NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Operating Reserves and Variable Generation

A comprehensive review of current strategies, studies, and fundamental research on the impact that increased penetration of variable renewable generation has on power system operating reserves.

Erik Ela, Michael Milligan, and Brendan Kirby

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

Technical Report
NREL/TP-5500-51978
August 2011

Contract No. DE-AC36-08GO28308

Operating Reserves and Variable Generation

A comprehensive review of current strategies, studies, and fundamental research on the impact that increased penetration of variable renewable generation has on power system operating reserves.

Erik Ela, Michael Milligan, and Brendan Kirby

Prepared under Task No. WE110920

Cover photo is NREL PIX 19498.

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

National Renewable Energy Laboratory
1617 Cole Boulevard
Golden, Colorado 80401
303-275-3000 • www.nrel.gov

Technical Report
NREL/TP-5500-51978
August 2011

Contract No. DE-AC36-08GO28308

Available electronically at http://www.osti.gov/bridge

Available for a processing fee to U.S. Department of Energy
and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/help/ordermethods.aspx

Cover Photos: (left to right) PIX 16416, PIX 17423, PIX 16560, PIX 17613, PIX 17436, PIX 17721, and

Printed on paper containing at least 50% wastepaper, including 10% post consumer waste.

Table of Contents

List of Figures

List of Tables

Abstract

With the advent of new technologies being introduced to the electric power system, the way in which the system is planned and operated may need significant changes. For instance, variable renewable generation, like wind power and solar power technologies, have some very different characteristics than traditional sources of generation technology that has historically met the electricity demand. The variability and uncertainty that is inherent in variable generation technologies adds to the variability and uncertainty in the existing system and can have significant effects on operations. Variability is the expected change in generation and demand balance (e.g., load changing throughout the day and wind power resource change). Uncertainty is the unexpected change in generation and demand balance from what was anticipated (e.g., a contingency or a load or variable generation forecast error). Power system schedulers often use scheduling techniques throughout the day to match generation and demand. When, because of variability and/or uncertainty, the total supply of energy is different than the total demand, system operators must deploy operating reserves to correct the energy imbalance. The way in which they do this and, especially, the way in which they plan for this can dramatically impact the reliability and efficiency of operating a power system with large amounts of variable generation. Systems today have developed their rules and practices based on a long-standing history of operations. Many systems are even now learning new ways to change these rules and practices where high penetrations of variable generation are becoming apparent. Renewable integration studies that are assessing the impacts of future high penetration systems are finding that systems must find alternative approaches for allocating and deploying operating reserves. This report tries to first generalize the requirements of the power system as it relates to the needs of operating reserves. A categorization of the various types of operating reserve is introduced to give a better understanding to the reader and to link similar reserve definitions in regions that have different naming conventions. It also includes a survey of operating reserves and how they are managed internationally in system operations today and then how new studies and research are proposing they may be managed in the future with higher penetrations of variable generation. The objective is to understand the differences, the commonalities, and to inspire thought on how new methods or metrics would be better suited in this future power system with larger amounts of variable generation.

Abbreviations and Definitions

Operating Reserves	Any capacity available for assistance in active power balance.
Operating Reserve Terms in Capital Letters	General operating reserve terms as defined by this paper as in section 2.
Operating reserve terms in italics	Operating reserve terms as defined specifically by a region or a study
ACE	Area Control Error
AGC	Automatic Generation Control
B/K	Bias (B:North America, K:Europe)
BA	Balancing Area or Balancing Authority
BAAL	Balancing Area ACE Limit
CAISO	California Independent System Operator
CPM	Control Performance Measure
CPS(1&2)	Control Performance Standard
DCS	Disturbance Control Standard
EENS	Expected Energy Not Served
ENTSO-E	European Network of Transmission System Operators for Electricity
ERCOT	Electric Reliability Council of Texas
EWITS	Eastern Wind Integration and Transmission Study
FAL	Frequency Abnormal Limit
FERC	Federal Energy Regulatory Commission
FRC	Frequency Response Characteristic
FRL	Frequency Relay Limit
FRR	Frequency Responsive Reserve
FTL	Frequency Trigger Limit
GW	Gigawatt
Hz	Hertz
ISO	Independent System Operator
ISO-NE	Independent System Operator – New England
LAaR	Loads Acting as a Resource
LESR	Limited Energy Storage Resource
LOLP	Loss of Load Probability
MISO	Midwest Independent System Operator
MW	Megawatt
NERC	North American Electric Corporation
NPCC	Northeast Power Coordinating Council
NYISO	New York Independent System Operator
NYSRC	New York State Reliability Council
QLPU	Quick Load Pick-Up
RSG	Reserve Sharing Group
RTO	Regional Transmission Organization
SPP	Southwest Power Pool
TSO	Transmission System Operator

UCTE	Union for Coordination of Transmission of Electricity
UFLS	Under Frequency Load Shedding
VG	Variable Generation
VOLL	Value of Lost Load
WECC	Western Electricity Coordinating Council
WITF	Wind Integration Task Force
WWSIS	Western Wind and Solar Integration Study

Acknowledgements

The authors would first like to thank Charlton Clark and his colleagues at the Department of Energy's Office of Wind and Hydropower Technologies for their financial support for this project. The authors would like to thank the NERC Integration of Variable Generation Task Force (IVGTF) and the WECC Variable Generation Subcommittee (VGS) for providing industry review of this paper. The team would also like to thank Charlie Smith of UWIG, Mark O'Malley, Damian Flynn and Michael Power of University College Dublin, and Matt Hunsaker of WECC for their review. The team would like to thank Hannele Holttinen of the VTT Technical Research Centre of Finland and the IEA Task 25 team for providing thoughtful and interesting discussion directly related to this topic area. Finally, the team would like to thank Bruce Green of NREL for his editorial assistance.

1 Introduction

Power system operators have a number of responsibilities that focus on maintaining reliability [1][2][3]. System generation must be as close as possible to the system load and electrical losses to ensure that system frequency is maintained at or very close to nominal levels (60 Hz in North America, 50 Hz in Europe and many other areas throughout the world). In areas that are a part of a larger synchronous interconnection it must be ensured that tie-line flows between areas are kept at their scheduled flows. This is achieved through numerous procedures on different time scales using both economic response and deployment of reliability reserves with both centralized control and autonomous response. Power flows must be kept below the maximum limits on equipment such as transmission lines and transformers through the dispatch of generating units and through direct flow control with phase shifting transformers and flexible AC transmission system (FACTS) devices. Voltage levels throughout the power system must also be kept within nominal levels at all locations on a network. This is achieved through reactive power management of generators as well as by controlling transmission system devices including voltage regulators, transformer taps, capacitors, reactors, static VAR compensators, statcoms, and synchronous condensers. Systems must also be able to withstand contingency events with preventive control actions holding reserves and limiting pre-contingency flows so that the system will survive the event and normal operations can be fully restored shortly following these events.

If system conditions could be easily predicted and were constant over all time frames, meeting these objectives would be relatively straightforward. However, many of the properties of the power system, including its generation output, load levels, and transmission equipment availability are both variable and unpredictable. Therefore, additional capacity (generation and responsive load availability) above that needed to meet actual load demands are made available either on-line or on-standby so that it can be called on to assist if load increases or generation decreases, due to unpredictability or variability of the conditions. Likewise, on-line generating capacity that can reduce supply or turn off is required if load decreases or generation increases.[1] This capacity, herein referred to as operating reserves, is utilized for many different reasons and comes in different shapes and sizes. It can be generalized for the purposes of this study as capacity used to maintain the active power (also referred to as real power, but active power will be used throughout the rest of this report) balance of the system.[2] Note that extra generation that is available to increase output (or load that is available to curtail) has historically been more of a reliability need since large generators and transmission lines can and do suddenly fail. It has been far less common for large loads to suddenly disconnect, so downward reserves have historically been less needed for power system reliability. Power systems with large amounts of variable generation/VG (both wind and solar), which can increase or decrease output unexpectedly, may raise the importance of both upward and downward reserves.

As discussed, the variability and uncertainty on systems is what causes the need for Operating Reserves. Variability is the expected changes in power system variables. Operating Reserves might be needed if this

[1] Responsive load and storage able to increase load upon command could also provide this reserve.

[2] The available capacity could be available as active power capacity or reactive power capacity, however, in this study we generally refer to active power reserve. Reactive power imbalance causes voltage differences but it is much more location specific, and requirements or utilizations are not as easily quantified for aggregated systems.

variability occurs at time resolutions that scheduling resolutions are not prepared for. For example, an hourly schedule might hold Operating Reserves for variability that occurs at 5-minute resolution since the hourly schedule is not prepared for that variability. This Operating Reserve may then be deployed by the 5-minute scheduling program to ensure balance of supply and demand. Similarly, the 5-minute scheduling program may hold Operating Reserve for variability at a time resolution faster than 5-minutes. Uncertainty is change in power system variables that is unexpected. Operating Reserves are needed for uncertainty since a different supply-demand profile is needed than what was scheduled. Figure 1 shows an example of variability and uncertainty for VG output. This variability and uncertainty can occur to some degree with all power system variables. These two terms and how we have defined them are very important throughout the rest of this paper.

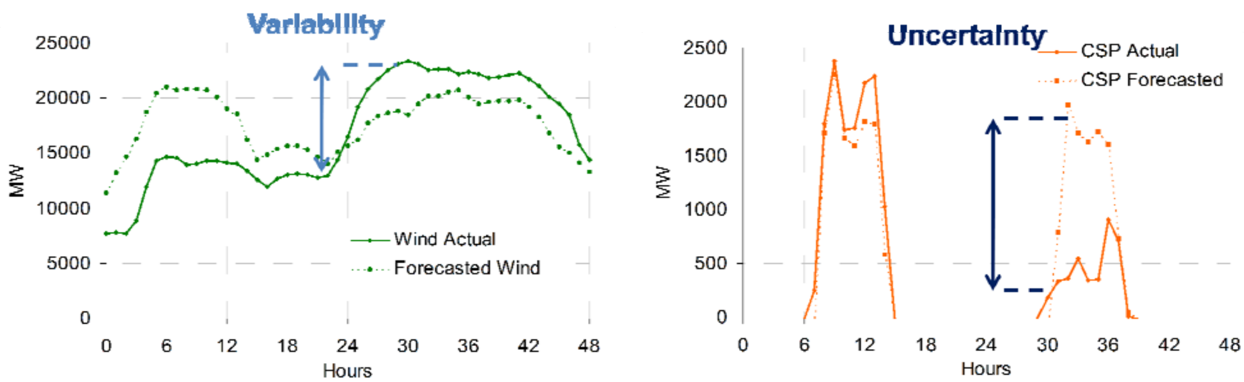

Figure 1: Example of Variability and Uncertainty.

Certain procedures are set forth by different entities on the amount of Operating Reserves required, who can provide them, when they should be deployed, and how they are deployed. The standards are generally based on certain reliability criteria and allowable risk criteria, but often differ, sometimes substantially, from region to region. Many studies have found that high penetrations of VG, such as wind and solar power, have such new characteristics on the power system that these standard rules must be accompanied by innovative methods and adjusted rules and policies to account for the increased variability and uncertainty that it introduces. It may be that using today's standards alone simply cannot capture these characteristics. Different studies have used a number of different methods, all attempting to answer the same questions.

This paper is a summary of operating reserves, their uses, and methodologies with emphasis on how their requirements may change with significant penetrations of VG. First, a brief overview is provided to describe the basics of power system operations, categorizing the many uses of operating reserve, describing each of the categories, and the impacts that VG may have on each category. The categories of operating reserves that we have defined have different naming conventions depending on the region, and effort must be made to understand a common term for each type. Section 3 discusses the rules and procedures of how system operations of today are defining and utilizing operating reserves in North America and in Europe, specifically looking at the North American Electric Reliability Corporation (NERC) and the European Network of Transmission System Operators (ENTSO) and each of their standards. We also emphasize some significant regional differences within North America. Section 4 discusses some of the new innovative methods that have been developed in large studies and fundamental research on both actual future systems and test systems with high VG penetrations. Finally, a summary is provided of the many differences and commonalities, and then recommendations for future research and current limitations are discussed. One objective of the report is to describe both commonalities and differences between actual system policies in different regions and differences between study methods used. Although there can be differences in each system based on its size, generating fleet, load characteristics and other factors, all of the regions and all of the studies are essentially requiring and utilizing operating reserves for the same reasons. Therefore, it is interesting to see requirements and methods that differ so much. This paper attempts to bring clarification between the policies of different regions and the methods of different studies.

2 Power System and Operating Reserves Overview

Figure 2 shows a general power system load pattern for a single day. It shows the different time frames where different strategies are used to ensure that the load is balanced. Forward scheduling of the power system includes schedules and unit commitment directions to meet the general load pattern of the day. Load following is the action to follow the general trending load pattern within the day. This is usually performed by economic dispatch and sometimes involves the starting and stopping of quick-start combustion turbines or hydro facilities. Regulation is the balancing of fast second-to-second and minute-to-minute random variations in load or generation. This is done by centralized control centers sending out control signals to generating units (and some responsive loads) that have the capability to rapidly adjust their dispatch set points. These strategies represent the balancing during normal conditions of the power system. The load is never constant and therefore each of these strategies helps correct the load balance. Also, conventional generation output may vary in different time scales as well and can further impact the generation and load balance. Lastly, the load forecast is never 100% accurate and each of these reserves is used to help mitigate the effects of load forecast errors.

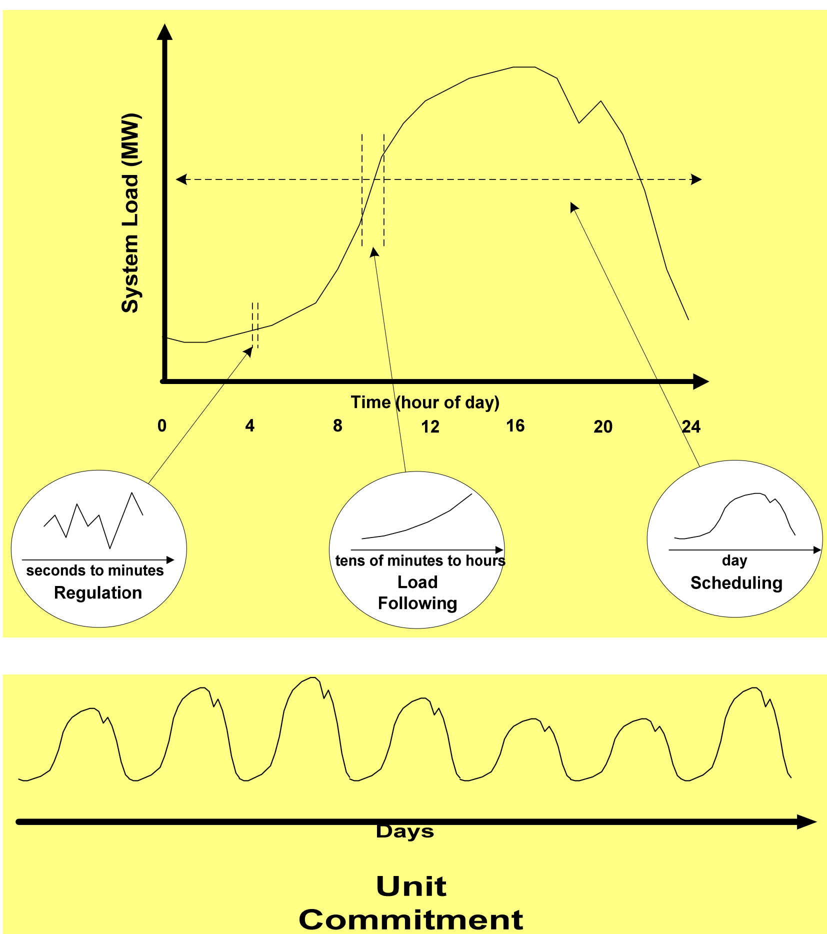

Figure 2: Power system operation time-frames.

Figure 3 shows how reserves are used in a coordinated way to respond to an emergency event and represents the North American procedures. These emergencies can be caused by generation forced outages or other component outages (e.g., transmission lines or transformers), both of which cause near-instantaneous changes on the power system. Generation outages cause a deficiency in generation to serve the load and will usually result in frequency decline and negative Area Control Error (ACE). Under-frequency can lead to involuntary load shedding, and can cause damage to machines if the frequency strays too far from its nominal level. ACE, given in MW, is the power balance error of a balancing authority (BA). It is calculated as the total interchange schedule error of the balancing authority area with consideration of the balancing area's response to frequency deviations (described in more detail further on). Transmission outages will result in a change in power flow that may cause overloading of lines or voltages that are not within their nominal levels. The common procedure for managing the possibility of transmission outages is by running security-constrained optimal power flow, which ensures that with no corrective action needed, power flows are still within their appropriate limits following transmission outages.

During loss of supply events, additional supply needs to respond to the disturbance immediately. As can be seen in Figure 3, this includes a number of different responses that vary by response time and length of time the response is sustained. Initially, when the loss of supply occurs, synchronous machines must supply kinetic energy to the grid, and by doing so, slow down their rotational speeds and therefore the electrical frequency. This inertial response that comes from synchronous generators and synchronous motors helps slow down the frequency decline. In other words, the more inertia on the system, the slower the rate of frequency decline. During this decline in frequency, generators will automatically respond to the change in frequency through governor response, and some load response will balance the generation and load at some frequency less than the nominal frequency. Spinning reserve that is synchronous to the grid and unloaded from its maximum rating and non-spinning reserve, which can be off-line but able to be synchronized quickly, are both deployed to fill the gap in energy needed from the loss and restore the frequency back to its nominal level. Furthermore, many areas have spot prices that may increase during supply shortages and incentivize response from resources that can assist in the event. Lastly, supplemental reserves are deployed with slower response to allow the other reserves to be unloaded once again so that the system can be again secure for a subsequent event. For over-frequency events, though not as common, a similar response might occur, but a reduction in output would be needed rather than an increase.

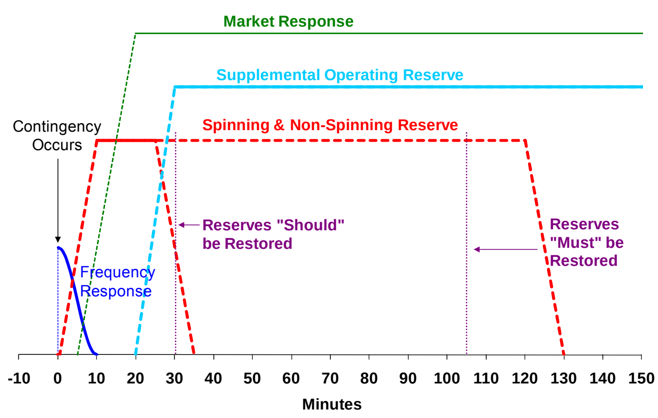

Figure 3: Reserve deployment.

In some cases, severe events are slower in nature. This may include constant increase or decrease in generation and/or load. Figure 4 shows one event where wind power steadily decreased for about four hours (yellow trace) [4]. When an event like this occurs where the wind power may be decreasing simultaneously with an increase in load, the effects are exacerbated. In most systems, the rules concerning how reserves can be used to respond to this type of event are not yet fully established.

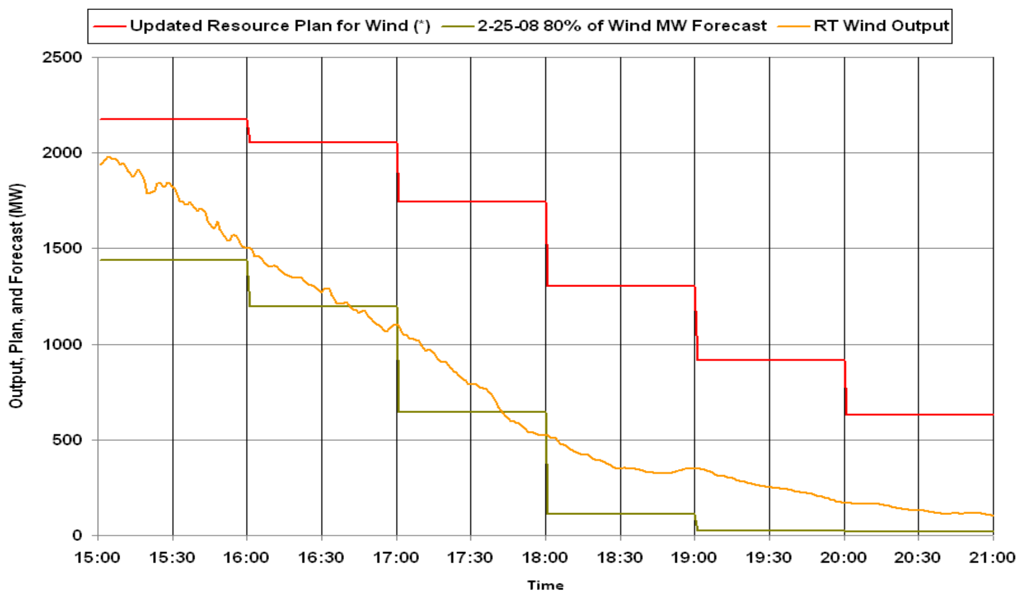

Figure 4: Wind ramp event, ERCOT February 26, 2008.

The characterizations of these different operating reserves can have different terminology and definitions depending on the system. Recent work has similarly categorized the different operating reserve types, for example [5] [6] [7]. In Robours et al. 2005 [5], the authors discussed the details of European TSO policies (including two U.S. regions) with a focus on the timing and the required amounts of reserved used for contingencies. The same authors expanded this in [6] with more details on the rules, additional regions and countries, and included voltage control services. A report including authors of the various TSOs throughout the world [7] included additional operating reserve services from a number of countries and some more detail on the different categories of operating reserve in existence today. In this paper, we attempt to do the same, but with consideration of future requirements with high penetrations of VG. We discuss the categorizing of operating reserve for selected reasons, but emphasize that our categorization is an example of how to distinguish between the different services and that the categorization can be done in many different ways. It also doesn't necessarily mean that requirements are enforced for each category and are completely decoupled from each other category. We separate the categories mainly for ease of understanding and demonstration.

Operating reserves can be characterized by their response speed (ramp rate and start time), response duration, frequency of use, direction of use (up or down), and type of control (i.e., control center activation, autonomous/automatic, etc.). Some operating reserves are used to respond to routine variability of the generation or the load. These variations occur on different time scales, from seconds to days, and different control strategies may be required depending on the speed of the variability. Other operating reserves are needed to respond to rare unexpected events such as the tripping of a generator. Another way of classifying the operating reserves could be based on whether they are deployed during normal conditions or event conditions. Normal conditions can be based on both variability and uncertainty, but are occurrences that are continuously taking place. Events can occur whether they can be predicted or not. There is a difference in the standby costs and the deployment costs for each reserve category based on how frequently they are used. This distinction along with the technical requirements leads to certain technologies being more suitable for different Operating Reserve types than others.

Both the normal and event response categories can be further subdivided based on the required response speed. As discussed, some events are essentially instantaneous and others take time to occur. Different qualities of these reserves are needed for different purposes. For instance, instantaneous events need autonomous response to arrest frequency excursions. The frequency then must be corrected back to its scheduled setting and the system's ACE must be reduced to zero during both instantaneous and non-instantaneous events. Lastly, there has to be some amount of reserves that can replace the operating reserve after they are deployed to protect the system against a second event. For the non-event reserves, they are typically meant to have zero net energy for a particular time period (over some time, they will have equal positive deployments and negative deployments in energy), and therefore we don't consider the need for the replacement of these operating reserves in our discussions. Figure 5 displays the classification system we have just discussed with an example of their naming convention. It should be noted again that this classification can be done a number of different ways. Our examples include names and classifications that we have used throughout this paper. As you can see, the classification is not tied to any one system terminology, and also includes a number of types that may not even yet be in existence in current systems.

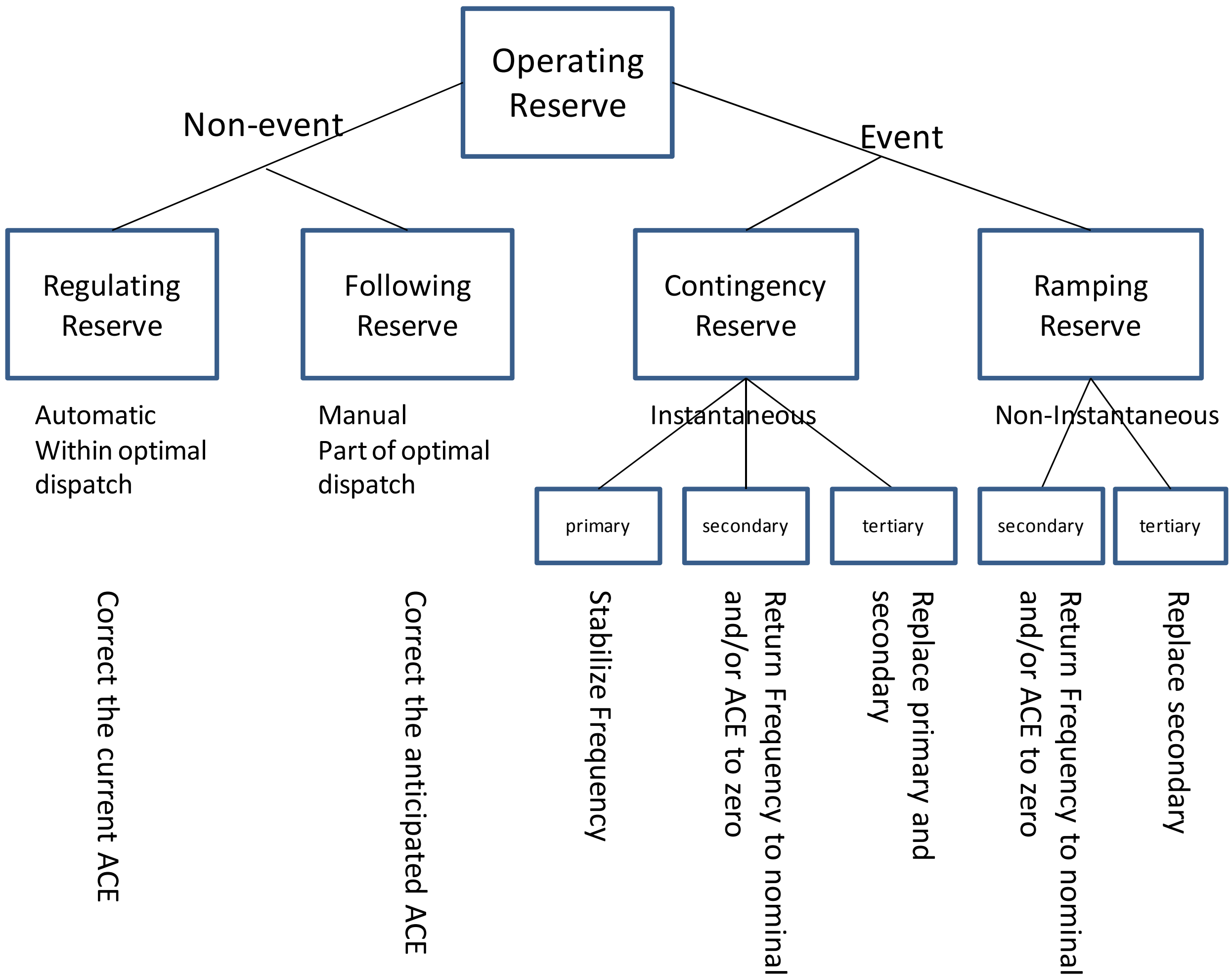

Figure 5: Example of operating reserve categories and how they are related.

The tree in Figure 5 explains how the Operating Reserve types relate to each other with the higher level categories including those below. From the highest level, we define Operating Reserves as any type of capacity being used to support active power balance. This is then separated into non-event reserve and event reserve. Events include things that are severe and rare and non-events are continuous events that happen so often they are not distinguishable from one another. The non-event reserve is separated by speed with Regulating Reserve being faster and Following Reserve being slower. The speed that separates these two reserve categories varies from system to system with Regulating Reserves operating automatically within the shortest optimal dispatch or market clearing interval and Following Reserve operating as part of the optimal dispatch or market clearing interval. Furthermore, their purpose can be differentiated with Regulating Reserve being used to correct the current imbalance whereas Following Reserve is used to correct the anticipated imbalance.

For event reserve, we classify Contingency Reserve and Ramping Reserve which are separated by speed as well. The speed that separates these two reserve categories is whether they are used for instantaneous events (or near-instantaneous, for example within a few cycles) or for non-instantaneous events. Further sub-categories are also added below the two event reserves.

For the instantaneous events, or Contingency Reserves, a certain portion of Primary Reserve must be automatically responsive to the event to ensure the system's frequency deviation is arrested and the load balance is maintained soon after the event. Primary Reserve must respond immediately following the event to avoid extreme frequency deviations that can cause damage or involuntary load shedding. Since this response stabilizes the frequency at some other level, Secondary Reserve is then deployed to return the frequency back to its scheduled setting.

Finally, Tertiary Reserve assists in replenishing the Primary and Secondary Reserve that was deployed for the event so that within some time following the event there is enough to respond to a second event. The actual time allowed for full response from these reserve types may vary from system to system, but generally the need for Primary Response is tens of seconds, Secondary is a few minutes, and Tertiary is tens of minutes. For Ramping Reserve, there are somewhat different needs. Due to the slowness of these events, the automatic frequency responsive need is not necessary. The Secondary Reserve here would be to correct frequency or ACE. Tertiary Reserve would be needed to protect against a subsequent event that were to occur in the same direction. The full response for these types might be quite different than their counterparts under Contingency Reserve. Table I shows all the defined reserves referenced throughout this report and the usage of each.

Table I: Operating reserves as defined throughout the rest of this report and their usage

Name	Use	Common Terms
Operating Reserve	Any capacity available for assistance in active power balance.	
Non-event Reserve	Capacity available for assistance in active power balance during normal conditions, or those that occur continuously.	
Regulating Reserve	Capacity available during normal conditions for assistance in active power balance to correct the current imbalance that occurs, is faster than economic dispatch optimization, is random, and requires automatic centralized response.	*Regulating reserve, regulation, load frequency control, secondary control.*
Following Reserve	Capacity available during normal conditions for assistance in active power balance to correct future anticipated imbalance, is not faster than economic dispatch optimization, and does not require automatic centralized response.	*Load following, following reserve, schedule reserve, dispatch reserve, balancing reserve.*

Event Reserve	Capacity available for assistance in active power balance during infrequent events that are more severe than balancing needed during normal conditions.	
Contingency Reserve	Capacity available for assistance in active power balance during infrequent events that are more severe than balancing needed during normal conditions and are used to correct instantaneous imbalances.	*Contingency reserve (spinning and non-spinning reserve).*
Ramping Reserve	Capacity available for assistance in active power balance during infrequent events that are more severe than balancing needed during normal conditions and are used to correct non-instantaneous imbalances.	*Ramping reserve.*
Primary Reserve - Contingency	Portion of Contingency Reserve that is automatically responsive to instantaneous active power imbalance and stabilizes system frequency.	*Primary control reserve, frequency responsive reserve, governor droop.*
Secondary Reserve - Contingency	Portion of Contingency Reserve that is not automatically responsive to the instantaneous active power imbalance and corrects frequency to nominal and/or ACE to 0.	*Secondary control reserve, spinning reserve.*
Tertiary Reserve - Contingency	Portion of Contingency Reserve that is available for assistance in replacing Primary and Secondary Reserve used during a severe instantaneous event so that they are available for a subsequent instantaneous event that occurs in the same direction.	*Tertiary control reserve, replacement reserve, supplemental reserve.*
Secondary Reserve - Ramping	Portion of Ramping Reserve that is used to correct the imbalance of a severe non-instantaneous event and corrects the frequency to nominal and/or ACE to 0.	*Ramping reserve.*
Tertiary Reserve - Ramping	Portion of Ramping Reserve that is available for assistance in replacing Secondary Reserve used during a severe non-instantaneous event so that eventually Secondary Reserve are available for a subsequent event that occurs in the same direction.	*Replacement reserve for ramping reserve.*

Different generation and responsive load technologies are better at providing different types of operating reserves. Conventional thermal and hydro generating units are inherently limited in the amount of spinning reserve they can provide by their ramp rates, although hydro generating units and combustion turbines may have faster ramp rates than steam turbine generators. These units can only supply as many MW of reserves as they can ramp in a specified time based on its regions' rules. Some internal combustion engine driven generators, aero-derivative combustion turbines, and hydro plants can start fast enough to provide non-spinning reserves even if they are not currently operating. Nuclear units have historically been built for base load and therefore usually do not provide operating reserves. Primary Reserve can be provided by any generator with a governor that can respond rapidly and that can maintain that response as frequency declines. Nuclear plant governors are typically blocked, preventing them from providing frequency responsive reserve. Large thermal plants operating with their valves fully open to maximize efficiency (sliding pressure or boiler follow mode) effectively disable the governor and do not provide frequency response either. Some combustion turbines' output declines with frequency since the compressors move less air as the speed decreases, also reducing their ability to provide frequency response. Some research is looking at the potential for wind power to provide certain types of operating reserve [8] [9].

Responsive load can be an ideal provider of many of the different types of Operating Reserves. Response can be faster than generation since full response is achieved as soon as the load control breaker opens; no ramping is required for most responsive loads. Response can be automatic as well if an under-frequency relay is used with a trip setting in the governor response range. Different loads can be set to trip at different frequencies, providing a frequency droop curve that mimics generator governor response. The contingency reserve characteristics of constantly standing ready with relatively infrequent and short actual deployments also match the response preferences of many responsive loads.

Regulating Reserve and Following Reserve are typically provided by units that are on the margin for providing energy economically. Regulating units must also have infrastructure set up so that they can receive and respond to automatic generation control (AGC) signals. Regulating Reserve resources also should have fairly fast and accurate response rates to adjust output to quickly changing signals. Recently, new market rules are allowing demand and limited energy storage resources to provide Regulating Reserve. This situation is continuously being evaluated so that more and more resources that are capable of providing any of these services are not precluded. The definition of the operating reserve should be technology neutral, meaning it should describe the desired response, not the type of resources that are allowed to provide that response.

