# The Evolution of Ancillary Services to Facilitate Integration of Variable Renewable Energy Resources

*A Survey of Some Changes to the Ancillary Services and Ancillary Service Markets*

**3002008987**

# The Evolution of Ancillary Services to Facilitate Integration of Variable Renewable Energy Resources

*A Survey of Some Changes to the Ancillary Services and Ancillary Service Markets*

**3002008987**

Technical Update, December 2016

EPRI Project Manager

A. Tuohy

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1338 ▪ PO Box 10412, Palo Alto, California 94303-0813 ▪ USA
800.313.3774 ▪ 650.855.2121 ▪ askepri@epri.com ▪ www.epri.com

## DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES



**THE ELECTRIC POWER RESEARCH INSTITUTE (EPRI) PREPARED THIS REPORT.**

**This is an EPRI Technical Update report. A Technical Update report is intended as an informal report of continuing research, a meeting, or a topical study. It is not a final EPRI technical report.**

## NOTE

# ACKNOWLEDGMENTS

The Electric Power Research Institute (EPRI) prepared this report.

Principal Investigators
E. Ela
A. Tuohy

This report describes research sponsored by EPRI.

---

This publication is a corporate document that should be cited in the literature in the following manner:

*The Evolution of Ancillary Services to Facilitate Integration of Variable Renewable Energy Resources: A Survey of Some Changes to the Ancillary Services and Ancillary Service Markets.* EPRI, Palo Alto, CA: 2016. 3002008987.

# ABSTRACT

The set of ancillary services used to ensure a reliable and efficient electric power system are in many ways the same types of services that have been provided for decades. The way in which power system operators and planners are achieving sufficient quantities of those services, and how system resources are scheduled and incentivized to provide those services, is evolving. In particular, relative to this report, increasing levels of Variable Energy Resources are impacting system operations due to the unique characteristics of these resources. The reliability services considered must therefore evolve in order to ensure reliability in an efficient manner, as well as providing for fair treatment of market participants.

This report provides a summary of some recent changes from balancing areas, independent system operators, and transmission system operators from the United States and abroad. An overview of some of the ancillary services used today is provided, as well as a brief introduction to the design of ancillary service markets. A number of changes to these ancillary services and ancillary service markets is then discussed, providing a variety of different approaches being taken around the world to effectively integrate variable energy resources into the power system.

**Keywords**

Renewable integration
Power markets
Ancillary services

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1 INTRODUCTION

Variable renewable generation, such as wind power generators and solar photovoltaic power generators, have grown in significant numbers in recent years. According to the American Wind Energy Association, the U.S. has over 73,000 MW installed capacity across the country, and represented over 40% of all new generator technology installations in 2015[1]. Solar power now includes over 24,000 MW installed in the U.S. with over 6,000 MW installed in 2015[2]. Variable energy resources (VER) have tremendous benefits, including emission-free energy, and zero fuel costs. However, due to particular characteristics of the technologies and their fuel sources, increased penetration of VER also creates unique challenges for system operators and planners in order to maintain reliability of the electric power system in an efficient manner. Because of these characteristics, many regions have adjusted the ways in which they require and procure reliability services, also referred to as ancillary services, to maintain reliability, ensure efficient least cost production to meet load, and, in certain restructured market regions, ensure that incentives are compatible with the directions that resources are being asked to follow.

It is important to recognize the unique characteristics of VER that may lead to the changes in these ancillary services. We list those that are most relevant to the discussion of this report below, with short descriptions followed. There are other unique characteristics not covered here, that may need special attention as well.

- Variability: the fuel source of VER, wind speed and solar irradiance, vary with time, and the total maximum capability of VER during one time period will differ from other time periods.
- Uncertainty: It is not possible to predict the power capability of VER with perfect accuracy. Thus, as it is changing from time period to time period, the system operators are also not able to predict exactly what the maximum capability of VER will be at different horizons ahead.
- Non-synchronous: VER, particularly variable-speed wind turbines and photovoltaic solar, are connected to the grid through an inverter, rather than a synchronous generator, as is the case with most thermal and hydro generation. Thus, VER do not provide inherent synchronous inertia to the system, nor can they sense frequency deviations through the speed of the turbine. Without additional controls, this means VER do not provide an automatic response to frequency deviations.
- Location: Utility-scale VER are typically located in windy and sunny parts of the country with significant land. These locations are often far from load and often in weak portions of the grid and require long-distance transmission. Other VER are located on the

[1] www.awea.org. Numbers are based on the end of 2015.

[2] www.solarelectricpower.org. Numbers are based on the end of 2015.

distribution system. These sources, although close to load, are often difficult to observe and control by the transmission system operator.

We describe a few of these characteristics in more detail. Figure 1-1 presents a 250 MW capacity wind plant power output for a one-day period. The output appears to vary substantially throughout the day, producing from zero to near its capacity and back. The variability is defined as the output changing through time. The variability of VER output is combined with the variability of the rest of system conditions, including the load, and other VER on the system, so that the overall variability that operators must manage is lower than the sum of each individual source of variability. The impact of variability can have adverse impacts on system reliability because the system resources must ramp up and down and change their commitment status at different time periods to accommodate this variability. If insufficient ramping capability is available, the system is at risk. This variability occurs at different timescales, from seasonal variability which can impact maintenance outages and hydro operation, to very-short-term variability which can impact system frequency control.

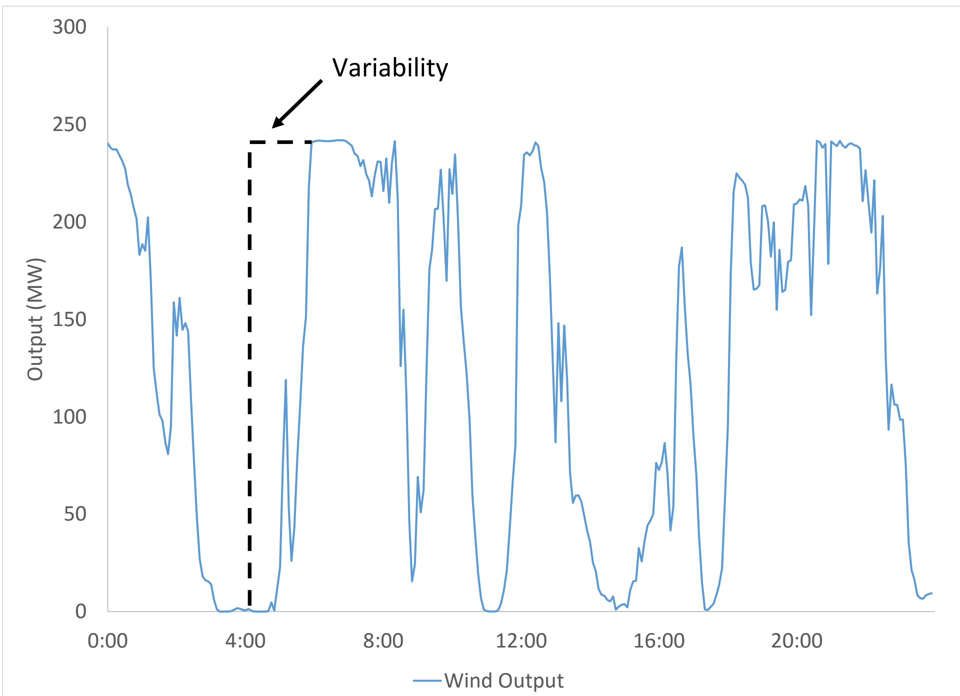


**Figure 1-1**
**Variability is the difference between output from one time period to another.**

Figure 1-2 shows a daily solar profile, including actual power output and day-ahead forecasted output. In this case, the forecast expected a clear sky, whereas in real-time, the output was much lower and more volatile due to increased cloud cover. Uncertainty is defined as the difference between expected output and actual output. The uncertainty of VER output can be combined with the uncertainty of the rest of system conditions, including load and generation availability. The impact of uncertainty can lead to potential reliability issues because the system resources are typically scheduled in advance of real-time and will be scheduled based on conditions that are different than their actual outcome. Many thermal plants require advance time to start and commit to being online, often several hours up to a day or few days ahead. With greater uncertainty, the commitment of these resources may be either more or less than what is truly needed, leading to both reliability and efficiency impacts. Shorter-term forecast uncertainty, like those of an hour-ahead up to several minutes ahead, can cause impacts as well, as the system resources are dispatched to energy schedules that are different than what may be required.

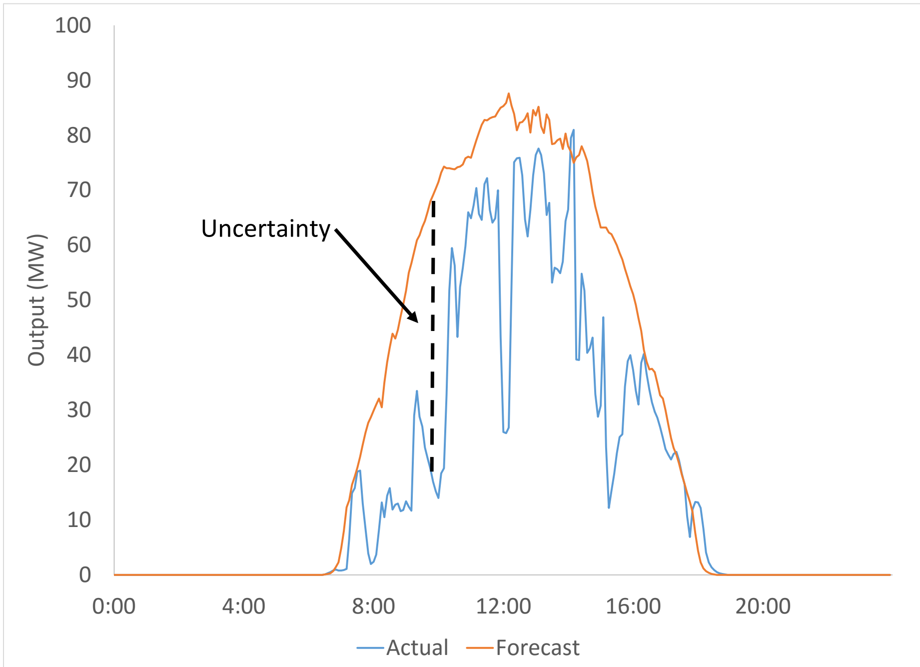


**Figure 1-2**
**Uncertainty is the difference between forecasted output and actual output.**

Figure 1-3 shows a system frequency following a loss-of-supply event (e.g., a large generator). Synchronous machines influence the rate of change of frequency directly following the event, with the frequency nadir (minimum point) occurring when the amount of primary frequency response coming from the turbine governor control, as well as load response, match the loss of supply event. Eventually the response reaches a new settling frequency. The chart shows the result of system frequency on the same system with the same loss-of-supply event, but with increasing percentages of non-synchronous resources (VER). Without any additional controls, increases in VER penetration will lead to displacement of synchronous resources, resulting in a higher rate-of-change-of-frequency, a lower frequency nadir, and a lower settling frequency. All of these conditions can lead to potential issues such as under-frequency-load-shedding (UFLS) relays, generator damage, and even potential for a wide-spread blackout.

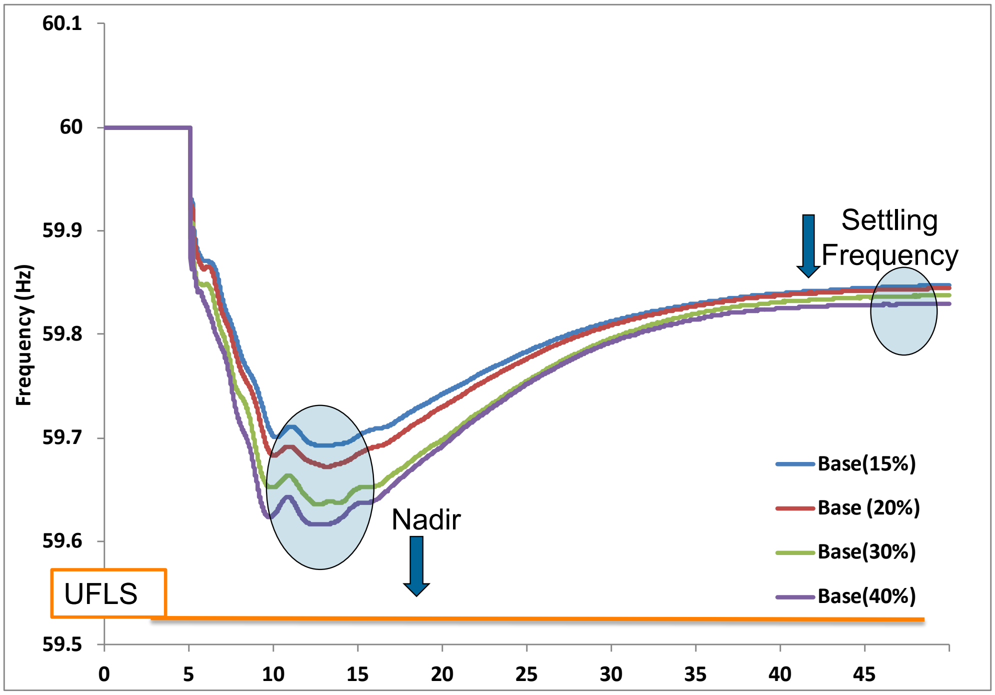


**Figure 1-3**
**System frequency following a loss of supply event at varying penetrations of non-synchronous generation. With more non-synchronous generation without additional controls, the frequency nadir decreases, eventually impacting the triggering of under-frequency load shedding.**

Finally, when large utility-scale VER are connecting to the system, they are typically connecting at locations that are far from load, either further loading up the transmission network, or requiring new transmission build-outs. This can impact the overall system congestion by causing greater potential for overloads, and can increase system losses as well. It also may lead to stability issues with long-distance transmission. On the other hand, distributed energy resources (DER), located close to load within the distribution network, have limited observability and controllability. In the current paradigm, the transmission system operator often does not have information on the current conditions of DER, and may not know how DER is impacting what is seen as load at the transmission level. In addition, the system operator typically does not have much control over DER either, and often cannot direct DER to provide reliability services to help the transmission system.

While many of the above characteristics have negative impacts on power system reliability, numerous strategies are also being evaluated such that these impacts have less impact and even such that VER can help provide reliability services to support power system reliability. Geographic diversity, by having VER spread out across a larger geographic area, can reduce variability impacts compared to a case where all resources are located in a concentrated area [1]. VER forecasting, where specific VER forecasting companies provide system operators with a prediction of VER output using meteorological information and statistical techniques, can greatly reduce the uncertainty of VER [2]. Finally, most modern VER technologies can have frequency control capabilities installed such that they can provide control very similar to the response seen from synchronous generation [3]. These strategies are just as important as those that involve changing and increasing requirements from the ancillary service products that are needed.

Other emerging technologies have unique characteristics as well. Demand response has emerged in many regions as a resource to provide energy during certain conditions as well as ancillary services. Demand response can respond quickly, but may have unique constraints on how often it can be called, how long the response can be sustained, and whether it will continue to be available into the future, depending on the technology. There has been a significant amount of energy storage technologies procured in certain regions in the United States and elsewhere. Energy storage can be used as a generator and a load, where energy stored can be generated at later time periods when needed, but at an efficiency loss. It also typically has extremely fast response capabilities, but may be limited in how long it can sustain energy output for. These technologies, and many other emerging technologies can provide tremendous benefits and contribute to the various types of ancillary services. They also can have an impact on the way that the ancillary service needs are determined and how the ancillary service procurements and markets are designed.

All of these characteristics can have an effect on system reliability and system economic efficiency. Ancillary services are those services beyond the provision of energy that are necessary to support power system reliability, and are attained by generating technologies and sometimes others (e.g., transmission technologies, demand response). These services combine with the provision of energy to ensure that load is being met constantly and that the system will not be put at risk following credible events. Section 2 provides an overview of some of the existing ancillary services that are common across all areas and that have been around for many years. Section 3 then provides an overview of the design of ancillary service markets, especially focusing on North American organized electricity market areas. Section 4 then provides a survey of some proposed changes to the existing ancillary services and some new ancillary service products and ancillary service market design changes. Section 5 provides a summary and concludes the report. The objective of the report is to provide readers with an understanding of the importance of these ancillary service products, including why they may be evolving due to the increasing levels of VER on various power systems or otherwise.

# 2
# OVERVIEW OF ANCILLARY SERVICES

In this section, we provide an overview to the existing ancillary services that are common across most regions. Most ancillary services are in place to support either frequency control or voltage control through the adjustment of active and reactive power. Typically, these services can be separated by this objective. However, a few other services are in place to support power system reliability in other ways.

## Operating Reserve for Active Power Control

Operating reserve defines the active power capacity that is held above (or below) the energy schedule of the system resources, to be used in case of an event or condition that occurs after the schedules are given [4]. Different operating reserve are needed for different reasons and terminology differs from region to region (see [5]). Figure 2-1 shows examples of operating reserve types with some of the categories based on common existing services while others are new or evolving services, such that operators may not necessarily have seen a need for these until recently. The former will be discussed in this section, while the latter will be discussed in Section 4. Operating reserves are typically held at a balancing area level, although it also can be shared across multiple balancing areas. Balancing areas are regions that contain generation, transmission and/or loads within a metered boundary area that must maintain the balance of generation and load within the metered boundary. In North America, these are managed by a Balancing Area Authority who maintains load/supply balance.

