Essential Reliability Services

Whitepaper on Sufficiency Guidelines

December 2016

RELIABILITY | ACCOUNTABILITY

Table of Contents

Preface

The North American Electric Reliability Corporation (NERC) is a not-for-profit international regulatory authority whose mission is to assure the reliability of the bulk power system (BPS) in North America. NERC develops and enforces Reliability Standards; annually assesses seasonal and long-term reliability; monitors the BPS through system awareness; and educates, trains, and certifies industry personnel. NERC's area of responsibility spans the continental United States, Canada, and the northern portion of Baja California, Mexico. NERC is the electric reliability organization (ERO) for North America, subject to oversight by the Federal Energy Regulatory Commission (FERC) and governmental authorities in Canada. NERC's jurisdiction includes users, owners, and operators of the BPS, which serves more than 334 million people.

The North American BPS is divided into eight Regional Entity (RE) boundaries as shown in the map and corresponding table below.
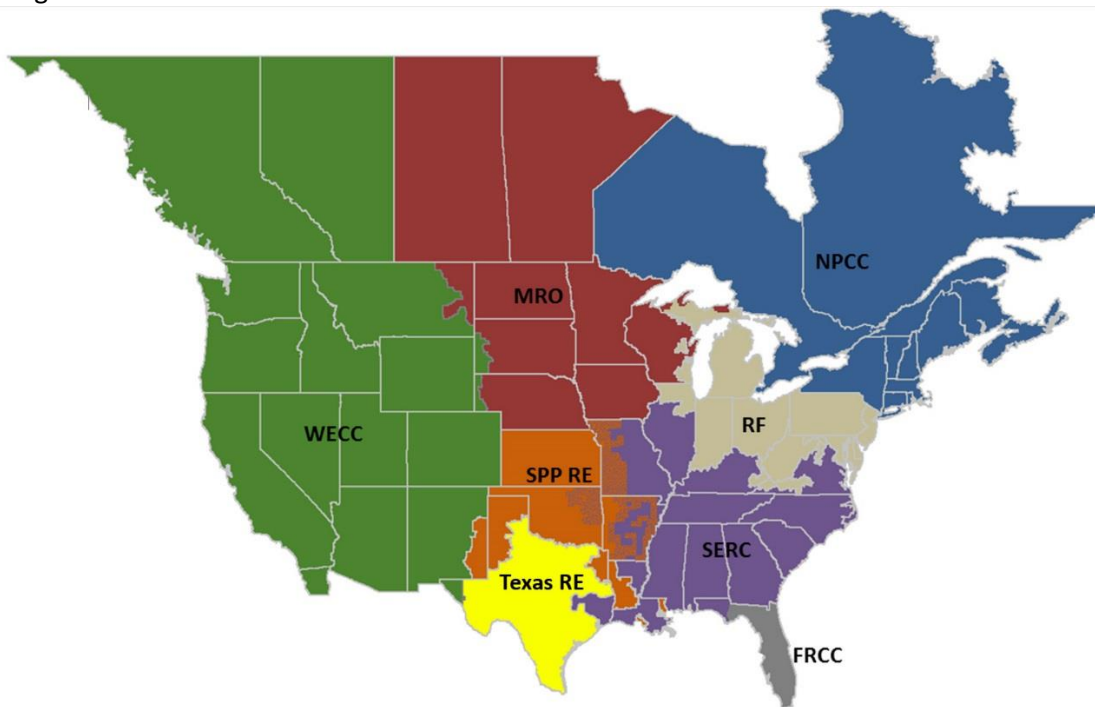
The North American BPS is divided into eight Regional Entity (RE) boundaries. The highlighted areas denote overlap as some load-serving entities participate in one Region while associated transmission owners/operators participate in another.

FRCC	Florida Reliability Coordinating Council
MRO	Midwest Reliability Organization
NPCC	Northeast Power Coordinating Council
RF	ReliabilityFirst
SERC	SERC Reliability Corporation
SPP RE	Southwest Power Pool Regional Entity
Texas RE	Texas Reliability Entity
WECC	Western Electricity Coordinating Council

Executive Summary

The North American BPS is undergoing significant changes in the resource mix and subsequent transmission expansion. Driven by a combination of factors, the rate of this transformation in certain areas is impacting planning and operations of the BPS. For example, environmental regulations are contributing to the acceleration of a significant amount of conventional coal-fired generation retirements while renewable portfolio standards and other factors are driving the development of variable energy resources (VERs). This has resulted in new generation being primarily natural-gas-fired as well as an increase in the penetration of wind and solar resources. Additionally, demand response and many types of distributed generation are increasing, leading to further BPS changes. These changes in the generation resource mix and technologies are altering the operational characteristics of the grid and will challenge system planners and operators to maintain reliability, thereby raising issues that need to be further examined.

The NERC Planning and Operating committees jointly created the ERSTF in 2014 to consider reliability issues that may result from the changing resource mix. The ERSTF proposed ERS measures in 2015 to examine and potentially monitor trends. The ERSTF was converted into the ERS Working Group (ERSWG) in 2016 and charged with identifying, evaluating, and developing Sufficiency Guidelines for each quantifiable measure.

The ERSWG developed this whitepaper that is focused on the methods used to create these Sufficiency Guidelines. The ERSWG will also produce a final report with evaluation and refinement of the Sufficiency Guidelines. The Sufficiency Guidelines are processes rather than specific values as the need for frequency response, generation ramping, and voltage tend to be specific to the characteristics of particular areas, balancing areas, or interconnections. The ERSWG recognizes that there are many possible ways to satisfy the needed system requirements identified as ERSs; however, it will take time and effort to implement changes to mitigate concerns, adjust operating practices, and obtain new sources of ERSs. By allowing trends to be monitored and concerns to be identified well in advance, the Sufficiency Guidelines should incent the appropriate planning and mitigation activities to address potential reliability issues on a timely basis.

This whitepaper further explores important technical considerations so the industry can understand, evaluate, and prepare for the increased deployment of VERs, the retirement of conventional units, advances in demand response and distributed technologies, and other changes to the traditional characteristics of generation and load resources. While the behaviors of conventional generators are well documented, newer technologies (e.g., wind, solar, battery storage, and other types of generators) offer capabilities that can contribute to reliability but often have different characteristics that are not yet as well documented as those of conventional generators. Therefore, the resource mix transition can and does have a profound impact on planning and operating both the BPS and distribution systems, and these impacts, while manageable, have to be accounted for in energy policymaking, regulation, system planning, and system operations.

Frequency Support

Frequency support is provided through the combined interactions of synchronous inertia and frequency response. Working in a coordinated way, these characteristics and services arrest the decline in frequency and eventually return the frequency to the desired level. New approaches that ensure frequency support with lower levels of synchronous inertia may be possible and are being studied. These approaches require proactive planning that adjusts the use of frequency controls and fast power electronics to acquire frequency support from nonsynchronous resources. This guideline proposes a four part process for evaluating inertia given the frequency control practices as they exist today:

    A. For the existing resource mix, determine a theoretical minimum level of synchronous inertial response (SIR) that is needed to provide sufficient time for existing resources that provide frequency response to

react and avoid triggering under frequency load shedding (UFLS). This is referred to as a "top down" approach.

B. Based on the theoretical value derived in Part A, further examine the actual dispatch of resources that would be on-line during such events. This will likely result in a more accurate value since it is probable that the calculated value is overly conservative. This is referred to as a "bottom up" approach.

C. Identify future SIR trends based on changes to the resource mix that are anticipated for coming years.

D. Evaluate alternatives for mitigation of low-inertia situations once the system starts trending towards minimum SIR level found in Parts A and B.