Since the advent of electric power industry restructuring, many regions throughout North America and Europe have developed ancillary markets for some Operating Reserves, along with their functioning energy markets. In many cases, these are hourly or faster markets that set prices based on the marginal cost to provide the type of Operating Reserve. In many markets, these Operating Reserves are co-optimized with those of the energy markets. This way, the most efficient scheduling is performed with all services. Prices in the Operating Reserve markets then can reflect the cost of providing the service as well as the lost opportunity costs. A lost opportunity cost reflects the fact that if a unit is asked to supply reserves, it may be losing out on profit to supply energy or possibly another reserve service [10][11]. The lost opportunity cost would normally be given back to the unit in this situation and often will set the Operating Reserve price for all resources supplying the service. Currently, Contingency Reserves, which may include Secondary and/or Tertiary Reserves, and Regulating Reserves, are the most common Operating Reserves to have dynamic hourly (or shorter) markets. The way the markets are designed is very important when evaluating the different operating reserve practices.

In recent years, much interest has been focused towards increased amounts of renewable generation on the power system. This includes a large portion of wind and solar power, both of which are referred to as VG. VGs have unique characteristics because their varying energy source can lead to a maximum generation potential that varies on different time scales and cannot be perfectly predicted. Many studies have shown that as these VG resources increase on power systems, it will impact the operation of the system [12] [13] [14].

Increases in VG can have impacts on many of the different Operating Reserve types mentioned. During normal conditions, the variability and uncertainty of VG will add to the variability and uncertainty of load and other generating resources in each of the different time scales. The increased variability that VG brings can increase the need for Following Reserve and Regulation Reserve, depending on the time scale. Also, a collection of VG units can contribute to a ramp event either by itself or with load. This may impact how Ramping Reserves are considered. Large VG forecast errors combined with load forecast errors are much more common and can be more severe compared to load forecasts alone. Such issues can be cause for reliability concerns, similar to those described as contingency events. However, their characteristics relating mainly to the speed at which they occur show some distinct differences. Therefore, it is important to hold the right type of Operating Reserves when responding. Next, we discuss in more detail each type of reserve discussed and relate how VG may have an impact on their requirements.

2.1 Regulating Reserves

Regulating Reserve covers the continuous fast and frequent changes in load and generation that create energy imbalance [15]. It is the finest scale of balancing done during normal conditions. It is used to correct the current imbalance caused load or generation varying within a period that has been scheduled by the shortest applicable market or economic dispatch interval. It is also used to correct the current imbalance from the total load or generation that differs from the forecasted condition. It corrects the current imbalance. In some areas, the shortest scheduling interval may be up to an hour and in others this may be as short as 5 minutes. What this means is that if the system operator dispatched units thinking the net load was moving in a certain direction, and the magnitude or direction is different than anticipated, the Regulating Reserve must be used to correct the mistake before the next economic dispatch cycle is finished. This could increase required Regulating Reserves if the very short-term windforecasts are not accurate.

In many isolated systems, this reserve is provided by governor response (e.g., U.K. and Ireland, other larger island systems). The resources with governors can then cover the normal balancing needs inside the economic dispatch interval automatically. In large interconnected systems with multiple balancing regions (e.g., North America and Continental Europe), normal imbalances (non-events) usually do not trigger frequency response due to the size of the system and the dead bands placed on the governor systems and therefore, the governor or frequency response control are only deployed during larger contingency events. Instead, these areas have units that have AGC in which they can be automatically dispatched centrally by the control center operator [16] [17]. The signal is not solely controlling frequency, but is in fact correcting the balancing area's imbalance based on its ACE. ACE is the MW imbalance at any given time and is calculated according to the equations below. The difference between the total actual and scheduled interchange flows ($NI_A – NI_S$) is the total imbalance of the individual balancing area's generation (Σ Pg_i) and load plus losses (Σ Pd_j + LOSSES) at any instant. However, since balancing areas are expected to respond to frequency deviations, which are based on the imbalance of the entire interconnection, a frequency component is added to the equation (F_A: actual frequency; F_S: scheduled frequency) to ensure that each balancing area helps restore frequency to the nominal value during a frequency excursion. The ACE calculation describes the individual balancing area's power imbalance while still expecting it to respond to frequency deviations, based on the area's frequency bias setting B, when the frequency excursion may have been caused by another area. This is called Tie-Line Bias Control and is generally the way units on Regulating Reserve are used in large interconnected systems. In an isolated system, the first component of ACE would be zero since there is no actual or scheduled interchange.

$$ACE = (NI_A - NI_S) - 10B\,(F_A - F_S)$$

$$NIA - NIS = \sum_{i=1}^{NG} Pg_i - \left(\sum_{j=1}^{ND} Pd_j + LOSSES \right)$$

Diversity inside the dispatch interval time frame reduces the power system Regulating Reserve requirements. Unlike the highly correlated daily load pattern, the minute-to-minute variability of individual loads is highly uncorrelated. The relative Regulating Reserves needed to achieve acceptable performance are consequently lower for larger balancing areas than for smaller balancing areas. Diversity is also important in understanding the impact VG has on Regulating Reserve requirements. For example, one wind turbine could vary by quite a bit on a minute-to-minute basis. However, because the minute-to-minute variability of individual wind turbines are not correlated, once aggregated with multiple turbines, the per unit output of the entire wind plant is much less variable in this time frame. When looking at an entire system, a large number of VG plants that are not geographically close to one another will achieve even further decreases in per unit variability. Figure 6 is based on [18] and shows that; comparing 138 turbines with 14 turbines (a subset of the same wind plant) reduces the 1-minute step change and standard deviation from 1.2 to 0.5 per unit and 2.1 to 0.8 per unit, respectively.[3] This is because the wind speed correlations occurring on a minute-to-minute basis are very location specific and therefore, dispersed wind plants generally have no correlation to one another on this time frame. The Regulating Reserve increases due to variability on this time frame may show increases, but these will depend on how many wind plants there are and their geographic proximity. It is important to understand the correlation of this variability with all sources that cause the need for Regulating Reserve.

[3] The average is the average of absolute differences from 1-minute to -hour steps, and that the % is % of installed capacity.

		14 turbines		61 turbines		138 turbines		250+ turbines	
		(kW)	(%)	(kW)	(%)	(kW)	(%)	(kW)	(%)
1-second									
	Average	41	0.4	172	0.2	148	0.1	189	0.1
	Std Deviation	56	0.5	203	0.3	203	0.2	257	0.1
1-minute									
	Average	130	1.2	612	0.8	494	0.5	730	0.3
	Std Deviation	225	2.1	1,038	1.3	849	0.8	1,486	0.6
10-minute									
	Average	329	3.1	1,658	2.1	2,243	2.2	3,713	1.5
	Std Deviation	548	5.2	2,750	3.5	3,810	3.7	6,418	2.7
1-hour									
	Average	736	7.0	3,732	4.7	6,582	6.4	12,755	5.3
	Std Deviation	1,124	10.7	5,932	7.5	10,032	9.7	19,213	7.9

Figure 6: Wind power step-change magnitude and standard deviation for different wind turbine sets.

The uncertainty of VG can also have an impact on Regulating Reserve requirements. Regulating Reserves are utilized to meet actual imbalances that occur while all units that are not providing Regulating Reserve are fixed (or ramping toward their next economic set-point). Therefore, if the forecasts used to schedule units not on Regulating Reserve are incorrect, units on Regulating Reserve must correct this error. For example, most U.S. ISOs and some large non-ISO regions have economic dispatch programs that are run every 5 minutes, and may take up to 5 minutes to compute the optimization and transfer both the input and output data. Although load forecast in these periods have been pretty accurate due to its trending nature, wind power forecasts usually are most accurate when using persistence forecasts (e.g., the last available output reading is the forecast for the next time period). Figure 7 shows this relationship of both variability and uncertainty impacts that can impact Regulating Reserve. The plot shows actual wind data for a twenty-minute period. The period of interest is 12:10 to 12:20. For the 5-minute economic dispatch period of 12:10 to 12:15, a forecast was used based on the actual wind data from 12:00 to 12:05, since the economic dispatch had to start at 12:05 due to the 5 minutes it needed to collect input, compute its optimization, and transfer the output signals, and this was the latest information it could use as the input. The interval 12:15 to 12:20 similarly used the average wind output from 12:05 to 12:10. Figure 7 shows these schedules, as well as what the actual average for all intervals were. Since the Regulating Reserve would attempt to follow the total system ACE (which includes the wind forecast imbalance), it can easily be seen how the Regulating Reserve is used to assist in imbalance caused by both variability and uncertainty. Figure 8 shows this more clearly.

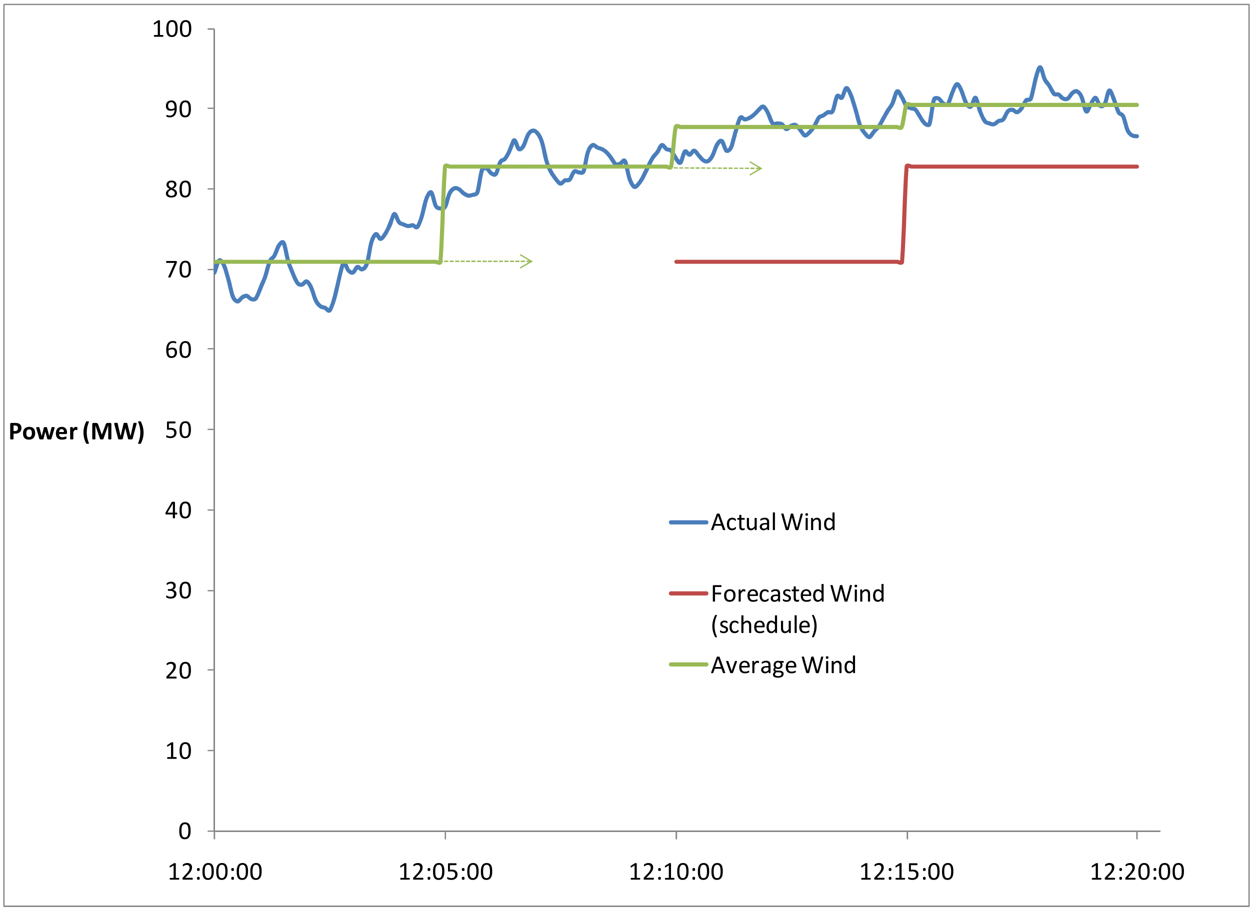

Figure 7: Example of wind's uncertainty and variability impacts on Regulating Reserve.

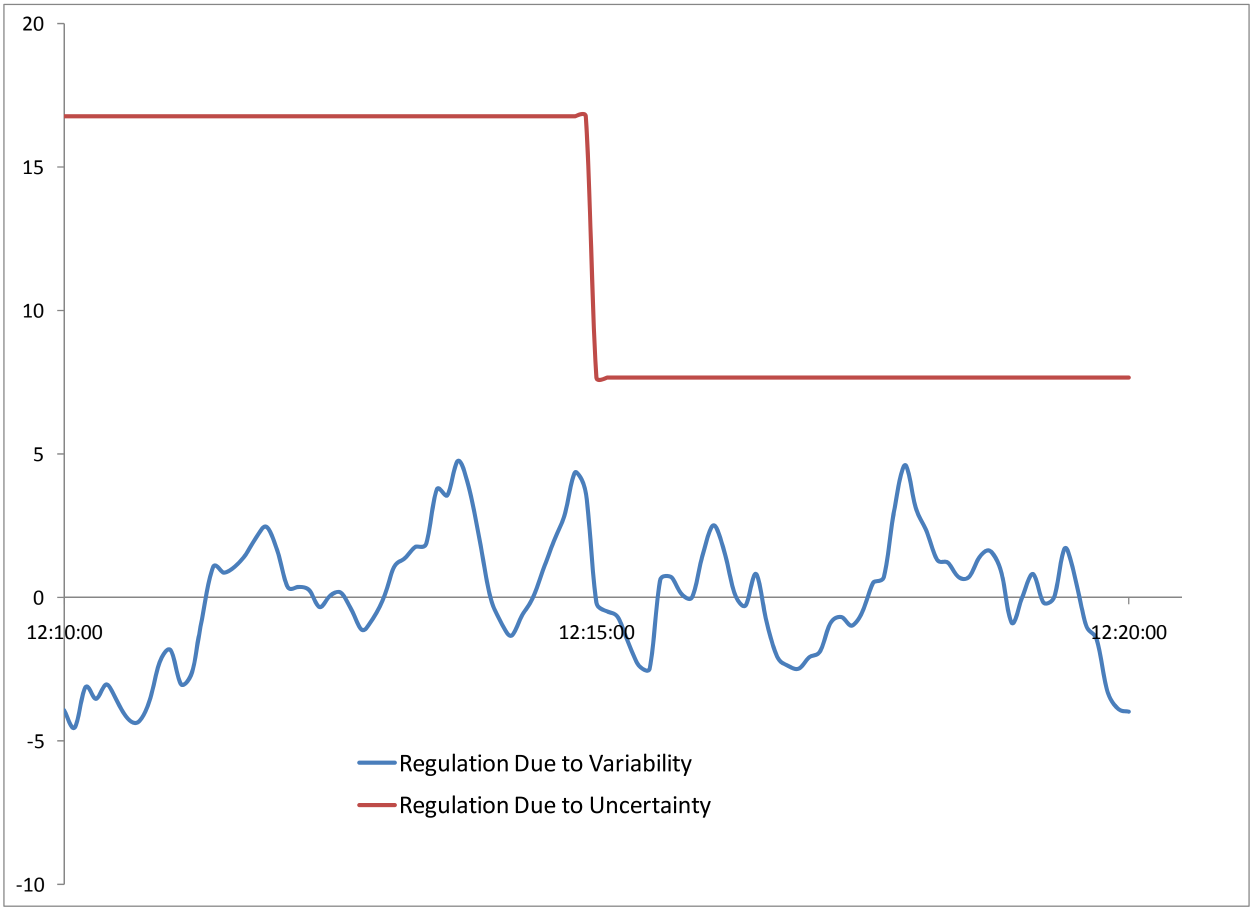

Figure 8: Actual Regulating Reserve needs due to variability and uncertainty.

It is very difficult (or impossible) for most resources to follow the instantaneous ACE values on very high time resolution. Figure 9 shows a typical fleet of large coal-fired generators following the load (shown as regulation up or down signals) in which a lag can be observed. It can also be seen that there is less high frequency noise in the generation signal compared to the load signal. It creates much unnecessary wear and tear impacts on the generators when rapid upward and downward movement is used. It also is arguable whether this type of behavior actually benefits system reliability. For these reasons, many AGC schedules will filter and integrate the ACE signal before distributing control schedules to Regulating Reserve units. For example, the ACE distributed to the Regulating Reserve units in many regions may be calculated as in the equation below where α is a proportional gain and T is the integral period. Typical values are α=(0,0.5) and T=(50,200)seconds. Recently, some energy storage devices have showed ability to follow faster, unfiltered ACE signals, with little or no wear and tear caused by doing so [19] [20].

$$ACE\_NEW(t) = \alpha * ACE(t) + \frac{1}{T}\int ACE(\tau)d\tau$$

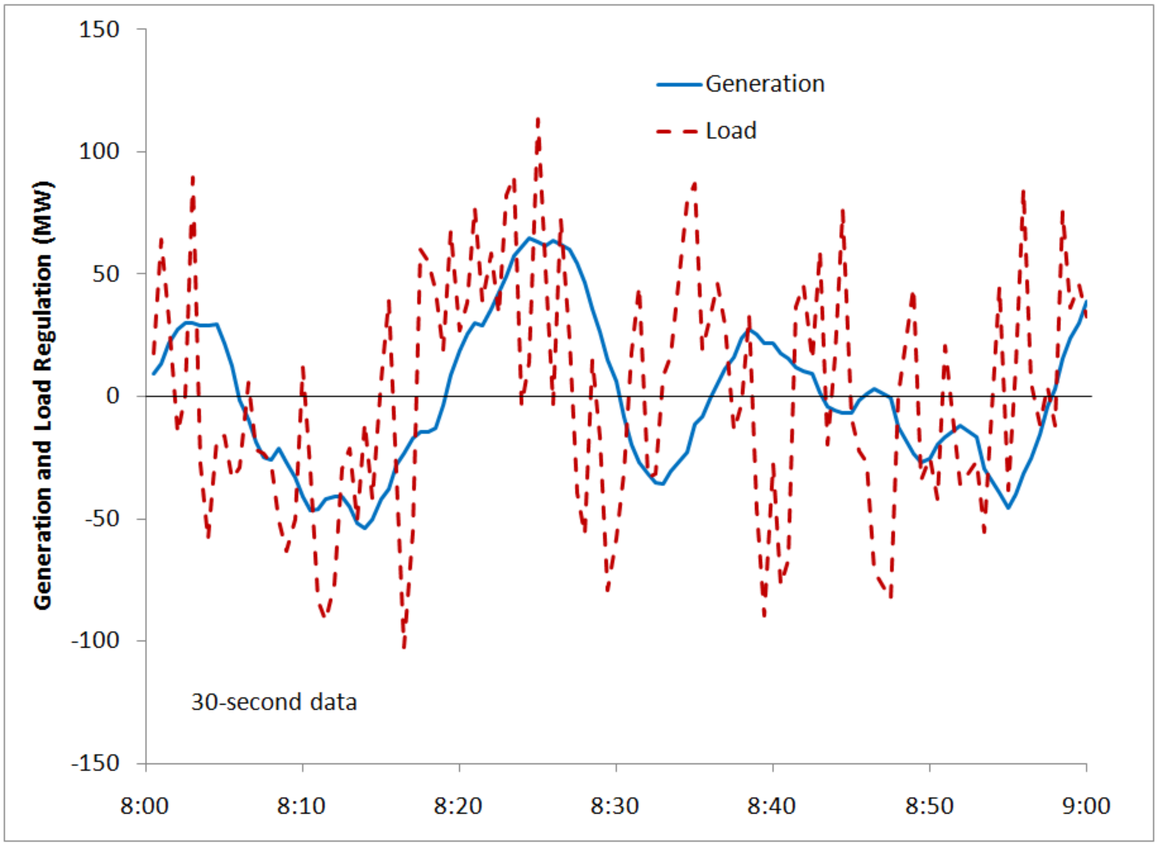

Figure 9: Typical thermal-generation fleet performance following Regulating Reserve signals.

2.2 Following Reserves

Following Reserve is very similar to Regulating Reserve, but on a slower time scale. It is needed to accommodate the variability and uncertainty that occur during normal conditions. Our definition of Following Reserve represents the movements that are reflected in the economic dispatch to correct an imbalance that will occur in the future. Variability and uncertainty requiring Following Reserve impacts are much less random, but larger in magnitude than Regulating Reserve. Following Reserve covers both typical load and VG patterns and inter-schedule interval variability. It also includes the uncertainty in forecasts from previous scheduling intervals to updated scheduling intervals, and with good information can be better predicted. For example, the uncertainty from a day-ahead or hour-ahead unit commitment would be met with Following Reserve in the real-time economic dispatch. In today's paradigm, Following Reserve is often a byproduct of the energy markets. Generally, the load follows a similar path every day, and therefore, the ramping and energy needed to follow this load can usually be easily supplied by energy markets. On some occasions, dedicated Following Reserves may be required. A hypothetical example of how this may occur can be seen in Figure 10 [21]. From 8:00, the load (or net load) has a very fast upward ramp. The cheap base load unit cannot meet the load itself because of its slow ramp rate. Therefore, an expensive peaking unit is dispatched with a fast ramp rate. The question in this case is whether the peaking unit is providing an ancillary service or if it is just providing energy.

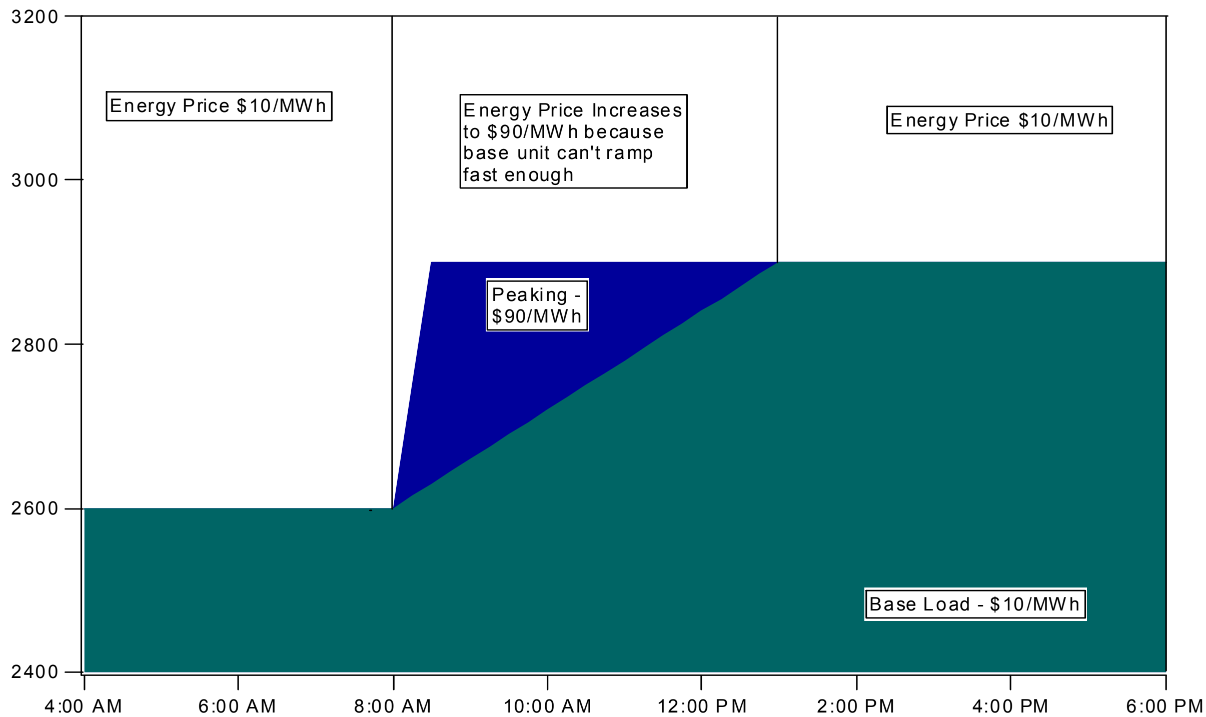

Figure 10: Insufficient ramping capability may result in the need for peaking response inflating the energy price.

It is still unclear if the trending patterns that are associated with load following are still intact with high penetrations of VG. The VG addition to the generation stack makes following the net load less predictable, and in some hours, units with higher ramp rates or more flexibility need to be available. Some of the large differences may come in how the initial unit commitment is made considering Following

Reserve with and without VG. For example, Figure 11 shows seven hours of a typical load ramp. The blue trace is the actual load and the red trace is the hourly average. The red trace might be the schedule that is used when making the unit commitment decision. The trend is continuously increasing and units can be committed sometime during the hour as prepared. For hour 17 for example, the first half of the hour is below the hourly average and the second half of the hour is above the hourly average. During real-time operations, the resources can be directed to start toward the middle of the hour rather than to the top of the hour based on the actual conditions.

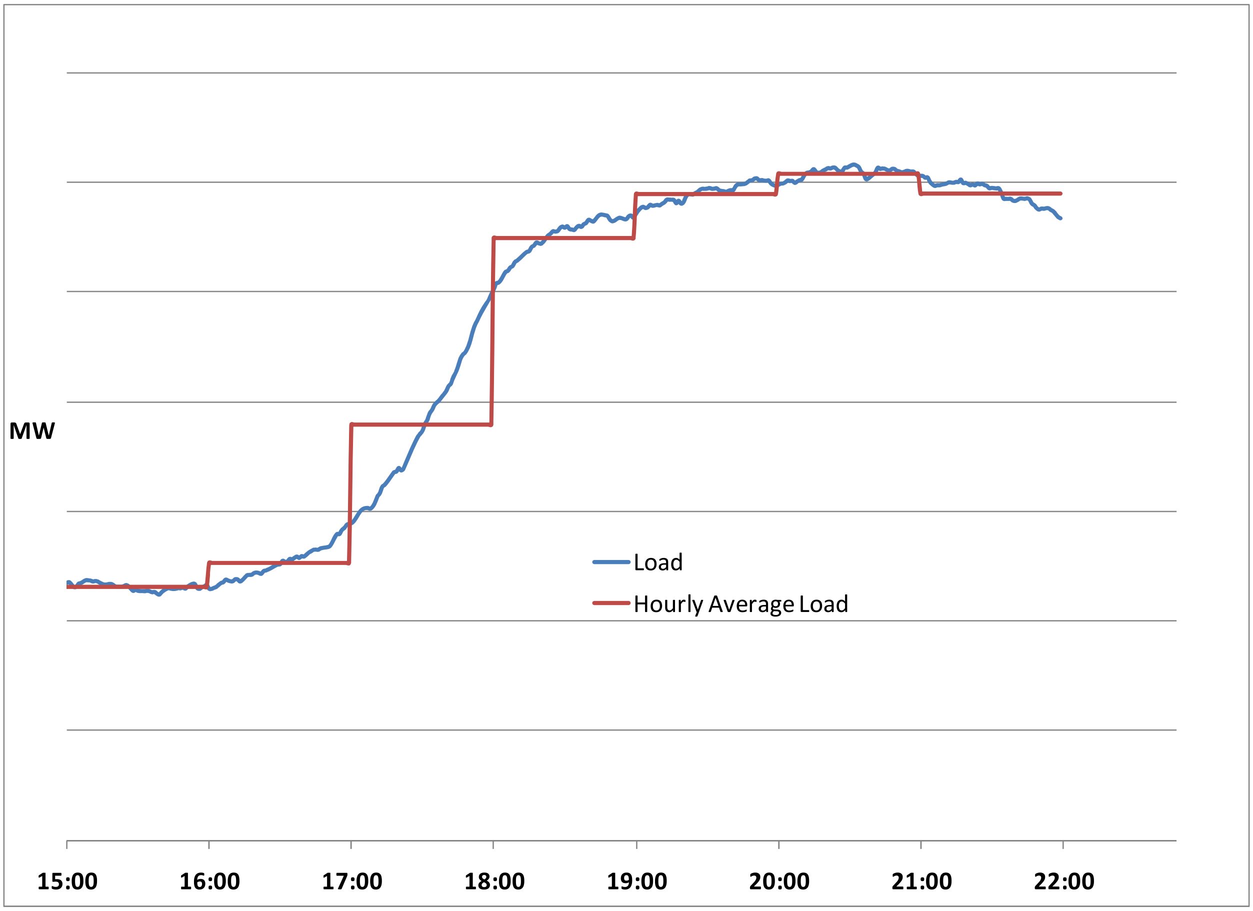

Figure 11: Typical load ramp and hourly average schedule to follow load.

Figure 12 shows the same plot with significant VG subtracted from the load to display the net load. The trends now are not monotonically increasing or decreasing as they were before. For example, in hour 18, the unit commitment would schedule to the hourly average and in hour 19, it might begin to turn off units. The big spike during hour 18 may not have been easily handled by the online units. This may show that certain Following Reserve capacity with specified ramping capability may be needed when making a unit commitment decision to be able to accommodate the sub-hourly variability. Alternatively, if the unit commitment used sub-hourly resolution when optimizing, this issue may not need to be dealt with explicitly. In other words, the Following Reserve is accommodated by advanced scheduling. It then becomes more of a market issue on whether the units performing the Following Reserve services deserve to be paid more than the ones who do not, and whether those units set the energy price that all other generators receive.[4] These same issues described for the variability can also be seen by the increased uncertainty of VG. Errors occur in the prediction of the net load and these errors will need to be accommodated with the Following Reserve. For example, the amount of resources committed from the day-ahead may need to be increased in case the day-ahead prediction of the net load is incorrect. The need for Following Reserve from the increased uncertainty may be required regardless of the length of the commitment or dispatch interval.

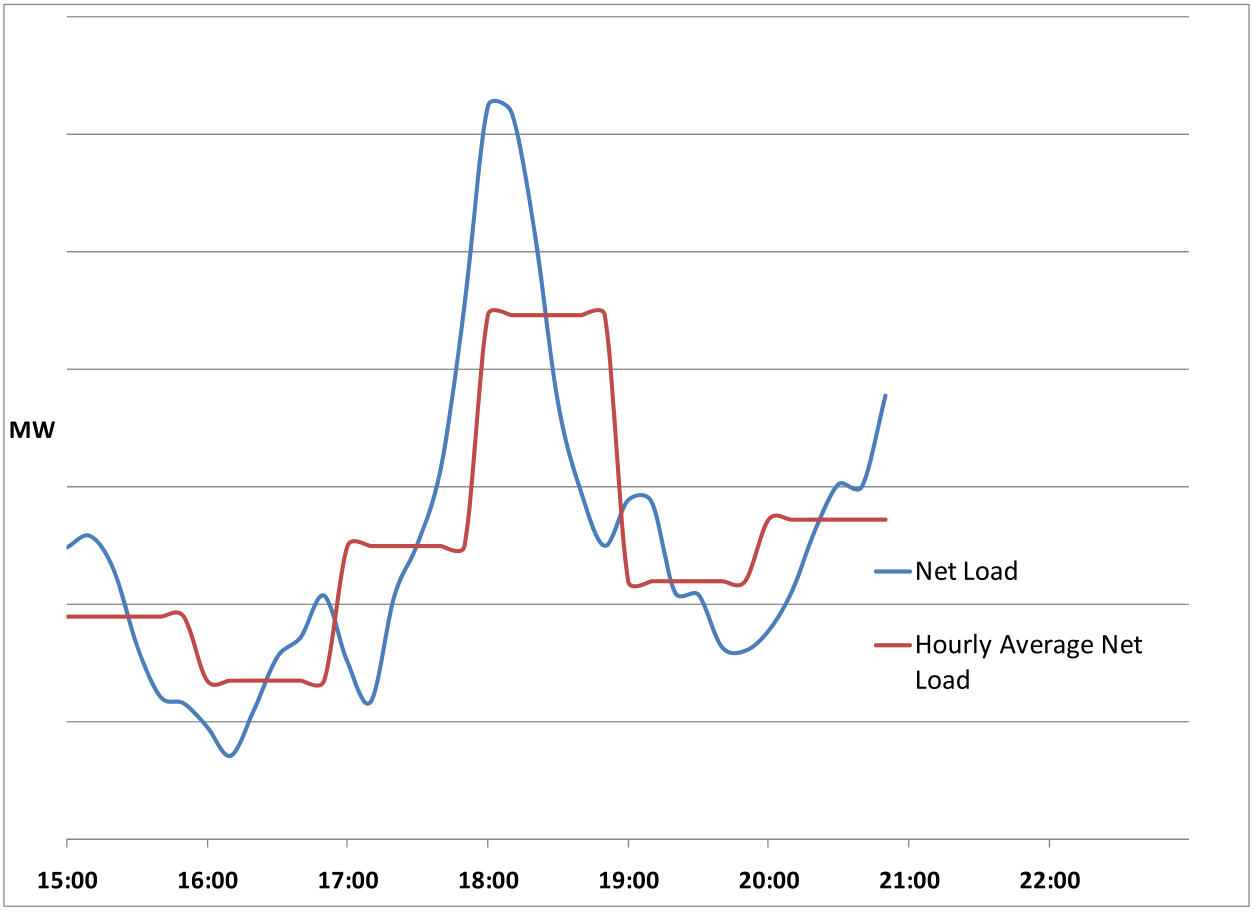

Figure 12: Load following needs for a more variable net load.

2.3 Contingency Reserves

Unlike Regulating and Following Reserves, Contingency Reserves are called upon during rare sudden events. The events usually considered are a large loss of supply either from generating resources or large transmission lines carrying imports, but more generally can consider loss of large blocks of load as well. Contingencies occur quickly and much of the reserves must act immediately. Figure 13 shows a typical response to a large loss of a generating unit. Immediately following the event, the inertia of synchronous rotating machines will either supply or absorb the energy difference, respectively for loss of supply and

[4] In two-settlement markets, real-time prices are paid only to resources that have differed from their forward agreed schedules. They will only get paid for the quantity of energy that is different from their day-ahead schedules.

load. Following this initial response, generator governors sense the frequency change and begin to adjust the input to increase or decrease the energy needed.[5] The governors respond to give more or less energy based on the frequency excursion.

