电力规划设计总院
China Electric Power Planning & Engineering Institute

Danish Energy Agency

3 Region Report

on flexibility measures for system integration of variable renewable energy

U.S.

Europe

China

Acknowledgements

This publication was produced by the Danish Energy Agency (DEA) and the Electric Power Planning and Engineering Institute (EPPEI) under the China Thermal Power Transition Program financed by Children's Investment Fund Foundation (CIFF). Special thanks to the Rocky Mountain Institute (RMI), COWI, and Ea Energy Analyses (Ea) for their chapter contributions and to the reviewers Regulatory Assistance Project (RAP) and Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ).

Contacts:

Sharissa Funk, Danish Energy Agency

ens@ens.dk

Shunchao Wang, Electric Power Planning and Engineering Institute

scwang@eppei.com

Table of Contents

List of Acronyms and Abbreviations

Part I USA

AS	Ancillary services
BAA	Balancing Authority Areas
CAISO	California Independent System Operator
CCNG	Combined cycle natural gas
CPUC	California Public Utilities Commission
DAM	Day-ahead markets
DER	Distributed energy resource
DR	Demand response
DRAM	Demand response auction mechanism
EIM	Energy Imbalance Market
ERCOT	Texas Independent Power Market System Operator
EV	Electric Vehicle
FERC	Federal Energy Regulatory Commission
FFR	Fast Frequency Response
FRP	Flexible ramping product
HVAC	Heating, ventilation, and air conditioning
IOU	Investor-owned utility
IPP	Independent Power Producer
ISO	Independent System Operator
MISO	Midcontinent Independent System Operator
NERC	North American Electric Reliability Corporation
NYISO	New York Independent System Operator
PDR	Proxy demand response product
PFR	Primary Frequency Response
PG&E	Pacific Gas and Electric, a utility
PJM	PJM Interconnection, regional transmission organisation in Eastern U.S.

PUC	Public Utility Commission
RDRR	Reliability Demand Response Resource
RoCoF	Rate of change in frequency
RRS	Responsive reserve service
RTO	Regional Transmission Organization
SDG&E	San Diego GAs and Electric, a utility
SIR	Synchronous inertial response
VRE	Variable renewable energy

Part II **Europe**

AC	Alternating current
BRP	Balancing Responsible Party
CHP	Combined Heat and Power
DAM	Day-ahead market
DC	Direct current
ENTSO-E	European Network of Transmission System Operators for Electricity
EU	European Union
IDM	Intraday market
IEA	International Energy Agency
MAE	Mean Absolute Error
m-FRR	Manual Frequency Restoration Reserve
NHPC	Net heat production costs
NOIS	List of bids in the Nordic regulating power market
NWE	North West Europe
TSO	Transmission System Operator
VRE	Variable renewable energy
XBID	European cross-border intraday market platform

Part III **China**

AS	Ancillary services
CHP	Combined Heat and Power
DEA	Danish Energy Agency
EPPEI	Electric Power Planing & Engineering Institute
LP	Low pressure
NEA	National Energy Agency
VRE	Variable renewable energy

List of Figures

List of Tables

Executive summary

Thanks to rapid expansion of renewable energy China has the largest installed capacity of variable renewable energy already today. Following the path towards a clean, low-carbon, safe and efficient Chinese energy system (13th Five-Year plan) this development will only be accelerated in the coming years. This development has also introduced new challenges on how to reliably operate the power system while maximizing renewable integration, particularly the need to increase system flexibility to effectively integrate variable renewable production.

High renewable energy regions in the U.S. and Europe have experience with similar challenges and can provide relevant examples of how these have been adressed. This report provides specific experiences from high renewable energy regions in the U.S., Europe, and China, that offer examples on different sources and measures that deliver or facilitate much flexibility in the system. Examples include improved flexibility of production units, demand flexibility, system flexibility and sector coupling, as well as the market set-up and regulation incentivizing and enabling more flexibility.

Based on these examples, the following key messages on flexibility measures for enhancing system integration of VRE in China can be extracted:

Main messages

- Flexibility is not a goal in itself.
- Reducing need for flexibility through e.g. better forecasting and enlargened balancing areas is most cost effective solution to VRE integration.
- In high VRE systems flexibility adequacy should be assessed in the planning proces.
- (Short term) power markets are essential for providing price signals for flexibility, the European intraday market has proven an effective market for adjusting day-ahead schedules to changes in production patterns especially for VRE.
- In Europe, the role of the balancing responsible party has further ensured that market parties are responsible for being in balance and finding the optimal market for trading their imbalances. This takes away burden and costs from the system operator but also makes the market more accessible for smaller generation units (decentral production).
- Ancillary service markets are in the spotlight when it comes to defining products that reflect the specific flexibility need of the power system in question. Correctly designed products meeting the specific needs can reduce power system costs substantially by reducing the need for large reserves.
- Integrating markets into larger market areas has significant economic benefits which can be achieved even without full harmonisation.
- Demand side response can play an important role in shifting load away from peak hours and therefore reducing need for expensive peak capacity.
- A complete transition to EVs cannot be supported by time-of-use tariffs but calls for more dynamic pricing schemes that ensure efficient and smart charging of vehicles.
- Successful technologies in providing flexibility have been flexible CHP, (industrial) demand response, power-to-heat, and hydro power. However, which technologies are succesful depends on many aspects than can differ between regions and will change over time. Therefore market price signals are the best way to ensure that the most effective technologies for providing flexibility are activated.

In the following, a summary of the findings of the report and their relevance for enhancing system integration of VRE in China is provided.

1 Fundamentals of enhancing power system flexibility

The rise of wind and solar power gives unprecedented importance to the enhancing of flexibility in electric power systems. Many different approaches and measures to increasing flexibility can be taken, therefore some fundamentals of power system flexibility should be highlighted to begin with:

- Flexibility is not free. Cost-effectiveness shall always lie at the heart of any technological innovation, policy reorientation and market design for enhancing system flexibility. Reducing the need of flexibility by means such as improving generation forecast and enlarging balancing areas should always be prioritized.
- It is important for decision-makers to focus on fundamental and proven measures first, such as improving the flexibility of thermal power plants and integrating adjacent power systems. These measures can provide the largest impact and are important to making future investments in flexibility more valuable.
- Beyond the fundamental measures, new options of flexibility enhancement should be further explored with caution, particularly when it comes to new, dedicated investments in flexibility assets, such as batteries. Flexibility is often given more attention than is necessary for the levels of VREs that grids are experiencing, as many grids have been able to integrate 20%+ of VREs with little new investments.
- Key characteristics of flexibility include ramp speed, ramp mileage, response time, and dispatch accuracy. Different power systems typically have different flexibility needs. Identifying the need of flexibility based on these four metrics could therefore be an important first step of enhancing power system flexibility.
- A holistic and pragmatic approach is needed. Introducing new market products and renewing market designs are effective measures which however need to be adapted to and take account of the circumstances and current environment in which they are to be applied. Delicate handling of existing issues can be required in the process. The power sector is a part of a wider energy system, looking at cross-sectoral synergies can provide very effective flexibility sources.

2 Experiences of Western U.S. and their relevance for China

The U.S. system is considered to be fragmented, with states and regions taking their own approach to electricity regulation and market operation. Among those states and regions, California is one of the frontrunners of actively promoting market reform to enhance power system flexibility. There are many of the strongest cases in the sphere of system flexibility occurring in California, particularly in regard to minimizing or procuring resources to manage inter-hour ramp, establishing a wider balancing area and also tapping into the potentials of demand side response and smart integration of EVs. Many of these efforts have strong relevance to the initiatives in China for greater system flexibility.

- Flexibility adequacy analysis in the planning period

Traditionally, adequacy analysis mostly concerns whether there is enough capacity to sustain the power supply throughout a year, and it is usually done in the planning stage. In the U.S., the adequacy analysis for utilities is carried out in the integrated resource planning (IRP) process. In recent years, flexibility adequacy is also included in the IRP of many utilities as a major component. Although this procedure is mainly a legacy of the old regulated institution, when it comes to the planning of investment with high externalities (such as many investments in flexibility), the effectiveness of it does not fade away in the new context of competitive power market. China regularly carries out 5-year plan for power system on the national level and included flexibility analysis in the last 5-year plan study for the period of 2016-2020. If the flexibility adequacy planning in China can permeate into the provincial and even municipal level, the effectiveness of flexibility analysis could be further improved.

- Integrating heterogeneous power markets

One of challenges for the integration of adjacent power markets is their heterogeneity in design and the entailing path dependence. Different power market setups result in different dispatching rules, different operation routines and different supporting IT systems. EIM in the Western U.S. is one of the prominent examples of integrating distinctive power markets successfully while limiting the need for harmonisation. EIM is a voluntary real-time energy-only market in which CAISO and surrounding utilities share their resources to find the most cost-efficient means to meet their demand for balancing energy in real time. EIM market draws on the learnings from the past efforts to merge the Western U.S. into one balancing area, which was proven to be futile, and takes a softer approach where key concerns of different stakeholders are respected. Particularly, it allows different balancing authorities (which are similar to local dispatching centres in China) to retain most of their control of scheduling. A key success factor was to place particular emphasis on optimizing software and processes to allow utilities to share information, reschedule, and coordinate dispatch as close to real-time as possible.

Many provinces in China are promoting spot markets. Although there exist some high-level overarching guidelines for provincial spot market design, the distinctive realities in different provincial power systems will inevitably lead to differences in provincial power market design. EIM provides an example for provinces in China which seek cross-provincial integration among different markets with distinctive designs. EIM has already showcased how short-term market integration, i.e. cross-jurisdiction electricity trading close to real time, could create significant economic benefits without bringing fundamental changes to the existing responsibilities. Therefore, more importance should be attached to the short-term market design in the future cross-provincial trading initiatives.

- Ancillary service markets tailored for high renewable energy shares

Since spot markets oftentimes are used as a reference point for long-term investment signals, the rules and regulations of the spot market should be kept as consistent and stable as possible. Ancillary service markets, on the other hand, provide the necessary agilities for policy makers to handle transient problems. California and Texas are looking to the future to identify new ancillary services products which can address their specific flexibility needs. In California, a flexible ramping product was designed to address inter-hour ramping needs. It improved scheduling of generators to ensure sufficient ramping capacity is available, without eating away regulation services needed to maintain system reliability. Texas tested a fast frequency response product which was designed to acquire enough synthetic inertial response to support a high wind system response to a contingency event.

Many regions and provinces in China has been actively promoting ancillary service market reform. In some cases, when the problem is very acute, the introduction of new ancillary service products could provide necessary economic incentives for new investment. Most notably, the down-regulation ancillary service market in Northeast China and frequency regulation ancillary service market in Shanxi province, both have been proved to be very effective in handling the pressing local problems. As solar power generation continues to increase at a fast pace, many provinces might be faced with a shortage of ramping up capacities in the near future. The ancillary service market in California with specific products for interhour ramp is one of the possible models to emulate so as to extract more flexibility from existing assets to support renewable integration.

- Tapping into the potential of demand side response

Using demand reponse as a main tool to balance the power system is not a new tool, and traditionally it is the utility companies that are the main buyers of demand response services. Utility companies procure demand response not only because of their balancing responsibility, but also because the economic benefits coming from reducing the demand at peak time in which the wholesale power price is usually very high. However, if the aim is to further scale up demand response, the utility-as-the-main-buyer model might become insufficient. The pioneering efforts in California have moved utility demand response programs to a market paradigm, in which demand response could participate in the wholesale market directly. But it was also observed that revenue from the wholesale market alone would not be enough for encouraging new investment needed for providing sufficient demand response services. Thus, California introduced a new program to further compensate demand response services by translating the value of demand response as a firm capacity to a new revenue of demand response aggregators (capacity payment). This approach has proven to be very successful in scaling up demand response in California.

In China, the need for demand response become more and more prominent. The fast urbanization process results in the fast growth of air-conditioning load. The simultaneous turn-on of air conditioners in early evenings of hot summer leads to a sharp peak demand. Currently, a peak-valley pricing system is used as the main tool to alleviate the peak demand. A few provinces are now piloting demand response programs in which grid company functions as the buyers. But still these methods are far from sufficient. The shortage of power in hot summer in some eastern provinces continues to grow. As China gradually introduces competitive wholesale market on the provincial level, including demand response in the wholesale market becomes a new option. However, it should be analysed whether the energy value of demand response, which is reflected in the wholesale market, is sufficient to sustain demand response business models or whether extra market programs, such as the DRAM program in California, might be needed.

- Managing EV integration for flexibility purpose

About half of the world's EVs are in China. China's system operators are now also facing both the challenges and opportunities brought by EVs to the power grid. On the one hand, uncoordinated EV charging would lead to unacceptable peak demand for utility companies. On the other hand, EVs could function as flexible loads to consume electricity surplus generated by wind and solar power. The most commonly used model for coordinated charging in China is the time-of-use rates model. But it is foreseeable that the time-of-use model would not be sufficient as the number of EVs further grow, as it is likely that a new peak demand will be induced at the low price moment. Shifting from the time-of-use model to a more interactive smart charging model, in which EVs charge according to the real-time needs and constraints of the power system, will be a necessity in the near future. Pilot programs in California have already proven the feasibility of using EVs as new flexibility resources. However, these programs are still operated at a relatively small scale. How to scale up the EV smart charging programs to a multi MW level still needs further explorations.

3 Experiences of European countries and their relevance for China

- Balance responsible parties as the intermediate layer for flexibility management

In the liberalized and decentralized European power market, system operators delegate financial balancing responsibility to market participants. China is now undergoing a new round of power market reform. One of the goals of this round of market reform is to build a competitive wholesale market. As for the dispatching model and balancing responsibilities, there seems to be no momentum toward decentralizing the current highly centralized model, as the centralised dispatching model in China has demonstrated its effectiveness on maintaining the stability and reliability of the national power system. Although the blunt copy of the balancing responsible party model to China is not necessarily realistic, this model does shed some light on possible new business models to cope with new trends in China's power sector. The power system is now gradually decentralized as distributed variable renewable energy generation (especially distributed PVs) continue to grow at a relatively fast pace. System operators face new challenges of keeping the system balance, as balancing of small units in remote areas of the grid is not only technically challenging but also administratively burdensome. Thus, introducing an intermediary layer between the system operator and distributed generation to aggregate small production units would be a possible solution to reduce the work load of system operators and also keep the operation of power system streamlined.

- Intraday market as an important remedy for wind and solar forecast error

The intrinsic uncertainties of weather render the forecast error of wind and solar power inevitable. The direct outcome from these forecast errors is the extra need for reserve capacity, contributing to the so-called system cost of integration of variable renewable energy. The accuracy wind and solar forecast of looking only a few hours ahead (usually referred to as ultra-short-term-forecast), compared to looking 24 or more hours ahead at the moment of scheduling, improves significantly. Thus, to reduce the need of reserves and the cost accordingly, there is a need to allow market participants to adjust their schedule compared to the day-ahead market. Fast reserves (e.g. activating batteries, ramping up/down electric boiler, etc.) are the most expensive ones, so the closer to physical delivery the adjustment is done, the better. Intraday markets in Europe provide market participants the opportunity to trade up to 30 minutes before physical delivery and thereby make adjustments to their day-ahead schedule. It should be noted that, even the two-settlement model (day-ahead + real-time) widely used in the U.S. does not contains an intra-day market explicitly, system operators still have various intra-day mechanisms, such as hourly residual unit commit in ERCOT and short-term unit commitment in CAISO. The advantage of an explicit intra-day market is that a liquid market can more cost effectively re-schedule the system, thereby also reducing the burden on variable renewable energy producers. A visible price signal on the intra-day market is is also important for putting a value on flexibility and thus guiding investments in flexibility assets.

China has launched several pilot projects for spot markets. Currently, the focus is mainly on the design of day-ahead market and real-time market/dispatch. The intra-day mechanisms are seldomly discussed. For provinces with high variable renewable penetration, the volume of intraday adjustment would be quite high, and even might challenge a compact system operator unit (dispatching centre) where flexibility resources are unavailable. To resolve this, one of the options is to introduce intraday trading. Compared to the real-time market, the intra-day market leaves a relatively longer time period for power plants to adjust the generation, which can prove an advantage for thermal power plants due to their large inertia. Intraday markets could provide a price signal for flexibility from traditional thermal power plants, where spot markets might not provide sufficient incentives for investment in flexibility for thermal power plants in China. Including an intraday market in the spot market could provide a necessary extra nudge.

- Operating hydro power in a deregulated power market

The deregulation of the power market in Nordic regions also led to the deregulation of hydro power stations. In this deregulated and liberalized market context, each of the power stations takes dispatching decisions on their own. This is in sharp contrast with the traditional centralized paradigm, in which all the generation programs are determined by one central dispatching centre. Although all the decisions are individually made, they collectively lead to a very efficient power system, as evidenced by the high level utilisation of renewable energy in this region.

Unlike wind and solar power, the marginal cost of hydro power plant with a reservoir is not zero, it equals to the opportunity cost of generating now compared to withholding the water for later generation. One of the key concepts in the decision making processes is the water value, which reflects the timely value of a unit of water in the reservoir. In a nutshell, the water value would be higher when there is less water in the reservoir, or a higher price expected in the future, and vice versa. When the current market price is higher than the water value calculated, the power stations will generate. Thus the revenue of a hydro power station largely depends on the accuracy of price forecasts, weather forecasts, etc.

Practices in Nordic regions shed some lights on the on-going power market reform in Southwest provinces, such as Yunnan and Sichuan, which also have abundant hydro power and increasingly more wind power plants. In a future Chinese market area with integrated regional markets, the Southern regions could develop a similar role as the Nordics as the “battery” of China. As for the hydro power station owners, the capability building on price forecasts on various time scales (hours ahead, days ahead, months ahead, even a year ahead) would be crucial for their financial performance.

- Unleashing the flexibility potential of industrial load

One of the observations from the development of demand-side management in Finland is that some of the industrial loads are more competitive than thermal and hydro power plants in providing fast frequency reserves to the system. As for industrial processes participating DSM, the cost is mainly opportunity cost, which is the industrial production reduction caused by reducing electricity consumption. Industry participates in demand-side management if the benefits of engaging in demand-side management outweighs the opportunity cost of reducing electrity consumption. For some industrial processes, a short term reduction in electricity consumption does not significantly affect their production. Industrial processes with high physical inertia, which can maintain stable industrial production when there is a short-term decrease of the electricity consumption, have a particularly large potential to provide short-term flexibility to the grid.

About 2/3 of electricity in China is consumed by industrial loads. Some of the industrial loads (e.g. electrolysis, chemical industry, etc.) could provide short-term responses without causing interruption of the industrial production. The potential of using industrial load for demand response is largely untapped mainly due to the inadequate market mechanisms. As for the North and West regions, energy-intensive industry coincides with the variable renewable generations. The reform of ancillary service market in these regions should take into account of the flexibility potential of industrial loads.

- Realizing manifold benefits of power-to-heat facilities

Sector-coupling is the new watchword in the field of system flexibility. The energy flow among electricity sector, heat sector and transportation sector, would provide an enormous pool of flexibility, while at the same time increasing utilisation of clean renewable electricity. Power-to-heat is one of the most matured and well-developed technologies for sector coupling. Denmark has been using electric boilers to handle variability of wind and solar for many years and is investing significantly in large scale heat pumps to provide clean heating to their district heating networks. Electric boilers not only consume the surplus from wind and solar power, thus reducing the renewable curtailment, but also support the power system by providing short-term flexibility.

The clean heating initiative has significantly promoted the deployment of electrical heating in China. Both heat pumps and electric boilers are used for this purpose. Both can be operated to consume variable electricity surplus and thus reduce the renewable energy curtailment. Currently, heat pumps are mainly installed on the consumer side while electric boilers are mainly integrated in power plants. Many provinces in China are recently considering to introduce both down-regulation products and fast frequency response products in an ancillary service market. Power-to-heat assets could provide both, thus the financial feasibility for investing in these assets would be significantly improved.

- Market oriented operation of flexible combined heat and power power plants

Flexibilisation of thermal power plants has been proven to be one of the most cost-effective measures in boosting system flexibility, both in Denmark and China. For combined heat and power (CHP) plant owners, physically retrofitting the power plants is only part of the story. As observed in Denmark, CHP plants need to continually adjust the heat and power production according to the market signal and also invest in new software and hardware. Better use of the flexibility to reap its benefits requires a systematic approach, by taking the economic cost of producing heat, with all relevant constraints considered, into account in all the major decision making processes including choosing investment options and doing daily operations.

According to a recent Sino-Danish study , increased thermal power plant flexibility could reduce Chinese CO_2 emissions by 39 million tonnes in 2030 due to less coal-based power or heat production and a reduction of VRE curtailment of 30 % compared to a scenario without flexible power plants. With Chinese spot markets implemented, increased thermal power plant flexibility will also raise the power prices for both VRE and other power plants, thanks to the marginal pricing principle, and reduce overall power system costs by an estimated 46 bn RMB in 2030. An essential precondition for developing enhanced power plant flexibility is a market framework that motivates both the development and utilisation of flexible characteristics in the system.

As more and more power plants have been retrofitted in China, the new market environment would also require the power plant owners to choose the right technologies and also be smart in the daily operation. The decision making concepts widely used in Denmark are described in this report, and can inform the decision making for China's flexible CHPs.

4 China's experiences

Since the release of Document #9 in 2015, China has embarked on a new round of market reform. The aim of the market reform is to move away from the governmental planning institutional setup to a competitive power market. But this process is far from completed. China is still in the midst of a transitional period of electricity market reform. In 2017, roughly 25% of the total electricity was traded on the market, another 75% was still transferred through grid companies with price and quantities largely determined by local governments. Trading today is mainly based on long-term (monthly and annual) bilateral contracts. In 2017, 8 provinces were announced by NEA as the first batch of provinces to try out spot market. However, spot market is inextricably linked with almost every facets of the power sector, ranging from end-consumer's electricity prices to the whole business model of grid companies. As already witnessed in the 8 provinces, the establishment of spot market is usually a protracted process.

Without spot markets, extracting flexibility from the existing assets and encouraging new investment on flexibility would require extra market design and new business models. One of the prominent and also successful efforts in China is the down-regulation market in Northern China. Due to the extra economic incentives from the new market, new business models have surfaced in traditional thermal power plants some of which are transformed into a hub for the integration of new flexibility assets.

- Down-regulation market in Northern China

Back in 2014, Northeast provinces, enduring the most acute RE curtailment problem, decided to inaugurate a new ancillary service market, delivering down-regulation services. This market provides a side payment mechanism involving only generation side. The end-consumer's price is not being influenced. The market effectively penalizes inflexible units while rewarding flexible units and has been proven to succesfully convert part of the previously curtailed energy to economic incentives for flexibility investment. The down-regulation market can coexist with a governmental planning paradigm, and does not require a thorough change to the power sector institutions. On the other hand it enlarges the social welfare, as the reduction of curtailment will lead to savings of fossil fuels. A sibling policy of down-regulation market is to exempt levies of electric boilers sitting behind the meter of CHP power plants, which has been essential for the scaling up of this market. Since the introduction of the down-regulation market in Northeast provinces, the curtailment rate has continued to decline from 20%+ in 2015 to less than 5% in 2015.

- Thermal power plant as a hub for integration of flexible assets

The installed capacity of thermal power plants in China reached 110 GW by the end of 2017, accounting for about 62% of the total installed generation capacity. Most of the thermal power plants are connected to the grid with a voltage level of 220kV and 500kV. There are various grid codes and technical standards to manage the integration of thermal power plants, requiring them to install various hardware and software which allow them to respond to the dispatching signal effectively. The cost of building up such a good connectivity with the grid is relatively high for small assets, but not that evident for large thermal power plants. The technological advancements also render other flexibility investments more and more attractive. Notably, lithium batteries and electric boilers, once considered to be too expensive for flexibility purpose, show increasingly promising prospects in China. The application of batteries and electric boilers on a utility scale (over MW) requires full-blown grid access hardware and software, to guarantee their observability and controllability. These expensive prerequisites are readily available in conventional thermal power plants. Thus, China has witnessed a new trend of integrating flexible assets, lithium batteries and electric boilers, behind-the-meter in the existing thermal power plants, which thereby have been transformed into a hub for integrating various flexible assets.

I. Introduction

1.1 Motivation for this report

Due to the rapid growth of variable renewable energy (VRE) in global electric power systems, grid operators are encountering new challenges to reliably operate the grid while maximizing renewable integration. A particular pain point has surfaced around the lack of flexibility throughout all aspects of operation: scheduling and dispatch, asset operation, and available technologies.

These challenges are particularly salient in China, where the largest installed capacity of variable renewable energy in the world is struggling to be fully integrated into the power system. China is deploying numerous strategies to deal with this problem, including setting aggressive targets, deploying new technology, and, perhaps most auspiciously, pursuing a new round of power market reform, which could play a major role in facilitating and incentivizing a more flexible power system.

The U.S. and Europe can provide relevant examples: each have regions with similarly high-shares of renewables, and have well-developed, mature power markets. Policy makers and regulators in each of these regions have introduced or updated market products and market designs, especially in short-term markets, to develop or unleash new flexibility potential on the grid.

But markets are only one element, technological innovation, policy and regulation also play an important role in these cases, especially in instigating reforms, pushing market operators to find solutions that span multiple energy systems and even the traditional sphere of the utility.

The objective of these policies and market redesigns is to ultimately provide the right context for assets of all types to perform more flexibly: revealing the value of more flexible operations, and where necessary, providing sufficient economic circumstances to invest in more flexible resources. That way, technical measures for enhanced flexibility can be deployed through a wide-array of business models, finding innovative ways to deliver this flexibility at least cost.

This report aims to understand these interconnections; how all of these elements work together to realize increased system flexibility, helping the reader to understand how the right market and regulatory mechanisms, nourish and inspire new business models and technologies to boost flexibility and address VRE integration issues. We will provide specific experiences from these 3 regions that offer examples on different sources and measures that deliver or facilitate much flexibility in the system. Examples include improved flexibility of production units, demand flexibility, system flexibility and sector coupling, as well as the market set-up and regulation incentivizing and enabling more flexibility.

The hope is that such experience can have value for different regulators, policymakers, industries, and researchers in China, at both the national, regional and provincial level, understanding how these experiences can be adapted to different situations and different stakeholders in China – to support China's ongoing power market reforms, and its quest to further its integration of VRE.

1.2 Report Structure

The objective of this report is to provide solutions, inspiration and insight for the different Chinese power sector stakeholders engaged in enhancing the flexibility of the Chinese power system. This report aims to do so by furnishing detailed case studies of different policies, mechanisms, and technical measures that enhance the flexibility of the system from each of the different regions (Europe (Nordic states), US (California, Texas), and China). For the US and China, a regional overview with the necessary context to understand the specific energy context (i.e. asset mix, demand level and profile etc.), as well as the specific market and regulatory mechanisms that shaped the measures implemented and provide rationale for why each region pursued their flexibility needs in this way, is provided. In an effort to summarize the findings and find applicability in the Chinese context, each example case is ended with a reflection on the relevance of this case in the Chinese context.
The remainder of this section provides a high-level flexibility framework for which to think about the types of flexibility required to operate a high-VRE power system. It will also inventory the most prevalent approaches to procuring flexibility and explore cross-cutting trends seen in all regions with high VRE. This section also aims to build upon previous collaborations between China's researchers and international partners, helping build connections with other foundational research on this topic.

1.3 Types of flexibility required for VRE integration

Although the electric power industry has coalesced around flexibility as a major need to enable a high VRE power systems, that flexibility need is often ill-defined. Flexibility takes many different forms, from having to ramp generators up and down to follow load to being able store energy at one time to be used during another. Even reducing the need for flexibility through better production forecasting and larger shared balancing areas is considered within this report.

At its broadest, flexibility is the ability to change energy output or consumption over certain timescales (seasonally, daily, hourly, sub-hourly to minute/second frequency control/regulation), in response to system conditions and signals. The ability for a certain resource to change within a certain time and to a certain dispatched state can be measured by the below metrics. While there is no standard set of defined metrics, these metrics collectively describe the attributes required by system operators across all regions. Furthermore, all flexibility products laid forth by these system operators specify requirements that include one or more of these metrics.

- Ramp speed, how fast a resource can change its output (MW/s)
- Ramp Mileage/deployment time, how long a resource can provide a change in output, including limitations to carrying capacity and minimum run-rates (seconds at a sustained MW/s)
- Response time, how fast a resource can change its operational state to the desired state (s)
- Dispatch granularity, how accurately a resource can provide the needed response (MW)

Characteristic	Challenge	Examples of VRE Impacts
Ramp speed	Traditional generators can only change output gradually, different reserves cannot come to full operation fast enough to cover frequency deviations	Large changes in solar and wind output or incidents of inverter trip off require generators to ramp significantly to cover gap.
Ramp mileage	Generators and T&D have minimum run rates and maximum capacity levels, batteries, flywheels and rotating motors have limited potential stored, all resources have limited capability to operate in an active state	Sustained declines in solar as evening demand increases requires sustained ramping up from dispatchable resources.
Response time	Start-up time limitations, delays in communication signals, inertia in the system makes changing directions of operations hard (generators ramping down have a delay before being able to ramp back up)	Small variations in solar and wind output create needs for regulation services or load-following. Increased moments of small unscheduled deviations require more on-call resources. High ramp periods create higher chances of generators tripping off and increased need for contingencies.
Dispatch accuracy	Large generators cannot provide very fine load-following or time-sensitive regulation services, critical peak demand response reduces demand as much as possible	Constant small changes in solar and wind output require accurate firming services so that by responding resources do not overshoot and exacerbate the problem.

Table 1: Description of Flexibility characteristics

These metrics are important to consider when understanding a system's flexibility challenge and which flexible resource can help address that problem. That is to say not all flexibility is created equal and is not always suited to the problem at hand. For example, California requires a high ramp rate over a sustained ramp to meet evening demand peaks that coincide with solar output declining with the sunset. Whereas the Denmark requires accurate, fast-responding resources to manage the moment by moment fluctuations in output in its system supplied at times by 100% wind.

The most economic solution at hand, however, is to minimize the need for these types of flexibility. In many high VRE regions this has been a key to successfully integrating VRE. For example, improving renewables fore production forecasting and thereby minimizing the need for costly real-time redispatch, or reducing the need for severe ramping by encouraging demand to shift to off-peak times through time-of-use retail rates. Defining the flexibility need, then identifying the magnitude of the service needed is an essential first step. All of the cases will highlight the current system conditions that created the need for more flexibility and how the mechanisms chosen were selected for their appropriateness to the flexibility challenge at hand.

1.4 Major sources of flexibility

In order to meet these new demands for flexibility, a series of standard measures have been deployed globally (table 2). Each of these measures contribute in their own way to addressing challenges sufficiently meeting VRE-induced requirements around ramp rate, ramp mileage, reaction time, or dispatch granularity. Although imprecise, for each standard measure an assessment is provided whether it substantially contributes to that metric. It is also worth emphasizing that a many of these measures do not require new physical assets, and in some cases do not require any changes to existing physical assets, merely a change to the institutions that dictate grid operation. Globally, these institutional measures have been pursued first, being, on average, cheaper and foundational to deploying future physical assets. Table 2 also provides an indication if these measures are institutional, physical or both. This list builds upon a list assembled for a study by NREL and China's National Renewable Energy Center[1].

Flexibility Measure	Physical or Institutional?	Ramp Rate	Ramp Mileage	Reaction Time	Dispatch Granularity	Discussed in this Report
Larger balancing areas: Increasing the size of geographic area where operators conduct resource planning and load-interchange-generation balancing	Both	X	X	X		X
Access to neighbouring markets: Physical interconnection via transmission networks and the institutional mechanisms to coordinate transactions with neighbouring power systems	Both	X	X			X
Faster energy markets: Shorten the time scale of scheduling, dispatch and settlement in power market	Institutional	X		X	X	X
Regional transmission planning for economics and reliability: VRE integration is considered in current transmission planning to minimize costs to interconnect and firm resources	Both			X		
Robust electrical grid: Transmission lines with adequate capacity to avoid binding constraints and redundancy to facilitate shifting patterns of power injection	Physical			X		
Improved energy market design: Create resource-neutral and performance-based energy market to select the best resources to provide what services, and avoid barring new resources because they cannot provide all services	Institutional	X	X		X	X
Demand response: Structuring markets to properly incentivize and utilize responsive load	Both	X	X			X
Geographically dispersed VRE: Build VRE resources across large geographic areas to smooth out the volatility of the aggregated supply	Physical	X	X	X		
Strategic VRE Curtailment: Create mechanisms to make economic choices to curtail VRE, evaluating the trade-off between the instantaneous value of the energy forgone and the value of procuring other ancillary services to fill the gap	Both	X	X	X	X	
VRE forecasting effectively integrated into operations: Improve VRE forecasting to reduce the system flexibility need incurred by the variability and uncertainty of VRE	Both	X		X		X

1 https://www.nrel.gov/docs/fy16osti/64864.pdf

New flexibility ancillary service products: Create new ancillary service product to incentivize the creation and utilization of system flexibility	Institutional	X	X	X	X	X
Sufficient reserves for VRE event response: Re-evaluate reserve margins so regulating reserves are not exhausted by responding to VRE fluctuations	Physical		X	X		X
Flexible conventional generation: Modify thermal plant operations from baseload to dynamic, changes in output, cycling on and off several times a day, reducing minimum run rates	Physical	X	X			X
Primary frequency response, inertial response, and response to dispatch signals with new VRE: Develop technical requirements or mechanisms to ensure during periods of high VRE that enough capacity exists to in a short-time frame arrest a contingency event	Both			X	X	X
Storage: Develop storage and integrate with system operation	Physical	X	X	X	X	X
Sector Coupling: Coordinating electricity production and consumption with heat provision, water rights, increases in transportation electricity demand, industrial electrification, etc.	Both		X		X	X

Table 2: Overview of standard flexibility measures

1.5 Procuring flexibility

By defining flexibility needs and identifying which set of resources are capable of providing those services, system operators can turn to understanding how to procure those resources (or reduce those flexibility needs). As load and RE output variability increases, dispatchable resources must change their output faster, more frequently, and with less notice in order to maintain reliability. Thus the procuring agent is usually the system operator , since it is the system operator's job to maintain reliability.

Therefore, it is useful to think about procuring flexibility in the terms a system operator thinks about reliability, spanning several timescales, typically paying for and procuring each via a different mechanism.

- Planning (>1 month): How do I minimize the need for new flexibility incurred by new demand/VRE, and how best to procure it?
- Scheduled dispatch (day-ahead): Did I schedule enough flexible capacity to meet expected ramps caused by changes in renewable output and demand?
- Real-time dispatch: Did I reserve enough capacity to cover unexpected flexibility needs caused by unexpected changes in VRE or demand?
- Event/contingencies: Do I have enough capacity available to recover from a loss event, especially if I have a large share of renewables interconnected at the time of loss?

The procurement mechanisms also look substantially different depending on the system operator's context. In the most basic sense, selling a flexible generator to a vertically integrated utility employs a very different business strategy than a merchant generator bidding flexibility services into an ancillary service model. This is further convoluted by the role of system operator frequently being shared by different parties, each with different mixed subsets of responsibility. This is why across many of the regions, similar flexibility measures were procured using very different models, and the resulting flexibility provided and the business models supporting those functions vary wildly.

This report will primarily focus on the scheduled (day-ahead) and real-time dispatch, specifically the interrelation between the role of day-ahead scheduling to meet expected flexibility, and the role of real-time dispatch to cover unexpected divergences.
While the planning horizon is important, we will only mention how payments for providing flexibility resource adequacy supported various flexibility business models. We will also cover some examples where flexibility contributes to meeting operational reserve requirements, especially where they cover the additional event contingencies associated with high VRE penetrations.

Although this systemization is in no way the only way to think about how flexible resources are paid, it is notable how many of the cases contained in this report see their revenue streams collating with this framework used by system operators, especially cases participating in fully liberalized electricity markets. Figure 1 connects the cases contained in this report to these timescales to indicate where they participate and receive compensation.

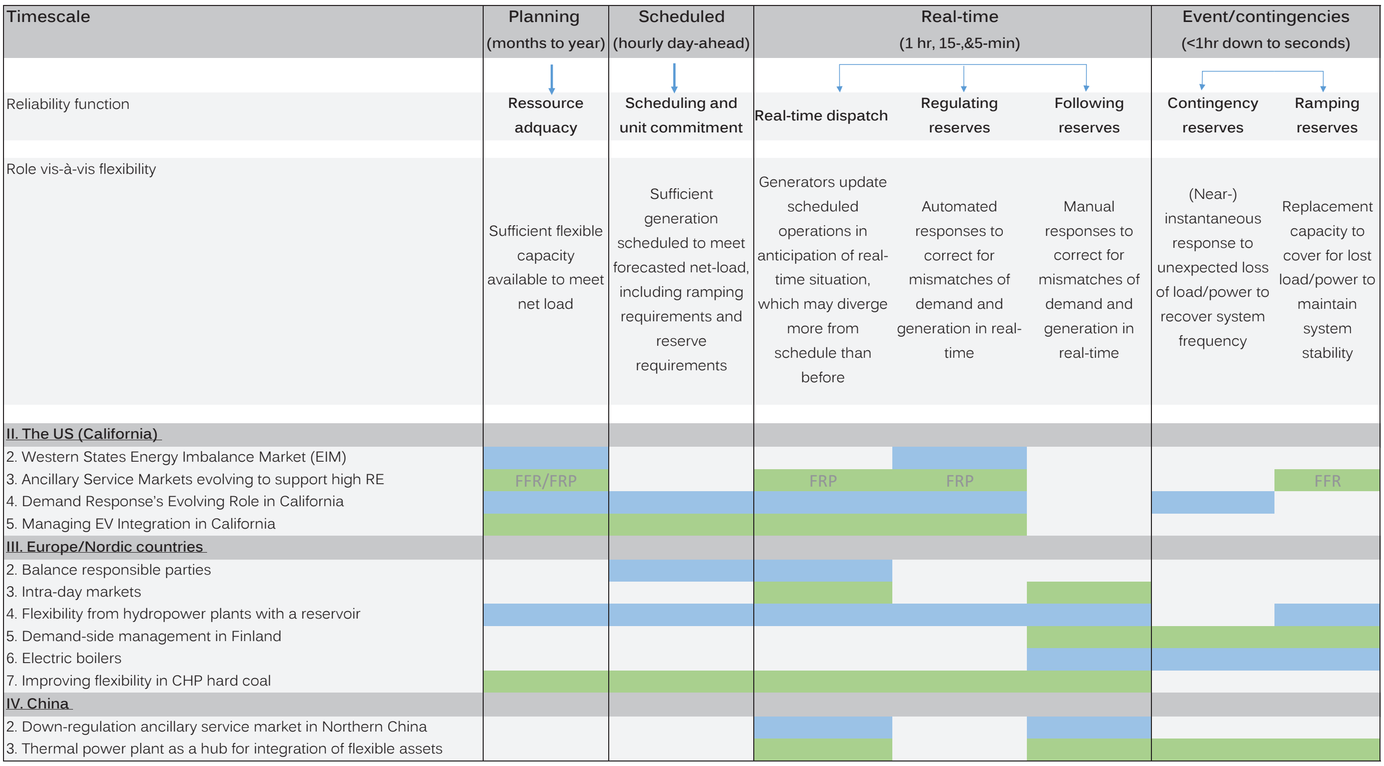


Figure 1: Timescales Flexibility is required and relevant cases applying to each

1.6 Regulatory perspective on flexibility

Most of the cases covered in this report look at what is the cutting edge of flexibility in electricity systems. Often the most impactful measures are the most mundane. It is important for decision-makers to focus on fundamentals first, as they can provide the largest impact and are important to making future investments in flexibility more valuable. These standard practices are not emphasized in the following cases, since they are well-documented in other resources already.

But all of the cases included assume that economic dispatch is in practice, dispatch schedules are hourly or less, and that reserve requirements are set according to reasonable risk of event.

Beyond these fundamentals, other cases with some of the largest impact potentials have been deemphasized in the report since they are straightforward, and the details do not warrant a full case. Examples include better communication between different balancing authorities and improved weather forecasting for VRE output predictions.
Flexibility is often given more attention than is necessary for the levels of VREs that systems are experiencing. Many systems have been able to integrate 20%+ of VREs with little changes to their operations. So the frenzy of building new dedicated markets for specific flexibility needs, providing retrofit incentives, and installing new assets are often times excessive and costly approaches. It is important to help identify stage gates when system operators must consider these new options, and not insist on adding these prematurely. One case included (and many others not included) provides an example where new institutional measures were tested out, proven to be successful, and then not adopted after regulators determined it was not essential to begin paying more for that flexibility at that point in time.

Finally, as new technology comes on board (e.g. IT-enabled DR, EVs, and battery storage) system operators need to be prepared for a more comprehensive rethinking of the models in which they dispatch, compensate, and interconnect those new options to ensure swift uptake and integration. Many current models for wholesale power markets are rapidly becoming antiquated in a high VRE system, since those methods were designed for a time when fossil-fuel based generators with high-marginal costs were setting real-time prices. These new models are being explored as we speak, and China, already on the vanguard of VRE adoption, may have an opportunity to leap ahead during this set of market reforms to new models better suited for VREs and highly flexible resource dispatch.

Chapter II

United States
California

Resume

This chapter provides examples of how California and Texas, two of the most advanced regions when it comes to renewable energy deployment and integration, have tackled the specific flexibility needs emerging in their power systems. California is known for the "duck curve" where a large evening peak in demand just when solar PV production declines causes pressure on the system. Texas has a large and increasing wind portfolio which can challenge system stability. The U.S. provides examples of how these challenges have been tackled in the context of a central dispatching and more vertically integrated system than Europe. This chapter shows how markets providing the right products and prices can provide the system operator with flexibility tailor-made to the systems needs.

II. The U.S. (California)

1. The U.S. power market

As with all cases presented in this report, it is important to understand the political, economic, and technical context of that system. The U.S. system is notoriously balkanized, with states and regions taking their own approach to electricity regulation and market operation. In order to provide sufficient context for the cases that follow, we provide a general overview of the U.S. system and its relevant actors, but then we focus on one region, California, providing deep context on their system. We selected California because many of the strongest cases around flexibility for renewable integration are occurring in California, particularly in regard to minimizing or procuring resources to manage inter-hour ramp. There are other regions with high renewables taking innovative approaches to integrating RE, and these will be used as points of comparison, understanding alternative pathways to a highly flexible system. These will be explored in sub-cases and call-out boxes throughout the main cases.

This first section of this overview will explain the system set-up in the U.S. while the second section will provide specific details on California's context.

1.1 Overview

The U.S., like all regions started with regulated utilities, with each state responsible for making sure utilities were providing just and affordable prices to the customers in the area they were designated to serve. As power systems advanced, utilities began connecting between different service areas, enabling exchanges of power between different utilities for economic and reliability reasons. Due to this broader interconnection, advances in technology and regulator desires to push for a more economically efficient power sector, competitive wholesale markets were introduced. Markets enabled more fluid transactions between different actors—generators, utilities, customers—which needed entities to facilitate, oversee, and help manage disputes.

Given this evolution pathway, there are different oversight overlays at work in the U.S. which are important to understand to interpret the motivations of different actors and the resulting solutions in the following cases. These layers are:

- **Regulation**: state-level entities overseeing utilities' provision of power to customers
- **Reliability**: standards that assign responsibility for maintaining operational reliability of the grid
- **Market**: markets for the trade of electric power across different regulatory jurisdictions (states)

These structures are not applied uniformly across all regions in the U.S.. Each state has their own regulator and policy objectives, markets were introduced only in some areas

and applied to different extents across those regions, and with those variations, the entities responsible for electric reliability coordination are also different.
The following sections will explore the general structure for each of these layers, identify the key actors, and describe their roles.

1.2 Regulatory Structure

The electric utility was a natural monopoly for a long-time, requiring regulatory oversight to approve investments and customer rates. In the U.S. today, many of the utilities are independent companies, not government run entities[1]. These **investor-owned utilities (IOUs)** are regulated at the state-level by the state's **Public Utility Commission (PUC)**. The IOUs have an obligation to serve all customers within their territory and determine what operational and investment expenditures are necessary to provide reliable and reasonable-cost energy. IOUs present these investment and operation plans to the PUCs who either approve or deny the proposal on the grounds of the request being just and reasonable. This approval process is the core of the PUC's responsibility, which are predominately focused on:

- maintaining affordable electricity prices;
- equitably assigning costs across customer groups, ensuring that proposed rate structures reflect the true costs to serve those customers and create the right incentives to encourage desired behaviour; and
- enforcing policies by evaluating the proposed investments against state goals (e.g. renewable targets) and ensuring that the utility meets its obligations under the law.

These guiding principles have also has shaped how utilities propose and justify investments, particularly the investments that directly increase investment (known as the rate-base) and, potentially, retail prices. Many of the cases discussed not only provide flexibility, but also generate many other values as a part of their deployment. This is in part due to the stringent requirements for IOUs to justify these investments.

On the wholesale level, markets are overseen by various entities. All regulated portions of the business are still overseen by PUCs and where a single ISO/RTO exists entirely in a state's boundaries, the CPUC has a role in regulating those markets. Since PUCs can only regulate within their state , regulatory approval of markets spanning state lines fall to the **Federal Energy Regulatory Commission (FERC)**, a U.S. federal agency that regulates:

- Rates for and the construction of interstate electric transmission
- Wholesale sale of electricity in interconnected regions, including construction and sales

1 There are publicly owned utilities in the U.S., most commonly at the municipal level, who are owned and operated by the local government in a non-profit manner. Publicly Most, though not all, publicly-owned utilities are not regulated by state PUCs, but rather are overseen by the PUC, but undergo a similar process with their local government entities governing bodies (e.g., elected municipal councils) or, in the case of consumer cooperative utilities, by their member-elected boards of directors. Given IOUs make up about 2/3 of all electricity sales in the U.S., we will focus on this model more than the municipal utility model.

Even in utilities that span multiple states, they are required to have different legal entities in each state and are each subject to the approval of that state's PUC.

of federally-owned hydro plants, and

- Adopting and enforcing reliability standards across the country

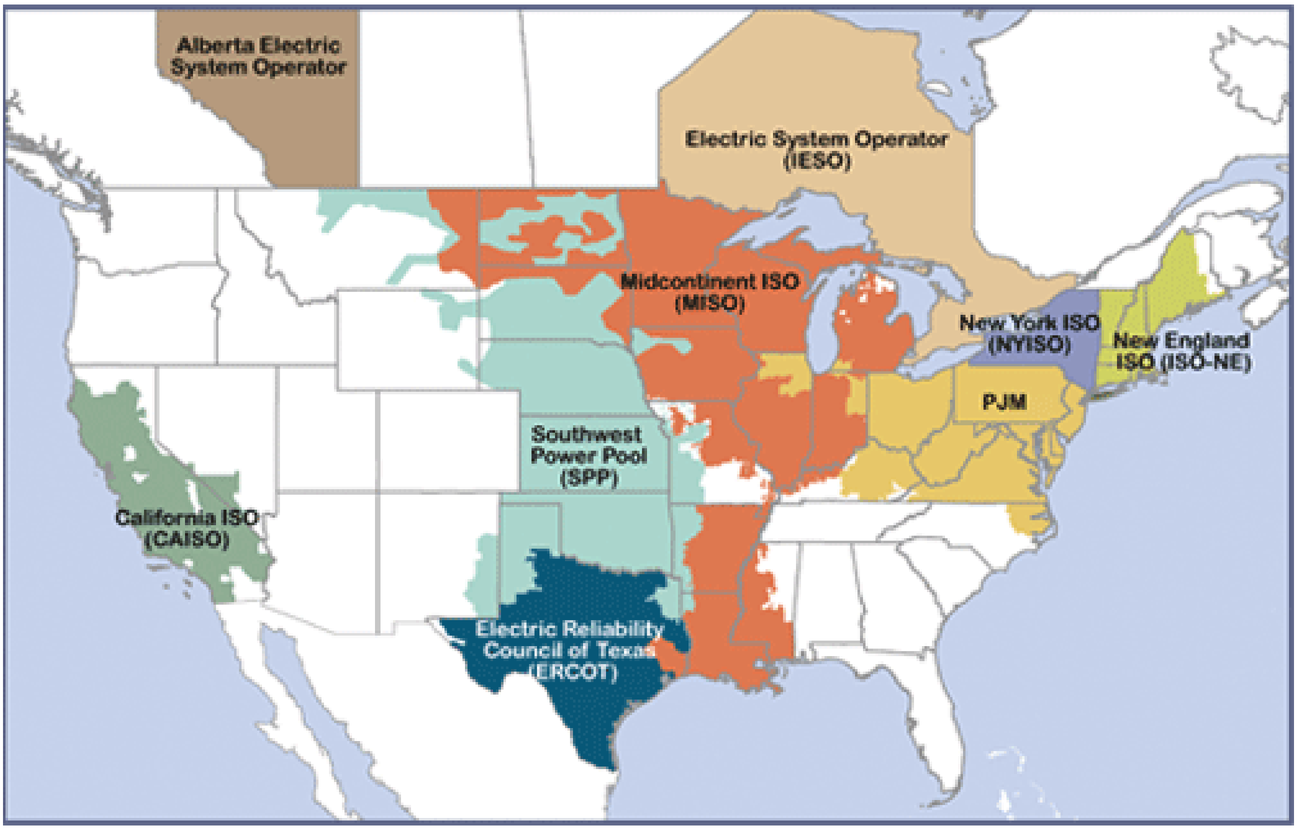


Figure 2: Regions with operating ISO/RTOs overseen by FERC[2]

1.3 Wholesale Electric Power Markets

Under the old regulatory regime, the financial risk of a non-performing asset was generally shared by shareholders and customers ("ratepayers"). Ultimately, evaluating if investments are justified is difficult, and this difficulty was one of the motivators for wholesale market reforms. Wholesale competition also aimed to reduce operational costs, which historically were readily passed through in rates, by submitting plants and utilities to competitive pressure to deliver energy more efficiently.

The primary models of deregulated power sectors in the U.S. today are described below[3] (figure 3).

2 https://www.ferc.gov/industries/electric/indus-act/rto.asp

3 Because different states and PUCs have different views on deregulation, the U.S. has some states that have moved to wholesale electricity markets and others that have not. And those that have arrived at very different formats for that deregulation.

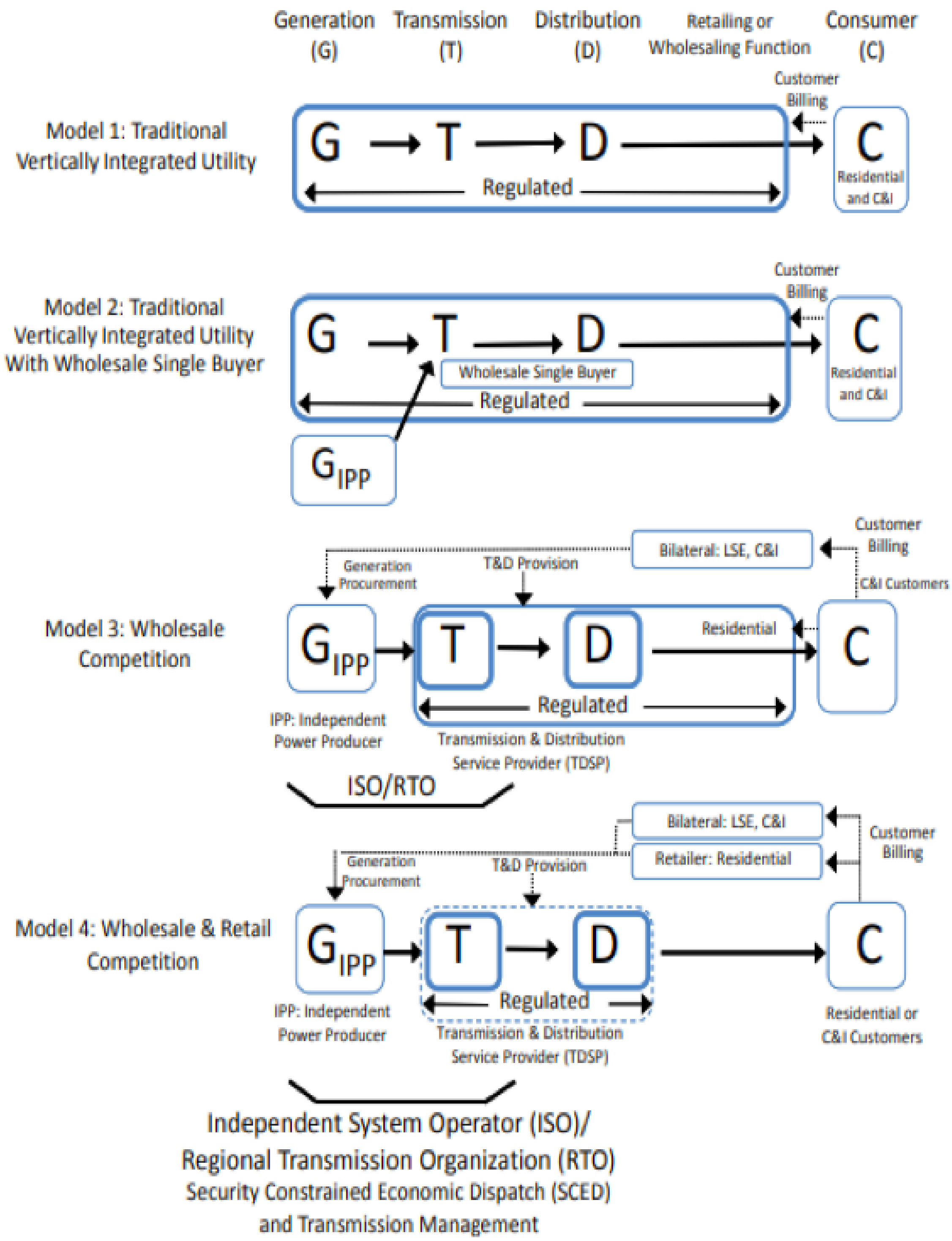


Figure 3: Primary models of electric deregulation in the U.S.[4]

4 http://sites.utexas.edu/energyinstitute/files/2016/09/UTAustin_FCe_History_2016.pdf

Regulated Utility (Model 1&2)
Vertically integrated utilities (VIU) own generation, transmission, distribution and retail. They are responsible for procuring power and delivering to the customer.

- VIUs generate power and also buy from independent power producers (IPPs).
- Utilities are the single buyer on the market, and have a monopoly on networks and sales to final consumers. They also take on the responsibilities of ensuring reliability.
- When making new investment (in G&T&D), the decision is subject to regulatory approval besides siting approval since the cost will pass to the consumers, increasing utility bill.

Generation competition (Model 3)
Wholesale market is facilitated by ISO/RTOs. Power generators compete in the market to be dispatched.

- New investment in generation capacity is a financial decision made by the corporations.
- Utilities provide both distribution and retail services to end users. Retail customers cannot not choose their power suppliers, thus regulator must approve rates that each electric utility charges its customers.
- ISO/RTOs are the independent entities that take the role of ensuring reliability.

Generation and retail competition (Model 4)

- Both generation and retail markets are deregulated.
- Retail customers can choose their suppliers. The utility still handles billing.
- Regulator must approve network rates that each electric utility charges.
- ISO/RTOs are the independent entities that take the role of ensuring reliabiltiy.

Figure 3b: Primary models of electric deregulation in the U.S

Most deregulated territories in the U.S. have focused on the supply-side, in one of the following ways:

- Unified central markets where all resources are dispatched and paid by the market (PJM, ERCOT, CAISO)[5]
- Power exchanges or power pools that facilitate transactions between different utilities and independent power producers (IPPs) in a region (SPP, MISO)
- Open access tariffs, which allow IPPs to sell directly to regulated utilities without fear of non-preferential dispatch treatment

Centralized markets and exchangers are operated by **Independent system operators (ISOs)** or **regional transmission organizations (RTOs)**, which were established to ensure

5 Dispatched by the market does not necessarily indicate that all markets are a gross pool model. The net pool vs gross pool framework is of great interest in China at the moment. Elements of gross pool and net pool exist across all markets in the U.S., where most generators are able to submit firm scheduling if they have secured transmission access rights. But given the challenge around this, and the efficiency/flexibility of virtual contracts most markets have moved toward more of a gross pool methodology.

non-discriminatory access to the transmission system. While ISOs/RTOs do not own physical assets, they are responsible for dispatching assets, maintaining operational control of the transmission grid, and planning for grid expansion[6]. Most of the cases presented occur in states with deregulated wholesale markets, in part because this model has been more conducive to renewable energy entry, prompting more focus around flexibility.

Far fewer areas have deregulated retail services, often due to an inability to overcome concerns around price exposure to customers. While none of our cases directly discusses how deregulated retail models have produced flexibility, many regulators are exploring this model in order to stimulate innovation in customer rate-making and leveraging demand-side flexibility.

1.4 Reliability and Dispatch

With all of these different entities owning and operating grid assets and all interconnected to a common system, it was important to have a cohesive paradigm for coordinating these different actors to ensure reliability. To this end, the North American Electric Reliability Corporation (NERC) was established to develop reliability standards. NERC is an international authority overseen by FERC and the Canadian Government.
Meeting these standards is the responsibility of a balancing authority (BA), typically a utility or RTO/ISO but also often a vertically-integrated utility, which is responsible for:

- developing balanced schedules to ensure adequate supply to meet demand
- dispatching assets in real-time to maintain balance between supply and demand
- maintaining adequate operational reserves to handle any event that may threaten reliability
- securing adequate generator and transmission capacity to meet future energy needs

These duties are carried out by the BAs, but the means by which they achieve these tasks vary significantly from territory. A general framework of how this looks in the U.S. is provided in figure §4, followed by brief descriptions of some of the major mechanisms employed.

6 "How the RTO Effects Market Efficiency"

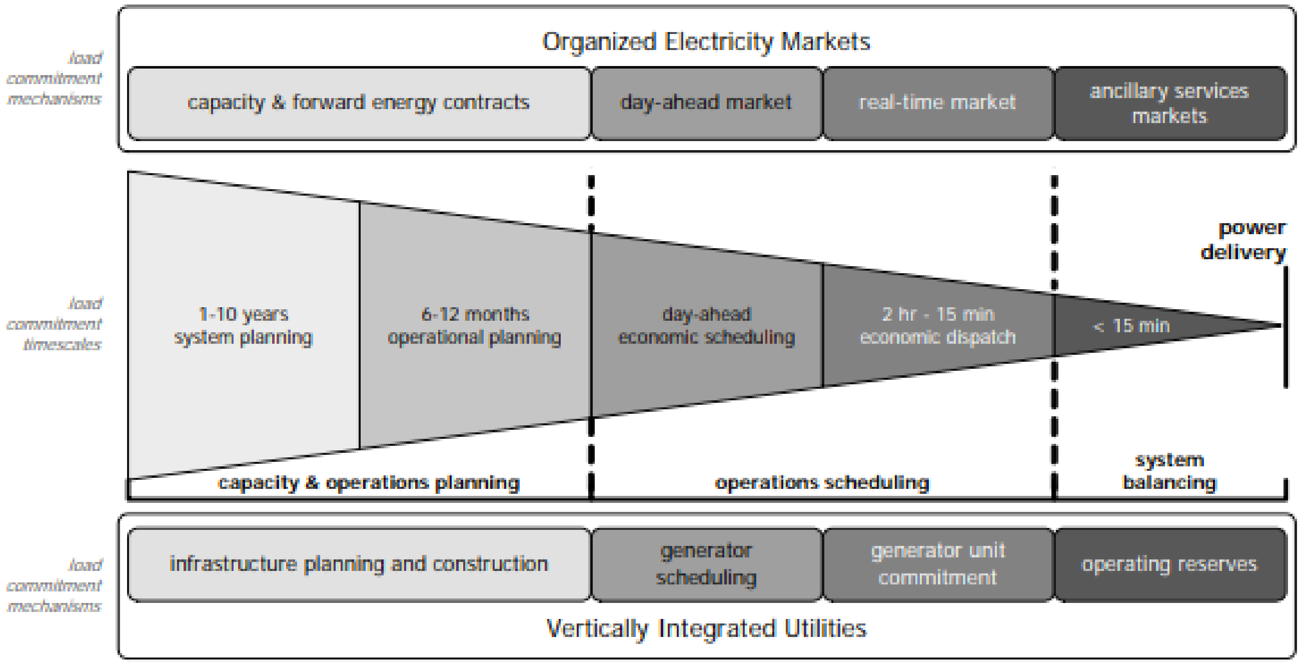


Figure 4: Electricity system planning and operation across markets and vertically integrated utilities[7]

Planning

- Regulated market: conduct an integrated resource planning to propose an investment plan to meet resource adequacy needs
- Deregulated market: establish centralized procurement mechanisms for the ISO/RTO or member utilities to procure adequate resources
    - » **Energy market**: power plants sign long-term bilateral contracts with enough generators to meet resource adequacy needs (called energy forward contracting)
    - » **Capacity market**: resources necessary to meet resource adequacy are cleared in a centralized bidding process, and receive payments to "reserve" their capacity for coming years[8].