It is important to start planning for the projected future resource mix rather than wait for synchronous inertia to reach the minimum value. For systems that are nearing the minimum inertia value, future inertial conditions should be forecast so that frequency control practices can be modified for the coming years. This report provides an overview of mitigation approaches for both the operating and planning time frames. Approaches for operating with lower levels of inertia will be discussed in future NERC reports.

Ramping and Balancing

For Balancing Authorities (BAs) with an increasing penetration of nondispatchable resources (defined to be any system resources that do not have active power management capability or do not respond to dispatch signals), consideration of system ramping capability is an important component of planning and operations. For BAs expecting changes in their load patterns or generation mix that could change ramping needs over time, the following process is proposed:

- As a prescreening evaluation, analyze system conditions, generation capability, and nondispatchable resources at various operating conditions. This evaluation should emphasize low-load conditions and hours with large ramping requirements to identify operating conditions that may lead to BA ramping capability shortfalls. A key focus will be to identify hours where dispatchable resources are a small percentage of total resources.

- If the prescreen indicates a concern, conduct a more detailed evaluation of historical ramps against real-time control performance standard (CPS1) hourly performance scores to determine whether there is a potential ramping concern. This information should also be communicated to the NERC Reliability Assessment Subcommittee (RAS) and Resources Subcommittee (RS).

Voltage Support

While voltage and reactive performance over a wide range of system conditions must be considered, there is significant justification for focusing the analysis on much smaller sub-areas of the BPS. Due to the inability to efficiently or effectively transport reactive power, the focus should be on defining sub-areas that have a unique set of voltage/reactive performance issues that can only be rectified by local resources and remedial actions. This guideline recommends:

- The Planning Coordinator, Transmission Planner, Reliability Coordinator, and Transmission Owner should collaboratively develop a criterion for defining "reactive power sufficiency" sub-areas within their study areas based on their knowledge of the unique characteristics of their systems.

- During this evaluation, these entities should consider the *NERC Reliability Guideline: Reactive Power Planning and Operations* for assistance in determining relevant sub-areas, appropriate voltage/reactive analysis techniques, and potential technical solutions.

Introduction

The NERC Planning Committee and Operating Committee jointly created the ERSTF in 2014 to consider the issues that may result from the changing generation resource mix, and the committees and the ERSTF released a concept paper in October 2014. The committees agreed that it was prudent to identify the ERSs, monitor the availability of these services, and develop measures to ensure the industry has sufficient awareness of the change in reliability services in the future. The concept paper noted that the key characteristics of a reliable grid could be categorized into two main categories: voltage support and frequency support. The changing generation resource mix raises a number of potential concerns, and the ERSTF was asked to identify measures that should be monitored, both in the operational and planning horizons, to ensure reliable operation of the BPS.

The ERSTF used data from across North America to develop and assess the validity of possible measures that could provide insight into trends and impacts of the changing resource mix. The analysis conducted by the ERSTF focused on measures that could be monitored by NERC to identify potential reliability concerns that may result from the changing resource mix. These measures were intended to provide NERC registered entities and industry with both a short-term operational view and a long-term planning horizon view that enable the identification of immediate reliability concerns and look into the future for needed adjustments. These measures were defined in the [ERSTF Measures Framework Report](#) in 2015 and are included in Appendix A.

The task force found that the most important ERSs for reliability largely focus on the topics of managing frequency, net demand ramping, voltage, and dispatchability. At the highest level, the recommendations can be categorized as follows:

- **Frequency:** These recommendations relate to restoring frequency after an event, such as the sudden loss of a major resource. The frequency within an interconnection will immediately fall upon such an event, requiring a very fast response from some resources to slow the rate of fall, a fast increase in power output (or decrease in power consumption) to stop the fall and stabilize the frequency, then a more prolonged contribution of additional power (or reduced load) to compensate for the lost units and bring system frequency back to the normal level. The task force recommended measures to track the minimum frequency and frequency response following the observed contingency events, track and project the levels of conventional synchronous inertia for each balancing area and the interconnection as a whole, and track and project the initial frequency deviation following the largest contingency event for each interconnection.

- **Ramping and Balancing:** Ramping is related to frequency, but more in an "operations as usual" sense rather than after an event. Changes in the amount of nondispatchable resources, system constraints, load behaviors, and the generation mix can impact the ramp rates needed to keep the system in balance. The task force recommended a measure to track and project the maximum one-hour and three-hour ramps for each balancing area.

- **Voltage:** Voltage must be controlled to protect the system and move power where it is needed. This control tends to be more local in nature, such as at individual transmission substations, in sub-areas of lower voltage transmission nodes, and the distribution system. Ensuring sufficient voltage control and "stiffness" of the system is important both for normal operations and for events impacting normal operations (i.e., disturbances). The task force recommended a measure to track and project the static and dynamic reactive power reserve capabilities to regulate voltage at various points in the system. The task force also recommended that industry monitor events related to voltage performance, periodically review the short circuit current at each transmission bus in the network, and do further analysis of short circuit ratios when penetration of nonsynchronous generation is high or anticipated to increase.

The task force also made a number of general recommendations, emphasizing that new resources may have different operating characteristics but can be reliably integrated with proper planning, design, and coordination.

Maintaining reliability is embodied in the predictability, controllability, and responsiveness of the resource mix. This report focuses on further examining the ERS measures and working to develop Sufficiency Guidelines.

This whitepaper is the result of the 2016 efforts of the ERSWG, and work will continue to provide the final report in 2017. As the ERSWG continues to examine the proposed measures, it is finding that some of them provide useful trends and insights while others are require some revision or reconsideration. Additional metrics are also being investigated and monitored as the ERSWG works with appropriate NERC subcommittees and working groups on data collection and analysis of the measures. The ERSWG expects to see ongoing enhancements to the measures and additional recommendations from the other working groups to provide NERC with greater clarity going forward.

Given the nature of ERSs and the significance of such services for energy policymaking, system planning, and system operations, NERC should anticipate the need for ongoing information sharing and support for a wide variety of stakeholders. Federal, state, and local jurisdictional policy decisions have a direct influence on changes in the resource mix and thus can affect the reliability of the BPS. Planning and operations analysis of these emerging changes must be done to ensure continued reliable and economic operation of the BPS.

Chapter 1: Frequency Support

Background

The recommendations of the ERSTF for frequency support relate to restoring frequency after an event, such as the sudden loss of a major resource. The frequency within an interconnection will immediately fall upon such an event, requiring a very fast response from resources to slow the rate of fall, a fast increase in power output (or decrease in power consumption) to stop the fall and stabilize the frequency, and then a more prolonged contribution of additional power (or reduced load) to compensate for the lost units and bring system frequency back to the normal level (see Figure *1.1*). The recommended ERS measures track the minimum frequency and frequency response following the observed contingency events, track and project the levels of conventional synchronous inertia for each balancing area and the interconnection as a whole, and track and project the initial frequency deviation in the first half-second following the largest contingency event for each interconnection.

For the current power grid, frequency support is provided through the combined interactions of synchronous inertia, primary frequency response, and secondary frequency response. Working in a coordinated way, these characteristics and services arrest the decline in frequency and eventually return the frequency to the desired level. This report addresses synchronous inertia (ERS Measures 1, 2, and 3). For ERS Measure 4 (Frequency Response at the Interconnection Level), the NERC Resource Subcommittee reviews frequency events by interconnection against the relevant criteria provided by the NERC Performance Analysis Subcommittee, and analysis of the selected events is included in the annual NERC State of Reliability report.

This report focuses on frequency support using frequency controls as they exist today, but approaches for supporting frequency with lower levels of synchronous inertia are under study in North American, Ireland, New Zealand, and the European Union. Future NERC reports will discuss these methods and the value of adjusting the use of frequency controls and fast-power electronics in nonsynchronous resources. Changing the use of frequency controls and new capabilities will take proactive planning, but it is also important to note that large North American interconnections have ample inertia today and are not approaching minimum inertial levels.