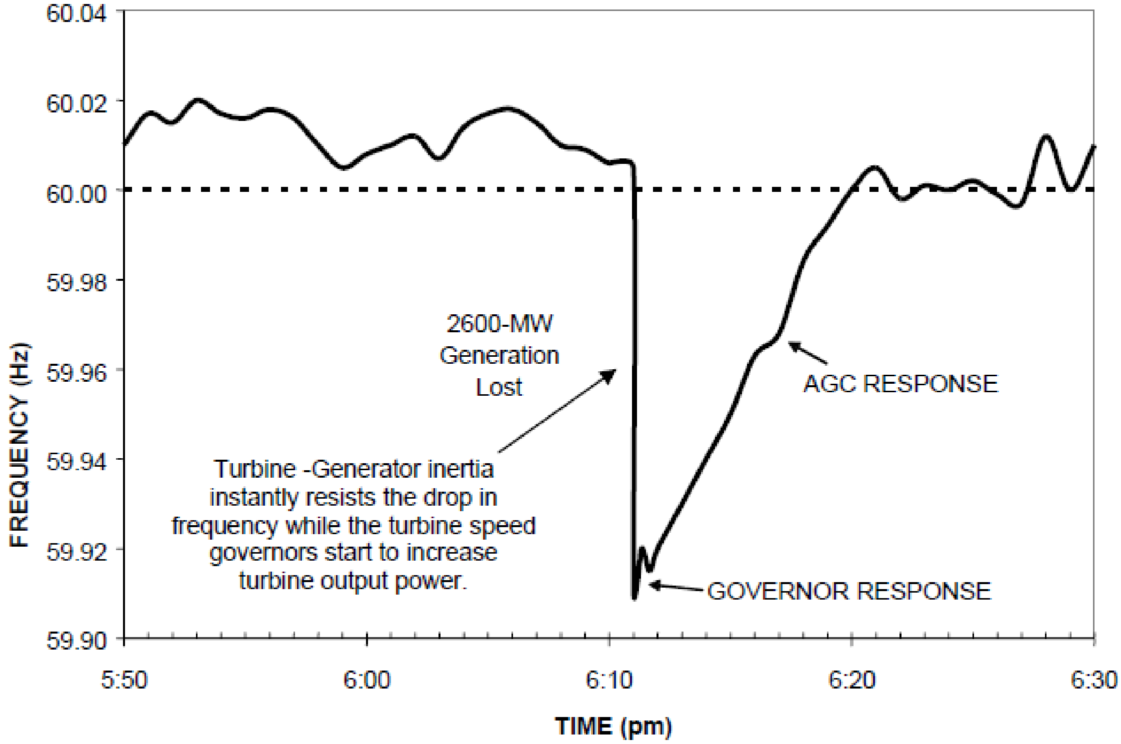

Figure 13: Example contingency (loss of supply) event and typical responses.

The inertial response and governor control (Primary Reserve) will stabilize the system frequency. However, it will be at an off-nominal frequency in which the system may be vulnerable towards collapse if another event were to occur. Therefore, Secondary Reserve response is required to return the frequency back to its nominal level.[6] Even the frequency levels that occur during single events can cause damage if sustained over a long period of time. Therefore, the Contingency Reserves should return frequency back to its nominal value fairly quickly. If nothing else occurs, at the time the Contingency Reserve has converted into energy, the system has less or no Contingency Reserves available if an additional event were to occur. Therefore, it is important for slower Tertiary Reserve response resources, which may mostly be off-line and starting soon after the contingency, to replace the Primary and Secondary Reserves. The response required of these reserves is therefore up to the operator's requirement (whether directly or via standards) of how soon following one contingency the system should be secure towards another.

[5] Motor-driven loads also naturally provide favorable frequency response with the mechanical load decreasing as frequency and rotational speed fall and increasing as frequency increases. This natural response is lost if the load is supplied through a solid-state variable frequency motor drive.

[6] Though this shows AGC on Figure 13, it can be a combination of units on AGC and manual response

Since we have defined Contingency Reserves as those that are needed for sudden near-instantaneous events, the impact of the variability and uncertainty of VG generally does not constitute a contingency event. Today's requirements for Contingency Reserve are usually based on the largest credible contingency that can occur on the particular system (which may include multiple generators or events). The system is then prepared to respond to any lesser event as well. Although faults on individual wind turbines (or PV cells) are negligible, a fault on the connection point disconnecting all the energy from the entire collection of individual facilities could be a major loss and should be considered as a contingency. If a single connection point serving VG to the grid could possibly have more capacity than the existing largest contingency, this may change the reserve requirement.

Large wind and solar ramping events are similar to conventional generation contingencies in that they are large and rare. They differ in that they are much slower (Figure 4 vs. Figure 13), last longer, and can be predicted. The extent to which conventional Contingency Reserves can be used for VG ramps is being discussed in certain areas.

2.4 Ramping Reserves

Ramping Reserve is probably the least well defined category of our list in current systems. Most, if not all balancing areas, do not consider this category today. This type of reserve is used for rare severe events that are not instantaneous in nature (Figure 4). Large load ramps occur every day, are predictable, and are met with Following Reserve and the action of the energy markets rather than with Ramping Reserve. Due to the greater unpredictability of wind and solar, infrequent large magnitude events may occur that require Operating Reserves. The way that Ramping Reserves are separated from Following Reserve is up to the particular balancing area. For example, Following Reserves may cover 95% of the possible deviations, and Ramping Reserves may cover the remaining 5%. Figure 14 shows a plot of a missed net load forecast and the actual net load. The plot shows a wind ramp that was not forecast to happen and therefore, the system operators were not prepared for the large increase in net load.

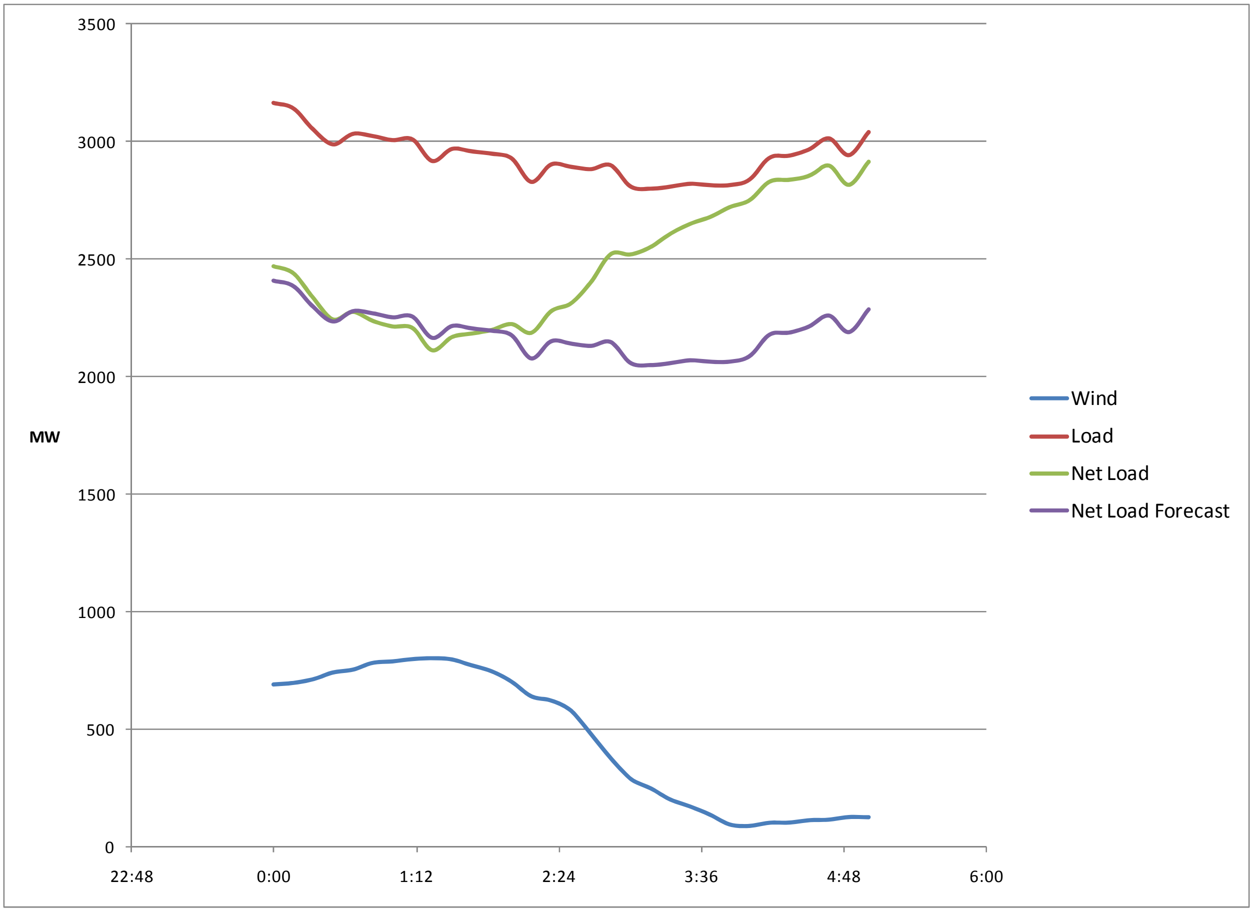

Figure 14: Net load forecast that does not forecast wind ramp.

With conventional practice, the system may have committed about 2100 MW to meet the forecast net load, plus additional Contingency Reserves, during hours 0 through 4. For a system of this size, an unexpected 600 MW increase in net load is pretty large (25% of the net load). If the Contingency Reserves were used to compensate for the drop in wind, there would be no more capacity available to respond if a conventional contingency were to occur during these hours and the system might have to shed load. On the other hand, the system would not need to respond to the 600 MW increase in net load in 10 minutes as is a typical response since the actual event took over 2 hours. It would have had access to much slower resources and offline resources during the entire event. Therefore, the optimal reserve to use for this event would not need to be as fast and a large portion could be offline.

Much research is currently being conducted on wind ramp forecasting [22]. Small changes of wind speed on the cubic region of a wind power curve can cause large relative changes in power, and the processes that cause these large changes can differ dramatically by region and season. Each different weather process that causes wind ramps requires different prediction techniques. The actual forecast that is received for wind ramps may be implicit or explicit, ramp rate or ramp event, and probabilistic or deterministic. This forecast should be the primary input (along with forecasts of load and other VG that cause large ramping) into the operator's decision on how much Ramping Reserve is needed, and the ramp rates and percentage of online capacity that is needed. The use of this information may be for economic reasons as well. Offline reserves cost very little while standing by, but usually cost a lot to be activated, which is the opposite of online reserve, which is expensive to procure. Therefore, the probability that the ramp will happen may influence how much online reserve is to be used compared to offline reserve. If there is a 5% chance of the ramp happening, offline reserve might be best for the event, whereas if there is a 75% chance of the ramp occurring, online reserve would make more sense to be held. Obviously, the predicted ramp rate and magnitude of the ramp event will also influence the characteristics of the resources providing the reserve that is held.

As shown in Figure 5, there is no Primary Reserve below Ramping Reserve. Because of the speed with which these events occur, there is no need for resources to automatically respond to stabilize the frequency since the time delays of the manual response used are quite sufficient. However, since the events can be large, there may be a certain type of Tertiary Reserve that is kept so that when the Ramping Reserve is used for the event, it can be replaced for a second event that may happen in the same direction (positive or negative ramp). This Tertiary Reserve is very different than that which would be used for Contingency Reserve. For instance, if a wind ramp occurs so that at the end of the ramp the wind plant is either near its rated capacity (for up-ramps) or close to zero (for down-ramps), it is not possible for an additional wind ramp event to occur in the same direction. Therefore, the type and the actual need will depend on what type of ramp event it is, as well as information on the probability of when the next event may occur.

2.5 Primary Reserves (Under Contingency Reserves)

Primary Reserve as shown in Figure 5 is a sub-category that is part of Contingency Reserve. Thereby, a certain portion of the Contingency Reserve must be automatically responsive to changes in frequency. During loss of supply contingency events, the frequency will decline based on the rotating machines slowing down to provide inertial energy (alternatively they will speed up during loss of load events). Figure 15 shows this inertial response and the frequency decline from point A to point C [23]. When supplying the energy needed during a loss of generation, other machines will supply kinetic energy to the grid and slow down, reducing the system frequency. When absorbing energy during a loss of load where there is too much generation, the machines will speed up and the frequency will increase. The amount of inertia on an interconnection will determine the rate at which the frequency deviation occurs. Generally, the more rotating synchronous machines, the higher the inertia and the slower the rate of change of the frequency deviation.

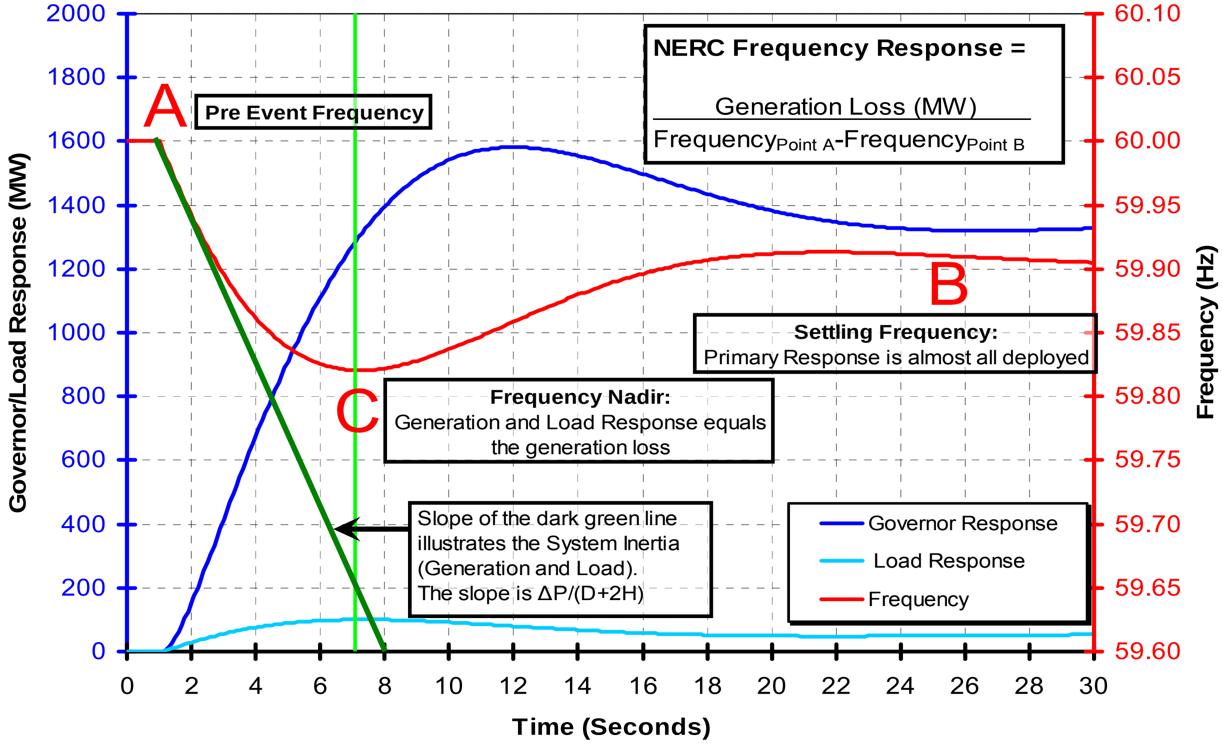

Figure 15: Frequency response to disturbance.

Soon after the event and frequency deviation occurs, generators and loads can then sense this change of frequency and start to adjust mechanical input to provide opposing response through its governor systems. This may include a combination of electrical, mechanical, and hydraulic means to adjust the input to the turbine (e.g., opening steam valves in a steam turbine generator). These governor's control equipment is provided a feedback signal, called droop characteristic, that allows a given frequency change to cause adjustment to a given power output change. Droop results in frequency stabilizing at point B in Figure 15, somewhat below the point A starting point. Figure 16 shows the droop curve. The slope of the line is considered the droop characteristic. The droop is usually somewhere between 4-6%, meaning for a 5% change in frequency (e.g., a 2.5 Hz for 50-Hz system and a 3 Hz for 60-Hz system), the generator controller would change its output by 100% of its capacity. Figure 16 shows a 5% droop curve with the unit's set point set at 80% of its output.

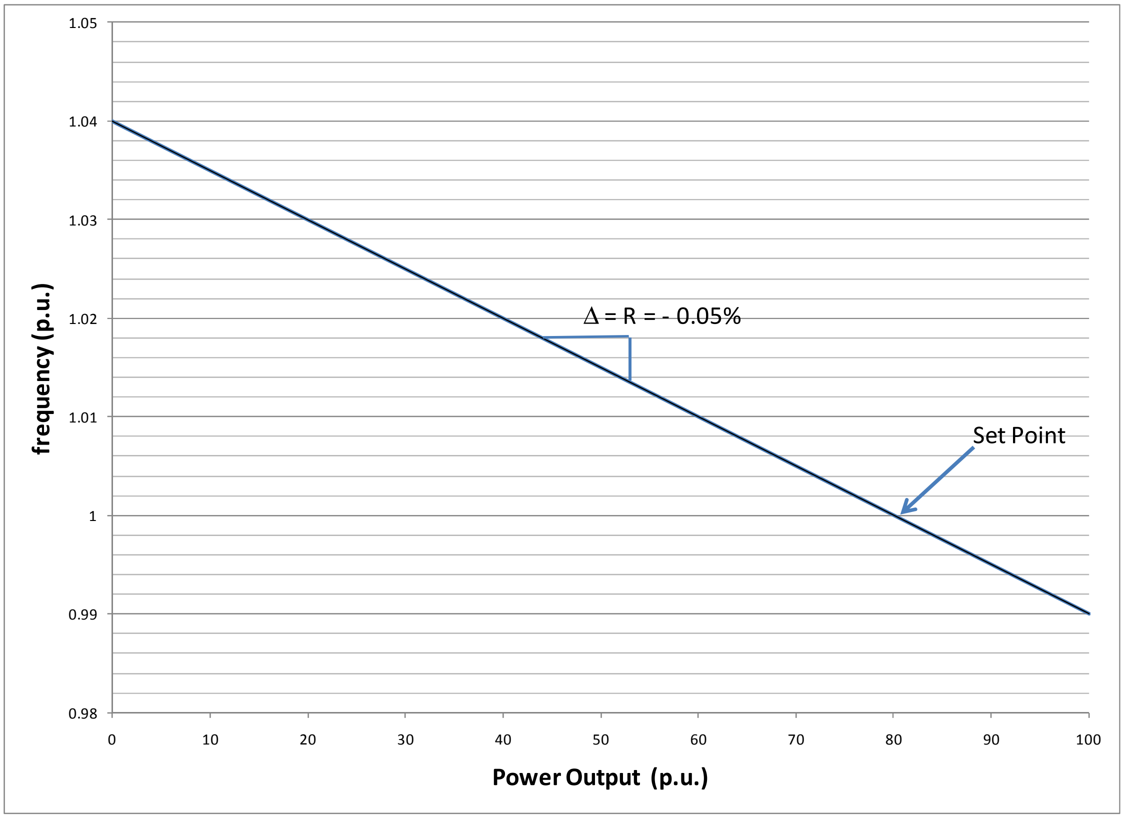

Figure 16: Droop curve with 5% droop.

There is one key difference in Figure 16 from governor droop curves in practice. In practice, there is normally a small governor dead band around the nominal frequency so that small frequency deviations are ignored to prevent unnecessary movements from the generating units. This is shown in Figure 17. A typical dead band might be in the range of 10-50 mHz. Therefore, depending on the size of the frequency deviation, the frequency response may differ depending on if the governors are triggered or not. Additionally, Figure 15 shows that load responds to frequency as well, usually reducing about 1 to 2 times the frequency deviation in per unit (i.e., a 1 or 2% change in load for a 1%, or 0.3 Hz change in frequency for 60-Hz system,). This total Primary Reserve assists in arresting the frequency deviation, balancing the generation and load, and stabilizing the frequency to a steady-state value. It allows time for the Secondary Reserve of the Contingency Reserve to make up for the power loss to return frequency back to its scheduled setting.

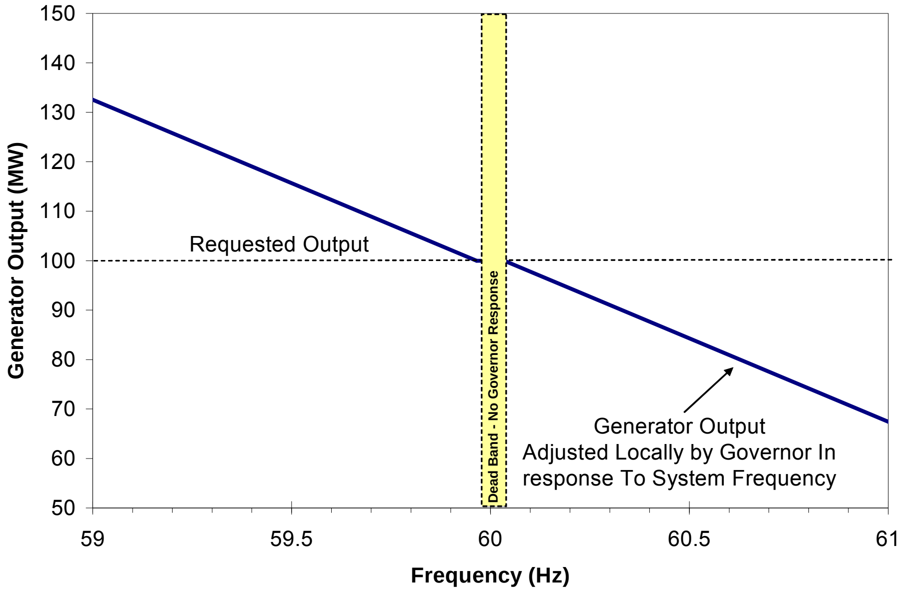

Figure 17: Governor droop curve with dead band.

Primary Reserve is needed to stop frequency deviations from becoming too large. This protects generators from excessive frequency deviations which can create conditions that may cause damage to the generators or set off under- or over-frequency relays which can shed system load or disconnect generators. An interconnection or balancing authority can define its total frequency response through historic analysis as well as analytics. This term is referred to as the 'frequency response characteristic,' with units usually of MW/0.1Hz, and defines how much response the system will give following a change in load balance. It should be noted that even though this value is often given as a single value in many interconnections and balancing areas, it is highly non-linear and conditional depending on what resources are on line, what loads the system currently has, and other things like for example governor dead bands. The frequency deviation for determining the frequency response characteristic usually is measured at the instant right before a disturbance and the time where the frequency settles to a steady-state value, usually about 30 seconds following the disturbance. A similar characteristic can measure the frequency deviation from the instant before the event and at the frequency nadir as is suggested in [24].

Because of their slow movements, VG events (specifically wind) involving these resources have very little impact on the need for Primary Reserve. Although more severe in the 5 to15-minute time frame, even PV plants can avoid extreme imbalances in the Primary Reserve time frame due to increased geographic diversity [25]. However, many of the VG resources that are supplying energy are displacing synchronous machines that have inertial response and governor response. While modern wind turbines can be designed to provide emulated inertia and governor response, most existing installed wind turbines currently do not provide either type of response. However, variable-speed wind turbines with power electronics can provide this type of service with a similar type of response to the conventional generation response [26][27][9].

2.6 Secondary Reserves (Under Contingency Reserves and Ramping Reserves)

Secondary Reserve for both Contingency Reserve and Ramping Reserve are used to return the frequency back to its nominal value and reduce ACE back to zero. This is the majority of the Contingency and Ramping Reserves and therefore, its descriptions and discussion about how VG can impact these type reserve types are covered under Section 2.3 and Section 2.4.

2.7 Tertiary Reserves (Under Contingency Reserves and Ramping Reserves)

Tertiary Reserve is unique because it is the only reserve category that is not deployed for energy imbalance, but is instead deployed for reserve imbalance. In other words, it is held in some manner so that when certain Operating Reserve types are used to correct the energy imbalance and converted into energy, it is used to restore that form of Operating Reserve. Tertiary Reserves do not need to be as fast as the reserves they are restoring. By increasing output (or decreasing consumption) slowly, Tertiary Reserves can allow the fast resource that was used to correct the imbalance to return to its reserve mode and thus restore the power system's response capability. The actual response time will depend on how soon the system must be ready for a subsequent event following the first event in the same direction (e.g., energy insufficiency or surplus). The Tertiary Reserve must sustain its response until the reserve it is replacing has been restored. Because Regulating Reserve and Following Reserve are used continuously in both positive and negative directions, they are essentially replaced continuously, which is why Tertiary Reserve is not applicable to these types.

Tertiary Reserve changes needed with increased penetrations of VG were briefly discussed in both the Contingency Reserve and Ramping Reserve sections (2.3 and 2.4, respectively). Basically, it will be a function of the probability of occurrence of an additional event following the original event, how soon it will happen, and what the acceptable level of risk is that the system is willing to allow.

3 Operating Reserves in Practice

We will now describe how different regions currently perform the operating reserve function. We will look at each of the categories as discussed in section 2 and evaluate how requirements are determined, what resources or technologies can provide them, when they should be deployed, how they are deployed, and any proposed changes in the future to each category. Throughout this section, reserve names with capital letters refer to our classification in section 2 and reserve names in italics refer to the naming used for the particular system being described. In some cases, reserve naming for a particular region might represent two of the categories discussed in section 2 and seen in Figure 5. Alternatively, two of the categories in Figure 5 might be represented by one name or definition in a particular region. It is important to note the differences between regions and in some cases we attempt to explain some of the possible reasons for those differences. Where applicable we emphasize any rules or procedures relating to operating reserves requirements that are impacted by VG. The paper will discuss both North America and European practices, with additional insight on some key regional differences in particular regions in North America.

3.1 North America (NERC)

The main organization for administering the standards in North America is the North American Electric Reliability Corporation (NERC). Although many regions inside North America may have specific regional requirements, NERC standards are overarching and can be seen in the NERC standards document [28]. The Federal Energy Regulatory Commission (FERC), which is the regulator for wholesale electric energy trading between states in the United States, designated NERC as the Electric Reliability Organization in 2007. NERC now has authority to issue penalties to organizations that violate reliability standards [29]. So far, no specific procedure or standard pertaining to operating reserve in NERC has been modified to account for increased VG. However, NERC is currently pursuing an initiative that is looking at how standards may have to change with high penetrations of VG [30]. Figure 18 describes the four interconnections and the NERC reliability regions in North America.

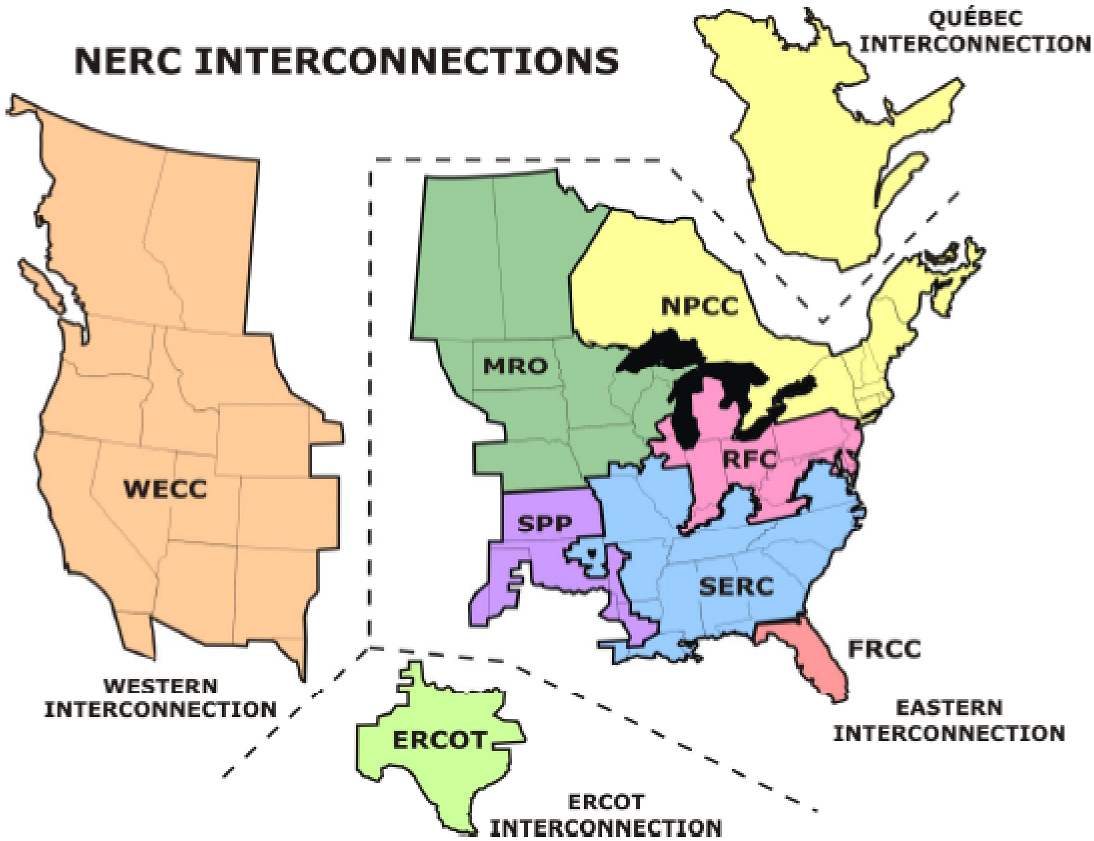

Figure 18: NERC Interconnections and Reliability Regions.

North America also has a unique distribution of power system structures. NERC defines a balancing authority as the authority that is responsible for balancing the generation and load for its balancing authority area. The two terms are often used somewhat interchangeably and shortened to balancing area, or BA. Each interconnection has a number of different BAs ranging from one (ERCOT and Quebec) to 68 (Eastern Interconnection). Figure 19 shows that there has been a trend towards balancing area consolidation, especially within the PJM and MISO regions. But there are also still numerous small balancing areas.

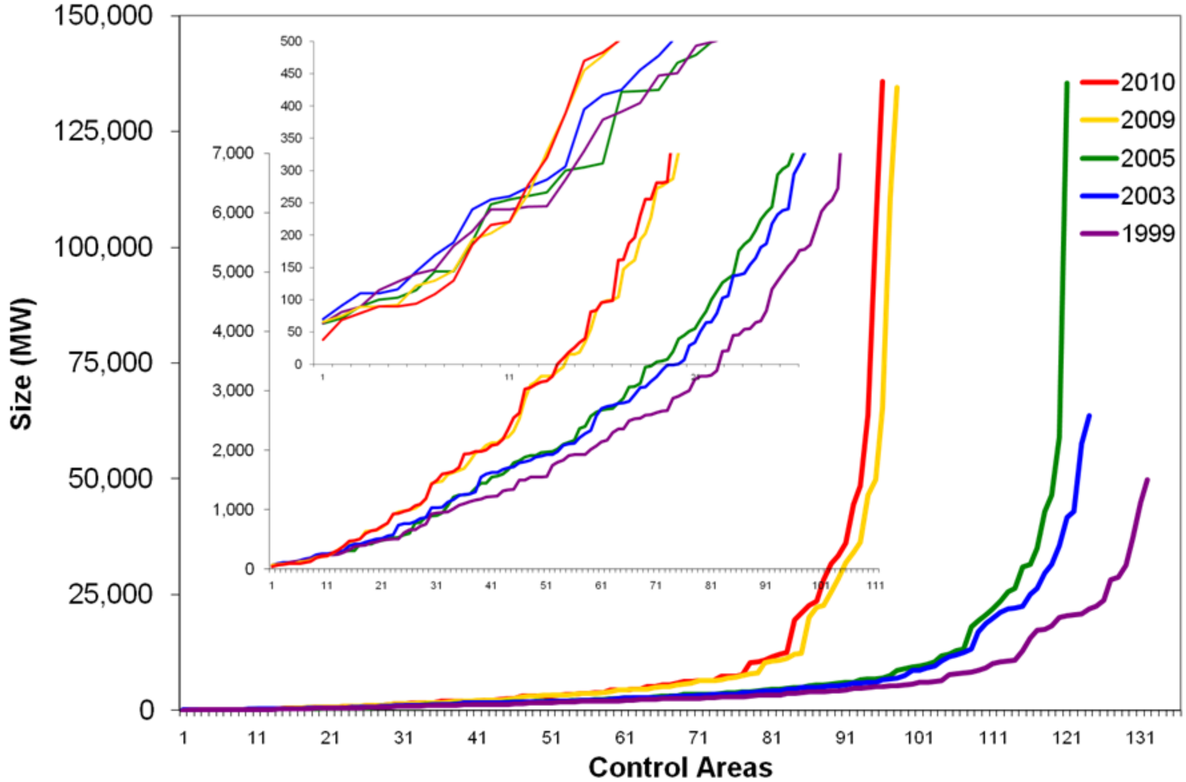

Figure 19: While there is a balancing area consolidation trend, there are also numerous small balancing areas.

The BAs in each interconnection vary on whether they are administrated in connection with wholesale energy markets with Independent System Operators (ISO) or Regional Transmission Organizations (RTO) or whether they are administered by regulated utilities [31]. The different administrative mechanisms affect how the policies, rules, and standards pertaining to operating reserve are developed. Figure 20 displays the different ISO and RTO regions that represent organized electricity markets in North America. Everywhere where there is not an ISO or RTO there are utilities which can range from a large regulated publicly traded corporation to a small cooperative or municipality.

