Advancing System Flexibility for High Penetration Renewable Integration

Michael Milligan, Bethany Frew, and Ella Zhou
National Renewable Energy Laboratory

Douglas J. Arent
Joint Institute for Strategic Energy Analysis

Advancing System Flexibility for High Penetration Renewable Integration

Michael Milligan, Bethany Frew, and Ella Zhou
National Renewable Energy Laboratory

Douglas J. Arent
Joint Institute for Strategic Energy Analysis

Prepared under Task No. ACTC.1330

NOTICE

Preface

China now installs more renewable electricity each year than any other country in the world. Much of this is *variable* renewable electricity, especially wind and solar generation. A growing body of experience exists from around the world on how to plan and operate electricity grids with high penetrations of variable renewable electricity. China is actively contributing to this body of experience given the rapid growth in renewable electricity deployment there, while at the same time digesting experiences from other countries.

This report is part of a series describing technical collaboration between the National Renewable Energy Laboratory (NREL), the China National Renewable Energy Center (CNREC) along with other key research institutes in China, and the Danish Energy Agency. The collaboration focuses on sharing experiences in the planning, deployment and operation of high-penetration renewable electricity grid systems. The Children's Investment Fund Foundation in the United Kingdom is funding this five-year collaboration.

The core element of the collaboration during this first year was a series of expert engagements in China to share technical knowledge and experience on four key topics:

1. Comprehensive energy scenario design and modeling
2. Renewable energy (RE)-friendly grid development
3. Power system flexibility
4. Boosting distributed generation of RE.

These engagements built on and significantly expanded existing collaboration between the Danish Energy Agency and CNREC experts.

This report summarizes some of the issues discussed during the engagement on the third topic listed above. By design, the focus is on flexibility options used in the United States. Exploration of whether and how U.S. experiences can inform Chinese energy planning will be part of the continuing project, and will benefit from the knowledge base provided by this report. We believe the initial stage of collaboration represented in this report has successfully started a process of mutual understanding, helping Chinese researchers to begin evaluating how lessons learned in other countries might translate to China's unique geographic, economic, social, and political contexts.

We look forward to continuing the collaboration for the remaining four years and building on these initial successes.

Acknowledgments

Acronyms

BA	balancing area
CAISO	California Independent System Operator
CIFF	Children's Investment Fund Foundation
CPP	critical peak pricing
DIR	dispatchable intermittent resource
DR	demand response
EIM	energy imbalance market
ELCC	effective load carrying capability
ERC	effective ramping capability
ERCOT	Electric Reliability Council of Texas
EUE	expected unserved energy
FERC	Federal Energy Regulatory Commission
GW	gigawatt
IRRE	insufficient ramping resource expectation
ISO	independent system operator
LOLE	loss of load expectation
LOLH	loss of load hours
LOLP	loss of load probability
min-gen	minimum generation
MISO	Midcontinent Independent System Operator
MW	megawatt
MWh	megawatt-hour
NREL	National Renewable Energy Laboratory
PFD	period of flexibility deficit
PJM	PJM Interconnection
PTC	production tax credit
RTO	regional transmission organization
RTP	real-time pricing
SPP	Southwest Power Pool
TOU	time of use
VG	variable generation

Table of Contents

List of Figures

List of Tables

1 Introduction

The goal of typical modern electricity systems is to ensure reliable delivery of electricity at an affordable cost to consumers. Flexibility is the ability of a system to respond to variability and uncertainty of demand and supply. Loads change, sometimes in unpredictable ways, and conventional generators may be unavailable due to unexpected events such as natural disaster or mechanical failure. Sources of variable generation (VG), such as wind and solar power, provide power that changes over time based on weather patterns and paths of the sun, which may introduce faster changes in aggregate supply than in systems without VG. As a result, VG increases the response requirements from conventional generators and load, even though it does not increase the overall capacity requirements (Milligan et al. 2011). The relationship between integration of renewable energy and flexibility has received considerable attention in recent years (see for example Holttinen et al. 2013). Grid integration studies in the United States have shown that system flexibility needs increase significantly when more than about 30% of a system's annual electricity demand is provided by VG, assuming certain operational improvements with the increase of VG toward this level (Denholm and Hand 2011). Similar conclusions are drawn in Chinese studies of wind integration, showing a jump in system flexibility issues after 30% penetration of wind (Li 2015). Due to operational and transmission constraints, however, regions such as Jilin and Xinjiang have high levels of wind curtailment even when the penetration level is far less than 30%. This report, along with the Children's Investment Fund Foundation (CIFF) project grid development report *'Renewables-Friendly' Grid Development Strategies* (Hurlbut et al. 2015)*,* will address these issues and lay out the strategies for accommodating a high penetration of renewables.

The need for flexibility applies to all time scales, ranging from the many years that comprise the planning and investment time horizon, to operational planning that may involve days to months, and to operation itself, which encompasses periods as long as a few days to as short as sub-seconds. The shortest time intervals are those in which inertial response provides the first line of defense against imbalance or frequency excursions—we will not address these issues in this report. We do note that wind turbines and solar inverters can now provide simulated primary frequency response and inertial response, along with automatic generation control, and can even respond to dispatch signals. Therefore, wind and solar power are capable of providing some of the flexibility needed by the system; however, these may not be sufficient nor the most economic sources of flexibility. In the discussion that follows, we focus on other sources of flexibility, noting that wind and solar power can provide some of this flexibility given improved power electronics and controls.

This report describes several potential sources of flexibility that can help maintain system balance with high levels of VG. The analysis of flexibility needs falls under the general task of planning for future power system needs, which is itself a broad and complex topic. With high levels of VG, the planning process does not fundamentally change, but rather is augmented so that the characteristics of this generation mix can be properly assessed. That is the focus of this report: to describe sources of flexibility that can be evaluated in the planning process to help the power system operator maintain system balance.

We describe both physical flexibility and institutional flexibility. Physical flexibility, which is the physical capability of power system components to respond to changes in demand and

supply, is a necessary but not sufficient condition to achieve flexible operations. The other required condition is institutional flexibility, which is the ability to deploy the physical flexibility when needed and when it is available through operational practices and/or market design structures. Most sources of flexibility include at least some component of physical and institutional flexibility. The importance of institutional flexibility must not be overlooked. In many cases, physical flexibility can be muted by institutional barriers. This is true even without the presence of VG. For example, as we discuss later, physical flexibility can be dampened by something as simple as the market settlement process, in spite of what might otherwise be a good market design. Table 1 summarizes the sources of flexibility that are discussed in this paper and the companion CIFF program paper entitled *'Renewables-Friendly' Grid Development Strategies* (Hurlbut et al. 2015), as well as additional options (see Milligan et al. 2009), and indicates their *dominant* categorization (physical, institutional, or both). We categorize flexibility measures as "physical" if their dominant flexibility provision is based on inherent physical characteristics of that technology or system component. For example, a robust electrical grid relies on transmission lines with sufficient capacity and redundancy; geographically dispersed VG involves building VG resources across large geographic areas to smooth out the aggregated supply.

Institutional measures provide flexibility primarily through market designs or operational practices that are generally technology-agnostic. Flexibility measures that require physical flexibility from the system components as well as proper operational, regulatory, or market structures are categorized as "both." For example, regional transmission planning for economics and reliability requires planning for a robust transmission network that connects flexible generators as well as the proper coordination and market signals to extract that flexibility within the network; VG forecasting effectively integrated into operations requires accurately forecasting the variability of VG and the operational practices to best utilize that information; and primary frequency response, inertial response, and response to dispatch signals with new VG technologies rely on operational and market structures to capture the flexibility from these physical attributes. Additional examples within each category are discussed in depth in Section 5.

Table 1. Flexibility Measures to Assist with the Integration of VG

Flexibility Measure	Physical or Institutional?	Discussed in this Paper	Discussed in *Grid Development Paper*
Larger balancing areas	Both	✓	
Access to neighboring markets	Both	✓	
Faster energy markets	Institutional	✓	
Regional transmission planning for economics and reliability	Both		✓
Robust electrical grid	Physical		
Improved market design	Institutional	✓	
Demand response	Both	✓	
Geographically dispersed VG	Physical		
Strategic VG Curtailment	Both	✓	
VG forecasting effectively integrated into operations	Both		✓
New flexibility ancillary services products	Institutional	✓	✓
Sufficient reserves for VG event response	Physical		✓
Flexible conventional generators	Physical	✓	
Primary frequency response, inertial response, and response to dispatch signals with new VG technologies	Both		
Storage	Physical	✓	

In this report, we first provide an overview of the current system planning process employed in the United States and additional considerations for higher penetration of VG. Then we discuss the process for assessing the overall system's need for flexibility, which includes (1) quantifying the system's flexibility requirements, (2) quantifying the existing system's ability to supply the needed flexibility, and (3) selecting sources of additional flexibility to satisfy any flexibility deficiency. Cost-benefit and additional considerations are also discussed. The report concludes with high-level lessons-learned for consideration by power system planners. There are many details of how the system is operated, coupled with potential market design elements that are complex and are not considered in this initial report. Instead, this report provides a high-level description of flexibility needs in the context of the general resource planning process. Here we do not consider other elements of planning such as power flow, dynamic stability, or transmission planning.

3

2 Planning for Variable Renewable Energy Sources

In many countries, resource planning has historically been accomplished by projecting future demand patterns, and evaluating one or more potential resource combinations to determine the resource mix that best accomplishes the competing objectives of maintaining reliability and minimizing cost, subject to various risk preferences and regulatory constraints. A central focus of this process is how to achieve and maintain resource adequacy—the level of installed capacity that is necessary to serve demand at all time periods. There are several competing approaches to assessing resource adequacy, including the use of planning reserve margins, which is the percentage by which installed capacity exceeds peak demand, and more rigorous probabilistic approaches based on loss of load probability (LOLP). Common probabilistic approaches include loss of load expectation (LOLE), which is often measured in days/year; expected unserved energy (EUE); and loss of load hours (LOLH). It is worth noting that LOLE and LOLH metrics only capture the number of events and do not reflect the size of the energy or capacity shortfalls. For this reason, the EUE metric is sometimes preferred. Examples of the large, uncorrelated differences that can be observed between the planning reserve margins and three commonly used LOLP-based metrics[1] are shown in Figure 1. The figure also shows that an LOLH of 2.4 hours per year is very different than an LOLE of 0.1 events per year, which indicates that rigorous benchmarking must be performed to determine the level of LOLH that corresponds to an LOLE reliability level of 0.1 events per year (see Ibanez and Milligan 2014). Additional examples, along with relevant discussions, can be found in North American Electric Reliability Corporation (NERC) (2011), Duignan et al. (2012), and Keane et al. (2011). It is important to understand and properly use these differing approaches to resource adequacy, as they can have significant consequences in the resource planning process that is discussed above.
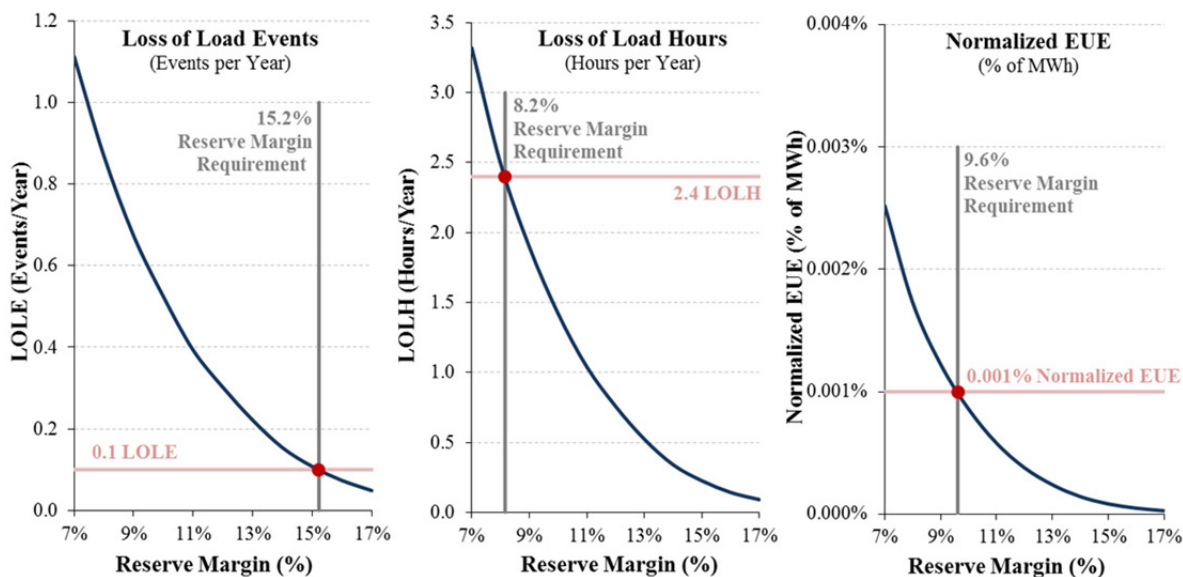
Figure 1. Planning reserve margins required to meet different physical reliability standards

Source: Pfeifenberger et al. 2013

---

[1] LOLH of 2.4 hours per year and LOLE of 0.1 events per year are different uses in the United States than the common "1 day in 10 years" standard; the 0.0001% normalized EUE standard is used in some international markets (Pfeifenberger et al. 2013).

4

Once a resource adequacy target has been adopted, the plant mix can be evaluated by calculating LOLP or a related metric. Assessments of the contributions of VG—or any resource type—can be carried out in this type of modeling framework.

The capacity contribution of a given resource or group of resources is called the effective load carrying capability (ELCC), and is graphically represented in Figure 2.
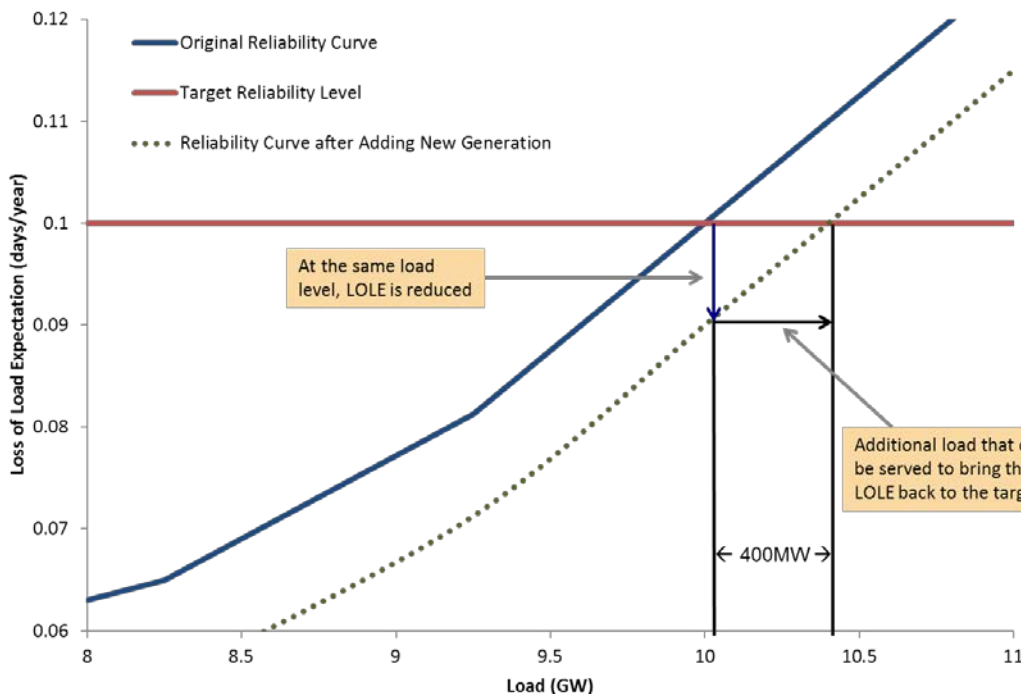
Figure 2. Graphical representation of effective load carrying capability

Source: adapted from Ibanez and Milligan 2014

The y-axis shows the reliability level in terms of days per year of LOLE; note that higher LOLE values denote worse reliability levels. The x-axis shows peak demand, and the original reliability curve (blue line) shows that, with a given resource mix, reliability gets worse at higher levels of peak demand. Assuming a one day in 10 years LOLE, this system can support about 10 gigawatts (GW) of peak demand at the reliability target (horizontal red line). When a new resource is added to the mix (e.g., 2000 MW wind plant), the entire reliability curve shifts right, as shown by the dotted green line. With this new resource, additional demand can be met, and at the new intersection of the reliability curve with the red target line, an additional demand of about 400 megawatts (MW) can be supported. Therefore, the capacity credit, or ELCC, of the new resource is 400 MW.

With high levels of renewable energy, this type of analysis is important so that sufficient resources can be developed in advance of the need. However, this analysis does not capture any of the flexibility needs or the attributes of flexibility solutions that must be addressed so that operational balance in the future can be achieved. New methods are now being developed to span this divide between resource adequacy and flexibility adequacy. For example, California Independent System Operator (CAISO) has implemented flexible capacity requirements and a

5

new ancillary market product to specifically incentivize flexible generator capability (CAISO 2014). These components are discussed in Section 5.7.

The impact of VG on system balancing needs —flexibility— is typically viewed through the "net load," which is load minus VG in each hour (other time increments may be used) and represents the load that must be met by the conventional generation fleet in each time step if all VG is utilized. Figure 3 shows how wind generation can impact system operations in an example week. Net load peaks are shorter in duration, resulting in fewer operating hours for conventional generators; this affects energy-based cost recovery and, consequently, may impede long-term security of supply (for further discussion of this issue see Milligan et al. 2012a and Ela et al. 2014a). Steeper ramps require a faster rate of increase or decrease of dispatchable generation. Lower turn-downs require dispatchable generators to turn down output to low levels (to accommodate high VG output periods) but remain available to rise again quickly (Cochran et al. 2014). Solar generation will result in characteristically similar impacts.

