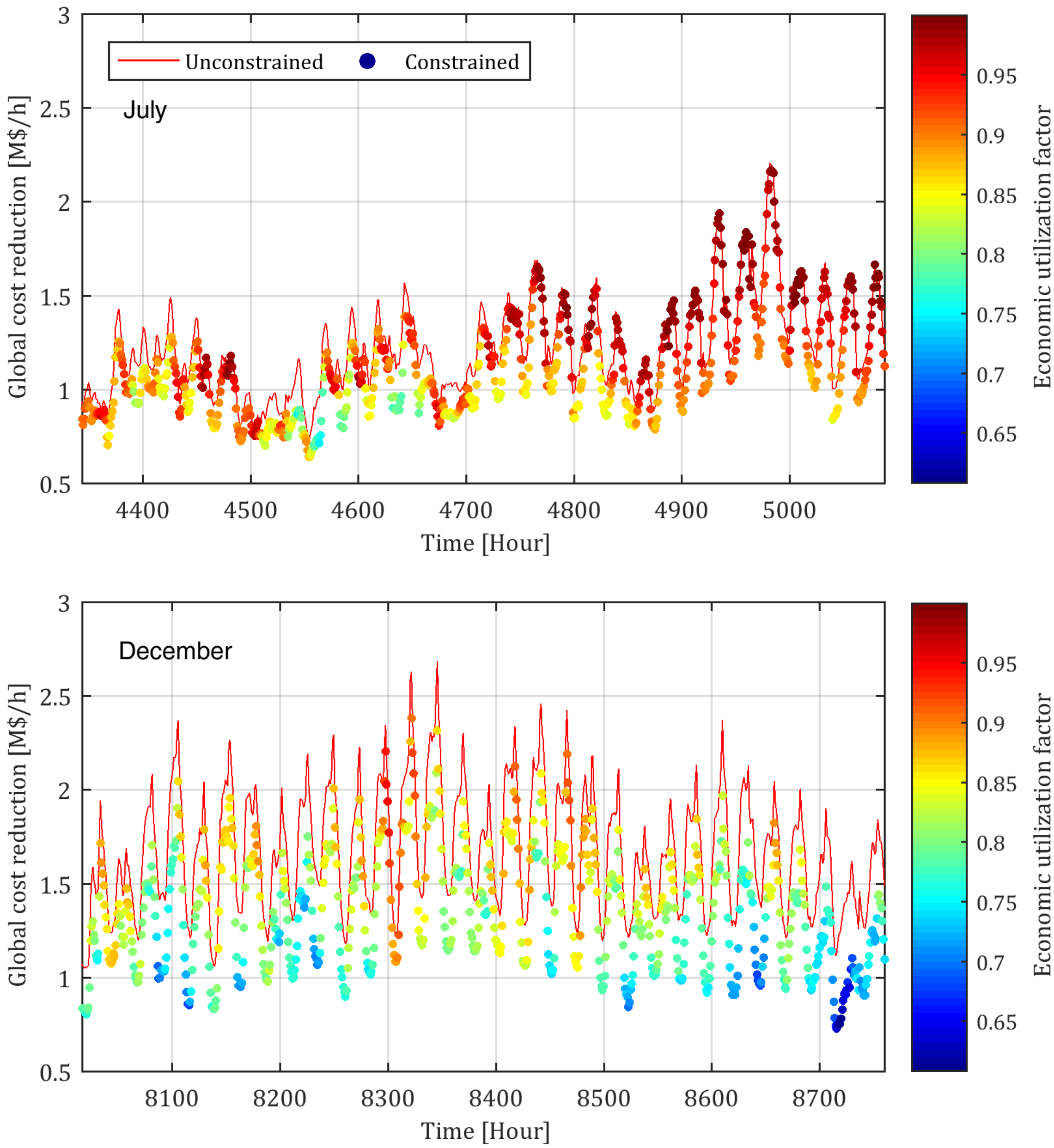


Fig. 4.10: Global hourly cost reduction.

generators decreases the overall producer surplus.

Table 4.7 compares the production cost per unit associated with the standalone and the constrained transfer solutions in every consolidated area for the 100% and the 95% inelastic demand conditions. In general, the production cost per unit decreases in importer-areas, and increases in exporter-areas. Table 4.7 indicates that demand response does not reduce the production cost per unit notably. Without

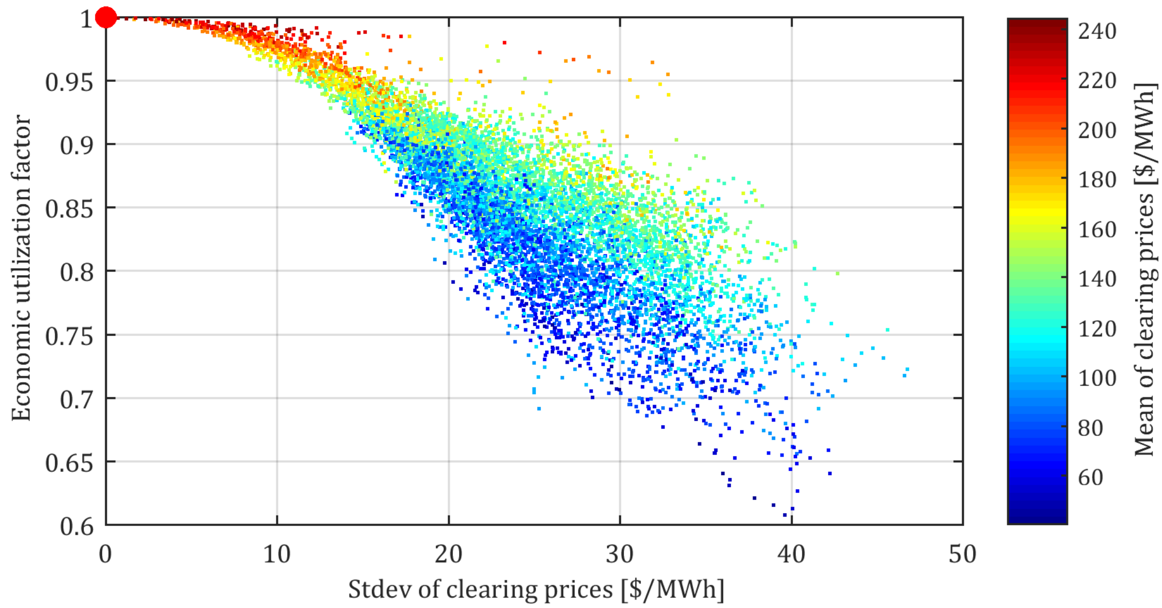


Fig. 4.11: Economic utilization factor vs. price standard deviation.

demand response, the average production cost under the standalone and the constrained transfer solutions are 41.96 $/MWh and 33.40 $/MWh, respectively. With 5% demand response, the average production cost under the standalone and the constrained transfer solutions are 41.95 $/MWh and 33.39 $/MWh, respectively. Notice that the electricity production cost is very low in consolidated areas with a significant amount of must-take generators in some hours.

4.5.3 British Columbia

We now illustrate the impact of optimal scheduling on British Columbia's export/import schedule. British Columbia's hydro-dominated grid connects Alberta's fossil-dominated grid to the rest of the WECC system. Hydroelectric storage assets in British Columbia can also provide the reserves needed to help WECC integrate a significant amount of intermittent renewables in the future.

Figure 4.13 illustrates the load duration curve of the Pacific Northwest–British Columbia tieline. The color represents the price difference. The positive direction is from Pacific Northwest to British Columbia, with positive price difference as the British Columbia price minus the Pacific Northwest price. The price difference is zero when the transfer flow is not at its limits, because the neighboring area prices will naturally equilibrate due to the power transfer. As shown, British Columbia exports a

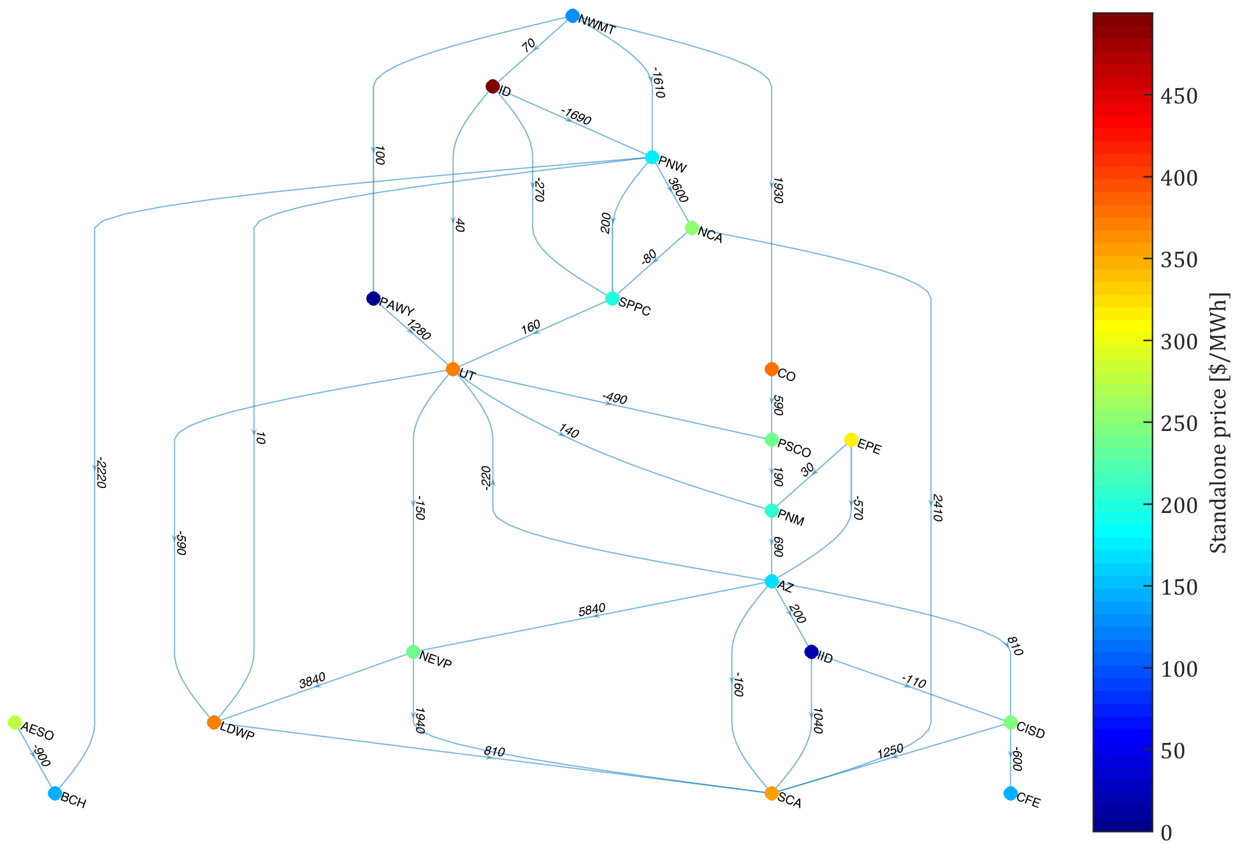


Fig. 4.12: Optimal transfer flow solution at the system peak hour.

significant amount of the imported electricity (820-900 MW) to Alberta. The positive direction of the Alberta–British Columbia tieline is from Alberta to British Columbia, with positive price difference as the British Columbia price minus the Alberta price. The optimal solution suggests that the Alberta–British Columbia tieline is usually operated under the full transfer capacity (900 MW), and as a result there is a price gradient on this tieline. Therefore British Columbia has an arbitrage opportunity buying relatively cheaper electricity from the US to sell to Alberta.

Figure 4.14 shows the surplus increase as a function of the net export in British Columbia. The shape of this curve should be a parabola according to Equation 4.7c, which is verified in the figure. The color represents the clearing price.

4.5.4 Sensitivity analysis

The model presented belongs to a class of agent-based models for which validation can be challenging [109]. Validation methods include the usual statistical, behavior,

Table 4.7: Production cost per unit in $/MWh

	Standalone		Constrained	
	100% inelastic	**95% inelastic**	**100% inelastic**	**95% inelastic**
AESO	78.48	78.47	67.51	67.50
BCH	32.72	32.71	31.31	31.30
PNW	26.33	26.31	24.64	24.63
NWMT	1.45	1.44	14.78	14.77
PAWY	0.00	0.00	11.11	11.10
NCA	32.14	32.13	31.68	31.67
SPPC	9.76	9.75	21.29	21.28
ID	109.06	109.07	32.83	32.79
UT	35.66	35.62	26.92	26.91
CO	50.93	50.92	19.19	19.18
LDWP	97.06	97.04	51.01	51.01
NEVP	49.45	49.44	49.91	49.90
PSCO	26.24	26.23	31.62	31.61
SCA	76.38	76.37	41.27	41.26
AZ	6.11	6.11	28.10	28.09
CISD	79.17	79.17	52.90	52.90
IID	0.01	0.01	13.82	13.81
PNM	16.58	16.57	27.71	27.71
CFE	21.43	21.43	33.10	33.11
EPE	90.20	90.19	57.33	57.33
Average	41.96	41.95	33.40	33.39

and structural validation, as well as so-called *face* validation in which experts observe the response of the system to various scenarios to consider whether the model exhibits the expected local and global properties and sensitivities. In spite of these challenges, agent-based models are particularly well-suited for the study of transient dynamics in highly complex hybrid econo-physical systems such as smart-grids [90]. If the model behaves in the same way as the physics of the problem dictates, the model can be a useful tool that provides insight to the system behavior [110].

For the proposed model we expect that (i) upgrading the transmission system and (ii) integrating demand response result in a greater cost reduction. Table 4.8 confirms this expectation: as the intertie constraints are relaxed the global cost decreases. The same trend is also observed when the fraction of responsive demand increases significantly. However, the economic benefit of introducting inter-hourly demand response is not substantial when the system is operated with the optimal inter-area transfer set. The rational behind this is that the optimal transfer set has already adjusted to price fluctuations (as much as the transmission system allows); hence the

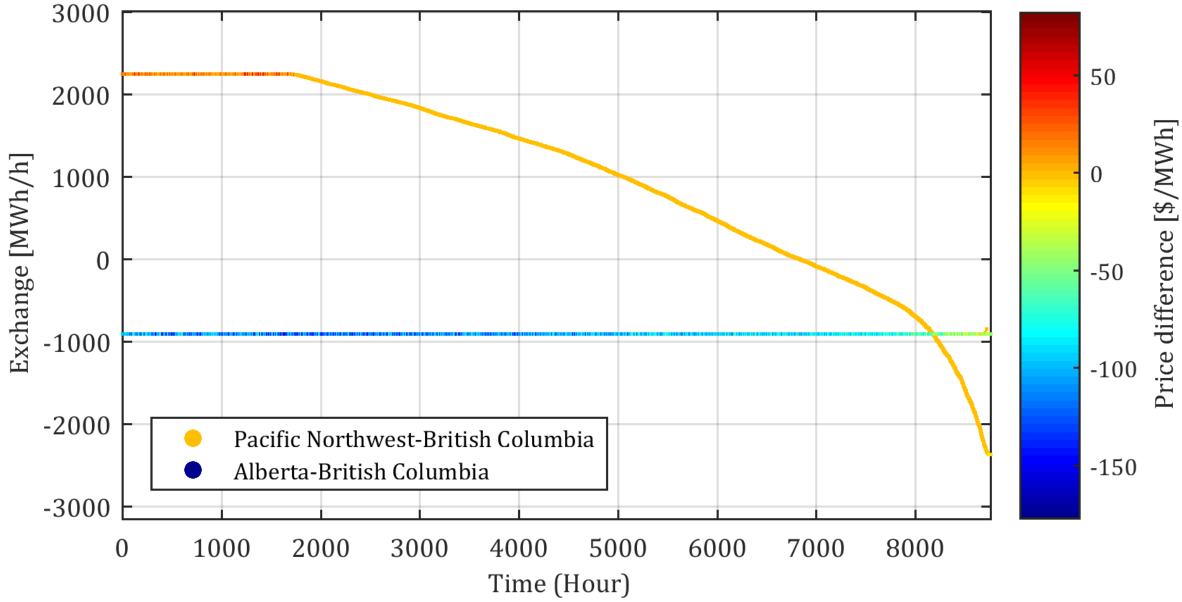


Fig. 4.13: Load duration curve of the Pacific Northwest–British Columbia and the Alberta–British Columbia tielines.

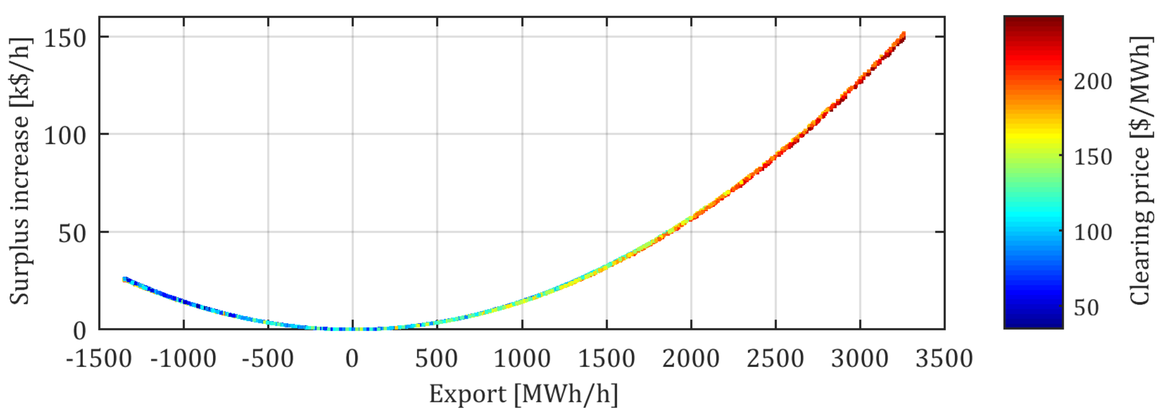


Fig. 4.14: Surplus increase in British Columbia.

remaining opportunity for demand response to react is very limited. If we used a more flexible demand curve such as that presented in previous studies by the authors [89, 111], the impact of demand response would be potentially greater, but probably still not comparable with the impact of the optimal interchange transfer.

Relaxation of the transfer limit from 75% to 90% (even to 105%) does not provide a significant additional cost reduction, in comparison to the achievable benefit from operating the system under optimal inter-area transfer set (relative to standalone case). Some tielines probably need major upgrades, but most have sufficient transfer capac-

Table 4.8: Global cost reduction

Demand response	75% transfer limit [B$]	90% transfer limit [B$]	105% transfer limit [B$]	Copper-sheet [B$]
0%	10.67	10.85	11.01	12.30
1%	10.68	10.86	11.02	12.31
5%	10.70	10.88	11.04	12.33
10%	10.72	10.91	11.06	12.35
20%	10.77	10.95	11.10	12.39
50%	10.85	11.04	11.19	12.48

ity at least for the WECC 2024 scenario studied. This finding highlights the great cost reduction opportunity that exists in operating the interconnection in an unified manner with the existing transmission capacity. This opportunity is more appreciable with greater penetration levels of intermittent renewables in the future. Operating the system in an economically efficient manner enables the interconnection to meet its greenhouse gas emissions reduction targets with a relatively moderate incremental cost. Although technical and non-technical barriers make a unified scheduling and dispatch control challenging, distributed dispatch control strategies exist that can provide sub-optimal inter-area transfer sets.