Operating reserve capacity can be defined by its reason for usage. For example, some operating reserve is used for large events, while others are used for normal balancing efforts that are not captured by energy schedules. The speed of response is also a characteristic in the type of operating reserve, some requiring rapid response, while others may require slower yet sustained response. Other characteristics include the direction of response to hold: upward, downward, or equal amounts of both, the technology requirements needed: for example autonomous frequency response capability, automatic generation control, online or offline. Table 2-1 describes these characteristics.

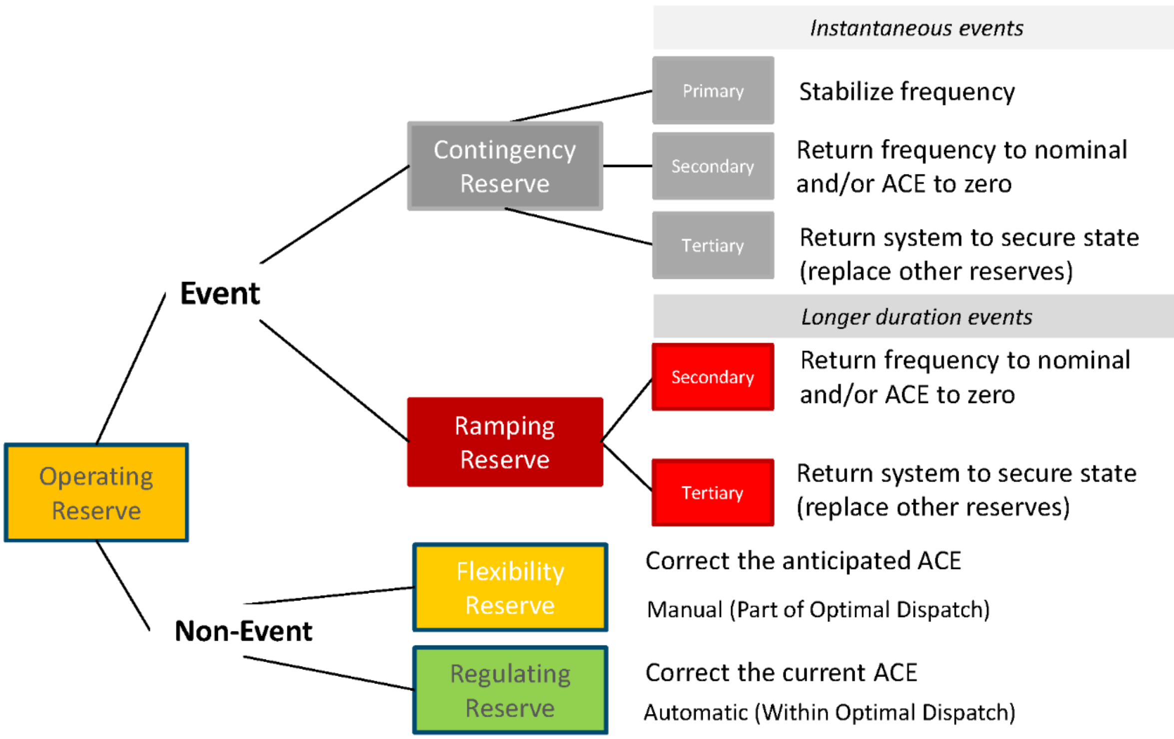


**Figure 2-1**
**Different forms of operating reserve used for different purposes**

**Table 2-1**
**Different characteristics that make up the different operating reserve types**

| **Operating Reserve Characteristic** | **Examples** |
|---|---|
| Condition for Deployment of Reserve | Nonevent, contingency event, renewable ramp event, ACE excursion |
| Reserve Direction | Upward “raise”, downward “lower”, both |
| Reserve Speed | Instantaneous response, non-instantaneous response, automatic control, manual control, response speed, delay allowance, sustainment period |
| Reserve resource status | Spinning, non-spinning, frequency responsive |
| Reserve Need (i.e., what does it help accomplish) | Stabilize system frequency, bring frequency to nominal level, replace other reserve, reduce area control error, reduce price spikes, reduce system costs |

### *Contingency Reserve – Spinning and non-spinning (Event > Contingency Reserve > Secondary on Figure 2-1)*

Contingency reserve is a type of operating reserve that is required for all balancing areas. It is a service required to be held, typically only in the upward direction, in the case of a loss of a large generator (or a large infeed from a neighboring area), in order to make up for the loss. The requirement is usually to have contingency reserve capacity that is greater than the largest unit, or infeed, within the balancing area. This is defined by the NERC BAL-002 as well as the ENTSO-E Policy 1 in North America and Continental Europe, respectively [6-7]. This reserve typically must be able to respond within some time frame to ensure that frequency and area control error (ACE- the difference between an area's net actual and scheduled interchange, taking into account frequency error and meter error correction) is brought back to nominal or to zero, respectively. In most systems, this is typically between ten and fifteen minute response times.

Contingency reserve also is typically split between a spinning requirement, where the resources must be online and synchronized in order to provide reserve based on unloaded capacity, or non-spinning, where the resources can be offline as long as they are able to start-up, synchronize and provide the necessary energy within the contingency time period (e.g., 10 or 15 minutes). A typical requirement in most U.S. balancing areas is for at least 50% of the total contingency reserve to come from spinning reserve.

### *Regulating Reserve (Non-Event > Regulating Reserve on Figure 2-1)*

Regulating reserve, also called regulation, frequency regulation, load frequency control reserve, secondary control, or nicknamed "Reg", is an operating reserve capacity that is held during the energy scheduling period in order to be deployed between scheduling periods for short-term imbalances between generation and load. The reason for regulating reserve deployments is to meet the changes that occur from variability of load, VER, or conventional generation (i.e., when they are not accurately following schedules), or the uncertainty from very short-term forecast errors (e.g., less than one hour ahead) from load or VER. Regulating reserve requirements vary by balancing area and often depend on the level of variability and short-term forecast uncertainty within the region. In North America, NERC's BAL-001 institutes the Control Performance Standards and the Balancing Area ACE Limit [6]. These standards are set forth to verify compliance with balancing using regulation (although not explicitly), rather than directing a balancing area with a specific requirement. In Continental Europe, the ENTSO-e Policy 1 has a specific equation for requirement of regulating reserve, often called secondary frequency control in Europe [7].

$$\mathbf{Regulating\ Reserve} = \sqrt{\mathbf{a} * \mathbf{L_{max}} + \mathbf{b^2}} - \mathbf{b} \qquad \text{Eq. 2-1}$$

Where $a$ and $b$ are constants (typically $a$ = 10, $b$=150), and $L_{max}$ is the maximum demand for the balancing area. Unlike NERC, however, there is no performance compliance standard in the European requirements.

Regulating reserve is online unloaded capacity controlled via automatic generation control (AGC), where the control room will pulse signals to all units that are regulating at time intervals ranging from 2 to 6 seconds. The signals are based on correcting the ACE, which calculates the MW imbalance within the area based on the error in interchange from its scheduled value and the error in frequency from scheduled value. Similar to contingency reserve, regulating reserve also has a fast response requirement. In North American ISO regions, this is typically five minutes, but can be longer response time in other regions. Unlike contingency reserve, regulating reserve is typically held in both upward and downward directions, such that it can response to both over- and under-generation conditions.

### *Primary Frequency Response (Event > Contingency Reserve > Primary on Figure 2-1)*

Primary Frequency Response (PFR), also called governor droop control, or frequency responsive reserve is the automatic autonomous response of increasing (or decreasing) power generation in response to a decrease (or increase, respectively) of frequency. Primary frequency response is required to arrest system frequency to a stable level before triggering of UFLS (see Figure 1-3). The response typically comes from the turbine speed governor response of conventional thermal and hydro plants, but depending on regional definitions can also include demand response, or non-synchronous resources. In North America, the revised NERC BAL-003-1 standard has just recently been instituted for minimum response requirement in MW/0.1 Hz for each balancing area [6]. In Europe, Policy 1 institutes a MW requirement (3,000 MW) for the interconnection and a MW and MW/Hz requirement for each TSO [7]. Each have some differences in how the MW/Hz requirement is calculated and what other requirements are included (insensitivity, speed of response, and frequency deviation where full response is required).

### Voltage Control and Reactive Power Support

Voltage control effectively requires the control of reactive power from generating units and transmission assets (e.g., capacitor banks, static VAR compensators, etc.). Voltage must be kept within between 5 and 10% of their nominal levels. Because of the nature of the transmission system, reactive power cannot be provided through far distances, such that the requirements are very localized. Transmission system operators typically require generators to provide reactive power support within specified ranges while in voltage control mode, such that the plant's reactive output is controlled to maintain a specified voltage level at the generator terminals or point of interconnection.

Similar to operating reserve, voltage control can also be further categorized by the speed of response. For example, during steady-state operations, voltages must be kept within normal limits such that reactive resources are committed and scheduled to levels where they can provide reactive support in advance. Then, dynamic reactive support requires an immediate control of reactive power to keep voltage levels within emergency limits following a contingency event (e.g., loss of line or generator).

### Black Start Service

Black start resources are defined by NERC as "generating unit(s) and its associated set of equipment which has the ability to be started without support from the system or is designed to remain energized without connection to the remainder of the system, with the ability to energize a bus, meeting the Transmission Operator's restoration plan needs for real and reactive power capability, frequency and voltage control, and that has been included in the Transmission Operator's restoration plan" [6]. Though black start service is seldom needed in practice, having sufficient resources that can be started without external power sources that can maintain voltage and frequency while load is energized is essential. It is critical that these resources can be relied upon to restore power after a complete or partial black out as quickly as possible. Because of the characteristics of this service, black start resources are typically combustion turbines or hydro generation.

## Long-term Installed/Unforced Capacity

Long-term capacity, though not typically classified as an ancillary service in the usual definitions, is a product secured in many regions separate from the provision of energy. From a long-term perspective, planners require a certain amount of available capacity within their system or external resources with import capability, such that they can serve load consistently. A reserve margin is used to set the amount of capacity that should be installed above the peak load. In most regions, this requirement is typically set at a 12-18% reserve margin above the expected peak load. The reserve margin is often based on the amount of capacity that will lead to having less than 1 day of involuntary load shedding across 10 years.

## Ancillary Services: Technology Contributions

Each of the above services can be provided by a multitude of different resources and technologies. However, not every technology can provide each service to the same level. Some technologies may have faster response times for active power operating reserve, and some technologies may have different capabilities and equipment needs for providing frequency control, voltage control, and black start service among other services. Figure 2-2 shows an example of how different technologies may provide different services at different levels with the colors representing zero as low (no provision) to a score of five as high (very good provision) [8]. Note that the numbers here are based on a number of assumptions and used mostly for example purposes, as many different versions of these technologies may be able to provide these services at differing levels of quality.

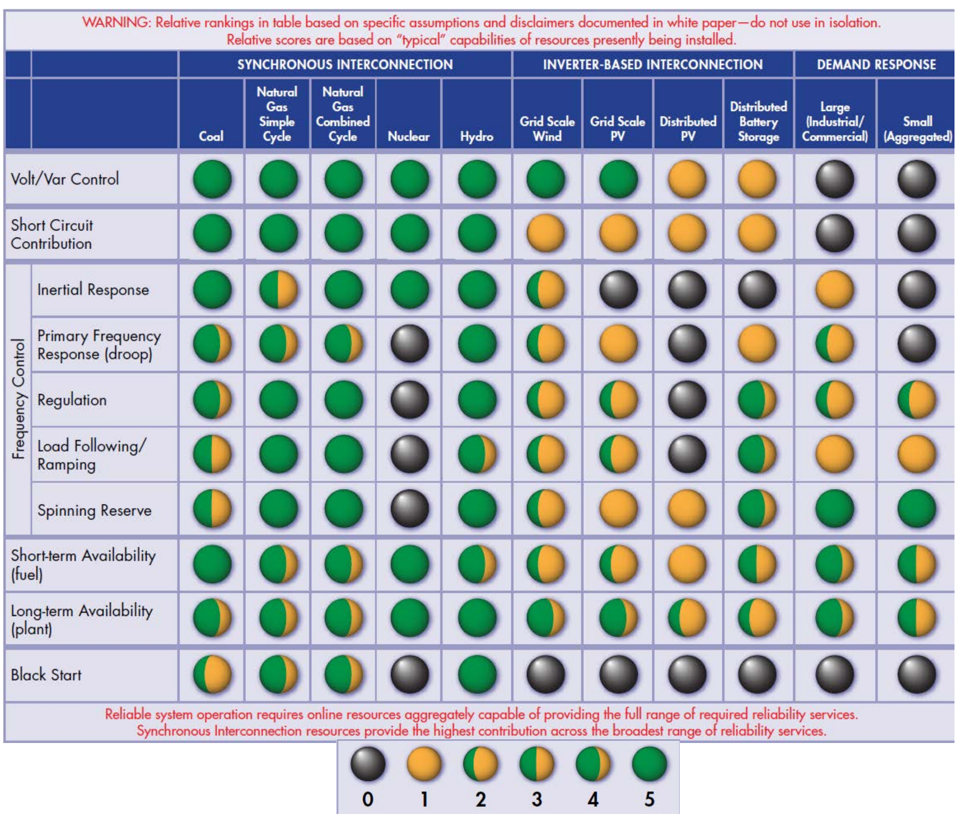


**Figure 2-2**
**Example of the capability of different resources contributing to various services [8].**

The level at which different technologies provide these services also may depend on the current requirements or regulations, compensation schemes in place, and the variation in costs that technologies may incur in providing those services. Regulations may guide which technologies in which resource must provide a service and the level at which it must provide that service. For example, many regulations require that historically conventional generating resources provide between 0.95 leading and 0.95 lagging power factor in the provision of reactive power [9]. Compensation schemes, either the rate at which a utility or balancing area pays for the service, or the prices at which an organized market sets for the service also provide the incentive for the level of service a resource may provide. This is also related to the different costs that a resource may have for providing each service. When the payment is higher than the costs, it makes sense for the resource to provide the service. More on different compensation schemes and market designs for ancillary services are discussed next.

# 3
# ANCILLARY SERVICE MARKETS

In regions with vertically integrated utilities that also act as the balancing area authority, specific rates are given for certain ancillary services. In organized electricity markets, there typically exists separate auctions or cost-based revenue rules for a select set of ancillary services. The following lists some of the current ancillary services as defined by the Federal Energy Regulatory Commission (FERC) and how it is typically compensated in organized electricity markets [10].

- **Scheduling, system control and dispatch**: Provided by the RTOs/ISOs as they schedule and control the resources on their system; not necessary applicable to ancillary service markets.
- **Reactive supply and voltage control from generation service**: Generally supplied as a static cost-based service without any competitive prices.
- **Regulation and frequency response service**: Regulation is typically supplied and priced by auction-based markets in RTOs/ISOs that update hourly or more frequently; it is used to correct area control error. However, frequency response, defined as the local droop response of governors autonomously responding to frequency, is generally not included in any dynamic markets nor is it given cost-based rates.
- **Energy imbalance service**: Typically, the service of the real time market (RTM) correcting the imbalance from the forward or day-ahead markets (DAM) and therefore a component of the real-time energy markets.
- **Operating reserve—synchronized reserve service**: Typically supplied and priced by auction-based markets that update hourly or more frequently in RTO/ISO regions.
- **Operating reserve—supplemental reserve service**: Typically supplied and priced by auction-based markets in RTO/ISO regions that adjust hourly or more frequently.

The ways in which ancillary service markets are designed and prices and schedules determined vary from market region to region in the United States. In areas where competitive markets do not exist, rates are set for the above services such that resources providing the service can earn revenue for doing so. Many international balancing areas, like most of Europe, while having restructured markets, often have similar processes for ancillary services as non-restructured U.S. areas. These countries will have rates for ancillary services based on anticipated costs. Markets in Australia and New Zealand are more similar to those in the U.S. In this section, we will mostly focus on the ancillary service market designs of the U.S. market regions.

## Market-based ancillary service products

Operating reserve is the ancillary service that is most commonly procured through competitive auction-based market mechanisms. Operating reserve ancillary services are sold in auction-based markets similar to energy markets, where the least bid-based costs are selected to meet the requirements and prices are based on the marginal cost to provide that service. The operating reserve types that have existing competitive markets—including synchronized reserve (contingency reserve—secondary), supplemental reserve (contingency reserve—tertiary), and regulation (regulating reserve)—are bought and sold in either day-ahead markets, real-time markets, or both in a similar manner to energy markets. In fact, the U.S. markets that have ancillary service markets currently co-optimize energy and operating reserve when clearing either the DAM, RTM, or both markets. This means that the markets are cleared simultaneously so that costs and requirements of both markets are considered when clearing the entire market.

A number of other features are fairly common in these market-based ancillary service products. The prices of ancillary services are based on the marginal cost of capacity reservation, and not based on the use of that reserve. In other words, the price is based on the cost of the marginal reserve provider to increase its reserve capacity, and is not based on the marginal cost of being asked to provide that reserve and convert it to energy. When called, the resources will then be paid the energy price for any reserve capacity that is converted to energy.

One of the additional components that are included in co-optimization of energy and ancillary service markets is the lost opportunity cost (LOC). The LOC is the cost of a resource's foregone profit in the energy market (or other ancillary service product), when needing to provide that ancillary service. For example, if a unit that costs $30/MWh for energy has its energy schedule reduced so that it can provide additional reserve, and the energy price is $40/MWh, that unit would have a LOC of $10/MWh. This is because the unit has lost the opportunity to make $10/MWh of profit over its costs in the energy market. This LOC of the marginal reserve provider would be included in the ancillary service price that is paid to all ancillary service providers. Therefore, the resources can be indifferent to what service they provide and provide the service that is most critical to the system operator.