Scheduling

- Regulated market: The utility schedules power plants in accordance to day-ahead expectations, and redispatches in real-time to maintain reliability
- Deregulated market: RTO/ISOs run **Day-ahead** and **Real-time energy markets**, the first to schedule demand, the second to dispatch any divergence from that schedule in real-time. All markets conduct centrally cleared, economic dispatch using locational marginal pricing to determine the economically optimal generation pattern without violating security constraints[9,10]
- There are significant differences in how various markets manage settlement, and the timescales in which they are conducted. For instance, NYISO schedule day-ahead, but pays all resources according to real-time dispatch, where PJM pays at Day-ahead, and

7 https://emp.lbl.gov/sites/all/files/report-lbnl-1252d.pdf

8 https://www.e-education.psu.edu/ebf483/node/646

9 https://www.pjm.com/-/media/training/nerc-certifications/markets-exam-materials/generation-itp/day-ahead-energy-market.ashx?la=en

10 https://www.aepenergy.com/2018/01/05/december-2017-edition/

any difference is bought and sold in the real-time markets.

Balancing and operational reserves

Grid operators make instantaneous adjustments to ensure grid stability when forecasts are wrong or when things do not go as planned on the power grid.

- Regulated market: Utilities take the full responsibility and shall include necessary capacity during resource planning
- Deregulated market: ISO/RTOs typically conduct via ancillary services market, which will be further discussed in section 3 Ancillary Service Markets evolving to support high RE

Many of the cases focus on the refinement of these procurement mechanisms and optimizing the systems by which planning, scheduling, redispatch, and contingencies are managed. As more renewable energy is added to the grid, and additional flexible operation is required, a good deal of those needs can be met by evolving these processes held by balancing authorities. The nuances of how these are handled will be discussed in the following cases.

1.5 Overview of U.S. Trends, and State Deep Dives

Across all regions in the U.S., some trends are shaping regional decision-making, with the most notable being:

- **Continually declining solar and wind prices.**
- **The shale gas boom**, which has made natural gas, cheap and plentiful. This has spurred investment in combined cycle natural gas (CCNG) across the U.S.. While open cycle natural gas combustion turbines are very good at providing flexibility, U.S. CCNGs are roughly as flexible as a coal plant.
- **Deployment of distributed energy resources** has increased rapidly, and is creating significant disruption to the traditional utility business model, which focused on electricity sales rather than on energy services.

California

These three trends are prevalent in California, largely driven by the state's aggressive emissions reduction goals (figure 5): renewable energy should contribute 33% of total electricity by 2020 and 50% by 2030. Its total electric energy demand is 206,336 GWh, with a peak demand of 50,116 MW. Peak demand and bulk demand have been growing year on year, despite some of the most aggressive energy efficiency programs in the country, in large part due to electric vehicle growth and a boom in manufacturing consumption.

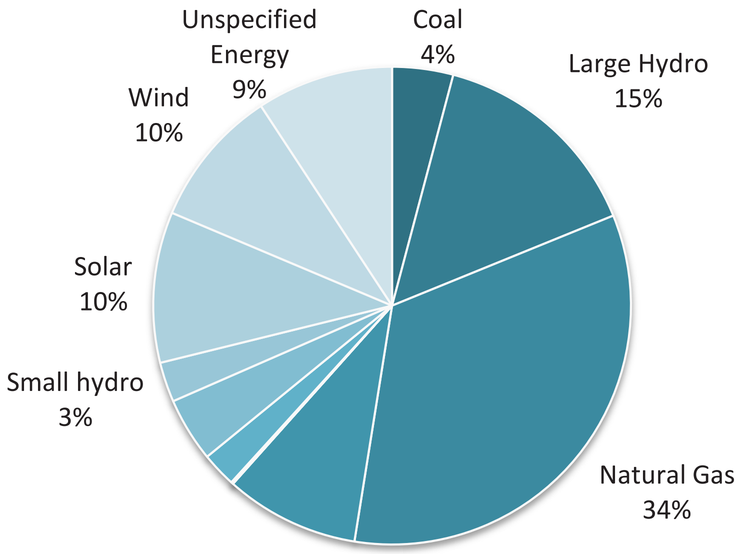


Figure 5: California's energy mix[11]

The California Public Utilities Commission (CPUC), the state's regulator, is very active shaping programs with the states IOUs and the ISO to advance California's renewable energy goals and wants California to become an incubator for innovative energy solutions.

The state has three IOUs, San Diego Gas and Electric (SDG&E), Pacific Gas and Electric (PG&E), and Southern California Edison (SCE). It also has a handful of municipal utilities, including two large one overseeing Los Angeles and Sacramento.

California deregulated wholesale market in mid 90s and formed **California Independent System Operator (CAISO)** to operate the wholesale market and manages the transmission grid[12]. Further deregulation in retail was halted after the California energy crisis in 2000, where unforeseen energy shortages and market manipulation derailed this early pilot of wholesale electricity markets. This resulted in one of the more unconventional market designs in the U.S. While CAISO operates all resources in the market, its 3 major member utilities still remain responsible for procuring adequate generation, either by owning the generation themselves, or contracting with IPPs and other suppliers such as other utilities. Also, as a consequence of the early deregulation experiments, CAISO operates a true gross pool model, because the schedules submitted by member utilities and their contracted generators frequently resulted in infeasible schedules and a lack of coordination during scheduling and real-time dispatch. CAISO operates several markets to meet this obligation: **energy markets (day-ahead and real-time), and ancillary services.**

- **Day-ahead market:** receives bids from all generators, and schedules for those generators under special dispatch conditions (e.g. hydro, geothermal, etc.), determines a feasible day-ahead schedule at least cost. This schedule is co-optimized with Ancillary Service markets when making the schedule.

11 http://www.energy.ca.gov/almanac/electricity_data/total_system_power.html

12 http://www.caiso.com/about/Pages/OurBusiness/Default.aspx

- **Real-time market:** dispatches every 15 and 5 minutes, and even for a single 1-minute interval under special conditions[13].
- **Ancillary services:** 4 types of products- regulation up, regulation down, spinning reserve and non-spinning reserve.

California's growing amounts of renewables, solar in particular, has driven increased needs for flexibility. The challenges they are facing are moments of oversupply during moments of low demand when firming resources are online at their minimum run-rates. The more infamous problem is the "duck curve" (figure 6), an operation situation when solar generation decline at sunset coincides with demand increasing toward system peaks. This creates a very long, sustained period of ramp that requires a lot of flexible generation to be "waiting in the wings" at minimum run rates to meet this ramp, and increased risk of a ramping generator tripping offline and creating a contingency event[14].

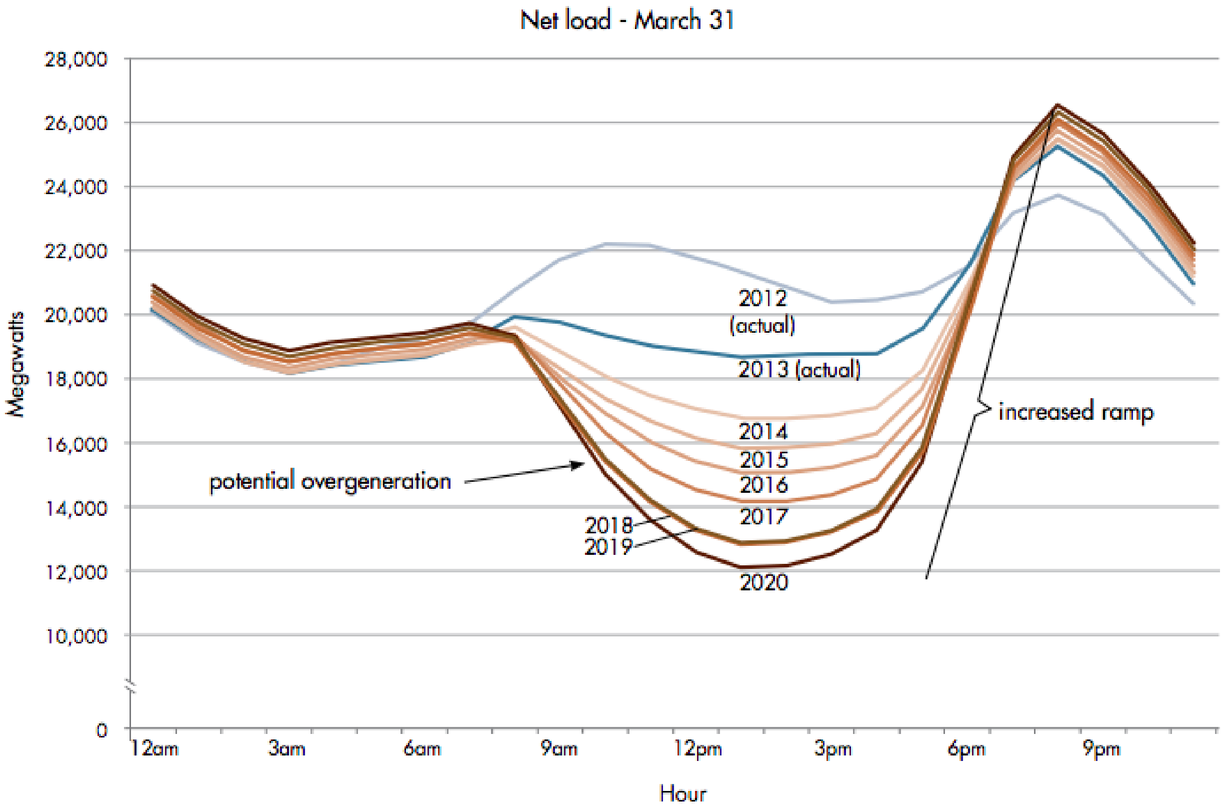


Figure 6: CAISO's duck-shaped net load curve[15]

13 http://www.caiso.com/market/Pages/MarketProcesses.aspx

14 https://www.vox.com/energy-and-environment/2018/7/31/17611288/california-energy-grid-regionalization-caiso-wecc-iso

15 https://www.energy.gov/eere/articles/confronting-duck-curve-how-address-over-generation-solar-energy

2. Western States Energy Imbalance Market (EIM)

2.1 Key messages and takeaways

EIM is a voluntary real-time energy-only market in which CAISO and surrounding utilities meet their real-time balancing needs more efficiently by pooling their resources and allowing a centralized dispatcher, in this case CAISO, to optimize exchanges to ensure balanced schedules.

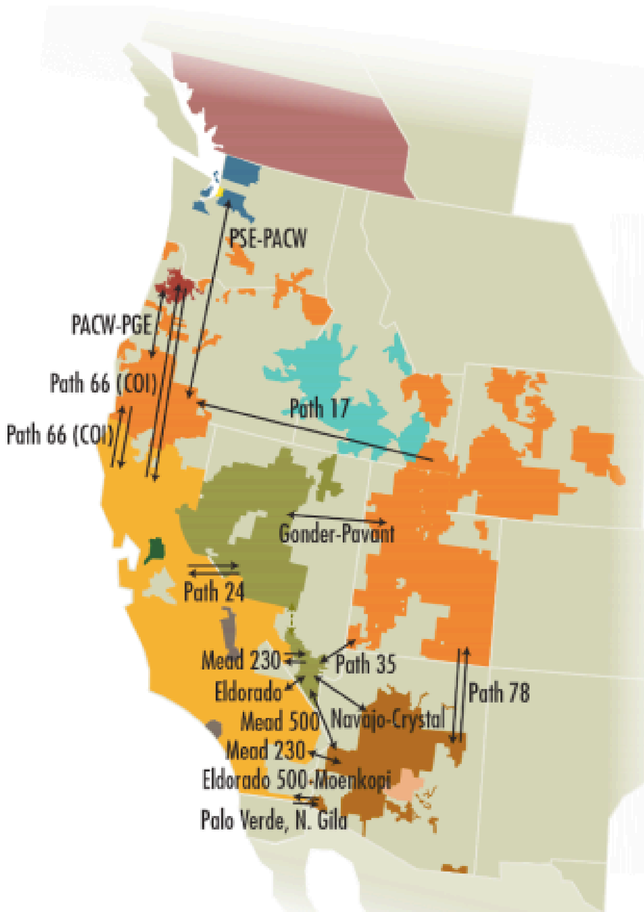
- The EIM limits the amount of balancing action necessary, since the net difference between day-ahead schedules and real-time is typically smaller across multiple regions than in an individual balancing area. EIM also allows participating utilities to use others' cheaper resources to fulfil their own demand and export surplus energy to others, including renewables that would otherwise be curtailed.
- CAISO tried to bring together the Western U.S. into a more coordinated balancing arrangement before, but other utilities were concerned about losing autonomy. EIM avoided this by allowing each participant to control their own scheduling decisions, and only in real-time pooling together to optimize any divergences across their schedules.
- Participating balancing areas are still responsible for scheduling: submitting balanced schedules, maintaining their own required operational reserves, and maintaining adequate ramping capacity to support real-time flexibility needs.
- CASIO, on the behalf of EIM, is responsible for real-time dispatch. To allow for CAISO to technically manage this coordination, all Balancing Authority Areas' (BAA) updated their dispatch to CAISO's standard, which also resulted in large savings for individual BAAs through improved dispatch.
- EIM has, as of 2017, reduced energy costs by $330 million since its implementation in 2014, curtailment fell by 161,097 MWh (more than 42% of total CAISO renewable curtailment) in 2017, and the need for additional ramping in EIM footprint was reduced by about 450 MW (roughly 38% of total ramping need) monthly in 2017 by allowing generators outside of CAISO to support intraday ramps.

2.2 Introduction

Challenges for renewable integration

With an increasing share of renewable energy, CAISO is already facing moments of curtailment where the cheapest way to balance the power system is to ramp down solar output. The primary challenge is moments of oversupply, where over 60% of instantaneous demand is being met by variable renewables, and other generators must remain on for firming, either to provide required reserves or to meet ramping needs later in the day when the sun sets. If these thermal plants remain online, they cannot ramp down below their minimum rates, and therefore cannot integrate unexpected surpluses of renewable. Under such situations, solar is the easiest to curtail, both technically and contractually, and typically CAISO pays them for the energy output they decrease. But as California plans to increase its renewable share from 29% in 2017 to 50% by 2030 , moments of curtailment will grow, as will reserves and fast-ramping capacity to support those penetrations, both adding to system costs.

CAISO identified enlarging its balancing footprint (figure 7) as the most cost-effective way to ensure adequate reserves and flexible going-forward, and also integrate more renewable energy to support their clean energy targets.



Path	Estimated Max Capacity (MW)
Path 24 *(west to east)*	100
Path 24 *(east to west)*	35-90
Eldorado	797
Path 35 *(west to east)*	580
Path 35 *(east to west)*	538
Gonder-Pavant	130
PACW to PGE	320
Path 66 *(ISO to PGE)*	627
Path 66 *(PGE to ISO)*	296
Path 66 *(ISO to PACW)*	331
Path 66 *(PACW to ISO)*	432
Path 17	400
PSE to PACW	300
Eldorado 500-Moenkopi	732
Palo Verde, N. Gila	3,151
Path 78 *(PACE to APS)*	625
Path 78 *(APS to PACE)*	660
Navajo-Crystal	522
Mead 500	349
Mead 230 *(APS <–> ISO)*	236
Mead 230 *(ISO to NVE)*	3,443
Mead 230 *(NVE to ISO)*	3,476

← one direction ↔ bi-directional

California ISO	Idaho Power Company *(entry April 2018)*
PacifiCorp	Powerex *(entry April 2018)*
NV Energy	BANC/SMUD *(planned entry 2019)*
Arizona Public Service	Seattle City Light *(planned entry 2020)*
Portland General Electric	Los Angeles Dept. of Water & Power *(planned entry 2020)*
Puget Sound Energy	Salt River Project *(planned entry 2020)*

Figure 7: Estimated maximum transfer capacity (EIM entities operating in Q1 2018)[16]

16 California Renewables Portfolio Standard (RPS), http://www.cpuc.ca.gov/RPS_Homepage/

Establishing the Energy Imbalance Market (EIM)

Enlarging balancing areas has long been a strategy to reduce operating reserves and be able to minimize the impacts of sudden changes in overall demand or supply. By aggregating load and variable wind and solar generation, the variability and forecast errors of the combined footprint is typically smaller than the sum of each BAA's[17] , the entities responsible for ensuring local supply and demand is in balance. This is because one BAA's positive imbalance may offset another BAA's negative imbalance, netting out to less divergence from the schedule, and less adjustment in real-time to match supply and demand. Historically, the benefits of combining regions together has been the reduction of uncertainty , which means less reserves would be needed to maintain the same operational standard . But in a high VRE system, it also allows a growing renewable base to be absorbed across a broader region by allowing for the export/import of unused renewable power, and spreads the ramping requirements to firm that generation across a larger fleet, reducing costs incurred by fast ramping. Other regions around the globe have also been able to integrate high levels of VRE by interconnecting with other regions.

In pursuit of this goal, CAISO together with PacifiCorp established the Energy Imbalance Market in late 2014. It is a voluntary real-time energy-only market where market participants (BAAs in the Western Interconnect who have selected to join, referred to as "EIM entities") are able to buy from other participants to settle any imbalance between their schedules and real-time conditions. They are buying from generators (known as "participating resources") located across the EIM territory. These EIM entities first conduct their own scheduling that ensure adequate supply to meet forecasted demand and procure enough reserves to meet reliability and ramping needs. Then in close-to real-time EIM entities provide the EIM with information on their current supply-demand mismatch, and the bids of any available generation capacity in their region. The EIM then takes over operation and finds the lowest cost resources to meet the net mismatch across the entire EIM territory.

Since its establishment, EIM has continuously expanded as other utilities within the western interconnection (the single synchronous territory in the Western US) have realized the benefits. EIM now includes 38 different BAAs . This links CAISO with 7 vertically integrated utilities, covering 8 states (and parts of Canada), with an additional 4 utilities planning to join by 2020 (figure 8)[18].

EIM represents a breakthrough in a decades long effort to bring the BAAs of the US's western interconnect together. Unlike prior efforts, EIM was able to overcome BAAs unwillingness to join by not pursuing a centralized ISO/RTO structure. This eased the sell-through process by 1) mitigating concerns of autonomy loss, and 2) lowering the hurdle for regulatory approval to join.

The energy imbalance structure was arrived at because it allowed BAAs to maintain the authority to make their own scheduling decisions, only using EIM to help them trade any difference between scheduled production and demand on EIM. While this helped overcome the political barrier, this structure poses unique scheduling and coordination challenges, requiring protocols to coordinate individual scheduling decisions with EIM-wide dispatch. This also requires clear definitions of roles and responsibilities to assure individual entities

---

17 PacifiCorp expanding regional energy partnerships fast fact, http://www.pacificorp.com/content/dam/pacificorp/doc/About_Us/Energy_Imbalance_Market/CAISO-PACFastFacts.pdf

18 Western Energy Imbalance Market website, https://www.westerneim.com/Pages/About/default.aspx

are not taking advantage of other EIM entities and maintaining their own obligations as BAAs.

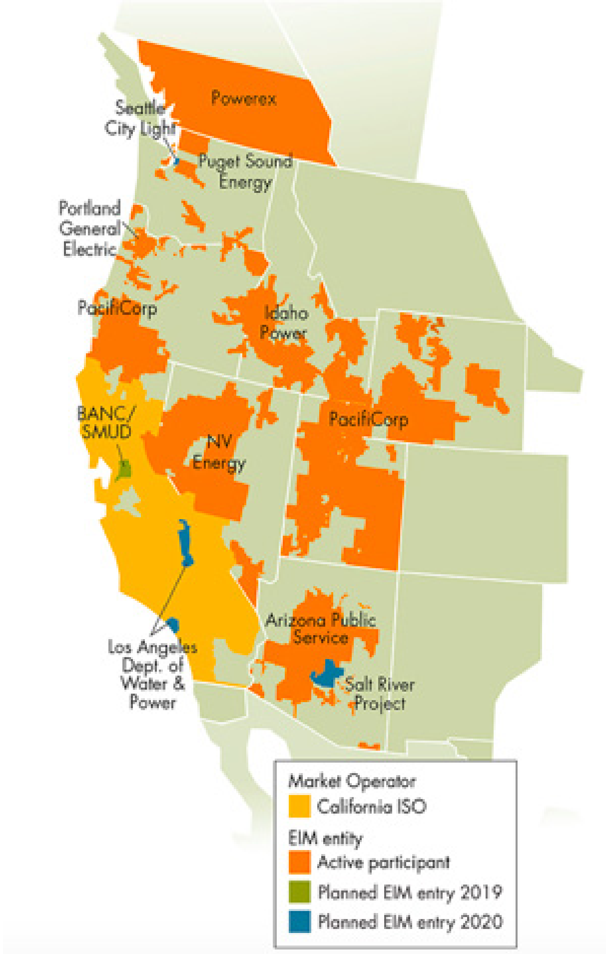


Figure 8: Western EIM active and pending participants

We focus on these two major points in this case, and explain why it may be an effective model for regions that need to:

- overcome regional protectionism and open up more cross-boundary transactions, and
- improve scheduling coordination across different balancing authorities.

2.3 Overcoming utility resistance to join EIM

The Western Interconnector's 38 BAA's have historically resisted joining a centralized market platform for fear of forfeiting some of their control over their energy decisions to California. Given California's size, surrounding utilities worried a unified market would result in their prices rising to California's high rates and that their generators would be indirectly affected by California's environmental policies. Overcoming this resistance dictated much of EIM's structure and the approach to expanding to new regions.

CAISO needed to design a system by which different system operators were able to maintain substantial autonomy over their scheduling and pricing, while still delivering the benefits of a unified balancing area. They also needed to be able to demonstrate to those actors the benefits of joining would far outweigh the costs and risks of joining.

This why EIM:

- Allows BAAs to maintain control of scheduling, allowing them to control their exposure to EIM dispatch
- Is a voluntary market, to reduce fears of lock-in
- Set up an independent governing body, to ensure larger entities did not overpower smaller participants
- Kept upfront costs low by building on existing market platforms, thereby reducing sticker-shock
- Focused on developing readily-quantified benefits, to attract new members, and ease their regulatory approval process

Scheduling Autonomy

EIM allowed its participants to define the terms of their engagement, which alleviated a lot of concerns around relinquishing dispatch control over to CAISO and the centralized price-setting methods used therein. In essence, EIM participants only participate when it makes economic sense to them. In EIM, an entity will still be in charge of creating its base schedule (similar to a day-ahead schedule in CAISO), and can select whether or not to adjust that day-ahead schedule to incorporate cheap EIM resources in real-time. The only power BAAs hands over to EIM is real-time energy dispatch to deal with deviations of load and output within the intra-hour time frame and the operation rights of reserve resources procured by BAAs. Additionally, EIM entities can determine which generation resources can be a participating generator, and how those resources offer into EIM. With this design, BAAs can make their own decisions on how much they integrate EIM into their standard operation.

Voluntary Participation

EIM was established as an opt-in market, where BAAs, or their carrier utility, were the ones responsible to decide whether or not to join EIM. This allowed them to conduct their own evaluations as to the benefits of joining, and decide the scope of their engagement. Furthermore, BAAs always had the option to leave without penalties, to eliminate fear of lock-in as a reason to not join. But one key component was if BAAs did not participate, neither could their merchant generators. This created pressure on BAAs to join since many of their generators wanted to be able to sell their unoccupied generation capacity to a wider market.

Governance

The EIM, like any ISO/RTO, requires an independent, unbiased operator that does not have assets or LSEs participating in the EIM. While CAISO is already an independent system operator, other BAAs would not be adequately represented in this structure. Therefore, the EIM has its own independent governance structure distinct from CAISO board. The selection process, overseen by EIM entities, focused on bringing in a diverse, well-qualified group to represent all perspectives, and to ensure no one entity or interest is elevated over another. This includes essential recourse functions in case certain EIM entities feel current rules need reformation to protect the collective EIM interest.

EIM does rely on CAISO's Independent Market Monitor, the authority charged with reviewing electric wholesale markets' performance, to review the EIM's function and provide this analysis to EIM's board of directors. This helps keep costs low by not having to build/pay for this function separately.

Builds on Existing Dispatch Program

The costs of expanding a current dispatch system is much lower compared to building a new market platform from scratch, and therefore it was determined they needed to use CAISO's dispatch system as a foundation. This helped lower this initial barrier to entry, as individual utilities/BAAs needed to demonstrate to their own regulatory authorities the net benefits of joining EIM.

To convince other BAAs to join, EIM needed to demonstrate clear cost reductions to participants. Developing a strong economic case was essential to convincing regulators in different regions to approve the upfront expenditure to join EIM.

Co-benefits from joining EIM

In order to join the EIM, utilities needed to update their dispatch systems to include locational information (at the nodal level) and reduce the timescales of scheduling and dispatch instruction. CAISO, as a market-based exchange, invested a lot of time and resources on improving system efficiency and grid operational systems. By joining EIM, BAAs are able to update their old systems and receiving training on CAISO's platform for a fraction of the cost of doing it themselves. Some utilities have found that the upfront cost paid for system updates and interconnection into EIM delivered material cost savings by optimizing their own dispatch, an update they would not have been able to gain regulatory approval for without having the other quantified EIM benefits to justify the investment. Additionally, many of these vertically integrated utilities did not have good visibility into their generator's true costs. By having to formulate bids to participate in EIM, individuals BAAs and utilities identified better patterns to operate their own generators, and exposed more opportunities for profitable exchange in the EIM beyond the expected benefits originally quantified.

2.4 Designing EIM's dispatch processes

While real-time power market operation is not new, EIM had to find a mode to coordinate across different BAAs—each having different scheduling processes and timelines—to support good information collection, active participation, and secure dispatch. To support this coordination, EIM needed to establish:

- A coordination timeline that occurs close enough to real-time delivery to have a more accurate picture of grid conditions, while leaving enough time for individual BAAs to redispatch to take advantage of cheaper energy within EIM.
- A clear division of roles and handoffs, to ensure individual BAA actions do not conflict with EIM's operation, particularly in regards to maintaining adequate reserves and flexible capacity.
- Dispatch systems that can coordinate many actors' information delivery across a wide-georaphic area, and quickly verify secure dispatch solutions

Scheduling and coordination timeline[19]

The market coordinate timeline has three main processes to coordinate between EIM's central market dispatch and BAA's scheduling, dispatch, and real-time situational awareness:

1. **EIM collects current situation data** (75 mins prior to delivery): BAAs submit their current base schedules, participating resources' availability and price points, and expected real-time demand. At 75 minutes out, BAAs can better estimate what the gap between their day-ahead schedules and real-time demand will be, and provide EIM enough visibility to deliver a redispatch solution.
2. **BAAs make redispatch decisions** (between 75 minutes to 40 minutes before delivery): EIM runs the system sufficiency evaluation, publishes the result, and lets BAAs adjust base schedules. The evaluation and reschedule process will be conducted twice so BAAs have an opportunity to optimize their base schedule. BAAs have option to ramp down their day-ahead scheduled generators to utilize cheaper energy resources that are available at other EIM entities. And EIM ensures there are enough flexible capacity standby for dispatch in real time frame later and BAAs do not over-rely on the EIM for balancing.
3. **EIM conducts real-time dispatch** (37.5 minutes before delivery through delivery): EIM takes control, runs the real-time market and sends dispatch signals directly to participating resources on a 15-minute basis, setting an average clearing price across those 15-minutes[20] for each node. From this time point, the decision-making power moves from BAAs to EIM, and BAA resources only follow their final base schedules and EIM's real-time dispatch signals.

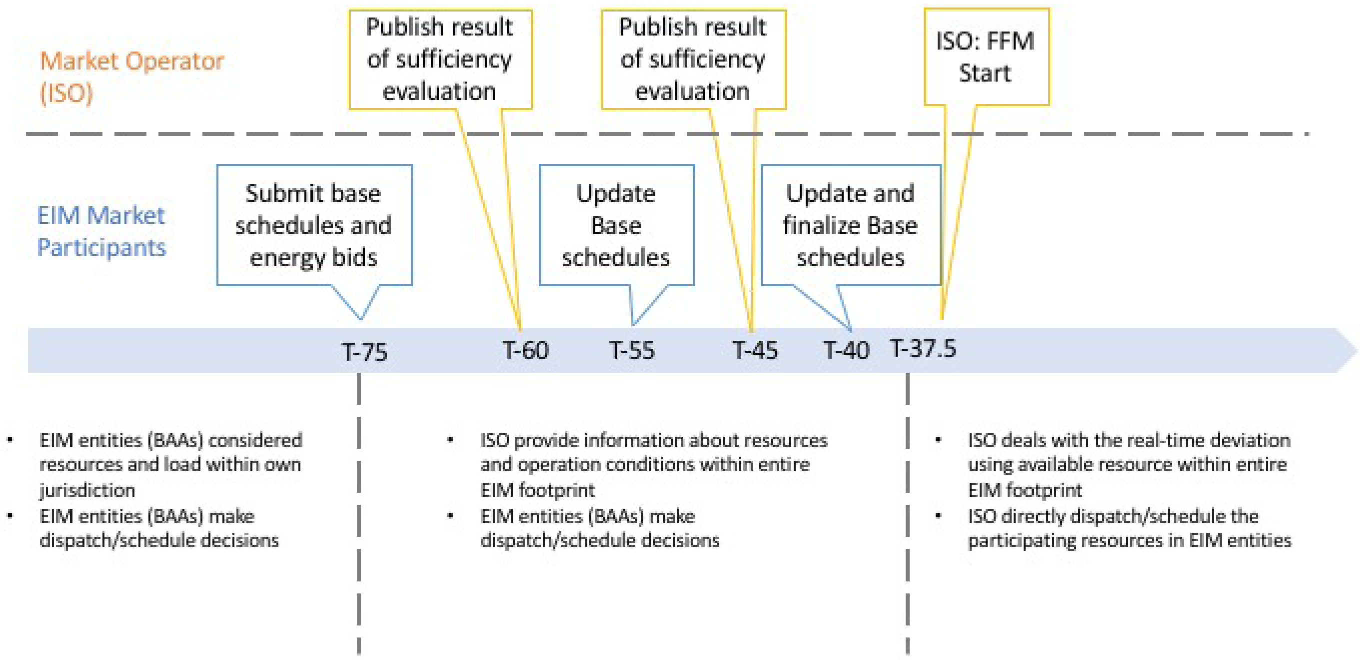


Figure 9: EIM Timeline.

19 The basics of the real-time market, https://www.westerneim.com/CBT/NEW%20The%20Basics%20of%20the%20Real-time%20Market/story_html5.html

20 CAISO runs 15-minute market in real-time market framework. The market clearing process running at 37.5 minutes before the delivery will decide the dispatch schedule of first 15 minutes of final delivery hour.

The EIM is not involved in the dispatch decision making process until 75 minutes before delivery and cannot directly make dispatch decisions until 37.5 minutes before delivery (figure 9). Before that, BAAs are wholly responsible for scheduling and dispatch decisions. This coordination process allows for some of the benefits of a centralized market to make its way into this system without requiring BAAs to cede their entire resource base to a central market for dispatch and scheduling.

Division of Roles and Responsibilities

In order to avoid uncoordinated dispatch, scheduling, planning, and reserve procurements, responsibility for most functions remain with BAAs until the EIM timescale described above. Each BAA still has the responsibility to:

- Establish and adjust any dispatch schedule before the start of real-time markets
- Manage their transmission system, including setting the transmission tariff
- Meet its own reliability obligations and operate its own ancillary service procurement processes
- Manage system planning including load forecasting, and setting power interchange plans with neighboring BAAs
- Carry adequate flexible capacity to cover uncertainty in load andVREforecast
- Collect and coordinate energy bids from participating resources to submit to EIM

EIM plays a role in ensuring that BAAs are maintaining these responsibilities, particularly enforcing that they submit balanced base schedules to EIM (not chronically under estimating), maintaining adequate operational reserves, and carrying adequate flexible capacity to meet expected intra-hour ramps and expected uncertainty factors. This monitoring is conducted by the market monitor, and there are penalties for non-compliance, including fines or limiting the volume of energy allowed to be traded in the EIM to the non-compliant BAA. In particular, the flexibility requirement uses an hourly evaluation, called the flexible ramping test, to verify EIM participants are carrying flexible capacity proportional to their contribution to uncertainty in EIM's overall load forecast, as well as its net import/export capabilities. Many BAAs have failed this test recently, which has pushed EIM to consider expanding ancillary service markets for flexibility reserves to the entire footprint.

Fast IT Systems

A good portion of start-up costs for the EIM is installing the right IT systems, which are able to gather appropriately granular data from BAAs, run grid analysis to ensure a timely and secure solution, and be able to coordinate many actors simultaneously adjusting their schedules.

It required a two-phase model run: one using each BAA's original base schedule, then a second based on updated base schedules. From the moment they receive the first result from CAISO (75 mins prior to delivery) to the moment CAISO takes over real-time dispatch (35 mins prior), BAAs must have time to:

- Run their own economic dispatch software to readjust their base schedule
- Transfer those updated schedules to EIM
- Distribute updated schedules to their plants

The time to complete these activities dictated the timescales for EIM, including the time needed to run the EIM's dispatch models. Other BAA's transmission network representation and reserve sufficiency evaluation methods needed to be improved, too. BAAs needed to advance to nodal network representation to better reflect transmission constraints and congestion during the EIM resource selection process. And while each individual BAA was still responsible for its reserve margin maintenance, a system sufficiency evaluation software was necessary to assess more globally to assure EIM it could safely operate under certain redispatch scenarios.

2.5 EIM's Impact: Decreasing costs, flexibility needs, and VRE curtailment

Research was conducted upfront to estimate the benefits to each participant by joining to help encourage them to sign-on to the EIM platform. Major benefits quantified were:

- Operation optimization: More advanced dispatch system and utilized cheaper resources available in the market
- Emissions reductions: Reduce renewable curtailment and use more cleaner resource
- Flexibility needs reductions: reduce the need for ramping and reserve

The upfront cost to conduct feasibility tests, set-up legal structures, and developing IT systems, as well as ongoing operation costs associated with EIM market operation and data management were far outweighed by the reductions in power and reserve procurement costs (figure 10)[21].

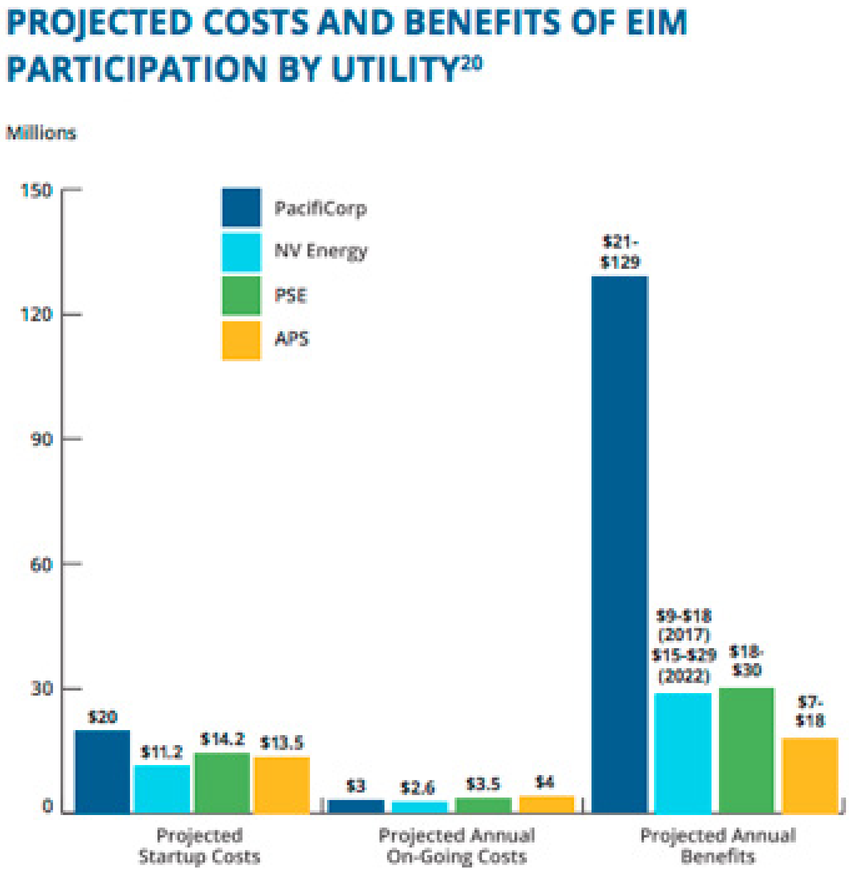


Figure 10: Actual gross benefits on an annual basis.

21 The growing western EIM: The economic, environmental, and energy security benefits of an expanding market, http://windenergyfoundation.org/wp-content/uploads/2017/11/ARA_WEF-The-Growing-Western-EIM.pdf

EIM's primary way of reducing costs is through participants rescheduling to take advantage of the cheapest resources available in real-time across the entire market footprint. Other EIM economic benefits that are not easy-to-quantify or directly attributed to EIM operation are not included explicitly in these benefits (e.g. improved BAA scheduling, deferred investment in new reliability assets, decreased wear and tear on assets having to operate more flexibly).

Although benefits for individual utilities may fluctuate due to weather conditions or other grid conditions for certain year, the effectiveness of EIM maintains consistent across all years. Since EIM was founded in Q4 2014 to Q2 2018, the cumulative gross EIM benefit was $401 million in cost savings (figure 11)[22].

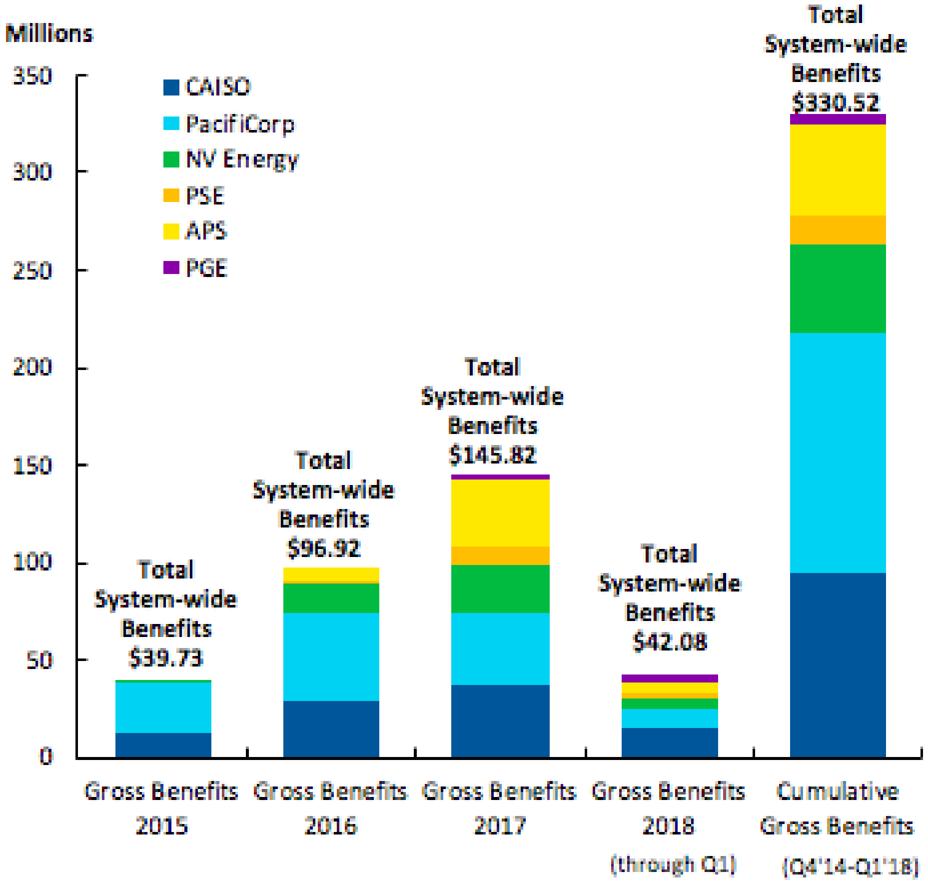


Figure 11: Actual gross benefits on an annual basis.

Fully utilizing renewable generation and reducing energy procurement cost

Frequently renewables are the available, cheap energy being traded on EIM, which means that many renewables that may not have been absorbed previously are now being integrated. EIM over its operation lifetime has reduced curtailment by 715,405 MWh and avoided an estimated 306,122 metric tons of CO_2 (table 3[23]). The amount of increased renewable integration is VRE that were not covered under BAA's base schedules, but dispatched under EIM[24]. Curtailment reduction, therefore, is high in years where there is high availability (and volatility) of renewables along with high levels of hydro availability (meaning that many VREs, typically wind, could not be used when hydro was operating under must-

22 Aggregate data from quarterly EIM benefits report

23 Ibid.

24 Western EIM Benefits Report First Quarter 2018, https://www.westerneim.com/Documents/ISO-EIMBenefitsReportQ1_2018.pdf

run). Unquantified cost savings includeVREeconomic curtailment payments (only done in CAISO) and the benefit of increased generation by these renewables.

Year	Quarter	MWh	Eq. Tons CO2
2015	1	8,860	3,792
	2	3,629	1,553
	3	828	354
	4	17,765	7,521
2016	1	112,948	48,342
	2	158,806	67,969
	3	33,094	14,164
	4	23,390	10,011
2017	1	52,651	22,535
	2	67,055	28,700
	3	23,331	9,986
	4	18,060	7,730
2018	1	65,860	28,188
	2	129,128	55,267
	Total	**715,405**	**306,112**

Table 3: Eq Tons of CO_2 avoided 2015-2018 from trading on EIM

Reducing operational cost incurred by flexibility

Even though EIM is an energy-only market, and BAAs still are responsible for procuring ancillary services to meet their reserve and reliability obligations within their operating area, EIM helps reduce the planning and operating reserves and flexibility that an individual BAA must carry in real time to manage load. It does this in three ways:

1. Reducing the overall need for flexibility and reserves on the system by balancing over a wider area.
2. Spreading ramping need over more assets, so the need could by satisfied by the aggregation of many plants performing small cheap ramping operations, instead of relying on few local assets conducting expensive, flexible operations.
3. Reducing the number of plants “waiting in the wings” at minimum output to fill the gap left by solar when the sun sets.

The reduction of system flexibility needed within EIM is substantial (figure 12)[25]. The percentage of flexibility ramping needs saved varies from 12%-26% of the total flexible capacity required for daily ramps in the initial period of EIM establishment, and increases to higher than 30% after December 2015, normally ranging between 35%-40%[26]. The percentage of flexibility needs reduced can reach as high as 50%, which was a historic maximum occurring on October 2016[27]. This reduction in flexibility ramping needs is equal to providing hundreds of MWs of flexibility resource to the system.

25 Aggregate data from quarterly EIM benefits report

26 Ibid.

27 Ibid.

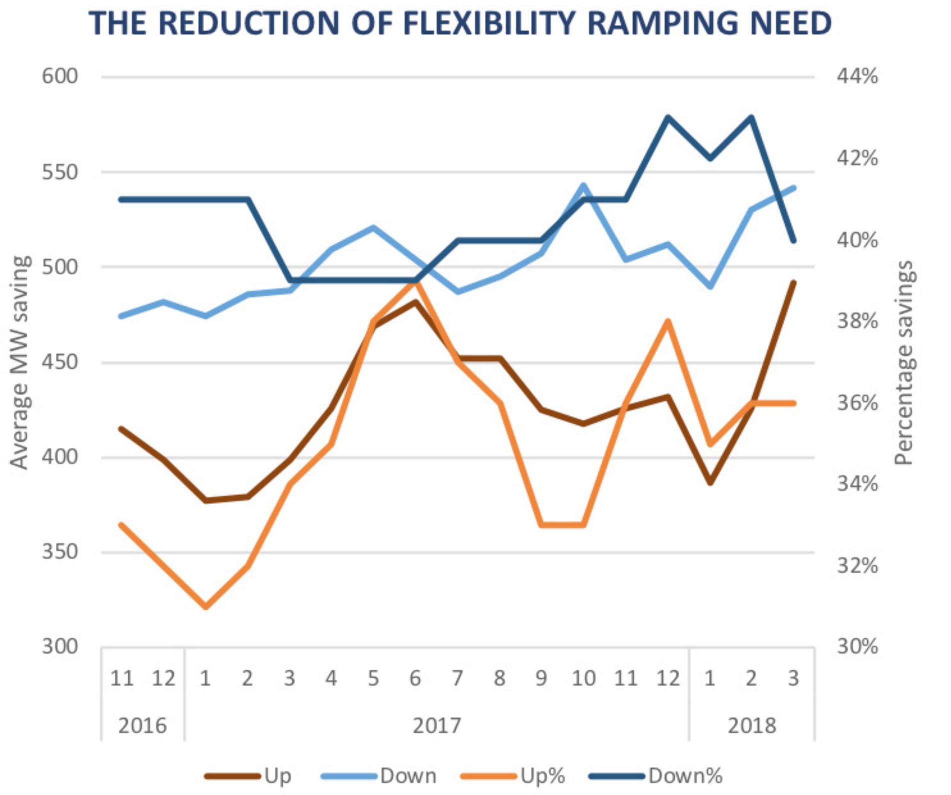


Figure 12: Flexibility ramping need reduction

2.6 Future EIM - Fully integrated regional market

While the benefits of EIM are substantial, the benefits of having a fully unified market are expected to be greater, particularly by reducing the cost to carry operating reserves (figure 13). This is why EIM is looking to pilot both day-ahead and ancillary service markets across the entire EIM footprint.

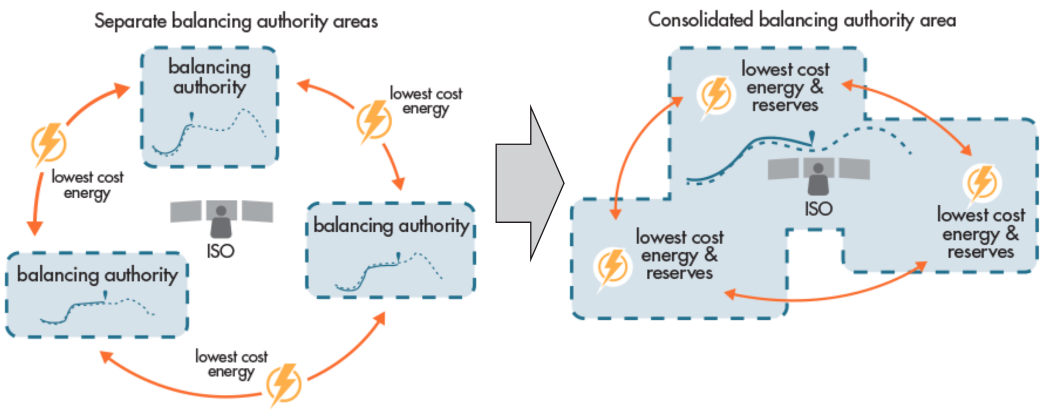


Figure 13: Current EIM and future full integrated market

BAA's would only purchase from the day-ahead market when it was economic, and likewise, when scheduling to meet their flexibility and reserve requirements, could select from a broader pool of resources.

In the case of EIM, many BAAs still have concerns that by joining a centralized market, their prices will rise to the levels of West Coast regions, especially those regions rich in cheap hydro, even though in most fully centralized markets direct bilateral trade is still allowed, providing a mechanism to keep low hydro prices local. While day-ahead markets are expected to be a slower move, ancillary service markets are expected to scale quickly, especially as California has a growing need to procure additional flexibility.

2.7 Relevance in a Chinese context

Path dependence has a clear impact on the integration of different power markets. Different power market setups result in different dispatching rules, different operation routines and different supporting IT systems. Without fully considering the status quo, the integration of power market would become a futile effort. The success of EIM partly attributes to the soft methodology it employs. EIM does not requires the overhaul of current market setups, and it respects the key concerns of different stakeholders. Particularly, it allows different balancing authorities (which are similar to local dispatching centres in China) to retain most of their control of scheduling. The benefits of EIM mainly thanks to the collective transfer of real time dispatching from different balancing authorities to EIM 75mins before the physical delivery. Due to the fact that the forecast errors and variability of generation and consumption in different regions usually cancel out each other, the integration of dispatching regions will lead to the reduction of the need to activate fast ramping up and down.

One of the efforts in China to integrate provincial power markets is the cross-regional incremental spot market (CRIPM). It leverages on the available transmission capacity on cross-provincial HVDC lines to transport extra clean power from the west to the east. CRIPM is different from EIM in many ways: 1) it is mainly operated day-ahead, and real-time imbalance is still adjusted inside each of provincial grid; 2) the provincial power grids retain all the control of scheduling. Thus, the benefits of CRIPM mainly come from using the margins on cross-provincial HVDC lines to transfer cheap wind and solar power to the load centres.

3. Ancillary Service Markets evolving to support high RE

3.1 Key messages and takeaways

- The traditional modes by which ancillary services were deployed to ensure reliability are evolving to accommodate rising shares of VRE, focusing more on flexibility and supporting new technologies, especially batteries, demand response, and renewables via smart inverters.
- Energy markets can provide some of this flexibility through specific flexibility ancillary service products, and should be optimized first, including co-optimizing energy and ancillary service markets to fully discover existing asset flexibility and the right prices for those services.
- California and Texas have some of the highest penetrations of VRE in the U.S. and are looking to the future to identify new products to address their challenges:
- California's Flexible Ramping product was designed to address inter-hour ramping needs. It improved the scheduling of generators to ensure sufficient ramping capacity is available, without eating away regulation services needed to maintain system reliability.
- Texas's Fast Frequency Response product was designed to acquire enough synthetic inertial response to support a high wind system response to a contingency event. While successfully demonstrating technical feasibility and effective market design, it was ultimately not adopted due to concerns of it not being necessary in the near-term.

3.2 Introduction

Definitions and Roles

In order to reliably and securely operate the grid, grid operators must constantly maintain the balance between power generation and consumption in real time, but also must maintain voltage stability, frequency stability, and be prepared to recover from major grid disruptions[28]. System operator must procure special services and functions beyond energy to achieve this balance, which are referred as ancillary services (AS). While definitions and terminology vary drastically across regions, we provide general definitions which we will use throughout this discussion[29]:

Non-event Reserves: corrects short timescale differences between supply and demand to maintain system frequency

- **Regulation:** automatic response to cover a limited range of output change. Typically required of generators.
- **Following:** manually dispatched resources to cover larger unscheduled changes in supply/demand. Participants are typically paid to have their capacity available, and paid for the service provided.

28 https://www.logicenergy.com/what-are-ancillary-services-and-why-do-power-grids-need-them/

29 Utilizes the framework from NREL's paper on operating reserves: https://www.nrel.gov/docs/fy11osti/51978.pdf

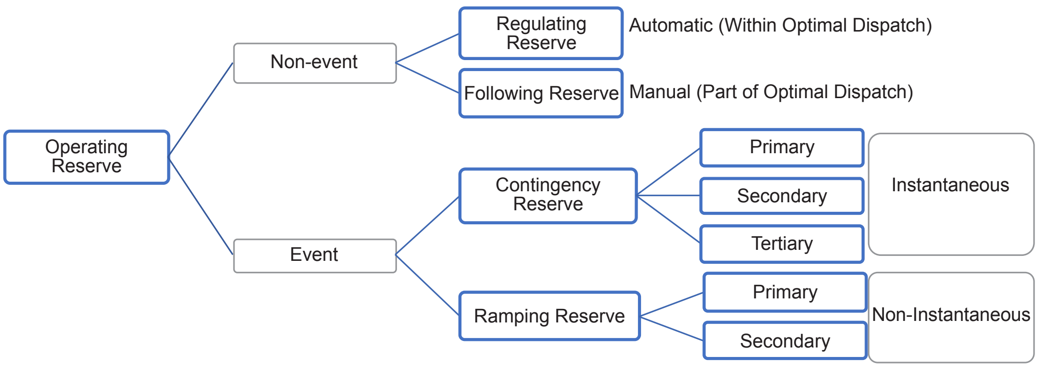


Figure 14: Event and non-event operating reserve

Event Reserves: resources that can be quickly dispatched in the case of an unexpected loss of supply

- **Primary Reserves** *(also known as frequency response reserves)*: Immediate response that helps stabilize the grid after a loss of supply for long enough for additional reserves to replace the lost supply
- **Secondary Reserves** *(usually further divided into spinning reserves and non-spinning reserves depending on response time)*: capable covering lost supply within 10 minutes of the event. Typically provided by generators already operating or able to start up quickly. Demand-side resources can also provide spinning reserves if they are able to reduce their load within the required time frame.
- **Tertiary Reserves** *(also known as supplemental reserves)*: more economic resources, but with slower response times (30- 60 mins), coming online to fill the capacity lost from the event

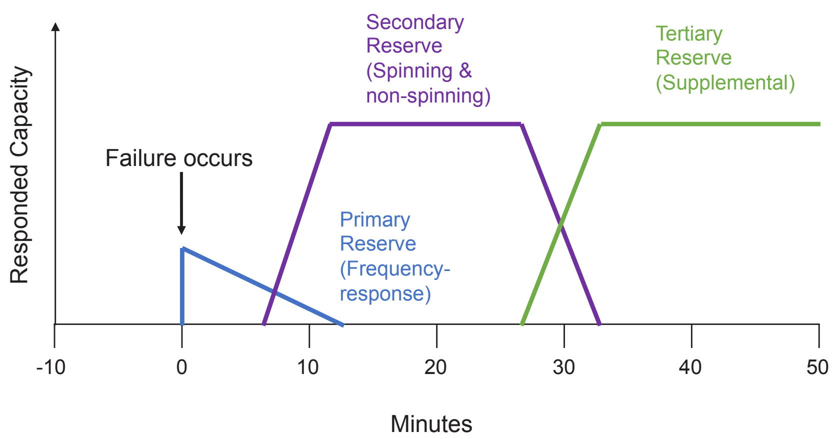


Figure 15: Illustration of event reserve response[30]

Increasingly, a new reserve classification is emerging in these markets, referred to generally as ramping reserve that can load-follow or ramp for sustained periods with substantial ramp mileage to cover for sudden losses in VRE.

These products are generally procured at volumes according to reliability standards established by NERC (typically 15%). These are required to be procured by the balancing authorities, who may procure them by several methods:

30 https://eta-internal-publications.lbl.gov/sites/default/files/spinning-reserves.pdf

- Require each generator to provide as a part of their normal dispatch (typical of a vertically integrated utility and some services also considered complimentary with DA & RT dispatch)
- Administratively set prices for the service and pay those generators contributing
- Creating a market for competitive selection of AS providers

We will predominately focus on the market-based approach, since the most aggressive renewable integration is typically happening in market territories and by defining AS market products, the role of these service definitions in maximizing flexibility is better exposed.

	Spinning Reserves	Non-spinning Reserves	Regulation
CAISO	Spinning	Non-spinning	Regulation-up Regulation-down Regulation Mileage-up Regulation Mileage-down
ERCOT	Responsive	Non-spinning	Regulation-up Regulation-down
ISO-NE	Ten-minute Synchronized	Ten-minute Non-synchronized Thirty-minute Operating	Regulation
MISO	Spinning	Supplemental	Regulation
NYISO	Ten-minute Spinning Thirty-minute Spinning	Ten-minute Non-synchronized Thirty-minute Non-synchronized	Regulation
PJM	Synchronized	Primary	Regulation
SPP	Spinning	Supplemental	Regulation-up Regulation-down

Table 4: Overview of the ancillary services offered by CAISO and ERCOT[31]

Increasing renewable penetration creates more fluctuations in grid supply, and therefore a greater need for ancillary services to maintain reliable operation[32,33]. The most typical operations discussed in this case are:

- Frequency Regulation: generators increase or decrease output (known as ”regulation up” and ”regulation down”) to maintain system frequency. As renewable output varies moment by moment, the need for frequency regulation increases, specifically the ramp mileage required to cover changes in renewable output. This has largely been addressed by shortening the timescales by which real-time markets clear resources, and by optimizing energy market and ancillary service scheduling (discussed in section 3.3). For sustained ramping, this is addressed by introducing new ramping products, discussed in section 3.4.
- Over-frequency events: this typically happens slowly and grid operators can respond by reducing output from some generators. Large, unpredicted surpluses of solar energy

31 Definitions per each service vary between different ISO/RTOs and although those nuanced differences do result in material differences in system operation, for the purposes of this discussion, the broad classifications of spinning, non-spinning, and regulation should be considered parallel with secondary contingency reserves, tertiary contingency reserves, and a combination of regulating and ramping reserves, respectively.

32 http://www.caiso.com/Documents/2019FinalFlexibleCapacityNeedsAssessment.pdf

33 https://www.caiso.com/Documents/FinalFlexibleCapacityNeedsAssessmentFor2017.pdf

create these situations, and grid operators must either ramp down of other generators or curtail RE. This has been dealt with by aligning demand with renewable availability (discussed in the DR case, section 4), and by implementing economic curtailment (discussed in section 3.3)

- Under-frequency events: often due to the loss of a large source of electricity. As renewables increase and more flexible operation occurs, the tendency to have generators trip off increases and causes more under frequency events[34]. The steps necessary to recover from such an event also become more difficult with higher amounts of renewables online. This is discussed later in section 3.4.

Many of the system operators seeing an increase in these types of events have taken a few steps to refine their market processes to improve both the economics and the effectiveness of their deployment. These include:

- Requiring higher levels of flexibility from energy markets, without calling on ancillary services
- Co-optimizing energy and ancillary service markets ensure sufficient flexibility at least cost
- Adjusting the definitions of AS to specifically define the type of flexibility required to operate the system

3.3 Energy Market Optimization

Energy markets already create flexibility

Since energy markets are based on generators marginal costs, renewables are almost always selected and other more expensive generators must adjust their output around them. Energy markets already require all generators to ramp up and down between each clearing period, which as real-time markets speed increases (most now clear every 5 minutes), this flexibility is naturally extracted. Generators typically submit their technical ramping restrictions with their bids, and system operators must simply abide by those during dispatch.

Generators' profits are always non-decreasing with shorter time-scale market: When there is a need for more power than scheduled in the system, the real-time price is typically higher than the day ahead price, so generators want to ramp up and sell more power. When there is a need to ramp down, the real-time price is normally lower than the day ahead price, and generators can ramp down and sell some of their scheduled power back at real-time price and keep the difference between day-ahead price and real-time price. But flexible operation is always more expensive than steady state operation, and unless these price differences are substantial enough, generators will prefer not to ramp.

This is why ancillary service markets were created, in order to find the resources most suited to providing this flexibility/reliability, and then paying them for that. Since AS are typically provided by reserving some of units' capacity, they are typically also paid for their opportunity costs in the energy market for providing those services.

34 https://www.nrel.gov/docs/fy12osti/55433.pdf

Co-optimizing energy and ancillary service markets

Historically, energy was procured first, then AS[35]. This resulted in situations when a unit committed in the energy market as a least cost supplier had no capacity available to bid into the AS markets. But another generator, offering only slightly more expensive energy, but offering much more expensive AS, ends up providing AS, and the entire system cost ends up being higher with this dispatch pattern. Instead, scheduling a slightly more expensive generator in the energy market, but saving the much cheaper ancillary service provided reduces total costs substantially.

Another consideration is how to cooptimize these dispatch decisions with future dispatch needs (similar to how start-up costs are factored into scheduling decisions). The cheapest resource mix to meet flexibility and energy in an isolated hour (n), may not be when you take its economic impact on the resources that then must be selected in hour n+1 (or 2 or 3) into account. For example, dispatchers may need to ramp generators up, but if significant upward ramp reserves are required in the following hour, it's more cost effective to ramp a more expensive generator up, to avoid the more economic generator being capped at its maximum output.

This co-optimization influences generator's bidding strategies for energy and AS markets. If generators expect they will be on the margin, they could bid more competitively in regulation up products to increase their chances of commitment. These strategies become more involved as renewables increase in prevalence. If they are a marginal generator anticipating wind and solar to be under forecasted, they may bid more competitive in energy markets to be fully dispatched, so they can still be eligible for downward ramp services. Renewables can also bid into these services, and frequently play a prominent role in downward ramp products by selecting to curtail their energy. In California, most downward regulation services during the day are provided by solar[36] and has been seen to reduce costs against other dispatch options in those circumstances.

While this may sound like market manipulation, it ultimately results in cheaper energy and flexibility through more competition to provide these flexible services, as well as real revelation of what plant's limits are. A 2012 research paper used an example plant with a minimum load of 40% that can fast start and fully load within ten minutes. Using 2012 NYISO day-ahead price, the paper showed that the generator could earn 17% more if it bid in such a way to maximize its participation in reserves, despite reducing its capacity factor from 80% to 64%. Despite these additional market payments, this resulted in net lower costs for NYISO, since it reduced its cost to maintain operating reserves[37].