Optimizing the interchange transfer effectively depresses the clearing price by sharing low-cost generation units over the interconnection. Consequently, the expensive peak units are dispatched less frequently and the annual producer surplus decreases. Table 4.9 indicates the global producer surplus reduction for different limits of transfer constraint and levels of demand response. For example, under the 75% limit of transfer constraint and 5% demand response, this reduction is 4% of the global producer surplus. In general we find that optimal scheduling decreases producer surplus which may have an adverse economic impact both on operators that rely on congestion revenues and zero marginal-cost generation units that must recover capital costs.

4.6 Conclusions

In this paper we have presented a method to model the limits and optimal generation and demand response allocation for the hourly schedule problem. The purpose is not to solve the interconnection resource allocation problem itself. Rather we seek to provide a method that facilitates impact studies of renewable generation and de-

Table 4.9: Global producer surplus reduction

Demand response	**75% transfer limit [B$]**	**90% transfer limit [B$]**	**105% transfer limit [B$]**	**Copper-sheet [B$]**
0%	3.96	4.07	4.26	6.07
1%	3.97	4.07	4.27	6.08
5%	4.00	4.10	4.30	6.11
10%	4.03	4.13	4.33	6.14
20%	4.08	4.18	4.38	6.19
50%	4.19	4.29	4.49	6.30

mand response interaction at the interconnection scale without fully implementing the various market designs extant.

Accordingly a new resource allocation modeling method that incorporates both energy supply and demand response resource is presented. This method is used to model the interconnection-scale scheduling problem in the Western Electricity Coordinating Council (WECC) system planning year 2024. We quantify the potential benefits of demand response integration in WECC using the proposed surplus-maximizing scheduling model considering tieline transfer limits on an hourly basis. The method determines the optimal inter-area transfer set that effectively makes the clearing price uniform over the interconnection's markets (as much as possible). The overall consumer surplus increases and the overall producer surplus decreases, while the increased magnitude is always greater than the decreased magnitude. At the same time, the magnitude of production cost decrease is greater than the magnitude of produce surplus decrease. Results show that inter-regional collaboration assist control areas in reducing production cost by $10.67B, equal to 22% relative to the standalone condition with inelastic demand. In addition, integration of 5% demand response provides an additional $27M cost saving. However, we also conclude that economic benefit of inter-hourly demand response is not substantial when the system is operated with the optimal inter-area transfer set, which suggests that the majority of the benefits for demand response arise from intra-hourly demand response resources which are not considered in conventional hourly resource scheduling mechanisms.

The hourly supply and demand resource commitments found in this paper can be used to establish a 5-minute redispatch market in each balancing authority to provide reliable and efficient short-term matching of generation and load resources. Suitable control systems are needed to enable demand response participation in redispatch markets and ultimately in the ancillary services markets for frequency [25] and voltage regulation [112].

Future work will apply the proposed scheduling method for energy planning purposes for renewable energy policy studies in the Western Interconnection beyond the year 2024. The objective will be to determine the most economically beneficial type and capacity of renewable energies in each region such that the interconnection can achieve a target level of greenhouse gas emissions reduction.

Appendix 1: Objective function

Under the standalone condition, supply and demand quantities are equal, so the clearing price in control area n is:

$$p_{0_n} = \frac{s_n\ p_{max} - d_n\ p_{min}}{s_n - d_n} + \frac{s_n\ d_n}{s_n - d_n}(q_{w_n} - q_{u_n}).$$

For a non-zero net export $e_n = q_{s_n} - q_{d_n}$, the demand price is equal to the supply price:

$$\underbrace{d_n\ (q_{d_n} - q_{u_n}) + p_{max}}_{p_{d_n}} = \underbrace{s_n\ (q_{s_n} - q_{w_n}) + p_{min}}_{p_{s_n}} = s_n\ (\underbrace{q_{d_n} + e_n}_{q_{s_n}} - q_{w_n}) + p_{min},$$

therefore the demand quantity is:

$$q_{d_n} = \frac{p_{max} - p_{min}}{s_n - d_n} + \frac{s_n\ q_w - d_n\ q_u}{s_n - d_n} - \frac{s_n}{s_n - d_n} e_n,$$

and the clearing price is:

$$p_{c_n} = \underbrace{\frac{s_n\ p_{max} - d_n\ p_{min}}{s_n - d_n} + \frac{s_n\ d_n}{s_n - d_n}(q_{w_n} - q_{u_n})}_{p_{0_n}} - \frac{s_n\ d_n}{s_n - d_n} e_n = p_{0_n} - \omega_n\ e_n,$$

so the price vector is:

$$\mathbf{p}\ = \mathbf{p}_0\ - \mathbf{\Omega}\ \mathbf{e}.$$

Accordingly, Equation 4.7a can be rewritten as:

$$\max \frac{-1}{2} \sum_{n=1}^{N} \omega_n\ e_n^2 \equiv \min \frac{1}{2} \sum_{n=1}^{N} \omega_n\ e_n^2.$$

Appendix 2: WECC intertie transfer capacity

Path number	**From area**	**To area**	**Minimum [MW]**	**Maximum [MW]**
P01	AESO	BCH	-1200	1000
P03	PNW	BCH	-3150	3000
P08	NWMT	PNW	-2150	3000
P09	NWMT	CO	-2573	2573
P14	ID	PNW	-2250	3400
P16	ID	SPPC	-360	500
P18	NWMT	ID	-256	337
P20	ID	UT	-2250	2250
P22	PNM	AZ	-2325	2325
P24	NCA	SPPC	-150	160
P26	NCA	SCA	-3000	4000
P27	UT	LDWP	-1400	2400
P30	UT	PSCO	-650	650
P31	PSCO	PNM	-690	690
P32	SPPC	UT	-235	440
P35	UT	NEVP	-580	600
P36	CO	PSCO	-1680	1680
P42	IID	SCA	-1500	1500
P43	LDWP	SCA	-4000	4000
P44	CISD	SCA	-2500	2500
P45	CISD	CFE	-800	400
P46N	NEVP	LDWP	-6000	6000
P46S	NEVP	SCA	-5000	5000
P47	EPE	AZ	-1048	1048
P48	EPE	PNM	-1970	1970
P49	AZ	NEVP	-10200	10200
P59	AZ	SCA	-218	218
P65	PNW	LDWP	-3100	3220
P66	PNW	NCA	-3675	4800
P76	PNW	SPPC	-300	300
P78	UT	PNM	-600	600
P79	AZ	UT	-300	265
P80	NWMT	PAWY	-600	600
PP1	PAWY	UT	-1700	1700
PP2	IID	CISD	-150	150
PP3	AZ	CISD	-1160	1650
PP4	AZ	IID	-160	260

Chapter 5

Electric Vehicle Participation in Transactive Power Systems Using Real-time Retail Prices

This paper was presented at IEEE HICSS-49 international conference on January 7th 2016:

Sahand Behboodi, David P Chassin, Curran Crawford and Ned Djilali. *Electric Vehicle Participation in Transactive Power Systems Using Real-time Retail Prices* 2016 49th Hawaii International Conference on System Sciences (HICSS). Available online at: http://ieeexplore.ieee.org/abstract/document/7427483/?section=abstract

Sahand Behboodi has done the major part of developing the methodology, performing the simulation, and writing the paper. David Chassin has helped Sahand to establish an understanding of the transactive control paradigm and the bidding strategy. David has also helped Sahand in coding the control strategies as well as writing the manuscript.

This chapter introduces a load management strategy for plug-in electric vehicles based on the transactive control paradigm. There exists a retail electricity market that provides an efficient use of the charging load flexibility. Depends on the available time to departure and the required time to reach full charge as well as grid real-time conditions, individual electric vehicles submit a buy bid (and also a sell bid under V2G scenario) to the double auction market. This strategy reduces energy costs practically by shifting the demand from super peak evening time to off peak after midnight or early morning in the presence of solar PV generation.

Abstract

Smart grids can help Plug-in Electric Vehicles (PEV) manage their load in a grid-friendly way. In this paper, we consider the case of PEVs participating in a retail double auction electricity regulation market, as in the so-called "transactive control" paradigm.

Price-responsive charging of PEVs is modeled in conjunction with real-time retail price signals from the utility. PEVs can defer charging or even discharge when the retail prices are high. Buy and sell reservation prices are based on expectations of future prices and opportunity costs of sold energy, respectively. Feeder capacity constraints also affect the retail price and are allowed to rise to the point at which supply equals demand. For the most advanced charging strategies, as the price rises, demand from PEVs drops, and if the constraint causes further price increases, the PEVs can begin to supply energy.

The results show that when rooftop solar energy is available transactive bid-response vehicle charging strategies significantly enhance short-term electricity demand elasticity and can reduce consumer energy costs by more than 75% in comparison to the unresponsive charge case.

Keywords

Demand response, plug-in electric vehicle, real-time price, transactive control, smart grid

Nomenclature

P	Price, in ¢/kWh.
Q	Demand, in kWh.
SOC	State of charge.
t	Time.
Greek symbols	
β	Battery capacity, in kWh.
γ	Degradation impact
η	Demand elasticity
κ	Battery capital cost, in \$/kWh.
ρ	Charging rate, in kW.
τ	Real-time
Subscripts	
A	Available
arr	Arrival

B	Buy
C	Cycle
D	Standard deviation
dep	Departure
M	Mean
O	Opportunity
R	Required
S	Sell
Abbreviations	
LMP	Locational marginal price
PEV	Plug-in electric vehicle
PV	Photovoltaic
Stdev	Standard Deviation
V0G	Uncontrolled strategy
V1G	Grid to vehicle
V2G	Vehicle to grid

5.1 Introduction

The adoption of plug-in electric vehicle (PEV) can displace petroleum use and tailpipe emissions, but will also impose an additional load on the power grid. PEV integration can have a disruptive impact on the power grid if not integrated smartly. The difference between the elapsed time required for charging and the time that the vehicle is plugged in allows timing flexibility that can be harnessed to provide grid services while at the same time meeting the needs of the consumer [13] [113]. The idea of smart vehicle charging is addressed in a large body of recent work [20] [114]. Smart charging also helps the grid balance the fluctuation of renewable resources [19, 23, 24, 30, 115, 22, 116], and as a result can contribute to greater GHG emissions mitigation. One important potential benefit of PEV integration is vehicle-to-grid technology [117] [118], which enables electric vehicles to offer regulation up/down service to the grid by making their on-board storage capacity available.

Perhaps the simplest approach to determine the PEV optimal charging is to consider a retail market environment and then distribute the aggregated PEV load on top of the total demand in a grid friendly way. This approach manages the charge schedule to minimize the charging cost. An example of centralized charging control algorithms is Nash certainty equivalence principle [119]. On the consumer side, financial incentives are necessary to encourage PEV owners to participate in the demand response programs [120]. In this paper, we examine a transactive control system where PEVs participate in a 5-minute real-time market that also has participating distributed renewable energy sources. The PEVs submit buy (charge) and sell (discharge) reservation (bid) prices to the market before clearing, and receive a price signal from the market after clearing. PEVs respond to the real-time price according to their original bid by charging, turning off, or discharging, as appropriate.

Without a controlling strategy the integration of a large number of PEV chargers poses challenges to reliable service because it may cause transformer overloads and feeder congestion [121], while of the generation and transmission level it may increase system operation risk [122] [123]. Capacity constraints can be managed under a real-time price computed using a retail double-auction, as shown in the US Department of Energy's 2006-2007 Olympic Peninsula [53], and 2013 American Electrical Power (AEP) gridSMART [54] demonstration projects. In the Olympic Peninsula project, distributed generation (DG) resources were dispatched using the same price signals as demand response resources, and it was shown that these signals gave rise to benefits that accrued to both the utilities and the consumers. In a similar manner, rooftop photovoltaic (PV) generation can be included in the retail market and displace feeder load.

A great deal is known about the demand response behavior of thermostatic loads such as waterheaters, heat-pumps and air-conditioners [26]. Demand elasticity has been extensively studied for these loads under various tariffs and pricing mechanisms, and relatively recent survey of the many load studies conducted over the years indicates that demand elasticity for electricity used by thermostatic loads can range from -0.07 to as high as -0.21, with a mean of about -0.14 [73]. In spite of the importance of this information to electricity tariff design in particular, much less is known about the response of electric vehicle charging to price signals, in part due to the relative novelty of the technology and in part due to its relatively slow adoption. Previous work on electricity pricing for electric vehicle charging suggests elasticity values of -0.10 [124] based on conventional electricity demand. This contrasts with the elasticity for the annual cost of a conventional vehicle of around -0.87 [125], which may be more appropriate given the increased disconnect between driving habits and electric bills. However, based on the experience with real-time price automation for thermostats, one expects that vehicle charger automation and better integration of charger control with real-time price utility dispatch will significantly impact short-term electric vehicle demand elasticity.

The paper provides a detailed description in Section 5.2 of alternative frameworks of the PEV charger bidding strategies for real-time price auctions and associated modeling assumptions. Results are presented in Section 5.3 with an examination of prices, load profiles, vehicle state-of-charge and elasticity, followed by a discussion of these results and the sensitivity of consumer cost to wholesale and retail price volatility in Section 5.4.

5.2 Model Description

Among other things, transactive control systems dispatch distributed generation and load using subhourly (5-minutes in this paper) real-time prices at which available supply capacity is expected to equal local demand. This strategy is extended to include loads such as price-responsive vehicle-chargers that may be operated as either "plug-and-go" loads (so-called V0G), as unidirectional price-responsive loads (so-called V1G), or as bidirectional price-responsive loads (so-called V2G). The same results have been achieved using other methods, such as distribution locational marginal pricing [124]. Specifically, charger demand is controlled using a bidding strategy that increases each charger's reservation prices as the expected vehicle departure time approaches and the probability of being able to achieve a full charge diminishes given its maximum charging capability. Such a bid/response system also allows sufficiently-charged units to discharge if the real-time price exceeds the expected opportunity cost of recharging later.

The efficient capacity allocation strategy arises naturally from the transactive control strategy, which integrates small-scale electric equipment with utility electric power distribution system operations as a first step toward integrating distributed generation and demand response into wholesale operations. These market-based paradigms are designed to find a Pareto-optimal allocation of supply capacity and demand response to resolve how much distributed generators should produce and customers consume in a future time interval. The transactive control systems we consider use distribution capacity markets to determine the price which minimizes the imbalance between supply and demand for electricity by participating equipment during the next operating interval. The system computes a 5-minute retail real-time price (RTP) that reflects the underlying wholesale locational marginal price (LMP), all other distribution costs and any scarcity rent arising from distribution capacity constraints, as well as any constraints on demand arising from consumer comfort preferences given the current state of loads.

Distributed generation, load shifting, demand curtailment and recovery are all induced by variations in real-time prices. In doing so the transactive control system can reduce the exposure of the consumers and the utility to price volatility in the wholesale market and the costs of congestion on the distribution system. The prices are discovered using a feeder capacity double auction that directly manages distribution, transmission or bulk generation level constraints, if any. Distributed generation is economically dispatched and consumer preferences are used to dispatch advanced load controllers that act as rational agents on behalf of the consumers. Devices such as electric vehicle chargers bid for energy when it is needed and respond to price changes when they occur. Using this system consumers can expect to realize greater savings on their bills when they show a preference for more savings, presumably at the expense of some otherwise deferrable or non-critical comfort.

5.2.1 Bidding Strategies

The retail capacity auction is the key component to determining the real-time price. For PEV units operating in *charge* mode, the thermostat bidding strategy employed by the Olympic and Columbus demonstration was used with some minor simplifications. The PEV *discharge* bidding strategy is based on the Olympic generator bidding strategy. However it is improved upon by including an opportunity-cost computed for recharging the battery that is derived from the generic transactive thermostat bidding strategy and accounts for the slower rate at which the desired state-of-charge is reached.

Using these strategies PEVs can participate in retail real-time electricity double auctions by offering buy and sell reservation prices at any particular time. Participating PEVs compare the real-time price with their buy and sell reservation prices. If the price is above a PEVs buy price it will forgo charging for the next 5 minutes. If the price is above a PEV sell price, it will discharge energy back to the feeder during the next 5 minutes. When there is a constraint on the load of a feeder, the clearing price will rise to the point that the total load (including unresponsive load and aggregate PEV load) is just below the supply limit. When there is an excess of PEV energy with an opportunity cost below the feeder supply price, the PEVs will discharge and help reduce utility net feeder load.

The PEV buy price is determined as

$$P_B = P_M + P_D\ K\ \frac{\Delta t_R}{\Delta t_A}$$

where P_M and P_D are the mean and standard deviation of the expected LMP over a time interval between the real-time and the departure; K is the consumer comfort control setting, which enables the consumer to control its charge/discharge behaviour. The forward price time horizon is typically between 2 and 8 hours and which is significantly less than the 24-hour forward price window used in the Olympic and Columbus demonstrations. For vehicle chargers we interpret the meaning of K as follows:

> *A customer with a higher value of K is more likely to depart with a fully charged vehicle than one with a lower value of K.*

A more vernacular definition can be thought of as the charging strategy aggressiveness, i.e., a more aggressive strategy has a higher value of K and will cost the customer more, but is more likely to fully satisfy the customer.

The required time to fully charge is:

$$\Delta t_R = (1 - SOC)\ \frac{\beta}{\rho}$$

where SOC is the battery state of charge, β is the battery capacity in kWh and ρ is the charging rate in kW. The available time until the departure is:

$$\Delta t_A = t_{dep} - \tau$$

where t_{dep} is the departure time, and τ is the real-time. The PEV sell price is expressed as:

$$P_S = \frac{P_O}{\eta^2\ \gamma} + P_C$$

where η is the round trip efficiency, and γ accounts for battery ageing. The cycling cost, P_C, accounts for the additional degradation costs of using V2G, which for $SOC < 80\%$ are estimated based on experimental data [126] as:

$$P_C = \frac{0.001\ \kappa}{(SOC + 0.4)^2}$$

where κ is battery capital cost in \$/kWh; if $SOC > 80\%$ the cycle cost is zero. P_O is opportunity cost for discharging during the next time increment δt:

$$P_O = P_M + P_D\ K\ \frac{\Delta t_R + \delta t}{\Delta t_A - \delta t}$$

The expectation price is estimated from the day ahead market LMPs.

