

Power System Balancing with High Renewable Penetration:
The Potential of Demand Response

by

D. Karl Critz

M.S. Mechanical Engineering
Massachusetts Institute of Technology, 1997

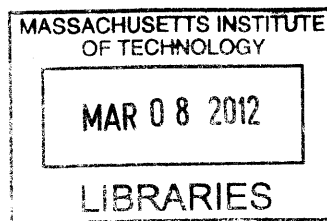
Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of

Master of Science in Management and Engineering
at the
Massachusetts Institute of Technology

September 2011

[February 2012]

©2011 David Karl Critz. All rights reserved.



ARCHIVES

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author: _____

Engineering Systems Division
September 2011

Certified by: _____

Stephen Connors
Director, Analysis Group for Regional Energy Alternatives, MIT Energy Initiative
Thesis Supervisor

Accepted by: _____

Patrick Hale
Director, System Design & Management

Acknowledgements

Thanks to the National Renewable Energy Laboratory for its support through the Joint Institute for Strategic Energy Analysis.

Thanks to K&K. You were always there for me.

Table of Contents

1	Introduction	7
2	Background	8
2.1	Demand Response	8
2.1.1	The Business of Demand Response	9
2.1.2	Load Clipping	10
2.1.3	Load Shifting	12
2.1.4	Price Response	13
2.2	Rolling Stochastic Unit Commitment	15
2.2.1	WILMAR Results	17
2.2.2	WILMAR and Demand Response	17
3	Method	19
3.1	Modeling Assumptions	19
3.2	WILMAR Operations	22
3.3	Modeling Demand Response	23
3.4	Parameter: Demand Response Quantity	23
3.5	Parameter: Demand Response Price	24
4	Results	26
4.1	Baseline: Stochastic Optimization with No Demand Response	26
4.1.1	Evaluation Metrics	27
4.1.2	Wind Harvesting	28
4.1.3	Thermal Unit Performance	31
4.2	Single Result with Demand Response	35
4.3	Demand Response Parametric Search – Price & Quantity	41
4.4	Business Impact	51
4.4.1	Pricing Demand Response Capacity	51
4.4.2	Effect of Demand Response on Electricity Price	54
4.4.3	Effect of Demand Response upon Wind Generator Profit	56
4.4.4	Effect of Demand Response on Coal Generator Profit	57
4.4.5	Effect of Demand Response upon Natural Gas Generator Profit	58
5	Conclusions	59
6	References	61

Table of Figures

Figure 2-1: Illustration of demand response load clipping	11
Figure 2-2: Illustration of demand response load shifting	13
Figure 3-1: Electrical demand and wind, Winter	20
Figure 3-2: Average electrical demand and wind supply by hour, all seasons	21
Figure 3-3: Load-duration curve, 8 selected weeks of 52	21
Figure 4-1: Winter generation profile, stochastic baseline.....	26
Figure 4-2: Wind forecasting and wind curtailment, stochastic optimization	29
Figure 4-3: Consolidated view of wind forecasting and curtailment.....	30
Figure 4-4: Load profile, coal unit, no demand response	32
Figure 4-5: Load profile, multiple coal units, no demand response	33
Figure 4-6: Load profile, multiple natural gas units, no demand response	34
Figure 4-7: Generation mix, with demand response.....	35
Figure 4-8: Demand response calls by time of day, wind and no wind.....	36
Figure 4-9: System cost, with and without demand response.....	37
Figure 4-10: Wind harvesting with demand response.....	38
Figure 4-11: Load profile, coal unit, with demand response	39
Figure 4-12: Load profiles, coal class, with demand response.....	40
Figure 4-13: Load profiles, natural gas class, with demand response	40
Figure 4-14: Demand response experimental space.....	42
Figure 4-15: System operation cost with demand response.....	43
Figure 4-16: Demand response price and customer impact.....	44
Figure 4-17: Demand response impact and system cost.....	45
Figure 4-18: Demand response and wind energy	46
Figure 4-19: Demand response and SO ₂	47
Figure 4-20: Demand response and CO ₂	47
Figure 4-21: Demand response and coal energy fraction	48
Figure 4-22: Demand response and natural gas energy contribution	49
Figure 4-23: Demand response capacity price.....	53
Figure 4-24: System operator's demand curve for DR capacity	54
Figure 4-25: Effect of demand response upon day-ahead electricity prices	55
Figure 4-26: Effect of demand response upon intra-day electricity prices	55
Figure 4-27: Impact of demand response upon wind generator profit and production ..	56
Figure 4-28: Impact of demand response upon coal generator profit and production....	57
Figure 4-29: Impact of demand response upon natural gas generator profit.....	58

Table of Tables

Table 2-1: Rolling Unit Commitment Example	17
Table 3-1: Selected Technical Capabilities of Generators	19
Table 3-2: Technical capabilities of demand response	23
Table 4-1: System operation evaluation criteria, no demand response	27
Table 4-2: Electrical load factors of generators, no demand response	31
Table 4-3: System operation evaluation criteria, with demand response	41
Table 4-4: Survey results, ratepayer compensation for disconnections (Lassen and Jensen 2005).....	50
Table 4-5: System operation evaluation criteria, with demand response	51
Table F-1: System operation evaluation criteria, with perfect information	70

Power System Balancing with High Renewable Penetration: The Potential of Demand Response

by

D. Karl Critz

Submitted to the Engineering Systems Division
on September 7, 2011 in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Management and Engineering

ABSTRACT

This study investigated the ability of responsive demand to stabilize the electrical grid when intermittent renewable resources are present. The WILMAR stochastic unit commitment model was used to represent a version of the Danish electricity and heat system with an enhanced level of wind generation. The study found that demand response reduced the marginal operating cost of the electrical system 3%. Demand response reduced CO₂/SO₂ emissions levels 3% by enabling 11% more generation of wind power. Demand resources representing 25% of nameplate wind power and priced at 150% of a gas turbine's marginal cost were a recommended combination that balanced maximum system improvement at minimal ratepayer impact. The system cost benefits of each study case enabled the calculation of a demand curve representing the system operator's willingness to pay fixed costs for capacity from the pool of operating savings. With demand response, wind generators increased profits, coal plants reduced profits slightly, and natural gas plant profit was cut to almost zero. With high levels of unpredictable renewable resources and limited ability to import power, demand response represents a promising technique to balance the grid at low cost.

Thesis Supervisor: Stephen Connors

Title: Director, Analysis Group for Regional Energy Alternatives, MIT Energy Initiative

Power System Balancing with High Renewable Penetration: The Potential of Demand Response

1 Introduction

The fundamental assumptions of the electric grid have remained mostly unchanged since its inception. An increased emphasis on renewable generation and new technologies are changing the way that utilities, regulators, and ratepayers will have to think about the system. Our Edison-era operating mindset assumes that demand is a given and that generation units can be scheduled ahead of time. Operating a system with significant amounts of unpredictable power from wind, sun, and other renewables will require new assumptions and tools. This study investigates the system benefits of modifying demand for electricity in response to changing wind conditions.

The utility industry developed "demand response" (DR) to help the system remain balanced in case of emergencies or extreme operating conditions. When demand is in danger of exceeding supply, the system operator has the option of reducing demand as well as increasing supply. In regions with significant electric heating, this will happen predictably in the winter months. Integration of renewable resources has the potential to create more times of system unbalance; if the wind does not blow as hard as predicted, the supply could decrease at any time. A traditional but expensive solution to the "intermittency" problem is to keep a pool of responsive generators operating at a low power level, ready to increase production to meet the lost wind power. Alternatively, the system could achieve balance by reducing load. Demand response is a quick-acting and controllable "generator". This study models the ability of demand response to respond to changing renewable power levels and ensure system balance.

A system operator needs to answer two fundamental questions about demand response as a resource: "how much?" and "when/how long?" Each power system will have a certain number of customers ready to reduce load, which can be thought of as a quantity of available demand response. The system operator must also decide when to dispatch the resource. By assigning a price for using the resource, an optimization can determine the ideal times and durations to call upon demand response. These two parameters (quantity and price) will influence system's cost, emissions, and reliability. The goal of the study is to quantify the relationship between demand response impact upon ratepayers and system operations. This can be used as an input to a policy or behavioral economics study to determine the optimal level of demand response.

The primary objective of this paper is to study the ability of demand response to encourage improved wind harvesting through system optimization. This will be investigated by modeling the Danish electrical grid and quantifying its behavior under a set of demand response quantities and prices.

2 Background

This investigation models the power system of Denmark, a nation with high penetration of wind power. The model is built on the principle of "rolling stochastic unit commitment", a system robust to unpredictable inputs. The experiment adds demand response to the model. This section will provide background on both (2.1) demand response and (2.2) rolling stochastic unit commitment.

2.1 Demand Response

The Department of Energy defines demand response as (Department of Energy 2006):

Changes in electric usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized.

Demand response may be managed directly by a utility or it may be administered by a third-party curtailment service provider (CSP). This investigation focuses on second half of the equation, in which ratepayers agree to reduce consumption during a time of high demand or an emergency. This response substitutes for high-cost peaking generation on a timescale of minutes to hours. The paper will not focus on long-term efficiency improvements, which substitute for baseload power and operate on a timescale of days to years.

Demand response enables a regional power system to run more efficiently. Demand resources can react very quickly (10 minutes using automation), have negligible startup costs, and have no minimum operating level. A peaking power plant, on the other hand, requires time and fuel to start up or must be kept operating at a low level for fast response. Without demand response, the optimization must keep units running at an inefficiently low output level or must incur startup costs for an uneconomically short duration of operation. Even without intermittent renewable resources on the grid, demand response stabilizes the system by offering a pool of highly-responsive resources that require little early decision-making. The remaining thermal units can then operate closer to their optimal efficiency levels or stay shut down. With intermittent resources, the value flexible demand response increases even more. The system's reserve and ancillary services requirements climb as the grid prepares for events that amount to frequent unplanned plant shutdowns. Demand response reduces unplanned curtailment of wind resources and enables thermal plants to run more efficiently, reducing costs system-wide.

2.1.1 The Business of Demand Response

Curtailement service providers arrange demand reductions for emergencies or other demand events with Independent System Operators (ISO). CSPs aggregate a pool of medium- to large-sized commercial customers (typically exceeding 200kW of load) willing to curtail and bids this capacity on the ISO markets. Customers receive a payment for the actual curtailed demand and are penalized for failing to curtail. Customers also get a participation payment for remaining ready to reduce demand. The CSP is a middleman between the customer, taking curtailment orders from the ISO and making the necessary changes at the customer site. The fastest responses stem from use of direct load control (DLC) software and hardware. Slower curtailment is handled manually as the provider calls the customer on the phone and requests a load drop within a fixed amount of time. The balance between supply and demand on the grid is then restored without any additional generation capacity.

Event-based curtailment has been used for decades in vertically integrated utilities. Utilities would contract with large customers to offer a lower electricity rate if the customer would agree to reduce load during a contingency, subject to penalties for failure to act. Interest in these interruptible/curtailable tariffs dropped when utilities were unbundled; generators, transmission agencies, and distributors had little incentive to solve systematic problems of availability. Bids of reduced load into supply markets are the restructured incarnation of these tariff programs.

The CSP business model funds itself through a percentage of incentive payments. CSPs such as Energy Curtailment Specialists, EnerNOC, CPower, Innoventive, Energy Spectrum, and Integrys Energy typically shoulder the cost of equipment installation upfront and do not charge customers a fee to participate. CSPs do not typically publish their provider/customer payment split, but it can be inferred from public documents. EnerNOC's annual report (EnerNOC 2009) shows \$191M in revenue and \$104M in cost of revenues. The company admits that cost of revenues consists mostly of payments to customers, though the figure also includes metering infrastructure and monitoring costs. If \$9M of cost of revenues is devoted to this overhead, it implies a typical 50/50 split of benefits between the CSP and its customer. With 3566MW under management, each MW is worth \$50k annually in reserve and incentive payments. A CSP helps customers capture this value and retains a portion.

Service providers also provide price response services. CPower offers customers on a fixed-price contract the ability to register a "strike price" at which the customer is willing to curtail load. Notification can be based on day-ahead or spot markets. The CSP bids this strike price into the market as a supply resource, then notifies the customer to curtail if the bid is accepted. The ISO pays at the market rate for curtailed load and the CSP funds operations from a portion of the payment. This is only available in regions with a "demand bidding" infrastructure.

CSPs target medium and large commercial as well as industrial customers. Each site's power consumption is large as compared to residential customers, so the installation of

equipment and use of staff time is invested into more power per site. A study of different business types (ISO-NE 2005) showed that manufacturing and education/government sites had the greatest elasticity of substitution due to flexibility in operations. Sites with customer-facing operations such as retail and those with mission-critical operations such as health care show the lowest ability to shift load. Few CSPs have attempted to enter the residential market, though some operate in large multi-family buildings. Due to scale of opportunity and investment per curtailable Watt, C&I customers remain the lifeblood of the CSP industry.

Competition for CSPs comes from a variety of sources. Utilities (especially vertically integrated ones) operate their own interruptible tariff programs for large industrial customers. Utilities and advanced metering infrastructure (AMI) companies automate equipment shutdown through direct load control (DLC). CSPs perform the same function and are therefore in direct competition with peaking generation, transmission towers, and other incumbent infrastructure. Simple apathy is also a conceptual competitor; with low power prices customers have little incentive to reduce loads. Curtailment has its place in this ecosystem, but must work to distinguish itself on its unique features.

There are three conceptual modes of demand response: (2.1.2) load clipping, (2.1.3) load shifting, and (2.1.4) price response.

2.1.2 Load Clipping

Load clipping occurs when a load is reduced from the system and is not replaced. If a building manager disables escalators or a storeowner reduces lighting during a demand event, that energy is not used at another time. Some larger customers may activate emergency generators rather than curtail energy use. Faruqui (2009) shows that in 15 limited-scale experiments demand response was able to reduce aggregate consumption by 13-36%. Tests that included some sort of automatic monitoring and shutoff technology reported the strongest results. King (2005) studied incentive-based reliability programs and concluded that they reduce total consumption by 0.2% on average, varying between a 5% increase and a 20% decrease.

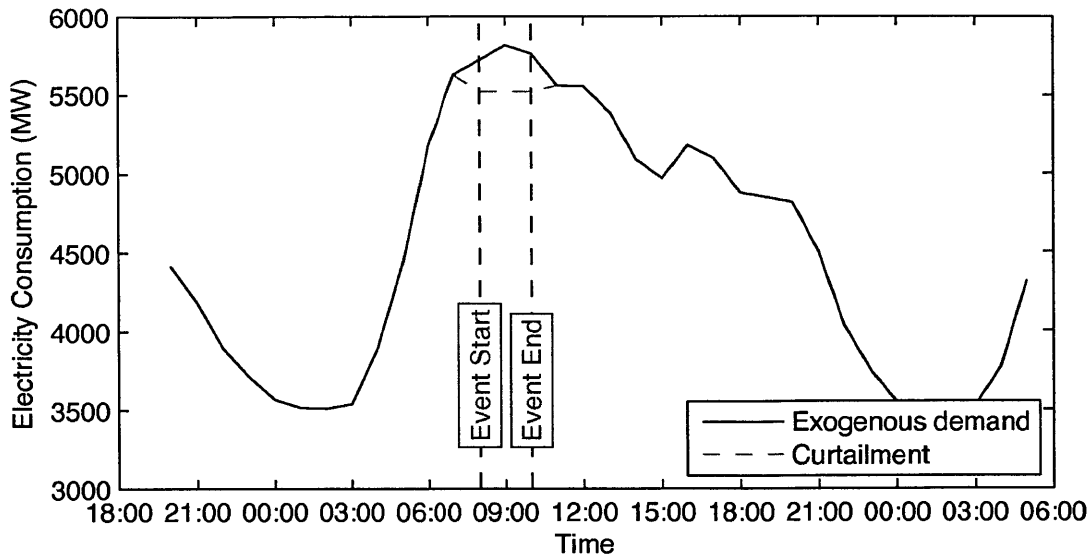


Figure 2-1: Illustration of demand response load clipping

Load clipping can be modeled as a fast-start generator. End-use customers receive an incentive payment when they are called to reduce demand. This payment is considered part of the marginal cost of running the resource, rather like fuel cost for a natural gas turbine.

Demand response is unlike a fully-dispatchable generator in that its response is not entirely predictable. A good curtailment services provider will manage response actively and deliver a load reduction very close to that requested by the ISO. It is not unusual to see response rates within 1% of the target (FERC 2009). ISOs accustomed to dealing with near-perfect dispatchability of resources are still becoming comfortable with a statistical resource such as DR.

The second aspect that differentiates demand response from a gas turbine is that its duration is limited. Once activated, a thermal generator can theoretically run until its fuel is exhausted. Demand resources are limited in the length of time that they can be used. End-use customers are limited in their willingness and ability to defer or curtail their energy use. Eventually, buildings will become uncomfortable and need to be cooled again. The water treatment plant will have to resume pumping.

The issues of less-predictable capacity and limited duration mean that ISOs impose varying rules about the participation of demand response:

- NYISO operates the Special Case Resources program (NYISO 2010), which requires a 4-hour 100kW reduction on 2 hours of notification. NYISO also maintains an Emergency Demand Response Program (NYISO 2008) in which reductions are voluntary in exchange for incentive payments. There is no recurring capacity payment since reductions are not considered firm.

- ISO New England's Real Time Demand and Profiled Response program requires 30-minute notification and curtailment duration of 2 hours. Incentive payments are the real time price with a \$.50/kWh floor . Capacity payments follow the ICAP market price (ISO NE 2005). A typical capacity payment is \$14/kW-month from the Winter Supplemental Program. (ISO NE 2009)
- Demand resources compete equally in PJM's forward capacity market and Synchronized Reserves Market. The frequency regulation market is also open to demand response, though rules prevent it from representing more than 25% of the available capacity in certain regions. There were 4.6GW of capacity available in 2008-2009. (PJM 2010)
- ERCOT's "Load acting as a Resource" (LaaR) program was one of the first to recognize demand response as an ancillary service. In 2008 the program registered 139 sites. Though rated capacity is 2016MW, planned participation is capped at 1150MW to recognize variable response rates. Participation rules were changed in 2007 to require certification and testing, which has improved performance. (Wattles 2008)
- California ISO (California ISO 2010) is drafting regulations to allow demand resource participation in ancillary services markets. These proposed rules relax requirements specifically for demand. Examples include reducing the continuous energy requirement, clarifying timing definitions, and halving the minimum capacity requirement.

Academically, it is practical to model load clipping demand response as a generator. Real-world operations are converging toward this view, but unpredictable response and limited duration are causing system operators to move cautiously.

2.1.3 Load Shifting

Load shifting occurs when a load is removed from the system during a demand event and is added at another time. A hotel may do its laundry at off-peak hours. A water utility may halt pumping and rely on reservoirs for a time. A water heater or air conditioner may be shut off during a time of high demand, then re-activate afterwards and use extra energy to restore its thermal state. Load shifting must not be entirely post-facto; with advance notice a building manager may expend extra energy before an event to pre-cool the air. Borenstein (2005) constructs a set of models that predict total energy consumption actually rises 2% with strong demand response. Holland (2005) similarly reports that a pilot demand response program in PJM increased average load. With load shifting, the "rebound effect" increases in off-peak demand potentially more than offset savings at peak.

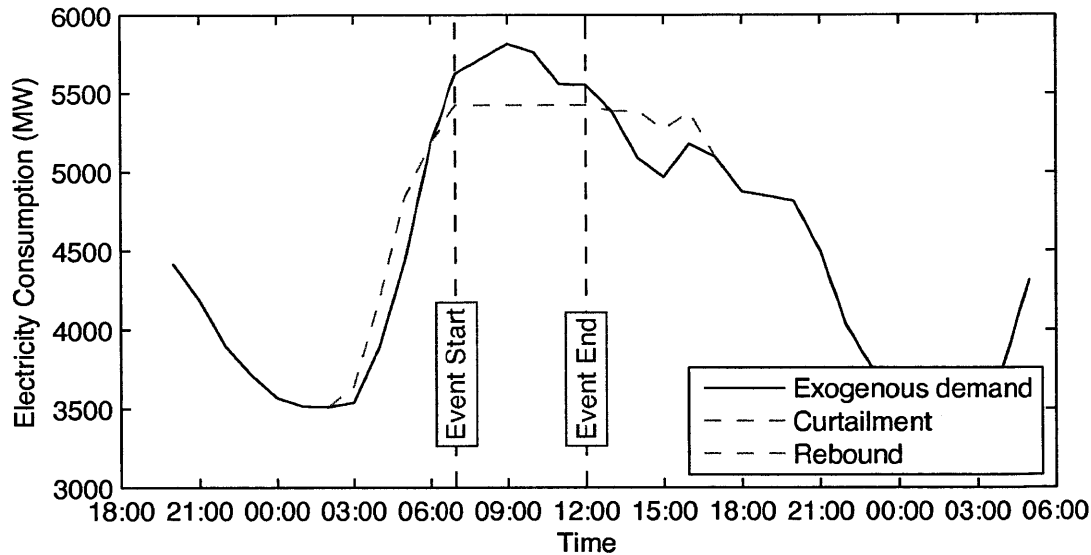


Figure 2-2: Illustration of demand response load shifting

Load shifting can be modeled as a form of storage. Curtailment "discharges" the available storage. Increased usage outside the event call will "charge" the storage. Livengood (2010) models responsive demand as storage, including constraints for temporal proximity of charging and discharging. Andersen et al (2006) model Danish demand response as 100% load shifting with the "Balmorel" unit commitment model. The virtual "storage" discharges during the day and charges at night, lowering and raising prices respectively. With no price associated with dispatching the resource, the authors find that it is either charging or discharging almost all hours of the week.

Treating demand response as storage implicitly acknowledges the duration constraint of load as a resource. If customers can curtail 100MW for 2 hours, they represent 200MWh of storage. The storage can be called at 50MW for 4 hours (or any other combination), but must then wait until the deferred load is consumed before the resource can be called again.

Due to the strength of the research suggesting rebound effects and the simplicity with which a storage model addresses duration constraints, this investigation primarily treats demand response as a load-shifting resource.

2.1.4 Price Response

Price response uses dynamic pricing and elastic demand to elicit a combination of load clipping and load shifting. Rather than explicitly dispatching demand resources from an ISO, ratepayers alter their consumption patterns as prices change. Electric customers may be exposed to a real-time price (PNNL 2007) or a day-ahead price (EDF 2009). Experiments have established parameters for elasticity of demand under these pricing conditions (Faruqui and Sergici 2009). Market pricing internalizes many of the activities

of a Curtailment Service Provider and provides many of the same results. Real-time pricing can have direct influence on wind integration. Sioshansi and Short (2009) find that price signals can smooth load and balance the system in response to spikes in wind levels.

Customers on flat-rate plans are currently charged according to the average cost of use in a given month. A user who runs an air conditioner during peak times is consuming an expensive resource below cost. The user who runs a laundry dryer at night is paying above-market prices for electricity consumed. Dynamic prices seek to end cross-subsidies and incentivize virtuous behavior. This idea is powerful and simple, but requires a new set of regulations, technology, and assumptions.