When flexibility needs are not met, the system may experience reliability and economic consequences. These include dropped load, VG curtailment, deviations from the schedule of area power balance, frequency and voltage excursions due to over- or under-generation, negative market prices, and price volatility.
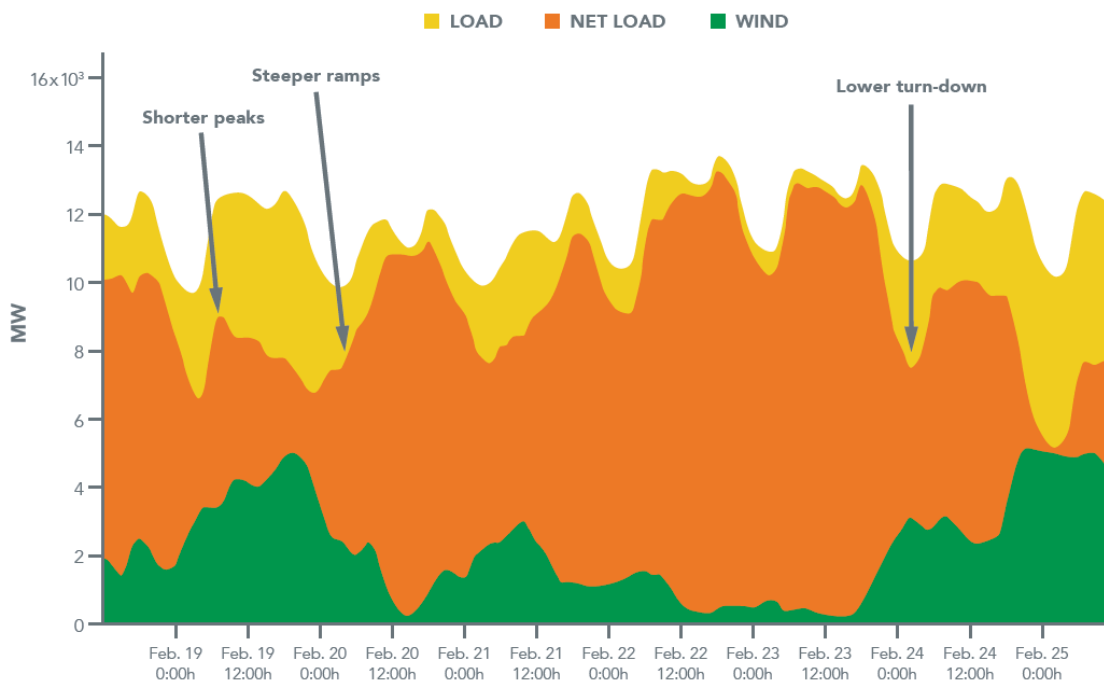
Figure 3. Wind (and solar) generation can lead to greater need for flexibility

Source: Cochran et al. 2014

As described above, flexibility can be provided by a suite of options, including physical and institutional intervention. Predominantly physical options include storage, flexible conventional generation (fast ramping and low output level capabilities), active power controls on VG, demand response through flexible load such as electric vehicles and programmable water and

6

space heating and cooling, and transmission networks with limited bottlenecks and sufficient capacity to access a wide range of balancing resources. Predominantly institutional options provide access to and best extraction of flexibility from the physical system, including large balancing areas (BAs); market designs that utilize centralized scheduling and fast dispatch; improved VG forecasting; and the implementation of demand response (DR) through a smart grid to allow customers to respond to market signals or direct load control (Cochran et al. 2014). It's worth noting that even for predominately physical options, such as storage, appropriate institutional framework needs to be in place to effectively utilize them. Figure 4 provides an alternative summary (compare with Table 1) of a subset of possible flexibility options, divided by category and relative costs. The flexibility options in this figure are ordered by relative cost on the vertical axis, with illustrative error bars indicating that there is a variation among the costs, which are very system-dependent.

RELATIVE ECONOMICS OF INTEGRATION OPTIONS
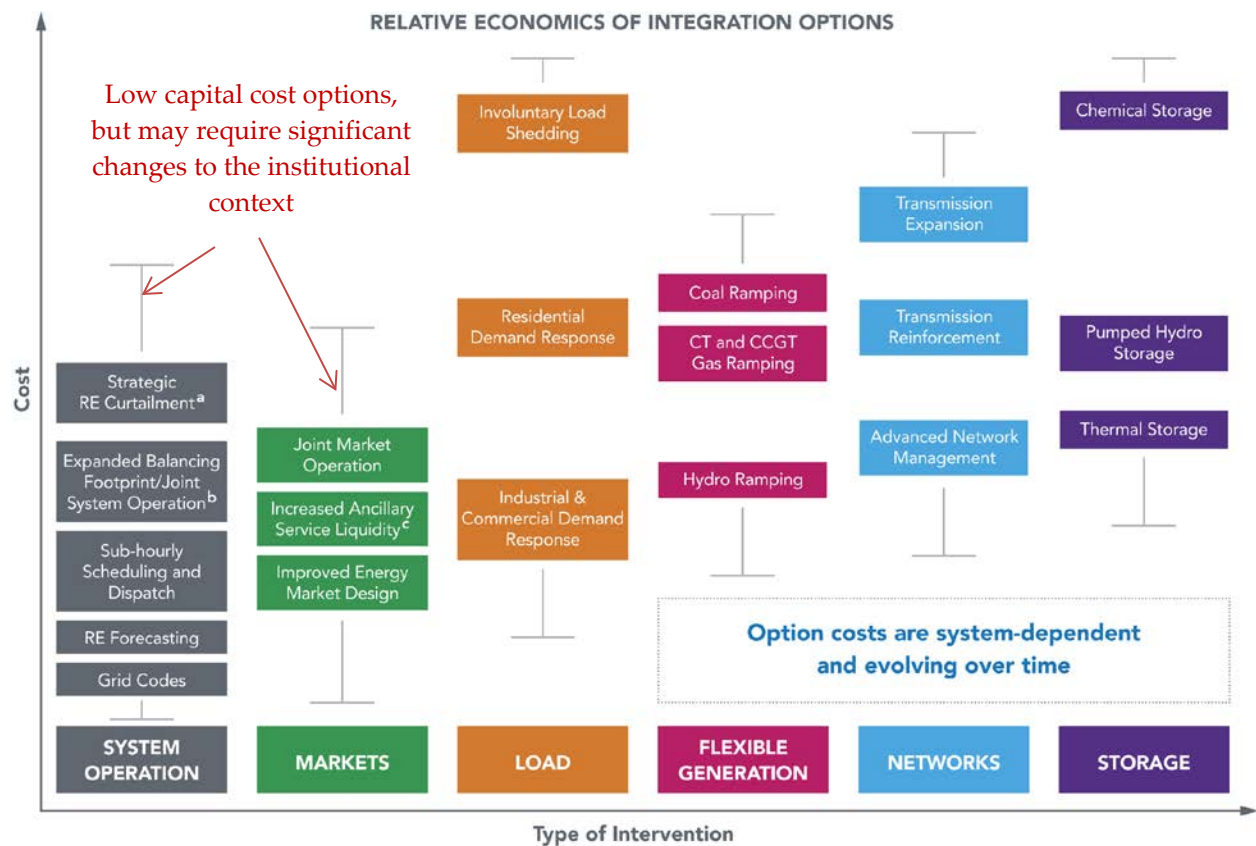
a There is a tradeoff between costs of flexibility and benefits of reduced (or no) curtailment, hence a certain level of curtailment may be a sign that the system has an economically optimal amount of flexibility.

b Joint system operation typically involves a level of reserve sharing and dispatch co-optimization but stops short of joint market operation or a formal system merger.

c Wind power can increase the liquidity of ancillary services and provide generation-side flexibility. Curtailed energy is also used to provide frequency response in many systems, for example Xcel Energy, EirGrid, Energinet.dk.

Figure 4. Flexibility options

Source: Cochran et al. 2014

7

3 Assessing the Need for Flexibility

This section describes data requirements and emerging modeling methods to quantify flexibility needs under future VG resources. The first step in assessing the overall system's need for flexibility is to understand and quantify the system's flexibility requirements. Various data are necessary to determine how much and what type of flexibility a system needs. These include installed capacities, locations of VG resources, and time series data of load and of those same VG resources. It is important that these VG and load time series data be time-synchronized to properly account for the underlying weather patterns (Milligan et al. 2012b).

Determining the need for flexibility begins with the development of a high-quality and fine-temporal-resolution data set, which, fortunately, is the same data set required for the backbone of integration analysis and modeling (NERC 2010). Alternative load profiles—hourly or sub-hourly demand curves for at least one year and covering the planning horizon—provide the first data set that is needed. Because this is a standard planning requirement, we do not describe this process here.

Accompanying this demand data is a complementary data set of wind and solar power, developed in a way that allows for multiple wind/solar penetrations, locations, and timing to be evaluated for alternative scenarios. The state of the art is to ensure that the wind power, demand, and solar power data are all based on the same meteorological year in order to capture accurate correlations between these datasets. Creating plausible scenarios from the data sets described above is the cornerstone of the flexibility needs analysis.

One common first step is to analyze the variability of the net load (load minus VG) constructed at hourly (or shorter) time intervals. Many types of statistical analyses can then be performed on the net-demand data series, which can also address multiple scenarios of renewable buildout, timing, mix of wind and solar, alternative demand scenarios, and many others.

A chronological analysis might analyze ramping needs based on the behavior of net load. An example is shown in Figure 5. This graph shows one week of demand and net load, based on a high wind penetration level. The increasing level of ramping can be discerned in the upper panel, and the lower panel of the graph quantifies the increase in ramping requirements based on the no-wind case.
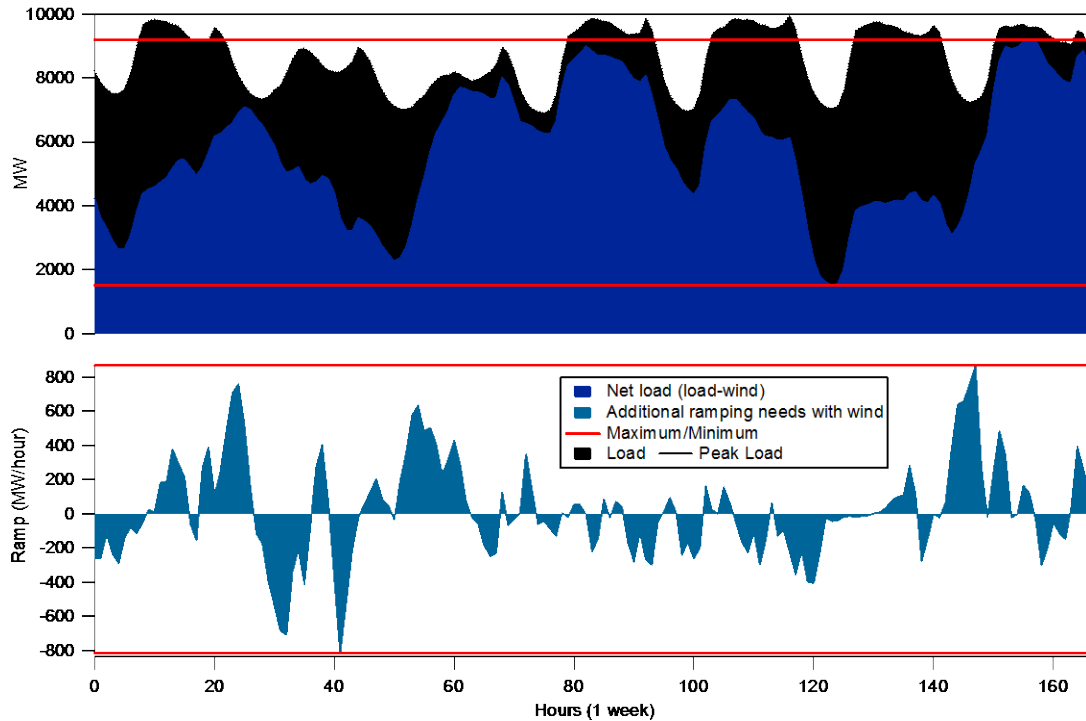
Figure 5. Example time series graph that shows the impact of high levels of wind energy on ramping needs and minimum generation levels (one selected week of data)

Source: Milligan 2014b

In the United States, flexibility needs are assessed using various approaches, which typically start from the net load using the time-synchronized load and VG data mentioned above. In CAISO, a recently approved measure incorporates flexibility needs into resource adequacy plans. Instead of only relying on peak load as an indicator of the required system installed capacity, CAISO will now also incorporate the forecasted net load maximum 3-hour ramp (in megawatts) for each month in its requirements for system capacity specifications (CAISO 2014). This measure assumes that the ramp event is constant over all 3 hours, which is often not the case. Figure 6 shows a distribution of the projected 2018 ratio of the maximum 1-hour net load ramp to the maximum 3-hour net load ramp by month. The different colors represent the percentage of the forecasted ratio each month, shown by quartile. These results show that the largest 1-hour net load ramps often comprises a significant portion of the maximum 3-hour ramp, reflecting shorter duration flexible capacity needs beyond the current 3-hour consideration. These 1-hour net load maximum ramps are projected to grow in size and occupy a larger share of the 3-hour ramp (CAISO 2014). The large single-hour ramp contribution in some months (e.g., March, October, and November) reveals that ramping rates, and not just magnitude, can be an important flexibility requirement.

9
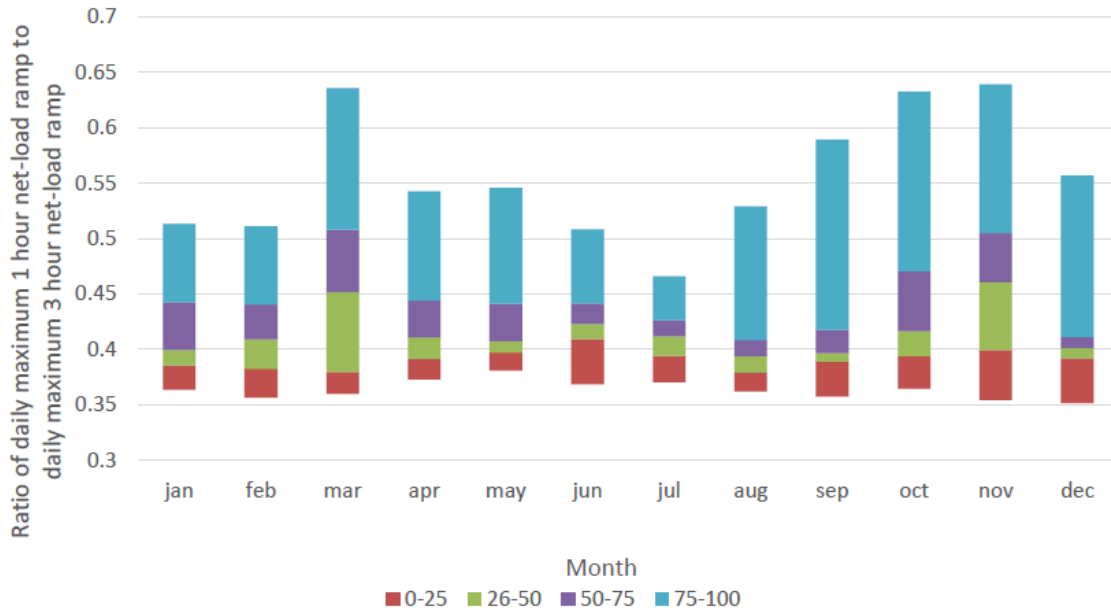
Figure 6. CAISO 2018 distribution of forecasted net load ratio of maximum 1-hr net load ramp to maximum 3-hr net load ramp with relative percentage of contribution (colors)

Source: CAISO 2015a

Annual data can be summarized statistically or can be arranged as duration curves. For example, ramp-duration curves can be constructed that can capture different levels of statistical containment. Figure 7 shows an example that is based on containment levels ranging from 90% to 100% and for time spans up to 12 hours. Comparing alternative ramp envelopes can help inform decisions regarding the type of resource(s) that may be capable of providing the service. For example, the relatively large but infrequent need for 4-hour ramping capability, comparing a 99% and 100% containment level, may suggest some form of DR as compared to a more expensive resource acquisition.

10

Probability of Net Ramp Magnitude and Duration
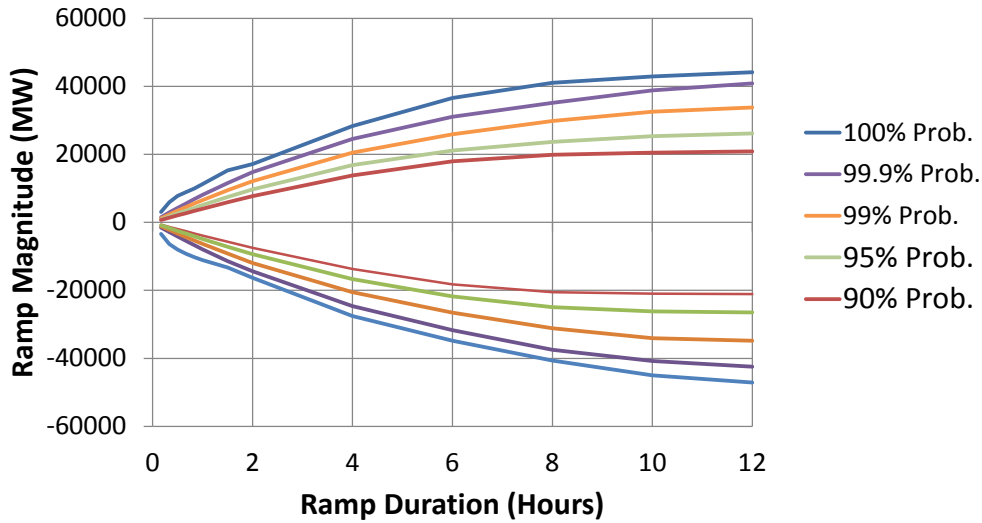
Figure 7. Example ramp duration curve that shows alternative ramping envelopes that correspond to different statistical exceedance levels

Source: King et al. 2011

Other approaches have also been utilized to capture ramp needs based on time of day/time of year. Figure 8 is a so-called "magic carpet" plot, which summarizes one year of hourly data into a visual representation that can be useful for system operators and planners to anticipate the times of day and year that ramping capability will most likely be needed (this particular plot is shown for the Energy Imbalance Market footprint in the western United States, which is discussed in Section 5.1). From the basic single-scenario data, statistical uncertainty bands could also be developed to capture potential impacts of forecasting uncertainty, both from the renewable resource and demand. Composite or multiple diagrams from additional years of data can also help characterize the impacts of both uncertainty and inter-annual variability on flexibility needs. This method, as well as other approaches, could also be extended to different time scales for additional support in identifying and quantifying the system's flexibility needs.