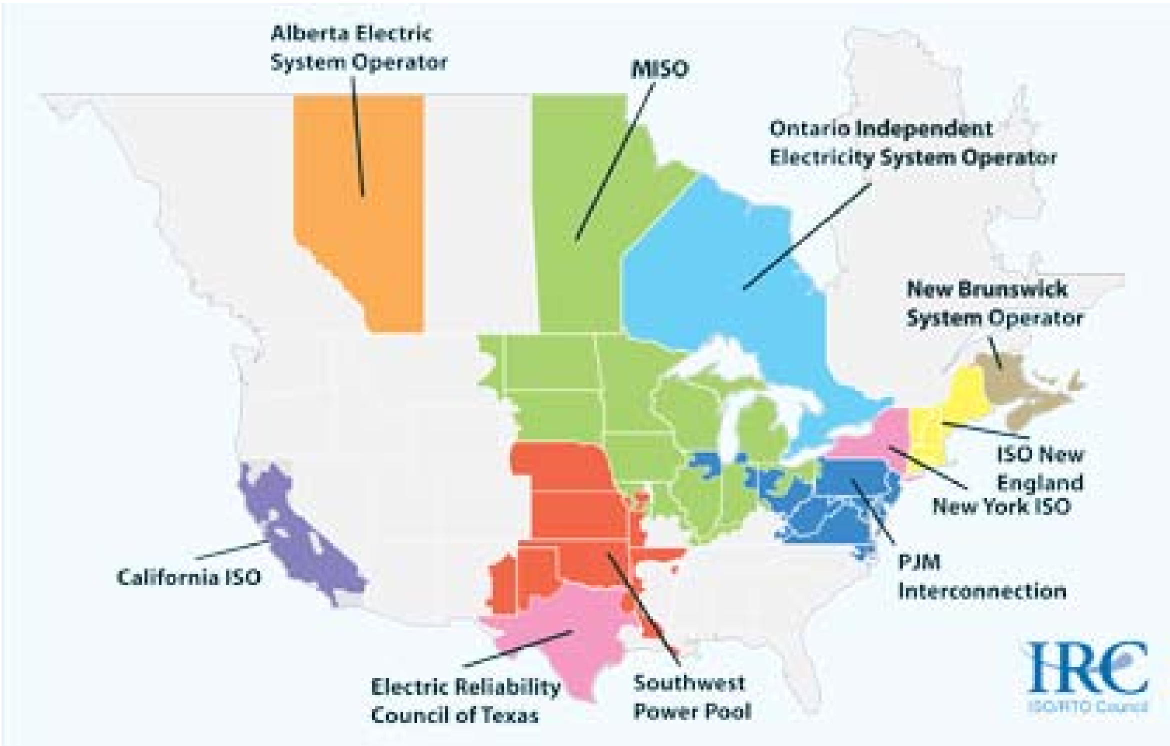

Figure 20: ISO and RTO regions in North America. (www.isorto.org)

The rest of this section will describe the different policies relating to the different categories for Operating Reserves under the NERC system.

3.1.1 Regulating Reserves

Terminology: NERC uses the term *regulating reserve* in its standards and documents. It may also refer to the service as *regulation service* or *regulation.*

Required Amount: NERC does not specify any direct requirement for *regulating reserve*. Instead, NERC standard BAL-005, requirement R2, states that "Each balancing authority shall maintain regulating reserve that can be controlled by AGC to meet the Control Performance Standard." The control performance standard is based on two separate metrics, CPS1 and CPS2, which are based on NERC standard BAL-001. Both of these measure how well the balancing area is balancing its load with generation based on its ACE. Note that ACE according to NERC is defined as the following:

$$ACE = (NI_A - NI_S) - 10B(F_A - F_S) - I_{ME}$$

Where NI is the sum of net interchange in MW, defined with exporting being positive, B is the frequency bias setting in MW/0.1Hz, F is the frequency in Hz, and I_{ME} is the meter error correction factor in MW. CPS1 is defined by the following equations:

$$CPS1 \geq 100$$

$$CPS1 = (2 - CF) * 100$$

$$CF = \frac{CF_{12-month}}{\epsilon_1{}^2}$$

$$CF_{clock-minute} = \left[(\frac{ACE}{-10B})_{clock-minute} * \Delta F_{clock-minute} \right]$$

CF is the compliance factor and $CF_{12\text{-month}}$ is calculated by taking the average of $CF_{\text{clock-minute}}$ for all available minutes for a 12-month period. ϵ_1 is a constant that is derived from a targeted 1-minute frequency bound calculated for each interconnection. The measure compares the square of the targeted bound with the product of the actual frequency deviation and the balancing area's ACE imbalance. Note that both ACE and frequency deviation are signed quantities. CPS1 rewards a balancing area if its ACE is in the direction that is trying to restore system frequency (e.g., over-generating when frequency is low) and penalizes a balancing area if its ACE is in the direction that is contributing to the system frequency error (e.g., under-generating when frequency is low). Note too that CPS1 is a statistical metric measured over a year; it does not require perfect compliance. CPS1 is also a 1-minute metric; it does not measure imbalances that are shorter than 1 minute. CPS2 is described next with the following equations.

$$CPS2 \geq 90\%$$

$$CPS2 = \left[1 - \frac{Violations_{month}}{TotalPeriods_{month} - UnavailablePeriods_{month}} \right] * 100$$

Violations = 1 if $\left| \frac{\sum ACE}{n_{samples\ in\ 10-minutes}} \right| > L_{10}$

$$L_{10} = 1.65 * \epsilon_{10} \sqrt{(-10B_i)(-10B_s)}$$

L_{10} is a MW value that is calculated for each balancing authority and is a desired upper limit of ACE for the balancing area. ϵ_{10} is a constant derived from the targeted 10-minute frequency bound. The balancing authority should keep its 10-minute ACE below its L_{10} for at least 90% of all 10-minute intervals for the month. This equates to an allowance of 14.4 violated 10-minute intervals per day.

While CPS1 strives to reduce energy imbalances that harm interconnection frequency, CPS2 strives to limit all energy imbalances. Both metrics are statistical with neither requiring perfect performance. Both apply to balancing areas, not to individual generators or loads. CPS1 addresses 1-minute imbalances while CPS2 addresses 10-minute imbalances.

CPS1 and CPS2 are balancing area performance metrics. Each balancing authority will determine its own way to require *regulating reserve* so as to meet the requirements of CPS1 and CPS2. Usually, the requirement is based on system peak or hourly load, time of day, weekday or weekend, and season. It may be that at different times, systems have more difficulty staying in compliance with CPS1 and CPS2 than other times.

Provider: NERC does not have any requirement as to what technologies are allowed to supply *regulating reserve*. They must be able to be controlled by AGC. While generators have been the traditional suppliers of *regulating reserve*, loads are beginning to enter this market. New storage devices (e.g., flywheels and batteries) are also starting to supply *regulating reserve*. In fact, the AGC term itself is now misleading and might be replaced with ARC (Automatic Resource Control), as proposed in a draft of NERC standard BAL-001-1 [32].

When deployed: *Regulating reserve* is required to be continuously controlled by AGC from the balancing authority. According to requirement R8 of NERC standard BAL-005, the control signal must be refreshed no less frequently than every 6 seconds. Each balancing authority will create its AGC signal as some function of the ACE so that high ACE values will direct downward *regulating reserve* movements and low ACE values will direct upward *regulating reserve* movements. The goal would be to comply with CPS1 and CPS2.

As seen in ACE equation above, the frequency bias setting is a major component. NERC discusses how this is set based on NERC standard BAL-003. Requirement R2 of the standard states that the bias setting be as close as possible to or greater than the balancing authority's actual frequency response. This number may either be a fixed value based on average frequency response or variable (either linear or non-linear) and should be calculated by reviewing historic response characteristics. Requirement R3 of NERC standard BAL-003 then states that balancing authorities shall operate their AGC based on tie-line frequency bias control as discussed in section 2.1. Lastly, requirement R5 states that the balancing authority's frequency bias setting should be at least as great as 1% of its yearly peak demand per 0.1-Hz frequency change or 1% of its maximum generation for the year if the balancing authority is a generation-only balancing authority. Table II shows examples of some area's frequency bias settings for different size systems and is taken from [33].

Table II: Balancing Authority frequency bias settings

Balancing Authority	Frequency Bias Setting	Estimated Peak Demand	Percent Bias/Peak
New York ISO	-334 MW/0.1Hz	33,441 MW	1%
Midwest ISO	-1038.6 MW/0.1Hz	103,864 MW	1%
Omaha Public Power	-25.5 MW/0.1Hz	2,546 MW	1%
Public Service Company of New Mexico	-35.7 MW/0.1Hz	2,625 MW	1.36%
Electric Reliability Council of Texas	-667 MW/0.1Hz	63,783 MW	1.05%
Lincoln Electric System	-7.8 MW/0.1Hz	776 MW	1.01%

Looking throughout all of the balancing authorities in NERC, many are exactly 1%, which mean they may be based on requirement R5 rather than matching their frequency response characteristic as closely as possible (requirement R3).

How deployed: *Regulating reserve* is continuously deployed through AGC to correct ACE. No further directions are given from NERC. Each generating type would use its own strategy based on its technology. It must, however, be automatic.

Proposed changes: A proposed NERC standard BAL-007, "balance of resources and demand" along with proposed standards BAL-008, BAL-009, BAL-010, and BAL-011 would replace the CPS1 and CPS2 [34][35]. In particular, the CPS1 will be replaced by an identical measure Control Performance Measurement (CPM). The CPS2 would be replaced with a new Balancing Authority ACE Limit (BAAL) requirement. The motivation behind the standard is an argument which says that during some instances, it is possible to impact interconnection frequency and yet still remain compliant under the current CPS2 standard. Balancing authorities could remain under their L_{10} values for excessive periods of time and be compliant while continuing to adversely impact the system frequency. The BAAL would be a frequency dependent limit. The limit would be unbounded when frequency is at its nominal level, but become more restrictive as frequency deviates from the nominal level. The calculations of the BAAL will be based on the frequency relay limits, the highest setting of under frequency load shedding in the interconnection. The BAAL will then be dependent on the interconnection frequency – the higher the frequency deviation, the stricter the BAAL will be.

To be compliant, the ACE could not exceed the BAAL for a time greater than the BAAL T_v (currently set at 30-clock-minutes). Figure 21 shows an example of how the BAAL would be set in terms of multiples of the frequency bias. The equation for how the BAAL is calculated for high and low limits is shown in the equations below. This shows that for small frequency deviations, the BAAL is essentially limitless, and at the frequency trigger limit, the BAAL is set at 10B*(FTL – Fs), where Fs is the scheduled frequency. The FTL is established for an interconnection and is defined as the frequency that is one single contingency margin above or below the lower and upper Frequency Abnormal Limit (FAL), respectively. The FAL is defined as that higher or lower frequency where any frequency above or below, respectively, leaves an interconnection with unacceptable risk as determined by the sum of MWs for the number of allowable contingencies in the interconnection. It is unclear as to how this will affect the amount of *regulating reserve* that a balancing authority would procure, but it will likely affect how and when they will be deployed to be in compliance with the new standard.

Proposed NERC standard BAL-008 and BAL-009 guide how reliability coordinators are to direct balancing authorities to help restore frequency when it exceeds frequency limits (especially if it exceeds the FTL for over 5 minutes). BAL-010 is very similar to BAL-003 on how frequency bias setting is calculated, but with additional details and procedural changes. Lastly, BAL-011 describes how the FTL, FAL, and Frequency Relay Limits (FRL) are calculated and determined for each interconnection.

$$BAAL_{low} = (-10B * (FTL_{low} - 60)) * \frac{FTL_{low} - 60}{F_A - 60}$$

$$BAAL_{high} = (-10B * (FTL_{high} - 60)) * \frac{FTL_{high} - 60}{F_A - 60}$$

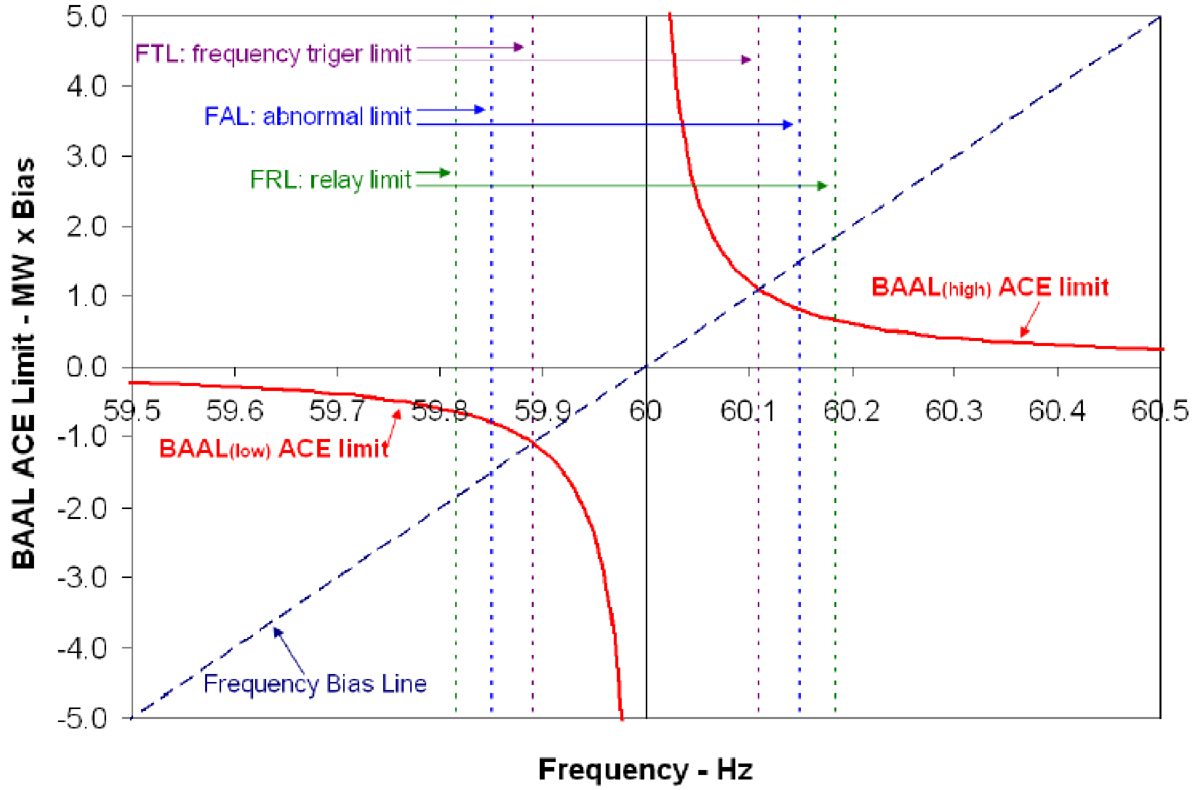

Figure 21: New frequency dependent BAAL requirement.

3.1.2 Following Reserves

NERC gives no standard or direction on Following Reserve.

3.1.3 Contingency Reserves

Terminology: NERC uses the term *contingency reserve* in its standards and documents. It also uses both *spinning reserve* and *non-spinning reserve* as different types of *contingency reserve.*

Required Amount: NERC standard BAL-002, the Disturbance Control Standard (DCS) requirement R3 states that "as a minimum, the balancing authority or reserve sharing group shall carry at least enough contingency reserve to cover the most severe single contingency." The same requirement also states that the balancing authority or reserve sharing group shall review the probable contingencies to determine the most severe contingencies at least once a year.

Provider: NERC standard BAL-002 requirement R4 requires that the ACE be returned to zero (or to its pre-disturbance level if its pre-disturbance level is originally negative) within 15 minutes as shown in Figure 22 and Figure 23. This drives the speed of response of resources that can provide *contingency reserve*. Requirement R1 of the same standard states that *contingency reserve* may be supplied from generation, controllable load resources, or coordinated adjustments to interchange schedules. NERC also defines both *spinning reserve* and *non-spinning reserve* to both be used to return ACE to zero during disturbances (or return ACE to a pre-disturbance level if its pre-disturbance level is originally negative). *Spinning reserve* is defined as unloaded generation that is synchronized and ready to serve additional demand. *Non-spinning reserve* is generating reserve not connected to the system, but capable of serving demand within a specified time, or interruptible load that can be removed from the system in a specified time.

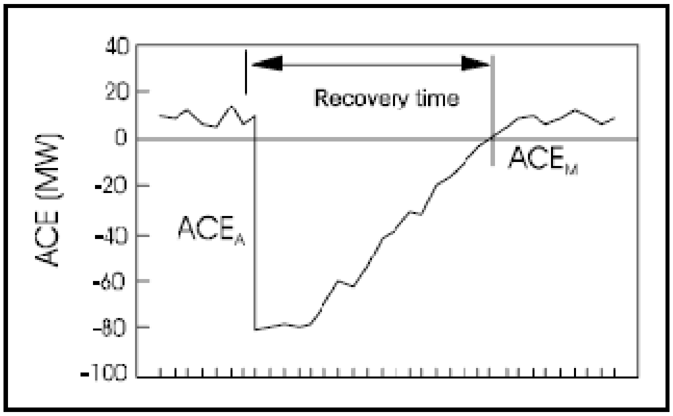

Figure 22: Example of DCS when pre-disturbance ACE is positive.

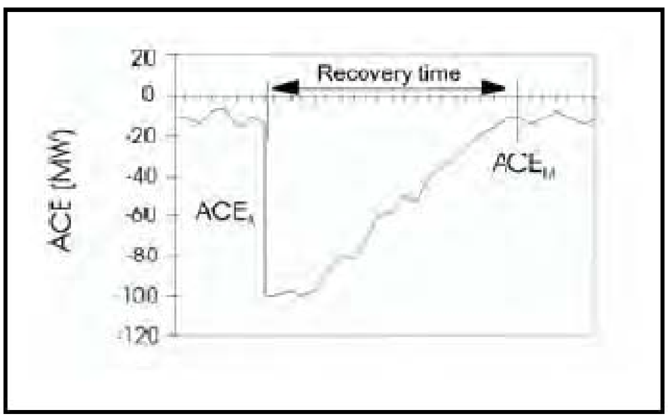

Figure 23: Example of DCS when pre-disturbance ACE is negative.

When deployed: *Contingency reserves* are deployed following a reportable disturbance. A reportable disturbance is defined by NERC in standard BAL-002, measure M1, as contingencies that are greater than or equal to 80% of the most severe single contingency in the balancing area. This number may be reduced by a regional entity, but it must be ensured that normal operating characteristics are excluded. It is important that all reportable disturbances are only characterized by sudden, unanticipated losses of supply-side resources.

How deployed: *Contingency reserves* are deployed by set points that are increased as directed by the balancing authority to make up the lost energy and return ACE to zero and frequency to its scheduled setting. The response can come from resources with AGC, online resources responding manually to the balancing authority's direction, and non-spinning reserve so that it can assist within the 15-minute DCS window.

Proposed changes: The proposed standards of BAL-007 through BAL-011 that are discussed in section 3.1.1 will also replace the DCS requirement. These reliability based requirements are supposed to give guidance on load balancing during both normal and abnormal events. Hence, the requirement for ACE deviations will depend on frequency deviation of the entire interconnection, and not on the size of the balancing area specific event that caused the ACE deviation (e.g., no distinguishing between normal vs. disturbance conditions). However, as it currently stands, proposed standard BAL-007 does not have any information on the required amount as is currently stated in BAL-002 requirement R3.

3.1.4 Ramping Reserves

NERC gives no standard or direction on Ramping Reserve. However, in a recent notice of proposed rulemaking, FERC asked for comments from industry stakeholders on the steps needed to resolve the confusion regarding the use of contingency reserves to manage extreme VG ramp events [36]. It is likely that this may lead to more specific rules and policies on how reserves should be used for ramp events.

3.1.5 Primary Reserves

Terminology: Although, there is no term used in the NERC standards, the term often used in North America is *frequency responsive reserves*.

Required Amount: NERC does not require a quantifiable amount for *frequency responsive reserve*. As stated in section 3.1.1, BAL-003 states the frequency bias setting to be 1% of the balancing area's peak demand, and that the balancing area's frequency bias be as close to possible or greater than its actual frequency response. However, this requirement is for frequency bias as part of the ACE equation, and therefore, it does not pertain to the actual frequency response of a balancing area.

Provider: Both generation and load that autonomously respond to changes in frequency.

When deployed: Deployed during frequency changes usually outside a frequency dead band.

How deployed: *Frequency responsive reserve* is automatically deployed when frequency changes from the nominal level (i.e. 60 Hz). Most North American governors use a 5% droop response and a 36-mHz dead band where a deviation below which it is not responsive. The droop or dead band is not seen in any NERC standard, however.

Proposed changes: In FERC Order 693, FERC ordered NERC to revise its BAL-003 standard so that it "defines the necessary amount of Frequency Response needed for Reliable Operation for each Balancing Authority with methods of obtaining and measuring that the frequency response is achieved." NERC is pursuing a frequency response initiative that is looking at what types of requirements are needed. Some analysis has shown that at least in the Eastern Interconnection, the amount of frequency response has been slowly declining in recent years [37]. This is a large reason for this initiative.

3.1.6 Secondary Reserves

Terminology: NERC uses the term *spinning reserve* or *non-spinning reserve* to refer to the Secondary Reserve used under Contingency Reserve. See section 3.1.3 for information regarding NERC's requirements. There is no term for Secondary Reserves used under Ramping Reserve.

3.1.7 Tertiary Reserves

Terminology: NERC uses the term *supplemental reserve* or *operating reserve – supplemental* to refer to the Tertiary Reserve used under Contingency Reserve. There is no term for Tertiary Reserves used under Ramping Reserve.

Required Amount: No quantifiable requirement for *supplemental reserve* is given by NERC.

Provider: No details are given on who is able to provide *supplemental reserve*.

When deployed: *Supplemental reserve* as discussed by NERC is described as reserve that can replace *contingency reserve*. Therefore, the *supplemental reserve* would be deployed soon after contingency events. In requirement R6 of NERC standard BAL-002 (DCS), it states that the contingency reserve should be restored 90 minutes following the start of the restoration period. The start of the restoration period is defined as the end of the disturbance recovery period, which as discussed in requirement R4 of the same standard and in section 3.3.3 of this document is 15 minutes following the disturbance event. Therefore, all contingency reserves must be restored and replaced90 minutes after the contingency should be recovered, and 105 minutes following the actual disturbance.

How deployed: The *supplemental reserve* is deployed as directed by the balancing authority. In most cases, these are resources that are off-line and must start up in time to help restore the *contingency reserve*.

Proposed changes: There are no known changes to the current standards around *supplemental reserve*.

3.2 Key North American Regional Differences

The regional systems inside North America will generally follow the NERC standards. However, in some cases, there are region-specific differences that are worth discussing. Since NERC rules apply to the whole continent, these differences should be at least as strict as NERC standards. Requirements follow a hierarchical scheme with NERC requirements overarching followed by regional reliability organizations as seen in Figure 18 followed lastly by the individual balancing authority requirements. Although, impossible to capture all important differences given the amount of different entities in North America, a number of key cases were identified and described that were of importance. The section will only present unique cases that are not identical to rules from NERC in reference to the different operating reserve types and future proposals specific to those regions.

3.2.1 Regulating Reserves

Requirement Definitions: Many different areas have their own ways of determining their *regulating reserve*. These are all based on the balancing authority's way of meeting its CPS1 and CPS2 requirements. Table III shows a few examples of areas and how they determine their requirements. Note that many of the requirements are based on ramp rate requirements. Many of the ISOs that run 5-minute dispatch intervals require that the full available *regulating reserve* requirement be available within five minutes. Others may be fifteen minutes.

Table III: *Regulating reserve* requirement definition from various North American balancing authorities

Region	Requirement Definition
PJM	Based on 1% of the peak load during peak hours and 1% of the valley peak during off-peak hours.
NYISO	Set requirement based on weekday/weekend, hour of day, and season.
ERCOT	Based on 98.8th percentile of regulation reserve utilized in previous 30 days and same month of previous year and adjusted by installed wind penetrations (described further below)
CAISO	Use a requirement floor of 350-MW up and down regulating reserves which can be adjusted based on load forecast, must-run instructions, previous CPS performance, and interchange and generation schedule changes.
MISO	Requirement made once a day based on conditions and before the day-ahead market closes.
ISO-NE	Based on month, hour of day, weekday/sat/sun.

Combined Up and Down vs. Separate: Some regions within NERC combine their upward and downward *regulating reserve* while others have separated requirements. Combined requirements imply that the upward and downward requirements must be equal and any resource providing upward *regulating reserve* must be able to provide the same amount of downward *regulating reserve*. Regions with separate services can have different requirements and can have different resources providing different amounts of upward and downward *regulating reserve.*

Table IV: Combined vs. separate *regulating reserve*

Region	Requirement Definition
PJM	Combined
NYISO	Combined
ERCOT	Separate
CAISO	Separate
MISO	Combined
ISO-NE	Combined

ERCOT: ERCOT's requirement for *regulation service* is a bit different than others. First, the 98.8 percentile is calculated for the up and down *regulation service* that had been deployed during the past 30 days and the same month of the prior year. This is done on an hourly basis. ERCOT will then calculate the increased amount of wind capacity compared to the previous months and use look up tables to add increased amounts of regulation based on a study performed for the region [38].

Table V shows the additional up-regulation for 1000-MW bins of additional wind capacity. If during the course of the last 30 days prior to the study period, the average CPS1 score was less than the 100% target, an additional 10% of regulation reserve will be procured during those hours where it was less. If the CPS1 score was less than 90%, an additional 20% of *regulation service* will be added. Also, for hours 0600 and 2200, ERCOT may increase the amounts further since these two hours are very ramp-constrained.

Table V: Additional up-regulation per 1000 MW of incremental wind generation capacity in ERCOT

Incremental MW Adjustment to Prior-Year Up-Regulation 98.8 Percentile Deployment Value, per 1000 MW of Incremental Wind Generation Capacity, to Account for Wind Capacity Growth																								
	Hour Ending																							
Month	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Jan.	2.8	4.2	3.1	3.7	2.5	0.4	2.3	2.2	4.2	5.9	7.6	5.7	4.7	3.3	2.8	2.3	4.0	8.6	4.2	2.7	1.6	2.7	1.4	1.6
Feb.	3.6	4.0	2.9	2.9	1.5	1.8	5.2	3.5	4.9	6.0	5.1	5.2	5.3	4.2	4.3	3.5	3.8	8.6	5.5	1.9	1.4	3.1	1.9	2.2
Mar.	5.5	5.3	4.6	4.2	2.6	3.3	7.1	7.9	6.8	5.7	4.2	3.4	2.8	2.6	2.7	2.3	2.9	7.7	6.8	2.1	1.1	3.0	1.5	2.8
Apr.	3.1	3.6	5.0	4.0	2.4	2.5	8.5	11.6	10.0	5.6	4.2	3.4	3.2	2.5	2.1	2.1	3.5	9.2	8.2	4.1	1.0	0.8	0.0	1.4
May	3.6	3.3	4.3	4.3	4.2	3.3	8.7	8.8	8.1	5.7	6.0	4.4	3.6	3.8	3.9	4.2	4.7	11.6	5.9	0.6	0.0	1.0	1.4	2.5
Jun.	2.3	2.6	3.3	3.7	3.9	2.4	8.5	8.2	6.6	4.5	4.2	3.1	2.5	2.5	0.7	0.2	1.3	7.5	3.3	1.7	0.7	0.3	0.6	1.3
Jul.	1.0	2.8	4.4	3.7	3.0	3.2	11.2	10.2	6.5	5.3	3.3	2.2	1.4	0.4	-0.9	-1.3	0.3	3.4	0.9	1.1	0.1	0.0	1.0	1.2
Aug.	1.4	3.8	4.5	4.5	2.2	0.9	6.3	6.8	6.6	6.6	3.2	2.6	2.1	1.2	1.4	1.3	1.3	4.6	1.2	0.9	0.7	0.8	1.1	1.3
Sep.	3.2	4.0	3.7	3.5	1.8	1.9	6.9	7.7	8.3	6.9	3.5	4.8	3.8	2.3	1.6	1.2	3.0	9.2	3.1	0.9	0.1	0.4	0.8	1.9
Oct.	3.4	2.8	2.4	2.2	1.7	1.8	5.0	5.8	6.1	5.9	4.0	5.4	3.2	2.2	1.2	1.7	3.1	6.8	0.8	2.1	0.0	0.2	1.8	2.5
Nov.	2.7	3.2	3.6	3.0	2.2	2.3	4.6	5.3	6.9	6.8	5.1	5.6	4.1	3.7	1.8	1.7	5.8	12.8	4.8	3.8	1.0	1.6	2.2	1.4
Dec.	2.8	2.4	1.4	2.1	1.2	0.4	2.8	2.7	3.8	4.6	6.8	7.0	6.0	4.4	3.3	3.0	5.0	9.9	4.3	2.6	2.1	4.3	2.0	1.5

Storage Resources Providing Regulating Reserve: Recently, the New York Independent System Operator/NYISO and Midwest Independent System Operator/MISO have also allowed certain storage resources (e.g., limited energy storage resources, or LESR) to provide *regulating reserve*. This required an initial rule change, accepted by FERC, to allow for continuous operation of *regulating reserve* for 15 minutes rather than sustained for one hour. The Energy Management System was enhanced so that when the LESR was low on energy, the dispatch would give it a negative energy schedule with therefore less regulating range so that it could store more energy [39]. An example is shown below in Figure 24. Usually, the energy schedule of the LESR would be at 0 MW, so that it can equally provide up and down regulation to its maximum levels (20-MW charging, 20-MW discharging for a 20-MW device). However, if it has been generating more than it has been charging and is low on the amount of energy stored, the dispatch schedule may give it a negative energy schedule as shown in the right example. This allows it to charge more so that it can eventually be brought back to operate at its optimal position. For example, if it has an energy schedule of -5 MW, and it is a 20-MW device, it can regulate between -20 and 10 MW and should consume 5 MW on average. This way, it can store more energy and allow it to eventually be used again at its most optimal point, 0 MW.

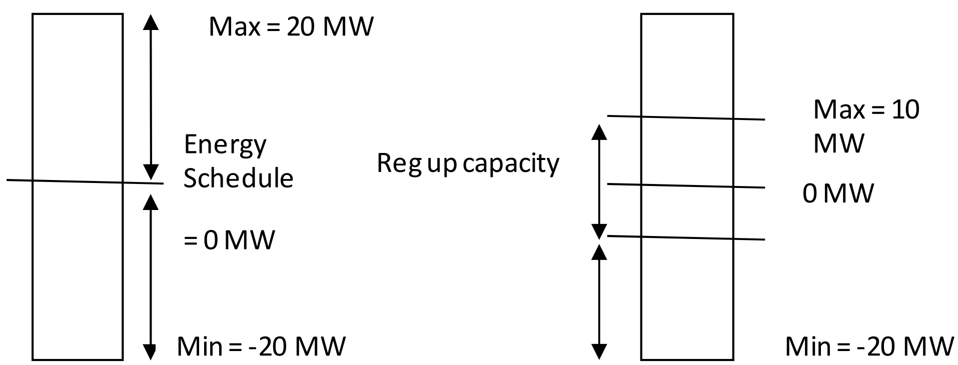

Figure 24: Example of a limited energy storage resource providing *regulating reserve* during normal (left) and during low storage (right) conditions.

In NYISO, the LESR are also deployed first whenever ACE requires regulation services. They are deployed up to their full amounts of regulation based on the need derived from the system ACE. The residual ACE recovery not covered by the LESR will be distributed among the other regulation resources in proportion to the amounts they were scheduled to provide.

Sub-Hourly vs. Hourly: In the definitions that we have used for Regulating Reserve in section 2.1, one of the ways we have separated it from load following is that it is faster than the shortest energy market clearing interval or economic dispatch interval. This is true in North America. The market intervals set a dispatch level for the resources based on economics, and then the *regulating reserve* resources are further adjusted automatically between one interval and the next to balance supply and demand while considering for random variations or the forecast error that may have occurred. This is important in that many areas within North America have different market interval lengths. All of the organized markets have 5-minute intervals, whereas some regulated utility areas without markets use *regulating reserve* to account for deviations from hourly generation and load schedules. Therefore, they have different requirements for what resources will need to be connected with AGC. Areas with hourly scheduling will usually keep all non-regulating resources fixed for the hour, allowing for a ramp to the next hour scheduling at 10 minutes before and after the top of the hour. Therefore, the same system with hourly scheduling will require more *regulating reserves* than if it had 5 or 15-minute economic dispatch intervals. An example of how the different scheduling procedures appear and how the *regulating reserve* gets deployed is shown in Figure 25 and Figure 26.

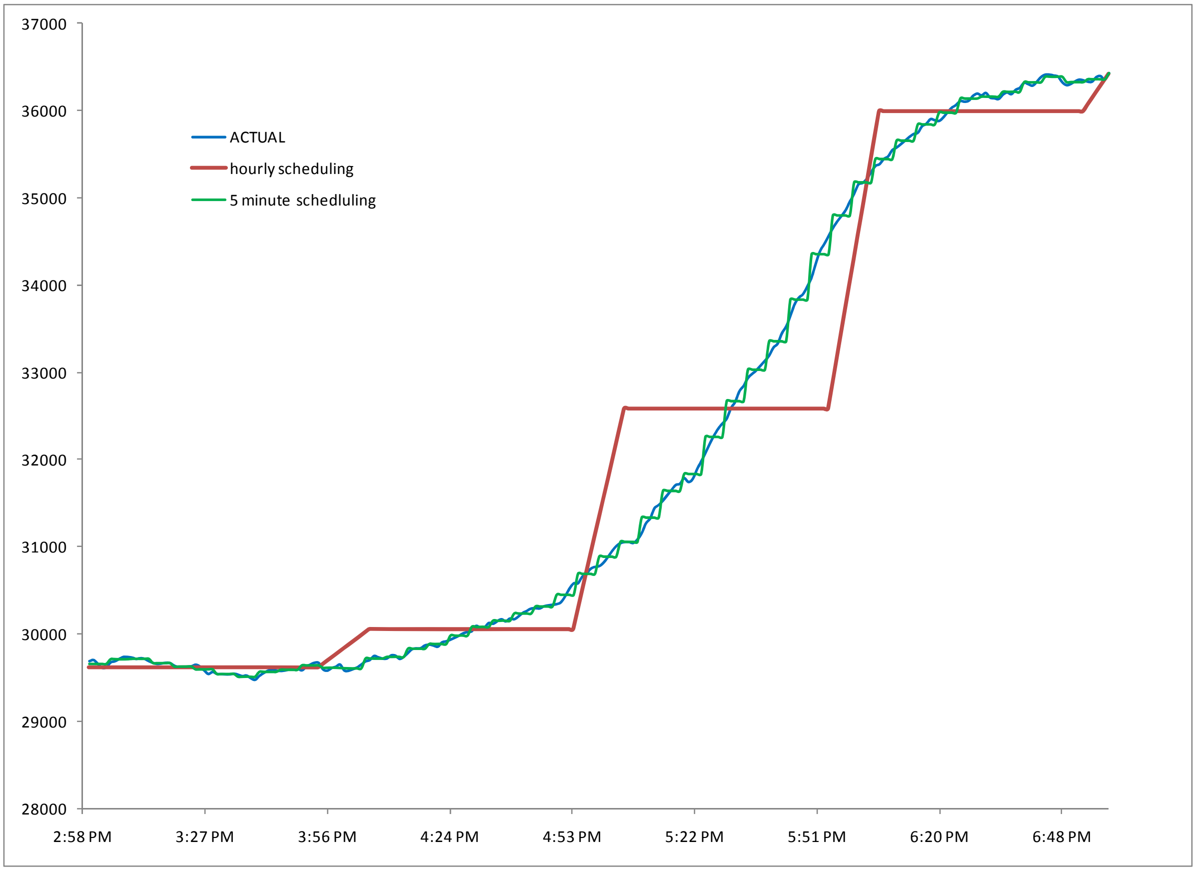

Figure 25: 5-minute scheduling vs. hourly scheduling.