Other common features in ancillary service markets include pricing hierarchy, shortage (or scarcity) pricing, and market power mitigation. Pricing hierarchy refers to the design of multi-product markets, where each product may be a higher valued service that others competing for the same capacity. For example, spin (online) reserve is generally considered more valuable than non-spin (offline) reserve. In these designs, pricing hierarchy makes it so that price for the higher valued service is always greater than or equal to the lower valued service. This ensures that the market participants providing ancillary services will have the incentive to provide the service that is most valuable to the system operator and for reliability. Shortage pricing refers to administratively-set prices when the system has insufficient ancillary service supply. For example, when the system is unable to commit enough resources to meet the spinning reserve requirement, there is no marginal provider, and the price will be set by the shortage price. Shortage prices vary by ISO and by ancillary service product. Regulating reserve may be in the few hundred dollars per MW-h range, whereas contingency reserve can be greater than $1,000/MW-h. Importantly, these shortage prices also impact the energy prices that are received by the entire energy-producing market due to co-optimization of energy and ancillary services. Finally, market power mitigation, which is in place for resources that have market power over energy market prices, can also be applied for ancillary services. Because there is less locational requirements for operating reserve and thus less risk of market power potential, market power mitigation tends to be less strict for ancillary services when compared to energy markets.

Table 3-1 shows the average ancillary service prices for various ancillary service products and regions for 2014 [11]. The California ISO (CAISO), Electric Reliability Council of Texas (ERCOT), Midcontinent ISO (MISO), Pennsylvania Jersey Maryland RTO (PJM), and New York ISO (NYISO) are included. It can be observed that on average regulating reserve is typically the highest priced service on average, followed by spinning contingency reserve then non-spinning contingency reserve. These prices vary year to year and hour to hour. For instance, some of the averages hide the fact that many of the ancillary service prices can range from many hours of prices at zero to a few at the shortage price of over $1,000/MW-h.

**Table 3-1**
**Average ancillary service prices for various products for 2014 [11].**

| Region | Regulating Up ($/MW-h) | Regulating down ($/MW-h) | Spinning reserve ($/MW-h) | Non-spinning reserve ($/MW-h) |
|---|---|---|---|---|
| CAISO | 5.41 | 3.90 | 3.34 | 0.14 |
| ERCOT | 12.48 | 9.77 | 14.15 | 5.48 |
| MISO | 11.24 | * | 2.58 | 1.34 |
| PJM | 43.70 | * | 4.21 | 0.95 |
| NYISO** | 13.76 | * | 4.07 | 0.49 |

*MISO, PJM, and NYISO regulation products are bi-directional with the same price for upward and downward service. ** Prices for spin and non-spin reserve are based on NYISO West zone, whereas the NYISO East zone is generally about 50-200% higher.

## Cost-based ancillary service products

Other ancillary services are not sold through competitive markets but are provided cost recovery based on certain pre-determined, non-competitive rates. For example, reactive supply and voltage control and black-start service generally do not have competitive markets. Neither of these services have prices and schedules that are market-based, and methods to determine the prices/schedules vary by ISO. Table 3-2 describes some of the most common reasons why cost-based recovery rather than competitive market-based pricing is used for some ancillary services.

**Table 3-2**
**Reasons why ancillary service markets for some ancillary services may not be a good option.**

| Reasons why a market product may not be justified | Example |
|---|---|
| Too complex to design | Volt/VAR support |
| too specific to certain local areas (little to no competition) | Volt/VAR support |
| System inherently has more than sufficient amounts of the service | Synchronous Inertia |
| Costs for the service are small, so cost of administrating market product may be overkill | Black start (restoration) service |
| A specific resource requirement rather than a system-wide need | Low Voltage Ride Through |

For these services, the individual market participant, the ISO, and in many cases FERC will agree on the rate for the particular resource and service. These may include both capability payments for the ability to provide the service, and provision payments, for actually being called upon and providing the service. Capability costs include the fixed capital cost of equipment needed to be able to provide the service. For example, excitation equipment and automatic voltage regulator costs are included in recovery of reactive support service. Training costs are often recovered for black-start service. Some initiatives are looking at whether there is any benefit or possibility to incorporate more competitive processes for some of these services.

### Markets for long-term capacity

Although we are not commonly referring to long-term installed capacity as an ancillary service in practice, it is important to briefly discuss the ways in which this service is incentivized in market regions. Installed capacity is procured in different regions by different mechanisms. The four most common mechanisms include the following:

- *Energy only markets*. Here, there is no additional payment for long-term installed or unforced capacity. These markets tend not to have any target reliability planning reserve margin. All revenue is expected to come through the energy and short-term ancillary service markets. Often, shortage pricing that triggers during reserve shortages is an important revenue stream for peaking resources to recover capital costs. Examples of this design include ERCOT, the Australian National Electricity Market (NEM), and the Alberta Electric System Operator (AESO).

- *Administrative capacity payments*. In this design, some administratively set payment is provided to resources to secure long-term capacity and to meet reliability planning reserve margins. These payments are determined through various means. The payments do not have any competitive process for what resources to build to meet the reserve margin. Rather, they provide additional revenue beyond that of the energy and ancillary services markets in order for resources to recover their fixed capital costs. Examples include Spain and Ireland.

- *Bilateral agreements*. Many market regions that do not have capacity markets still have reserve margins for the various load serving entities to meet within the ISO. These load serving entities will have some bilateral agreements and payments with capacity suppliers in order to secure the capacity to meet the reserve margins in advance. Examples include Southwest Power Pool (SPP) and CAISO.

- *Centralized capacity markets*. Centralized capacity markets are in place that operate similarly to energy markets, where the ISO selects the least cost set of capacity bidders to meet the reserve margin. These markets typically take place either a few months or a few years ahead of the planning time frame. Prices are based on the marginal capacity supplier as well as a capacity demand curve. Examples include ISO New England (ISO-NE), NYISO, and PJM.

# 4

# NEW AND EVOLVING ANCILLARY SERVICES

In this section, we describe how some of the various ancillary services and ancillary service markets are evolving and some new services that have been introduced. This may include new services that were never necessarily needed previously, services that were needed but never explicitly called out as they were being provided inherently by the incumbent technologies of the system, and those services that were always needed but may not have had market mechanisms or regulations to procure the services from individual resources. In many cases, the requirements of existing ancillary services are also evolving. While the focus of this report is on the changes to ancillary services that are occurring because of the increased penetration of VER and the characteristics of VER, we also discuss how some of the ancillary services are evolving for other reasons. For example, reasons may include other new emerging technologies, changing reliance on different technologies (e.g., more gas), computational improvements in the software used to schedule the power system, and market efficiency improvement goals.

Numerous changes to the various ancillary services discussed earlier are occurring around the world. Throughout this section, we discuss some of the overall themes and new or evolving services and the motivation behind why these services are now being recognized or introduced. We then provide a few examples of certain regions that have been leading efforts on introducing or modifying these services. These set of areas that are included are in no way exhaustive of the world-wide changes that are occurring, but provide the reader with a better understanding of how these changes are occurring in practice.

## Ancillary Service Redesigns

A few entities and balancing area regions have begun or recently completed a full redesign of the set of ancillary services that they consider within their region. In these cases, either new ancillary services were added to the set that they consider necessary for maintaining reliability, or existing services are further separated to better align actual needs with service definitions. In some cases, this includes an expansion, where the number of recognized ancillary services has increased substantially. In other cases, it may simply be a renaming or rebranding of the existing services to put greater emphasize on how the needs are evolving. We focus on three entities: the North American Electric Reliability Corporation's Essential Reliability Services Working Group and its "essential reliability services", ERCOT's future ancillary service redesign, and the Irish TSO Eirgrid's DS3 program.

### *North American Electric Reliability Corporation (NERC) Essential Reliability Services Working Group*

The NERC Essential Reliability Services Task Force was started in 2014 to examine the evolving reliability risks in power system operation, determine how these risks should be tracked, and how these risks might be evaluated. This task force, which evolved into the Essential Reliability Services Working Group (ERSWG) in 2016, consisted of members from NERC, utilities, ISOs and researchers. The group reported in late 2015 with the Measures Framework Report and a set of educational videos aimed at describing a number of different essential reliability services [12]. These include services that are currently included in the ancillary services defined by FERC, but also focus on other areas. In late 2016, the group will be releasing a set of "sufficiency guidelines" that describe how the relevant entities (system operators, planners, etc.) should assess whether there are sufficient resources providing these services in a given area [13]. This included detailed description of data and methods to assess the different services described below.

The aim of this working group was not to define new ancillary services, as much as define methods to track particular essential services. It also doesn't define particular values for sufficiency, as much as the process for assessing the availability of services. However, the issues investigated by this group may also inform future market developments. Among the services investigated were:

- *Frequency Support*: This includes both inertia and frequency response, and recognizes that with new resources such as VERs, system inertia may be reduced, but new resources may be able to replace this if equipped with the right controls. Four measures are proposed for this area to examine the sufficient of frequency support. This includes a measure that determines the minimum level of synchronous inertia needed and a measure of the amount that would actually be available given the actual commitment of resources. These are interconnection wide, to benchmark the current year. The third measure looks at synchronous inertia in future years, and whether these are close to the minimum amount identified in the first two measures. The final measure is about evaluating alternatives (e.g. the use of VERs, energy storage, etc.)
- *Ramping*: With increased variability due to VERs, there may be concerns about insufficient ramping over multiple hours. The process proposed starts with pre-screening for identifying whether there are hours when dispatchable resources make up a small portion of overall system resources during times when there are large ramping requirements. If there are, further studies are required to study the hourly performance of the Control Performance Standard (CPS1) score.
- *Voltage Support*: Voltage and reactive power control needs to be considered over different conditions related to increased VER penetration. However, as this is a very local issue, the main aim in the NERC report was to describe a means to identify areas where steady-state and/or dynamic reactive power provision may be insufficient. These can then be studied by the relevant local reliability entities based on a number of studies described in the report.

In any of the three areas described, if there is a sufficiency concern, there will be a need to alter operating practices. This may involve increasing ancillary service requirements, adding new services, requiring capability of these services, or other potential means to incentivize these services, such as providing competitive prices for resources that provide the service.

The ERSWG will also release a report on Distributed Energy Resources (DER) in 2016. The aim is to identify and describe planning and operating concerns related to increased penetration of resources on the distribution system. As NERC is focused on bulk system reliability, the main aim of this was to identify how planning and operating practices, and eventually standards, may need to evolve to recognize the unique characteristics of DER. This included aspects such as data needs and modeling approaches for planning, and operational information needed to ensure system operators have the visibility and control required to maintain bulk system reliability.

DER will alter the value of different services and these resources may eventually be providing ancillary services themselves. In order to do this, data and models will need to capture DER characteristics in enough detail to reliability plan and operate the system. As such, the ERSWG and follow-on work is likely to inform future ancillary services market development. It also means new tools and business processes may be needed in operations.

### *Electric Reliability Council of Texas*

Over the last few years, ERCOT has been going through a large-scale redesign of its ancillary service products and markets. ERCOT introduced the future ancillary services team, or FAST, which sought to look at the current set of ancillary services and whether the definitions and types are still valid on today's system, as the system where the existing ancillary services were designed was mostly based on one with large steam generators as the predominant generation type [14]. The evaluation of the revised ancillary services is therefore mostly motivated by new resources, which have different characteristics and performance from traditional units. ERCOT now has about 16,000 MW of installed wind capacity within its region, with a peak load of about 70,000 MW. It has had an instantaneous penetration of close to 50% of its load being provided by wind generation[3].

Figure 4-1 shows the current ancillary services and those proposed in the redesign. While regulation up and down services are relatively unchanged, the proposal may include separating out traditional regulation from fast responding regulation as different products. Responsive reserve, which refers to any online capacity used to restore system frequency, is proposed to be disaggregated to five distinct ancillary services. Fast frequency response refers to the immediate response to frequency deviations. It is separated further by the trigger point, and would be typically provided by load resources or any technology that can provide a full response within 0.5 seconds following the frequency event. Primary frequency response (PFR) is also separated from contingency reserve (secondary reserve). PFR would respond proportional to system frequency, quickly after the event, but slower than FFR, to stabilize system frequency. The new service of contingency reserve, used to restore frequency to nominal level, would then be divided into reserve capacity that is dispatched through the system's economic dispatch or whether it responds through manual action. Non-spin reserve is similarly divided by whether it is dispatched through economic dispatch or manually. Last, synchronous inertia is a new service that is part of the new design that would use only synchronous generators to support reducing the rate of change of system frequency instantly after the frequency event occurs.

---

[3] On March 23, 2016, 48.28% of load was being supplied by wind power. See http://www.ercot.com/gridinfo/generation/windintegration/.

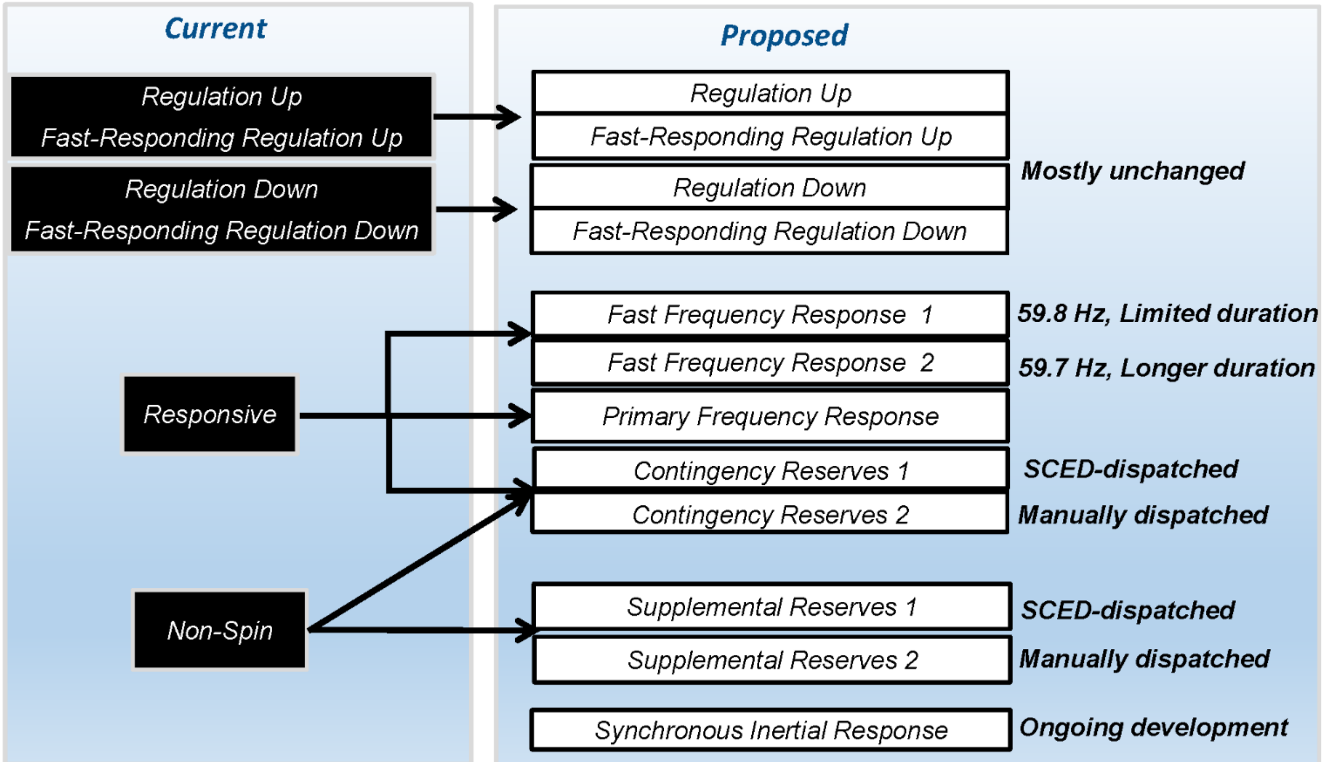


**Figure 4-1**
**ERCOT's current and future ancillary service products.**

Each of these services will have requirements and each will have ancillary service markets. The requirements may be specific to the one service, or they may be combined with others. As an example, the requirement for PFR may depend on how much FFR and synchronous inertia the system has. The ancillary service market design would be similar to existing ancillary service market designs for each of the different services. Resources would bid the cost of the service and the quantity in which they can provide. ERCOT would select the least-cost set of resources to provide each service using co-optimization with the energy market in the DAM. Prices from each service would be based on the marginal cost of providing each service. Though, some complexities are also involved in the pricing that can impact prices of services that have relationships with other services.

Although the overall proposal for future ancillary services was voted down by stakeholders in 2016, ERCOT and some stakeholders are still reviewing the possibility of specific services and designs going forward.

### *Eirgrid*

The Irish power system, which consists of both the Republic of Ireland and Northern Ireland, is a synchronous island system, with only DC connections to the rest of the European continent. With over 20% of energy coming from wind, it has one of the highest penetrations of non-synchronous penetration, if not the highest, in the world. As such, the system operators (Eirgrid and System Operator of Northern Ireland, or SONI) setup the "Delivering a Secure, Sustainable Electricity

System" (DS3) program.[4] One of the key aims of that program is a redesign of the system services procured and deployed to maintain reliability [15]. The program also focuses on grid codes, operational tools and interconnection requirements, which are not described here.