Dynamic pricing has 3 different forms, but this discussion will predominantly consider “real-time pricing” (RTP).

- Time of use pricing does not gather actual market data but rather charges users based on historical average prices at a certain type of day. This has simpler metering requirements since 2-way communication is not required between the retailer and customer. Historical trends rarely reflect actual market need, especially in a contingency event. As time of use pricing does not reflect actual market conditions, it is receiving little academic or practical interest.
- Critical peak pricing imposes a higher tariff during times of need, be it a standard capacity issue or a contingency. The time scale may use the day-ahead or real time market prices to trigger the event. This model has been implemented in France with a simple day-ahead indicator on the meter. Meters require 2-way communication, but can use simple data transmission.
- Real time pricing exposes the customer to the actual spot price of electricity at any given time. This requires complicated meters with advanced display capabilities.

The potential peak reduction achievable with dynamic pricing dwarfs that of curtailment. FERC (2009) estimates that by 2019, pricing-based demand response could represent almost 150GW in peak load reduction in the United States. This compares to a 50GW potential for all other forms of demand response combined.

The role of customer price responsiveness is a key factor to the success of dynamic pricing. The differentiator of many advanced metering infrastructure (AMI) or “smart meter” designs is the manner in which customers are informed of price swings. The simplest meters may not communicate at all; the customer is left to guess at usage. More advanced meters will show spot and day-ahead rates so that the customer can plan activities and avoid peak rates. However, peak hours often occur in the afternoon when residential customers are at work and unavailable to manage appliances. Even corporate customers may not have staff assigned full-time to watch and respond to rate changes. The most advanced designs communicate with on-site equipment to reduce or cease operations in response to customer-programmed price signals. In all cases, responsiveness depends on an educated customer base that understands price fluctuations and is willing to spend the time and attention to respond to them.

Real-time pricing faces a number of barriers to implementation.

- Advanced Metering Infrastructure deployment is problematic from a technical and economic level. Despite multiple international demonstration sites, there is no agreement on the desired capabilities of a smart meter. The costs of the meters will be great and a debate exists over whether those costs should be included in rates, amortized over time, or expensed in the year of deployment (FERC 2006).
- Legislation exists in California, New York, and the United Kingdom, preventing a “dynamic by default” policy. Default settings have strong influence over consumer behavior; research shows that fewer than 20% of consumers would switch to dynamic pricing if given the option but that only 20% would switch away if defaulted in.
- Faruqi and Sergici (2009) suggest that these laws only protect consumers from dynamic generation charges and that retailers may legally charge realtime prices for transmission and distribution. This seems too clever and narrow a reading, one likely to be plugged by lawmakers just as fast as it can be attempted.
- Even if customers are placed on dynamic rates, state regulators may place caps on maximum tariffs and limit the ability of the market to respond. New York and Delaware already have such restrictions and others are likely to follow as dynamic pricing spreads.

Real-time pricing has the potential to change the electrical industry and the operations of CSPs, but technical and legal factors will continue to slow adoption. Ratepayers (especially residential customers) are unfamiliar with thinking about electricity as having a variable price. Achieving desired levels of elasticity might require smart appliances or other technology to automate consumer preferences. Half of all Danish electricity consumption uses interval meters with the option of purchasing power at realtime rates. However, few customers choose this option (Andersen, et al. 2006). Given longer time horizons required for broad implementation of dynamic pricing, this investigation will not model elastic demand.

2.2 Rolling Stochastic Unit Commitment

Stochastic rolling commitment represents an evolution of traditional unit commitment techniques. It increases the frequency of decision making in order to modify decisions as information quality improves. It also evaluates a range of possibilities and chooses the best generation mix that optimizes across all eventualities, rather than optimizing the most likely scenario or even the expected value of all scenarios. This technique is suited to environments with uncertainty, such as systems with high intermittent generation capacity.

Traditional unit commitment strategies deploy generators according to their marginal cost and constraints related to ramping rates as well as minimum/maximum operation levels. There is little uncertainty in these models. Small variations in demand are handled through up/down regulation. Large loss of supply through the sudden unavailability of transmission or generation capacity is handled through N-1 contingency planning. These decisions are made once per day with some intra-day fine-

tuning. The same process can be conducted at a vertically integrated utility in a liberalized market; with complex bids the main difference between the two is the source of the variables used for the optimization. Vertically integrated utilities obtain the parameters through audits and analysis, while in a market the parameters are decided by the generators in a bid submission. This system has proven effective for low-uncertainty environments with good demand forecasting and predictable thermal generation.

The Risø DTU and University of Stuttgart developed a rolling stochastic unit commitment model called "Wind power Integration in Liberalized electricity MARKets" (WILMAR) (Barth, et al. 2006). This model includes the entire Nordic combined heat and power system, but is extensible to represent other grids. It consists of two modules:

- The Scenario Tree Tool generates forecasts of statistical events such as wind, sun, load, generation/transmission contingencies, forced outages, and hydro filling. It also calculates reserve needs based upon these variables. These are based upon Monte Carlo simulations from historical data. Wind data also uses spatial correlation between towers. (Söder 2004). This is based on an Auto Regressive Moving Average time series model and assumes that forecast error is Gaussian (Lange and Focken 2006). Note that demand and wind predictions are assumed to be independent, but replacement reserve depends upon the 90th percentile error of both. The Scenario Tree Tool generates 1000 paths at each decision period, then compresses them to six weighted representations of the most likely and similar outcomes for efficiency's sake.
- The Scheduling Model uses GAMS to optimize the system for lowest cost according to the scenarios and a description of the system's transmission/generator characteristics. These parameters include fuel costs, operation & maintenance costs, startup costs, maximum/minimum power, ramping rate, efficiency curves (piecewise linear), lead-times. The model is solved each hour. Every decision period solves for the generation mix that will produce the best outcome without knowing which probabilistic scenario will be used in the future. In addition to scheduling generator output, the model also respects the need for up/down regulation, contingency reserves, and replacement reserves.

Rolling unit commitment involves frequent changes to the daily unit commitment plan. Forecasts of wind, cloud cover, and demand improve throughout the day so the decisions can be updated to reflect greater certainty. As long as a unit has not taken a binding start-up or shutdown action at the time of re-evaluation, it can be adjusted during the rolling interval. Because some units have slow startup ramps or long lead times, it is necessary to make decisions well ahead of when these units will actually be needed. The frequency of the decision cycle is variable in the tool, but most studies use a 3-hour interval for 8 decisions per day. Each decision cycle receives a tree in two stages (12 and 36 hours) with six terminal branches. Actions taken in the immediate decision cycle are treated deterministically based upon actual values, then optimized across all future scenarios. Forecast hours in the immediate 3-hour period are handled by replacement reserves.

For example, a timeline of the rolling decision process could resemble the following:

Day	Time	Name	Action
1	12:00	Loop 1	Optimize across 6 scenarios covering the next 36 hours, until 23:59 on day 2. Commit to day-ahead schedule.
	15:00	Loop 2	Optimize across 6 improved scenarios covering the next 33 hours, until 23:59 on day 2. Adjust schedule from Loop 1 where no hard startup/shutdown states exist.
	18:00	Loop 3	Optimize across 6 improved scenarios covering the next 30 hours, until 23:59 on day 2. Adjust schedule from Loop 2 where possible.
1-2	21:00-09:00	Loops 4-8	Continue to optimize scenarios ending on Day 2 23:59. Adjust where possible.
2	12:00	Loop 9	Optimize 6 scenarios covering the next 36 hours, until 23:59 on day 3. Commit to day-ahead schedule
2-3	15:00-09:00	Loops 10-16	Optimize until 23:59 on day 4, adjust where possible.
3	12:00	Loop 17	Optimize until 23:59 on day 5, commit day-ahead.

Table 2-1: Rolling Unit Commitment Example

2.2.1 WILMAR Results

Ela et al. (2011) modeled the entire United States eastern interconnect in WILMAR. The results were compared to the traditionally-dispatched Eastern Wind Integration and Transmission Study (Enernex 2010). The authors found that rolling unit commitment could save 2.4% cost and stochastic unit commitment could save 0.95%. This represents a potential \$1,502M and \$598M benefit, respectively. Tuohy et al. (2009) similarly quantify a 0.9% improvement in costs due to stochastic unit commitment in Ireland.

Note that system cost savings could be even greater than the amounts quoted in these studies. The cost savings in the above studies occur at the hours-to-weeks unit commitment time scale. By running the grid more efficiently, it is possible that longer-term savings may enable deferral of capacity expansion in generation or transmission.

Morales Gonzales (2010) applied WILMAR to Spain. Due to the Iberian peninsula's weak ties to the rest of the European grid, it is almost an island system. The paper concludes that a stochastically-based short-term trading system for wind producers achieves a great decrease in profit variability for a relatively small decrease in profit.

2.2.2 WILMAR and Demand Response

Keane et al. (2011) used WILMAR to model demand response in Ireland. The team investigated demand response through three different mechanisms. Dynamic pricing

relies upon demand elasticity to curtail in response to price signals. Load shifting considers demand response to be a form of "storage" by increasing consumption before and after an event while reducing consumption during an event (40 €/MWh). Load clipping reduces consumption during the event (80 €/MWh). The test case most relevant to this investigation is "Case 1", which includes load shifting and clipping units but no price response. The simulation dispatched demand response mostly as reserve due to its flexibility. This case reduced the number of hours in which spinning and replacement reserve requirements were not met by 50%. Demand response maintained wind production, slightly increased coal contribution, and cut into base load gas.

The case was run in both stochastic and deterministic mode. Stochastic optimization lead to increased energy production from mid merit gas plants for both demand response scenarios. This is attributed to uncertainty in wind output. The presence of demand response lead to dramatically lower capacity factors at peaking units, though the stochastic optimization still used them to counter variable wind conditions.

3 Method

This investigation customized the WILMAR rolling stochastic unit commitment model by adding demand response. It searched a parameter space based upon the quantity and price of available demand response resources. All cases studied an isolated version of the Danish heat/electricity system with enhanced wind capability.

3.1 Modeling Assumptions

The simulation models the Danish heat and electrical system over 8 weeks from 4 seasons. The primary inputs to the model are demand (heat and electricity), wind (forecast and actual), and the technical characteristics of the generation pool.

All simulation runs use the same assumptions regarding the generator mix in the Danish electrical system. Most thermal generation assets are combined heat and power (CHP) units, able to adjust their generation mix to provide heating or electricity. A few are heat-only units. District heat supplies 61% of Danish homes, and 80% of the heat comes from CHP plants (Danish Energy Agency 2010). Note that even without any wind generation, the system is capable of handling peak demand.

Fuel	Generator Count	Nameplate Capacity¹ (total, MW)	Lead Time (mean, hr)	Min Load Factor (mean)	Min Operation (mean, hr)
Natural gas	14	2611	1	0.5	1
Coal	12	4358	4	0.4	4
Fuel oil	19	903	1	0.2	1
Light oil	1	196	1	0.2	1
Muni waste	5	300	12	0.75	4
Straw	7	297	4	0.4	3.57
Wood	2	317	4	0.4	4
Wind	2	4100	--	--	--
<i>Total</i>	<i>62</i>	<i>13082</i>			

Table 3-1: Selected Technical Capabilities of Generators

The energy mix and load factors of each generator are an output of the simulation. Each are different depending upon the amount of demand response available. See Table 2-1 for typical load factors by fuel type.

All simulations used 1200MW of nameplate wind capacity in the east and 2900MW of nameplate wind capacity in the west for a total of 4100MW. In 2009, wind power accounted for 18.3% of electricity supply (Danish Energy Agency 2010). This is

¹ Electricity only. Heat is not reflected in this table.

represented by 2,821 MW nameplate capacity of onshore turbines and 661 offshore for a total of 3482MW. The increase in simulated capacity is intended to represent a 20% wind scenario.

The study simulates 8 weeks of system operation with their attendant wind and demand levels. These 8 weeks are a concatenation of 2 winter, 2 spring, 2 summer, and 2 autumn weeks. The following chart shows demand during the peak winter months. See Figure A, Figure A-, and Figure A- for spring, summer, and autumn behavior. Also displayed is net demand, which represents demand minus available wind.

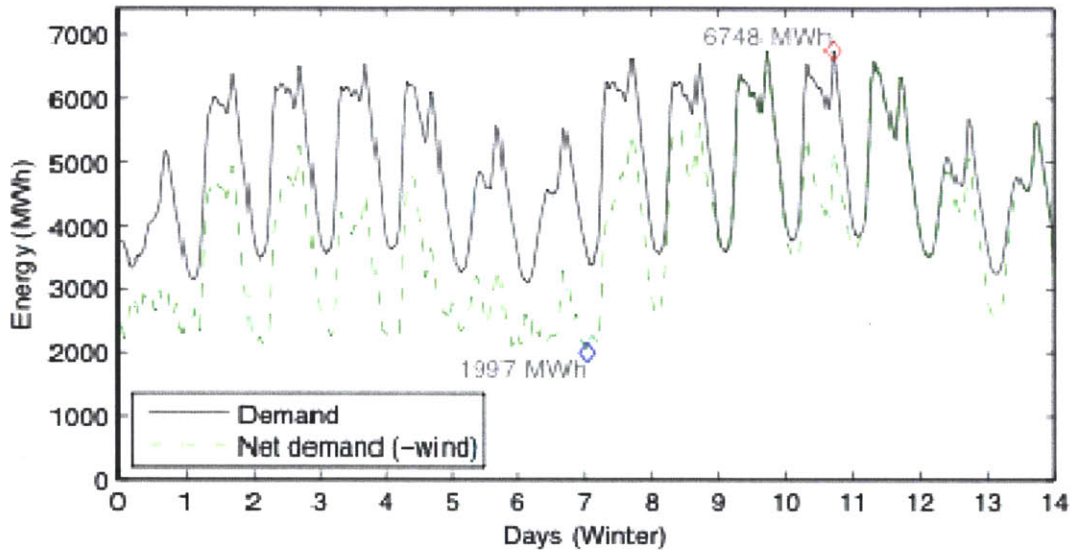


Figure 3-1: Electrical demand and wind, Winter

Demand peaks in the afternoon and evening. Wind is relatively constant, though there is a mid-day maximum. The following chart shows average energy intensity across the entire 8-week study period broken down by hour. Error bars represent one standard deviation of variation.

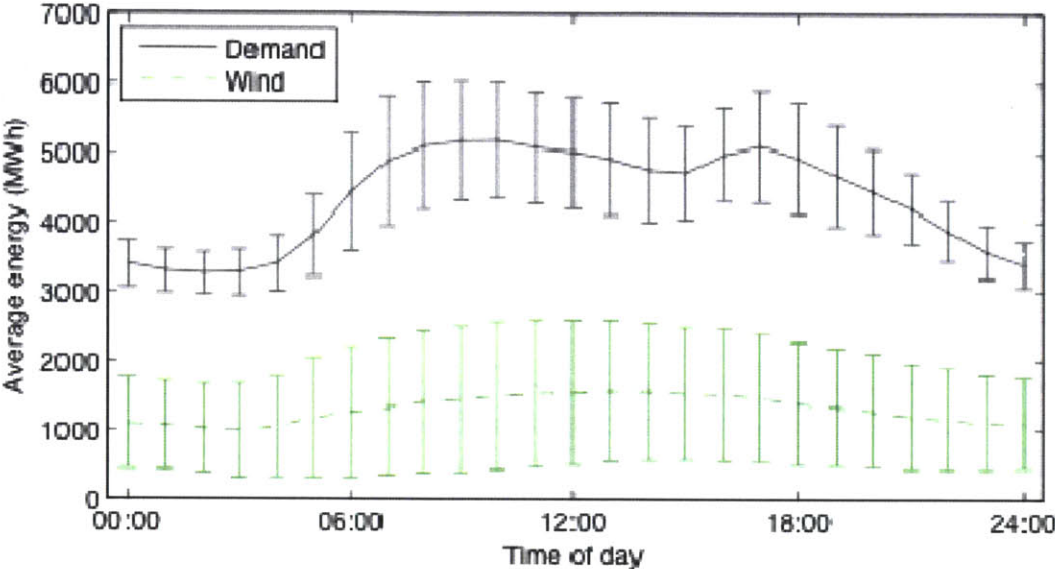


Figure 3-2: Average electrical demand and wind supply by hour, all seasons

Sorting each hour's energy use (absolute and net) produces a load-duration curve. The shape shows how much time the system spends at peak and minimum demand.

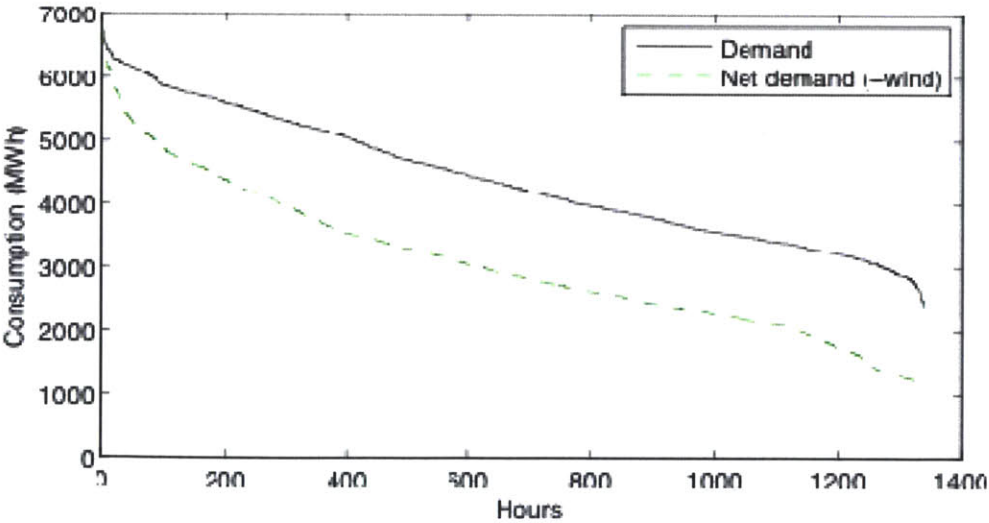


Figure 3-3: Load-duration curve, 8 selected weeks of 52

The load-duration curve for the 8 week study period is relatively flat. The system is above 90% of peak load 4.5% of the time. California, for example, is above 90% of peak load for only 57 hours per year - 0.65%(California ISO 2007). After subtracting wind generation, the net load-duration curve shows more peak behavior requiring units to be active for short amounts of time. Noteworthy is that there are some periods of very low

net load (1100MW) below the combined minimum operation level of all coal generators (1743MW = 4358*0.4, from Table 3-1). This implies that either some coal must be shut down when the wind is blowing at night or the wind must be curtailed. Note also that the net load-duration curve will not be identical for all subsequent cases. Realized wind power will vary each time the simulation is run.

All simulations used the same heat demand. Heat demand is highest in Winter and nearly nonexistent in Summer. Heat is relevant to the optimization because of the prevalence of CHP units.

The simulation breaks demand and supply into regions and monitors electrical flows between them. Transmission capacity is limited between regions, though it is rare for transmission to be the constraining factor in the simulations. The transmission system does incur losses when used. All data is presented here in aggregate for both regions.

To prepare for a follow-up study of island electric grids, all imports and exports of energy were ignored. This alters the nature of the Danish electrical system, which typically imports power from Norway or Germany when wind drops and exports power at night.

3.2 WILMAR Operations

All simulations used a rolling commitment window of 3 hours. (Tuohy, et al. 2009) studied the influence of window frequency on the Irish electrical system with high wind penetration. As the decision window lengthens from 3 to 6 hours, replacement reserve requirements and associated costs increase. As the decision window shortens from 3 hours to 1, costs increase for up- and down-regulation. The investigation concluded that a stochastic optimization delivers lowest-cost results with a 3-hour window.

Most simulations were run in "stochastic" mode. Stochastic simulations optimize scheduling across a portfolio of 12 scenarios for each 36-hour decision window. The resulting decisions are therefore almost guaranteed not to be optimal for the resulting realized conditions. For comparison's sake, a few simulations were run in "deterministic" mode. Deterministic simulations assume perfect forecasting and therefore represents an ideal for system operation.

Wind curtailment was allowed in all simulations. WILMAR has an option that forces wind generators to bid their expected power output into the day-ahead market. This option is used if maximum wind production is an explicit target of system operation. The simulations allowed wind generators to bid less than their maximum.

There is no dynamic pricing or elastic demand. Consistent with Andersen et al. (2006), few customers in Denmark use dynamic pricing plans. Note that WILMAR has the capacity to simulate with dynamic pricing, so this could be a variable in a follow-up study.

3.3 Modeling Demand Response

This study models demand response as a load shifting generation resource. For the purpose of simulation, the actual demand curve remains unaltered. The net demand curve after demand resource calls and rebound effects represents the realized consumption curve. This allows for a comparison of exogenous demand with demand response contributions.

Demand response is a flexible resource. It has zero startup/shutdown cost and zero minimum operation duration. It can run 100% efficiently at any load factor. If the simulation operated on a timescale of less than 1 hour, demand response would have some ramping constraint as resources came online. With the long timescale, demand response is considered to have an unconstrained ramp rate.

Attribute	Units	Fast-DR	Slow-DR	Natural Gas Turbine ²
Power capacity	MW	205-4100	205-4100	391
Minimum load	MW	0	0	196
Lead time	hours	0	3	1
Shutdown time	hours	0	0	1
Min operation	hours	0	0	1
Max operation	hours	2 ³	2	N/A
Spinning reserve?	-	Yes	No	Yes
Fuel cost	€/MWh	28.67-71.67	22.93-57.34	27.00
O&M Cost	€/MWh			1.67
Startup Cost	€/MWh	0	0	5

Table 3-2: Technical capabilities of demand response

The main technical constraint of demand response is its lead-time. The pool is divided equally into "fast-DR" and "slow-DR" to reflect differences in lead-time. Slow-DR has a 3 hour lead-time. Real world fast-DR has a 10-minute lead-time and is modeled here as 0 hours. Fast-DR is permitted to operate as spinning reserve, while slow-DR is not.

In this simulation, there is no demand response load clipping. Consistent with Borenstein (2005), the rebound effect demands more energy than was originally deferred. The "storage" is assumed to "charge" with an efficiency of 95%. See "Appendix E: Demand Response Modeled as Load Clipping" for a comparison of load shifting and load clipping in the WILMAR model.

3.4 Parameter: Demand Response Quantity

² Unit identifier is "DK W Trefor ST-E9-NGn"

³ Each MW of capacity can operate for a maximum of 2 hours.

The study varies levels of demand response quantity as an input to each simulation. Demand response quantity can be thought of the number of customers who are committed to reduce load when asked to do so. The baseline simulation has zero demand response capability. All other cases vary the amount of available demand response resources to investigate the effect upon grid operations and customers.