5.2.2 Scenarios

The model considers the case where 100 homes on a capacity-constrained feeder have both rooftop PV panels and PEVs with predominantly night-time charging needs. Three charge scenarios are considered with a feeder constraint 40% of the total nameplate connected load, as shown in Figure 5.1. The V0G scenario assumes vehicle chargers begin charging as soon as the vehicle is plugged in, unless the real-time price exceeds the customer's maximum reservation price, with comfort setting considered. V1G assumes the vehicle chargers only charge when the real-time price is below the expected average price for the remaining time to departure, with comfort setting considered. V2G assumes that charging is like V1G but will also discharge when the real-time price is above the opportunity cost of recharging later given the expected average price for the remaining time to departure, with comfort setting considered.

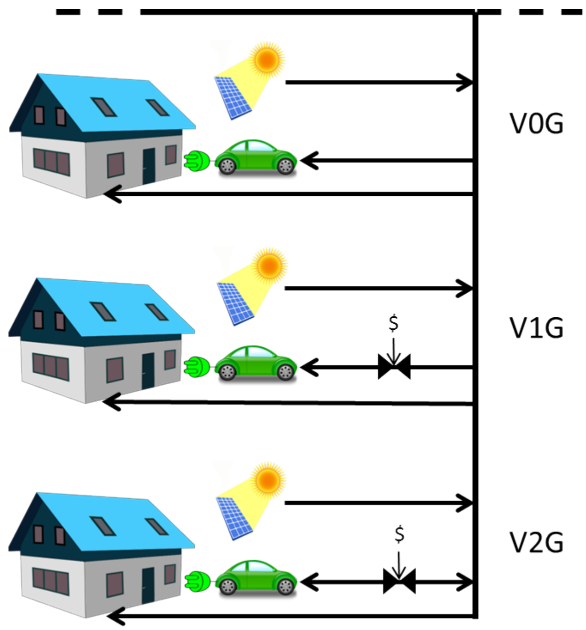


Fig. 5.1: Household load and photovoltaic distributed generation with vehicle/grid integration scenarios: dumb charger (V0G/top), unidirectional price-responsive charger (V1G/middle), bidirectional price-responsive charger/discharger (V2G/bottom)

5.2.3 Assumptions and inputs

Driving Pattern

Driving pattern data were analyzed to get insight into how vehicles are actually driven. Driving Diary data are extracted from the Canadian Plug-in Electric Vehicle Survey 2013 [127]. A normal distribution is fitted on the home arrival and departure times. The mean and standard deviation used are shown in Table 5.1. In addition, a normal distribution is assumed for the battery SOC level at arrival time.

Table 5.1: Driving pattern parameters

Variable	Mean	Stdev
Arrival time [HH:MM]	18:00	2:00
Departure time [HH:MM]	8:00	2:00
Arrival SOC level [%]	70	10

Charge characteristics

The charge/discharge rate depends on the battery SOC level, but for simplicity here the rate is assumed to be uniform during the charge period. We assume the charge/discharge

Table 5.2: Modeling inputs and assumptions and inputs

Assumption	Mean	Stdev
Customer comfort [%]	25	10
Round trip efficiency [%]	92	1
Ageing factor [%]	100	5
LMP [$/MWh]	80	24
Battery size [kWh]	50	10
LMP peak time [HH:MM]	16:00	
Unresponsive load [MW]	0.200	
Feeder capacity [MW]	0.344	
Solar capacity [MW]	0.202	
Number of customers	100	

rate is 6.6 kW for Level 2 and 16.8 kW for Level 3 to compare the impact of charge rate [128].

The battery technology status will be a strong determinant of PEV future success in the marketplace. The auto industry needs cheap, high energy density, fast charging and long life time batteries [129]. Battery ageing is managed by adjusting the bid price as a function of the actual duty cycle relative to expected duty cycle for the battery's age. The strategy is derived from the Olympic demonstration's license usage premium for DG bidding so that the ageing factor is:

$$\gamma = \frac{N}{N-n} \times \frac{M-m}{M}$$

where N is the battery's lifetime, M is the number of expected cycles in the battery's lifetime, n is the current age of the battery (in units of N), and m is the number of cycles used so far.

The general assumptions for this model are shown in Table 6.1.

Solar panel

The effect of solar power is examined using insolation data from the Victoria School-Based Weather Network (www.victoriaweather.ca) on a July day in Victoria, BC (approx. 48°N latitude) with intermittent cloudiness. The residential rooftop panels have power output normally distributed about a mean of 2 kW with 0.1 kW standard deviation truncated at ±3 stdev. We assume 100% penetration of solar PV so that every home with an unresponsive peak load of 2 kW has a PV panel and the bid price for solar PV is zero.

5.2.4 Demand Elasticity

We evaluate the arc elasticity of vehicle operating cost [130]

$$\eta = \frac{P}{Q}\frac{dQ}{dP} = \frac{d\log Q}{d\log P} = \frac{\log(Q_{after}/Q_{before})}{\log(P_{after}/P_{before})}.$$

where the two values for P and Q are found for the unconstrained feeder (before) case and for the constrained feeder (after) case. In their study of mixed-logit models of alternative-fuel vehicles in the UK, Batley et al. [130] found that operating cost elasticities were between -0.15 for a 20% fuel cost reductions and -0.25 for a 5% fuel cost reductions, depending on the model used. However, in many cases feeder congestion can cause price volatility far beyond the range studied. Increased short-term purchasing flexibility can be expected to yield elasticity results that are not very comparable to previous studies, particularly for the V1G and V2G cases.

The analysis of elasticity distinguishes between evening charging, which is dominated by inelastic consumers who prefer the "plug-and-go" option and morning charging which is dominated by consumers who are generally more flexible but have less time to departure. The delineation in this paper uses noon and midnight, with all morning charging during the AM hours and all evening charging during the PM hours.

5.3 Results

The aggregated load of a 100 PEV fleet, with a 5 minutes time resolution, is determined for three charging scenarios with the third scenario evaluated for Level 3 charger instead of Level 2. In all scenarios, LMPs, feeder capacities, unresponsive load, vehicle arrival and departure times and state-of-charge, consumer comfort settings, and PV outputs are identical.

5.3.1 V0G Scenario

We evaluate first the "V0G" scenario in which vehicle chargers consume electricity as soon as the vehicles return home and do so until the batteries are fully charged, unless the retail price exceeds a reservation price set by the consumer based on how likely the vehicle is to be fully charged by the departure time. This scenario corresponds to the minimum demand elasticity case with an overloaded feeder, where all the PEV demand elasticity arises from the consumer's maximum reservation price. The LMP and RTP prices resulting from this method of operation, as well as the total and feeder load profiles for a single day are shown in Figure 5.2 (left). The corresponding PEV state-charge profiles for all 100 vehicles and the corresponding aggregate demand elasticity is also shown in Figure 5.2 (right).

The instantaneous elasticities are shown in Figure 5.2 (bottom) and from these we estimate the average evening and morning demand elasticities of the PEV load. Note that because charging is never postponed, there is no measurable secondary elasticity for the morning hours.

5.3.2 V1G Scenario

The second scenario we examine is the "V1G" scenario in which vehicle chargers submit bids to the retail double-auction for the chargers' capacities at the consumer's reservation price, which is based on how likely the battery is to be fully charged by the departure time. The buy price defers charging based on the expected average price until departure instead of the maximum reservation price. Figure 5.3 shows the impact of V1G on the clearing price and the aggregated load. This scenario corresponds to the moderate elasticity case, where a significant fraction of the total demand elasticity arises from the willingness of highly charged batteries to postpone demand until hours when the price is likely to be lower. The LMP and RTP prices resulting from this method of operation, as well as the total and feeder load profiles for a single day are shown in Figure 5.3 (left). The corresponding PEV state-charge profiles for all 100 vehicles and the corresponding aggregate demand elasticity is also shown in Figure 5.3 (right).

The instantaneous elasticities are shown in Figure 5.3 (bottom) and from these we can estimate the average evening and morning demand elasticities of the PEV load. Note that because charging can be postponed, there is now a secondary elasticity for the morning hours when chargers still needing energy need to top off the batteries, in spite of potentially higher prices.

5.3.3 V2G Scenario

The final scenario we examine is the "V2G" scenario in which vehicle chargers not only submit bids for the consumer's reservation demand price but also offer to sell energy from the batteries if the opportunity cost is sufficiently low relative to the retail price. This scenario corresponds to the high elasticity case, where a significant fraction of the total demand elasticity arises from the willingness of highly charged batteries to not only postpone demand until hours when the price is likely to be lower but also are willing to sell-back energy if the price is high enough. The net effect is to increase demand for "free" solar energy when it is available and store it in the batteries until it can be sold in the middle of the night to batteries that need energy. The LMP and RTP prices resulting from this method of operation, as well as the total and feeder load profiles for a single day are shown

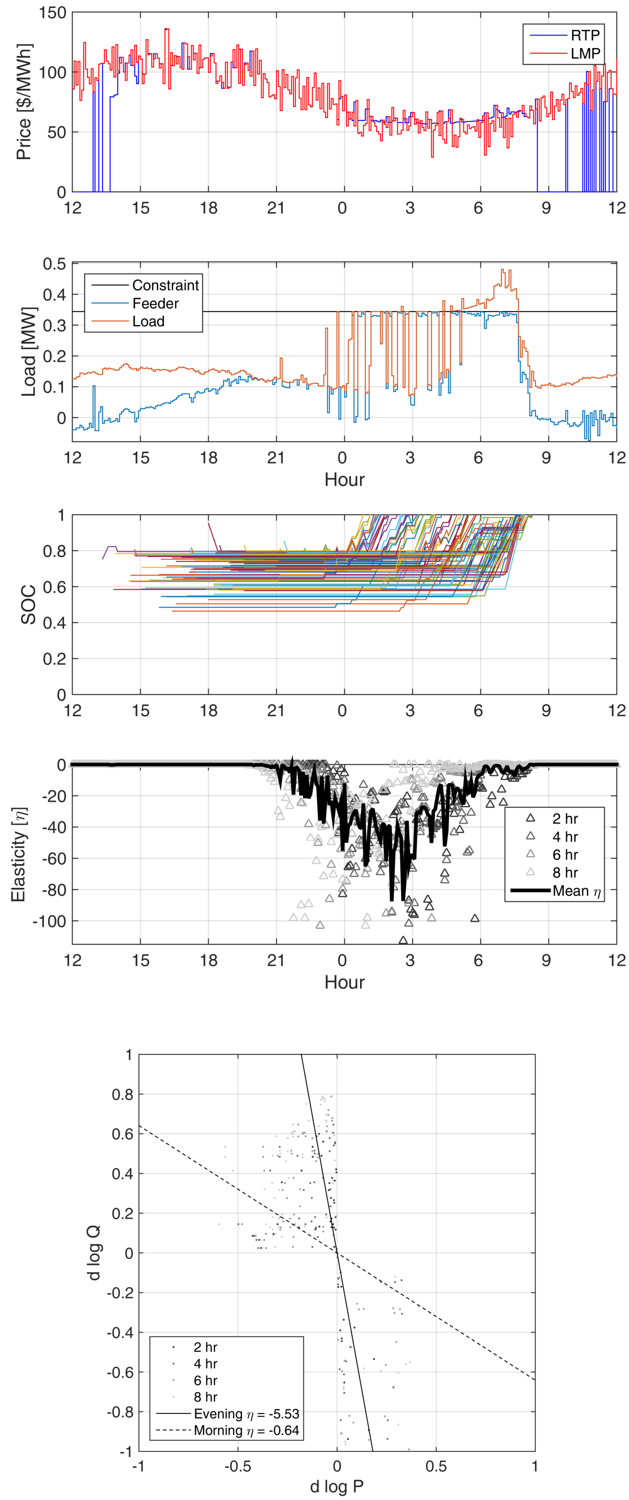


Fig. 5.2: Price, load, state-of-charge and elasticity for a single day of combined "V0G" PEV chargers and rooftop PV

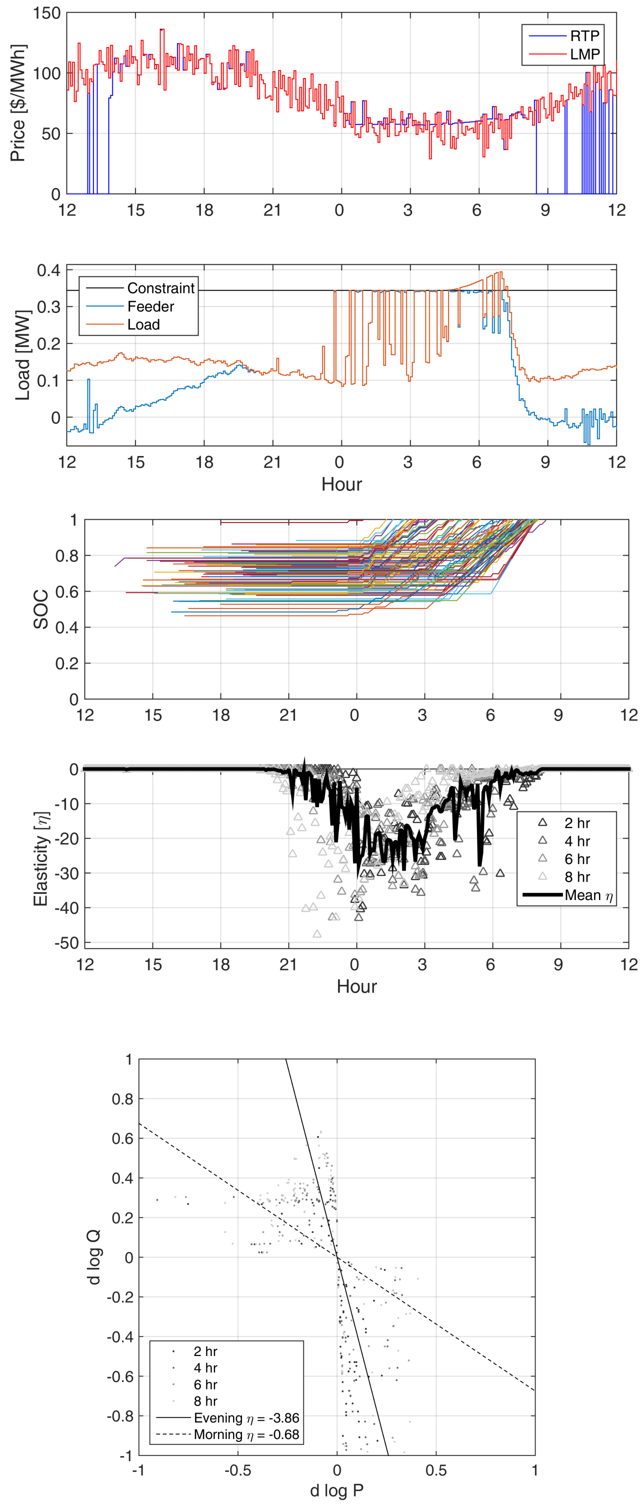


Fig. 5.3: Price, load, state-of-charge and elasticity for a single day of combined “V1G” PEV chargers and rooftop PV

Table 5.3: PEV prices, revenues and elasticity results

Output	**V0G**	**V1G**	**V2G-L2**	**V2G-L3**
RTP mean [$/MWh]	120.29	66.80	66.82	67.36
RTP stdev [$/MWh]	144.57	37.58	37.56	36.88
Peak price time	19:50	16:10	16:10	16:10
Energy [kWh/PEV.day]	16.63	15.96	15.66	15.73
Payment [$/PEV.day]	4.46	0.94	0.91	0.93
P effective [¢/kWh]	26.82	5.90	5.80	5.89
η_D (evening)	-0.41	-3.86	-3.86	-5.53
η_D (morning)	(na)	-0.68	-0.59	-0.64

in Figure 5.4 (left). The corresponding PEV state-charge profiles for all 100 vehicles and the corresponding aggregate demand elasticity is also shown in Figure 5.4 (right).

The instantaneous elasticities are shown in Figure 5.4 (bottom) and from these we can estimate the average evening and morning demand elasticities of the PEV load. The V2G case is examined for both Level 2 and Level 3 charging rates, but the results for prices and loads for V2G-L2 are very similar to V1G and therefore not illustrated.

5.4 Discussion

The results summarized in Table 5.3 suggest that the mean clearing price and the price volability are generally reduced by about 50% when V1G of V2G are used. The peak price time is shifted to later in the evening under V0G but not under V1G or V2G charging. PEV energy consumption by customers is reduced about 4% using V1G and reduced about 5.8% using V2G. Net payment from PEV charging with PV supply is significantly reduced using V1G and V2G, with a corresponding significant reduction in the effective price paid by PEVs. With V0G charging all charging is completed in the evening as soon as the vehicles return home. Thus there is no morning elasticity and evening elasticity is relatively low, albeit greater than Batley's annual cost elasticity of between -0.15 and -0.25. With V1G and V2G charging, a significant amount of charging is deferred and evening demand elasticity increases significantly. However morning elasticity is introduced and it remains relatively low, although somewhat higher than V0G evening elasticity.