11
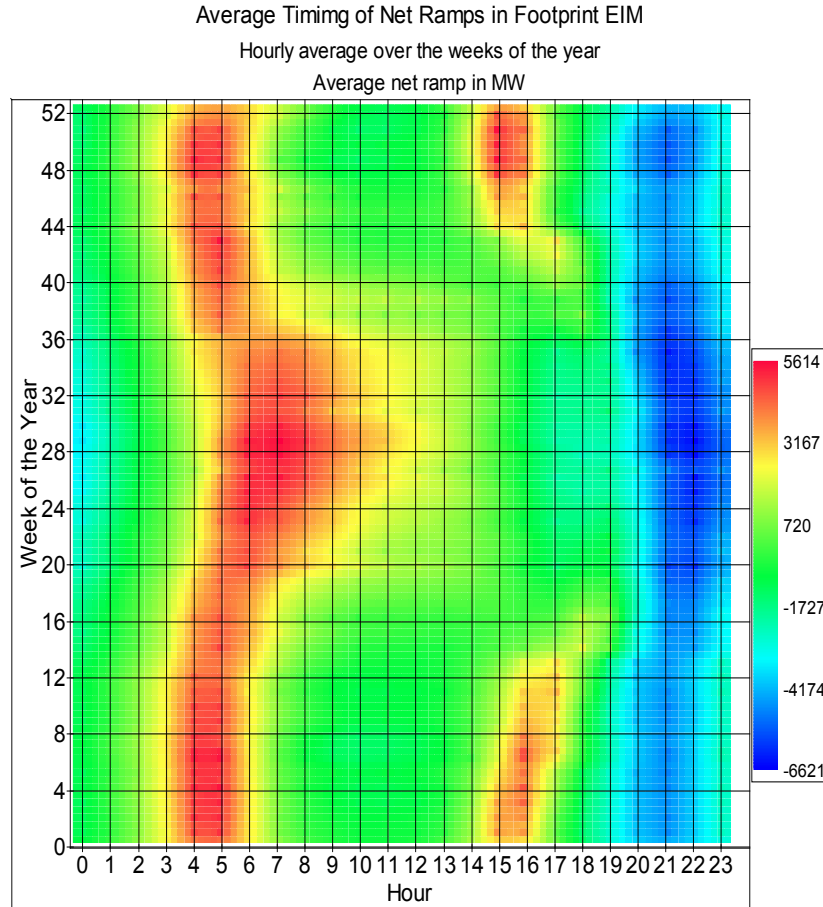
Figure 8. Example net ramp behavior by hour of day and week of year in the Energy Imbalance Area (EIM) footprint

Source: King et al. 2011

When interpreting the results from net-demand analyses such as these, it is important to remember that these analyses assume that wind and solar generation are totally passive and unable to provide ramping or turn-down capability. This assumption is increasingly at odds with industry practice in the United States, where VG resources are being equipped with active power control. These expanded capabilities can be appropriately represented in production simulations by allowing all capable resources to provide some level of required ancillary services, especially balancing services that can be analyzed by production simulation modeling. In such a framework, the economic provision of these services can be robustly calculated, and based on the results of these analyses, mitigating measures can be evaluated (Ela et al. 2012b).

12

4 Assessing Existing Flexibility Resources

Once the system's flexibility needs have been estimated, the next step is to assess the existing system's ability to supply flexibility by characterizing the flexible resources available to it. Simple approaches compare generation and demand-side resource data against the quantified need from the previous section. Appropriate data include generation characteristics, such as minimum generation (min-gen) levels and ramping rates, and existing alternative sources of flexibility, such as DR resource profiles. Any relevant institutional constraints should also be considered at this point. A simplified spreadsheet tool can be used to estimate the dispatch stack and resulting flexibility (min-gen, ramping magnitude and speed, etc.) for the comparison against the quantified system need; an example of this type of approach can be found in Kirby and Milligan (2005).

A more complex and recommended approach to assess the existing flexibility resources and needed flexibility is to simulate production, using modeling tools such as Plexos, GE-MAPS, Gridview, or Pro-Mod. These grid simulation tools model the operation of the entire bulk power system. These tools are sometimes referred to as "production cost" and "security-constrained unit commitment and economic dispatch" models. "Security-constrained" reflects the inclusion of transmission constraints in the economic dispatch and unit commitment processes. Production simulation requires input data on various costs (of generator fuel; variable operations and maintenance; generator start cost; contract purchase and sale price; transmission wheeling; energy, ancillary services, fuel; market prices, etc.), system load, plant characteristics, transmission capability, and generation uncertainties of VG resources. Assumptions regarding the potential future state of the power system are critically important and will have a significant impact on the model outputs. Some models take a deterministic approach, using a single year of load, wind, solar, and hydro conditions, and incorporating regulation and load following requirements in economic commitment and dispatch decisions. Some models stochastically simulate different conditions with a set of scenarios of different weather years while assuming perfect foresight in unit commitment decisions. Some stochastic models develop an initial commitment considering the uncertainty at that time and adjust commitment or dispatch of resources as needed (Kiviluoma et al. 2010; CPUC 2014; CAISO 2015b).

The flexibility of the simulated system can be assessed by examining several key outputs of the production simulation model. Such outputs include: total up/downward reserve shortfall, max up/downward reserve shortfall, number of hours of up/downward reserve shortfall, total renewable curtailment, maximum renewable curtailment, number of hours of curtailment, total dump energy, maximum dump energy, number of hours of dump energy, along with reliability metrics such as LOLP, LOLE (often expressed as days/year), LOLH-LOLE in units of hours per year, or EUE.

Some relatively new methods for assessing flexibility in the power system include insufficient ramping resource expectation (IRRE), periods of flexibility deficit (PFDs), effective ramping capability (ERC), and CAISO's ramp-based flexible capacity method (discussed above). IRRE uses a probabilistic approach to determine the number of periods when a power system cannot meet each net load ramp. The operational characteristics for each generator, the energy production time series—historical or simulated—of each flexible resource, and each resource's maximum and minimum rated output, start-up time, ramp up and down rate, forced outage

probability, and production levels are required for IRRE calculation (Lannoye et al. 2012b). PFD differs from IRRE in that it makes a direct comparison between the available flexibility from a simulated production time series and the net load ramps in the chosen direction. It identifies the time horizons associated with the flexibility deficit so that different solutions might be devised for each flexibility issue (EPRI 2014). ERC describes a unit's contribution to the system's ability to ramp upward or downward over a specified period of time (Lannoye et al. 2012a). ERC is similar to ELCC, except that instead of calculating contributions toward meeting overall capacity needs, ERC uses the unit's maximum ramp in a given direction and time period to indicate contributions toward meeting ramping needs.

14

5 Options to Increase Flexibility

If the available system flexibility is not sufficient to cover the need as determined by the previous steps, then sources of additional flexibility should be evaluated based on their technical and economic merits. The best solutions are system specific and include both the necessary physical flexibility and the institutional access to that flexibility. In this section, we discuss the most-selected flexibility options from experience in the United States, as summarized in Table 1. The availability of these measures is not uniform throughout the United States.

5.1 Larger Balancing Areas

In the United States, a key physical and institutional flexibility mechanism is increasing the size of BAs.[2] This typically involves the physical interconnection of adjacent regions through an enhanced transmission network. However, there are alternative approaches to achieve some or all of the benefits of such operational consolidation. They include dynamic scheduling, intra-BA scheduling at sub-hourly time steps, or other wide-area economic dispatch concepts that do not require physical consolidation of BAs (Milligan and Kirby 2010a; Denholm and Cochran 2015). Larger BAs provide greater access to load and generation diversity and a larger pool of reserves. This results in numerous operational efficiency benefits. For example, the ramping capability of generation adds linearly, whereas the ramping need of large areas increases less than linearly (Figure 9).
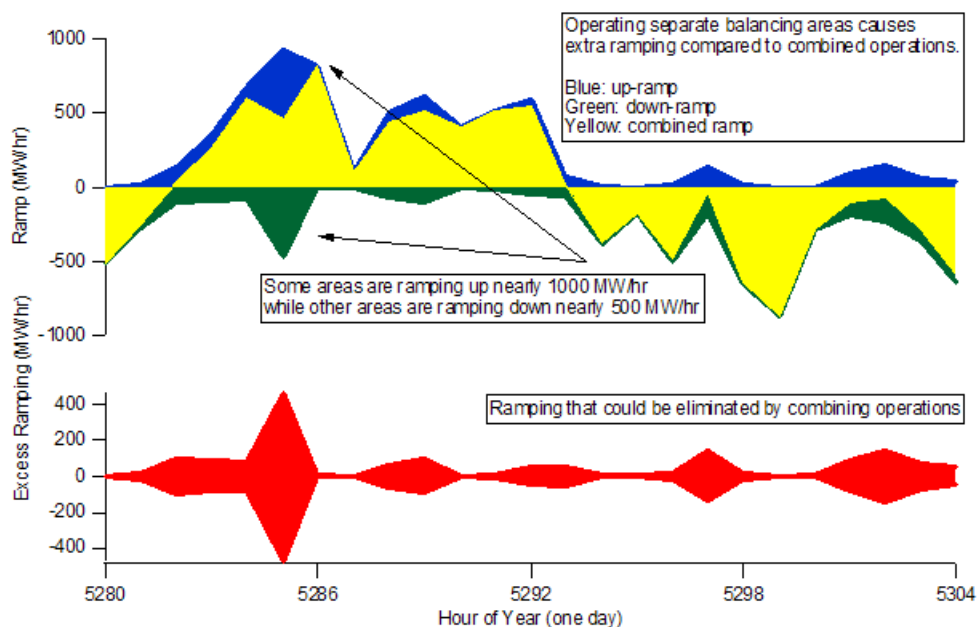
Figure 9. Large BAs result in less total ramping need

Source: Milligan and Kirby 2007

---

[2] Balancing (Authority) area is the collection of generation, transmission, and loads within the metered boundaries of an entity (Balancing Authority) that integrates resource plans ahead of time and maintains load-interchange-generation balance of that area. The Balancing Authority supports interconnection frequency in real time.

The institutional benefits of larger BA and faster scheduling are shown in Figure 10.[3] The aggregate regulation reserve requirements (and resulting system costs) across this interconnection decrease as the BA footprint grows from small to medium to large (right to left in the figure). These reserve requirements (and costs) also decrease as the dispatch interval and forecast lead times decrease (colored bars). Smaller dispatch intervals correspond to faster energy markets, and smaller forecast lead times correspond to more frequently updated (and therefore accurate) VG and load forecasts.
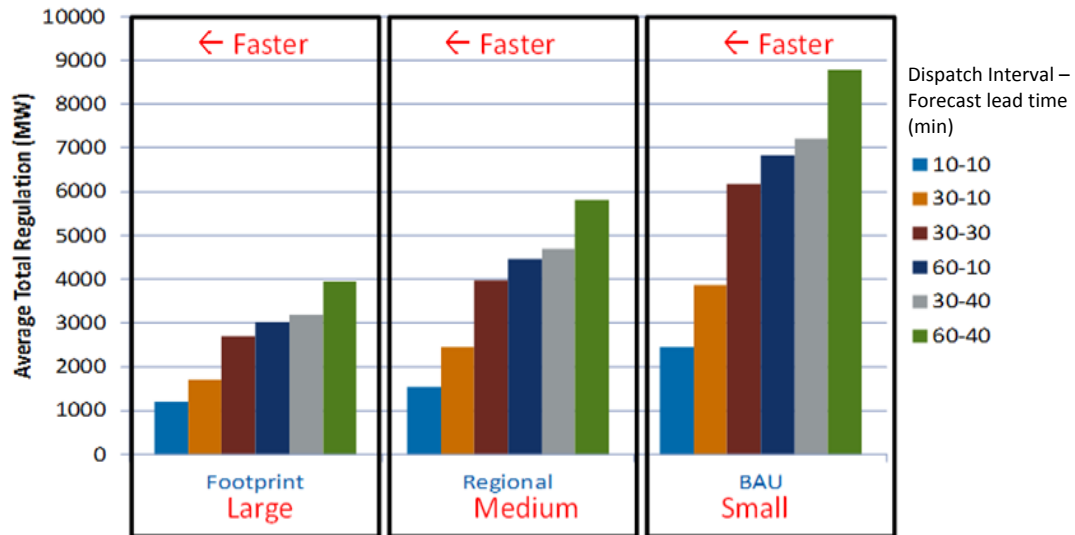
Figure 10. Faster energy scheduling (colors) and larger BAs (panels) greatly reduce aggregate regulation requirements and wind integration impacts

Source: Milligan et al. 2011

Midcontinent Independent System Operator (MISO), Southwest Power Pool (SPP), and PJM Interconnection (PJM) are examples of regional transmission organizations (RTOs) in the United States that have physically expanded to capture these BA size benefits. MISO created and integrated its South Region in 2013, citing benefits of improved reliability and reduced regulation and spinning reserve requirements (see MISO's website[4]). SPP integrated portions of Nebraska in 2009 and was approved in 2014 to add large portions of the upper Great Plains, citing an estimated $334 million in net system benefits from increased access to generation into and out of Nebraska and availability of lower-priced hydro generation (SPP 2009; FERC 2014). Various utilities have joined PJM since 2004, expanding its footprint from North Carolina to Illinois (see PJM's website[5]). In the western United States, a new EIM has been formed (see

---

[3] This graph includes all of the U.S. Western Interconnection except for California and Alberta.
[4] https://www.misoenergy.org/WhatWeDo/StrategicInitiatives/ SouthernRegionIntegration/Pages/SouthernRegionIntegration.aspx.
[5] http://www.pjm.com/about-pjm/who-we-are/territory-served.aspx and http://www.pjm.com/about-pjm/who-we-are/pjm-history.aspx.

CAISO's EIM website[6]), and although initial participants include only the CAISO and PacifiCorp, other utilities have announced plans to join the EIM in the next 1 to 2 years, which will increase the effective balancing size of this market. The EIM is a real-time, security-constrained economic dispatch on imbalances, running every five minutes. It therefore does not include the other aspects of the large RTO markets in the United States and is an approach to pooling the economic dispatch of imbalances.

While no BA operator in the United States has ever decided that it is too big, there may be extra cost considerations of a large BA. These include the cost of additional computational requirements for monitoring and operating the system, including security-constrained unit commitment and security-constrained economic dispatch, as well as power system state estimation. Additionally, transmission congestion can prevent the realization of the full set of benefits.

5.2 Access to Neighboring Markets

Access to neighboring markets is both a physical and institutional flexibility measure and is closely related to the larger BA option. It requires physical interconnection via transmission networks and the institutional mechanisms to manage the coordinated operation of those transmission lines and the markets that they connect. This latter coordination across borders can include a wide variety of options, including dynamic scheduling or allowing interchange schedules to change at relatively short time steps.

This flexibility measure captures the interplay between transmission and generator capacity needs. Building new lines allows multiple areas to share generator resources, thereby reducing the total required generating capacity among all constituent regions. The resulting economic and reliability benefits are realized regardless of the VG penetration level. As shown in Figure 11, for modest levels of VG (roughly 7%–20% wind and solar capacity-based penetration in all constituent regions), the total system costs (production and reliability, red dots) decrease as the access to neighboring markets increases (moving from "Island Case" down to "Long Neighbor Case"). At the same time, resource adequacy (0.1 LOLE target, blue dots) benefits are achieved with greater access to neighboring markets. The cases in this figure reflect different levels of neighbor assistance: Base Case, where the neighbors have 15% reserve margins and the Study RTO has 11,000 MW of intertie capacity; "Long Neighbors Case," where the neighbors' reserve margins are increased to 20% compared to 15% in the Base Case (and intertie capability equal to the Base Case); "50% Transmission Case," with interties at 50% relative to the Base Case (and neighbors' reserve margins at 15%); and an "Island Case" with no interties (Pfeifenberger et al. 2013).