3.4 New Ancillary Service Product Definitions

Once these more fundamental pieces are in position, system operators typically turn to refining their AS suite to dial in the products to provide the reliability and flexibility services they need. Here, we focus on two regions in the U.S. where ancillary service markets are being reconsidered and revised to address their specific flexibility challenges, both cause by increasing shares of renewable energy on the grid, but also serve the other three objectives

35 AS commitment is still done sequentially in parts of Europe

36 Solar is more likely to provide downward ramp services than wind, because wind's subsidy is paid per MWh generated, where solar is an investment tax credit paid at the time of construction.

37 http://www.consultkirby.com/files/PowerGen-2013_The_Value_of_Flexible_Generation_Nov_2013.pdf

for system operators as well. These cases are California's Flexible Ramping product designed to address inter-hour ramping needs, and Texas's procurement of more fast-acting resources for frequency control purposes due to a decline in system inertia due to wind.

California's Flexible Ramping Product

Need for fast ramping

California's high penetration of solar energy along with the high demand peak in the late afternoon has created a challenge known as the "duck curve" in which solar production begins to taper exactly as demand is on the rise (see section 1.5). This has resulted in assets having to ramp at a faster rate and over greater ranges of output. Current[38] solar and wind penetration levels have only necessitated about 10 to 12 GW of flexible procurement in CAISO. However, with California's rapidly scaling renewable portfolio standard which reaches 50% in 2030, flexible resource adequacy requirements would rise by roughly 17 GW under a scenario of equal growth in wind and solar[39].

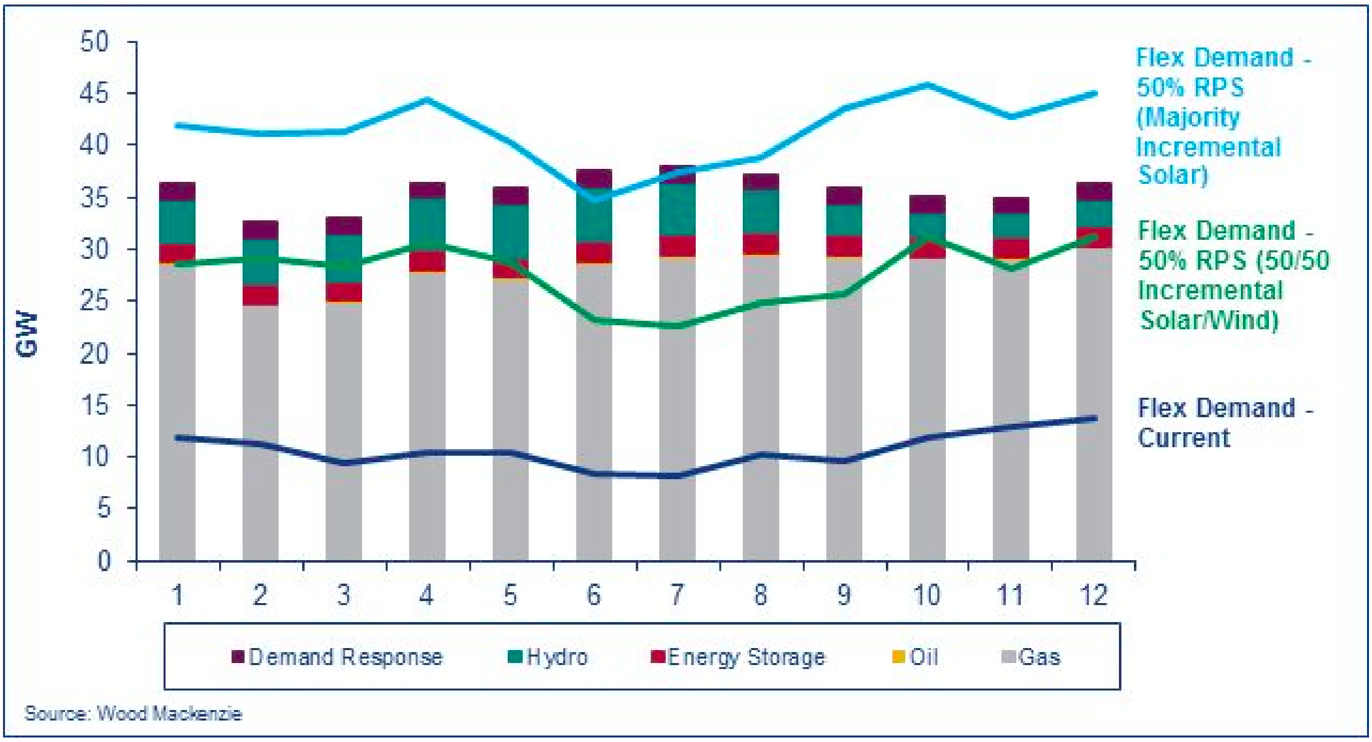


Figure 16: CAISO Monthly Flexible Demand Scenarios vs. Current Flexible Supply. Horizontal axis depicts months from January to December.

Dispatchers have begun scheduling differently to accommodate, because often the units scheduled under day-ahead markets (DAM) lack sufficient ramping capability and flexibility to handle the afternoon ramp up, and the capacity available to be dispatched in real-time is also insufficient to cover fluctuations in solar/wind between different 5-minute dispatch periods.

Flexible Ramping product

California created their flexible ramping product (FRP), which allowed CAISO to co-optimize

38 Reference based on year 2017

39 https://www.greentechmedia.com/articles/read/can-california-conquer-the-next-phase-of-renewables-integration#gs.FpkaZKw

their flexible capacity scheduling with their energy market. First, CAISO implemented a flexible ramping sufficiency test to identify how much flexibility was required, considering things like renewable and demand forecast errors and contingencies that resulted in generators previously scheduled being unable to ramp. After several years trial operation using an administrative pricing mechanism, on November 1st 2016, CAISO implemented two market products in the 15- and 5-minute markets: Flexible Ramp Up and Flexible Ramp Down. In addition, CAISO extended the existing flexible ramping sufficiency test to the entire EIM footprint, to ensure feasible ramping capacity for all real-time interchange schedules.

The result of this increased flexibility has been less out-of-merit order dispatching (lower cost), minimizing the use of reserve margins for planned excursions (greater reliability), and less solar curtailment (less emissions). While the flexible ramping product may be used to attract new flexible resources when the system is approaching a shortfall, given today's sufficiency levels, these flexible incentives also help prevent retirement of existing critical flexible capacity in the face of solar-induced revenue erosion in wholesale markets[40]. Plants that are typically scheduled for these flexibility ramps are often plants already on the margin and have been struggling to meet their going forward costs. Due to the new incentive provided by the flexible ramping product, they have further refined their operation to be as flexible as possible, even two-shifting operation (starting up and shutting down twice in a day).

Total net payments to generators in the ISO and energy imbalance market areas for providing flexible ramping capacity in 2017 were about $25 million. As market players better understood the products and markets became more competitive, monthly total payments decreased during the year from around $3 million per month between January and May to around $1.4 million between June and December. Although flexible ramping payments increased with the implementation of the flexible ramping products, payments per megawatt-hour of load remained low, with average net payments per megawatt-hour of load were about $0.07/MWh during 2017 (figure 17)[41]. For comparison, payments for ancillary services in the ISO were about $0.75/MWh of load during the same time period.

Whether FRP reduced the overall AS-energy market cost is hard to quantify due to the difficulty of controlling for all the variables influencing electricity prices, FRP implementation issues, and an increase in operating reserve requirement in 2017. Ancillary service costs increased to $0.75/MWh of load served in 2017 from $0.52/MWh in 2016. This represents an increase from about 1.6% of total wholesale energy costs in 2016 to 1.9% in 2017, which are the highest yearly values since 2011, both as a percentage of wholesale energy costs and per megawatt-hour of load[42]. While the increase in reserve requirement drove up the volume and price for procuring regulation, the need for regulation seemed to decrease with the implementation of FRP in 2017 (figure 18), however this result suffers from a calculation error which resulted in an underestimate for the need of upward flexible ramping capacity during some ramping intervals[43]. The error was corrected in February 2018 and these future results will be necessary to verify FRP's effectiveness.

40 https://www.greentechmedia.com/articles/read/can-california-conquer-the-next-phase-of-renewables-integration#gs.FpkaZKw

41 http://www.caiso.com/Documents/2017AnnualReportonMarketIssuesandPerformance.pdf

42 Ibid.

43 http://www.caiso.com/Documents/2017AnnualReportonMarketIssuesandPerformance.pdf

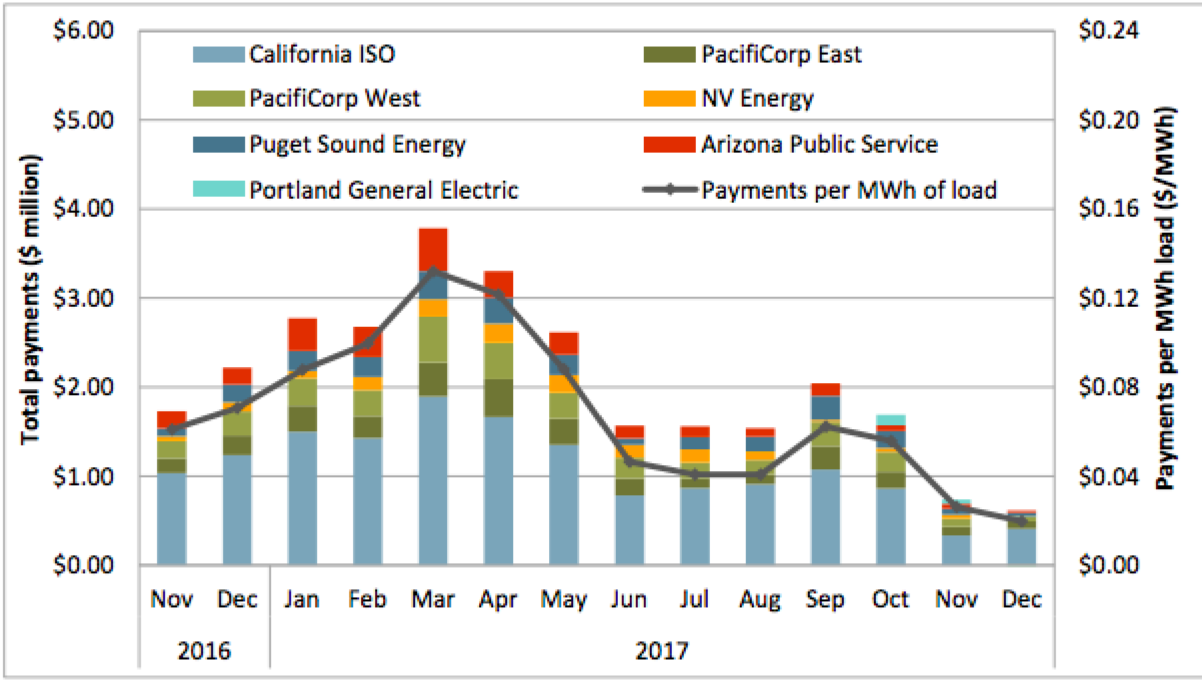


Figure 17: Monthly flexible ramping payments by balancing area

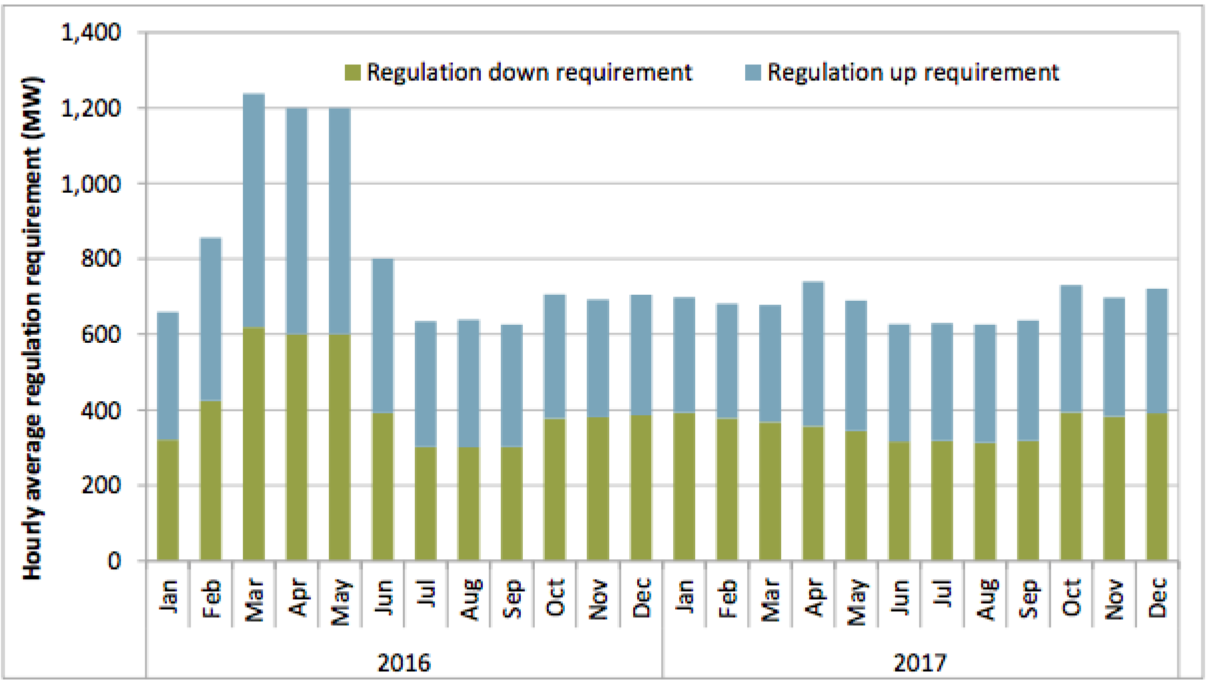


Figure 18: Monthly average day-ahead regulation require.

Beyond FRP's integration into real-time markets, CAISO also incorporated flexibility needs assessments into its resource planning process. It requires utilities to procure sufficient flexible resources for future delivery years and ensures flexible resources meet technical criteria for ramp speed and response time and are required to be available and bid into the CAISO energy markets[44].

44 https://www.greentechmedia.com/articles/read/can-california-conquer-the-next-phase-of-renewables-integration#gs.zHycEoc

RE generation forecast improvement to transit flexible ramping need

As FRP is specifically designed to deal with the flexibility incurred due to forecast errors, increasing renewable forecast accuracy will directly reduce the need of FRP, and more broadly reduce the need for other flexibility. The benefit of wind forecast improvements are well proven by various U.S. researchers[45,46]. Generally, more accurate forecasts can[47]:

- Reduce reserve levels: Regulation reserve; Flexible/load following reserve
- Improve unit commitment and dispatch efficiency: Utilize the least expensive units; Less "mileage" on operating units; Less starting of gas turbines and other fast acting units,
- Reduce VRE curtailment

The case of Xcel

At the end of 2017, Xcel Energy Colorado's (a large vertically integrated utility) renewable portfolio was comprised 2560 MW of wind capacity with about 310 MW of solar capacity, which supply about 28% of total energy consumption in Colorado area[48]. In 2017, wind contributed 24% to Xcel Energy Colorado's annual load, with certain hours reaching nearly 70% of total load[49]. The increasing need for other resources to quickly ramp up and make up the supply has been identified as a major reliability concern[50]. Periods when wind production increases unexpectedly also create challenges.

In 2008, Xcel teamed up with National Center for Atmospheric Research (NCAR) to improve its wind forecast accuracy. The new forecast system used a variety of inputs from satellites, planes, radars, ground-based weather stations, and sensors on the wind turbines themselves to delivers high-resolution, wind energy forecasts for up to a 168-hour period, producing a new forecast every 15 minutes[51,52]. The 168-hour predictions allows for better unit commitment and scheduling decisions, but the 15 minutes by 15 minutes updated forecasts have made a substantial difference in determining when plants needed to be regulating, and when they needed to be making large redispatch for longer-term ramp-downs. The new forecast system improved the wind forecast accuracy by 39% relative to previous forecast methods. Mean Absolute Error (MAE) dropped from 18.01% in 2009 to 11.04% by 2013 in

45 https://www.nrel.gov/docs/legosti/old/7803.pdf

46 https://www.nrel.gov/docs/fy12osti/50907.pdf

47 http://greeningthegrid.org/trainings-1/presentation-an-introduction-to-wind-and-solar-power-forecasting-1

48 http://investors.xcelenergy.com/Interactive/NewLookAndFeel/4025308/Xcel_Energy_Inc-Hosting_Page_2018_ClientDL/ar/HTML1/xcel_energy-ar2017_0036.htm

49 https://www.xcelenergy.com/energy_portfolio/renewable_energy/wind/co_wind_power

50 http://info.aee.net/hubfs/EPA/AEEI-Renewables-Grid-Integration-Case-Studies.pdf?t=1440089933677

51 Ibid.

52 https://www.xcelenergy.com/Energy_Portfolio/Renewable_Energy/Wind/Wind_Operations

Colorado[53].

They discovered wind production takes anywhere from tens of minutes to several hours to ramp down instead of suddenly disappearing (as some variability forecasts assume in other reserve setting models). Therefore, Xcel filed with FERC a request to introduce a Flex Reserve Service product that has a longer time period, on the order of 10 minutes to an hour, to ramp up, instead of having to ramp up instantaneously to cover losses[54]. In 2011, Xcel carried a 30-minute flexibility reserve that provides 1 megawatt of replacement backup power for every megawatt of wind power to cover any unexpected wind output decreases. This new method maintains reliability while significantly reducing the costs associated with reserve[55].

And with the increased confidence in handling wind integration, Xcel plans to add nearly 3700 MW new wind capacity by the end of 2021[56].

Texas addressing frequency response challenges

Renewable integration challenge for operational frequency

As of 2017, ERCOT[57] had the highest installed capacity of wind in the U.S. at 20,682MW, contributing over 18% of Texas's energy. Many hours experience over 40% of energy coming from wind, including almost day-long periods at these levels. While these levels are seen in other regions, those regions rely heavily on exporting that power and other regions for balancing. Texas is essentially isolated from other grids and has to balance entirely on its own. Renewable capacity in Texas continues to grow, nearly 9 GW more wind is expected to come online in the next two years, meaning VRE installed capacity would be 40% of ERCOT's peak demand[58].

From these high penetrations, one of challenge ERCOT identified is frequency control during times of high penetration. As the share of variable renewables in Texas keeps increasing, the share of synchronized thermal generators with rotational inertia has declined. Research showed that for the same capacity loss, a system with higher renewable penetration sees the frequency decline much faster and results in a larger frequency delta. As a result, the grid's ability to respond to changes in frequency has significantly diminished. Accordingly, ERCOT has been actively testing new frequency response standards and reinventing its AS products to more efficiently meet these needs.

53 http://info.aee.net/hubfs/EPA/AEEI-Renewables-Grid-Integration-Case-Studies.pdf?t=1440089933677

54 https://www.xcelenergy.com/Energy_Portfolio/Renewable_Energy/Wind/Wind_Operations

55 http://info.aee.net/hubfs/EPA/AEEI-Renewables-Grid-Integration-Case-Studies.pdf?t=1440089933677

56 https://www.xcelenergy.com/energy_portfolio/renewable_energy/wind

57 ERCOT is the independent power market system operator for Texas

58 http://www.ercot.com/content/wcm/lists/89476/FAS_TwoPager_April2016_FINAL.pdf

Dissection of a Generator Loss Event

Rotational momentum describes the ability for traditional generators with rotating motors to absorb some of the frequency drop during a contingency event. During an event, frequency quickly drops off and the grid must quickly respond to stop further decline and buy enough time for spinning (secondary) reserves and then replacement (tertiary) reserves to come online and bring frequency back to standard 60 Hz. If the frequency drop is not slowed and stopped, blackouts can occur.

The process could be divided into different phases:

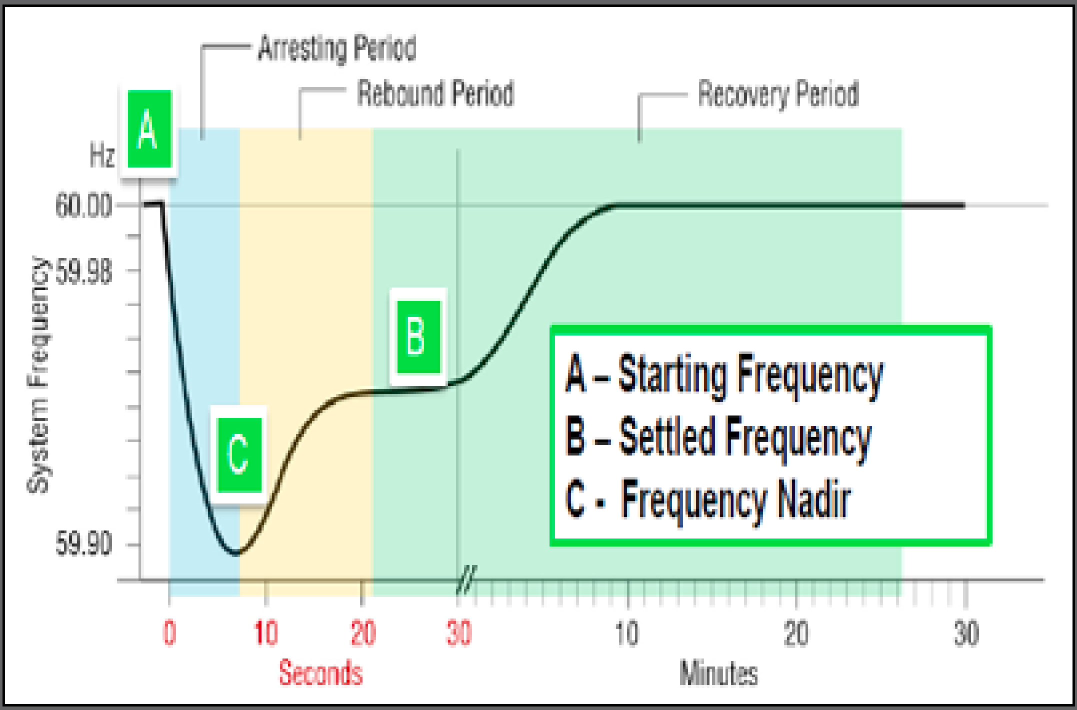


Figure 19: Typical frequency response following a generator trip

- Arresting period (blue in figure 19): Frequency declines as generator loss. System need to slow the decline in frequency, known as the rate of change in frequency (RoCoF). If not slowed in the first seconds after an event (blue in figure 19), will result in frequency declining too rapidly for any other resources to respond, and result in blackouts.
- Rebound period and recovery period (yellow and green in figure 19): Frequency is stabilized by primary frequency reserves, and then is returned to normal levels as secondary frequency reserves react to the event and come online.

Traditionally, rotating generators automatically and instantaneously resist frequency deviations[59,60] known as synchronous inertial response (SIR), which is a service generators automatically provide for free. BAAs procure responsive reserve service to cover the reserves for the latter two periods to help restore the frequency.

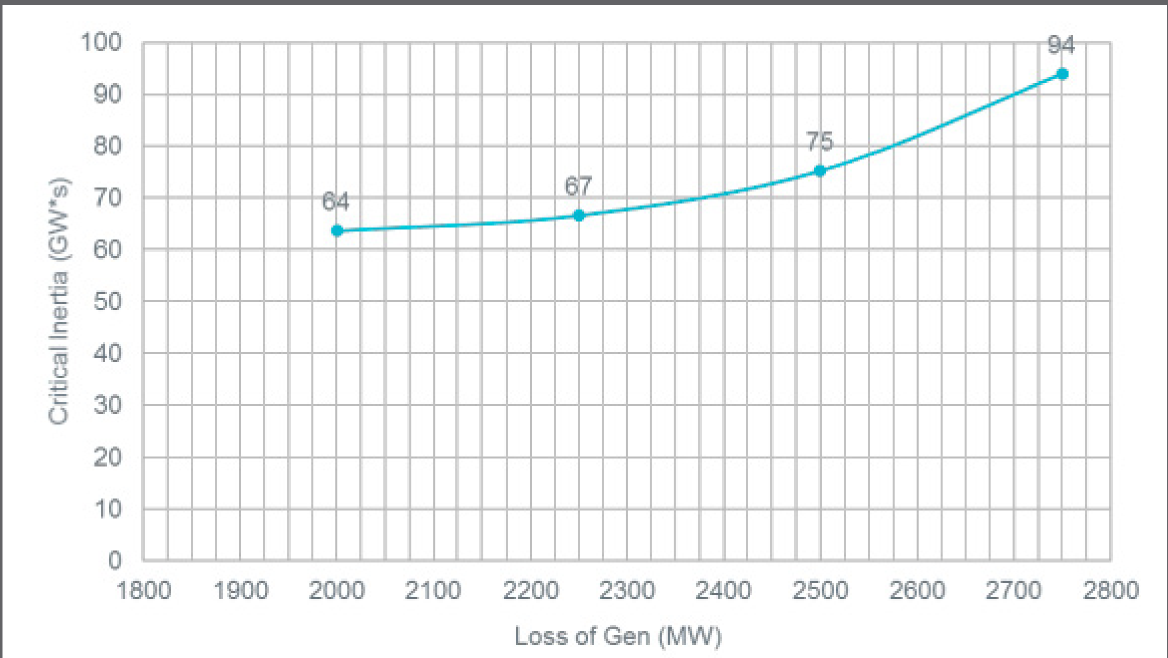


Figure 20: Critical Inertia at Resource Contingencies of different sizes from 2,000 to 2,750 MW

However, as less conventional generators are online, the free sources of SIR declines. Grid operators need to find an alternative, often described as synthetic inertia, which typically uses inverter-based or demand response technologies that can respond on a sub-second basis to help SIR reduce the RoCoF and help recover the frequency.

59 This is through electric governors, a device used to maintain the stability of the turbine/generator speed.

60 https://www.utilitydive.com/news/energy-storage-good-option-for-frequency-response-group-says/506329/

Approaches for acquiring more FFR

ERCOT took two parallel approaches to procure more synthetic inertia and lower its reserve carrying costs:

- Set up an AS market product to procure synthetic inertia via FFR
- Change the requirements of PFR to assure frequency performance

Market Approach

Creating FFR as a separate AS market product was part of a more holistic AS redesign effort in ERCOT. Fast Frequency Response (FFR) was defined recognizing the technical limitations of current synthetic inertia—inverters need roughly 30 cycles (0.5 seconds) to confirm a frequency drop and automatically deploy—and recognizing the speed required to meaningfully slow the RoCoF. Therefore FFR's product definition requires resources to respond within 0.5 seconds and sustain its response for 10 minutes, where Primary Frequency Response (PFR) could follow to restore the frequency, responding within 1-1.5 seconds and sustaining for 1 hour. Wind is capable of providing a 9 seconds of increased output, similar to the 10-12 seconds of SIR provided by thermal generators to arrest the RoCoF[61,62].

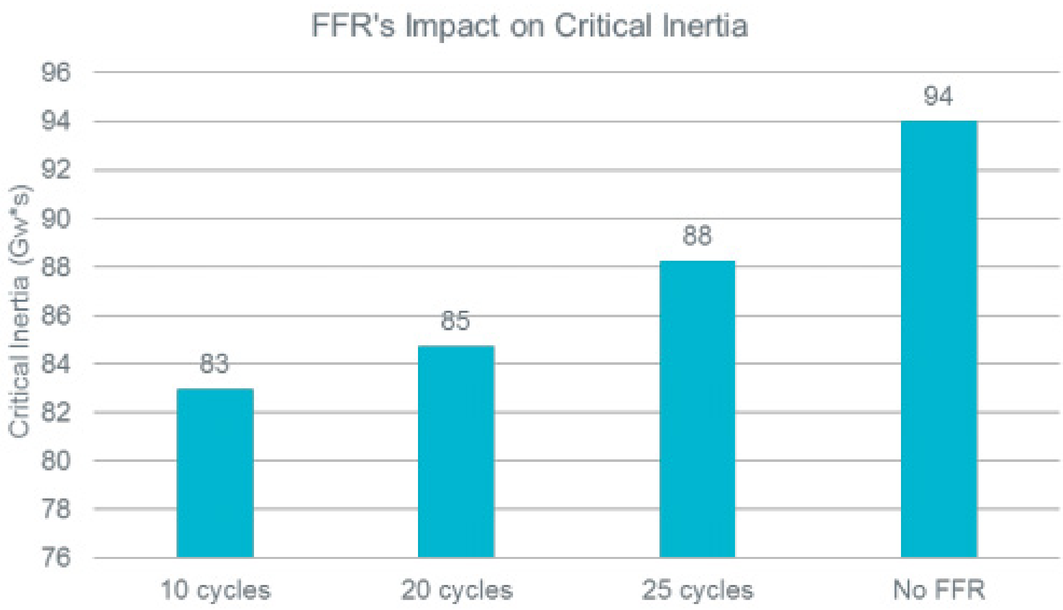


Figure 21: Critical Inertia w/ FFR (different response time).

There were some criticisms of the FFR definition, because the 10 minute response requirement limited wind energy's participation. Therefore the argument was FFR and PFR should be completely divided into two, one solely focused on RoCoF, the other on stabilizing the frequency until spinning reserves are ramped up. Therefore, ERCOT bifurcated its current responsive reserve service into FFR and PFR, with FFR focused on arresting the RoCoF and PFR buying time for secondary reserves to fully dispatch. With more FFR, the

61 https://www.aemc.gov.au/sites/default/files/content/661d5402-3ce5-4775-bb8a-9965f6d93a94/AECOM-Report-Feasibility-of-FFR-Obligations-of-New-Generators.pdf

62 Quebec found 9 seconds was usually enough time to bring other resources online that could respond within the 0.5 second limit but needed longer to ramp up to full output.

frequency decline slows down and the less rotational inertia will be needed (table 5)[63].

ERCOT only procures FFR when there is insufficient synchronous generators to arrest a RoCoF event. When it is determined FFR is necessary, the needed volume of PFR does not change, but it can be substituted for additional FFR, since FFR is defined to cover a long enough time frame to give time for secondary reserves to come online (spinning reserves are typically required to be fully available within 10 minutes)[64]. Since FFR's faster response raises the frequency nadir (the lowest the frequency drops during an event), FFR is more valuable to grid dispatchers than PFR, and therefore PFR requirements can be discounted if FFR is procured instead.

In scenarios modelled by ERCOT, the interrelation of PFR and FFR is explored under various scenarios of net load provision:

Net Load (GW)	PFR_{min} (MW)	FFR_{min} (MW)	PFR/FFR Substitution Ratio*
65	1400	0	0
35	1400	700	1.5
18	1400	1400	2.35

Table 5: PFR and FFR minimum requirements under different grid scenarios[65].

ERCOT's product was largely designed with demand response and utility-scale batteries in mind, and these resources contributed substantially to the pilot and demonstrated their ability to provide this service.

The resulting lower costs driven by reducing reserve requirements and minimizing curtailment of VRE were estimated in a cost-benefit analysis estimated savings at $20 million a year. This figure is based on 2016 and 2024 scenarios that included natural gas prices of $4.35 per million Btu (MMBtu) and about $11-16 million using a lower natural gas price of $2.36 per MMBtu[66].

Standards Approach

Despite the benefits, regulators questioned whether it was necessary. The project fueled a lot of discussions between stakeholders and regulators and contributed to the creation of a new NERC reliability standard which helped to answer some of these questions in the absence of FFR markets.

In 2014, FERC approved a new reliability standard developed by NERC and had BAs and generators in ERCOT pilot the requirements for improved frequency control service. The standard defines various responsibilities for power system stakeholders to better plan for,

63 http://www.ercot.com/content/wcm/lists/144927/Inertia_Basic_Concepts_Impacts_On_ERCOT_v0.pdf

64 FFR and PFR do not help bring the frequency back to normal frequency, which is the role of secondary reserves. This creates a confusing naming convention, where in the overall framework presented at the beginning of this chapter, both an event), FFR is more valuable to grid dispatchers than PFR, and therefore PFR requirements can be discounted if FFR is procured instead.

65 http://www.ercot.com/content/meetings/fast/keydocs/2014/0328/PFR_FFR%20Assessment_FASTworkshop_03282014.pdf

66 http://www.ercot.com/content/wcm/lists/89476/FAS_TwoPager_April2016_FINAL.pdf

record, and operate under frequency deviation events[67]. The decision was not least fueled by the discussions during ERCOT's development of a new FFR product, which revealed the need for a better understanding of how often frequency deviations were occurring and the synchronous inertia actually available during these events.

Upon implementation of the standard in 2015, ERCOT determined it already have enough frequency control capability and synthetic inertial response within the redefined PFR standard, and that there was no need to create a seperate AS product to procure it. Up to a half of ERCOT's PFR was able to be provided by load resources with under frequency relays, which provided the same functionality as FFR without having to pay more[68]. Stakeholders could solve this frequency control problem with little investment, and therefore did not need a market to bring more investment in. in addition, FFR did not need to be co-optimized with other markets; complying with this standard did not limit a resource's ability to participate in energy markets and the existing market PFR product. Despite this decision, other regions, including California, Ireland and Hawaii, are experimenting with different models to address this synthetic inertia challenge.

3.5 Relevance in a Chinese context

Unlike the energy market, which is universal and fundamental, ancillary service products are usually designed on an ad hoc basis and some of them are ephemeral. Ramping ancillary service product is the result of duck curve in California, and FRR in Texas originated from the worry of inertia shortage due to the increasing uptake of non-synchronous wind power. Ancillary service markets are important as they provide the necessary agilities to policy makers to handle transient problems. In some cases, when the problem is very acute, the introduction of new ancillary service products could provide necessary extra economic incentives for new investment. China also has many different ancillary service products, which will be illustrated in Chapter IV.2. Many of these products are designed to target very specific problems and they are usually characterized by high economic incentive in the early stage and thus attract quick investment in new areas. Most notably, the down-regulation ancillary service market in Northeast China and frequency regulation ancillary service market in Shanxi province, both have been proved to be very effective in handling the pressing local problems.

67 https://www.nerc.com/pa/Stand/BAL001TRE1/BAL-001-TRE-1%20Redline.pdf

68 https://www.nerc.com/comm/Other/essntlrlbltysrvcstskfrcDL/ERS_Forward_Measures_124_Tech_Brief_03292018_Final.pdf

4. Demand Response's Evolving Role in California

4.1 Key messages and takeaways

- California wants demand response to play a larger role in its system, and from the late 1990s on through today, California has progressively reformed the models by which DR is procured and participates. These changes have continuously pushed for more DR capacity and more diverse functionality from that DR. The evolution of programs and products also reflects the changing system needs (peak reduction to solar integration) and technological capabilities of DR over time.
- Utility programs were the main model of DR deployment in the beginning, focusing on manual curtailments of large commercial and industrial (C&I) customers during critical peak hours. This model presented great economics to the utility, grew substantial, and is the mature, dominant model for most DR in California.
- As demand response grew, its integration into wholesale market operations was prudent to utilize it as a reliability resource. California's Reliability Demand Response Resource (RDRR) product integrated the current emergency demand response within the utilities into the wholesale market. Alongside RDRR, California created a proxy demand response (PDR) product that would allow demand response to participate outside of the narrowly defined terms for the reliability product, which aimed to create opportunities for other types of demand response to be bid into the market. Ultimately, this product did not attract much attention, not enough money being available to bring new, non-critical peak products into the market.
- To resolve this issue, CAISO piloted a demand response auction mechanisms (DRAM), which asked DR providers to submit offers that utilities could accept and pay them for their contribution to meeting resource adequacy requirements. This opened opportunities for non-C&I models, which were a majority in DRAM. DRAM also required that DR to participate in markets under PDR and RDRR, and while most still focused on critical peak power provision, some winning DRAM bids provided non-critical peak demand response, typically using automated, residential programs that successfully engage customers on changing their demand behaviour.
- California wants to further expand the role of non-critical peak demand response. Evolved rate design, including demand charges and time-of-use (TOU) pricing, have increased load-flexibility at customer sites, and could be a more successful model than centralized market products for deploying comprehensive demand response, provided regulators are willing to expose customers to some price volatility.

4.2 Introduction

This case, although focused on a handful of wholesale market products for DR, ultimately tells the story of how DR products have been refined over the years, updating in response to system need, technical potential, and policy objectives. The corresponding demand response deployments and the associated business models evolved largely in response to these regulatory and market redesigns. Getting these program and product designs right required many pilots, engagement with DR service providers, and much tinkering to hone in on models that delivered the right types of response at least cost. Despite these shifts and changes in direction, California's overall objectives have remained unchanged, and can be seen as a guiding force throughout:

1. Grow the volume of DR by compensating sufficiently and fairly for the values it provides
2. Expand the role of DR by allowing it to be dispatched in more circumstances

DR's Evolution in California participation

Demand response has a long history in California. During the California energy crisis, electricity scarcity, real or fabricated, caused many hours in the market to become very expensive. In order to reasonably hedge against those high prices, California's Public Utility Commission (CPUC) required its utilities to run **utility DR programs**, where utilities could reduce demand during those high-priced hours. The economics of these programs are very favorable: the avoided market payments far exceeded the payments made to customers.

As these models matured, demand response grew to a sizable contributor of mitigating system peaks. But this presented operational challenges in emergency situations, where having the utility dispatching these resources presented a risk to system stability. Therefore, the California ISO (CAISO) requested that demand response contributing to resource adequacy be integrated with market operations. This new product did not create the market entrance of new DR as expected, in part because without a capacity payment, the economics for deploying these new DR resources outside of utility programs did not pen out. This, coupled with overly generous compensation within the utility programs, prompted exploring a centralized procurement model where DR could provide offers, and the utility could decide how much to procure. These resources would count toward the utilities' **resource adequacy requirement**, and would be required to bid into CAISO's wholesale dispatch markets.

And while CAISO's current set of products and procurement mechanisms support some of these aggregators, there is a notion that aggregated wholesale products are likely not the right way to deploy these types of DR at scale. More aggressive deployment of **TOU rates or retail models** that incentivize customers to respond to price signals, will better align customer action with real-time system needs.

These models are summarized at a high-level in the following figure (figure 22).

We will discuss each of these major evolutions in California's DR market, discussing some of the notable business models therein, and briefly discuss where California may be heading to further create a role for DR in providing flexibility and renewable integration.

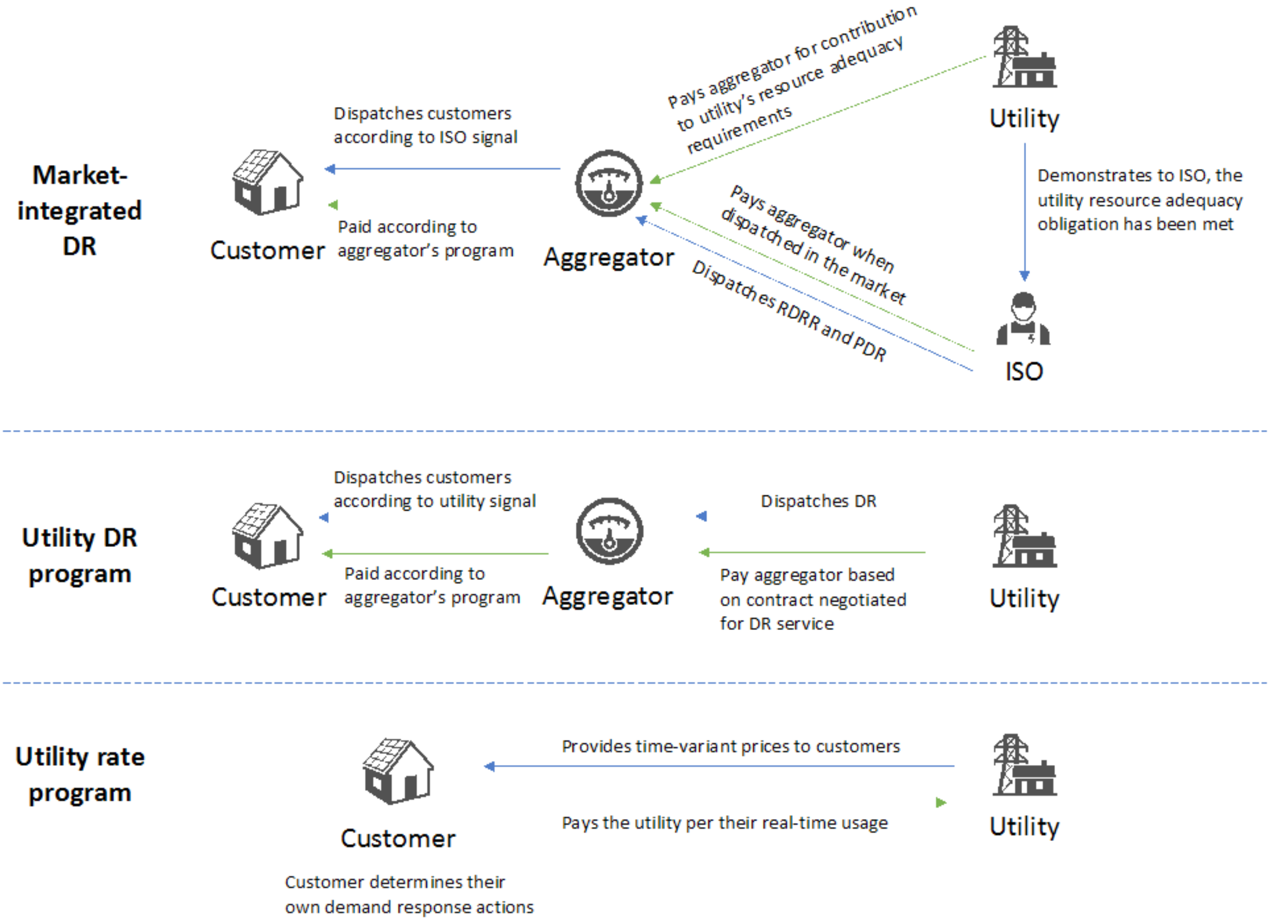


Figure 22: Overview of general DR frameworks discussed in this case

4.3 Services DR can provide

Before diving into the details of each phase in DR's evolution in California, it is helpful to orient around what major grid services DR has the capability to provide, and in particular, which services CAISO wants to capture through its DR product refinements:

1. Peak energy demand reduction
2. Reducing system ramp requirements
3. Firming intermittent renewable resources
4. Relieving network congestion costs
5. Reducing resource adequacy requirements

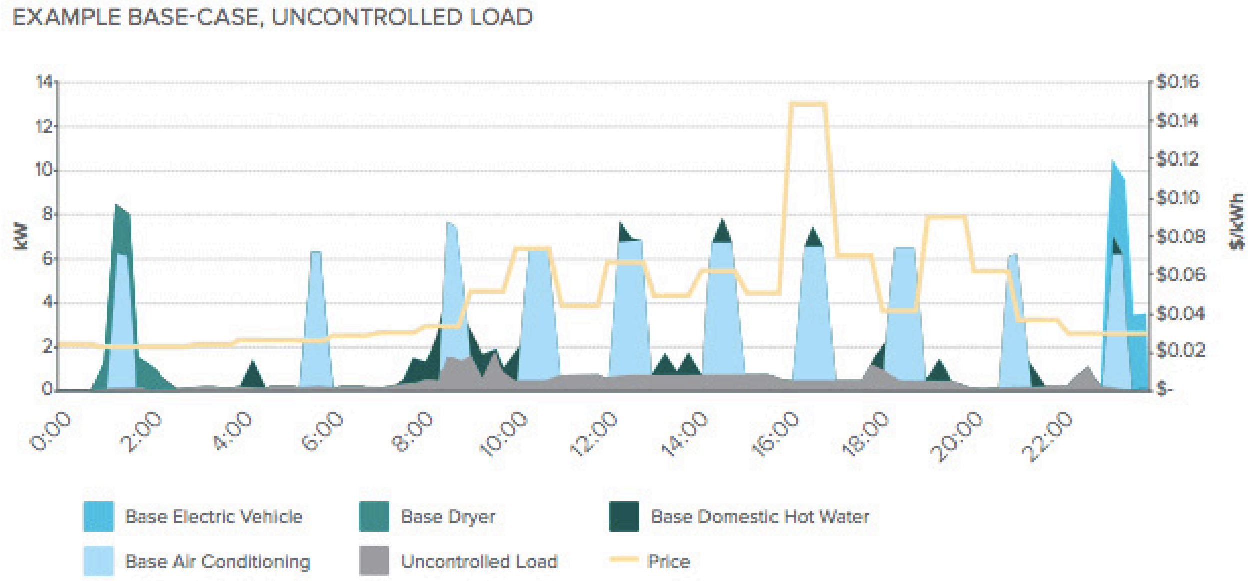
Below we give a brief description of what those services are, how DR can provide them and in what markets, and finally what types of DR can likely participate.

1. **Peak energy demand shaving** remains the "bread and butter" of demand response, just as moments of scarcity are the bulwark of generator's earnings in a year. This payment also aligns with where the most value is to the grid[69], but as that changes, so too will the market prices for other services, which may attract DR elsewhere. This service requires substantial demand reductions, and has historically involved many large industrial/commercial users, but most types of DR can provide this service. Given the scale of reductions, participant fatigue (when participants become less likely to respond after many events in quick succession) is substantial, and most DR programs require participants to only reduce their load a set number of hours every season.
2. **Flattening increasingly steep system ramps:** California's system requires large amounts of ramping capacity in the late afternoon, when demand reductions coincide with solar generation ramping down. Demand response could contribute to ramp-up, ramp-down services during ramp periods to lessen their severity. DR could also load-shift demand from peak to off peak, reducing the need for system ramps, but that service cannot currently be captured in California's existing markets. By doing so, DR would also help decommit generators scheduled on to meet ramps later in the day, saving unit commitment costs. Finally, during fast ramp periods the probability of tripping a generator is very high, and demand response, if dispatched instead of generator ramping, also reduces this risk. This kind of service would require automated or controllable loads (likely heating, ventilation, and air-conditioning (HVAC), hot-water heaters, and EV/stationary battery storage), and would need to have low latency in response time.
3. **Firming intermittent energy resources**: Variable renewable resources diverge from their forecasts and create regulation and balancing challenges. DR can serve as a firming resource in balancing markets to eliminate these divergences. This is currently being done in some microgrid demonstration projects with solar, DR, and batteries. Since this service would require DR to act on faster timescales, it would need to be automated loads, and the latency would need to be low. Hot water heaters, certain types of HVAC, and LED lighting aggregated at scale could help play these functions, as well as EV charging and behind the meter batteries.
4. **Relieving network congestion stress:** When transmission and distribution assets become congested, the cheapest sources of power cannot be utilized, and demand reduction is a more useful tool to manage congestion rather than altering power flows. This would look similar to peak load reduction, since that congested area would be paid at the locational marginal price , but revenues from financial transmission rights (FTR) could also be accrued by DR aggregators. Furthermore, DR could help avoid new transmission and distribution infrastructure investment, and part of that money could go to support localized DR deployment efforts[70]. Most types of DR could provide this service, although the requirements for reliable delivery must be higher when binding congestion is the price of under delivery.
5. **Reducing resource adequacy requirements:** Every utility is required to demonstrate they have procured sufficient energy and capacity contracts to reliably cover peak

69 Peak shaving potentials from DR is significant, according to analysis, if half of California's Investor-Owned Utilities (IOU) customers switched to their electricity company's existing time-of-use (TOU) rates, the reduced peak demand is enough to save the need for about 30 "peaker" 100 megawatt power plants. This same transition to TOU rates could save electric utilities and customers nearly $500 million annually, a nearly 20% system-wide cost reduction

70 https://www.edf.org/sites/default/files/demand-response-california.pdf

demand in the coming years. Dispatchable demand response plays a key role in achieving this resource adequacy. Therefore, the same way other energy resources and generators are paid to ensure their availability for future years (in California, typically long-term bilateral contracts), DR should get paid similarly. Most types of DR would be eligible provided they could be reliably available during the expected hours of system peak (between 4pm and 7pm for California).



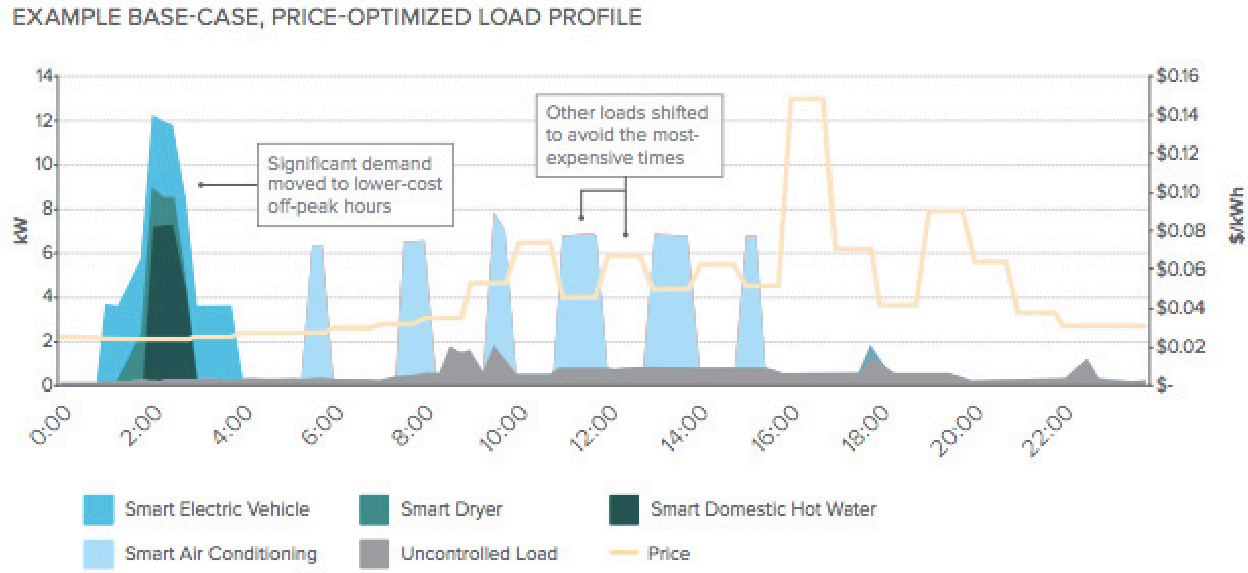


Figure 23: Load profile comparison under real-time pricing.

4.4 Utility DR Programs

Problems for traditional utility DR programs

Traditional utility-procured models existed long before wholesale DR products were implemented. They focus largely on C&I customers for critical peak reduction, where utilities dispatch the events (not the ISO) anytime they anticipate prices would be high. This reduces utility procurement costs significantly, thus utilities are willing to pay for these programs without adding these program costs to customer rates. Originally, utilities were intended to run and develop these programs, but designing and dispatching these programs required specialized skills, and aggregator models emerged to run those programs on the behalf of the utility. Aggregators are responsible for gathering together participants for the program, identifying how much each can contribute to demand reductions, creating terms under which those customers are willing to participate, and implementing the systems necessary to call events and get those sites to respond.

We explore the traditional model below. Since this model is already well-documented, we focus on explaining how this model for demand response creates a profile for demand resource that is very economic for energy cost reductions, but ultimately not suited for the other types of DR values California is looking to capture. It is important to emphasize though that these critiques are not of the utility DR model, which for good reason remains prominent, but rather an explanation of why California felt the need to open up other models for DR to participate.

Focused on utility cost reductions, not resource adequacy
Utility programs typically pay aggregators for delivering a specified number of DR events during a certain season (typically 80-100 per season) and pay both a program-based payment and event-based payment, the first paid per kW reductions available in the program, the second for the per kWh volume of response during each event. Since the marginal cost of calling events is lower than the cost of buying from markets above a certain price level, the utility will typically call every event they are allowed that season. This results in participant fatigue.

Dispatch timescales not aligned with system dispatch, creating inefficiencies
When dispatching utility program DR, the aggregator receives signals from the utility when they anticipate spot market prices will be high, and then relays a load reduction request to customers. But utilities have an incentive to not call an event until they are as certain as possible it will be an expensive hour on the grid, thereby not wasting one of their events. Therefore, aggregators developed their own predictive systems to give participants a 24-hour heads up when events may take place. While this 24-hour “heads up” is roughly the same timescale as the day-ahead market, these processes are not linked. Furthermore, utility dispatch is not linked with real-time market action and therefore there is a lack of coordination by having different actors dispatching different events.

Dispatch modes do not result in granular dispatchability
At peak times, utilities want as much DR as possible, since over delivery of load reduction is less harmful than under delivery. To minimize the risk of under delivery, contracts with the utility have ”price cliffs” where if the response is not above a certain MW threshold, the utility will not pay the aggregator for that event. To avoid this, aggregators typically call upon all demand in a zone during an event. Therefore, these programs and DR technologies are not focused on dispatch accuracy, something that would need to evolve if DR would serve a balancing role on the grid.

These minimum response requirements are also passed onto participants in their contracts, so when participants respond to events, it’s not as a matter of degrees, it’s a matter if they respond or not. Most levers, like AC, lighting, and other non-critical functions, are not ramped down in these programs, but entirely cut using pre-assigned settings, or in some cases automatic switches operated by aggregators that just shut loads off completely. While this means these programs typically enjoy very high response rates, 90-80%, this type of DR misses an opportunity for these measures to respond granularly and be of more use to dispatchers outside of critical peak contexts.

Moving Beyond the Utility Programs

While traditional DR programs remain very economically successful, California began seeing the opportunities for DR left on the table and began its DR programs. CPUC wanted to drive DR-based reliability to cover resource adequacy, and therefore mandated separate programs for reliability. This type of DR was to be procured but not dispatched unless an emergency was called, meaning it could not participate in economically-determined, utility-called events. This meant the savings from these programs did not cover their costs. Therefore, each utility was allowed to offer packages to facilities instead of DR offering in prices.

This resulted in a few challenges:

- **Not dispatched by the balancing authority.** CAISO is responsible for maintaining the reliability of the grid, and by not having access to dispatching utility DR meant they could not use this DR for reliability dispatch. The dispatch of reliability DR was turned over to CAISO, discusses in section 4.5.
- **Overcounting of DR's Reliability Contribution.** Utilities needing to meet their reserve requirement saw DR as an option cheaper. They started maxing out these programs and counting their traditional DR programs as a part of their reserve contribution. This ended up reducing the amount of firm dispatchable units available for frequency regulation (which this type of DR cannot provide due to the lack of granular dispatchability). The reliability DR programs were eventually capped at 2% of the reserve margin, and utility programs were considered load-modifying and could not count as reserves.
- **Overpayment.** Waiting until extreme emergency situations to dispatch DR resulted in higher costs. Deciding to dispatch earlier and as a part of economic dispatch could reduce costs. Additionally, facilities developed a price expectation from these reliability DR utility programs (generally in the $120-$130/kW-year range) and utilities with targets had little room to negotiate, resulting in little competitive price pressure[71]. Addressing this will be discussed in section 4.6.

4.5 CAISO DR Market Participation

Given these challenges and resulting decisions, CAISO created market products to integrate existing reliability DR. Two products were created, first, one that replicated the way utility critical peak programs defined DR's participation to give aggregators the option to be dispatched through the ISO instead of the utility, and then a second that allowed maximal flexibility to the DR aggregator to define their resource. These products are:

- Reliability Demand Response Resources (RDRR) is an emergency demand response dispatched directly by CAISO in cases when reliability is at risk
- Proxy Demand Response (PDR) where DR participates directly in energy markets and ancillary service markets, and is dispatched when economic

We will explore the mechanics of each of the products, and the rationale behind their design, and then understand how DR is participating in each of those products.

71 ftp://ftp2.cpuc.ca.gov/PG&E20150130ResponseToA1312012Ruling/2014/08/SB_GT&S_0350923.pdf

Design	Acronym	Services	Market dispatch	Description
Proxy Demand Resource	PDR	Energy, non-spin, residual unit commitment (RUC)	Economic day-ahead and real-time	Bids into ISO markets as supply
Reliability Demand Response Resource	RDRR	Energy	Economic day-ahead, reliability real-time	Bids into ISO markets; used for reliability purposes

Figure 24: Comparison of PDR and RDRR design[72].

RDRR's product definitions are compatible with existing emergency DR programs under utilities, with the ISO dispatching the resource instead of the utility. The program targeted the same types of DR signed up under utility programs, including Investor-Owned Utilities' interruptible load programs, direct-load control programs and agriculture and interruptible pumping program[73]. These resources are characterized by their large size, and their ability to respond to reliability event in real-time.

In the RDRR program, resources are not expected to respond (or be ready to respond) all of the time. The ISO must first call an emergency response situation, in which RDRR resources are added to the merit order line-up requiring them to bid in near the price cap. This way, RDRR resources only need to be ready to be called under these situations, and even then, their bids help distinguish which resources and how much the volume will actually get called. During 2017, RDRR was dispatched during 4 intervals during the year. The less frequent and more accurate dispatching helps keep payments low during these critical priced hours.

If RDRR resources (like water pumping) can easily schedule their operations, they can also elect to bid their spare resources into the real-time market as they want (just like any other resource), provided they are still able to deliver according to RDRR expectations. This is a helpful workaround for RDRR, which otherwise is prescriptive enough that it otherwise could limit the participation of some of California's largest DR providers from experimenting with new models for DR integration. To participate in RDRR, the resource must be aggregate at least 0.5 MW.

PDR is intentionally designed to provide as few restrictions as possible. New technology such as smart meters, advanced energy management systems, and customer-facing apps have enabled more demand types to participate in DR and more modes in which that demand participates. Demand response has become more automated and less intrusive, increasing customer acceptance while allowing grid operators to more reliably dispatch it. These evolutions have opened up the small commercial and residential markets and the associated aggregator business models. And with more sophisticated aggregation, comes

72 http://www.caiso.com/Documents/ReliabilityDemandResponseResourceParticipationOverview.pdf

73 Pumping loads are from the California Aqueduct in which large amounts of water are pumped over mountains and around the state to meet water requirements.

more sophisticated dispatch, where aggregated, controllable loads can precisely follow dispatch signals. These evolutions present a case for DR to be economically deployed at much higher volumes and dispatched outside of peak demand hours, aligned with CAISO's goals for these reforms.

These technology changes informed the design of PDR. The theory is that aggregators can aggregate new types of DR participants or participating loads, assess what market products those resources could meet, and bid in accordingly. PDR can participate in day-ahead, real-time, and non-spinning reserves. If a resource cannot respond reliably during certain hours, then they do not need to submit bids during those hours[74]. Or if the resource can only respond if given 24-hour notice, then the resource will only bid into day-ahead markets[75]. In that spirit, PDR is what its name connotes: a proxy generator, and accordingly can bid economically into day-ahead energy market, day-ahead and real-time non-spinning reserve market, and 5- minute real-time energy market. PDR requires a low-threshold for capacity to participate (aggregated to 0.1 MW to participant in energy bid and 0.5 MW to participate in non-Spinning Reserve bid). In each market, PDR participates as an economic resource and is dispatched in merit order, and is paid at the market clearing price when utilized.

Arguments for compensating DR in wholesale market

Followed by CAISO's efforts to integrate DR into wholesale market as a demand side resource, in 2011 the Federal Energy Regulatory Commission (FERC) issued a landmark ruling, FERC Order 745, standardizing the compensation of demand response (DR) in competitive wholesale markets. According to this order, demand response resources participating in competitive wholesale energy markets must, like generators, be paid full locational marginal price (LMP). Many economists opposed this ruling and argued that the most efficient method is to offer dynamic prices and naturally, demand reductions are rewarded with the avoided cost of the energy not used[76]. But demand response companies like EnerNOC and Viridity Energy have praised the ruling as a fair way to compensate customers for the costs of implementing real-time power-down technology in their buildings and factories[77]. And environmentalist applauds the ruling as they envision a bigger role for DR in embrace clean energy future[78].

74 http://www.caiso.com/Documents/PDR_RDRRParticipationOverviewPresentation.pdf

75 http://www.caiso.com/Documents/RevisedDraftFinalProposalVersion2-ReliabilityDemandResponseProduct.pdf

76 https://www.cei.washington.edu/wordpress/wp-content/uploads/2014/07/DR-_Negash_Kirschen.pdf

77 https://www.greentechmedia.com/articles/read/equal-pay-for-demand-response-goes-to-court#gs.QBt3Gjo

78 https://www.powermag.com/supreme-court-revives-ferc-order-745-on-demand-response/?pagenum=2

4.6 Demand Response Auction Mechanism (DRAM)

After integrating dispatchable DR into wholesale markets, two challenges for deploying DR remained unresolved:

- Market-participating DR was still largely procured by utilities at very high prices
- There had not been additional market entrances of new DR, since market revenues were insufficient on their own to create new entrants.

California looked to PJM, another major ISO in the U.S., where DR had flourished in the past few years, reaching 6.5% of total system peak, totaling over 11000 MW of DR in 2012[79]. A notable difference had been that DR participated directly in capacity markets. This additional revenue was seen as necessary to bring high volumes of DR into the market, especially residential DR.