Demand response power is presented here as a normalized ratio relative to the nameplate capacity of all wind generation. The quantities explored range from 10% of all wind power (410MW) 200% of all wind power (8200MW). As peak demand is less than 7000MW, the maximum DR case is a thought experiment representing full participation and fully-curtailable load. True response capability will be a function of time and date. Real-world summertime curtailment would be lower due to the lack of electric heating. Real-world nighttime curtailment would be lower since many energy-drawing operations are not happening. (Lassen and Jensen 2005) surveyed Danish ratepayers and determined that residential weekend curtailment is 2-3x greater than weekday load reduction. These details are ignored for the purpose of the study.

Demand response also has an energy component relating to the duration of curtailment. All resources are assumed to have a maximum duration of 2 hours, giving an energy capacity (MWh) double the maximum power capacity.

Each demand response pool is divided equally into slow-DR and fast-DR resources.

The smaller demand response quantities are achievable. A Danish demonstration project with electric heating (Kofod and Trogeby 2004) showed that it was possible to obtain demand response of up to 5 kW per involved household. There are 125,000 households with electric heating in Denmark, implying 625MW of full-participation wintertime capacity from residential use alone. A small demonstration project (Nordel 2004) signed up 31MW of industrial users with generators in a single factory park. Multiplying this by 20 similarly-sized parks results in 620MW of industrial demand response. This 1GW of demand resources alone represents the 25% of wind case. The achievability of the higher cases may require further study and new technologies.

3.5 Parameter: Demand Response Price

The study investigates 3 levels of demand response price: low, medium, and high. This price is not necessarily related to the cost of operating the demand response resource or the incentive payments granted to subscribers. The price is an input to the optimization function. It represents the bid price of a unit of demand energy (€/MWh). The theoretical CSP will be paid the (day-ahead or intra-day) market price of the actual energy curtailed at the time it is delivered.

Walawaker et al. (2008) found that the optimal range for demand response payments in the PJM area was \$67-77 \$/MWh. This value is system-dependent. Demand response

price here is represented as a normalized quantity relative to the marginal cost of the most expensive peaking unit on the system.

The optimization considers only the marginal price of activating demand response. There is also an annual fixed cost to maintaining a demand response pool. These capacity payments are usually greater than the actual energy payments. Since capacity payments do not influence actual unit commitment decisions, they can be ignored by the optimization. Selecting the optimal quantity of DR will depend on the fixed costs. The relationship between optimal quantity and capacity costs will be discussed in the conclusion.

Demand response is not a uniform resource. Demand that responds very quickly (<10 minutes) to system interruptions is priced at a premium to responses with more lead time. Longer durations of response also command higher prices; a customer may not mind a 30-minute interruption but a 6-hour loss can be invasive. As mentioned in section 3.4 Parameter: Demand Response Quantity, duration is assumed to be the same for all generators. The pool is divided equally into "fast-DR" and "slow-DR" to reflect differences in lead-time. Slow-DR has a 3 hour lead-time and uses the reported price of demand response. Real world fast-DR has a 10-minute lead-time and is modeled here as 0 hours. Fast-DR commands a 20% premium on the price of slow-DR.

Note that the price of demand response is not fully independent of quantity in the real world. There is a distribution of ratepayer willingness to curtail in response to varying demand levels. A small desired quantity of demand response may be achievable at relatively low prices by targeting ratepayers with flexible energy needs. The last customer to participate may only be willing to do so in exchange for large incentive payments. (Lassen and Jensen 2005) show that the bottom 8% of ratepayers require incentive payments that are 60% lower than the mean incentive for a given curtailment profile. The marginal cost of demand response will always increase the more is needed.

The benefits of demand reduction do not require that all customers participate. By lowering prices for everyone, curtailers cause system benefits greater than their own incentive payments (if on a fixed-rate plan) or savings from deferred cost (if on dynamic pricing). This raises the possibility of a free rider problem; a customer on a flat rate plan may enjoy lower fixed generation charges due to reduced system costs while not paying premiums for peak use. The converse argument is that some participants consider incentive payments an inefficient subsidy. Richard Tempchin of the Edison Electric Institute declared (FERC 2006) that

[A]ny payment to a customer for demand reduction should never exceed the wholesale price minus the retail price that the customers would have otherwise paid to own the power. Any payment above this level would be a subsidy, that is, a nonmarket payment that has to be recovered through a tax or charge on all customers.

In a market where so many customers are shielded from the true cost of their consumption, cross-subsidies are already inherent in the system and may need to be

corrected with other targeted payments. There is no one correct model for incentives and pricing for demand response.

4 Results

The experiment simulated the Danish heat and power system over 8 weeks representing 4 seasons. The output of the experiments was a characterization of system operation with varying quantity and price of demand response as compared to the baseline scenario of no demand response. With these results, the investigation aimed to describe the tradeoffs between demand response's impact on ratepayers and its ability to improve electrical system operations.

4.1 Baseline: Stochastic Optimization with No Demand Response

The stochastic optimization with no demand response represents the control group. All demand response experiments will compare to this result. This section details the behavior of the stochastic optimization.

The output of the simulation is a time series representing each hour. Some data such as demand, electricity price, heat price, or transmission quantity are given on a per-region basis. Each generator unit has an electricity output, heat output, fuel usage. The following chart aggregates electricity generation by fuel type and shows two weeks of winter output. Generation profiles for the Spring, Summer, and Autumn weeks are available in Figure B-1, Figure B-2, and Figure B-3 respectively.

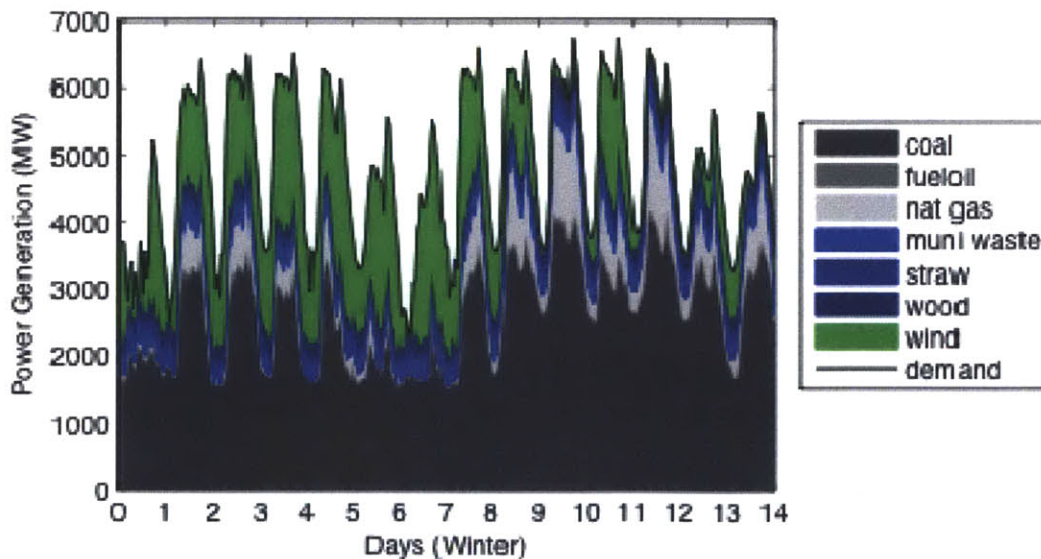


Figure 4-1: Winter generation profile, stochastic baseline

Note the prevalence of coal power (darkest grey). Wind is steady and strong at the beginning of the time period but there are two days at the end with almost no wind at all. Natural gas fills the void left by decreased wind power. Biofuels and municipal waste

remain constant throughout due to long lead times and extended minimum operation times.

4.1.1 Evaluation Metrics

The performance of the system was evaluated based upon operating cost, emissions levels, generation mix, and ratepayer impact of demand response.

Criterion	Unit	Value (No DR)
Cost	€/MWh	27.97
CO ₂ emissions	kg/MWh	166.2
SO ₂ emissions	kg/MWh	1.09
Wind energy	fraction	.28
Coal energy	fraction	.55
Natural gas energy	fraction	.05
Demand response impact	MWh/day	N/A

Table 4-1: System operation evaluation criteria, no demand response

- **Cost:** System cost represents the objective function of the optimization. It represents fuel cost, operation & maintenance cost, startup/shutdown cost, transmission losses, taxes, and subsidies. Systems have costs on longer timescales such as forward capacity markets and transmission planning; these are not represented in the optimization. Cost varies over time, so the value presented here is an average. Cost is measured in €/MWh of demand. Note that the normalizing energy term (/MWh) represents exogenous energy demanded rather than energy produced. Energy lost to transmission or storage cycling is waste and should not be considered when normalizing.
- **CO₂ & SO₂:** This represents the average amount of emissions produced by the system over time. This is influenced primarily by the aggregate generation mix. However, thermal plants will consume more fuel per unit of output electricity when starting up or when operating close to their minimum operating load. The system can produce fewer emissions for the same power level by running fewer plants closer to their optimal efficiency point rather than several plants at the minimum. Each fuel has a level of CO₂ and SO₂ emitted per GJ burned, so calculating emissions is a matter of multiplying fuel consumed at each power plant by the constants. Emissions are measured in kg/MWh of demand.
- **Wind energy:** This is the energy contribution of wind over the study period. The actual amount of wind available at any given hour is determined by nature, but the amount of wind harvested by a wind farm will depend upon unit commitment decisions. If all thermal units are running near their minimum levels, a wind farm may be directed to feather blades and curtail output. The primary objective of this paper is to study the ability of demand response to encourage improved wind

harvesting through system optimization. Wind energy is measured in MWh, though it is presented here as a fraction of overall grid energy generation.

- **Natural Gas energy:** This is the total contribution of natural gas to the electricity mix. Natural gas is a highly responsive resource, with fast ramping rates and short advance notification requirements. It is a natural complement to variable wind power.
- **Coal energy:** This is the total fraction of all electricity that is produced by coal plants. Coal is an inexpensive fuel, but coal generators are slow to start up and cannot react to fast changes in wind availability.
- **Demand response impact:** This number represents the average amount of energy requested of demand resources per day. Not all participating ratepayers will be called for every demand event, so this does not reflect the impact of demand response upon every participating individual. It is a sum value for impact across the entire system. The value for the baseline case is of course zero.

Other metrics were calculated but will not be discussed further:

- **Fuel oil energy:** Fuel oil is used in Denmark primarily to provide heat, not electricity. It can provide some electricity as a peaking resource. The plants ramp quickly and have short lead times. Fuel oil plants have a high marginal cost like natural gas plants. The effects upon fuel oil generators are similar to those of natural gas generators.
- **Straw/Wood/Municipal Waste energy:** The system also contains straw, wood, and municipal waste generators. These make up a small portion of the overall energy mix. These generation units are relatively inflexible and will be used 1-2% more with demand response availability.
- **Reliability:** Reliability is measured as the number of hours that the system fails to meet load or spinning reserve requirements. The baseline case with no demand response has zero hours with reliability issues. There are still zero hours of inadequate supply or reserve after adding demand response.

4.1.2 Wind Harvesting

Uncertainty in the simulations comes mainly from wind power availability. The simulation also contains a chance of generation unit failure. Other probabilistic events such as run-of-river hydro power, hydro reservoir filling, and solar power collection are not used in this simulation. Demand forecasting is assumed to be perfect. Uncertainty can lead to significant "wind shedding", the process of feathering blades and collecting less than 100% of available wind energy. Wind shedding will be an important factor in the upcoming discussion of demand response and its effects on grid operation.

This figure characterizes the quality of wind forecasting and the ability of the system to harvest available wind:

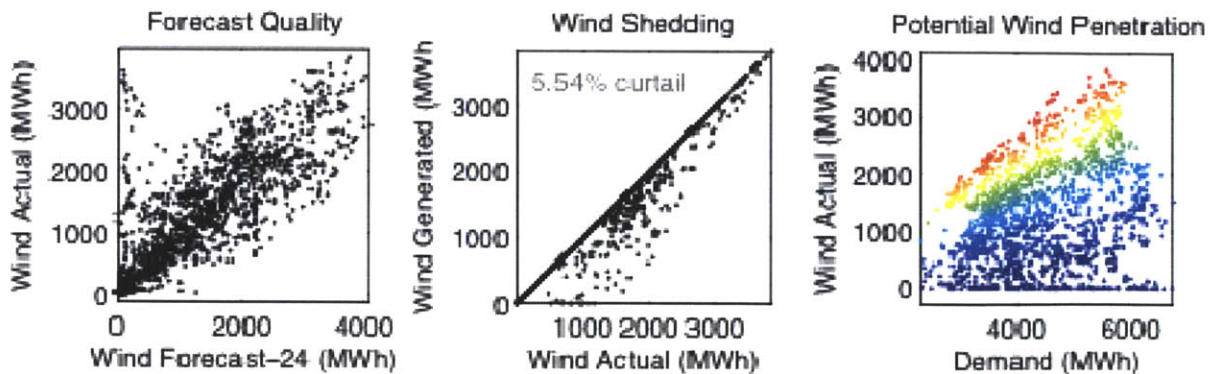


Figure 4-2: Wind forecasting and wind curtailment, stochastic optimization

The first plot shows the forecast of wind strength 24 hours ahead as compared to the actual realized wind strength. Wind strength is shown in units of energy rather than velocity to account for nonlinearities in the turbine's power curve. The actual wind is usually within a factor of 2 of the 24-hour forecast. However, there are notable deviations such as the cluster of time points on the left of the chart indicating a low forecast and much stronger realized value.

The second plot shows the ability of the wind generators to harvest the available wind energy. The X-axis shows the actual available wind energy and the Y-axis shows the energy produced. The strong diagonal line shows that wind producers usually achieve full wind production. However, there are times when wind operators must feather their blades and collect less than the entire available wind energy. Feathering is not ideal, but is a valid unit commitment strategy. During the night when wind is high and demand is low, it may be preferable to run a wind generator at less than its peak power output in order to keep a baseload generator above its minimum operating level. Increases in forecast error will increase shedding; an earlier unit commitment decision may not be able to harvest energy from a sudden midday gust of wind if there is no down-regulation capacity on the grid. In this case, wind operators were forced by system constraints to shed 5.54% of available wind energy.

The final plot shows the relationship between actual wind energy and the system demand. This shows the grid constraints that may lead to wind shedding. When the ratio of available wind power to demand is high (red colors), thermal generators are more likely to be at their minimum levels and require reduced wind production. This is related to the discussion of the load-duration curve from Figure 3-3. This chart uses a rainbow scheme to display the ratio between the two factors in order to prepare the reader for an identical color affordance in chart Figure 4-3.

Factoring each of the three above relationships as a ratio enables all data sets to be plotted on one chart. Each dot is one hour of simulation.

- The X-axis shows the wind forecast as a ratio of actual wind to the day-ahead forecast. Because the error is two-sided, the ratio is displayed logarithmically. 1.0 represents perfect wind forecasting.
- The Y-axis shows the wind feathering behavior as a ratio of generated wind to the actual wind available at realtime. 1.0 represents perfect wind harvesting.
- Color shows the maximum wind component that could be realized during that hour in relationship to the total demand. 1.0 would represent wind supplying 100% of demand, a condition never achieved in any simulation.

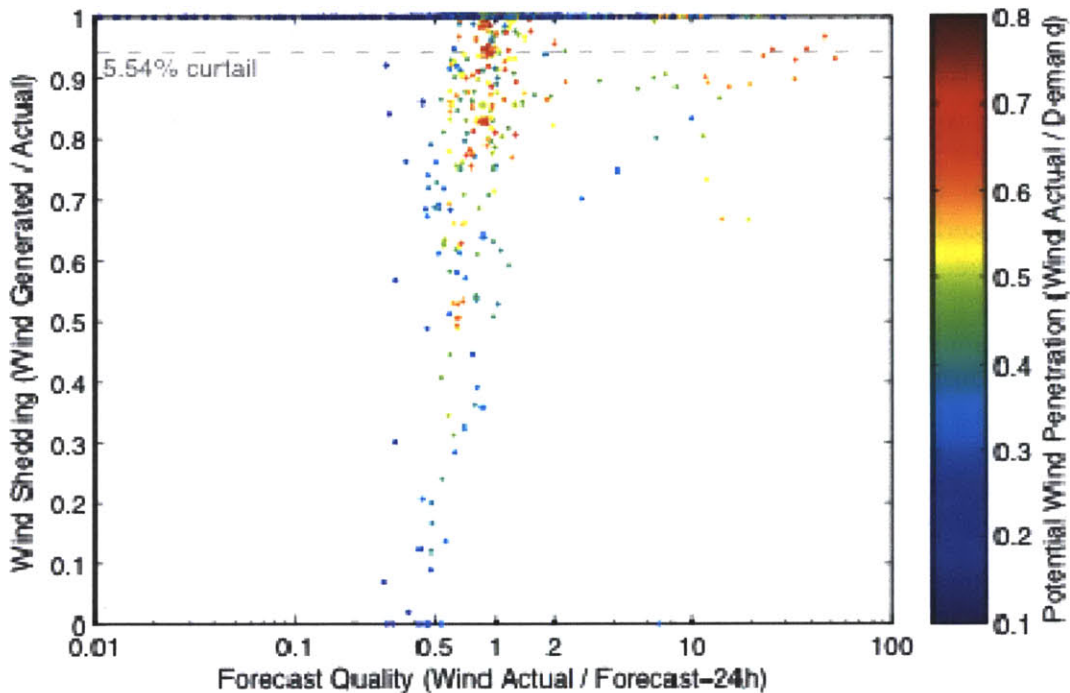


Figure 4-3: Consolidated view of wind forecasting and curtailment

When actual wind is much weaker than the forecast wind (left side of the graph), generators tend to harvest 100% of the available energy. The system is expecting more energy from the generators, so there is no need for feathering.

The center of the graph shows the typical case in which actual wind is within a factor of two of the forecast. Though the grid can generally work around these errors and allow wind to produce its theoretical maximum, there are some cases in which wind operators must make deep cuts to operations. Note the preponderance of blue marks in the lower region; this shows that the deepest cuts happen when wind generation is a relatively small amount of overall demand. By bringing a new thermal unit online, the wind is curtailed or even disabled to keep the thermal unit above its operating minimum.

The right region of the graph represents actual wind much stronger than the forecast. In most cases, wind operators can increase their production and thermal operators can

curtail levels. Notice the clusters of orange and red points in the 80-95% harvesting range; when wind represents a very high level of system power there may not be any room for the thermal generators to reduce capacity. However, there is so much wind that shedding 200MW of power is still a relatively small percentage of the overall production.

4.1.3 Thermal Unit Performance

Thermal generation units are those that burn fuel to produce electricity and (often) heat. With no hydro or solar capacity on the Danish grid, all non-wind generators are thermal units. Most thermal units use fossil fuels like coal or oil, but some use renewable fuels like wood. Thermal units have costs associated with starting up, as well as an amount of lead time before they reach operational capacity. The optimization must work around the technical constraints on thermal units to balance demand when wind supply varies.

The "capacity factor" of a unit shows how frequently and at what output level that unit is used. The nameplate power rating of a generator limits its maximum energy production during the study period. In reality, few units are run at their maximum level all the time. Wind generation is limited by available resources and thermal units are dispatched according to system need. The fraction of realized energy produced divided by potential energy is the capacity factor. The following table shows capacity factors for each generator, aggregated by fuel type. The capacity factor for the least-used and most-used generator is also shown. Note that the table shows only electrical capacity factor, not heating capacity factor. Note also that generator counts and nameplate power numbers may be different than Table 3-2, as this table only includes units that generate nonzero levels of electrical energy during the study period.

Fuel	Generator Count	Electrical Power (MW)	Electrical Energy (GWh)	Capacity Factor (mean)	Capacity Factor (min)	Capacity Factor (max)
Natural gas	9	2567	306	0.09	<0.01	0.26
Coal	12	4358	3163	0.54	0.12	0.75
Fuel oil	5	903	42	0.04	<0.01	0.19
Muni waste	5	300	333	0.83	0.63	0.98
Straw	7	297	93	0.23	0.01	0.70
Wood	2	317	206	0.48	0.02	0.57
Wind	2	4100	1671	0.30	0.28	0.37

Table 4-2: Electrical load factors of generators, no demand response

Municipal waste and coal plants spend more time running close to maximum than other fuel types. These plants have long lead times and slow ramp rates, so the optimization prefers not to vary their output over the course of the simulation. Natural gas and fuel oil units react quickly and are frequently cycled. These units are offline almost the entire time with only a few hours of operation, hence the low capacity factor and existence of units with load factors less than 1%. Wind achieves a capacity factor of 30%, which is in the reported range of 25.7-46.7% for Danish wind farms (Smith 2011).

The load profile of a unit shows how much time it spends at a given percentage of its maximum output. The following chart is a histogram showing number of hours on the Y-axis and levels of output on the X-axis. Black bars represent electricity output and white bars are heat output. The dashed line shows the minimum stable electrical output of the plant. The plant may spend some time below this level while starting up or shutting down. The far left bar (zero load factor) represents time that the unit spends shut down. The maximum electrical capacity is shown in the upper right corner of the graph.

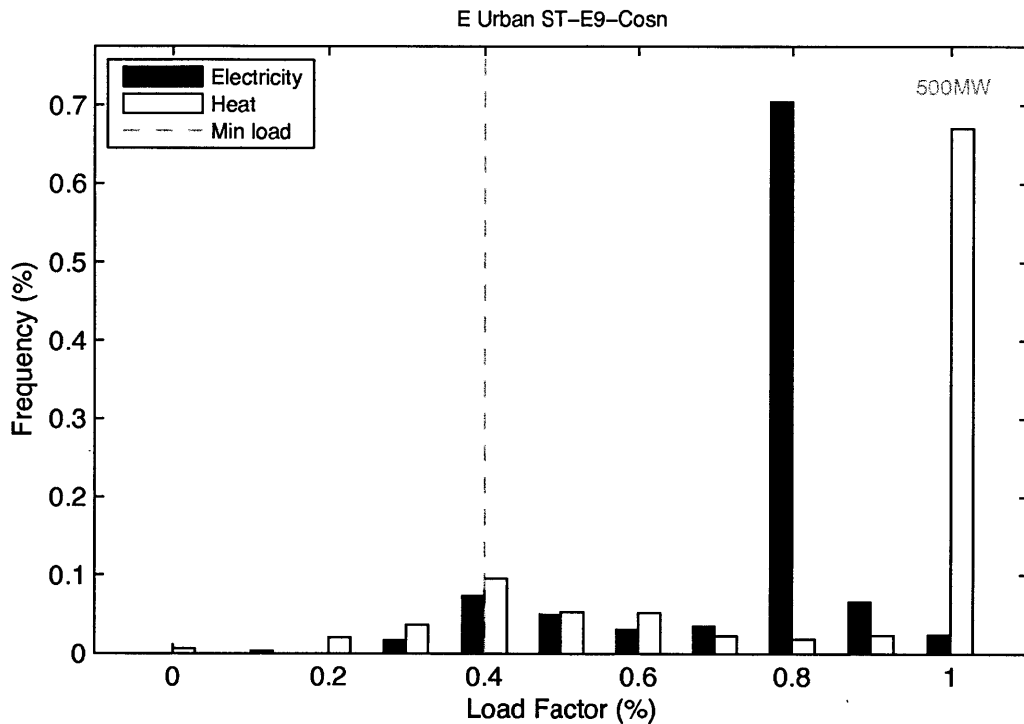


Figure 4-4: Load profile, coal unit, no demand response

This generator runs consistently close to its 80% of its maximum electrical capacity and 100% of its heat capacity. It is one of the more consistent units in the coal fleet.