---

[6] "Energy Imbalance Market." 2015. Folsom, CA: California ISO. Accessed September 2015, http://www.caiso.com/informed/pages/stakeholderprocesses/energyimbalancemarket.aspx.
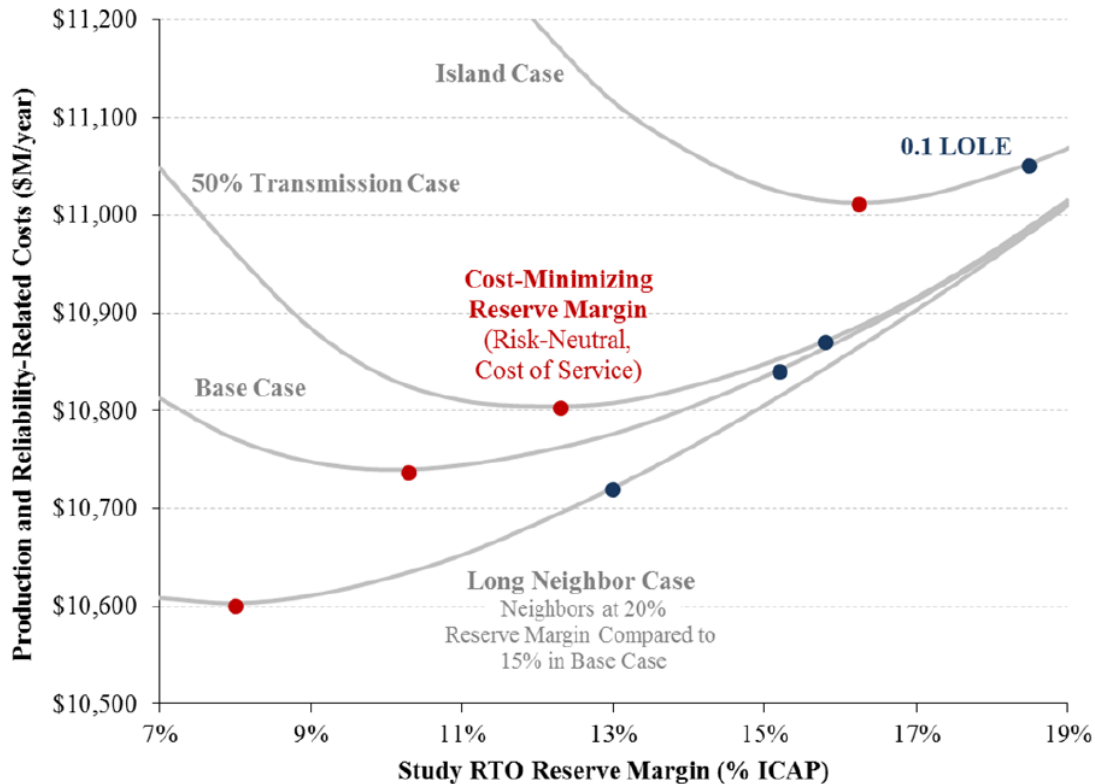
Figure 11. Total system cost versus reserve margin with varying intertie assumptions
Source: Pfeifenberger et al. 2013

These trade-offs are further shown in an analysis of the Western Interconnection in the United States (Ibanez and Milligan 2012), where the authors analyzed alternative wind/solar build-outs from *The Western Wind and Solar Integration Study Phase 2* (Lew et al. 2013). A reference case had 8% annual energy from wind and 3% from solar. Alternative cases had 33% of annual demand supplied by wind and solar split evenly, and high-wind/low-solar and high-solar/low-wind combinations. Ibanez and Milligan evaluated how much effective installed generator capacity could be replaced by transmission based on an assessment of resource adequacy. Key results are presented in Figure 12. The figure shows the reduction in effective capacity—the ELCC of the transmission additions and subsequent reduction in the need for resources—made possible by perfect transmission[7] within each subregion and by perfect transmission across the interconnection. Although copper sheet transmission is unlikely to ever be built, the example shows the trade-off between transmission and generation and the impact that transmission can potentially have on the need for new resource additions.

---

[7] Assuming perfect transmission—with no congestion or transmission constraints (i.e., perfect energy transfer between any two points in the system)—is often referred to as a "copper sheet" scenario.

18
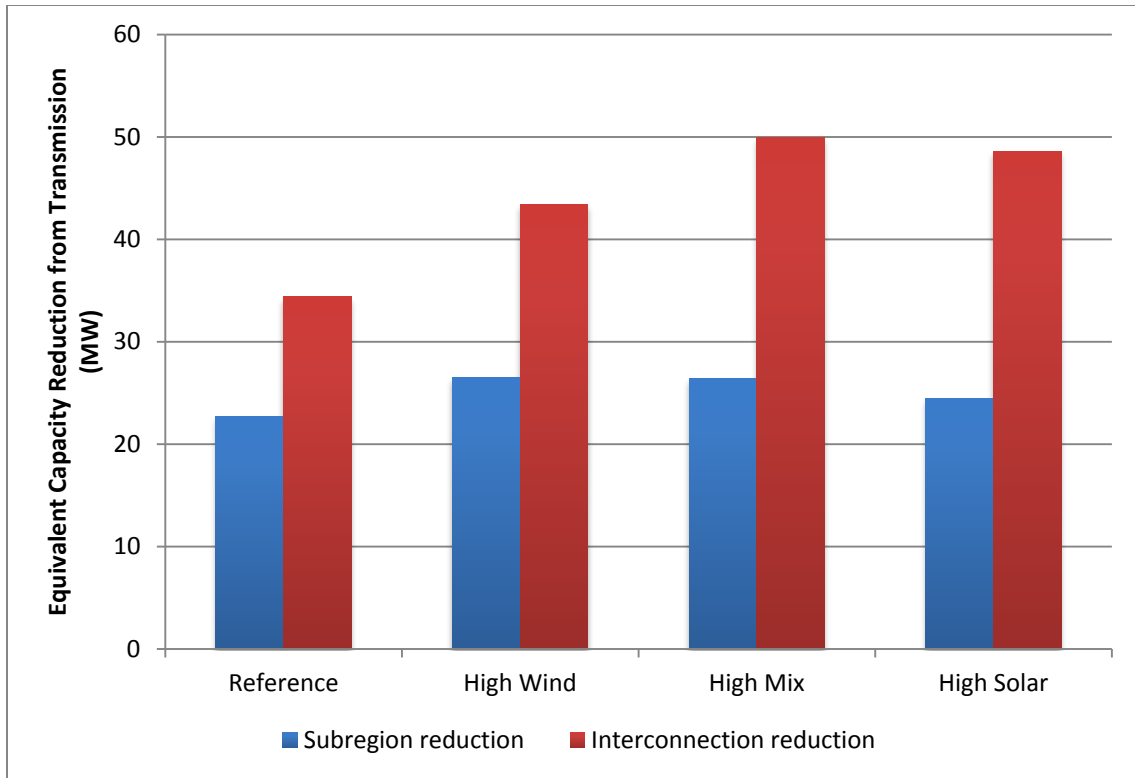
Figure 12. Impact of interconnection on resource adequacy in the western United States

Source: adapted from Ibanez and Milligan 2012

5.3 Fast Energy Market

Fast scheduling and dispatch is a key institutional flexibility mechanism that enables the system to access available physical flexibility that would otherwise remain locked within coarser operational time steps. This concept is summarized in Figure 13 for a case study of the Bonneville Power Authority (BPA) operating area in the Pacific Northwest of the United States. Moving from hour scheduling (top, blue line) to 10-minute scheduling (bottom, red line) results in a significant reduction of spinning and supplemental reserve requirements. This reduction affects both the frequency and magnitude of the reserve needs, because faster scheduling and dispatch can more accurately follow actual system conditions.

19
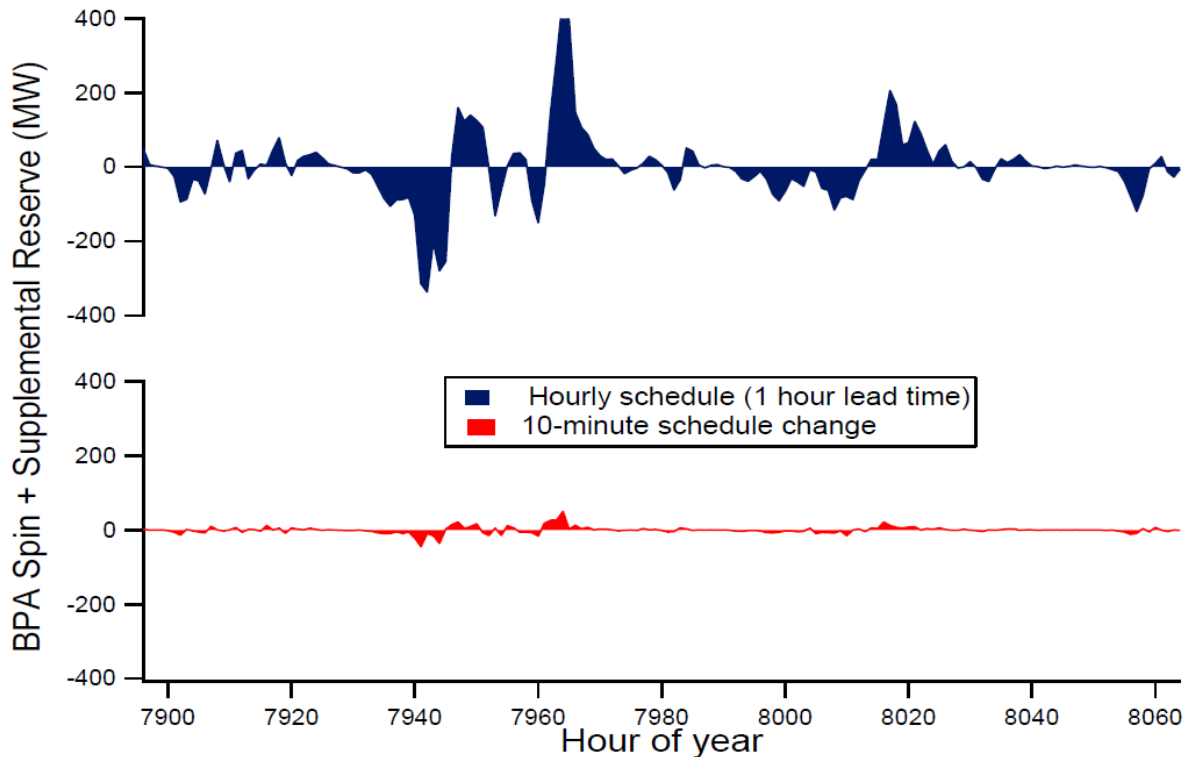
Figure 13. Dynamic/fast scheduling of wind out of BPA

Source: Milligan et al. 2011

Although all RTO markets in the United States operate a 5-minute economic dispatch, market settlements are not necessarily based on 5-minute prices. The temporal resolution of the settlement period determines the price that the generators are actually paid. In some cases, the settlement is carried out every 5 minutes at the 5-minute prices (locational marginal prices are used in all of the U.S. RTO/ISO markets). In other cases, settlement is carried out hourly and is based on an average of all 5-minute prices within the hour.

To explore the impact of settlement on flexibility incentives, Ela et al. (2014a) describe three different operating strategies and how the profit to suppliers differs for each strategy depending on whether an hourly or 5-minute settlement is implemented. These scenarios are shown in Figure 14. The dashed lines are different operating strategies: in the "5-min Sched" scenario, the supplier follows a schedule based on the 5-minute locational marginal prices; in the "Moving Hrly Avg Sched" scenario, the supplier follows an output based on the current hourly moving average locational marginal prices; and the "Perf Knowl Hrly Avg Sched" scenario is a hypothetical example if the supplier had perfect knowledge of only the final average hourly price. The graph shows the difference in flexibility that can be extracted by using the 5-minute settlement (5-minute price), which is significantly greater than hourly settlements (moving hourly average price).

The importance of this conclusion cannot be overemphasized. Frequent scheduling and shorter settlement intervals allow for better pricing of actual conditions and provides incentives for resources that can follow the prices. This example shows that flexibility can be muted by something as simple as the market settlement process in spite of what might otherwise be a good market design.

20
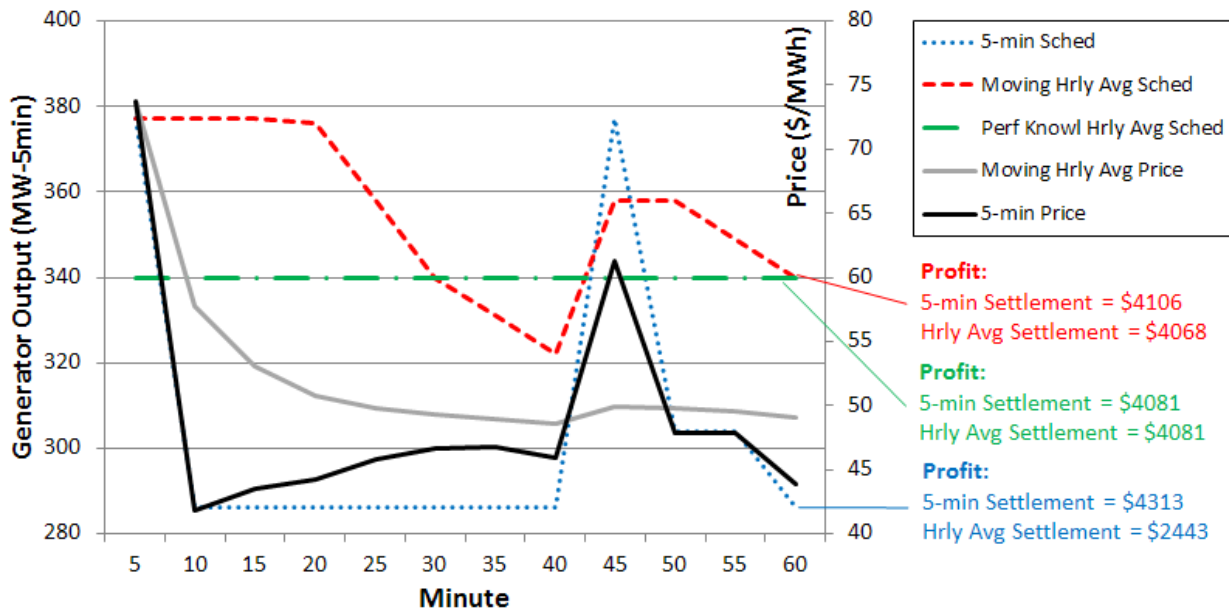
Figure 14. Profits of different operating scenarios with 5-minute settlements versus average hourly settlements

Source: adapted from Ela et al. 2014a

In the United States, different regions have different scheduling and dispatch time step sizes. All restructured markets (i.e., ISOs/RTOs) in the United States have a 5-minute economic dispatch with at least 15-minute interchange scheduling,[8] per Federal Energy Regulatory Commission (FERC) Order 764 as discussed below. Regulated markets have varying time intervals for dispatch and scheduling (e.g., hourly in much of the western United States). Many regions are moving to shorter time intervals to capture the flexibility and market efficiency benefits discussed above. No system has ever decided to go to a longer time interval dispatch. For example, the ERCOT region in the United States moved from a 15-minute to 5-minute dispatch (along with other market improvements) to improve system operation, resulting in a significant reduction of VG curtailment (Bird et al. 2014). Recent federal regulation is also assisting with the push for faster scheduling. FERC Order 764, issued in 2012, requires (among other details) transmission providers to allow customers the option of scheduling at 15-minute intervals instead of hourly to specifically assist with the integration of VG.

5.4 Improved Market Design

Improving market design is an institutional tool to greatly improve power system flexibility. Four major market principles have proved to work in the United States: large, fast, resource-neutral, and performance-based. The first two principles are discussed above. A resource-neutral principle means that all resources, regardless of technology, should compete to supply ancillary services on an equal footing, based on their different reliability contributions. A new potential ancillary service product could be defined by the required notification period, response speed, response depth, or length of performance—such specifications are technology-independent in the

---

[8] This refers only to scheduling with neighboring regions. In some areas, the conversion to a 15-minute interchange may not yet be complete.

21

market design. In addition, the payment should be performance-based to ensure consistency: two suppliers who provide different levels of service should be compensated accordingly—greater product provision should lead to higher payment; two suppliers of the same quantity of the same product, regardless of sources, should be compensated in the same way. This principle can also be applied to customers, demand, or demand-response. FERC Order 755 details the rationale for "resource-neutral" and "performance-based" payment and removed the previous "unjust, unreasonable, and unduly discriminatory or preferential" rates that resulted in economically inefficient dispatch of frequency regulation resources (FERC 2011).

Some market designs can have unintended consequences for flexibility. Two examples can be provided from the United States. The first example is the interaction of large energy markets with frequency response as alluded to above. During the past several years, frequency response in the U.S. Eastern Interconnection has been declining. Although there are likely multiple reasons that contribute to this decline, market design is likely contributing to the problem (see Ela et al. 2012a). The eastern energy markets provide for various penalties and/or costs if energy delivery schedules are not met by market participants. Yet if a generating unit responds to a frequency event, causing it to deviate from its energy delivery schedule, there is an economic penalty in many markets, causing the withdrawal of governor response and an overall decline in frequency response. This problem is not insurmountable; markets for frequency response could be designed and co-optimized in the same way that other ancillary service markets are included. For example, see Ela et al. (2014b; 2014c).

The second U.S. example results from the energy-only provisions with the current production tax credit (PTC) for wind energy. This subsidy provides a financial credit for each megawatt-hour generated by a wind plant that has qualified for the credit. However, if a wind plant were to provide the regulation ancillary services (which is still not allowed in many U.S. markets), it would result in a financial loss of the PTC. One possible way to remedy this disincentive would be to broaden the PTC to remunerate for the opportunity cost of providing regulation services. More detailed discussion of active power controls on wind turbines can be found on NREL's website.[9]

Other essential aspects of market design, such as the fundamental value of economic dispatch, the combination of faster markets with effective and widespread renewable energy forecasting, and nodal or zonal markets are discussed in the companion paper '*Renewables-Friendly' Grid Development Strategies* (Hurlbut et al. 2015).

5.5 Demand Response

Structuring markets to properly incentivize and utilize responsive load is a promising flexibility option that requires physical flexibility from responsive loads and the institutional structures to appropriately incentivize the desired response. Such DR serves as a reliability resource by reducing load during critical periods. DR has the potential to provide balancing capability on multiple time scales, ranging from seconds to seasons, by offering energy, capacity, and/or ancillary services (regulation, load-following, contingency). Market designs that emphasize performance requirements, such as notification period, time to start, time to run, etc., can often

---

[9] http://www.nrel.gov/electricity/transmission/active_power.html.

easily accommodate DR that is technically capable of providing the service. Incorporating DR into the economic dispatch allows for cost-effective utilization of DR.

Figure 15 summarizes how DR (and the related, but different, mechanism of energy efficiency) can adjust the load shape. These responses vary by speed, duration, magnitude, and frequency. Energy efficiency reduces total energy consumption (during the respective time of demand for related services such as lighting) through more efficient end use technologies, such as compact fluorescent or light-emitting diode lighting instead of incandescent light bulbs. Price response programs and peak shaving are energy services that shift load from a more system-constrained time period to a time period with more available supply; pre-cooling with air conditioning or pre-heating water are commons examples of this. Reliability and regulation responses are ancillary services that respond very quickly to deviations in scheduled net load (regulation) or loss of supply (contingency). DR can also serve as a capacity resource by being available to supply "generation" (i.e., negative load) during certain high load hours.