System Services in Ireland are defined as those, other than energy, that are required for the continuous, secure operation of the bulk system (and are thus very similar to FERC-defined Ancillary Services). The redesign initiated under the DS3 program has resulted in a number of new System Services, as well as alteration to the size of the System Services compensation pool. The redesign was motivated by a number of reasons, including the need for recognizing the value of new technologies, ensuring existing services providers will receive appropriate remuneration, and ensuring that system operators will have sufficient reliability services available with increasing amounts of energy coming from renewables.

Under the proposal, the number of services has grown from seven to fourteen. From October 2016, an interim period was established to examine the newly proposed arrangements and ensure the process could be managed in an efficient, fair manner. These arrangements include a tariff for each of the new services, and a performance scalar to incentivize greater performance by the providers of the services [16]. The tariffs are calculated in a manner different than the method used for most of the services provided in the US. The concept used in Ireland is that they start with a 'pot' of compensation, roughly equal in size to the existing pot used for the current services. Then, a pot is calculated for each of the system services, based on the relative value of each of these services, and a tariff is determined based on the pot size and volume required. The relative size is then adjusted until the total pot size is close to the existing pot of funding for compensating ancillary services. Once this total pot has been determined, service providers can bid on each service and are paid based on the volume procured and the tariff.

While the means to compensate the providers is thus quite different than the US, the set of actual services are similar to those of NERC, ERCOT, and other U.S. entities. Existing services used in Ireland include the following:

- Operating Reserve (Primary, Secondary, Tertiary 1 and Tertiary 2), which ensure the system can maintain frequency within desired bounds, and restore to nominal after an event on the system. Different categories are used for different response times.
- Replacement Reserve (Synchronized and De-Synchronized), which are used to restore the operating reserves back to their desired level.
- Steady State Reactive Power, which is used to ensure that resources on the system can provide steady-state reactive power control, and both synchronous and non-synchronous resources can provide this.

These existing services generally stay the same under the new DS3 arrangements, although replacement reserve time frames are reduced to avoid overlap with the ramping product described below, and reactive power product was restructured to provide reactive power across the widest possible active power range.

New services that are being procured include the following:

[4] More information is available at http://www.eirgridgroup.com/how-the-grid-works/ds3-programme/

- Synchronous Inertial Response, which is a response that is immediately available from synchronous generators, synchronous condensers and some synchronous demand loads (when synchronized) because of the nature of synchronous machines and is a key determinant of the strength and stability of the power system.
- Ramping Margin (1, 3 and 8 hours) is the margin that a resource can provide to manage longer term ramps with a good degree of certainty. One, three and eight hours were determined to be important ramping horizons based on system operator analysis.
- Fast Frequency Response, which is MW response from synchronous or non-synchronous resources faster than Primary Operating Reserve, available within 2 seconds and sustained for at least 8 seconds.
- Dynamic Reactive Response, which is a reactive current response to voltage dips on the local network.
- Fast Post-Fault Active Power Recovery, which are units that can recover their MW output quickly following a voltage disturbance.

As expected, most of these new products are primarily driven by the fact that wind power is variable, uncertain and asynchronous to the system. However, they are also proposed to recognize that new types of resources (demand side, battery or flywheel storage, and wind plants) can provide certain services well, but not in the same manner as traditional synchronous generation. Note that, at present, while a tariff has been proposed, the last three of the above products are not being procured under the interim arrangements, as more testing and verification needs to be done to determine how these are being provided by different resources. Also, wind and other newer resource types are still being tested for inclusion in most of the ramping and reserve services, rather than participating fully. Both of these issues will likely be resolved in the next phase, when the interim arrangements become permanent.

The interim arrangements are being used to test how the system services redesign is working. It is expected that, with experience, some of these services may evolve. This could be in the size of the tariff, the requirements for those who are cleared to provide the service, or it may even require additional products. The fact that the overall size of the ancillary services compensation has increased relative to energy and capacity is also illustrative of how services become more important with increasing renewable penetration.

### Changing regulating reserve requirements

In most regions in the United States, regulating reserve requirements are based on a specific policy that historically did not incorporate any impacts of VER. The requirements may be a fixed level that is constant for all hours and days, a level that is different for on-peak hours and off-peak hours, or a variable requirement that depends on hour of day, season, and weekday vs. weekend. Thus, these requirements do not necessarily depend on the anticipated impact of VER.

The need for regulating reserve can be impacted by the variability and uncertainty of VER. We first explain what causes the need for regulating reserve. Regulating reserve is used to correct the ACE that occurs when either load is different from its dispatch forecast, or generation is different from the schedule determined from the economic dispatch procedure. In non-restructured regions, the economic dispatch may last for fifteen minutes to an hour. In ISO regions, the economic dispatch procedure is done between 10-20 minutes prior to the dispatch interval and lasts for a five-minute interval. In the latter case, the load and VER forecasts are thus made 10-20 minutes in advance, and schedules of other supply resources are determined based on those forecasts. They are either constant, or linearly interpolated across the five-minute interval length. These 10-20 minute ahead forecasts cannot be perfectly accurate, and thus regulating reserve from resources that are on AGC must be used to correct for that uncertainty. It is also not likely that the load, VER output, or conventional generation output (from resources not providing regulating reserve) can remain constant for the full five-minutes. Regulating reserve must also be used to correct for this variability.

Figure 4-2 shows an example of how a wind plant may affect the regulating reserve need for both variability and uncertainty. In this example, a persistence forecast is used, such that the current average output is used for the 10-minute-ahead forecast. The output then varies within each five-minute interval due to variability. Figure 4-3 shows how the uncertainty and variability each impact the regulating reserve need, where the regulating reserve need equals the difference between actual and schedule, either due to variability or uncertainty.

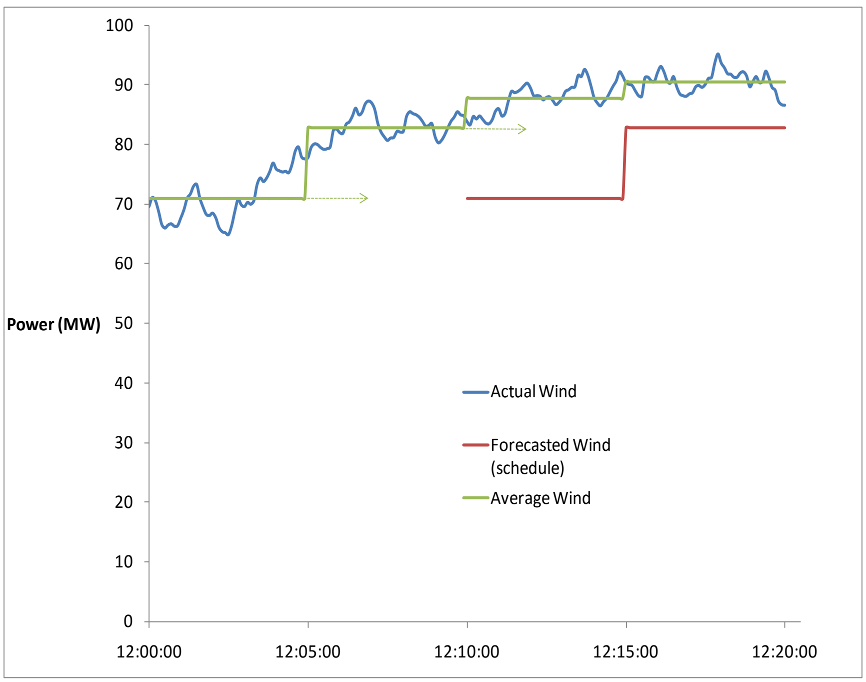


**Figure 4-2**
**Impact of variability and uncertainty on the need for regulating reserve.**

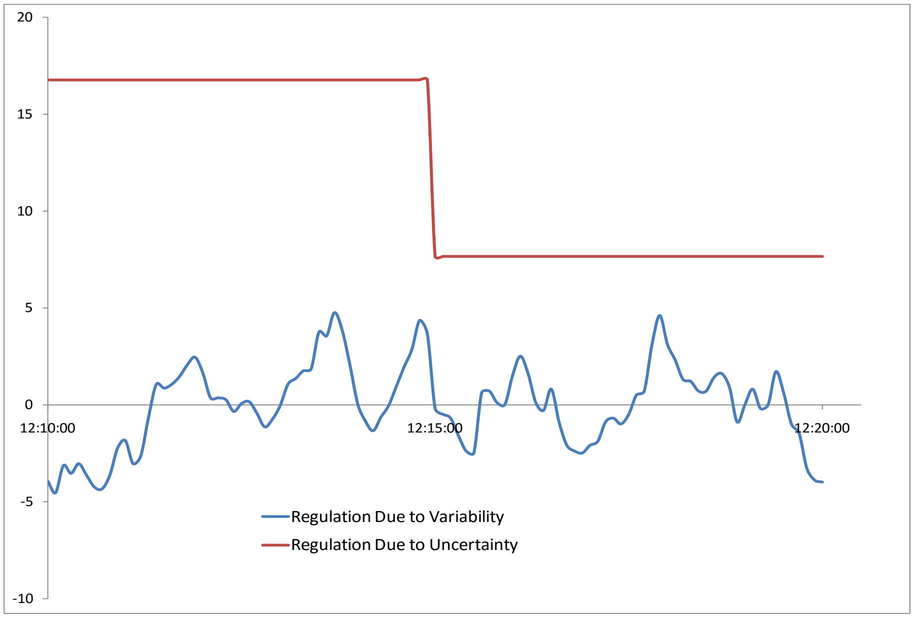


**Figure 4-3**
**Regulating reserve need due to variability and uncertainty.**

Numerous studies have been performed over the last decade to understand the impact of increasing levels of VER on power system operations, including a determination of what integration costs may be incurred. Many of these studies evaluate the increased regulating reserve requirements on the system with increased VER [17]. In some of the earlier studies, the regulating reserve requirement on the increased VER system was determined based on the increased standard deviation of the net load at some pre-defined timescale. Then, based on operating experience, a multiple of standard deviations was used to calculate the total regulating reserve to ensure a high percentile of regulating reserve sufficiency. For example, the first NYISO wind integration study used the standard deviation of net load increase at a six-second resolution, the timescale at which its AGC runs, and determined the requirement to be an increase of 36MW, based on multiplying its standard deviation increase of 12 MW at that timescale by a factor of three [18]. A study in Minnesota found that the standard deviation increased by 2 MW per 100 MW of wind power added to the system, and it multiplied this standard deviation by a factor of five, for increased performance [19]. Since these early studies, more recent studies have enhanced these methodologies by separating out impacts of wind, solar and load [20], incorporating NERC compliance standards to determine the appropriate multiplier [21], utilized ramp, duration, and capacity impacts simultaneously [22], and incorporated probabilistic methods for reserve requirement determination [4, 23]. EPRI previously summarized some recent advances in [24].

Some areas have modified their regulating reserve in practice as well, in many cases based on some of the above-mentioned study methodologies. Most of the areas that have made some changes are those that have significant levels of VER and each make some increase to their requirement based on anticipated VER conditions. We briefly go through a few examples of areas that have made these changes and the methods that they now employ, including ERCOT, Southwest Power Pool (SPP), and Hawaii Electric Company (HECO).

### *Electric Reliability Council of Texas*

As discussed earlier, ERCOT has had a significant amount of wind power on its system over the last several years. ERCOT modified its regulating reserve requirement methodology a number of years ago based on a study that it had conducted in 2008 [25], though it has evolved since then. ERCOT bases its regulating reserve requirement to match the maximum of the following four values: (1) the 95$^{th}$ percentile of actual deployments of regulating reserve from the same month of the previous year, (2) the 95$^{th}$ percentile of regulating reserve deployments from the same month of two years prior, (3) the 95$^{th}$ percentile of net load changes during the same month of the previous year, and (4) the 95$^{th}$ percentile of net load changes during the same month of two years prior [26]. Each of these factors will be increased based on the increased level of installed wind capacity compared to the period for which was studied (e.g., if there is more wind installed in the current month compared to the same month of the previous year) by a specific rate (MW increase per MW incremental installed wind capacity). ERCOT may then make changes based on anticipation of large block schedule changes (i.e., during the 0600 and 2200 hour time periods), when the regulation exhaustion rate exceeds a certain level, or when CPS1 scores were deficient for the last 30 day period.

### *Southwest Power Pool*

The Southwest Power Pool (SPP) also has a significant level of wind generation on its system. At the end of 2015, SPP had 12,397 MW of installed wind generation on its system, with a record production of 10,780 MW and a record instantaneous penetration of 45.1% of load being supplied by wind generation[5]. SPP made some changes to its regulating reserve requirement to account for the increased variability and uncertainty that comes from wind power. Specifically, SPP uses a methodology that incorporates the impact of load magnitude, load changes, wind magnitude, and wind variability and adds these components together to get the total requirement for regulating reserve [27]. As an example, the below equation shows the requirement methodology:

$$\begin{aligned} RegUpRequirement \\ &= a^{up} * MTLF(t) + b^{up} * \big(MTLF(t+1) - MTLF(t)\big) + c^{up} * IRF(t) + d^{up} \\ &\quad * (IRF(t+1) - IRF(t)) \end{aligned}$$

$$\begin{aligned} RegDownRequirement \\ &= a^{down} * MTLF(t) - b^{down} * \big(MTLF(t+1) - MTLF(t)\big) + c^{down} * IRF(t) \\ &\quad - d^{down} * (IRF(t+1) - IRF(t)) \end{aligned}$$

Where *a*, *b*, *c*, and *d* are coefficients that help determine the linear relationship of magnitude and variability to regulating reserve requirement need, *MTLF* is mid-term load forecast, and *IRF* is intermittent resource forecast. The variability component cannot be negative for the RegUpRequirement and cannot be positive for the RegDown Requirement. As an example, SPP uses a=0.005, b=0.02, c=0.01, and d=0.03 for both upward and downward coefficients in its requirements document (though this may be different than what is used in practice). These

---

[5] http://www.platts.com/latest-news/electric-power/houston/us-southwest-power-pool-sets-new-wind-peak-record-21139345.

requirements are calculated for every hour for the next seven days and are updated daily based on new information.

### *Hawaii Electric Company*

HECO has made several adjustments to its regulating reserve requirements as a system with a very high penetration of both wind and solar power. The most recent method is to backup the wind and solar power one for one with spinning reserve capacity up to a maximum aggregate amount [28]. In the day time the maximum is cut at 18% of nameplate capacity and in the night at 23% of nameplate capacity. These numbers were determined based on scatter plots that showed that the likelihood of a 30-minute ramp event dropping to zero was very unlikely at levels above these maximum values, but that at points lower, it was observed to occur.

## Regulating reserve pay for performance market design

In addition to the changes to regulating reserve requirements, there has also been some significant changes to the market design for regulating reserve (or regulation). One of the most significant changes that has occurred to the ancillary service markets within the United States over the last several years is the "pay-for-performance" design for the regulating reserve market. This design was implemented by the six FERC jurisdictional ISOs (all but ERCOT), and was based on FERC's Order 755 which was issued in Oct of 2011 [29].

FERC Order 755 had three primary directives to it. First, it required all the ISOs to incorporate lost opportunity costs that were incurred by the marginal resource into the regulating reserve clearing price. Many of the ISOs had already done this but a few provided lost opportunity costs only specific to the resources that incurred those costs rather than as part of the clearing price that goes to all regulating providers. The second major component was the use of a market-based price and payment for performance of regulating reserve, which addresses a payment for the amount of movement that a regulating resource provides. Third, FERC directed that the ISOs accounted for accuracy when determining the level of payment and eligibility for regulating resources. This accuracy was based on how well a resource followed the control signal given through the AGC. These three directions led each ISO to make design changes to its regulating reserve ancillary service market; though, as we discuss next, each having its own unique design characteristics. Although, each of the six ISOs that made changes all have unique designs, we focus on two for this report, NYISO and PJM.

### *NYISO*

The NYISO had implemented numerous changes to its regulating reserve market to comply with FERC Order 755. In addition to the regulation capacity bid and the regulation capacity price, a regulation movement bid and regulation movement price are added. The regulation movement bid is based on wear and tear cost from additional movements within the interval. It is only paid in the real-time market. The calculation of the regulation movement price is based on the value of the marginal regulation capacity provider's regulation movement bid. The assumed amount of movement is about 10 times the regulation capacity – in other words, NYISO expects that resources would have absolute movement from its minimum to its maximum regulation capacity limits (or vice-versa) ten times within an hour. Finally, the overall revenue received from regulation movement is equal to the regulation movement price multiplied by the regulation movement of the resource multiplied by the resources performance index. The performance index is a percentage that is calculated based on the resource's accuracy in responding to the signals that it is given.

### *PJM*

PJM had some similar and some different implementations to comply with FERC Order 755. The primary differences include its accuracy calculations, its additional regulation signal, and the regulation benefits factor.

PJM uses a three-part accuracy calculation to determine how well a regulating resource follows its signal. The parts include accuracy scores in the ability to reduce the amount of delay in responding, the ability to correlate its response with the signal, and the precision of following the signal. These scores are equally weighted and will impact whether the resource qualifies for providing the service, whether an existing regulating resource should be disqualified, and how much revenue the resource is provided when regulating.

