DOCKETED

Docket Number:	15-IEPR-06
Project Title:	Renewable Energy
TN #:	205828
Document Title:	Large-scale Solar Association Comments: Report: Integrating Higher Levels of Variable Energy Resources in California
Description:	N/A
Filer:	System
Organization:	Rachel Gold/Large-scale Solar Association
Submitter Role:	Public
Submission Date:	8/25/2015 2:19:14 PM
Docketed Date:	8/25/2015

Comment Received From: Rachel Gold, Large-scale Solar Association
Submitted On: 8/25/2015
Docket Number: 15-IEPR-06

Report: Integrating Higher Levels of Variable Energy Resources in California

Please see attached report by GE Energy Consulting - Integrating Higher Levels of Variable Energy Resources in California.

Additional submitted attachment is included below.

Final Report:

Integrating Higher Levels of Variable Energy Resources in California

Prepared for: **Large-scale Solar Association**

Prepared by: **Debra Lew**

Mark Schroder

Nicholas Miller

Matt Lecar

June 15, 2015

Imagination at Work

Legal Notices

This report was prepared by General Electric International, Inc. as an account of work sponsored by the Large-scale Solar Association. Neither the Large-scale Solar Association, General Electric International, Inc., nor any person acting on their behalf:

1. Makes any warranty or representation, expressed or implied, with respect to the use of any information contained in this report, or that the use of any information, apparatus, method, or process disclosed in the report may not infringe privately owned rights.

2. Assumes any liabilities with respect to the use of or for damage resulting from the use of any information, apparatus, method, or process disclosed in this report.

Foreword

This paper was prepared by General Electric International, Inc. acting through its Energy Consulting group based in Schenectady, NY, and submitted to the Large-scale Solar Association. The team gratefully acknowledges the discussions and contributions of Susan Schneider, Clyde Loutan, Jim Blatchford, Shucheng Liu, Don Trethaway, Mark Ahlstrom, Brendan Kirby, Erik Ela, Greg Brinkman, Todd Strauss, Antonio Alvarez, Benyamin Moradzedah, Larson Plano, Will Dong, Beth Vaughan, Dan Consie, Rudy Ravens, Doug Davie, Gary Ackerman, Michael McMullen, Kris Ruud, Sundar Venkataraman, Jovan Bebic, Derek Stenclik, and Gary Jordan. Questions and any correspondence concerning this document should be referred to:

Debra Lew

Technical Director

GE Energy Consulting

1 River Road

Schenectady, NY 12345

(303) 819-3470

debra.lew@ge.com

Table of Contents

List of Figures

List of Tables

List of Acronyms

BAA	Balancing Authority Area
CAISO	California Independent System Operator
CHP	Combined Heat and Power
CPUC	California Public Utilities Commission
CSP	Concentrating solar power
DA	Day-ahead
DAM	Day-ahead market
DIR	Dispatchable Intermittent Resources
EIM	Energy Imbalance Market
ERCOT	Electric Reliability Council of Texas
EV	Electric vehicle
FERC	Federal Energy Regulatory Commission
FMM	Fifteen-minute market
FRAC	Flexible Resource Adequacy Capacity
GE	General Electric International, Inc.
GW	Gigawatt
GWh	Gigawatt hour
IFM	Integrated Forward Market
LMP	Locational Marginal Price
LTPP	Long-term Procurement Plan
MISO	Midcontinent Independent System Operator
MW	Megawatt
MWh	Megawatt hour
OTC	Once-Through Cooling
PG&E	Pacific Gas and Electric
PHEV	Plug-in Hybrid Electric Vehicle

PSCO	Public Service of Colorado
PV	Photovoltaics
PURPA	Public Utility Regulatory Policies Act
RPS	Renewables Portfolio Standard
RT	Real-time
RTM	Real-time market
RUC	Residual Unit Commitment
VER	Variable Energy Resource

1 EXECUTIVE SUMMARY

1.1 Introduction

California has long been a leader in renewable energy. With a 33% Renewables Portfolio Standard (RPS) by 2020 and Governor Brown's recent executive order directing a reduction of the state's greenhouse gas emissions (GHG) by 40% by 2030, California is examining higher levels of renewables, including a 50% target. Shifting to higher levels of renewable resources necessitates evaluation and potential changes to how the grid is managed. The California Independent System Operator (CAISO) has identified a number of issues that need to be addressed in order to effectively manage the grid and balance the system with increasing levels of variable energy resources[1] (VERs), such as wind and solar. Critical to addressing these issues is an understanding of the aspects of California's current system that make it unique:

1. It relies on a large amount of imports that cannot be controlled in the same way that internal resources can.
2. It has a large amount of self-scheduling that reduces CAISO's ability to balance the system.

Recognizing these challenges, this paper examines potential mitigation options that can be implemented today or in the longer term (out to 2030) to support higher penetrations of VERs in California.

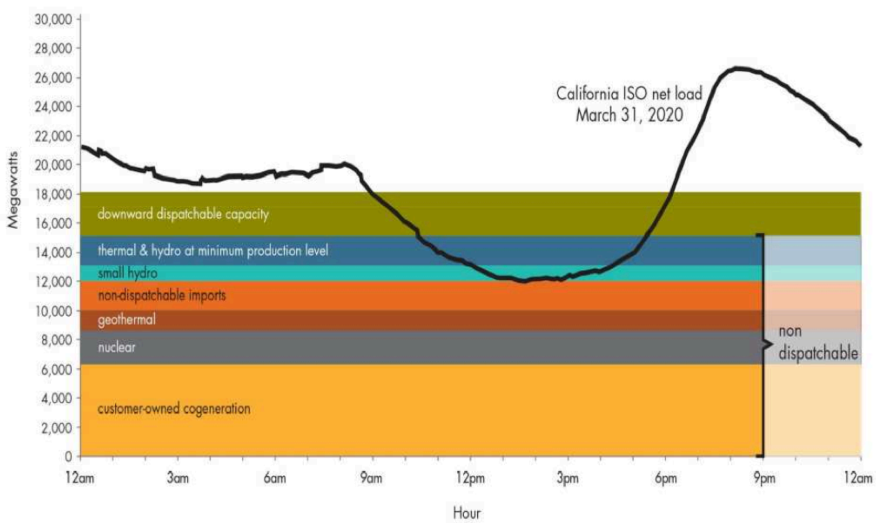


Figure 1 – Projected CAISO spring net load (black line) in 2020 and non-dispatchable stack of generation (colored layers).
Source: FERC Docket AD14-9-000, Testimony of Brad Bouillon, CAISO (June 10, 2014).

1 FERC defines a VER as "a device for the production of electricity that is characterized by and energy source that: 1) is renewable, 2) cannot be stored by the facility owner or operator, and 3) has variability that is beyond the control of the facility owner or operator".

California leads the country in solar generation, with over 5% of California's electricity from utility-scale solar in 2014 plus additional electricity from over 2,300 MW of behind-the-meter solar. Solar generation tends to coincide with demand. In other words, solar plants produce power while the sun is shining, which is also when demand for energy is high. This is especially helpful during the summer when air-conditioning increases the system load. During the *non-summer* months, solar output can serve a large portion of the load and can result in a duck-shaped, net load[2] curve (Figure 1 shows this "duck curve"). The net load curve, combined with CAISO's large "stack" of non-dispatchable resources (colored layers in Figure 1) can make balancing the system difficult. One of CAISO's concerns is that if more VER output than was scheduled shows up in the real-time market, there can be too much supply. When there is too much supply, the market responds and prices drop, sometimes to negative levels below zero. When prices go negative, generators are incentivized to reduce output, because if they are exposed to negative prices, they must pay to produce power. This is an economic response that is part of the overall management of the system. It is not a reliability issue. However, if output does not drop enough, downward reserves may be depleted. If downward reserves are depleted, generation may exceed load, which is a reliability issue.

This paper investigates mitigation options to help California integrate higher levels of VERs into its grid. It examines options that address both the economic manifestations of the duck curve as well as the reliability implications of these options. These options include extracting flexibility and real-time price-responsiveness from the following:

- Interties
- Thermal generators
- Combined Heat and Power (CHP)
- Economic dispatch of VERs
- Load

In order to shut down thermal generators during low load, high VER output hours, the following resources can provide ancillary services that the thermal generators may have been providing:

- Grid-friendly VERs
- Reserves from storage
- Reserves from demand response

1.2 Positioning the System

In typical utility operations, the day-ahead time-frame is when the system is positioned using the best available forecast information on load and VERs. It is positioned not only to meet the load, but also to survive any contingencies, to handle forecast errors, and to manage VER variability. CAISO runs a day-ahead market that clears supply bids against demand bids. They also run a reliability procedure that uses day-ahead VER and load forecasts and commits[3] more capacity if

[2] The net load is defined as load minus VER output.

[3] Commitment of a generator for a particular hour means that the generator will be online and generating power during that hour. Dispatch of a generator refers to the output level of an online generator.

needed. However, it does not decommit capacity that is not needed. Some of the VERs are scheduled into the day-ahead market, similar to conventional generators. Other VERs are not scheduled in the day-ahead market, which can result in over-commitment of generators in the day-ahead. While existing processes help to reduce over-commitment, a better solution would be to encourage VERs participation in the day-ahead market or to allow the reliability process to decommit based on forecasts. Improving how VER forecasts feed into reserve requirements and VER forecast accuracy itself will further help.

1.3 Flexibility and Price-responsiveness from the System

In real-time, CAISO needs to balance the system, despite forecast errors, short-term variability of VERs, and any potential contingencies. Integrating high levels of VERs requires a system that is as flexible and responsive to real-time prices as possible. Flexibility includes not just technical capability (e.g., to stop and restart or turn down) but also the ability to use that flexibility. Price-responsiveness requires a unit to be exposed to the negative prices and to be capable of backing down in response. The non-dispatchable stack of generation in Figure 1 has "slices" that may be able to provide additional flexibility and price-responsiveness. Some slices such as nuclear, geothermal, small hydro, etc. cannot easily provide flexibility, however, the following are worthy of further investigation:

- **Interties** – California has a large amount of imports and these imports have been increasingly self-scheduled in the real-time market in the last year, reducing their flexibility and real-time price-responsiveness. An immediate potential mitigation is for utilities to add "negative" slices to the stack of non-dispatchable resources, i.e., during low load, high solar output days, utilities could schedule several hours of exports to neighboring utilities. Expansion of the Energy Imbalance Market will help in the near-term, but the extent to which it can help will depend on transfer capacities and imbalances in the real-time. In the longer term, other utilities such as PacifiCorp joining the CAISO can help much more, because that will increase the diversity of the VERs, load and generators for both day-ahead positioning and real-time balancing. Another option to encourage price-responsiveness and provide ramping capability would be to find a way to allow interties to help meet some of the new flexible resource adequacy capacity requirement.
- **Thermal generators** – To integrate high levels of solar midday, the thermal units may need to turn off or down midday and restart in the evening. There are uprate options for the combined cycle gas plants to reduce minimum generation levels and increase maximum generation levels. For example, uprating all the existing F-class combined cycle plants in CAISO would provide an estimated 2,000 – 2,500 MW of additional range to meet the evening ramp. Uprate options for a hot combined cycle plant, such as one that was turned off after the morning load ramp, can allow it to restart in 75 minutes, which is plenty of time to help meet the neck of the duck curve.
- **CHP** – The 2011 Qualifying Facilities/CHP settlement with California's CHP fleet of about 8,000 MW has resulted in some increased flexibility. About one-quarter of the fleet has moved to a new tariff that allows utilities to schedule their output, making them

dispatchable. It is estimated that another quarter of the fleet may also be able to provide flexibility if incentivized.

- **VER economic dispatch** – Flexibility from new VERs can also be harnessed to help balance the system by requiring them to be economically dispatched in real-time. VERs can be curtailed manually, as they are today in CAISO. However this gross level of control may result in more curtailment than the finer control of putting VERs on economic dispatch, because the system operator may be conservative about the amount of curtailment requested. Adding a finer level of control by economically dispatching VERs is also more likely to result in VERs being released from curtailment as soon as there is room on the system. The financial implications of VER curtailment are complex due to various requirements, incentives and contract structures, and depending on the reasons for VER curtailment, make-whole provisions may be needed. But here there is an opportunity to structure new VER development and transactions so that VERs continue to contribute to GHG reduction goals while maintaining system reliability.
- **Load** – Over-supply and negative prices present a great opportunity for load, but load must be exposed – and able to respond accordingly – to negative prices. As a first step, efforts should focus on finding ways to make energy available to loads at real-time prices. This includes structuring electric vehicle charging, water pumping tariffs, etc. to enable real-time price-responsiveness. Here, the idea of peak and off-peak pricing needs to evolve as today's negative prices sometimes occur during peak times. For example, existing business time-of-use and electric vehicle charging rate structures charge peak or partial-peak rates during the midday in the spring when over-supply and negative prices are likely to occur. Shifting the rate structure to instead incent energy use when supply is high, as indicated by low prices, would help to resolve multiple issues to the benefit of both the electricity market and the consumer, while helping the state to meet its GHG reduction goals.

1.4 Replacing Services from Thermal Generators

During low load, high solar output days, CAISO may want to turn off generators midday to accept as much solar power as possible and then restart these generators to meet the evening peak. In addition to enabling the system to accept as much solar as possible, this approach provides the maximum amount of range from these generators. It is especially important to turn off thermal generators because they have a minimum generation level (when they are "on", they must run at some minimum output level such as 40% of rated capacity). But in order to turn off thermal generators, CAISO needs to replace not only the energy but also the ancillary services that they provide:

- **Grid-friendly VERs** – VERs can provide a number of essential reliability services to the grid including reactive power support to regulate voltage (even when the sun is not shining and/or wind is not blowing), regulating reserves, synthetic inertia, and primary frequency response. New VERs could provide (and be compensated for) down reserves. This can displace not only the amount of down reserves that a thermal generator is providing but also the minimum generation level of that thermal generator. In order for VERs to provide upward responses for regulating reserves or primary frequency response, they need to be pre-curtailed, but this could be limited to an ancillary service provided during times of over-supply. Not all of these services are needed from all VERs, and it is cheaper and easier to ask for any of these services from new VERs than to renegotiate an existing contract and retrofit a VER over its 20+ year lifetime. For that reason, CAISO should consider requiring one or more of these services, as appropriate, as compensated elements of new VER installations.

- **Reserves from storage** – One MW of storage can be leveraged to displace more than one MW of thermal generation when there is excess VER supply. For example, 20 MW of storage could decommit a thermal generator that is operating at 60 MW and providing 20 MW each of up and down regulation (ignoring losses), providing 3:1 leverage. Half of the 1,325 MW by 2020 storage mandate could displace the +/- 600 MW regulation requirement and associated 1,400 MW of thermal generation.
- **Reserves from demand response** – Demand response can provide fast up-reserves which can displace thermal generators that are currently providing those services. For example, the Electric Reliability Council of Texas (ERCOT) uses load resources to provide up to half of its spinning reserve requirements. It is currently examining the use of demand response for fast frequency response.