The following chart displays the load profile for selected coal generators. Most are combined heat and power units, though one (E Rural ST-C8-COs) is electricity-only.

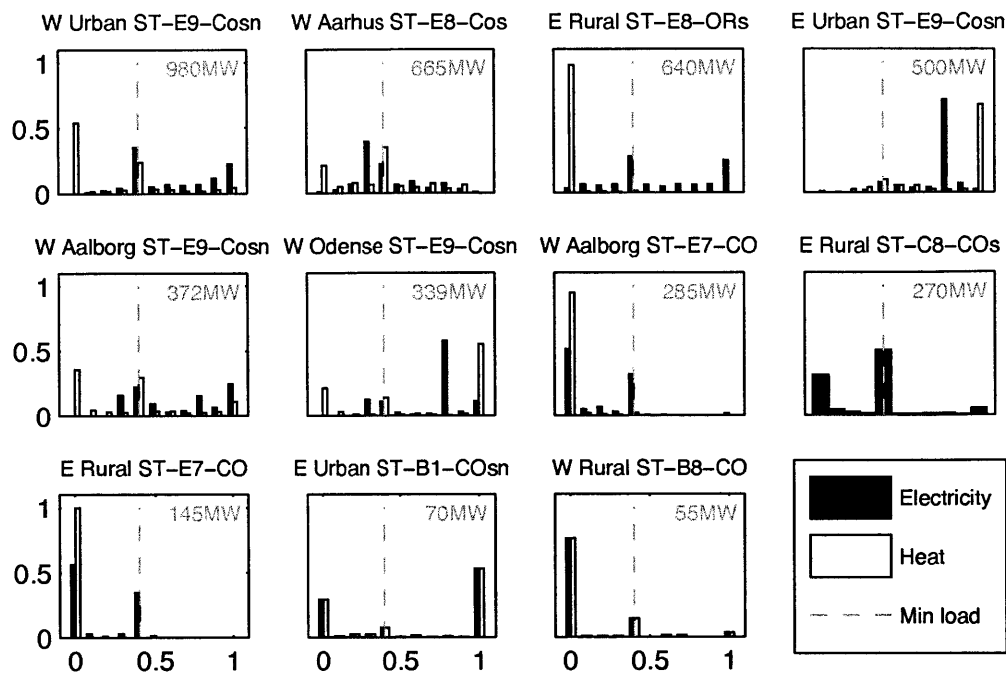


Figure 4-5: Load profile, multiple coal units, no demand response

The optimization favors running multiple units near their operational minima. All units are shut down at some time or another; this does not look like traditional baseload power. (The generator from Figure 4-4 is in the upper right here.)

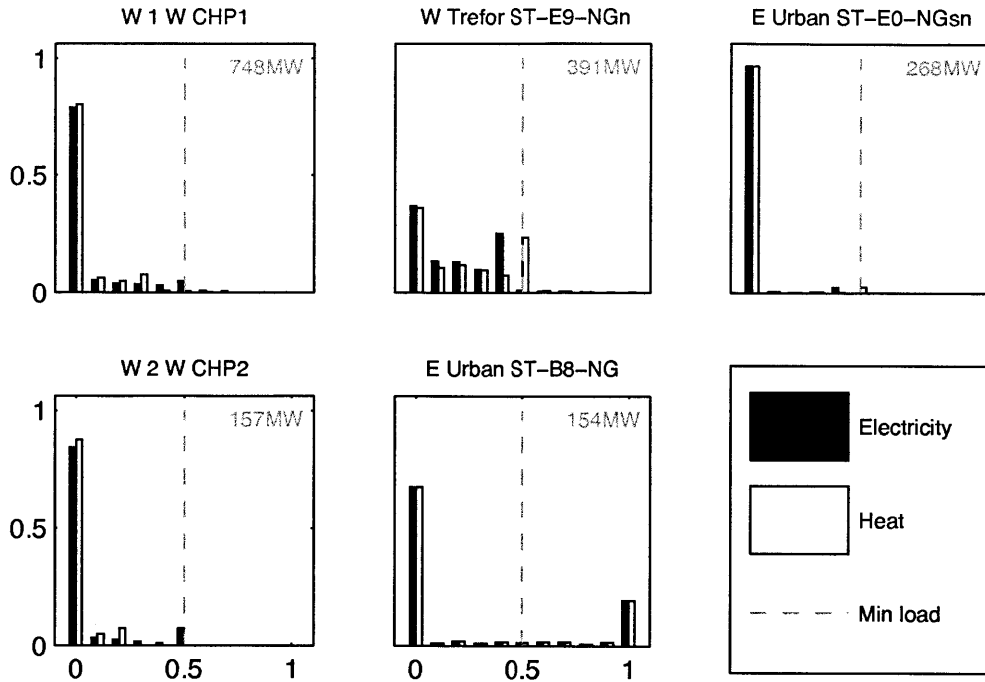


Figure 4-6: Load profile, multiple natural gas units, no demand response

The optimization barely dispatches natural gas units at all. This accounts for the typical 9% capacity factor seen in Table 4-2. Most of the generators spend their time shut off, then ramp up for a short burst at minimum capacity, then ramp back down. This is not the desired mode of operation for any generating unit, as it runs the plant in its inefficient range and wastes fuel on startup/shutdown.

4.2 Single Result with Demand Response

The primary purpose of this investigation was to explore the impact of demand response price and quantity upon system operation. This section's results show the effect of a single combination of these two demand response parameters upon system cost, emissions, and operational decisions.

The optimization uses demand response to fill gaps when thermal units are not available. The following figure shows demand response being dispatched on a windy summer day. Similar to Figure 4-1, the colored bands show the electricity produced by generators in different fuel classes. This figure adds red bands to show times when demand response is active. Note that total energy consumed can now exceed the exogenous demand levels, indicated by a magenta line for the "rebound effect." This case was run with 2100MW of available demand response, priced at 120% that of a natural gas unit.

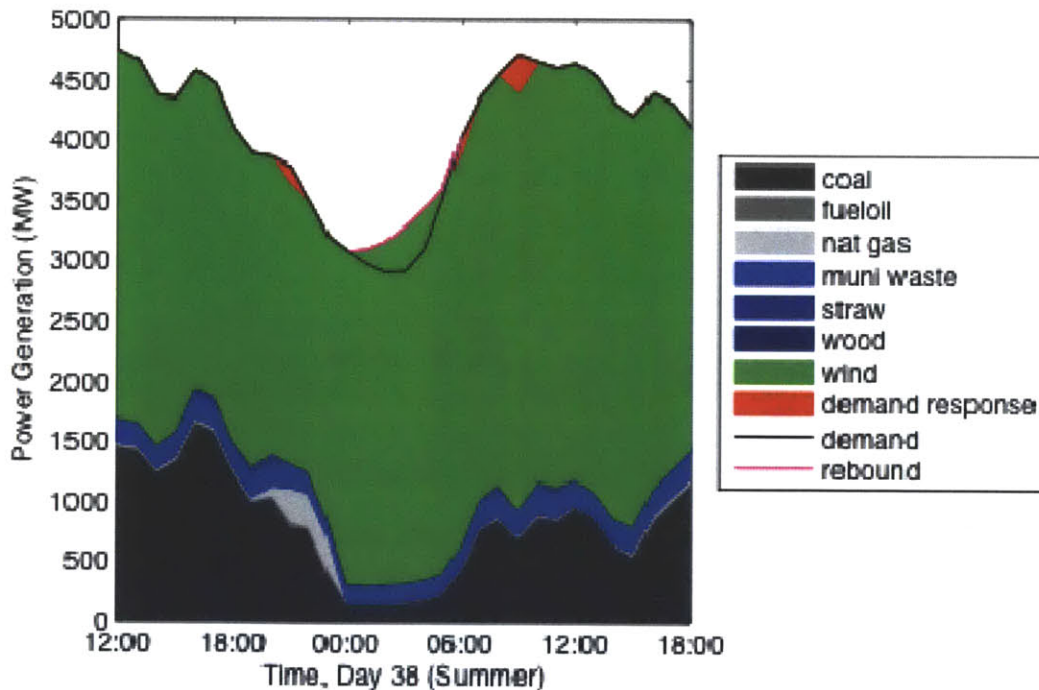


Figure 4-7: Generation mix, with demand response (Q=0.5, P=1.2)

In the simulation, strong winds overnight lead to a near-zero net demand. Building owners pre-cooled their facilities and shifted consumption into the night to prepare for a demand curtailment the next day. Demand response stepped in during the rapidly rising demand and at peak on the next morning. The slow-moving municipal waste plant did not have to shut down overnight.

Demand response addresses both wind uncertainty and demand peaking. Even with zero wind contribution, the system will still call upon demand resources to optimize operational cost. The following figure compares a full wind (4100 MW) case and a zero-wind case. Both use 2050 MW (quantity = 0.50) of load shifting demand response with a price ratio of 1.2 as compared to a natural gas generator.

The differences between the two operational philosophies are best illustrated by inspecting the use of demand resources by time of day. The left chart compares the with-wind demand response energy calls paired hourly with their corresponding hours with no demand response. The right chart shows the average demand response call by hour for both wind cases. Color on both charts indicates time of day, with red at noon and blue at midnight. Note that this graphic does not include the "rebound effect"; all values represent only demand that is lower than exogenous system normal.

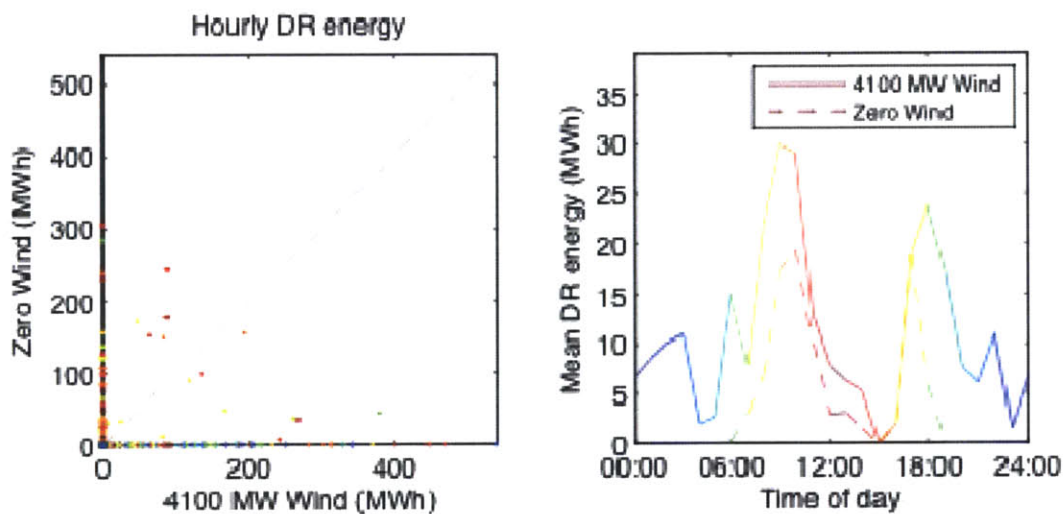


Figure 4-8: Demand response calls by time of day, wind and no wind

When the system has no wind, demand response is only used to manage consumption during the mid-day and evening peaks. With unpredictable wind, demand response may be dispatched at any time of day. The mean energy called per day is 3x as large with wind than without. Event length is 2x larger with wind than without. Roughly 2/3 of the value of demand response is in service of handling wind uncertainty.

Demand response acts as a cushion to prevent high system costs. The baseline no-DR scenario is not capacity constrained, even in the winter peaking weeks. It therefore contains no extremely large cost spikes. The comparison of hourly prices between the two cases (left) shows no large deviations. Viewing average system cost by hour (right) shows that demand response depresses cost by 1-2 €/MWh at all times of the day. Average system cost without demand response is €27.97/MWh. With demand response, system cost drops to €26.92/MWh.

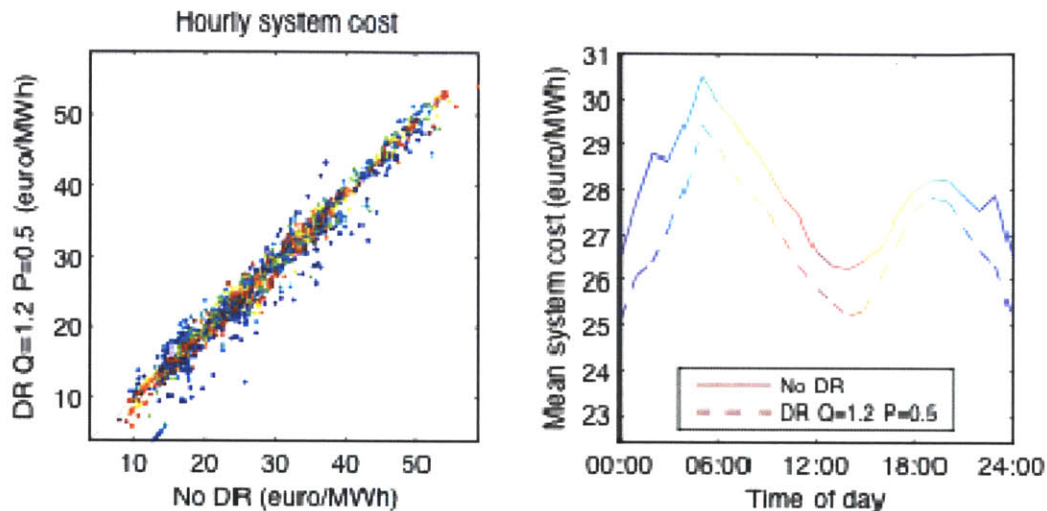


Figure 4-9: System cost, with and without demand response

Note that in all stochastic runs, a different actual wind pattern will appear. Values that are summed over the entire study period are roughly comparable since differences will even out in the long run. Any differences at the hourly level will depend upon wind behavior and system response to that wind. At the fine time scale presented here, individual differences cannot be distinguished between wind and demand response.

Demand response can react faster to changing conditions faster than a natural gas or fuel oil unit. This enables the optimization to schedule more wind in times when a less-responsive system would conservatively plan to feather wind power. Similar to Figure 4-3, the following chart shows wind load shedding as a function of forecast quality and net demand levels.

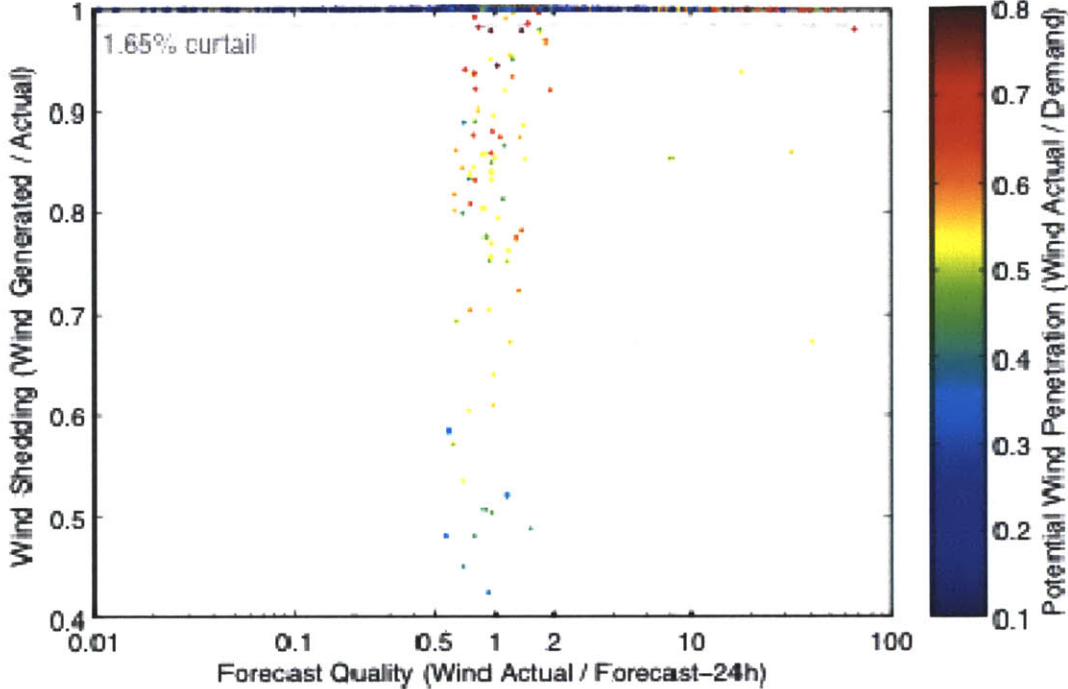


Figure 4-10: Wind harvesting with demand response

Curtailment levels are only 1.65% with demand response as compared to 5.54% without. There are zero hours in which wind had to be completely curtailed. The case of higher-than-expected wind with low net demand (red dots on the upper right of the graph) has more hours with 100% wind harvesting.

Demand response causes the optimization to run thermal plants differently. As in Figure 4-4, this chart shows the load profile for Denmark's largest coal plant. Black bars show the baseline case with no demand response. Red bars represent a system with demand response.

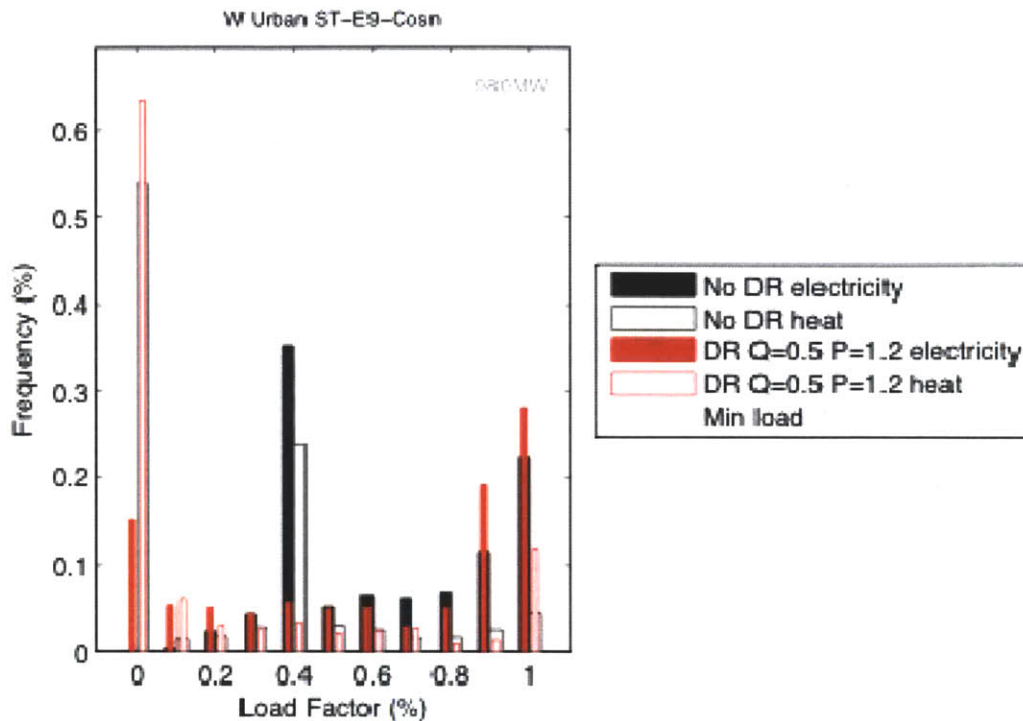


Figure 4-11: Load profile, coal unit, with demand response

With demand response, this coal unit runs at its minimum generation much less. While it was never shut off without demand response, demand response causes the plant to be shut down about 15% of the time. Total time spent at 90% and 100% load is greater.

This effect is consistent across units. This plot shows the load profile for the coal and natural gas generator classes. Heat has been removed for readability's sake.

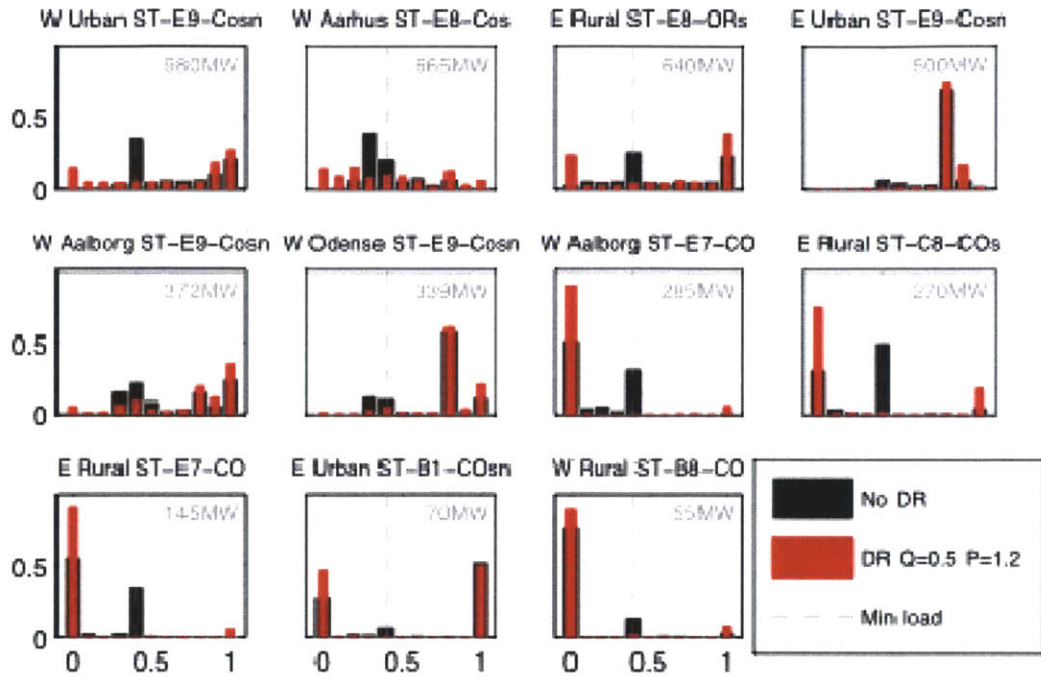


Figure 4-12: Load profiles, coal class, with demand response

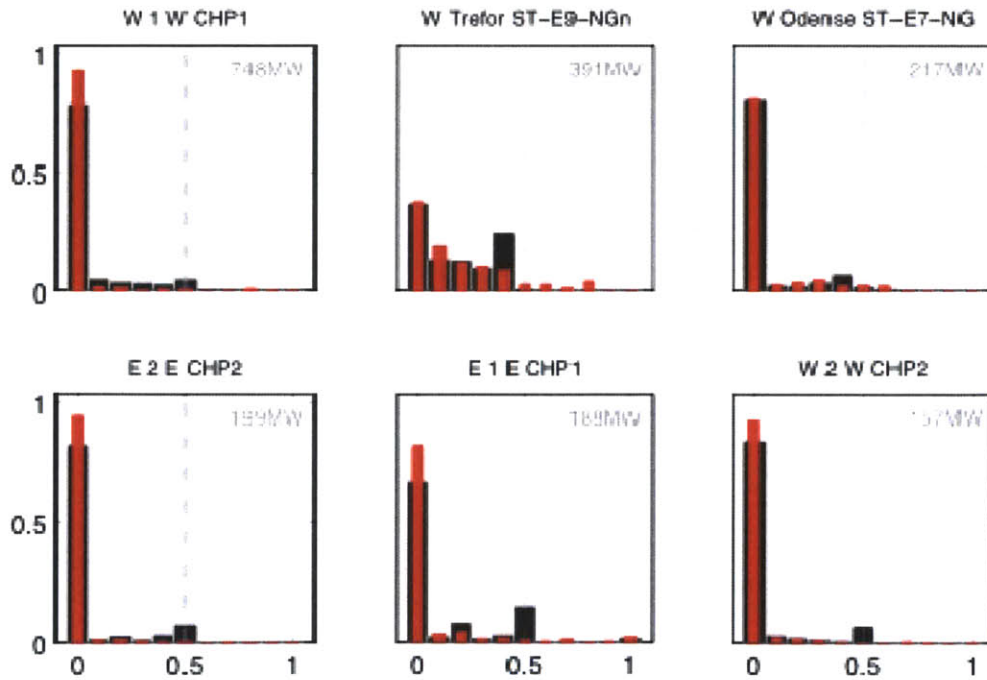


Figure 4-13: Load profiles, natural gas class, with demand response

Demand response enables the electrical system to run with a more efficient use of resources. Adding demand resources improves cost, lowers emissions, enables more wind, and reduces use of fossil fuels.