Figure 4-4 shows an example of the two different signals used for PJM regulating resources. The RegA signal is the traditional regulation signal sent to resources before Order 755. This signal is filtered with an integral term such that resources do not chase rapidly changing positive and negative values of ACE. The RegD signal is designed to correct for the higher frequency changes in ACE. It also has an energy neutrality component, where it attempts to keep the positive and negative signals about equal over a shorter time frame. This signal is designed for faster resources including energy storage resources who can also take advantage of the energy neutrality.

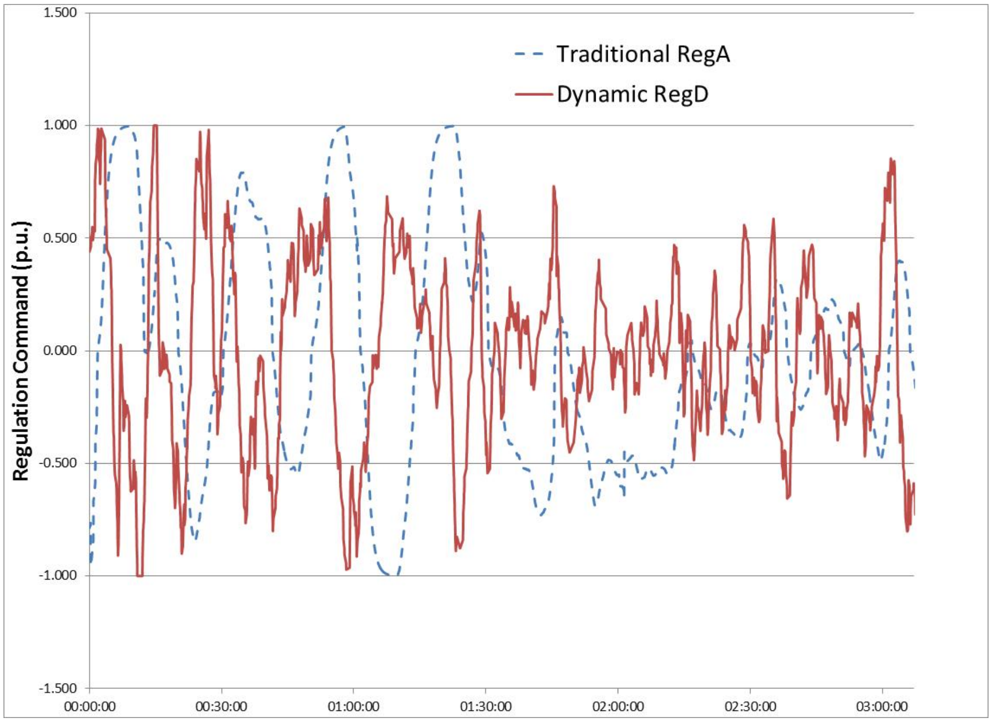


**Figure 4-4**
**PJM traditional RegA signal compared with new dynamic RegD signal used for different resources providing regulating reserve.[6]**

Finally, the PJM regulation benefits factor is a component of the revised regulation market that describes the additional benefit of faster responding resources, i.e., those that follow the RegD signal. It is based on a study performed in 2011 [30]. The benefits factor is higher when there are less resources providing the faster signal, with a saturation point where the benefit of RegD resources is negligible. This allows the regulation market to substitute these resources to meet the overall requirement. It also provides for a greater payment for the faster resources. They will get paid a multiple of the regulation capacity price as determined by the benefits factor.

Table 4-1 shows average prices for the two ISOs for regulation capacity and regulation mileage. Note that the NYISO regulation movement price can be multiplied by the movement multiplier (e.g., 10) to provide a better comparison to the regulation mileage price of PJM. This new component of the regulating reserve market can provide greater revenue streams particularly for those resources that provide a more accurate response that has greater absolute movement as requested by the ISO.

[6] www.pjm.com/markets-and-operations/ancillary-services.aspx

**Table 4-1**
**Average regulation capacity and regulation mileage (or movement) prices for PJM and NYISO for 2015.**

| | Regulation Capacity Price | Regulation Mileage Price |
|---|---|---|
| NYISO | $8.95/MW-h | $0.21/MW movement |
| PJM | $28.25/MW-h | $3.36/MW-h |

## Primary frequency response and fast frequency response services and markets

Although primary frequency response is a service that has been provided by generating technologies for many years to support adequate frequency control and to avoid tripping of under-frequency load-shedding relays, it has not been described as clearly as an ancillary service nor has it had an ancillary service market in most power systems, including all of the United States. Some industry members have been observing the level of frequency response, measured in MW/Hz on the Eastern Interconnection, and have observed that the levels have been declining over the last two decades (see Figure 4-5) [31]. This shows the amount of response during large loss of generation under-frequency events. Since load has been increasing and more generators have been installed on the system since the start of the decline, overall frequency response should have been increasing. However, researchers believe that much of the decline is due to governor dead bands that are very high and insensitive to normal frequency deviations, blocked governors not providing any response, or generators operating in modes that provide energy more efficiently yet provide primary frequency control poorly [32]. Although this decline predates any significant penetration of VER or DER, these resources are nonsynchronous and do not traditionally provide primary frequency response. Thus, with the current trend of declining frequency response, and with VER or DER displacing conventional synchronous machines, the trend may continue to decline without any additional action.

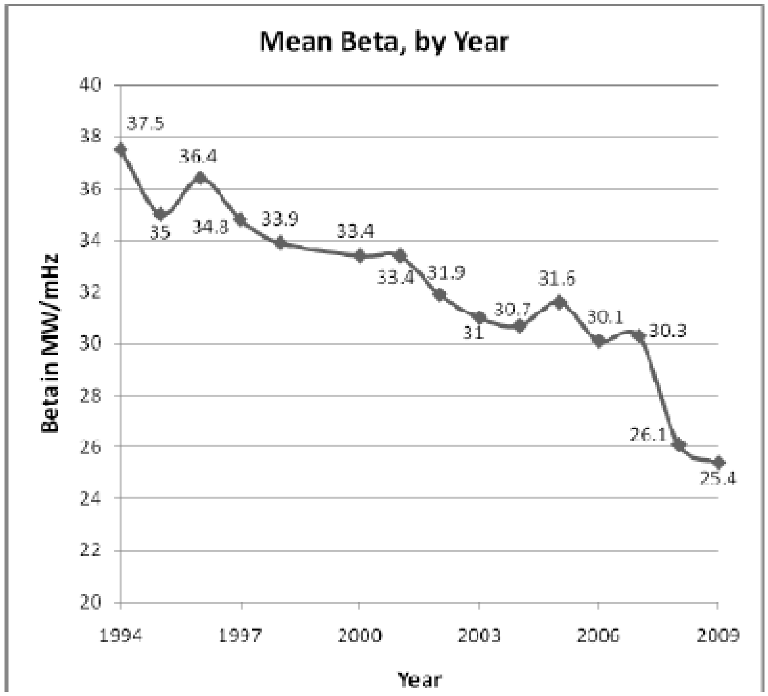


**Figure 4-5**
**Frequency response (Beta) has been declining since the 1990s in the U.S. Eastern Interconnection [31].**

Although VER is nonsynchronous, many VER technologies, particularly wind but also solar PV, have mechanisms for providing primary frequency response in a manner very similar to conventional synchronous generators. Variable-speed wind plants can provide primary frequency control through pitch or torque control procedures [5]. However, just as in synchronous machines, they have to be operating at a power level below the maximum available power in order to provide response during under-frequency events. Due to the lack of fuel costs, it is much rarer for VER to be operating at these levels. When they are, they are able to provide an upward response based on similar droop settings and response speeds to conventional generation. Figure 4-6 shows an event in ERCOT, where wind plants regularly provide primary frequency response, where the wind plant (in blue) responds immediately to a decline in system frequency (in red). Recent studies have shown solar PV provide this response through a power electronics control as well [33].

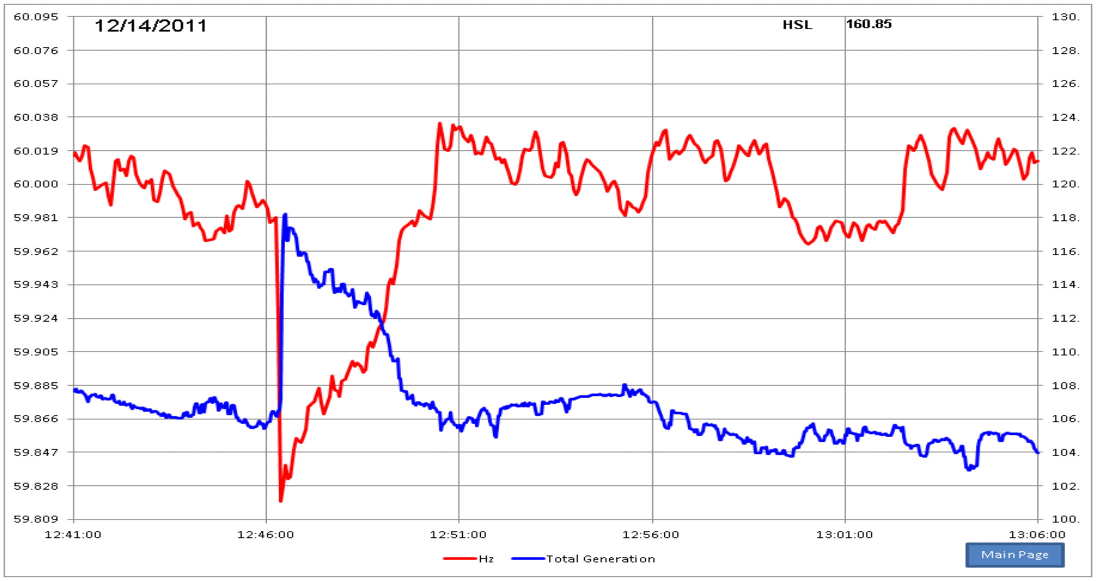


**Figure 4-6**
**A wind plant (in blue) responds immediately to an under-frequency event (in red).** [7]

As discussed in Section 2, standards in both North America and Europe direct balancing areas and Transmission System Operators on the level of primary frequency response that should be maintained within their system. The North American requirement was put in place largely due to the frequency response declines that were being observed. However, these standards do not direct which technologies or resources should provide the service, only that the area must maintain minimum levels. In regions which the generation owners are separate from the responsible balancing area authorities, it is possible that either payments or individual requirements are necessary to ensure enough primary frequency response is available. In the United States, a recent notice of inquiry was issued by FERC, requesting information on whether there should be requirements for interconnecting resources to provide primary frequency response, requirements for all resources including those already existing to provide primary frequency response, or some form of compensation for resources to provide the service [34].

The NERC BAL-003-1 standard introduced the frequency response obligation (FRO) [6]. The FRO directs each BA to maintain a minimum amount of frequency response in MW/0.1Hz when responding to contingency events. The requirement is determined first as an interconnection requirement, where the response should avoid automatic involuntary under-frequency load-shedding following a NERC Category C (N-2) event in the interconnection. This requirement is then prorated for each BA based on its generation production and load share within the interconnection. This requirement starts for all FERC jurisdictional BAs in 2016, and is the first time that they will maintain a frequency response minimum within North America.

We discuss a few examples of regions which have included primary frequency response and fast frequency response as an ancillary service or who have proposed to do so. Australia has a 6-

[7] Sandip Sharma, ERCOT, "Frequency control requirements and performance in ERCOT ISO," presented at EPRI/NREL/PJM Inverter Generation Interconnection Workshop, Apr 11-12, 2012.

second frequency control ancillary service, while National Grid United Kingdom (NGUK) recently implemented a fast frequency response service. ERCOT has an ongoing proposal to introduce both[8].

### *Electric Reliability Council of Texas*

As discussed earlier, ERCOT proposed the PFR and FFR ancillary services as part of its future ancillary services redesign [14]. PFR is defined as a service that acts without operator action to provide an arresting and/or counter response proportional to frequency deviations to maintain the steady-state frequency of the interconnection. ERCOT states that PFR operates within the first few seconds following the disturbance event and is fully delivered within 12 to 16 seconds. That response must also be sustained for an additional 30 seconds.

The PFR service had several other features as part of the requirements of resources that were able to participate in its provision. A provider must have a governor droop setting that does not exceed 5%, where the droop percent is equal to the percent frequency deviation that would cause a 100% change in power output for the resource. It also included a requirement that the governor dead band setting, which is a level of frequency deviation below which no response is provided, should be no lower than 36 mHz. The proposed service also limited the amount of PFR that a single resource could provide by limiting it to the response that would be provided through a 1% change in frequency outside the dead band. For a 5% droop setting, this equates to about 20% of maximum capacity.

The FFR service is a related, but slightly different new service being proposed at ERCOT. The response provided is during large frequency events, but the quantity provided is a block of energy rather than an amount proportional to the frequency deviation. Although it is not specifically referenced by technology, the service would most commonly be provided by demand response resources who voluntarily trip off at certain frequency levels to provide what is similar to a boost in power (though storage may be a future provider as well). The FFR service is further split into two sub-products: FFR1 and FFR2. Both types of FFR would be expected to provide the full response within half a second. FFR1 would have a higher trigger point at 59.8 Hz, such that it would be expected to be activated more frequently than FFR2 which has a trigger point at 59.7 Hz. However, FFR1 would only be obligated to provide its response for up to 10 minutes while FFR2 would be expected to provide its response for longer periods of time.

Figure 4-7 shows some examples of how PFR and FFR would contribute depending on the magnitude of the event and of the frequency deviation. FFR is either all or nothing while PFR provides a response that is proportional the frequency deviation that occurs.

---

[8] Although ERCOT's proposal has been voted down by stakeholders, there is still an initiative to move forward with certain parts of the proposed ancillary services.

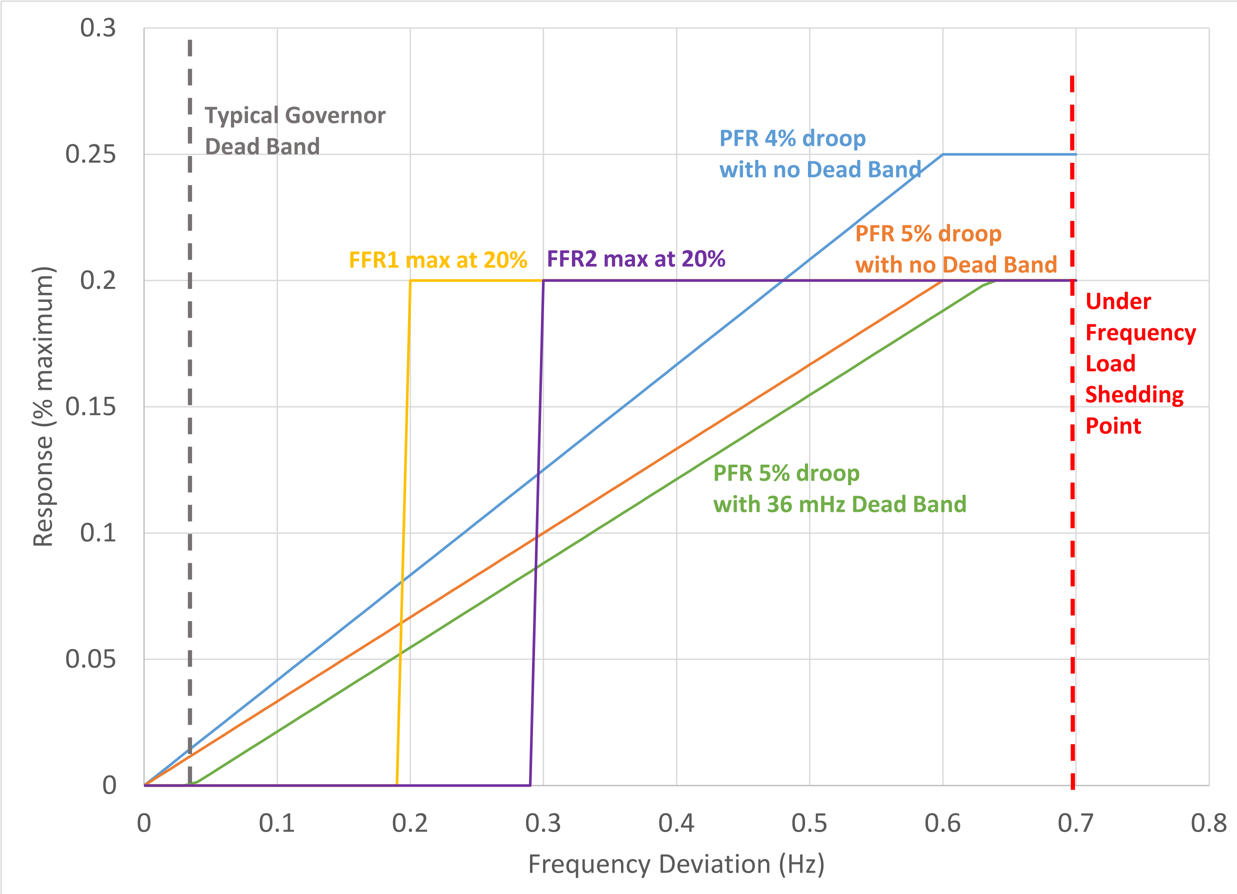


**Figure 4-7**
**Different response expectations of PFR and FFR based on frequency deviations.**

The minimum requirement of PFR is determined based on the FRO for ERCOT, which is about 286 MW/0.1Hz (although ERCOT uses a more conservative requirement than the FRO). In addition to the PFR requirement, the full requirement is a shared requirement between PFR and FFR, where certain combinations of PFR and FFR can be used together to meet the full requirement. The two are linked through an equivalency ratio, which determines how much FFR is equivalent to 1 MW of PFR. These services would have ancillary service markets as well, with prices that are calculated to represent the marginal cost of the service. During the discussion on these two services, there was a lot of debate on whether PFR and FFR should be paid the same rate or whether the FFR providers should be paid something different based on the equivalency ratio of that service compared to PFR.