1.5 Conclusion

There are a number of operational, technical, market, and institutional changes that can be made in California to unlock additional flexibility in the system to help integrate more VERs. The key is to find flexibility and price-responsiveness in as much of the system (including conventional generators, VERs and load) as possible. Encouraging VERs in the day-ahead market or decommitting generation in the day-ahead reliability process will allow CAISO to optimally schedule generators in the day-ahead. Finally, reliability services can be provided from other sources besides thermal generators, which can allow thermal generation to be decommitted as necessary.

2 INTRODUCTION

California has long been a leader in renewable energy. It leads the country in solar generation: over 5% of California's electricity was produced by utility-scale solar in 2014 plus additional electricity from over 2,300 MW of behind-the-meter solar. California also has a 10% wind energy penetration.

Wind and solar are variable energy resources (VERs): their variability is beyond the control of the facility owner or system operator. High penetrations of VERs can be operationally challenging because of their variability and uncertainty (forecast error). With a 33% Renewables Portfolio Standard (RPS) target for 2020 and the Governor's proposal of 50% renewables by 2030, the California Independent System Operator (CAISO)[4] predicts challenges in balancing the system, absent significant changes in the system or operations. This paper investigates potential solutions to grid operational challenges in California with increasing levels of VERs. It examines issues in today's system and how the system can be prepared for high levels of variable energy resources that are key to meeting the state's long-term greenhouse gas (GHG) reduction goals in 2030.

Production simulation studies that model generator operation and transmission flows across the entire Western Interconnection for a year, have been conducted to study high penetrations of VERs in California. For example, the Western Wind and Solar Integration Study Phase 2[5] studied VERs serving up to 33% of the electricity demand across the West, and the Low Carbon Grid Study[6], examined VERs serving up to 45% of the demand in California. These studies have found that technically, the system can be balanced with much higher levels of wind and solar than are present today.

Solar output tends to coincide with demand. This is especially helpful during the summer when cooling loads are high. However, during the *non-summer*, low-load months, solar output can be a significant fraction of the load. The net load (defined as load minus wind minus solar) shape is similar to a duck profile (see Figure 2). The duck curve has two main challenges. First, as the belly of the duck drops lower with increasing solar generation, it can bump into the "stack" of non-dispatchable resources (colored layers in Figure 3) which are regarded as the minimum generation in the middle of the day because they typically are not or cannot be dispatched down or decommitted. Second, the evening load rise that makes up the neck of the duck becomes steeper with increasing solar generation. The CAISO forecasts that this evening load rise in 2020 will result in a 3-hour ramp of up to 13,000 MW[7].

---

[4] This paper focuses on CAISO which is the Balancing Authority Area that balances most of California's load, but other Balancing Authority Areas within California such as the Sacramento Municipal Utility District are also managing large amounts of VERs.

[5] D. Lew et al, "Western Wind and Solar Integration Study Phase 2", NREL/TP-5500-55588, Sep. 2013. http://www.nrel.gov/docs/fy13osti/55588.pdf

[6] G. Brinkman et al, "California 2030 Low Carbon Grid Study (LCGS): Phase 1", Oct. 2014. http://www.lowcarbongrid2030.org/wp-content/uploads/2014/10/LCGS_PhaseI_NRELslides.pdf

[7] The duck curve shown here underestimates the challenge because behind-the meter solar generation has not been taken into account but is expected to increase significantly between now and 2020.

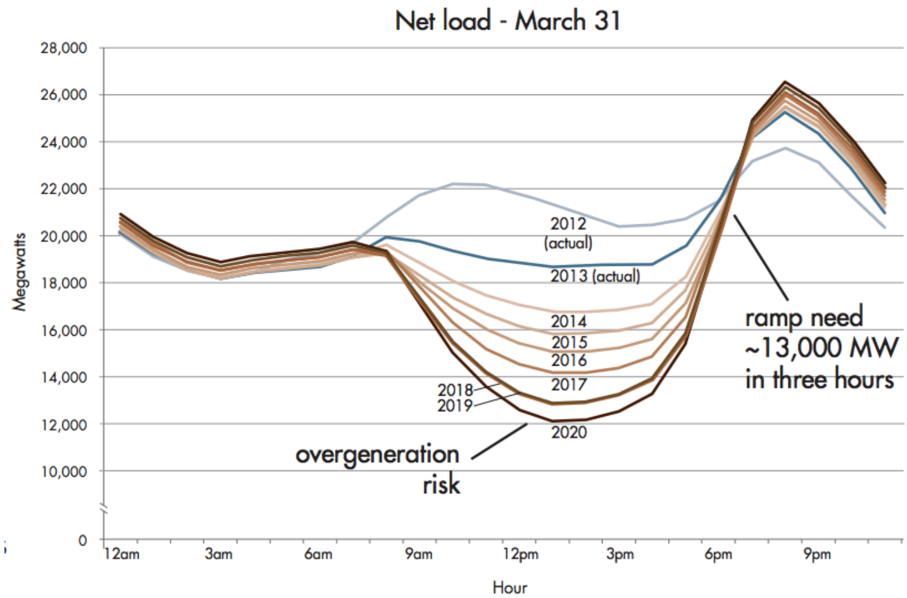


Figure 2 – Projected CAISO net load in 2012-2020 during the low load spring season. Note that the y-axis is truncated.
Source: CAISO, "Fast Facts," October 2013.

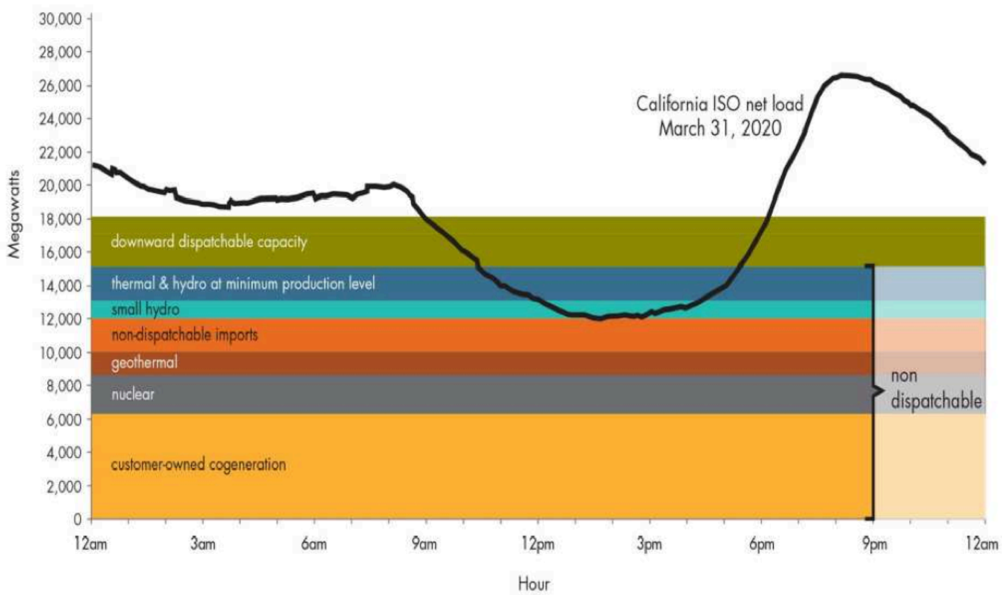


Figure 3 – Projected CAISO spring net load (black line) in 2020 and non-dispatchable stack of generation (colored layers).
Source: FERC Docket AD14-9-000, Testimony of Brad Bouillon, CAISO (June 10, 2014).

CAISO has limited generation resources available to back down in the middle of the day, with increased levels of negative prices in the real-time market, occasional manual renewable energy curtailments, and occasional over-generation, and has undertaken a number of efforts to address this issue including a recent reduction of the bid price floor to -$150/MWh with provisions to further drop it to -$300/MWh to encourage generators to back down. In addition, in order to meet the evening load rise, CAISO has instituted a new Flexible Resource Adequacy Capacity (FRAC) requirement, which requires generator ramping, stopping and restarting capabilities.

Why are the models showing that high VER penetrations are technically possible when there are difficulties in reality? What are institutional, contractual, market or other barriers to integration of VERs? This paper seeks to investigate mitigation options to these barriers distinguishing between options that can be implemented today versus in the longer-term future. The operational challenges of today can be more difficult during this transition period in that they occur occasionally and are hard to predict. In a high VER penetration future, these same challenges may be easier to manage because they will occur regularly and system operators and market participants will be accustomed to the mechanisms for managing them.

2.1 California is unique

California is blessed with some attributes to help it reach its renewables and GHG emissions targets: abundant and high quality renewable energy resources; gas and hydro capacity (as opposed to coal) for balancing; and a large Balancing Authority Area (BAA) footprint with good geographic diversity of load and resources. Unfortunately, California also has attributes that make it difficult to integrate VERs.

First, California relies on imports to provide one-third of its electricity. CAISO cannot control these imports in the same way that it can control its internal resources. This is an institutional, not physical limitation. For example, an internal generator can be dispatched at 5-minute intervals. Fast dispatch allows it to help balance variability of VERs. An identical, but external, generator, which is exporting to CAISO, is typically scheduled at hourly intervals (15-minute scheduling was recently made available but most imports continue to be scheduled hourly).

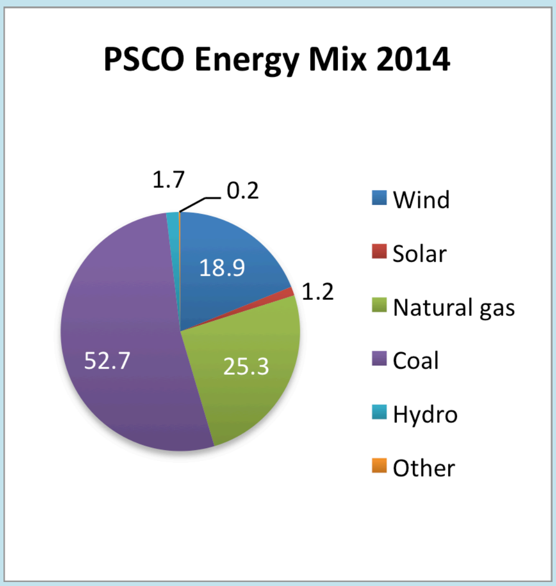
Second, and somewhat related to the first, much of the energy within CAISO is self-scheduled, including that on the interties. For example, Load-Serving Entities frequently schedule their own generation to meet their own load and ancillary service needs as opposed to economically offering it into the market. When utilities self-schedule their own generation from within CAISO or on the interties from external sources, it reduces the economic efficiency of the overall system and leaves CAISO in the position of balancing the system using only a small subset of total system resources to do so.

Complicating matters, generation is increasingly decentralized in California. There is an estimated 2,300 MW of distributed, behind-the-meter photovoltaics (PV) in California now with significant expected additions driven by customer demand. CAISO currently has neither visibility nor control of behind-the-meter resources such as distributed PV. As a result, both load forecasting and system balancing are more challenging.

In addition, there is high risk involved with renegotiating existing contracts to modify a generator's services, such as requesting existing generators to provide new ancillary services or be more responsive to price. This enhanced risk further complicates CAISO's work.

High VER penetrations have successfully been integrated

Like California, Xcel/Public Service Company of Colorado (PSCO) has an aggressive RPS (30% by 2020). In 2014, PSCO had significant penetrations of wind (19% by energy) and solar (1%) on their system (see chart below from 2014 PSCO data). They reported that wind served over 50% of the demand during 6% of the hours in Oct 2014. What makes this particularly impressive is that they have a small BAA with a peak load of 7 GW (compared to CAISO's peak load of 50 GW). Their system is different from California's in that they manage higher wind penetrations and lower solar penetrations, with higher coal and less gas generation for balancing. At times, they curtail wind at night when coal is backed down to minimum generation levels. They conducted a study to determine that coal cycling costs and wind curtailment costs were similar, but because there are uncertainties around coal cycling impacts, they prefer to curtail wind. When wind is curtailed, they allow it to provide up regulating reserves (PSCO has 1,678 MW of wind on Automatic Generation Control). To help integrate additional VERs, they are initiating a Joint Dispatch Agreement (similar to an Energy Imbalance Market) with two other entities.



However, PSCO is not finding high levels of VERs as difficult to manage. They recently told their Public Utilities Commission that they did not need new gas generation or new storage to balance the system. A major difference is that PSCO is able to control (commit and dispatch) all of their resources.

However, other BAA's have successfully integrated high levels of VERs into their grids. The text box on page 2-14 provides an example of this success[8]. A notable difference in BAA operations is the ability of the system operator to control and dispatch all of its resources.

8 Xcel Energy Services, "An investigation of Potential Electric Storage Options on the Public Service Company of Colorado System", Response to CO PUC Decision No. C13-1556, Proceeding No. 14M-1160E, Dec. 8, 2014.

2.2 Grid Challenges

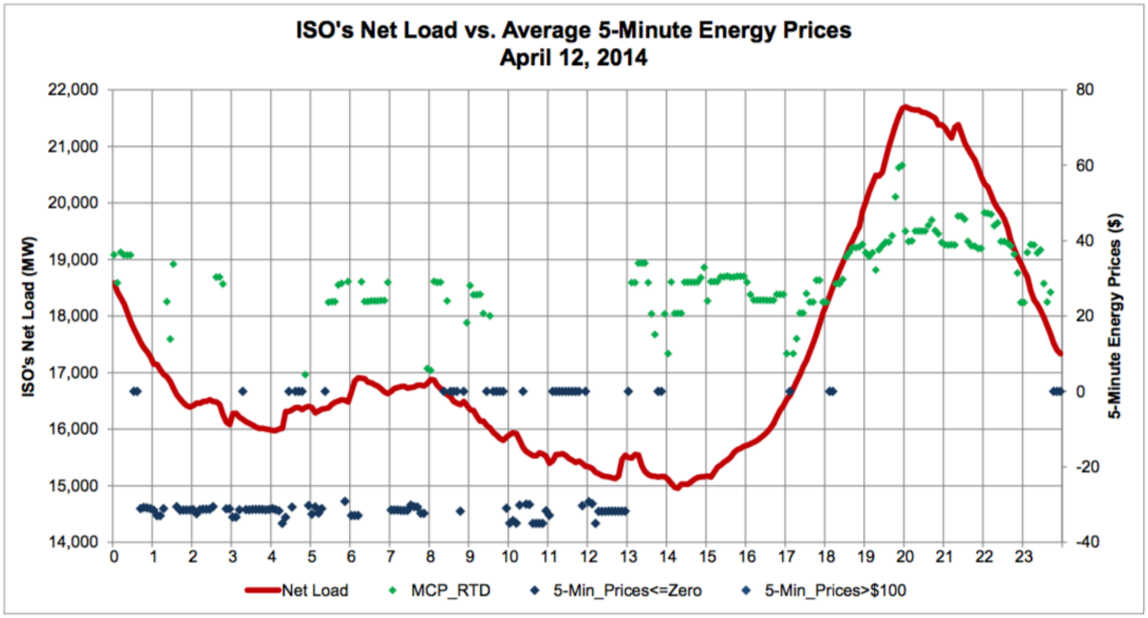


Figure 4 – CAISO net load and average 5-minute energy prices on April 12, 2014. Green dots show positive 5-minute prices and blue dots show zero or negative 5-minute prices. Source: FERC Docket AD15-4-000, Technical Conference on Environmental Regulations and Electric Reliability, Wholesale Electricity Markets, and Energy Infrastructure, Statement of Mark Rothleder, CAISO (Feb. 19, 2015).