In addition, an example of the Electric Reliability Council of Texas (ERCOT)'s tool for operational monitoring and forecasting of synchronous inertia is included Appendix D. This provides an example of real-time monitoring of inertia that was recommended as an industry practice in Measure 5 of the *ERS Measures Framework Report.*

Understanding Synchronous Inertia

Synchronized rotating generators and motors interconnected to a power system store kinetic energy. Stored kinetic energy of a machine at nominal rotating speed can be determined based on the following relationship:

$$Kinetic\ energy\ = \frac{J \cdot \omega_o^2}{2} \qquad (1)$$

In this equation, *J* is the combined moment of inertia of the synchronous machine and turbine prime mover in kg·m² (based on characteristics of its size and weight), and ω_o is the nominal machine rotor speed in rad/s.

The inertia constant *H* is defined as the kinetic energy divided by the machine's rated capacity in MVA, P_{MVA}:

$$H = \frac{J\omega_o^2}{2P_{MVA}}, [s] \qquad (2)$$

The *H* constant is the time in units of seconds that it would take for the synchronous machine to deliver all of its stored kinetic energy to the power grid, assuming that it is producing at rated power and continues to rotate at rated synchronous speed.

As long as the machine is synchronized to the grid, the *H* constant (and therefore the amount of kinetic energy) is the same regardless of the machine's current power output level. Equation (1) can be rewritten in terms of inertia constant *H* as shown below, making it clear that stored kinetic energy of a synchronous machine is independent of the machine's power output level and can be evaluated based solely on machine parameters (*H* and rated MVA) and machine status (on-line/off-line).

$$Kinetic\ energy\ = \frac{J \cdot \omega_0^2}{2} = H \cdot P_{MVA} \qquad (3)$$

Immediately after a contingency event (e.g., a generator trip), this stored kinetic energy is drawn from all remaining synchronous machines to maintain the power balance between production (that has changed due to the generator trip) and consumption (that still remains the same). This withdrawal of kinetic energy is called the synchronous inertial response (SIR). As stored kinetic energy is drawn from the generators, they slow down and system frequency therefore declines, as shown in Figure *1* The initial rate at which system frequency declines depends on the amount of inertial response (stored kinetic energy) available at the time of the event, (i.e., the number and size of generators and motors synchronized with the system). For any given instance, the synchronous inertial response of the system can be calculated as the sum of individual inertial responses from all on-line synchronous machines. The red line in Figure 1.1 shows the initial rate of change of frequency due to inertial response from synchronous machines.
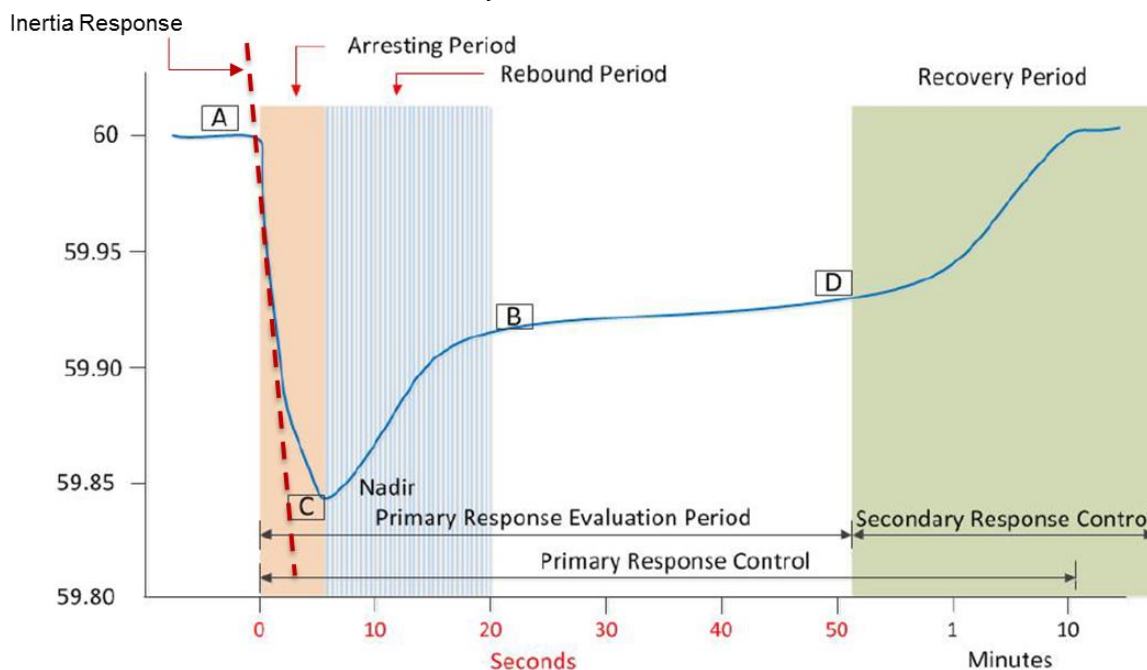
$$H_{sys} = \sum_{i \in I} H_i \cdot MVA_i \qquad (4)$$

Figure 1.1: Typical Frequency Excursion and Recovery

With increasing use of nonsynchronous generation and other electronically coupled resources (both generators and loads), the level of SIR is reduced. Particularly in areas with a high penetration of renewable resources, this leads to a need to consider both the amounts of SIR and the available amounts of primary frequency response based on expected SIR conditions. The critical issue is the coordination between the level of SIR and the level of primary frequency response that is delivered in the arresting period as the delivery of additional power within the arresting period must be sufficient to regain the balance between generation and demand. When this power balance is re-established, additional kinetic energy is no longer drawn from the synchronous machines, so the frequency stops falling (i.e., the "nadir" point in Figure 1.1 is established). Once the nadir is established, primary

and secondary response resources can provide additional power to eventually return the system to the normal frequency.

For frequency support, the important issues are the nadir frequency and frequency recovery (ERS Measure 4), with the level of SIR being just one component of the frequency response relationship. While a reduced level of SIR will result in a faster initial rate of frequency change immediately after a disturbance, the critical reliability issue is to keep the nadir above the UFLS level for the interconnection. For systems where the amount of SIR is decreasing, the nadir frequency can be maintained in various ways (this will be discussed in the section on mitigation options). So while this report recommends monitoring of SIR values to identify when other actions may be prudent, this is not intended to imply that reduced SIR is a current or near-future problem for most interconnections; there are many options for maintaining or improving the nadir if and when it becomes a concern.

Interconnections with growing amounts of nonsynchronous generation or electrically coupled resources should project future SIR trends based on historical SIR information and planned projects in the interconnection queue (e.g., signed interconnection agreements and financial commitments). These projections will help a BA anticipate decreasing interconnection SIR conditions, which will increase the challenges associated with meeting the interconnection frequency response obligations[1] (IFROs), assuming that frequency response contributions from those resources in the arresting period remain unchanged. The ability to anticipate the changes in the SIR will help the BAs develop measures to offset any decline in SIR to meet their IFRO required in BAL-003. Understanding how changes in SIR affect and interact with primary frequency response will be crucial in preserving reliability and required to determine if a minimum SIR requirement is necessary or other actions may be needed during the planning and operating of the bulk power system to sustain reliable operation following a disturbance.

The ERSTF Measures Framework Report (December 2015) previously proposed several measures for trending synchronous inertia. The ERSWG developed a spreadsheet for Measure 1 and 3 as described in Appendix B. Data collection for this measure started in June 2016 by the NERC Resources Subcommittee (RS).