1. **Energy Efficiency** programs reduce electricity consumption and usually reduce peak demand

2. **Price Response** programs move consumption from day to night (real time pricing or time of use). **Peak Shaving** programs require more response during peak hours and focus on reducing peaks every high-load day

3. **Reliability Response** (contingency response) requires the fastest, shortest duration response. Response is only required during power system "events" – *this is slowly developing*

4. **Regulation Response** continuously follows the power system's minute-to-minute commands to balance the aggregate system – *this is just beginning to be done and may have the potential to dramatically change production costs, especially for aluminum*

Figure 15. Basic types of demand response with selected notes on their availability

Source: adapted from Milligan and Kirby 2010b

Programs that encourage responsive load modifications in the United States consist of two broad categories: price-based and incentive- or event-based mechanisms (Goldman et al. 2010). Price-based programs (also sometimes referred to as economic DR) motivate end-users to adjust their electricity usage by varying the price of electricity. Incentive- or event-based DR programs (also sometimes referred to as emergency DR) provide financial compensation to end-users for reducing their load upon request or for giving the program administrator direct control over certain consuming equipment (historically, this has primarily focused on air conditioners and water heaters). Event-driven programs usually have upper limits on the duration of individual events and total number of event-hours per year, often no more than 40 to 100 hours per year,

which corresponds to times when reserve margins drop below threshold conditions or when wholesale prices spike (Goldman et al. 2010). More utilities in the United States have historically offered price-based rates over incentive-based programs to customers, but it is unclear which category will become the preferred option from both the utility and customer perspectives. A recent study with 10 U.S. utilities revealed nearly double average peak load reduction with price-based DR programs than with incentive-based plans (21% vs. 11%), but this difference was largely eliminated when programmable communicating thermostats were installed (Cappers et al. 2015). Table 2 provides examples of programs within each of these two categories.

Table 2. Common Types of Demand Response Programs

Price Options	Incentive- or Event-Based Options
TOU rates: Rates with fixed price blocks that differ by time of day.[a]	**Direct load control:** Customers receive incentive payments for allowing the utility a degree of control over certain equipment.
CPP: Rates that include a pre-specified, extra-high rate that is triggered by the utility and is in effect for a limited number of hours.	**Demand bidding/buyback programs:** Customers offer bids to curtail load when wholesale market prices are high.
RTP: Rates that vary continually (typically hourly) in response to wholesale market prices.	**Emergency demand response programs:** Customers receive incentive payments for load reductions when needed to ensure reliability.
	Capacity market programs: Customers receive incentive payments for providing load reductions as substitutes for system capacity.
	Interruptible/curtailable: Customers receive a discounted rate for agreeing to reduce load on request.[b]
	Ancillary services market programs: Customers receive payments from a grid operator for committing to curtail load when needed to support operation of the electric grid (i.e., ancillary services).[c]

CPP = critical peak pricing; RTP = real-time pricing; TOU = time of use.

[a] Some analysts do not consider TOU rates to be a dynamic demand response option because the rating periods and prices are fixed, and utilities typically do not regard customers on TOU as a resource that can be dispatched similar to a generator when needed to support grid operations. A well-designed TOU rate, however, may induce customers to make long-term investments that reduce peak demands.

[b] Some utilities also regard interruptible tariffs as a "price-based" option, particularly if their interruptible tariff includes dynamic pricing provisions during emergency events (e.g., some tariffs give customers the option of "riding through" a curtailment event by paying higher real-time prices and still receiving electricity).

[c] Ancillary services demand response arrangements can also be viewed as a pricing program, because real-time pricing signals can be set up under a tariff to trigger event-specific customer behavior.

Source: Goldman et al. 2010

DR programs can be designed in many different ways. The best portfolio for a given system will depend on the system's existing physical and institutional structures and its expected load growth

24

and portfolio deployment. One example for business customers in one California utility, Pacific Gas and Electricity (PG&E), is shown in Table 3.

Table 3. Example Demand Response Programs for Business Customers in California

Program Name	Description (Partial)
Incentive- or Event-Based Demand Response	
Base Interruptible Program	Provides monthly or per-event incentives for curtailing load with either 30-minute or 4-hour advance notification. Curtailments limited to 120 hours per year.
Capacity Bidding Program	Offers monthly payments from May through October for curtailing nominated load on either a day-ahead or day-of basis, up to 24 times per month, when load conditions require the use of generators with heat rates of 15,000 Btu/kWh or greater.
Demand Bidding Program	Provides payments of $0.50–0.60 per kWh to customers who submit day-ahead or day-of offers to curtail load.
Optional Binding Mandatory Curtailment Plan	Allows customers to be exempt from rotating outages in return for agreeing to reduce load by 5 to 15 percent within 15 minutes of notification.
Peak Choice	Allows customers to customize their demand response by selecting from a range of advance notice, timing, load reduction, and number of day options, where incentives are determined by option combinations.
Scheduled Load Reduction Program	Pays $0.10 per kWh for commitments to reduce load one to three times per week, 4 hours at a time, from June through September.
SmartAC (Air Conditioner Load Control)	Offers $25 to customers who allow installation of either an air conditioner compressor switch or a smart thermostat that can be used to control load up to 100 hours per year when CAISO declares emergency or near-emergency conditions. This program is also offered to residential customers.
Price-Based Demand Response	
CPP	Provides lower rates during on-peak and partial-peak hours in exchange for higher (three to five times normal) rates on up to 12 days between May and October when system demand is high.
Demand Response Technical Assistance	
Technical Assistance and Technology Incentive Programs	Provide engineering assistance and cash incentives to support installation of equipment or software supporting demand response.
Integrated Energy Audit	Offer audits that comprehensively address opportunities in energy efficiency, time-of-use management, demand response, self-generation, and renewables.

Sources: <http://www.pge.com/mybusiness/energysavingsrebates/demandresponse/>;
<http://www.pge.com/mybusiness/energysavingsrebates/analyzer/integrated/>.

Btu = British thermal unit; CAISO = California ISO; CPP = critical peak pricing; kWh = kilowatt-hour; TOU = time of use.

Source: Goldman et al. 2010

In the PJM region, DR resources perform and are paid like traditional "supply-side" resources (generation). PJM currently allows responsive load to participate, just as a generating unit, in its forward capacity market (peak shaving), "price responsive demand" (as negative load) in the day-ahead and real-time energy markets, and frequency regulation and synchronized reserve in the ancillary services markets (PJM 2014). The energy DR programs are further divided into

emergency (voluntary load shedding) and economic (dispatched load shedding) categories. Capacity payments make up the vast majority of all DR revenue in PJM (Figure 16).
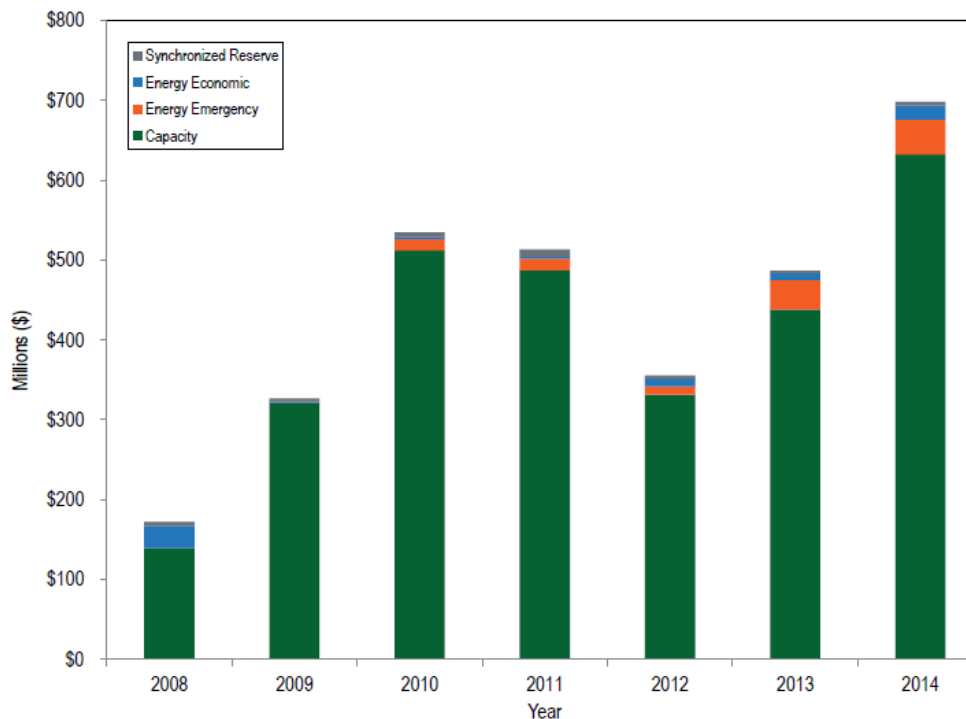
Figure 16. PJM demand response revenue by market, 2008–2014

Source: Monitoring Analytics 2015

However, the exact way in which DR market structures will evolve in the United States is highly uncertain because of current legal issues (EPSA v. FERC 2014) with how DR may or may not compete with generating units. This is largely an issue of revenue sufficiency for existing generators: DR suppresses capacity prices and the frequency of energy and ancillary services scarcity events, which reduces the revenue received by generators to cover both their fixed and variable costs. The PJM region, for example, is establishing alternative approaches for how DR could be implemented to both retain system reliability and meet all legal obligations. These approaches would treat DR as a demand-side resource, where the compensation would shift from payment for energy or capacity (supply-side) to one of avoided costs from avoided energy or capacity (demand-side) (PJM 2014).

A key reliability and economic benefit of these DR programs is a reduction in peak load, which corresponds to a reduction in required system capacity and associated costs. It is cheaper to turn down load than to build new generating capacity. The load duration curve for CAISO shown in Figure 17 illustrates the relationship between system peak capacity requirements and the frequency and size of DR events needed to achieve a given level of peak capacity reduction. For this case, a 5% reduction in peak load would require 20 hours per year of well-targeted DR. Similar observations were noted for the Public Service Company BA in Colorado, with about 3.5% of peak load reduction potential in only 12 hours of the year (Denholm et al. 2015). The 40 to 100 hours per year of individual DR events previously listed is therefore not a strict rule-of-thumb, but depends on the individual system's desired reliability and peak reduction levels.

26
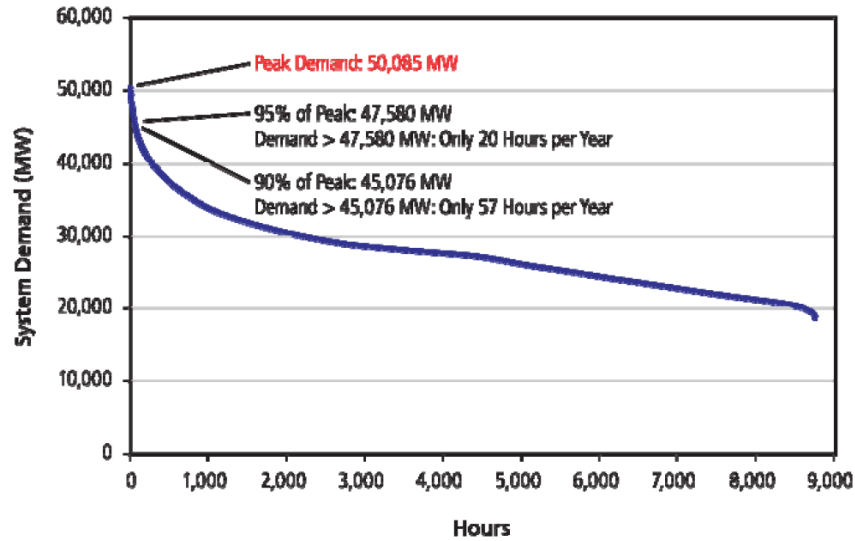
Figure 17. CAISO opportunity for reliability-based demand response

Source: Goldman et al. 2010

Another way for DR programs to benefit the system is through ancillary services, such as contributing to the regulation reserve requirement. Figure 18 shows how the aggregate of numerous responsive load resources can meet the regulation reserve requirement (an ancillary service) in response to the power system operator's automatic generation control signals. This results in an energy-neutral balancing of the minute-to-minute net load deviations. The red line shown on the graph represents the regulation signal sent by the ISO to the DR system server; in this case, the DR system server is Enbala Power Network's real-time control optimization engine "R3OE." Enbala is a DR aggregator for industrial, commercial, and institutional end-users. The green line represents the response of all resources and thus the amount of regulation being provided. Any individual responsive load (each line at the top of the figure) has a very small contribution to the overall regulation requirement, but in aggregate, they can provide a significant and effective resource.

27
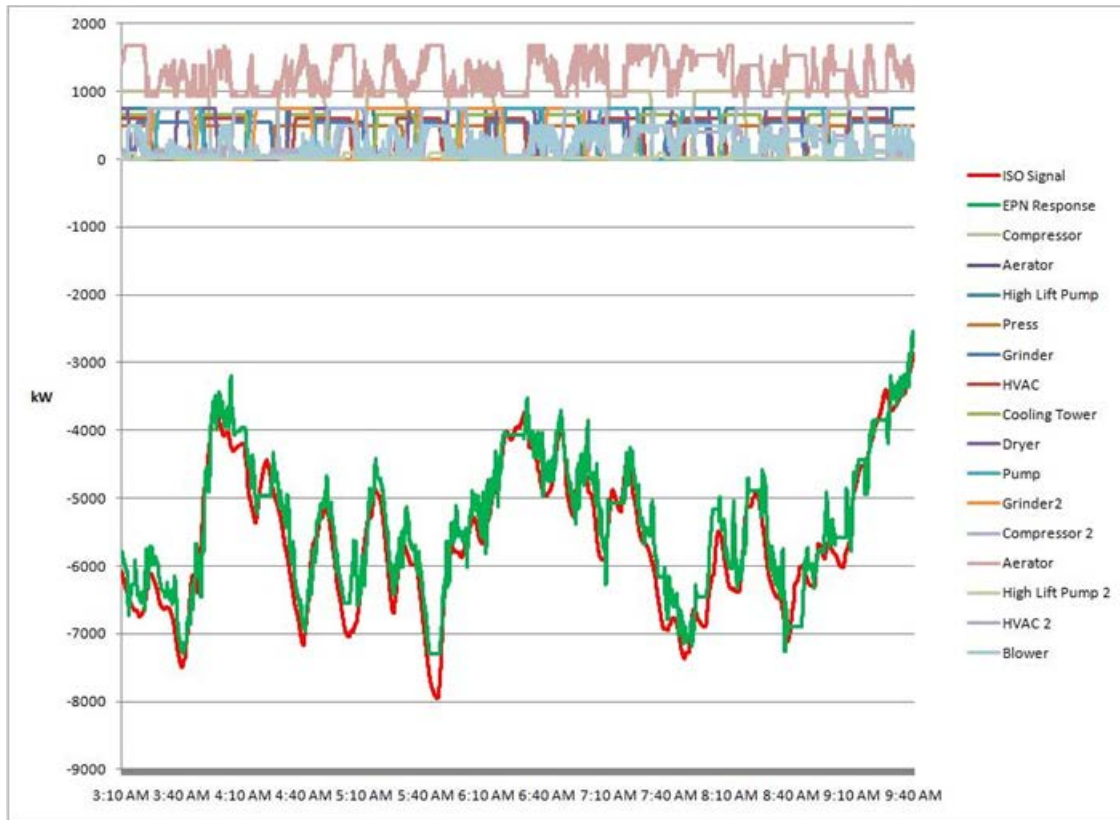
Figure 18. Aggregate demand-side response can meet regulation signal
Source: Milligan 2014b

Hummon et al. (2013) laid out the methods for quantifying these benefits of DR, especially during times when the system is most constrained, in an analysis of DR in the Colorado test system. The study quantified the value of DR to provide energy, capacity, and operating reserves, based on production cost model simulations with aggregate DR resource profiles from Olsen et al. (2013). Results of the study are summarized in the Appendix. The value of DR can be expressed as an annual availability factor[10] (similar to a generator capacity factor) or as the cumulative availability across the year[11] (dollars per megawatt-hour) (Hummon et al. 2013). This value is based on the amount of available resource and its coincidence with times of high production costs, which are generally reflective of times when the system is in greatest need of flexibility.

5.6 Strategic Renewable Energy Curtailment

Curtailment of VG provides flexibility through the physical reduction in supply when more generation is available than the system can utilize. Institutional mechanisms to incentivize and manage that curtailment are critical to accessing this flexibility. The presence of curtailment is an

---

[10] The annual availability factor is the sum of the maximum capacity available during each time period divided by the peak available capacity times the number of time intervals (i.e., total hours per year).

[11] The cumulative annual availability is the fraction of electricity from an aggregation of end uses that is flexible through DR, which reflects the correlation of each resource's availability to times of high market prices for operating reserves as well as its ability to take advantage of large energy price differences across hours of the day.

indicator of inflexibility in the system, typically resulting because of transmission congestion or lack of transmission access, or less often due to excess generation during low load periods. Usually the most expensive VG plant is curtailed first to alleviate system congestion or to maintain system balance. The economic compensation provided to the VG plant varies and depends on the specifics of the power contract (Bird et al. 2014). In the United States, this economic trade-off is further complicated for projects that utilize production-based subsidies. Some of this curtailed VG can be used for other important grid services, such as regulation up reserves[12], frequency, and ramping services. The economic choice to curtail VG at any given moment reflects the trade-off between the instantaneous value of the energy produced and the value of these other services.