CAISO is experiencing increasing levels of negative prices in the real-time market the middle of the day. However, it is important to note that negative prices are an *economic*, not *reliability* issue. Negative prices are the market's way of telling generators to reduce output. To maintain *reliability*, the system operator can issue exceptional dispatches and curtail generation.

To manage the duck curve, the system operator needs to be able to decommit (shut down) or dispatch down (reduce output) generating units in the middle of the day. In a well-functioning market, generators should respond to negative prices by backing down. This backing down of generators should reduce generation sufficiently to avoid over-generation situations. Of course, for generators to respond to negative prices, two conditions must be met. The generator:

1. Must be "exposed" in some economic fashion to the negative prices, and
2. Must be capable of backing down.

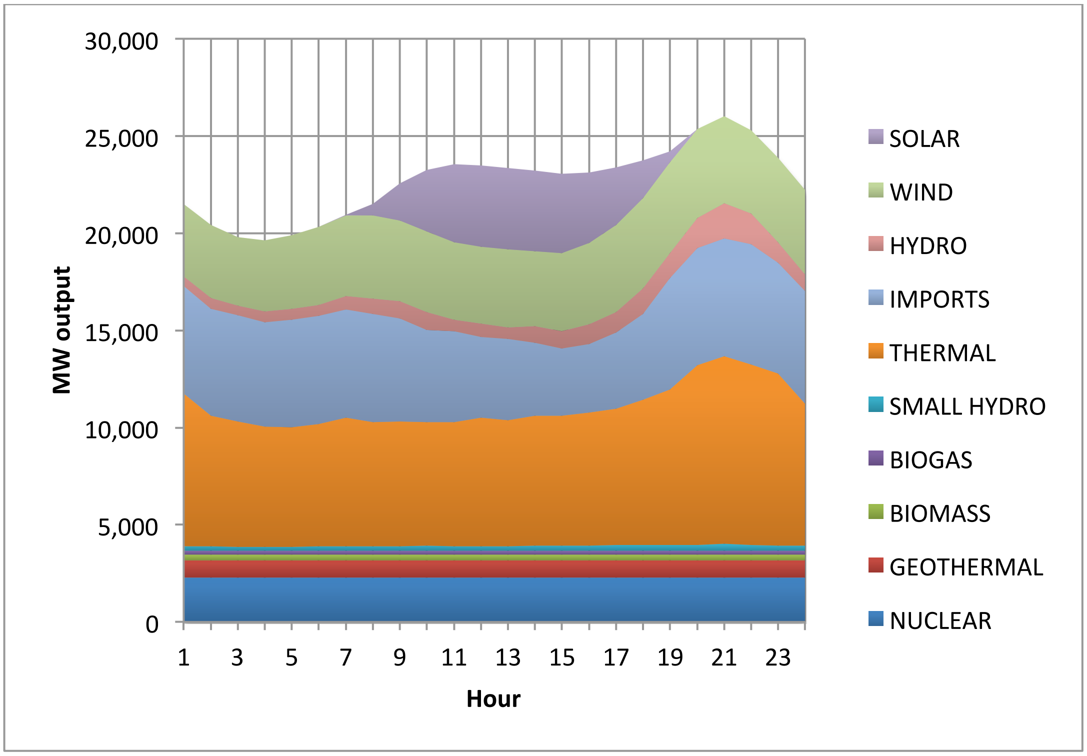


Figure 5 – CAISO dispatch on April 12, 2014.
Data: CAISO Daily Renewables Watch, Apr. 12, 2014.

Figure 4 shows the net load and 5-minute energy prices on a day in which 43% of the 5-minute intervals had zero or negative prices. On this day, negatively priced bids (as opposed to transmission congestion or a power balance issue) were the main reason for the negative prices. However, between 12:00 and 13:00, CAISO exhausted its down regulating reserves and this culminated in an over-generation event. CAISO reported this was due to nearly 10 GW of renewables that were not scheduled in the day-ahead market and the fact that it was a weekend so loads were lower than normal[9]. Wind was under-forecast in the day-ahead, but the forecast errors were similar or less than the errors earlier that week. Net load did not go below 15,000 MW that day. The dispatch of generation on that day (Figure 5) shows that despite the fact that prices were zero or negative during 9:30 – 13:00 thermal generation did not back down, and hydro and imports backed down slightly. While thermal generation is not expected, for example, to back all the way down to zero during negative prices, because those units may be needed later in the day, the degree of their responsiveness to price is poor.

It is the multi-colored stack of non-dispatchable resources in Figure 3 that combine with the net load's duck curve shape to create balancing issues. Increasing VER levels cause the belly of the net load duck curve to bump into this stack of non-dispatchable units. CAISO reports that this non-dispatchable stack is 10 – 12 GW today, based on the maximum capability of those

[9] CAISO Market Update Call Meeting Minutes April 17, 2014.

resources. An additional factor that needs to be accounted for is that the current drought conditions in California have led to less hydro power than normal. In a normal or "wet" hydro year, higher levels of non-dispatchable resources may create more pressures on the duck curve.

As VER penetrations increase, curtailment will increase unless current practice evolves and other measures are taken. Increasing curtailment means that renewable capacity will need to be overbuilt in order to meet RPS targets. While curtailment can be brought down to zero or near zero with additional storage or other measures, the cost of these different approaches needs to be evaluated and as explained below, in some cases curtailment may be the most effective, and even the most economic, use of resources.

How much curtailment is manageable is a question that factors into new project feasibility. The California Public Utilities Commission (CPUC) has approved a variety of different curtailment risk-sharing clauses for new renewable energy contracts. Moving forward, developers may need to adjust prices for new projects so that they can recover costs, and some changes may impact financing and necessitate changes to future contract terms depending on how curtailment levels and risks are treated.

2.3 Solar Versus Wind Integration

While other BAAs (ERCOT – Electric Reliability Council of Texas, MISO – Midcontinent Independent System Operator, and PSCO) have successfully integrated high penetrations of wind, California (and perhaps Hawaii) is paving new ground in the U.S. in integration of high solar penetrations. In Europe, Germany has the most extensive experience with integrating high solar penetrations. The recent solar eclipse on March 20, 2015 proved Germany's ability to balance high solar penetrations, when they managed a 4,370 MW ramp in PV output over a 15-minute interval, without curtailing[10].

While integration of wind and solar into the power system share some of the same challenges in terms of variability and uncertainty, they also differ in some respects. Solar integration is more about ramp management while wind integration is more about ramp prediction. Managing the solar profile requires the ability to turn down or shut down generators midday, in accordance with the sunrise ramp, having room on the system to accept a large amount of solar generation, and the ability to ramp up or start up generators in the evening, in accordance with the sunset ramp. Managing wind requires the ability to hold adequate reserves to manage wind down ramps and the ability to turn off or turn down generators when wind output is high.

The Western Wind and Solar Integration Study Phase 2 investigated the differences between wind and solar integration by comparing a high solar, high wind and balanced solar/wind scenario. While it found that overall production costs and emissions were similar across the three

[10] B. Ernst, "Solar Eclipse and PV Impact on European Power Systems," Utility Variable Generation Integration Group Spring Technical Workshop, Minneapolis, MN, April 2015.

scenarios, it did note that the balanced mix of solar/wind resulted in the lowest production costs and curtailment[11].

2.4 Mitigation Options

There are many mitigation options to help California integrate higher penetrations of VERs. In this paper, a handful of options, that can be reasonably implemented while having a significant impact either today or in the longer term (our to 2030) with higher VER penetrations, are explored. Section 3 discusses how CAISO can best position the system in the day-ahead. Section 4 explains how flexibility and price-responsiveness may be extracted from the following resources in real-time:

- Interties
- Thermal generators
- Combined Heat and Power (CHP)
- Economic dispatch of VERs
- Load

Section 5 investigates shutting down thermal generators during low load, high VER output hours. The thermal generators may have been providing both energy and ancillary services. While the energy may be sourced from the VERs, the ancillary services may be sourced from the following resources, allowing the thermal generators to shut down:

- Grid-friendly VERs
- Reserves from storage
- Reserves from demand response

Finally Section 6 summarizes the various options and their ability to help the system better integrate VERs.

[11] D. Lew et al, "Western Wind and Solar Integration Study Phase 2", NREL/TP-5500-55588, Sep. 2013. http://www.nrel.gov/docs/fy13osti/55588.pdf.

3 POSITIONING THE SYSTEM

In typical utility operations, the day-ahead time-frame is when the system is positioned using the best available forecast information on load and VERs. It is positioned not only to meet the load, but also to survive any contingencies, handle forecast errors, and manage VER variability. This section describes options that can help integrate VERs in this time-frame, including:

- VER participation in the day-ahead market
- Use of VER forecasts to determine reserve requirements
- VER forecast improvements

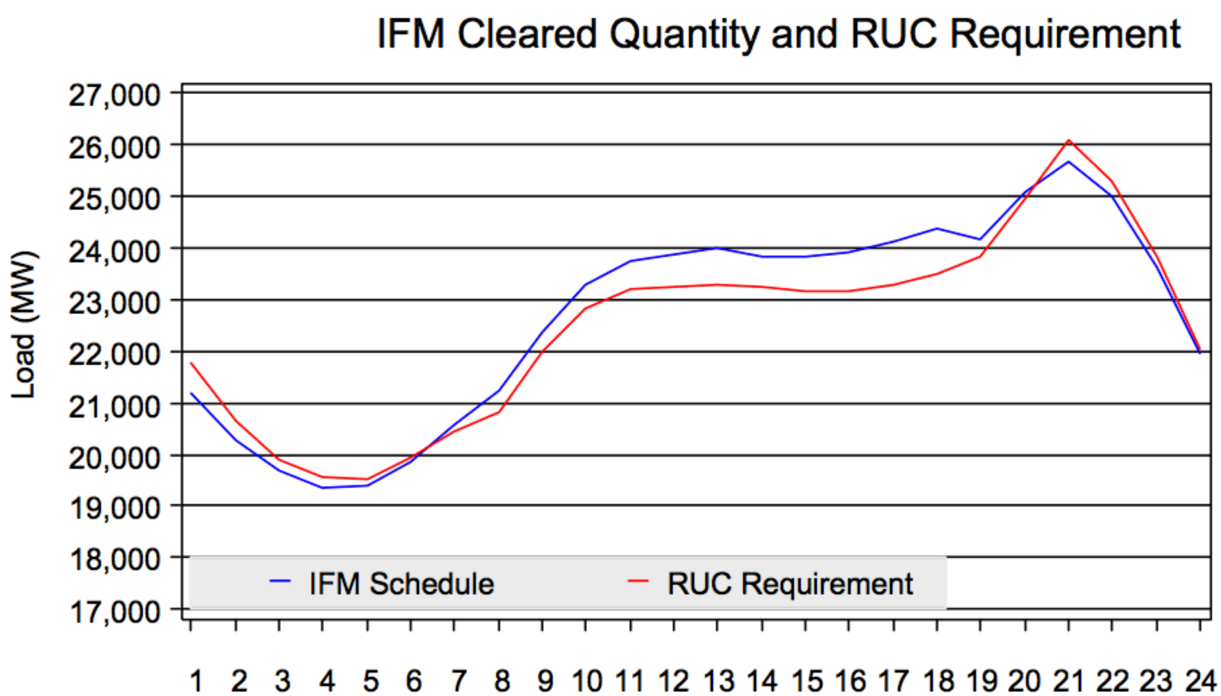


Figure 6 – CAISO day-ahead market processes for April 12, 2014, showing the IFM schedule and the RUC requirement.
Source: CAISO, Daily Market Watch, April 12, 2014.

CAISO has a day-ahead (DAM) and real-time (RTM) market. The DAM includes an integrated forward market (IFM) to clear bids and a residual unit commitment (RUC) to ensure adequate capacity to meet demand. The IFM clears bid-in supply against bid-in demand; in contrast the RUC clears physical supply against CAISO's forecasted demand (Figure 6 shows the IFM and RUC for April 12, 2014, the same day shown in Figure 4 and Figure 5). The RUC is a reliability process. It uses the day-ahead (DA) VER forecast and CAISO's DA load forecast and ensures adequate capacity if the IFM is low. However, the RUC does not decommit if the IFM schedules are high. This means that CAISO will schedule more generators than the IFM if CAISO's load and VER forecast show a need but it won't remove generators from the IFM results.

In the RTM, VER output can be higher than was scheduled in the DAM. This may be due to VERs that did not participate in the DAM or DA forecast errors. Flexibility will need to be extracted from the rest of the units that are online in accordance with the bid stack. A power balance issue in which there is too much supply, can lead to negative prices and the use of down reserves. This can degrade into an over-generation condition.

The mechanism that reduces over-commitment is convergence bidding, which helps to converge DA and real-time (RT) prices. Virtual bids for supply may be made in the DA. This appears to the system as a load in the RT. As mentioned above, the RUC checks the IFM results, so if there is too much virtual bidding of generation, the RUC will commit additional capacity. If the opposite occurs and there are more VERs in RT than were scheduled in the DA, and there is little virtual bidding of generation, then flexibility issues may arise in RT.

While much of the solar is generally scheduled in the DAM today, a significant amount of wind is not. Depending on VER levels and virtual bids on a given day, there could be over-commitment of generation that in turn could lead to operating challenges. Therefore, this paper recommends that CAISO investigate options to encourage market participants to either self-schedule or economically offer as much of their VER output into the DAM as possible. For example, MISO has found ways to reduce risks for VERs in the DAM such as allowing VERs to give a 4-hour-ahead update to their DA bid/offer (see text box on page 3-22). Because 4-hour-ahead forecasts are more accurate than DA forecasts, this allows for more accurate scheduling and helps to mitigate risks associated with forecast error. VER participation in the DAM allows for transparency, and for the market to do its job. Another option (or one that could be implemented today while VER participation in the DAM is encouraged) could be to have the RUC fulfill an economic as well as reliability function, so that it decommits generation in the DA based on forecasts. Being willing to decommit in the RUC and intra-day[12], based on updated VER forecasts, can help make flexibility available.

Another area to investigate is how VER forecasts feed into reserve requirements. The industry as a whole is learning how to more effectively integrate VERs based on forecasts. For example, it may be fairly easy to forecast that a wind ramp-down will occur but more difficult to predict the exact timing of the ramp. Managing a ramp phase error by holding extra reserves a couple hours before and after the forecasted ramp could be a mitigation option. Similarly, it may be fairly easy to forecast partly cloudy, windy conditions (that lead to high solar variability), but more difficult to predict the exact solar output at any given hour. Holding extra reserves on high solar variability days (and fewer reserves on clear, cloudless days) could be a solution. Today, assessment of forecast errors for the California system and reserve levels or operational steps to mitigate those errors could be useful. In the future, incorporating stochastic forecasts and tools into operations may be helpful. Analysis on the interplay between forecasts, look-ahead in the intra-day unit commitment, and the asymmetric ramping capabilities of the system would help to optimize positioning of the system in a high VER future.