Inertia Sufficiency Guideline

Purpose of the Inertia Sufficiency Guideline
The purpose of this guideline is to examine the system capability under low-inertia conditions to arrest frequency decay and avoid involuntary UFLS after a large generator trip based on each area's existing primary frequency control capabilities and practices. Once system inertia, as based on historical data (Measures 1 and 3), starts approaching a minimum value as described in this document, each area should consider revising the existing frequency control practices and capabilities and introduce additional measures to ensure that system frequency is arrested above the prevailing first stage of involuntary UFLS after the largest contingency.

> **Fast frequency response** is active power injection automatically deployed in the arresting phase of a frequency event aimed at providing full response before the frequency nadir is reached.

At a high level, the recommended approach consists of four parts:

  a. For the existing resource mix, determine a theoretical minimum level of SIR that is needed to provide sufficient time for currently assumed frequency response resources to avoid UFLS (top down).

---

[1] http://www.nerc.com/pa/Stand/Project%20200712%20Frequency%20Response%20DL/BAL-003-1_clean_031213.pdf

b. Based on this value, further examine the actual dispatch of resources that would be on-line during such times since it is likely that the calculated value determined in Part A is overly conservative (bottom up).

c. Identify future SIR trends based on changes to the resource mix that are anticipated for coming years (trends).

d. Evaluate alternatives for mitigation of low-inertia situations once the system starts trending towards minimum SIR level found in Parts A and B.

Part A: Method for Determining Minimum Synchronous Inertia under Existing Frequency Control Practices (Top Down Approach)

The flowchart for the proposed method is shown in Figure 1.2. The procedure is as follows:

1. Calculate system inertia for every hour in a year in MVA*seconds. This is calculated as a sum of individual inertial contributions from all on-line generators. Find the lowest inertia instance in a year. This is covered under Measures 1 and 3 as described in the *2015 ERSTF Framework Report.*

2. At that lowest inertia condition, calculate the rate of change of frequency (RoCoF, in Hertz per second or Hz/s) based on the interconnection's resource contingency criteria (RCC), which is the largest identified simultaneous category C (N-2) event, except for the Eastern Interconnection, which uses the largest event in the last 10 years. This is a part of the Measure 2 calculation and the interconnection RCC values are shown in Table 1.1. There is no standard requirement for a system to operate without shedding firm load after an RCC event, but this is considered to be the best practice for system design.

Table 1.1: Resource Contingency Criteria for each Interconnection [i.e., Largest Category C Event (N-2)[2]]			
ERCOT	**EI**	**WECC**	**HQ**
2750 MW	4500 MW	2740 MW	1700 MW

3. With the RoCoF from the previous step, calculate how long it takes to reach the first stage of UFLS after the RCC event. If this time is sufficient (e.g., 1-1.5 seconds[3]) for the existing means of frequency response (fast frequency response and some portion of the primary frequency response) to start deploying and to help arrest frequency above the first stage of UFLS, then this RoCoF is not critical.

4. Gradually scale down the inertia found in Step 1, such as in steps of 10%. Repeat Steps 2 and 3 until the resulting RoCoF is such that a prevailing UFLS first step is reached before the existing frequency response capabilities can become effective (e.g., for ERCOT this time is 1-1.5 s during low inertia conditions). Return to the last inertia value that still is sufficient. This is the first approximation of the theoretical minimum system inertia.

---

[2] In the future, a higher MW contingency may need to be considered due to significant amounts of distributed energy resources disconnecting at low frequency after an initial generation trip event on the BPS.

[3] The purpose here is to make sure the RoCoF does not result in UFLS within 1-1.5 seconds (i.e., before frequency response (fast and/or primary) can become effective). The time needed for fast and primary frequency response to become effective may vary for different systems and different synchronous inertia conditions. These times can be verified from historic event analysis and dynamic simulations. For example, in ERCOT, load resources with underfrequency relays are providing a portion of responsive reserve service (an ancillary service with the function to arrest, stabilize, and recover frequency after an event). These load resources trip offline within 0.5 s after system frequency falls at or below 59.7 Hz. Therefore, in ERCOT, the RoCoF that leads to 59.7 Hz within a few hundred milliseconds after an event and then reaches the first stage of UFLS within the next 0.5 s becomes important. If UFLS is not reached within this time frame, load resources will trip and are very effective at arresting frequency above the UFLS threshold.

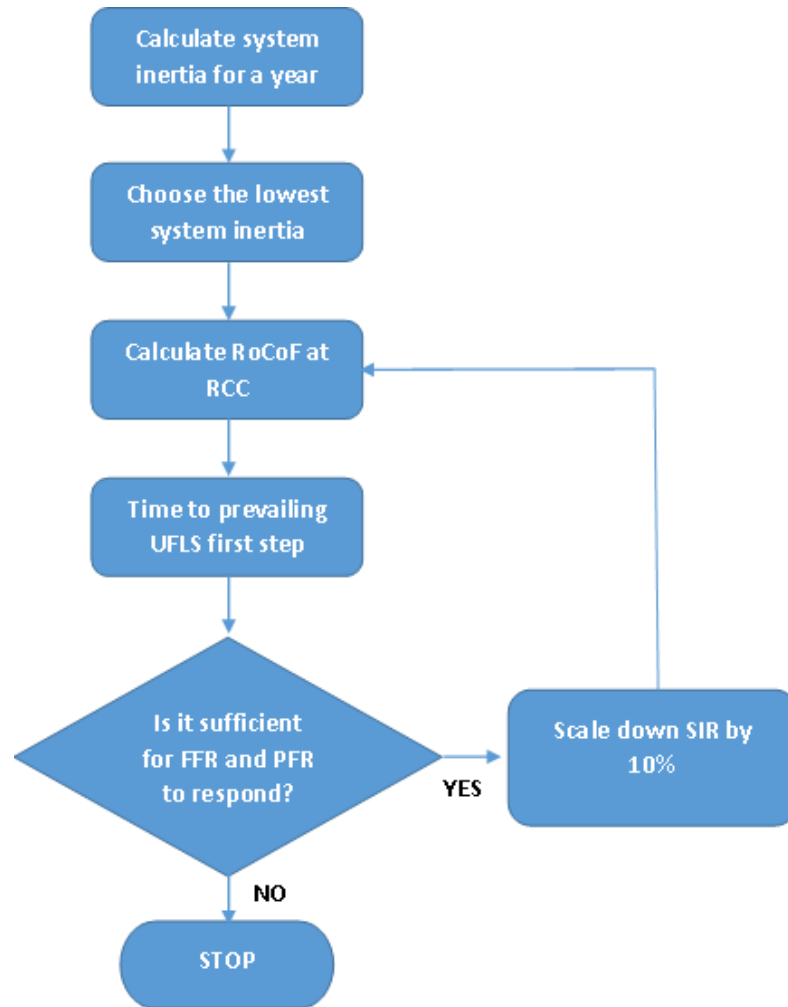The system inertia value determined here is a **theoretical** top down approximation of minimum inertia.

Figure 1.2: Top-Down Approach to determine minimum synchronous inertia

Part B: Verification of Actual Lowest Achievable Synchronous Inertia under Existing Frequency Control Practices (Bottom-Up Approach)

In practice, there is often more synchronous generation on-line at any given time due to other considerations. The flowchart for the proposed method is shown in Figure 1.3. To obtain a more realistic value based on expected dispatch, a bottom-up approach can be used:

1. This approach starts with the minimum synchronous inertia that will always be on-line for a system. For example:

    a. Are there any reliability-must-run units (e.g., for voltage support or transmission reliability) that would need to be on-line in these conditions?

    b. Are there any synchronous condensers or generators in synchronous condenser mode that would need to be on-line in these conditions?

    c. Are there any behind-the-meter industrial generators that are always on-line?

    d. Are there nuclear units? Of those, how many units could simultaneously be on maintenance during low-inertia times?

e. How many synchronous generators are on-line to provide required reserves and what are those generators (what are their typical inertia values)? Here it is important to consider if there are any physical limitations or regulatory restrictions on how much a single resource can or is allowed to contribute towards a reserve requirement. In ERCOT, for example, a generator is not allowed (by protocols) to offer more than 20% of its highest sustainable limit towards responsive reserve service.[4] Calculate the total inertia contribution for all these "must run" units that need to be on-line to contribute to reliability, capacity, or the ancillary service requirements.