CAISO created the DRAM pilot in 2015 to provide additional revenues to push additional DR to join the market. CAISO aimed to[80]:

- Compensate DR for its contribution to resource adequacy to help stimulate enough revenue for market entry
- But while these resource adequacy payments were necessary, the current levels ($120/kW) were believed to be too high, and they wanted a different mechanism to bring these resources online

DRAM is a pay-as-bid program that allows distributed energy resource (DER) aggregators to offer their service to utilities to meet the utility's resource adequacy requirement and receive a "capacity payment" from utilities, who in are California the ones responsible for procuring adequate capacity for system peaks[81]. Once qualified for DRAM, DER aggregators are obligated to participate in CAISO as a supply resource (via PDR) or reliability energy (via RDRR) and receive "energy payment" and "auxiliary payment" from CAISO. This opened up non-utility DR programs, since it provided an upfront payment for non-utility DR aggregators, providing sufficient money to deploy the basic infrastructure for their business. Furthermore, utilities who acquire the capacity have no claim to the revenues the winning bidders may receive from the energy market, making sure the energy payments can be used to support sustainable DR models.

In order to qualify for DRAM, all resources must meet an availability requirement and a ramp-up requirement. All qualified DRAM resources must be able to reduce use or add energy for up to 4 hours at a time for up to three consecutive days during the state's late afternoon and evening peaks, over the course of California's peak month (August) to qualify[82].

The resource must also be available within 10 minutes of dispatch request. Any winners are required to bid the winning capacity directly into the CAISO day-ahead energy market

79 https://www.pjm.com/-/media/markets-ops/dsr/2017-demand-response-activity-report.ashx

80 https://cpowerenergymanagement.com/maximize-demand-response-earnings-californias-dram-program/

81 CAISO's utilities are responsible to procure resources to meet their resource adequacy needs, usually by building new generation resources or signing forward energy contracts.

82 https://www.utilitydive.com/news/what-to-expect-from-california-utilities-new-aggregated-demand-response-of/412614/

during the hours specified by DRAM. And even in no event of dispatch, they must respond to several test events during the course of the year to maintain their award.
CAISO required utilities to procure a minimum amount (22 MW in total) of DR via DRAM to get the pilot running. These utilities procured well above these minimums, with DR presenting a less expensive option. Although DRAM capacity levels are required of utilities, CAISO leaves plenty autonomy for utilities to define their own DRAM contractual obligations, such as what other hours those resources need to be available.

	2016	*2017*	*2018*
Resource adequacy (capacity) products	System	System, local and Flexible Capacity	System, local and Flexible Capacity
Type of demand response	PDR—Day-Ahead Market ≥100kW	PDR—Day-Ahead and Real-Time ≥100kW RDRR—Real-Time ≥500kW	PDR—Day-Ahead and Real-Time ≥100kW RDRR—Real-Time ≥500kW
Operating months	June–December 2016	January–December 2017	January–December 2018 and 2019
Budget	$4 million each for PG&E and SCE; $1 million for SDG&E	$6 million each for PG&E and SCE; $1.5 million for SDG&E	$12 million each for PG&E and SCE; $3 million for SDG&E
Capacity	40 MW	82 MW	200 MW

Table 6: DRAM evolution from 2016 to 2018[83].

California's model evolved from PJM's DR Market design

While PJM's capacity market was successful in bringing more DR online, it has been criticized for overcompensation its DR since it pays all qualifying resources a single clearing price. CAISO DRAM, on the other hand, is pay-as-bid, and set utilities as the responsible parties for procurement. This design should theoretically function as PJM's capacity market by increasing DR revenue and maintaining competitive pressure, while avoiding overcompensation.

83 https://www.greentechmedia.com/articles/read/californias-dram-tops-200mw-as-utilities-pick-winners-for-distributed-energ#gs.8YNcY=w

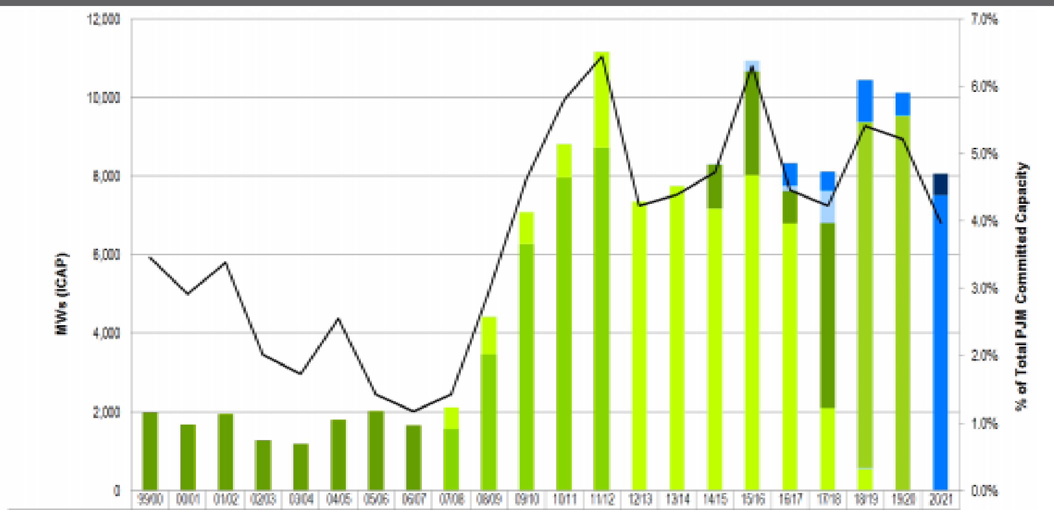
PJM's DR Market

Introduction

PJM started its DR pilot program in 2000, and since has achieved the most progressive and successful development of DR in the US, serving as an important model for other markets. PJM pioneered DR aggregators directly participate in markets, bypassing utilities and retailers, which helped to enable DR utilization in more situations other than critical peak demand response (e.g. load shredding). Also, DR resources were treated as generating assets and were able to participate in PJM's capacity market, receiving compensation for their contribution to planning reserves. Access to capacity markets has been the primary driver DRs substantial participation in PJM's markets.

DR Market Overview

Many of the fundamentals in PJM are similar to those rules aligned on in California. All DR resources are able to participate in energy, capacity and ancillary service markets. DR is dispatched in merit order based on submitted bids (for economic DR), or manually by system operators as needed (for emergency DR). Most DR resources are provided through aggregators that are registered in PJM as Curtailment Service Providers . Curtailment Service Providers receive payments, and then compensate end-customers based on their private agreement.



	99/00	00/01	01/02	02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10	10/11	11/12	12/13	13/14	14/15	15/16	16/17	17/18	18/19	19/20	20/21
PRD																		0	0	0	0	558
Annual - CP																		574	493	1075	563	7508
Base - CP																		0	0	8774	9563	0
Annual - DR																22	266	144	822	44	0	0
Extended Summer - DR																1,126	2,635	836	4894	0	0	0
Limited - DR									555	940	815	869	2,425	7,361	7,753	7,166	8,027	6783	2111	555	0	0
Limited - ILR									1,585	3,489	6,274	7,961	8,731					0	0	0	0	0
Limited - ALM	2,005	1,699	1,962	1,292	1,207	1,806	2,042	1,677										0	0	0	0	0
DR as % of Total PJM Committed Capacity	3.5%	2.9%	3.4%	2.0%	1.7%	2.5%	1.4%	1.2%	1.4%	2.9%	4.6%	5.8%	6.4%	4.2%	4.4%	4.7%	6.3%	4.5%	4.2%	5.4%	5.2%	4.0%

Figure 25: DR demand response capacity awards 1999-2021[84].

84 https://www.pjm.com/-/media/markets-ops/dsr/2017-demand-response-activity-report.ashx

In energy markets, DR participants have day-ahead and real-time options and receive payments at respective prices, and DR does contribute to locational marginal price setting[85,86].

In capacity markets, DR resources bid in their expected reduction potential for delivery three-years from the auction date, with a per MW bid price, that reflects its revenue shortfall not expected to be covered in the delivery year. Capacity markets have been the biggest revenue source for DR since the implementation of capacity markets, accounting for 97.6% of total DR revenue, at $126.3 million in Q1 2018 (figure 25). All DERs, not just dispatchable DR, can participate in PJM's markets as an aggregated resource, and include behind-the-meter generators/batteries, energy efficiency measures installed in lighting, HVAC, etc. For ancillary service markets, DR resources submit bids to contribute to reserves. This is typically a very small revenue source for DR, but increasingly relevant as more types of technologies are joining into the DR market, particularly those with power electronics, which could enable demand to play in some of these shorter-timescale services.

Major Points of Difference with CAISO's Design

CAISO's market differs in a few ways. CAISO still has a mixed model, combining vertically integrated utilities with a market model. So while PJM's goal was to fully integrate DR into its markets, CAISO has decided to leave some of the DR procurement functions to utilities, and the rest to markets. PJM took the all-market approach to fully leverage the 3rd party Curtailment Service Provider model. PJM believed this model has proven to produce more innovation and realize much higher levels of DR potential than utility programs, since utilities ultimately still have a profit motive to not reduce demand.

In PJM, most DR revenues are from the capacity market and many have argued that although it is a very good way to attract DR participation, a single clearing price overcompensate participants. FERC has also expressed concern about PJM's capacity market not resulting in just and reasonable prices . Whereas in CAISO, they choose to use long-term contracts to meet planning reserves, and accepted resources in DRAM are paid-as-bid. The auction mechanisms (DRAM) is CA's answer to needing to provide some revenue to DR providers to make the business model work, but not overpay them, and truly discover at what costs they can deliver this service.

CAISO has also placed more emphasis on economic demand response, instead of other products that are manually dispatched by market operators. In PJM, most of the procurement of DR early on was oriented around emergency demand response, thereby only being dispatched during rare, high-priced hours. California's requirements for bidding may produce more granular dispatch of DR in non-extraordinary circumstances (daily ramps, etc.). Also, California's emergency DR dispatch is still market-based, where RDRR is added to the merit order during emergency conditions. In PJM, emergency dispatch is manual and pays resources at an administratively set rate, which makes it difficult to distinguish which resources to dispatch and could lead to higher prices in the long-run by not setting those prices appropriately.

85 https://www.pjm.com/-/media/markets-ops/dsr/2017-demand-response-activity-report.ashx

86 http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2018/2018q1-som-pjm.pdf

Resulting Changes in DR composition

These changes in DRAM had a profound impact on volumes of market DR procured, especially PDR. The DR capacity procured via DRAM keeps increasing, from 40 MW in DRAM I, to 82 MW in DRAM II and 200 MW in DRAM III. This indicates a growing number of DR aggregators able to offer lower price points than other generators coming online, and increased acceptability of these resources with utilities. One concern is that as 6,200 MW of CAISO's flexible capacity is set to retire in the coming years, peak demand response cannot fill potential energy gaps left outside of DRAM hours. There are preliminary conversations on how to modify resource adequacy requirements away from peak demand coverage, to a wider, coverage-based metric that ensures adequate dispatchable capacity to cover net-load across all hours of the year[87].

Many of those new DRAM procurements selected to participate in the PDR market, with the total amount of PDR capacity registered in 2017 increasing to about 270 MW from about 160 MW during 2016. While more DR was procured, not much of that DR was dispatched. This is in part because many DRAM qualifying resources had the strategy of qualifying for DRAM and then bidding high to minimize their dispatch. This ultimately is an acceptable outcome, since DRAM is only focused on resource adequacy, and these resources are able to meet the dispatch/test requirements for DRAM, but no more. But given PDR's flexibility in defining its participation, there are some aggregators that have found providing short dispatches of energy in California's 5-minute markets profitable.

Almost all PDR capacity was bid at the cap of $1,000/MWh, and therefore only a fraction of PDR was dispatched in real-time markets. In 2017, there were 525 intervals in the 5-minute market calls for resources to deal with momentary capacity shortages. In 46 percent of these intervals, proxy demand response resources were dispatched on $1,000/MWh energy bids, an increase from 13 percent of intervals in 2016, showing PDR's role in real-time dispatch increasing. During each of these intervals, the quantity dispatched from each resource was very small, averaging about 0.5 MW, similar to 2016's 0.3 MW. This was likely a contributing factor to their dispatch, since other generators may have requirements on the magnitude and sustained direction of their ramp, or may already have been obligated to provide resource adequacy.

PDR resource aggregators seem to be becoming increasingly comfortable with this type of dispatch, because the frequency and volume in which PDR was bid in also increased significantly in 2017. About 17 MWh of PDR bid in on an average of 9 hours during weekday periods and 11 hours during weekend periods, compared to 2016's 10 MWh bid in during 4 hours of weekday periods. This growth began in 2016 and continued in 2017, resulting in 6 times higher PDR capacity bidding economically in the real-time market (figure 26)[88]. These increases came most from PDR's real-time bidding, which does align with the desire to increase the flexible dispatch of DR over time.

87 Interview with John Goodin and Jill Powers. Conducted August 20th.

88 http://www.caiso.com/Documents/2017AnnualReportonMarketIssuesandPerformance.pdf

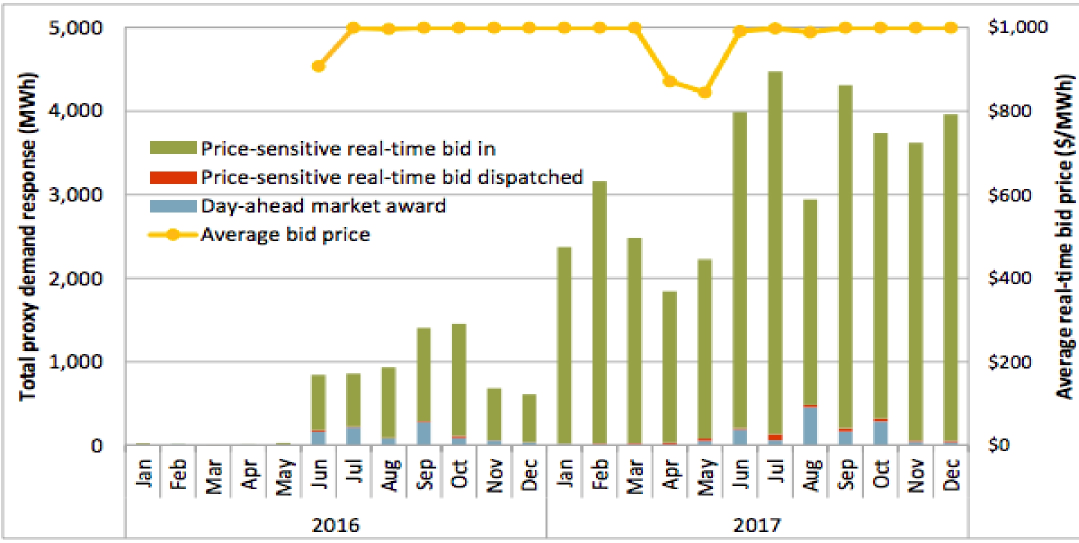


Figure 26: Proxy demand response awards and bids[89].

This shift in demand response capability is partially attributed to DRAM, which allowed a good deal of flexibility to the types of programs that could qualify. Many of the programs, like that of OhmConnect (discussed in the following section), needed the DRAM payment to make their models successful, but the solution they deployed allowed for a lot more flexibility in the types of dispatch it could perform.

DRAM has also resulted in shifts in the overall volumes and types of DR procured. The number of utility operated DR programs has generally declined with the increasing participation in DRAM/PDR/RDRR (table §7). But the total DR capacity from utility supply-side DR is still far greater than the current DRAM pilot. Reliability-based programs[90] accounted for 64 percent of capacity from utility-managed demand response resources in 2017, basically remains at the same level in recent years[91]. But price-responsive utility-managed programs[92] have decline substantially, down from more than 50 percent three years ago, and accounted for about 36 percent of total demand response capacity in 2017.

Given all of these changes, it is hard to assess the implications of each market revision. For instance, the total DR capacity available in 2017 decreased slightly from the previous year, with total measured capacity of about 1,023 MW compare to 2016's 1320 MW. It is hard to determine if this is caused by momentary fluctuations as various actors, utilities and independent providers, reshuffle their business models between these options, or if there has been a substantial market exit. It may also reflect the economically rational exit of DR resources that were being paid too much under previous reliability-based utility programs.

89 Ibid.

90 Reliability-based programs consist primarily of large retail customers under interruptible tariffs and air conditioning cycling programs

91 http://www.caiso.com/Documents/2017AnnualReportonMarketIssuesandPerformance.pdf

92 Price-responsive programs includes Day-ahead price-responsive programs which are triggered on a day-ahead basis in response to market or system conditions that indicate relatively high market prices and Day-of price-responsive programs which can be dispatched during the same operating day for which the load reduction is needed.

Utility/type	2013 Estimated MW	2014 Estimated MW	2015 Estimated MW	2016 Estimated MW	2017 Estimated MW
Price-responsive					
SCE	706	790	690	595	456
PG&E	404	418	285	193	125
SDG&E	61	61	83	54	46
Sub-total	1,171	1,269	1,058	842	626
Reliability-based					
SCE	684	733	767	770	731
PG&E	332	313	334	383	401
SDG&E	1	0	1	1	1
Sub-total	1,017	1,046	1,102	1,155	1,132
Total	2,187	2,315	2,160	1,997	1,759
Resource adequacy allocation	2,582	2,299	2,047	1,831	1,778
With 15 percent adder	2,970	2,644	2,354	2,105	2,045

Table 7: Utility operated demand response programs

FERC ordered markets to create a standard model for energy storage market participation

Energy storage, increasingly deployed as back-up power at customer sites, has begun to participate as demand response in wholesale markets. Paying for storage as demand response creates challenges in establishing a baseline, since it could charge during baselining periods to over inflate their demand. Various markets undertook rule revisions to address storage, but FERC wanted to create a standard across the entire country, thereby avoiding the challenges experienced during DR's development, having to have different business models in each region. On February 15, 2018, the FERC issued Order No. 841 requiring ISO/RTOs to revise market rules to remove barriers for storage to participate in wholesale markets. FERC leaves each individual market to determine their specific rules, but requires those rules must:

- Enable electric storage resources are eligible to provide all capacity, energy, and ancillary services that the resource is technically capable of providing in the RTO/ISO markets.
- Specify that the sale of electricity from RTO/ISO markets to a storage resource must be at the wholesale locational marginal price (if the storage asset resells that energy to markets)
- Require storage dispatch and procurement to be bid based, requiring markets to set up the appropriate auction mechanisms for each market, where storage can set market clearing prices as both a buyer and a seller.
- Establish a minimum size for participation in RTO/ISO markets that does not exceed 100 kW (commonly at 1MW before).

These guidelines helped accelerate battery adoption.
Enabling **participation in all markets** ensures access to revenue streams essential to grow battery business models, particularly ancillary services and capacity/resource adequacy mechanisms.

- Although AS is not be a huge source of revenues now, a little extra revenue can make a big difference in making deployment economics pencil out. A study found wholesale market revenues alone would result in only 1,000 MW (3,000 MWh) of storage in ERCOT, but if distribution system benefits are included, especially frequency regulation, the market could support as much as 5,000 MW (15,000 MWh) of energy storage.
- Participation in capacity markets or forward energy contracts, where applicable, helps integrate storage and DERs into centralized planning processes, particularly as planning begins to explicitly consider flexible capacity sufficiency.

Specifying batteries could **buy energy at wholesale rates** ensured distributed batteries could perform energy arbitrage at the wholesale level, a feature essential in the future to integrate higher levels of renewables.

By allowing **batteries to set market pricing**, short timescale services that batteries and DR are more adept at providing, like the 5-minute real-time energy, can rise to prices that reflect scarcity, thereby encouraging more market entrance.

Elimination of **minimum size requirements** has greatly lowered capital hurdles for developers. Getting to 1MW starved many businesses models out before those batteries could access market revenues. Previously, batteries had to rely on utility programs and behind-the-meter arbitrage models for their earnings.

As individual states reviewed their rules, it became obvious that many of the existing rules were not explicitly designed for storage or DERs, but rather built for traditional generators and had the inadvertent effect of barring customer-side resources. Therefore, a systematic review of all interconnection and dispatch standards was often necessary to identify any and all rules that may have this effect. Many credit FERC's timely intervention as essential to help standardize battery participation in a way that allowed for tremendous growth under the wholesale market system.

4.7 New Residential and Market-based models

DRAM has brought numerous new resources into the market that previously had high barriers to entry. It was hard for newcomers and innovative business models to enter when almost all business went through utility programs. But with the broader definitions of PDR and DRAM allowing for higher levels of procurement, new aggregators entered, focused predominately on:

- Residential energy consumption: particularly enabled by smart thermostat technology which could control users' HVAC, water heating, and other loads with little customer intervention
- Behind-the-meter batteries: Residential or C&I energy storage, installed for reliability purposed, could be aggregated to participate in markets (discussed at greater lengths in

battery storage call-out box)

- EV charging: including smart charging technology and EV charging as a demand response resource (this is discussed at greater lengths in section 5)

Numerous companies were contracted under DRAM, including smart EV charging providers like eMotorWorks and residential energy management service companies like EnergyHub, which were both major winners for DRAM[93]. DRAM also saw companies entering the market whose core business was not energy management, like the EV & battery company Tesla and residential rooftop solar installation companies[94]. Although revenue numbers for DRAM cannot be disclosed[95], the winning providers list implies a strong business case for these new residential models.

OhmConnect

OhmConnect, a San Francisco-based residential DR start-up aggregator, won contracts with all three utilities for more than 30% of the total DRAM capacity for both years of DRAM III[96]. Founded in 2013, OhmConnect currently has 300,000 users who provide 100 MW DR service, equivalent to 4 peaking plants and has delivered 860,000 kg of avoided CO_2 emissions[97].

Like other DR aggregators, OhmConnect closely monitors wholesale market prices to determine when they may dispatch. They send text messages to customers calling for load reduction during expected high price period, called OhmHours. Once customers agree to reduce their load for that time period, the company will bid these aggregated reductions into the market as energy products and receive wholesale payments for each MWh delivered. This revenue is then divided between OhmConnect and responding customers, where the company takes a cut around 20%[98]. The average OhmConnect customer earns between $100-300 per year, with some more active participants earning in the thoU.S.nds[99]. To date, OhmConnect has paid 4.8 million dollars in total to their customers[100], implying 6 million revenues from CAISO wholesale market.

OhmConnect accesses its customers power consumption data via utility smart meters, which is critical for them to calculate the actual load reduction, which needs to be based on a baseline determined from actual consumption and forecasted load[101]. This meter data also helps OhmConnect better understand their user's U.S.ge patterns, so they can send out load reduction request in a more targeted fashion.

---

93 https://www.utilitydive.com/news/what-to-expect-from-california-utilities-new-aggregated-demand-response-of/412614/

94 https://www.greentechmedia.com/articles/read/californias-dram-tops-200mw-as-utilities-pick-winners-for-distributed-energ#gs.5l_VpBY

95 The evaluation of the demand response auction mechanism (DRAM) pilot is undergoing and requires more time than originally expected. So most of DRAM data especially cost and dispatch related are strictly confidential

96 https://www.greentechmedia.com/articles/read/californias-dram-tops-200mw-as-utilities-pick-winners-for-distributed-energ#gs.5l_VpBY

97 https://www.ohmconnect.com/about-us

98 http://energypost.eu/three-new-energy-companies-finding-value-in-three-new-business-models/

99 https://www.ohmconnect.com/blog-post/pocket-445-per-year-by-connecting-your-ohmconnect-account-with-a-smart-thermostat

100 https://www.ohmconnect.com/

101 https://www.ohmconnect.com/wiki/getting-started

OhmConnect also serves as a platform for deploying more DR and efficiency, like smart thermostat installation. OhmConnect advertisements for smart thermostats increase adoption (and may also provide them advertising revenues) and once installed allows the company to remotely adjust the settings of customers' AC systems. More and more customers, currently around 20%, give thermostat control rights to OhmConnect and the small modulations during the day go largely unnoticed by customers[102]. It creates a win-win situation where OhmConnect is more confident in their response rate, and customer enjoys increased payments with almost no customer disruption. Beyond smart thermostats enabling demand response, they also represent a huge energy efficiency opportunity by optimizing heating and cooling set points (customers are also often willing to purchase smart thermostats for the automated comfort control benefits it provides). On average, smart thermostats could save 10% to 12% energy consumption on heating and 15% on cooling. Customers could receive a $75-100 rebate at the time of installation[103] and on average earn $131-145 in annual energy savings.

Future non-dispatchable DR – Price triggered load flexibility

While wholesale participation has created a lot of opportunities for new DR to enter, there is also a growing push from system operators, DR providers, and academics to improve the effectiveness of load-modifying DR by evolving Time-of-use (TOU) rates to real-time pricing[104]. With current smart appliance technology, it's easy to reshape a customer's demand profile continuously in ways that are invisible to the customer and minimally affect their behaviour. But this requires a more-granular rate structure to direct smart appliances to respond appropriately. Studies show that smart appliances can automatically shift customers' energy consumption to off-peak hours with dynamic, real-time, hourly pricing to a net economic benefit to both them and the grid (see figure 23)[105].

While California continues to revise the various mechanisms and definitions for procuring DR, they typically have not disbanded previous options for DR procurement, only recalibrating and coordinating to ensure the mechanisms work synergistically, and do not, in aggregate, result in higher costs than necessary. This approach affirms the central philosophy that different DR types require different procurement approaches, and that system operators should not view one approach as a complete solution, but rather one approach used to target a very specific value that DR can provide to cost-effective, flexible system operation.

4.8 Relevance in a Chinese context

The aspiration of using DR as a main tool to balance the power system has been there for a long while. Utility companies are usually main buyers for DR services. Utility companies procure DR not only because of their balancing responsibility, but also because the

102 http://energypost.eu/three-new-energy-companies-finding-value-in-three-new-business-models/

103 Rebates amount varies by utilities. https://www.ohmconnect.com/blog-post/pocket-445-per-year-by-connecting-your-ohmconnect-account-with-a-smart-thermostat

104 Time-of-use rate is a static rate plan better align the price of energy with the cost of energy at the time it is produced. The defining of time slots for different rates is based on the historical load data. An easy-to-confused concept is real-time pricing, which is dynamic price of energy that reflects the cost of energy at more granularity time scale such as hourly. To notice the real-time pricing is retail rate, so it may not be directly linked to the wholesale price in real-time market.

105 https://rmi.org/wp-content/uploads/2017/05/RMI_Document_Repository_Public-Reprts_RMI-TheEconomicsofDemandFlexibilityFullReport.pdf

economic benefits coming from reducing the demand at peak time in which the wholesale power price is usually very high. However, if the aim is to further scale up DR, the utility-as-the-main-buyer model becomes insufficient. The pioneering efforts in California is the shift from utility DR programs to a market paradigm, in which DR could participate in the wholesale market directly. But it is also observed that revenue from the wholesale market alone would not be enough for encouraging new investment on DR services. Thus, California introduced a new program DRAM to further compensate DR services. DRAM basically translates the value of DR as a firm capacity to a new revenue of DR aggregators and proved to be very effective on scaling up DR in California.

In China, the need for DR become more and more prominent. The fast urbanization process results in the fast growth of air-conditioning load. The simultaneous turn-on of air conditioners in early evenings of hot summer leads to a sharp peak demand. Currently, a peak-valley pricing system is used as the main tool to alleviate the peak demand. A few provinces are now piloting DR programs in which grid company functions are the buyers. Still these methods are far from sufficient, as the shortage of power in hot summer in some eastern provinces continues to grow. As China gradually introduces competitive wholesale market on the provincial level, including DR in the wholesale market becomes a new option. However, the energy value of DR, which is reflected in the wholesale market, might not be able to sustain DR business models. In order to attract more investment, extra market programs, such as DRAM in California, which reflects the capacity value of DR, might still be needed.

5. Managing EV Integration in California

5.1 Key messages and takeaways

- California has set aggressive targets for EV adoption. Growth in electricity demand from electric vehicles (EVs), if left uncoordinated, would significantly add to peak demand, inducing further intraday flexible ramping needs. Californian utilities are implementing pilot projects to 1) shape the way customers charge their EVs in order to avoid new EV charging load adding further to grid peaks and flexibility costs, and 2) understand how EVs can be used as a dispatchable demand-side resource.
- In SDG&E Power Your Drive Program, California puts the responsibility for coordinating EV charging behaviour on the utility. SDG&E owns and operates the charging stations, ensuring the location of chargers and the charging pattern will reduce system stress, not add to it. The program passes real-time prices to its customers, and the payment platform allows those customers to select which prices they are willing to charge at. The pilot has been successful in shaping EV charging load and supporting the growth of EVs in SDG&E's territory, but there is debate if the utility ownership model is necessary to achieve this success.
- PG&E-BMWi ChargeForward pilot aggregates EVs to participate in demand response markets as dispatchable demand resources. The pilot revealed smart charging deteriorates EVs' ability to serve as a dispatchable resource in the current system, since EVs already participating in smart charging programs do not have availability during critical peak events when dispatchers would need them. But the pilot confirms the possibility of using EVs as dispatchable grid resources from a customer acceptability and technical coordination perspective. This indicates the possibility of dispatch outside of critical peak reduction, especially in higher renewable systems.

5.2 Introduction - EVs' potential system impact and Utilities involvement

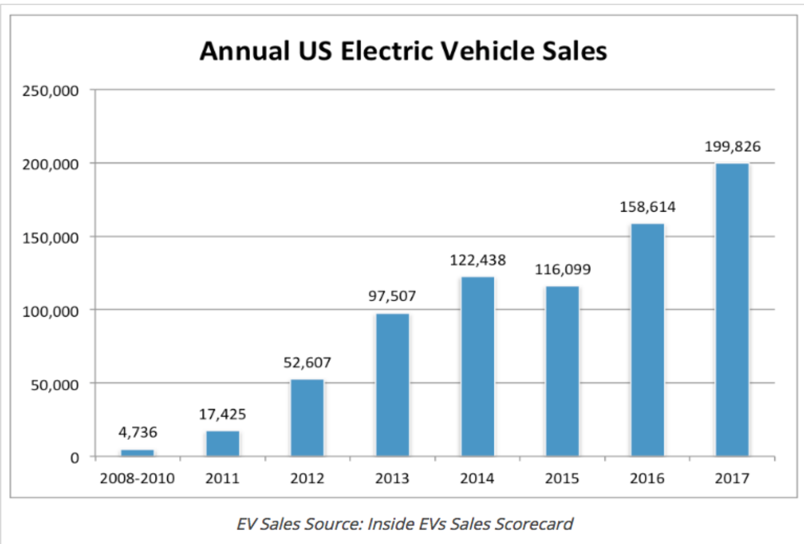


Figure 27: Annual EV sales in U.S.

EV demand growth

The EV market is continuously growing in the U.S. (figure 27), accounting for 764,516 vehicles in total by the end of 2017[106]. California is the leading state in terms of adoption, representing more than 50% of the total U.S. EV market, reaching 5.31% EV market share, well above the national average of 1.16%[107]. California's ambitious targets for decarbonizing transportation mandates 1.5 million zero-emission vehicles are on the road by 2025[108] and 5 million zero-emission vehicles by 2030[109], which will further drive EV adoption to unprecedented levels. Forecasts show that by the end of 2018, 1 in every 10 new cars sold in California will be EVs[110].

EV-related electricity demand represents a significant share of new electricity demand. Under moderate scenarios, there could be 2.9 million EVs on the road in the U.S. within five years, bringing over 11,000 GWh of load to the U.S. power grid, or about $1.5 billion in annual electricity sales[111]. A car with a 30 kWh battery stores as much electricity as the average U.S. residence daily consumption. This means introducing one more EV is roughly the same as adding one new household to the system[112]. If left to develop ad hoc, EVs will likely add to the utility cost base by increasing peak demand. California Energy Commission's research shows the load impact of EV charging could potentially increase peak demand by 1 GW by 2025[113]. This potentially has two impacts: 1) utilities will have to procure more generation capacity to meet this higher peak load, and 2) as California's peak gets higher, so too does the ramping requirement for generators, given the duck curve characterizing CAISO's net load.

Current utilities' EV rate schedules

The utility can (and arguably has the responsibility to) play a key role in shaping EV charging, to minimize the impacts to EV and non-EV customers alike[114,115]. As utilities have an interest in increased electricity sales, utility-based programs to encourage EV adoption have existed for a while[116], but have primarily focused on adoption, not load shaping. But as EV demand grows, load-shaping, typically through time-of-use (TOU) rates[117] could play a major role in shifting EV demand from peak hours to off-peak hours, thus increasing their electricity sales without substantially increasing their wholesale energy payments.

---

106 http://evadoption.com/ev-statistics-of-the-week-historical-us-ev-sales-growth-market-share/

107 http://evadoption.com/ev-market-share/

108 https://phys.org/news/2018-05-electric-vehicles-billions-energy-storage.html

109 https://www.utilitydive.com/news/cec-california-ev-chargers-will-add-1-gw-of-peak-demand-by-2025/519517/

110 http://evadoption.com/will-california-reach-10-ev-sales-market-share-by-december-2018/

111 https://www.rmi.org/wp-content/uploads/2017/10/RMI-From-Gas-To-Grid.pdf

112 https://rmi.org/wp-content/uploads/2017/04/RMI_Electric_Vehicles_as_DERs_Final_V2.pdf

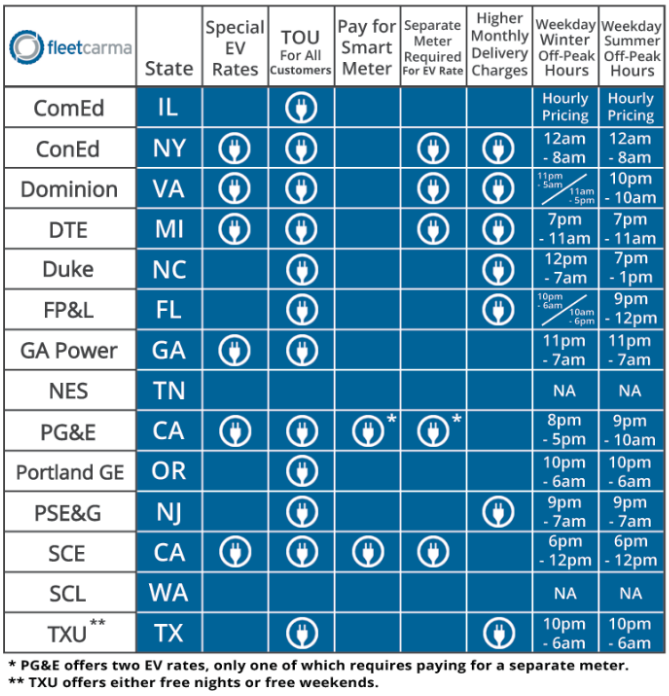
113 https://www.utilitydive.com/news/cec-california-ev-chargers-will-add-1-gw-of-peak-demand-by-2025/519517/

114 Utilities are direct rebates to those who purchase EVs, or special EV charging rates which offer cheaper, typically flat rates

115 https://www.swenergy.org/data/sites/1/media/documents/publications/documents/How_Leading_Utilities_Are_Embracing_EVs_Feb-2016.pdf

116 https://www.afdc.energy.gov/laws/state

117 Time-of-use rate is a static rate plan better align the price of energy with the cost of energy at the time it is produced. The defining of time slots for different rates is based on the historical load data. An easy-to-confused concept is real-time pricing, which is dynamic price of energy that reflects the cost of energy at more granularity time scale such as hourly. To notice the real-time pricing is retail rate, so it may not be directly linked to the wholesale price in real-time market.



fleetcarma	State	Special EV Rates	TOU For All Customers	Pay for Smart Meter	Separate Meter Required For EV Rate	Higher Monthly Delivery Charges	Weekday Winter Off-Peak Hours	Weekday Summer Off-Peak Hours
ComEd	IL		🔌				Hourly Pricing	Hourly Pricing
ConEd	NY	🔌	🔌		🔌	🔌	12am - 8am	12am - 8am
Dominion	VA	🔌	🔌		🔌	🔌	11pm - 5am / 11am - 5pm	10pm - 10am
DTE	MI	🔌	🔌		🔌	🔌	7pm - 11am	7pm - 11am
Duke	NC		🔌			🔌	12pm - 7am	7pm - 1pm
FP&L	FL		🔌			🔌	10pm - 6am / 10am - 6pm	9pm - 12pm
GA Power	GA	🔌	🔌				11pm - 7am	11pm - 7am
NES	TN						NA	NA
PG&E	CA	🔌	🔌	🔌*	🔌*		8pm - 5pm	9pm - 10am
Portland GE	OR		🔌				10pm - 6am	10pm - 6am
PSE&G	NJ		🔌			🔌	9pm - 7am	9pm - 7am
SCE	CA	🔌	🔌	🔌	🔌		6pm - 12pm	6pm - 12pm
SCL	WA						NA	NA
TXU**	TX		🔌			🔌	10pm - 6am	10pm - 6am

* PG&E offers two EV rates, only one of which requires paying for a separate meter.
** TXU offers either free nights or free weekends.

Figure 28: EV rates example in US

Based on a number of factors, power companies in different areas have taken different approaches to make charging rates for EVs. Some provide rate plans tailored specifically for EV owners, while others offer the same TOU rates as residential customers. Some require customer install a separate meter for EV charging to take advantage of special EV rates (figure 28). For example, TXU Energy gives power away for free between 10pm and 6am.

EV charging behaviour is very responsive to these rates, and all of California's utilities currently operate some form of TOU rate for their EV customers. California's TOU rates work substantially well, minimizing charging during peak hours and see a significant ramp-up in EV charging whenever their lowest rates go into effect (figure 29).

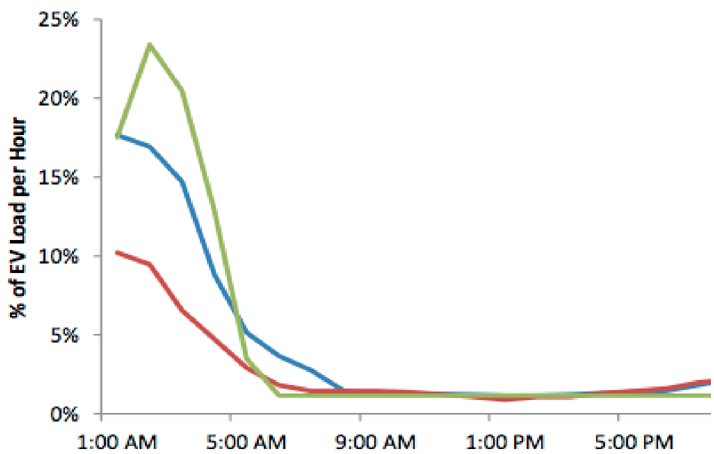


Figure 29: Daily EV load for major California utilities using TOU rates

Utilities like NV Energy and Arizona Public Service have already seen substantial avoided costs by employing TOU rates and have been able to share those savings with

EV customers, who, on average, saved $245 and $446 each year respectively by taking advantage of special TOU rates.
While the TOU rate are beneficial to the system currently, at scale, TOU rates could present new challenges for utilities to grapple with in terms of EV integration:

- TOU rates are static, and asVREpenetration increases, excess solar and wind may be available during times when TOU pricing is high, thereby preventing demand to shift to absorb excessVRE
- Sudden changes in pricing between periods could exacerbate ramp periods or create new ramp periods as all EVs jump online at the same time when off-peak prices come into effect

Recently, in areas like California where EVs are already becoming a substantial load, utilities are seeing some of the limits of the simple TOU model and want to test new approaches to shape EV deployment. Contained here are two cases of utility-based EV programs: the first case is focused on deploying EV smart charging at scale and shaping EV charging profiles through real-time pricing, while the second case is focused on testing EVs as dispatchable demand respond resource.

5.3 Power Your Drive: SDG&E's EV charging station Pilot

The Power Your Drive (PYD) program moves from existing TOU rates to real-time pricing in an attempt to more precisely and significantly shift customer charging behaviour.
While the shift to real-time pricing is substantial, the pilot attracted more attention due to the expanding role of the utility as an agent to drive EV adoption in pursuit of California's aggressive climate goals.

The program allows SDG&E, one of California's regulated utilities, to build its own smart chargers to support state goals of decarbonized transportation sector. SDG&E was approved a 3-year budget of $45 million by the regulatory commission to install 3500 electric vehicle (EV) charging stations at 350 multi-family or workplace locations in January 2016[118].
A total of 170 EV drivers are currently enrolled in the program accounting for 6,777 unique charging sessions[119], 62,705 kWh in volume and $10,442 in power bills. Although the current user pool is relatively small, initial results have been promising and will be studied further in the coming phases of the project.

In the following sections we explore how SDG&E set up the **real-time pricing** program in order to elicit significant customer behavioural change, how charging was coordinated through a centralized IT platform and a user-friendly app, and the detailed mechanisms this program employed to ensure the utility ownership model maximized benefits for all customers.

Real-time pricing

SDG&E's choice to use real-time market prices, not time-of-use (TOU), for its EV rates made a big difference in reducing system flexibility needs and costs, by:

- Coordinating charging withVREavailability
- Sending significant price spikes to create customer response

118 https://www.sdge.com/sites/default/files/regulatory/FINAL_Power_Your_Drive_Semi_Annual_Rpt.pdf

119 Ibid.

- Ensuring prices still resulted in lower fuel costs for customers who responded to those rates
- Enabling EV charging behaviour to be predictable and coordinated through a common payment app

In the pilot, day-ahead locational wholesale market prices are passed through to customers for each particular charging station location, thus revealing the true real-time cost or value of charging the EV in every moment, and creating more substantial pricing swings to guide user's charging behaviour, compared with the traditional time-of-use (TOU) pricing. TOU pricing stays static over the course of an entire season, and usually has only 3 pricing periods, which correspond with demand, not day to day changes in renewable availability and system stress. (table 8)[120].

	On-peak	Super off-peak	Off-peak
Time of day	4 PM – 9 PM (Every day)	Midnight – 6 AM (Workdays) Midnight – 2 PM (Weekends & Holidays)	All other hours
Summer rate (Jun 1 – Oct 31)	$0.54/kWh	$0.22/kWh	$0.28/kWh
Winter rate (Nov 1 – May 31)	$0.24/kWh	$0.22/kWh	$0.23/kWh

Table 8: SDG&E Electric Vehicle Time-of-Use Rates

This difference pales in comparison with real-time pricing seen in Power Your Drive, where rates start around $0.13/kWh and can soar to over $0.5/kWh during very hot days where the grid is heavily constrained due to high U.S.ge[121]. Figure 30 is an example of the Power Your Drive price schedule for a site in San Diego on August 12, 2018[122]. The highest price point reached more than $1.28/kWh, more than 2 times of TOU rate in peak hour.

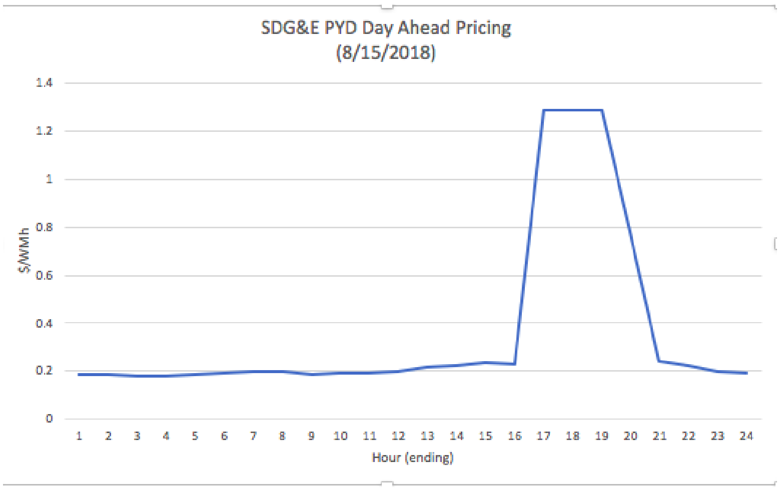


Figure 30: SDG&E Power Your Drive Day Ahead Pricing

Customer benefits

Based on the pilot, estimated fuel savings for all EV users charging through Power Your

120 http://webarchive.sdge.com/clean-energy/ev-rates

121 http://webarchive.sdge.com/clean-energy/electric-vehicles/power-your-drive-faq

122 https://www.sdge.com/pyd-day-ahead-pricing?id=1071

Drive is total $13,764, about 55% cheaper compared with using a gasoline car. The average charging price in Power Your Drive is below $0.167/kWh, cheaper than the lowest EV TOU rate that SDG&E current holds. Even if EV users who charged through Power Your Drive are now charging only during the super off-peak pricing period for the current SDG&E rate, they will still spend 32% more money than the Power Your Drive average.

Customer Charging Coordination Platform

In order to give customers the control necessary to respond to real-time price fluctuations, a coordinating IT platform was necessary, including a utility payment app that can be used across all chargers in the program. This app allows customers to set maximum prices they are willing to pay, and when prices exceed that level, their cars stop charging[123]. This is set per car, so regardless of which charger that customer is using, the same rules will apply. However, customers always maintain the option to override any rules set in the app, as opposed to other programs for demand response that may require firm compliance during critical hours.

Since the pricing system uses locational marginal prices on the market, having foresight into which areas may have congestion and therefore may have high prices on that day is also essential to create customer response. Therefore, the app sends customers day-ahead locational marginal market prices to help them plan when to plug-in their vehicles and which chargers to use. The locational component is integrated with Google Maps to assist drivers to find available chargers and their current/expected prices (figure 31)[124].

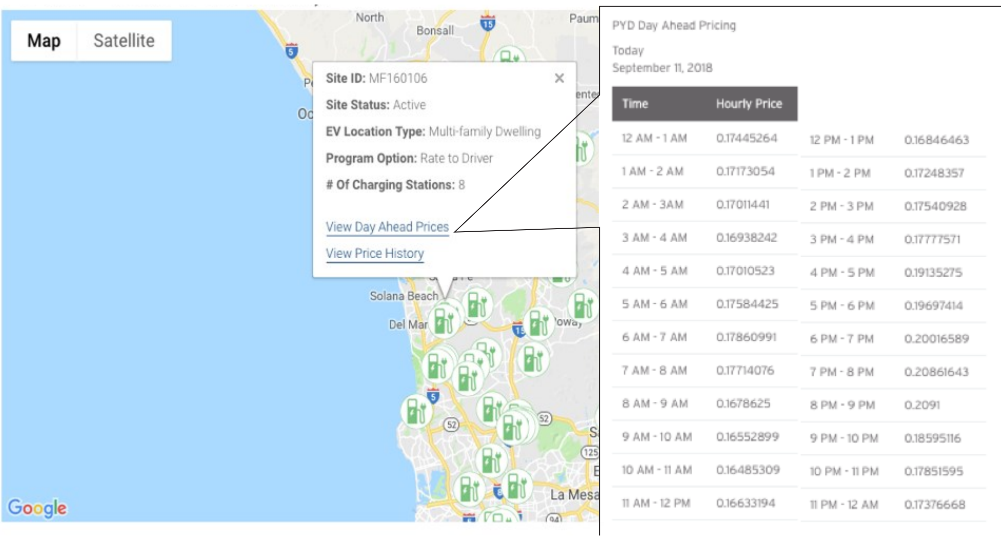


Figure 31: Power Your Drive day ahead rates map

Utility ownership model

The program allows San Diego Gas and Electric (SDG&E), one of California's regulated utilities, to build its own smart charging infrastructure. Instead of permitting the utility to invest in new fossil fuel assets to meet peak growth, regulators and SDG&E instead initiated the Power Your Drive pilot to assess if appropriately rolled-out EV charging infrastructure investment could effectively defer traditional grid investment (and further support EV

123 http://www.sandiegouniontribune.com/sdut-sdge-charging-stations-2016may16-story.html

124 https://www.sdge.com/pyd-map

adoption).

The decision to allow the utility to own and operate charging infrastructure was nevertheless a contentious one. The primary arguments made in favor of the utility-owned model are:

- Access to cheap capital upfront and having a large, central procurement agent is essential for ensuring smart charging is available on day-one and can be quickly scaled up.
- Utilities are in the best position for a coordinated and equitable roll out.
- Utilities can play a key role in standardization and providing transparent and verifiable information to develop the industry.

Access to cheap capital

Access to utility rate-based capital was essential to deploy charging infrastructure at scale. 3rd parties only deploy chargers when there are enough vehicles to earn a return, creating a chicken-and-egg problem where the lack of EV infrastructure discourages people from buying EVs.

Coordinated and equitable roll out

State regulators were particularly focused on EV infrastructure siting decisions, ensuring that they:

- Met the objectives of the program (e.g. alleviating car-related criterion pollutants, addressing access concerns for under-privileged customer bases, etc.)
- Worked to minimize the need for further utility investment, or actually enhanced utility economics (e.g. new EV charging sites do not contribute to distribution system stress)

Since regulators see access to EV charging as a public good, they believed a regulated entity needed to be involved in siting to ensure equitable distribution. This was especially important from a social equity perspective, where low-income neighborhoods lack of access to charging prevented them from benefiting in the proven fuel savings associated with EVs, or sometimes they did own EVs, but charging may be done through unsafe, uncoordinated means (e.g. extension cords run out of windows). The program also targeted workplaces to allow for more charging during the day when solar is available. This helps ensure EVs are connected to the system when renewables are available and occasionally curtailed, as most of CAISO's renewable curtailment happens during work hours between 9 AM - 1 PM[125].

SDG&E is also in a position to optimize site selections from a distribution grid perspective, avoiding concentrating chargers on certain already strained distribution system equipment, and thereby avoiding new distribution upgrades.

Standardization and data transparency

In the Power your drive program, a standard information platform was developed that allows utilities to directly collect EV charging infrastructure U.S.ge and charging data. Making data accessible and transparent to all market participants is essential to support further industry development. In this case, SDG&E's majority ownership of chargers early on enabled and incentivised their investment in a central standardised coordination platform. SDG&E is also in a good position to develop the payment app for such EV chargers since they already have the data privacy protocols in place to handle customer electricity bills.

---

125 https://www.caiso.com/Documents/Wind_SolarReal-TimeDispatchCurtailmentReportApr24_2018.pdf

Concerns around the utility ownership model
The utility ownership model has drawn substantial debate: some argue that it was essential to deploy the EV charging at scale and in an optimized way, while others argue that regulators were too hasty to give this business to the utility, eliminating opportunities for the free market to participate. The main objections to utility ownership were:

- It falls outside the regulated role of the utility and that other uses of that capital could have been used to help keep electricity costs down, if that was the primary justification for the utility owner-operator model.
- It violates principles of fair competition, and perhaps outside vendors, if given the chance, could have provided better solutions than those arrived at by SDG&E

5.4 BMW iChargeForward: PG&E's Electric Vehicle Smart Charging Pilot

While the SDG&E program tests the foundation of EV load shifting through smart charging pricing, EV charging programs can go beyond load shaping, to using EVs as dispatchable resources. PG&E and BMW's iChargeForward are testing exactly that: how EVs can participate directly in dispatch as demand resource.

Pilot overview

From July 2015 to December 2016, Pacific Gas & Electric and BMW worked together on a vehicle-to-grid (V2G) demonstration pilot to prove the technical feasibility of using electric vehicle charging as a flexible and controllable grid resources. This pilot aimed to understand the potential of using Electric Vehicles (EV) for grid services, particularly their availability, response level, and response speed.

The EVs, 97 BMW i3 vehicles in the San Francisco Bay Area, participated as a 100kW aggregated proxy demand response (PDR) resource in CAISO's day-ahead and real-time energy. When EVs received grid signals (via telematics embedded in the vehicle) they would delay their charging for one hour.

Since a critical piece of this pilot was to see how EVs could respond to ISO dispatch signals and be assemble as current wholesale market DR product, the full 100kW needed to be available for dispatchers as other traditional DR resources follows the same requirements. But the pilot could not sacrifice EV charging to respond to DR dispatch signals, therefore back-up batteries were also installed to one-for-one cover for EV battery capacity. This additional 100kW of second-life (used) EV batteries were located at BMW Group's Technology Office U.S. site was used to cover any portion of the dispatch signal EVs could not cover. These batteries were already built to support BMW's campus microgrid system integrating on-site renewable energy, managing customer demand charges, and supporting main building functions in the case of a power outage.

Findings

Overall, the pilot proved out the concept that EVs could technically act as a grid resource. The resource responded to 209 demand response events called over the 18-month pilot period, successfully meeting expected dispatch results 90% of the time, totalling to 19,500 kWh of energy services provide[126]. The pilot also confirmed the expectation that most EVs are not available during critical peak demand response events. In large part this is due to TOU rates already limiting their contribution to peak demand. But given EVs were able to respond to grid signals within the accepted timeframes (under 4 minutes for participation in real-time and day-ahead markets), the pilot shows promise for EVs to be dispatched via market signals[127]. And as latency[128] is reduced further, grid operators could explore using EVs for regulation services, which requires smaller, but faster responses from resources. The pilot also showed that delays in charging were highly acceptable to EV owners. Further details of the major findings are below:

Availability
Given EVs already shift charging away from peak hours, EVs can contribute little to critical peak demand response which is most of what CAISO currently calls for DR today. Across all the DR events called during the pilot, on average 20% of the total contribution was attributed to the vehicle pool and 80% from the 2nd life stationary battery system.

The pilot helped develop a clearer picture of when vehicles were plugged in, charging, and be able to delay charging (figure 32)[129]. If an event is called from 11 PM to 12 AM, the vehicle pool could contribute over 35% of the 100-kW dispatch signal since many EVs were already charging during that time and could reduce power withdraw. Whereas during morning ramp-ups as people drive to work, EVs could only contribute an average of 5% as fewer EVs are connected to the grid. The highest hours of EV contribution during the pilot was 50%, during a dispatch between 12m-2am, albeit this dispatch is infrequently called for in day to day grid operation.

While under CAISO's proxy demand response product aggregators could just offer EV capacity during hours of known availability, the frequency of dispatch would be minimal during those hours under current system conditions. Therefore, it is critical to aggregate EV with other supplement energy DR or battery systems to meaningfully participate in today's DR markets, especially when the EV pool is relatively limited in size.

126 BMWi ChargeForward: PG&E's Electric Vehicle Smart Charging Pilot Report, http://www.pgecurrents.com/wp-content/uploads/2017/06/PGE-BMW-iChargeForward-Final-Report.pdf

127 As general supply resources that could change the system's net load, not only be used to shave the peak demand

128 The reaction time between dispatch and the resource are actually fully delivered

129 BMWi ChargeForward: PG&E's Electric Vehicle Smart Charging Pilot Report, http://www.pgecurrents.com/wp-content/uploads/2017/06/PGE-BMW-iChargeForward-Final-Report.pdf

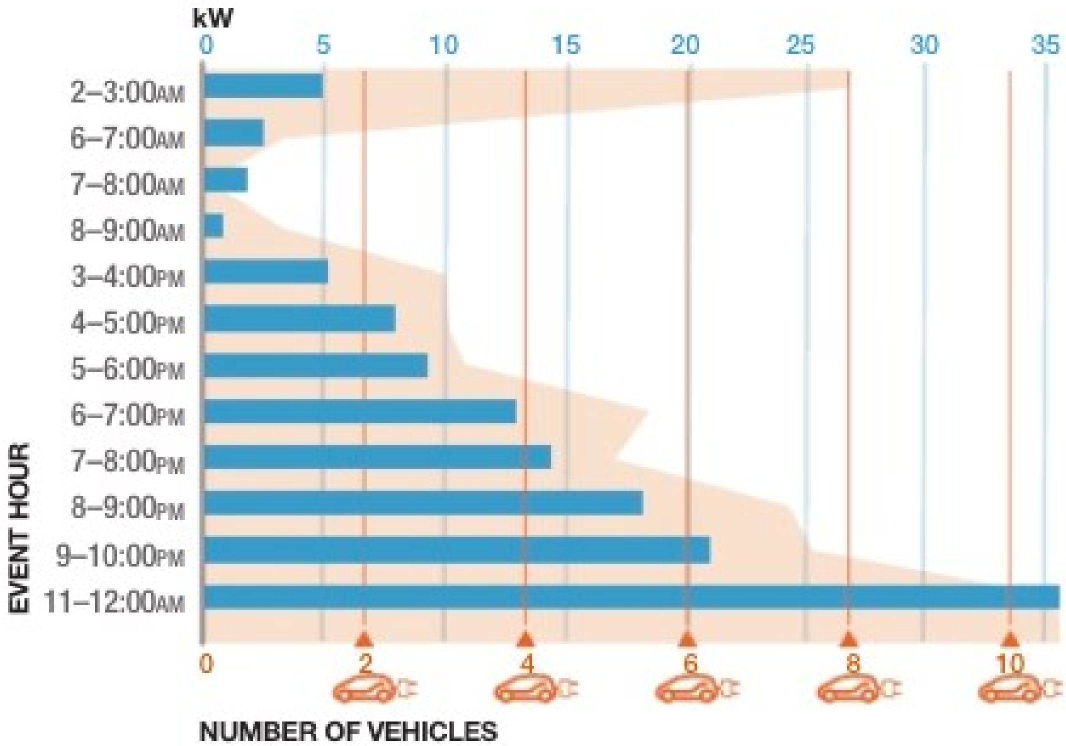


Figure 32: Aggregated Power Draw by Driver Archetypes

Customer response

95% of customers indicated in a survey they saw little to no impact of their charging and driving behaviour, indicating a high willingness to continue participating. Over the course of the study only 6% of users ever selected to opt out, which is below levels of opt-out associated with most residential DR programs. A customer's willingness to respond to a DR event had little to do with customer acceptability, but was almost entirely determined by availability. EVs that participated in the most events (some participating in over 50 events) were users who were not already enrolled in an EV TOU charging rate. These customers also drove their EVs frequently and on a very consistent schedule, resulting in longer periods of charging and consistent coincidence with DR events (figure 33). This reinforces the takeaway that as EVs become more prevalent, system operators do have to determine whether EVs at scale are best to just respond on their own to system conditions and pricing (e.g. through TOU tariffs or real-time pricing), or if they are more valuable as dispatchable resources in critical events, and how to effectively coordinate between those two modes of EV load shaping.

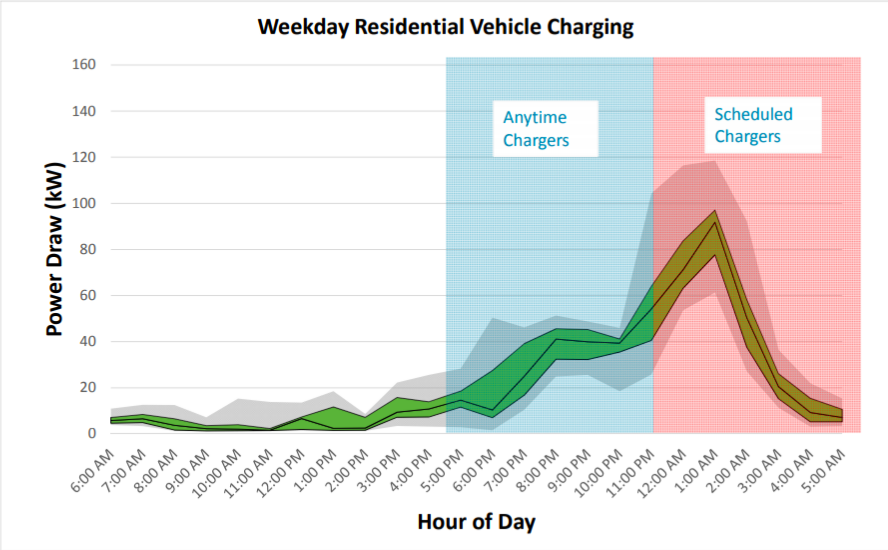


Figure 33: Weekday Residential Vehicle Charging profile for PG&E territory

Latency and responsiveness
While the pilot moves into further stages, a promising indication is that the technology of sending signals to cars over cellular systems can feasibly respond in less than 4 minutes required by wholesale energy markets. The average latency per event was only 2 to 3 minutes.

While this pilot was small in scale, it shows promise for EVs to go beyond simply coordinating their charging with the grid, toward increased levels of dispatchability, including some vehicle-to-grid (V2G) pilots where electric vehicle batteries could even inject energy back onto the grid to provide highly lucrative services in critical situations.

5.5 Weighing the Trade-off: EVs as demand-side or supply-side resources

These two pilots reflect two, conflicting approaches to EV integration: whether EVs should be integrated as demand-side resources (SDG&E) or supply-side resources (PG&E). In most contexts today, the demand-side approach creates greater value to the grid and the EV customer (in terms of deferred investment and avoided energy costs) and is simpler to implement. The supply-side approach, even if EVs charged with no grid coordination, would likely see less frequent dispatch, and not capture the full benefits of peak demand reduction, presenting a less sure-fired approach to deferring new peaker plant capacity.

In the future, there may be moments where EVs will be plugged in, and charging at low-rates, but a 5-minute dispatch interval could come up to manage an unexpected ramp of RE, and that EV could interrupt its charging for that 5-minute interval, getting paid for this balancing service. As more is understood about EV charging patterns, customer behaviour and perceptions, this balance and coordination between supply-side and demand-side approaches to EV coordination will continue to evolve.