[12] Short-term unit commitment process that looks ahead 3 hours beyond the trading hour.

Improvements in DA and intra-day VER forecasts would also be helpful. CAISO reported that the mean absolute error of DA aggregate wind and solar forecasts in 2014 was 7%[13]. VER forecasting, especially solar power forecasting, is a relatively new science, especially in comparison to the decades of experience that the utility industry has in load forecasting. Additionally the numerical weather prediction models were originally designed to provide information about temperature and precipitation at the ground level, not cloud movement nor wind speeds hundreds of feet above the ground. Accurate forecasting of behind-the-meter, distributed PV is also essential, because that generation is automatically netted from load. Efforts are continuing at the national as well as state level to evolve models and techniques to improve VER forecasting. These should be monitored and incorporated into CAISO operational practices as appropriate.

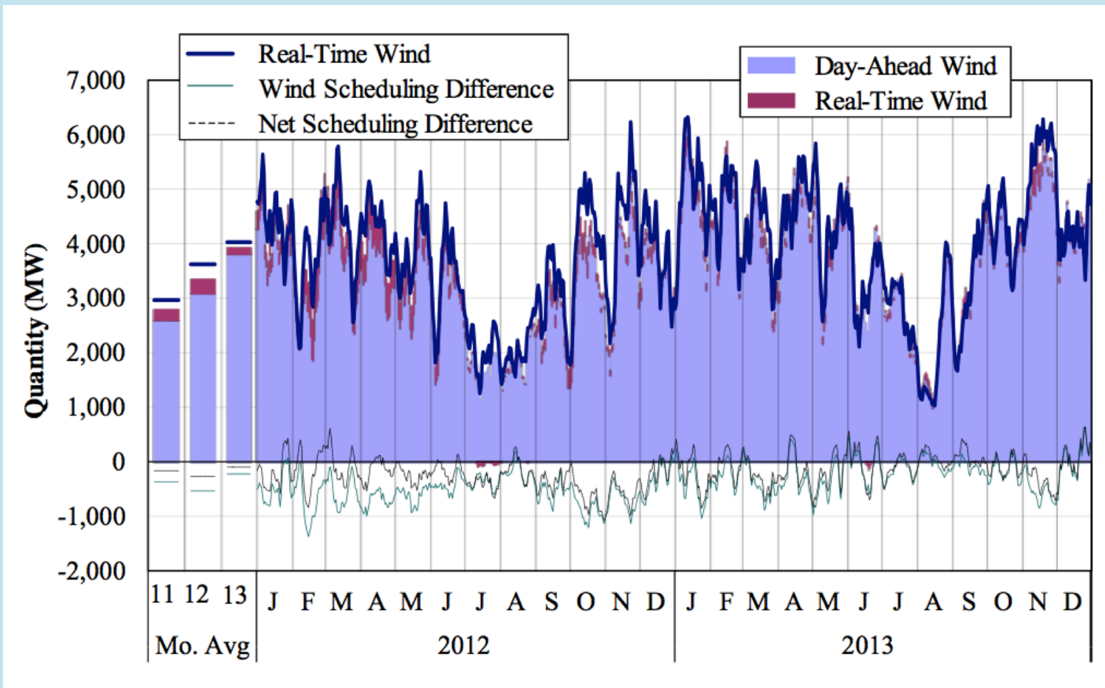
In summary, these are key recommendations to the day-ahead operational processes that position the system to better integrate VER generation:

- Find ways to encourage the scheduling of as much VERs as possible in the DAM. Also investigate decommitting generation in the RUC based on VER forecasts.
- Investigate conditions under which VERs stress the system (e.g., ramps, forecast errors, etc) and assess how reserve requirements or look-ahead in the unit commitment can better position system in the future.
- Continue to support, monitor, and incorporate improvements in DA, intra-day and behind-the-meter VER forecasts.

[13] California Public Utilities Commission Rulemaking 13-12-010, Order Instituting Rulemaking to Integrate and Refine Procurement Policies and Consider Long-Term Procurement Plans, Reply testimony of Shucheng Liu, CAISO, Oct. 22 2014.

VERs in MISO and the Dispatchable Intermittent Resources Program

MISO has successfully integrated wind into their DAM and RTM. Wind is required to submit DA forecasts so that MISO can maintain reliability. These DA forecasts are not financially binding. Wind may provide a 4-hour-ahead update to its DA bid/offer, using 4-hour-ahead forecasts which are more accurate. This helps to mitigate DA forecast error risk. The reduced risk for wind in the DAM combined with slightly higher prices in the DAM have resulted in the fact that nearly all of the real-time wind is scheduled in the DAM (see graph below from the 2013 State of the Market Report conducted by MISO's Independent Market Monitor, Potomac Economics).



With a few exceptions, new wind plants in MISO are required to participate in the Dispatchable Intermittent Resources (DIR) Program. In real-time, DIR are required to submit 12 forecasts in 5-minute intervals for the next hour. The DIR can modify its forecast up to 10 minutes prior to the start of each interval, which allows it to use high accuracy forecasts. DIRs can either self-schedule (be price-takers) or submit offers into the RT market. The DIR is dispatched at 5-minute intervals by MISO at or below the forecast maximum limit. The DIR tariff has make-whole provisions if the actual energy dispatched does not fully compensate the DIR for its costs. There are penalties if the DIR performance is more than +/-8% from its dispatch set point for at least four intervals in a row. DIR has been hailed as a successful means to reduce manual curtailments that previously managed congestion or over-generation, and to control wind resources and respond to wind variability.

4 FLEXIBILITY AND PRICE-RESPONSIVENESS FROM THE SYSTEM

In the DA, CAISO positions the system based on forecasts and market bids. During the day of operation, CAISO needs to refine its positioning based on several-hour-ahead forecasts through the short-term unit commitment process. Then in real-time, CAISO needs to balance the system, despite forecast errors, short-term variability of VERs, and any potential contingencies. This section assesses slices of the stack from Figure 3 for RT flexibility. The two guiding principles to address the duck curve are:

- *As much of the system as possible should be responsive to RT price, and*
- *As much of the stack as possible should be as flexible as possible.*

Each slice of the non-dispatchable stack from Figure 3 was examined to see how it could provide more flexibility and RT price-responsiveness. To accurately quantify the contribution from increased flexibility of each slice of the stack, more data, analysis and modeling is necessary. But insight can be gained by examining the drivers behind inflexibility and how much flexibility may be attainable with institutional or technical fixes. The most promising areas identified were:

- Interties
- Thermal generator uprates
- CHP
- VER economic dispatch
- Load

The RTM includes a fifteen-minute market (FMM) to commit fast start units, and determine financially binding FMM schedules and locational marginal prices (LMPs), and a 5-minute RT economic dispatch. The FMM is helpful for VERs in that VER scheduling is now based on a 15-minute forecast (which is calculated as the average of three 5-minute forecast intervals) instead of hourly. VERs with an economic bid have an upper dispatch limit set to that 15-minute forecast and self-scheduled VERs receive self-schedules determined by that forecast. The forecast data used in this process is determined at 37.5 minutes before the interval. VER forecasting typically uses a combination of weather models and statistical methods in the longer (day-ahead) time frames but in the shorter (5-10 min ahead) time-frames, persistence[14] forecasts are often better than statistical and other methods[15]. The 37.5 minute-ahead time-frame *does not* benefit from the weather models nor the higher accuracy of very short-term persistence forecasts. Shortening this 37.5 minute-ahead lead time would result in a more accurate schedule for the VERs which could ease balancing issues for the CAISO. In the RT dispatch, 7.5 minute-ahead forecasts are used to determine upper bounds for the VER output, unless there is an economic bid to reduce the level to below the forecast. This 7.5 minute lead time *does* benefit from the higher accuracy short-term forecast.

---

14 Persistence forecasts assume that the weather in the next interval will be the same as the weather in the current interval.

15 Mark Ahlstrom, NextEra Energy Resources, personal discussion, April 3, 2015.

4.1 Interties and Self-schedules

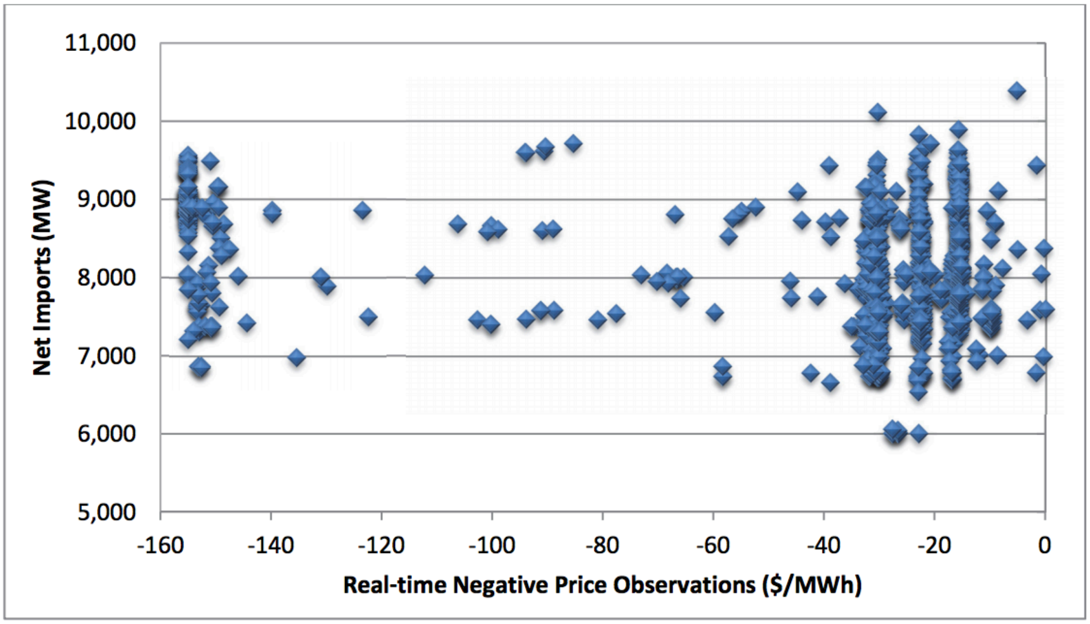


Figure 7 – Net imports vs RT negative prices, May 1 – Sep 9, 2014. Source: California Public Utilities Commission Rulemaking 13-12-010, Order Instituting Rulemaking to Integrate and Refine Procurement Policies and Consider Long-Term Procurement Plans, Reply testimony of Shucheng Liu, CAISO, Oct. 22 2014.

California currently relies heavily on imports. For example, in 2014, electricity generation serving California was nearly 300,000 GWh, of which one-third was net imports[16]. Imports help to keep California's energy costs low. Imports vary from jointly owned units, imports from other utilities via bilateral contract, renewable energy resources to help meet RPS targets, economic bids in the market, dynamically scheduled units, etc. and historically, exports have been far smaller than imports.

Today, CAISO can have significant amounts of net imports even during periods of negative prices as seen in Figure 7, showing that imports are largely insensitive to RT prices. This data can also be viewed from the perspective that part of the reason that negative prices are occurring is that net imports are so high. Certainly there are some imports that can't provide flexibility, e.g., the Palo Verde nuclear plant. And while hydro can be a very constrained resource due to irrigation, environmental factors, recreation and other needs, it could be explored further to see if more flexibility can be extracted. However, most thermal plants should have the capability to be sensitive to price. Lack of price-sensitivity of imports is explored below.

16 Imports minus exports

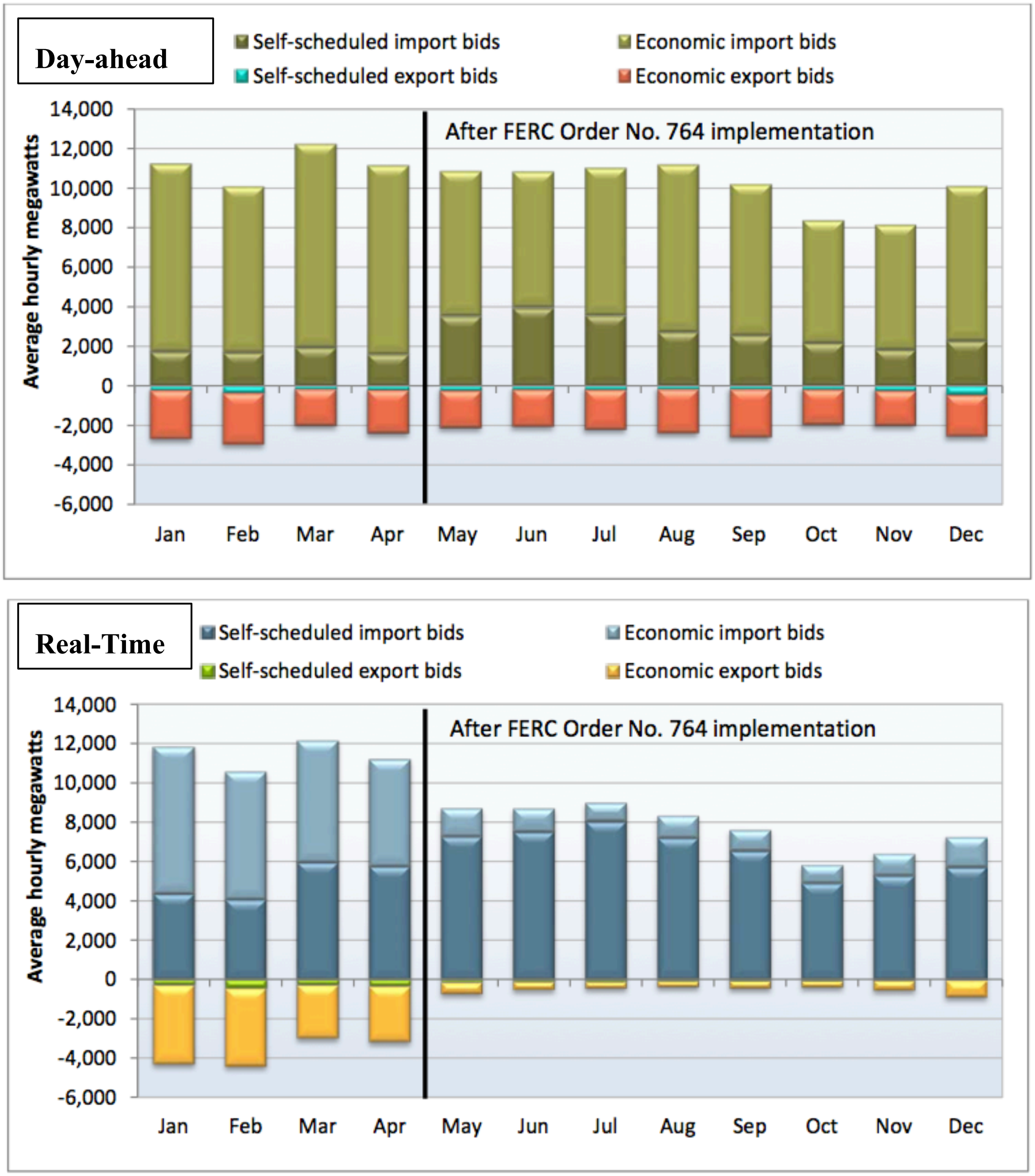


Figure 8 – CAISO imports/exports that are self-scheduled versus economically bid into the DA (top) and RT (bottom) markets in 2014.
Source: CAISO Department of Market Monitoring, "Q4 2014 Report on Market Issues and Performance," Mar. 3, 2015.