2. Compare the inertia calculated in Step 1 and the theoretical minimum inertia in Part A.

3. If the calculated inertia from Step 1 is higher or equal to the theoretical inertia in Part A, then the system will have sufficient synchronous inertia at all times, unless:

4. Any operations principles of the "must run" units change

5. Reserve requirements decrease

6. Contribution from a single resource towards any of the on-line reserve requirement changes

7. New reserves are introduced, or entry of new resources (not providing synchronous inertia) into the ancillary services market becomes possible

8. If the calculated inertia from Step 1 is lower than theoretical inertia in Part A, then starting from the unit commitment and total "must-run" inertia value obtained in Step 1, bring additional synchronous units on-line one by one, based on unit merit order. Stop once the total system inertia value is close to the theoretical minimum value determined in Part A.

9. The results can be verified with dynamic simulations using the adjusted unit commitment and simulating an RCC event. Since the above calculation does not take into account load damping and primary frequency response, it is possible that the true minimum synchronous inertia value is slightly lower than this theoretical estimate. However, note that load damping and governor response do not significantly affect RoCoF in the first few seconds of an event. There can also be additional concerns with low synchronous inertia apart from frequency events, such as for voltage oscillations and stability issues due to insufficient synchronizing torque.

---

[4] This is because with a 5 percent governor droop setting, a generator is not able to provide more than 20% of its capacity in response to system frequency change of 0.6 Hz (i.e. from 60 to 59.4 Hz, with 0.1 Hz margin above first stage of UFLS).
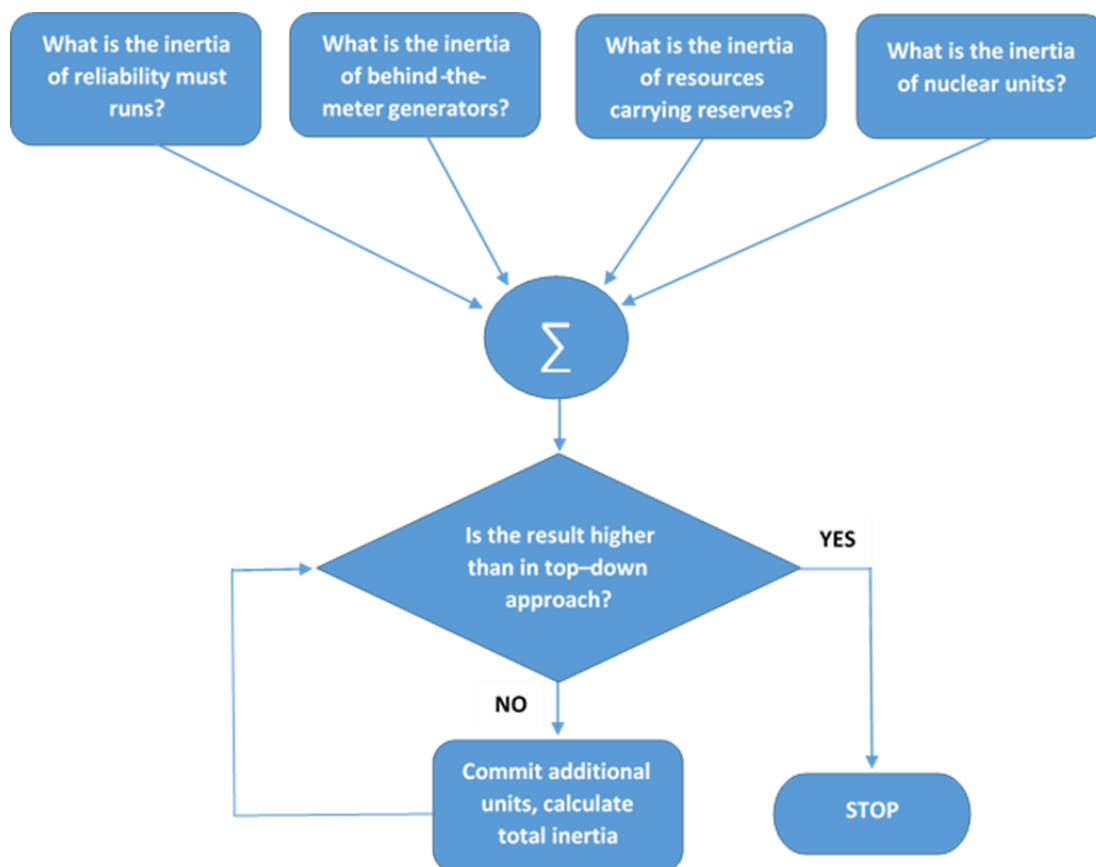
What is the inertia of reliability must runs?

What is the inertia of behind-the-meter generators?

What is the inertia of resources carrying reserves?

What is the inertia of nuclear units?

Σ

Is the result higher than in top–down approach?

YES

NO

Commit additional units, calculate total inertia

STOP

Figure 1.3: Bottom-Up Approach to Verify Minimum Sufficient Synchronous Inertia Determined via Top-Down Approach

Part C: Trending of Synchronous Inertia versus Minimum Inertia Value

Once a minimum inertia value is identified, actual synchronous inertia of the system (Measure 1) has to be monitored against this value. However, it is important to start planning ahead of time rather than waiting for synchronous inertia to reach the minimum value. For systems that are nearing the minimum inertia value, it would be practical to start forecasting future inertial conditions for the coming years.

A method for forecasting future minimum inertia conditions was proposed in ERSTF Measures Framework Report. The method uses historic trends between system inertia and net load, as well as planned nonsynchronous generation, to forecast minimum system inertia in a future year. However, applying this approach to the past years was found to yield results that were overly conservative. The example in Figure 1.4: Inertia Forecasting Methodology Applied to Historic Years in ERCOT demonstrates this overly conservative result by showing the historic inertia in ERCOT (box plots) and predicted minimum inertia (red dots) in the following year using the proposed methodology[5]. At any given time, system inertia depends on factors other than just load and nonsynchronous generation (e.g., fuel prices, environmental policies, commitment practices, etc.). More accurate prediction of future inertia requires detailed modeling of the future generation mix, unit commitment, and dispatch. The ERSWG will be proposing further improvements to the inertia forecasting methodology in 2017.