Criterion	Unit	Value (no DR)	Value (DR)	Δ (%)
Cost	€/MWh	27.97	26.93	-3.72
CO ₂ emissions	kg/MWh	166.2	160.1	-3.81
SO ₂ emissions	kg/MWh	1.09	1.05	-3.81
Wind energy	fraction	.280	.313	+11.8
Coal energy	fraction	.550	.533	-5.5
Natural gas energy	fraction	.053	.039	-26.4
Demand response impact	MWh/day	N/A	265	N/A

Table 4-3: System operation evaluation criteria, with demand response

4.3 Demand Response Parametric Search – Price & Quantity

This investigation has now explored the baseline behavior with no demand response and looked in-depth at the system with one combination of demand response price and quantity. This section will look at a family of demand response parameters and view the system behavior.

The following several charts will display the results of the system evaluation criteria as a function of the size/quantity input parameters. In most cases demand response increases wind power harvesting and reduces thermal plant use, which lowers system cost and emissions. The quantity of demand response generally has more system impact than the price; it is generally valuable enough that the optimization uses as much of the resource as possible.

This graph shows the search space used for all following simulation runs.

The Y-axis represents the bid price of demand response as a normalized fashion relative to the most expensive peaking unit on the system. Price is expressed on the plot through color, blue cheapest and red most expensive.

The quantity (MW) of DR is expressed on the X-axis as a normalized factor relative to the nameplate capacity of the amount of wind on the system (MW). The smallest amount of demand response power capacity (shown by a plus) is 5% the nameplate power capacity of wind. For ease of tracking, all equal-quantity values are connected by a thin line. Note that in all cases the energy capacity (MWh) of the demand response scales equally with respect to the power availability.

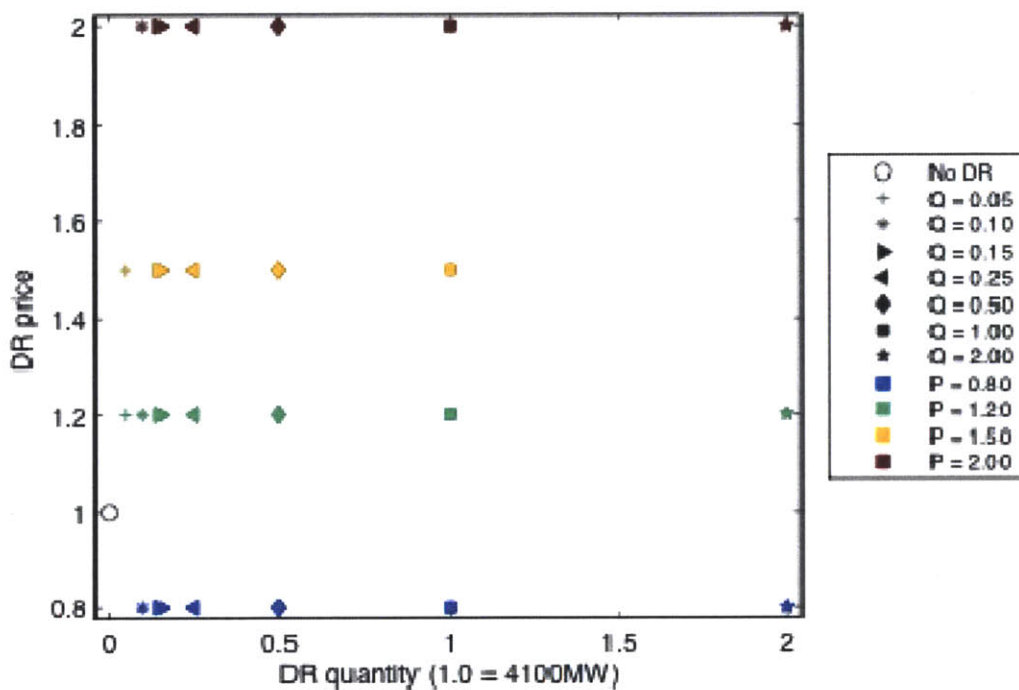


Figure 4-14: Demand response experimental space

This graph shows no results, only the relationships of the inputs. Not every possible combination of parameters was simulated. Some parameter pairs were simulated multiple times to test the repeatability of the results. The following figures show the system's operational metrics within the framework of this study space.

The objective function of the optimization is system operational cost (€/MWh). This is normalized relative to the baseline case with no demand response. Cost varies over time, so the value presented here is an average over the 8-week study period.

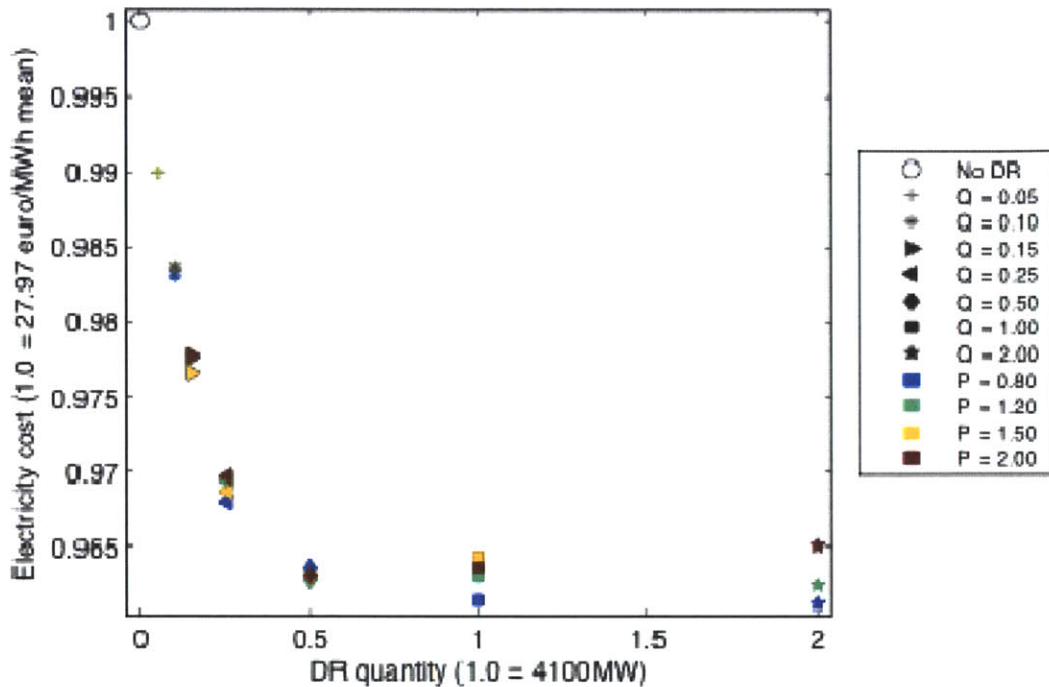


Figure 4-15: System operation cost with demand response

Extremely small quantities of demand response have little impact upon system cost. With little capacity available, the resource is not deployed as frequently or as deep as the optimization would prefer. A 0.25 quantity rating shows a 3% cost savings for all price levels. Further quantity levels asymptotically approach a improve performance of 3.5%. Because the rebound effect requires more net energy to be generated than is curtailed, increasing energy production demands match the decreasing marginal value of the next unit of demand response energy. Quantity has much more effect on system cost than does price.

Demand response enables cost savings through enabling more wind power (near-zero marginal cost) and through more efficient deployment of thermal power generators. These aspects will be explored in greater detail later.

Unlike other energy sources, deploying demand response has a direct effect upon ratepayers' lives. The level of this impact will influence the ability of the system operator to deploy demand response as a solution. In the following chart, the X-axis represents the normalized price of demand response. The Y-axis represents the impact that demand response has upon ratepayers. The metric represents the average energy dispatched per day. (MWh/day) DR impact internalizes both quantity and price while reflecting the experience of participating end-users.

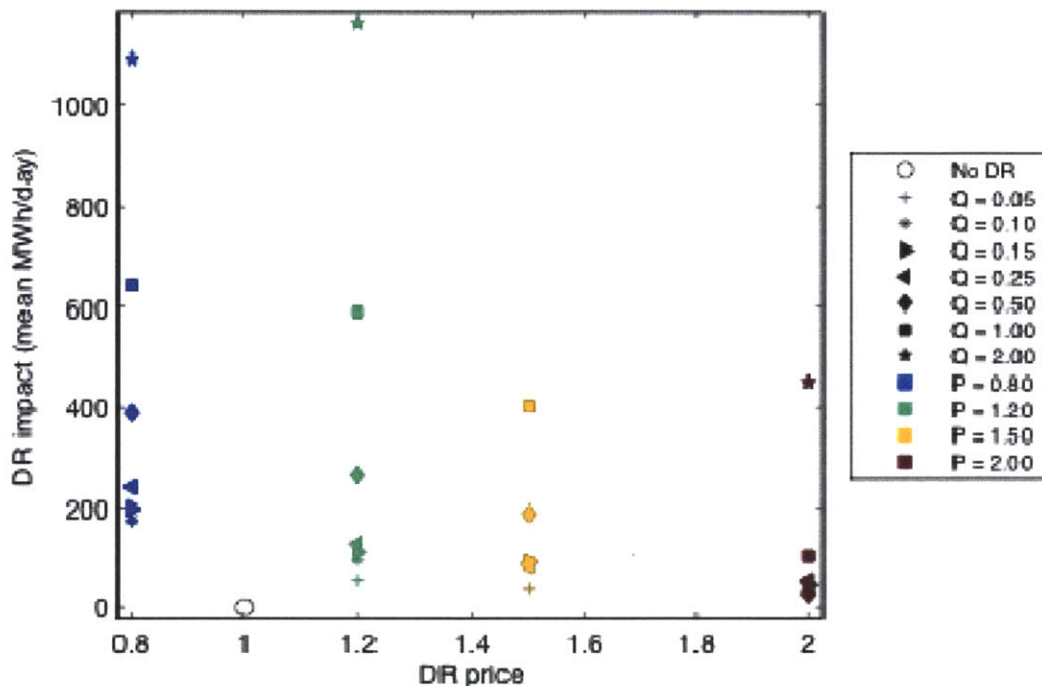


Figure 4-16: Demand response price and customer impact

When demand response is plentiful and cheap, it is called frequently. Case (P=0.80 Q=2.0) sees an average of 4.3 hours of demand response calls per day with some extreme cases up to 16 hours. These are not necessarily large in terms of energy demanded; even though the Q=2.0 cases have an unrealistically large 8200MW power rating, the maximum hourly call is 500MW. Larger pools of demand response resources create larger use not through power capacity but through energy capacity. Each individual resource is limited in the duration that it may be called. By chaining multiple customers' curtailments, a demand event can last much longer than any given customer's ability to participate. The generous quantities available mean that the pool of curtailable load is never close to being exhausted.

In contrast, small quantities of demand response are energy-limited. The (P=0.80 Q=0.10) case has 40 days with some sort of demand call as compared to 45 days at Q=2.0. Each of these calls are shallower (max 200MW) and shorter (2.7 hours) in order to avoid emptying the pool of demand resources. In this case, there are 5 occasions on which demand response is unable to sustain operations at current rates.

Demand response impact will be displayed as the X-axis in most upcoming figures.

This information can now be used to trade off efficiency of system operation and customer impact of demand response. This graph does not put either of the input variables (demand response price of quantity) onto the X- or Y-axis. Use color, shape, and the legend to map individual cases to the outcomes. The X-axis contains demand response impact, as in Figure 4-16. The Y-axis represents normalized system cost, as in Figure 4-15.

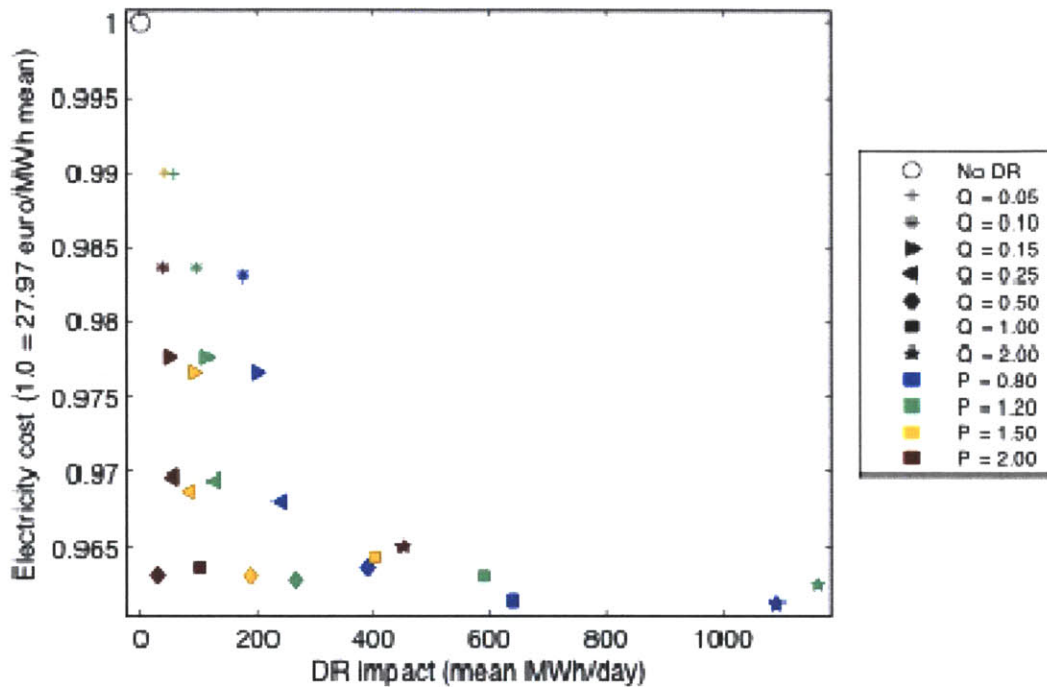


Figure 4-17: Demand response impact and system cost

As long as demand response quantity is at or above 50% of wind power, system costs remain relatively constant for all possible levels of demand response impact. Lowering the price of demand response serves to increase perceived customer impact (by up to an order of magnitude) without improving costs. Until $Q=0.5$, the system is capacity-constrained and is willing to use demand resources at any price. Beyond $Q=0.5$, the optimization will use demand resources at all price points without improving system costs significantly.

At the same price ($P=2.0$), perceived impact varies between 25-100 MWh/day, a small variation on the scale of results shown here. However, system operation costs span the entire range of results received.

From Figure 4-8, 2/3 of the function of demand response on the grid is in cushioning the unpredictable effects of wind power. This figure shows the relationship between demand response and wind penetration. The Y-axis here represents the amount of energy contributed to the grid from wind power. The factor is normalized relative to the contribution of wind in the no-DR baseline case. Demand response changes wind scheduling by providing more emergency up- and down-regulation to the system. Wind does not need to be curtailed as much in order to keep thermal generators running at their operational minimum levels.

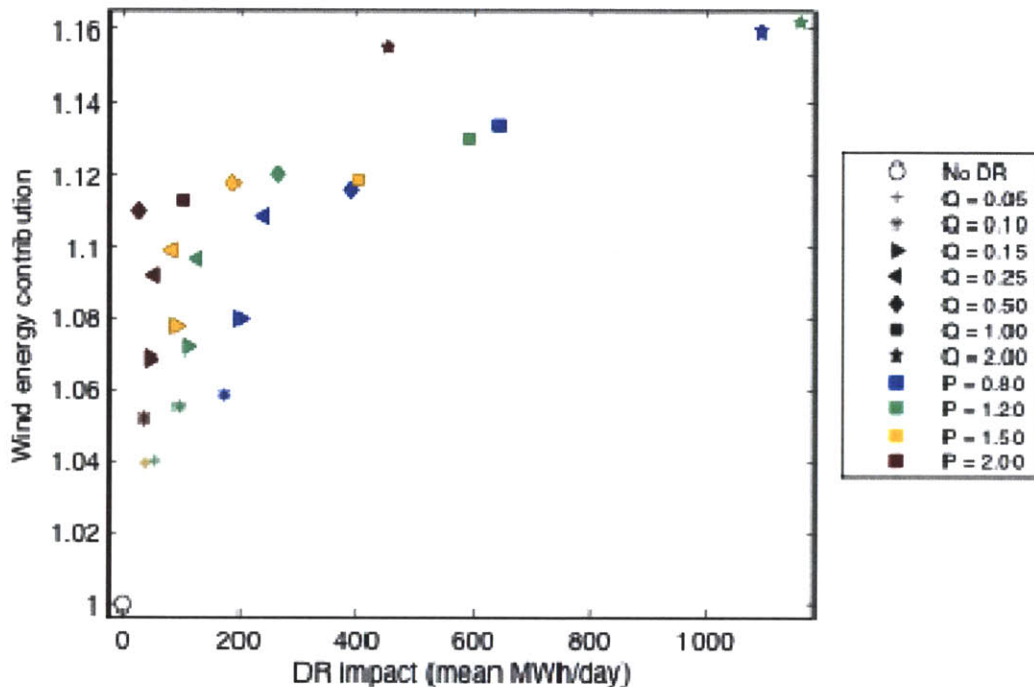


Figure 4-18: Demand response and wind energy

Even with small amounts of demand response ($Q=0.10$), wind contributes 5% more to the overall energy mix than with no demand response at all. Feasible quantities of demand response ($Q=0.25-1.0$) show improvements in the 10% range. The infeasible $Q=2.0$ case is able to increase wind penetration by over 15%. Price has a smaller but still noticeable influence on wind generation.

All demand response parameters have <1% wind shedding during the 17:00-20:00 hours when electrical consumption is high and wind is low. Demand response influences wind shedding before and after this peak time, enabling a greater harvest of wind energy. At 12:00, ($P=0.80$ $Q=0.5$) averages 1.8% wind shedding while wind producers must shed 4.0% with no demand response. Most of the wind energy is added to the system during off-peak hours due to increased demand from load shifting.

The simulations also enable an investigation of the system's emissions. In the following figures, the X-axis uses the same demand response impact measure as above. The Y-axes show the normalized CO_2 and SO_2 emissions for the grid, relative to the baseline case with no demand response.

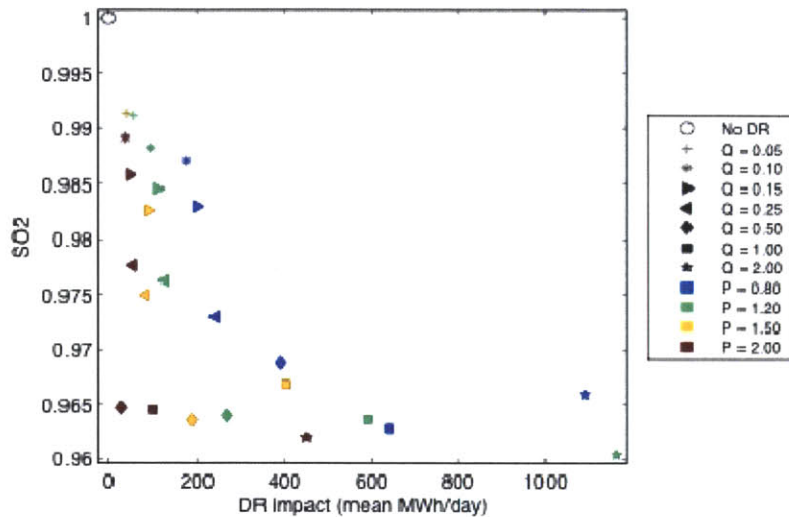


Figure 4-19: Demand response and SO_2

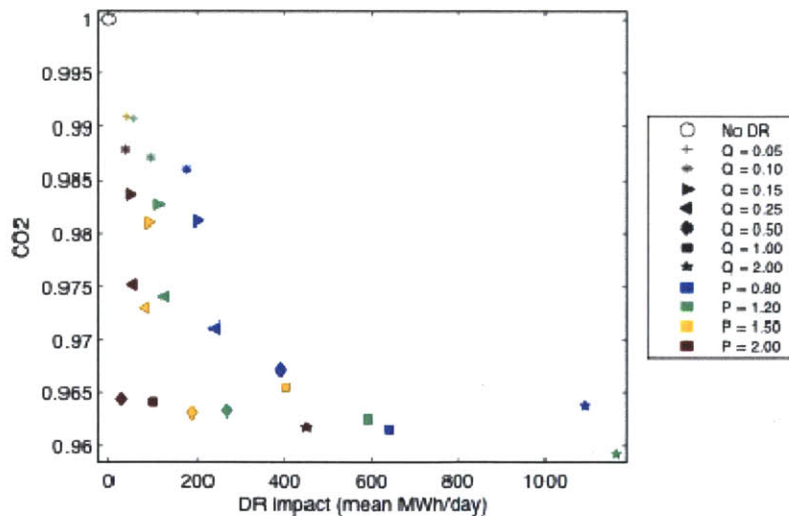


Figure 4-20: Demand response and CO_2

Similar to the operating cost numbers, emissions decrease as the size of the demand response pool increases from $Q=0.05$ to $Q=0.50$. Beyond $Q=0.50$, emissions remain relatively flat. Decreasing demand response price has some influence on emissions, but not as much as quantity.

Since emissions are proportionate to fuel burned and demand response does not significantly alter the fuel mix, the shapes of the SO_2 and CO_2 graphs are quite similar.

The fuel mix is the primary driver of emissions and cost. Thermal generator operations (startup/shutdown, load factor) contribute as well. The contributions of each fuel type therefore explain the variation in system performance. This chart shows coal plants' contribution to the grid's energy needs, normalized by the amount of coal used in the no-DR baseline case.