Imports/exports are bid into the DAM as self-scheduled or economic bids. Once the bids are accepted, they can be self-scheduled or re-bid in the RTM. The FMM began in May 2014, to comply with FERC Order 764 to offer 15-minute transmission scheduling to better accommodate VERs. An unintended outcome of the FMM is that starting in May 2014, most of the import and export bids in the hour-ahead market have been self-scheduled as opposed to economically bid into the market. This is shown by the dark blue bars in Figure 8 which represent the self-scheduled imports in the RTM and comprise most of the imports in the RTM. A possible reason for the increase in self-scheduling of imports in the RTM is that for those resources that cannot change within the hour, the new FMM and the fact that RT interties settle against 15-minute prices exposes them to potentially lower prices during some of those 15-minute intervals. Instead, those who clear the DAM choose to self-schedule in RT and clear the RTM consistent with their DA award.

In addition to the old options for RT intertie bids (self-schedule, economic hourly block bid, dynamic schedule) the new market also allows for an economic hourly block bid with a single intra-hour economic schedule change and a 15-minute economic bid. The faster interval bids should help with integrating VER, but CAISO reports that in May and June 2014, 96% of intertie bids were hourly blocks, most of which were self-scheduled. Only 1.5% of bids were 15-minute economic bids.

Another part of the issue is that transmission was traditionally scheduled hourly across the West, so CAISO's trading partners who may still be on hourly schedules may not be able to change schedules on 15-minute intervals. However, Bonneville Power Administration, which exports to California, did change to 15-minute scheduling in October 2014 but this does not seem to have had much of an impact.

The increase in self-scheduling of imports also occurred, but less dramatically, in the DAM as can be seen by the higher dark green bars starting in May in Figure 8. In January and February, the average amount of imports clearing the DAM is slightly less than those clearing the RTM. In April and May, however, the average amount of imports clearing the DA is slightly greater, showing that on average, in RT, fewer imports are economic than were thought to be in the DA. If the import bids that are accepted in the DA turn into self-schedules in the RT then not only is there less liquidity in the RTM, less economic bids can get locked in as well.

Similarly, self-scheduling of exports also rose significantly, from 8% of the RTM in January-April 2014 to 42% in May-June 2014. Because export volumes are small, this can't be discerned in these figures.

The end result of all of this is that there is significantly reduced liquidity in the import/export market. This lack of flexibility reduces CAISO's ability to balance the system. Ideally imports/exports would be dispatched as close to RT as possible and as often as possible to accommodate VER variability. However, the significant market change in May 2014 which tried to achieve such, resulted in further insensitivity to RT prices and very few participants in the faster FMM. This underscores the difficulty of establishing new markets and rules to incent a particular outcome.

The very high levels of self-scheduling are a key reason for CAISO's difficulty in system balancing and in accommodating higher levels of VERs. As noted earlier, imports may clear the DAM and avoid price volatility in the RTM by self-scheduling in RT. As a result of the California energy crisis, the CPUC encouraged the investor-owned utilities to hedge costs with self-supply and bilateral contracts. Understanding what exactly needs to change in CAISO to reduce the levels of self-scheduling requires understanding the motivations of market participants and further research. It may also be worth examining whether overall prices might decline if the market, rather than self-schedules, were allowed to determine generator dispatch.

This is a complex issue but worth investigation because of the very large amounts of self-scheduling. Real-time self-scheduled imports during May-September were about 7 GW. Table 1 shows current imports from jointly owned units, dynamic schedules and RPS imports of 3,639 MW. If imports continue to be about 7 GW and no flexibility can be extracted from this 3,639 MW, that leaves about 3,400 MW of potential flexibility from this slice of the stack.

Resources	Capacity (MW)	Type
Palo Verde	811	Jointly Owned Unit
Hoover	768	Jointly Owned Unit
Dynamic Schedules to MUNIs	627	
Pseudo Ties (Renewables)	40	Renewables
Other Dynamic Schedules (wind/solar)	676	Renewables
Dynamic Schedules from IID (Biomass & Geothermal)	518	Renewables
LSEs contractual agreement with BPA	200	Renewables
Total Expected Import	**3,639**	

Table 1 – CAISO imports of jointly owned units, RPS imports, and dynamic schedules. Source: California Public Utilities Commission Rulemaking 13-12-010, Order Instituting Rulemaking to Integrate and Refine Procurement Policies and Consider Long-Term Procurement Plans, Reply testimony of Shucheng Liu, CAISO, Oct. 22 2014.

4.1.1 Energy Imbalance Market

The Energy Imbalance Market (EIM) between CAISO and PacifiCorp recently went live. The EIM is a potentially useful tool because its 5-minute economic dispatch on the interties provides a lot of flexibility. The planned expansion of the EIM to NV Energy, Puget Sound Energy, and Arizona Public Service will help because larger EIM footprints result in more diversity to help all participants more efficiently balance their systems. This diversity is not just the geographic diversity of the wind and solar resources across the participant BAAs, but also the diversity of the load and of the generators that balance the load. To the extent that EIM participants have a flexible system, they could benefit from backing their generation down and taking negative-priced energy from CAISO. In this way, EIM expansion can help mitigate some of the duck curve issues today. The levels to which EIM expansion can help will depend on the volume of flows between participants, which in turn will depend on transfer capacity between participants and levels of imbalance at any given interval.

However looking into the future, it is noted that over-supply in a 40% RPS scenario is projected to be thousands of MW deep (see Figure 9). By definition, the EIM is a RT market and deals with RT imbalance, not unit commitment. Because EIM participants are committing their own units in the DA, they are highly unlikely to be holding imbalances on the order of thousands of MWs. However, the possibility of PacifiCorp joining CAISO in the full DAM would help significantly by expanding the size of CAISO's BAA, thus increasing the diversity in load and VERs as well as increasing the number of generators available to balance the system. Expansion of CAISO's BAA can help mitigate the duck curve in the longer term.

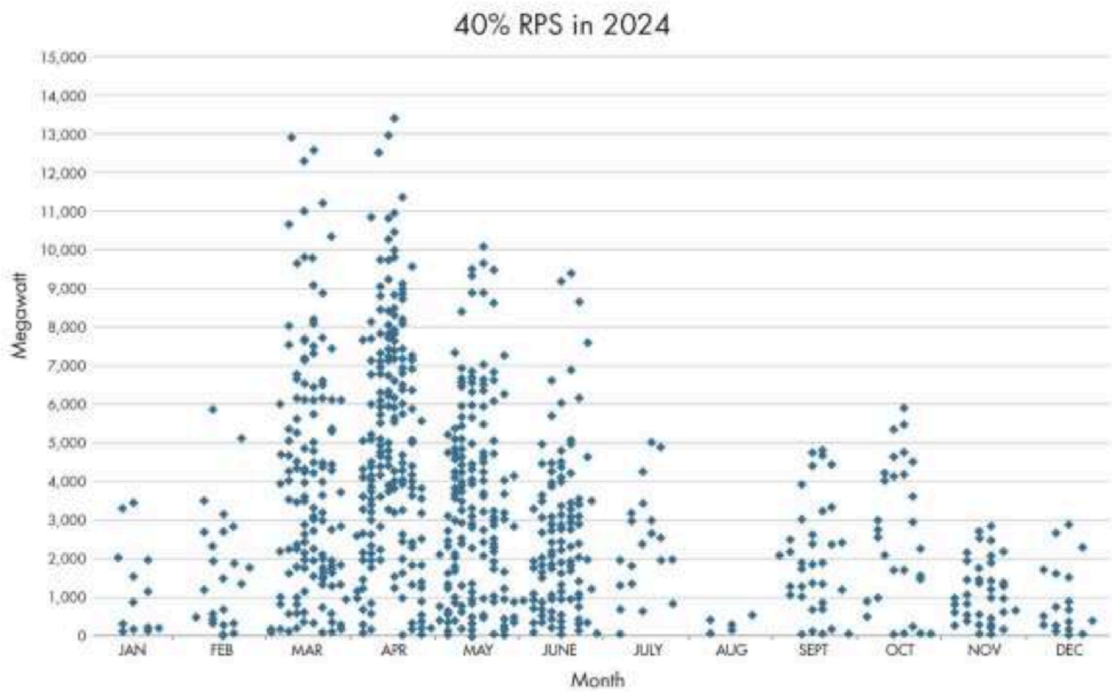


Figure 9 – Curtailment from CAISO modeling of 40% RPS Scenario for 2024. Source: FERC Docket AD15-4-000, Technical Conference on Environmental Regulations and Electric Reliability, Wholesale Electricity Markets, and Energy Infrastructure, Statement of Mark Rothleder, CAISO (Feb. 19, 2015).

4.1.2 Exports

Over-supply from high VER output tends to be a greater issue during periods, which simultaneously have low loads and high VER output. Today, those periods tend to be weekends during the spring. Figure 9 shows that in a future with high VER penetrations, the fall, winter and early summer may also start to experience over-supply. A potential mitigation option is to add in negative slices to the stack from Figure 3 during the middle of the day. Essentially, this would be scheduling in of exports during likely over-supply conditions to reduce the floor of the stack. Market participants could do this today for spring weekends when clear skies are forecast, but it may require the ability to quickly start a generator if the forecast is inaccurate. The system is likely to have asymmetric ramping capabilities and constraints during different time periods, i.e., during the spring, CAISO may be more likely to run out of downward capability midday and upward capability in the evening. Analysis of these capabilities would help to determine whether the risks of actions, such as scheduling exports midday, are worth the potential benefits.

In the future, when negative prices are low enough and frequent enough, scheduling exports during this time may become a regular business opportunity for market participants.

4.1.3 Flexible Resource Adequacy Capacity Requirement

To meet CAISO's recent FRAC requirement, utilities must offer into the market a certain amount of capacity to provide flexibility especially to meet ramping needs. There are three categories of FRAC capacity: base, peak and super-peak. Base ramping resources must be able to start at least twice a day, provide at least six hours of energy at their effective flexible capacity value, and submit economic bids for energy and ancillary services from 5 am to 10 pm. FRAC resources need to be dispatchable in the 5-minute real-time dispatch.

Currently interties are not included in the FRAC. Interties cannot be dispatched at 5-minute intervals and these resources first need to qualify for the Resource Adequacy requirement. However, imports provide a significant amount of energy and their insensitivity to price can provide balancing challenges. Finding ways to allow interties to contribute to some part of the FRAC requirement could help to pull more flexibility from these resources and is worthy of further analysis.

4.2 Thermal generators

As the net load curve increasingly has a double peak, especially during low load winter and spring seasons, CAISO will need several, if not all, of the following features and functions from existing (and new) thermal units:

- **Ability to start/stop with short minimum downtimes** – Dispatchable generation may be called upon to operate in the morning to meet the morning load rise, turn off midday if high VER levels are forecast, and then turn back on in the late afternoon/evening to meet the evening load rise.
- **Startup reliability** – System operators may be reluctant to decommit units within the day, because of low confidence in start reliability. This challenge can be overcome with implementation of rigorous re-start procedures and the demonstration of reliable restarts. Start-ups need to be *reliable and repeatable*.
- **Minimum generation levels** – Generators may need the capability to turn down to lower levels to accommodate high levels of VERs, but stay online to provide ancillary services or meeting the load ramp later in the day. This also allows units to remain online to provide frequency response while simultaneously reducing contribution to the midday stack.
- **Range of ramping capability** – The duck curve does not necessarily require fast ramping capability (since the ramp is over a 2-3 hour period) but rather standard ramping rates, with expanded range over the multi-hour load ramp.
- **Reduced startup and part load emissions** – California has several air quality zones with ozone and other emissions issues.
- **Reduced carbon dioxide emissions** – Features or capabilities that improve thermal generator efficiency are useful in meeting CA's GHG reduction targets.

Combined cycle plants can ramp from minimum to maximum generation well within the time of the neck of the duck curve. A conventional 2 x 1 (2 combustion turbines x 1 steam turbines) F-class combined cycle plant installed in the early 2000's could provide its full range of capability from minimum generation (defined by emission permit requirements and after-treatment capabilities) with both combustion turbines to maximum generation in less than ten minutes. For example, a 530 MW combined cycle plant may have a minimum generation level of 50%, or 265 MW. Each combustion turbine provides a very modest 15 MW/minute ramp rate, plus the steam turbine adds about half of that again, resulting in a total ramp rate of 37 – 38 MW/minute. This ramp from minimum to maximum generation takes about seven minutes.

Since CAISO allows multi-state offers, the minimum generation level of this type of 2 x 1 combined cycle plant with the steam turbine and only one combustion turbine operating would be about 35% of the plant's maximum generation level. Therefore the MW ramp range of this single plant would jump to 345 MW. This ramp capability would be well within the two hour time-frame needed since the first combustion turbine and steam turbine would already be operating.

Combined cycle plants are intrinsically designed to stop and restart much more often than many other thermal generators. The primary hurdle in the plant start-up is the thermal soak required within the heat recovery steam generator (or boiler). Since their installation in the early 2000's, utilities across North America have learned from their experience with over 300 existing F-class combined cycle power plants, which meticulous procedures are required to optimize their plant start-up from "off" and "cold" up to maximum generation in well less than four hours. The collective experience, along with enabling software, could be applied to all F-class combined cycles within CAISO to enable multi-day, reliable and repeatable starts for the seasonal time period of the "duck curve". A combined cycle plant may be "off" but "hot", such as after turning the plant off during the morning solar generation ramp. New procedures and control software now enable this same combined cycle plant to reach full load in approximately 75 minutes, which is more than adequate to serve the 2 – 3 hour load ramp in the evening.

Thermal plant cycling has impacts on costs and emissions. Generic cost ranges for hot, warm and cold starts and ramps of combined cycle, combustion turbine and coal plants have been estimated, based on in-depth cycling studies of over 400 plants[17]. Emission rates of CO_2, NO_x, and SO_2 from starts, ramps, and operation at minimum generation levels have been extracted on a unit-by-unit basis for nearly all of the thermal generators in the US[18]. In 2013, the Western Wind and Solar Integration Study Phase 2 simulated system operations across the West with these datasets to examine impacts of VER-induced cycling. The results showed that cycling had very little impact on emissions. Cycling costs, if attributed to the thermal generators, increased by $0.47 – 1.28/MWh, which is relatively small when compared to total fuel and O&M costs of $27 – 28/MWh[19].

---

17 N. Kumar, et al, "Power Plant Cycling Costs", NREL/SR-5500-55433, April 2012. http://www.nrel.gov/docs/fy12osti/55433.pdf

18 D. Lew et al, "Western Wind and Solar Integration Study Phase 2", NREL/TP-5500-55588, Sep. 2013. http://www.nrel.gov/docs/fy13osti/55588.pdf.