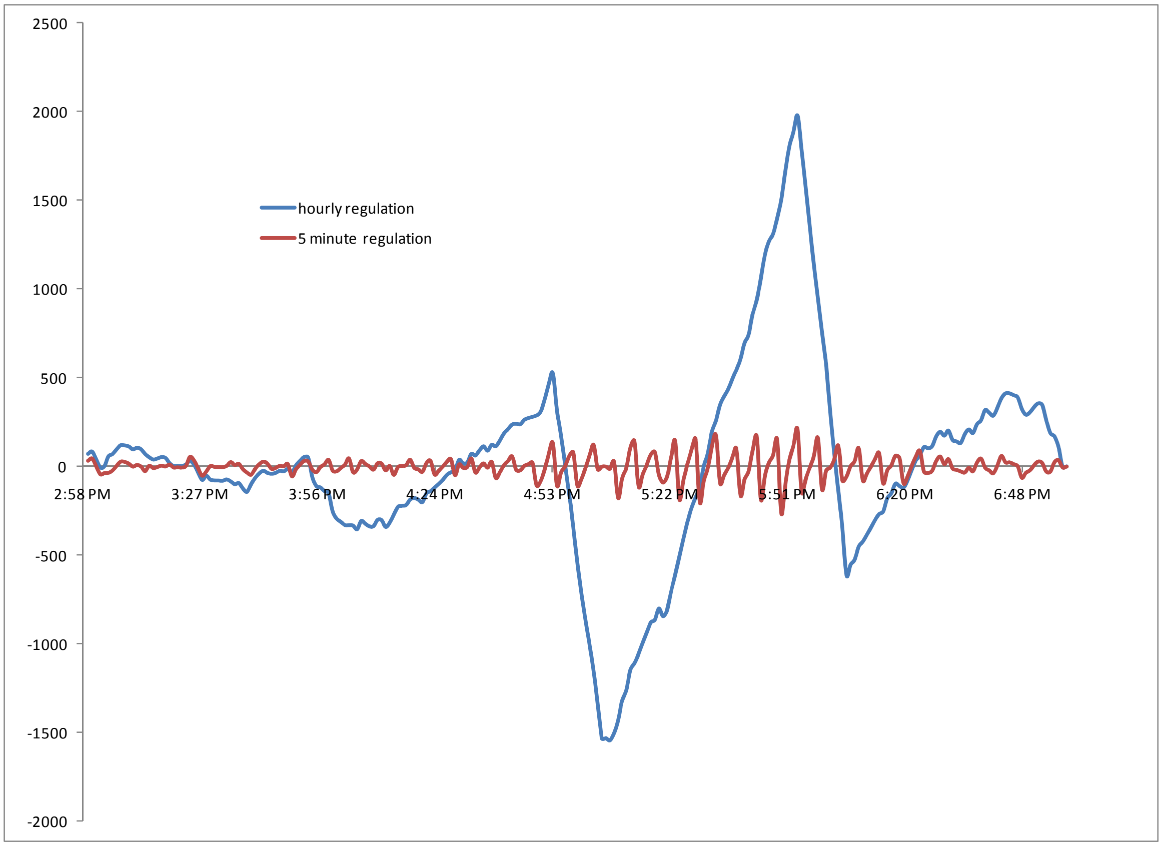

Figure 26: Regulating reserve needs for 5-minute scheduling vs. hourly scheduling.

Alberta Electric System Operator: In the Alberta Electric System Operator [40], there is a requirement for *regulating reserve* resources to have a delay to the AGC signal that is no longer than 28 seconds during normal conditions and 40 seconds during short-ramp control signals – AGC signals that reverse from a raise to a lower signal or vice versa. They also require a minimum of 15 MW of reserve to be provided by an individual unit, as well as a (MW/minute) ramp rate of 1/10 times the maximum regulation range (MW) the resource is qualified to provide. AESO also requires *regulating reserve* resources to be equipped with operating speed governors.

Demand response: Responsive load is beginning to provide Regulation Reserve. Alcoa's aluminum smelter in Warrick, Indiana, provides *regulation service* to MISO by participating in MISO's ancillary service markets. Response is fast and accurate [41]. Loads are expected to begin providing *regulation* to the NYISO in the near future [42].

Ace Diversity Interchange: Northern Tier Transmission Group, a group of transmission providers and customers cooperating to improve operations of the grid has implemented the ACE Diversity Interchange (ADI) program [43]. Based on the principle of load diversity, some balancing areas experience a surplus of generation at the same time that others have a deficit. The ADI pools the *regulation reserve* signal from participating BAs, and performs a simple netting of signals before sending out revised control signals to participating AGC generating units. The result is that overall *regulation reserve* needs are pooled, resulting in less needed response. Initial field trials have, at the time of this writing, been limited to a maximum of 25

MW of reduced regulation to ensure that ADI does not cause any un-foreseen reliability issues. The ADI was suspended during field trials of BAAL, but has recently been restored. The impact of the ADI is to reduce overall regulation, and thus, *regulating reserve*, from the participating utilities.

3.2.2 Following Reserve

ERCOT: In the ERCOT system [44], *non-spinning reserve service* seems to closely align with our definition of Following Reserve. The *non-spinning reserve* requirement is calculated based on the 95th percentile of the observed hourly net-load forecast error from the past 90 days. ERCOT will then subtract its *regulation up* requirements from this 95th percentile to determine the full requirement for *non-spinning reserve*. This essentially says that between the *regulation up* and *non-spinning reserve*, it can cover 95% of the net load forecast errors. The requirement also says it can be used for loss of generating units. The requirement for participation in providing this service is to be either on-line or off-line, with full response required within thirty minutes.

3.2.3 Contingency Reserve

Western Electricity Coordinating Council: The Western Electricity Coordinating Council (WECC) oversees operations and planning procedures performed in the Western Interconnection of North America [45]. It is the regional reliability organization that is responsible for ensuring reliability standards in this interconnection. As a NERC Reliability Region, WECC imposes an additional requirement that supplements the NERC *contingency reserve* requirement. The WECC rule requires an amount of *contingency reserves* equal to the greater of (1) the most severe single contingency, and (2) the amount equal to five percent of the total load served by hydro generation, and seven percent of the total load served by thermal generation in the balancing authority or reserve sharing group. A reserve sharing group (RSG) is a collection of balancing areas that collectively maintain, allocate, and supply *contingency reserves* required for each balancing authority's use when recovering from a contingency event within the group. WECC also requires that at least half of the balancing area's (or RSG's) reserves must be spinning, which is defined as being automatically responsive to frequency deviations and being able to fully respond within 10 minutes. The rest of the *contingency reserve* may be met by off-line non-spinning reserve, interruptible load, or interchange transactions. All of this reserve must be able to fully respond within 10 minutes.

WECC has proposed changing this rule to require the greater of (1) the most severe single contingency and (2) the amount equal to three percent of the net generation (generation minus station service power) and three percent of the total load being served in the balancing authority or reserve sharing group.

Demand Response: Appropriately responsive load can provide any of the *contingency reserves* in all regions with the exception that WECC rules currently do not allow demand response to provide spinning reserve [46]. Technically, demand can provide better reliability response than generation since full response is usually achieved immediately by tripping the load. Loads can respond autonomously to frequency deviations with under-frequency relays set to operate in the generator governor range. Different loads can trip at different frequencies, providing a droop curve. Response speed is only limited by the communications when responding to system operator command, and this is typically much faster than the 10 minutes allowed for full generator response. Demand response is typically a good match for *contingency reserve* requirements because the interruptions are typically shorter than those required for peak reduction, and response speed is no longer a major technical challenge [10].

Both PJM and ERCOT limit the amount of *contingency reserves* which can be supplied by demand response. PJM currently limits demand to 25% of the total reserve requirement and states that there should be at most one level of operator intervention to drop load. ERCOT has loads acting as a resource (LAaR), which can provide contingency reserve either manually or by the triggering of under frequency relays (below 59.7 Hz) during emergency events. No more than 50% of the total *Responsive Reserve Service* can come from LAaR. LAaRs typically offer to supply more than their allotted 50% share and typically at a cost well below that of generation.

Locational Requirements: Many of the regions in NERC, and especially those with large geographic and electrical size, will have locational requirements within their balancing authority. These requirements are based on transmission limitations. For instance, if a certain location of the balancing area is import constrained and a contingency occurs within the area, there may not be transmission capacity available for *contingency reserve* outside the area to help in the recovery. Therefore, many of these import-constrained areas will require that some portion of the *contingency reserve* be located within the constrained area. Some examples of areas that have these locational requirements for their Contingency Reserves are MISO, PJM, CAISO, NYISO, and ISO-New England.

PJM*:* In PJM, zones are separated since there are multiple reliability regions within the PJM region [47]. Most of PJM is in the Reliability First Corporation, but the southern part is in the SERC Reliability Corporation. Therefore, these requirements will be different based on separate rules as well as when transmission is limited from one balancing area to the other. In PJM, the *synchronized reserve requirement* is separated into *tier 1 synchronized reserve* and *tier 2 synchronized reserve*. *Tier 1 synchronized reserves* are defined as marginal units that would be part- loaded based on energy and includes demand response that can automatically curtail load. The VACAR *synchronized reserve* requirement within PJM actually bases its contingency reserve on 15 minutes of response time and is calculated by taking 1.5 times the largest unit in VACAR. However, the *synchronized reserve* in that area must respond within ten minutes.

NYISO*:* The NYISO will also deploy *10-minute operating reserve* when its ACE is less than -150 MW, rather than rules strictly on contingency events [39].

3.2.4 Ramping Reserve

There are no known regional rules regarding Ramping Reserve.

3.2.5 Primary Reserve

WECC*:* As discussed in the *contingency reserve* requirement, at least half of the contingency reserve requirement of WECC must be spinning, which is specifically defined as being responsive to frequency [45]. WECC also discusses in its reliability criteria that governor droop shall be set at 5% and that governors shall not be blocked unless required by regulator mandates. In some recent proposals, a governor droop setting criterion has been developed, which suggests governor droop between 3% and 5% instead of strictly at 5%, and stricter rules requiring generator owners to operate with their governors in service [48].

A *frequency responsive reserve* (FRR) procedure has been proposed in a WECC whitepaper [49]. This is also related to contingency reserve in that it describes the amount of *contingency reserve* that must be responsive to frequency. The proposal and white paper are strictly for loss-of-supply resource events, and not for loss-of-load events. The proposal argues that the quality of the current *spinning reserve* (Contingency Reserve) is only monitored by the DCS, as discussed in section 3.1.3. The proposal allocates an amount of *FRR* that is proportional to the size of the balancing authority system, and that is technically defensible for the need of *FRR*. The proposal suggests a *FRR* amount for all of WECC that is allocated to each balancing authority based on size and a characteristic of response to frequency as a percentage of the system on-line capacity per 0.1 Hz. This is the frequency response characteristic (FRC). The goal is that all balancing authorities will assist in arresting frequency decline no matter where the cause occurred, and that this will result in a more balanced response across the system. The method looked at how different units may contribute individually to different FRC and FRR requirements. What it also shows is that higher FRR requirements may result in more concentrated contributions from individuals, whereas higher FRC requirements would result in more distributed response. Therefore, careful selection of these requirements will be required to receive the desired system response requirement, as well as individual balancing area contribution to the system requirement.

ERCOT: ERCOT protocols have more detail on the required response [50]. The combined response of all generating resources on the ERCOT system shall be at least 420 MW/0.1 Hz. ERCOT is the only balancing area that guides its minimum frequency response. The protocols also discuss the required *primary frequency response* from wind powered generation resources with standard generation interconnection agreements signed after January 1, 2010. The wind generation plants should have adjustable dead bands to match those of other conventional resources or that which is provided in the operating guides, and a similar droop to the other resources of 5%. It also says that wind generation resources with interconnection agreements signed on or before January 1, 2010, shall have *primary frequency response* capabilities by December 1, 2010, if ERCOT believes this is physically practical.

ERCOT has also been working on requirements for inertial response in its interconnection [51]. An on-line tool was developed to estimate how much *inertial frequency response* is available in the balancing area during scheduling processes. *Inertial frequency response* is calculated by dividing the total MW loss during a disturbance by the frequency difference from the instant before the disturbance and frequency nadir (divided by ten to get in terms of MW/0.1Hz). The *inertial frequency response* estimator tool (IFRET) was introduced in February 2010 to estimate the inertial frequency response by observing the system load, total on-line conventional generation capacity, spinning reserves available, and ratio of wind generation to total generation. If the tool shows insufficient *inertial frequency response*, the system operator then has the ability to change the commitment of units.

Hawaii: In the Hawaiian Electric Company [52], a 3-second *quick load pick-up* (*QLPU*) requirement is in place. This requirement states that sufficient generating capability must be available such that, upon the trip of any unit, the remaining units will have sufficient QLPU to restore system frequency to 58.5 Hz within 3 seconds.

Hydro Quebec: Hydro Quebec is a single balancing area synchronous interconnection in the province of Quebec, Canada. Hydro Quebec has passed recent standards that apply specifically to wind power plants [53]. The requirement is for all wind power plants larger than 10 MW to be equipped with a frequency control equipment to provide emulated inertial response. The requirement does not apply to steady-state frequency control (i.e., governor control). The requirement states the entire wind plant to have an inertia constant similar to that of a conventional synchronous generator, of 3.5s. The original requirements also stated the wind plant must give an increase of power of 5% above its current output during under-frequency, for duration of 10 seconds. Subsequently, Hydro Quebec ran some simulations to understand what parameters the wind plant must attain in order to get similar frequency response to a system with all synchronous generators [54]. The study assumed two different types of control, a step function, where the response would come immediately, and a proportional function where the response would come slower. The results are shown below in Table VI.

Table VI: Hydro Quebec parameters for wind power plants providing emulated inertial response

Parameters	Parameter Description	Proportional Function	Step Function
Deadband	The frequency deviation at which the emulated inertia response will activate.	0.3 Hz	0.5 Hz
Active power contribution	The minimum amount of energy above the current power output the plant must produce.	6%	6%
Duration of the active power contribution	The minimum time at which the plant must sustain its active power contribution before going into recovery phase.	10 seconds	10 seconds
Activation time	The maximum time after the event occurs before the response is given.	1 second	1 second
Transition time	The minimum time at which the active power contribution goes into a recovery phase in order to recover the energy and bring turbine speed back to its initial speed.	3.5 seconds	3.5 seconds
Maximum generation reduction during recovery phase	The maximum amount of power reduction when the plant goes into recovery phase.	20%	20%

3.2.6 Secondary Reserves

See Section 3.2.3 and Section 3.1.4.

3.2.7 Tertiary Reserves

WECC: WECC requires that reserves shall be restored no later than 60 minutes following a disturbance event [45].

NPCC: The Northeast Power Coordinating Council (NPCC) has more stringent rules than NERC for Tertiary Reserves of Contingency Reserves [55]. In its "Operating Reserve Criteria," NPCC states that its *30-minute operating reserve* shall at least equal one-half its second highest contingency loss. However, the same standard says that the *10-minute reserve* (Contingency Reserve) shall have the same recovery restoration as the requirement of NERC's DCS, 105 minutes following a disturbance. The 30-minute reserve can be synchronized or offline but available within 30 minutes.

NYISO: The New York State Reliability Council (NYSRC) [56], which oversees reliability criteria of NYISO, which is also within the NPCC, directly states in its reliability manual that its 10-*minute reserves* (Contingency Reserve) shall be restored within 30 minutes following a disturbance. The criteria also states that NYISO shall have sufficient *30-minute operating reserve* equal to one-half of the *10-minute operating reserve*, which is based on the most severe contingency, and therefore, more strict than the NPCC requirement of one-half its second highest contingency loss.

3.3 Europe: ENTSO-E (UCTE)

The European Network of Transmission System Operators for Electricity (ENTSO-E) was fully functional in July 2009 and coordinates security, adequacy, markets, and sustainability for the 42 transmission system operators (TSO) in Europe. The regional groups are based on the synchronous interconnections and they are shown below in Figure 27. Its vision is to remain the focal point for all European technical policy issues related to the TSOs. Therefore, many of the reliability standards that guide the use of operating reserves in European TSOs are maintained within the ENTSO-E. In this section, we will focus on the policies that were set forth for the prior organization of the Union for Coordination of Transmission of Electricity (UCTE), which is the major portion of the ENTSO-E and represents the Continental European Interconnection [57]. UCTE was the predecessor of ENTSO-E, specifically, for the TSOs of that interconnection.

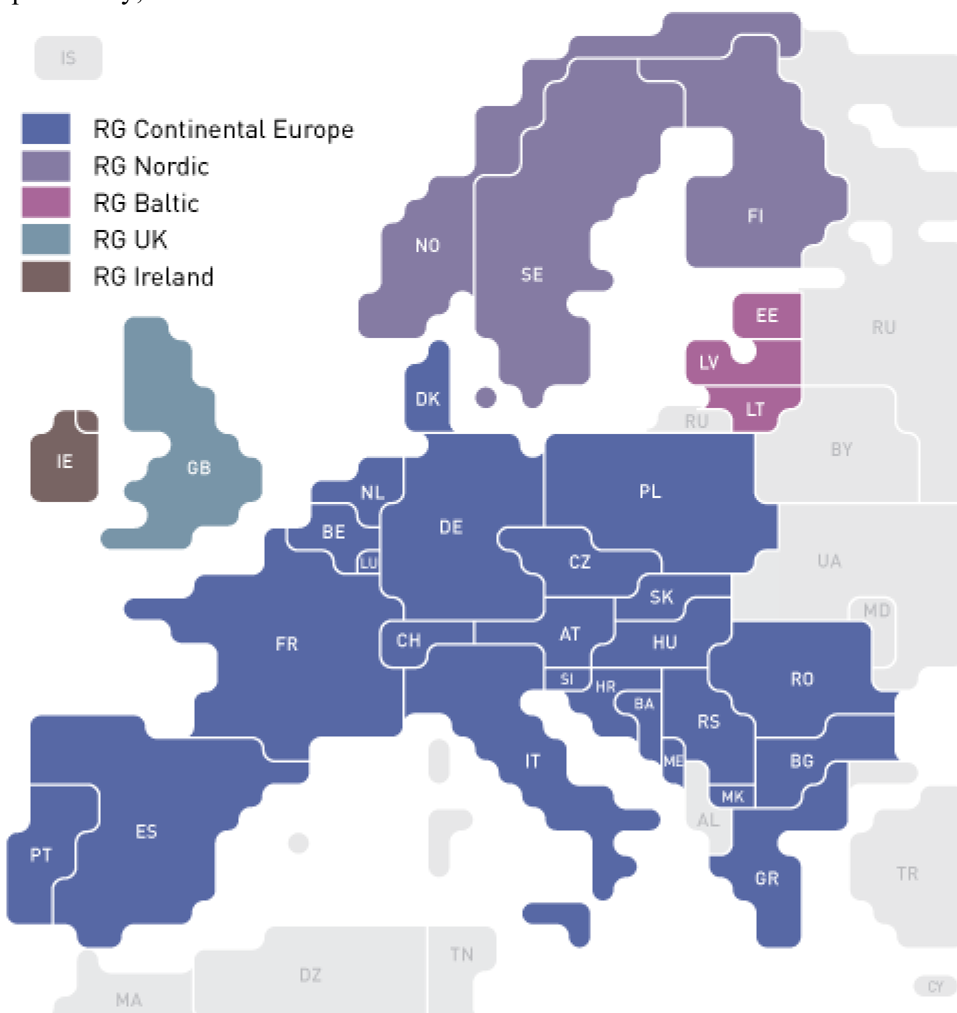

Figure 27: Countries and Interconnections under the ENTSO-E.

3.3.1 Regulating Reserves

Terminology: The term *secondary control reserve* is used for control of ACE. This same term is used for control of normal conditions and disturbance conditions, and therefore, the same term is described herein with both Regulating Reserve and Contingency Reserve, and is separated where appropriate based on its application.

Required Amount: *Secondary control reserve* is held by control areas to balance normal deviations of frequency and interchange schedules that occur based on load and generation changes. The requirements can either be based on empirical analysis or by probabilistic methods. The empirical method determines a minimum requirement for "noise signals" and is calculated for each control area with the following formula:

$$R = \sqrt{a * L_{max} + b^2} - b$$

a is currently set at 10 MW and b at 150 MW. The L_{max} is defined as the hourly maximum of the load of the day. The technique is called the "control capability for variations."

There is a second recommended requirement for sizing of *secondary control reserve*, known as the "probability for reserve deficits" method. This is a probabilistic sizing approach which directs the requirement to enable the control of the ACE to zero in some percentage of all hours during the year. The UCTE also recommends how ACE should be corrected by *secondary control reserve*. ACE is calculated similarly to how it is in North America.

$$ACE \; = \; P_{calc} - P_{sched} + K(F_a - F_s)$$

K is equivalent to B in the NERC ACE equation, i.e., the frequency bias constant. The UCTE policy recommends K be calculated as the product of the contribution coefficient c_i of the area and the overall network power frequency characteristic. The overall network power frequency characteristic is the total amount of frequency response including load response and the contribution coefficient is the share of frequency response for the balancing area based on the control area's share of total generation in the interconnection for the year. Also, UCTE recommends following ACE with PI (proportional integral) type control, as can be seen in the following equation:

$$P_{control} = -\beta * ACE - \frac{1}{T_n}\int ACE * dt$$

where β is the proportional factor and T_n is the integration time constant. There is no recommendation on values for these constants. However, the policy discusses typical values ranging from 0 to 0.50 for β and from 50 to 200 seconds for T_n

There is no formal compliance measure set forth by the UCTE. However, during committee sessions, the TSOs will present their ACE performance for certain time periods, and areas with bad performance will be pressured into either procuring more *secondary control reserve* or improving performance by some other means.

Provider: The UCTE does not recommend anything for what specific technologies are supposed to provide *secondary control reserve*. It also does not recommend the ramping capability required. It does however require that 66% of the *secondary control reserve* used for a control area be geographically located within that control area.

When deployed: UCTE recommends that *secondary control reserve* must start its control action at most 30 seconds after the control direction was given. The cycle time is recommended to be between 1 and 5 seconds.

How deployed: *Secondary control reserve* is under automatic control. The UCTE describes the organization of control that each control area should follow. There are three different schemes that it defines. Centralized control is defined by control that is performed centrally by only one control area in the designated control block. Pluralistic control is a decentralized control scheme that has all control areas in a block regulate their own control area on their own. Hierarchical control has one control area operate the control block, and then the subordinate control areas will regulate their areas based on the control given from the directing control area.

Proposed changes: There are no known proposed changes to the *secondary control reserve* requirement for the UCTE.

3.3.2 Following Reserves

A form of *Tertiary control reserve* labeled as *schedule-activated tertiary control reserve* is activated with relation to the predefined time frame of the energy schedule or energy market intervals. The reserve would then be used for slower normal variations that occur in a control area to minimize ACE. Requirements for this specific purpose are not defined in any detail, however.

3.3.3 Contingency Reserves

Terminology: Contingency Reserves in Europe are distributed among three different terms. These are defined as *primary control reserve*, *secondary control reserve*, and *tertiary control reserve*. Each is faster responding than the next, and the slower reserve is supposed to replace the faster reserve as shown in Figure 28.

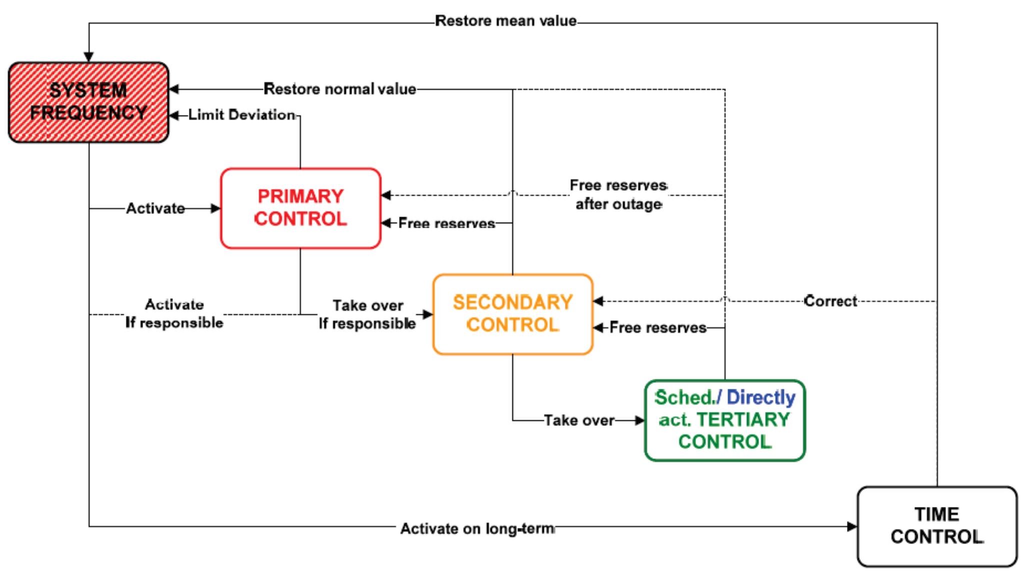

Figure 28: Role of different types of Contingency Reserve in the UCTE region.

The definitions of each of these reserve types are stated below:

Primary Control Reserve: Local automatic control based on frequency deviations that immediately respond to oppose the frequency deviation, and is shared by all control areas in the interconnection where the deviation occurs. (Also see section 3.3.5 for further information).

Secondary Control Reserve: Centralized automatic control that responds immediately, but from direction of the control center to oppose area control error, which consists of frequency deviation and interchange error. The obligation to provide *secondary control reserve* lies with the control area where the loss of supply occurred. (Also see section 3.3.1 for the separate use of this reserve type and section 3.3.6 for further information).

Tertiary Control Reserve: Manual control that is used to either replace the *secondary control reserve* or complement it in the case that the *secondary control reserve* is not sufficient for purposes of restoring frequency and interchange schedule deviations on its own following a disturbance. (Also see section 3.3.6 for further information).

Required Amount: In UCTE, there are specific requirements for both *primary control reserves, secondary control reserves*, and *tertiary control reserves*. We discuss the requirements for *primary control reserve* in the Primary Reserve section. *Secondary control reserve* is used for normal random variations as well as to bring frequency back to nominal level and ACE to zero during large disturbances. *Tertiary control reserve* is used to replace the *secondary control reserve* so that the *secondary control reserve* can be ready for a second disturbance. However, *tertiary control reserve* is also available to supplement the *secondary control reserve* during larger disturbances where *secondary control reserve* cannot alone balance the disturbance.

Provider: UCTE policies provide no requirement on what resource type is allowed to provide these reserves. It does require that a fixed share of 50% of the total needed *secondary control reserve* and *tertiary control reserve* be kept inside the control area.

When deployed: *Primary control reserve* is deployed almost immediately following the disturbance. *Secondary control reserve* is deployed through AGC soon afterward. Once the *secondary control reserve* is deployed, the *primary control reserve* can be released. *Tertiary control reserve* is then deployed later to replace the *secondary control reserve*. This is depicted in Figure 29. During disturbances the ACE should be recovered within fifteen minutes.

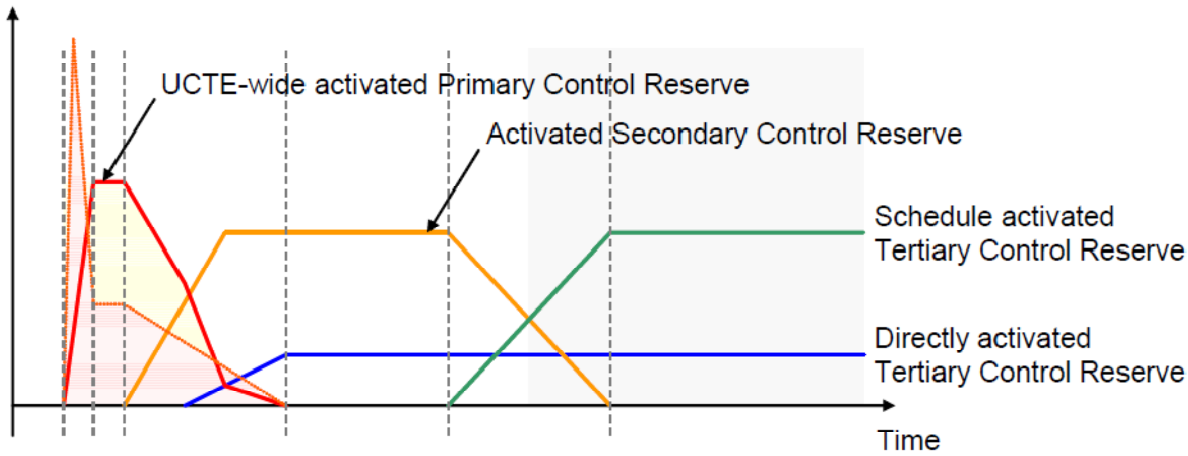

Figure 29: Time line for primary, secondary, and tertiary control reserve following contingency event.

How deployed: *Primary control reserve* is deployed automatically, responding to frequency deviations through governor systems. *Secondary control reserve* is deployed through AGC signals that are from the control center. *Tertiary control reserve* is deployed with manual changes to set points either from notification from the system operator or through schedule changes.

Proposed changes: There are no known proposed changes to any of the Contingency Reserves requirements for the UCTE.

3.3.4 Ramping Reserves

There is no requirement with UCTE on the reserve type corresponding to Ramping Reserve.

3.3.5 Primary Reserves

Terminology: The UCTE uses the term *primary control reserve*.

Required Amount: In the UCTE, *primary control reserve* is used immediately following disturbances to stabilize the system frequency at a stationary value within a few seconds. The total required *primary control reserve* for all of UCTE is 3,000 MW. This value is calculated based on what the UCTE considers the maximum instantaneous power deviation in the synchronous system, which is 3,000 MW. Each control area must provide a share of this reserve as described by contribution coefficients that are determined each year as shown in the equation below. The contribution coefficients (c_i) are based on the control area's share of total energy generated within one year in proportion to the total energy generated in the interconnection. Therefore, the sum of all contribution coefficients would be equal to one.

$$PrimaryControlReserve_i = c_i * 3000$$

The reserve amount given to the interconnection and to each control area must be fully activated when a frequency deviation occurs of greater than ± 200 mHz. Therefore, the minimum network power frequency characteristic of primary control is 3,000 MW divided by 200 mHz or 15,000 MW/Hz (1,500 MW/0.1Hz). The average network power frequency characteristic of primary control ends up being larger at 19,500 MW/Hz (1,950 MW/0.1Hz). This response includes only the response of *primary control* *reserve* from generators and not the self regulating affect of load. The overall network power frequency characteristic, which includes load response to frequency and some surplus-control of generation, is higher than 19,500 MW/Hz.

The UCTE policy also states the minimum total frequency deviation (frequency nadir) is 0.8 Hz. This includes a safety margin of 200 mHz above the highest under-frequency relay setting. This safety margin is maintained due to controller insensitivity, a possible stationary frequency of 50 mHz below nominal, and larger dynamic frequency deviation at the site of the incident not in the model.

Provider: The UCTE policy gives no reference as to who can provide *primary control reserve*. A control area is allowed to increase its *primary control reserve* by 30% up to 90 MW by offering to cover part of the obligations of other control areas.

When deployed: *Primary control reserve* is activated a few seconds following a disturbance event and must be delivered until the power deviation is fully offset by the *secondary* and *tertiary control reserve*. The UCTE recommends a dead band (insensitivity) of no more than 10 mHz. For instances that are less than 50% of the maximum disturbance considered, 1,500 MW, primary reserve must be fully deployed within 15 seconds. For the full disturbance of 3,000 MW, primary control reserve must fully respond within 30 seconds. For disturbances greater than 1,500 MW but less than 3,000 MW, the full response must be within a linear interpolation from 15 seconds to 30 seconds. For example, a 2,000-MW disturbance should have full response within 20 seconds.

How deployed: *Primary control reserve* is deployed through governor control responding directly to frequency deviations.

Proposed changes: There are no known proposed changes to *primary control reserve* from the UCTE.

3.3.6 Secondary Reserves

Terminology: The term *secondary control reserve* is used by the UCTE for control of ACE. This same term is used for control of normal conditions as well as disturbance conditions, and therefore, the same term is described herein with both Regulating Reserve and Contingency Reserve and is separated where appropriate based on its application.

Required Amount: The requirement of *secondary control reserves* can be seen in section 3.3.1 but also a rule of enough *secondary control reserve* and *tertiary control reserve* to cover the largest credible contingency is recommended.

Provider: The UCTE policy gives no reference as to who can provide *secondary control reserve* for contingency events. The correction to the initial ACE must happen within 15 minutes at the latest.

When deployed: The UCTE policy recommends the *secondary control reserve* to start at most 30 seconds following the disturbance period, and must return correction to the initial ACE within 15 minutes at the latest.

How deployed: *Secondary control reserve* is under automatic control. The UCTE describes the organization of control that each control area should follow. There are three different schemes that it defines. Centralized control is defined by control that is performed centrally by only one control area in the designated control block. Pluralistic control is a decentralized control scheme that has all control areas in a block regulate their own control area on their own. Hierarchical control has one control area operate the control block, and then the subordinate control areas will regulate their areas based on the control given from the directing control area.