---

[5] For example, ERCOT used data from 2013 and knowledge about all nonsynchronous generation projects that came in-service in 2014. Based on this information, ERCOT predicted the minimum inertia in 2014 and compared the predicted value to the actual inertia in 2014.
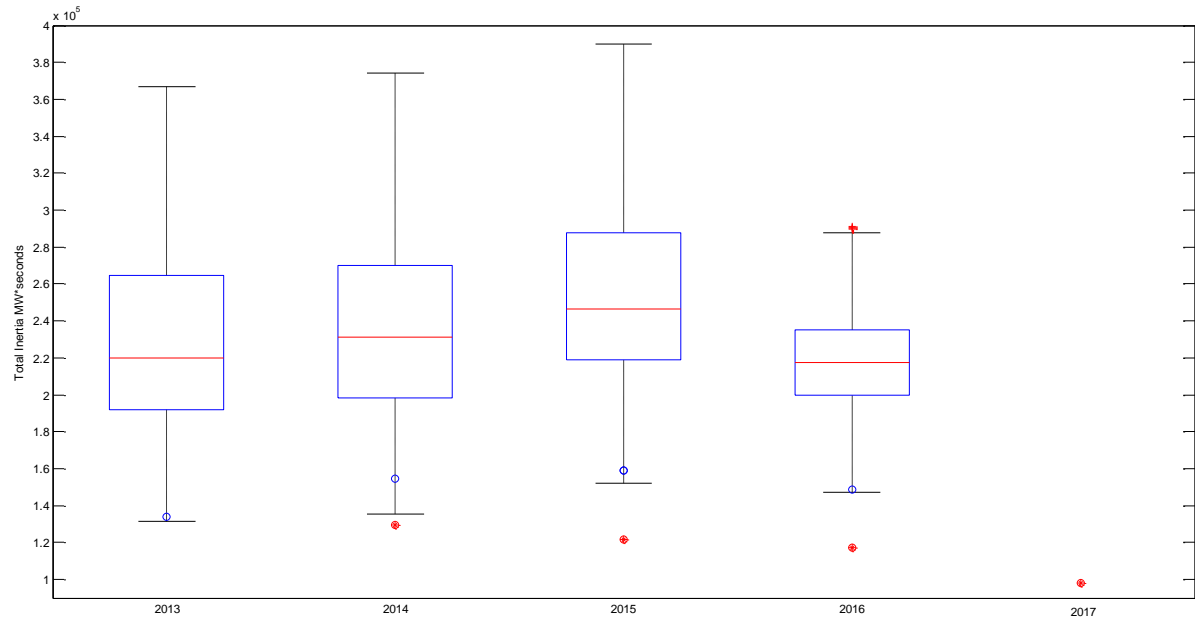
Figure 1.4: Inertia Forecasting Methodology Applied to Historic Years in ERCOT

As the synchronous inertia is approaching the minimum value, frequency control measures need to be revised and additional means of fast frequency support (e.g., from load resources, storage, synthetic inertia from wind turbines, etc.) should be put in place. Note that fast frequency response may also be introduced to address other issues such as resource adequacy, the need for flexibility and to improve energy dispatch efficiency. This may be effective by allowing certain generating resources to be available for energy production while allowing other resources, such as load or storage, to provide frequency reserves. In this case, the synchronous inertia calculation must be revised by repeating Parts A and B and including characteristics of the fast frequency response resources in the analysis.

A case study of ERCOT calculation of minimum inertia is shown in Appendix C.

Part D: Mitigation Alternatives at Extremely Low System Inertia
There are multiple solutions to mitigate the impact of low system inertia:

- Committing additional units for their synchronous inertia, committing different units that have higher inertia, and/or using synchronous condensers

- Slow down the rate of change of frequency by increasing the rate of primary frequency response of the system in MW/s per Hz

- Slow down the rate of change of frequency by increasing the speed of frequency response, such as by adding fast frequency response from load resources, storage, synthetic inertia from wind generation, and so forth

These solutions can be implemented as described below.

Approaches to Bring Additional Synchronous Inertia On-line

- **Start-up more synchronous generation units**: If system inertia is low, system operators may choose to bring more synchronous generation units on-line and thus increase system inertia. One benefit of this approach is that it would only be implemented during conditions when insufficient inertia was otherwise committed on the system and therefore it would not affect solutions during other periods. The generating units, once started, would at least run at their minimum sustainable level, producing energy that is not needed otherwise. Considering the relative inertia provided by the different units that could be brought on-line and their minimum generation levels would then be important. This is only recommended when the system inertia falls below the minimum inertia level and causes a concern for reliability.

- **Include the need for Inertia in the Planning of Reserves**: Inertia can be included as part of reserves in the unit commitment process. The commitment of reserves would take into account the expected level of inertia of the system.

- **Install Synchronous Condensers**: A synchronous condenser is similar to a synchronous generator but without a turbine. Its main purpose is to provide dynamic reactive power support. Being that it is a rotating machine, it can also contribute to total system synchronous inertia. While this option may be expensive if used for inertial support alone, synchronous condensers also provide voltage support. As one of the long-term solutions, use of synchronous condensers for multiple purposes may prove to be more cost efficient. The contribution to inertia from synchronous condensers that are installed for transmission system support should be included in the assessment of how much inertia is available.

Approaches to Increase the Rate of Primary Frequency Response

- **Increase Reserves for Frequency Response:** By considering frequency responsive capacity in the unit commitment process, it is also possible to obtain a higher rate of response (MW/s per Hz) for mitigating the risk of severe under frequency events. This mitigation measure is easy to implement since it can be accommodated by the existing procedures and framework. However, this option would be inefficient to some extent since large amount of reserve capacity will be reserved and not available for energy production.

Approaches to Add Fast Frequency Response

- **Use Fast Frequency Response from Other Technologies:** As was pointed out above, what is critically needed at low-system-inertia conditions is faster response to counteract the higher rate of change of frequency (if the voltage oscillation is not a problem). Fast frequency-responsive load resources (e.g., large industrial loads, heat pumps, industrial refrigerator loads, storage devices, etc.) can provide full response in a few hundred milliseconds of the under frequency events. Recent research studies[6] suggest that frequency response can be provided by nonsynchronous generators that reserve operating margin for providing frequency support. This approach will be further discussed in the future NERC reports.

---

[6] S.A.P. Meliopoulos, M. Saeedifard, V. Vittal, Y. Lu, S. Choi, E. Polymeneas, D. Ramasubramanian, Z. Yu, Stability, Protection and Control of System with High Penetration of Converter Interfaced Generation, Final Project Report, Power Systems Engineering Research Center, March 2016

- **Synthetic Inertial Response from Wind Generation**: Another example of fast frequency response is synthetic inertia provision from wind generation resources (from Type 3 and Type 4 wind turbines). When the wind turbine plant controller senses a drop in system frequency, it can very quickly extract kinetic energy from the rotating mass of the wind turbine. This results in an increase in active power injection during the arresting period. The effectiveness of the response and recovery of the wind generation resource to its predisturbance state depends on operating conditions of that resource. While synthetic inertial response capability is already included as a part of interconnection requirements in Hydro Quebec, this technique has not yet been commercially used in other interconnections.

Chapter 2: Ramping and Balancing

Background

With the increasing penetration of generation resources for which the Balancing Authority (BA) may have limited ability to control the level of output, consideration of system ramping capability is an important component of planning and operations.

For example, the California Independent System Operator (CAISO) has started to experience ramping and over-supply concerns. High penetrations of nondispatchable resources are meeting a large portion of their customer's energy needs during some times of the day, resulting in the need for additional flexibility and ramping. This is not a totally new concern for BAs as some resources and imports have long had output levels that are not easily altered through the dispatch process to meet system ramping needs. Additionally, newer resources may or may not be incorporated into the dispatch process or they may be considered "must take" resources, so these resources can significantly contribute to increasing ramping needs for some BAs. The combination of all such factors can result in increased periods of over generation, ramping scarcity, and other situations that cause an overreliance on the interconnection for balancing.

There are many ways to mitigate ramping and balancing concerns, but mitigation methods may take significant time and effort. The focus of this guideline is to ensure that emerging concerns are identified early so that changes can be accomplished in a timely and reliable way. Therefore, as a best practice, BAs should regularly examine their fleet make-up to determine if changes are needed in their supply procurement and unit commitment practices to adequately balance generation and load to help meet the shared responsibility of supporting interconnection frequency and to meet their frequency response obligation following a contingency.

To address the ramping concern, this guideline has been developed in collaboration with the NERC Essential Reliability Services Working Group (ERSWG), NERC Resources Subcommittee (NERC RS) and NERC Reliability Assessment Subcommittee (NERC RAS) to provide BAs with methods to identify trends and indications of potential ramping concerns within their footprint.