19 Ibid.

4.2.1 Uprates of thermal generators

One technical option which may enable many, if not all, of the features and functions noted above would be to "uprate" the existing combined cycle fleet in CAISO. There are many uprate options for combined cycle plants that were installed in the early 2000's, which enable creative methods to provide flexibility, including lower minimum generation levels. After the uprate, a combined cycle plant would have a 45% minimum generation level running in 2 x 1 mode, and a 30% minimum generation level running in 1 x 1 mode. Uprates can reduce the minimum generation level and increase the maximum generation level to help accommodate the duck curve, and increase overall generator efficiency, to help meet state GHG reduction targets.

Generators with new power factor designs can help mitigate some of the challenges associated with high levels of VER's. For example, the thermal generator could increase reactive power output to further decrease real power output as the solar output increases midday. This can reduce real power output while maintaining a reasonable loading on the turbine to keep emissions within permit levels. This could be reversed in the late afternoon when the system needs real power. Reactive power provision is a very localized need, so trading real for reactive power would need to be carefully analyzed to see if it is a viable option.

As discussed above, starts need to be reliable and repeatable in a high VER future. This can be achieved through fleet experience and process improvements. The midday period when combined cycle plants may be shut down to accommodate the duck curve, is at the edge of the "hot start" time frame. This means less "cyclic stress" on the plant and start times to minimum generation approach one hour. With high solar penetrations, two starts per day are not necessarily detrimental, as long as maintenance regimes are upheld.

Emissions are also reduced with the uprates. New and more advanced dry low NO_x combustion systems installed on the combined cycle fleet not only enables many of the features identified above, but could also address some of California's local ozone issues. Power latent heat rate improvements can also be undertaken to increase plant efficiency and reduce GHG emissions.

The existing F-class combined cycle fleet residing within the CAISO footprint is slightly over 15,000 MW. The average minimum generation level of this fleet in 2 x 1 mode is 55% of the maximum generation level. The fleet's current ramping range (from minimum generation to maximum generation) is about 7,000 MW. With potential uprates, the expanded ramping range, if the fleet were running in 2 x 1 mode, could increase to 9,500 MW. In 1 x 1 mode, the range increases from 9,500 MW to 11,500 MW. *This potentially extra 2,000 – 2,500 MW of range helps to address the neck of the duck curve.* The expanded maximum output would increase to about 17,000 MW, which helps to meet the head of the duck curve. Also, uprating an existing combined cycle power plant is typically faster (due to permitting and component delivery process) and less expensive than procuring and constructing any new generation site.

While this 2,000 – 2,500 MW of expanded range may be technically possible and while the institutional challenges may less formidable than other options in this paper, the challenge is that the generator that bears the cost will not necessarily benefit economically from the uprate. In fact, reducing minimum generation levels and increasing start-up reliability, could result in the generator being run less. An assessment of costs and benefits of flexibility retrofits for combined cycle plants undertaken for a sample system in the West showed that while retrofits made sense on a system-level, they did not necessarily make sense on a unit-by-unit level.[20] Whether the existing FRAC requirement and energy and ancillary service markets are adequate to incent thermal generator upgrades needs to be assessed.

Finally, there does not appear to be much opportunity to take advantage of Once-Through Cooling (OTC) retrofit downtime to do thermal uprates, as most of the OTC retrofits will occur at very few combined cycle power plants, which do not have proven, available uprate options.

4.3 Qualifying Facilities/Combined Heat and Power

California's 8,000 MW CHP fleet represents a portion of the non-dispatchable stack that can be difficult to back down midday. Emissions issues at part-load have plagued the CHP fleet in particular, making flexibility from CHP difficult. Many of the CHP units are small but some can be in the 100 – 300 MW range. California's CHP fleet has already added a significant degree of flexibility in recent years under a 2011 negotiated settlement agreement between the parties. This settlement introduced an option for the CHP owner and their utility off-taker to switch to an approved Utility Prescheduled Facility tariff option, which includes both baseload and peak/off-peak rates. This was the first major change to contract conditions since the original long-term PURPA Standard Offers from the 1980s and was intended to incentivize more economically rational operation for units with operational flexibility. Approximately 25% of the fleet (2,000 MW) has migrated to this new tariff option to date as a result – a report from CPUC staff is expected soon that will provide more precise figures. *An additional potential of up to a further 25% of fleet flexibility (2,000 MW) may exist to be unlocked through further modification of contract incentives*.

However, even with better pricing, technical and operational constraints limit the ability of most CHP facilities to ramp up and down widely within their operating range. In addition to power, the facilities are designed to provide baseload steam to a thermal host. Replacement steam for the thermal operation is therefore required in the event the output of power generation is constrained. Where auxiliary boilers exist to provide an alternative source of steam, these must burn natural gas, adding both variable expense and emissions, which reduce the environmental attractiveness of the CHP resource. Moreover, in many cases, sufficient boiler capacity is not in place to meet the entire steam requirement, because the CHP is sized to provide the primary source of steam, meaning the thermal host would not be able to operate normally without additional capital investment.

4.4 Economic Dispatch of VERs

Economic dispatch of VERs in the RTM exposes VERs to the negative prices which can help CAISO balance the system when there is an over-supply of generation. There may be a concern, however that moving to economic dispatch of VERs from "last resort" manual curtailments may lead to more frequent curtailment. However by having this granular level of control, CAISO may

---

[20] S. Venkataraman, et al, "Cost-Benefit Analysis of Flexibility Retrofits for Coal and Gas-Fueled Power Plants, NREL/SR-6A20-60862, Dec. 2013. http://www.nrel.gov/docs/fy14osti/60862.pdf

actually curtail VERs less and release the VERs sooner than they would if they were manually block-curtailing output. Other ISOs have moved to 5-minute economic dispatch of wind with satisfactory results. The text box on page 3-22 explains MISO's DIR program, which economically dispatches wind on a 5-minute interval and includes make-whole provisions.

Because tax credits, renewable energy credits and other incentives are outside the CAISO market, VERs may still be able to earn revenue at slightly negative prices, in contrast to conventional generators which pay to generate. This means that VERs may be the last to back down and may not be incented to back down until prices are sufficiently negative. Backing down of VER production is essentially curtailment and while VER curtailment needs to be carefully considered (VER generation helps California meet its GHG reduction goals; curtailment hinders cost recovery), economic VER dispatch is also an important tool to balance the system and to facilitate integration of additional VERs.

At low VER penetrations, it is less important for the system operator to be able to control VER output. Curtailment may be zero or extremely low at those penetrations as well. At high penetrations, it is very important for the system operator to be able to control VER in order to balance the system and some curtailment will necessarily be part of the solution. Existing VERs may not necessarily have the capability to be economically dispatched. *Requiring new VERs to be economically dispatched can help CAISO with its balancing challenges as California moves to increasing levels of VER penetration.*

4.5 Load

The final piece examined is not really a slice of the non-dispatchable stack but rather the role of load. For load, the duck curve is a huge opportunity. Low-cost energy in the middle of the day was a utility dream just a decade or two ago, when peak demand was in the middle of the day and meeting peak demand was sometimes challenging. The most promising and relatively near term option is to bring customer behavior and load attributes into better alignment with grid needs. Deferrable loads do not necessarily need to be served immediately, upon plugging into the grid, but rather need to be served within a certain time frame. Water pumping loads such as irrigation may fit this bill. Heating and cooling loads such as chillers or cold storage are another option. And this is a great opportunity for electric vehicles (EV), plug-in hybrids, and potential new desalination loads.

The key is to have these loads be responsive to price. Consider the case of EVs: the Governor of California has set a goal of 1.5 million EVs by 2025. Charging algorithms will need to consider forecasted and real time prices. EV owners coming home and plugging their car in at 6pm will exacerbate the duck curve. EV owners allowing the grid to determine when their car is charged as long as it is charged some time over the next eight hours will help mitigate the duck curve. EV owners who don't necessarily need a full charge midday but who plug in anyway and charge only if prices go to zero, may benefit from a win-win situation in which the grid is better balanced and their car is topped off for "free". While the technology and infrastructure to have a sufficient number of smart charging stations connected to CAISO RTM information isn't available today, in the longer-term, this should be possible.

Using the estimates of 3 million EV's on the road in 2030[21], consuming 12 kWh/day, there could potentially be an additional load of 36,000 MWh/day, or 13,000 GWh/year. In comparison, curtailment in CAISO's Long-term Procurement Plan (LTPP) 2024 production modeling was 153 – 407 GWh/year[22]. The worst hour was 13,402 MW of curtailment. So while EVs can't eliminate the worst hour of curtailment, they can reduce overall curtailment using smart charging.

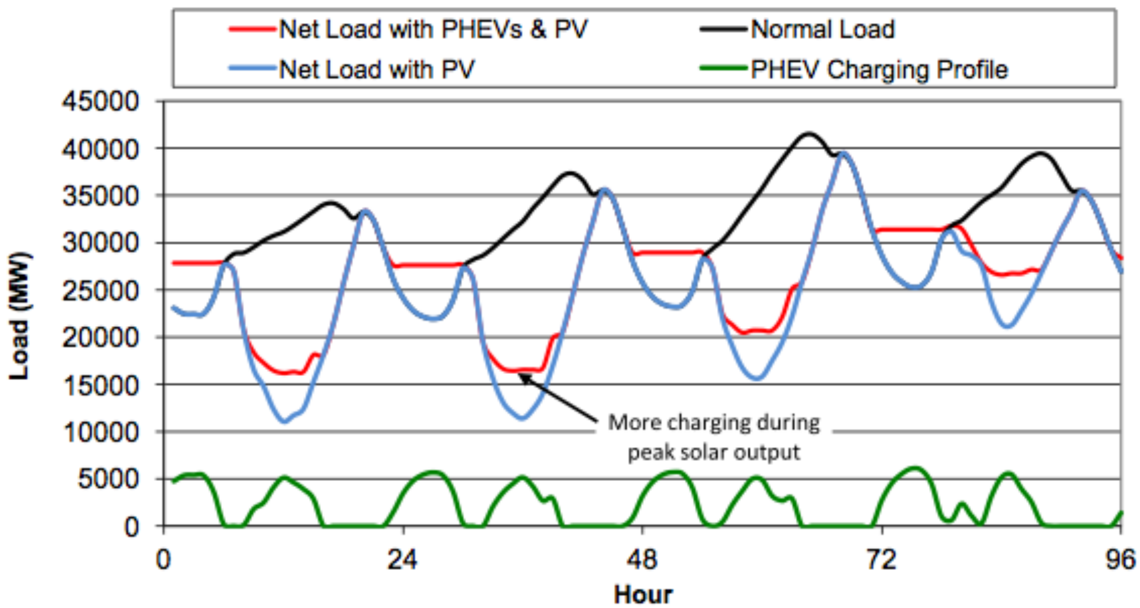


Figure 10 - Scenario of 15% PV penetration with 30% PHEV's during April in ERCOT. Mid-day charging of PHEV's is controlled with a delay of up to 2 hours. Source: P. Denholm, M. Kuss, and R. Margolis, "Co-benefits of large scale plug-in hybrid electric vehicle and solar PV deployment," *Journal of Power Sources*, 236 (2013) 350-356.

In order to determine the mitigation capability of the additional EV load, production simulations of the CAISO system with and without the EVs would need to be run. This could not be undertaken for this paper, but insights can be gleaned from a similar analysis that was conducted to examine the benefits of plug-in hybrid electric vehicles (PHEVs) on high PV scenarios in ERCOT, focusing on reducing PV curtailment[23]. Actual driving patterns and distances in TX were used to create charging profiles and needs. Figure 10 shows the profile of PHEV charging and the resulting net load. The 3 million EVs in California that result in a 4% of load penetration

[21] The Low-Carbon Grid Study extrapolated from 2012 Executive Order goal of 1.5 million zero emission vehicles by 2025 and the California Energy Commission's "more aggressive vehicle electrification scenario" for 2024 to reach the 3 million EV estimate in 2030.

[22] Shucheng Liu, CAISO, Phase I.A Stochastic Study Testimony to California Public Utilities Commission on Order Instituting Rulemaking to Integrate and Refine Procurement Policies and Consider Long-Term Procurement Plans in Rulemaking 13-12-010, Nov. 20, 2014.

[23] P. Denholm, et al, "Co-benefits of large scale plug-in hybrid electric vehicle and solar PV deployment," *Journal of Power Sources*, 236 (2013) 350-356.

discussed above are similar to approximately 18% PHEV penetration in Figure 11. For that PHEV penetration, and assuming controlled charging of vehicles, PV curtailment in a 20% PV scenario, is reduced by nearly 40%. Production simulation analysis of the California power system and driving patterns would be very different than Texas, but this simply illustrates that significant reductions of PV curtailment are possible with EV's on the order of 4% penetration. Vehicle-to-grid would provide even greater benefits when that technology is available.

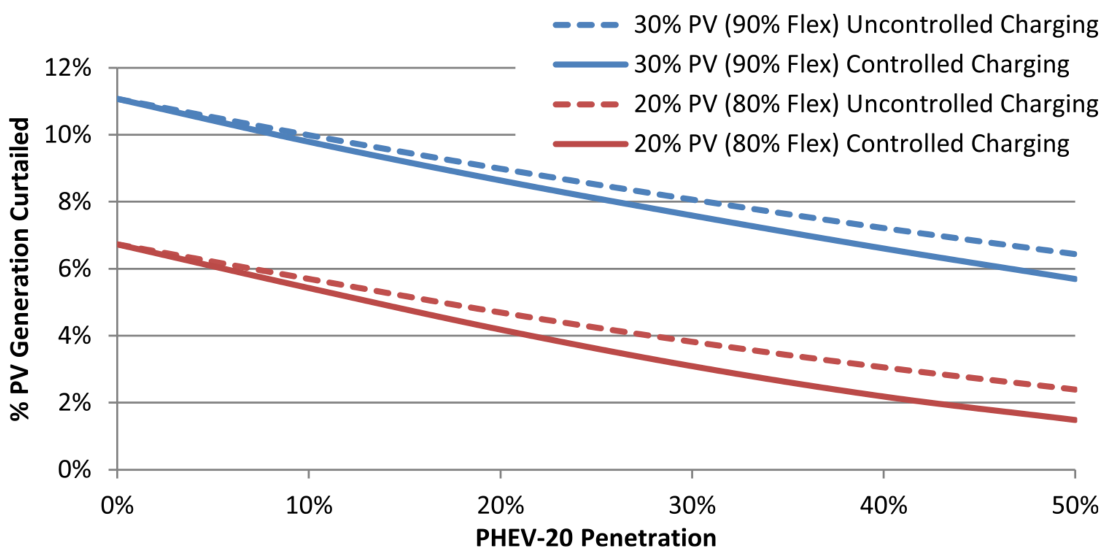


Figure 11 – Decrease in PV curtailment as a function of PHEV penetration. Source: P. Denholm, M. Kuss, and R. Margolis, "Co-benefits of large scale plug-in hybrid electric vehicle and solar PV deployment," *Journal of Power Sources*, 236 (2013) 350-356.