Table 5.4 compares the final SOC level under different charging control scenarios. The first column is the percent of full charge achieved at the departure time. The second column is the result of a consumer PEV satisfaction survey [131] showing the acceptance rate of PEV owners to different levels of battery charge. The results suggests that high charger capacity allows the charging strategy to take better advantage of customer charging flexibility by

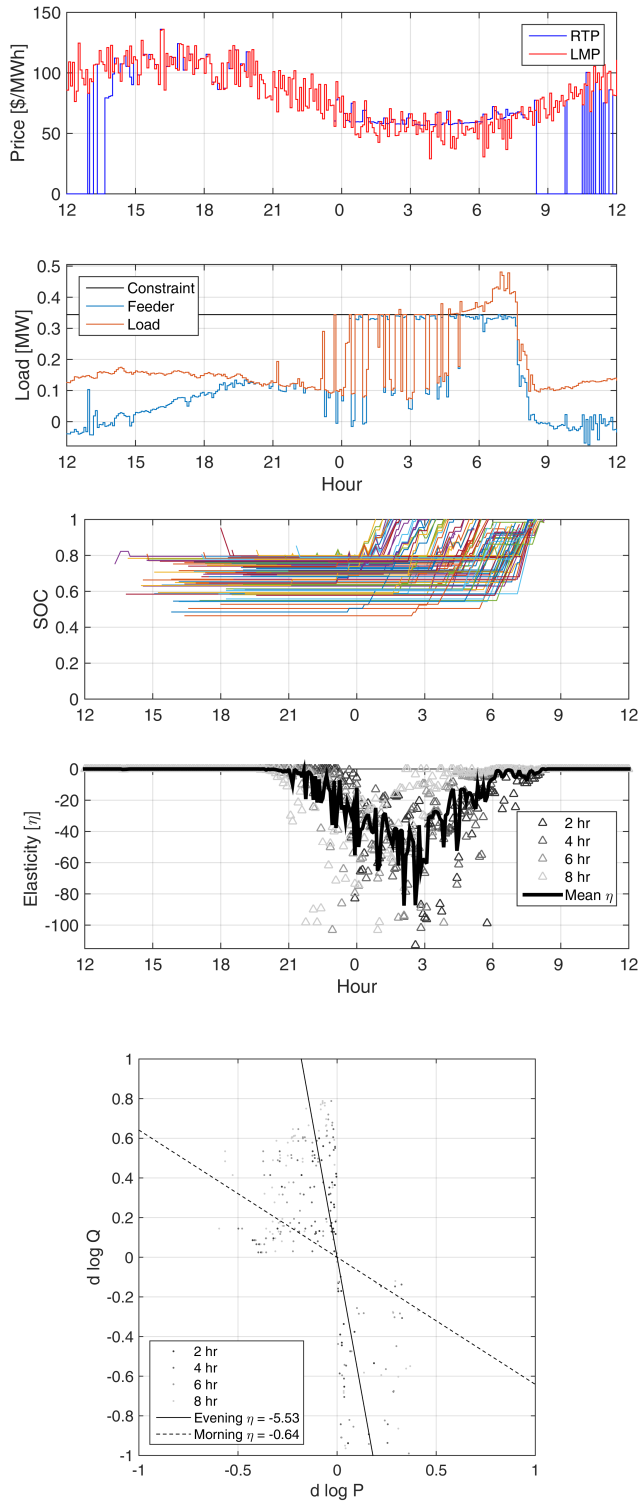


Fig. 5.4: Price, load, state-of-charge and elasticity for a single day of combined "V2G-L3" PEV chargers and rooftop PV

Table 5.4: Final SOC level

Final SOC	Survey	V0G	V1G	V2G-L2	V2G-L3
99–100	35%	100%	64%	59%	57%
80–99	45%	0%	36%	41%	43%
60–80	20%	0%	0%	0%	0%

Table 5.5: Retail price sensitivity to wholesale price volatility

	Effective Retail Price [¢/kWh]			
LMP Volatility	**V0G**	**V1G**	**V2G-L2**	**V2G-L3**
Subhourly	**Daily volatility: 37.5%**			
6.25%	27.98	5.80	5.71	5.76
12.5%	26.82	5.90	5.80	5.89
25%	24.44	5.91	5.82	5.84
50%	19.68	6.01	5.68	5.99
Daily	**Subhourly volability: 12.5%**			
18.75%	26.64	7.11	6.93	6.95
37.5%	26.82	5.90	5.80	5.89
56.25%	26.86	4.55	4.41	4.45
75%	27.22	3.12	2.93	3.05

discharging some batteries "back" into other batteries in the morning hours. This has two important effects. First it allows more solar energy stored overnight to be delivered in the morning. Second, it mitigates price volatility when solar energy would otherwise depress prices or PEV demand would increase it. This results in slightly higher prices when Level 3 charging is in use, but the money is primarily being transferred from less flexible customers directly to more flexible customers, rather than to the utility.

The sensitivity of the effective retail price was examined relative to both the subhourly and daily wholesale price volatility. The results are shown in Table 5.5. Generally consumer costs increase as price volatility increases because they have insufficient elasticity. However the availability of advanced price-sensitive charger strategies increases their elasticity to such a degree that it reverses this trend for daily price volatility. Interhour PEV charger storage capacity has a significant impact the effective prices consumers pay for energy, particularly when photovoltaic energy sources are available.

Table 5.6 tabulates the impact of the customer comfort setting on the departure SOC level and the effective price. Customers with a higher willingess to pay incur higher costs but are also more likely to achieve full charge at the time of departure in the morning.

Table 5.6: Customer comfort setting impact on SOC and retail price under V1G scenario

Customer comfort (K)	**0.25**	**0.50**	**1.00**	**2.00**
Fully charged [%]	64	71	91	100
SOC 80% to 98% [%]	36	29	9	0
Effective price [¢/kWh]	5.90	5.98	6.17	6.25

5.5 Conclusions

In this paper we demonstrated how advanced electric vehicle charging strategies that consider real-time prices can be used to improve inter-temporal coordination between charging needs and charging opportunities. The results are beneficial both to the consumer and the utility: consumers can take advantage of inexpensive renewable sources normally only available at other times of day, and utilities can reduce the risk of overloading distribution assets by allowing customer to exchange energy within the distribution system. When using real-time retail price to reduce the impact of distribution constraints, utilities can also help connect PV to PEV and reduce the customer's net vehicle energy costs by more than three quarters when compared to the simple "plug-and-go" charging, thus significantly bolstering the economic case of electric vehicles over internal combustion engine-powered vehicles.

The results also suggest that there is a qualitative difference between the impact of short-term subhourly price volatility and long-term daily price volatility. Automated load control strategies increase consumer demand elasticity and respond easily to long-term price volatility by deploying energy storage capacity for inter-hour load shifting. These impacts can be evaluated using agent-based simulations of plug-in vehicle charging operated in a real-time pricing environment in conjunction with solar photovoltaic distributed generation.

This results suggest that the modeling approach in this paper can be used to evaluate the combined interaction of PV, PEV and HVAC equipment all operating together as well. This is particularly important given that the PEV can help shift energy from PV to HVAC when the peak times are not coincident, a challenge that will certainly be explored in future work.

Overall consumers would realize significant savings when PEVs participate in a retail electricity markets and receive real-time price signals which they compare to buy/sell reservation prices to make charge/discharge decisions. This subhourly bid-response strategy requires knowledge of price expectations based on day-ahead hourly wholesale prices, which suggests that the integration of wholesale and retail markets must consider not only topological but also inter-temporal exchange to induce sufficient demand response participation at the wholesale level.

Chapter 6

Transactive Control of Fast-Acting Demand Response Based on Thermostatic Loads in Real-Time Electricity Markets

This paper has been submitted to Applied Energy journal in March 2017:

Sahand Behboodi, David P Chassin, Ned Djilali and Curran Crawford. *Transactive Control of Fast-Acting Demand Response Based on Thermostatic Loads in Real-Time Electricity Markets*.

Sahand Behboodi has developed the theory, coding the simulation, and writing the text. David Chassin has helped Sahand in analyzing the simulation results and writing the manuscript.

This chapter addresses the topic of demand response control for thermostatic loads. We present a new operation strategy for HVAC loads that minimizes the energy costs and prevents the grid congestion, while maintaining the room temperature in the comfort bound set by the consumer. In the this strategy, thermostats participate in a 5-minute retail electricity market based on transactive control paradigm. An agent-based approach is used to simulate the collective behaviour of HVAC loads to price changes in both heating and cooling modes.

6.1 Abstract

Coordinated operation of distributed thermostatically controlled loads (TCLs) such as heat pumps and air conditioners can reduce energy costs and prevents grid congestion, while maintaining room temperatures in the comfort range set by consumers. This paper furthers efforts towards enabling thermostatically controlled loads to participate in real-time retail electricity markets under a transactive control paradigm. An agent-based approach is used to develop an effective and low complexity demand response control scheme for TCLs. The proposed scheme adjusts aggregated thermostatic loads according to real-time grid conditions under both heating and cooling modes. A case study is presented showing the method reduces consumer electricity costs by over 10% compared to uncoordinated operation.

Keywords

Agent-based modeling, market-based control, smart grid, thermostatically controlled loads

Highlights

- Market-based control for flexible loads based on transactive paradigm
- Load aggregation of thermostatically controlled loads using an agent based approach
- Demand response with fast reaction to market price fluctuations
- Reduced electricity costs in both cooling and heating modes

6.2 Introduction

Demand response automation allows customers to adjust their electricity usage in response to changes in energy prices or to incentive payments. Price changes and incentives can induce lower electricity use at times of high market prices or when system reliability is jeopardized [14]. Demand response can be a valuable resource for system operators, particularly when significant levels of intermittent renewables are present [89, 132]. Loads that are well suited for demand response include heat pumps [133], air conditioners [28], domestic hot water tanks [134], plug-in electric vehicles [72], water distribution systems [135], electrolyzers [56] and smart appliances such as laundry machines and dishwashers [136].

The inherent flexibility of thermostatically controlled loads (TCLs) such as heating, ventilation and air conditioning (HVAC) allows consumers to use their devices in a grid-friendly way [137], reduce energy costs [138] and potentially mitigate CO_2 emissions [139]. The benefits of demand response are numerous and varied but must be structured carefully to be fair to all participants. An integrated model of demand and supply is presented in [140] to evaluate the overall cost benefit of introducing demand response for electric heating systems (building interior and domestic hot water) coupled with thermal energy storage. This model shows that (i) demand response can assist the grid operator in dispatching the generation fleet in a more economically-efficient manner, and that (ii) the higher the renewable generation, the greater the benefit of demand response.

In recent years several operational strategies have been suggested using TCLs to mitigate the imbalance between power supply and demand. Many involve managing TCLs by varying the setpoint temperature within a limited range [17, 141]. Understanding the behavioral characteristics of the user is crucial to designing an effective control mechanism for demand response. This is a very broad area of research, with opportunities for many differing approaches. These include a logit form of discrete choice model to represent the aggregate load behavior under critical peak pricing in [142] and under real-time pricing [111].

Incentive-based strategies dispatch flexible loads through economic means [143]. Among this set of operational strategies, price-based [144, 145] and market-based control methods [146, 53, 54] provide significant and useful responses to real-time grid conditions. In the former method, the utility sends a price signal to flexible loads which respond in accordance to their states and settings. The price is such that the collective response of loads is ideally as much as the grid operator needs to maintain the energy balance in an economically-efficient manner or to track an optimal load schedule. There are concerns regarding fairness [147] and stability of such a load control strategy [148, 21]. In the latter method (market-based), flexible loads submit their willingness-to-pay price for a particular level of demand for electricity to the market operator. This buy bid is computed based on the load state, consumer comfort objective as well as an expectation of future market price and its volatility. The market operator collects buy and sell bids to clear a double auction market, and broadcasts the clearing price, while scheduling its operations for the cleared load quantity. If the buy bid of a participant is above the clearing price, the participant will consume electricity (and pay the clearing price); otherwise it will forgo consumption until the next market clearing. This indirect control operation method is favored among customers in comparison with direct control methods because it does not restrict customers autonomy or invade their privacy [53].

Price-based demand response for large population of loads has long been used in conjunction with time-of-use rates and critical-peak tariffs, as well as peak-time rebates. More direct aggregate TCL control methods are based on linearized models of open loop control, and some include a feedback mechanism to track a target aggregate load. Open loop aggregate TCL control models based on first principles [149], parameter identification [150], and model identification [151] have long been available. Most recently, methods that close the aggregate TCL control loop were proposed by Perfumo et al. [152], Kok et al. [146], Hammerstrom et al. [53], and Widergren et al. [54]. The latter three were demonstrations of market-based mechanisms in pilot projects using real customers who bid for energy use and were presented with real-time prices in utility operations.

Market-based control using bidding loads can be implemented by retrofitting existing thermostats, such as used in previous demonstrations [146, 53, 54]. An alternative approach is to design a new type of thermostatic controller that updates only when the market clears, and has no prescribed deadband. The advantage of this approach is that demand response better tracks the quantity corresponding to the clearing price [75].

Real-time electricity markets serve as a platform for matching the supply and demand of electricity on a least cost basis. The bid (buy) quantity of individual flexible loads is substantially smaller than the bid (sell) quantity of generation units. At the same time, the number of participating loads is much larger than the number of suppliers. Therefore, it is challenging to include these loads individually in the wholesale market in conjunction with large generation units. This challenge is addressed by aggregating loads so that the load aggregator (similar to a retailer) participates in the wholesale market on behalf of them. Load aggregation reduces the communication complexity and also eases the market settlement process, although it may influence efficiency of the resource use. An immediately important research question to investigate is the suitable size of a load aggregator that participates in a wholesale market.

Recent advances in Information Technology enables smart grids to include individual distributed generation units and flexible loads in retail markets so that they can together reduce the energy costs [113]. Retail transactive control systems are similar to energy wholesale markets, but the price signals are applied at much finer temporal and physical granularity than is possible in wholesale markets. However, the need to deploy suitable infrastructure remains a significant barrier to implementation of market-based control methods. Flexible loads must receive real-time information about the power grid condition to quantify their willingness-to-pay price, considering their flexibility and comfort. For instance, this information can be a price signal representing the estimated mean and variance of electricity price over a time period in near future (e.g. 1 hour), based on which price-sensitive loads compute their willingness-to-pay prices. Integrated closed-loop and bid/response market mechanisms (e.g. transactive control) require two-way communications infrastructure, which may come at higher costs than one-way open-loop dispatch mechanisms, but may enhance system stability and efficiency.

The value of demand response in supporting grid operation is more pronounced when ancillary services are also included along with basic temporal load shifting [153]. Accordingly, a control scheme is required that enables thermostatic loads to provide a fast-acting, energy neutral response to real-time grid conditions. We propose a market-based (indirect and centralized) load control method that includes TCLs in real-time retail markets under transactive control, using agent-based modeling approach. The main contributions of this work are: (i) introduction of a new market-based control scheme for TCL, (ii) reduction in user input and model complexity, and (iii) demonstration of effectiveness of the transactive control scheme.

This paper begins with an introduction to the proposed bidding strategy and market structure in Section 6.3. A case study to demonstrate the performance of the presented transactive control method in terms of cost saving against uncoordinated operation is provided in Section 6.4 for both heating and cooling HVAC operation modes. Finally, conclusion and suggestions for future work are presented in Section 6.5.

6.3 Model description

In this section, we present the thermostatic transactive control setup. A new thermostat design is presented that reduces the complexity of implementing the control method. A bidding strategy for thermostatic loads is proposed based on the expected mean and variance of the market price over a future time window. At the end, a simple thermal model is provided to represent energy balance in buildings and serving as the simulation physics model.

6.3.1 Temperature state

In contrast to regular thermostats, demand responsive thermostats do not have one setpoint temperature with deadband-based control. Instead they allow for a range of temperatures that may vary according to both the thermal comfort and cost savings desired by the consumer. An aggregated thermostatically controllable load consists of a non-trivial number of thermostats that control heat pumps or air conditioners by maintaining room temperature (T) in a customer-specified temperature range of $[T_l, T_u]$. We define a dimensionless temperature (state):

$$\theta = \frac{T - T_l}{T_u - T_l}$$

that varies over the range of $[0, 1]$. In heating mode, the absolute value of the temperature upper bound is greater than the absolute value the temperature lower bound ($|T_l| < |T_u|$), while in the cooling mode the opposite is true ($|T_l| > |T_u|$). When $\frac{1}{2} < \theta < 1$ the space temperature is closer to the upper temperature bound than to the lower bound and there is *less* pressure for the load to be in the ON mode. Conversely when $\theta < \frac{1}{2}$ the room temperature is closer to the lower bound temperature and there is *more* pressure for the load to be ON. Notice that T_l and T_u are specific to each load. Without any coordinated control strategy, as long as room temperature is in the range of (T_l, T_u) the heat pump or the air conditioner does not operate. As soon as the room temperature reaches T_l or $\theta = 0$, the thermostat commands the device to operate until the room temperature reaches T_u or $\theta = 1$. However, with a coordinated control strategy, individual thermostats command differently, although they maintain the room temperature in the same comfort range. Consumers can change their temperature range throughout the day according to their comfort and savings needs, although in the simulation in Section **??** we do not consider such variations.

6.3.2 Bidding strategy

We consider a hypothetical real-time retail electricity market for a distribution feeder in which each thermostat submits a buy bid based on its value of θ and expected future prices. The retail market operator gathers these buy bids as well as sell bids of distributed generation units to clear the market every τ minutes (market update cycle), considering the locational marginal price (LMP) and unresponsive load quantity. On behalf of loads and generation units in the retail market, the operator also participates in a separate wholesale market to determine the hourly LMP, thereby acting in the wholesale market as an aggregator. We could have alternatively considered that thermostats provide ancillary services to the grid; then they would be sellers of ancillary services as well as buyers of energy.

The retail market operator broadcasts the clearing price (P_c) in \$/MWh as well as the estimated mean price ($\bar{P}$) and standard deviation price ($\hat{P}$) over a future time window of duration (t_{fw}) to all TCLs. Active thermostats provide their willingness-to-pay per unit of energy price by computing the bid price:

$$\phi = \bar{P} + \sqrt{2}\ \mathbf{erf^{-1}}(1 - 2\ \theta)\ \hat{P}, \tag{6.1}$$

where $\mathbf{erf^{-1}}$ is the inverse error function. If the bid is above the clearing price the thermostat consumes electricity, otherwise it forgoes until the next market cycle (e.g. $\tau = 5$ minutes $< t_{fw} = 1$ hour). The proposed bidding strategy is based on the quantile function (inverse cumulative distribution function) of the Gaussian distribution. The magnitude of θ determines the probability of its associated bid price being above the clearing price over the future time period of t_{fw}. The bid will be greater than the expected mean price when $\theta < \frac{1}{2}$ and below the expected mean price when $\theta > \frac{1}{2}$.

In the simulation presented in the next section, every household is assumed to be equipped with a rooftop PV solar panel. The magnitude of the net load on the feeder (the total unresponsive load plus the total responsive load minus the total PV generation) should not exceed the feeder capacity. Scarcity rent from any lack of PV generation will raise auction clearing prices and induce thermostats to reduce load to match the available feeder capacity.

Figure 6.1 shows the structure of the control process. The wholesale market operator provides a forecast of LMPs over a future time window of t_{fw}. The participating thermostats compute their bids based on the mean and standard deviation of these LMPs as well as their temperature states. A model to predict the temperature state is described in the next subsection. When the price becomes steady over t_{fw} (zero or otherwise), the effect of θ on the bid price is ignored since $\hat{P}$ is zero, and as a result every bid will be equal to $\bar{P}$. To avoid this issue, a minimum price (e.g. 1 \$/MWh) can be specified for $\hat{P}$.

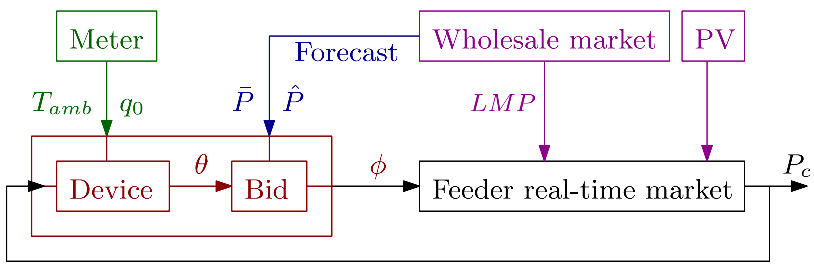


Fig. 6.1: Control process diagram.