### *Australia and New Zealand*

Both the Australian National Electricity Market in Australia and Transpower in New Zealand include ancillary services that correspond to primary frequency response for contingency events. In Australia, eight different frequency control ancillary services are included: regulation raise, regulation lower, six-second raise, six-second lower, sixty-second raise, sixty-second lower, five-minute raise, and five-minute lower [35]. The six-second service is used for arresting frequency decline and the sixty-second service is used for stabilizing system frequency. The six-second service corresponds to the support that synchronous inertia (discussed next) would provide, while the sixty-second service is closer to the support that primary frequency response provides. In both cases these services are not tied to technology, and they don't even require the service to be online; only that the injection (or reduction in terms of lower service) to be provided within the associated time frame. Certain monitoring and compliance to ensure the resources providing these services are providing sufficient response. The response for six-second service must be provided automatically and locally at the plant, without any direction from the system operator, the Australian Electricity Market Operator (AEMO). Each of these services have ancillary service market constructs that are similar to U.S. markets, with resources providing price-quantity bids to provide the service and prices based on the marginal cost to provide the service.

In New Zealand, an "instantaneous reserve" is a required service to arrest the declining frequency during disturbance events[9]. It includes a fast instantaneous reserve and a sustained instantaneous reserve. The fast instantaneous reserve needs to sustain output within sixty seconds, while the sustained instantaneous reserve must respond for long enough for the frequency to be brought back to nominal level. These services can be provided by generating units or interruptible loads and must respond automatically to these events. Although the market is similar to those of the U.S. ancillary services, the costs of the service are actually paid by generators (and HVDC owner), rather than by loads. The overall cost of this service in 2015 was about $19M New Zealand dollars (about $13M USD with current exchange rates), which is a relatively high cost for the service for a system that consumes about 10% of the energy that ERCOT does.

### *United Kingdom*

In the UK, the System Operator, NGUK, is only responsible for balancing supply and demand over the last 15 minutes. Prior to that, bilateral and exchange markets are used for participants to trade energy supply and demand. For the last 15 minutes before delivery, NGUK is required to balance supply and demand, primarily through the Balancing Mechanism, where they accept offers to buy and sell electricity from loads and generation.

System frequency control is procured through a number of categories in the UK [36]. Dynamic response provides pre-fault control as well as responding to contingencies, similar to AGC. Non-dynamic sources provide frequency response post-contingency based on changing consumption through automatic relays. Post contingency response is divided into primary (up to 30 seconds), secondary (30 seconds to 30 minutes) and longer duration reserves, with NGUK procuring enough of these to maintain frequency within acceptable limits. In the first category, all transmission connected generators over 100 MW must provide "Mandatory Frequency Response" as part of the UK Grid Code; this is a dynamic response, which controls frequency pre-contingency and responds to frequency changes post-contingency. Firm Frequency Response is the second type of response, and this can be provided by dynamic and non-dynamic sources of frequency response. Resources are paid to provide this service, in contrast to mandatory provision, and it can be provided by balancing and non-balancing units alike. Mandatory providers can provide this service, as well as demand response.

NGUK procures and pays for frequency responsive services on a monthly basis through tendering. The generators provide information on cost to hold frequency response and get paid

[9] https://www.transpower.co.nz/system-operator/electricity-market/instantaneous-reserve

on a monthly basis. Requirements for dynamic and static response are determined based on the size of largest infeed loss at risk, system demand, and system inertia. They then procure amounts sufficient to be able to manage frequency pre- and post-contingency. Often, this means most of the requirement is covered by dynamic response capable resources, but at some times of the day or year, static requirements are also significant. Mandatory Frequency Response is paid based on a holding fee (£/h held based on monthly bids from the generator) and energy payment (£/MWh based on energy delivered and consumed). Firm Frequency Response is paid an availability fee for hours they are available, initiation fee when they are selected to provide frequency response, notification fee when used and an energy fee. As such, there is already a primary frequency response market in the UK. However, it overlaps significantly with frequency regulation in the US, with payments for both pre- and post-fault behavior.

With increasing renewables and DER on the system, as well as more HVDC interconnection with the rest of Europe, there are concerns that inertia is decreasing due to displacement of synchronous generation. Additionally, new technologies like energy storage can provide a very fast frequency response that may reduce the overall burden. NGUK therefore determined that it would be beneficial to add a new frequency responsive service, called Enhanced Frequency Response (EFR), with the aim to improve pre-disturbance frequency control through changing power output or consumption by pre-defined amounts based on frequency, as well as provide benefits post-contingency as it would be a dynamic service [37]. Providers would need to be able to quickly respond and sustain response for 15 minutes. For example, for frequency deviations of 0.25 Hz, 44% of the cleared EFR would need to be activated within 1 second. As storage is a primary provider of this type of service, dead bands were identified at which they would not need to operate, in order to manage state of charge.

NGUK determined that no more than 200 MW of total frequency response (which is over 1100 MW currently) would be provided by EFR as it is introduced, with the option not to procure any if uneconomic. However, in the first auction, all 200 MW was procured as energy storage. NGUK have determined that the capability cleared in this market will save over £200M ($250M USD) over the four-year contracts, with storage being paid a little over £9/MWh ($12/MWh). This came in lower than expected – it is likely that storage developers have other potential sources of revenue beyond the EFR market.

### Synchronous Inertia

Synchronous inertia is another service that has always been provided by system resources, but that was never necessarily considered as a separate ancillary service nor has it been paid for through market-based or cost-based rates. As an inherent feature to synchronous machines, it is not something that can be increased or decreased by these technologies, nor is it something that can be turned on or off. The ability of a synchronous generator to provide inertia is determined only by whether it is synchronized (online) or not (offline). Generally speaking, there is also not lost opportunity costs as exists for many other ancillary services, as the generator does not have to adjust its active power output to provide a specific amount of inertia. Finally, most large interconnections had sufficiently large amounts of synchronous inertia, such that there was historically no reason to incentivize or request additional inertia from the resource fleet.

VER are non-synchronous, and do not provide instantaneous injections of energy based on spinning mass as the rotational speed is not synchronized with that of electrical frequency. Thus, it does not provide any synchronous inertia. Variable-speed wind technologies do have a spinning mass, albeit not synchronous. By use of controls, these technologies can provide fast, yet not instantaneous, injections of power by extracting the kinetic energy of the spinning wind turbine blades, after controls sense the change in frequency. However, the energy must be "paid back" to prevent the turbine from stalling. PV can only provide a synthetic inertia if it is not providing full available power beforehand, which may be more similar to FFR described earlier. Synchronous motors provide inertia similar to synchronous generators.

### *Electric Reliability Council of Texas*

In the redesign of ancillary services by ERCOT, as described previously, the ISO has also proposed a service for synchronous inertia [14]. ERCOT had seen various conditions over the past six years where the amount of synchronous inertia on its system has been significantly low. This is primarily during light spring-time load conditions at night, where large amounts of nonsynchronous wind generation are producing, and low energy prices may have caused conventional synchronous resources to turn off. ERCOT believes that at inertia levels below 100 GW-s, the system will be at risk of involuntary load shedding after the loss of its two largest units. In addition, at levels less than 120 GW-s, it may lead to greater requirements for spinning reserve.

As a service, synchronous inertia is different from other active power ancillary services in that there is no capacity reservation. Once a synchronous generating unit is online, it provides all of its synchronous inertia. There is no way that the resource can produce less. An offline resource cannot provide any synchronous inertia.

As part of this proposal, ERCOT will publish hourly synchronous inertia requirements as part of its daily ancillary service plan. The requirement will be based on the amount of inertia that would not exceed a pre-defined maximum allowable rate of change of frequency. This requirement will then be met as part of the day-ahead market, with qualified scheduling entities offering in the cost offers of resources that can provide synchronous inertia. The day-ahead market will then procure the required amount based on a least cost selection. Even though synchronous inertia is a discrete quantity service, where a resource can only provide zero or full capability amount of its synchronous inertia, ERCOT will make awards between its range of zero and the full amount. The price for synchronous inertia will then be calculated based on the marginal cost of meeting the synchronous inertia constraint. So far, the proposal was strictly for synchronous inertia from synchronous generators. Synthetic inertia from nonsynchronous resources would not contribute towards this service, and would likely better fit ERCOT's proposed FFR service.

### *Hydro Québec*

Hydro Québec (HQ) is a single balancing area interconnection in North America. As a peak load of 38,900 MW, it is the smallest interconnection on continental North America. This led HQ, a predominantly hydro system, to require its generation interconnecting to the system to provide inertia or inertia emulation [38]. This included all wind generation with a rated output of greater than 10 MW. These plants were required to provide an emulated inertia response during frequency disturbances, when frequency deviations are significant (i.e., greater than 0.5Hz). The response is directed to provide a similar characteristic to that of conventional synchronous generation [39]. For example, the response should be similar to a 3.5 second inertia constant of a conventional synchronous generator. In order to do this, the response should inject an additional 5% of power for at least 10 seconds. At that time, the wind turbines would require a recovery mode and would reduce output from the initial stage. There is also a requirement that the response must start within 1 second, and must reach its 5% injection level within 1 second. An example of the contribution that wind would be providing for emulated inertia is shown below in Figure 4-8.

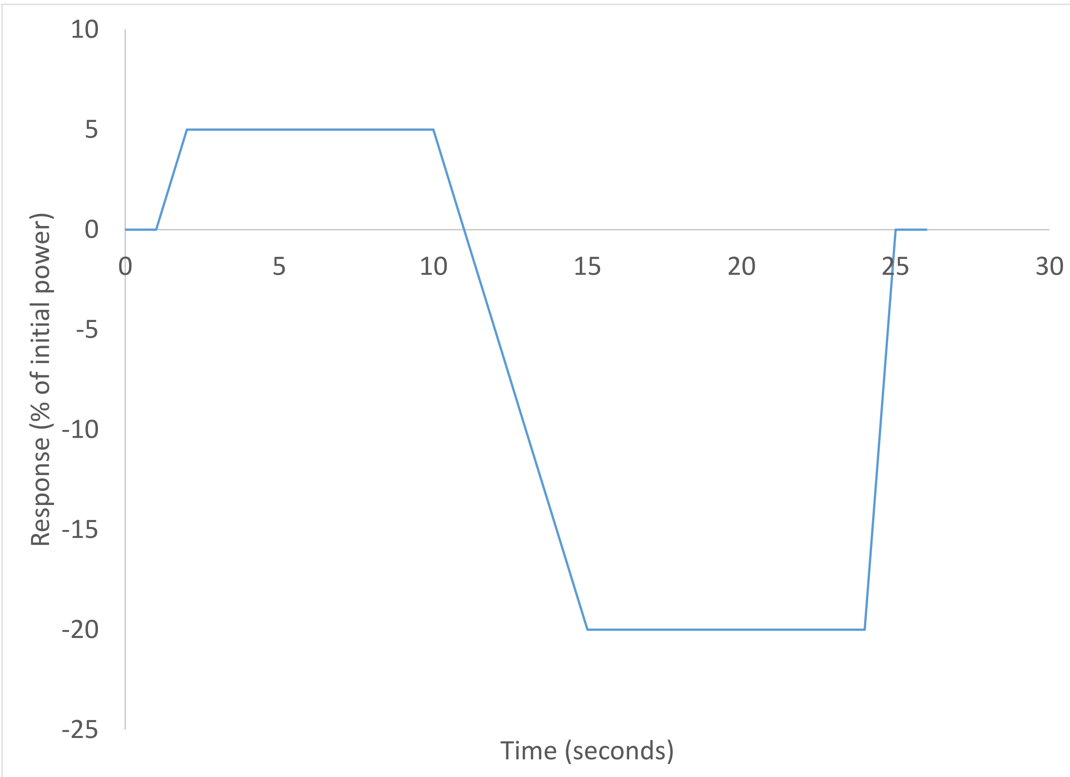


**Figure 4-8**
**Hydro-Quebec's emulated inertia required response from wind generators (derived from [39]).**

This mandated requirement for a nonsynchronous resource to provide a service provided by synchronous resources is fairly unique across the industry. The requirement is enforced to assist in short-term frequency recovery. However, no system-wide inertia requirement has been determined nor is there any compensation schemes for providing this service.

### *Eirgrid*

As described earlier, Eirgrid is currently in the process of revising ancillary services that they procure for operating the Irish power system. One of these new services is Synchronous Inertia. As described above, this is the response that is immediately available from synchronous generators, synchronous condensers and some synchronous demand loads (when synchronized) because of the nature of synchronous machines. As an island system, inertia is a key determinant of the strength and stability of the power system.

Based on the methodology described earlier, the system portfolio for the next several years is examined to determine the overall system needs. Different sensitivities around flexibility of the portfolio is also examined. Simulations are run with and without different services to determine their relative value. Portfolios are adjusted until the amount of services procured meets system requirements. The amount actually procured for each is then based on a methodology developed by Eirgrid, where they determine the minimum amount needed to maintain reliability across the year.

The synchronous inertial component is calculated by ensuring that the hourly real time requirement is met in every hour, with the least number of SIR providers needed to meet the requirement. SIR requirement is assumed to be 17,500 MW-s in the next few years. It was determined that payments for SIR would be approximately 2% of the total system services payments, which ends up being about €0.46/MW-$s^2$-h. This is a relatively small payment at present, but is paid to those generators that are online and have been contracted to provide the service. Obviously, this is different than a US ISO, where procurements would be cleared on a daily or hourly basis; in Ireland these are done for a yearly basis. In the first year of this product, SIR performance is not being tested. In future years, resources providing SIR will be tested, once more experience has been gained in how to test. In the future, other resources such as energy storage and wind could potentially provide this service, but currently do not.

### Short-term flexible ramping products

Looking at Figure 2-1, flexibility reserve is reserve capacity that is held for normal, non-event conditions, but based on anticipated imbalance impacts that may or may not be deployed through automatic control. It can also be referred to as load following reserve, following reserve, or ramp capability. There is no national-level standard for this type of reserve either through NERC or ENTSO-e. However, a few regions have seen a need for this type of reserve (or more likely, a benefit rather than a need). This product is mostly defined as rampable capacity reserved within the real-time dispatch interval that can be used at a subsequent dispatch interval when ramping is needed. Figure 4-9 shows a dispatch schedule with reserve for both regulating reserve and flexile ramping reserve (“Flex up” and “Flex Down”) that is held in case conditions are different than expected. As long as the dispatch between t and t+5 is with the shaded regions, there is enough ramping capacity to meet it. If the next dispatch forecast at t+5 is within the red-shaded region, then there is enough ramping capacity to meet the uncertainty without an impact to regulating reserve quantities.