Mitigating the duck curve with EVs is a longer term option. However, several pilots are already underway in California, including a U.S. Department of Defense and Lawrence Berkeley National Laboratory collaboration at the Los Angeles Air Force base.

Today, making energy available to some loads at RT prices could be a win-win situation for CAISO, utilities and loads. End-users can be very creative when it comes to low-priced or free energy (let alone negatively priced energy). Water pumping, pre-cooling or pre-heating of buildings, pre-heating of water heaters, and chillers are only a few of the types of loads that might want to take advantage of this opportunity. While industrial loads are often thought of as the "low-hanging fruit" for demand response, there are likely to be opportunistic loads in the commercial and even residential sectors, as ERCOT is finding with rapidly increasing enrollment in their time-sensitive pricing programs[24].

[24] P. Wattles, ERCOT Market Design and Development, "Retail DR and Price Response 2014 Product Headcounts," Demand Side Working Group Meeting, Feb. 3, 2015.

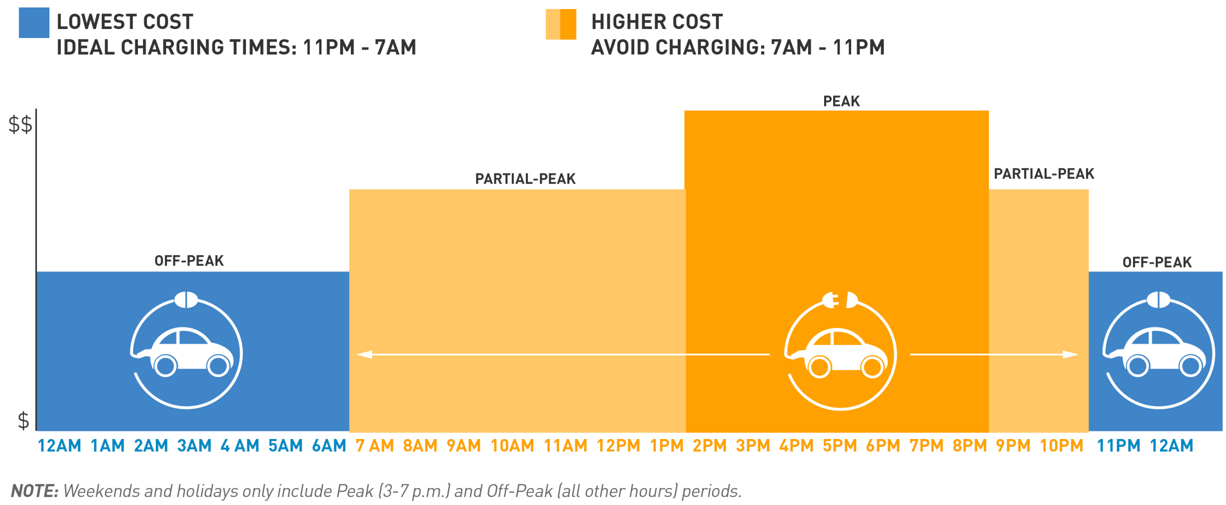


Figure 12 – EV charging rate structure offered by PG&E.
Source: Pacific Gas & Electric, "Electric Vehicles: Making Sense of the Rates," Accessed Mar. 31, 2015. http://www.pge.com/en/myhome/saveenergymoney/pev/rates/index.page

Current business time-of-use or EV charging rate structures charge peak or partial-peak prices during middays in the spring, when over-supply and negative prices are likely to occur (see Figure 12). It will take time and effort to change the definition of peak versus off-peak time periods, especially in terms of re-educating consumers as to the redefined time periods. The grid of the future, with higher levels of VERs, EVs, and demand response will lead to less predictable peak and off-peak periods. And because weather will have an impact on when peak and off-peak occur, defined time periods will be less consistently valid. Therefore, it is important to move to RT price signals, which reflect actual conditions. This will also take time, effort, and new technology, but in the longer term will help greatly to integrate high VER levels and manage a changing grid.

4.6 Key recommendations

To better integrate VERs:

- As much of the system as possible should be responsive to RT price.
- As much of the system as possible should provide flexibility.

Price-responsiveness and flexibility can be sourced from the following parts of the power system:

- Reducing the levels of self-scheduling can help with system balancing and in accommodating higher levels of VERs.
- Expanding the EIM would help provide markets for excess energy that may otherwise be curtailed.
- Expanding CAISO to include other BAAs would help significantly by providing greater diversity (of load, VERs and generators) in both the DAM and RTM.

- Scheduling exports during midday hours during likely low-load, high VER output days could help reduce curtailment and potential system constraints.
- Finding ways to allow interties to contribute to some part of the FRAC requirement could help to pull more flexibility from imports.
- Uprating thermal generators would allow them to provide greater range to contribute to ramping needs and lower minimum generation levels to accommodate more VER output. Higher confidence in start-up reliability would enable more frequent shut downs which allows for additional VER output to be integrated.
- Incentivizing the remaining non-dispatchable CHP fleet for flexibility could unlock additional generation that would otherwise be baseloaded.
- Requiring new VERs to be economically dispatched would help with system balancing.
- Aligning customer behavior and load attributes with grid needs would help significantly. Exposing some loads to real-time prices would help in the near term. In the longer term, movement away from conventional peak/off-peak prices will help, as will new, controllable loads such as EVs.

5 REPLACING SERVICES FROM THERMAL GENERATORS

In order to accommodate high VER output during low load hours and minimize curtailment, thermal generators will need to be decommitted while retaining a reliable, secure system. Decommitting thermal generators is especially useful because thermal generators have a minimum generation level, which adds to the non-dispatchable stack from Figure 3. If thermal generators are providing energy and ancillary services, then both of these need to be replaced. Ancillary services support the reliable operation of the electric power system. They include services such as regulating reserves, spinning reserves, non-spinning reserves, primary frequency response, and reactive power support. High VER output may be able to replace the energy that thermal generators were providing during those hours. Key options to replacing these ancillary services from thermal generators include:

- Grid-friendly VERs
- Reserves from storage
- Reserves from demand response

5.1 Grid-friendly VERs

In order to integrate higher penetrations of VERs, VERs need to be part of the solution. VERs can provide some essential reliability services, in some cases, better than conventional generators can. And by providing these services to enable decommitment of thermal generation, VER curtailment can be reduced. VERs may cost more with these features enabled, so VERs would need to be compensated appropriately. While all of these services are not needed from every VER on the system at all times, it would be good planning today to consider having new VER installations provide these services. It is cheaper and easier to install grid-friendly VERs rather than to retrofit and potentially renegotiate contracts.

It is important to distinguish here between VERs that are connected to the grid through an inverter, such as PV and wind, and VERs that are connected to the grid through a conventional turbine such as concentrating solar power (CSP). CSP plants use solar energy to heat steam that runs a steam turbine. With respect to ancillary services, CSP plants are similar to conventional generators. On the other hand, inverter-based generators, such as PV and wind, use fast power electronics to provide these responses, and their capabilities are explained further in this section. Because the utility industry has more experience with wind, provision of ancillary services from wind is more common than from PV. However, the mechanisms to implement most of these services using PV should be relatively similar to wind.

It would be useful to review CAISO's production simulations of future scenarios and evaluate the need for additional reliability services based on which plants are decommitted and what services they were providing. Production simulation analyses of high VER penetrations across the West show a large number of hours when VERs are serving 50% of the load[25]. The text box

[25] D. Lew et al, "Western Wind and Solar Integration Study Phase 2", NREL/TP-5500-55588, Sep. 2013. http://www.nrel.gov/docs/fy13osti/55588.pdf

on page 2-14 discusses times in which 50% of the load in PSCO is served by wind. When instantaneous penetration levels become this high, grid reliability and stability need to be re-assessed and the need for grid-friendly PV and wind grows. PV and/or wind can provide inertia, down reserves, up reserves, over-frequency response, under-frequency response and voltage regulation.

5.1.1 Inertia

Spinning rotors of wind turbines have inertia, and wind turbines can provide an inertia-based response that is a very fast response to a drop in system frequency, such as that due to loss of a generator. This controlled response is of the type required for emerging "fast frequency response" ancillary services (e.g. in ERCOT, Ireland). From the perspective of system reliability (particularly to avoid under-frequency load shedding), this response can, in some cases, provide twice the impact of inertia from a conventional generator[26]. Inertial response is not required in the US, but Quebec, Ireland and Ontario require synthetic inertia on their new wind installations. Wind turbines do not need to be curtailed to provide this response but there is a recovery period following deployment of the response. Today, it is worth analyzing whether this response should be required by new wind installations.

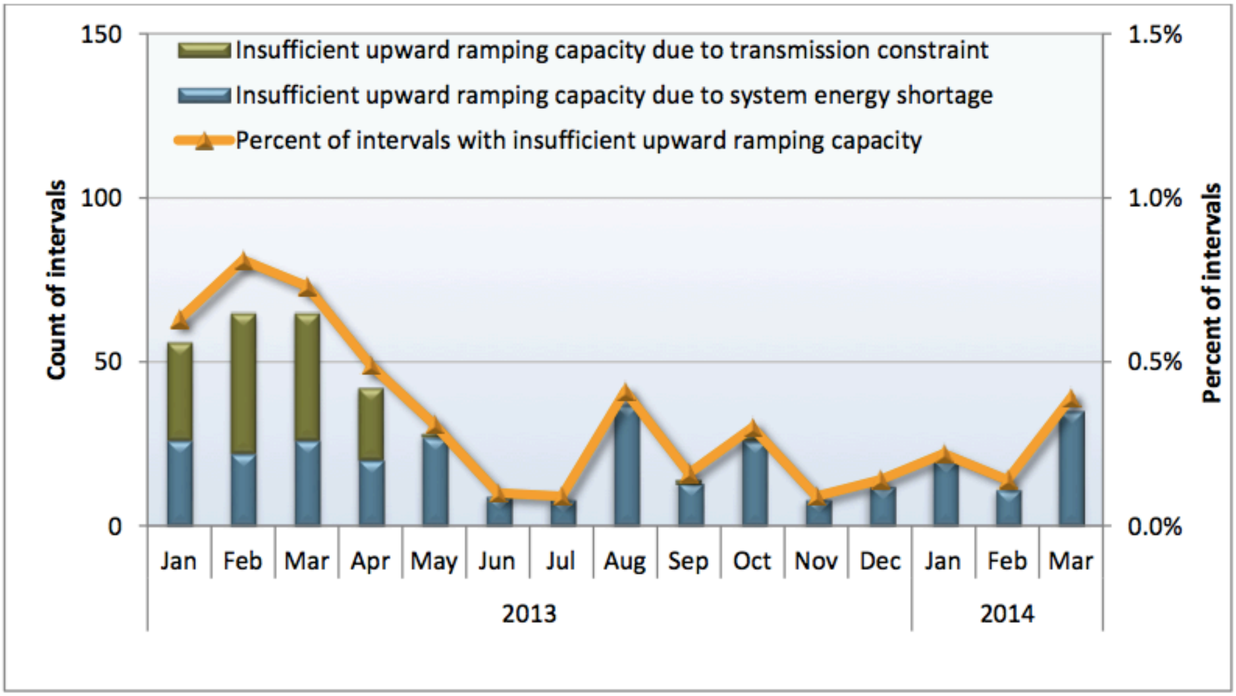
PV has no inertia and as such cannot provide an inertia-based response, but it can provide other fast responses as discussed below.

5.1.2 Down reserves

Thermal units that provide down reserves must operate at their minimum generation level *plus* the level of down reserve that they are supplying. For example, a 100 MW thermal unit may be providing 60 MW of down regulating reserve. In order to do this, it needs to operate at its minimum generation level (say 40 MW) plus 60 MW, or 100 MW. It would be providing 60 MW of down regulation reserve and 40 MW of energy. If there is excess supply midday, VERs may be replacing the 40 MW of energy from the thermal unit. Additional VERs can provide the 60 MW of down regulation reserve. Depending on how many hours this 100 MW thermal unit can be displaced, and when it is needed again, it may potentially be decommitted. *Every MW of down reserve from VERs can potentially displace more than a MW of the stack from* Figure 3.

Both PV and wind can provide down reserves. They may have limits to how much downward capability is available at any time. For example, a 20 MW PV plant producing 10 MW during a given hour cannot provide 20 MW of down response. However, if the PV plant is only producing 10 MW then the other generation that is serving the load should be able to provide the down reserves.

[26] N. Miller, et al, "Impact of Frequency Responsive Wind Plant Controls on Grid Performance," DOE/NREL Active Power Control Workshop, Boulder, CO, Jan. 27 2011.



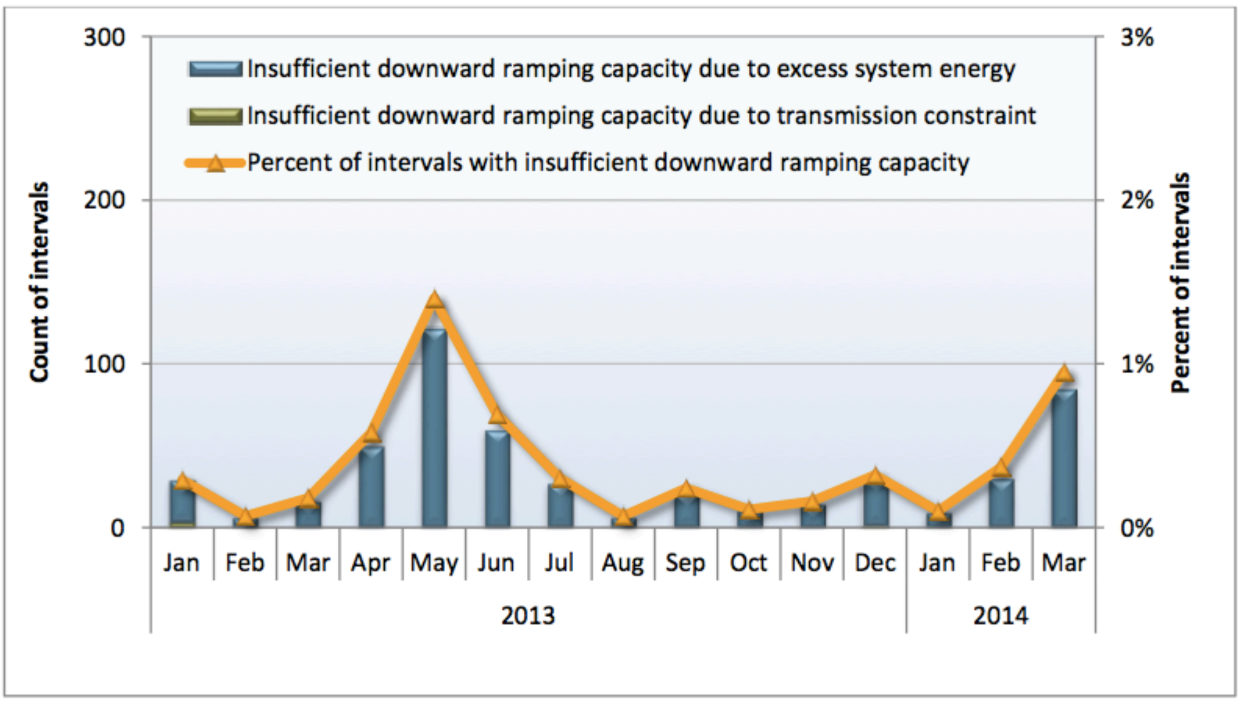


Figure 13 – Insufficient upward (top) and downward (bottom) ramping capacity in 2013-2014. Note the difference in the y-axis scales.
Source: CAISO Department of Market Monitoring, "Q1 2014 Report on Market Issues and Performance," May 22, 2014.