Proposed changes: There are no known proposed changes to *primary control reserve* from the UCTE.

3.3.7 Tertiary Reserves

Terminology: The term *tertiary control reserve* is used by the UCTE for use of replacing the *primary* and *secondary control reserve* following a contingency.

Required Amount: It is recommended that enough *tertiary control reserve,* including *directly activated tertiary control reserve* and *schedule activated tertiary control reserve,* is available to cover the largest loss of supply in the individual control area.

Provider: The UCTE policy gives no reference as to who can provide *tertiary control reserve* for contingency events. *Tertiary control reserve* is split between *directly activated tertiary control reserve* and *schedule-activated tertiary control reserve. Directly activated tertiary control reserve* is manual reserve used during contingency events at any time, independent from a time-frame of exchange schedules. *Schedule-activated tertiary control reserve* is activated with relation to exchange schedule time frames, which may normally be on 15-minute intervals.

When deployed: *Tertiary control reserve* which also is called *fifteen-minute reserve*, is used to replace the *secondary control reserve* soon after the *secondary control reserve* are deployed following a disturbance. However, there is no recommendation on the required time of when these reserves must be fully replaced to withstand the subsequent disturbance.

How deployed: *Tertiary control reserves* are either deployed manually through operator direction or through directed schedule changes.

Proposed changes: There are no known proposed changes to *tertiary control reserve* from the UCTE.

3.4 Summary

The rules for operating reserves used for active power balance have been developed by years of experience in each region. Systems differ in the types of resources they have available, the system load characteristics, the size and frequency response characteristics, and the transmission network. The philosophy on how each system operator deals with risk can also be quite different. These differences may have led to the different requirements that ensured that each system could maintain reliability. Figure 30 shows the terminology connection between the UCTE operating reserves and NERC operating reserves. Table VII shows some of the policy differences between the two entities.

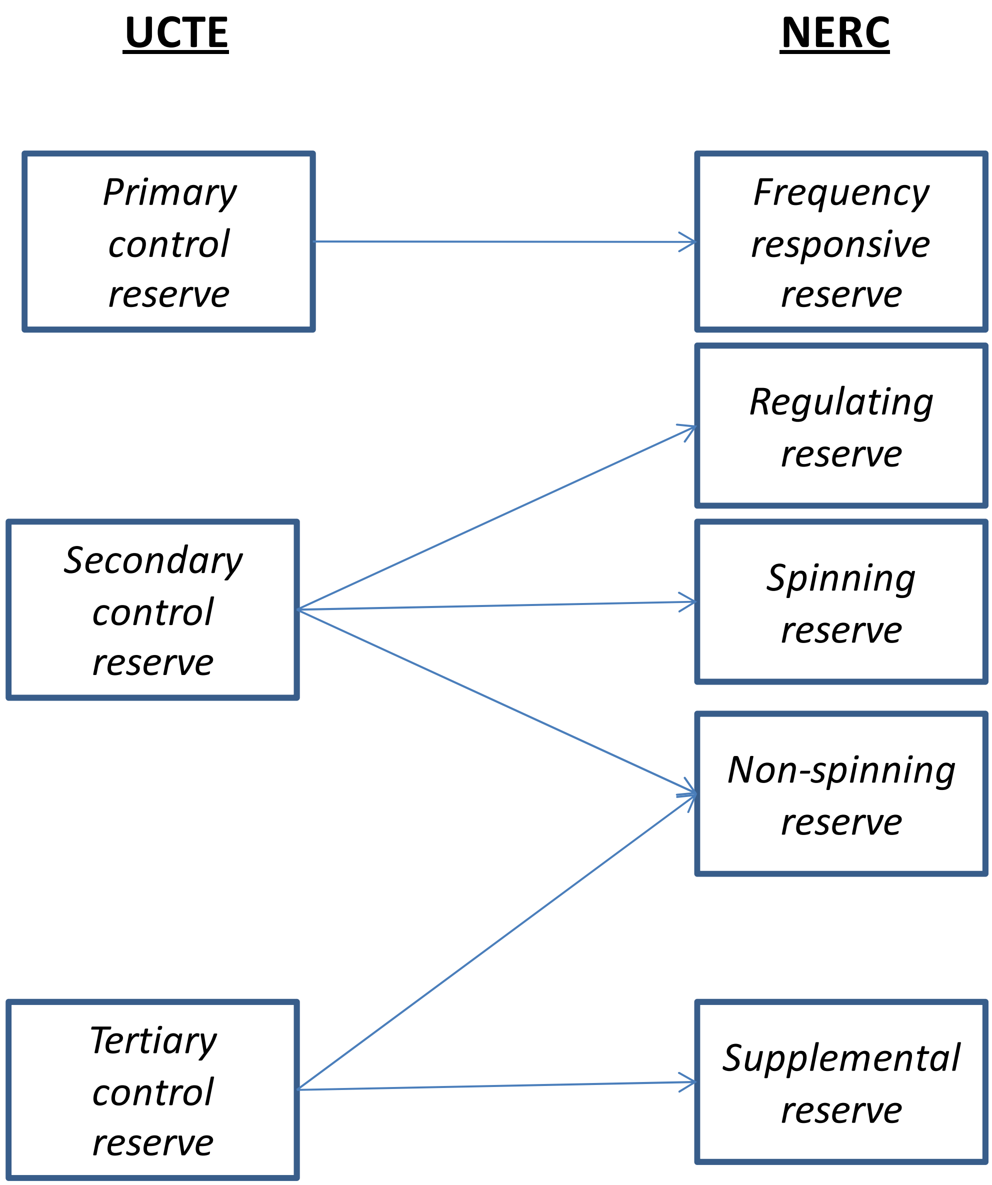

Figure 30: UCTE and NERC terminology.

Table VII: NERC and UCTE policy differences

	N. America (NERC)	Europe (ENTSOE/UCTE)
Regulating Reserve	NERC enforces CPS1 and CPS2 compliance measures but has no policy on what the actual *regulating reserve* requirement quantity should be. The CPS1 and CPS2 drive the requirements which are mostly based on time of day and season.	UCTE does recommend a *secondary control reserve* requirement which is based on a statistical equation and mostly comes from load variability. However, secondary reserve is used for both contingencies and normal variations. There are no compliance measures.
Following Reserve	No requirements	No requirements
Contingency Reserve	DCS requires ACE to be returned to 0 or its pre-disturbance level if negative within 15 minutes. Enough *contingency reserves* required to recover largest contingency. Many regions require at least 50% to be spinning.	Similar requirement to DCS. Return ACE to zero within 15 minutes. Split between *primary*, *secondary*, and *tertiary*. Enough of these reserves should be available to cover the largest contingency.
Primary Reserve	No requirement. Some discussions on a future requirement. Only a frequency bias requirement as part of ACE equation of 1% peak load. Governor dead bands mostly set at 36 mHz and droop at 5%, but not required.	*Primary control reserve* (3000 MW) split between TSOs based on energy contribution. 3000 MW based on largest credible Interconnection-wide event. Full Response at 200 mHz. Response characteristics based on UFLS relay setting and 200-mHz safety margin. 20-mHz maximum insensitivity.
Ramping Reserve	No requirements	No requirements
Tertiary Reserve	No quantifiable requirement but contingency reserve must be replaced within 105 minutes following contingency.	*Tertiary control reserve* requirement is larger than the largest contingency. There is no requirement on how soon any reserves should be replaced.

A key difference between North American and European policies is that NERC has policies that distinguish normal balancing from disturbance events, whereas UCTE does not. However, even though it is used for normal and event conditions, the method for determining the *secondary control reserve* requirement for UCTE is based on normal variability. UCTE also does not have a policy for measuring compliance during normal conditions as NERC does. It will be interesting to see if this difference remains once higher penetrations of VG enter the system. Lastly, the UCTE has a much more stringent requirement for its *primary control reserve* than NERC does for its corresponding *frequency responsive reserve*. It may have been that NERC received the sufficient response without ever having a defined requirement, and therefore, any standard or policy was not needed. We are now seeing NERC evaluating this, and some of the initial discussions look like the policy may be very close to the one that has been established for UCTE. Both regions have very similar requirements for Contingency Reserve, both in the amount required, and timing on how it should respond. Neither regions have enacted any type of Ramping Reserve requirement as of yet.

Very few regions have begun to change any of their operating reserve policies due to the increased amount of VG. ERCOT has begun to consider the possibility of wind forecast errors and the increased variability of wind in its *up and down-regulation reserve* and its *non-spinning reserve.* Many NERC regions have made enhancements to allow newer resources like demand response and energy storage devices to provide different types of reserves. Very few regions in North America are allowing VG to provide any type of reserve. However, ERCOT does require wind power to provide over-frequency *frequency response.* Hydro-Quebec is also requiring wind power to provide an emulated inertial response. In many European TSOs, many of the offshore wind power plants are beginning to be asked to provide active power control as well. Many of the high penetration systems including Denmark and Ireland, are looking at ways in which wind power can assist in reliability by providing different types of Operating Reserves. Overall, many of these policies have been around for a number of years, and significant changes are just starting to be enforced. With higher penetrations, we may see this trend continue.

4 New Proposed Methodologies and Renewable Integration Studies

In recent years, many utilities, independent system operators, transmission system operators, government researchers, and academic researchers have been looking at how systems with large penetrations of VG impact reliability and costs. Some studies include as part and others as the focus, the new requirements of operating reserves that must be kept in order to maintain the same reliability level as without VG. The research has been split between large renewable integration studies and fundamental research projects. The large-scale renewable integration studies are evaluating real systems with simulated increased VG penetration and practical specific changes needed for that particular system, while the research papers are often modeled with test systems and results that are more general to the industry than a specific system. The majority of the large-scale renewable integration studies in 4.1 are based in the United States with the vast majority of the research studies in 4.2 being from Europe. Therefore a broad perspective of U.S. and European research on new methods of operating reserve requirements with high penetrations of VG is presented. Both types of studies are discovering new innovative methods that are advancing the understanding of this area. We will now discuss various studies and some of the new methodologies that have been reported.

4.1 Large-Scale Renewable Integration Studies

In recent years, a number of different entities have initiated or completed studies specifically aimed at analyzing the impacts and costs of operating power systems with large penetrations of variable renewable generation. So far, this has mostly included wind energy, but other technologies (e.g., solar) are being included now as well. The studies usually make assumptions about the nature of the power system at some time in the future. This may include additional load, additional conventional generation, additional transmission, and particular market and operational structures. The studies then use future wind power outputs and analyze the power system using simulations and statistical techniques. The wind power data may be developed using mesoscale modeling techniques to recreate wind power output time-series data of past years' weather as input. The objectives of these studies vary, but most focus on both the operating costs and operating impacts. One of the major operating impacts and major contributions to increased costs of integrating wind and other VG is the increased amount of operating reserve that the power system must hold in order to maintain a system with reliability equal to today's system without VG. We focus on the different methods that these studies have applied to answer these questions. More information on the state of the art of these studies and summaries of the studies can be found in [12][13][14]. An entire collection of these studies can be found on the following website: www.uwig.org/opimpactsdocs.html.

NYISO/NYSERDA: The Effects of Integrating Wind Power on Transmission System Planning, Reliability, and Operations (2005)

The NYISO/NYSERDA wind power integration study titled "The Effects of Integrating Wind Power on Transmission System Planning, Reliability, and Operations" was one of the first major large-scale wind integration studies performed in the United States [58]. The study evaluated the operational impacts of 3,300 MW of wind power in the New York Control Area with a projected peak load of about 33,000 MW. One of the main investigations of the report was to evaluate the needs for increased operating reserves on the system due to the additional variability of wind power. For *spinning reserves* (Contingency Reserves), it was decided that abrupt loss of wind power from all wind plants throughout the state was not a credible contingency, and therefore, the single largest contingency was unchanged. This meant the required *spinning reserves* were also not changed. The study also evaluated *regulation* (Regulating Reserve). The amount of *regulation* did increase with the 3,300 MW of wind based on the increased standard deviation of 6-second variability of net load. The standard deviation was shown to increase by 12 MW. The study team used three standard deviations as was currently practiced to ensure 99.7% of all of these changes would be met by the required *regulation* reserves. This meant that the *regulation* increased by a total of 36 MW to the existing requirement of between 225 MW and 275 MW. Lastly, the study evaluated *load following*. The analysis looked at 5-minute changes in net load and how the standard deviation of these changes increased compared to that of load alone. Five minutes is the current dispatch interval of the NYISO's security-constrained economic dispatch program and corresponds to the load following time frame for the area. The statewide standard deviation increases by 1.8 MW from 54.4 MW to 56.2 MW with the added wind generation. However, it was concluded that this minor change could be accommodated by the existing processes and resources in the N.Y. Control Area without any new requirements.

Minnesota: 2006 Minnesota Wind Integration Study (2006)

The "2006 Minnesota Wind Integration Study" was also a groundbreaking study in the United States, evaluating high penetrations of wind power [59]. This study evaluated up to 25% (also 15% and 20%) wind energy (as percentage of total electric energy for the year, which is approximately 5,700 MW of wind capacity with a peak load of 20,984 MW). This was the highest penetration of wind power studied in the United States up to that time. Therefore, the study had to develop new ways of determining the operating reserve requirements as traditional requirements are almost entirely based on load and conventional generation.

The study focused on the state of Minnesota, but balancing area operation was considered to be controlled by MISO. This meant that the *regulating reserves* would be considered for the larger footprint rather than smaller utility areas. Generally, the *regulating reserve* requirement as a percentage of the load decreases as the size of the balancing area increases, as can be seen in Figure 31, which is taken from the report. The *regulating reserve* requirement evaluated the added variability of wind, but calculated it to be a 2- MW standard deviation for every 100-MW wind plant installed. This calculation was based on operational data from existing wind plants. The ratio was used to calculate the *regulating reserve* requirement as seen in the following equation:

$$Reg\ Req = k\sqrt{\sigma_{load}^2 + N(\sigma_{W100}^2)}$$

where k is a factor relating *regulating reserve* capacity requirement to the standard deviation of the variations (assumed to be 5 in this study, reflecting current practices); σ_{load} is the standard deviation of variations from load; σ_{W100} is the standard deviation of variations from a 100-MW wind plant; and N is the wind generation capacity in the scenario divided by 100. The results showed increases of 12, 16, and 20 MW for the 15, 20, and 25% cases, respectively.

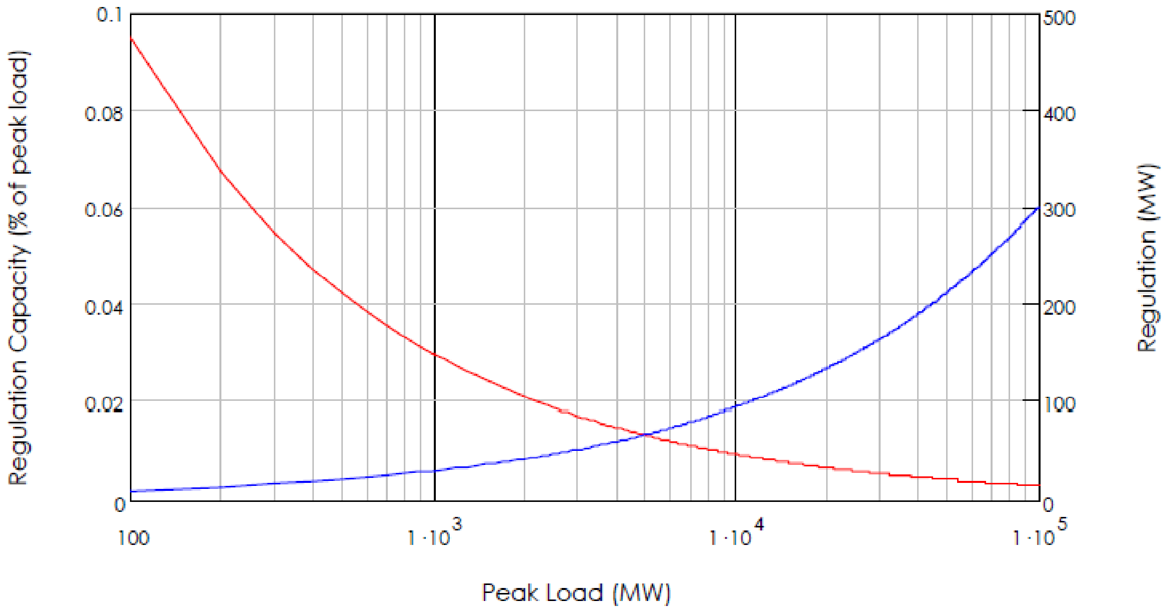

Figure 31: Approximate regulating requirements for a balancing authority as a function of peak demand.

The Minnesota study quantified two other reserve categories that the New York study did not. In the Minnesota study, these are defined as *load following* and *operating reserve margin* (Following Reserve and Ramping Reserve, respectively). *Load following* was calculated as twice the standard deviation of the 5-minute changes in the net load, and increases ranged from 10 to 24 MW for the three cases. The *operating reserve margin* was allocated specifically for hourly forecast errors in the net load. The forecast of net load was a forecast of load, which is generally pretty accurate minus a forecast of wind, which is assumed to use persistence. This analysis assumed a dynamic requirement, one that was not constant for all hours, but in fact was a function of the amount of expected wind capacity. The consideration was based on the fact that the variability of wind is highest when the wind capacity is in the middle range (i.e., 40-60% of total capacity) due to the wind turbines being on the steepest parts of their wind-speed-to–wind-power conversion curves. Therefore, more reserves were needed during times when wind generation was at its middle range compared to times of very low wind generation or times of very high wind generation. The requirement was also based on two standard deviations. For the 25% scenario, this *operating reserve margin* was determined to be on average 538 MW, with a maximum of 704 MW. Lastly, the study also concluded that no additional *contingency reserves* would be required due to the wind because the largest contingency was unchanged.

Arizona Public Service Wind Integration Cost Impact Study (2007)

The "Arizona Public Service Wind Integration Cost Impact Study" was completed in 2007 and evaluated 1-10% wind energy penetrations as a percentage of total annual energy, with some different sensitivity on geographic diversity [60]. The study modeled the Arizona Public Service (APS) transmission system with modeled wind power data used as input. The study used a similar approach to that of the 2006 Minnesota study for calculating the *regulating reserves*. It used a standard deviation of 1.5 MW for a 100-MW wind plant and added enough 100-MW wind plants to reach the desired penetration of wind. The square root of the sum of the standard deviation of load squared with that of each wind plant's standard deviation squared was calculated to get the total standard deviation of 1-minute changes in net load. Again, wind and load were determined to be uncorrelated on this time frame, as were outputs of wind plants with each other. A multiple of five times the standard deviation was used in this study to capture virtually all of these changes with the *regulating reserves*. The increases due to 4% and 10% wind energy were calculated as 2.4 MW and 6.2 MW of additional *regulating reserves*, respectively. This amount was used as a constant addition to *regulating reserve* requirements for all hours of the year.

An additional reserve defined as *hour-ahead firmness* (Following Reserve) was calculated. The reserve was calculated by evaluating the combined errors of load and wind in the hour-ahead time frame. The study showed that the increases in these reserves due to wind were highly dependent on the accuracy of the current short-term load forecast. The more accurate the load forecast, the more impact the errors of the wind forecast would have on the net load forecast errors based on the modeled forecast data from the study. For the 4% wind case, if the hour-ahead load forecast had a standard deviation of 0.5% of the peak, the increase in these reserves due to wind would be 58.6 MW, whereas if the load forecast error had a standard deviation as large as 2% of the peak, the increase of this reserve due to the wind was only 22.4 MW. These cases assumed the reserve was calculated by taking two standard deviations of the net load forecast errors based on analysis for the entire year. For the 10% wind case and lowest load forecast error, the increase of this reserve was 174 MW, due to a standard deviation increase of the wind case compared to the load alone case of 87.2 MW.

Unlike the Minnesota study, the APS study used the same approach for the day-ahead forecast reserve as it did for the hour-ahead reserve. For both of these, the terminology used was "firmness" rather than operating reserve. It meant that only a certain percentage of the associated wind forecast would be "counted on," which is essentially identical to adding additional reserve while counting on 100% of the wind forecast. This analysis took results of typical day-ahead wind forecast errors and the study used a conservative approach for the day-ahead firmness factor. It assumed a 20% error as being the average day-ahead wind forecast error for the most difficult wind sites. Taking two standard deviations, this turned out to be using a 60% firmness factor for the day-ahead. This is equivalent to adding 40% of the forecast of wind as operating reserves to be ready for the error in real-time. The day-ahead forecast analysis was less sophisticated than the hour-ahead forecast as it did not take the load forecast errors in to account and assumed the 20% error even for wind plants that were not as difficult to forecast (and may have only had errors equal to 10% of the wind plant).

California ISO: Integration of Renewable Resources (2007)

In late 2007, a study called "Integration of Renewable Resources" was completed by the California ISO [61]. The study evaluated the impacts of wind and solar on the California ISO. The study evaluated 6,700 MW of wind generation on the system.[7] The study focused on *regulation* (Regulating Reserve) and *load following reserves* (Following Reserve). It also evaluated how ramp requirements may be increased, but did not specifically quantify a Ramping Reserve requirement. In order to determine some of the amounts of these operating reserves, it required a detailed assessment of the scheduling protocols of the California ISO. This can be seen in Figure 32. At the time of this study, the California ISO creates hour-ahead schedules based on market closures that occur 75 minutes prior to the hour. Load forecasts are actually

[7] The study evaluated other renewables including solar and geothermal, but the integration impacts focused on wind generation.

created 2 hours before the hour for the hour-ahead scheduling. The hour-ahead schedules are hourly blocks with transitions between the hours scheduled with ramps that start ten minutes before the hour and end ten minutes after the hour. Then every 5 minutes, the economic dispatch program is run. This uses forecasts that are from 7.5 minutes prior to the start of the interval, and dispatch instructions are sent 5 minutes prior to the start of the interval. For wind forecasts, persistence forecasts are used, and therefore, the output of each wind plant at 7.5 minutes prior to the start of the interval is the forecast for the entire 5-minute interval. Obviously, the generation and load is changing within each 5-minute interval as well. With this procedure in mind, the study determined the difference in adjustment of the economic dispatch and the hour-ahead schedule as *load following* and the difference between the realized values and the economic dispatch as *regulation*.

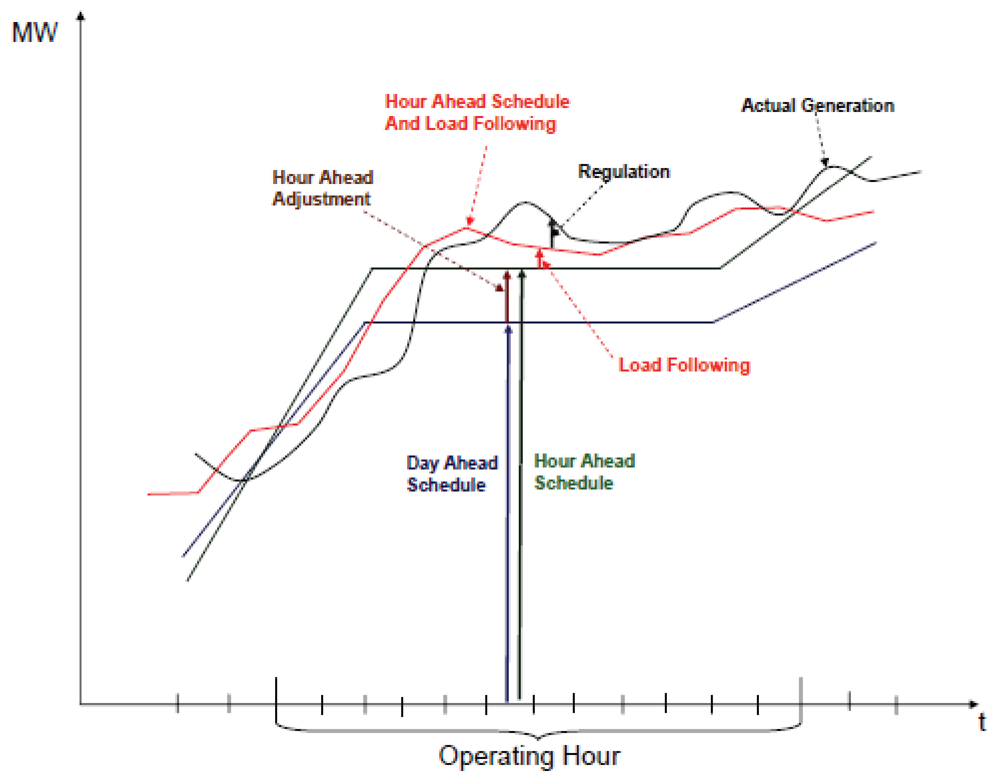

Figure 32: Load following and regulation based on CAISO scheduling process.

This analysis meant that the *load following* and *regulation* are highly dependent on the hour-ahead and very short term (e.g., 7.5 min ahead) wind forecast accuracy. Generally, in the previous studies the variability was the only source of information used to calculate these requirements. For this analysis, the capacity, ramping capability, and ramp duration were calculated for both *load following* and *regulation* using the “swinging door algorithm.” The swinging door algorithm evaluates high resolution data and determines turning points where consistent signals make it so a previous point is not within the ramping range [62]. From this algorithm, a three dimensional scatter plot can be created with capacity, ramp rate, and ramp duration as the coordinates. A box representing a probability so that a certain number of these points are within the box and others are outside is drawn. Once the probability is determined, each of the three components can be determined to define the *regulation* and *load following* requirements in terms of capacity (MW), ramp rate (MW/min), and ramp duration (seconds). These were all calculated as hour-specific requirements. The increases were much more significant than some previous studies with *load following* capacity requirement maximums increasing as much as 800 MW (2,700 MW existing increased to 3,500 MW) and *regulation* capacity requirement maximums increasing as much as 230 MW (250 MW existing increased to 480 MW) for upward directions. The downward *load following* and *regulation* maximum increases were 600 MW (2,450 MW existing increased to 3,050 MW) and 500 MW (250 MW existing increased to 750 MW), respectively. These are the maximum increases by hour, and the average increases were lower. For both *load following* and *regulation,* the ramp rate requirements also increased quite significantly with the added wind generation.

Multi-hour ramps were also an issue that was evaluated in the study. Although the impacts of these large ramps were analyzed, no Ramping Reserve requirement was discussed in the study. This ramping requirement was shown to increase with added wind generation, depending on season. Figure 33 shows the maximum ramp increases for morning and evening ramps due to increased net load ramping.

Seasons	*2006 Morning Ramps MW*	*20% RPS Expected Morning Ramps MW*	*Change due to Intermittency MW*	*2006 Evening Ramps MW*	*20% RPS Expected Evening Ramps MW*	*Change due to Intermittency MW*
Spring	6,860	8,494	955	7,962	9,788	984
Summer	10,090	12,664	1,529	10,589	12,135	427
Fall	7,229	8,995	1,023	11,511	13,483	740
Winter	6,979	8,631	926	7,856	9,293	603

Note: Morning Ramps – Spring & Fall: HE7 through HE9; Summer: HE8 through HE10; and Winter: HE6 through HE8. Evening Ramps – All seasons: HE22 through HE24

Figure 33: Increases in multi-hour ramps due to the added wind generation.

ERCOT: Analysis of Wind Generation Impact on ERCOT Ancillary Services Requirements (2008)
In early 2008, "Analysis of Wind Generation Impact on ERCOT Ancillary Services Requirements" was performed for ERCOT that specifically focused on the impacts that up to 15,000 MW of wind power would have on ERCOT's ancillary services [38]. The study evaluated this wind penetration on 2008 load levels which is about a 28% penetration of wind capacity as a percentage of peak load. The study also ignores present day transmission constraints as it assumes through the Competitive Renewable Energy Zone (CREZ) process that sufficient transmission will be built to support this amount of renewable expansion [63]. As discussed in section 3.2, ERCOT has five main categories for operating reserves: *regulation service up* (Regulating Reserve), *regulation service down* (Regulating Reserve), *responsive reserve service* (Contingency Reserve – Primary and Secondary Reserve), *non-spinning reserve service* (Following Reserve), and *replacement reserve service*. *Replacement reserve service* is actually accounting for a difference between market bids and ISO forecasts (i.e., discrepancy between operations and markets), and therefore, not really defined in the definitions of section 2. The requirements for each of these services were evaluated using state-of-the-art techniques.

The *regulation service up* and *regulation service down* is defined in the study as the deviation of the net load from the economic dispatch set point. The analysis of both up and down services showed that the general nature of the *regulation service* requirements did not change with the increasing amount of wind power, but that certain hours did have new peaks driven almost entirely by wind power. Figure 34 shows the changes in both *regulation up* and *regulation down service* for the hourly requirements. The average increase for both up and down service is about 18 MW. As can be seen by the minimum of up service and the maximum of down service, the added wind doesn't necessarily always increase the deployment. The study determined the requirements by statistically capturing the 98.8th percentile of deployments by hour by month. On average, this increased the requirement from 232 MW to 285 MW for *up regulation service* and -233 MW to -281 MW for the *regulation down service* when increasing to 15,000 MW of wind power. The study also recommended considering the increased amount of capacity of wind power from the previous year where the deployments are studied to determine the requirements so that increases (or decreases) can be made with the consideration of the higher wind penetration. Lastly, the study recommended using the wind forecast to add or subtract from the *regulation service* requirement based on what the predicted wind output will be.

	+REG	-REG
Mean	17.7	-18.2
Sigma	64.9	65.1
Maximum	444.2	265.3
Minimum	-287.2	-453.1

Figure 34: Summary of changes in hourly maximum regulation deployments of 15,000-MW wind capacity.

The study did not go into great detail on the requirements related to *responsive reserve service* and *non-spinning reserve service*. The maximum magnitude of net load ramp events with 15,000 MW of wind power would increase from 3,101 MW to 4,502 MW in a 30 minute period. It is likely that these ramp periods would have to be covered by some combination of the *responsive reserve service* and *non-spinning reserve service* (some combination to equal the Ramping Reserves). They discussed the tradeoffs and that certain extreme weather events could change the requirements because of significant changes in wind power. The current *non-spinning reserve service* requirement requires a startup time of offline resources of 30 minutes. It was suggested that if the startup time is 10 minutes, this may reduce the requirements of the *responsive reserve service* since the energy from the offline resources can be deployed more quickly. It was stressed that while *responsive reserve service* needs to be able to cover generation outages, it only needs to cover large drops in wind generation to the extent the *non-spinning reserve service* cannot. Lastly, it was discussed that the wind power forecast used in the day-ahead commitment should be able to guide ERCOT when determining the procurement of the *non-spinning reserve service* just as it does with the load forecast.

Ireland: All Island Grid Study (2008)

The "All Island Grid Study" in Ireland was published in 2008 and examined the Irish system's ability to integrate various penetrations of wind generation [64]. Six plant portfolios were examined to meet the load forecasted for 2020. Portfolio 1 contained 2 GW of wind; 2, 3, and 4 contained 4 GW; portfolio 5 contained 6 GW; and portfolio 6 contained 8 GW of wind generation. This is in the context of a projected peak load of 9,618 MW and a load factor of 63.9%. The study involved hourly scheduling of the system with the WILMAR system planning tool [65].

The study incorporated a refined implementation for reserve provision with only two categories specified in the model: spinning and replacement reserve. The definition of a unit capable of meeting the replacement reserve standard was an off-line unit with a start-up time of less than 60 minutes and online units whose capacity was not allocated to the spinning reserve requirement. This is a highly simplified model given the existing structure of reserve provision in the Irish system. The requirements for spinning and replacement reserve were based on a mixture of existing and proven requirements and newer techniques for the provision of reserve for wind generators.

The spinning reserve requirement is calculated as being the size of the largest on-line unit plus an additional contribution for wind generation, calculated based on the work in [66] (and described in detail later in section 4.2. Ireland is an island system with one 500-MW interconnector in operation and a 500-MW interconnector under construction, both to Great Britain. System modeling for the year 2020 assumed that 100 MW of spinning reserve can be obtained through interconnection. Another 50 MW of reserve is assumed to be provided from interruptible contract loads. Of the remainder, a constraint of a maximum of 50% of reserve demand can be provided by pumped storage. Wind generators are allowed to provide spinning reserve when curtailed.

The demand for spinning reserve is illustrated in Figure 35 on a weekly averaged basis. Spinning reserve is required more frequently as the amount of wind increases in the portfolio, significantly so in portfolio 6. The scheduled outage of the largest unit on the system (i.e., 480-MW combined cycle gas turbine) is seen to reduce the spinning reserve requirement significantly during weeks 31 to 34. While the VG requires additional spinning reserve, the largest contributing factor remains the loss of the largest conventional unit.

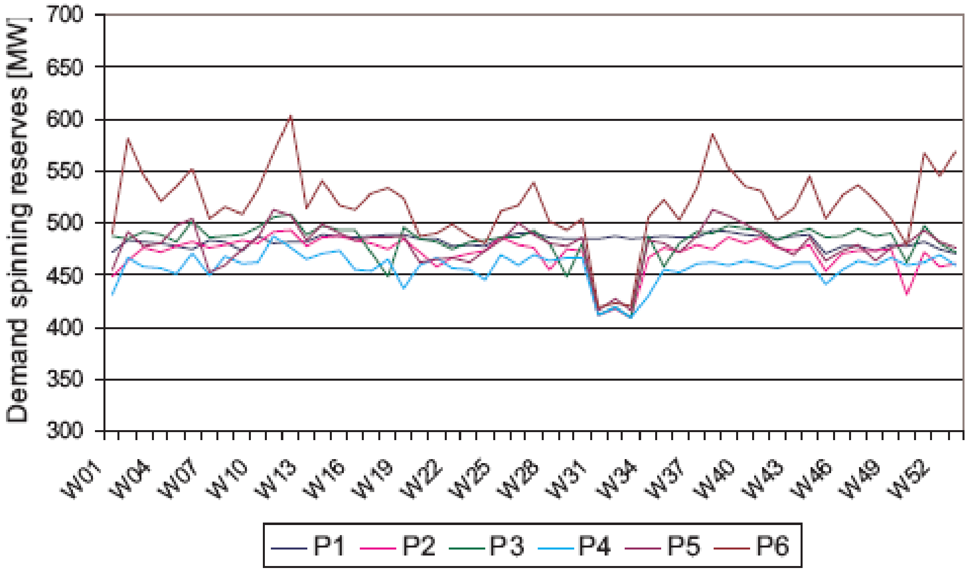

Figure 35: Weekly demand for spinning reserve for each generation portfolio.