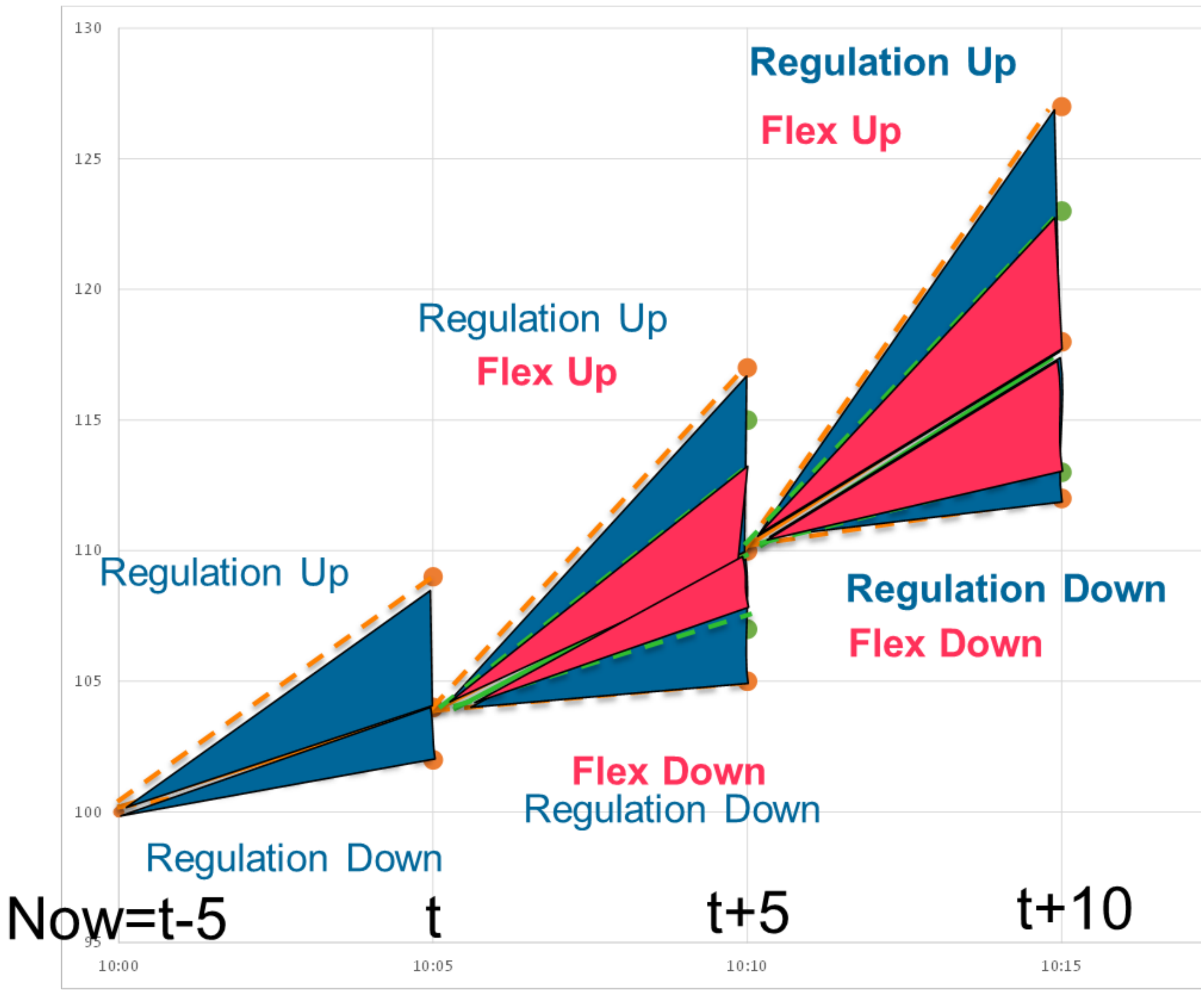


**Figure 4-9**
**Flexible ramping product is capacity reserved beyond regulation for uncertain conditions.**

Flexible ramping reserve has some key differences to the other reserve types. It is for these reasons that regions that have introduced flexible ramping products did not simply increase other reserve products like contingency or regulating reserve. Table 4-2 shows some of the characteristics of flexible ramping products compared to regulating reserve and spin and contingency reserve (spinning or non-spin reserve). First, the type of process that guides or directs the deployment of that reserve is shown. Regulating reserve is typically guided by the AGC which provides control signals for resources to respond to, based on the current ACE on the system whereas contingency reserve is guided by the operators directing the resource to respond based on an event. Flexible ramping reserves is deployed by the economic dispatch model, economically just as energy is being scheduled. Regulating reserve is used every interval and after the dispatch interval is complete and between each dispatch interval, contingency reserve is rarely used, and flexible ramping is used often though likely not every interval. In terms of what the reserve is used for: regulating reserve for short-term, minute-to-minute changes in net load, contingency reserve by large loss of supply contingency events, and flexible ramping by errors in forecast error that are typically at least five-minutes in advance. In terms of ancillary service markets, the table also shows the level of penalty (or shortage) prices (See Section 3), with contingency reserve being the highest, following by regulating reserve, and flexible ramping typically having the lowest penalties. Finally, for bidding purposes in market areas, regulating reserve almost allows allow bids that represent the costs of wear and tear or efficiency impacts, contingency reserve sometimes allow bid costs, and so far, none of the markets allow bids for the flexible ramping reserve product.

**Table 4-2**
**Difference between flexible ramping products and other traditional active power ancillary services.**

| Operating Reserve Characteristic | Regulating reserve | Contingency reserve | Flexible ramping product |
|---|---|---|---|
| ***What guides response*** | Automatic generation control | Operator-directed, often manual | Security-constrained economic dispatch directions |
| ***Frequency of use*** | Every interval | Rarely, only after large events | Fairly often |
| ***What is reserve used for*** | Short-term (less than five-minutes) changes in load and VER production | Loss of supply contingencies | Forecast errors and (several minutes timeframe) ramp events |
| ***Penalty price values*** | \$80-\$600/MW-h (medium) | Typically larger than \$500/MW-h (highest) | Between \$5 and \$250/MW-h (lowest) |
| ***Non-zero bids allowed?*** | Yes due to generators reflecting wear and tear and efficiency costs from fast response | Sometimes, depending on the region | No, costs are only based on the lost opportunity costs |
| ***When is it deployed*** | After dispatch interval (in between real-time dispatch market runs) | After dispatch interval (sometimes through a new dispatch run, e.g., real-time dispatch-corrective action mode (CAM) or real-time contingency dispatch (RTCD)) | Part of dispatch interval |

Of the regions that have added this new service to their list of reserve products, most claim the difference of this service from contingency and regulating reserve as the main reason for needing resources to provide the service, as well as providing compensation for those resources to do so. The primary motivator was increased reliability by using this capacity to ensure the load balance is met during long-duration net load ramps while still maintaining regulating and contingency reserve for their true purposes. Two ISOs that have implemented this new service this year, MISO and CAISO, also saw benefits in the way that the service can be used for reducing productions costs and reducing real-time price spikes. We will discuss the design of the service from these two ISOs as well as the flexibility reserve that was introduced by Xcel Energy's Public Service of Colorado, a regulated utility balancing area.

### *Midcontinent Independent System Operator and California Independent System Operator*

Both MISO and CAISO have introduced short-term flexible ramping products to their operations and ancillary service markets in 2016 [40], [41]. MISO's product, ramp capability product, began May 2016. CAISO's product, flexible ramping product, was introduced in November 2016. Much of the motivation for these designs came from instances where shortage pricing, set when ancillary services were not sufficient and therefore the ancillary service and energy price were set at very high levels, were occurring solely based on five-minute periods that had insufficient ramp available. When the ISO systems made commitment and dispatch decisions to meet the expected conditions for the next period, two things could occur. First, if the decisions were only made for the single period, it is possible that those decisions would not be able to meet the ramp conditions that were in future time periods. This could lead to shortage prices. Second, if the commitment and dispatch decisions are made to meet the expected conditions, if the conditions (e.g., net load) were much different than those expected, there may not be enough ramp to meet those conditions. By procuring ramp products that can be released for energy, the ISOs anticipated that this additional ramping capacity would mitigate these unnecessary transient price spikes. So far in MISO's case, they have found this to be true.

Both products have a lot of similarities. They both reserve capacity in the current dispatch periods in order to hold back ramping capability for future dispatch periods. Both procure capacity in both the upward (for increases in net load) and downward (decreases in net load) directions. The maximum requirements for both products are also very similar. They are based on the expected variability within the time horizon plus a high confidence interval (e.g., $95^{th}$) of uncertainty conditions based on statistical net load forecast error characteristics.

There are also several differences in the products that MISO and CAISO implemented. We summarize these in Table 4-3. First, while CAISO is procuring the flexible ramp product for the next five minutes so that it can be used in the next five-minute interval, MISO procures the ramp capability for two intervals ahead at 10-minute horizon. Second, the shortage price, the price that is set when there is insufficient flexibility reserve, is also different. In MISO, a single penalty value of $5/MW-h is used, while in CAISO a stepped curve with penalty prices ranging from about $11 to $250/MW-h are used (for upward flexibility). Thus, MISO allows the product to be short at relatively low costs. The different markets and scheduling processes also differ. MISO procures ramp product in all of its markets and scheduling processes while CAISO does not procure flexible ramp in the day-ahead process. Finally, MISO also includes post-deployment deliverability constraints for its ramp capability, as it does for other ancillary services. These constraints ensure that the ramp product can be delivered given transmission constraints.

**Table 4-3**
**Comparison of MISO and CAISO short-term flexibility products.**

| Characteristic | MISO ramp capability product | CAISO flexible ramp product |
|---|---|---|
| Ramp horizon time | 10 minutes (2 RTSCED intervals) | 5 minutes (1 RTSCED interval) and 15 minutes (FMM) |
| Insufficiency cost (scarcity price for ramp product) | $5/MW-h | Stepped demand curve ($11 to $250/MW-h for upwards) |
| Minimum Requirement | N/A, MISO uses a single-period economic dispatch, so there is no variability requirement and only one requirement | Expected variability |
| Maximum Requirement | Expected Variability + 2.5s (uncertainty) | Expected variability + 95$^{th}$ percentile (uncertainty) |
| Markets | DAM, LAC, and RTM | FMM and RTM (not DAM) |
| Deliverability | Post-deployment deliverability constraints | |

The short-term ramp products have prices associated with the services such that the resources providing those services can get paid. In both ISOs, the prices are purely based on the lost opportunity cost to provide energy, as nonzero offers for the service are not allowed. Thus, resources will get paid when the marginal provider of the flexibility product has a lost opportunity cost and is backed down from what it would have provided for energy in order to meet the flexibility reserve product. In MISO, as of the end of September 2016, prices for ramp capability up are $0.55/MWh and $0.13/MWh for day-ahead and real-time markets, respectively [42].

### *Xcel Energy – Public Service Company of Colorado*

The Public Service Company of Colorado (PSCo), a subsidiary of Xcel Energy, is a vertically integrated utility and the Transmission Service Provider for transmission and ancillary services under the Xcel Energy Open Access Transmission Tariff. As a balancing authority, it has one of the largest wind penetrations in the country, with over 20% of its annual energy requirements coming from wind. It has exceeded 67% of energy from wind on an instantaneous basis, and more than 55% on a daily basis. As such, PSCo has to manage the variability and uncertainty associated with wind power, in a relatively small system with over 2 GW of wind on a 6.5 GW peak. A number of different mechanisms are employed to efficiently manage wind. For example, wind forecasting is deeply integrated into operations, and wind generation can, and often does, provide AGC, allowing fossil generation to be decommitted during times of high wind penetration. As solar penetration increases, similar approaches are being used.

In 2014, PSCo filed a new schedule with FERC (Schedule 6A) to provide for Flex Reserve Service [43]. This is a supplemental category of reserves to address large reductions of online wind power over tens of minutes. PSCo needs to cover these down ramps with available generation capacity, which cannot come from resource assigned to meet contingency reserve requirements, as these are for specific types of disturbances, not down ramps of wind. Flex Reserve is comprised of excess contingency reserve as well as online and offline generation available within 30 minutes. Under the proposed service, PSCo recovers the cost for procurement of the capacity from the VERs on the system – while this is currently mainly wind power, more solar power is now coming online and contributing to the requirements.

Requirements for the Flex Reserves are based on studying historical 30-minute ramp data [44]. First, the data is binned into production levels of 100 MW (i.e. all 30-minute ramps for which production at start of the period is between 0 MW to 100 MW, 100 MW to 200 MW, and so on). Then, for each bin, the 95$^{th}$ percentile of the largest ramp in that bin is determined as the ramping requirement. This produces an inverted 'u' shape, with the largest requirements at around 50% capacity factor (i.e. 1,200 MW output for a 2,400 MW wind fleet), due to the fact that the wind power-speed curve is most variable at around 50% of installed capacity. The actual requirement then goes from approximately zero MW at zero MW of wind output (as wind cannot drop off further) to approximately 800 MW at 1,200 MW of wind output. Rather than reduce the requirement for higher wind power output, it is kept constant at a little over 800 MW for all higher wind power outputs. When solar is also included in the examination, it does not significantly add to the requirement, as there is significantly more wind power currently installed. This requirement is then procured in the commitment and dispatch process, and the cost for doing so assigned to VERs.

### Reactive power and voltage control provision and compensation

Reactive power support, or VAR support, which is the primary mechanism for controlling voltage magnitudes, is another ancillary service that has been provided by resources for decades. It has mostly been a service provided by conventional generation technologies and static and dynamic transmission assets (e.g., capacitor banks or FACTS devices). In many regions, generator technologies receive a cost-based payment to recover fixed capital costs involved with the equipment that is needed to provide dynamic VAR support. In addition, in ISO/RTO markets when a generator must back down energy provision in order to provide reactive power outside of its normal range, the generators will typically get paid a lost opportunity cost. While this service is not new, due to some changes in the resource mix as well as motivation from computational improvements and efficiency benefits, the way in which resources are paid for this service may be evolving.

Today, there are two main mechanisms for paying for volt/VAR support. First, many resources are paid for the capability to provide this service. That is, resources recover costs that are incurred for having the equipment installed that allows them to provide dynamic reactive support. One of the most popular mechanisms for doing this is the so-called American Electric Power (AEP) method, which started in the PJM region and now is used in a few other ISOs. This describes the types of costs that are incurred including the exciter, accessory electric equipment that supports the exciter and some additional investment that is used as part of the revenue requirement. The second way in which resources are paid for reactive power service is through provision payments, when actually providing reactive support. The payment here is typically based on any lost opportunity costs that a resource may incur. A reactive power providing resource can incur a lost opportunity cost when it provides reactive power outside of its normal range such that it has to reduce active power (energy) in order to do so [45]. When it does this, it is possible that it is missing out on additional profit in the energy market that it could have made if it were to provide its full amount of energy in the energy market. This cost is paid to the individual reactive power resource so that it is incentivized to provide the reactive power support as directed by the transmission operator.

The way that resources are scheduled in order to provide reactive power support is similar across most United States ISO regions. Because the scheduling models are made through direct current (DC) power flow, voltage constraints cannot be handled explicitly. Instead, the ISO will use offline voltage stability tools to look at certain interfaces to see where stability issues can potentially be an issue. This limit is converted to an approximated MW-based import limit into or out of the area. While this constraint cannot optimize the optimal dispatch of reactive power on generators or reactive control devices, it can ensure that a certain amount of generating units are online that can provide that support within the area. This is usually done in the day-ahead time frame to make sure the units are committed to be able to provide reactive power support. Then, in real-time, operators make adjustments to correct for any changes to system conditions.

There has been substantial efforts by the research community on better scheduling and even marginal cost pricing for reactive power. This includes the use of full alternating current (AC) power flow models within the software that clear the markets for energy (they currently use approximate DC power flow models which ignore voltage constraints and reactive power). In addition, FERC held a workshop in 2016 on the ways of compensating reactive power[10]. In the workshop, FERC speakers presented on the costs to provide reactive power support for both traditional and non-traditional resources, how those costs are most typically recovered, how the ISOs compensate for reactive supply, and whether there are more efficient ways to compensate for reactive supply. There was also some discussion over whether compensation for reactive power should be done in a similar manner to how energy is compensated. However, while there has been lots of discussion, there has not been significant changes by many market regions recently. MISO has been one market area that has been looking at improved mechanisms to schedule and price reactive power. We discuss the concepts and proposals that they have had.

### *Midcontinent Independent System Operator*

When MISO integrated the south region, it found a significant increase in the amount of uplift the system incurred due to voltage constraints and local reliability requirements [46]. Uplift refers to costs that are not paid through the market prices but to specific resources for various reasons. The voltage and local reliability (VLR) requirements forced a substantial amount of out-of-market commitments in certain load pockets. Because prices are only paid for energy and active power ancillary services and the resources needed for VLR would not have been used otherwise, they did not recover all of their costs through the energy and ancillary service markets. Thus, they were owed revenue sufficiency guarantees (also referred to as make-whole payments). The total increase in uplift due almost entirely to the VLR commitments in MISO South went from about $2.4M per month to $11.5M per month. In addition, these costs are not always paid by the region that requires the generation but can be spread out across MISO. This

---

10 https://www.ferc.gov/EventCalendar/EventDetails.aspx?ID=8283&CalType=%20&CalendarID=116&Date=06/30/2016&View=Listview.

substantial increase in out-of-market costs led MISO, stakeholders, and its independent market monitor (IMM) to evaluate some other options.

Two recommendations were given by MISO's IMM on the subject of improving uplift allocation and incentives for VLR service. First, it was recommended that MISO better identify in the day-ahead market when resources are committed solely due to VLR service, such that uplift required to make those resources whole can be allocated and are paid by the loads that are within the region that requires that service. Second, it was recommended to introduce a new ancillary service product, similar to other reserve services, that was local to the areas that required VLR service. The product would be scheduled and priced similar to other services. The objective is that it would provide prices for the resources that are committed solely for VLR in local areas such that the uplift would be reduced and the prices and value of that service would be more transparent. Even though an active power service would be used to ensure resources are able to provide reactive power and voltage support, this design could work in theory. It would bring the units online and pay them to be online, or, because it would be a 30-minute off-line reserve as it is mostly for second contingency conditions, incentivize 30-minute quick-start generation to build in that area, so that they can be used to turn on only when needed to provide the reactive support.

MISO has also studied improved ways of scheduling for resources to provide reactive power support [47]. Since using proxy interface constraints are based on MW limits, they will ensure a certain amount of MW may be provided within an area. However, the energy output of the resource has little to do with ensuring enough reactive power control resources are committed. It is more about having the right set of resources committed to ensure sufficient reactive power support will be available. Therefore, MISO has proposed some enhancements to its commitment process that include a complex set of binary constraints to ensure that sufficient units are committed to provide reactive power support. For example, a constraint may be that at least two of a set of four generators must be online. Various combinations are used for different load pockets and different load levels with critical VLR needs. This method can provide more efficient commitment of resources to provide reactive power and voltage control. However, it does not solve the issue of pricing, incentives, and uplift. Binary constraints, by definition, cannot be reflected within the marginal cost pricing paradigm.

### Long-term flexible capacity

As discussed in Section 2, long-term capacity is bought and sold in some regions as a product, either through auction-based markets or through bilateral agreements. This provides assurance that the system will have sufficient capacity installed to meet future peak load, considering uncertainties in the load and the supply resource availability. However, this does not assure any other attributes of the suppliers other than MW capacity. With increasing VER output, it may be just as important to ensure that the resources installed have enough flexibility characteristics as well, to meet the increasing variability and uncertainty that is anticipated on these future systems when making planning decisions. This flexible capacity paradigm is similar though not identical to the flexible ramping product discussed earlier. While the flexible ramping product was to ensure that flexibility was being committed and dispatched in a way that enables sufficient flexibility for operational time periods, this flexible capacity refers to the assurance of flexible capacity being installed on the system in time for planning horizons.