VERs providing regulating reserves would bid into the ancillary service market and be compensated for this service. PV and wind, which are controlled by power electronics rather than thermal processes, can provide down reserves more quickly and effectively than thermal generators. Wind is permitted to provide regulating reserves but must be telemetered and tested in order to do so. No wind plant has yet been tested in CAISO[27].

VERs that are not providing regulation but are economically dispatched, will provide downward ramping capability as determined by the market. CAISO tends to lack downward capability more frequently than upward capability (see Figure 13). This insufficient downward capability is nearly always due to system-wide excess supply conditions, such as high wind output during low demand hours, rather than transmission constraints, as can be seen by the comparison of blue to green bars in bottom of Figure 13.

CAISO regulation requirements are up to 600 MW of upward and 600 MW of downward response. If thermal resources are providing regulation down and their minimum generation is 40%, then replacing 600 MW of downward response with VERs during midday over-supply conditions could displace up to 1,000 MW of thermal generation.

5.1.3 Provide up reserves

Both PV and wind can provide up-reserves as appropriate, such as when curtailed due to over-supply. As shown in the text box on page 2-14, Xcel/PSCO has wind plants on AGC, providing both up and down regulation when they are being curtailed. If significant solar output is forecasted such that negative prices are probable, then solar may want to consider pre-curtailing and bidding into the ancillary service markets to provide up reserves. This is not a trivial proposition in that ancillary services in the CAISO market need to sustain output for 30 minutes. This is likely not a cost-effective option today but at some point in the future with high VER penetrations, it could become cost-effective. By providing up-reserves, similarly, VERs can potentially displace *more* than simply the up-reserve quantities, because a thermal unit has some minimum generation level that can be eliminated if that unit is decommitted.

5.1.4 Over-frequency response

In response to over-frequency events, both PV and wind can provide primary frequency response and reduce output. In fact, PV and wind can provide a sharper response than the governor response from thermal units, and this response could be tuned to address a particular system's reliability needs. Requiring this response of new PV and wind today would be helpful for overall grid reliability.

5.1.5 Under-frequency response

As appropriate, such as when they are curtailed, both PV and wind can provide primary frequency response to under-frequency events and increase output. Again, this response can be sharper than the response from thermal units and can be tuned appropriately. ERCOT currently requires wind to provide this response when wind plants are curtailed.

[27] K. Porter, et al, "Variable Generation and Electricity Markets," Utility Variable Generation Integration Group, April 2015.

The CAISO Frequency Response Study simulated the use of primary frequency response from wind plants in California and found that by adding 1,812 MW of wind plant headroom, the response was better than the case with 12,000 MW more headroom from conventional generators[28]. Therefore under some low load, high wind conditions, it showed that primary frequency response from wind could be over six times as effective as that from conventional generators.

There is less experience with PV providing under-frequency response but similar to wind, this can also be done by pre-curtailing the PV, and allowing PV to increase output in response to a decrease in frequency. The Western Wind and Solar Integration Study Phase 3 simulated the use of primary frequency response from PV (the response was tuned in this particular case to be aggressive) and found that this very fast response could provide significant benefit to the system[29].

5.1.6 Voltage regulation

Both PV and wind can regulate voltage by providing or consuming reactive power. PV and wind can provide good reactive power support to the system but they do have hard limits to this provision, as opposed to conventional generators that will typically have high overload capabilities. It is important to note that the power electronics for PV and wind systems can provide reactive power even when the sun is not shining or the wind is not blowing, by drawing real power from the grid and converting it to reactive power.

5.2 Reserves from Storage

California has a mandate to procure 1,325 MW of storage by 2020. Storage can be used to provide various services – peak shifting, ancillary services, etc. These services, in turn, will determine the type of storage technology used, e.g. pumped hydro storage or compressed air energy storage may be good at peak shifting, while flywheels may be good at providing fast reserves. While electrical storage is currently relatively expensive, thermal storage, such as CSP thermal storage, can provide significant benefits with low energy losses[30].

Storage could be used to store excess solar energy midday and discharge during peak evening hours. For example, a 20 MW battery could store 20 MW midday and discharge 20 MW during peak evening hours (for simplicity, this ignores losses). This would be 1:1 leverage of the storage.

Storage can provide up and down reserves that can potentially leverage displacement of more MW than the storage capacity. The same battery could be used midday to help decommit an 80 MW thermal unit which is operating at 60 MW and providing 20 MW of up regulation and 20

[28] N. Miller, et al, "CAISO Frequency Response Study," Nov. 9, 2011. http://www.caiso.com/Documents/Report-FrequencyResponseStudy.pdf

[29] N. Miller, et al, "Western Wind and Solar Integration Study Phase 3 – Frequency Response and Transient Stability," NREL/SR-5D00-62906, December 2014 http://www.nrel.gov/docs/fy15osti/62906.pdf

[30] P. Denholm et al, "An Analysis of Concentrating Solar Power with Thermal Energy Storage in a California 33% Renewable Scenario," NREL/TP-6A20-58186, March 2013. http://www.nrel.gov/docs/fy13osti/58186.pdf

MW of down regulation and 40 MW of energy. In this case, the same 20 MW of battery storage displaces the regulating reserve and the excess PV midday displaces the 60 MW of energy. *This allows the battery (combined with excess PV midday) to provide 3:1 leverage.* In this example, the PV capacity providing the excess supply is ignored because it is assumed that this excess supply would otherwise be curtailed.

Similar to the order-of-magnitude estimate on down reserves from VERs and ignoring losses, half of California's storage mandate, if targeted towards displacing minimum generation levels of thermal generators and if used in conjunction with excess PV midday, could roughly displace 600 MW of upward and downward regulation, or thermal generation operating at 1,400 MW with a maximum output of 2,000 MW and a minimum output of 800 MW. The other half of the mandate could be targeted towards reducing load following or spinning reserves or other ancillary services. Even if that were leveraged 1:1 displacing thermal generation, the 1,325 MW mandate would displace a total of about 2 GW of generation in the stack.

Detailed studies would need to be undertaken to assess economics, losses, operations, etc., but these are meant to be illustrative examples of how to use storage more effectively in addressing minimum generation levels. Today, careful consideration should be given to how new storage facilities should be designed and operated to help the grid. Using new storage facilities in a way that leverages their ability to more effectively help decommit thermal generation than simple day/evening shifting should be examined and clearly communicated to the market.

5.3 Reserves from Demand Response

Load resources can provide fast up-reserves which can displace thermal generators that are currently providing those services. For example, ERCOT procures a responsive reserve service (similar to CAISO's spinning reserve) to respond to contingency events such as loss of a large generator. ERCOT allows demand response to provide half of their 2,800 MW responsive reserve requirement, or 1,400 MW, and load resources provide close to that limit[31]. They have 3,056 MW of registered load resource capacity, most of which is automatically activated when frequency drops too low. Most of their capacity comes from large industrial electro-chemical process loads, but they also have small- and medium-sized industrial and commercial facilities that range from oil field equipment, cement plants, and manufacturing to compression, pumping and data centers. *Demand response is well-suited to provide contingency reserves*: load can react more quickly than a conventional generator which depends on thermal processes to increase output; contingency events are relatively rare so it may be expensive for a thermal generator to run continuously to provide this reserve if it is seldom called upon; these loads tend to be running continuously and therefore available to be called upon.

Use of demand response to provide spinning reserves is an option that can be implemented in the near term. CAISO's new operating reserve requirement is the greater of either the single most severe contingency or 3% of the sum of load, internal generation, and net pseudo and dynamic imports. Half of this is met by spinning reserve and half by non-spinning reserve. During the

[31] ERCOT, "Annual Report of Demand Response in the ERCOT Region," Version 1.0, March 2015.

fourth quarter of 2014, when this new requirement was implemented, CAISO held about 600 MW – 900 MW of spinning reserves, depending on the time of the day[32]. During midday hours, this requirement was about 800 MW. For those middays when excess supply occurs, demand response could be used to displace thermal generators that would otherwise be providing this service. If thermal generators are providing spinning reserves and their minimum generation is 40%, then replacing 800 MW of spinning reserves with demand response during midday over-supply conditions could displace about 530 MW of thermal generation.

ERCOT is in the process of changing their ancillary services, and they are examining the use of demand response for fast frequency response, which responds faster than primary frequency response. Simulations of fast frequency response in ERCOT under some high wind, low load conditions show that one MW of fast frequency response can provide the same impact as 2.35 MW of primary frequency response in terms of helping the system recover from a contingency event[33]. In the longer term, examining options such as demand response in reliability studies could help CAISO adjust their ancillary service requirements to take advantage of new, fast capabilities as well as meet the needs of an evolving generation portfolio.

5.4 Key recommendations

To decommit thermal generators without sacrificing system reliability, their energy and ancillary services must be replaced. Grid-friendly VERs, storage, and demand response are all options for providing these ancillary services. In some cases they can provide responses more effectively or faster than conventional generators. Key recommendations include:

- Consider requiring new VER installations to provide some or all of the following grid-friendly services:
    - Inertia
    - Down reserves
    - Up reserves
    - Over-frequency response
    - Under-frequency response
    - Voltage regulation
- Instead of using storage for price arbitrage, leverage storage during VER over-supply hours to decommit thermal generators, and reduce pressure on the duck curve.
- Examine use of demand response to meet spinning reserves especially during VER over-supply hours to decommit thermal generators.
- In the longer term, investigate how CAISO can take advantage of some of the faster responses that load, storage and VERs can provide to adjust and possibly reduce ancillary service requirements while maintaining reliability.

32 CAISO Department of Market Monitoring, “CAISO Q4 2014 Report on Market Issues and Performance,” March 3, 2015.

33 J. Matevosyan, ERCOT, “Future Ancillary Service Developments in ERCOT”, NERC Essential Reliability Services Task Force, August 7, 2014.

6 CONCLUSION

Integrating high levels of VERs in California is possible but may require changes to institutional practices, operating practices and existing physical infrastructure. In the DA, it will help CAISO optimally position the system, and avoid over-commitment, if as much VERs as possible can be scheduled in the DAM and/or the RUC can decommit, based on VER forecasts, as appropriate. Investigating conditions under which VERs stress the system, and assessing how that can feed into reserve requirements may help to better position the system. And improving VER forecast accuracy over all time-scales will be helpful.

California's duck curve problem is a combination of increasing VER levels that reduce the net load and a large stack of non-dispatchable generation that serves this net load. In the RT, it is important to have as much of the system as possible be flexible and responsive to real-time prices, so this paper examined the non-dispatchable generation, VERs, and load for flexibility.

The interties appear to be the largest slice of this stack that may be able to provide flexibility but they are also the most difficult, institutionally, to address. Reducing levels of self-scheduling can help with system balancing by making generation responsive to RT prices. Expanding the EIM will help by providing more markets for excess energy in RT, but a much bigger impact will occur if other BAAs will join the CAISO's full DA and RT market. Scheduling exports during the middle of days when loads are likely to be low and VERs are likely to be high, may help relieve duck curve pressures. And finding ways to allow interties to contribute to the new FRAC requirement could help imports provide more flexibility.

Uprating thermal generators can provide greater range to contribute to ramping needs, and ensuring that starts are repeatable and reliable will enable decommitment of thermal units as appropriate. About a quarter of the CHP fleet may have flexibility that could be unlocked with further incentives. Requiring new VERs to be economically dispatched would help with system balancing. In the future, significant flexibility is likely to be extracted from load, especially in new, potentially deferrable loads such as EVs. Moving away from conventional peak/off-peak pricing to RT prices will facilitate the ability of load to integrate VERs. An option that could be undertaken today would be to expose some loads to RT prices and allow those loads to find creative ways to take advantage of low or negative prices.

In a high VER future, thermal generators will need to be decommitted at times to accommodate high VER output levels. To decommit thermal generators, not only does their energy need to be replaced but also their ancillary services. These ancillary services can be provided from other sources such as grid-friendly VERs, storage, and demand response. Today, California should consider requiring new VERs to have the capability of providing certain ancillary services. Retrofits can be difficult and costly to implement and these systems have lifetimes of 20 years or more. So while a response such as an under-frequency response may be unlikely to be deployed today, it may be a response that the system will need in five or ten years when levels of VER are much higher. Analysis of the costs and benefits of providing grid reliability services and how the future grid may evolve, should all be part of this consideration. Both storage and demand response can provide reserves such as regulation or spinning reserves and help to decommit thermal generation.

Table 2 is a high-level summary of the options evaluated to relieve duck curve pressures and integrate higher levels of VERs. The levels of flexibility that may be unlocked are not additive. For example, VER down reserves and reserves from storage can both mitigate the same slice of the stack, i.e., decommitting thermal generators whose main purpose is to provide those ancillary services.

	How much can it help?	How much does it cost?	How hard is it?	How long will it take?
CHP	2 GW	$	✓	✓
Interties	3.4 GW	$	✓✓✓	✓✓
Combined cycle uprates	2 – 2.5 GW	$$	✓✓	✓
VER Down reserves	1 GW	$	✓✓	✓✓
Storage	2 GW	$$$	✓	✓✓
EVs	significant	*	*	✓✓✓
Demand response	0.5 GW	$	✓	✓

Table 2 – Summary of mitigation options and their ability to help, cost, difficulty and timeframe (✓= a little and ✓✓✓ = a lot).
Note that these are rough estimates and that further assessment of each option is needed.
***While the absolute cost and difficulty of deploying a large number of EVs is significant, the *marginal* cost and difficulty of making EVs (that are deployed for other reasons) grid-friendly is relatively small.**

This paper has taken a high-level view of VER integration challenges and options in California, based on CAISO market information and a number of studies previously conducted on the California system. There are a number of next steps that could be undertaken to further assess the impacts and interaction between these different options:

- Analyze CAISO operational data in more detail to understand what leads to over-generation conditions, under what conditions the system is most vulnerable, and how curtailment may be reduced.
- Examine the costs of mitigation options and the interplay between them because their ability to help unlock flexibility is not simply additive. Production simulations could quantify the ability of any or all of these options to increase flexibility and reduce curtailment.
- Gaining a better understanding of which thermal generators tend to be online and why, during potential over-supply hours, in order to identify the types of ancillary services that may be useful to source from other parts of the system.

- Investigate system reliability for future scenarios and analyze how the fast response capabilities of VERs, demand response and storage can contribute to system reliability and potential evolve ancillary service requirements. Assess costs and benefits for VERs to provide various ancillary services to help determine if these capabilities should be required/incentivized.