Curtailment of VG is achieved either through manual directives by the system operator or through automated market-based mechanisms, such as special economic dispatch protocols. The first curtailment method is command and control. During over-generation conditions, the system operator can make a decision to curtail VG regardless of whether it is the optimal or economic choice. The level of curtailment ordered by the operator may be too high or too low and may respond too quickly. The second curtailment method is economic curtailment. This incorporates VG generation (and any subsequent over-generation) directly into the economic dispatch process, which (under normal conditions) requires no direct operator intervention and can be done very quickly and cost effectively. In the United States, this method is working today in NYISO and MISO and has attracted significant interest by other regions. Because the use of wind energy is more widespread than solar energy in the United States at the time of this writing, some of the market changes have specified wind energy and not solar energy. In the near future, these approaches may also apply to solar energy. VG resources are bid into the market like conventional generation. The difference is that usually wind will only bid into the dispatch-down market[13]. If this is economical, then that wind will be dispatched downward to help avoid over-generation. As carried out by the real-time market, this can be an optimal solution because (a) only the needed level of dispatch is activated, and (b) the dispatch of wind will only occur if it is the least-cost option.

In 2011, MISO implemented an economic curtailment program called the Dispatchable Intermittent Resource (DIR) protocol, which effectively places the 5-minute dispatch optimization of wind power plants on automatic generation control under command of the MISO real-time market systems. The result is an overall reduction of curtailed VG with a much higher level of operational efficiency and transparency. MISO manages more than 14 GW of wind in its market footprint and has been experiencing local transmission congestion issues during certain periods that were traditionally managed by manual curtailments of specific wind plants. Figure 19 shows monthly total DIR-dispatched wind and the percent that was dispatched downward

---

[12] Regulation up service is an ancillary service that provides capacity that can start responding to signals within a set time period (5 seconds in ERCOT). Such capacity is the amount available above any base point but below the high sustainable limit of a generation resource and may be called on to change output as necessary throughout the range of capacity available to maintain proper system frequency. This typically involves setting aside a portion of output for frequency that would otherwise be used for energy (or that is otherwise curtailed). See more details in ERCOT (2014).

[13] A dispatch down market is a market in which, at times of excessive electricity generation, eligible generators would be offered payments to reduce their output, compensating the unit for its opportunity cost in the energy market.

29

(below economic maximum) under this DIR protocol. According to MISO, approximately 95% of wind energy's potential can be captured through economic dispatch. All new wind generation facilities in MISO must register as DIRs, and more than 80% of wind generation in MISO is dispatchable.
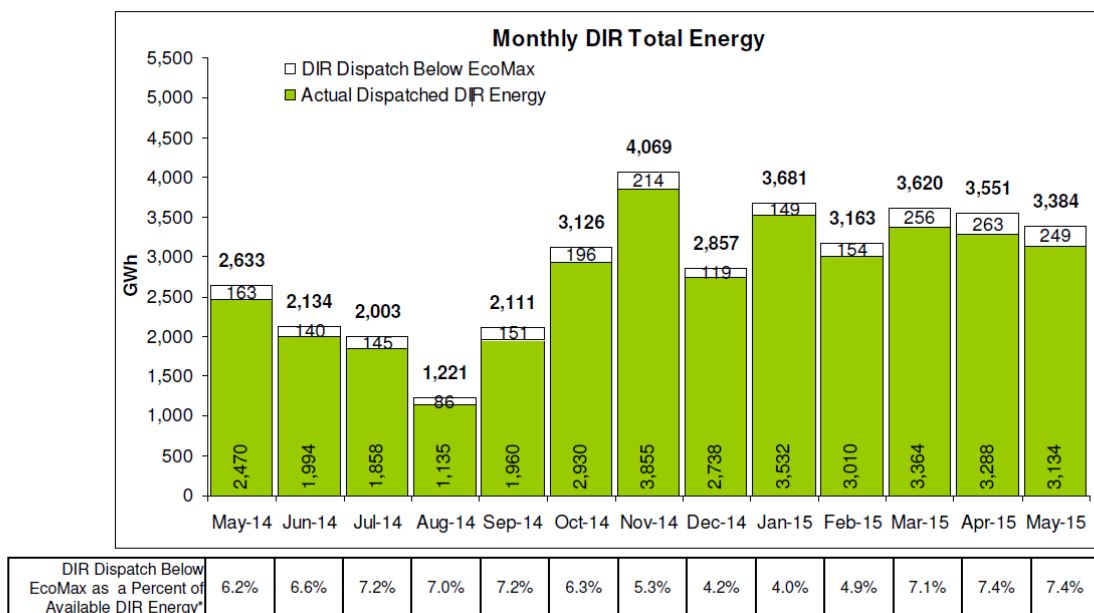
Figure 19. Wind power plant dispatch in MISO with DIR protocol

Source: MISO 2015

5.7 New Ancillary Services Products

Some market areas in the United States are now investigating the effectiveness of incentivizing energy flexibility as a new ancillary service product, commonly referred to as "FlexiRamp." Such ancillary service products are institutional flexibility measures. Since 2011, CAISO has included a flexibility constraint in its market-clearing engine that ensures that sufficient ramping capacity is committed and available in the real-time commitment and real-time dispatch process (CAISO 2011, Abdul-Rahman et al. 2012). The amount of ramp capability that is required in this constraint is determined by the CAISO operators based on (1) the expected level of variability for the interval, (2) the potential uncertainty as a result of load and VG forecast error, and (3) the differences between the hourly, 15-minute average net load levels and the actual 5-minute net load levels. Units that incur a lost opportunity cost by withholding their capacity from other ancillary services in order to meet this ramping constraint are compensated at an amount equal to the system's incremental cost of increasing the ramping need by one unit. Currently, the constraint is only for upward ramp capability needs. However, CAISO is now proposing to extend this market product to include downward ramping, use the 5-minute real-time dispatch interval rather than the 15-minute real-time pre-dispatch model, include the product in the day-ahead market, and implement a flexibility demand curve to account for both variability (known ramps) and uncertainty (unforeseen ramps).

NREL analysis compares this proposed FlexiRamp product to a look-ahead dispatch with and without freezing the advisory prices. Figure 20 summarizes the financial results for a test system

30

with two generators: G1 is very flexible (100 MW/min ramp rate capability), while G2 is less flexible (10 MW/min ramp rate capability). The positive cash flows (blue bars) are the revenue, negative cash flows (red striped bars) are the costs, and the resulting net profit or loss (diamonds) are shown within each bar, with the value noted above each set of bars. Results with the proposed FlexiRamp product for a single-period dispatch are shown on the left. Results for a multi-period (look-ahead) economic dispatch without the FlexiRamp product are shown in the middle and right sets of bars: the middle set updates the settlement price at each time interval from the look-ahead advisory price, but the right set uses the prices of those future (advisory) time intervals as the final (binding) prices paid to the generators. Results from this analysis reveal that flexible ramping constraints, such as FlexiRamp, perform similar functions as security-constrained economic dispatch from an *operational* standpoint but not from a *pricing/incentive* standpoint. Units that provide reserves for a future (advisory) time interval may not get paid for that service if the binding price is not set to the advisory value, thus removing the incentive of the unit to participate in this market. Thus, other institutional mechanisms, such as utilizing advisory intervals as pricing for the binding interval, can achieve similar outcomes as a flexible ramping product.
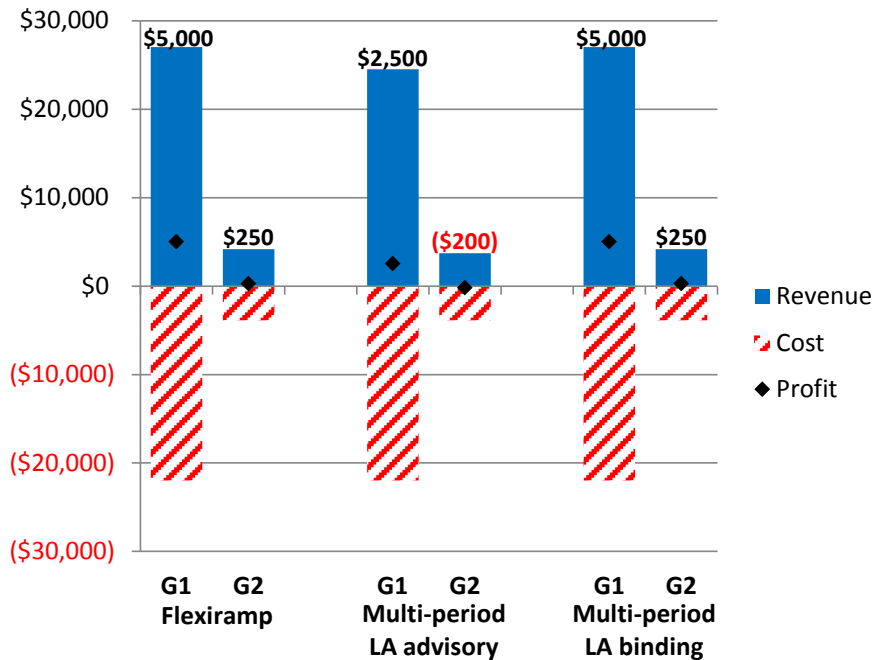
Figure 20. Comparison of CAISO FlexiRamp ancillary service product for a flexible generator (G1) and less flexible generator (G2) for a single period with FlexiRamp, multiple period for current time step (binding), and multiple period for next time step (advisory).

LA = look ahead. Source: Milligan 2014a

5.8 Flexible Conventional Generation Units

Flexible conventional generators, such as coal and natural gas combustion turbines with the ability to cycle on and off and run at lower output in order to follow changes in output from VG, are another important source of physical flexibility. However, the economic viability of these traditionally inflexible plants, particularly coal, which was intended to run at annual capacity factors of around 80%, is uncertain under such operating conditions (Cochran et al. 2013). This

31

is because hardware and extensive operational modifications are required to make these plants flexible (minimum run conditions below 40% of capacity). In China, where coal-fired power plants contribute to over 70% of total electricity generation, a cost-benefit analysis would be of particular importance to determine if the most economical option for these plants is to (a) retrofit, (b) retire, or (c) change operational processes to enhance flexibility. Another option would be to replace or supplement these less flexible generators with fast-start reciprocating engines and combustion turbines with lower minimum loads and higher efficiencies, which can be started within minutes without incurring startup costs and can help provide ramping and non-spinning reserves.

Practical experience in North America has revealed that it is possible to modify a traditionally designed coal plant that was intended to run only at baseload into one that can meet peak demands, cycling on and off up to four times a day to meet morning and afternoon electricity demand (Cochran et al. 2013). Key to this specific success was changing operational practices, as well as inherent design features that facilitate cycling. The main operational changes included monitoring and managing temperature ramp rates, creating a suite of inspection programs for all affected equipment (large and small), and continual training to reinforce the skills needed in monitoring and inspections.

The more frequent start and turndown, or cycling, of traditionally inflexible plants to achieve a higher level of flexibility causes equipment damage, reduces the life expectancy, and impacts emissions. These were modeled and assessed in a comprehensive study of the western United States, which found that up to 33% wind and solar energy led to relatively small cycling impacts on overall emissions and production costs (Lew et al. 2013). Figure 21 summarizes how the additional costs from cycling are significantly outweighed by avoided fuel costs from renewables. However, from the perspective of individual thermal generators, increased cycling costs and lower utilization due to energy displaced by VG may erode profitability. The generator may need higher prices in the ancillary service or capacity markets or other institutional structures to stay financially viable. Market designs and other approaches to incentivize flexibility can help with the transition to a high-renewables future by incentivizing the generation and DR characteristics that are beneficial. Whether this future includes large amounts of retro-fit coal units or alternatives will therefore be a function of the most cost-effective technologies that can provide the needed services under the market structure.
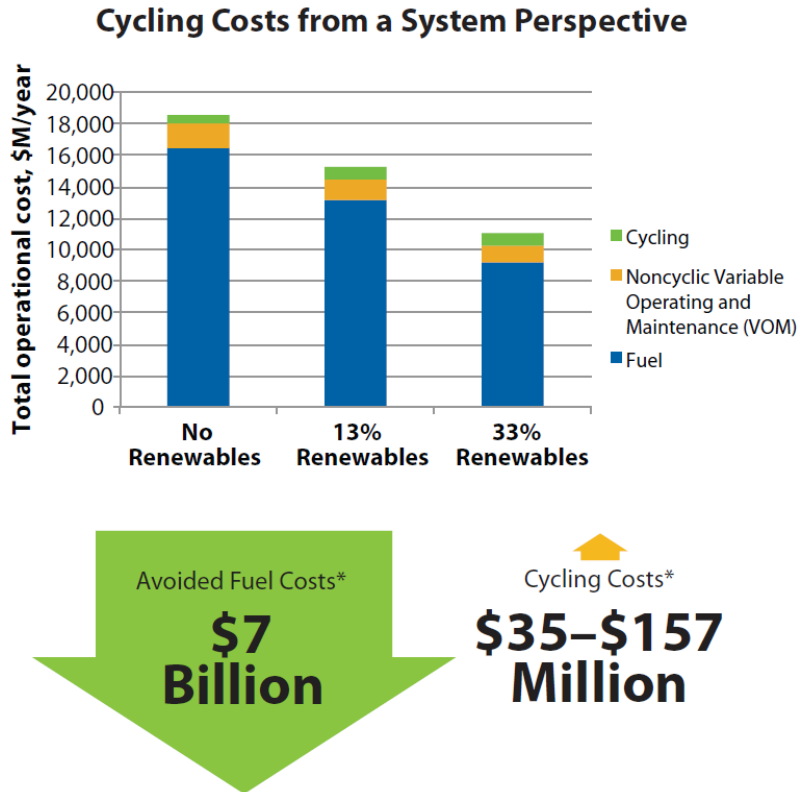
Figure 21. Cycling costs, though can be significant for the plant owner, have a small impact on the overall system operating cost savings due to the integration of renewables

Source: Lew et al. 2013

5.9 Storage

Storage is a physical flexibility option that allows energy produced in one time step to be used, minus efficiency losses, at a later time. Similar to DR, storage can provide firm capacity, energy shifting, and ancillary service benefits. These benefits are determined primarily by the discharge time, as shown in Figure 22. Storage technologies that can respond to changes in demand on short time scales, from minutes to fractions of seconds, are better adapted for power management (bottom left section in this figure). These rapidly responding technologies provide transient stability, frequency regulation, and other ancillary services to maintain voltage and frequency levels within prescribed bounds. Technologies in this regime include flywheels, super-capacitors, and a variety of batteries, which often have smaller capacities. Storage technologies that are better adapted for energy management (upper right section in this figure) provide continuous discharge for extended periods of time to balance changes in load over longer time scales, from days to weeks to seasons. These larger capacity technologies offer firm capacity and energy shifting (arbitrage) services and include pumped hydroelectric storage and compressed air energy storage. Figure 22 does not show thermal energy storage, which would cover a power range from 1 kW (building) to more than 100 MW (concentrating solar power plants), with a discharge time of minutes to several hours (Denholm et al. 2010). An additional cost comparison of select storage technologies is shown in Figure 4.
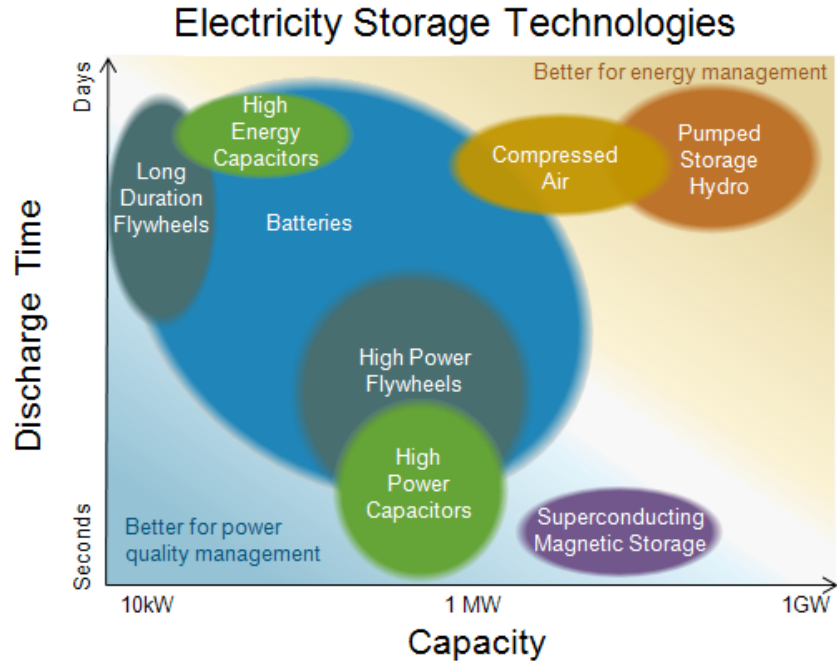
33

Figure 22. Energy storage technologies and ratings

Source: EIA 2011

The value of storage to the system depends on multiple factors, and—like DR—quantifying this value has been a key challenge for developers in the United States, especially because some benefits are not fully realizable in many markets. Simple metrics, such as levelized cost of energy, are not sufficient; detailed time-series analysis using software tools to co-optimize multiple services provided by different storage technologies are required to properly value storage (Denholm et al. 2013).

The effectiveness of storage depends on many factors, including physical characteristics of the storage technology used (discharge time, efficiency, cost, and storage size on both an energy and capacity basis) and characteristics of the power system (VG penetration, portfolio of VG technologies, amount of existing flexibility). Figure 23 shows one example[14] of how the value of storage (blue, left axis) increases as the penetration of solar generation increases, but this value is diminished when the existing system is more flexible (right versus left plot). Additionally, curtailment of solar generation increases as the penetration of solar increases (right axis), but this curtailment is dampened when storage is present (red versus black lines).

---

[14] This example is based on the CAISO system with a high diversity of solar generators across the southwest United States. The storage resource assumes 10% of peak capacity with 1 week of storage. The current resource mix case represents a medium-term system that contains about 80% incumbent generators and assumes some load growth and generator retirements. The flexible resource mix case represents a long-term system that has reached a generation equilibrium.