5.6 Relevance in a Chinese context

Roughly, half of the world's EVs are in China. China's power system is now also facing both the challenges and opportunities brought by EVs. On one hand, uncoordinated EV charging would lead to unacceptable peak demand for utility companies; on the other hand, EVs could function as flexible loads to consume electricity surplus generated by wind and solar power. The most commonly used model for coordinated charging in China is the TOU rates model. But it is foreseeable that, the TOU model would not be sufficient as the number of EVs further grow (a new peak demand will be induced at the low price moment). Shifting from the TOU model to a more interactive smart charging model would be a necessity in the near future. Pilot programs in California have already proved the feasibility of using EVs as new flexibility resources. However, these programs are still operated at a relatively small scale. How to scale up the EV smart charging programs to a multi MW level still needs further explorations.

Chapter III
The Nordic and European Power market

Resume

The European internal market for energy is unparalleled. In particular the Nordic region is an example of how regional markets and the right price signals have allowed for some of the highest renewable energy levels being integrated into a highly flexible and highly interconnected power system. This chapter provides examples of how the most important flexibility resources - hydro power, flexible combined heat and power plants, power to heat, and industrial demand response – have been incentivised to actively provide flexibilility to the Nordic power system.

III. Europe/Nordic countries

1. The Nordic and European power market

1.1 Historic development

Finland, Norway, Sweden and Denmark (app. 27 million inhabitants) deregulated and integrated their respective national electricity markets during the 1990s into a joint Nordic electricity market.

The deregulation enabled the dissolution of vertically-integrated monopolies, through a separation of ownership of grids and system operation from power production (generators) and suppliers. The transmission and distribution of electricity remained a controlled monopoly with regulated tariffs; while the production and trading of electricity became subject to competition (without price regulation), facilitated by the Nord Pool power exchange. The deregulation opened up the market by giving, to any power producer, a non-discriminatory access to the transmission network. The Transmission System Operators (TSOs) are responsible for the security of supply, i.e. the balancing of the system; and for providing a well-functioning transmission (high-voltage) grid. They are a non-commercial organization (generally state-owned), neutral and independent of commercial actors.

Another important choice was to implement a zonal market design based on price areas with one common electricity price, where transmission and distribution costs are not included. This is in contrast to the U.S. model, where nodal prices enable a very high geographical resolution of price signals.

The integration and coupling of the national electricity markets into one common Nordic market involved the removal of barriers to cross-border trade, such as border tariffs; and started with the establishment of the Nord Pool power exchange. The power exchange was established as a department in the Norwegian TSO in 1993; and in 1996, when Sweden joined, it became the world’s first international day-ahead power market. Finland and Denmark respectively joined in 1998 and 2000. The Baltic States joined during the period 2010 – 2013.

The common Nordic electricity market is the most harmonized cross-border electricity market in the world, as a result of many years of consolidation and merger activities. Strong interconnectors enable high-level trading within the Nordic countries, as well as with neighbouring countries (e.g. Holland, Germany, and Poland).[1]

Market organization

The Nordic electricity market and most of the markets in Europe are based on the self-dispatch principle: this means that, while Transmission System Operators (TSOs)are

1 The existing Nordic regulating power market” ‚EA Energy Analyses, http://www.ea-energianalyse.dk/reports/1027_the_existing_nordic_regulating_power_market.pdf

responsible for balancing real-time electricity production and consumption, it is the market actors (i.e. wholesale power producers/sellers and buyers) that determine day ahead (i.e. the day before the day of operation) their own generation and consumption schedules and their hourly dispatching. Since the market actors decide their dispatch position based on their own economic criteria, i.e. profit optimization, the optimizing strategy and decisions are handed over to each market actor. They are combining in-house information with sophisticated market models to take decisions that would maximize their profits in both the short term (daily dispatch), as well as in the longer run (investment decisions). The alternative to the self-dispatch principle is a central dispatch model, where the system operators are responsible for the dispatching and unit commitment.

The deregulation of the power market led to the development of some new central market actors. Compared to central planning, in the market-based power market there are two central actors in charge of Balancing Responsible Party (BRP) and of physical power exchanges.

Balancing Responsible Party (BRP)

A BRP is aggregating expected daily consumption and production on behalf of wholesale electricity consumers and power generators. A BRP is a market actor who has concluded an agreement on balance responsibility with the TSO, and which then becomes financially liable for imbalances. BRPs thus function as the intermediate between the producers and consumers on one hand, and the power exchange and TSO on the other hand. Large power generators typically undertake the BRP role themselves, while smaller market actors let a BRP access the power market and trade on behalf of them to minimize their transaction costs. BRPs, therefore, buy and sell power on the power exchange (Nord Pool) on behalf of the power generators and electricity consumers they represent. BRPs are economically responsible to the TSO for any imbalance between their forecasts and the actual daily production/consumption. In case of wrong predictions, the TSO will still secure real-time balancing, and the cost associated with compensating imbalances will be charged to the BRPs.

Power exchanges

The Nord Pool is a physical power exchange and service market actor in the wholesale electricity market. The Nord Pool power exchange is owned by the Nordic and Baltic TSOs. Its most central roles are the development of a merit-order (least marginal cost) curve based on quantities, the ranking of the BRPs' bid prices, and the setting of the day-ahead prices. This creates high price transparency and continuously signals the value of production on an hourly basis; hence creating clear incentives for flexible operations. The power exchange thus facilitates flexibility but also introduces a price volatility risk that can be difficult to handle. Physical power exchanges are hence often developed together with financial power exchanges for risk management. The day-ahead prices are used as the underlying price reference for financial contracts and hedging products enable market actors to trade with their desired risk profile. In figure 34 the general organizational setup for the Nordic electricity market is depicted.

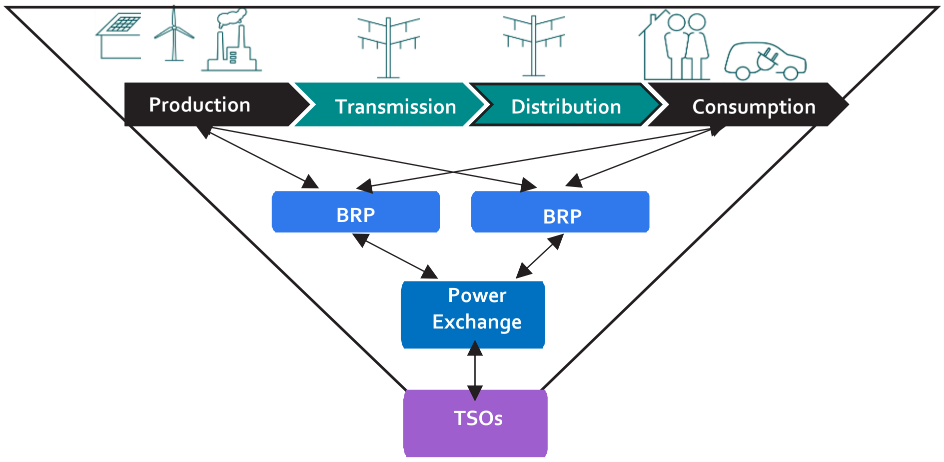


Figure 34: General organizational setup for the Nordic electricity market. Source: Energinet

1.2 The Nordic short-term electricity markets

The Nordic short-term physical electricity markets consist of three markets with different time scheduling: the day-ahead market, the intraday market, and the balancing market (figure 35).

The day-ahead and intraday market are operated by Nord Pool and opened for trading to market participants; while the balancing market is under the responsibility of the TSOs (single buyer market). Electricity is traded on the day-ahead market. In addition to the three short-term markets, the long-term financial market is supporting long-term risk management with monthly, quarterly and annual forward price run by a separate financial exchange (NASDAQ)[2]

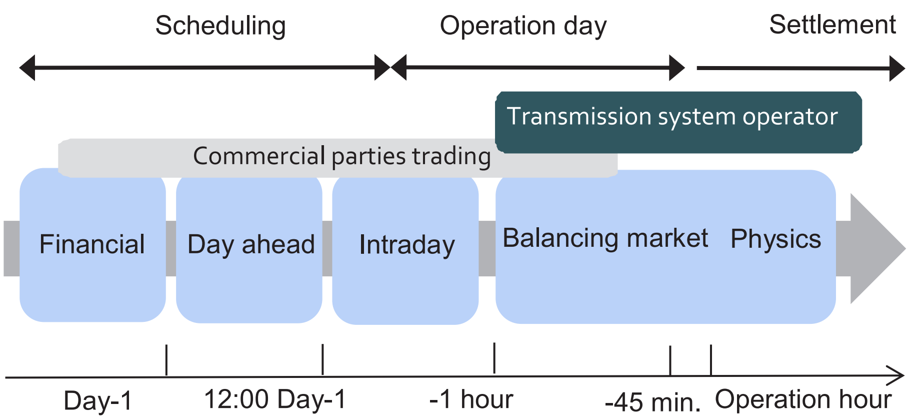


Figure 35: Main phases of the Nordic power market. Source: Energinet

Day-ahead market

Not later than the day before the day of operation BRPs trade via the Nord pool day-ahead

2 For more information see: https://www.nordpoolspot.com/the-power-market/ and http://www.nasdaqomx.com/commodities/markets/power/nordic-power

market in order to cover their consumption and sell their output for every hour of the 24 hours of the subsequent days. The BRPs submit economically binding notifications for production, consumption, and trade for the subsequent 24-hour period to the TSO. This trading is finished 12 hours before midnight (so at 12 AM) and constitutes the day-ahead market.

The day-ahead market is thus a daily auction where the cheapest production offers for each hour will be used to fulfil the given demand for each hour in the coming 24-hour period (midnight to midnight). This results in a least-cost dispatch where the marginal cheapest production is prioritized and unique hourly electricity prices for the upcoming 24 hours are determined. The settlement price in each hour is thus the marginally most expensive offer that is accepted to meet the demand in the given hour. All market actors receive/pay the same settlement price in the given hour (i.e. 'pay as cleared' pricing). The regulatory minimum and maximum price cap are currently negative 500 EUR pr. MWh and 3,000 EUR pr. MWh.

The volatility of the day-ahead prices reflects the dynamics between demand level, supply level and the marginal cost of delivering the supply. The variation of the day-ahead market prices over a 24-hour period creates a very clear economic price signal for producer and consumers. The future expectations on the prices are a decisive factor to choose the investments in capacity and flexibility that are most optimal to make. The Nordic day-ahead market is, therefore, the primary price signal for efficient dispatching and balancing, as well as decisive information for management of investments.

Intraday market

Intraday markets for electricity allow for continuous trading every day around the clock until one hour before delivery. The intraday market enables market actors to make up for any changes in their production/consumption forecasts after the day-ahead market has closed. It offers to market actors the opportunity to compensate for unexpected imbalances and to offer their own unused flexibility. In 1999, the world's first intraday market started in Finland and Sweden and has since spread to the other Nordic and Baltic countries. This common cross-border intraday market is named Elbas.

BRPs are financially responsible for ensuring that the notifications for production, consumption, and trade sent to the TSO are in balance. The approved notifications can be changed up to one hour before the hour of operation through trading in the intraday market or through bilateral trade. After the day-ahead market is closed (12 hours before operation hour) the market actors can trade in the intraday market. The intraday market is opened for trading from 12 hours before production until 60 minutes before production.

There can be significant changes between the day-ahead estimated production, and what market actors can actually deliver in operation. This could be due to changes in wind forecast or an unplanned outage. When the trading in the intraday market is closed, and market actors have submitted their final schedules, the TSO takes the balancing responsibility and procures energy in the Nordic market to secure system balance - at a possibly higher cost[3].

3 For more information see: Data is public available for day ahead prices, volumes, export/import flows as well as intraday volume and export/import flows etc. at: www.nordpoolgroup.com/Market-data1#/nordic/table

Balancing market (and ancillary services (reserves))

After intraday gate closure the Nordic TSOs takes over the responsibility for the physical balancing of the electricity system. The balancing reserves are procured by the TSO (single buyer) among the market actors and consist of products that cover a different spectrum of time scale.

The fastest frequency stability of the transmission system is secured by Frequency Containment (primary) Reserves and Frequency Restoration (secondary) Reserves. Capacity on these markets is being paid to be ready to potentially become activated within 15 seconds to 2-5 minutes. They are activated automatically in accordance with frequency deviations (when the system needs regulation), but are generally expensive and have limited capacity.

To minimize the use of expensive automatic reserves, slower replacement reserves (manual reserves) are activated through a request from the TSO. Manual reserves are called regulating power in the Nordics and there is a joint regulating power market (NOIS list) managed by the TSOs. BRPs can make bids to the regulated power market with their amount (MW) and price (DKK/MWh) for down- and up-regulation power. All bids are collected in the joint Nordic NOIS-list and are sorted with increasing prices for up-regulation and decreasing prices for down-regulation; thus creating a merit order list of separate bids of down- and up-regulation. Just as with the day ahead market, the regulation price is settled and paid as cleared i.e. all market actors face the same price.

To secure sufficient manual reserves TSOs rely both on advanced purchasing of capacity and on the market participants' voluntary bids. TSOs then use the reserves and the voluntary bids to balance the system. Currently, the minimum bid size on the regulation power market is 5 MW, which means that only larger industrial consumers or producers can place bids or alternatively in a portfolio of bids. The required time from activation to full power delivery is 15 minutes.

Often the payment for the different types of reserves consists of a capacity payment (i.e. payment for being available) and an activation payment for the actual delivery of the service. In table 9 an overview is given on the general balancing categories in the Nordics.

General balancing categories in the Nordics	Activation time	Nordic market size	Payment
Frequency containment (primary reserves)	Automatic activation Full effect within 15-30 sec.	1,200 MW	Reserve payment only
Frequency restoration (secondary reserves)	Automatic activation	300 MW	Reserve + activation payment
Replacement (manual reserves)	Manual activation Full effect within 15 min. of activation	N-1 (load frequency control area)	Common Nordic market for regulating power (NOIS) Reserve + activation (voluntary bids) payment

Table 9: Overview of general balancing categories in the Nordics

Nord Pool handles approximately 75 % of the total market volume. The remaining share is traded bilaterally over-the-counter outside of the Nord pool. This high penetration and market liquidity in the Nord Pool are a guarantee for market actors that both price signals are reliable and flexibility is rewarded.

The yearly traded volume in the Nordic financial market is in the range of 700 TWh, which is almost twice the total physical production of 400 TWh. It shows the high liquidity and importance of the financial market. The power traded in the day-ahead market is roughly 300 TWh, while in the intraday market size it is roughly 5 TWh, which is comparable in size to the regulated power market (4 TWh). In table 10 an overview of the three separate but interrelated short-term markets in the Nordics is given.

The high share of flexible hydro production in the Nordics results in relatively low imbalance costs in the balancing markets, and reduces the need and incentive to use the intraday market compared to other regions in Europe. The remaining available interconnector capacity after the day-ahead market is made available to the intraday market[4].

Nordic Market overview	Day ahead (power exchange)	Intraday (power exchange)	Regulated power market
Market type	Market actors through joint auction	Market actors through continuous and bilateral trading	Single buyer (TSO)
Pricing	Pay as cleared	Pay as bid	Pay as cleared
Annual volume (Nordic)	400 TWh	4 TWh	3 TWh
Interconnector capacity allocation	First priority	Second priority	Third priority

Table 10: Overview of the three separate short-term markets in the Nordics

1.3 Nordic production mix and power prices

The electricity production mix in the Nordic countries is highly dominated by hydro (roughly 60 %), but covers large national differences as illustrated in figure 36. Denmark is characterized by a share as large as almost 50 % of its production coming from wind power, while the remaining comes from thermal power. Norway is characterized by being hugely dominated by hydropower just as Sweden is, but Sweden is also using nuclear and has a growing share of wind and solar power. Finland has a relatively equal share of hydropower, nuclear and thermal power.

4 For more information see: “The Nordic Electricity Exchange and The Nordic Model for a Liberalized Electricity Market” at https://www.nordpoolgroup.com/globalassets/download-center/rules-and-regulations/the-nordic-electricity-exchange-and-the-nordic-model-for-a-liberalized-electricity-market.pdf

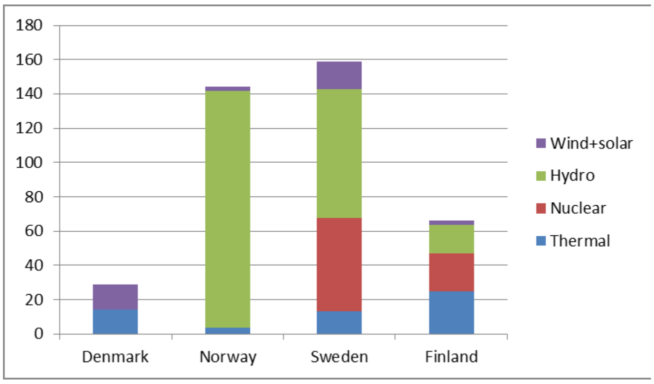


Figure 36: Nordic power production in 2016, TWh.[5]

On the one hand, the high share of hydro-power leads to a large yearly difference in hydro-based power production due to large differences in the level of precipitation from year to year. On the other hand, this large share of hydropower (the majority is reservoir hydro) allows for a very flexible production. Thanks to this substantial flexible hydropower the volumes in the Nordic intraday market are relatively small compared to other European countries and regions. Hydropower represents 90% of the volume in the balancing market in the Nordics, while thermal producers represent 8% and consumption is approximately 1%.

Due to the low temperatures during the long winters in the Nordics, heating represents a large part of the energy consumption and plays an important role in the power sector. The heating system in Norway and Sweden is mostly based on both electric and water-based individual heat distribution and is able to shift from electricity to other alternatives if power prices are very high. Contrary to Norway and Sweden, the majority of the heat in Denmark is supplied by district heating produced by both large and small scale combined heat and power (CHP) plants.

On the one hand, the high share of hydro-power leads to a large yearly difference in hydro-based power production due to large differences in the level of precipitation from year to year. On the other hand, this large share of hydropower (the majority is reservoir hydro) allows for a very flexible production. Thanks to this substantial flexible hydropower the volumes in the Nordic intraday market are relatively small compared to other European countries and regions. Hydropower represents 90 % of the volume in the balancing market in the Nordics, while thermal producers represent 8% and consumption is approximately 1%.

Due to the low temperatures during the long winters in the Nordics, heating represents a large part of the energy consumption and plays an important role in the power sector. The heating system in Norway and Sweden is mostly based on both electric and water-based individual heat distribution and is able to shift from electricity to other alternatives if power prices are very high. Contrary to Norway and Sweden, the majority of the heat in Denmark is supplied by district heating produced by both large and small scale combined heat and power (CHP) plants.

Wholesale power prices

The day-ahead power price development in the Nordics over the last 15 years is shown in figure 37 (lhs). The large yearly difference has been driven by the amount of hydropower available in the given year in the Nordic region.

5 https://www.nordpoolspot.com/historical-market-data/

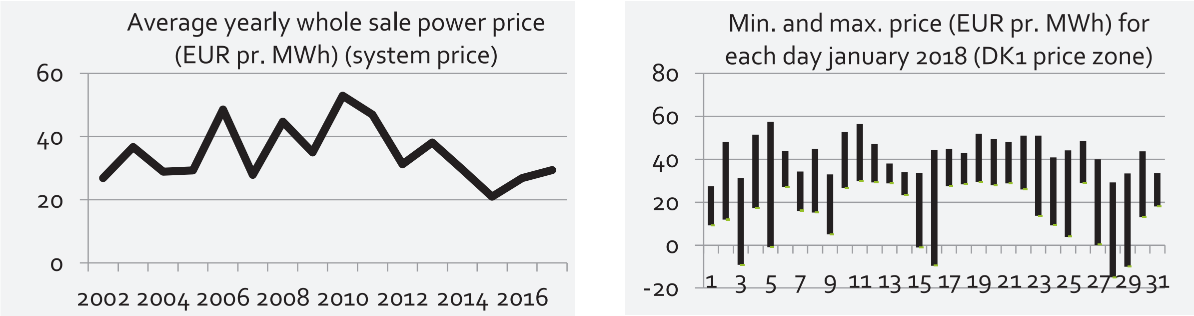


Figure 37: Average Nordic wholesale price (lhs.) and minimum and maximum prices in DK1 for each day in January 2018 (rhs.)

While the yearly average price reflects the average price/value of electricity production, the main result or effect of the day-ahead market is the general volatile prices, not only over the year but also within 24 hours. In figure 37 (rhs) the respective minimum and maximum price for each day in January 2018 for DK1 price area is depicted, clearly showing both the large variation within each day and the variation between days.

1.4 The Nordic market's bidding zones and market coupling

The power market in the Nordics, as well as the rest of Europe, is based on a zonal pricing model. This implies that all market actors within a given bidding/pricing zones get the same market prices. The Nordic market itself is divided into several pricing zones. Finland constitutes a single price zone, while Denmark, Sweden and Norway are divided into several pricing zones, generally reflecting internal congestions in the transmission grid. Figure 38 shows the price zones in the Nordic power market. Since 2010-2013 the three Baltic countries have been integrated.

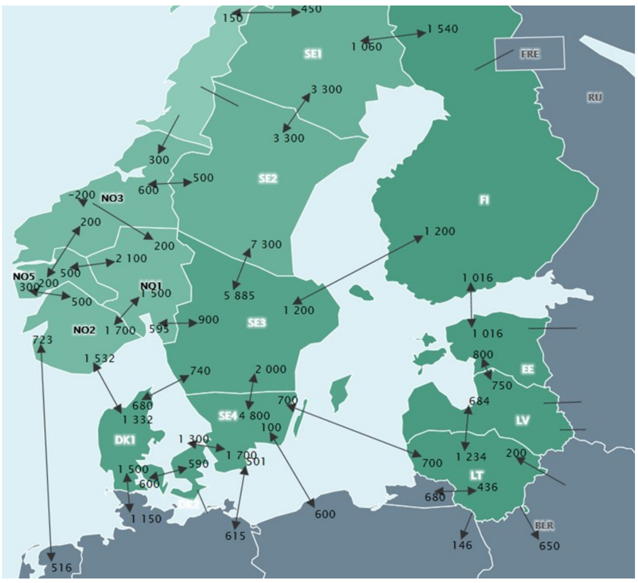


Figure 38: The 15 price zones in the Nordic Power Market[6]

6 Source: http://driftsdata.statnett.no/Web/map/snpscustom (updated with prices and power exchange continuously)

A prerequisite for coupling neighbouring power markets is to allow third-party access to the interconnectors through the market. In Europe, two methods have been identified and used:

- Explicit allocation of capacity (transmission capacity is allocated to the market separately and independently from the marketplaces where electrical energy is traded)
- Implicit allocation (capacity and energy are auctioned together)

The coupling of the Nordic day-ahead markets is based on the concept that TSOs "give" all Available Transmission Capacity on the interconnectors between the Nordic pricing zones to Nord Pool. Thus the only possibility to trade between pricing zones is to trade on Nord Pool. Nord Pool collects all bids and offers in all bidding zones for the following 24 hours and calculates, through an algorithm that optimizes social welfare, clearing price, traded volumes of all participants and flows on the interconnectors based on. More specifically Nord Pool uses implicit allocation of the interconnectors: interconnector capacity is allocated concurrently with electricity being traded (this is the case for both the day ahead market coupling as well as the intraday market coupling).

Doing this secures that the direction of electricity flow on the interconnectors practically always flows from low price zones (exporting) to high price zones (importing), thus optimizing the overall social welfare. In a situation with unlimited interconnector capacity between price zones, the price will be the same in all prices zones. This theoretical price is termed the system price. In reality, interconnector capacity is limited for many hours of the year. Theoretically marginal cheaper production in one price zone could in principal substitute more expensive production in a neighbouring price zone, but in practice, it is not possible to do so due to the limited capacity of the interconnectors (the export/import is already at its fullest capacity). In these situations - when the algorithm shows congestion (i.e. it is not physically possible to transport all the electricity that else would minimize overall production costs) then the bidding zones will have different power prices.

Take the case of Denmark, which has 2 price zones (DK1 and DK2) connected by a 600 MW interconnector. In the first half of 2018, the prices were the same in the two areas in approximately 75% of the hours, illustrating that there exists congestion between the two areas in 25% of the hours. A larger interconnector capacity would reduce congestion leading to a higher social welfare, as the aggregated demand in the two price zones would be fulfil with an overall cheaper production. However the economic gain might be too low to support an additional investment in increasing the interconnector capacity.

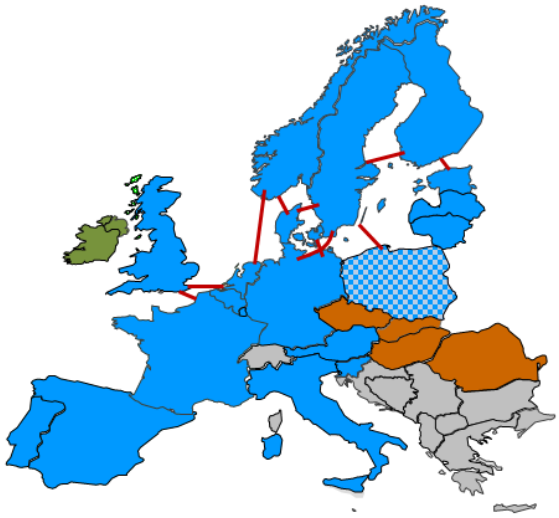
1.5 European power market

In the last 20 years the European Commission has pushed forward a European liberalization of the electricity production and consumption, and has defined the common target model with one interconnected European power price based on the Nordic market model.   The European system size is about 1,000 GW and represents about 3,000 TWh production.

The first European example of market coupling was the trilateral coupling between Belgium, France, and the Netherlands in 2006 which developed into the Central West Europe (CWE) market coupling when also Germany and Austria joined the market coupling in 2010. In 2009, the market coupling between the Nordic countries and Germany was introduced.

The day-ahead market has been harmonized with the European market coupling in 2016,

and is today literally one integrated day-ahead market with a common price setting. The intraday markets will be harmonized and market coupled in 2018. Common European rules for harmonization of the balancing markets were agreed at the end of 2017 and will be implemented in the course of the next 4 years with final harmonization of the balancing markets in 2022. For now, mid-2018, market coupling covers most of the western part of the EU and there are plans to cover the remaining part of EU (see figure 39).
The European electricity market over the last 20 years has developed on a step-by-step basis, but with a common target model as the objective. The development has thus been through the combination and incremental development of national and regional solutions.



Day-ahead market coupling status in January 2018

REGIONAL DAY AHEAD IMPLICIT AUCTIONS		
(blue)	North West Europe (NWE)	Price coupling
(blue checkered)	Poland	Poland price coupled within NWE through SwePol-link
(green)	Ireland and Northern Ireland	All Island market, single price zone
(orange)	Czech – Slovak – Hungary-Romania	Price coupling

Source: APX, updated by Matti Supponen

Figure 39: Day-ahead market coupling status, January 2018. Source: APX, updated by Matti Supponen.

During the first years of opening the markets, it became clear that different ways of implementing the electricity market in the Member States would prevent further integration of markets. A need for a target model became evident. The process to agree on a target model was launched in 2006 and it was finalised in 2009 taking in total 3 years to develop and agree upon.

The work on the target model was led by the National Regulatory Authorities, with an active participation of all stakeholders. The main elements of the target model are:

- Forward markets, in which prices linked to day-ahead markets are the main hedging tool for electricity trading.
- Long-term (yearly, monthly) cross border capacity is sold by TSOs through an auction of financial or physical products for hedging the cross-border position of energy traders.
- Cross-border trade optimisation based on day-ahead market coupling in which all remaining cross-border capacity is allocated together with the calculation of prices in each price zone.
- Since the start of 2018 a pan-European intraday platform ”XBID” allows cross-border intraday optimisation using remaining capacity in the grid.
- Final cross-border optimisation through TSOs exchanging balancing energy using remaining capacity in the grid is on the horizon.

2. Balance responsible parties

2.1 Key messages and takeaways

- BRPs act as middlemen reducing the number of interfaces of the TSOs and the power exchange. That is, there is no interface between the TSO or power exchange and the individual producers, retailers or suppliers
- The BRP model allows consumers and smaller producers to contribute to the operational flexibility in the system by giving them access to the market
- BRPs can take advantage of netting of imbalances. The larger and more diverse portfolio of customers the more netting of imbalances.
- The BRP model is based on extensive interface and online communication between actors.
- BRPs are private companies optimizing their profits and the regulatory regime and settlement systems shape the way they act
- The competition between the BRPs drives the development of the forecast and production planning systems and user interfaces
- Regulation of BRPs is necessary, but additional requirements also constitute entry barriers, reducing competition.

2.2 Background and motivation for the BRP role

In a European context, defining the role of the BRPs has been one of the cornerstones of liberalizing the power sector. The BRPs are private intermediaries between the TSO and the power exchange on the one side and the producers and consumers on the other. If not for the BRPs, the TSO and power exchange had to handle a huge number of customers instead of a limited number of professional BRPs. Thus, the BRPs handle the interfaces to producers and consumers and act as their connection to the power markets.

The customers of the BRPs can roughly be divided into two categories: flexible and inflexible.

- Inflexible customers: Portfolios of households (most often via electricity retailers), institutions and small private enterprises, solar power plant and wind turbines . To offer the best possible prices to these customers requires precise forecasts to minimize imbalance costs.
- Flexible customers: CHP plants (some with electrical boilers or large heat pumps) and heavy consumers. These customers often have the best information on their own expected production/consumption schedule, but the BRP offer them day-ahead market (day-ahead market) price forecasts and easy access and advisory on how to optimize utilization of flexibility

Thus, one of the motivations for BRPs is that it is more efficient to have BRPs make forecasts for portfolios of producers and consumers, instead of each producer and consumer having to make forecasts for themselves.

The BRPs charge fees to their customers, for example,

- fixed monthly fee
- fee pr. MWh consumption/production
- balance settlement
  - » If the BRP and the customer makes a ”variable customer balance settlement contract”, the BRP has a net income from the total imbalance settlement as more customers' imbalances to some degree always net out before the BRP settles imbalances with the TSO. This is sometimes referred to as the ”portfolio effect”
  - » If the BRP and the customers make a ”fixed price for imbalances contract” the BRP still have a portfolio effect, but it is uncertain if the imbalance costs encountered by the BRP are less than the fixed price received from the customers

Also, the BRPs gather competencies in understanding the rather complicated electricity markets consisting of day-ahead market, intraday market and ancillary services for which the regulatory framework is often changed. Thus, The BRPs are professional intermediaries or agents for the customers in utilizing opportunities in the electricity markets by facilitating a flexible behaviour in different markets. The competition for customers between the BRPs is extremely important as it forces the BRPs to keep improving forecasts, access to markets and advisory.

2.3 Balancing Responsible Parties (BRP) model

Balance responsible parties buy and sell electricity on behalf of electricity suppliers and consumers. On a daily basis, BRPs send forecasts for how much they expect to be produced and consumed by the producers and consumers they have the balance responsibility for. BRPs are financially responsible for imbalances between expected and actual production and consumption.

The Nordic balance responsible parties can be separated into:

- Balance responsible parties for portfolios of power producers and
- Balance responsible parties for portfolios of power consumers
- Balance responsible parties for power trade

A big power plant or a big consumer can choose to be BRP for themselves if they estimate that the fees paid to an external BRP exceed the costs of establishing and operating the communications systems, etc. themselves.

The focus of the case study will be on production and consumption – i.e. with less focus on trade. The case study will focus broadly on the framework in which the BRPs operate[7].

Requirements to BRPs

In order for a company to become a BRP, it must enter into an agreement with the TSO and fulfil a number of requirements, including:

7 The regulatory setup is not fully aligned in Europe and not even in the Nordic countries. Therefore this report will focus on the overall setup in the Nordic countries omitting some of the details that make the setups a little bit different

- Must be registered for value added tax (VAT) in Denmark or another EU member state or European Economic Area country.
- Must, if necessary, provide adequate security for its obligations, i.e. a credit insurance
- Must have completed a test approved by the TSO of its communication with the TSO
- Must submit the required master data electronically to the TSO.[8]

2.4 Regulatory setup

There has not been established, so far, European consensus about what is the best regulatory setup and thus, this field requires some choices to be made by policymakers. This section describes some of the alternatives with a strongest focus on the Nordic setup.

The description of the setup will include interfaces between BRPs, TSOs the power exchange and the producers and consumers. Moreover, it will include a description of the settlement systems between TSO and BRPs, and between BRPs and customers, impacting the choices and costs of the producers and consumers.

Dispatch and balancing system

The dispatch system is a key part of the regulatory setup as it defines how producers are "ordered to produce". The choice to be made is between central dispatching and self-dispatching modes.

Definition

Self-dispatching model' is a scheduling and dispatching model where the production schedules and consumption schedules as well as dispatching of power generating facilities and demand facilities are determined by the scheduling agents (BRP) of those facilities.
In a 'central dispatch system', the role of the scheduling agent (BRP) is carried out by the TSO.

In the self-dispatching system, the market signals in terms of prices on the day-ahead market, intraday market and ancillary services market give both the BRPs and their customers incentives to utilize their flexibility. When ancillary services are actually needed – that is when an imbalance is detected - the TSO get an active role, and still they only order activation based on bids from the BRPs not looking at what BRP and what plant is behind the bid. As such, in the self-dispatching system, the TSO only reacts to detected but not predicted imbalances.

8 See link for example of master data: http://osp.energinet.dk/SiteCollectionDocuments/Engelske%20dokumenter/El/DataHub%20-%20Regulation%20I.pdf

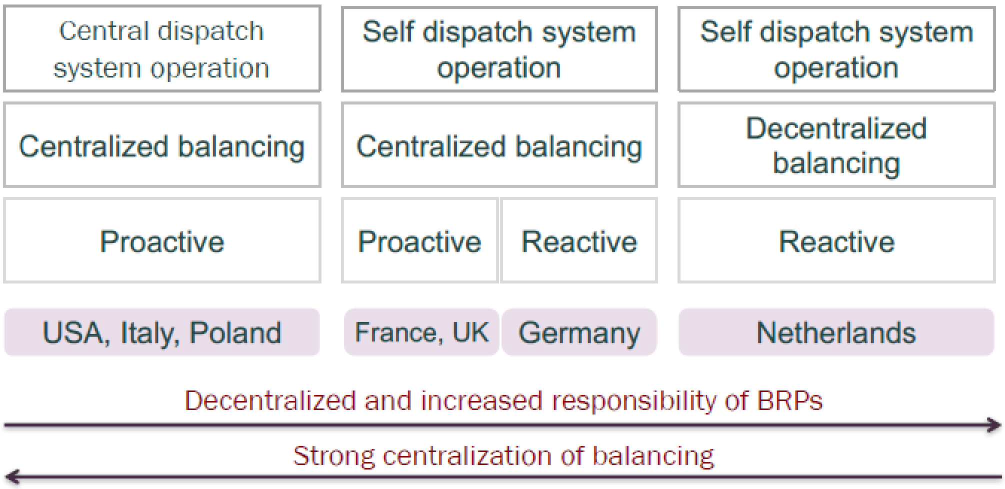


Figure 40: Different national balancing models[9]

Figure 40 shows different market design choices in the USA and some European countries. The further to the right in the figure the more responsibility is left to (private sector) BRPs and the further to the left, the more responsibility to the (most often state-owned) TSOs. When the responsibility and the initiative are left to the BRPs it is important to have an imbalance settlement system that gives the adequate incentives. Therefore this issue elaborated in the next chapters.

In the Nordic countries, BRPs are responsible for generating and sending production and consumption schedules to producers and consumers based on day-ahead market signals from Nord Pool Spot and regulating power signals from the TSOs. Also, the BRPs settle their imbalances with the TSOs making it a decentralized balancing system. The TSOs activate regulating power only when imbalances are detected. Thus, the proactive minimization of imbalances is left to the BRPs who can send new schedules to their portfolio of flexible producers or they can use the intraday market to minimize expected imbalances by trading with counterparties.

As illustrated in figure 40 this implies an increased responsibility of the BRPs in balancing the system. As the BRPs are private companies their overall purpose is profit maximization. Thus, the imbalance settlement system must be designed to give the BRPs financial incentives to act in accordance with the interests of the overall power system.
As the BRPs have more responsibility and financial incentives to act flexibly it is interesting to investigate how these incentives work, especially the link between risks and opportunities. The imbalance settlement system is one of the most important factors in shaping BRPs' behaviour.

However, there are several ways to design the imbalance settlement system. In the Nordic countries cross-border marginal pricing and a combination of both one-price and two-price balance settlement systems has been chosen.

9 ENTSO-E has made this illustration as a part of a roadmap for French electricity balancing, that is without particular focus on Nordic countries. The Nordics would be placed in the right side of the illustration. Source: https://docstore.entsoe.eu/Documents/MC%20documents/balancing_ancillary/2017-12-07/20171207_French_Roadmap_on_Electricity_Balancing_3.pdf

One-price and two-price balance settlement systems[10]

After the operational hour, deviations between the volumes sold by BRP for production (BRPP) and volumes bought by BRP for consumption (BRPC) are measured. These deviations (known as imbalances) are settled with the TSO on a monthly basis.
For BRPC these imbalances are settled under a 'one-price' settlement system. It is referred to as a one price system because regardless of whether a BRPC contributes to an overall system imbalance, or aids in reducing a system imbalance, the regulating power price for that hour is the settlement price for the imbalance. As a result, it is actually possible for BRPC to profit from their imbalances. For example in an hour where the BRPC uses more electricity than planned, but there was an excess of electricity in the system as a whole (thus resulting in down regulating power being activated) the settlement price used to pay for the excess electricity will be less than the spot price. Table 11 below depicts the price used and direction of payment under the one-price system for BRPC.

Electricity System Situation	BRP_C bought too little	BRP_C bought too much
Up-regulation (not enough electricity in the system)	Pay Regulating Power Price (higher than day-ahead market price)	Receive Regulating Power Price (higher than day-ahead market price)*
Down-regulation (excessive electricity in the system)	Pay Regulating Power Price (lower than day-ahead market price)*	Receive Regulating Power Price (lower than day-ahead market price)

Table 11: Settlement prices for the imbalance for a BRPC under a one-price system. *Indicates situations where the BRPC can profit from imbalances. (Adapted from Energinet.dk, 2010)

The intuition behind the profits is, that if the BRPC has bought too power much and the system is in need of power, then the BRPC sell the excess power at a high price because power is scarce.

Almost the same applies if the BRPC has bought little power and the system has too much power, then the BRPC buys the deficit at a low price because there is excess power.
BRPP, on the other hand, is settled under a two-price settlement system, thus not allowing them to profit from their imbalances. If a BRPP imbalance contributes to the system imbalance (a BRPP , for example, produces more electricity than planned in an hour where the system had an excess of electricity), then the settlement price is the regulating power price (in this example a price lower than the day-ahead market price). If however, the BRPP imbalance is in the opposite direction of the system imbalance, then the day-ahead market price is used as the settlement price. Table 12 depicts the settlement situation for a BRPP.

10 This chapter is an slightly edited version of a chapter in: http://www.ea-energianalyse.dk/reports/1027_the_existing_nordic_regulating_power_market.pdf

Electricity System Situation	BRP$_P$ sold too little	BRP$_P$ sold too much
Up-regulation (not enough electricity in the system)	Receive day-ahead market Price	Pay Regulating Power Price (higher than day-ahead market price)
Down-regulation (excessive electricity in the system)	Receive Regulating Power Price (lower than day-ahead market price)	Pay day-ahead market Price

Table 12: Settlement prices for the imbalance for a Production Balance Responsible under a two-price system.

No matter if the imbalance settlement system is one-price or two-price, the imbalances are settled for the net imbalance of the BRPs portfolio of consumption and production, respectively. Therefore it is important for the BRPs to have large portfolios of production or consumption as the imbalances of the individual customers have a tendency to – to some degree – net out. This happens when some customers' imbalances in one direction are partly offset by other customers' imbalance in the other direction. This tendency is stronger the more customers and the more diverse customers the BRP has. Therefore, large BRPs have, when it comes to this cost component, an advantage over small BRPs.

Linking back to figure 40 the one-price settlement system indicates more responsibility to the BRPs than the two-price system as the BRPs are given stronger incentives to take positions that reduces system imbalances. In the two-price system, there is less incentive to do so because in that system a BRP only have an incentive to the BRPs own imbalance. Some would say that the one-price model gives the BRPs incentive to speculate and create imbalances and that is also true. If that is a problem, all comes down to if BRPs are believed to have competencies to do it profitably as a profitable speculation in normal circumstances helps the system.

The concern of the TSO is to always have enough available regulating power to cover the net imbalances of all the BRPs. Thus, they worry about both the volume of available reserves and the BRPs net imbalances. The BRPs, on the other hand, are concerned about maximizing profits and minimizing risks. If there is a scarcity of reasonably priced regulating bids, imbalances tend to become more expensive and profits on imbalances helping the system increases. If, for instance, a BRP suspects scarcity of upregulating bids, then the BRP has a tendency to buy more power in the day-ahead market and intraday market . If there actually is a scarcity of upregulation bids, this maneuver helps the TSO because it reduces the need for upregulating bids. Therefore, if the BRP has the information and competencies to foresee the situation on the market for regulating power, then the behaviour of the BRP helps the TSO. But if the BRP is wrong, and the situation is the opposite from expected, then the BRPs maneuvers increase net imbalance in the wrong direction making a potential scarcity of regulating power worse. Therefore, if policymakers trust the BRPs ability to build competent organizations, then they could very well choose the one-price-model.

The one- vs. two-price model is not the only choice policy makers have to make. Also, the geographical areas in which imbalances are settled plays a role and are important to the utilization of flexibility across price areas borders.

The next chapter elaborates on how the regulating power prices used for imbalance settlement are defined.

Marginal pricing example

Referring to figure 41; say there are two price areas, A and B and there is a lack of power in the system. Therefore, the TSOs orders manual upregulation. For simplicity assume that the lack of power is only due to less than expected wind production for the two BRPPs; BRPA and BRPB. In price area, A BRPA has a deficit of power dA and up-regulation bids A1-A4. In price area, B BRPB has the deficit dB. and up-regulation bids B1-B4.

Individual price area pricing

With individual price area pricing, the balancing prices will be MPA for BRPA and MPB for BRPB. Thus BRPB pays less for its imbalance due to relatively cheap upregulation bids in country B whereas the situation is opposite for BRPA. With individual price area pricing, there are fewer up regulation bids available to the TSOs and there is a risk that BRPs in imbalance will have to pay balancing price equal to expensive bids to settle imbalances. BRPs will pass on the costs to their customers meaning that BRPs in price areas with high upregulation bids (and corresponding low down regulation bids) will have to increase the price they charge their customers for their services.

Cross-border pricing – no congestion

With cross-border pricing, the pool of bids available for regulating imbalances is larger. If there is no congestion (enough free capacity on interconnectors) the TSOs will have the largest pool of upregulation bids available to regulate the imbalance and only the lowest bids (B1-4 + A1) will be activated. This implies that both BRPA and BRPB will pay MPAB to settle their imbalance. Therefore settling the imbalance of BRPB is more expensive as MPAB is higher than MPB. However, on average it is much cheaper to settle the imbalances this way as the TSOs avoid activating the highest bids A2-4.

Cross-border pricing – congestion

With cross-border pricing but limited capacity on interconnector, the B4 bid cannot be used to help regulate the imbalance in price area A. Therefore, the marginal bid used in price area A is A2 and the marginal bid used in price area B is B3 and the result is that MPA* > MPB*. However, the TSOs still avoid activating the very high bids of A3-4 still making the balancing more cost-efficient than with individual price area pricing.

Thus, cross-border pricing on average makes BRPs imbalances cheaper to regulate and with increasing competition between BRPA and BRPB they will on average pass on fewer costs to producers and consumers compared to the case of individual price area pricing.

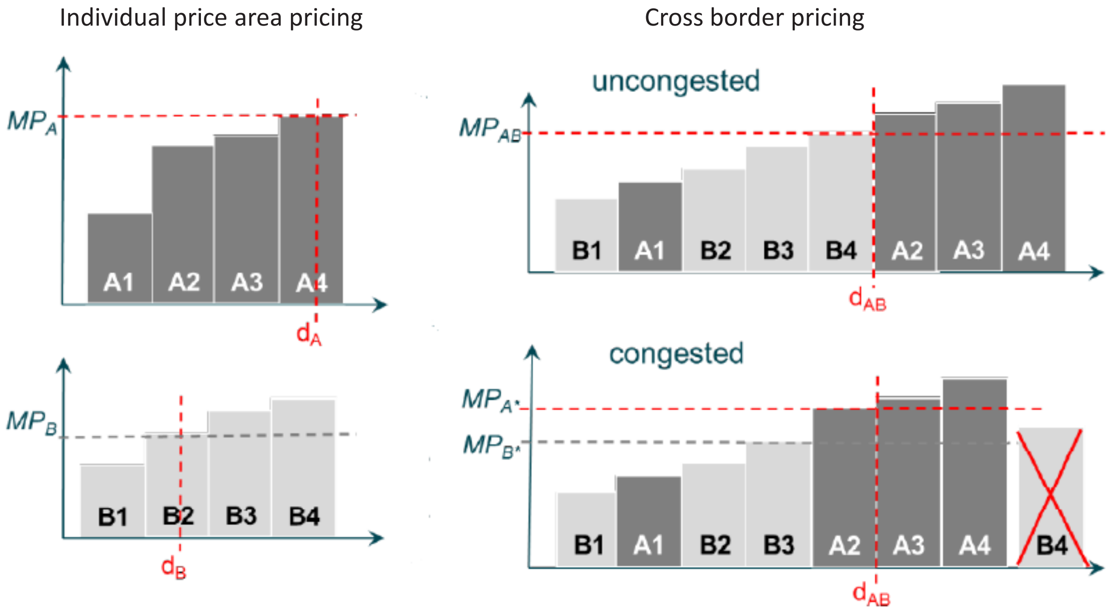


Figure 41: Pricing of imbalances via prices of bids for manual regulating power

Contracts and communication signals

In the common Nordic balancing market, cross-border marginal pricing is applied, meaning that regulating bids go into one pool that the TSOs can choose from as long as there is free capacity on the interconnectors. As figure 42 shows, BRPs are responsible for the proactive balancing; TSOs only start activating regulating power when imbalances are actually detected. Bids and signals for manuel up- and down regulation goes through the BRPs as well as signals with the results of day-ahead and intraday market.

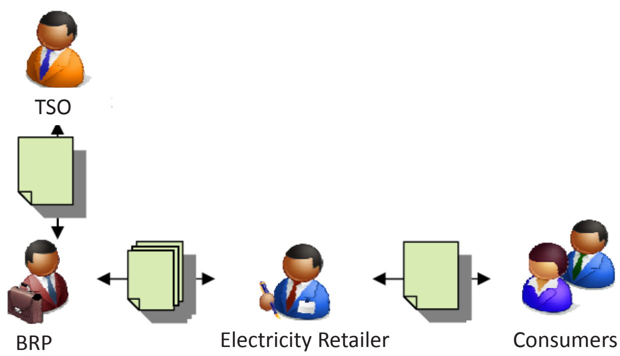


Figure 42: Contractual setup for consumers[11]

Electricity suppliers typically act as middlemen between BRP and consumers, but for very large consumers, such as industries, data centres, coldstores, etc., contracts can be made directly between BRPs and consumers.

11 http://energitilsynet.dk/fileadmin/Filer/0_-_Nyt_site/EL/Tilsynsafgoerelser/2013/Bilag_praecisering_af_balanceansvar.pdf

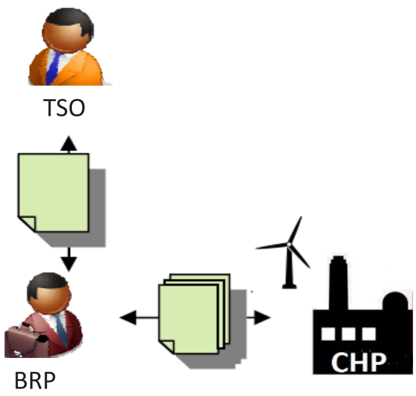


Figure 43: Contractual setup for producers

Producers such as wind turbines owners and CHP stations typically enter directly into contracts with the BRP. If a CHP station has an electric boiler or large heat pump, they typically enter directly into a consumer contract with the same BRP that handles the production.

Like the BRPs have imbalances against the TSO, the wind turbines, CHP stations, electricity retailer (or the sum of consumers he supplies), etc., have imbalances to the BRPs. In most cases, the BRPs make the forecasts of production and consumption. BRPs typically make portfolio forecasts for wind turbines, solar power plants, and small and medium-sized consumers and sell and buy the expected volume on the power exchange. The BRP can also buy forecasts on wind and solar from external providers to improve their performance.

Large consumers (heavy industry, electric boilers, hyperscale datacentres), CHP stations, and Other producers such as hydropower, emergency power producers typically submit individual bids via BRPs to the power exchange and the BRP makes individual production and consumption schedules for these customers. The large consumers, CHP stations etc. have a lot of information about their own operations that the BRP does not have. Therefore, large consumers, CHP stations, and other producers prefer to make their own bids (price and volume) to the power exchange via the BRP.

Types of contracts

Most often a BRP contract has three price components:

- X EUR/month
- Y EUR/MWh
- Z EUR for imbalances
    - » Fixed i.e. the same amount each month, or
    - » Variable dependent on the actual imbalances of the particular customer

The components are independent of each other and each can be negotiated.

The contracts between BRPs and *electricity retailers, wind turbine, and solar power plants* typically imply that the BRP is responsible for making forecasts and submitting bids to the power exchange. However, when actual production or consumption is different from the forecast, the BRP settles imbalances with the TSO and encounters imbalance costs. The fixed price for imbalances passes on the risk to the BRP and therefore a risk-averse BRP might charge a fixed price that on average is higher than the actual imbalance costs.

Therefore, customers often get a better price when they choose a variable balance cost. In this case, it is required that the BRPs publish their variable imbalance costs charged to customers if the customers should be able to compare the performance of the different BRPs. Except for some of the of the wind turbines these customers typically have no flexibility or their capacity is too small to make any effort to utilize flexibility. Therefore, these customers typically focus only on reducing costs. The focus on reducing cost implies that the customers choosing a variable imbalance cost to monitor the performance of the BRP, when they receive the monthly settlement and the customers choosing fixed price on imbalances compare with fixed prices offered by other BRPs.

The contracts between BRPs and *CHP plant* or heavy consumers often implies a variable price on the customers' individual imbalances as these customers have SCADA systems to monitor and adjust for imbalances themselves. When the BRP or the customer detects the imbalance, the customer can adjust for the imbalance themselves (adjust how they operate different units) or they can ask the BRP for a price on trading out the imbalance. Also – based on the advice from the BRP – they can also leave the imbalance and wait for the market to set the imbalance price in the regulating market. If the advice is to wait for the balancing market and the customer's imbalance is in the opposite direction from the market imbalance, the customers' imbalance is "for free". Therefore, these customers typically focus on low price per month and MWh, on a safe and user-friendly software, and good advisory. These parameters are not necessarily straightforward to evaluate and therefore, larger customers often have periodic meetings with the BRP to discuss the performance on the electricity markets and also, the customers can seek independent advice from their industry association. The industry association also arranges courses and conferences to educate their members.

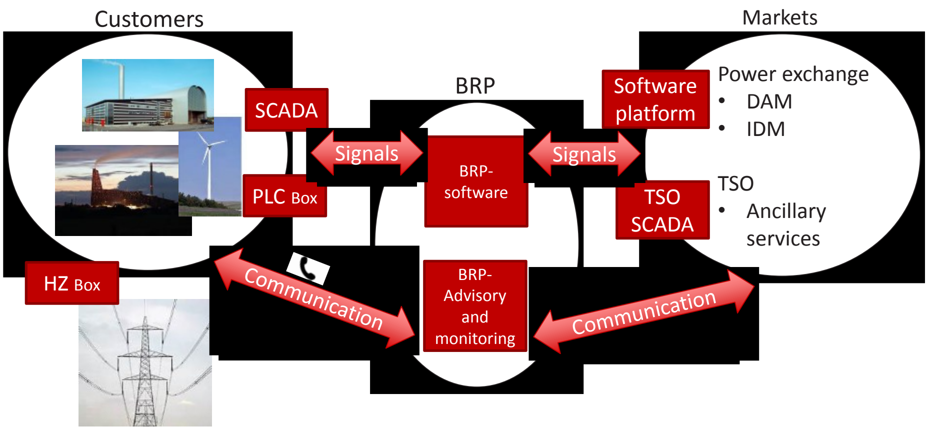


Figure 44: Communication setup for BRPs regarding producers and heavy consumers

Figure 44 illustrates the concept of the BRP model as it works for flexible customers, producers, and heavy consumers. That is in the BRP software or in SCADA software automatically communicating with BRP software, they enter their bids and receive power exchange results in form of production and consumption schedules from the BRP software. The same applies to ancillary services where the BRP sends and receives signals to/from the TSO.

Consequently, a prerequisite for BRPs to be successful is usually to invest in a safe, fast and user-friendly software (interface between BRP-software and customer's SCADA software) and establish a capable 24/7 advisory unit (e.g. on bidding strategies) at a low cost.

Large customer taking on BRP function

If a large consumer or producer can live up to the requirements mentioned in section 2.3 there is nothing to preclude them from being BRP to own production or consumption. In this make-or-buy decision, the following issues are important to consider.

Cons of taking on BRP function:

- Cost of buying or developing IT systems[12]
- Cost of applying BRP tasks to the control room and analysis staff[13]
- Cost of not having BRP-portfolio advantage

Pros of taking on BRP function

- External BRPs fees per MWh and per month can be avoided
- For consumers, the possibility of making profits on imbalances

If the producer or consumer wishes to have total confidentiality about consumption/production schedules, marginal costs etc. taking on the BRP function is a way to achieve this. Another motivation for taking on the BRP function could be to start out with being BRP to own production/consumption and start as BRP for other parties after establishing the required organization.

If a large consumer or producer is not able to negotiate a satisfactory BRP contract it is always an alternative to take on the BRP function themselves as long as they can build the required organization.

Whether or not a large consumer or producer chooses to take on the BRP function themselves, they need to dedicate resources to understanding the electricity market. If they choose to buy full external BRP services, typically significant resources need to be dedicated to understanding and negotiating the contract with the BRP. This is especially true for flexible customers. Therefore, the BRP function can be seen a route to the electricity markets and also a valuable advice on how best to utilize flexibility, nevertheless it requires an effort for the flexible customer to take advantage of the possibilities of the electricity markets.

2.5 Relevance in a Chinese context

The BRP role is a result of a liberalized and decentralized power market, in which system operators delegate part of the balancing responsibility to the market participants. China

12 24/7 staffed control room and staff of analytics are necessary to take on BRP function

13 Consumers/producers with several consumption/production units can – to some extent - benefit from a portfolio effect

is now undergoing a new round of power market reform. One of the goals of this round of market reform is to build a competitive wholesale market. As for the dispatching model and balancing responsibilities, there currently seems to be little momentum toward decentralizing the currently highly centralized model, as this model does provide more control for dispatchers and operators maintaining the stability of the power system. Although the blunt copy of BRP model to China might not be realistic, the BRP model does shed some lights on possible new business models to cope with new trends in China's power sector. The power system is now gradually decentralized as distributed generations (especially distributed PVs) continue to grow at a relatively fast pace. System operators face new challenges of keeping each level of the system balance. Most of the DGs have a much lower observability and controllability compared to traditional power plants, the balancing at the low voltage level of power system will become not only technically challenging but also administratively burdensome. Thus, introducing an intermediary layer between the system operator and DGs to aggregate dispersed DGs would be a possible solution to reduce the workload of system operators and keep the operation of power system streamlined.

3. Intra-day markets

3.1 Key message and takeaways

- The intra-day market allows the market participants to trade any deviations from their scheduled production/consumption and thereby minimise their imbalance costs prior to the operating hour. This facilitates integration of wind and solar production, because the predictability of their production increases significantly when getting closer to real-time.
- Intra-day markets may reduce the imbalances in the operating hour, leading to a more secure system operation, as reserves to a higher degree are available for emergency response in cases of contingencies (e.g. large generator trips).
- The intra-day market offers flexible units a second chance to sell their flexibility, after the day-ahead market trading has concluded.
- As the share of wind power in the Nordic system has increased, the volume traded on the intraday market has increased correspondingly the past few years. Today, the volume traded in intra-day markets is still limited (around 1% of total demand). Data shows that much of the trade takes place very close to the operating hour, when VRE producers have higher certainty about their production.
- The design of the intraday market is continuously evolving to enhance trading and improve flexibility. New auction-based trade and coupling of the European intraday markets are set to increase trading. The settlement period in the Nordic intra-day market will be changed from hourly, as it is today, to 15 minutes in 2020. This will give new possibilities for units with short-term flexibility, and increase flexibility for adjusting for short-term structural imbalances, e.g. in the morning and evening ramps of demand.
- The granularity in the Nordic intra-day market will in 2020 be changed from 1 hour to 15 minutes. This will increase flexibility and can reduce some structural imbalances, e.g. when demand is known to increase/decrease in the morning/evening and give new possibilities for units with short-term flexibility.
- Balancing costs are low in the Nordic system – primarily because balancing reserves are shared through the joint market, enabling the large share of hydropower plants in the North to deliver low cost regulation to the entire Nordic system.

3.2 Intra-day markets – a response to imbalances

The need for intraday markets

In the Nordic countries, and many European electricity markets, the first step to achieving a balance between scheduled supply and demand is delivered through the day-ahead market. The day-ahead market is the anchor market in the European electricity market system. Transmission capacity is allocated for this market and bids compete in a large geographical area – extending from Norway to Portugal. The bids must be delivered to the market operator before 12.00 (noon) the day before the day of operation. This results in 12-36 hours until the operating hour (to the first and the last hour of next day).

As the hour of operation approaches several things can change, e.g. power plant units can fall out, or electricity demand can change because of unexpected changes in temperature.

The intra-day market then gives market participants the opportunity to correct and update schedules before the TSO takes over real-time dispatch and the final balancing during the operating hour. This chapter presents the Nordic intraday market and how it facilitates market actors to balance any deviations from production/consumption schedules. The Nordic intraday market spans four Nordic countries and three Baltic countries. Since June 2018 the Nordic intraday market has been coupled with the European XBID intraday market platform since summer 2018. This chapter will therefore also present the role of increased integration of electricity markets in Europe in facilitating VRE integration through enhancing the power systems' flexibility.

Focus will be on Denmark, as it has some of the regions highest wind shares (close to 40 %) and thus a lot of experience with VRE integration challenges.

Predictability of wind

As variable renewable energy (VRE) penetration increases, an increasingly important aspect is the predictability of these sources. Updated weather forecasts can result in changes to the estimated solar and wind production. While it is still complicated to accurately predict the hourly generation for the next day, significant improvements have been made related to the short-term prediction (1-5 hours), and new methods to improve predictability are being constantly developed. A combination of real-time measurement of electricity generation of wind and solar power and meteorological information can deliver a solid prognosis for the upcoming few hours.

Figure 45 illustrates how forecasts for wind improve as the hour of operation nears. Energinet, the Danish TSO, regularly compares the difference in the settlement data from forecasted production, relative to installed wind capacity. The figure displays the Mean Absolute Error (MAE) as a function of the time horizon up to 40 hours before the hour of operation. There are roughly 5,000 MW of installed wind capacity in Denmark, thus an MAE of 3% one hour before the operation hour corresponds to 150 MW. The figure clearly illustrates the reduction in the prediction error as the hour of operation nears, and thus provides a good example of the benefits associated with trading closer to the hour of operation, particularly as it relates to facilitating the improved integration of VRE power. Allowing for a trade closer to the hour of operation has value for all market participants as it provides them with additional time to trade out of any potential imbalance situations. As shown by the strong trend for increased error in forecasting with increasing time lag in figure 45, this is particularly the case for wind producers.

It is worth mentioning, that forecast uncertainty is highly dependent on the expected wind speeds. The uncertainty is largest at medium wind speeds, and the higher the expected wind speed, the more likely that generation will be less than expected (negative deviation). The story is similar when the geographical scope of the trading area is expanded. All market participants benefit from a larger trading area (i.e. if a traditional power generator in one price area has difficulties, expanding the number of price areas it can trade with increases the probability that it can trade out of its potential imbalance at a lower cost), but VRE producers in particular benefit. For wind, this relates to how wind fronts roll in, and how electricity generation from wind parks correlate with each other depending on the distance between the generators. Figure 46 illustrates this for 6 different time horizon averages (i.e. wind power production summed over the previous 5 minutes, up to 12 hours). If the 1-hour average line is taken as an example, there is a relatively high correlation between generators within 20 km, but the correlation is greatly reduced when the distance is over 100 km, and when it is over 500 km there is almost no correlation between generators output. When there is a surplus of wind in one area, it is likely that there is less wind in an area some 100 km away.