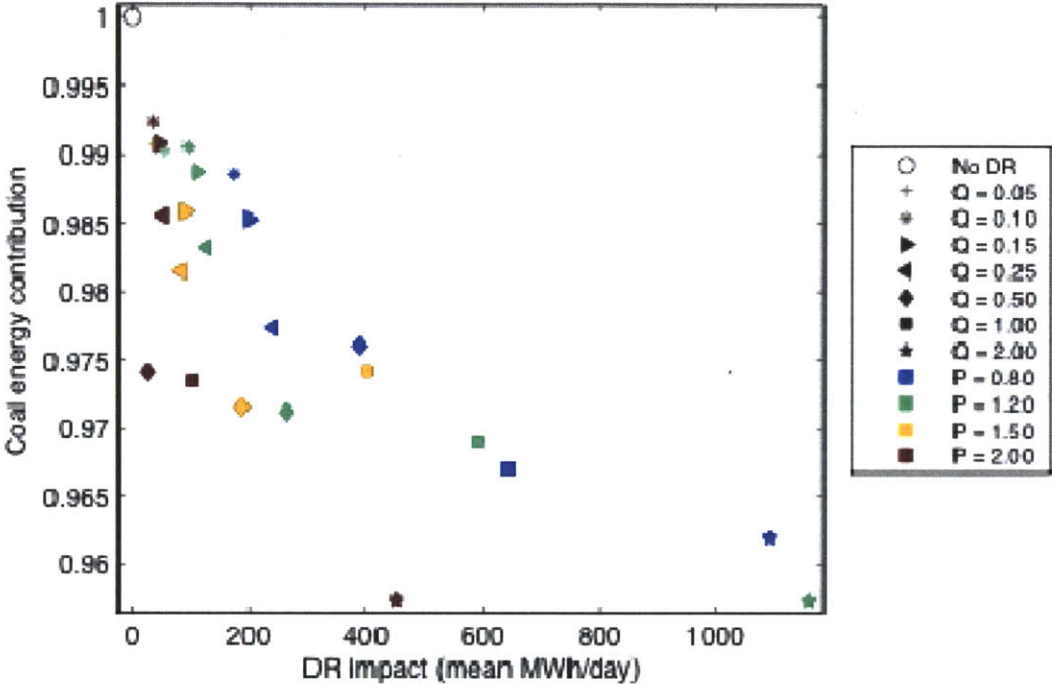


Figure 4-21: Demand response and coal energy fraction

Coal represents the largest source of energy on the Danish grid. Wind is the second largest. As demand response enables wind power to increase (Figure 4-18), coal power decreases.

When demand response capability is small (Q=0.10), startup counts of coal plants can actually increase by 1-2% (Figure D-1). Otherwise, demand response enables coal plants to start up about 10% less often.

Demand curtailment substitutes for natural gas generation, which already makes up a relatively small (less than 10% that of coal) amount of the system's energy contribution. As a highly responsive resource with high marginal cost, natural gas and demand response behave similarly when handling forecast uncertainty from wind. The substitution effect is not perfect since demand response is duration-limited and does not produce heat. This graph shows the quantity of demand response resources (X-axis) against the total energy contribution of natural gas generators (Y-axis). The natural gas energy contribution is normalized relative to the amount of natural gas generation in the base case with no demand response.

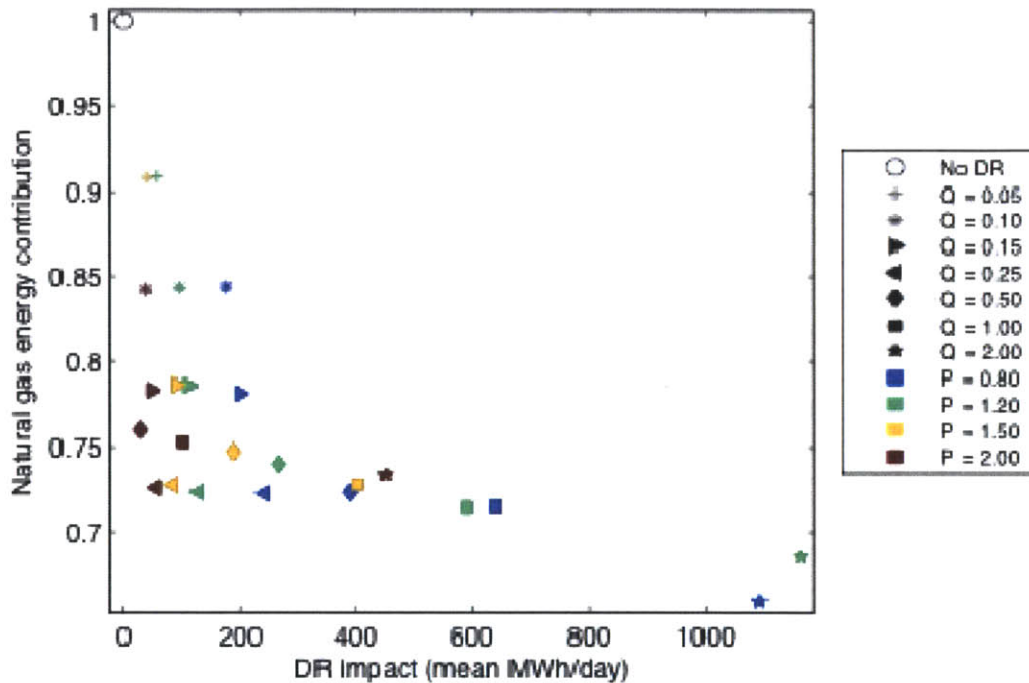


Figure 4-22: Demand response and natural gas energy contribution

In most cases, natural gas generation represents about 25% less energy contribution than the base case. A regression (Figure C-2) of generation from flexible fuel sources (natural gas & fuel oil) shows a coefficient of -0.84 ($r^2=0.58$). Every 100 MWh of demand response called reduces the amount of flexible fuel generation by 85MWh. The elasticity of this substitution (and the quality of the fit) increases by ignoring the $Q=2.0$ data. Demand response is definitely substituting for natural gas generation.

When $Q>0.10$, startups of natural gas plants are reduced by 40% (Figure D-2). Reducing startups causes less waste from startup costs and enables the plants to run in a more efficient operating range.

The preceding investigation has defined system response in terms of abstract demand response quantity and pricing. The work of selecting an ideal price and quantity depends upon customers' tolerance for electrical interruption. Determining the social cost of

demand response and selecting an optimum parameter pair is beyond the scope of this study. Other studies can show the shape of the social utility curve and point the way to future directions. Lassen and Jensen (2005) surveyed Danish residential electricity customers to determine their willingness to have appliance use interrupted. The survey covered both event duration and frequency. It found that customers placed a high per-hour price on any interruption at all, but that as frequency and duration increased customers demand marginally less compensation. This result implies that the lowest-cost solution per unit of energy curtailed is to have a smaller pool of demand response customers and to call them regularly for longer individual events.

€/hr compensation ⁴		Frequency (events/appliance-year)		
		2-3	12	30
Duration (hr)	0.25	18.27	6.04	2.55
	1	6.04	1.88	0.80
	3	2.68	0.94	0.53

Table 4-4: Ratepayer compensation for curtailment (Lassen and Jensen 2005)

This result may be dependent upon technology. Manual curtailments require that the ratepayer change the daily patterns of his or her life by deferring laundry or dishwashing. Once the customer has had to change consumption patterns, it may matter less to the customer how long the deferment lasts. The survey did not look at passive curtailments such as temporarily changing the setpoint on a water heater. In this case, customers may prefer more frequent and shorter interruptions. Customer preferences and available technology will influence the ideal dispatch of demand resources.

Without a complete study of consumer preferences, it is still possible to identify a preferable quantity and price for demand response resources. Improvements to system cost are more dependent upon quantity than price, flattening out at $Q=0.25$. At each quantity level, higher prices lead to lower demand response use but also cause more wind curtailment and higher emissions. A price ratio of $P=1.5$ favors low customer impact as an enabler of implementation. This case ($Q=0.25$ $P=1.5$) represents favorable tradeoffs between cost, customer impact, and emissions goals. In comparison to the ($Q=0.5$ $P=1.2$) case discussed in section 4.2, this case achieves within 1% of the cost and emissions outcomes at 1/3 the customer energy impact (MWh/day).

⁴ Survey conducted in Danish Krone at $\text{€}_{2004} = \text{DKK } 7.44$

Criterion	Unit	Value, no DR	Value, DR Q=0.5 P=1.2	Value, DR Q=0.25 P=1.5	DR Δ (%)
Cost	€/MWh	27.97	26.93	27.10	-0.7
CO ₂ emissions	kg/MWh	166.2	160.1	161.7	+0.9
SO ₂ emissions	kg/MWh	1.09	1.05	1.06	+1.0
Wind energy	fraction	.280	.313	.307	-1.9
Coal energy	fraction	.550	.533	.539	+1.1
Natural gas energy	fraction	.053	.039	.039	--
DR impact	MWh/day	N/A	265	85	-67.9

Table 4-5: System operation evaluation criteria, with demand response

4.4 Business Impact

The system operator or regulator may be primarily concerned with the overall efficiency or emissions of the grid, but the inclusion of robust demand response will also affect the fortunes of generation unit owners and operators. Segmenting stakeholders into beneficiaries and those who would be hurt helps system planners identify sources of resistance and make accommodations. Demand response is a new business with its own implications. Wind operators are beneficiaries of demand response, coal operators see small changes, and natural gas generators are hurt significantly.

4.4.1 Pricing Demand Response Capacity

The business of demand response is based on bidding "generation" in the form of reduced demand into energy markets. The curtailment services provider (CSP) pays the end-use customer an incentive payment and retains the rest for operations and profit. The optimization considered only a timescale on which marginal cost for dispatched energy ("energy payments"). System operators also make "capacity payments" to have the resource standing by and ready to curtail. Capacity payments are fixed costs paid whether the resource is used or not (€/MW). As demand response calls can be rare, capacity payments represent most of the income received by CSPs and their end-use customers.

This study has heretofore focused upon the days/weeks timescale of unit commitment in which demand response dispatch is considered only as a function of its marginal cost. The optimization also contains enough information to calculate an annual value of demand response capacity. A system operator must pay for every megawatt of demand response. Since all cases with demand response saved system costs, there is a breakeven point at which the operator is willing to pay for that capacity.

The break-even payments are expressed as:

$$\Delta\text{SystemCost}_{\text{nodr-dr}} = \text{EnergyPayments}_{\text{dr}} + \text{CapacityPayments}_{\text{dr}}$$

or

$$\begin{array}{cccccc} \Delta\text{SystemCost}_{\text{nodr-dr}} & = & \Sigma\text{EnergyPrice}_{\text{dr}} & * & \text{Energy}_{\text{dr}} & + & \text{CapacityPrice}_{\text{dr}} & * & \text{Capacity}_{\text{dr}} \\ (\text{€}) & & (\text{€/MWh}) & & (\text{MWh}) & & (\text{€/MW}) & & (\text{MW}) \end{array}$$

All terms are known for each case except $\text{Capacity}_{\text{dr}}$, which can be calculated from the others. Each generator may have a different price and a different amount of use; recall that fast-acting demand resources enjoy a price premium over those with longer lead times. The resulting capacity calculation is an average for all generators. Capacity payments for fast response will realistically prove to be higher and slow response will have lower payments. Note that the calculation shows the value of capacity only during the 8-week study period. Reported capacity payments are annualized by multiplying by 52/8. The resulting quantity represents the maximum that the system operator is willing to pay per year for standby demand response capacity; most likely actual payments will be less.

Plotting against the relative energy price for each case, the capacity price is surprisingly uniform. For any given quantity of demand response resources, capacity prices remain the same for all energy prices. The high end for capacity prices is approximately €40k/MW-yr. This is representative for the capacity price of a peaking advanced combustion turbine, which is about \$₂₀₀₇60k/MW-yr at a 10% carrying cost. (EIA 2009)

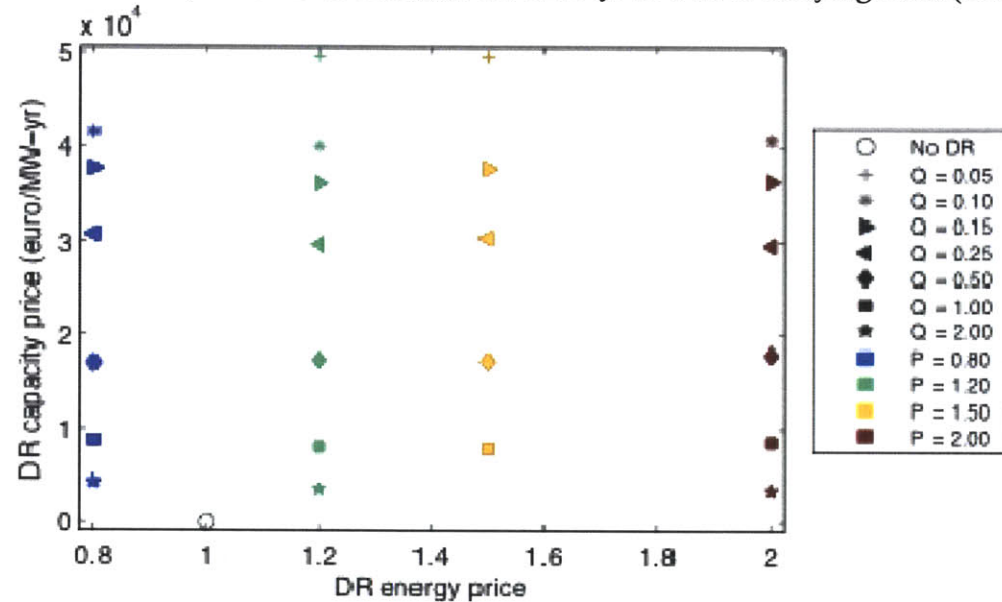


Figure 4-23: Demand response capacity price

Any capacity payment below these amounts represents an improvement to system efficiency. Capacity price is a function of quantity. A small pool of demand resource resources is valuable for every unit of capacity. Larger amounts of demand response capacity are less valuable.

This quantity/capacity-price relationship defines a demand curve for the system operator. The system operator is willing to pay a certain amount for a certain quantity of the service. The regression below eliminates the unrealistic $Q=2.0$ and the null $Q=0$ cases.

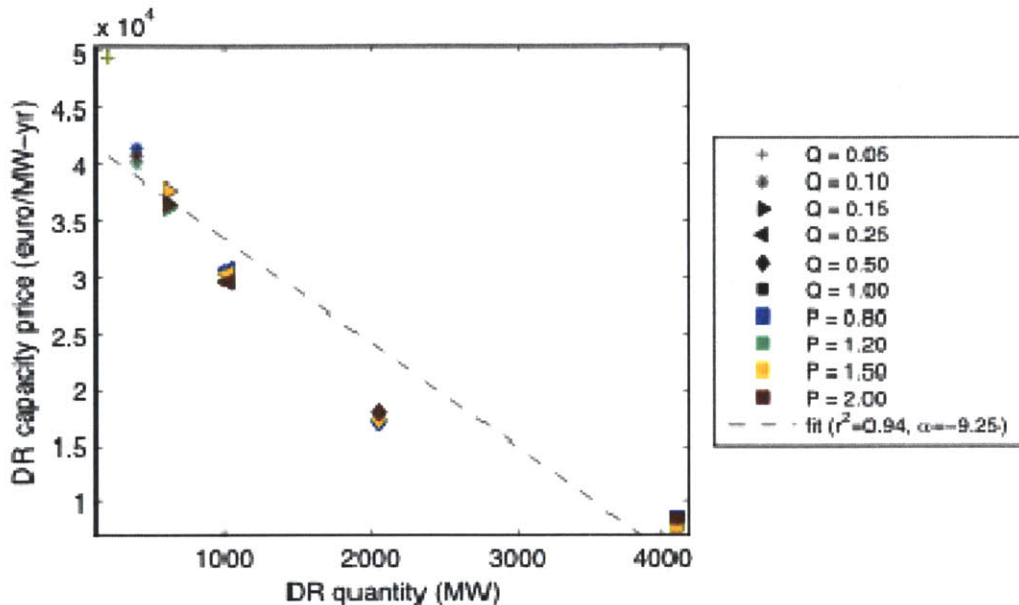


Figure 4-24: System operator's demand curve for DR capacity

The first few units of demand response are of great value to the system operator in terms of reducing overall system cost. Savings are moderate, but energy payments are very low due to lack of actual utilization. The system operator is therefore willing to pay the small number of committed end-use customers a great deal to sign up.

The system will also display an upward-sloping "supply" curve. There will be a small number of customers who are willing to curtail their demand for relatively small incentive payments. As the pool of demand response grows larger, increasingly reluctant customers must be enticed with larger capacity payments. Defining the shape of this supply curve is beyond the scope of this study. With both the supply and demand curves available, the optimal quantity and capacity price for demand response could be calculated.

4.4.2 Effect of Demand Response on Electricity Price

Demand response flattens the wholesale price of electricity. When prices are high, demand response creates a pool of "generation" resources that prevent excessive peak prices. During off-peak hours, demand response enables fewer generators to run at their absolute minimum levels so that prices do not become extremely low as inflexible generators try to avoid shutdowns. Shifting load into the off-peak hours has only a small influence on flattening the price. See Figure E-4 for a comparison of hourly day-ahead prices with load shifting and load clipping demand response; note that prices are still elevated at night even without an increase in load.

In the following figure, the left chart compares the hourly day-ahead price for each hour of the "No DR" simulation with the corresponding hour in the (Q=1.2 P=0.5) simulation. The right chart shows averages for each hour. Time of day is represented in both the left and right charts as color; red is noon and blue is midnight.

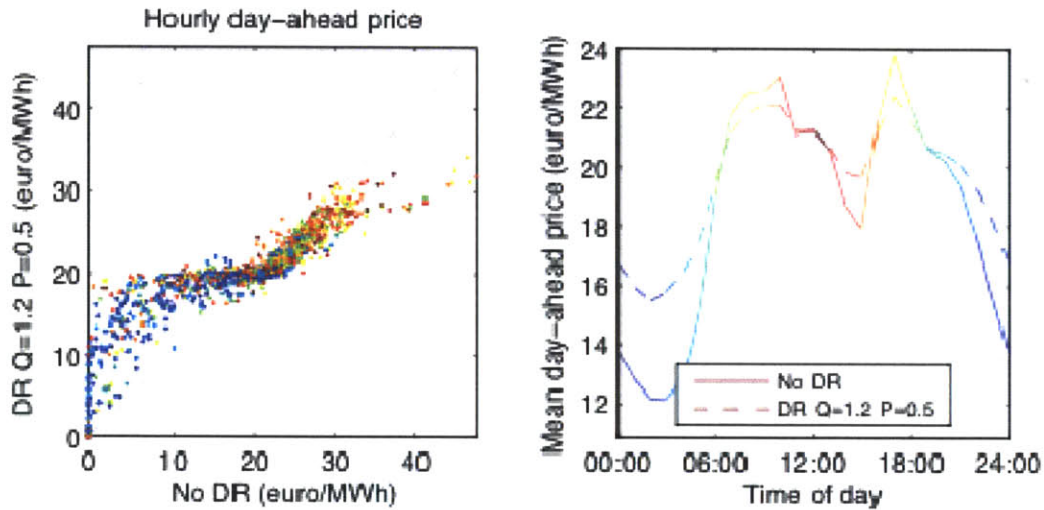


Figure 4-25: Effect of demand response upon day-ahead electricity prices

Demand response has a similar flattening effect on intra-day prices as well. Note the left edge of the left chart. It has a row of zero intra-day prices primarily in the evening hours. These are almost entirely eliminated as demand response leaves fewer thermal units operating at minimum capacity overnight.

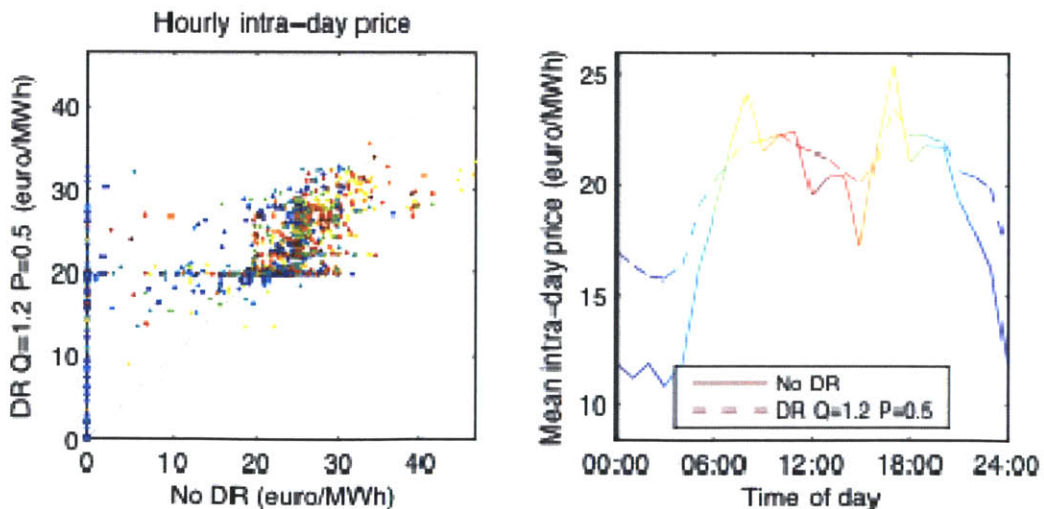


Figure 4-26: Effect of demand response upon intra-day electricity prices

The price behavior of the system has a great impact on the ability of thermal generators to realize profits, especially those that depend upon peak midday prices. This will be shown in detail in a discussion of natural gas generator profitability.

4.4.3 Effect of Demand Response upon Wind Generator Profit

The following analysis compares profit for generator owners across different fuel classes, starting with wind farms. It calculates profit by summing all revenue sources and subtracting cost sources for each individual generator unit. Revenue sources are determined by day-ahead electricity payments, incoming intra-day electricity payments, and ancillary services revenue. Combined heat and power (CHP) units also receive revenue from heat production. Costs are a function of outgoing intra-day electricity payments, fuel cost, operation & maintenance, and startup/shutdown cost. Results are then averaged to obtain profit by generator fuel class. Note that heat-only units are excluded from the analysis. In the following charts, profits and energy production are expressed as a ratio relative to the No DR base case.

The following figure shows the amount of wind energy produced and the profit realized for wind energy producers under each of the demand response scenarios. Figure 4-18 showed that increased quantity and decreased cost of demand response leads to less curtailment of wind. This chart also shows the effect that increased production has upon profits.