6.3.3 Building thermal model

A lumped system thermal energy balance equation for a household unit is used to define the relevant physics of the system:

$$M\ \frac{dT}{dt} + U\ (T - T_{amb}) - q_0 - q_{hvac} = 0, \tag{6.2a}$$

where M, U, q_0, q_{hvac} are respectively thermal mass in kJ/°C, overall heat transfer coefficient in kW/°C, total thermal load in kW$_{th}$, heat pump or air conditioner output in kW$_{th}$, specific for each unit. The total thermal load (q_0) encapsulates multiple sources such as internal loads. We assume M, U, T_{amb}, $q_0 + q_{hvac}$ do not vary over time τ. Solving this differential equation, the temperature state of load j in the next market cycle $(k+1)$ will be:

$$\theta_j(k+1) = e^{-\frac{U_j \tau}{M_j}}\ \theta_j(k) + (1 - e^{-\frac{U_j \tau}{M_j}}) \times \left(\frac{T_{amb} - T_l}{T_u - T_l} + \frac{q_0 + q_{hvac}}{U\ (T_u - T_l)} \right)_j. \tag{6.2b}$$

The value of q_{hvac} is positive and negative, respectively, for heating (obviously $T_{amb} < T_{l_j}$) and cooling ($T_{amb} > T_{l_j}$) modes, and zero during an OFF cycle ($\phi_j < P_c$).

6.4 Performance analysis

In this section, we present a simulation test to demonstrate the performance of the proposed transactive control method for thermostatic loads in a real-time market environment.

6.4.1 Case study

General assumptions are given in Tables 1 and 2. The market update cycle (τ) is set to 5 minutes to be similar to the redispatch time of energy imbalance markets in CAISO [154]. Thermal model parameters in Equation 6.2b are randomly and uniformly distributed between the minimum and the maximum values given in Table 6.2, and most of them are taken from [152].

A detailed energy balance simulation of buildings is not in the scope of this work, and as noted earlier we use a lumped system approach to represent the thermal behavior of each building. A significant portion of the energy consumption of devices that are not involved in the demand response program (unresponsive electric loads) is eventually converted to heat in conditioned spaces. Another important heat source that should be taken into account is the solar gain (through windows and/or walls), which depends on the location, orientation and architecture of the buildings. For simplicity, we assume the magnitude of solar gain is linearly related to the magnitude of the solar PV panel output for the building. Accordingly, we assume the magnitude of q_0 is equal to the magnitude of the PV output plus 86% [134] of the magnitude of the unresponsive load, ignoring the ventilation and non-electric internal loads.

Table 6.1: Inputs

Parameter	**Unit**	**Value**
Feeder capacity limit	kW	6000
Market cycle (τ)	min	5
Mean non-HVAC load	kWh/h	2.0
Number of participants	–	1000

Figure 6.2 illustrates the a typical profile of the LMP that comes from the underlying wholesale energy market that changes on an hourly basis. The magnitude of variation of the total non-TCL (unresponsive) demand is also shown. We assume LMP and non-TCL demand profiles are the same for the winter and summer days. In addition, the total solar

Table 6.2: Model parameters

Parameter	Unit	Min	Max
Air conditioner capacity	kW_e	3.00	5.00
Air conditioner COP	–	1.88	3.12
Ambient temp. summer	°C	26.0	32.0
Ambient temp. winter	°C	-2.0	4.0
Building thermal mass	MJ/°C	54.1	90.0
Comfort range size	°C	0.85	1.15
Heat pump capacity	kW_e	3.00	5.00
Heat pump COP	–	2.63	4.37
Heat transfer coefficient	kW/°C	0.38	0.62
Indoor temp. summer	°C	15.5	25.4
Indoor temp. winter	°C	14.6	24.4
Solar PV capacity	kW	2.25	3.75

PV generation in a mid January day and also a mid June day based on historical data in Victoria, BC (approximately 48°N latitude) are shown. From the assumed profiles, the daily solar PV output per household is 2.64 kWh and 24.29 kWh on a mid January day and on a mid June day, respectively. It should be mentioned that all these profiles are inputs to the simulation.

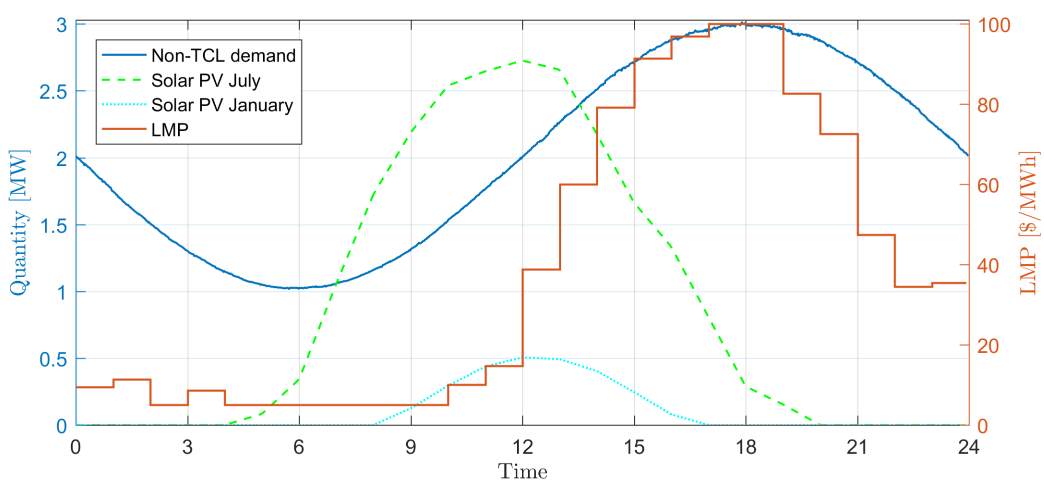


Fig. 6.2: LMP on the feeder and total unresponsive load.

6.4.2 Heating mode

Figure 6.3 illustrates the temperature state distribution (horizontal axis) and the associated bid price distribution (vertical axis) at 12 AM, 9 AM, 3 PM and 9 PM on the second day of operation with the proposed control method. In this simulation, we assume the initial temperature states (12 AM day 1) are random-normally distributed around 0.5. Performing the simulation with different initial temperature state distributions leads to similar results indicating insensitivity to initial conditions. The bid prices are normalized with respect to the mean LMP, $\bar{P}$, over the future time window of an hour. The horizontal and vertical boxes show the distribution density of the temperature states and the associated bid prices respectively at the aforementioned time slots. In each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25^{th} and 75^{th} percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the "$\times$" symbol. We note that the temperature state of most loads are very close to the lower bound at 12 AM and 9 PM, suggesting loads are saturated probably because they have experienced a relatively expensive electricity period prior to these hours. On the other hand, the temperature state distribution at 3 PM shows an opposite profile where more temperature states are close to the upper bound. This behaviour indicates loads are anticipating more expensive electricity hours by pre-heating. Finally, the 9 AM temperature states are well distributed, with a median closer to the upper bound though.

Figure 6.4 illustrates the market settlement conditions at the aforementioned time slots. The small circle marks the clearing point at the intersection of the supply curve (red) and the demand curve (blue), associated with each time slot. The demand curve includes an unresponsive part (not-shown) and another part consisting of responsive TCL bids that depend on room temperature, comfort range and also price average and volatility over a future time window of an hour. The supply curve consists of two segments and a vertical component, which enforces the feeder constraint. The first segment represents the total solar PV generation at zero marginal cost, and the second segment represent the feeder (with a constrained capacity) at a given LMP reflecting the underlying wholesale market situation. PV generation at 12 AM and 9 PM is zero. As mentioned, a significant proportion of TCL demand is fully saturated at 12 AM and 9 PM and becomes unresponsive to electricity price. At 9 PM, the clearing price rises to the point that the responsive portion is such that the cleared load quantity does not exceed the feeder load constraint.

Figure 6.5 shows the evolution of temperature states over a day of operation (right side vertical axis) using boxplots. We note that the temperature state of most loads that are fully depleted at midnight ($\theta \rightarrow 0$) increases so they are fully "charged" by noon ($\theta \rightarrow 1$). The dispatched load that is at its maximum at 1 AM gradually decreases during the flat LMP period until 5 AM at which the expected price for the next time slot (6–7 AM) goes up, so the dispatched load increases to its maximum again to pre-heat the room. Subsequently, loads maintain their high charge state (high θ) in anticipation of the

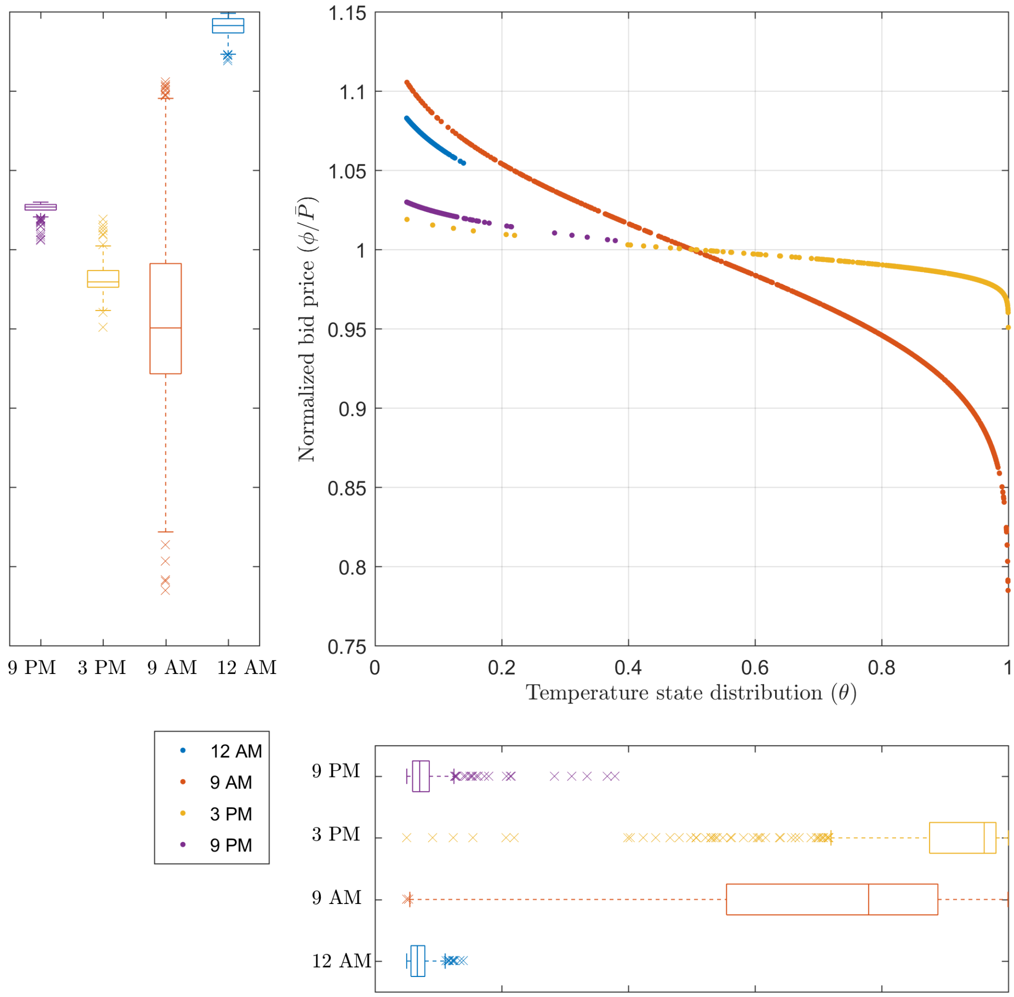


Fig. 6.3: Temperature state distribution vs. bid price distribution at 12 AM, 9 AM, 3 PM and 9 PM on a mid January day: whiskers "$\|$" and outliers "$\times$".

expensive evening period during which most switch to the OFF mode. At night time, loads consume a minimal amount of electricity (enough to maintain the room temperature in the comfort range) despite their low θs. This behaviour is because loads benefit by waiting for the after-midnight inexpensive electricity period. A dot indicates the clearing price at time t and its color represents the associated dispatched responsive load (left side vertical axis). Clearing prices are the same as the hourly LMPs, except when the clearing price rises to prevent feeder overloading. The net load on the feeder should always be smaller than the feeder capacity limit.

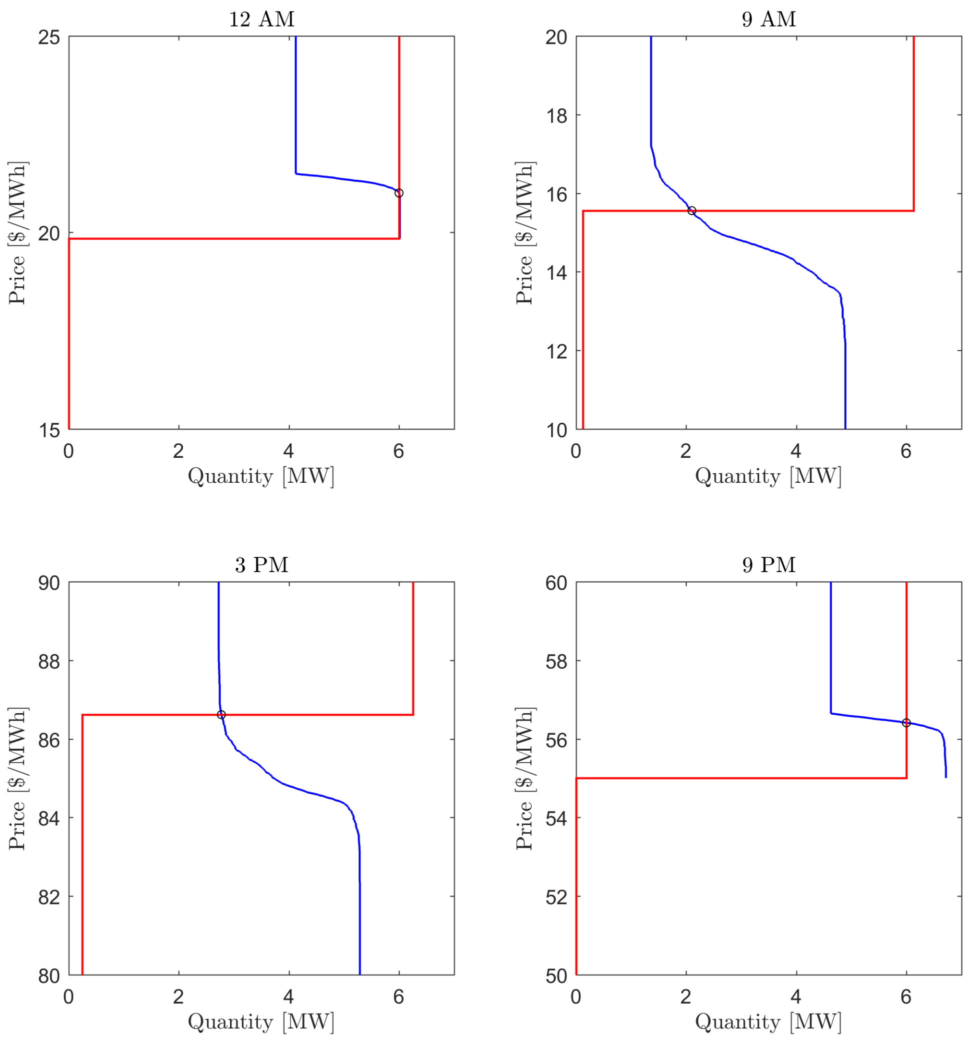


Fig. 6.4: Market settlement at 12 AM, 9 AM, 3 PM and 9 PM on a winter day: supply curve (red) and demand curve (blue).

Figure 6.6 illustrates the clearing price as well as the total load before and after implementation of the proposed method of operation on a mid January day. The clearing price in the uncoordinated case is the same as the input LMP profile since the uncoordinated load does not exceed the feeder capacity limit in this simulation. We notice that the implementation of this strategy shifts a significant portion of the TCL energy requirements from relatively expensive periods to inexpensive periods. As a result the blended clearing price (weighted average price also equal to daily cost divided by daily consumption) decreases from 42.43$/MWh to 38.50$/MWh. On the other hand, it introduces some oscillations and

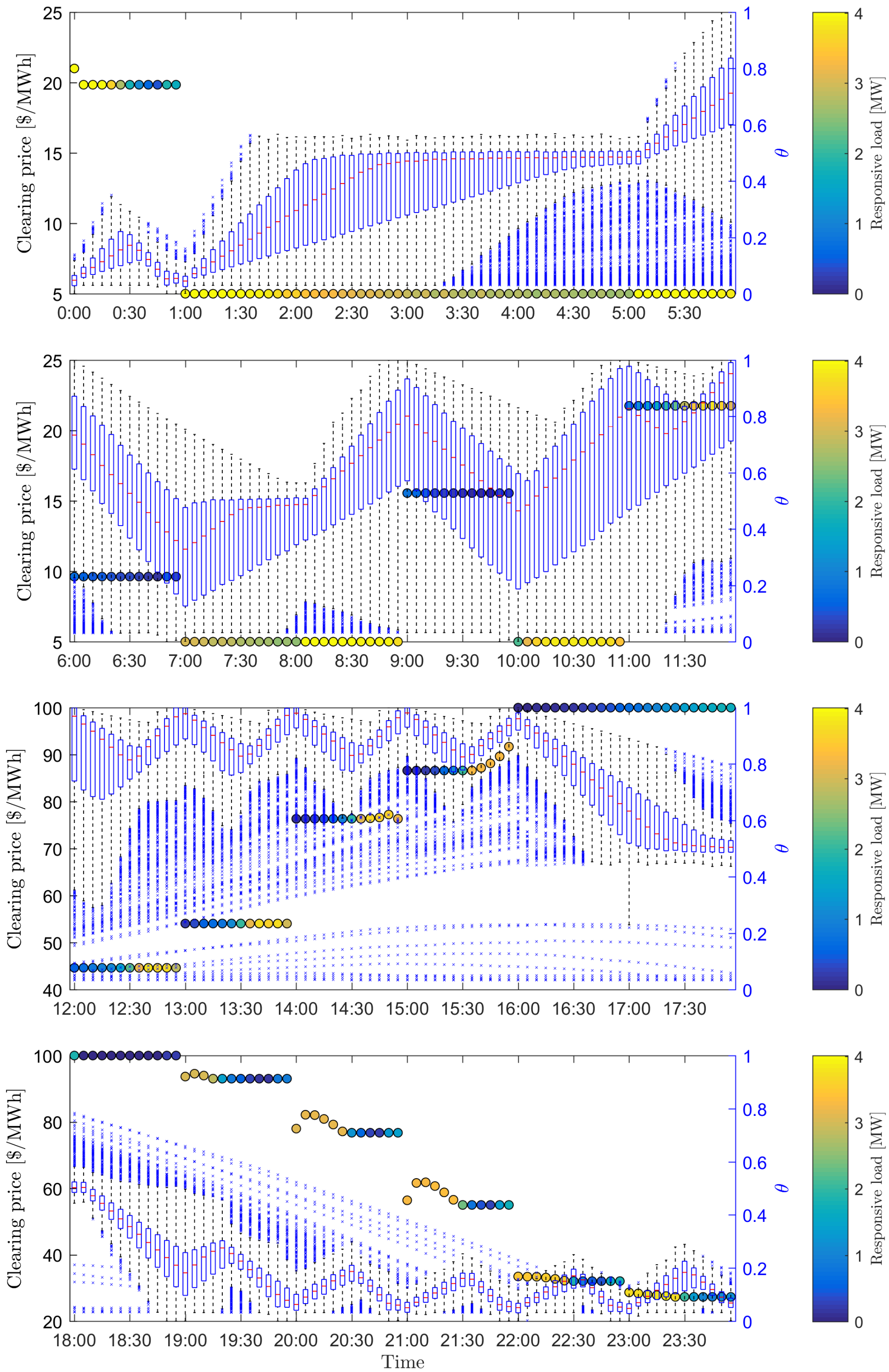


Fig. 6.5: Temperature state evolution (heating mode): temperature state distribution (blue boxplots), clearing price (circles) and cleared responsive load (colorbar).

spikes in the demand profile, which can cause operational problems such as harmonic voltages and currents. There exist solutions to avoid this issue, principally employing lockout times and bid compensation [54]. Alternatively, the aggressive reaction of loads to LMP changes can be mitigated to some extent when the grid operator provides a longer-time forecast of $\bar{P}$ and $\hat{P}$, thereby reducing the impact of sudden price changes.