Replacement reserve is calculated as a function of the possible forced outages of units and an additional margin which is a function of the 90th percentile of the net load forecast for each particular scenario. The 90th percentile was chosen as it most closely matches experience with proven reserve standards. WILMAR implements rolling unit commitment and has stochastic optimization functionality, which requires probabilistic forecasts and thus, replacement reserve is required for each potential probabilistic scenario. Demand for replacement reserve is a function of the installed wind power and the forecast error over longer horizons, as shown in Figure 36 and Figure 37. Figure 36 shows how the replacement reserve requirement is a function of how far ahead the optimization is evaluating. In other words, generally, errors will be larger further out and therefore more replacement reserve is required compared to more immediate horizons. In Figure 37, where portfolio 5 contains 6 GW of wind generation, the requirement for replacement reserve is seen to exceed 3 GW in one instance. This is due to a 1-GW load rise at the same time as a 1 GW decrease in wind.

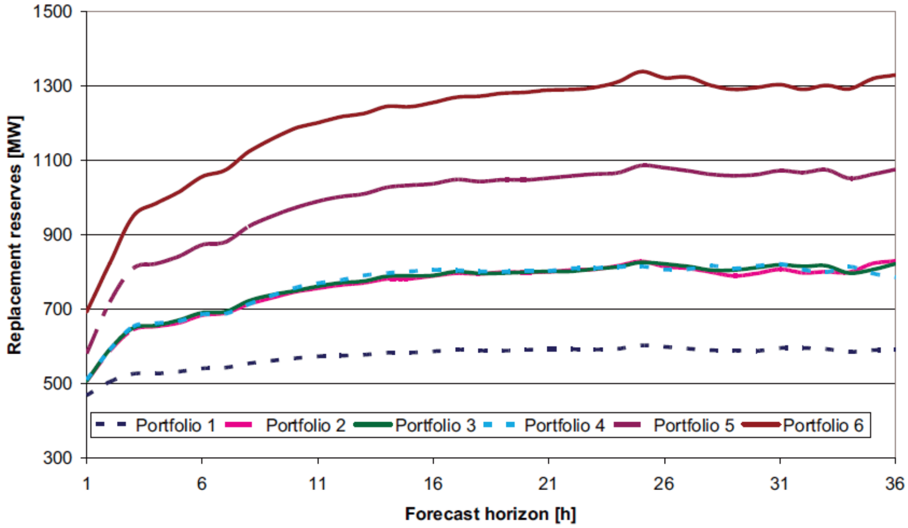

Figure 36: Average requirement for replacement reserve by time horizon.

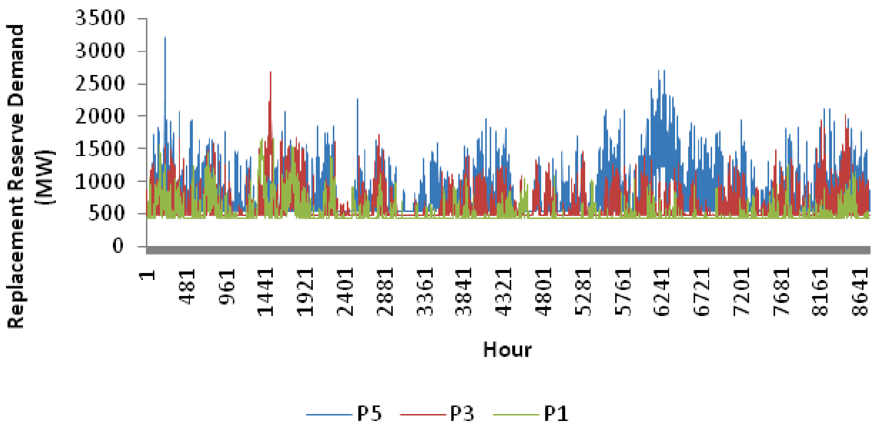

Figure 37: Hourly requirement for replacement reserve.

Eastern Wind Integration and Transmission Study (2010)

The "Eastern Wind Integration and Transmission Study" (EWITS) evaluated the operational impacts of various wind penetrations, locations, and transmission build-out options for most of the U.S. Eastern Interconnection [67]. The study included three scenarios of 20% wind energy with each representing different primary locations of the wind, and one 30% wind energy scenario. The majority of this region is currently operated by ISOs and RTOs who administer wholesale electricity markets. These markets have evolved since their inception in the late 1990s. The further evolution of the rules and procedures that the markets will follow was a key assumption on how operating reserve requirements are determined in the study, with the boundaries between balancing areas and what is extracted from sub-hourly energy markets also having an impact on the method used.

The first procedure of the study was to determine the *contingency reserves* required. As many previous U.S. studies have done, these assumed the current rule and determined that the largest contingency was not affected by the large amounts of wind generation. One and a half times the single largest hazard in each operating region determined the amount of *contingency reserves* for that region.

Many prior studies in the United States concluded a slight, but not insignificant, increase in the amount of required *regulating reserve* due to the increased variability of wind added to that of the load. In EWITS, a similar methodology to the prior studies was performed. The minute-to-minute variability separated from a 20-minute rolling average of a 100-MW wind plant was used for the analysis, and the standard deviation was determined to be 1 MW. It was assumed that there is no correlation between wind plants for power output deltas in this time frame, and therefore, the total standard deviation for a balancing area was calculated by geometrically adding the 1-MW standard deviation for all 100-MW wind plants on the system. For load-only, the *regulating reserve* requirement was assumed to be 1% of the total load, and assumed to be equal to three times the standard deviation of the load variability. Since load and all wind variability on this timeframe were also considered to be independent of one another, the standard deviations of all wind and load were then geometrically added together by calculating the square root of the sum of their squares. The total standard deviation was increased by less than 1 MW when the wind was added to the load, and therefore the variability of wind was not considered as part of the *regulating reserve* for the study.[8]

Unlike other studies, it was determined that the uncertainty in the wind forecasts used for economic dispatch would impact the *regulating reserve* much more than what was shown for the variability. Economic dispatch programs that run every 5 minutes would use information from at least 10 minutes before the operating interval. Since it is too late to adjust the economic dispatch for any deviations, these deviations would all be met by units providing *regulating reserve*. Assuming a 10-minute-ahead persistence forecast, the additional *regulating reserve* was determined by looking at the standard deviation of 10-minute changes in wind output (load forecast for 10-minute ahead was assumed to be quite good and load forecast error was ignored). Figure 38 shows the standard deviation of the 10-minute-ahead wind forecast errors as a function of the average hourly production of the total wind. The highest variability is near 50% production, where the anticipated 10-minute change can be up or down, and also relates to wind turbines being at the steepest part of the wind speed to power conversion curve. The function was used for the hourly wind-related standard deviation of the *regulating reserve* requirement and was geometrically added to the load *regulating reserve* requirement discussed above. The equation is shown below, where σ_{st}(HourlyWind) is the standard deviation of wind forecast errors that is a function of the predicted wind portrayed in Figure 38.

$$Reg\ Req\ =\ 3 * \sqrt{\left(\frac{1\% Hourly\ Load}{3}\right)^2 + \sigma_{ST}(HourlyWind)^2}$$

[8] Calculations based on a balancing area with 100-GW load and 60 GW of wind power, which was about the average for the largest ISO balancing areas that were a part of the study.

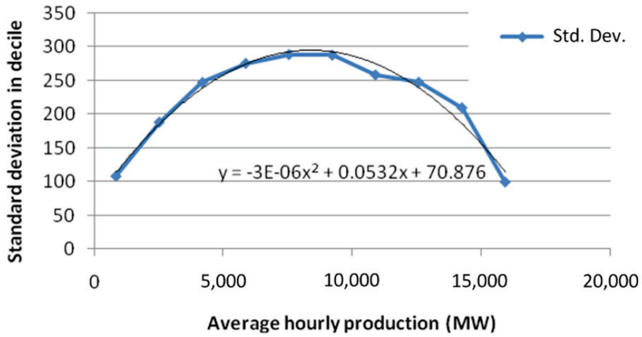

Figure 38: 10-minute-ahead wind generation forecast errors as function of wind production.

A similar approach was used for the hour-ahead wind forecast error. However, in this case it was assumed that the errors that were not occurring often could be compensated for with off-line reserve. Therefore, one standard deviation of the hour-ahead forecast error was required to be spinning, and two standard deviations could be non-spinning. Also, since the reserves were used in the production cost simulations for the study, it was ensured that if the reserves had to be used for the hour-ahead forecast error of the hour in question, those reserves did not have to be kept in real-time. In other words, if reserves were needed because less wind was available than forecast, the model would release that amount of reserves in real-time since the reserves were used for the forecast error and not needed further. The total amounts of all reserves used in the study are shown in Figure 39. The reserve requirement was an hourly value that was a function of wind levels.

Reserve Component	**Spinning** (MW)	**Nonspinning** (MW)
Regulation (variability and short-term wind forecast error)	$3 \cdot \sqrt{\left(\frac{1\% \cdot HourlyLoad}{3}\right)^2 + \sigma_{ST}(HourlyWind)^2}$	0
Regulation (next-hour wind forecast error)	$1 \cdot \sigma_{NextHourError}\ (PreviousHourWind)$	0
Additional Reserve		2 × (Regulation for next hour wind forecast error)
Contingency	50% of 1.5 × SLH (or designated fraction)	50% of 1.5 × SLH (or designated fraction)
Total (used in production simulations)	Sum of above	Sum of above

Figure 39: Summary of reserve methodologies for EWITS.

Southwest Power Pool: Wind Integration Task Force Integration Study (2010)

In the Southwest Power Pool (SPP) "Wind Integration Task Force (WITF) Integration Study," operational requirements for 10%, 20%, and 40% wind energy were studied for the SPP footprint [68]. The study assumed that SPP was operating as a single balancing authority with a co-optimized energy and ancillary service market (i.e., Day-2 Market). In this study, the reserve determination methodologies were unique once again. The study evaluated reserve requirement needs for *regulation reserves* (Regulating Reserve), *load following reserves* (Following Reserve), and *contingency reserves*. For *regulation-up* and *regulation-down* reserves, the team explicitly used the NERC CPS2 standard to determine the increases. Therefore, to equate to the 90% compliance requirement of CPS2, the 5th and 95th percentiles were used as the bounds to the requirement. Also, because the ACE is out of compliance only if the 10-minute average is above the balancing area L_{10} (see section 3.1.1), this value is used in the overall equation of determining the requirement. The equations are shown below:

$$R_{up} = \sqrt{(0.01 l_{peak} + L_{10})^2 + a\Delta {W_{95}}^2} - L_{10}$$

$$R_{down} = \sqrt{(0.01 l_{peak} + L_{10})^2 + a\Delta {W_5}^2} - L_{10}$$

Where l_{peak} = peak load; a is a calculated coefficient, and ΔW are the respective percentile of 10-minute deltas. Figure 40 shows the *regulation reserve* requirements for peak days. Like in the other two studies, the project team proposes that *regulation reserve* requirements be dynamic, and are time varying based on system conditions. The study also recommends the possibility of a *load following reserve* requirement and provides analysis on some of the ramping requirement increases in this time frame. Lastly, the *contingency reserve* requirement was not changed for the study, but it was recommended that it be reevaluated if extensive high-voltage transmission expansion occurs in order to transfer high penetrations of wind power from remote locations.

Season	Peak Load	Case	Wind Short-Term Variability			Regulation Requirement	
			ΔW_{95}	ΔW_5	Std Dev	Up	Down
Winter	33,237	Base	51	49	31	338	338
		10%	114	107	67	357	360
		20%	216	206	129	417	425
		40%	381	362	228	562	581
Spring	37,169	Base	62	59	37	379	379
		10%	130	122	78	401	405
		20%	237	225	141	465	474
		40%	407	390	242	617	635
Summer	45,822	Base	68	64	41	465	466
		10%	140	131	84	487	491
		20%	251	235	148	546	557
		40%	422	397	249	684	710
Fall	36,621	Base	52	50	31	371	372
		10%	116	110	68	390	393
		20%	218	208	129	448	455
		40%	383	363	228	586	606

Figure 40: Total regulation requirements for seasonal peak loads (MW).

4.1.1.1 Western Wind and Solar Integration Study (2010)

The "Western Wind and Solar Integration Study" (WWSIS) was conducted for a set of the Westconnect utilities as shown in Figure 41 [69]. Although it focused on this region, the entirety of the U.S. portion of the Western Interconnection was modeled with wind, load, generation, and transmission data. The study analyzed the integration impacts of up to 30% wind energy and 5% solar energy in the study region, with 20% wind energy and 1% solar energy in the rest of the Interconnection. A large part of the study analyzed the impacts of operating reserves that the increased penetrations of wind and solar would have on the system.

Figure 41: WWSIS WestConnect Study Footprint.

As discussed in section 3.2.3, the WECC region is proposing to require its balancing areas and reserve sharing groups hold 3% of its generation plus 3% of its load as *contingency reserve*, with half of this as online *spinning reserve*. The study used this rule and an analysis indicating that the standard deviation of 10-minute load variability is about 1% of the load, to make the assumption that 3 times the standard deviation of the 10-minute variability would be held for the *variability reserve* (Following Reserve). The study recommended that one standard deviation of wind variability be held for *regulation reserve* (Regulating Reserve) and have AGC capability. The report also discussed the need for a dynamic requirement that depends on both the wind and load levels. The team then analyzed the standard deviation for combinations of wind and load levels with each represented by 10 bins. An example of the standard deviation of the 10-minute net load delta for the study footprint is shown in Figure 42.

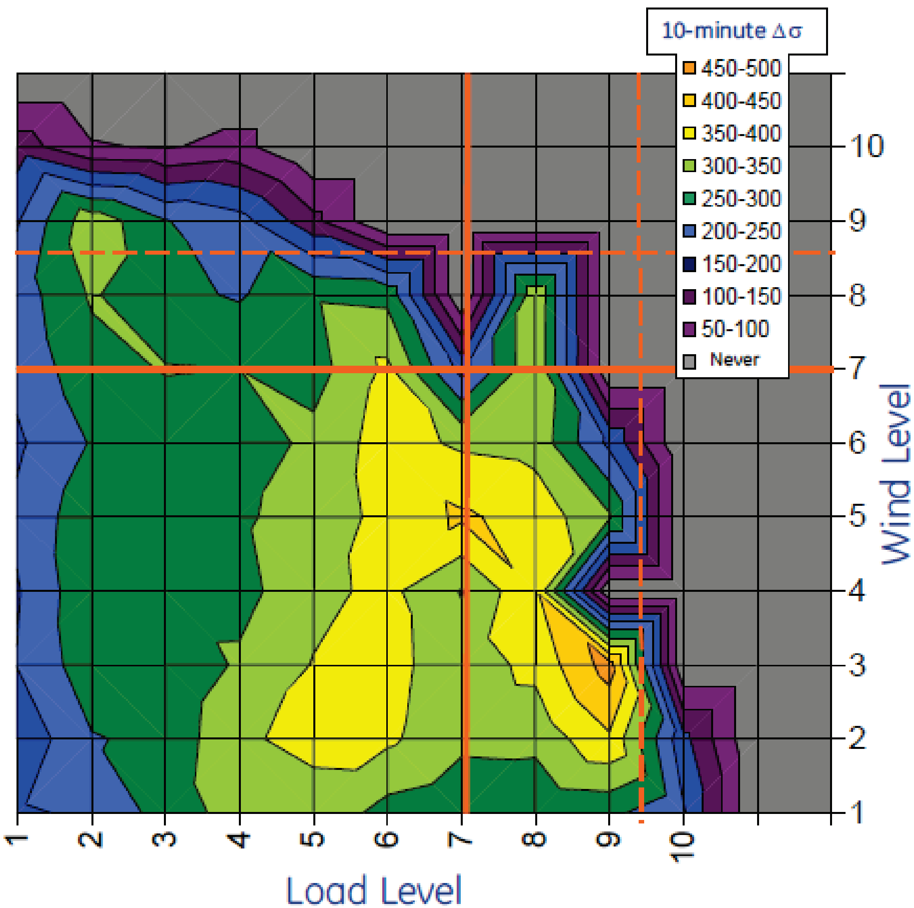

Figure 42: standard deviation of 10-minute delta for different wind and load levels.

The figure shows that the requirements would change based on the predicted wind and load levels. For example, low load levels generally have little changes, and therefore, would need less variability reserve. High variability can be seen around the load level 7 and wind level 5, as well as load level 9 and wind level 3. At times where these levels were predicted, the system operator may require additional reserves to accommodate this variability. Figure 43 shows the requirements based on the 3σ rule for all the individual regions in the study footprint. The rule can be compared with using the 3% load, and then with the 3σ for load variability alone. In areas with high wind and relatively low load (i.e., Wyoming), some significant changes in the *variability reserves* can be seen. Figure 44 also shows the same methodology using one standard deviation applied to the *regulation reserve* need.

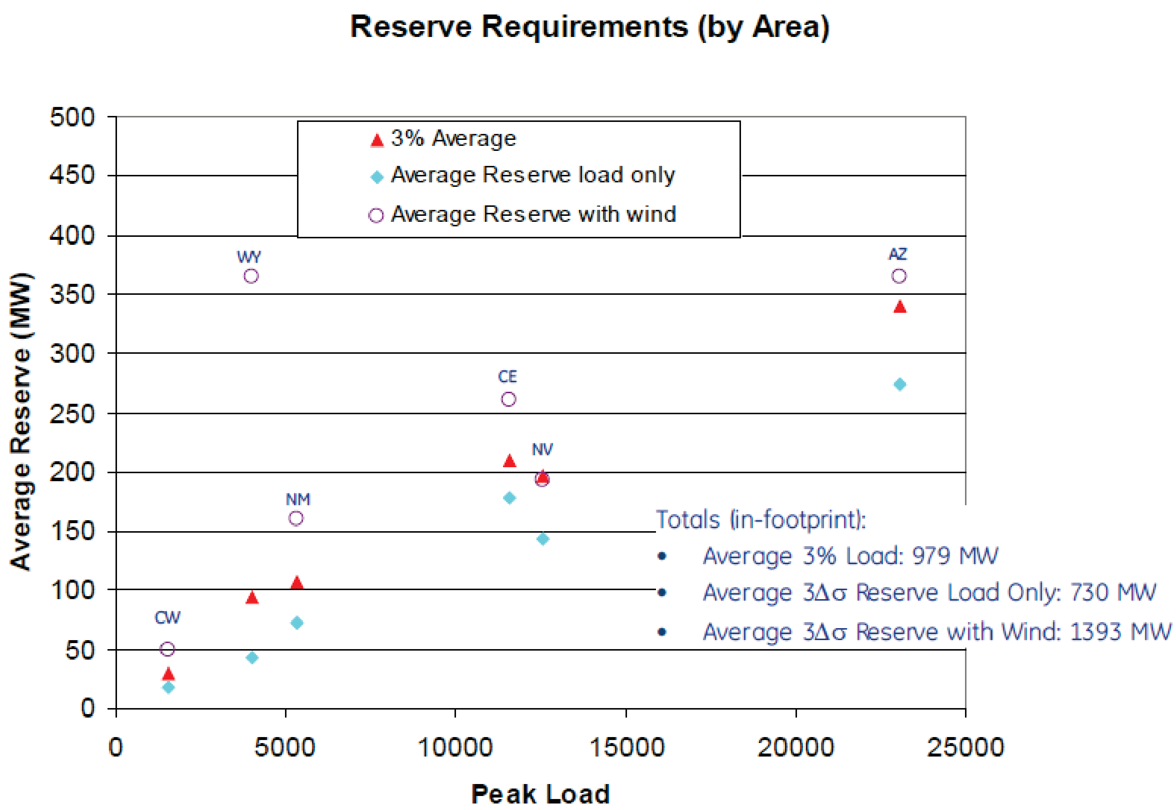

Figure 43: Reserve requirements using 3% or using 3 sigma rule for each area.

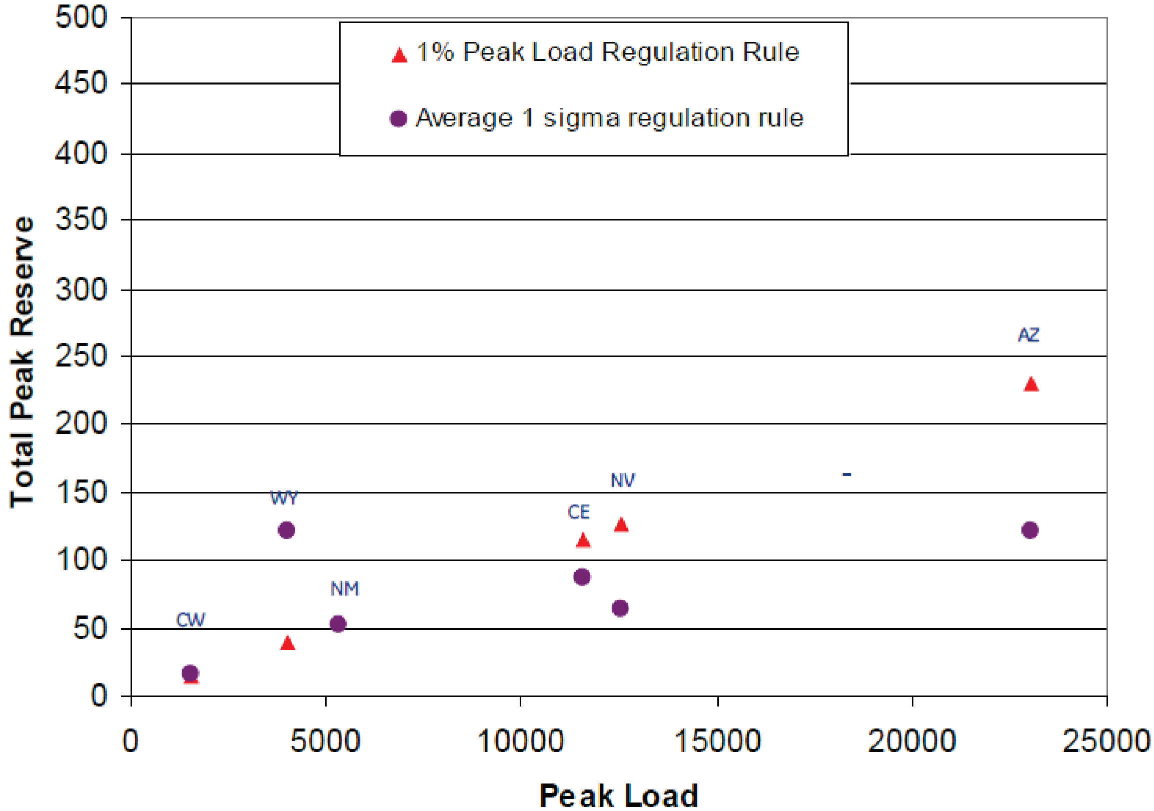

Figure 44: One % peak load *regulation* rule and average 1 sigma *regulation* rule for all regions.

This method was then analyzed further so that simple rules that didn't need detailed look-up tables could be used by system operators when determining the needs for these *variability reserves*. At first a rule of 3% load plus 5% wind was analyzed with satisfactory results, which can be seen in Figure 45. Using this rule it was found that the quantity of reserves held was always greater than or equal to the 3σ rule. However, it was also seen to be over-carrying the reserve in many of the wind and load level scenarios. Furthermore, in individual regions like Wyoming, this simple rule actually led to many instances of being short of the 3σ requirement.

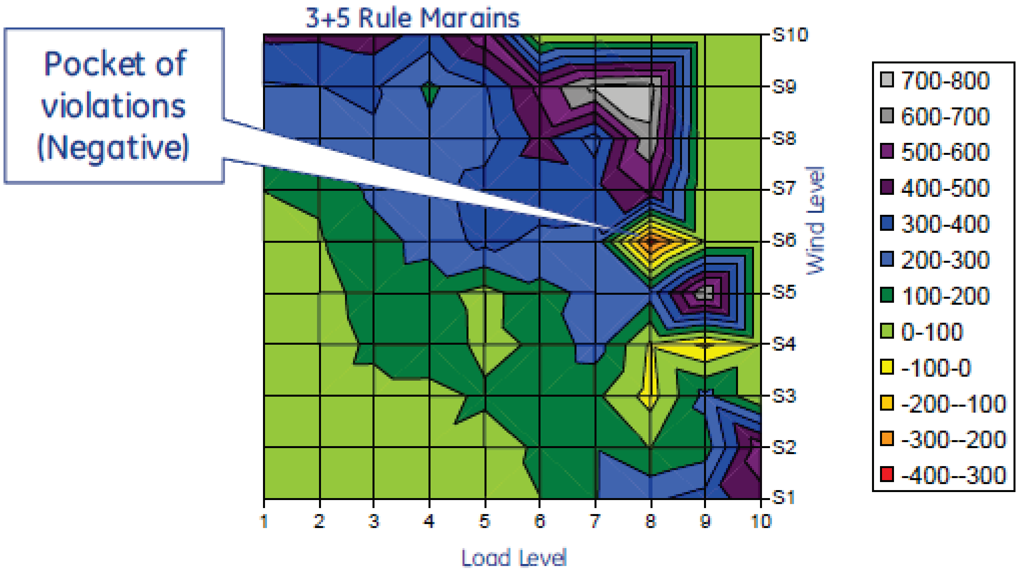

Figure 45: Three + 5 rule for operating reserve requirements for different wind and load levels.

To further improve on the use of simple rules to carry the needed *variability reserve*, the team then used a best fit using three degrees of freedom. These included the coefficient to load level, the coefficient to wind level, and an amount of wind capacity where further incremental increases in reserve were not needed. In other words, X*Load + Y*Wind, until Wind > Z(nameplate capacity). It was determined that this rule was a good compromise between accuracy of capturing the 3σ need for variability reserve and simplicity. Table VIII shows the requirements for the study footprint and for each region inside the study footprint.

Table VIII: WWSIS reserve rules for 30% local priority scenario

	Load Only	30% LP Scenario		
	(% of load)	Load Term (% of load)	Wind Term (% of wind production)	up to (% of wind nameplate)
Footprint	1.3	1.1	5	47
Arizona	2.2	2.2	5.6	36
Nevada	2.1	1	10.7	54
Colorado East	2.4	2	5.7	68
New Mexico	2	3.1	3.5	70
Wyoming	1.3	2.7	8.7	33
Colorado West	1.8	3.1	7.3	100

4.2 Academic Research

In addition to the integration studies that are being performed by utilities, transmission system operators, and other organizations representing future penetrations on existing power systems, there has been substantial work from the academic institutions in the area of power system operations with high penetrations of VG and in particular, determining operating reserve requirements. The work has been done in parallel with many of the studies and many of the methods have shared some common traits. It is important to note that researchers have been investigating optimal operating reserve requirements for some time before the introduction of large penetrations of VG and how VG could change these requirement methods. As discussed earlier, most of the operating reserve requirements have been based on rules of thumb that compute the requirements based on the largest units or a percentage of the load or generation. Even if the only uncertainty is the failure of conventional units, the deterministic criteria used today do not take into account the failure rates of the generating units, the value of not shedding load, or the possibility of two or more near-simultaneous outages.

In [70], the authors proposed a probabilistic method where the unit commitment problem would determine the optimal amount of *spinning reserve* (Contingency Reserve and Following Reserve) for different load levels based on the probability of having insufficient generation due to generator outages or load forecast errors. The method evaluated the load probability distribution and evaluated the risk involved with the current unit commitment using a capacity outage probability table by the convolution of the cumulative probability distributions. This shows the amount of risk by looking at the probability that the generation online is less than the load. The process would iterate until the risk level was met for the time period. The paper discussed briefly the idea that the risk index may be based on arbitrary concepts, and introduced the concept of determining the appropriate risk value based on the cost of load shedding.

In [71] and [72], this idea was further researched and methods were discussed that used the reliability of units, probability of load forecast errors, and different values of lost load to determine the appropriate reserve requirement (mostly Contingency Reserve). In [71], the authors evaluated this in a market setting with two different types of pricing mechanisms. In [72], they performed the calculation offline from the formal unit commitment solution by minimizing both the cost of reserve and the cost of load shedding ignoring inter-temporal constraints. These methods showed promise on how operating reserve requirements could be determined based on the reliability of units, the possibility of multiple outages, and the social cost of load shedding, which in theory, all showed benefits over the current rule of thumb procedures.

Many researchers have been realizing that the uncertainty and variability impacts that come from significant amounts of VG are different from that of conventional generation outages, and that the methodologies in place may change significantly in order to determine the optimal requirements to maintain a reliable, yet cost-effective system.

Dany 2001: In analysis performed by [73], the impacts on *primary control reserve* (Primary Reserve), *secondary reserve* (Regulating Reserve and Following Reserve), and *long term reserve* (a planning reserve and not Operating Reserve) were evaluated with high penetrations of wind power on four test power systems representative of the German system. Because of the small variations in the time frame of *primary control reserve* when compared to the variations of load and outages of large conventional units, there was determined to be negligible impacts to the requirements of *primary control reserve*. For *secondary reserve*, there was a separation into two components: *secondary control reserve* (Regulating Reserve) and *manual secondary reserve* (Following Reserve). The *secondary control reserve* is essentially utilized for the variations in the net load, while the *manual secondary reserve* is utilized for the increased forecast error and is deployed through an optimized dispatch (fluctuations greater than 15 minutes). The assessments of these reserve increases are based on calculating the standard deviation of the ACE with and without the wind power. The paper also emphasizes the need for negative (downward) control reserve since the minimum generation points of conventional generating units becomes binding more often at high wind penetration levels. Two programs were developed to calculate these requirements. The first models the dynamics of the system along with the adjusted reference points (AGC set points) to determine a distribution of ACE. The second models the dynamics of the system in a simplified manner so that outages of conventional units are also modeled. The first model evaluates the change in ACE while the second model will determine the probability that there is a deficit of *manual secondary reserve*, $Pr_{D,MR}$, and the probability that there is a surplus, $Pr_{S,MR}$, meaning that there is not enough *negative manual secondary reserve* due to the minimum generation limits of conventional units. These probability values are mostly due to the outages of units and forecast errors of load and wind power.

Since criteria of sufficient *secondary reserve* is not as concrete as other reliability metrics (e.g., loss of load probability), the analysis simply compares the increases in σ(ACE), $Pr_{D,MR}$, and $Pr_{S,MR}$ compared to the same systems without wind power. The wind penetrations analyzed were up to 100% wind capacity by the peak load of different systems, with different wind locations and different conventional generating sets. The analysis of *secondary control reserve* evaluated the change in the standard deviation of ACE with increasing wind capacities compared to a system with no wind. Most of the systems showed about a 35% increase in the standard deviation of ACE at a total wind capacity equal to the peak load. A direct conclusion from this analysis is that the increase in the *secondary control reserve* capacity does not reduce the area control error as much as the increase of the reserve power gradient (ramp rate of the reserve). It was concluded that the systems needed to increase the power gradient of the *secondary control reserve* three times as much in the 100% of peak load scenario to get same ACE as the system without wind.

In the second simulation, the required *manual secondary reserve* was evaluated by assessing the increase in both $Pr_{D,MR}$ and $Pr_{S,MR}$ with the increasing wind penetrations. Since the main attribute to these values increasing was wind power forecast errors, which at the time of this writing there was little experience in, the simulations were performed for different errors with standard deviations ranging from 10% to 20% of installed wind capacity. The increase of the need for *manual secondary reserve* did not change significantly until at a level of wind at about 20% capacity of peak load. From 20% to 40% of peak load, the increase in this reserve was very near linear at an increase of about 300 MW per GW of installed wind (10% forecast error) to 600 MW per GW of installed wind (20% forecast error) with the system of a peak load of about 6 GW. For the larger system of 20-GW peak load, this linear increase happens at a wind penetration of 10% to 20% of peak load. This increase grows more exponentially at larger penetrations. For the 6-GW system, the required reserve is between two and three times the amount without wind at a penetration of 100% of the peak load. For the 20-GW system, this is between three and six times. Although the paper did not show examples, the author discussed that the requirements for the negative *manual secondary reserve* to be very similar to that which came out of the analysis for the positive reserve.

Doherty 2005: The analysis of [66] evaluated the need for *system reserves* (Contingency Reserve and Following Reserve) due to wind and load forecast errors either by themselves or in addition to contingency (loss of supply) events. A model evaluated the reliability of a system based on the full outage probability (i.e., the probability of an outage occurring in any hour, which is different than forced outage rate, the probability that the unit is out during the hour), and wind and load forecast errors. Its risk level is determined by the load shedding incidents (LSI) tolerated per year. For example, if one LSI occurred during the year and lasted 24 hours, this would equate to an LSI of one and a loss of load expectation of 24 hours. The load forecast errors were combined with the wind forecast errors as independent and uncorrelated normally distributed Gaussian values.

The probability of load shedding during any hour would be equal to the probability of load shedding during normal instances when the system is holding its full reserve requirement, plus the probability of load shedding following a time period Hr after a disturbance event, when the system is using reserve to respond to the initial disturbance, and is therefore more vulnerable to load shedding because of the lesser reserve amount. A load shedding incident here is defined as any incident in which the supply deficiency is more than the quantity of reserve. Both full outages and partial outages of conventional units are discussed. The event of three conventional unit outages within the reserve replacement time period is ignored due to its very small probability. Therefore, the probability of load shedding occurring is seen in the following equation:

$$PLS_h = PLSNO_h + \frac{1}{2}(Hr) * [FOP_{1,h} \quad FOP_{2,h} \quad \dots \quad FOP_{G,h}] \times \begin{bmatrix} PLSFO_{1,h} - PLSNO_h \\ PLSFO_{2,h} - PLSNO_h \\ \vdots \\ PLSFO_{G,h} - PLSNO_h \end{bmatrix} + \frac{1}{2}(Hr) * [POP_{1,h} \quad POP_{2,h} \quad \dots \quad POP_{G,h}] \times \begin{bmatrix} PLSPO_{1,h} - PLSNO_h \\ PLSPO_{2,h} - PLSNO_h \\ \vdots \\ PLSPO_{G,h} - PLSNO_h \end{bmatrix}$$

Where PLS is the probability of load shedding, PLSNO is the probability of load shedding during the normal case (no prior generator trip), Hr is the reserve replacement time in hour, FOP is the full outage probability, PLSFO is the probability of load shedding following a full outage of a generator, POP is the partial outage probability, and PLSPO is the probability of load shedding following a partial outage of a generator. G is the set of generators. The ½ coefficient of Hr represents the integral of the linear response of the replacing reserve (Tertiary Reserve) that replaces the reserve used for the generator outage.