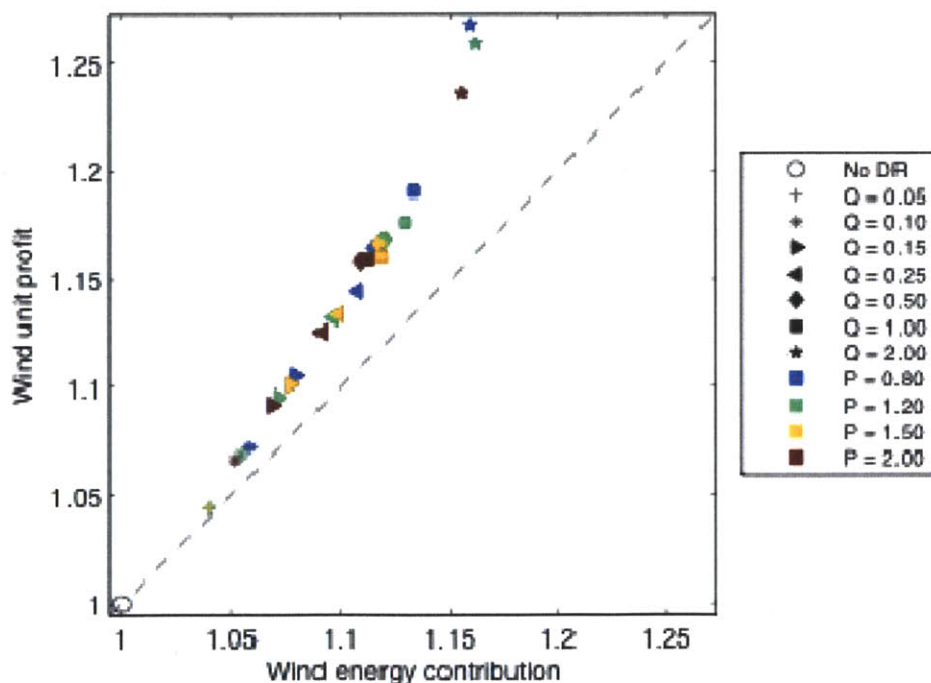


Figure 4-27: Impact of demand response upon wind generator profit and production

Wind plants have no startup cost, no fuel cost, and no operation & maintenance cost in the model. Their profits are simply a function of electricity quantity sold and the price at the time. Demand response increases production of wind power, especially at night. It

also increases energy prices at night through increased efficiency of thermal unit dispatch⁵. Wind generator owners benefit from both effects, as shown by the fact that profit rises faster than quantity as demand response enables greater production.

4.4.4 Effect of Demand Response on Coal Generator Profit

Coal represents the main dispatchable resource on the grid. As wind increases, coal production drops. However, demand response enables coal plants to avoid shutdowns during times of low net demand at night.

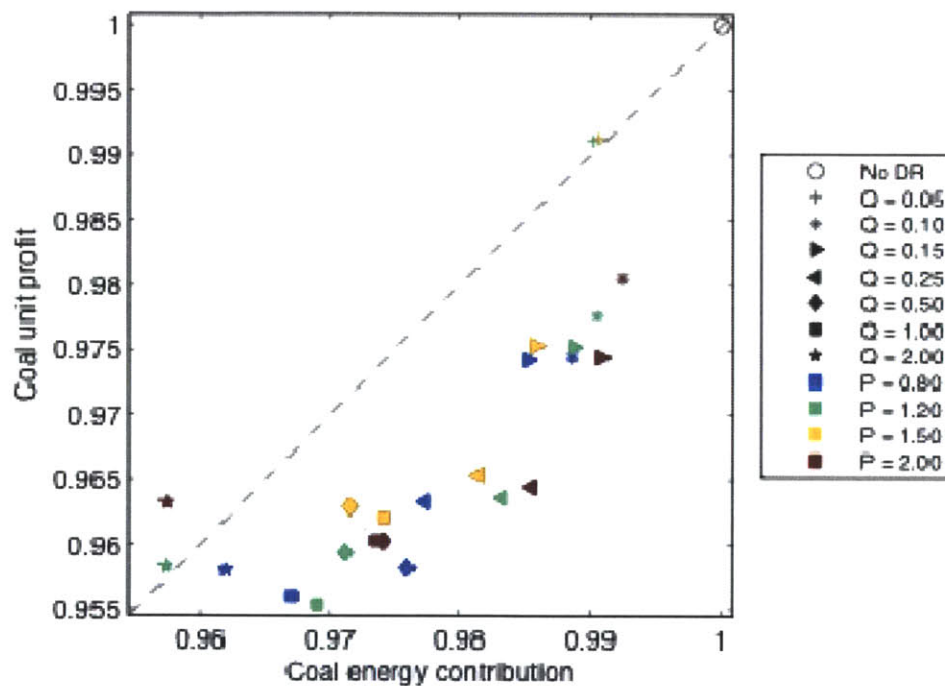


Figure 4-28: Impact of demand response upon coal generator profit and production

Profits for coal operators track production levels from Figure 4-21. For cases for $Q=0.25$ and $Q=0.50$ gross production drops 2.5% while profits drop 4.0%. In most cases, coal generators benefit from reduced start/stop cycling. Coal generators tend to operate 24 hours per day, so the price smoothing effect of demand response provides benefit and harm. As stakeholders, owners of coal generation plants may not gain with demand response, but will not lose much either.

⁵ Note that load clipping (as opposed to load shifting) demand response still elevates prices during off-peak hours. This upward-bending curve should persist across demand response modeling classes. Increased off-peak demand represents a relatively small part of price differences during off-peak hours.

4.4.5 Effect of Demand Response upon Natural Gas Generator Profit

Natural gas and demand response are substitutable for each other in the short run. Natural gas generation has a high marginal cost and is only called when prices are elevated. The profit calculation here takes place only on the day-to-day timescale and does not include the annual capacity payments or ancillary services payments from which peaking unit operators make most of their revenue.

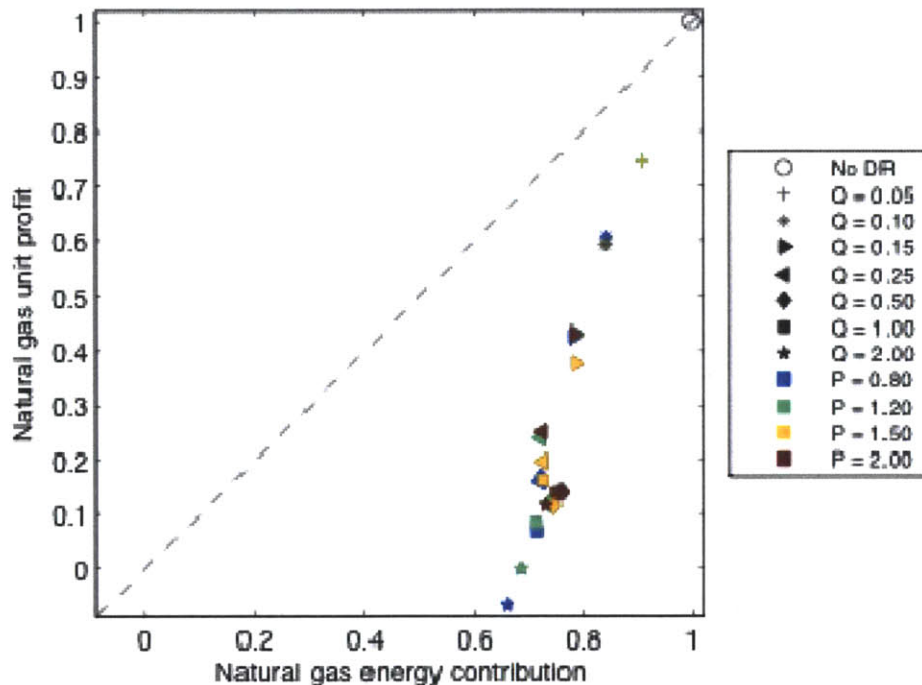


Figure 4-29: Impact of demand response upon natural gas generator profit

Electricity production profits from were never high in the natural gas generator industry even before the introduction of demand response. With demand response, natural gas operators suffer under the double effects of reduced supply calls (Figure 4-22) and reduced peak prices (Figure 4-25). Since gas plants do not tend to operate during off-peak hours, they do not benefit from the price increasing effects of demand response at night. The compounded effects of a 20-40% volume reduction and reduced peak prices reduce profits to near or even below zero.

Generators may suffer economically if demand response creates a more even load profile. Generators may have high prices from peak-price days built into their revenue structure. With these windfall days gone, generators may be required to bid higher prices to preserve profit margins. This situation was validated experimentally in PJM (Holland and Mansur 2005) when demand response reduced profits for oil (59%) and gas (34%) generators. (Ruff 2003) characterizes this as a wealth transfer from generators to consumers and speculates that the generators may take market and lobbying action to preserve rents. In the long term, the presence of demand response on the system may depress investment and lead to less peaking unit capacity. The remaining peaking unit

operators may be able to operate profitably in an environment with less thermal peaking unit competition.

Transmission, distribution, and retail providers may also experience disincentives to support demand response. The 10% of generation infrastructure that serves peak demand is also reflected in transmission and distribution. If demand response causes loads to level, transmission and distribution companies would see a one-time reduction to their required capital expenditures until growth catches up with the amount saved by peak shaving. Since transmission and distribution companies are remunerated primarily by capital expenditures, they may see demand response as a threat to their revenue.

Demand resources may be used longer than the 10-minute runtime of spinning reserve but cannot run indefinitely like a quick-start plant; customers will eventually want to turn their air conditioning back on. The addition of demand response to the grid may make it unprofitable to run these quick-start plants and harm long-term stability. These flexible generation resources must be a continued part of the energy mix, as they have capabilities that demand response does not. As stakeholders, flexible plant operators can be expected to oppose demand response since it harms their direct interests. Demand response is an imperfect substitute for this capacity; their concerns must be addressed.

5 Conclusions

The purpose of this study was to investigate the ability of responsive demand to balance the electrical grid when intermittent renewable resources are present. The WILMAR stochastic unit commitment model was used to represent a version of the Danish electricity and heat system with an enhanced level of wind generation. The study found that demand response reduced system operation cost by enabling less curtailment of wind power. This was accomplished through more efficient use of thermal resources; fewer thermal plants ran at their operating minimums during off-peak hours, which allowed turbines to harvest extra wind instead of shedding it. Greater wind penetration and more efficient thermal unit operation also reduced CO₂ and SO₂ emissions. Thermal units such as coal and natural gas ran more efficiently, with fewer startups/shutdowns and less time spent at minimum operating levels.

The operating metrics were more influenced by the quantity of demand response resources than their pricing. Quantity was expressed as a fraction of nameplate wind power. Demand resources representing 25% of nameplate wind power were a promising recommended quality level. At 10% the grid did not see much effect. At 50% and above the improvements were small but the customer impact is great. A marginal operating price of 150% that of a gas turbine provided a good balance of system improvement and customer impact. The system cost benefits of each study case enabled the calculation of a demand curve representing the system operator's willingness to pay fixed costs for capacity.

The business impact on incumbent power plant operators was mixed. Wind farm operators saw a ~10% increase in output and a ~15% increase in profits. Day-ahead prices rose at night and fell at peak hours. Coal plant operators experienced a slight decline in output and profit. Demand response impacted operators of flexible generation resources such as natural gas especially hard. By substituting for their generation and depressing peak prices, natural gas operators suffered 35% drops in output and over 90% decreases in profit. Flexible generators will be important to the future stability of the grid, as demand response is limited in ways that generators are not. Implementing demand response will require measures to ensure the continued investment in and availability of peaking generators.

This study focused on the producer side of the equation, quantifying the benefits of demand response to system operators and generators. It estimated the impact of demand response on customers, but does not provide tools to value that impact. A follow-up study should use behavioral economics to determine ratepayer willingness to curtail and determine the optimal level of demand response. This abstract model can then drive a political analysis of the practical steps needed to implement robust demand response.

The demand response model studied here was strictly based on a predictable load-shifting model. In reality, demand response is not as controllable as the study suggests. If demand response does not supply as much load reduction as planned, the optimization may need to have additional standby resources ready to supply power. The influence of stochasticity in demand response for unit commitment is a parameter worthy of additional study.

With high levels of unpredictable renewable resources and limited ability to import power, demand response represents a promising technique to balance the grid at low cost.

6 References

- Andersen, Frits Møller, et al. *Analyses of Demand Response in Denmark*. Roskilde: Risø National Laboratory, 2006.
- Barth, R, H Brand, P Meibom, and C Weber. "A Stochastic Unit Commitment Model for the Evaluation of the Impacts of the Integration of Large Amounts of Wind Power." *9th International Conference on Probabilistic Methods Applied to Power Systems*. Stockholm, Sweden, 2006.
- Borenstein, S. "The Long-Run Efficiency of Real-time Electricity Pricing." *The Energy Journal* 26, no. 3 (2005).
- California ISO. *Operational Needs from Demand Response Resources*. Corporate Communication, California ISO, 2007.
- California ISO. "Revised Draft Final Proposal for Participation of Non-Generator Resources in California ISO Ancillary Services Markets." 2010.
- Danish Energy Agency. *Energistatistik 2009*. Copenhagen: Danish Energy Agency, 2010.
- Department of Energy. "Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them." 2006.
- EDF. *Financial Report 2009*. Financial Report, EDF, 2009.
- EIA. *Assumptions to the Annual Energy Outlook 2009*. Energy Information Administration, 2009.
- Ela, Erik, Peter Meibom, Rüdiger Barth, Aidan Tuohy, and Michael Milligan. "Advanced Unit Commitment Strategies for the US Eastern Interconnection." 2011.
- Enernex. *Eastern Wind Integration and Transmission Study*. National Renewable Energy Laboratory, 2010.
- EnerNOC. *Annual Report 2009*. Annual Report, Boston, MA: EnerNOC, 2009.
- Faruqui. *Unlocking the €53 Billion Savings from Smart Meters in the EU*. The Brattle Group, 2009.
- Faruqui, A, and S Sergici. *Household Response to Dynamic Pricing of Electricity: Survey of the Experimental Evidence*. San Francisco: The Brattle Group, 2009.
- FERC. *National Assessment of Demand Response and Advanced Metering*. Docket AD-06-2-00, FERC, 2006.
- FERC. *National Assessment of Demand Response Potential*. Staff Report, FERC, 2009.
- Holland, S, and E Mansur. *The Distributional and Environmental Impacts of Time-Varying Prices in Competitive Electricity Markets*. Working Paper #42, Yale School of Management, Yale School of Management, 2005.
- Holland, S. and E. Mansur. *The Distributional and Environmental Effects of Time-Varying Prices in Competitive Electricity Markets*. Working Paper, Yale School of Management, Yale School of Management, 2005.
- ISO NE. *An Evaluation of the Performance of the Demand Response Programs Implemented by ISO-NE in 2005*. ISO New England, 2005.
- ISO NE. *Technical Feasibility and Value to the Market of Smaller Demand Response Resources Providing Ancillary Services*. ISO New England, 2009.
- ISO-NE. *An Evaluation of the Performance of the Demand Response Programs Implemented by ISO-NE in 2005*. ISO New England, 2005.

Keane, A, et al. "Demand Side Resource Operation on the Irish Power System with High Wind Power Penetration." *Energy Policy*, 2011.

King, C. & D. Delurey. "Energy Efficiency and Demand Response: Twins, Siblings, or Cousins." *Public Utilities Fortnightly*, 2005.

Kofod, C, and M Trogeby. "Demand Response Offered by Households with Direct Electric Heating." *Demand Response in Energy Markets*. Nordic Energy Research, 2004.

Lange, M, and U Focken. *Physical Approach to Short-Term Wind Power Prediction*. Berlin: Springer, 2006.

Lassen, C, and K.L. Jensen. *Estimating Welfare Loss in Controlled Power Interruptions*. Den Kongelige Veterinær - og Landbohøjskole, 2005.

Livengood D, F. C. Sim-Sim, C. Ioakimidis & R. Larson. "Responsive demand in isolated energy systems." *Islands*, 2010.

Morales Gonzales, Juan Migel. *Impact on System Economics and Security of a High Penetration of Wind Power*. Tesis Doctoral, Departamento de Ingeniería Eléctrica, Electrónica, Automática y Comunicaciones, Universidad de Castilla-La Mancha, Ciudad Real: Universidad de Castilla-La Mancha, 2010.

Nordel. *Demand Response in the Nordic Countries, A Background Survey*. Nordel, 2004.

NYISO. "Emergency Demand Response Program Manual." 2008.

NYISO. *Installed Capacity Manual*. NYISO, 2010.

PJM. *Load Management Performance Report 2009-2010*. PJM, 2010.

PNNL. *GridWise Demonstration Project Fast Facts*. Richland, WA: Pacific Northwest National Laboratory, 2007.

Ruff, Larry. "Demand Response: Reality versus "Resource"." *The Electricity Journal* 15, no. 10 (2003).

Söder, L. "Simulation of Wind Speed Forecast Errors for Operation Planning of Multi-Area Power Systems." *Proc. 8th International Conference on Probabilistic Methods Applied to Power Systems*. Ames, 2004.

Sioshansi, R, and W Short. "Evaluating the Impacts of Real-Time Pricing on the Usage of Wind Generation." *IEEE Transactions on Power Systems* 24, no. 2 (2009): 516-524.

Smith, A. *Capacity factors at Danish offshore wind farms*. 2011 йил 02-04.
<http://www.energynumbers.info/capacity-factors-at-danish-offshore-wind-farms>
 (accessed 2011 йил 04-04).

Tom, MG. "File: Demand Response.png." *Wikimedia Commons*. 2006 йил 03-10.
http://commons.wikimedia.org/wiki/File:Demand_response.png (accessed 2011 йил 04-04).

Tuohy, A, Peter Meibom, Eleanor Denny, and Mark O'Malley. "Unit Commitment for Systems With Significant Wind Penetration." *IEEE Transactions on Power Systems* 24, no. 2 (2009).

Walawakar, R, S Blumsack, J Apt, and S Fernands. "An Economic Welfare Analysis of Demand Response in the PJM Electricity Market." *Electricity Policy* 36, no. 10 (2008): 3692-3702.

Wattles, P. "ERCOT DR Overview." *ISO New England 2008 Regional Energy Conference*. 2008.