Understanding the need for flexibility is a challenging process. Unlike the need for adequate capacity, flexibility includes more than one attribute. Ramp rates, start-up times, absolute active power range, ability to sustain output, and minimum online and offline times are just a few attributes that differ between technologies that change the ability of the technology to be flexible when certain conditions change and need response over different timescales. Research has been conducted on how to evaluate the amount of flexibility needed on a system and how to assess the amount of flexibility available on a particular system [48]. These assessments are not always easy to translate into the actual need for flexible resources that need to be installed on the power system [49]. Another unanswered question is how to incentivize a resource to build with enhanced flexibility attributes. Many of these unanswered questions may lead to more significant evolution of the need for this long-term service in the future, as utilities and ISOs begin to determine what their needs are on their respective system. While long-term flexibility is a service that most regions would agree is something that is needed, few have explicitly labeled it as a service or made any significant changes to either require that long-term flexibility needs are met or ensure that compensation is provided for providing long-term flexibility. The CAISO is one of the few regions that has made some substantial changes to identify this as a service over the last few years. We discuss their design next.

### *California Independent System Operator*

In order to ensure sufficient flexibility in the resource adequacy time frame, CAISO undertook a process to develop a new procurement target related to resource adequacy [50]. This went in place in 2015 for the resource adequacy showings of that year. Load-serving entities were required to not only procure sufficient capacity to meet forecasted peak load but also meet additional flexibility requirements with their capacity. The objective is that the system should have sufficient flexible capacity available to meet forecasted system needs. The need is determined based on a minute-by-minute dataset of actual load from the previous year, together with minute-by-minute data for variable generation; projected future variable generation installations are included by scaling wind and load appropriately. The overall requirement is calculated based on the largest 3-hour ramp in each month (since 3 hours is the period of most concern to CAISO), plus the maximum of either the largest contingency or 3.5% of the peak demand in that month, plus an error term to adjust for load forecast error. The error term is based on actual outcomes in practice and is adjusted annually. The requirements vary by month for each of the next three years. While California has identified 3-hour ramping issues as the key issue for its particular circumstances, this may not be the same for all regions at all times.

Once the total procurement requirement has been calculated, the next step is to allocate to each LSE. This is done for each of the three individual components (maximum ramp, contingency or 3.5% of peak demand and error term), based on the contribution of the LSE to that component. They must procure sufficient effective flexible capacity to meet their seasonal requirement. The Effective Flexible Capacity (EFC) is calculated for each resource based on their capability to ramp in 3 hours. The contribution from quick start units is their maximum ramp from cold start in 3 hours, whereas for longer start units, the EFC is calculated as the 3-hour ramping capability when the resources are online and at minimum generation. The requirements calculated in late 2015 for the next 3 years are shown in Figure 4-10.

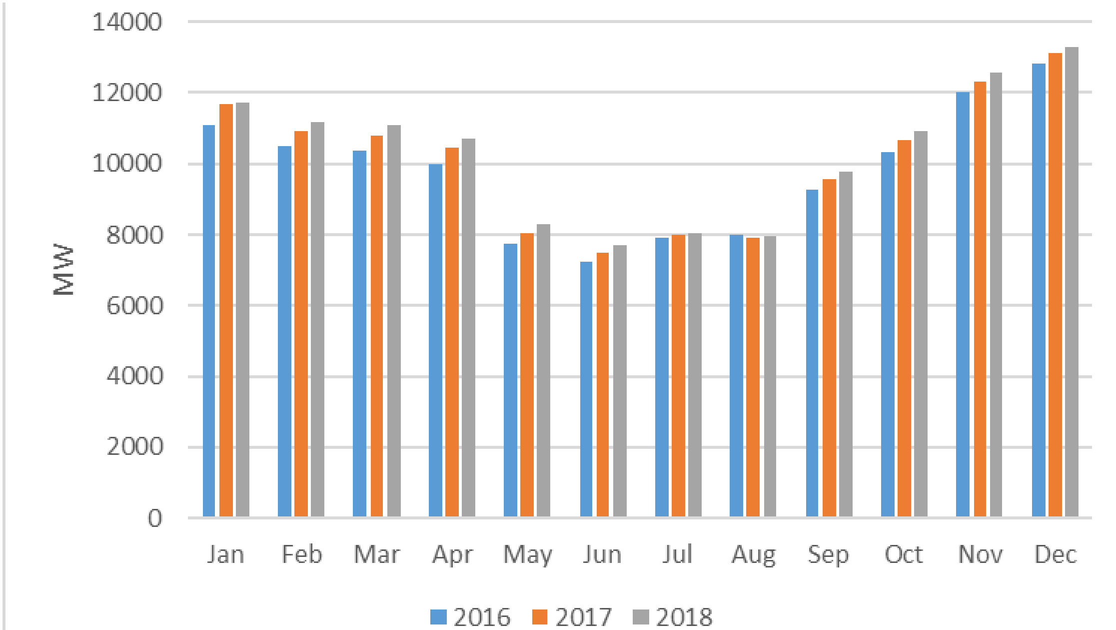


**Figure 4-10**
**Flexible capacity procurement requirements in CAISO [50]**

The CAISO established multiple must-offer categories that are needs-based rather than technology-based, to ensure a technology-agnostic design. Different categories are established to allow for participation of energy-limited resources and other types of resources. They also reflect the fact that the largest monthly ramp does not need to be met every day. As such, there is a baseload flexible capacity requirement, available from 5am to 10pm every day to manage largest secondary ramps in each month (secondary being the largest daily net load ramp that doesn't correspond with the largest ramp), a peak requirement to manage 95% of the largest ramp available for 5 hours a day, and a superpeak requirement to meet the last 5% of ramp needs, and only needed for 3 hours per weekday, a maximum of 5 times per month. In summer months, the baseload flexibility and superpeak flexibility typically make up nearly all of the requirements, while in other months, peak requirements also make up a large portion. This is due to the relative size of the secondary ramp compared to the primary ramp.

Resources can provide Effective Flexible Capacity toward this requirement based on their contribution to three-hour ramping needs. For units with start times over 90 minutes, their EFC is based on the smaller of the operational range (capacity – minimum generation) or the 3-hour ramp rate. For fast-start units, the EFC is the maximum output that the unit can reach from a cold start in 180 minutes. At present, resources such as conventional generators, storage, demand response, and renewables can contribute to this need. CAISO is revisiting the product based on the first few years' experience. This includes, for example a proposal to increase availability of super peak resources to 7 days a week, and revisiting of how EFC is calculated to better account for actual unit ramp rates.

# 5
# SUMMARY

The set of ancillary services that are used to ensure a reliable and efficient electric power system are in many ways the same types of services that have been provided for decades. The way in which power system operators and planners are achieving sufficient quantities of those services, and how we schedule and incentivize technologies and resources in order to provide those services, are evolving. Increasing levels of VER are impacting the way in which we must ensure reliability, efficiency, and fair treatment of compensating resources to provide the services that lead to reliability and efficiency.

This report provides a summary of some recent changes from balancing areas, independent system operators, and transmission system operators from the United States and abroad. An overview of some of the ancillary services used today worldwide is provided as well as a brief introduction to the design of ancillary service markets in the United States. A number of changes to these ancillary services and ancillary service markets is then discussed, providing a variety of different approaches being taken around the world.

Variable energy resources bring increased variability to the system, increased uncertainty to the system, and their displacement of synchronous resources leads to a reduction in the frequency responsiveness of the system. Other new technologies, like various types of energy storage, demand response, distributed energy resources, electric vehicles, and flexible transmission assets, also impact the way in which ancillary services are needed, either positively, negatively, or both. These technologies can help provide some of the services needed to maintain reliability, but may do so in a way that is different than the conventional steam, hydro, and combustion turbines that have been a part of the electric power system for decades. This, in itself, requires changes to the services, to ensure that operators and planners are requesting a capability, rather than a specific technology-centric response. Finally, the changes that we have seen in increased computational capabilities and increased motivation for efficiency due to restructured electricity markets have each led to evolution in the definitions of ancillary services and how we procure those services. When new methods can be used to achieve services more efficiently, and when software now allows us to solve large problems in order to use these new methods, as long as there are no drawbacks, these new methods should be explored.

We see a continuing evolution of the way that ancillary services are defined, scheduled, and incentivized on power systems to come. If the resource mix continues to evolve, we are likely to find other services that were inherently being provided, that may need additional recognition. As the demand becomes more and more responsive, consumers may provide more of those services themselves rather than having to pay separately for supply-side resources to provide them. On evolving, smarter grids, it may be that communication may be the most important service to the reliability of the power system. We will continue to monitor the evolution of the services that are used to maintain reliable power system operations, and provide updates when significant changes are on the horizon.

# 6
# REFERENCES AND BIBLIOGRAPHY

## References


1. International Energy Agency, Variability of wind power and other renewables: Management options and strategies, IEA / OECD, Paris, 2005.
2. Ahlstrom, M., Bartlett, D., Collier, C., Duchesne, J., Edelson, D., Gesino, A., Keyser, M., Maggio, D., Milligan, M., Möhrlen, C. and O'Sullivan, J., 2013. Knowledge is power: Efficiently integrating wind energy and wind forecasts. IEEE Power and Energy Magazine, 11(6), pp.45-52.
3. E. Ela, V. Gevorgian, P. Fleming, Y. Zhang, M. Singh, E. Muljadi, A. Scholbrook, J. Aho, A. Buckspan, L. Pao, V. Singhvi, A. Tuohy, P. Pourbeik, D. Brooks, and N. Bhatt, Active power controls from wind power: Bridging the gaps, Tech. Rep. NREL/TP-5D00-60574, National Renewable Energy Laboratory, 2014.
4. Operating Reserve Determination. EPRI, Palo Alto, CA: 2015. 3002006112.
5. E. Ela, M. Milligan, and B. Kirby, Operating Reserves and Variable Generation, Technical Report NREL/TP-5500-51978, August 2011.
6. North American Electric Reliability Corporation, *Reliability Standards for the Bulk Electric Systems of North America,* May 2016.
7. ENTSO-E, UCTE Operational Handbook Policy 1, Load-Frequency Control and Performance, March 2009.
8. Contributions of Supply and Demand Resources to Required Power System Reliability Services. EPRI, Palo Alto, CA: 2015. 3002006400.
9. Federal Energy Regulatory Commission, *Standard Large Generator Interconnection Agreement (LGIA),* Accessed December 2016.
10. Ela, Erik, Brendan Kirby, Nivad Navid, and J. Charles Smith. Effective ancillary services market designs on high wind power penetration systems. In 2012 IEEE Power and Energy Society General Meeting, pp. 1-8. IEEE, 2012.
11. Z. Zhou, T. Levin, G. Conzelmann, Survey of U.S. Ancillary Services Markets" ANL/ESD-16/1, January 2016.
12. North American Electric Reliability Corporation, Essential Reliability Services Task Force Measures Framework Report, November 2015.
13. North American Electric Reliability Corporation, Essential Reliability Services White Paper on Sufficiency Guidelines, December 2016.
14. Electric Reliability Council of Texas, Future Ancillary Services in ERCOT, ERCOT Concept Paper, Draft Version 1.1, Accessed November 2016. Latest version: 11/1/2013.
15. Single Electricity Market Committee, DS3 System Services Procurement Design and Emerging Thinking Decision Paper, SEM-14-108, Dec 2014.
16. Eirgrid, DS3 System Services Contracts for Interim Arrangement, Decision Paper, August 2016.

17. E. Ela et al., Evolution of operating reserve determination in wind power integration studies, IEEE PES General Meeting, Minneapolis, MN, 2010, pp. 1-8.
18. New York ISO, Growing Wind: Final Report of the NYISO 2010 Wind Generation Study, NYISO, Sept. 2010.
19. Enernex, Final Report - 2006 Minnesota Wind Integration Study, Nov 2006.
20. Enernex, Eastern Wind and Transmission Study, NREL/SR-5500-47078, Golden, CO: National Renewable Energy Laboratory, 2010.
21. Charles River Associates, SPP WITF Wind Integration Study, January 2010.
22. Makarov YV, C Loutan, J Ma, and P De Mello. May 2009. Operational impacts of wind Generation on California Power Systems. IEEE Transactions on Power Systems 24(2):1039-1050.
23. M. A. Matos and R. Bessa, Setting the operating reserve using probabilistic wind power forecasts, IEEE Transactions on Power Systems, vol. 26, no. 2, pp. 594–603, May 2011.
24. Reserve Determination Methods for Variable Generation: Industry practices and the current research. EPRI, Palo Alto, CA: 2014. 3002004242.
25. GE Energy, Analysis of Wind Generation Impact on ERCOT Ancillary Services Requirements, Prepared for Electric Reliability Council of Texas, March 2008.
26. Electric Reliability Council of Texas (ERCOT), ERCOT Methodologies for Determining Ancillary Service Requirements. Effective Date of 01/01/2016.
27. SPP Operations, "Calculating Regulating Reserve Requirements in the SPP Integrated Marketplace," Published 2/27/2014. Available: https://www.oasis.oati.com%2FSWPP%2FSWPPdocs%2FCalculating_Regulating_Reserves_SPP_IM.docx.
28. Docket No. 2014-0183 – Instituting a Proceeding to review the Hawaiian Electric Companies Power Supply Improvement Plans, Order no. 33320 Compliance Filing, Feb. 16, 2016.
29. Federal Energy Regulatory Commission, Docket Nos. RM11-7-000 and AD10-11-000; Order No. 755, issued October 2011.
30. Kema, KERMIT Study Report, PJM Interconnection, Ref Number: Statement of Work 11-3099, December 2011.
31. J. Ingleson and E. Allen, Tracking the Eastern Interconnection Frequency Governing Characteristic, IEEE Power and Energy Society General Meeting 2010.
32. IEEE Task Force on Large Interconnected Power Systems Response to Generation Governing, Interconnected power system response to generation governing: Present practice and outstanding concerns, IEEE Special Publication 07TP180, May 2007.
33. V. Gevorgian and B. O'Neill, Advanced Grid-Friendly Controls Demonstration Project for Utility-Scale PV Power Plants, Technical Report NREL/TP-5D00-65368, National Renewable Energy Laboratory, January 2016.
34. Federal Energy Regulatory Commissions, Notice of Enquiry on Essential Reliability Services and the Evolving Bulk-Power System—Primary Frequency Response, Docket No. RM16-6-000, issued Feb 2016.
35. J. Riesz, I. MacGill, "Frequency Control Ancillary Services: Is Australia a Model Market for Renewable Integration?", 2013.

36. National Grid, Frequency Responsive Services, Available at http://www2.nationalgrid.com/uk/services/balancing-services/frequency-response/, Accessed Dec 2, 2016.
37. National Grid, Enhanced Frequency Response, Available at http://www2.nationalgrid.com/Enhanced-Frequency-Response.aspx, Accessed Dec 2, 2016
38. Hydro Quebec, Transmission provider technical requirements for the connection of power plants to the Hydro-Quebec transmission system, Direction Planification des actifs, February 2009.
39. J. Brisebois, N. Aubut, Wind Farm Inertia Emulation to Fulfill Hydro-Quebec's Specific Need, IEEE Power and Energy Society General Meeting, 2011
40. California ISO, Flexible Ramping Product, accessed December 2016. Available: https://www.caiso.com/informed/Pages/StakeholderProcesses/FlexibleRampingProduct.aspx.
41. Midcontinent ISO, Ramp Management, accessed December 2016. Available: https://www.misoenergy.org/WHATWEDO/MARKETENHANCEMENTS/Pages/RampManagement.aspx.
42. Midcontinent ISO, Ramp Capability Product Performance Update, presented to MISO Market Subcommittee, November 2016.
43. Federal Energy Regulatory Commission, Docket No ER-14-1969, Dec 2014
44. Drake Bartlett, Xcel Energy Colorado's Flex Reserve for Variable Generation, UVIG User's Group Meeting, Denver, October 2016.
45. K. Abdul-Rahman, J. Wu, E. Haq, and P. Ristanovic, "Considerations of Reactive Power/Voltage Control in CAISO Market Operations," *Proceedings of the IEEE Power and Energy Society General Meeting*. 2011.
46. Midcontinent ISO, 2015 State of the Market Report, available www.misoenergy.org.
47. Y. Chen, D. Savageau, F. Wang, A. Casto, Voltage and local reliability commitment under electricity market operations, IEEE PES General Meeting, 2016.
48. Flexibility Methods and Guidelines: Assessing Operational Flexibility in Systems with High Penetrations of Variable Generation. EPRI, Palo Alto, CA: 2016. 3002008380.
49. Strategic &Flexible Planning: Deliverable Flexibility and Resource Adequacy. EPRI, Palo Alto, CA: 2016.3002008379.
50. California ISO, Flexible Resource Adequacy Criteria and Must-Offer Obligation, Nov 30, 2016. Accessed from www.caiso.com.

## Bibliography

*Wholesale Electricity Market Design Initiatives in the United States: Survey and Research Needs*. EPRI, Palo Alto, CA: 2016. 3002009273.

**Export Control Restrictions**

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

**The Electric Power Research Institute, Inc.** (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI members represent 90% of the electric utility revenue in the United States with international participation in 35 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; and Lenox, Mass.

Together…Shaping the Future of Electricity



3002008987

**Electric Power Research Institute**
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com