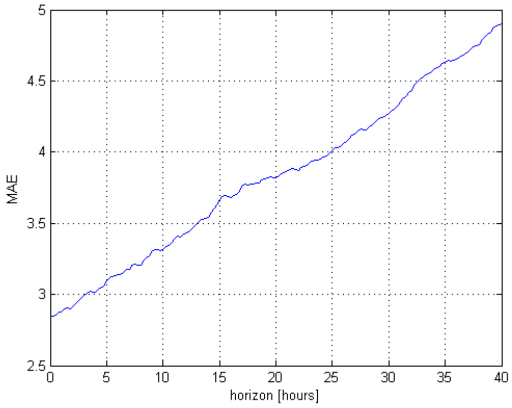


Figure 45: Absolute error (Mean Absolute Error, MAE) for Energinet.dk's Danish wind power forecasts in % relative to total installed capacity over a 40-hour time horizon[14].

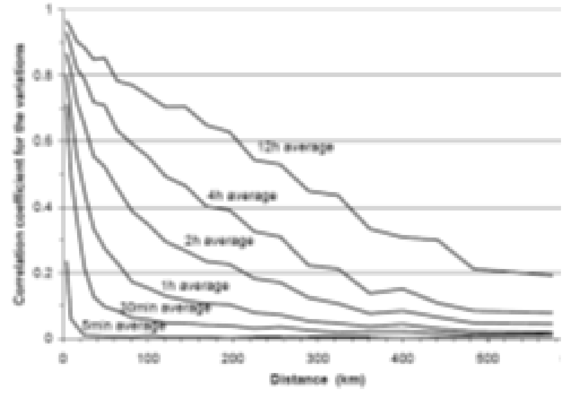


Figure 46: The smoothing of the correlation that takes place for wind power over distances[15].

The stochastic nature of imbalances costs

The key argument for using the intra-day market is to avoid penalties for having imbalances. If actual generation deviates from the final schedules submitted to the TSO when the intraday market closes, any real-time deviations are managed by the TSO and the costs incurred are distributed on the individual balance responsible parties according to their imbalance. The TSO thus has to activate regulating power (manual Frequency Restoration Reserve, m-FRR[16]) to mitigate the total imbalance.

The penalty, i.e. the cost of being in imbalance, varies and can be difficult to predict. On average, the amount of hours where the Danish TSO activates upward regulation, down-

14 Energinet.dk and the Danish Energy System - Hami project", Energinet.dk, 2016

15 "IEA Wind", IEA,

16 As defined by the European Network of Transmission System Operators for electricity ENTSO-E

regulation, or no regulation at all, are equal (each activated 1/3 of the time). For a balance responsible party (BRP), this means that for 1/3 of all hours there are no penalties for the imbalances and the imbalances are traded at the spot price. The penalty is usually highest when there is need for upward regulation, as more expensive generators are forced to ramp up.

The size of the penalty depends on the volume of the total imbalance, the availability of transmission capacity and the capacity balance in the market. As the Nordic market is interconnected, and a lot of the upregulation capacity is provided by hydropower located in Northern Norway and Sweden, transmission capacity becomes a key factor for the price that BRPs have to pay for the imbalance (see also chapter 2 Balance responsible Parties). Activation is coordinated with neighbouring TSOs, as in some cases positive imbalances can be netted with negative imbalances between neighbouring balancing areas.

Balancing responsible parties can thus choose whether to trade in the intraday market and hand in a balanced portfolio to the TSO, or whether to pay for their imbalances being settled by the TSO in real time. Whether the cost of being in imbalance will be high or not can be difficult to predict at the time when the balancing responsible party has to deliver the final schedule to the TSO. Imbalances from changes in weather predictions (temperature or wind speeds) may be evaluated by all balance responsible parties. However, contingencies such as failures in power plants, are first known to the power plant owner but hard to predict for other market participants. The intra-day market on the other hand can deliver a known cost of correcting any deviations from the generation or demand schedule.

3.3 The Nordic intra-day market: Nord Pool Elbas

The Nordic intra-day market, known as Elbas, was introduced in 1999 in Sweden and Finland, and today covers all the Nordic and Baltic countries, as well as Germany. Eastern Denmark joined in 2004, Western Denmark in 2007, Norway joined in 2009, and Latvia and Lithuania in 2013.

Elbas overview

The Elbas market allows market participants to trade power until up to one hour before the operating hour (real-time). Thereby market participants can reduce their expected imbalances, i.e. the deviation between scheduled and real-time production, and flexible units get a second chance to sell their flexibility, after the day-ahead market trading has concluded.

The trade system functions as a stock market. Trade takes place when bid and ask prices from two traders match. The result is a pay-as-bid system. This is in contrast to the day-ahead market where bids are brought together in daily auctions with pay-as-clear, also known as marginal pricing. Trade on Elbas is anonymous, with the Nord Pool platform serving as the middleman. As with any markets, the potential trade depends on the available transmission capacity. With available capacity, trade can take place over large areas (e.g. between a seller in Demark and a buyer in Finland). When there is no available capacity (in the relevant direction) trade cannot take place.

Imbalances can thus be moved between price areas for market participants that are active in several areas. Imbalances may cancel out, or the market participants may expect the cost of imbalances to be smaller in another area.

Due to uncertainty regarding e.g. imbalance costs, and price formation in the intra-day market, there is no single optimal strategy for pricing bids in the intra-day market. In some hours the imbalance costs are low, and the benefit of trading in the intra-day market may be low. Therefore, bidding in the intra-day market must be based on expectations about imbalance prices. Expectation can be built, among other factors, on historical information, the status of major power plants and transmission lines, and meteorological information about changes in weather systems. See (Scharff, Egerer, & Söder, 2014) for a discussion of the decision-making in the different markets.

Since imbalance costs are different for generation and demand, market participants may also use the intra-day market to move imbalances between the two systems to reduce imbalance costs (see chapter 2.3 Balance Responsible Parties for more information on pricing of imbalances).

Trade volume on the intra-day market

The traded volume on the intra-day market is much smaller that the trade on the day-ahead market. In 2017, Nord Pool traded 394 TWh in the day-ahead market (Nordic and Baltic) – and 6.7 TWh in intra-day market (Nordic, Baltic and German), i.e. less than 2% of volume in the day-ahead market. However, the volume is continuously increasing (see figure 47).

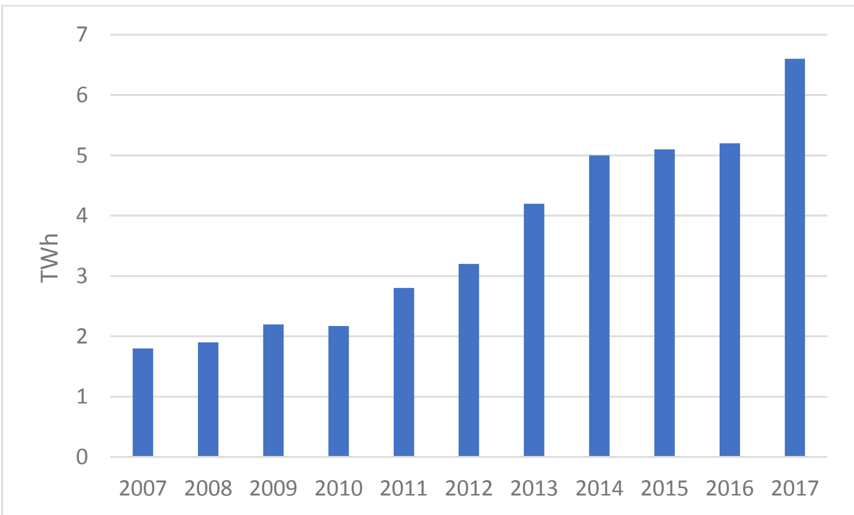


Figure 47: Traded volume on Nord Pool intra-day market[17].

The share of trading varies significantly among the Nordic countries. The relative share traded in Denmark is more than 10 times higher than in Norway (see figure 48). An explanation could be that Denmark has high shares of wind power, and thus a large need for balancing energy that can even out fluctuations in wind production. At the same time, Denmark has fewer resources with high flexibility compared to the other Nordic countries, and therefore usually buys balancing services from Norwegian or Swedish hydro power, which are typically highly competitive in the intraday market.

The low potential gain of undertaking a trade, , due to relatively low costs of imbalance, coupled with the complex bidding strategy, can be part of the reason for the relatively low trade on the intra-day market. An analysis of the value of considering both the day-ahead

17 "2017 Annual Report – Investing for the Customer", Nord Pool, 2018.

and the intra-day markets for a Norwegian hydropower plant found that the additional value of considering the intra-day market is less than 0.65%. The analysis included a stochastic description of the price, the future price formation and three cascading hydro plants as an example[18].

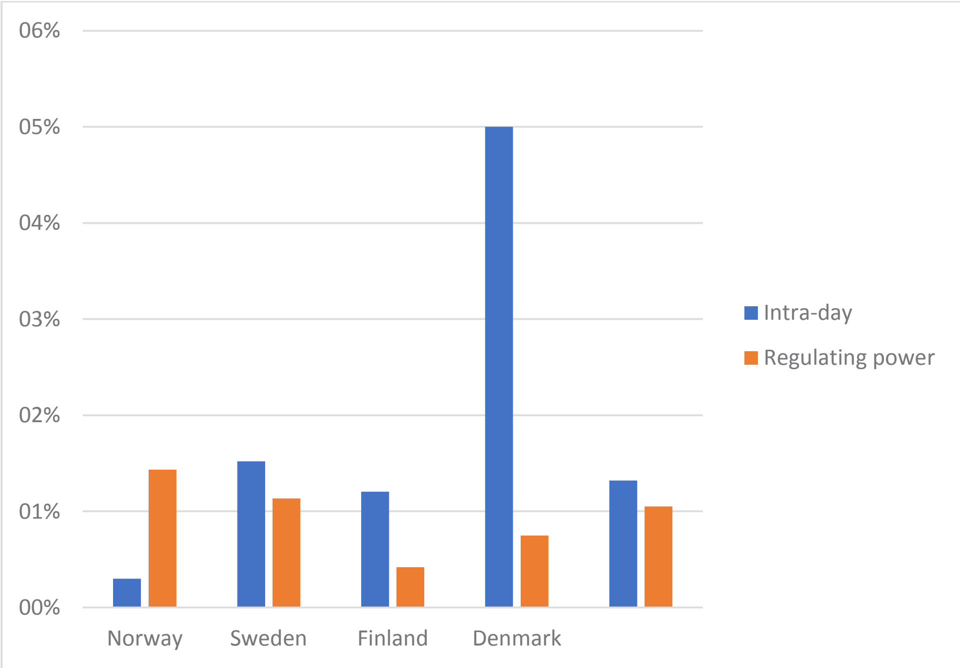


Figure 48: The trade on the intra-day market and the activation of regulating power, % of yearly demand. Nord pool, 2017.

As the intra-day market can be seen as a measure to reduce the amount of regulating power activated by the TSOs, it can be relevant to compare the two. In figure 48 it can be seen that for the Nordic countries the volume of the two markets are similar, and that the volume of intra-day and regulating power for the whole Nordic area both are around 1% of demand. The volume of regulating power has been quite stable the last five years, while the intra-day volume has increased. More wind power has been added to the system during this period, and this indicates that the additional imbalances introduced by the system have been compensated by the use of the intra-day market.

Intra-day auctions

The bilateral nature of the trade results in a large number of trades and varying prices. The European regulator has reasoned that a higher degree of transparency could be achieved if an auction system was developed. Nord Pool is now offering two daily auctions for Germany: an auction at 22:00 for each of the hours of the next day, and an auction at 10:00 for the last 12 hours of that day. These auctions are similar to the auctions used in the day-head market. The larger volume of bids is expected to give more competition and make it more attractive to use the intra-day market. The auction system is a supplement to the traditional setup.

So far, the adoption rate has been very slow. No trade was recorded in any of the daily

18 “Day-ahead market bidding for a Nordic hydropower producer: taking the Elbas market into account” in Computational Management Science, Faria, Eduardo; Fleten, Stein-Eri, 2009.

auctions in July of 2018, and only a few trades took place in June of 2018. This is likely because market participants prefer to trade as late as possible (figure 49), and therefore use the traditional bilateral system. This illustrates how difficult it can be to design market systems. There are good arguments for adding daily auctions to the intra-day market, however, if it is not used by market participants, it has little value.

Timing of trade on the intra-day market

Trade on Elbas can start at 15:00 the day before the operating day, and in the Nordic system trade ends one hour before the operating hour. Based on the timeline for bidding in the day-ahead market, the first hour of the operating day is 9 hours away from the deadline for submitting bids (see chapter 1.2 The Nordic short-term electricity markets). For the last hours of the operating day the deadline is 33 hours away. Figure 49 and figure 50 illustrate that more trade takes place in the intra-day market in the hours that are far from the day-ahead deadline, as for these hours the uncertainty about their final balance is particularly high. 1/3 of the trade takes place in the last trading hour (between one and two hours before the operating hour) (Nord Pool, 2017). This highlights the importance for balancing responsible parties to be able to adjust their scheduled production and demand closer to real-time, as with increasing amounts of VRE, predicting production and demand 33 hours ahead in time becomes increasingly inaccurate.

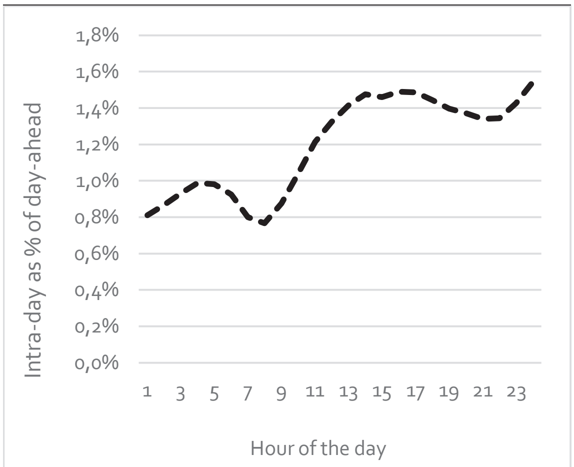


Figure 49: Trade on intra-day as % of trade on day-ahead (2014-2018).

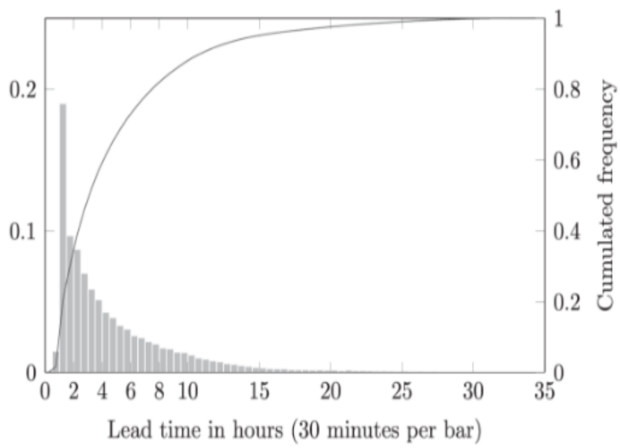


Figure 50: Leadtime for trades on Elbas from March 2012 to February 2013. (Scharff & Amelin, Trading behaviour on the continuous intraday market Elbas, 2016)

Cross-border trade on the intra-day market

In 2017, 75% of the intra-day trade was cross-border through the Elbas intraday market. Bilateral trade between two parties is possible, however only within one price area and not cross-border.

Figure 51 provides a detailed picture of trade on the Nordic intra-day market. For most price areas trade is dominated by a trade to/from other price areas. The two exceptions are Germany and Finland. For these two countries, the majority of trade is within the country. This is likely to manage internal bottlenecks in the transmission grid. The two countries only have one price area, so domestic bottleneck cannot be managed in the day-ahead market. It is also clear from the figure that Norway has relatively little trade. An interview with the Norwegian generator Statkraft indicated that the current low activity in Elbas is because many hydro-based generators can adjust imbalances internally by use of other hydro plants and that the additional flexibility is sold directly to the TSO as regulating power. Statkraft expects that the use of Elbas will increase in the future – as exporting flexibility from their hydro power plants to other countries becomes more and more attractive.

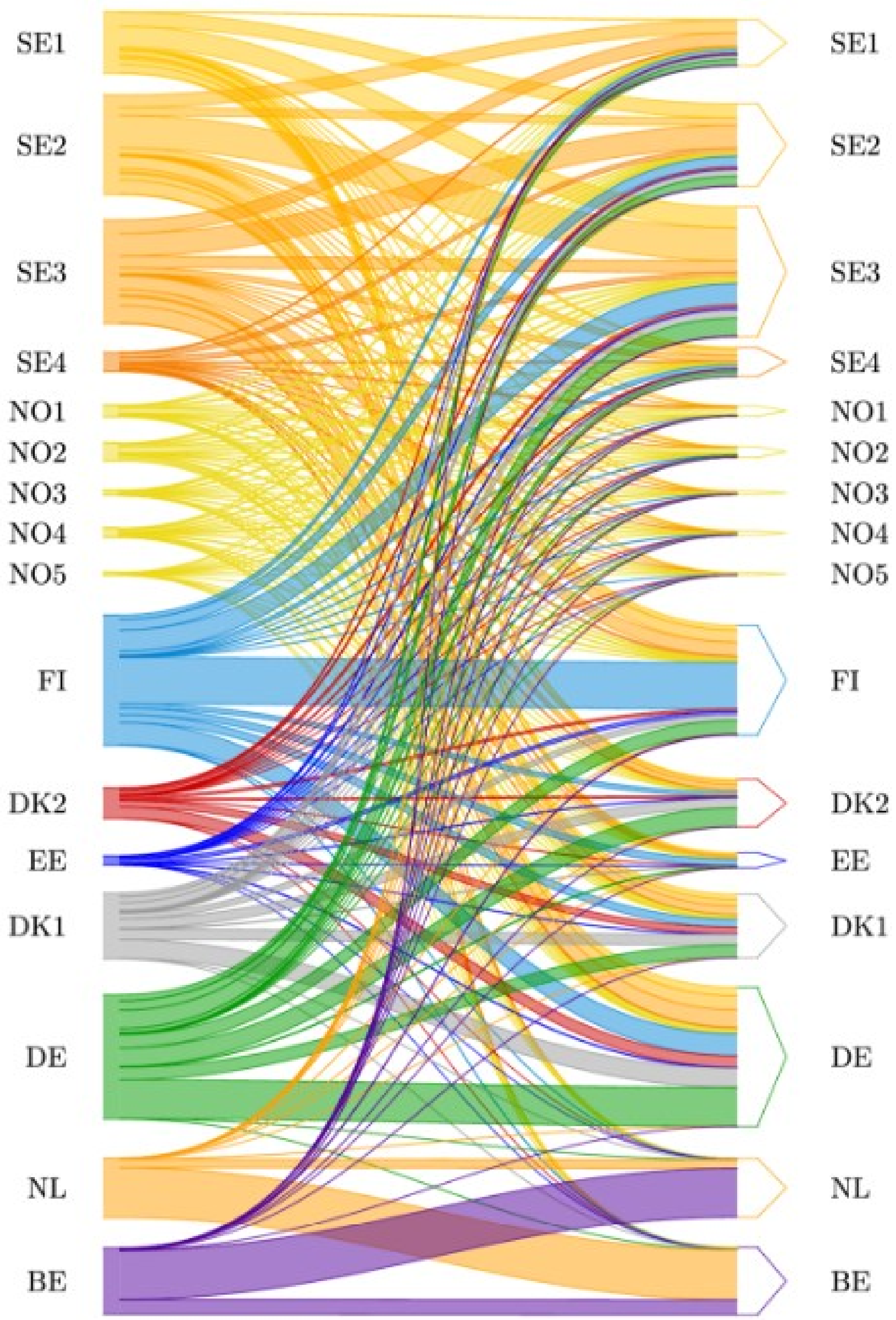


Figure 51: Trade on Nord Pool Intra-day market Elbas from March 2012 to February 2013. Left side is selling, right side buying[19,20].

19 Trading behaviour on the continuous intraday market Elbas" in Energy Policy, Scharff, Richard; Amelin, Mikael, 2016)

20 SE = Sweden, NO = Norway, DK = Denmark, FI = Finland, DE = Germany, NL = Netherlands, BE = Belgium, EE = Estonia. When a country name is followed by a number it refers to different price areas.

Future Development

Going forward, the plan is to reduce the settlement period in the Nordic system from 1 hour to 15 minutes. This is expected to make it easier for demand resources to participate – e.g. more industrial processes can be disconnected for 15 minutes than for 1 hour. In addition, some structural imbalances can be avoided with the finer granularity, e.g. during morning/evening when demand is increasing/decreasing. In the first phase of this transition, the 15-minutes time step will be introduced in the intra-day market – and in the long-term potentially also in the day-ahead market. For the intra-day market, the 15-minute settlement is expected to be introduced by 2020[21].

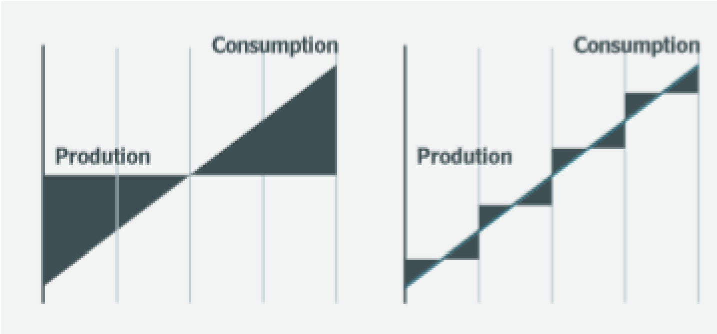


Figure 52: Simplified illustration of how structural imbalances can be reduced by using a finer granularity, e.g. 15 minutes instead of hourly resolution[22].

3.4 Emergence of XBID

The trend towards further integration of markets, enabling sharing of reserves and netting of imbalances over wider geographical areas, continues. A significant milestone towards the long-term vision of a single integrated European intraday market was reached on June 13th of 2018 with the first trading day for XBID, a joint initiative by the power exchanges EPEX SPOT, GME, Nord Pool and OMIE and the North Western and South Western European and Baltic TSOs. In this first step, XBID extends intraday trading to cover 14 countries, reaching from Portugal in the Southwest of Europe, to Finland in the Northeast (figure 53). The majority of remaining European countries are expected to join in 2019.

With respect to XBID, ENTSO-E, the European Network of Transmission System Operators, which represents 43 electricity transmission system operators from 36 countries across Europe stated that: "With the rising share of intermittent generation in the European generation mix, connecting intraday markets through cross-border trading is an increasingly important tool for market parties to keep positions balanced. As the intraday market develops it will enable increased optimisation of the use of generation - especially variable Renewable Energy Sources – and will also enable demand response products to develop. It will also lead to welfare benefits. The purpose of the XBID initiative is to increase the overall efficiency of intraday trading." [23]

21 The way forward - Solutions for a changing Nordic power system", Nordic TSOs, 2018

22 Ibid.

23 European Cross-Border Intraday (XBID) Solution and 10 Local Implementation Projects successful go-live", ENTSOE, 2018, https://www.entsoe.eu/news/2018/06/14/european-cross-border-intraday-xbid-solution-and-10-local-implementation-projects-successful-go-live/

Figure 53: The bidding zone borders of the countries (marked in orange) that were coupled together with the June 2018 launch of EXBID[24].

XBID overview

The June 2018 introduction involved 14 countries, reaching from Portugal in the Southwest of Europe, to Finland in the Northeast (figure 53).

The core components of XBID are a:

- Common IT system with one Shared Order Book (SOB)
- Capacity Management Module (CMM)
- Shipping Module (SM)

XBID involves collecting all orders submitted by market participants in each of the participating countries in one centralised order book referred to as the Shared Order Book. Meanwhile, information regarding the available intraday cross-border capacities are provided by the various TSOs in the Capacity Management Module. As long as transmission capacity is available, orders from market participants in one bidding zone can be met by orders from market participants in other bidding zones.

Trades are executed on a first-come-first-serve basis, wherein the highest offered bid to buy is matched with the lowest offer to sell. After a trade is made, both the Shared Order Book and Capacity Management Module are immediately updated, which for the order book essentially means removing the matched bids, while the updated Capacity Management Module now reflects the new transmission capacities on the affected lines. This process of continuously matching buyer and seller bids, while simultaneously updating the Shared Order Book and Capacity Management Module continues for the entire XBID trading period.

Lastly, information regarding completed trades is provided to all relevant parties via what is referred to as the Shipping Module. The Shared order book provides this Shipping Module with data regarding both trades made between two different price areas, and also trades undertaken within a price area but between two different Exchange Members[25].

It is interesting to note that the intraday market for these various countries is interlinked, despite having varying product availabilities. For example, Germany and Austria have

24 “XBID Launch Information Package”, Nordpoolgroup.com, 2018

25 Cross-Border Intraday: Questions & Answers”, Epexspot, 2018

15-minute products, France also has 30-minute products, while Portugal and Spain are the only countries without User Defined Blocks (table 13). 15-minute continuous trading will be introduced in Belgium and in The Netherlands starting on July 10th, 2018. However, more countries are expected to adopt shorter products, as the clear trend, also supported by European legislation, is to trade closer and closer to the delivery time and in shorter time intervals.

	Germany	Austria	France	Nordic, Baltic, Netherlands and Belgium	Spain and Portugal
15 min	X	X			
30 min	X		X		
Hourly	X	X	X	X	X
Block bids	X	X	X	X	

Table 13: Product availability in the different market areas (Nordpoolgroup.com, 2018)

Trade on XBID is listed on the Nord Pool market information. Figure 54 reveals that the introduction of XBID in June 2018 has not yet lead to the expected increase in traded volume on Elbas. However, it should be noted that it takes time for market actors to determine whether it is cost-effective to enter new markets (particularly if this involves updating software, etc.), so it may take some time before the launch of XBID is reflected in increased trade volumes.

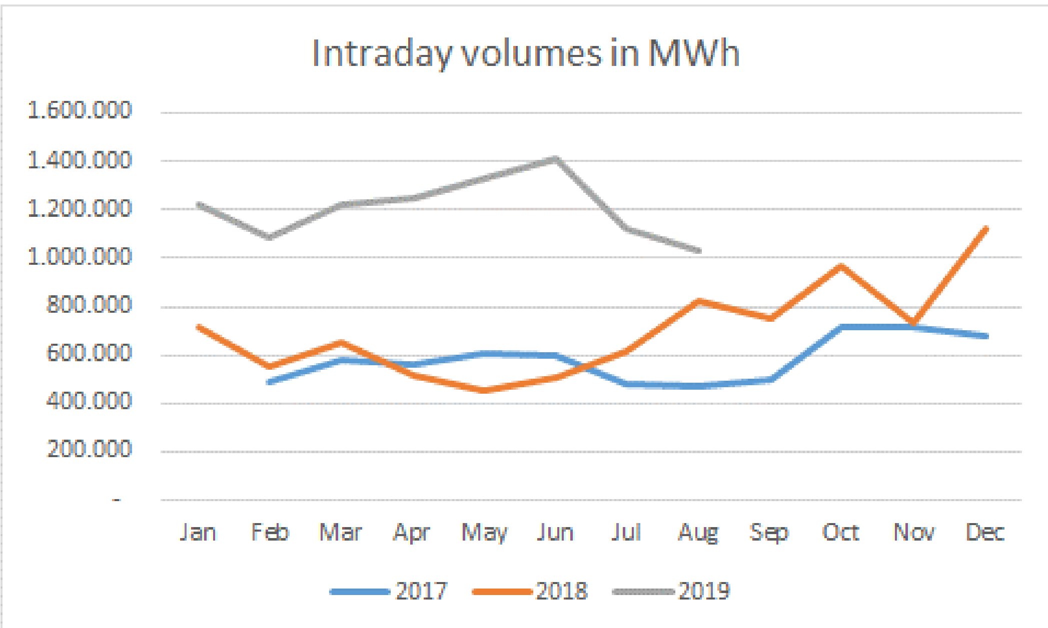


Figure 54: Trade on Elbas (1.1.2017-2.8.2018). Source: Nord Pool.

3.5 Socioeconomic and market participant perspectives

From a socioeconomic perspective, the elimination of imbalances via the intraday market prior to the hour of operation, rather than relying on more expensive reserves activated by the TSO during the hour of operation results in lower overall system costs. For a study of the socio-economic benefit of market integration see Booz & Company, David Newbery, Goran Strbac, Danny Pudjianto, Pierre Noël, LeighFisher, 2013[26]. Here the total benefit of European integration of electricity markets is estimated to be €2.5–4 bn per year, or about €5 to €8 per capita per year.

From a TSO perspective, allowing market participants to minimise their imbalances prior to the hour of operation does not impact the TSOs financially, as imbalance costs are passed on to the market participant responsible for the imbalance. As the primary function of the TSO is to maintain overall system security, reducing imbalances prior to the hour of operation is of interest to TSOs. Large imbalances may result in a less secure operation, as fewer reserves may be available when a large failure takes place. Reducing imbalances by use of the intra-day market may, therefore, increase the level of secure operation. Data from the Nordic system indicates that the volume of activated regulating power has not increased during the last five years – even after expansion of the wind power generation. This may be partly because of the increased use of the intra-day market.

VRE producers have an economic incentive to participate in the intraday market as it potentially provides a lower cost option to reduce likely imbalances during the upcoming hour of operation. The greater the liquidity of the market, and the broader the geographical scope, the higher the probability that the VRE producers will be able to find a counterparty that can improve their balance position at a lower cost than would occur via the implementation of reserves during the hour of operation.

3.6 Relevance in a Chinese context

The intrinsic uncertainties of weather renders the forecast error of wind and solar power inevitable. The direct outcome from these forecast errors is the extra need of reserves on different time scales. This also eventually contributes to the so-called system cost of integration of variable renewable energy. Another fact of wind and solar forecast is that its accuracy of looking only a few hours ahead (usually referred to as ultra-short-term-forecast), compared to looking 24 hours ahead, is largely satisfactory. Thus, to reduce the need of reserves and the cost accordingly, the adjustment to the day-ahead schedule is necessary. As the fast reserves (direct curtailment of wind and solar, activating diesel engines or batteries, etc.) are the most expensive ones, the adjustment is better to be done 3~4 hours before the physical delivery. Intraday market in Europe plays a significant role for this type of adjustment. It should be noted that, even the two-settlement model (day-ahead + real-time) widely used in the U.S. does not contains an intra-day market explicitly, system operators still have various intra-day mechanisms, such as hourly residual unit commit in ERCOT and short-term unit commitment in CAISO. The advantage of an explicit intra-day market is that the invisible hand could be more cost effective in terms of re-scheduling the system and guide the investment on flexibility assets through providing indications of the value of flexibility.

China has launched several pilot projects on spot market. Currently, the focus is mainly on the design of day-ahead market and real-time market/dispatch. The intra-day mechanisms are seldomly discussed. For provinces with high variable renewable penetration, the workload of intraday adjustment would be quite high, even might be too much for a

26 Benefits of an integrated European energy market", Booz & Company, David Newbery, Goran Strbac, Danny Pudjianto, Pierre Noël, LeighFisher, 2013

compact system operator unit (dispatching centre). In these scenarios, one of the option is to introduce intraday trading, which allows power producers to gradually correct their generation programs. Moreover, compared to the real-time market, the intra-day market leaves a relatively longer time slot for power plants to adjust the generation. Due to their large inertia, thermal power plants are not fast enough to respond to the real-time market. Intraday would leave a space for flexibility from traditional thermal power plants. As many argue that a spot market could not provide sufficient incentives for investment in flexibility for thermal power plants in China, including an intraday market in the spot market could provide a necessary extra push.

4. Flexibility from hydropower plants with a reservoir

4.1 Key messages and takeaways

- Nordic hydro production has a profound effect on both annual day-ahead electricity prices, but also balancing costs.
- In particular, hydro power plays an important role in balancing wind power production.
- There is four times more hydro capacity than wind power capacity in the Nordic area today.
- In the future, installed capacities of wind and solar power will increase significantly, while hydropower capacity is expected to only incur limited growth, as the best resource locations are already developed.
- Based on a wide-ranging source of inputs, hydro power operators utilise complex models in an attempt to optimise hydro production and maximise the value of their flexibility in various electricity and ancillary service markets.

4.2 Background

Reservoir-based hydropower can act as electricity storage, thus helping the balancing of electricity demand and generation in the power system. Due to merit order-based pricing systems, when variable renewable energy (VRE) such as wind or solar power are generating electricity at very low marginal cost, other sources will typically reduce their generation. The sites featured in this chapter do not involve pumping, but solely adjustment of generation from the hydro plants. The water saved by not generating in hours with low electricity prices can be stored for hours, day, weeks or months – without incurring significant losses. The primary losses in relation to utilising hydro as storage are the transmission losses related to the extra flow in the transmission system between the wind power area and the hydropower area, e.g. in the DC lines from Denmark to Norway. However, these losses relate to the system as a whole, and are not borne by the generator. Additional – but small – potential losses that are felt by the generator, relate to evaporation of water from the reservoir.

For the operators of hydro facilities without pumping the most important considerations are how best to utilise their annual water resources, i.e. if electricity prices are slightly above average today, there is a potential opportunity cost of utilising water today that cannot be utilised in the future if even higher prices are realised, whereas there is also a risk that higher prices may not come, and the operator will then have forgone some profits. A number of the factors that play into this optimisation will be discussed.

4.3 Nordic power system

The Nordic electricity system has extensive amounts of hydropower, as 48% of generation capacity is hydro-based, and 56% of annual generation comes from hydro. The active volume of the reservoirs is equivalent to 90 TWh, corresponding to 41% of the yearly hydro generation, or 23% of the total yearly generation (see table 14). Nearly half of the European reservoir capacity is located in Norway.

Hydro plants are very flexible and a large amount of hydro in the Nordic system is the major reason for low regulating power prices, and therefore low imbalance costs in the Nordics. This is a clear benefit of wind and solar power.

	Capacity (MW)	Generation (GWh)
Hydro	50,166	220,300
Fossil fuel	20,526	30,586
Wind	13,789	33,366
Nuclear	11,858	82,824
Other	7,010	25,366
Solar	854	744
Total	**104,203**	**393,186**

Table 14: Installed capacity and generation in the Nordic countries (2016).

The reservoir levels of the Nordic hydro system have a clear seasonal variation (figure 55), with a large inflow in spring and early summer (roughly speaking, weeks 16-32). In comparison, it can be noted that it would require an entire month of full wind power generation to produce a sizable impact on the storage level (i.e. 10,000 GWh). The hydro plants with reservoir can always absorb short-term variations (e.g. hours), and only in very rare situations when the reservoir is nearly full or empty is the short-term flexibility limited. For practical purposes, storage is not restricted by the reservoir size but is instead restricted by the limited transmission capacity form wind production areas to the hydro sites. E.g. there is over 30,000 MW of hydro capacity in Norway, but only 6,205 MW transmission capacity to/from Norway, hereof 3,800 MW of alternating current (AC) connection to Sweden and 2,355 MW DC connection to Denmark and the Netherlands[27]-

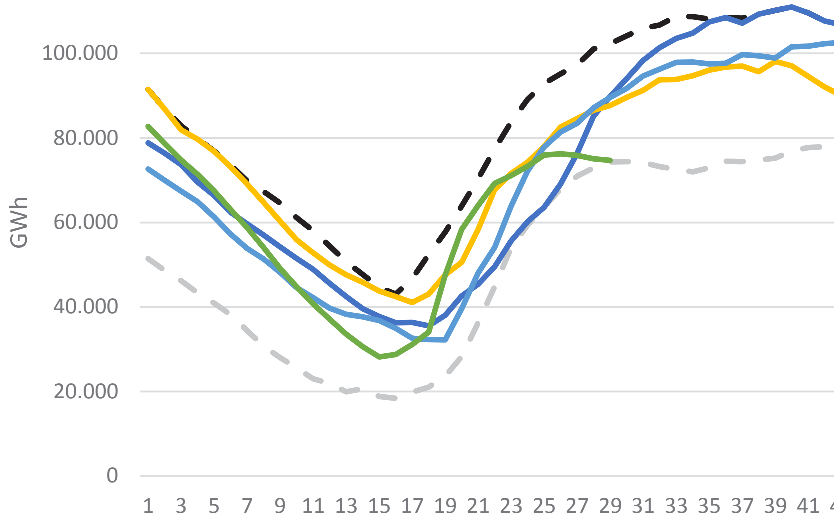


Figure 55: Hydro reservoir levels in the Nordic countries per week. Maximum and minimum are based on

27 Capacities are from Nord Pool: www.nordpoolspot.com/globalassets/download-center/tso/max-ntc.pdf. Average value of import/export capacity has been used. Hydro capacity from ENTSO-E: www.entsoe.eu/data/power-stats/inventory-of-generation/

data from week 1, 2001, to week 31, 2018. Nord Pool.

The total level of the reservoirs can go from minimum to maximum in a relatively short time (what determines the minimum and maximum will be discussed later). In 2015, there was a record low level in week 25, but already by week 37, the level was at the maximum (figure 55). From week 21 to week 29 in 2018 the opposite development took place, i.e. from a high level to a minimum level over the course of just 2 months. This dynamic makes it complicated to predict electricity prices – and for the hydro plant to plan its operation.

The reservoir level in the Nordic hydro system is one of the most important parameters that influence the electricity price. When reservoir levels are high, hydro plants will seek to generate more, e.g. to avoid spillage of water, and vice versa when the levels are low (i.e. save water for more profitable hours). The seasonal effect of snow and rainfall on electricity prices is so significant that those that work with electricity markets refer to years with large amounts of rain/snow as ‘wet’ years, while the opposite is ‘dry’ years.

When reviewing long-term electricity price series, it is often quite clear when these extreme years have taken place. The extent to which Norwegian, and to a lesser extent Swedish, hydro production affects electricity prices in the entire Nordic area (and beyond) is illustrated in figure 56, which displays historic power prices for Western Denmark. Despite having essentially no hydro capacity in Denmark, the figure clearly illustrates the large impact that water levels in the Nordics have on power prices throughout the region.

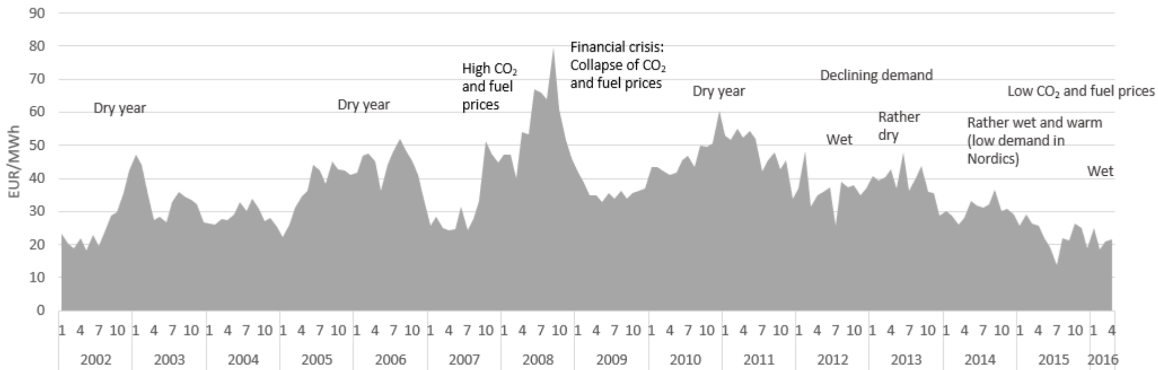


Figure 56: Historical average power prices in Western Denmark from 2002 to 2016 (nominal EUR/MWh).

4.4 Optimal dispatch of hydropower – a simplified setup

In order to determine the optimal dispatch of hydropower, a simplified method would involve assuming full foresight[28]. Electricity demand per hour, fuel prices, generation capacities, hydro inflow and wind and solar profiles are therefore known. In this simplified world there is one optimal way to utilise the hydropower. Each hydro plant would have a threshold value – called the water value – and each plant should generate when the electricity price is above this value. Within the boundary of minimising the overall system cost, in this way the value for each plant owner is maximised.

The water value changes over time and depends on the price area that the plants are located in, on the reservoir size, the inflow of water, etc. In this system, the coordination between wind and hydro is realised through a merit order-based price mechanism. When

28 As is the case in a number of power sector models such as the Balmorel/EDO model.

there is a lot of wind power (which has a very low marginal cost) the electricity price will be relatively low and water in the reservoir would be left for later use.
A simple example: Assume that a hydro plant owner expects that it can generate at full capacity half of the time (based on the expected inflow), and that that (all) future prices will be the same as the last month. Then the water value can be found as the median of the last month's prices. Generation should, therefore, take place during all hours with an electricity price higher than this price. Generally speaking, this would, for example, mean producing when there is little wind in the Nordics and holding back hydro production when wind production is high.

4.5 Real-life aspects

A long list of aspects will make real-life operation of hydro plants more complex than in the simple example above. There will be uncertainty about many aspects related to future inflow and electricity prices, including:

- Future water inflow to the reservoir, which is dependent on aspects including future rainfall and temperature (i.e. due to the melting of snow)
- Future electricity prices, which are dependent on a number of factors, such as:
  - » Future electricity demand
  - » Future generation capacities (outage on plants and transmission lines).
  - » Generation from wind and solar power

The plant owner would still use a water value to guide the operation, but this would need to be updated frequently, e.g. daily or weekly as new information becomes available. The computed water values thus are based on the long term forecast of inflow and electricity prices and govern the daily dispatch from the individual plant. The water values could also reflect strategic decisions (i.e. an aggressive or conservative bidding strategy).

Many other aspects may be relevant for the practical operation of hydro plants. Some plants have restrictions on:

- Minimum water flows in the downstream river.
- Maximum flow down-stream (e.g. relevant in case of overflow)
- Minimum reservoir level
- Maximum change per day in reservoir levels

These restrictions exist in most of the Norwegian sites and are meant to protect fish, the stability of the river banks as well as recreational aspects. These restrictions must be respected in the computation of water values. In some cases, the plant would need to generate a certain minimum to respect minimum downstream water flow. In this way, only a part of the generation is controllable.

Some hydro plans are located on the same river as other hydro plants. In this case, the plants are using the water in what is referred to as a cascading fashion. The water available for down-stream plants is dependent on the generation at up-stream plants. If the lower reservoir capacity is small, the generation of the lower plants will be dictated by the generation of the upper plant. If the reservoir is sufficiently large, the impact of the cascading system can be ignored in the short-term operation.

With respect to maintaining the reservoir levels, it is important to avoid spoilage of water. Spillage represents a waste of the resource but can in extreme cases be necessary due to a need to maintain the security of the day-ahead market or elements downstream.

An empirical study comprising eight years of electricity generation from Norwegian hydro with storage showed that if the market price increases with 1%, the generation increases with 17 GWh/week (100 MW)[29]. Other parameters, such as temperature (which influences electricity demand) and the hydrological balance also influence the generation.

Case: Agder Energi

In a normal hydro year, Agder Energi generates 8.1 TWh/year from 39 hydropower plants located on five rivers in southern Norway. The hydro plants have a total reservoir capacity of 5.3 TWh. The largest reservoir (the largest I Norway) can contain water corresponding to three years of inflow.

Many of the hydro plants are cascading plants located on the same river, and the largest reservoirs are located high in the mountains. These mountain-based reservoirs also typically have the greatest height (resulting in largest generation per m3). The plants downstream have smaller reservoirs but can still contain water corresponding to production for weeks or month. Nevertheless, the hydro units on the five rivers are interconnected and the use of water from the upper reservoir will also determine generation at the lower plants. Only on a short time scale, e.g. within a week, can the units be considered as individual units.

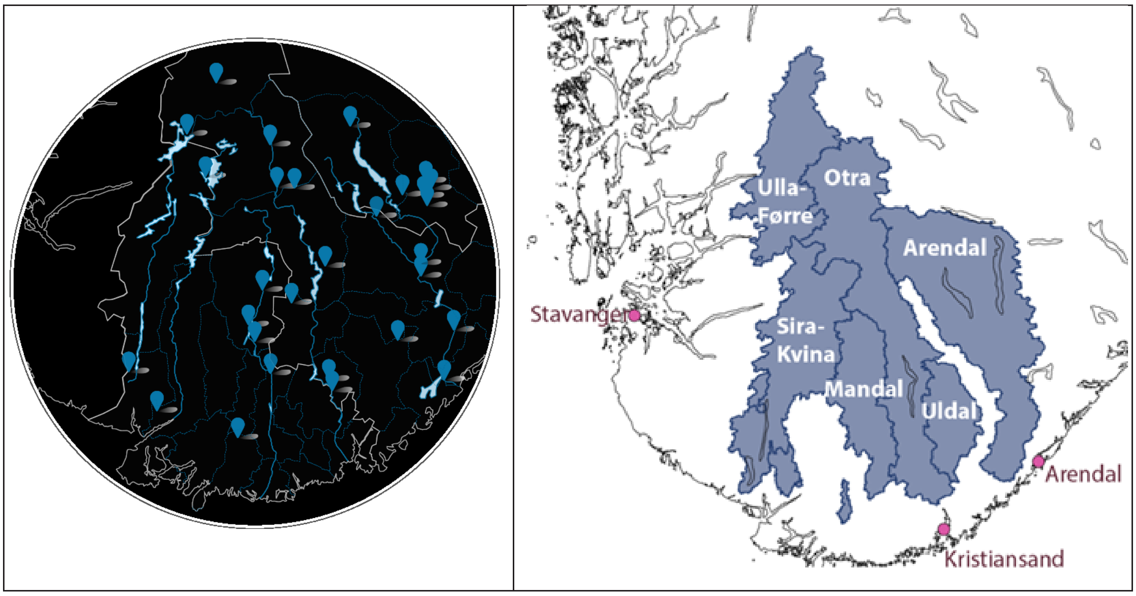


Figure 57: The catchment area (right) and the location of the Agder Energi hydro plants.

Reflections from Kristian Wiik Ravnaas, Agder Energy, Senior Portfolio Manager:

> *We compute water values for each individual hydro plant in our system. We use many sources of information and update the water values as new information becomes available (such as temperature, rain, and coal and natural gas prices). Temperature affects both the melting of snow and electricity demand. The water*

29 As is the case in a number of power sector models such as the Balmorel/EDO model.

values are typically re-evaluated every day. For the first two upcoming weeks of power price predictions, the information regarding expected temperature and inflow is important. Beyond this time horizon, average historical values are often used – in combination with the historical distribution of wet and dry situations.

The electricity price predictions are becoming more complicated due to the expansion of wind and solar power. Maintenance of the hydro plants typically takes place during summer, when both electricity demand and electricity prices are low. However, with increasing shares of wind power and interconnectors, low and high prices is expected to occur at any time during the year, which is challenging for the maintenance planning in cascading river courses.

The hot and calm summer of 2018 has been extreme. There has been little wind, and high temperatures have increased evaporation from the hydro day-ahead markets.

Case: Statkraft

Statkraft is owned by the Norwegian state and is Norway's largest, and the Nordic region's second-largest, power producer. Statkraft has 353 power plants with a total installed capacity of 19,080 MW. This includes 12,899 MW hydro in Norway, and 1,267 MW hydro in Sweden.

Reflections Bjørn Rasmus Dah, VP Spot sales, Energy Management, Statkraft:

We are working with long-term price prognoses. Several of our reservoirs are so large that water could be stored for more than one year, meaning production planning is a long-term decision. We have to evaluate the situation many months ahead to manage the water correctly, and we have to plan this management based on many uncertain factors. A risk of flooding must be weighed up against the risk of emptying the reservoirs prematurely, restrictions must be complied to, and we should have water left if the prices rise. Based on the price prognoses and updated fuel prices, reservoir levels and weather prognosis, water values are calculated each day for each plant. To make the detailed operational plan we use the SHOP model. This program is run several timers per day (see box below for the models utilised).

Actual reservoir levels, weather forecasts, fuel prices, import and export capacities as well as wind generation in Sweden and Denmark are important for the formation of prices in Norway.

The flexibility of our hydro is primarily activated in the day-head market and as regulating power. It can also be used in the intra-day market. The use of the intra-day market has traditionally been low, but the volume traded in Norway is increasing. For a generator with hydro plants, it is often possible to correct imbalances within our own system, e.g. with other hydro plants. In the future, it is likely that Norwegian hydro will increase the supply of flexibility to other countries with high shares of wind and solar.

Examples of models that Statkraft uses in its hydro operational planning

EMPS[30] - multi-area power-market simulator - is a tool for forecasting and planning in electricity markets. It has been developed for optimisation and simulation of hydrothermal power systems with a considerable share of hydropower. It takes into account transmission constraints and hydrological differences between major areas or regional subsystems. The objective is to minimise the expected cost in the entire system subject to all constraints. In principle, this solution will coincide with the outcome in a well-functioning electricity market. The simulated system can e.g. be the Nordic system or Northern Europe. The basic time step in the EMPS model is one week, with a horizon of up to ten years. Within each week, the time-resolution is 1 hour or longer. In the strategy evaluation, incremental water values (marginal costs for hydropower) are computed for each area using stochastic dynamic programming. A heuristic approach is used to treat the interaction between areas. In the simulation part of the model total system costs are minimized week by week for each climate scenario (e.g. 1931 – 2012) in a linear problem formulation.

SHOP[31] (Short-term Hydro Operation Planning) is a program for short-term hydro operation planning. The program is based on an optimisation formulation where complex hydraulic configurations of water courses can be modelled. The program is able to handle any number of cascaded watercourses. The optimisation is based on successive linear programming and may include mixed integer formulation. The general objective of the program is to utilise the available resources, and to maximise the profit within the period in consideration, by exploiting the options for buying and selling in the spot market, while fulfilling firm load obligations. The study period is flexible, typically 7 to 14 days. The time resolution is typically 1 hour or 15 minutes.

30 "EMPS - multi area power-market simulator", SINTEF, 2018, www.sintef.no/en/software/emps-multi-area-power-market-simulator

31 "Short-term Hydro Operation Planning", SINTEF, 2018, www.sintef.no/en/software/shop

4.6 Additional sources

For more details regarding Nordic and European hydro capacity (including pumped hydro capacity) see Euroelectric; VREB Powertech, 2018[32]. In addition, numerous publicly available sources provide historical hydro data[33]. For a review of studies regarding European wind and solar production balanced with Nordic hydropower, see Ingeborg Graabak, 2016 .

4.7 Relevance in a Chinese context

The deregulation of the power market in Nordic regions also led to the deregulation of hydro power stations. In this deregulated and liberalized market context, each of the power stations makes decisions on their own. This is in sharp contrast with the traditional centralized paradigm, in which all the generation programs are determined by one central dispatching centre. Although all the decisions are individually made, they collectively leads to a very efficient power system, as evidenced by the high level utilisation of renewable energy in this region.

Unlike wind and solar power, the marginal cost of hydro power plant with a reservoir is not zero, it equals to the opportunity cost of generating now compared to withholding the water for later generation. One of the key concepts in the decision making processes is water value, which reflects the timely value of a unit of water in the reservoir. In a nutshell, the water value would be higher when there is less water in the reservoir, or a higher price expected in the future, and vice versa. When the current market price is higher than the water value calculated, the power stations will generate. Thus the revenue of a hydro power station is largely depends on it accuracy of all kinds of prognoses: price prognoses, weather prognoses, etc.

Practices in Nordic regions shed some lights on the on-going power market reform in Southeast provinces, such as Yunnan and Sichuan, which also have abundant hydro power and increasingly more wind power plants. One of the learnings for the hydro power station owners is that the capability building on price prognoses on various time scales (hours ahead, days ahead, months ahead, even a year ahead) would be crucial for their financial performance.

32 Facts of Hydropower in the EU", Euroelectric; VGB Powertech, 2018.

33 Examples of sources include:

- The Norwegian Water Resources and Energy Directorate, NVE, (www.nve.no/hydrology) publish real time data and weekly data about hydrology. This includes reservoir level and information about the volume of snow on the mountains (in water equivalents).
- Nord Pool publishes weekly data about reservoir levels in the Nordic system (www.nordpoolgroup.com/historical-market-data).
- Data for European hydro can also be obtained from ENTSO-E's Transparency platform (transparency.entsoe.eu).

5. Demand-side management in Finland

5.1 Key messages and takeaways

- Demand-side management (DSM) can serve as a cost-effective flexibility resource in all the different power markets (spot, intra-day, balancing and particularly ancillary services)
- It is necessary to undertake contractual and technical requirement updates in order to allow demand side to participate in existing markets
- Stakeholder consultations and pilot projects are important for the development of DSM
- The use of aggregators can facilitate the introduction of small units in the market
- Participating in DSM allows individual stakeholders to realise value for their flexibility

5.2 Background

Finnish electricity system

As was outlined in the Nordic and European power market overview, Finland is part of the Nordic synchronous grid and Northern Finland has two 400 kV AC connections with Sweden, and a 220 kV AC connection to Norway. There is also a total of 1.2 GW of DC interconnector capacity between Rauma in Southwestern Finland and Dannebo and Finnböle in Sweden (price area SE3), as well as 1 GW of DC capacity between Southern Finland and Estonia. Lastly, Finland is connected to Russia via three 400 kV lines to Vyborg, and two other 110 kV lines[34].

In 2017, Finland's electricity demand was met by 4 nearly equal shares of electricity imports, nuclear generation, non-thermal generation (hydro, wind and solar), and thermal generation (roughly ½ of which was biomass, ¼ coal, and the remainder primarily comprising natural gas and peat).

	Capacity (GW)	Generation (TWh)
Hydro	3.2	14.6
Wind	2.0	4.8
Solar	0.0	0.0
Nuclear	2.8	21.6
Thermal	9.4	24.0
Net imports		20.4
Total	**17.5**	**85.5**

Table 15: Installed capacity, generation, net imports and for Finland (2017 year-end preliminary data).[35]

Demand-side management

Demand-side management (DSM) has garnered increasing attention in recent years largely due to the increasing levels of fluctuating renewable energy production, which therefore

34 Nordic power system and interconnections with other systems", Fingrid, 2018, https://www.fingrid.fi/en/grid/electricity-system-of-finland/nordic-power-system-and-interconnections-with-other-systems/

35 Statistics Finland's PX-Web databases", Statistics Finland, 2018, http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin__ene__ehk/?tablelist=true

require sources of system flexibility. DSM is one of those sources, and involves reducing electricity demand when the system is in short supply of electricity (which, in the Nordic system, is reflected via high electricity and/or ancillary service prices) and increasing electricity consumption when the system has abundant electricity (i.e. when electricity prices are lower). According to the the International Energy Agency (IEA), "Demand response potential typically amounts to around 15% of peak demand."[36] Widespread cost-effective DSM development requires communication and data management, which have seen large advancements and cost reductions in recent years, thus also contributing to an increased focus on DSM.

There are essentially four main types of demand response:

- Shifting of electricity demand from one period of time to another. Prime examples include water treatment plants or industry with a number of pumping processes that do not need to run continually and can, therefore, run during times with low electricity prices, or heating and cooling elements that can adjust their temperature setpoints according to electricity prices and thereby provide additional cooling/heating during lower price periods and less during high price periods.
- Fuel shift. This applies to end-users that can rely on more than one fuel input in order to satisfy their heat and/or electricity needs. An industrial customer may, for example, have both an electric boiler and a natural gas-based boiler to produce process steam and can switch between the two depending on the electricity and natural gas prices.
- Utilisation of backup systems. A number of public, commercial, and industrial sites (e.g. hospitals, airports, etc.) have backup generators in case of situations when electricity supply is interrupted. These units can act as system reserves and also provide demand response.
- Peak shaving/valley filling. Peak shaving refers to reducing electricity demand when prices are high, without using the corresponding amount of electricity at another point in time, while valley filling is the opposite. This last form of demand response is only relevant in extreme scarcity situations.

5.3 DSM from a system and market actor perspective

Fingrid's reserve obligations

The Finish TSO is one of the world leaders in the utilisation of Demand-Side Management. Fingrid's reserve obligations and contributions thereto from DSM are outlined in the table below.

36 Re-Powering Markets", Chapter 6, p. 154, IEA, 2016

Reserve product	Procurement channel	Maximum contracted capacity	Obligation	From DSM
Frequency Containment Reserve for Normal Operation (FCR-N)	- Yearly market - Hourly market - Other Nordic countries - Vyborg DC link - Estonia, Estlink 1 & 2	55.0 MW 134.7 MW - 90 MW 35 MW	Roughly 140 MW	4 MW
Frequency Containment Reserve for Disturbances (FCR-D)	- Yearly market - Hourly market - Other Nordic countries	455.7 MW 471.4 MW -	220–265 MW	430 MW
Automatic Frequency Restoration Reserve (aFRR)	- Hourly market - Sweden	- -	70 MW (certain morning and evening hours only)	0
Manual Frequency Restoration Reserve (mFRR)	- Balancing energy and balancing capacity markets - Fingrid's reserve power plants - Leasing reserve power plants	- 935 MW 262 MW	880–1100 MW[205]	100-300 MW

Table 16: Fingrid's reserve obligations and contributions from DSM[37,38].

Day-ahead and intraday markets[39]

In addition to DSM participation in the reserve markets, Fingrid estimates[40] that 200-600 MW participate in the day-ahead markets, while 0-200 MW participate in the intraday markets. To put these figures into perspective, from 2014 to 2017, the amount of electricity purchased in Finland on the day-ahead market averaged 57 TWh, while the amount sold averaged 42 TWh, thereby reflecting the large annual electricity imports. During this same period, annual trade on the intra-day market averaged 800 GWh, with the total purchase and sale volumes being quite similar.

37 Demand Side Management", Fingrid, 2018, https://www.fingrid.fi/en/electricity-market/demand-side-management/

38 Reserves and Balancing Power", Fingrid, 2018, https://www.fingrid.fi/en/electricity-market/reserves_and_balancing/#reserve-products

39 As chapter 1 outlined the relevant electricity and ancillary services markets in the Nordic countries, which included a description of the day ahead and intraday markets, and the various ancillary services, only brief descriptions of the specific reserves are included in the text box below.

40 The day-ahead and intraday markets are run by Nord Pool Spot, and as such the ranges provided are estimates provided by Fingrid experts

Fingrid's reserve products

Frequency Containment Reserve for Normal Operation (FCR-N)
FCR-N is an automatic reserve that is used to maintain the frequency at 50 Hz. When the frequency deviates by more than +/- 0.1 Hz, then these automatic reserves must deliver up/or down-regulation. The minimum bid size for this reserve is 0.1 MW. Providers of this reserve receive both a capacity payment and energy payment, however, as it is a symmetric product, the net energy bought/sold is quite small.

Frequency Containment Reserve for Disturbances (FCR-D)
FCR-D is a reserve that is acquired by TSOs to ensure that the power system can withstand a significant disconnection in the grid (for example a large generation unit). FCR-D reserve providers must deliver up-regulation when the frequency falls below 49.5 Hz. The minimum bid size in Finland is 1 MW, and it is an asymmetrical product, i.e. only up-regulation is required. FCR-D providers only receive a capacity payment.

Automatic Frequency Restoration Reserve (aFRR)
According to Fingrid, aFRR is a centralised automatically activated reserve also designed to maintain a frequency of 50 Hz., but it is only purchased for particular morning and evening hours if given advance notice.

Manual Frequency Restoration Reserve (mFRR)
mFRR, also referred to as regulating power, is a manual reserve that is activated in order to restore the availability of the fast-acting automatic reserves. The minimum bid size in Finland is currently 10 MW (5 MW if bids are activated electronically) and payment is for energy only according to the regulating power price for that hour.

DSM background/development in Finland

In many countries and regions, growing levels of variable renewable energy in have recent years increased the need for system flexibility. In Finland this problem was exacerbated by the fact that a) the new VRE was pushing the historic providers of this flexibility (i.e. traditional power plants) out of the market, and b) a new large nuclear unit was soon to come online and is anticipated to run close to 100% load, thereby providing little flexibility[41].

As a result, roughly 5 years ago, Fingrid, facing rising reserve costs and growing needs for flexibility, decided to investigate and promote DSM through the launch of a special project. This involved consultations, meetings, and various forms of communication targeting a broad and large number of stakeholders. Finland already had more than 15 years of experience with large-scale industry participation in reserve markets, as wood, chemical and metal industries delivered both FCR-D and regulating power. The goal was now to

41 "DSR in Finland" Interview with Jäppinen, Jonne, Fingrid, August 2018.

expand this participation to small, medium and even household end-users[42].

Fingrid determined that the aim was to not have separate reserve markets for demand, but instead make alterations to the existing markets so that all reserve markets in Finland allowed for demand participation. This required undertaking changes in a number of areas, for example the wording of contracts had to be updated (i.e. an FCR-D supplier was no longer referred to as a power plant), and technical requirements had to be updated so they also reflected the technical characteristic of demand units. As one of Fingrid's DSM experts explained: "The requirements had to be fair to all participants, but the demand side also needed a little push"[43].

An example of a necessary technical adjustment was for the FCR-D reserve product, where additional disconnection and reconnection steps are outlined. For example, on Fingrid's website it stipulates that relay-connected resources can deliver FCR-D as power plants do (50% of effect after 5 seconds and 100% after 30 seconds after a frequency step change of -0.50 Hz, or via immediate disconnection when the frequency has been at 49.7 Hz or less for 5 seconds, at 49.6 Hz or less for 3 seconds, or 49.6 Hz or less for 1 second. Under this 2nd option, reconnection of the load is allowed when the frequency has been at least 49.9 Hz for 3 minutes[44].