The PLSNO, PLSFO, and PLSPO are determined by calculating the *system reserve*, or remaining reserve, due to full and partial outages, being below the positive net demand forecast error (i.e., the actual net demand was higher than the forecast net demand by an amount greater than the reserve or remaining reserve). By setting an LSI target, the amount of *system reserve* can be quantified by analysis of these steps. A case study using this methodology was performed for the Ireland electricity system with an installed capacity of 7500 MW and wind penetrations up to 2000 MW. Full outage probabilities of conventional units in the system ranged from 0.003 to 0.006, the standard deviation of load forecast errors was taken to be 75 MW, and standard deviation of wind forecast errors was taken to be 15% of the installed capacity for a 3-hour forecast horizon. With an LSI of 3 incidents per year, the reserve was found to be about 625 MW compared to 475 MW without wind.

Other important conclusions came from this work. It discussed the impact that forecast horizon had on wind forecast error. The standard deviation of wind forecast errors increased with forecast horizon and therefore, the quantity of *system reserve* would increase for time horizons further ahead than for those closer to the operating hour. It then evaluated the different reserve types based on the time frame of response that they are expected to operate at. In each time frame, a different standard deviation of the net demand forecast error drives the requirement. At very short time frames, this standard deviation is very close to that of the standard deviation of variations of the net demand, since persistence forecasts are considered to be the most accurate type of forecast in these periods. Figure 46 shows results from the study with an LSI of 3 per year and a reserve replacement time of 2 hours for each of the reserve categories.

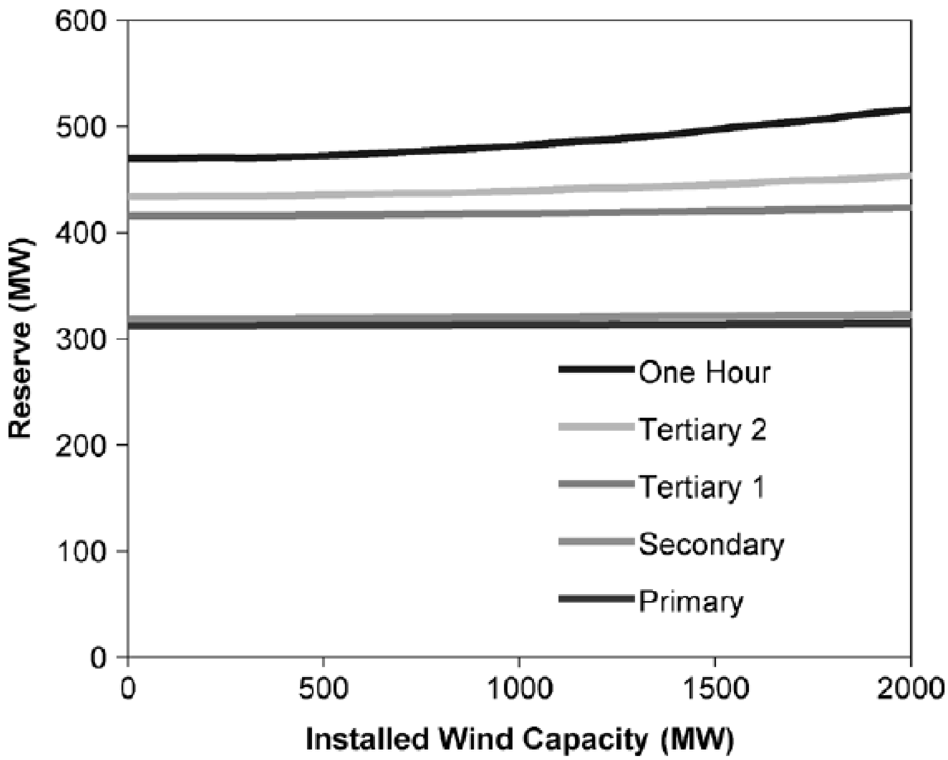

Figure 46: Reserve requirement of different reserve types based on their forecast horizon.

Ortega-Vazquez 2009: In [74], a novel approach was used in determining the optimal amount of spinning reserve for penetrations of wind power on a system. The methodology is based on [72] mentioned above that was not evaluating impacts of wind generation, but more on conventional generation outages and load forecast errors. The method determines the *spinning reserve* (Contingency Reserve and Following Reserve) requirement in an offline study that equals the marginal cost of supplying the *spinning reserve* with that of the socioeconomic cost associated with load shedding. The authors assume that a large geographically distributed set of wind plants will lead to a wind forecast error that is normally distributed with a standard deviation calculated from the following:

$$\sigma_w^t = \frac{1}{5} w_F^t + \frac{1}{50} W_I$$

Where w_F is the wind forecast and W_I is the installed capacity of the wind. This error is combined with that of the load, and uncertainties of conventional generation outages determined from a capacity outage probability table. The uncertainties along with the *spinning reserve* requirement can determine the expected amount of load shedding, and with a VOLL, an expected socioeconomic cost associated with load shedding. The optimal *spinning reserve* required is determined for different net load levels by giving a requirement that has a marginal cost equal to that of the additional cost reduction of load shedding it ensures. The method was tested on the IEEE Reliability Test System with an installed capacity of 3105 MW, installed wind capacity of 620 MW, and VOLL equal to $1000/MWh. The result was compared with ones with the methods of using the largest unit as reserve, using 3.5 times the standard deviation of net demand forecast, and the sum of both. The 3.5σ method performs very similarly to the optimized *spinning reserve* method at low to moderate wind levels. Using the largest unit as reserve and using the sum of the largest unit and 3.5σ increased the total costs by about 1.44% and 2.66%, respectively.

Morales 2009, Yong 2009, Bouffard 2008, and Wang 2008: In Morales et al [75], Yong et al [76], Bouffard et al [77], and Wang et al [78] the *operating reserve* (Contingency Reserve and Following Reserve) requirement was inherently managed in the unit commitment or economic dispatch problem. Each of these formulated a two-stage stochastic unit commitment problem, where multiple scenarios of wind power were modeled so that the unit commitment solution was able to meet each of these scenarios reliably and in some cases cost effectively. More information on stochastic unit commitment scheduling can be found in [79]. In [78] the second-stage wind outcomes were not optimized, but the model ensured that they could be met reliably with respect to load balance and transmission limits. The other models would calculate probability distribution functions for wind and load, including a discretized normally distributed wind and load [77], a truncated normally distributed wind and load [76], and distribution taken from observed values with applied scenario reduction techniques [75].

The methods of [75]-[77] modeled the second stage scenarios and associated probabilities as part of the objective function, minimizing the total expected costs. In [76], generation contingencies were additional stochastic inputs along with the wind and load, in [77] the wind and load were combined for a stochastic net load input, and in [75], wind was the only stochastic variable. In [76] and [75], the transmission network was also included so that different scenarios wouldn't violate transmission constraints. In [76], it also included AC constraints in the network modeling so that voltage and MVA flows were considered. In [76] and [75], both *spinning* and *non-spinning reserve* were modeled based on the contingencies that can occur and the possible wind scenarios. In [76] this stochastic security-constrained optimal power flow was compared with a deterministic optimal power flow with exogenous reserve requirements to evaluate the difference in the reserve quantities. In [75], tradeoffs were made between spinning and non spinning reserves by scheduling wind power at higher and less probable outputs in the 1st stage market, to allow for cheaper *non spinning reserve* compared to *spinning*. All of these methods considered an hourly resolution model for 24-hour, 1-hour, and 4-hour study horizons for [75], [76], and [77], respectively. A summary of the methods and results can be seen in Table IX.

Table IX: Comparison of different stochastic methods

	Morales et al [75]	Yong et al [76]	Bouffard et al [77]
Network Model	DC	AC	None
Stochastic Variables	Wind	Wind, load, and generator contingencies	Wind and load
Study Period	24 hour	1 hour	4 hour
Test System	73-bus IEEE Reliability Test System, 118-branch, aggregated generator groupings, peak load of 2850 MW	179-bus, 263-branch, 29-generator, 60785 MW load,	1-bus, 0-branch, 3-generator, Load (H1:30 MW, H2: 80 MW, H3: 110 MW, H4: 40 MW)
Wind Power	150-MW wind plant at single bus	600 MW, 1800 MW, and 3000 MW at 1 bus	0-20% wind power penetration
Scenario conditions	Wind scenarios from actual measurements, scenario tree was reduced to 20 scenarios	1. Wind forecast scenarios: (probability of occurrence, %deviation from output) 0.1, -12; 0.2, -6; 0.4, 0; 0.2, 6; 0.1, 12 2. Generator outages: 8 units with probability of outages ranging from 0.01 to 0.015 3. Load forecast scenarios: unclear	1. Wind forecast scenarios: $\sigma_{wt} = 0.02 + 0.2\hat{w}_t$, where σ_{wt} is the standard deviation, and $\hat{w}_t$ is the wind forecast. 2. Load forecast scenarios: Standard deviation = 2% of hourly demand Both are combined to a scenario tree.
Results	Total cost depending on which bus the wind was sited on: Energy from \$408,704 to \$411,917 Reserve from \$13,631 to \$15,049 Cost per MWh depending on bus: Energy from \$5.12/MWh to \$5.16/MWh Reserve from \$3.79/MWH to \$4.18/MWh	600-MW wind pen.: Spin = 26.93 MW Non-spin = 7.01 MW 1800-MW wind pen.: Spin = 26.05 MW Non-spin = 8.69 MW 3000-MW wind pen.: Spin = 27.1 MW Non-spin = 8.11 MW	Cost of reserves at 0%: \$175 Cost of reserves at 20%: \$250

Ruiz 2009: In [80], the authors present a similar stochastic framework for modeling uncertain wind power scenarios. The methodology used was proposed in [81]. They include uncertain wind and generation outages as the stochastic input and model hourly resolution for a 24-hour period. In this case, the unit commitment decision for all but fast start units must be fixed in the first stage. The key contribution from this work was the notion of how additional deterministic *operating reserves* (Contingency Reserve and Following Reserve) are required when operating a stochastic unit commitment program. The research compared a stochastic case where reserve was required in each of the stochastic scenarios and one where the reserve was only required in the base (median) scenario. They analyzed how different deterministic reserve requirements supplemented with the stochastic framework could increase efficiency of system operations. A case study was performed on the Public Service of Colorado system. The stochastic case with reserve only required in the base scenario had a 0.23% cost reduction when compared to the traditional deterministic case, whereas the case with reserve requirement in each scenario had a 0.18% cost reduction. Even though these savings are very low, they were positive for all 10 simulated periods. When the same systems were tested without a flexible pumped hydro storage plant, the percentage savings of the stochastic scenario reserving the base scenario doubled. The work emphasized the notion that because of computational difficulties of stochastic programs, not all uncertainty can be captured within the scenarios, and deterministic procedures must complement the stochastic programs to manage the uncertainty.

Matos 2011: In [82], a different approach of determining the *operating reserve* (Contingency Reserve, Regulating Reserve, and Following Reserve) was researched. The authors determined probability mass functions for conventional generator outages and discretized probability functions for wind power scenarios with a discretized probability distribution function for load to create a distribution function of the random variable representing the system generation margin. This describes the probability of having more or less generation than load. Negative values would represent occurrences where the generation was short and load shedding would be required. The area under the curve that is left of zero would equal the total loss of load probability (LOLP). By adding reserve to the system, the entire curve would essentially be shifted to the right by the amount of the reserve. Therefore, adding 700 MW of reserve would give a loss of load probability that is equal to the probability of the original system generation margin being less than -700 MW. Finally, the authors describe that the amount of required reserve can be obtained by using these probability distribution figures and using decision making processes. The decision making process discussed included the following: a maximum threshold of risk (e.g., maximum EENS), an equivalent reliability cost (using VOLL or something similar), and a value function approach (similar to a reliability cost, but more of a non-linear cost function).

The method and tool used for the analysis was called the reserve management tool. A case study was performed for the Portuguese power system with a total installed capacity of 10,395.8 MW and 2,742 MW of installed wind capacity. This method was compared to (A), the UCTE method described in section 3.2.1 for secondary reserves, plus tertiary reserves greater than the largest unit online (B), a method using $6*\sqrt{(Load)}$ during fast load variation and $3*\sqrt{(Load)}$ otherwise for secondary and the largest unit plus 2% of the forecast hourly load for tertiary, and (C) a method based on taking the square root of the sum of variance of wind, load, and conventional generation and multiplying by a constant of 3. These three methods compared with the reserve management tool showed different results of reserve requirements throughout the day. Rule A and Rule B were insensitive to risk and to the wind power levels. A Monte Carlo simulation was performed for one of the hours. Rule A showed incredible risk and rule B showed excessive reserve levels. Rule A had 19.21% loss of load occurrences, rule B had 1.07% loss of load occurrences, and the reserve management tool had 1.37% loss of load occurrences. The key being that the reserve management tool keeps the risk level that it was defined with, whereas rule A and rule B do not have that capability. Rule C has a risk threshold given to it, but the assumption of a Gaussian distribution of errors causes the actual risk to be different. The work in this paper emphasized that different decision makers have different perspectives on how much risk they are willing to take and how much they value reliability, and allowed for their methodology to be flexible for different decision making policies.

4.3 Summary

The industry and academic perspectives of studying the optimal operating reserve determination problem have provided significant advancement from the way that the current standards have been developed. It is generally agreed that increased VG will create an increase in the total amount of operating reserves held in power systems. The main questions are how much and of what type. Most of the studies broke out the reserve categories in the ways that they are broken out today in the NERC, UCTE, or other jurisdictions. With these in mind, all of the studies have determined that no additional Contingency Reserves will be needed, as total wind power loss from the wind resource in the time frame of contingencies is not a possible scenario. There has been much debate on how wind power affects the normal condition reserves, like Regulating Reserve and Following Reserve. Some studies suggested that the increase in Regulating Reserve would be increased significantly, whereas others would say that the required response to variability and uncertainty is much slower and should be added to the Following Reserve. Some further studies try not to segregate the reserve categories and determine a total Operating Reserve that can be used for any imbalance on the system. Each study seems to learn from those in the past, and these debates can be resolved in the future with improved models and more detailed data.

Many of the initial wind power integration studies basically used the methods that were developed in industry, while incrementally adding the impacts of wind power to the calculations. More recent studies realized that entirely new methods had to be developed to better capture the changing system needs while still keeping similar reliability metrics. The probability distribution of VG unavailability is much different than that of conventional generation's unavailability. Also, the way that systems manage the scheduling of the power system will have a significant impact on how operating reserves should be held. This includes the time delays and ramping requirements of actual dispatch intervals that are apparent today, as well as more futuristic scheduling strategies like stochastic techniques. A key conclusion of many of the more recent studies in both industry and academia is that these operating reserve requirements need to be dynamic. Rather than keeping the same requirements all times of the day and all days of the year, system operators should be using the information available to them to understand when the system is at low or high risk, and then schedule operating reserves accordingly. Some of the more recent research discussed the need for dynamic requirements not only based on time, but time horizon as well. As uncertainty is not only a function of time, but time ahead.

Although there are many similarities between the many methods described in this section, there are also many differences. Many of the methods reference those from past studies and refine these methods with more innovative techniques. Some focus on different VG impacts more than others (e.g., variability, uncertainty, ramp events). Others may be limited by what kinds of data are available. High time-resolution VG data is very difficult to obtain, especially for researchers. Furthermore, in order to study higher penetrations, those that are not yet in existence, much of the data must be modeled and how realistic the modeled data is may not be easily validated. It makes it difficult for system operators to adopt these study methodologies and know which are the most optimal for their systems. Further research is needed to connect these methods together and emphasize which may be more advantageous for specific system characteristics.

5 Conclusions and Discussion

This report summarizes the many different policies, standards, rules, and analytical methods for determining the operating reserve requirements for various systems. The requirements include the capacity amounts, speed at which they should respond, how they should respond, when they should respond, and how soon after they respond should they be ready for subsequent usage. The policies around current system operations have been developed after many years of experience of reliably operating a particular system and input of the stakeholders of that system. Some policies have shared some very similar methods while others are very different. Many of these systems have very small amounts of VG and therefore, have not seen the need to adjust these methods and policies due to the increased variability and uncertainty of VG. However, most are very aware of the impacts that they are starting to see with the current amounts now on their system. This makes the topic of operating reserve requirements with VG a topic with lots of interest from system operators around the world, and leads to much of the study work that has been performed in recent years.

Studies have been performed for existing systems to estimate the needs of the future for these systems if they were to have large penetrations of VG. Other studies have been performed for test systems, where the goal was more on testing methodologies rather than specific answers for a particularsystem. Each of these studies also may have had different limitations as well. For example, some studies may have been required to keep the current operating or market structure as constant for the future study. Others may have not had the data available for them to study the same types of impacts or same categories of Operating Reserve. The large-scale wind integration studies from section 4.1 seem to use the current terminology of the Operating Reserve on their current systems, with increases to each type and definitions of what each type can and cannot be used for. The academic studies of section 4.2 mostly seem to use more generic terms deciding that a requirement be kept and shared between all forms of variability and uncertainty. On one hand, separating out the requirements may allow for the best form of operating reserve to be used for its particular purpose. For example, specific ramp rates or online/offline contributions kept for certain types of events or non-events. On the other hand, keeping reserve as one requirement exemplifies the fact that a single risk metric is being used, and that the many reasons for loss of load should be considered together to understand the full risk levels. Further work will likely use a tradeoff of these two advantages.

Many of the rules and standards of current operations use a static reserve requirement. Either a single reserve value is used for all the time or simple time-of-day rules are built in to different reserve requirements. Some may use single variables as input to their reserve rules (e.g., load in UCTE *secondary control reserve* requirement). However, many of the newer studies are determining that the variables that cause variability and uncertainty should be used in an intelligent manner to determine a dynamic reserve requirement. It should be understood how all these variables impact the reserve requirements how these variables can be used as part of the reserve determination problem. It may be that there is a function that has all power system variables that can calculate the most optimal amount of Operating Reserves on the system. Since more Operating Reserve can always increase the reliability, the main part of this equation is the minimum risk level allowed for the system or the value the system operator gives to losing load or having ACE or frequency deviations.

As some of the studies showed, new innovative techniques can be done to inherently schedule Operating Reserve without having deterministic requirements set aside. It might be said that a stochastic scheduling program that had infinite computing power so that it can use all possible power system scenarios as part of its solution at a very high time resolution would not need to have any externally input Operating Reserve requirements. The tradeoff between these methods and the reality of finite computing power may lead to newer techniques on how systems schedule for uncertainty and variability and how Operating Reserves requirements are determined.

Some of the uncertainty into how exactly systems will be impacted with high penetrations of VG is because of the lack of data. The data needed to truly understand the impacts are not abundantly available, and modeling of this data is not always perfectly realistic. The more that this data becomes available as larger penetrations of different technologies enter the market and share this data, the better some of these studies and methods may become. It may just be that systems can incrementally understand better the needs of their system as more VG resources enter the grid.

New technologies have recently been shown to provide many of our Operating Reserve types very satisfactorily. For example, demand response, storage resources, and even some VG resources have shown capabilities of providing different types of Operating Reserves for different systems. Each of these resources has different capabilities, limitations, and costs that may change how systems prepare for further uncertainty and variability. It is important though, that the power system needs are what is used when determining the Operating Reserve requirements, and should enforce rules limiting resources that are physically able to provide similar capabilities to historical resources.

In conclusion, further work is needed to find commonality between different methods and policies to determine the most optimal amount of Operating Reserve on different systems. The events that cause the need to use Operating Reserve with higher penetrations of VG need to be better understood so that they can be translated into Operating Reserve requirements. More work on what types of power system phenomena affect the Operating Reserve requirements is needed so dynamic requirements can be tried in current systems. Better sophisticated modeling of uncertainty and high-resolution scheduling programs are needed that try to simulate operator actions rather than only the automatic scheduling programs. Better quality data and more of it will be needed for advancing this research. Better understanding of resource technologies and how they can each contribute differently to different needs will truly advance the research. It is easy to see due to the interest in researchers and power system industry members that the topic is very important and may have a significant impact on how systems integrate more VG technologies onto the grid cost effectively.

6 References

[1] A.J. Wood and B.F. Wollenberg, Power Generation Operation and Control, New York: John Wiley & Sons, 1996.

[2] Prabha Kundur, *Power System Stability and Control*, New York: McGraw-Hill, 1994.

[3] Roy Billinton and Ron Allan, *Reliability Evaluation of Power Systems*, 2nd ed. New York: Plenum Press, 1996.

[4] ERCOT Operations Report on the EECP Event of February 26, 2008. Electricity Reliability Council of Texas. March 2008. Available: http://www.ercot.com/meetings/ros/keydocs/2008/0313/07._ERCOT_OPERATIONS_REPORT_EECP_022608_public.doc.

[5] Yann Robours et al., "A survey of definitions and specifications of reserve services," Technical Report, University of Manchester, December 2005.

[6] Yann Rebours et al., "A survey of frequency and voltage control ancillary services – Part I: technical features," *IEEE Transactions on Power Systems,* vol. 22, no. 1, pp. 350-357, February 2007.

[7] CIGRE Working Group C5.06, "Ancillary services: an overview of international practices," Technical brochure number 435, October 2010.

[8] Brendan Kirby et al., "Providing minute-to-minute regulation from wind plants," *Proceedings of the 9th Annual Large-Scale Integration of Wind Power into Power Systems and Transmission Networks for Offshore Wind Power Plant*, Quebec, Canada, October 2010.

[9] N.W. Miller et al., "Impact of frequency responsive wind plant controls on grid performance," *Proceedings of the 9th Annual Large-Scale Integration of Wind Power into Power Systems and Transmission Networks for Offshore Wind Power Plant*, Quebec, Canada, October 2010.

[10] Brendan Kirby, "Demand Response for Power System Reliability: FAQ," ORNL/TM 2006/565, Oak Ridge National Laboratory, December 2006.

[11] Deqiang Gan et al., "Energy and reserve market designs with explicit consideration to lost opportunity costs," *IEEE Transactions on Power Systems,* vol. 18, no. 1, pp. 53-59, February 2003.

[12] Erik Ela et al., "The evolution of wind power integration studies: past, present, and future," *Proceedings of Power & Energy Society General Meeting*, Calgary, Canada, July 2009.

[13] J. Charles Smith et al., "Utility wind integration and operating impact state of the art," *IEEE Transactions on Power Systems,* vol. 22, pp. 900-908, Aug. 2007.

[14] Michael Milligan et al, "Advancing wind integration study methodologies: implications of higher levels of wind." *Proceedings of American Wind Energy Association Windpower* 2010, Dallas, TX, May 2010.

[15] Eric Hirst et al., "Ancillary-service details: Regulation, load following, and generator response," ORNL/CON-433, Oak Ridge National Laboratory, October 1996.

[16] Nasser Jaleeli et al., "Understanding automatic generation control," *IEEE Transactions on Power Systems*, Vol. 7, no. 3, pp 1106-1122. Aug 1992.

[17] Noureddine Bekhouche, "Automatic generation control before and after deregulation," *Proceedings of the Thirty-Fourth Southeastern Symposium on System Theory,* Huntsville, AL, March 2002.

[18] Yih-Huei Wan, "A primer on wind power for utility applications," NREL Technical Report NREL/TP-500-36230, December 2005.

[19] Matthew Lazarewicz et al., "Grid frequency regulation by recycling electrical energy in flywheels," *Proceedings of Power Engineering Society General Meeting,* Denver, CO, June 2004.

[20] Federal Energy Regulatory Commission, "Notice of Proposed Rule: Frequency regulation compensation in the organized wholesale power markets" Docket Nos. RM11-8-000, AD10-11-000, February 17, 2011.

[21] Michael Milligan et al., “Analysis of sub-hourly ramping impacts of wind energy and balancing area size,” *Proceedings of American Wind Energy Association Windpower 2008*, June 2008, Houston, TX.

[22] William Grant et al., “Change in the Air,” *IEEE Power and Energy Magazine*, vol. 5, pp 47-58, Nov.-Dec. 2009.

[23] FERC Technical Conference on Frequency Response in the Wholesale Electric Grid, Washington, DC, September 2010.

[24] Joseph Eto et al., “Use of frequency response metrics to assess the planning and operating requirements for reliable integration of variable renewable generation,” Prepared for the Federal Energy Regulatory Commission, LBNL-4142E, December 2010.

[25] Andrew Mills et al., “Implications of wide-area geographic diversity for short-term variability of solar power,” Technical Report LBNL-3884E, September 2010.

[26] H.T. Ma et al., “Working towards frequency regulation with wind plants: combined control approaches,” *IET Renewable Power Generation,* vol. 4, no. 4, pp.308-316, 2010.

[27] I. Erlich et al., “Primary frequency control by wind turbines,” *Proceedings of the IEEE Power and Energy Society General Meeting*, Minneapolis, MN, July 2010.

[28] North American Electric Reliability Corporation, “Reliability Standards for the Bulk Electric Systems of North America,” June 2010.

[29] Federal Energy Regulatory Commission, Order 693: Mandatory Reliability Standards for the Bulk-Power System, Docket No. RM06-16-000, March 16, 2007.

[30] North American Electric Reliability Corporation, “Accommodating High Levels of Variable Generation,”April 2009.

[31] “The Value of Independent Regional Grid Operators,” a report by the ISO/RTO Council, November 2005.

[32] NERC Standard Authorization Request, BAL-005-1 – Automatic Resource Control, June 2010.

[33] North American Electric Reliability Corporation, “2009 CPS2 Bounds,” February 2009.

[34] North American Electric Reliability Corporation, Standard BAL-007-1, BAL-008-1, BAL-009-1, BAL-010-1, BAL-011-1, BAL-012-1, Standards authorization request, July 2004. Available: http://www.nerc.com/docs/standards/sar/BAL_007-012_03-11-05.pdf.

[35] Western Electricity Coordination Council, “WECC white paper on the proposed NERC balance resources and demand standards,” December 2006. Available: http://www.wecc.biz/Documents/2007/News/WECC%20White%20Paper%20on%20the%20BRD%20Standards-Dec2006-v3.pdf.

[36] Federal Energy Regulatory Commission, “Notice of Proposed Rule: Integration of Variable Energy Resources” Docket No. RM10-11-000, November 18, 2010.

[37] Jim Ingleson et al., “Tracking the Eastern interconnection frequency governing characteristic,” *Proceedings of the IEEE Power and Energy Society General Meeting,* Minneapolis, MN, July 2010.

[38] GE Energy. “Analysis of Wind Generation Impact on ERCOT Ancillary Services Requirements,” Prepared for Electric Reliability Council of Texas, March 2008.

[39] New York Independent System Operator, “Ancillary Services Manual,” November 2010.

[40] Alberta Electric System Operator, “Operating Reserves Technical Requirements,” December 2004.

[41] DeWayne Todd et al., “Providing Reliability Services through Demand Response: A Preliminary Evaluation of the Demand Response Capabilities of Alcoa Inc.,” ORNL/TM 2008/233, Oak Ridge National Laboratory, January 2009.

[42] Brendan Kirby et al., “NYISO Industrial Load Response Opportunities: Resource and Market Assessment,” Prepared for the New York State Energy Research and Development Authority, Oak Ridge National Laboratory, June 2009.

[43] Northern Tier Transmission Group, “ACE Diversity Interchange overview,” January 2008.

[44]Electric Reliability Council of Texas, "ERCOT Methodologies for Determining Ancillary Service Requirements," 2009.
[45]WECC, "Western Electricity Coordinating Council minimum operating reliability criteria," April 2005.
[46]ISO-RTO Council, North American Wholesale Electricity Demand Response Program Comparison, April 2009, www.isorto.org.
[47]PJM Interconnection, "PJM Manual 11: Energy & Ancillary Services Market Operations," Revision: 45, June 2010.
[48]Governor Droop Setting Criterion, PRC-003-WECC-CRT-1, January 2011.
[49]Western Electricity Coordinating Council, "White Paper on Frequency Responsive Reserve Standard," April 2005.
[50]ERCOT Protocols Section 5: Dispatch. June 1, 2010.
[51]Sandip Sharma et al., "System inertial frequency response estimation and impact of renewable resources in ERCOT Interconnection," presented at ERCOT Emerging Technologies Working Group Meeting, March 2011.
[52]Hawaiian Electric Company, "HECO System Overview," presented at Technical Conference, March 14, 2008.
[53]Hydro-Québec TransÉnergie, Transmission Provider Technical Requirements for the Connection of Power Plants to the Hydro-Québec Transmission System, February 2009. Available: http://www.hydroquebec.com/transenergie/fr/commerce/pdf/exigence_raccordement_fev_09_en.pdf
[54]J. Brisebois and Noel Aubut, "Wind farm inertia emulation to fulfill Hdyro-Quebec's specific need," *Proceedings of the IEEE Power and Energy Society General Meeting,* Detroit, MI, July 2011.
[55]Northeast Power Coordinating Council, A-06 Operating Reserve Criteria, December 2008, section 3.2, accessed 9/21/2010. Available: http://www.npcc.org/documents/regStandards/Criteria.aspx
[56]New York State Reliability Council, "NYSRC Reliabilitiy Rules, for planning and operating the New York State Power System, section D-R4, D-R2, July 2010. Available: http://www.nysrc.org/pdf/Reliability%20Rules%20Manuals/RR%20Manual%2027%20final-2%20July%2010-10.pdf.
[57]Report: ENTSO-E, UCTE Operational Handbook Policy 1, Load-Frequency Control and Performance, March 2009.
[58]GE Energy "The Effects of Integrating Wind Power on Transmission System Planning, Reliability, and Operations: Report on Phase 2," Prepared for The New York State Energy Research and Development Authority, March 2005.
[59]Enernex Corporation, "2006 Minnesota Wind Integration Study Volume I," November 2006.
[60]Northern Arizona University. "Arizona Public Service Wind Integration Cost Impact Study," September 2007.
[61]California ISO. "Integration of Renewable Resources," Nov. 2007.
[62]Yuri Makarov et al., "Operational impacts of wind generation on California power systems," *IEEE Transactions on Power Systems,* vol. 24, no. 2, pp. 1039-1050, May 2009.
[63]Warren Lasher, "The development of competitive renewable energy zones in Texas," *Proceedings of IEEE Transmission and Distribution Conference and Exposition,* Chicago, IL, April 2008.
[64]Peter Meibom et al., "All Island Grid Study Workstream 2B: Wind Variability Management Studies," July 2007.
[65]Peter Meibom et al., "Stochastic optimization model to study the operational impacts of high wind penetrations in Ireland", *IEEE Transactions on Power Systems*, e-published, (ISSN: 0885-8950) (DOI: 10.1109/TPWRS.2010.2070848)

[66]Ronan Doherty and Mark O'Malley, "A new approach to quantify reserve demand in systems with significant installed wind capacity," *IEEE Transactions on Power Systems,* vol. 20, no. 2, pp. 587-595, May 2005.
[67]Enernex Corporation, "Eastern Wind Integration and Transmission Study," Prepared for the National Renewable Energy Laboratory, January 2010.
[68]Charles River Associates, "SPP WITF Wind Integration Study," Prepared for Southwest Power Pool, January 2010.
[69]GE Energy, "Western Wind and Solar Integration Study," Prepared for the National Renewable Energy Laboratory, May 2010.
[70]H.B. Gooi et al., "Optimal scheduling of spinning reserve," *IEEE Transactions on Power Systems,* vol. 14, no. 4, pp. 1485-1492, November 1999.
[71]Jianxue Wang et al., "Operating reserve model in the power market," *IEEE Transactions on Power Systems,* vol. 20, no. 1, pp. 223-229, February 2005.
[72]Miguel Ortega-Vazquez and Daniel Kirschen, "Optimizing the spinning reserve requirements using a cost/benefit analysis," *IEEE Transactions on Power Systems,* vol. 22, no. 1, pp. 24-33, February 2007.
[73]Gundolf Dany, "Power reserve in interconnected systems with high wind power production," *Proceedings of IEEE Porto Power Tech Conference,* Porto, Portugal, September 2001.
[74]Miguel Ortega-Vazquez and Daniel Kirschen, "Estimating the spinning reserve requirements in systems with significant wind power generation penetration," *IEEE Transactions on Power Systems,* vol. 24, no. 1, pp. 114-123, February 2009.
[75]Juan Morales et al., "Economic valuation of reserves in power systems with high penetration of wind power," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 900-910, May 2009.
[76]T. Yong et al., "Reserve determination for system with large wind generation," *Proceedings of Power & Energy Society General Meeting,* Calgary, Canada, July 2009.
[77]François Bouffard and Fransisco Galiana, "Security for operations planning with significant wind power generation," *IEEE Transactions on Power Systems,* vol. 23, no. 2, pp 306-316, May 2008.
[78]J. Wang, M. Shahidehpour, Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Transactions on Power Systems,* vol. 23, no. 3, pp. 1319-1327. August 2008.
[79]Lei Wu et al., "Stochastic security-constrained unit commitment," *IEEE Transactions on Power Systems,* vol. 22, no. 2, pp 800-811, May 2007.
[80]Pablo Ruiz et al, "Wind power day-ahead uncertainty management through stochastic unit commitment policies," *Proceedings of Power Systems Conference and Exposition,* Seattle, WA, March 2009.
[81]Pablo Ruiz et al., "Uncertainty management in the unit commitment problem," *IEEE Transactions on Power Systems,* vol. 24, no. 2, pp. 642-651, May 2009.
[82]Manuel Matos and Ricardo Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Transactions on Power Systems,* vol. 26, no. 2, pp.594-603, May 2011.