34
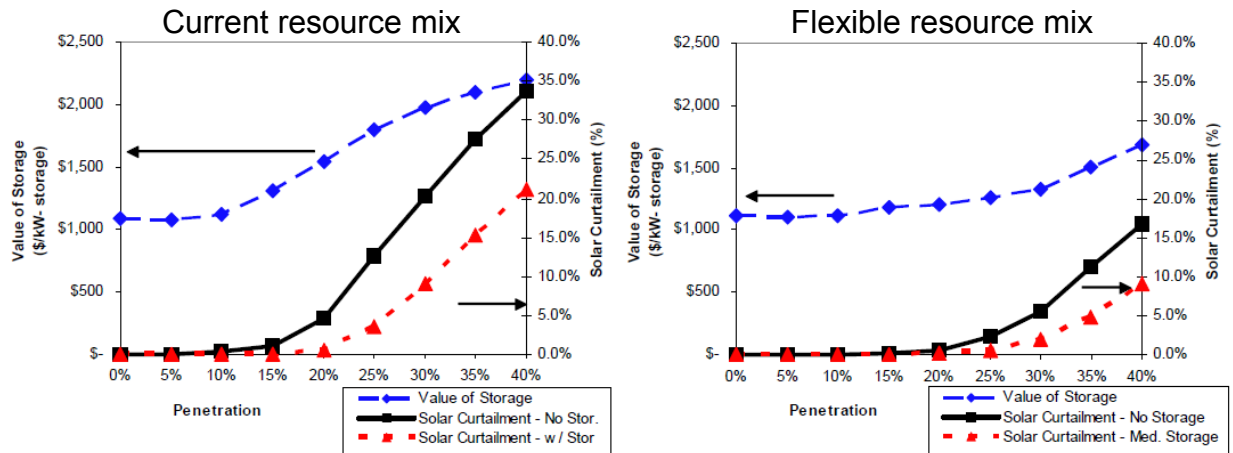
Figure 23. Value of flexibility (shown here for solar) depends on system flexibility and VG penetration

Source: Mills 2008

In the United States, the value of energy storage in restructured markets has historically included energy shifting (arbitrage) and ancillary services (regulation and contingency reserves). Table 4 summarizes these values. In these markets, energy and ancillary services are co-optimized so that the least-cost mix of generators is chosen to provide the necessary energy and ancillary services. Thus, the use of energy storage technologies previously mentioned and shown in Figure 23 (e.g., pumped hydro, batteries, etc.) must compete with other sources of flexibility.

35

Table 4. Historical Values of Energy Storage in U.S. Restructured Electricity Markets

Market Evaluated	Location	Years Evaluated	Annual Value ($/kW)	Assumptions
Energy Arbitrage	PJM[a]	2002–2007	$60–$115	12 hour, 80% efficient device. Range of efficiencies and sizes evaluated. Also considers price difference suppression effect in a market setting using price/load relationships.
	NYISO[b]	2001–2005	$87–$240 (NYC) $29-$84 (rest)	10 hour, 83% efficient device. Range of efficiencies and sizes evaluated.
	USA[c]	1997–2001	$37–$45	80% efficient device. Evaluates ISO-NE, CAISO, PJM
	CA[d]	2003	$49	10 hour, 90% efficient device.
	CA[f]	2010–2011	$25–$41	4 hour, 90% efficient device.
	CA[h]	2011	$46	16 hour, 75% efficient pumped storage device.
Regulation Reserves	NYISO[b]	2001–2005	$163–$248	
	USA[e]	2003–2006	$236–$$429	PJM, NYISO, ERCOT, ISO-NE.
	CA[f]	2010–2011	$117–$161	Co-optimized arbitrage and regulation, most value is derived from regulation.
Contingency Reserves	USA[e]	2004–2005	$66–$149	PJM, NYISO, ERCOT, ISO-NE.
Combined Services	CA[f]	2010–2011	$117–$161	Arbitrage and regulation, most value is derived from regulation.
	CA[h]	2011	$62–$75	Arbitrage, regulation, and contingency. Included operational constraints of pumped storage.
	USA[g]	2002–2010	$38–$180	Arbitrage and contingency. CAISO, PJM, NYISO, MISO.

[a] Sioshansi et al. 2009
[b] Walawalkar et al. 2007
[c] Figueiredo et al. 2006
[d] Eyer et al. 2004
[e] Denholm and Letendre 2007
[f] Byrne and Silva-Monroy 2012
[g] Drury et al. 2011
[h] Kirby 2012
Source: Denholm et al. 2013

6 Cost-Effectiveness Evaluation of the Options

Multiple approaches to obtain flexibility may be available, in which case it is usually desirable to perform an analysis to compare the costs and benefits of these approaches. In some cases, institutional improvements are most cost-effective because they provide access to flexibility that currently exists, but is otherwise not accessible to the power system operator. Improvements such as large BAs and fast economic dispatch have value even in systems with no VG, and this value can be significantly enhanced with large VG levels. Once implemented, these features will have a long useful life, and therefore deliver benefits over the very long term. Similarly, new transmission interconnections, when coupled with efficient means to pool the operating requirements of two or more regions, have a long life and can therefore provide lasting benefits.

Other flexibility improvements may be more short-lived or may perhaps depend on external factors such as fuel prices. As a general rule, alternative flexibility options can be analyzed in the context of production simulation modeling. The value of each alternative can be calculated by adding it to the mix, adjusting other modeling parameters accordingly, and calculating the production cost savings. This can be done over the long-term, or suitable estimates of the long-term value based on short-term evaluations can be made (but are not as robust). Sensitivities to external factors such as fuel prices, alternative resource mixes, different rates of demand growth or composition, and many others may change the value of any mitigation option.

The value of a given flexibility option will also depend upon what other options have already been adopted. This is illustrated in Figure 10. Using the example data in the diagram, suppose the flexibility option that is under evaluation is the size of the BA. First, let us suppose that all of the regions are operating with a 60-minute dispatch with a 40-minute notification period. The reduction in regulating needs can be estimated as the difference between the green bars in the "small" grouping and in the "large" grouping: approximately 5,000 MW. Alternatively, suppose that all of the regions are already operating at a 10-minute dispatch with a 10-minute notification period (blue bars in the figure). In this case, the reduction in regulating needs is approximately 1,400 MW. Thus, the benefit of a large BA depends on (among other things) the dispatch time step. Because of the non-linear nature of the power system, it is likely that the value of any given source of flexibility depends on other sources that have already been, or have been assumed to have been, adopted.[15]

---

[15] A similar argument can be made based upon Figure 23, which compares the value of storage in an inflexible system as compared to an already-flexible system.

7 Other Planning Considerations/Issues

Even in regions with large electricity markets, including the United States and Europe, the markets are regulated to ensure that the well-known shortcomings of electricity markets can be mitigated. Ela et al. (2014a) provide a very detailed examination of wholesale electricity market characteristics and impacts, showing some key areas that have not been well thought out. Established text books on power system economics (e.g., Stoft 2002) provide broader frameworks that do not include the consideration of renewables. Therefore, the interplay between the regulatory environment and any market mechanisms must be thoroughly analyzed. For example, in parts of the United States and Europe, there is increasing concern regarding the ability of the energy and ancillary service markets to sufficiently incentivize the development of new capacity that has the needed flexibility characteristics. The extent to which this is a problem is not yet well-known, and possible solutions include several versions of forward capacity markets (with or without flexibility requirements, or tranches). The market structures provide the enabling environment for ensuring both resource adequacy and sufficient flexibility, alongside the appropriate reimbursement mechanisms to incentivize needed attributes of both current and future resources.

8 Conclusions

Worldwide experience with integrating renewables, along with results from high-quality wind/solar integration studies (see Milligan et al. 2015 for a U.S.-based summary) show that integration can be done effectively if the system operator has sufficient tools to manage the increase in variability and uncertainty that will occur with VG. Flexibility is important, and multiple approaches exist for measuring the need for flexibility and the extent to which it is available. Similarly, multiple approaches may be considered to ensure sufficient flexible capacity will be available in the future when it would be needed. Long-term resource adequacy and flexibility are critical considerations for system planning. Multiple approaches and combinations of physical and institutional components, including forward capacity markets, ancillary services markets and other market constructs have proven effective. Assessments of future flexibility needs and evaluation of various market (and technical) constructs utilizing state of the art approaches to security-constrained economic dispatch analysis, combined with sophisticated reliability analyses offer important insights into the discussions for both physical and institutional market developments. Unintended consequences of market design, such as the misalignment of the incentives for generators to provide frequency response in energy markets (Ela et al. 2012a), should be carefully considered and mitigated when they are identified.

Market solutions for economic dispatch and for the acquisition of ancillary services have been shown to be effective in many parts of the world. The key factors that energy markets provide include (1) large BAs, (2) fast energy markets, and (3) incentives for generators to operate efficiently. Markets are not perfect, however, and in the United States and elsewhere there is a significant body of literature that supports the public regulation of electricity markets. The path forward for China may include the development of energy and ancillary service markets, but it may instead comprise other means of providing system flexibility to deliver reliable, affordable power that accommodates the renewable energy goals of the country. Options to consider include a market-like economic dispatch run by State Grid of China Corporation and China Southern Power Grid, some form of EIM,[16] or a full RTO-like structure.

---

[16] We note that the EIM as implemented in the United States does not include coordination of unit commitment, which likely results in economic inefficiencies. However, the EIM may be a good first step that results in jointly optimizing dispatch, setting the stage for future coordination of unit commitment that could result in additional cost savings and efficiencies (e.g., E3 2015).

References

Abdul-Rahman, K., H. Alarian, M. Rothleder, P. Ristanovic, B. Vesovic, and B. Lu. 2012. "Enhanced System Reliability Using Flexible Ramp Constraint in CAISO Market." *IEEE Power and Energy Society General Meeting Proceedings*, July 22–26, 2012, San Diego, California.

Bird, L., J. Cochran, and X. Wang. 2014. *Wind and Solar Energy Curtailment: Experience and Practices in the United States*. NREL/TP-6A20-60983. Golden, CO: National Renewable Energy Laboratory.

Byrne, R.H. and C.A. Silva-Monroy. 2012. *Estimating the Maximum Potential Revenue for Grid Connected Electricity Storage: Arbitrage and Regulation*. SAND2012-3863. Albuquerque: Sandia National Laboratories.

CAISO (California Independent System Operator Corporation). 2011. "Opportunity Cost of Flexible Ramping Constraint: Draft Final Proposal." Folsom, CA: CAISO. Accessed May 2014. http://www.caiso.com/Documents/DraftFinalProposal-FlexibleRampingConstraint.pdf .

———. 2014. "Flexible Resource Adequacy Criteria and Must-Offer Obligation: Market and Infrastructure Policy Draft Final Proposal." Folsom, CA: CAISO.

———. 2015a. "Reliability Services Initiative – Phase 2 and Flexible Resource Adequacy Criteria and Must Offer Obligation – Phase 2: Issue Paper." Folsom, CA: CAISO.

———. 2015b. "Before the Public Utilities Commission of the State of California: California Independent System Operator Corporation Deterministic Studies." Folsom, CA: CAISO.

Cappers, P., L. Hans, and R. Scheer. 2015. *American Recovery and Reinvestment Act of 2009: Interim Report on Customer Acceptance, Retention, and Response to Time-Based Rates from the Consumer Behavior Studies*. LBNL-183029. Berkeley, CA: Lawrence Berkeley National Laboratory. Accessed September 2015. http://emp.lbl.gov/sites/all/files/lbnl-183029.pdf.

Cochran, J., D. Lew, and N. Kumar. 2013. "Flexible Coal: Evolution from Baseload to Peaking Plant." NREL/BR-6A20-60575. Golden, CO: National Renewable Energy Laboratory. Accessed June 2015. http://www.nrel.gov/docs/fy14osti/60575.pdf .

Cochran, J., M. Miller, O. Zinaman, M. Milligan, D. Arent, B. Palmintier, M. O'Malley, S. Mueller, E. Lannoye, A. Tuohy, B. Kujala, M. Sommer, H. Holttinen, J. Kiviluoma, and S.K. Soonee. 2014. *Flexibility in 21st Century Power Systems. 21st Century Power Partnership*. NREL/TP-6A20-61721. Golden, CO: National Renewable Energy Laboratory.

CPUC (California Public Utilities Commission). 2014. "Collaborative Review of Planning Models." San Francisco: CPUC.

Denholm, P. and J. Cochran. 2015. "Balancing Area Coordination: Efficiency Integrating Renewable Energy into the Grid." NREL/FS-6A20-63037. Golden, CO: National Renewable Energy Laboratory.

Denholm, P. and M. Hand. 2011. "Grid Flexibility and Storage Required to achieve Very High Penetration of Variable Renewable Electricity." *Energy Policy*, 39:1817–1830.

Denholm, P. and S.E. Letendre. 2007. "Grid Services From Plug-in Hybrid Electric Vehicles: A Key to Economic Viability?" *Electrical Energy Storage – Applications and Technology Conference*, September 25, 2007.

Denholm, P., E. Ela, B. Kirby, and M. Milligan. 2010. *The Role of Energy Storage with Renewable Electricity Generation*. NREL/TP-6A2-47187. Golden, CO: National Renewable Energy Laboratory.

Denholm, P., J. Eichman, T. Markel, and O. Ma. 2015. *Summary of Market Opportunities for Electric Vehicles and Dispatchable Load in Electrolyzers*. NREL/TP-6A20-64172. Golden, CO: National Renewable Energy Laboratory.

Denholm, P., J. Jorgenson, M. Hummon, T. Jenkin, D. Palchak, B. Kirby, O. Ma, and M. O'Malley. 2013. *The Value of Energy Storage for Grid Applications*. NREL/TP-6A20-58465. Golden, CO: National Renewable Energy Laboratory.

Drury, E., P. Denholm, and R. Sioshansi. 2011. "The Value of Compressed Air Energy Storage in Energy and Reserve Markets." *Energy* 36:4959-4973.

Duignan, R., C.J. Dent, A. Mills, N. Samaan, M. Milligan, A. Keane, and M. O'Malley. 2012. "Capacity Value of Solar Power." *Proceedings of the 2012 IEEE Power and Energy Society General Meeting*, 22-26 July 2012, San Diego, California. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE). NREL/CP-5500-54832. Accessed July 2015. http://dx.doi.org/10.1109/PESGM.2012.6345429.

E3 (Energy and Environmental Economics). 2015. *Regional Coordination in the West: Benefits of PacifiCorp and California ISO Integration*. San Francisco, CA: E3. Accessed September 2015. https://www.caiso.com/Documents/StudyBenefits-PacifiCorp-ISOIntegration.pdf.

EIA (U.S. Energy Information Administration). 2011. "Electricity storage technologies can be used for energy management and power quality." *Today in Energy*, December 14, 2011. http://www.eia.gov/todayinenergy/detail.cfm?id=4310.

Ela, E., A. Tuohy, M. Milligan, B. Kirby, and D. Brooks. 2012a. "Alternative Approaches for Incentivizing the Frequency Responsive Reserve Ancillary Service." *Electricity Journal* 25(4): 88-102; Accessed March 2015. http://dx.doi.org/10.1016/j.tej.2012.04.015.

Ela, E., B. Kirby, N. Navid, and C. Smith. 2012b. *Effective Ancillary Services Market Designs on High Wind Power Penetration Systems*. NREL/CP-5500-53514. Golden, CO: National Renewable Energy Laboratory. Accessed June 2015. http://www.nrel.gov/docs/fy12osti/53514.pdf.

Ela, E., M. Milligan, A. Bloom, A. Botterud, A. Townsend, and T. Levin. 2014a. *Evolution of Wholesale Electricity Market Design with Increasing Levels of Renewable Generation*. NREL/TP-5D00-61765. Golden, CO: National Renewable Energy Laboratory.

Ela, E., V. Gevorgian, A. Tuohy, B. Kirby, M. Milligan, and M. O'Malley. 2014b. "Market Designs for the Primary Frequency Response Ancillary Service – Part I: Motivation and Design." *IEEE Transactions on Power Systems.* Vol. 29(1):421-431; NREL/JA-5D00-55356. Accessed May 2015. http://dx.doi.org/10.1109/TPWRS.2013.2264942.

Ela, E., V. Gevorgian, A. Tuohy, B. Kirby, M. Milligan, and M. O'Malley. 2014c. "Market Designs for the Primary Frequency Response Ancillary Service - Part II: Case Studies." *IEEE Transactions on Power Systems.* Vol. 29(1):432-440; NREL/JA-5D00-55357. Accessed June 2015. http://dx.doi.org/10.1109/TPWRS.2013.2264951.

EPRI (Electric Power Research Institute). 2014. *Metrics for Quantifying Flexibility in Power System Planning*. Palo Alto, CA: EPRI. Accessed June 2015. http://www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000003002004243.

EPSA (Electric Power Supply Association) v. FERC (Federal Energy Regulatory Commission). 2014. United States Court of Appeals for the District of Columbia Circuit, No. 11-1486. Decided May 23, 2014. Accessed March 2015. http://www.ferc.gov/legal/court-cases/opinions/2014/11-1486.pdf.

ERCOT (Electric Reliability Council of Texas). 2014. "ERCOT Regulating Reserve Service Working Document – Work Session (4-11-2014)". Taylor, TX: ERCOT. Accessed March 2015. www.ercot.com/content/meetings/fast/keydocs/2014/0411/Regulation%20Discussion%20Paper%20as%20of%2004082014_1140.doc.

Eyer, J., J. Iannucci, and G. Corey. 2004. *Energy Storage Benefits and Market Analysis Handbook: A Study for the DOE Energy Storage Systems Program*. SAND2004-6177. Albuquerque, NM: Sandia National Laboratories. Accessed December 2014. http://www.prod.sandia.gov/cgi-bin/techlib/accesscontrol.pl/2004/046177.pdf.

FERC (Federal Energy Regulatory Commission). 2011. "Frequency Regulation Compensation in the Organized Wholesale Power Markets." Order No. 755. Washington, DC: FERC.

———. 2014. Docket Nos ER14-2850-000 and ER14-2851-000. Issued November 10, 2014. http://ferc.gov/CalendarFiles/20141110183532-ER14-2850-000.pdf.

Figueiredo, F.C., P.C. Flynn, and E.A. Cabral. 2006. "The Economics of Energy Storage in 14 Deregulated Power Markets." *Energy Studies Review* 14:131–152.