Most relevant products provided by DSM today

FCR-D

The most relevant product for DSM is FCR-D, and Fingrid typically requires 260 MW of this fast-acting frequency reserve. The reason that it is so well-suited to demand is that it is relatively easy to set up and implement, i.e. essentially only a frequency meter is required, and perhaps most importantly, it is activated quite rarely. For example, Fingrid notes that FCR-D is generally only activated a few times per year (perhaps up to a few times per month), and it is not necessarily the entire load that is activated, because depending on how large the frequency drop is, it may only be necessary to activate a portion of the total.

According to Fingrid, a perfect example of the supply growth from DSM sources in recent years for FCR-D is greenhouses.[45] Greenhouses require a lot of light, but if this light is turned off occasionally for a short period of time it is unlikely to affect the quality of the tomatoes for example, and therefore this form of electricity demand is perfectly suited to providing FCR-D. In fact, the growth in DSM supply potential for FCR-D as of January of 2018 was roughly 430 MW. This now exceeds Fingrid's typical hourly requirement, and over 70% of the FCR-D procured by Fingrid in 2018 came from DSM.

FCR-N

As was highlighted in Table 16, a rough estimate of the required FCR-N is 140 MW, with this coming from an annual market, hourly market, DC interconnectors and other Nordic countries. As of January 2018, roughly 4 MW was supplied by DSM. There are a number of reasons that the DSM contribution is currently much lower for FCR-N. Firstly, because FCR-N

42 Ibid.

43 Ibid.

44 "Reserves and Balancing Power", Fingrid, 2018, https://www.fingrid.fi/en/electricity-market/reserves_and_balancing/#reserve-products

45 "DSR in Finland" Interview with Jäppinen, Jonne, Fingrid, August 2018.

is a synchronous product, electricity demand must also be regulated up on short notice (whereas it only has to be decreased for FCR-D), which is often a difficult task for demand. Secondly, FCR-N has a requirement that the regulation must be delivered in a linear fashion, which is also currently challenging for most demand sources as they are better suited for immediate disconnection (which as mentioned above is permitted for FCR-D). Lastly, the large number of activations (i.e. numerous times within an hour) for this product are less-suited for many forms of demand. Taking the greenhouse as an example, having the lights constantly turning on and off would unlikely to be acceptable.

Regulating power

Fingrid purchases mFRR both in the form 800-1,100 MW of fast disturbance reserves (i.e. reserves that are only used for disturbance purposes), and 'traditional' regulating power from the Nordic regulating power market. The vast majority (roughly 1,000 MW) of the fast disturbance reserves are provided by gas turbine units. According to Fingrid, of the roughly 100-300 MW of DSR capacity that can deliver mFRR, the bulk of this comes from traditional industrial sources such as wood, chemical, and metal industries. These facilities undertake load shifting, utilise fuel shift (some have CHP plants on site), while some also bid in with their back up capacity.

Effects of DSM participation

Given the interconnected and complex nature of electrical systems, it is difficult to say with certainty what the direct effects of implementing DSM have been, however, it is interesting to reflect on the development in price and quantities of the most relevant reserves. Fingrid procures both FCR-N and FCR-D via international transmission links, and annual and hourly domestic markets. The quantities and prices from both the annual and hourly domestic auctions are publicly available.

FCR-N and FCR-D domestic auction process

The annual FCR auctions are held each autumn for the upcoming year. For each of the two products, suppliers bid in with their capacity and price, and the price of the highest accepted bid becomes the fixed hourly price for all accepted bids for the entire year. Each day during the course of the year, winning suppliers submit plans prior to 18:00 the day before the day of operation outlining how much of their awarded reserve capacity will be available the following day. I.e. an operator that was awarded 10 MW of capacity in the annual auctions and informs the TSO the day prior to the operation that it has 6 MW available for the following day will receive the annual hourly payment for 6 MW for each hour of that day.

Based on these inputs, and updated prognoses, Fingrid determines how much additional reserve shall be purchased in the hourly markets. Bids for hours in the following 24-hour period must be submitted by 18:30. A reserve provider that was awarded an annual contract may also submit bids, but only if all of its annually awarded capacity was first made available by 18:00. The price for each hour is the most expensive winning bid for the respective hour.

FCR-D

Figure 58 below displays the maximum awarded annual capacities and corresponding prices for FCR-D in Finland from 2014-2018 (light blue), along with the average hourly quantities and corresponding prices (orange) from the hourly markets. It is clear from the figure that the maximum quantity of capacity purchased on the annual market has been increasing since 2011, while the corresponding price in the annual markets grew steadily from 2011 to 2017, before falling significantly in 2018. While it is too early to establish whether this is a new trend or a one-year anomaly, given the large increases of DSM capacity that have entered the FCR-D market during the past two years it is reasonable to assume that this new supply has contributed to the fall in 2018 prices.

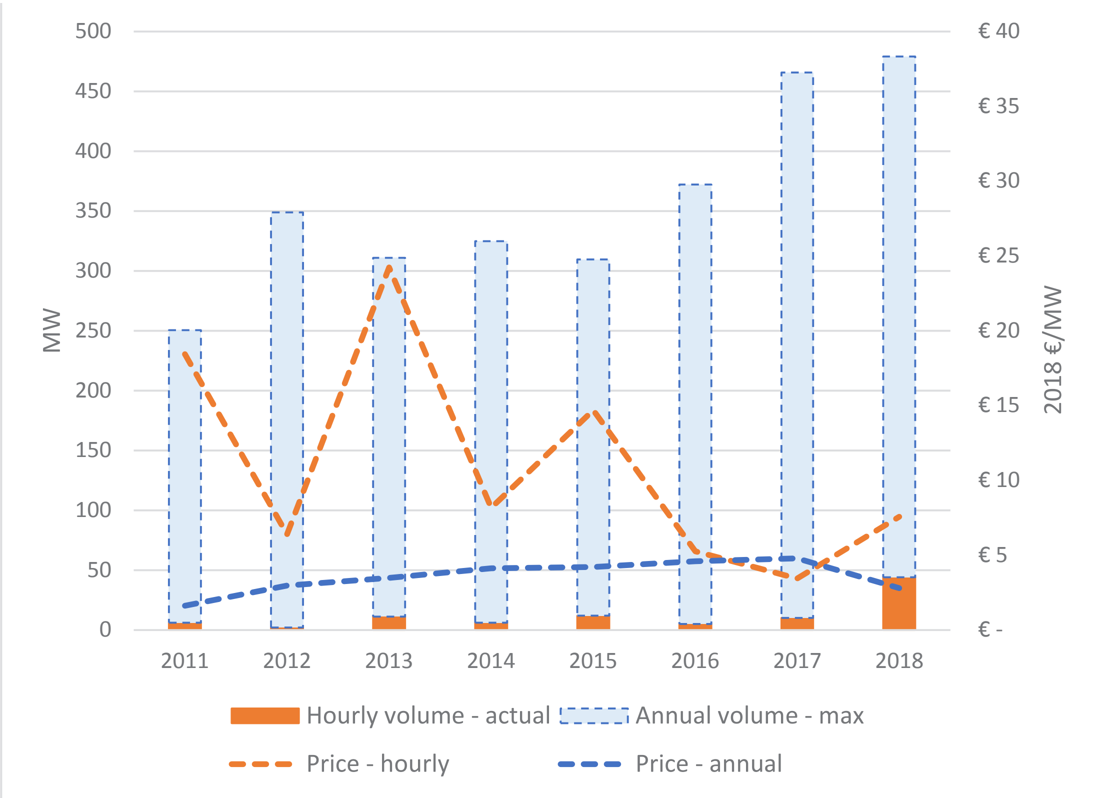


Figure 58: Annual average real prices (€2018) and volumes of FCR-D purchased by Fingrid in the annual and hourly auctions from Jan 1st, 2014 to August 7th, 2018. Note that the annual volume is the maximum capacity was awarded for reserve provision in the yearly market. The actual volume varies from hour to hour depending on how much free capacity the reserve providers make available.

According to Fingrid, the large increase in FCR-D purchased on the hourly markets in 2018 is due to the fact that the purchase of reserves is an optimisation process involving how much to purchase from different markets, while also attempting to keep both markets ”alive”. In addition, the usability of the purchased capacity also plays a role, i.e. how many hours per year are expected[46].

Effects on market actors

For various market participants, the changes undertaken by Fingrid to enable greater DSM participation outlined above (such as alteration of the contract language and technical requirements), allows these market actors additional avenues through which they can offer their flexibility to the market. A perfect example above is the growth in FCR-D supply from

46 “DSR in Finland” Interview with Jäppinen, Jonne, Fingrid, August 2018.

greenhouses and other industrial demand that can withstand occasional short-term power interruptions.

With DSM overtaking the vast majority of the FCR-D market, a relevant question is what earlier suppliers of this product have been pushed out of the market? According to Fingrid, some of what now is provided by DSM was capacity from power plant units that are no longer in the market. However, the majority of the replaced capacity was previously provided by hydropower, and it is likely that this capacity now bids in on the FCR-N market now instead. However, Fingrid requires only roughly 140 MW of this reserve, and it is a market that is already highly dominated by hydro[47].

Figure 59 below illustrates why hydro producers would prefer to receive a capacity payment for the FCR-N (prices are typically 3-5 times higher), but also that the amounts purchased by Fingrid are considerably lower.

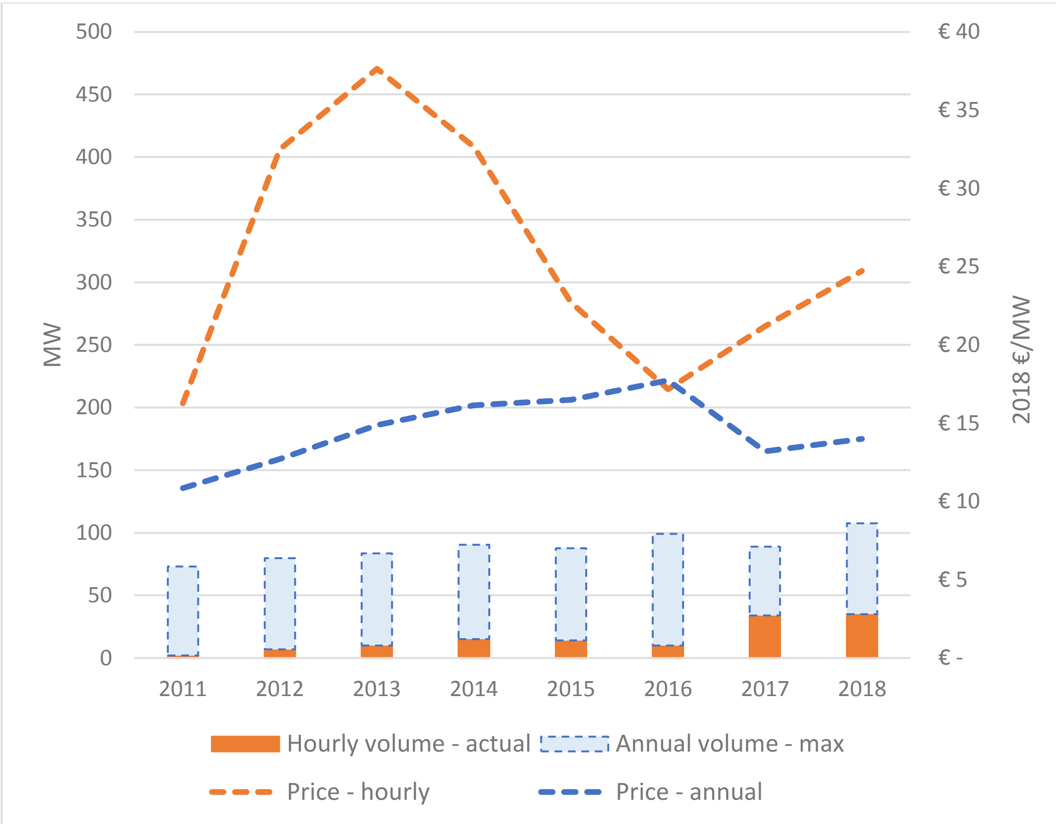


Figure 59: Annual average real prices (€2018) and volume of FCR-N purchased by Fingrid in the annual and hourly auctions from Jan 1st, 2014 to August 7th, 2018. Note that the annual volume is the maximum capacity was awarded for reserve provision in the yearly market. The actual volume varies from hour to hour depending on how much free capacity the reserve providers make available.

The following table illustrates hypothetical examples of annual revenue potential for a few demand response actors on the FCR market in Finland. Hydro is assumed to be able to provide FCR during nearly all hours, while it is assumed that traditional industry in Finland such as wood, chemical or metals can provide FCR-D in roughly 6,000 hours. Greenhouses are meanwhile assumed to provide FCR-D reserve during 4,000 hours during the year (i.e.

47 Ibid.

during hours when there is not enough sunlight).

DR provider	Market	Hours	Price per MW	Revenue per MW
Hydro	FCR-D annual	8,000	€ 4.10	€ 32,800
Industry - wood, chemical or metal	FCR-D annual	6,000	€ 4.10	€ 24,600
Greenhouse	FCR-D annual	4,000	€ 4.10	€ 16,400
Hydro	FCR-N annual	8,000	€ 15.00	€ 120,000

Table 17: Illustrative examples of potential revenue from the FCR market in Finland for DR service providers.

The prices per MW utilised in the example are based on an average of the latest 3 years prices in the annual market as displayed in figure 58 and figure 59. The marginal costs of an activation are quite small, and therefore the primary costs associated with these revenues are the upfront investment costs for the frequency measurement and decoupling equipment and the annual costs associated with bidding into the market. These annual costs may form part of the agreement with the BRP and could be reflected in a slightly lower hourly payment after the BRP has received its portion (see section 2 for more on BRPs).

Outlook for DR going forward

In the future, it is anticipated that growing amounts of both FCR-N and regulating power will come from the demand side in Finland. Current examples of pilot projects include:

- Fortum, a Finish energy company, is undertaking aggregation of a large number of household water heaters (electric boilers), with this capacity to be utilised in the FCR markets. With a large number of users, it is possible to deliver a more linear response.
- Helen, a Finish electricity retailer, and producer, will be using reserve power (i.e. back up), for example from data centres, hospitals and shopping centres, to participate in the balancing market.
- Voltalis, a France-based company, will aggregate a number of demand-side resources in Finland, including households for use in the regulating power market.

5.4 Relevance in a Chinese context

One of the observations from the development of DSM in Finland is that, some of industrial loads could out-compete traditional power plants (thermal and hydro power plants) in providing fast frequency reserves to the system. This is mainly due to the physical inertia of some of the industrial processes, i.e., the short-term decrease of the electricity supply would not result in the significant reduction of production. Thus, the opportunity cost for responding to the power grid is lower than the market revenue.

About two thirds of electricity demand in China is consumed by industrial loads. Some of the industrial loads (e.g. electrolysis, chemical industry, etc.) could provide short-term responses without causing interruption of the industrial production. The potential of using industrial load for demand response is largely untapped mainly due to the inadequate market mechanisms. As for the North and West regions, energy-intensive industry coincides with the variable renewable generations. The reform of ancillary service market in these regions should take into account of the flexibility potential of industrial loads.

6. Electric boilers

6.1 Key messages and takeaways

- Changes of balancing market design have been driving the development of electric boilers in Denmark
- Electric boilers have generally been profitable investments based on income from flexibility services.
- First movers often make the most profit in new markets, but increased competition might dilute profits in smaller balancing markets
- Investing in flexibility can lead to profits in markets not foreseen at the time of investment decision
- Operation and participation in the flexibility markets can be partially automatic when using software provided by Balance Responsible Party (BRP) or SCADA system provider

6.2 Background

Today the Danish power generation assets are made up by roughly 5,000 MW wind and 1,000 MW solar power, about 4,000 MW of large CHP extraction plants with an installed electric capacity of typical 200-400 MW pr. plant and then around 2,300 MW of mainly 1-50 MW small-scale backpressure CHP plants (about 450 plants typical with a capacity below 5 MW and then app. 50 plants with capacity above 10 MW).

Electric boilers are typically installed in combination with CHP plants. Around 300 MW of electric boilers (typical between 2 and 12 MW) are installed today at the small-scale backpressure CHP plants. In a total of around 200 MW is installed at some of the large CHP extraction plants who operate electric boilers up to 80 MW of installed capacity. Network capacity and heat demand limit the size of the electric boilers.

Short technical description

An electric boiler offers very high flexibility at a relatively low investment cost. The nominal investment is estimated to be 70,000 EUR per MW for electric boilers larger than 10 MW. However, the power to heat efficiency is about 95%-98% whereas for instance, heat pumps offer efficiency rates between 300%-500%. Therefore, electric boilers should not be considered baseload units, but rather reserve capacity units suitable for offering flexibility to electricity markets.

Electric boilers have very low standby electricity consumption and maintenance costs. They can increase and reduce their electricity consumption very quickly making them suitable to deliver flexibility at all short time scales ranging from day-ahead and intraday day markets to the different products in the Nordic balancing market.

Role of electric boilers in the different short-term markets

Electric boilers can deliver down-regulation (through increased consumption) when there is ”too much electricity” in the market, either in the day-ahead market, intraday market or balancing the market. In Table 18 an overview is given indicating the relevance of each of the different markets for activation of electric boilers, the timing of bidding and the function of the electric boiler.

Market		Relevance of each market	Bidding process	Function
Day ahead		Limited, but highly dependent on: - Regulation related to taxes and tariff - Number of hours with low prices	First	Power use & heat production
Intraday			Second	
Balancing reserve	Primary	Key market – importance change over time due to competition and regulation	First	Availability of capacity ready for primary regulation
	Secondary	Market design does not allow activation	N.A.	N.A.
	Manual	Key market – market has been boosted since Denmark started to down regulate German excess production	Third	Power use & heat production

Table 18: Overview of the relevance of each of the different market for activation of electric boilers, the timing of bidding and the function of the electric boiler.

Bidding principles for activation in the Day-ahead, intraday and manual down-regulation market

When a district heating supplier operates both a CHP unit, gas fired boilers and an electricity boiler the operator has several options in the overall power/heat optimization. The electric boiler should be activated when the electric boiler has net heat production cost (NHPC) that is lower than for alternative heat production units. The concept of NHPC shows how the coproduction of power and heat can be optimised given different production options. In figure 60 (a simplified illustration) the downward sloping red line shows that as the electricity price increases, the net cost of producing heat is reduced (NHPC is reduced). Once other power/heat production/consumption options (i.e. electric boiler, heat-only boiler etc.) are at the plants disposal, the optimal operation (lowest NHPC) changes.

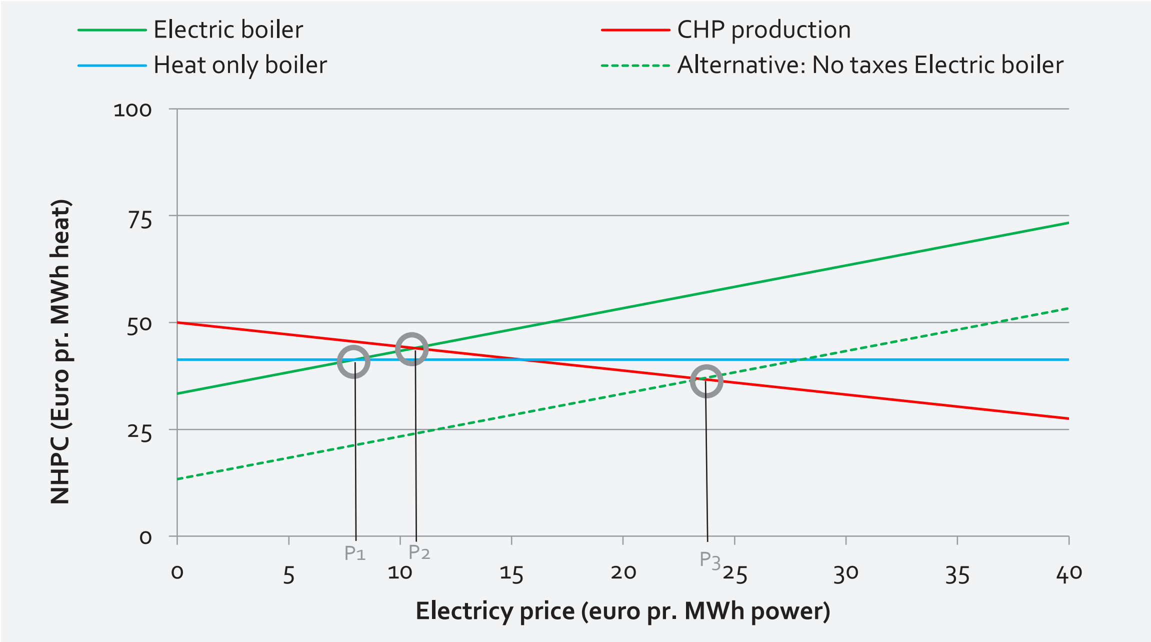


Figure 60: Illustration of NHPC and the activation price points of an electric boiler

- In a CHP plant with only an electric boiler the electric boiler should be activated when day-ahead market, intraday market prices or manual down-regulation price is lower than P2 as it will more economical to do so than to just deliver heat and power at low prices from the CHP plant.
- If a CHP plant also has a heat-only boiler (e.g. a smaller gas engine) then the electric boiler should be activated when day-ahead market, intraday market prices or manual down-regulation price is lower than P1 else the heat-only boiler will be the cheaper alternative.

The solid green line shows that with a relatively high current taxation of electricity consumption for heating the electric boilers are only competitive with other alternatives in situations with relatively low electricity prices. However, if taxes on electricity consumption on electric boiler for heating are removed the situation would be quite different. In this situation, it would be economically optimal to activate the electric boiler up to power prices reaching P3 so the number of operating hours of electric boilers would increase dramatically (numbers in the graph are just illustrative). It is worth noticing that the heat-only boiler in this situation is not competitive at all and thus would only have the role of a backup and peak demand unit only.

This example illustrates that taxes are a very important factor for determining marginal prices and thus also the number of operating hours of the different production/consumption options on the plant. It has recently been decided to reduce taxes on electricity consumption for heating, which will lead to a stronger coupling between power and heat sector and towards an energy system with a higher degree of electrification.

These relatively simple principles are enough to operate the electric boilers in the day-ahead market, intraday market, and manual reserves markets. The owner of electric boilers are simply bidding for buying electricity at prices below their marginal costs – that is the price that can be seen on the x-axis..

Bidding process

Every day the electric boiler operator has to make a trade-off of which market to participate in as the electric boiler cannot offer consumption in day-ahead market, intraday market, manual down-regulation or primary reserves at the same time. The BRP provides indicative day-ahead market price prognoses, but not for the other markets. Therefore, the electric boiler operator each day has to make a pricing strategy that reflects his expectation in all relevant markets that day. If – for example – the electric boiler operator expects attractive price levels in manual or primary markets, the operator should lower his bid in the day-ahead market (and potentially intraday market) market to make sure, that the electric boiler will only be activated in the day-ahead market or intraday market if the activation of the electric boiler will generate a substantial profit margin (higher than the expected earnings in the balancing market i.e. manual and primary reserves).

An electric boiler reduces the NHCP of a district heating supplier if it can produce the heat at a lower net cost than alternative units. The value of the heat is the same for all heat producing units , but the electric boiler can reduce net costs if the electricity can be bought at a low price (even negative) or if the flexible capacity can be sold to the TSO. Therefore, the operator of an electric boiler places bids in several markets to either buy electricity for consumption cheaply or selling flexible capacity for a high price.

When operating in the day-ahead market, the electric boiler is notified at about 13:00 about what hours to run the next day if their bid results in activation of their electric boiler. The electric boiler owner then places bids for down-regulation in the primary reserves market since electric boiler is able to start consumption within 30 seconds that is a requirement in the primary reserve market. Later in the afternoon the operator get the result of the primary reserves auction and know what hours the next day the electric boiler should be ready for delivery of primary reserves .

For the hours in which the electric boiler will not be activated in neither in the day-ahead market or primary reserves then bids can be places in intraday market and if these bids do not lead to activation, bids can then be placed in the manual reserve market. Bids in the manual reserves market can be changed or deleted until 45 minutes before the hour of operation.

The electric boiler is not able to consume the same MW in more markets. So if a 10 MW electric boiler consumes 10 MWh/h in the day-ahead market, then the electric boiler is not able to offer consumption in the intraday market, manual reserves or primary reserves market. However, it is able to offer upregulation (lower its consumption) as it is able to stop consuming with very short notice. This, however, will not be elaborated furtherly as it accounts for only very limited extra profits to the electric boiler owner.

There are some seasonal challenges in utilizing the electric boiler, as the heat demand in summer is low. Therefore, the operating hours of the electric boiler might be limited in the summer, and the operator must be careful to only place bids for the most attractive hours and markets. However, as delivery of frequency controlled down-regulation does not imply a large heat production, this market is also attractive in the summertime.

Data flow for activation

When operating in the power market the owners of the electric boilers use software systems of their BRP or SCADA provider to place bids in the day-ahead market and intraday market and for balancing power – due to market design especially primary and manual reserves. The BRP receives spot market results from the power exchange Nordpool and sends signals to the CHP plants and electric boilers via the software. Regarding balancing power, the system is the same, only the signals come from the TSOs. Most CHP stations have a Programmable Logic Controller (PLC) box installed receiving signals from BRP software and connecting the SCADA system of the CHP stations, thus automating the process. When the electric boiler is ready for activation of primary reserves, it receives signals from a ”frequency sensor box” placed on the grid. This frequency sensor box measures the frequency on the grid and sends a signal to the electric boiler to start consuming electricity (negligible amounts, but enough to assist with frequency control) when frequency on the grid is increasing above a certain limit.

6.3 Operation of electric boiler on a CHP unit with a heat storage tank

If a heat storage tank exists on the CHP plant the plant should optimize it operation to minimize or avoid producing heat when the NHPC is high and instead produce more heat when it is low. The highest NHPC is seen around the interception of the red and the green line in figure 60. The plant should avoid or minimize producing heat when electricity prices are in this ”middle range”.

Instead, the plant should produce more (excess) heat when electricity prices are outside of the middle electricity price range and store the excess heat production in the heat storage tank. A heat storage tank does not impact the lines in the NHPC diagram, but the possibility to store heat allows for decreasing/increasing power/heat production during high/low NHPC periods and thus reducing the average NHPC. Having a heat storage tank thus implies:

- Minimize/avoid producing heat when NHCP is high
- Increase heat production to fill the heat storage tank when NHCP is low
- The heat storage tank can absorb some deviations from the planned heat production
  - » Relevant for heat production associated with the delivery of ancillary services – as results from the ancillary auction are not yet known when day-ahead market bids are submitted
  - » Relevant for errors in prognoses of for instance heat consumption

Optimizing the production on a CHP plant with electric boiler and heat storage tank and often also other production/consumption sources (e.g. heat-only boiler) is a complex planning task as it is not a trivial task to make prognoses of heat demand, electricity prices and free capacity of the heat storage tank. However, it is a task that Danish CHP plants undertake every day using planning tools offered by a variety of different providers. Some planning tools are integrated into BRP software while others are integrated within SCADA systems.

See an example of utilizing the heat storage tank at Skagen Varmeværk in section 6.7. This example also illustrates some of the information available in the SCADA system.

6.4 Issues regarding ownership of heat producing units by different stakeholders

Operation of an electric boiler becomes a complex task especially with the existence of a heat storage tank and several other heat production sources in the system, but the task can be undertaken with support from prognoses and planning systems. Moreover, when the electric boiler and CHP plant delivers balancing services the production plans have to be adjusted quickly as the activations of balancing services is not known before the hour of operation. Therefore, it is important that the production plan for all units – except maybe for certain baseload units such as solar heat and heat pumps – can be changed without any notice to reduce the total NHCP as much as possible. If the different heat production units are owned by different legal entities the contract between the different legal entities in the district heating system must be very flexible and allow for continuous optimization reducing the NCHP. This is, however, easier said than done, and the Danish experience is that the best and most frictionless optimization can be achieved when the flexible heat production units are owned and operated by the same legal entity. If other legal entities own heat producing units it is easiest to handle if they own inflexible baseload units such as industrial excess heat with heat pumps or heat exchangers.

6.5 History of Danish electric boilers and their participation in different markets

The Danish electric boiler market first began taking off when there was an expectation that the further installation of wind power would lead to a lot of hours with very low electricity prices. The market got a boost after Nordpool in 2009 changed the regulation and allowed negative day-ahead prices instead of having a price floor at zero.

The market for primary reserves was from 2009 to 2012 driving the market for investment in electric boilers. Before the change in regulation in 2009 for participation in the primary market, the Danish TSOs had already informed market actors about the plans to open the market. From 2009 the operators of the electric boilers were allowed to make block bids of 4 hours duration with asymmetric products (i.e. operators were allowed to just offer just down-regulation and not forced by regulation to offer a capability of both up- and down-regulation). This change in regulation opened up the market for electrical boilers and implied that electric boilers of almost all sizes could participate in this market. As shown in figure 5 both the number and the installed capacity increased slowly in 2006-2009, but then the installed capacity doubled in three years during 2010-2012.

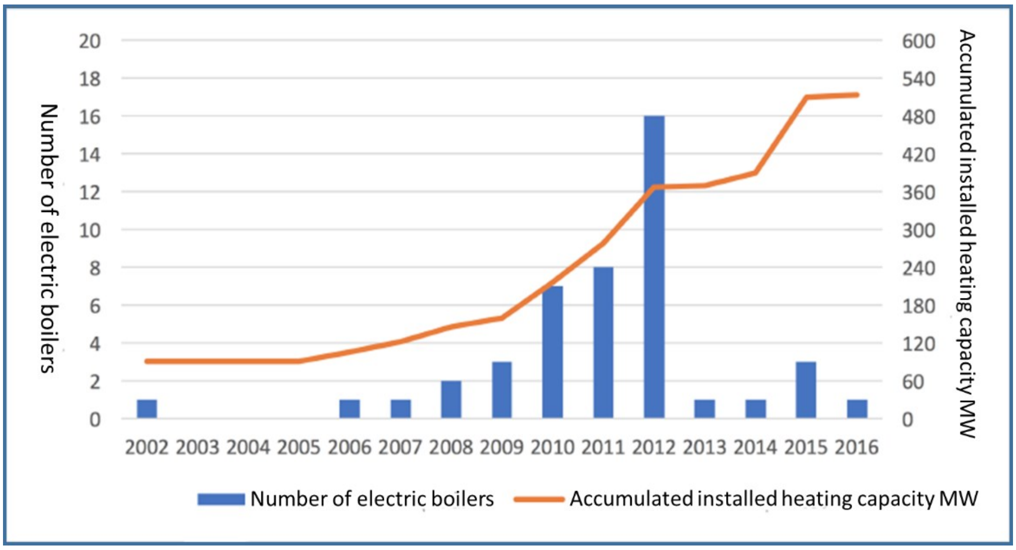


Figure 61: New electric boilers and accumulated installed capacity. Source: Danish Energy Agency and Danish District Heating Association.

Electric boilers impact on the power system

The installation of electric boilers – or for that sake any other flexibility enhancing measures - have not been able to reduce the number of hours where day-ahead market and manual down-regulation prices are zero or negative in Denmark's price areas DK1 as shown in Table 19 (Denmark consists of two price areas i.e. DK1 and DK2) This is due to the fact, that both day-ahead market and manual down-regulation prices in most hours are influenced by other price areas such as Germany or the price areas in Norway and Sweden. In that context, the installation of approximately 500 MW electric boilers does not have a large influence. However, when there is no free capacity on the interconnectors to neighbor price areas the Danish electric boilers are able to influence the prices. That might be the explanation for the decrease in the number of negative down-regulation prices in the years after the peak in 2010. However, after 2013 the number of hours began to increase again reflecting further installation of wind and solar power and also German demand for manual down-regulation in Denmark.

Year	Day-ahead market	Manual down-regulation
2006	29	230
2007	86	195
2008	29	81
2009	56	159
2010	13	302
2011	19	167
2012	35	69
2013	41	131
2014	47	133
2015	68	144
2016	65	174
2017	90	235

Table 19: Number of hours with day-ahead market or manual down-regulation prices at zero or negative[48]

The installation of electric boilers have not been able to eliminate the number of hours with negative prices, but on the other hand day-ahead market prices from 2017 when wind power represented over 40% of the Danish power consumption showed only 90 hours of prices at zero or negative corresponding to 1% of the total number of hours in a year.
In Table 20 consumption from electric boilers compared to wind production and total consumption is shown for Western Denmark[49]. Electric boiler consumption does not account for large shares, but it is worth noticing, that electric boilers are able to increase the consumption with 7 % when day-ahead market or manual down-regulation prices are negative (i.e. when the electric boiler are fully activated). This clearly shows that the primary role for electric boilers is the balancing of the electricity system and not base load heat production.

48 The number of hours with zero or negative prices gives a good indication of the number of hours an electric boiler could operate with appreciable profits. Source: http://osp.energinet.dk/_layouts/Markedsdata/framework/integrations/markedsdatatemplate.aspx

49 Denmark has two price areas, where Western Denmark has the most wind turbines and the most electric boilers.

Year	Electric boiler consumption in percent of wind power production	Electric boiler consumption in percent of total consumption	Electric boiler consumption in percent of total consumption when negative prices
2014	0,9%	0,4%	6%
2015	2,3%	1,3%	7%
2016	1,6%	0,8%	7%
2017	2,1%	1,2%	7%

Table 20: Electric boiler consumption in percent of wind production, total consumption and total consumption during periods with negative prices (Western Denmark)[50].

The relatively low level of electric boiler consumption compared to total consumption should also be seen in the light of the high Danish taxes on electricity consumption. In another tax regime, the electric boiler consumption in the percentage of total consumption would be somewhat higher. The electric boiler consumption in percent of total consumption in case of negative prices would be the same as more or less all electric boilers are activated in these hours.

The primary reserve market – the main driver for investment in electric boilers

Since electric boilers were able to participate in the primary reserves market, this has been a very important factor driving the investments in electric boiler in Denmark, especially in the period 2009-12.

50 Source: http://osp.energinet.dk/_layouts/Markedsdata/framework/integrations/markedsdatatemplate.aspx

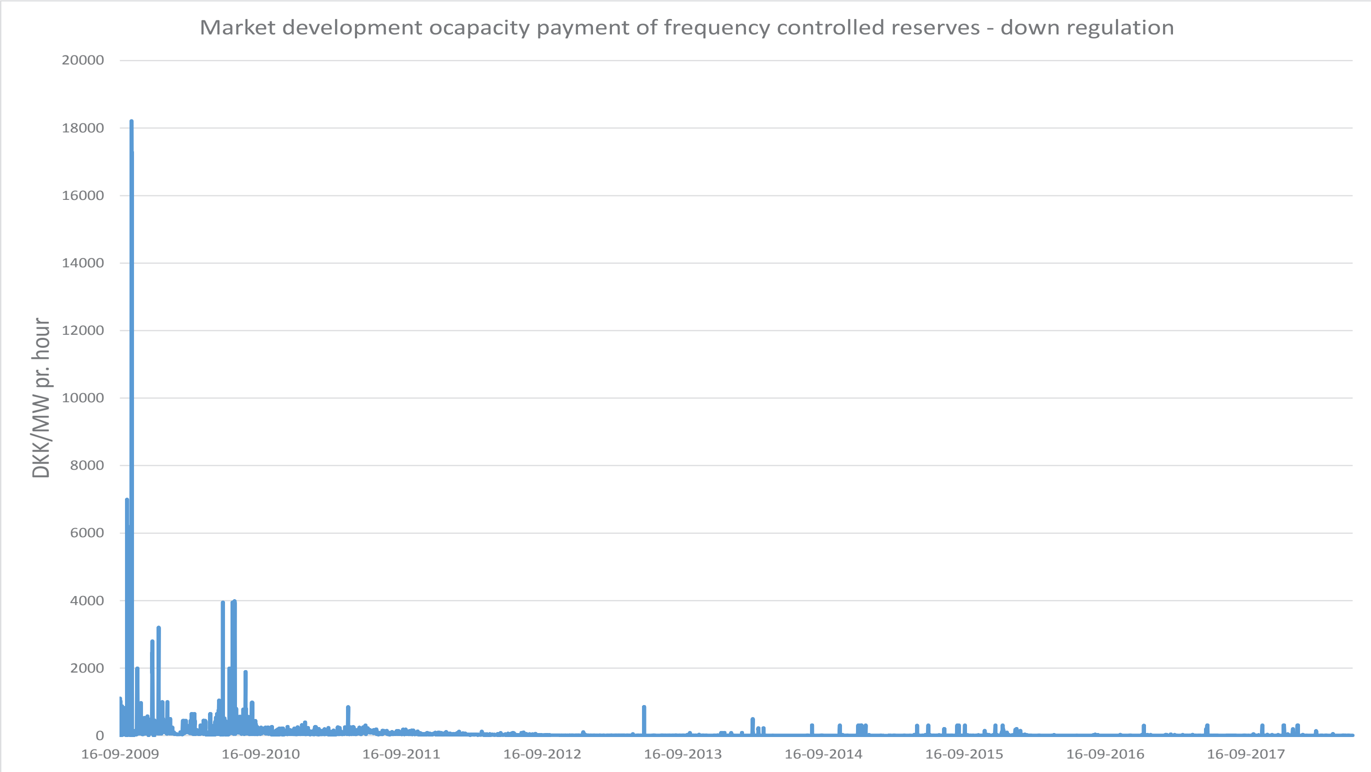


Figure 62: Availability prices of primary reserves for down-regulation[51].

Figure 62 shows that the first movers in the market were able to get a very attractive availability payment for the first couple of years (2009-2010). There were some extreme prices (several thousand DKK/MW pr. hour), but also an availability payment of 500 DKK/MW pr. hour is very attractive when the cost of delivery is close to zero. A contribution margin of 500 DKK/MW per hour is very rare at the day-ahead market and manual reserves market. However, in 2011 and 2012 when the installed capacity of Danish electric boilers increased the market prices of primary reserves decreased substantially making the manual down-regulation market the most attractive market again.

Consequently, the boilers installed in 2012 never got to sell primary reserves to the prices that earlier installed boilers did.

A new market for the services of the electric boiler emerged

However, investing in flexibility in a constantly changing electricity market can give unexpected payoffs. After the 2011 nuclear incident in Japan Germany announced phasing out all its nuclear power capacity partly substituted by investing in wind power. The wind power capacity was installed rapidly and German transmission capacity was not expanded at the same pace. This development paved the way for a Western Danish (DK1 price area) down-regulation of Northern German excess production. This was organised via a special agreement between the Danish TSO and the German TSO.

51 One DDK approximately equals one yuan.

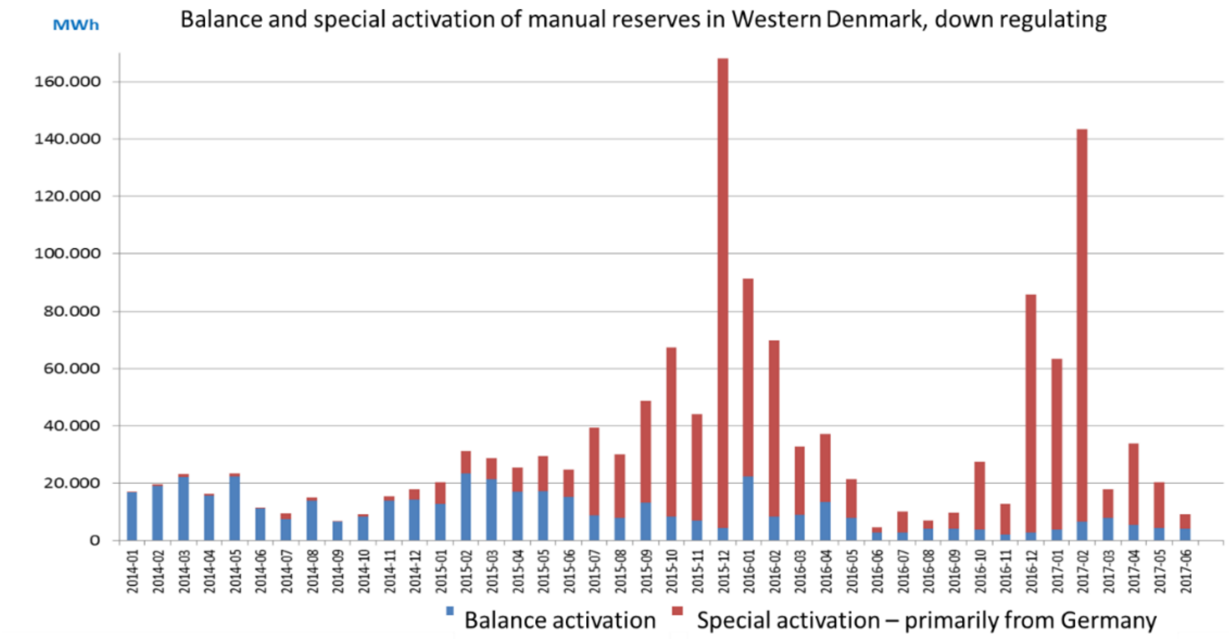


Figure 63: Balance and special activation of manual reserves in Western Denmark, down-regulating

6.6 Outlook and perspectives for Danish electric boilers

Recent changes and debate on taxes on electricity consumption for heat production may give the electric boilers a more important role as heat production units, but the main focus has been on improving the conditions for heat pumps as heat pumps are necessary for making use of industrial excess heat and avoid waste of resources. In the debate on electrification of the heat sector and using excess VRE production, electric boilers and heat pumps are often mentioned in the same sentence.

Year	COP	Baseload	Investment costs	Planning process	Primary dependency	Reduce waste of excess heat	Balancing electricity system (flexibility)
Electric boilers	1:1	Expensive on average	low	short	Electricity connection	poor	good
Heat pumps	4:1	Almost always competitive	high	long	Source of heat	good	poor

Table 21: Attributes of electric boilers and heat pumps

It is, however, important to distinguish. As seen from the Table 21 electric boilers and heat pumps are very different regarding their attributes, the only common denominator being that they consume electricity and produce heat. Given the high efficiency factor (4:1) and low flexibility potential, heat pumps are more useful as baseload heat production units.

The economics of scale and the flexibility of a heat storage tank can improve the business case when electric boilers that can store excess heat production in the tank, and incentivize investment in larger electric boilers. Also the connection point can affect the business case, e.g. in 6.8 The case of the electric boiler at Studstrupværket, which has been connected to the transmission grid together with a CHP plant. When connected to the transmission grid, the electric boiler operator does not have to pay distribution tariff for electricity consumption. This might become increasingly relevant in the future, when electricity consumption taxes are expected to be lowered.

Overall we have identified the following factors influencing the future development of electric boilers in the Danish electricity system.

Factors **positively** influencing the investment in and profitability of electric boilers in Denmark:

- Reduction of tax on electricity consumption
- Future increases in renewable power production will result in more hours with low power prices
- Shut down of coal and nuclear power stations in Europe

Factors **negatively** influencing the investment in and profitability of electric boiler:

- Closer cross-border market integration of balancing services as especially hydropower production in neighboring countries offer very competitive balancing prices
- Increased interconnector capacity to neighboring countries reduce the number extreme events on electricity markets – and the potential for electric boilers are as a flexible unit able to make profits on extreme events and prices.

In the following, two cases of electric boilers in Denmark will be presented that will be affected by the abovementioned factors. The first is Skagen Varmeværk[52] that already has a relatively long history. The second is an electric boiler at a large CHP plant owned by Orsted.

6.7 The case of the electric boiler in Skagen

The Skagen CHP plant is supplying 2.600 households (60.000 MWh) with heat. The heat is produced in mixture with waste incineration plant, gas boilers, a 12 MW electric boiler (since 2008), gas fired engines and also supplied with excess heat from a fish oil factory.

The gas fired engines, the gas boilers, the electric boilers, the heat storage tank and the grid is owned by a cooperative of consumers. The waste incineration plant is owned by the municipality and the fish oil factory is privately owned.

The objective of the investment in the electric boiler was to participate in manual down-regulation, but the primary down-regulation was also a part of the principal business case as the TSO had published plans to change the market design in favour of electric boilers. The investment in the electrical boiler had a very good business case from autumn 2009, paying itself back three times over a two-year period, based on income from primary reserve market from autumn 2009 to autumn 2011. Thus, the investment decision was made on an expectation of future profitable markets, but the decision was also made under a large degree of uncertainty as the price developments could not be predicted accurately.

Participation in the day ahead market and manual reserve market implies a substantial heat production, limiting the heat production of other units when the heat storage tank is full.

52 The small scale CHP plant supplying district heating in the town of Skagen in Denmark.

Therefore, the daily planning of heat production is relatively complex and the operator of Skagen CHP plant must take into account:

- expected heat demand from town based on statistical model and weather forecasts
- level of free capacity in heat storage tank
- gas prices
- expected electricity prices in the day ahead and manual reserve markets
- impacting both electric boiler and gas engines

To optimize daily operations most BRPs and some SCADA providers, therefore, offer optimization software to the CHP stations to able them to take all the above-mentioned factors into account.

Figure 64 shows a simplified figure from the SCADA system of Skagen CHP's production the 1. march 2018. The figure illustrates how the gas boiler, fish oil factory and waste incineration plant are supplying most of the heat demand from the town.

The fish oil factory and the waste incineration plant are baseload units most of the year. The gas boilers can easily be turned off, but as March 2018 was cold and the heat consumption high, the gas boiler was also operated as a baseload unit this day with relatively stable production. The flexible units such a gas fired engines and electric boiler are activated when electricity prices are favourable. As we did not see high electricity prices that particular day, the engines were not activated. Low electricity prices incentivised that the electric boiler was activated for some hours of the day. This can be seen from figure 64b. In hours with low down-regulation prices, the electric boiler is producing and in those hours the heat is stored in the heat storage tank and the heat storage content is increasing.

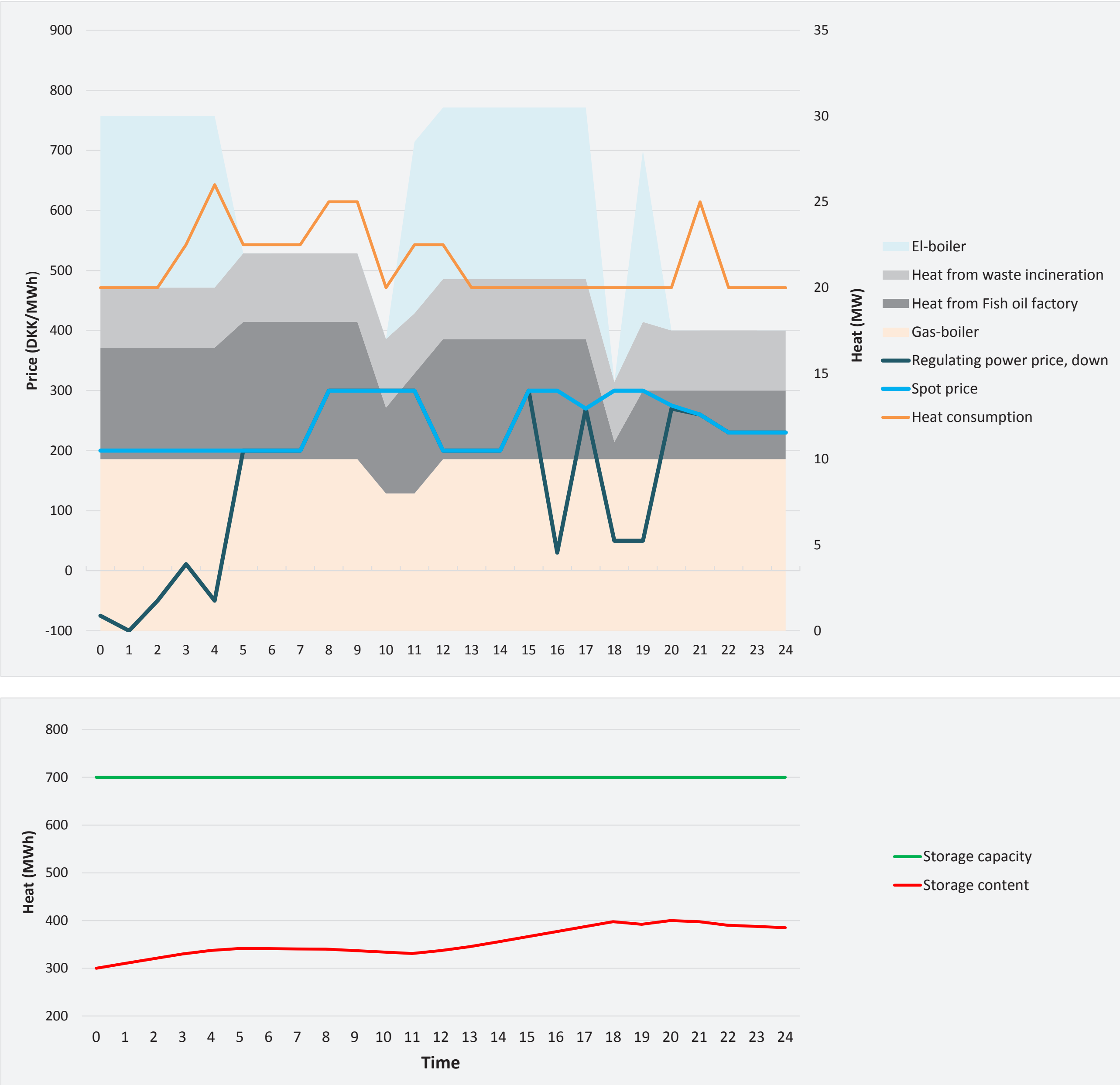


Figure 64: SCADA figures for the production at Skagen CHP, March 1, 2018.

Figure 64 also shows that down-regulation prices between 11 – 15 hours that day are not low. That indicates that the activation of the Skagen electric boiler in those hours was due to 'special down-regulation' ordered by the German TSO via the Danish TSO as a redispatch measure to overcome constraints in the German grid. Data on volumes for the 'special down-regulation' are available, and prove that indeed 'special down-regulation' was activated from hours 11 – 15 that particular day.

Hour	11	12	13	14	15
MWh	292	438	484	437	484

Table 22: MWh special regulation down, March 1, hours 11-15 , 2018[53].

Table 22 shows that the German TSO ordered considerable amounts of downward special regulation (compared to approximately 500 MW installed capacity of electric boilers)[54]. This indicates that the prices of special regulation down could have been very low these hours leading to very low NHPC.

6.8 The case of the electric boiler at Studstrupværket

The electric boiler has been established in cooperation between Orsted, the owner of the largest CHP plant "Studstrupværket" in the area and AVA, operating both the heat distribution network and some CHP plants. The electric boiler with 2x40 MW capacity[55], and a capex of approximately 10 million EUR, is placed next to the CHP plant, which has an installed capacity of 350 MWel and 455 MWheat.

The electric boiler is included in Orsted's real time optimization system used at the CHP plant. Having an asset able to deliver in more than one market 24/7, makes Orsted able to shift between units depending of various prices in various markets. In particular, the electric boiler is used in the following situations:

1. For heat production when a) the CHP plant has to be running due to forced production when there is high demand for district heating, and b) the day-ahead spot market prices are low. Hence the CHP plant and electric boiler are co-optimized against the spot price. Co-optimising assets with different fuels (power and coal, in this case), provide maximum flexibility to Orsted to optimize in the day-ahead market with fluctuating prices.

2. For heat production when the CHP plant suffers full or partial breakdown. That is, the Orsted load dispatcher will be ready to start the electric boiler if supply of heat is threatened. The electric boiler cannot, cover the full production from the CHP plant, but it will buy time for starting emergency production. When the situation is stabilized, the electric boiler can be part of the heat production, given the intraday market price is attractive.

3. The heat production of the electric boiler can prolong the windows for maintenance of the CHP plant. This is a benefit if day-ahead market or intraday market prices are low. Then being able to postpone the startup of the plant, can be a benefit to Orsted.

4. Selling flexibility on the ancillary service market, mainly manual down regulation. The electric boiler does not supply frequency controlled reserves at the moment.

53 Source: https://www.nordpoolgroup.com/Market-data1/Regulating-Power1/Special-Regulation1/Special-regulation-volumes/ALL/Hourly/

54 Please note that also other units provide regulation down. That could be shut down of wind turbines or scaling down coal fired CHP electricity production.

55 2X40 MW was chosen because 40 MW electric boilers are off-the-shelf products.

The electric boiler is operating if the electricity price is attractive compared to alternative fuels for heat production in the day-ahead market, intraday market and ancillary services markets. The payback time for the electric boiler has not been calculated, but it can be noted that several sources of income are necessary to make the investment financially sustainable.
The tariff system in the Danish electricity grid is designed such that an electric boiler can avoid paying the distribution tariff by connecting directly to the transmission grid. The TSO might agree on this setup if the electric boiler has a sufficiently large installed capacity or if the electric boiler can be connected to an existing high voltage transformer station such an the transformer station for the Studstrupværket CHP plant. The connection to the transmission grid saves Orsted from having to pay (expensive) distribution tariffs, and is therefore important to the operating expenditures of operating the electric boiler and can lead to more operating hours for the electric boiler.

Overall, there are important synergies between the CHP plant and the electric boiler, and the installation of the electric boiler has increased the overall flexibility of the plant. Orsted is convinced that the electric boiler will be financially sustainable for the years to come, and Orsted is open to invest in more electric boilers if it is possible to establish good business cases for both Orsted and the heat customers.

6.9 Relevance in a Chinese context

Sector-coupling is the new watchword in the field of system flexibility. The energy flow among electricity sector, heat sector and transportation sector, would provide an enormous pool of flexibility. Power-to-heat is one of the most matured and well-developed technology for sector coupling. Denmark has been using electric boilers to handle variability of wind and solar for many years. The electric boilers not only consume the surplus from wind and solar power thus reduce the renewable curtailment, but also support the power system by providing various flexibility.

The clean heating initiative has significantly prompted the deployment of electrical heating in China. On the demand side, heat pumps are usually used to generate heat for households in the vicinity. On the generation side, electric boilers are installed in CHP plants which are usually connected to the municipal heat grid. Heat pumps do not have much potential on providing flexibility to the system at large scale, but can provide some flexibility when paired with heat storage. Electric boilers, especially electrode type boilers, on the other hand, are extremely flexible and able to provide all kinds of power regulation needed. Recently, many provinces in China are considering introducing both down-regulation product and fast frequency response product to the ancillary service market. Since the electric boiler could provide both products, the financial feasibility for investing electric boilers in power plants in these provinces would be significantly improved.

7. Improving flexibility in CHP hard coal

7.1 Key messages and take aways

- When investing in flexibility, the net heat production costs (NHPC) is reduced for many years to come
- The actual NHCP reduction varies a lot from year to year
- If the unit has to run (forced operation) due power or heat supply stability, low minimum load becomes even more important.
- The single most important investment is in reduced minimum load, but better use of overload and more regulating power is also important.

The large Danish coal fired CHP units still operating today were established during the 1990's. The primary focus of the plants was base load electricity production with heat as an excess product and during the summer time most of the heat was not being utilized. However, large volumes of renewable electricity production have reduced the average electricity prices, and the EU emissions trading system (ETS) has made production based on high-emission technology more expensive. Therefore, base load electricity production has become less attractive to the large CHP units, and heat production and the selling the flexibility of a dispatchable unit on the electricity markets has become more important as VRE production causes fluctuating day-ahead market market prices and a need for ancillary services.

Short technical description

The large coal fired combined heat and power (CHP) units in Denmark have electricity production capacities between 300 and 400 MW and for this case we use a capacity of 375 MW as an example. The plants have all invested in increased flexibility, but the investments made at each plant are not exactly the same. This example will assume a combined heat and power plant with super critical steam cycle with reheat and a once-though boiler with pulverisers.

Minimum load can be reached at 20% of maximum with one pulveriser in operation and heavy fuel oil on stand-by. Fluegas recirculation and eco-bypass control flue gas temperature before deNOx. After DeNOx operation below 300°C for limited periods regeneration takes place at high temperature.

The investments in retrofitting are

- Automated bypass of high-pressure heaters
- Automated ancilliary services (throttling and condensate stop)
- Eco-bypass
- Denox-catalyst
- Optimised burner control for optimal flame stability at single pulveriser operation in min load.

With retrofitted flexibility the minimum electricity production can be reduced and the maximum electricity production can be increased. Moreover quicker receipt of regulation power signals and automated implementation in SCADA system has led to a larger manual regulating capacity where the plant has to change load within 15 minutes.

The coal fired CHP plant in Denmark assumed in this realistic example has the following technical data:The table shows a maximum load increased from 100% to 110%: Before the investments in flexibility overload was also possible, but after the investments, the operations in overload became more flexible and it became easier operate precisely on the load schedule to come back to normal operation schedule. Therefore, operating in overload became more frequent after investing in flexibility. The plant can operate a long time in overload, but at lower efficiency rates, hence operating in overload is only profitable with high electricity prices in day-ahead market, intraday market or regulating power markets.

	Standard flexibility	Retrofitted flexibility
Installed electricity capacity	375 MW	375 MW
Maximum load electricity	100%	110%
Minimum load electricity	20%	30%
Regulating capacity in 15 minutes	53%	60%
Marginal cost of electricity [2]	36.6 EUR/MWh	36.6 EUR/MWh
Operation period	24/7 April-October	24/7 April-October

Table 23: Simplified assumptions for the purpose of this flexibility case
*For simplicity, 36.6 EUR/MWh is assumed to be the marginal cost of the plant throughout the analysis unless other assumptions are clearly stated. In the actual daily planning at large CHP plants, they operate with numerous marginal costs of electricity production. These marginal costs are continuously updated.

7.2 Reduced minimum and increased maximum

When operating during a day with day-ahead market prices both above and below the marginal cost, the retrofitted flexibility improves the operating profits of the plant. Based on the assumptions mentioned in above the principles of daily operations in day-ahead market in case of both standard and retrofitted flexibility are shown below.

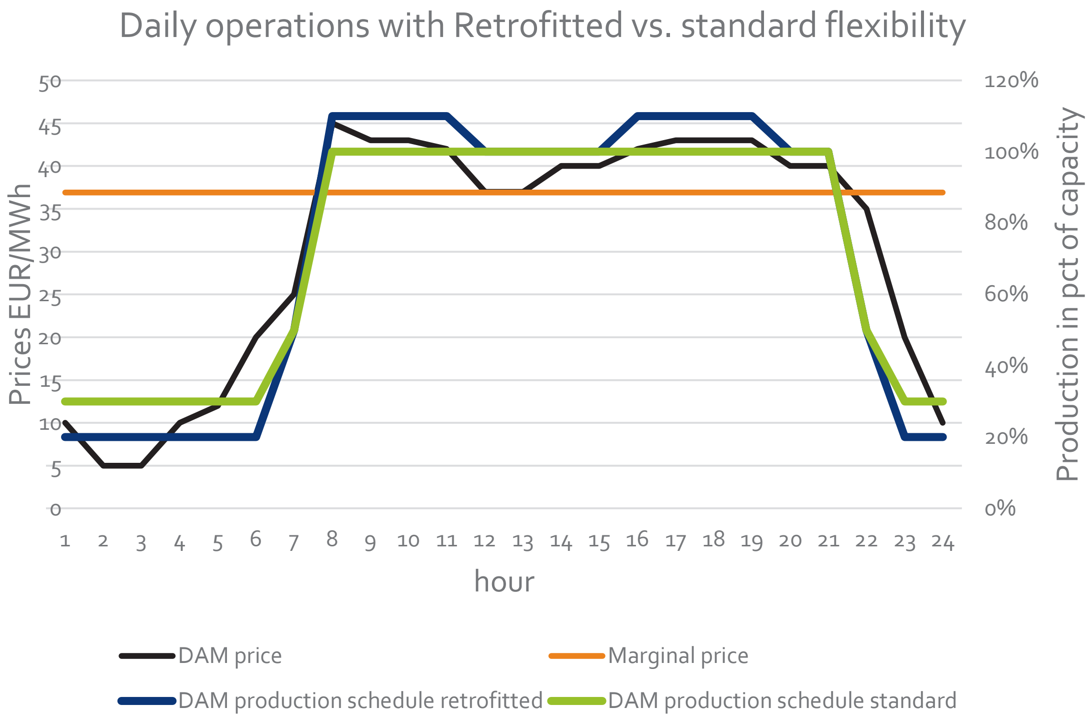


Figure 65: Example of daily operations in day-ahead market

The day-ahead market price (dark blue line) is below the marginal price (green line) in the morning and in the evening. However, the plant does not shut down because of large start-up costs and because of the heat demand from town. It can be seen from the example that the retrofitted plant production (pale blue line) has lower minimum and higher maximum than the standard plant (red line). Thus, when multiplying the electricity production with the day-ahead market prices the retrofitted flexibility plant has a revenue from the electricity market that is 9.450 EUR higher than in the standard flexibility case. That is, investing in retrofitted flexibility improves the financial performance of the plant on days where electricity prices are fluctuating above and below marginal costs of the plant. As the lowest day-ahead market prices are much lower than the marginal prices the increased revenue is mainly due to reduced minimum production. If the example was from a day with very high electricity prices the increased maximum would have been very valuable.