Appendix A: Wind/Demand Profiles

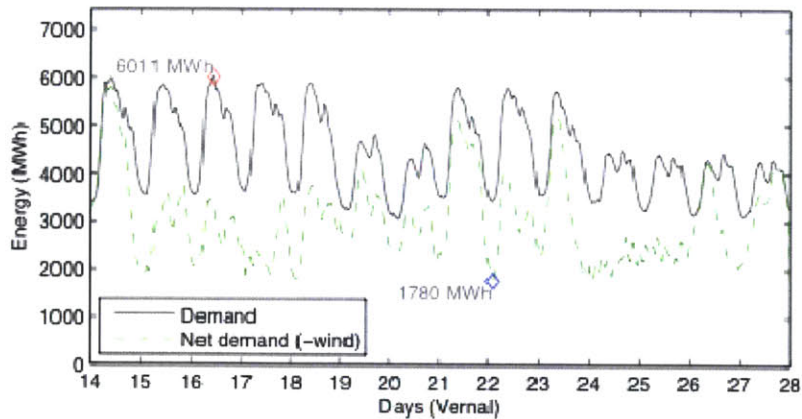


Figure A-1: Wind/demand profile, 2 weeks of spring

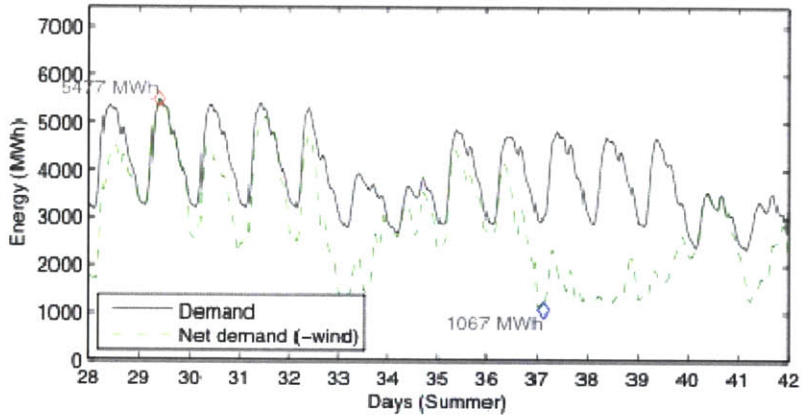


Figure A-2: Wind/demand profile, 2 weeks of summer

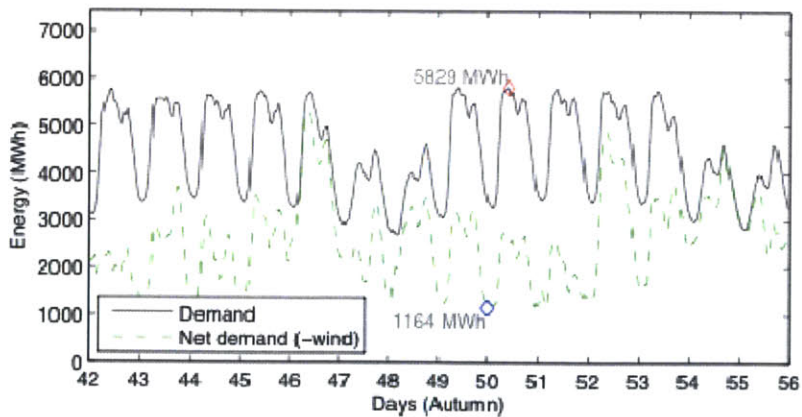


Figure A-3: Wind/demand profile, 2 weeks of autumn

Appendix B: Generation Profiles

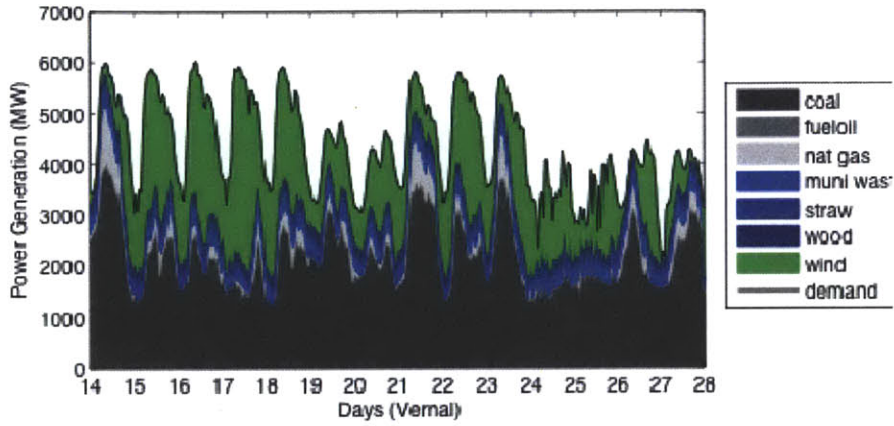


Figure B-1: Spring Generation Profile, Stochastic Baseline

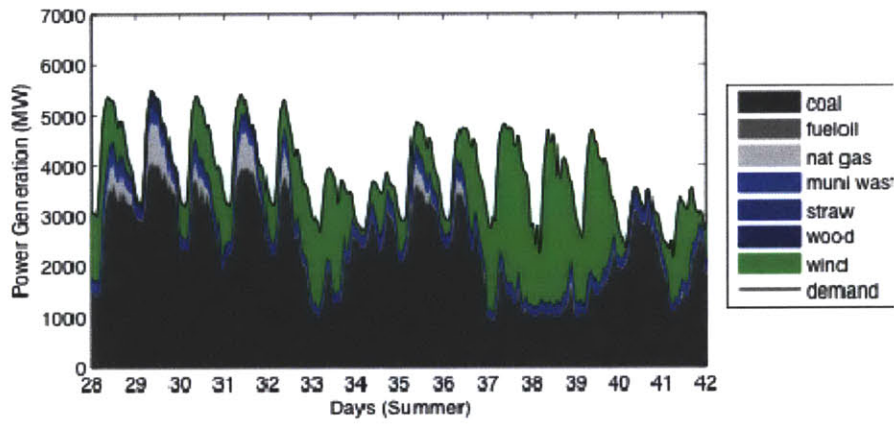


Figure B-2: Summer Generation Profile, Stochastic Baseline

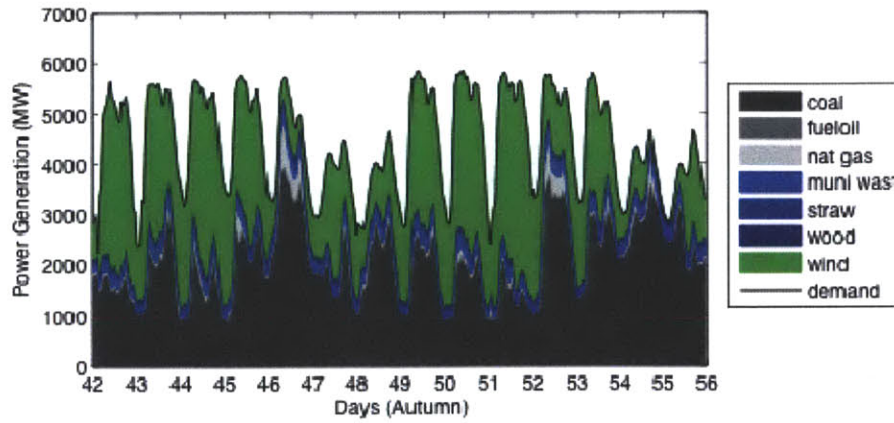


Figure B-3: Fall Generation Profile, Stochastic Baseline

Appendix C: Generator Type Elasticity

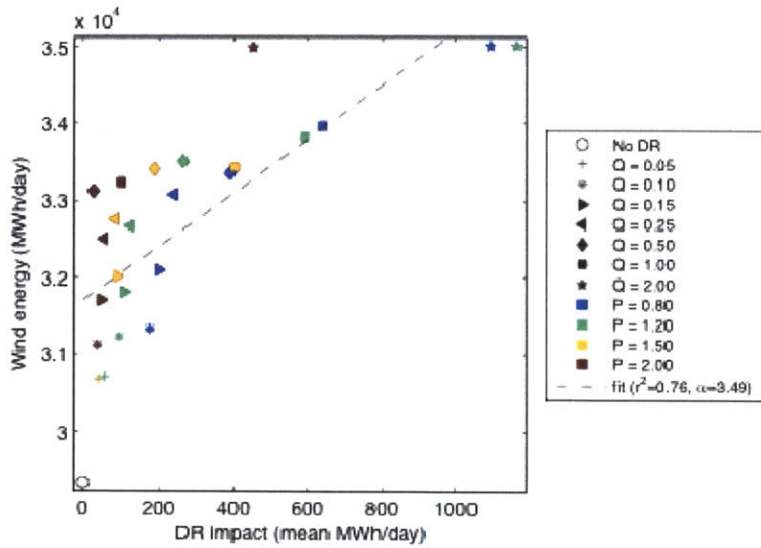


Figure C-1: Demand response and wind energy

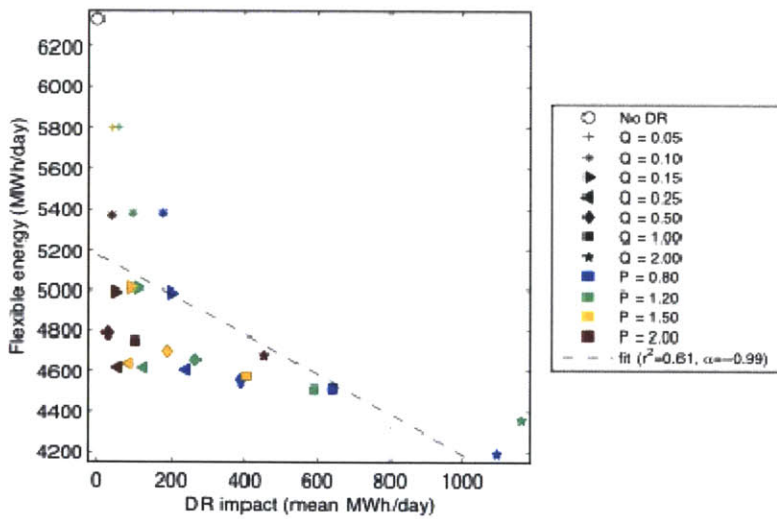


Figure C-2: Demand response and flexible power, substitution elasticity

Appendix D: Startup Frequency by Fuel Type

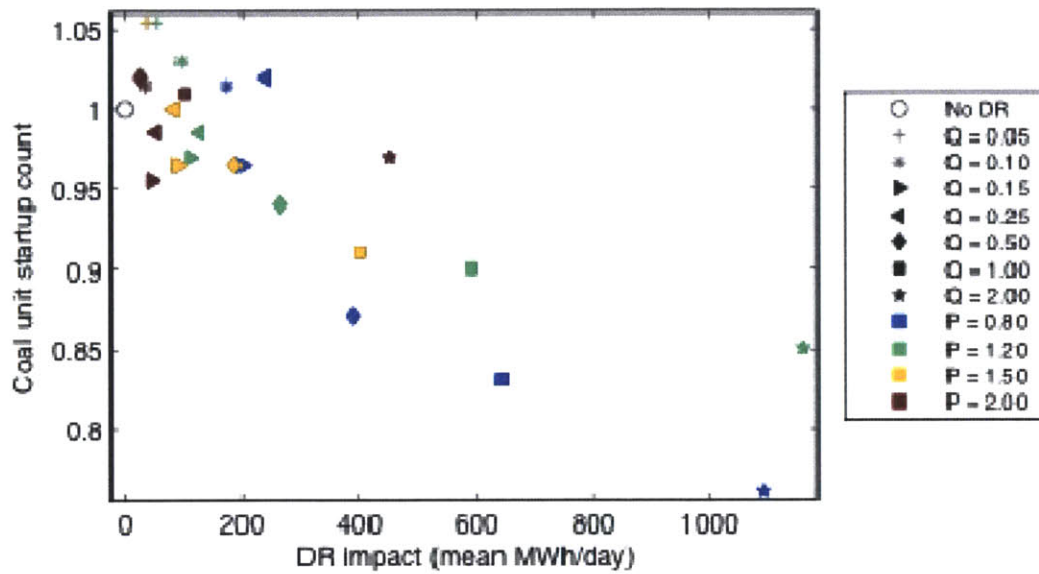


Figure D-1: Coal unit startup counts with demand response

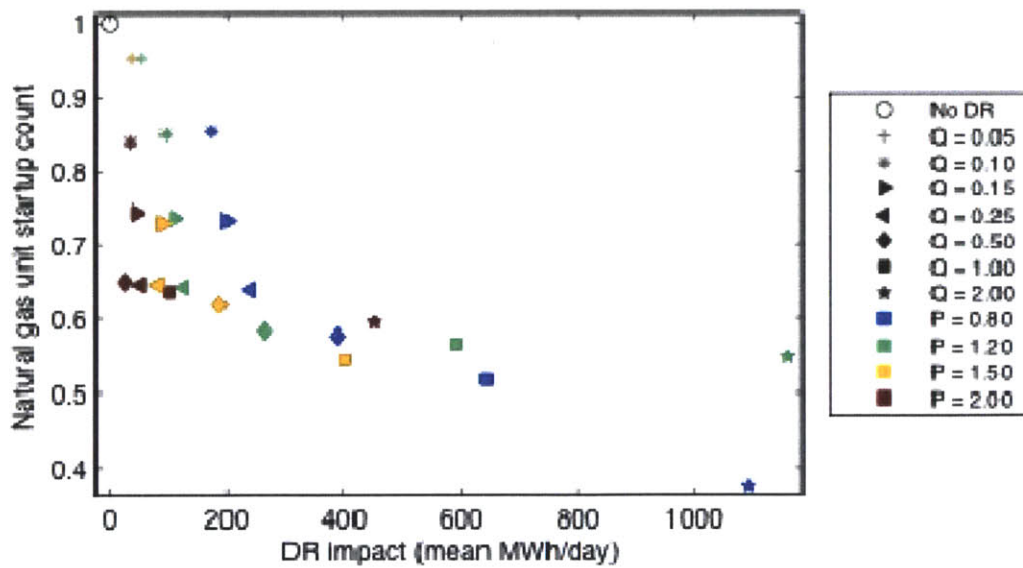


Figure D-2: Natural gas unit startup counts with demand response

Appendix E: Demand Response Modeled as Load Clipping

The experiments described in section 4.2 all use demand response as a load shifting resource (see section 2.1.3 for more information). Demand response may also be modeled as a load clipping resource, as described in section 2.1.2.

For the purposes of comparison, a set of simulations were run with demand response modeled as a clipping rather than a shifting resource. No external constraints were placed upon duration or frequency of calls. Andersen et al. (2006) modeled both load clipping and load shifting demand response in the same simulation, recommending that load clipping DR have a higher price. This comparison investigates a $Q=0.5$ case with $P=1.2$ for load shifting DR and $P=1.7$ for load clipping DR.

Plotting the generation mix with load clipping DR, there is no "rebound" effect. Demand response is primarily used during times of peak demand.

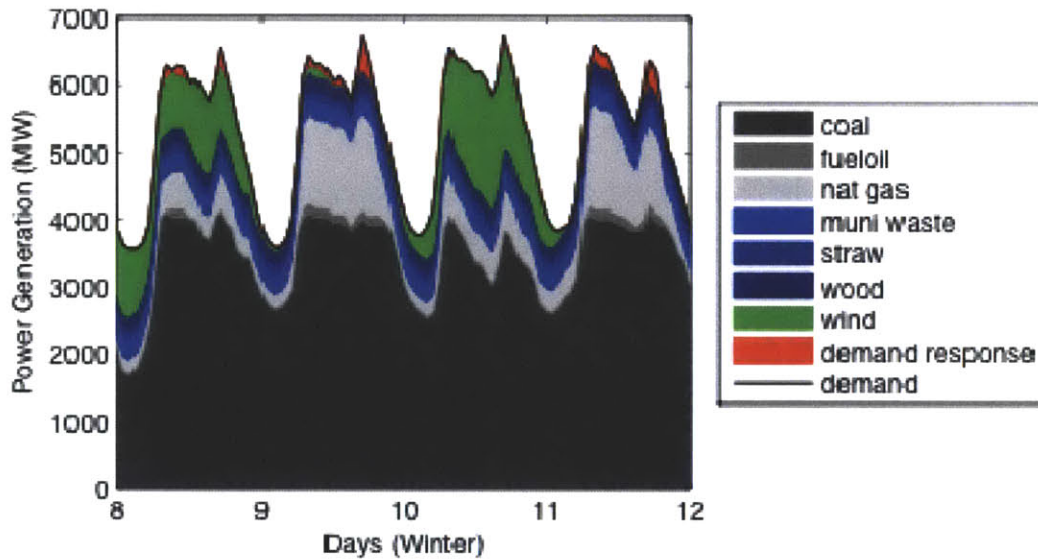


Figure E-1: Unit commitment with peak clipping

This chart plots average energy dispatch by hour. This verifies that load clipping is used more during midday peak load and less at night.

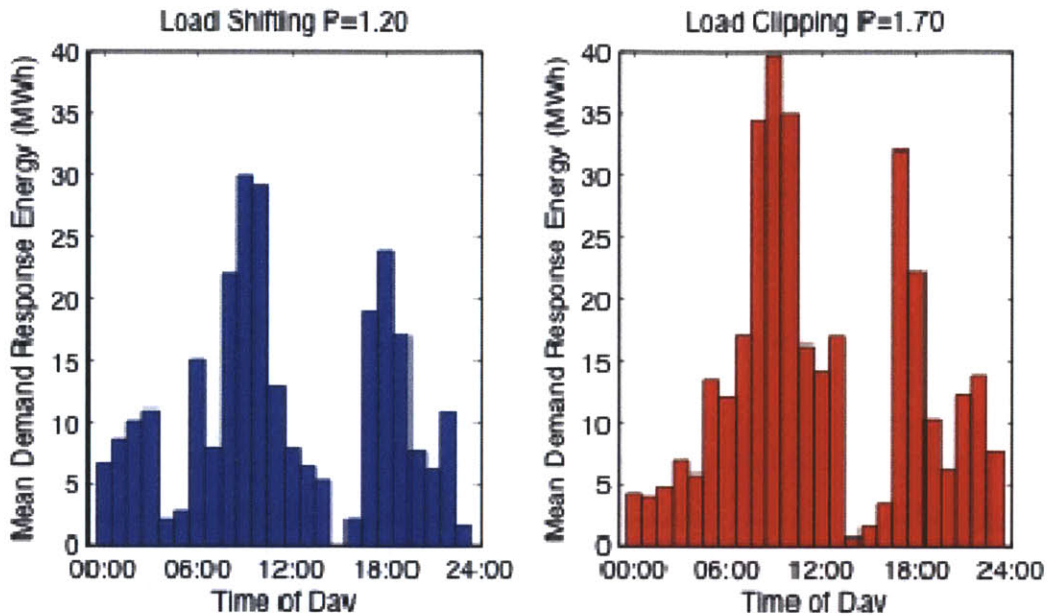


Figure E-2: Load shifting vs load clipping by hour

Rather than averaging by hour, the following chart plots comparable per-day averages. Each circle represents a single day within the study period. The left chart displays the energy intensity of demand calls. The right chart shows the duration of those same calls. Dashed lines are averages.

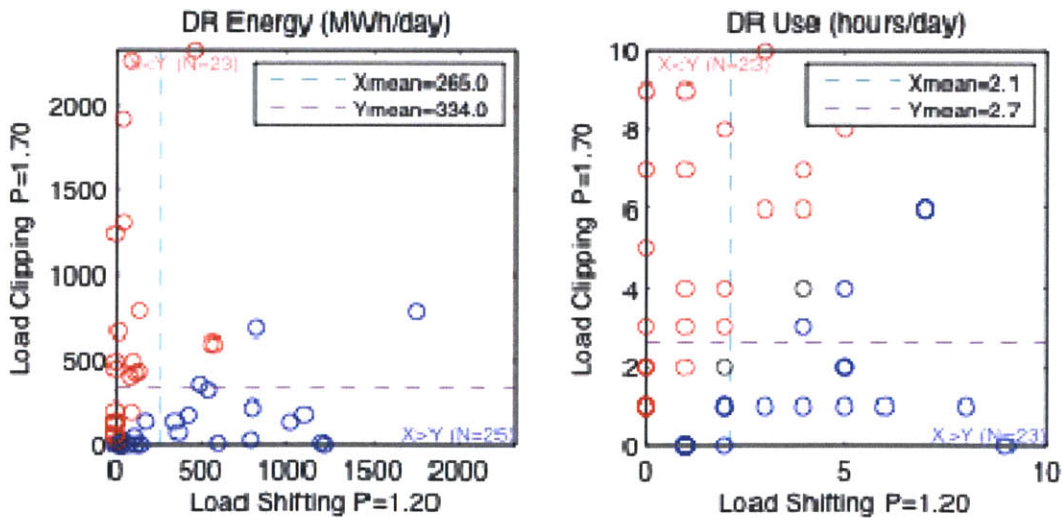


Figure E-3: Load shifting vs. load clipping by day

Peak clipping curtails 26% more energy each day, and events are 29% longer. The same days with high load shifting intensity have low load clipping intensity. Conversely, days with high load clipping intensity use low amounts of load shifting. There are no days on

which both both demand response methods are used with great intensity. The optimization may be using the resources for different ends. It is possible that both demand response methods may coexist in the same system and address different problems.

Other demand response prices were simulated, but not investigated in depth. With $P=2.0$, load clipping demand response is only called one hour per day on 3 separate days. This is so infrequent that it is not compelling to study. With $P=1.2$, load clipping demand response is called an average of 13.4 hours per day and is active for 53 of the 56 simulation days. This is so frequent that it violates the duration restrictions and is not valid to study.

Demand response as load clipping has a similar effect on day-ahead prices as load shifting. This shows that the increased nighttime prices with load shifting DR are more an effect of improved thermal unit utilization and less an effect of increased off-peak demand. The minimum price with load shifting is about 16€/MWh, the minimum price with load clipping is about 15€/MWh, and the minimum price with no demand response at all is about 12€/MWh.

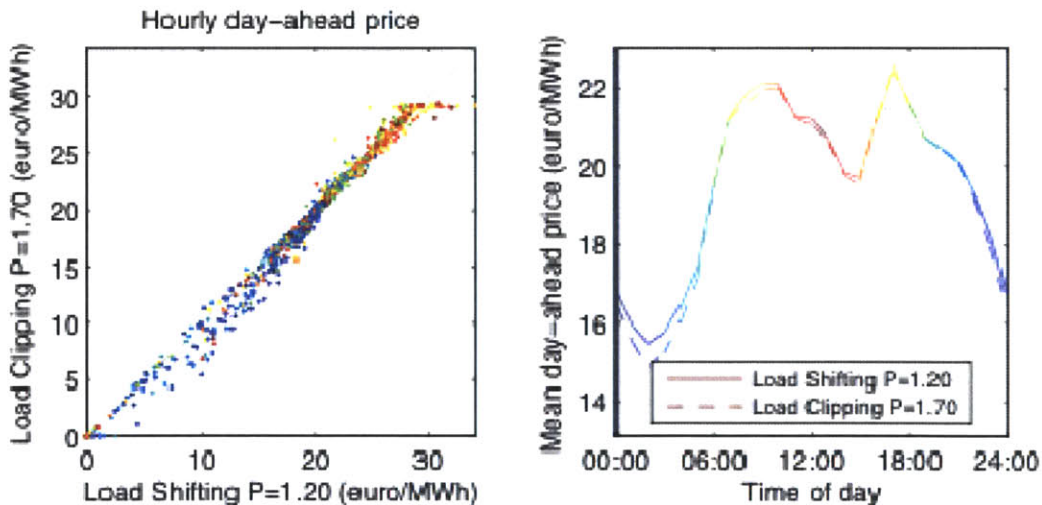


Figure E-4: Effect of peak shifting/ clipping DR on electricity price

Appendix F: The Value of Perfect Information

WILMAR's stochastic unit commitment model optimizes response for 12 possible future scenarios at each 3-hour rolling decision period. Generation choices are robust to forecast error, but are rarely optimal for the realized wind level. With perfect information, it is necessary only to optimize the system for the known future. This section compares the portfolio of demand response cases in the stochastic and perfect-information deterministic operating modes.

Criterion	Unit	Value, Stochastic, no DR	Value, Perfect Info, no DR	Value (DR), Stochastic, (Q=0.25 P=1.5)
Cost	€/MWh	27.97	27.57	27.10
CO ₂ emissions	kg/MWh	166.2	163.68	160.2
SO ₂ emissions	kg/MWh	1.09	1.07	1.06
Wind energy	fraction	.28	.30	.31
Coal energy	fraction	.55	.54	.54
Natural gas energy	fraction	.05	.05	.04
Demand response impact	MWh/day	N/A	N/A	85

Table F-1: System operation evaluation criteria, with perfect information

Perfect information leads to system improvements for cost and emissions, but demand response shows even better results. Comparing the demand response portfolio in both the perfect-information and stochastic case, it is seen that system performance approaches equivalence as more demand response is added. The 1.5% cost advantage of perfect information with no demand response shrinks to less than 0.1% with ample demand response.

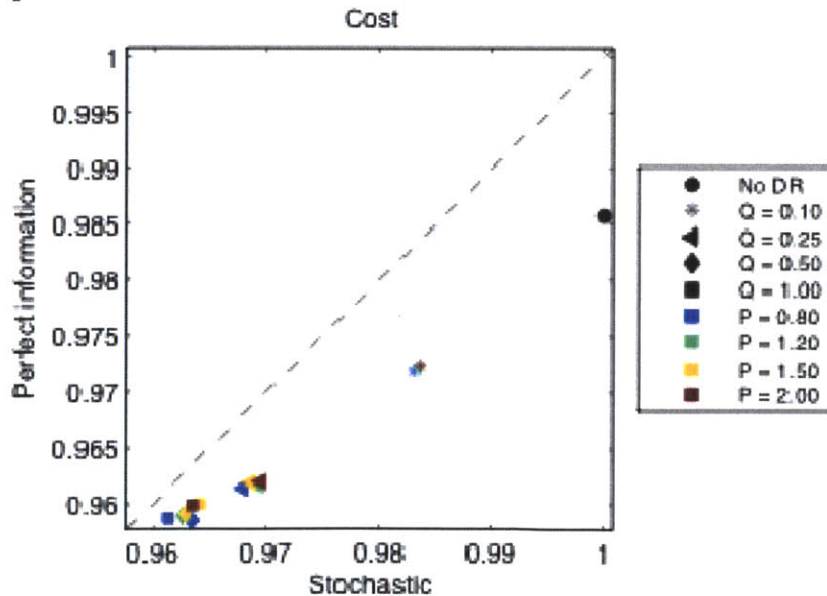


Figure F-1: System cost with and without perfect information

Similarly wind penetration is higher in all cases with perfect information, but the advantage narrows with greater quantities of demand response:

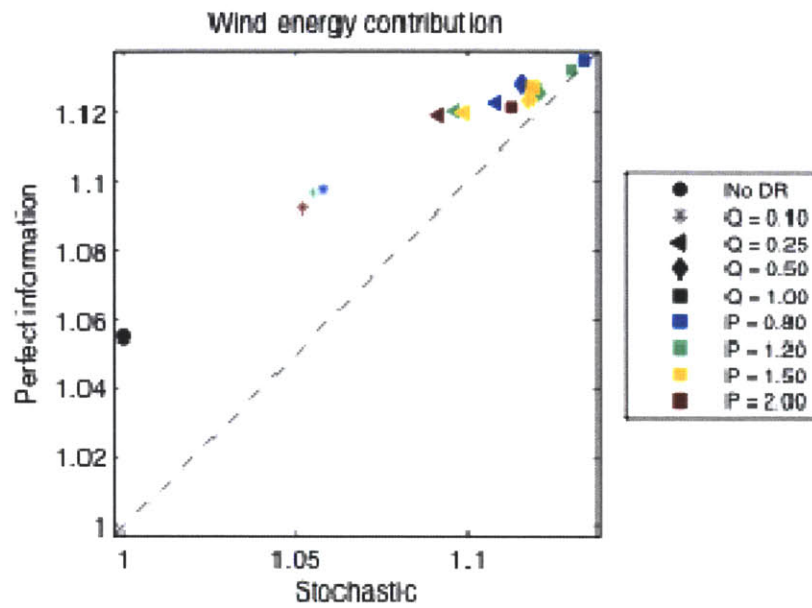


Figure F-2: Wind energy with and without perfect information