A set of screening methods for BAs expecting changes in their load patterns or generation mix that could change ramping needs over time was created. BAs should also examine their system ramping capabilities under various operating conditions, especially during light load conditions. Screening methods will also be used on a periodic basis by the NERC RAS and RS to aid BAs in monitoring potential issues. In addition, if the evaluation indicates concerns in future years that need to be addressed, a method is proposed to monitor trends and conduct further analysis in the framework report. Examples of how CAISO and ERCOT are doing this analysis can be found in Appendices F and G.

Prescreening Evaluation Method

BAs can use the following prescreening process as a simple high-level review to determine if additional detailed evaluations are prudent to identify potential shortages of ramping capability. The first step in performing this Pre-Screening is to identify those future operating conditions that may lend themselves to the BA's ramping capability shortfalls. While determination of "constrained operating conditions" is somewhat subjective, these operating conditions typically correspond to operating hours where the available dispatchable system resources constitute a smaller percentage of the total system resources. These conditions normally correspond to operating hours where system ramp capability is nearing required ramp or system load is expected to be low and the magnitude of operating nondispatchable system resources, as defined below, is meeting a large portion of the system demand.

To perform the prescreening process, a BA will need to determine the following for each of the forecasted constrained operating conditions:

1. The expected minimum load[7] for the constrained operating condition (Y_{CL}).

2. The amount of nondispatchable resources (Y_{ND}) that are expected to be operational for the same year during the screening hour.

> **Nondispatchable resources** refer to those system resources that do not have active power management capability or do not respond to dispatch or Automatic Generation Control (AGC) signals. These may include nuclear generators, geothermal generators, and older utility scale renewable generators and DERs.

3. The maximum regulation up (R_u) and regulation down (R_d) required during the constrained operating condition.

4. The maximum load following up (LF_u)[8] and load following down (LF_d)[9] required for the next one to three hours during the constrained operating condition.

5. The maximum net load[10] increase (MNL_u) or maximum net load decrease (MNL_d) forecasted in the next one to three hours.

6. The contingency reserve (**CR**) that is needed to cover the BA's Most Severe Single Contingency (MSSC) during the constrained operating condition.

7. The evaluation percentage (X%) should be a value between 30% and 50%, with the value selected by the BA based on their generation mix[11].

The actual screen is summarized as follows:

Level 1 Ramp Penetration Screen:

For current constrained operating conditions, if $Y_{ND} \geq$ **X% of Y_{CL},** then perform the Level 2 Ramp Screen, otherwise go directly to the CPS1 evaluation.

---

[7] Minimum load, not minimum net load, is used to avoid double counting of resources.
[8] LF_U is the intra-hour upward capacity reserved to meet uncertainty and variability associated with VERs and load.
[9] LF_D is the intra-hour downward capacity reserved to meet uncertainty and variability associated with VERs and load.
[10] Net load is calculated by subtracting nondispatchable generation (including net behind-the-meter Distributed Energy Resources) from total system load.
[11] Elements to consider include the storage capabilities, hydro resources, quick start resources, demand side management, etc.

Level 2 Ramp Screen:

For constrained operating conditions where $Y_{ND} \geq X\%$ of Y_{CL}, if Level 1 Ramp Screen indicates concerns, the BA should use the following hourly screening process to determine if additional analysis is warranted:

1. Verify that the downward ramping capacity of the committed dispatchable resources (resources that are on-line with downward ramping capability or can be brought off-line within one to three hours) can cover ($R_d + LF_d + MNL_d$).

2. Verify that the upward ramping capacity of the committed dispatchable resources (resources that are on-line with upward ramping capability or can be brought on-line within one to three hours) can cover (**CR** + $R_u + LF_u + MNL_u$).

A more detailed analysis of ramping shortfall must be performed if both level 1 (operating) and level 2 (planning) ramp screens fail. If a BA believes the level 1 and 2 screenings indicate ramping shortfalls, the BA can contact the NERC RS and RAS for assistance and further monitoring.

CPS1 Evaluation:

In either case above, perform a CPS1 Evaluation. The CPS1 Evaluation described in this document provides an evaluation method that will be used by the NERC RS, but the BA may use their preferred method.
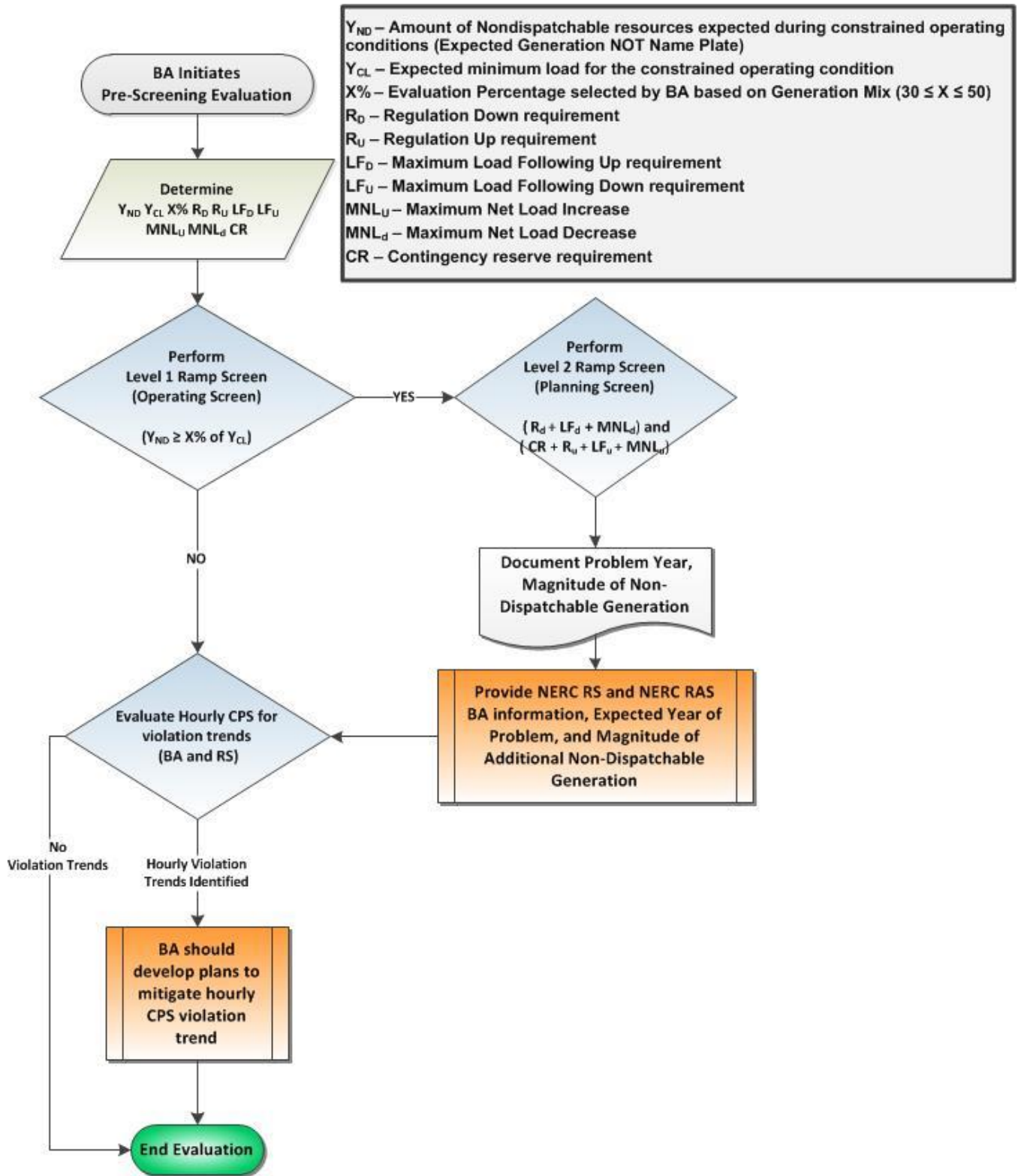
Figure 2.1: Prescreening Evaluation Process Diagram for Ramping

CPS1 Evaluation Method

Historical CPS performance can be an indication of potential ramping issues. A BA can evaluate its historical ramps against its real-time control performance standard (CPS1[12]) to determine whether it is beginning to experience ramping problems. This can be accomplished by evaluating hourly CPS performance data for trends.