When these simplified assumed plant characteristics are applied to ten years of historic day-ahead market prices in the winter period of October to April the following illustration of the value of retrofitted flexibility can be seen.

DAM prices vary a lot from year to year and these variations have a major impact on the value of flexibility[56]. For the last four mild winters with relatively low day-ahead market prices, the value of reduced minimum operation have by far exceeded the value of increased maximum operation. The years before that the situation was the opposite, as cold winters lead to a high electricity demand for heating. In the cold 2010/2011 winter Norwegian water reservoirs froze, leading to significantly higher DAM prices in the Nordics. Besides from the Norwegian hydrobalance, also increased volumes of other variable renewables in northern European countries have led to increased seasonal fluctuations. In total – and with all the simplifying assumptions – the plant could have reduced the NHCP by 13 million EUR due to reduced minimum load and 12 million EUR due to increased maximum, based on the day-ahead market prices illustrated in figure 66.

56 In the Nordic region the Norwegian hydro balance has a massive impact on overall price levels

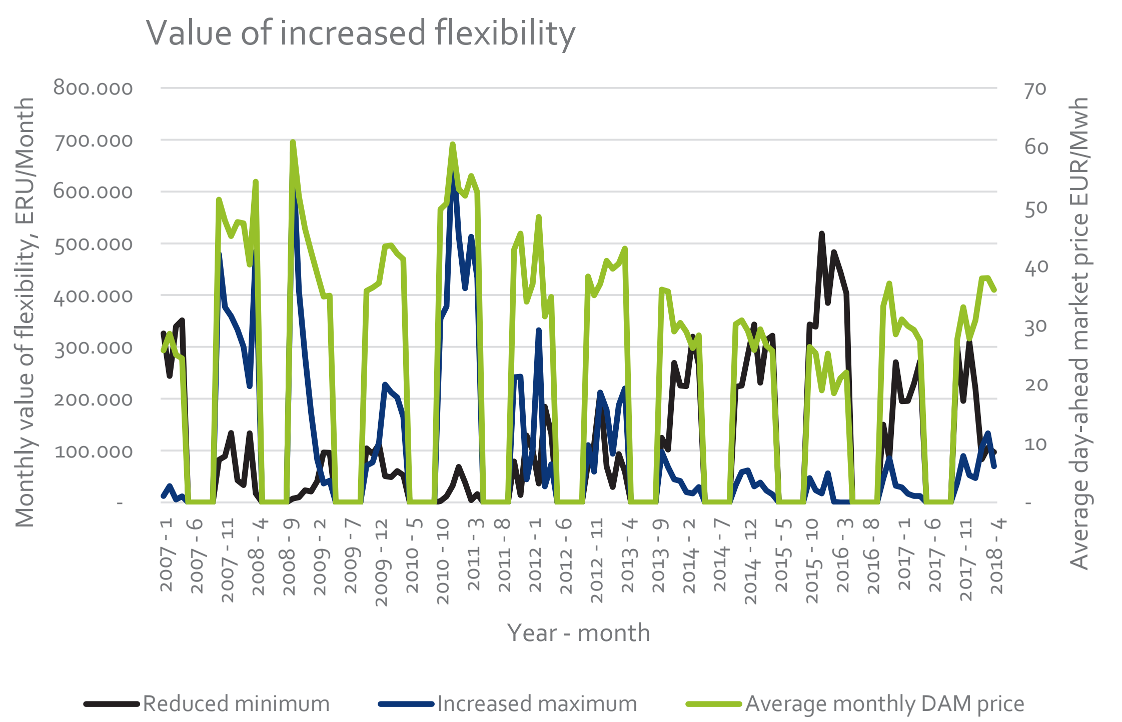


Figure 66: Value of increased flexibility through reduced minimum and increased maximum operation.

7.3 Quicker receipt of regulating signals

Previously, CHP plants received a phone call to inform them in case of activation of their regulating power, which their staff then manually executed[57]. Today, they receive a signal in their SCADA system. For some years the Danish TSO has been using an automated IT system sending out signals to the BRPs who on their side also have automated IT systems passing on the signals to the plants in their portfolio. Recent changes have made it possible for these signals to automatically trigger programmed operations leading to actual activation of the regulating power. The staff in the control room only have to monitor the process.

These changes in operations have led to quicker activation of regulating power and therefore the plant can also make larger adjustments to the production within 15 minutes than before. The figure below illustrates how the volumes of the bids can be increased with retrofitted flexibility. The example is based on the same day-ahead market scheduled production as seen from figure 65.

57 The term manual regulation therefore refers to the situation more than ten years ago, when the process was actually manual.

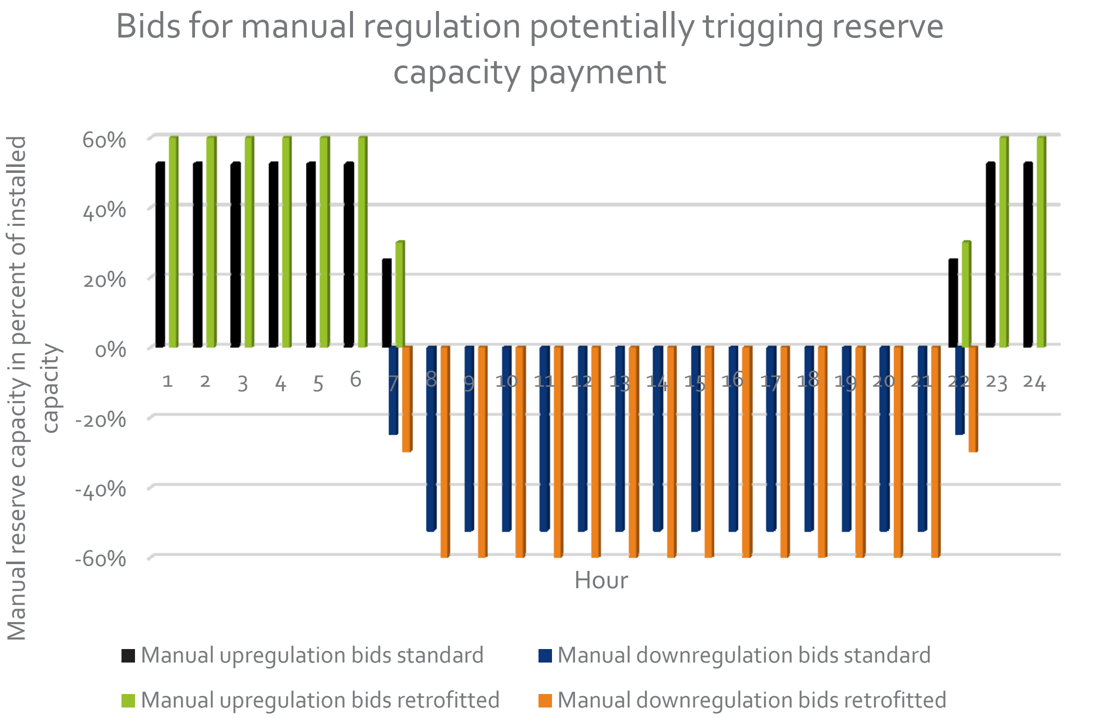


Figure 67: Volume of manual regulating bids with standard and retrofitted operation.

Figure 67 illustrates that the plant can increase the volumes of bids for all operating hours of the plant and thus improve earnings when retrofitted. Thus, quicker receipt and handling of activating signals lead to an increased capacity of delivering regulating power.

The last ten years have seen different market changes when it comes to how the TSO procures manual reserves. Together with changing price levels, this has led to varying value of the capacity payment for regulating bids. This is illustrated in figure 68.

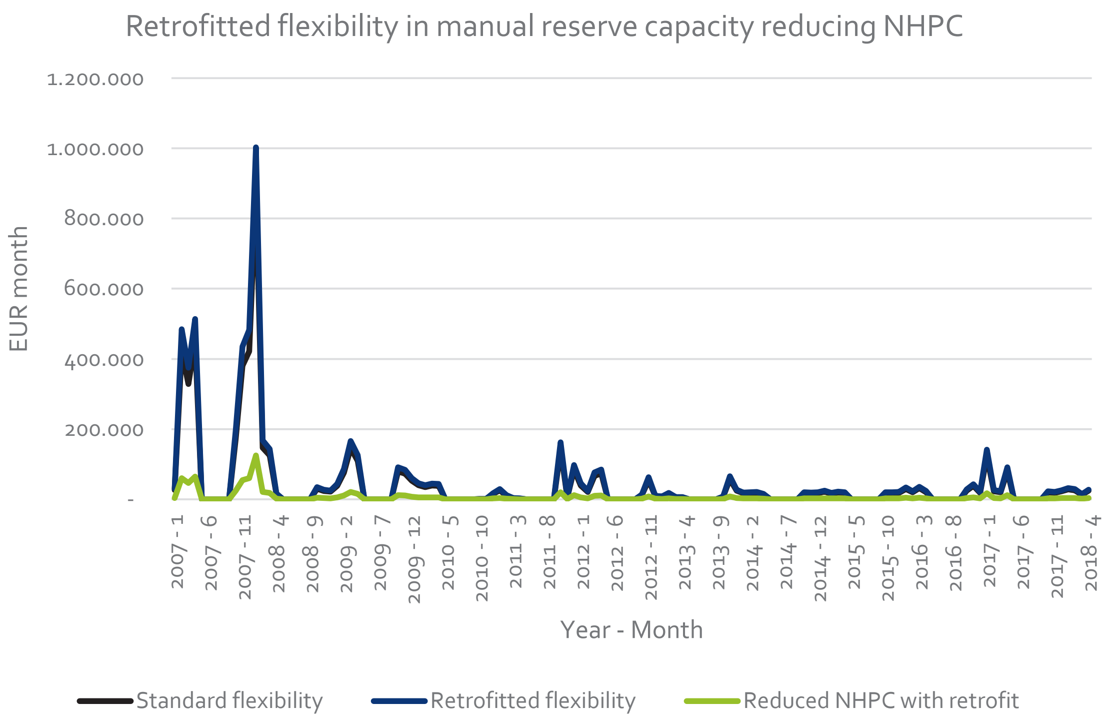


Figure 68: Retrofitted flexibility reduces NHPC due to increased capacity payments for regulating bids

The figure shows how the revenue from reserve capacity payments changes from year to year. The black line is the revenue with standard flexibility and the blue line is the revenue with retrofitted flexibility. The green line constitutes the difference or in other words the extra revenue of retrofitted flexibility. Especially the winter 2007-2008 was very profitable, whereas the market has been less profitable for the last ten years. Adding up for the whole period, the retrofitted flexibility has enabled increased capacity payments for manual reserves by 812.000 EUR and thus reduced the NHPC of the plant correspondingly.

However not only the volumes of the capacity payment bids can be increased with retrofitted flexibility, also the actual activated regulating power can.

The principles in activating larger volumes of regulating power can be seen from the following figure.

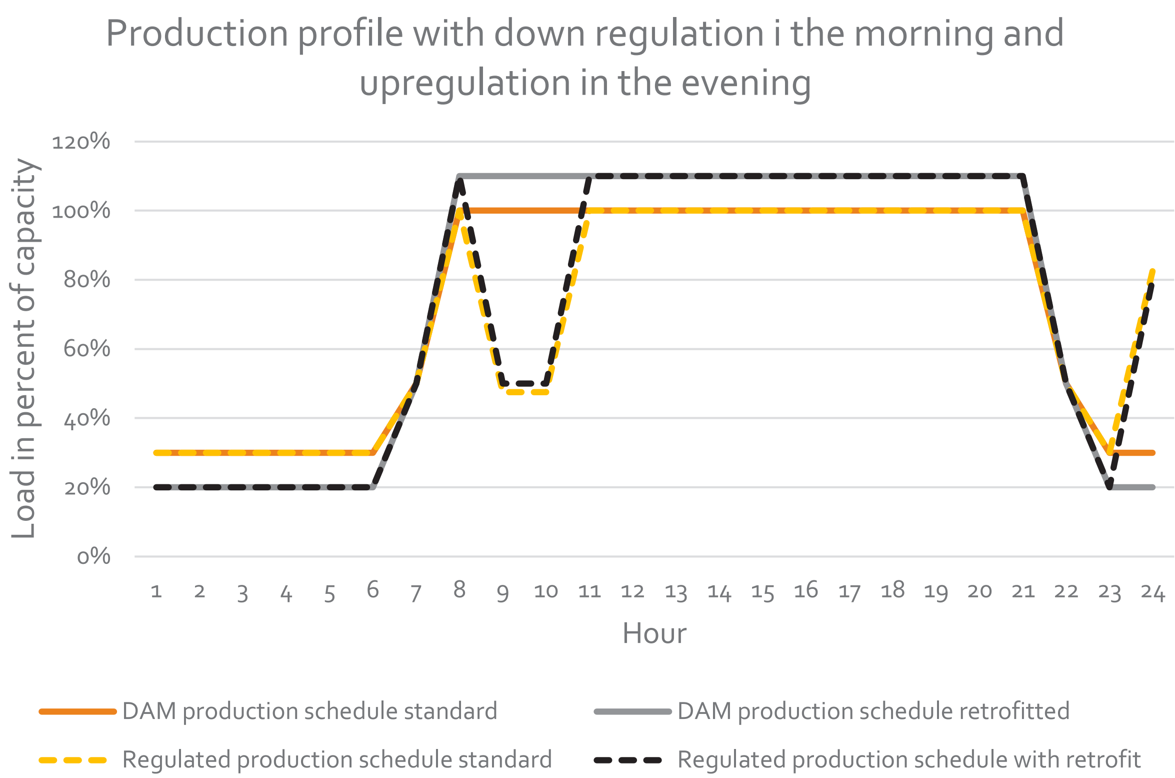


Figure 69: Production profile with activated manual regulating power.

Figure 69 shows the production profile of a day with activated regulating power (dotted lines) in the case of both standard (yellow dotted line) and retrofitted flexibility (black dotted line). With retrofitted flexibility the difference of DAM production profile and regulated production profile is larger than for standard flexibility – meaning that the regulation capacity is larger. The figure shows an example with down regulation from 8 to 10 and upregulation from 23 to 24. Even though the activated volumes seem to differ only by a few percentage points from the standard to the retrofitted case, when the contribution of activated manual regulating power is summarized over more than ten years, the extra volumes do make a difference in reducing the NHPC.

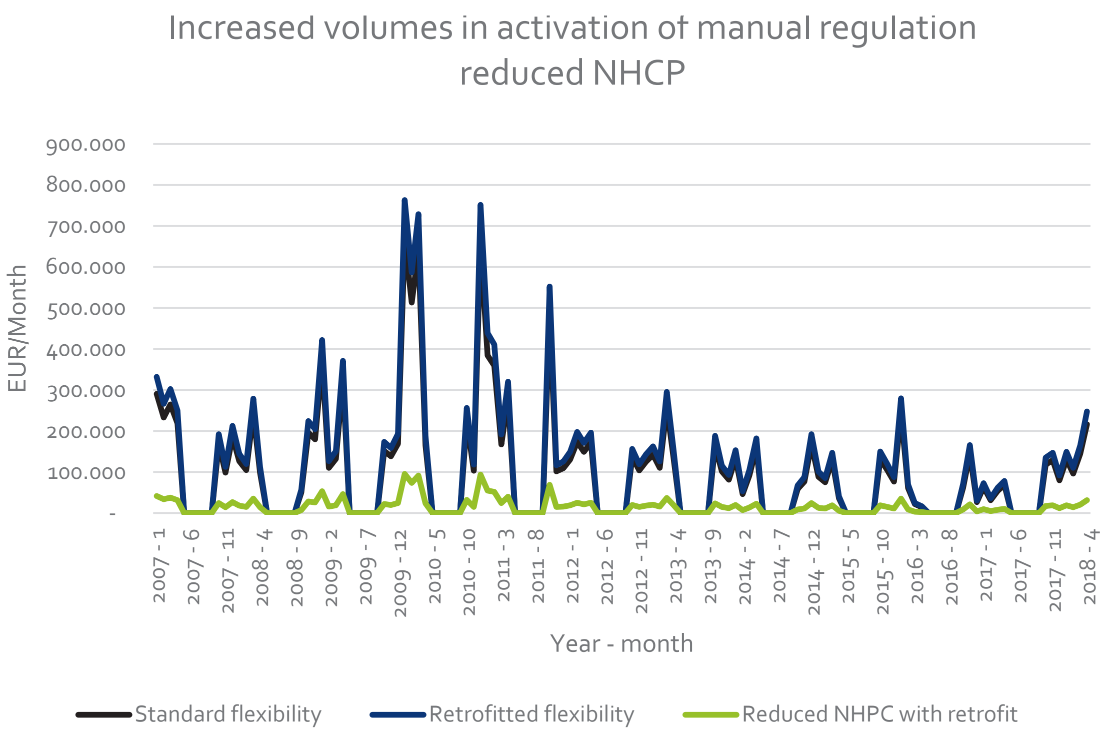


Figure 70: Increased volumes in activation of manual reserves reduces the NHCP

Figure 70 above shows how the activation of regulating power is more valuable with retrofitted flexibility (the blue line) than with standard flexibility (the black line). The green line shows the extra value of retrofitted flexibility compared to standard flexibility (difference between blue and black line). Also in the case of activated manual regulating power the value of the extra flexibility varies a lot over the years. In some months the value is 100.000 EUR, while in others it is only 2.000 EUR. In total, the increased flexibility has been worth 2 million euro for the last ten years based on the published prices of activated manual regulating power. Since the Danish regulation of German excess wind power ('special regulation') began, there have been large volumes of additional down regulation where extra volumes of flexibility have improved the profits of the plant even more. Prices in this market are not publicly available, but the volumes of these special downward regulations are so large that the revenues likely have contributed substantially to reducing the NHCP of the plant[58].

7.4 Other ancillary services

The plant has also invested in automated ancillary services (throttling and condensate stop) that enable the plant to deliver primary and secondary reserves. In these markets the BRP of the plants sends bids to the TSO and receives the results of the auctions before passing signals on lost and won auctions back to the plant. In the case of the primary frequency controlled reserves the plant gets the activation signals from a frequency measuring box placed on the grid and activation of secondary reserves is ordered directly from the TSO on a high priority communication connection. Receiving these signals has required investment in the IT systems and communication lines at the plant but also enabled the plant to generate a significant turnover on these products.

58 See also the electric boiler case for an explanation of Danish down regulation of German excess wind power production

7.5 Relevance in a Chinese context

Flexibilisation of thermal power plants has been proven to be one of the most cost-effective measures in boosting system flexibility, both in Denmark and China[59]. For CHP owners, physically retrofitting the power plants is only part of the story. Better use of the flexibility to reap its benefits requires a systematic approach. As observed in Denmark, CHP power plants need to continually adjust the heat and power production according to the market signal and also invest in new software and hardware. All of these should be built on solid decision making processes. The core concept for power plants owners is NHPC. NHPC measures the economic cost of producing heat with all relevant constraints considered. Its value is factored into major the decision making processes including choosing investment options and doing daily operations.

As more and more power plants have been retrofitted in China, the new market environment would also require the power plant owners to choose the right technologies and also be smart in the daily operation. The concept of NHPC, which is widely used in Denmark, could shed some lights on the decision making of China's flexible CHPs.

59 Thermal Power Plant Flexibility (2018). https://ens.dk/sites/ens.dk/files/Globalcooperation/Publications_reports_papers/thermal_power_plant_flexibility_2018_19052018.pdf

Chapter IV

China

Resume

China is still in the midst of a transitional period of electricity market reform. Trading today is mainly based on long-term (monthly and annual) bilateral contracts. Without spot markets, however, extracting flexibility from the existing assets and encouraging new investment on flexibility requires alternative market design and new business models. This chapter will describe one of the prominent and also successful efforts in China: the down-regulation market in Northern China. Due to the extra economic incentives from the new market, new business models have surfaced in traditional thermal power plants some of which are transformed into a hub for the integration of new flexibility assets.

IV. China

1. Down-regulation ancillary service market in Northern China

1.1 Key messages and takeaways

- Down-regulation market is a side payment mechanism involving only generation side. It is proven to be an effective mechanism to convert part of the curtailment waste to economic incentives for flexibility investment.
- Down-regulation market can coexist with a governmental planning paradigm, and would not require a thorough change to the power sector institution.
- Certain overdrive in economic incentives is usually necessary to attract prompt investment in flexibility.

1.2 Background

At the end of 2017, the installed capacity of wind power and photovoltaic reached 163 GW and 130 GW, respectively. Variable renewable energy (VRE), i.e. excluding hydro power, produced roughly 7% of the total annual electricity consumption in China. The VRE penetration levels are much higher in northern regions, where 2/3 of VRE capacity is installed. As reported by State Grid Corporation of China, VRE penetration ratios of four provinces in the northern regions, namely, Ningxia, Inner Mongolia, Gansu and Xinjiang, exceeded 20% in 2017. Gansu, one of the provincial grids with the highest VRE penetration levels, experienced in 2017 that VRE production at a peak moment reached 67% of the provincial production.

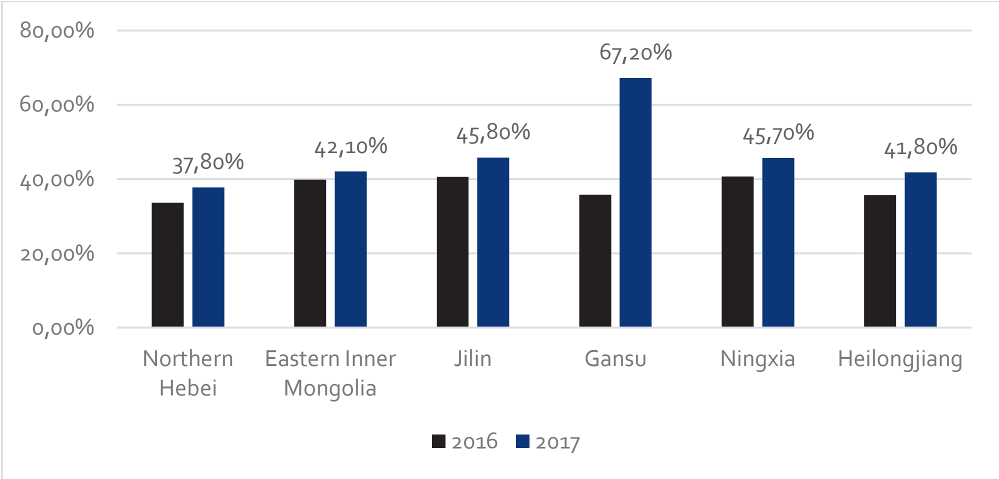


Figure 71: Recorded highest ratio of VRE production to total production.

Following the fast deployment of VRE nationally, was the acute wind and solar curtailment. In 2016 and 2017, the wind and solar production curtailed was close to 50TWh. Multiple measures have been taken to reduce theVREcurtailment, including introducing red-flag warning mechanisms to slow down investment in regions with high curtailment, prompting cross-region/provinces electricity trading, and most importantly, deploying down-regulation ancillary service market in many northern provinces.

Conflict between heat and power

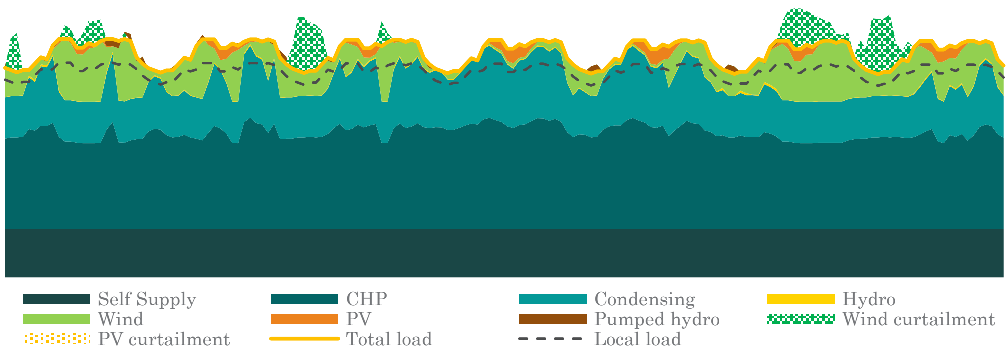


Figure 72: Typical operation of a power system in a Northern Province in a winter week (simulation).

One of the root causes of the high curtailment rate is the conflict between heat and power. Northern regions of China usually have a 4~5 months winter with temperature ranging from -5~-30°. Ensuring safe heat supply in the winter time is prioritized over power system economic operation. CHP units, as the backbones of district heating systems in northern regions of China, are given priority to generate in the winter time. CHP units are originally not designed to be base-load units, as they are the most efficient among fossil fuel generations, thus not very flexible. In the winter time, the forced generation from CHP units squeezes the room in the power system for accommodating wind and solar generation. Even during times with full wind and solar curtailment, the total forced generation from CHP power plants can exceed the valley consumption. The wind curtailment rate in Northeast China, one of the coldest regions in China, exceeded 50% in the winter of 2014 and 2015. How to incentivize the CHP units to voluntarily reduce their generation became the most pressing problem in Northeast China. Sometimes, the surplus and the scarcity are just two sides of a coin. In this case, the dual of the electricity surplus is the scarcity of down-regulation of fossil units. One of the pillars supporting the down-regulation market is the scarcity nature of down-regulation.

Electricity market reform in China

Since the release of Document #92 in 2015, China has embarked on a new round of market reform. The aim of the market reform is to move away from the governmental planning institutional setup to a competitive power market. But this process is far from completed. China is still in the midst of a transitional period of electricity market reform. In 2017, roughly 25% of the total electricity consumption/generation was traded on the market, another 75% was still transferred through grid companies with price and quantities largely determined by local governments. Trading today is mainly based on long-term (monthly and annual) bilateral contracts, there has not taken into account short-term flexibility. In 2017, 8 provinces were announced by NEA as the first batch of provinces to try out spot market, in which electricity is usually traded on sub-hourly basis and thus the value of flexibility could be embodied. However, the implantation of spot market models to an existing large power system is not straight-forward at all. Back in early 2000s, many provinces tried spot market but had to call it off after a short-time period of pilot run. Spot market is inextricably connected to almost every facets of the power sector, ranging from end-consumer's electricity prices to the whole business model of grid companies. As already witnessed in the 8 provinces, the establishment of spot market is usually a protracted process. Northern provinces, enduring the acute curtailment problem, do not have the time for a full-blown spot market to cultivate flexibility in the system. One of logics behind the down-regulation market is that, if the problem is only about encouraging fossil fuel units to free room for wind and solar power, a side payment mechanism among them will be enough. There is no need for a bottom-up reform, as the opportunity cost of establishing such full-blown competitive market is too high in such a context.

1.3 Features of down-regulation market

Here are some of the features of the down-regulation market:

- The side payment only flows among different generators, thus the end-consumer's price will not be influenced.
- Down-regulation market is an auction-based market with sub-hourly uniform clearing price.
- Down-regulation market penalizes those inflexible units while rewards flexible units.
- The economic incentives in down-regulation market is designed to be evidently larger than the cost of flexible operation to attract prompt investment in flexibility.
- Down-regulation market is not Zero-Sum Game and it enlarges the social welfare, as the reduction of curtailment will lead to savings of fossil fuels.
- A sibling policy of down-regulation market is to exempt levies of electric boilers sitting behind the meter of CHP power plant, is essential for the scaling up of this market.

Side payments in down-regulation market

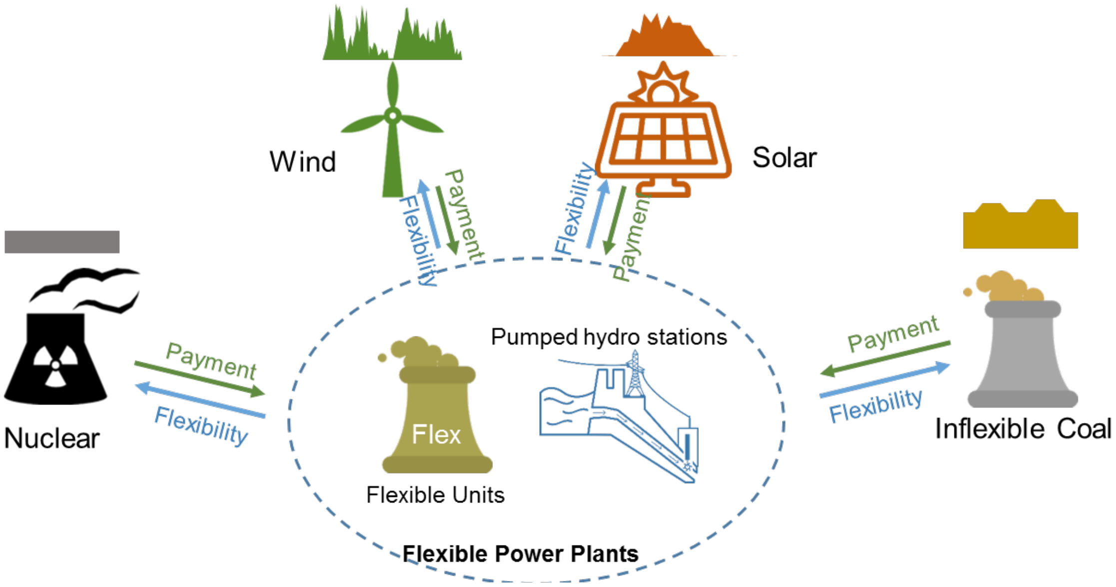


Figure 73: Payment flows in down-regulation market

In down-regulation market, a baseline of down regulation capability is drawn, which in the northeast region is 50%. When power plants operate under the baseline, they are considered to provide ancillary service to the system, while those plants operating above are considered to consume this type of ancillary service. When the system has a generation surplus, power plants operating above the baseline pay power plants operating under the baseline. When the down-regulation was first introduced in Northeast, about 50% of the payment was allocated to inflexible coal power plants. This number is lowered to less than 30% last year, as more and more coal power plants being retrofitted for greater flexibility.

Operation of down-regulation market

Down-regulation market is operated by the provincial and regional dispatching centres (system operators). In northeast, the dispatching centre will ask power plants to submit the price-quantity bid before 10 am every day and will build up an order book accordingly. In the daily operation, the dispatching centre will activate the activate power plants based on each power plant's price bid (lowest got activated first). The bidding price of last unit activated determines a uniform clearing price for settlement next day, and all power plants activated in this time interval will be paid based on this clearing price. The settlement is carried the day after on a 15-min basis. It should be noted that, this is single-side market, with only generation side involved. Price caps and floors are also used in the down-regulation market.

Load rate		Price floor(Yuan/kWh)	Price cap(Yuan/kWh)
Condensing units	40%~48%	0	0.4
CHP units	40%~50%	0	0.4
All	<40%	0.4	1.0

Table 24: Price caps and floors in Northeast down-regulation AS market

Incentives in down-regulation market

The cost of participating down regulation market is mainly two-fold 1) the cost of retrofitting and adding new hardware to enhance flexibility of existing power plants; 2) the opportunity cost of reducing production. The second part of the cost is around 0.2 Yuan/kWh. The first part of cost depends on the upfront investment in the power plant flexibilization, but it usually less than 0.1 Yuan/kWh. The average clearing price of the first two-year operation is around 0.43 Yuan/kWh, which is evidently higher than the cost invoked due to down-regulation. According to many power plants, the payback time of retrofitting existing power plants is usually a few months. For power plants investing new hardware, the payback time is estimated less than 3 years. One observation is that, the decision making behind investing in new technologies is different from that behind investing matured technologies. The payback time should be evidently shorter because of the intrinsic uncertainties and risks of new technologies. To attract prompt investment in new areas, such as flexibilization of power plants, certain overdrive in economic incentives is usually necessary.

Coexist with the existing governmental planning institution

In Northeast, most of generation and consumption is still scheduled by local governments. The thermal generation is still largely priced using the fixed benchmarking pricing mechanism. The down-regulation market is like an add-on to the existing paradigm. Without overthrowing the whole power sector institution, down-regulation markets provide a short-term price signal to boost flexibility.

Levies exempt for electric boilers installed behind the meter

Since the launch of down-regulation market, the investment of electric boilers in CHP power plants has boomed in China. Besides the down-regulation market itself, the removal of all levies on electric boilers sitting behind the meter has also contributed to the rapid growth. Different from other electricity consumptions in the power grid, which is usually charged with various levies (i.e. taxes, grid tariffs and renewable surcharges), electric boilers installed behind-the-meter are exempted from all these levies in China. One of the market rules in the down-regulation ancillary service (AS) markets deems electric boilers as ordinary auxiliary equipment, same as fans and mills that consume electricity without incurring any levy. This rule also raises many discussions. Many argue that this exemption is to encourage the profligate use of electricity which is largely converted from fossil fuels with a factor of 30~40%, therefore it is against the principles of energy saving. Another strong argument supporting this exemption is that, in the time period when the electric boiler is invoked, there is always surplus of electricity and curtailment of wind. Therefore, consuming electricity inside CHP power plants will create room for wind power and reduce curtailment, and thus is equivalent to consuming wind power directly. The down-regulation market rules also require that those power plants installed with electric boilers should also install heat storages, which enable them to further optimize the operation to save fuels.

1.4 Thermal power flexibility enhancement

The wide-spread enhancement of thermal power plant flexibility in China is the direct result of down-regulation market. But when the down-regulation markets was first introduced in Northeast China late 2014, the reaction from the power plants was not as active as it is now. Many power plant owners considered the potential of further reducing minimum load minimal, and were unwilling to invest in retrofitting the assets. EPPEI and DEA organized a study tour to Denmark in early 2016. NEA also participated this study tour and saw the real-life flexible power plants. After this study tour, NEA decided to launch demonstration projects on thermal power flexibility. In total 22 power plants, with a total capacity of 17 GW, joined the demonstration project. The minimum load of many of the coal-fired units has been substantially reduced (to around 30% or even less) and therefore left more space for RE. Many of the demonstration power plants, along with other power plants not in the demonstration projects, have made notable progress on flexibilization of the existing units. The minimum load of some of the condensing units have been lowered from about 50% to 30%. As for CHP units, with some minor retrofitting, the minimum load in winter season has been reduced from 70% to 40%. The net output of those power plants installed with a new electric boiler has even been reduced to nearly zero.

Zhuanghe 600MW unit

Reduction of minimum load on condensing units is usually constrained by two factors: flame stability and emission control. To overcome these two obstacles, the operation mode and control logic need to be optimised. New investments in the emission control system is also required in many cases.

One of the successful examples in the 22 demonstration project power plants is the Guodian Zhuanghe power plant. This power plant has two 600 MW units commissioned in 2007. The 600 MW units used to have a minimum load above 280 MW. After the refurbishment in the last two years however, the minimum load dropped to 180 MW. The main technical solutions utilised at the Zhuanghe plant included:

- Using low heat-value coal in the low load region to keep more mills and burners in operation to maintain the flame stability.
- Bypassing the economiser to increase the flue gas temperature before the de-NOx facilities.
- Systematic optimization of control logic.

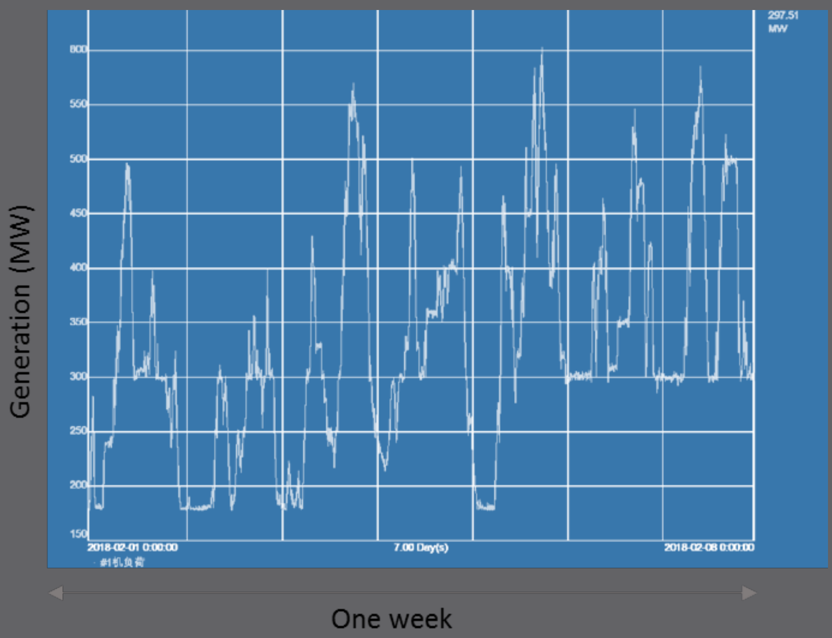


Figure 74: Operational profile of Guodian Zhuanghe 600 MW condensing unit (One week)

Another major achievement of the Zhuanghe power plant is that in the range from 30%-100% load, the emissions are well below the very strict Ultra Low Emission standard (Dust< 5mg/m^3, SO_2< 35mg/m^3, NOx< 50mg/m^3).

The cost of using this technology is highly dependent on the situation in each power plant. In the demonstration power plants, the cost of retrofitting was between 40~100 Yuan/kW.

Jinshan 200MW unit

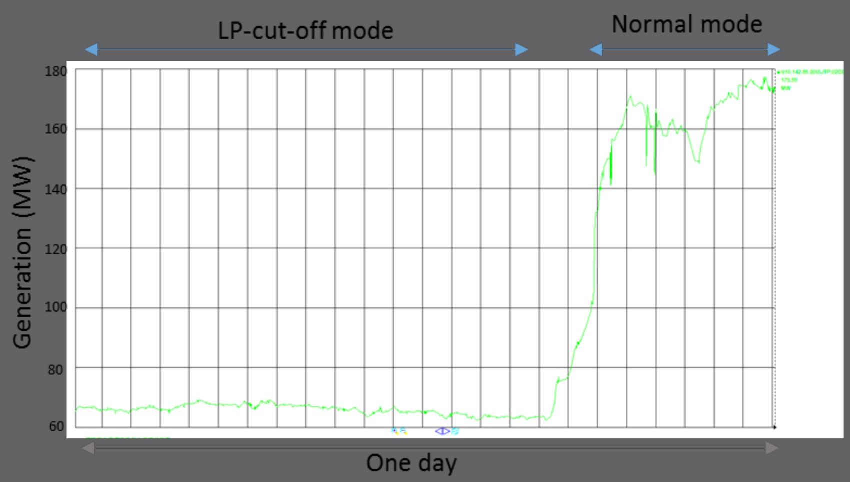


Figure 75: Operational profile of Huadian Jinshan 200 MW extraction unit (Transition from LP-cut-off mode to normal mode)

As outlined, the reason that CHP units (usually extraction units in China) must maintain a 60% or 70% minimum load rate during the winter season is due to heat demand from the district heating system. If the technical constraints for reducing the electricity output are further explored, issues related to the minimum cooling steam of the LP turbine will present themselves. Due to the fast rotation of blades in turbine there is always heat generated from friction. To prevent over-heat and blast, a certain amount of cooling steam needs to flow into the low pressure (LP) turbine. To reduce the electricity output, the minimum cooling steam must be reduced. This could be achieved through optimization of control logic and valves. After the steam flowing to the LP turbine is reduced to a minimal value, the extraction unit will operate almost as a back-pressure unit. Under this mode (LP-cut-off mode), the CHP unit will be able to produce more heat than under normal mode (therefore, with the same amount of heat demand, the electricity output can be reduced). The LP cut-off mode used to be considered technically impossible in China.

In August of 2016, the DEA and EPPEI organized a study tour a to a CHP plant (Fynsværket) in Odense in Denmark where the participants, including senior technical experts from 16 demonstration power plants, noticed that the Danish CHP plant used this mode during the heating season. The delegation had a thorough discussion with the operations manager of the power plant and they realized that LP-cut-off mode could also be achieved in China. After the study tour, Huaneng Linhe, Huadian Jinshan and a number of power plants had successful pilot runs in 2017. One of technical barriers is that the LP turbine will have a transitional blast and over-heat operation, and the key to success is thus how to safely slide from the normal mode to LP-cut-off mode.

A successful example using this technique is Huadian Jinshan power plant. Through invoking LP-cut-off mode, the forced electricity output of the 200 MW unit in Huadian Jinshan is reduced from 170 MW to roughly 70 MW (see figure 75). This has freed up roughly 100 MW for wind and solar power production in Liaoning province. The cost of using this technology is relatively low because little hardware investment is required. The cost is estimated to be less than 50 Yuan/kW.

1.5 Impact of down-regulation market

Since the introduction of this new market, the renewable curtailment in Northeast China has been reduced tremendously. As for the first half of 2018, the curtailment rates for Liaoning, Jilin and Heilongjiang, which used to have the highest curtailment rates, are 1.3%, 6.10% and 5.00%, respectively. The renewable curtailment problem in Northeast is close to be fully resolved. The relative success of the down-regulation market pilot means several other provinces in China are setting up this mechanism. Up to this point, another 8 provinces, including Gansu, Xinjiang, Ningxia, Shanxi, Shandong and Fujian, have established a similar market. Many other provinces are considering launch the down-regulation market in the near term.

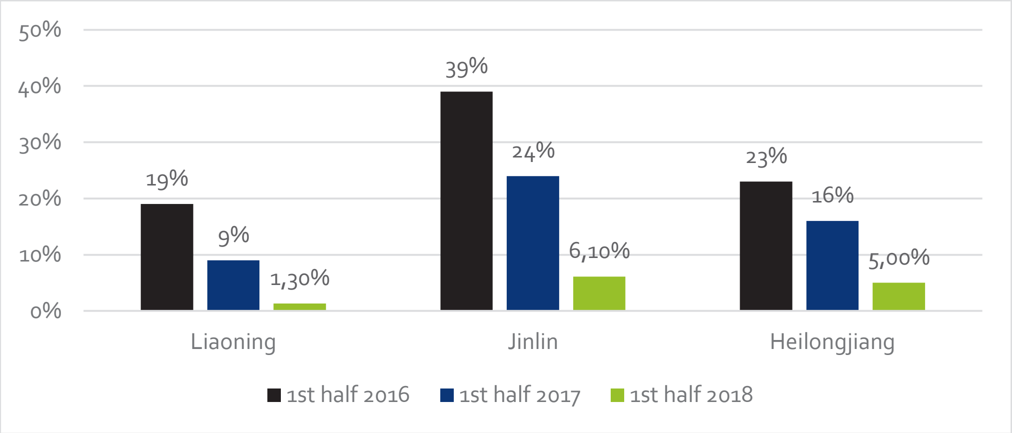


Figure 76: Reduction of curtailment rate since the launch of down-regulation market in Northeast China

1.6 Future development of down-regulation market

The down-regulation market in China is still largely operated on the provincial basis. Since the flexibility potential is limited for one provincial power system, for regions with high VRE penetration (i.e. Northwest China), the coupling of provincial down-regulation markets, which leverages on the complementary nature of adjacent provincial power systems, becomes a necessity. Although the unification of adjacent provincial markets will maximize the social welfare, a welfare distribution mechanism should be designed deliberately to guarantee no province leave the coalition. The coupling/integration of provincial down-regulation markets are in the pipeline already in Northeast, Northwest and North China. Another direction the down-regulation heading is to be smoothly integrated into the future spot market. Fujian province, which already has the down-regulation market in operation, is also one of the 8 provinces trying out spot market. One spot market model Fujian province is considering now, is to combine the down-regulation market with the current long-term trading mechanism. The long-term trading locks price, quantity and generation/ consumption time profile, the deviation is settled through trading down- and up-regulation on the spot market.

2. Thermal power plant as a hub for integration of flexible assets

2.1 Key messages and takeaways

- Adding flexible resources in existing thermal power plants is very cost effective.
- Third-party investment in flexible assets in existing thermal power plants require new business models.

2.2 Background

Thermal power plants in China

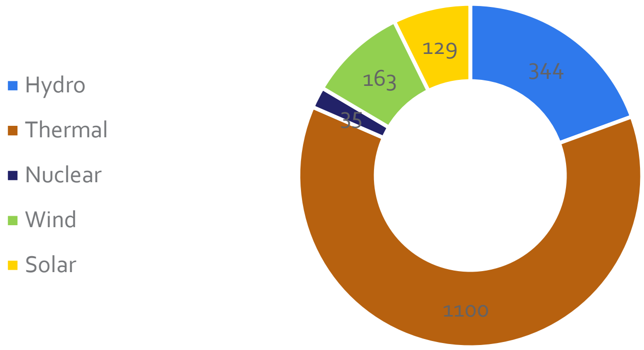


Figure 77: Generation mix by the end of 2017 (GW)

Thermal power plants are the backbones of China's power system. The installed capacity of thermal power plants in China reached 110 GW by the end of 2017, accounting for about 62% of the total installed generation capacity (China Electricity Council statistics). About 89% of thermal units in China are coal fired units. About 80% of the coal fired units are 300MW units and above. Most of the thermal power plants are connected to the grid with a voltage level of 220kV and 500kV.There are various grid codes and technical standards to manage the integration of thermal power plants to the grid. Thermal power plants are required to install various hardware and software, such as automatic generation control (AGC), automatic voltage control (AVC), environmental management system (EMS), wide area management system (WAMS), etc., which allow them to response to the dispatching signal effectively. The cost of building up such a good connectivity with the grid is relatively high for small assets, but not that evident for large thermal power plants.

As the flexibilization of thermal power plants proves to be one of the most cost-effective measures to increase flexibility of the power system, the technological advancements also render many other options more and more attractive. Notably, lithium batteries and electric boilers, once considered to be too expensive for flexibility purpose, show increasingly promising prospects. However, the application of batteries and electric boilers on a utility scale (over MW) requires full-blown grid access hardware and software, to guarantee their observability and controllability. On the other hand, these expensive prerequisites are readily available in conventional thermal power plants. Thus, China has witnessed a new trend of integrating flexible assets, lithium batteries and electric boilers, behind-the-meter in the existing thermal power plants.

Ancillary service markets in China

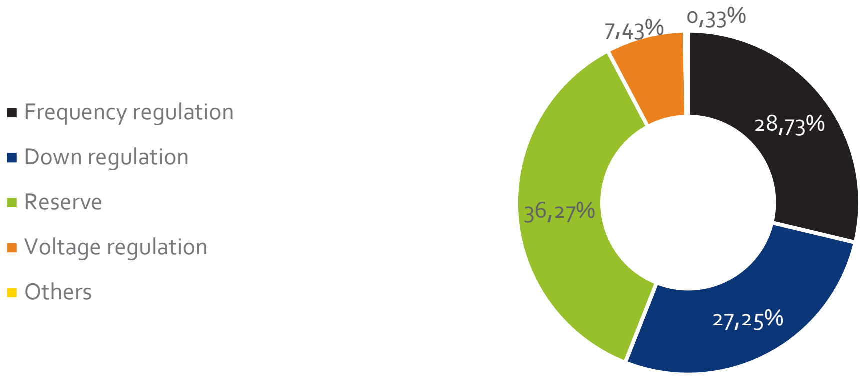


Figure 78: Total payment in ancillary service market in the 2rd half of 2017.

Ancillary service markets, mainly focusing on down-regulation and fast frequency response in China, has boomed for the last few years. In the second half of 2017, the total payments for ancillary services in China is about 6.6 billion RMB (roughly 1 billion USD). The payments were all allocated to units which underperformed in the sense of production response compared to average.

The down-regulation ancillary service market, as mentioned in the last chapter, has tremendously boosted the installation of electric boilers in thermal power plants. The frequency response market, which was also introduced in several provinces in China, has boosted the installation of battery in thermal power plants. Now, thermal power plants have become hubs for the addition of flexible resources in the system.

Frequency response ancillary service market in Shanxi province

Shanxi province was the first to introduce the frequency regulation market in China. Before this, the grid company managed the frequency response (which is usually referred as the AGC response in China) using a scoring-board scheme, which adds and deducts points according to the performance of each power plant. The designated point for each indicator of performance is fixed, thus this scheme could not reflect the short-term scarcity of frequency regulation. Shanxi's frequency response market is a day-ahead single-side auction market. The market is organized as follows:

- Determine the need for the frequency regulation of the next day. The default value is set to be 15% of the total capacity of units to be committed next day.
- Calculate the coefficient of performance of each bidding units based on the nearest historical data. Three indicators are used to characterize the historical performance of each unit, including response time, regulation speed and response deviation.
- Recalibrate the bidding price using the coefficient of performance. The bidding price of units under-performed historically will be raised, thus decreasing their possibility of being activated.
- Determine a uniform clearing price based on the bidding price of the marginal unit.
- The units are activated according to their bidding price.
- Payments are allocated to units that are not activated.

The bidding price in the market is capped and also floored. The price range was initially set as 12-20 Yuan/MW, and then was adjusted to 5-10 Yuan/MW because the fast deployment of batteries in the system pushed the bidding price down. Battery installed behind the meter in a power plant is viewed as an auxiliary equipment and is not subject to different levies, such as grid tariff, taxes or renewable surcharges.

In the 4th quarter of 2017, the total payment in Shanxi's frequency response market was around 180 million RMB. The frequency response market in Shanxi has boosted the investment of utility scale batteries in Shanxi province. By end of 2018, there are three battery facilities are in operation in Shanxi, amounting to a total capacity of 27MW/13.5MWh. Another 11 projects are in pipeline in Shanxi. All of those projects are behind-the-meter of coal power plants.

Third party investment

The introduction of market also brings in certain uncertainties, which were not seen in the traditional governmental planning paradigm. Most of the power plants are state-own enterprises in China. These power plants are generally risk-adverse, thus unwilling to engage in new businesses with high uncertainties. Moreover, due to the rise of coal price and the reduction of utilization hours, coal power plants in China cannot support a large investment financially. Banks are also reluctant to provide large loans to traditional power plants. It is estimated by China Electricity Council that the thermal power sector of the largest 5 power producers in China lost 13.2 billion Yuan in 2017.

On the other hand, many private sector companies, owning some of the state-of-art technologies, are more willing to take on the market risks. Therefore, almost of all the new flexible assets in the traditional power plants, including large-scale electric boilers and utility-scale batteries, are invested by third-party private companies.

2.3 Case study I: Utility scale batteries in thermal power plant

Installing batteries has become an attractive option to boost the frequency response capability of thermal power plants for their owners, as the cost of lithium battery continues to drop. One of the advantages of using battery for frequency regulation is its much better response rate compared to thermal units. It is estimated that lithium battery is better than thermal units by a factor of 20-30. The marginal cost incurred by charging and discharging battery in the process of frequency response is comparable to thermal units. Compared to thermal units operating part load, lithium battery usually has a lower marginal cost.
There are also some cases of installing batteries in thermal power plant in Europe. For example, STEAG, one of the largest power producers in Germany, put a 90MW battery station in Weiher thermal power plant in 2016. Kilroot coal power plant in U.K. installed a 10MW battery station in 2016. China started to see a burst of battery installation in thermal power plants in early 2017, due to the revision of market rules on frequency regulation ancillary services. Unlike the cases in Europe, most of the battery stations in thermal power plants in China are not invested by the hosts, (i.e. power plant owners), but owned by third party investors.

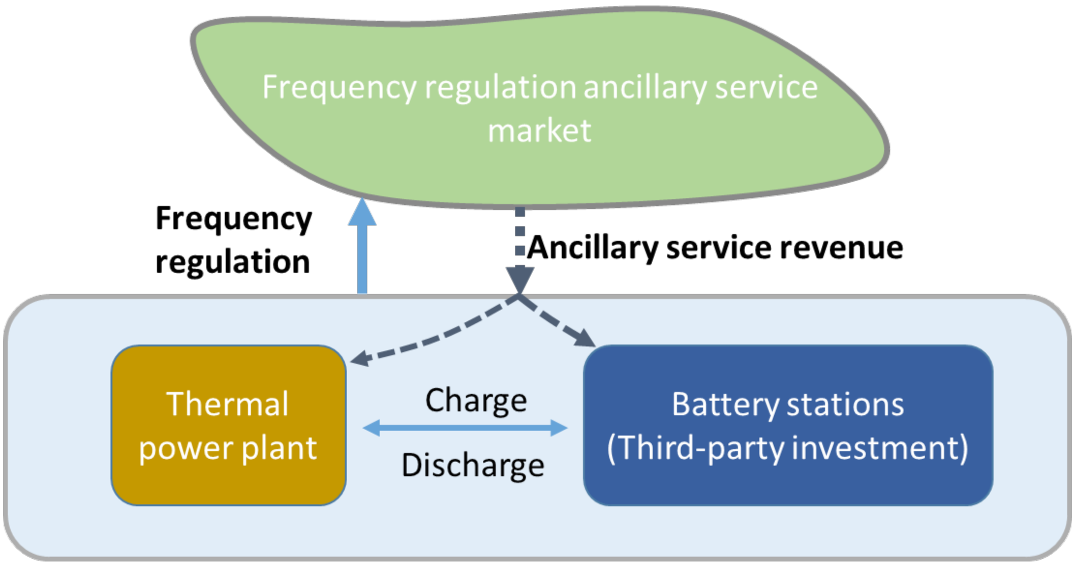


Figure 79: Business model for battery investment in power plant.

The core part of the contract between the power plant owner and the third-party investor is the distribution of earning from the frequency regulation ancillary service market. In most of the cases, the third-party investor will take 70%~80% of the earning, while the power plant owner, who provides the space for battery instalment, will take the remaining. As the competition among third-party investors becomes increasingly intense, while the number of thermal power plants suitable for battery installation is limited, the bargaining power starts shifting to the power plant owners. A lower sharing of earnings for battery investors has also been witnessed in several recent projects.

According to public available information, the payback time for such projects is usually in the range of 4~6 years. Thus, these projects are not of strategic purpose, but also good businesses.

2.4 Case study II: large scale electric heater/boiler in CHP power plant

The down-regulation ancillary service market attracts more and more investment in the electric boiler/heater installed behind-the-meter of CHP power plant. Four CHP power plants in northeast China have installed large scale electric heater and heat storage. The electric heater usually has a capacity of 300MW, and the heat storage has a capacity of 1500MWh to 2000MWh. The medium used in the heat storage is magnesium oxide (MgO) brick, which can be heated up to 500° when there is electricity surplus in the grid. The energy density of MgO bricks, in terms of kJ/L, is about 3 times of that of hot water storage. With the help of this extra electric load, the net output of the CHP unit can reach almost zero net output, without significantly influencing the district heating temperature. In one winter season, each of this large storage facility could absorb more than 200GWh electricity surplus.

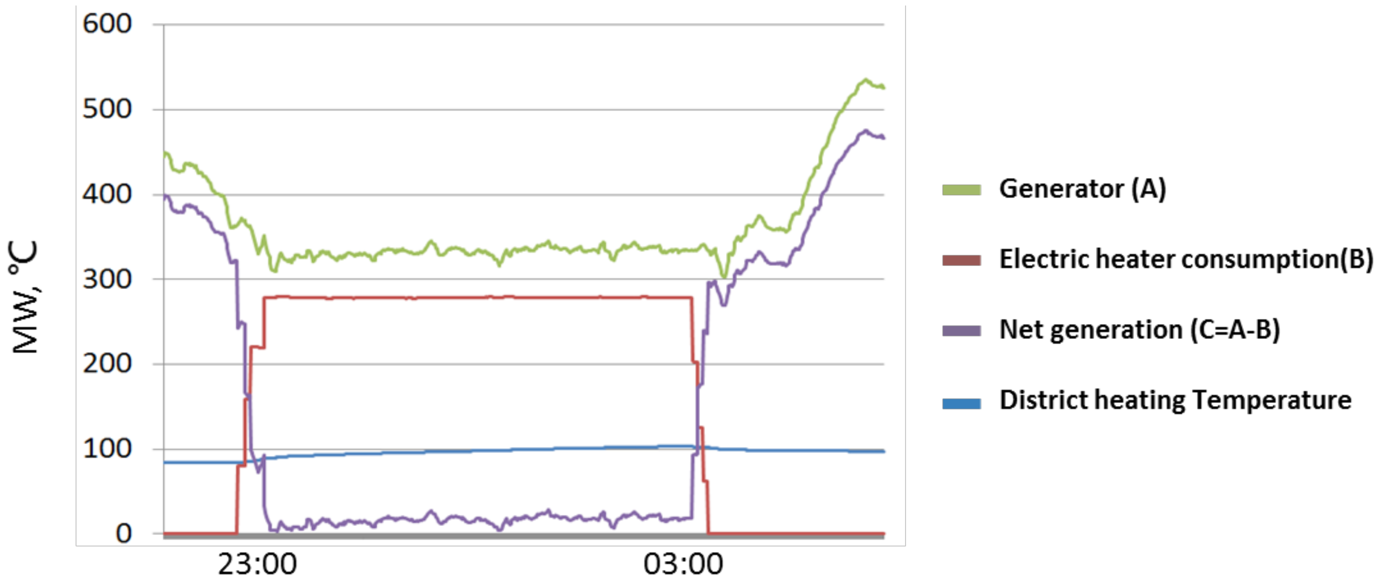


Figure 80: Operational profile of Diaobingshan CHP power plant (280MW electric heater + 1960MWh MgO heat storage).

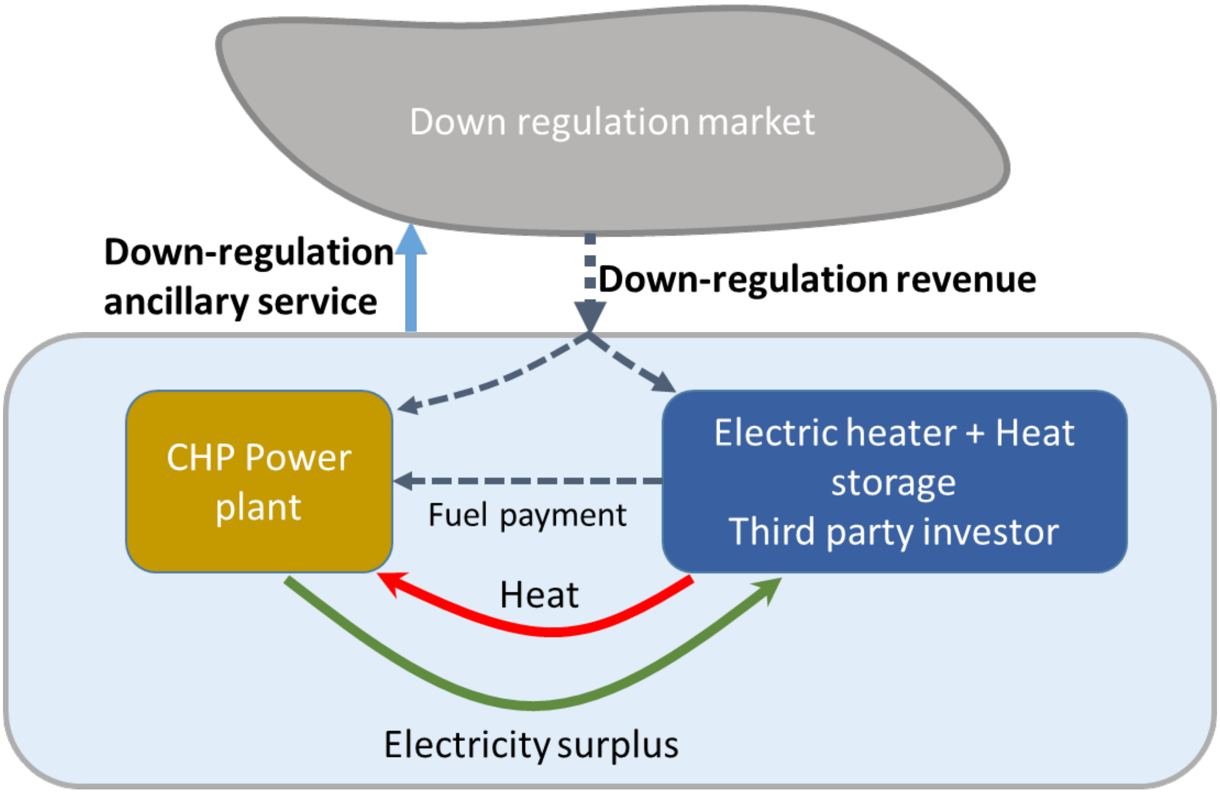


Figure 81: Business model for electric boiler in CHP power plant.

Almost of all the large heat storage facilities are invested by a third-party private company. These companies are usually have more capital at hands and are willing to take risks. The business model is illustrated in figure 81 when the system needs down-regulation service (usually in time period with strong wind in late night), CHP power plant will sell some of its generation to heat storage facility investor, and the heat storage investor will pay the power plant based on the fuel cost. The revenue they got from the down-regulation market will be distributed according to predefined contract. The heat will be stored and transferred back to the CHP power plant as the power plant required.

Danish Energy Agency

电力规划设计总院
China Electric Power Planning & Engineering Institute