Goldman, C., M. Reid, R. Levy, and A. Silverstein. 2010. *Coordination of Energy Efficiency and Demand Response*. LBNL-3044E. Berkeley, CA: Lawrence Berkeley National Laboratory. Accessed December 2014. http://emp.lbl.gov/sites/all/files/REPORT%20lbnl-3044e.pdf.

42

Goodin, J. 2008. "California Roundup: Summary of DR Activity in California." 2008 National Town Meeting on Demand Response, June 8.

Holttinen, H., A. Tuohy, M. Milligan, E. Lannoye, V. Silva, S. Muller, and L. Soder. 2013. "The Flexibility Workout." *IEEE Power and Energy Magazine,* Nov/Dec.

Hummon, M., D. Palchak, P. Denholm, J. Jorgenson, D.J. Olsen, S. Kiliccote, N. Matson, M. Sohn, C. Rose, J. Dudley, S. Goli, and O. Ma. 2013. *Grid Integration of Aggregated Demand Response, Part 2: Modeling Demand Response in a Production Cost Model*. NREL/TP-6A20-58492. Golden, CO: National Renewable Energy Laboratory. Accessed December 2014. http://www.nrel.gov/docs/fy14osti/58492.pdf

Hurlbut, David, Ella Zhou, Kevin Porter, and Doug Arent. 2015. *Renewables-Friendly' Grid Development Strategies*. Golden, CO: National Renewable Energy Laboratory.

Ibanez, E., and M. Milligan. 2012. "Impact of Transmission on Resource Adequacy in Systems with Wind and Solar Power." *Proceedings of the 2012 IEEE Power and Energy Society General Meeting*; July 22–26, 2012, San Diego, California. Piscataway, NJ: Institute of Electrical and Electronics Engineers. NREL/CP-5500-57685. Golden, CO: National Renewable Energy Laboratory. Accessed December 2014. http://dx.doi.org/10.1109/PESGM.2012.6343955.

Ibanez, E., and M. Milligan. 2014. "Comparing Resource Adequacy Metrics." 13th International Workshop on Large-Scale Integration of Wind Power into Power Systems. Berlin, Germany. Nov. 11-13, 2014. NREL/CP-5D00-62847. Golden, CO: National Renewable Energy Laboratory. Accessed March 2015. http://www.nrel.gov/docs/fy14osti/62847.pdf.

Keane, A., M. Milligan, C. D'Annunzio, C. Dent, K. Dragoon, B. Hasche, H. Holttinen, N. Samaan, L. Söder, and M. O'Malley. 2011. "Capacity Value of Wind Power." *IEEE Transactions on Power Systems*, Vol 26, No. 2, May (IEEE Wind Capacity Value Task Force paper).

King, J., B. Kirby, M. Milligan, and S. Beuning. 2011. *Flexibility Reserve Reductions from an Energy Imbalance Market with High Levels of Wind Energy in the Western Interconnection*. NREL/TP-5500-52330. Golden, CO: National Renewable Energy Laboratory. Accessed December 2014. http://www.nrel.gov/docs/fy12osti/52330.pdf.

Kirby, B. 2012. "Co-Optimizing Energy and Ancillary Services from Energy Limited Hydro and Pumped Storage Plants." Palo Alto, CA: EPRI, HydroVision.

Kirby, B. and M. Milligan. 2005. *A Method and Case Study for Estimating the Ramping Capability of a Control Area or Balancing Authority and Implications for Moderate or High Wind Penetration*. NREL/CP-500-38153. Golden, CO: National Renewable Energy Laboratory. Accessed December 2014. http://www.nrel.gov/docs/fy05osti/38153.pdf.

Kiviluoma, J., M. O'Malley, A. Tuohy, M. Milligan, B. Lange, H. Holttinen, M. Gibescu, P. Meibom, A. Vigueras-Rodriguez, and E. Gomez-Lazaro. 2010. "Impact of Wind Power on the Unit Commitment and Power System Operations." Ackermann , T., ed. *Proceedings of*

the 9th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants, 18-19 October 2010, Quebec, Canada. Langen, Germany: Energynautics GmbH. NREL/CP-5500-54611.

Lannoye, E., D. Flynn, and M. O'Malley. 2012. "Evaluation of Power System Flexibility," *IEEE Transactions on Power Systems*, 27(2):922–931, May 2012b, doi:10.1109/TPWRS.2011.2177280.

———. 2014. "Transmission, Variable Generation, and Power System Flexibility," *IEEE Transactions on Power Systems*, 30(1):57–66, Jan. 2015. doi:10.1109/TPWRS.2014.2321793.

Lannoye, E., M. Milligan, J. Adams, A. Tuohy, H. Chandler, D. Flynn, and M. O'Malley. 2012a. "Integration of Variable Generation: Capacity Value and Evaluation of Flexibility." IEEE Power Engineering Society (PES); 2012. Summer Meeting. 2012a.

Lew, D., G. Brinkman, E. Ibanez, A. Florita, M. Heany, B.-M. Hodge, M. Hummon, and G. Stark. 2013. "The Western Wind and Solar Integration Study Phase 2." NREL/TP-5500-55588. Golden, CO: National Renewable Energy Laboratory. Accessed December 2014. http://www.nrel.gov/docs/fy13osti/55588.pdf.

Li, H., Z. Lu, Y. Qiao, and P. Zeng. 2015. "Assessment on Operational Flexibility of Power Grid with Grid-Connected Large-Scale Wind Farm." *Power System Technology,* Vol. 39 No. 6: 1672-1678.

Milligan, M. 2014a. "Market Designs for Efficient Wind and Solar Integration into the Bulk Power System." EUCI: Fast Ramp and Intra-Hour Market Incentives, San Francisco, CA.

———. 2014b. "Sources of Flexibility for Wind and Solar Integration." UVIG Short Course.

Milligan, M., and B. Kirby. 2007. "Impact of Balancing Areas Size, Obligation Sharing, and Ramping Capability on Wind Integration." NREL Report No. CP-500-41809. http://www.nrel.gov/docs/fy07osti/41809.pdf

———. 2010a. "Market Characteristics for Efficient Integration of Variable Generation in the Western Interconnection." NREL/TP-550-48192.

———. 2010b. "Utilizing Load Response for Wind and Solar Integration and Power System Reliability." NREL/CP-550-48247. http://www.nrel.gov/docs/fy10osti/48247.pdf.

Milligan, M., B. Kirby, J. King, and S. Beuning. 2011. "The Impact of Alternative Dispatch Intervals on Operating Reserve Requirements for Variable Generation." *10th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants Proceedings*, Oct. 25–26, 2011, Aarhus, Denmark.

Milligan, M., B. Kirby, R. Gramlich, and M. Goggin. 2009. *Impact of Electric Industry Structure on High Wind Penetration Potential.* NREL/TP-550-46273. http://www.nrel.gov/docs/fy09osti/46273.pdf.

Milligan, M., B. Kirby, T. Acker, M. Ahlstrom, B. Frew, M. Goggin, W. Lasher, M. Marquis, and D. Osborn. 2015. *Review and Status of Wind Integration and Transmission in the United States: Key Issues and Lessons Learned.* (Supplement to the U.S. DOE Wind Vision). NREL/TP-5D00-61911. Golden, CO: National Renewable Energy Laboratory. Accessed September 2015. http://www.nrel.gov/docs/fy15osti/61911.pdf .

Milligan, M., E. Ela, B. M. Hodge, B. Kirby, D. Lew, C. Clark, J. DeCesaro, and K. Lynn. 2011. "Cost-Causation and Integration Cost Analysis for Variable Generation." NREL Report No. TP-5500-51860. http://www.nrel.gov/docs/fy11osti/51860.pdf

Milligan, M., E. Ela, D. Lew, D. Corbus, Y. Wan, B. Hodge, and B. Kirby. 2012b. "Assessment of Simulated Wind Data Requirements for Wind Integration Studies." *IEEE Journal on Sustainability* 3(4):620-626. http://dx.doi.org/10.1109/TSTE.2011.2160880.

Milligan, M., H. Holtinnen, L. Söder, C. Clark, and I. Pineda. 2012a. "Markets to Facilitate Wind and Solar Energy Integration in the Bulk Power Supply: An IEA Task 25 Collaboration." NREL/CP-5500-56212. Golden, CO: National Renewable Energy Laboratory. http://www.nrel.gov/docs/fy12osti/56212.pdf.

Milligan, M., K. Clark, J. King, B. Kirby, T. Guo, and G. Liu. 2013. "Examination of Potential Benefits of an Energy Imbalance Market in the Western Interconnection." NREL/TP-5500-57115. http://www.nrel.gov/docs/fy13osti/57115.pdf.

Mills, A. 2008. "Valuation of Variability Mitigation Strategies for Wind and Solar Generation: Preliminary Analysis and Results." UVIG Operating Impacts and Integration Studies Users Group Meeting, October 1, 2008, Denver, CO. Accessed December 2014. http://www.uwig.org/members/Denver2008/OIIS/Mills.pdf

MISO. 2015. *MISO Monthly Operations Report to the Board of Directors Market Committee.* June 17th, 2015. Carmel, IN: MISO. Accessed July 2015. https://www.misoenergy.org/Library/Repository/Meeting%20Material/Stakeholder/BOD/Markets%20Committee/2015/20150617/20150617%20Markets%20Committee%20of%20the%20BOD%20Item%2004%20Monthly%20Operations%20May%202015.pdf.

Monitoring Analytics. 2015. *2014 State of the Market Report for PJM: Volume 2.* Carmel, IN: MISO. Accessed September 2015. http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2014/2014-som-pjm-volume2.pdf.

NERC (North American Electric Reliability Corporation). 2010. *Flexibility Requirements and Metrics for Variable Generation: Implications for System Planning Studies.* Princeton, NJ. Accessed December 2014. http://www.nerc.com/files/IVGTF_Task_1_4_Final.pdf.

Olsen, D.J., N. Matson, M.D. Sohn, C. Rose, J. Dudley, S. Goli, S. Kiliccote, M. Hummon, D. Palchak, P. Denholm, J. Jorgenson, and O. Ma. 2013. *Grid Integration of Aggregated Demand Response, Part I: Load Availability Profiles and Constraints for the Western Interconnection*. LBNL-6417E. Berkeley, CA: Lawrence Berkeley National Laboratory. Accessed December 2014. http://eetd.lbl.gov/sites/all/files/lbnl-6417e.pdf.

Pfeifenberger, J.P., K. Spees, K. Carden, and N. Wintermantel. 2013. *Resource Adequacy Requirements: Reliability and Economic Implications*. The Brattle Group and Astrape Consulting. Accessed December 2014. http://www.ferc.gov/legal/staff-reports/2014/02-07-14-consultant-report.pdf.

PJM. 2014. *The Evaluation of Demand Response in the PJM Wholesale Market*. Accessed July 2015. http://www.pjm.com/~/media/documents/reports/20141007-pjm-whitepaper-on-the-evolution-of-demand-response-in-the-pjm-wholesale-market.ashx.

Sioshansi, R., P. Denholm, T. Jenkin, and J. Weiss. 2009. "Estimating the Value of Electricity Storage in PJM: Arbitrage and Some Welfare Effects." *Energy Economics* 31:269–277.

SPP (Southwest Power Pool). 2009. "Three Nebraska Utilities Complete Transition to Southwest Power Pool" [Press release]. Accessed December 2014. http://www.spp.org/publications/Nebraska_Utilities_Complete_Transition_to_SPP.pdf.

Stoft, S. 2002. *Power System Economics: Designing Markets for Electricity*. New York, NY: IEEE Press & Wiley-Interscience, John Wiley & Sons, Inc.

Walawalkar, R., J. Apt, and R. Mancini. 2007. "Economics of Electric Energy Storage for Energy Arbitrage and Regulation in New York." *Energy Policy* 35:2558–2568.

Appendix

Table A-1 shows the DR resource ability during the top 20 hours of load (as a proxy for capacity provision). Figure A-1 shows the value of DR services across a full year.

Table A-1. Availability of Demand Response to Provide Energy in the Colorado Test System on an Annual Basis as Well as in the Top 20 Hours of Greatest Demand

| Demand Response Resources Providing Energy[a]	Annual			Top 20 Load Hours	
	Capacity (mean/min/max) [MW][b]	Annual Available Energy with Constraints [GWh][c]	Annual Hours Available[d]	Capacity (mean/min/max) [MW]	% of Load (mean/min/max)
Residential Cooling	10.9 / 0 / 108.7	38.8	5,390	55.1 / 41.3 / 72.2	0.4 / 0.3 / 0.5
Residential Water Heating	1.8 / 0.5 / 3.9	15.7	8,784	1.3 / 1.1 / 1.6	0 / 0 / 0
Commercial Cooling	2.1 / 0 / 46.4	10.7	4,338	8.2 / 4.2 / 12.9	0.1 / 0 / 0.1
Commercial Heating	2.2 / 0 / 25.5	3.8	8,390	0 / 0 / 0	0 / 0 / 0
Municipal Pumping	1.7 / 0.4 / 3.8	2.1	8,784	2.1 / 1.7 / 2.7	0 / 0 / 0
Wastewater Pumping	1.5 / 1.5 / 1.5	1.6	8,784	1.5 / 1.5 / 1.5	0 / 0 / 0
Refrigerated Warehouses	0.2 / 0 / 0.4	0.3	8,685	0.3 / 0.3 / 0.4	0 / 0 / 0
Agricultural Pumping	17 / 1.7 / 41.2	49.9	8,784	36.6 / 32 / 40	0.3 / 0.2 / 0.3
Data Centers	8 / 8 / 8	11.7	8,784	8 / 8 / 8	0.1 / 0.1 / 0.1
Total[e]	45.4 / 14.5 / 227.8	134.6	8,784	113.1 / 91.8 / 137.3	0.8 / 0.7 / 1

Source: Hummon et al. 2013
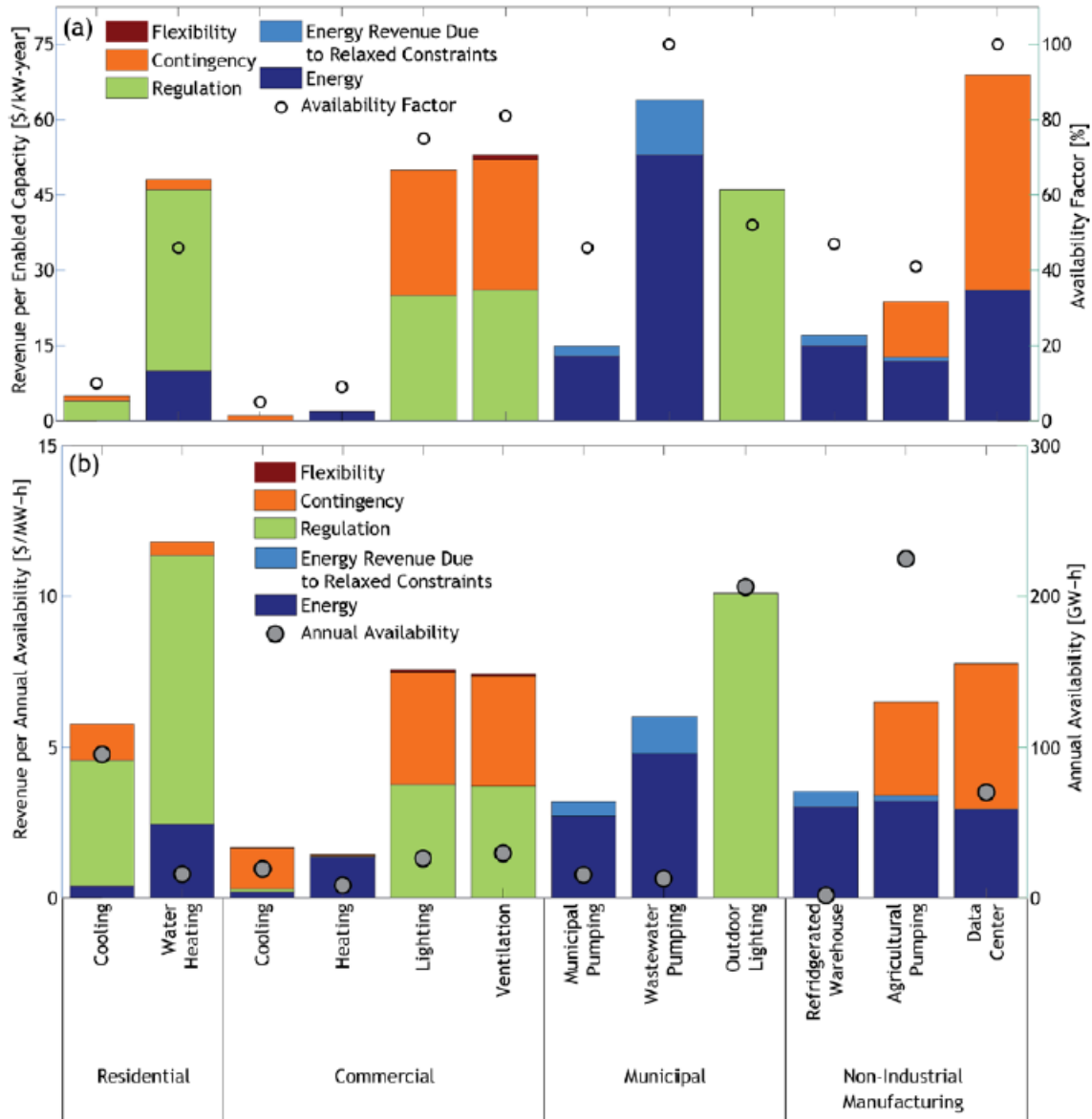
Figure A-1. Average annual revenue (left axis) from the day-ahead market per (a) total enabled capacity and (b) annual availability for each type of DR resource in the Colorado test system

Annual DR resource availability (right axis) is expressed as the (a) annual availability factor or as (b) total annual availability.
Source: Hummon et al. 2013