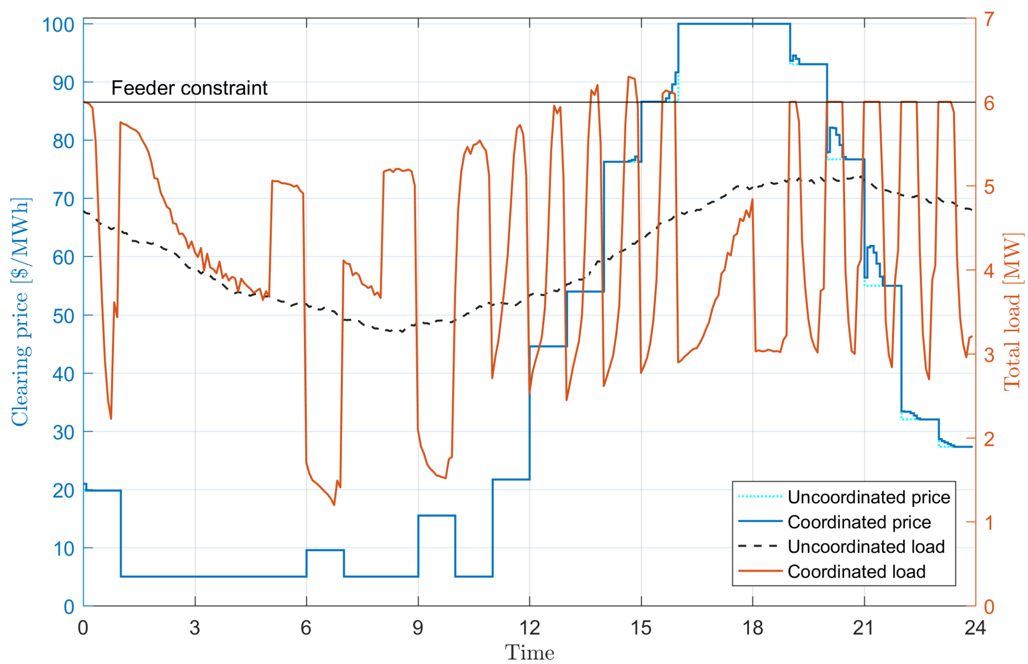


Fig. 6.6: Clearing price and total load profile (heating condition).

Another approach to mitigate demand fluctuations is to modify the bidding strategy to include the ON→OFF and OFF→ON switch costs. This would also reduce cycling wear out of HVAC systems. Accordingly, we increase the bidding price for those thermostatic loads that are in the ON state by ϕ_s ($ per switch), and decrease the bidding price of OFF state loads by ϕ_s. With the switching price included, thermostats only respond to substantial price variations. Including the switching cost effectively shifts bid prices from the marginal point, which makes the settlement process less efficient and as a result decreases the potential cost reduction, but smooths out the load profile. The magnitude of ϕ_s can be either fixed or variable (e.g. a constant percentage of the original bid over time). It can also vary from customer to customer.

Figure 6.7 illustrates the shape of the responsive load under different control strategies. Case A is the profile shown in Figure 6.6. Case B is similar to case A with $\bar{P}$ and $\hat{P}$ values obtained in a six hour future time window instead of in a one hour window. For simplicity, we do not consider any forecast error in estimating $\bar{P}$ and $\hat{P}$ values. In case C, each bid price is adjusted by including/deducting a percentage ($\pm r$) of the bid price for the switching cost, which is selected randomly and uniformly between 0 and 50%. The sign is positive for loads in the ON state, and negative for loads in the OFF state. There are periods during which the bid price is zero or small so that this increase does not make a difference. Thereby, whenever the original bid (ϕ) is below a certain value (ϕ_0) we use ϕ_0 (here 33$/MWh) to compute the adjusted bid (ϕ^*):

$$\phi^* = (1 \pm r) \times \max(\phi, \phi_0).$$

Finally, case D combines both modifications suggested in cases B and C.

Table 6.3 summarizes the electricity blended price as well as the daily energy consumption associated with each case. Including the switching cost (case C) slightly increases the daily energy consumption while estimating $\bar{P}$ and $\hat{P}$ over a six hour time window (case B) reduces the consumption. In addition, we note that the blended price in case C is significantly higher. Large load changes may require additional voltage control action within the distribution system, however coordination of the load control system and the voltage control system is not in the scope of this work.

Table 6.3: Performance comparison

Parameter	**Unit**	**Case A**	**Case B**	**Case C**	**Case D**
Electricity blended price	$/MWh	38.50	39.44	45.10	37.41
Daily energy consumption	MWh	1204	1072	1215	1204

Table 6.4 compares the energy cost and the blended price with the control strategy (case D) for heat pumps versus the uncoordinated scenario. The energy cost is reduced by 10.2%, while the energy consumption of the heat pumps decreases by 1%. Note that this cost saving happens on a day that the system utilization factor (ratio of the time that a equipment is in use to the total time that it could be in use) is around 54%. If the average ambient temperature increases by only 1°C, the utilization factor will decrease to 51% and the cost saving percentage will increase to 10.6%. If the average ambient temperature increases by 5°C, the utilization factor will decrease to 38% and the cost saving percentage will increase to 11.2%. Note that a significant portion of the energy cost belongs the unresponsive loads. It should be mentioned that this cost reduction does not reflect the entire achievable benefits of demand response integration. Potentially, demand response can reduce the reserve requirements and assist with primary frequency regulation.

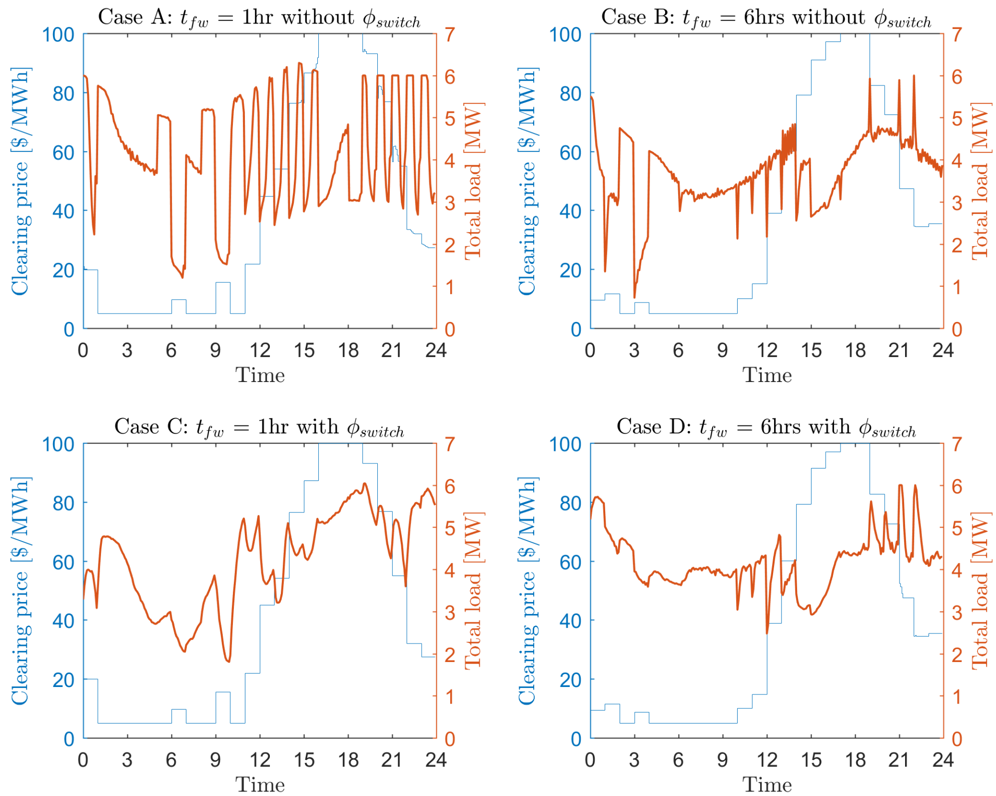


Fig. 6.7: Total load profile under different control strategies (heating condition).

Table 6.4: Results (heating mode)

Parameter	Unit	Coordinated	Uncoordinated
Daily electricity cost	k$	45.02	50.11
Blended price	$/MWh	37.41	41.42

6.4.3 Cooling mode

In cooling mode, solar gain no longer works in favour of the HVAC system, although solar PV panels provide a significant portion of customers demand in summer. Figure 6.8 illustrates the market settlement at 12 AM, 9 AM, 3 PM and 9 PM on a mid June day. We should mention that all results reported in this subsection are related to the case *D*: six hours future time window and a switching cost as a percentage of the original bid specific for each consumer. Including the switching cost in the bidding spreads out the demand curve compared to those in Figure 6.4. Moreover, we note negative bid prices at 12 AM and 9 AM.

Figure 6.9 shows the clearing price as well as the total load before and after implemen-

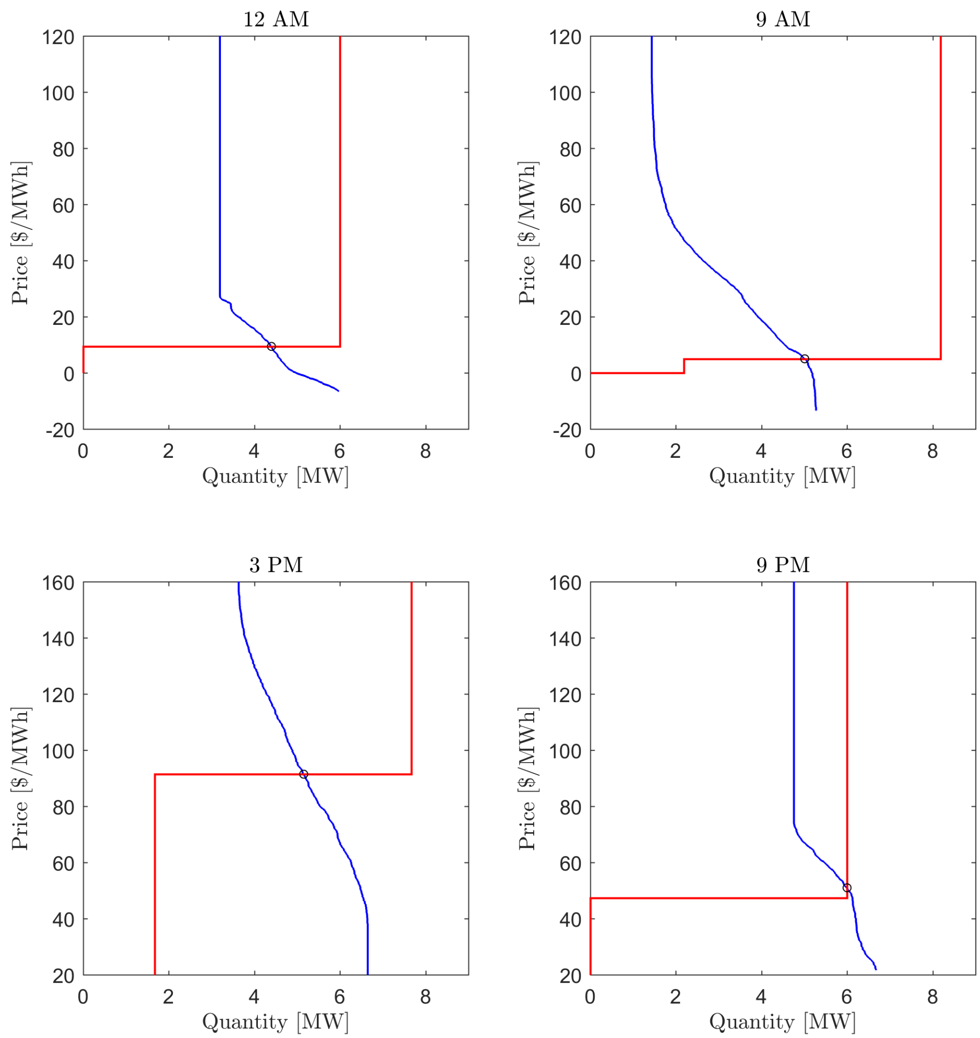


Fig. 6.8: Demand curves at 12 AM, 9 AM, 3 PM and 9 PM on a mid June day: supply curve (red) and demand curve (blue).

tation of the proposed method of operation on a mid June day.

Rooftop solar PV panels supply a significant portion of demand on the mid June day so that the electricity price is lower than the mid January day. Table 6.5 compares the energy cost and blended price associated with the control strategy for air conditioners versus the uncoordinated case. The energy cost is reduced by 10.6%, while the energy consumption of the air conditioners does not change. Note that this cost saving happens on a day that the utilization factor of air conditioners is around 70%. If the average ambient temperature decreases by only 1°C, the utilization factor will decrease to 66% and the cost

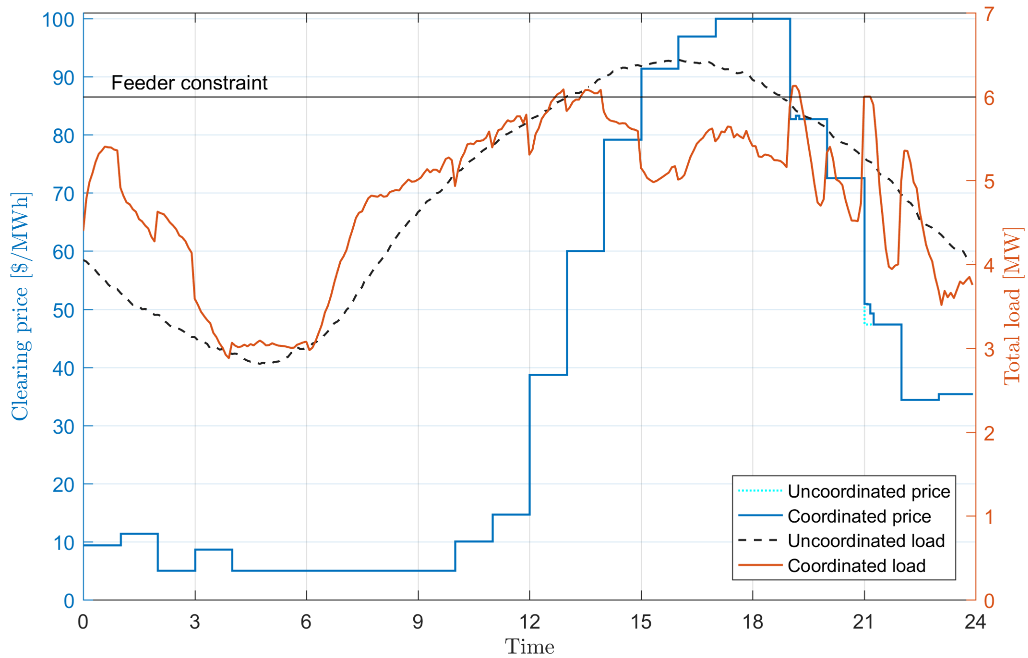


Fig. 6.9: Clearing price and total load profile (cooling mode).

saving percentage will increase to 11.9%.

Table 6.5: Results (cooling mode)

Parameter	Unit	Coordinated	Uncoordinated
Daily electricity cost	k$	48.13	53.87
Blended price	$/MWh	34.92	39.09

6.5 Conclusions

In this paper, a new market-based control method for thermostatic loads is presented based on transactive control paradigm. Implementation of the method requires real-time two way communication between the loads and the market operator. An agent-based modeling approach is used to aggregate loads. The demand response aggregator is highly responsive to market price fluctuations. The demand profile is smoothed by introducing a switching cost as well as providing price forecasts (mean and standard deviation) over a longer future time window. This control method reduces the electricity cost by 10.2% on a cold day, and by 10.6% on a hot day in comparison with the uncoordinated operation, for a case study with given inputs.

Flexible loads can assist the system operator maintain the grid frequency and voltage within the admissible range of their nominal value. Therefore, a future objective is to design a control method that incorporates energy balancing service with frequency/voltage regulation. Another essential extension of this work is to investigate the stability of the proposed operation strategy.

In addition to thermostatically controlled loads, plug-in electric vehicles and smart appliances can also be used to provide ancillary services if the smart grid infrastructure is available. However, the characteristics (magnitude, duration, frequency and availability) of these loads are different from HVAC loads; thus their operation strategy will be different. Accordingly, a more advance control scheme will be needed to optimally control an aggregate of different loads.

A physical demonstration, similar to the Decentralised Energy Exchange (deX) in Australia [155], is eventually required to gather field data, and to have a better understanding of consumers' behavior and technical barriers such as communication and metering response, particularly for tariff-based incentives.

Chapter 7

Conclusions and Future Work

Generation variation is a challenge to the seamless integration of renewable energies. The traditional approach to dealing with generation variation is to operate additional spinning reserve units, which often is too costly and can lead to increased overall emissions. Super grids and demand response programs are lower-cost solutions to mitigating generation variation. This dissertation suggests (1) taking a unified approach to performing resource allocation at the interconnection scale, and (2) operating distributed energy resources, including flexible electric loads, in a grid-friendly way to deal with power fluctuations. Both suggestions lead to a lower overall electricity price that can facilitate renewable integration.

In summary, the main contributions presented in this dissertation are:

(1) **Demand response integration into wholesale markets**
A new model of demand response integration in bulk electric systems is presented. With price-responsive loads included, the market settlement is such that the economic surplus is maximized. Thereby, the resource allocation objective function is reformulated to maximize surplus rather than minimize cost to include demand response.

(2) **Optimal export/import schedule to maximize the interconnection surplus**
A centralized scheduler is developed to determine the unconstrained and constrained power transfers within an interconnection in order to maximize the overall surplus. This resource allocation method can potentially lead to a similar generation and load dispatch schedule that would result from operating the entire assets (in the interconnection) under a hypothetical single giant control area. Although neither of these approaches are practically implementable, they set upper bounds to lowering the operational cost that distributed control approaches, i.e. agent-based control methods, can achieve. Also, a new parameter, economic utilization factor, is defined and used to evaluate the performance of system interties.

(3) **Super grid operation of the Western Interconnection**
The benefits of operating the Western Interconnection in a unified manner is explored, using a reduced model of the WECC system consisting of 20 consolidated areas, each with a hypothetical wholesale market for the planning year 2024. A key result was that integrating demand response in the super-hourly time scale does not contribute to a significant cost reductions. This suggests sub-hourly scale examination is required.

(4) **Transactive control strategy for plug-in electric vehicles (smart charging)**
A load management scheme is suggested that includes electric vehicles in a 5-minute retail double auction market. Every vehicle submits its buy and sell bids to the market operator, considering its demand urgency and the future price opportunity (transactive control paradigm). Comparing the cost reduction associated with either unidirectional charging (V1G) or bidirectional charging (V2G) scenarios with the uncoordinated charging (V0G) scenario, it is concluded that smart charging can significantly reduce the energy cost, in particular in the presence of substantial amount of variable renewable generation. Simulation results indicated V2G does not provide significant additional benefits, given big impact on battery life and complexity.

(5) **Market-based control approach for thermostatically controlled loads**
An agent-based demand response model is developed for thermostatically controlled loads. The proposed operation method reduces complexity by introducing the temperature comfort range. Every thermostatic load participates in the retail market, considering its temperature state (load flexibility) and expected price fluctuation (mean and standard deviation) in a future time window. This method of operation prevents overloading the supply feeder, and also provides an efficient short-term matching of distributed generation and flexible load resources.

7.1 Summary

The key conclusions of this dissertation are as follows:

In Chapter 2, we introduced a simple cost model of intermittent renewable and demand response integration that can be used to determine the optimal level of variable generation resources in an electric system. The model incorporates production, uncertainty, emission costs, as well as capacity expansion and mothballing costs, and considers the impacts of generation variability and load flexibility on the hourly price of electricity. We found that for a typical power grid, the cost function is minimized with about one third variable renewable generation. The results showed that the optimal level of renewable resources is highly sensitive to a carbon tax and the resource's capital cost. We also noticed that integrating inter-hourly demand response does not significantly impact the optimal level.

In Chapter 3, we modified the unit commitment objective to maximize the economic surplus instead of minimizing the operational cost, in the presence of price-sensitive loads. We also discussed the impact of electricity imports/exports (within an interconnected system) on the economic surplus. Next, we presented a new interconnection-wide resource allocation modeling method that incorporates both energy supply and demand response resources, and then tested it on a hypothetical system consisting of two consolidated areas that loosely represents North America's Western Interconnection. The optimal generation schedule associated with the stand-alone, copper-sheet (unconstrained-flow) and constrained-flow models of the interconnection were compared. Furthermore, we assessed the impact of expanding the system tieline capacity on inter-change scheduling. We found that, as the level of variable generation increases, the optimal use of tielines becomes more beneficial.

In Chapter 4, we presented a method to determine the optimal resource allocation hour-ahead schedule in an interconnection. The purpose was not to solve the interconnection resource allocation problem itself; rather, we sought to provide a method that facilitates impact studies of renewable generation and demand response interaction at the interconnection scale without fully implementing the various existing market designs. The method determines the optimal inter-area transfer set that effectively makes the clearing price uniform over the interconnection's markets (to the extent possible). We demonstrated the method on a 20-area reduced model of the WECC system (each with a hypothetical wholesale market) for the planning year 2024. Simulation results illustrated that inter-regional collaboration assists control areas in reducing production costs by 22% relative to the standalone condition with inelastic demand. In addition, results indicated that the overall consumer surplus increases and the overall producer surplus decreases, while the increased magnitude is always greater than the decreased magnitude. Thereby, the global surplus (total consumer and producer surplus) increases. At the same time, the magnitude of production cost decrease is greater than the magnitude of producer surplus decrease. We also quantified the potential benefits of demand response integration in wholesale markets, and noticed that the economic benefit of inter-hourly demand response is not substantial when the system is operated with the optimal inter-area transfer set. We concluded that the majority of the benefits for demand response arise from intra-hourly demand response resources which are not considered in conventional hourly resource scheduling mechanisms.

In Chapter 5, we demonstrated how advanced electric vehicle charging strategies can be used to improve inter-temporal coordination between charging needs and charging opportunities to lower charging costs. When using real-time retail price to reduce the impact of distribution constraints, utilities can also help connect solar PV panels to charging stations and reduce the customer's net vehicle energy costs when compared to the uncoordinated scenario, thus significantly bolstering the economic case of electric vehicles over conventional vehicles. We evaluated these impacts using agent-based simulations of vehicle chargers operating in a real-time pricing environment in conjunction with PV generation. Overall, consumers would realize significant savings when electric vehicles participate in a retail electricity markets and receive real-time price signals which they could compare to buy/sell reservation prices to make charge/discharge decisions. This intra-hourly bid-response strategy requires knowledge of price expectations based on day-ahead hourly wholesale prices, which suggests that the integration of wholesale and retail markets must consider not only topological but also inter-temporal exchanges to induce sufficient demand response participation at the wholesale level.

In Chapter 6, we proposed a new market-based (indirect and centralized) operation method for thermostatically controlled loads based on the transactive control paradigm. Implementation of this method requires a real-time two way communication between these loads and the retail market operator. Similar to the smart charging method, we used an agent-based approach to model the HVAC load control strategy. The simulation results indicated that implementing this control strategy would enable the load fleet to provide a fast collective response to market price fluctuations that reflects underlying technical real-time energy balance requirements at the wholesale market level. The control method reduces energy costs, prevents grid congestion and facilitates renewable integration.

In summary, we concluded that, considering capability of loads to provide a fast response to price changes as well as their flexibility restrictions, market-based demand response programs are useful to mitigate sub-hourly power fluctuations rather than shifting energy demand for super-hourly time periods.

7.2 Future Work

The following studies and modifications are recommended as extensions of the current work to provide a more accurate insight into benefits and limitations of super grids and demand response programs.

(I) **Incorporating reserve requirements in the interconnection-wide economic dispatch problem:** Future work could be done to include reserve scheduling in the resource allocation problem. A co-optimization of energy and reserve allocation is required to provide an efficient use of resources, in particular in the presence of significant amounts of variable generation. In an interconnection, reserve sharing can lower the total operational cost. In addition, the availability of demand response can reduce the need to operate costly spinning reserve units necessary to back up generation and load variations.

(II) **Considering operational constraints for supply and demand resources:** There are a number of operational constraints that are excluded from the proposed scheduling model for simplicity, including ramping limits as well as start-up and shut-down restrictions (a history of events happening in previous hours). It is also required to delve deeper into system reliability and survivability matters e.g., N-1 contingency constraint. Including operational constraints in the resource allocation process would improves on the functionality of the model. In addition, with these constraints included, the total system cost would be described more accurately than the model presented in Chapter 2.

(III) **Investigating the impact of including energy storage systems on the optimal operation:** In conjunction with super grids and demand response programs, energy storage systems can be used to mitigate generation and load fluctuations. A future objective should be exploring the benefits of integrating energy storage systems in both wholesale (e.g. pumped hydroelectric storage) and retail (e.g. battery storage) electricity markets.

(IV) **Developing an interconnection optimal resource planning model:** The cost reduction associated with operating an interconnection in a unified manner is more appreciable with greater penetration levels of renewables. Accordingly, optimal resource planning should also be done at the interconnection-scale. One essential extension of this work is to apply the scheduling method for energy planning purposes. The objective would be to determine the most economically beneficial type and capacity of renewable energies in each area so that the system can achieve the targeted greenhouse gas emissions reduction with the minimal added cost. In addition, we can introduce an interconnection-wide carbon tax to incentivize clean energy resources that can accelerate the phase out process of fossil fuel power plants. As an aside, it is necessary to consider the system long-term resource adequacy in regions with (energy-only) markets.

(V) **Evaluating performance of a demand response program consisting of electric vehicles, smart appliances and HVAC loads:** A model of a demand response program combining technologies with different characteristics (magnitude, duration, frequency and availability) is necessary to understand their collective behaviour. Sometimes these flexible loads can be complementary, and sometimes incompatible. Therefore, a more advanced control scheme should be developed to optimally control an aggregation of different types of loads. In designing load management strategies, robustness of the operation control method should also be considered.

(VI) **Using demand response to provide primary frequency control service:** Flexible loads can assist the system operator in maintaining grid frequency within the admissible range of its nominal value. Therefore, a future objective is to design a control method that incorporates energy balancing services with frequency regulation. In addition, an essential piece of this work would be to examine the stability of the control method.

Bibliography

[1] G Sinden. Assessing the costs of intermittent power generation. *UK Energy Research Centre, London, UK, Tech. Rep*, 2005.

[2] MA Ortega-Vazquez and DS Kirschen. Estimating the spinning reserve requirements in systems with significant wind power generation penetration. *IEEE Transactions on Power Systems*, 24(1):114–124, Feb 2009.

[3] MA Cohen and Callaway DS. Physical effects of distributed PV generation on California's distribution system. *arXiv preprint arXiv:1506.06643*, 2015.

[4] AE MacDonald, CTM Clack, A Alexander, A Dunbar, J Wilczak, and Y Xie. Future cost-competitive electricity systems and their impact on US CO_2 emissions. *Nature Climate Change*, 2016.

[5] G Kalcon, GP Adam, O Anaya-Lara, G Burt, and K Lo. HVDC network: Wind power integration and creation of super grid. In *10th International Conference on Environment and Electrical Engineering*, pages 1–4. IEEE, 2011.

[6] J Aghaei and MI Alizadeh. Demand response in smart electricity grids equipped with renewable energy sources: A review. *Renewable and Sustainable Energy Reviews*, 18:64–72, 2013.

[7] DP Chassin. Load control analysis for intermittent generation mitigation. In *46th Hawaii International Conference on System Sciences (HICSS)*, pages 2305–2311. IEEE, 2013.

[8] JW Feltes, BD Gemmell, and D Retzmann. From smart grid to super grid: Solutions with HVDC and FACTS for grid access of renewable energy sources. In *Power and Energy Society General Meeting*, pages 1–6. IEEE, 2011.

[9] N Samaan, R Schellberg, D Warady, S Williams, R Bayless, S Conger, R Brush, T Gossa, M Symonds, K Harris, et al. Analysis of benefits of an energy imbalance

market in the NWPP. *Pacific Northwest National Laboratory, Richland, WA, PNNL-22877*, 2013.

[10] A Battaglini, J Lilliestam, C Bals, and A Haas. The supersmart grid. In *European Climate Forum*, 2008.

[11] TB Nguyen, N Samaan, and C Jin. Evaluation of production cost savings from consolidation of balancing authorities in the US Western Interconnection under high wind and solar penetration. In *Conference on Technologies for Sustainability*, pages 9–14. IEEE, 2014.

[12] G Wang, Q Zhang, BC Mclellan, and H Li. Multi-region optimal deployment of renewable energy considering different interregional transmission scenarios. *Energy*, 2015.

[13] A Brooks, E Lu, D Reicher, C Spirakis, and B Weihl. Demand dispatch. *IEEE Power and Energy Magazine*, 8(3):20–29, 2010.

[14] U.S. Department of Energy. Benefits of demand response in electricity markets and recommendations for achieving them. *Washington, DC, USA, Tech. Rep*, 2006.

[15] JE Price and J Goodin. Reduced network modeling of WECC as a market design prototype. In *Power and Energy Society General Meeting*, pages 1–6. IEEE, 2011.

[16] M Hunsaker, N Samaan, M Milligan, T Guo, G Liu, and J Toolson. Balancing authority cooperation concepts to reduce variable generation integration costs in the Western Interconnection: Intrahour scheduling. Technical report, March 2013.

[17] S Bashash and HK Fathy. Modeling and control of aggregate air conditioning loads for robust renewable power management. *IEEE Transactions on Control Systems Technology*, 21(4):1318–1327, 2013.

[18] H Wu, M Shahidehpour, A Alabdulwahab, and A Abusorrah. Demand response exchange in the stochastic day-ahead scheduling with variable renewable generation. *IEEE Transactions on Sustainable Energy*, 6(2):516–525, 2015.

[19] D Dallinger and M Wietschel. Grid integration of intermittent renewable energy sources using price-responsive plug-in electric vehicles. *Renewable and Sustainable Energy Reviews*, 16(5):3370–3382, 2012.

[20] W Gu, H Yu, W Liu, J Zhu, and X Xu. Demand response and economic dispatch of power systems considering large-scale plug-in hybrid electric vehicles/electric vehicles (PHEVs/EVs): A review. *Energies*, 6(9):4394–4417, 2013.

[21] W. Zhang, J. Lian, C. Y. Chang, and K. Kalsi. Aggregated modeling and control of air conditioning loads for demand response. *Power Systems, IEEE Transactions on*, 28(4):4655–4664, Nov 2013.

[22] J Druitt and WG Früh. Simulation of demand management and grid balancing with electric vehicles. *Journal of Power Sources*, 216:104–116, 2012.

[23] J Wang, C Liu, D Ton, Y Zhou, J Kim, and A Vyas. Impact of plug-in hybrid electric vehicles on power systems with demand response and wind power. *Energy Policy*, 39(7):4016–4021, 2011.

[24] A Schuller, CM Flath, and S Gottwalt. Quantifying load flexibility of electric vehicles for renewable energy integration. *Applied Energy*, 151:335–344, 2015.

[25] V Lakshmanan, M Marinelli, J Hu, and HW Bindner. Provision of secondary frequency control via demand response activation on thermostatically controlled loads: Solutions and experiences from Denmark. *Applied Energy*, 173:470–480, 2016.

[26] D Wang, S Parkinson, W Miao, H Jia, C Crawford, and N Djilali. Online voltage security assessment considering comfort-constrained demand response control of distributed heat pump systems. *Applied Energy*, 96:104–114, 2012.

[27] J Mathieu, M Dyson, D Callaway, and A Rosenfeld. Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations. *Proceedings of the ACEEE Summer Study on Buildings, Pacific Grove, CA*, 1000(2000):3000, 2012.

[28] JH Braslavsky, C Perfumo, and JK Ward. Model-based feedback control of distributed air-conditioning loads for fast demand-side ancillary services. In *52nd Conference on Decision and Control*, pages 6274–6279. IEEE, 2013.

[29] N Ruiz, I Cobelo, and J Oyarzabal. A direct load control model for virtual power plant management. *IEEE Transactions on Power Systems*, 24(2):959–966, May 2009.

[30] M Vasirani, R Kota, R Cavalcante, S Ossowski, and NR Jennings. An agent-based approach to virtual power plants of wind power generators and electric vehicles. *IEEE Transactions on Smart Grid*, 4(3):1314–1322, 2013.

[31] JM Griffin and SL Puller. *Electricity deregulation: choices and challenges*, volume 4. University of Chicago Press, 2009.

[32] Mohamed H Albadi and EF El-Saadany. A summary of demand response in electricity markets. *Electric power systems research*, 78(11):1989–1996, 2008.

[33] LA Greening. Demand response resources: Who is responsible for implementation in a deregulated market? *Energy*, 35(4):1518–1525, 2010.

[34] WECC's System Adequacy Planning Department. WECC 2024 common case. www.wecc.biz/SystemAdequacyPlanning/Pages/Datasets.aspx. [Online; accessed 11-Nov-2016].

[35] Energy Information Administration. Electricity supply, disposition, prices and emissions. http://tinyurl.com/qdg2wj2, May 2015.

[36] MS Markoff and AC Cullen. Impact of climate change on Pacific Northwest hydropower. *Climatic Change*, 87(3-4):451–469, 2008.

[37] MK Chandel, LF Pratson, and RB Jackson. The potential impacts of climate-change policy on freshwater use in thermoelectric power generation. *Energy Policy*, 39(10):6234 – 6242, 2011. Sustainability of biofuels.

[38] G Blackburn, C Magee, and V Rai. Solar valuation and the modern utility's expansion into distributed generation. *The Electricity Journal*, 27(1):18–32, 2014.

[39] MJE Alam, KM Muttaqi, and D Sutanto. Mitigation of rooftop solar PV impacts and evening peak support by managing available capacity of distributed energy storage systems. *Transactions on Power Systems*, 28(4):3874–3884, Nov 2013.

[40] PL Joskow. Competitive electricity markets and investment in new generating capacity. *AEI-Brookings Joint Center Working Paper*, (06-14), 2006.

[41] M Hildmann, A Ulbig, and G Andersson. Empirical analysis of the merit-order effect and the missing money problem in power markets with high RES shares. *IEEE Transactions on Power Systems*, 30(3):1560–1570, 2015.

[42] I Dobson, BA Carrera, VE Lynch, and DE Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2):026103, 2007.

[43] P Pinson, H Madsen, et al. Benefits and challenges of electrical demand response: A critical review. *Renewable and Sustainable Energy Reviews*, 39:686–699, 2014.

[44] F Shariatzadeh, P Mandal, and AK Srivastava. Demand response for sustainable energy systems: A review, application and implementation strategy. *Renewable and Sustainable Energy Reviews*, 45:343–350, 2015.

[45] DS Callaway. Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. *Energy Conversion and Management*, 50(5):1389–1400, 2009.

[46] CL Su and D Kirschen. Quantifying the effect of demand response on electricity markets. *IEEE Transactions on Power Systems*, 24(3):1199–1207, 2009.

[47] S Parkinson, D Wang, C Crawford, and N Djilali. Comfort-constrained distributed heat pump management. *Energy Procedia*, 12(0):849 – 855, 2011. The Proceedings of International Conference on Smart Grid and Clean Energy Technologies.

[48] D Wang, S Parkinson, W Miao, H Jia, C Crawford, and N Djilali. Online voltage security assessment considering comfort-constrained demand response control of distributed heat pump systems. *Applied Energy*, 96:104–114, 2012.

[49] S Koch, JL Mathieu, and DS Callaway. Modeling and control of aggregated heterogeneous thermostatically controlled loads for ancillary services. In *Proc. PSCC*, pages 1–7, 2011.

[50] T Broeer, J Fuller, F Tuffner, D Chassin, and N Djilali. Modeling framework and validation of a smart grid and demand response system for wind power integration. *Applied Energy*, 113:199–207, 2014.

[51] FC Schweppe, RD Tabors, JL Kirtley, HR Outhred, FH Pickel, and AJ Cox. Homeostatic utility control. *IEEE Transactions on Power Apparatus and Systems*, PAS-99(3):1151–1163, May 1980.

[52] BA Huberman and SH Clearwater. A multi-agent system for controlling building environments. In *ICMAS*, pages 171–176, 1995.

[53] DJ Hammerstrom, R Ambrosio, J Brous, TA Carlon, DP Chassin, JG DeSteese, RT Guttromson, GR Horst, OM Järvegren, R Kajfasz, et al. Pacific northwest gridwise testbed demonstration projects. *Part I. Olympic Peninsula Project*, 2007.

[54] SE Widergren, K Subbarao, JC Fuller, DP Chassin, A Somani, MC Marinovici, and JL Hammerstrom. AEP Ohio gridSMART demonstration project real-time pricing demonstration analysis. *PNNL Report*, 23192, 2014.

[55] K Subbarao, JC Fuller, K Kalsi, RG Pratt, SE Widergren, and DP Chassin. *Transactive control and coordination of distributed assets for ancillary services.* Pacific Northwest National Laboratory, 2013.

[56] D Wang, S Parkinson, W Miao, H Jia, C Crawford, and N Djilali. Hierarchical market integration of responsive loads as spinning reserve. *Applied Energy*, 104:229–238, 2013.

[57] B Dupont, K Dietrich, C De Jonghe, A Ramos, and R Belmans. Impact of residential demand response on power system operation: A Belgian case study. *Applied Energy*, 122:1–10, 2014.

[58] D Neves, A Pina, and CA Silva. Demand response modeling: A comparison between tools. *Applied Energy*, 146:288–297, 2015.

[59] Q Wang, C Zhang, Y Ding, G Xydis, J Wang, and J Østergaard. Review of real-time electricity markets for integrating distributed energy resources and demand response. *Applied Energy*, 138:695–706, 2015.

[60] HG Kwag and JO Kim. Optimal combined scheduling of generation and demand response with demand resource constraints. *Applied Energy*, 96:161–170, 2012.

[61] P Sreedharan, D Miller, S Price, and CK Woo. Avoided cost estimation and cost-effectiveness of permanent load shifting in California. *Applied Energy*, 96:115–121, 2012.

[62] D Dallinger, S Gerda, and M Wietschel. Integration of intermittent renewable power supply using grid-connected vehicles–a 2030 case study for California and Germany. *Applied Energy*, 104:666–682, 2013.

[63] N Mahmoudi, TK Saha, and M Eghbal. Modelling demand response aggregator behavior in wind power offering strategies. *Applied Energy*, 133:347–355, 2014.

[64] T Rajeev and S Ashok. Dynamic load-shifting program based on a cloud computing framework to support the integration of renewable energy sources. *Applied Energy*, 146:141–149, 2015.

[65] E Heydarian-Forushani, MEH Golshan, M Shafie-khah, and JPS Catalão. Optimal behavior of demand response aggregators in providing balancing and ancillary services in renewable-based power systems. In *Technological Innovation for Cloud-Based Engineering Systems*, pages 309–316. Springer, 2015.

[66] M Fripp. Switch: a planning tool for power systems with large shares of intermittent renewable energy. *Environmental science & technology*, 46(11):6371–6378, 2012.

[67] P Santoro, V Galdi, V Calderaro, and G Gross. Quantification of variable effects of demand response resources on power systems with integrated energy storage and renewable resources. In *Eindhoven PowerTech*, pages 1–6. IEEE, 2015.

[68] M Joung and J Kim. Assessing demand response and smart metering impacts on long-term electricity market prices and system reliability. *Applied Energy*, 101:441–448, 2013.

[69] G Artač, D Flynn, B Kladnik, M Pantoš, AF Gubina, and R Golob. A new method for determining the demand reserve offer function. *Electric Power Systems Research*, 100:55–64, 2013.

[70] J Valenzuela, PR Thimmapuram, and J Kim. Modeling and simulation of consumer response to dynamic pricing with enabled technologies. *Applied Energy*, 96:122–132, 2012.

[71] DS Kirschen, G Strbac, P Cumperayot, and D de Paiva Mendes. Factoring the elasticity of demand in electricity prices. *IEEE Transactions on Power Systems*, 15(2):612–617, 2000.

[72] S Behboodi, DP Chassin, C Crawford, and N Djilali. Electric vehicle participation in transactive power systems using real-time retail prices. In *49th Hawaii International Conference on System Sciences (HICSS)*. IEEE, 2016.

[73] A Faruqui and S Sergici. Household response to dynamic pricing of electricity: a survey of 15 experiments. *Journal of Regulatory Economics*, 38(2):193–225, 2010.

[74] P Barooah, A Bušic, and S Meyn. Spectral decomposition of demand-side flexibility for reliable ancillary services in a smart grid. In *48th Hawaii International Conference on Systems Science (HICSS)*. IEEE, 2015.

[75] DP Chassin, J Stoustrup, P Agathoklis, and N Djilali. A new thermostat for real-time price demand response: Cost, comfort and energy impacts of discrete-time control without deadband. *Applied Energy*, 155:816–825, 2015.

[76] S Stoft. Power system economics. *Journal of Energy Literature*, 8:94–99, 2002.

[77] U Focken, M Lange, K Mönnich, HP Waldl, HG Beyer, and A Luig. Short-term prediction of the aggregated power output of wind farms- a statistical analysis of the reduction of the prediction error by spatial smoothing effects. *Journal of Wind Engineering and Industrial Aerodynamics*, 90(3):231–246, 2002.

[78] S Borenstein, M Jaske, and A Rosenfeld. Dynamic pricing, advanced metering, and demand response in electricity markets. *Center for the Study of Energy Markets*, 2002.

[79] CF Manski. The structure of random utility models. *Theory and Decision*, 8(3):229–254, 1977.

[80] BC Hydro. BC Hydro's greenhouse gas intensities, 2004, and 2009 to 2013. http://tinyurl.com/n2dq2by.

[81] PR O'Donoughue, GA Heath, SL Dolan, and M Vorum. Life cycle greenhouse gas emissions of electricity generated from conventionally produced natural gas. *Journal of Industrial Ecology*, 18(1):125–144, 2014.

[82] BC Hydro. 2013 integrated resource plan, chapter 3: Resource options, appendix 3a-4: 2013 ROR update resource options database summary sheets. http://tinyurl.com/khn3vjt, Nov 2013.

[83] M Bolinger and R Wiser. Understanding wind turbine price trends in the US over the past decade. *Energy Policy*, 42:628–641, 2012.

[84] British Columbia Government. Facility-reported GHG emissions in BC. http://tinyurl.com/nytp9gv.

[85] BC Hydro. Independent Power Producers (IPPs) currently supplying power to BC Hydro and with projects currently in development. http://tinyurl.com/n6gjrd6, May 2015.

[86] C Zhao, J Wang, J Watson, and Y Guan. Multi-stage robust unit commitment considering wind and demand response uncertainties. *IEEE Transactions on Power Systems*, 28(3):2708–2717, 2013.

[87] GA Stern, JH Yan, PB Luh, and WE Blankson. What objective function should be used for optimal auctions in the ISO/RTO electricity market? In *Power Engineering Society General Meeting*, pages 10–pp. IEEE, 2006.

[88] N Samaan, M Milligan, M Hunsaker, and T Guo. Three-stage production cost modeling approach for evaluating the benefits of intra-hour scheduling between balancing authorities. In *Power & Energy Society General Meeting*, pages 1–5. IEEE, 2015.

[89] S Behboodi, DP Chassin, C Crawford, and N Djilali. Renewable resources portfolio optimization in the presence of demand response. *Applied Energy*, 162:139–148, 2016.

[90] DP Chassin, S Behboodi, C Crawford, and N Djilali. Agent-based simulation for interconnection-scale renewable integration and demand response studies. *Engineering*, 1(4):422–435, 2015.

[91] MC Hu, SY Lu, and YH Chen. Stochastic–multiobjective market equilibrium analysis of a demand response program in energy market under uncertainty. *Applied Energy*, 182:500–506, 2016.

[92] C Zhang, Q Wang, J Wang, M Korpås, and ME Khodayar. Strategy-making for a proactive distribution company in the real-time market with demand response. *Applied Energy*, 181:540–548, 2016.

[93] WW Hogan. Contract networks for electric power transmission. *Journal of Regulatory Economics*, 4(3):211–242, 1992.

[94] EPEX Spot. European Power Exchange. www.epexspot.com. [Online; accessed 11-Nov-2016].

[95] S Newell and K Madjarov. Economic evaluation of alternative demand response compensation options. *The Brattle Group*, 2010.

[96] JR Pierce. Primer on demand response and a critique of FERC order 745, a. *Geo Wash. J. Energy & Envtl. L.*, 3:102, 2012.

[97] J Krutilla. The international Columbia river treaty: An economic evaluation. *Water Research, Baltimore: Johns Hopkins*, pages 69–97, 1966.

[98] GB Sheblé. *Computational auction mechanisms for restructured power industry operation*. Springer Science & Business Media, 2012.

[99] P Siano and D Sarno. Assessing the benefits of residential demand response in a real time distribution energy market. *Applied Energy*, 161:533–551, 2016.

[100] JE Price. Benchmarking a reduced test-bed model of WECC region for unit commitment and flexible dispatch. In *Power and Energy Society General Meeting*, pages 1–5. IEEE, 2013.

[101] A Zakariazadeh and S Jadid. Smart microgrid operational planning considering multiple demand response programs. *Journal of Renewable and Sustainable Energy*, 6(1):013134, 2014.

[102] M Mazidi, H Monsef, and P Siano. Robust day-ahead scheduling of smart distribution networks considering demand response programs. *Applied Energy*, 178:929–942, 2016.

[103] MathWorks. Moore-Penrose pseudoinverse of matrix. www.mathworks.com/help/matlab/ref/pinv.html. [Online; accessed 11-Nov-2016].

[104] J Nocedal and S Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[105] MathWorks. fmincon SQP algorithm. `http://tinyurl.com/jycokm9`. [Online; accessed 11-Nov-2016].

[106] S Boyd and L Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[107] WECC's System Adequacy Planning Department. Integrated transmission and resource assessment– summary of 2015 planning analyses, March 2016.

[108] KP Schneider, E Sortomme, SS Venkata, MT Miller, and L Ponder. Evaluating the magnitude and duration of cold load pick-up on residential distribution feeders using multi-state load models. *IEEE Transactions on Power Systems*, 31(5):3765–3774, 2016.

[109] F Klügl. A validation methodology for agent-based simulations. In *Proceedings of the 2008 ACM symposium on applied computing*, pages 39–43. ACM, 2008.

[110] P Windrum, G Fagiolo, and A Moneta. Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, 10(2):8, 2007.

[111] DP Chassin and D Rondeau1. Aggregate modeling of fast-acting demand response and control under real-time pricing. *Applied Energy*, 181:288–298, 2016.

[112] O Homaee, A Zakariazadeh, and S Jadid. Real time voltage control using emergency demand response in distribution system by integrating advanced metering infrastructure. *Journal of Renewable and Sustainable Energy*, 6(3):033145, 2014.

[113] F Rahimi and A Ipakchi. Demand response as a market resource under the smart grid paradigm. *IEEE Transactions on Smart Grid*, 1(1):82–88, 2010.

[114] P Zhang, K Qian, C Zhou, BG Stewart, and DM Hepburn. A methodology for optimization of power systems demand due to electric vehicle charging load. *IEEE Transactions on Power Systems*, 27(3):1628–1636, 2012.

[115] TK Kristoffersen, K Capion, and P Meibom. Optimal charging of electric drive vehicles in a market environment. *Applied Energy*, 88(5):1940–1948, 2011.

[116] A Zakariazadeh and Sh Jadid. Integrated scheduling of electric vehicles and demand response programs in a smart microgrid. *Iranian Journal of Electrical and Electronic Engineering*, 10(2):114–123, 2014.

[117] W Kempton and J Tomić. Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *Journal of power sources*, 144(1):268–279, 2005.

[118] H Liang, BJ Choi, W Zhuang, and X Shen. Optimizing the energy delivery via V2G systems based on stochastic inventory theory. *IEEE Transactions on Smart Grid*, 4(4):2230–2243, 2013.

[119] Z Ma, DS Callaway, and I Hiskens. Decentralized charging control for large populations of plug-in electric vehicles. In *49th Conference on Decision and Control*, pages 206–212. IEEE, 2010.

[120] M Mallette and G Venkataramanan. Financial incentives to encourage demand response participation by plug-in hybrid electric vehicle owners. In *Energy Conversion Congress and Exposition (ECCE)*, pages 4278–4284. IEEE, 2010.

[121] S Shao, M Pipattanasomporn, and S Rahman. Grid integration of electric vehicles and demand response with customer choice. *IEEE Transactions on Smart Grid*, 3(1):543–550, 2012.

[122] Z Liu, D Wang, H Jia, and N Djilali. Power system operation risk analysis considering charging load self-management of plug-in hybrid electric vehicles. *Applied Energy*, 136:662–670, 2014.

[123] Z Liu, D Wang, H Jia, N Djilali, and W Zhang. Aggregation and bidirectional charging power control of plug-in hybrid electric vehicles: Generation system adequacy analysis. *IEEE Transactions on Sustainable Energy*, 6(2):325–335, 2015.

[124] R Li, Q Wu, and SS Oren. Distribution locational marginal pricing for optimal electric vehicle charging management. *IEEE Transactions on Power Systems*, 29(1):203–211, 2014.

[125] SL Mabit and M Fosgerau. Demand for alternative-fuel vehicles when registration taxes are high. *Transportation Research Part D: Transport and Environment*, 16(3):225–231, 2011.

[126] SS Williamson. EV and PHEV battery technologies. In *Energy Management Strategies for Electric and Plug-in Hybrid Electric Vehicles*, pages 65–90. Springer, 2013.

[127] J Axsen, HJ Bailey, and G Kamiya. The Canadian plug-in electric vehicle survey (CPEVS 2013): Anticipating purchase, use, and grid interactions in British Columbia. Technical report, Oct 2013.

[128] W Su, H Eichi, W Zeng, and MY Chow. A survey on the electrification of transportation in a smart grid environment. *IEEE Transactions on Industrial Informatics*, 8(1):1–10, 2012.

[129] S Chu and A Majumdar. Opportunities and challenges for a sustainable energy future. *nature*, 488(7411):294–303, 2012.

[130] RP Batley, JP Toner, and MJ Knight. A mixed logit model of UK household demand for alternative-fuel vehicles. *International Journal of Transport Economics*, pages 55–77, 2004.

[131] J Bailey and J Axsen. Anticipating PEV buyers acceptance of utility controlled charging. *Transportation Research Part A: Policy and Practice*, 82:29–46, 2015.

[132] S Behboodi, DP Chassin, N Djilali, and C Crawford. Interconnection-wide hour-ahead scheduling in the presence of intermittent renewables and demand response: A surplus maximizing approach. *Applied Energy*, 189:336–351, 2017.

[133] G Bianchini, M Casini, A Vicino, and D Zarrilli. Demand-response in building heating systems: A model predictive control approach. *Applied Energy*, 168:159–170, 2016.

[134] B Alimohammadisagvand, J Jokisalo, S Kilpeläinen, M Ali, and K Sirén. Cost-optimal thermal energy storage system for a residential building with heat pump heating and demand response control. *Applied Energy*, 174:275–287, 2016.

[135] R Menke, E Abraham, P Parpas, and I Stoianov. Demonstrating demand response from water distribution system through pump scheduling. *Applied Energy*, 170:377–387, 2016.

[136] R Dhulst, W Labeeuw, B Beusen, S Claessens, G Deconinck, and K Vanthournout. Demand response flexibility and flexibility potential of residential smart appliances: Experiences from large pilot test in belgium. *Applied Energy*, 155:79–90, 2015.

[137] S Koch, M Zima, and G Andersson. Potentials and applications of coordinated groups of thermal household appliances for power system control purposes. In *2009 PES/IAS Conference on Sustainable Alternative Energy*, pages 1–8. IEEE, 2009.

[138] D Patteeuw, K Bruninx, A Arteconi, E Delarue, Dhaeseleer, and L Helsen. Integrated modeling of active demand response with electric heating systems coupled to thermal energy storage systems. *Applied Energy*, 151:306–319, 2015.

[139] D Patteeuw, G Reynders, K Bruninx, C Protopapadaki, E Delarue, W Dhaeseleer, D Saelens, and L Helsen. CO2-abatement cost of residential heat pumps with active

demand response: demand-and supply-side effects. *Applied Energy*, 156:490–501, 2015.

[140] A Arteconi, D Patteeuw, K Bruninx, E Delarue, W Dhaeseleer, and L Helsen. Active demand response with electric heating systems: impact of market penetration. *Applied Energy*, 177:636–648, 2016.

[141] R Yin, EC Kara, Y Li, N DeForest, K Wang, T Yong, and M Stadler. Quantifying flexibility of commercial and residential loads for demand response using setpoint changes. *Applied Energy*, 177:149–164, 2016.

[142] G OBrien. *Scheduling, Revenue Sharing, and User Behavior for Aggregated Demand Response*. PhD thesis, Stanford University, 2014.

[143] MH Albadi and EF El-Saadany. Demand response in electricity markets: An overview. In *Power Engineering Society General Meeting*, pages 1–5. IEEE, 2007.

[144] MA Fotouhi Ghazvini, J Soares, N Horta, R Neves, R Castro, and Z Vale. A multi-objective model for scheduling of short-term incentive-based demand response programs offered by electricity retailers. *Applied Energy*, 151:102–118, 2015.

[145] S Nojavan, K Zare, and B Mohammadi-Ivatloo. Optimal stochastic energy management of retailer based on selling price determination under smart grid environment in the presence of demand response program. *Applied Energy*, 187:449–464, 2017.

[146] JK Kok, CJ Warmer, and IG Kamphuis. Powermatcher: multiagent control in the electricity infrastructure. In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 75–82. ACM, 2005.

[147] S Borenstein. Wealth transfers among large customers from implementing real-time retail electricity pricing. *The Energy Journal*, pages 131–149, 2007.

[148] M Roozbehani, MA Dahleh, and SK Mitter. Volatility of power grids under real-time pricing. *IEEE Transactions on Power Systems*, 27(4):1926–1940, 2012.

[149] S Ihara and FC Schweppe. Physically based modeling of cold load pickup. *IEEE Transactions on Power Apparatus and Systems*, (9):4142–4150, 1981.

[150] S El-Ferik and RP Malhame. Identification of alternating renewal electric load models from energy measurements. *IEEE Transactions on Automatic Control*, 39(6):1184–1196, 1994.

[151] DS Callaway. Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. *Energy Conversion and Management*, 50(5):1389–1400, 2009.

[152] C Perfumo, E Kofman, JH Braslavsky, and JK Ward. Load management: Model-based control of aggregate power for populations of thermostatically controlled loads. *Energy Conversion and Management*, 55:36–48, 2012.

[153] U.S. Department of Energy-Quadrennial Energy Review (QER) Task Force. Transforming the nation's electricity system: The second installment of the quadrennial energy review. Technical report, Jan 2017.

[154] S Lenhart, N Nelson-Marsh, EJ Wilson, and D Solan. Electricity governance and the Western energy imbalance market in the United States: The necessity of interorganizational collaboration. *Energy Research & Social Science*, 19:94–107, 2016.

[155] Greensync. Decentralised Energy Exchange (deX): a new renewable energy digital marketplace that will transform Australia's energy industry. https://tinyurl.com/zg5ajor. [Online; accessed 3-Mar-2017].