To perform the CPS1 Evaluation with the method that will be used by NERC RS, a BA will need to collect and format the following data. The BA may select a different historical period, but the RS method and the examples below assume the use of three years of historical data:

1. Collect hourly CPS1 data for historical years.

2. Calculate the total number of hourly CPS1 excursions that were less than 100 percent[13].

3. Calculate the total number of occurrences where the BA had three or more consecutive hours of excursions.

4. Group CPS1 excursions by BA, Year and Month to evaluate trends over the historical period.

5. Group CPS1 excursions by BA, Year and Hour of day to evaluate trends over the historical period.

The following steps are recommended. Again, this is based on the method that the NERC RS intends to use for their analysis (see Appendix E) with three years of historical data, but the BA can evaluate CPS1 based on their own preferred methodology and adjust the screening tests accordingly. Note that while conducting any of the screening tests below, if an increased rate of excursions is observed, the BA may be experiencing ramping issues. *Figure 2.2* provides a decision tree for the CPS1 evaluation.

1. **Screening Test #1:** Review total number of hourly CPS excursions that are less than 100% to determine if the BA has experienced more than 3000 hours of excursions.

2. **Screening Test #2:** Review number of 3 or more consecutive hourly CPS excursions that are less than 100% to determine if the BA has experienced more than 300 instances of 3 hour consecutive excursions.

3. **Screening Test #3:** Using CPS1 data (hourly and three-consecutive hour data) that has been grouped by year by month, review trends to determine if the BA is experiencing an increased excursion rate.

4. **Screening Test #4:** Using CPS1 data (hourly and three-consecutive hour data) that has been grouped by year by hour of day, review trends to determine if the BA is experiencing an increased excursion rate.

If a BA fails any of the screening tests, the NERC RS will contact the BA and recommend they perform detailed reviews of CPS1 past and potential future performance by evaluating hourly CPS1 trends. Examples of detailed reviews are presented in Figures 2.1 and Figures 2.2 below. By reviewing detailed hourly CPS1 performance, the BA can identify time periods where it is currently experiencing, or may experience, issues meeting ramping up or ramping down requirements.

---

[12] CPS1 is a statistical measure of a BA's area control error (ACE) variability in combination with the interconnection frequency error from scheduled frequency. It measures the covariance between the ACE of a BA and the frequency deviation of the interconnection, which is equal to the sum of the ACEs of all of the BAs. CPS1 assigns each BA a share of the responsibility for controlling the interconnection's steady-state frequency. The CPS1 score is reported to NERC on a monthly basis and averaged over a 12-month moving window. A violation of CPS1 occurs whenever a BA's CPS1 score for the 12-month moving window falls below 100 percent.

[13] RS may select values other that 100% for evaluation purposes. Value is for evaluation only and is not associated with any compliance requirements.
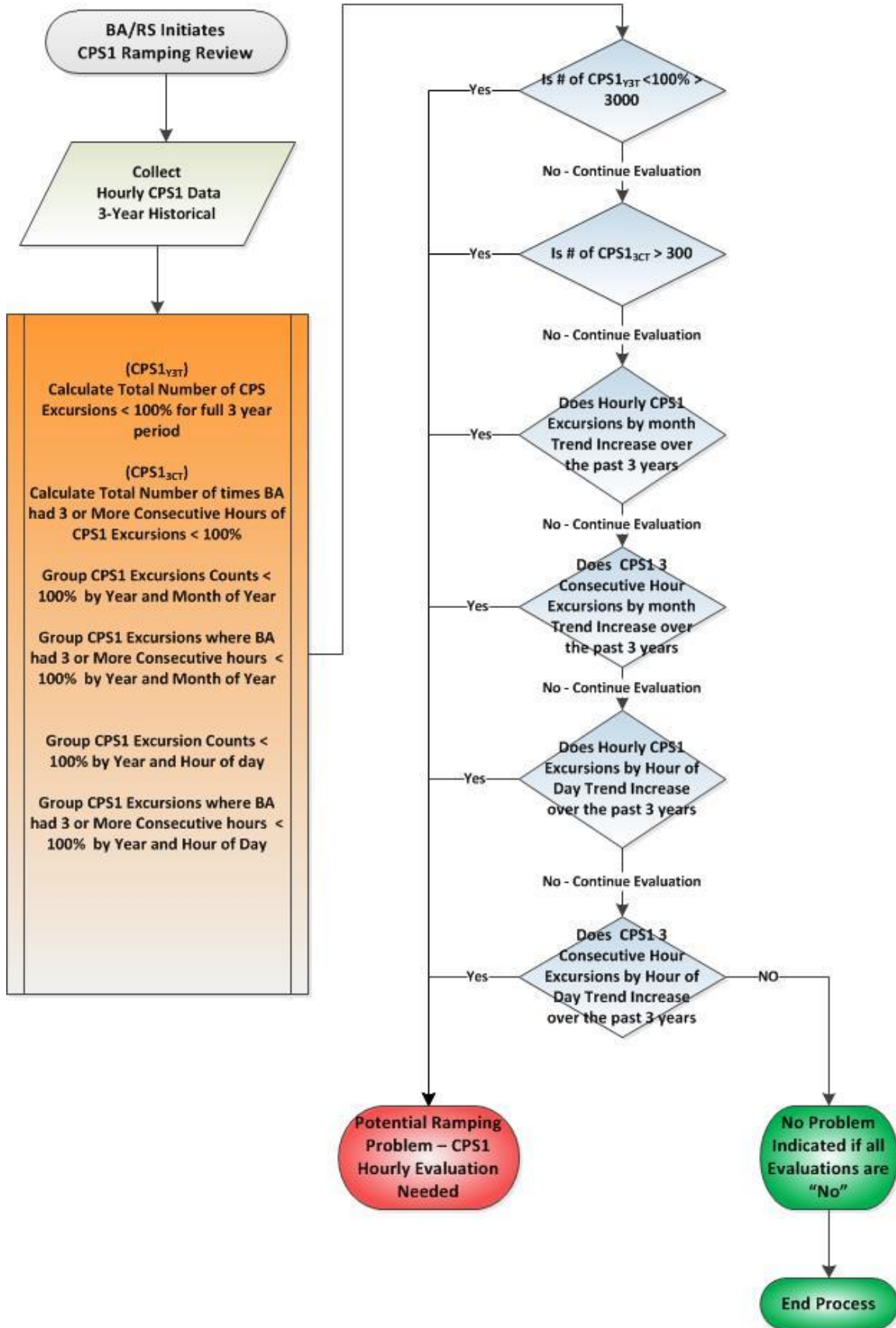
Figure 2.2: CPS1 Evaluation Decision Tree for Ramping

Example of CPS1 Hourly Evaluation

As an example of the evaluation, the plot in Figure 2.3 for a BA shows that the average CPS1 score for March was 122.5 percent, which is very good performance. However, the CPS1 scores for hours 22, 23 and 24 (red bars) were consistently below 100 percent for the entire month. This is an indication that the downward ramping capability is insufficient and the BA has an overreliance on the interconnection for balancing during these hours.
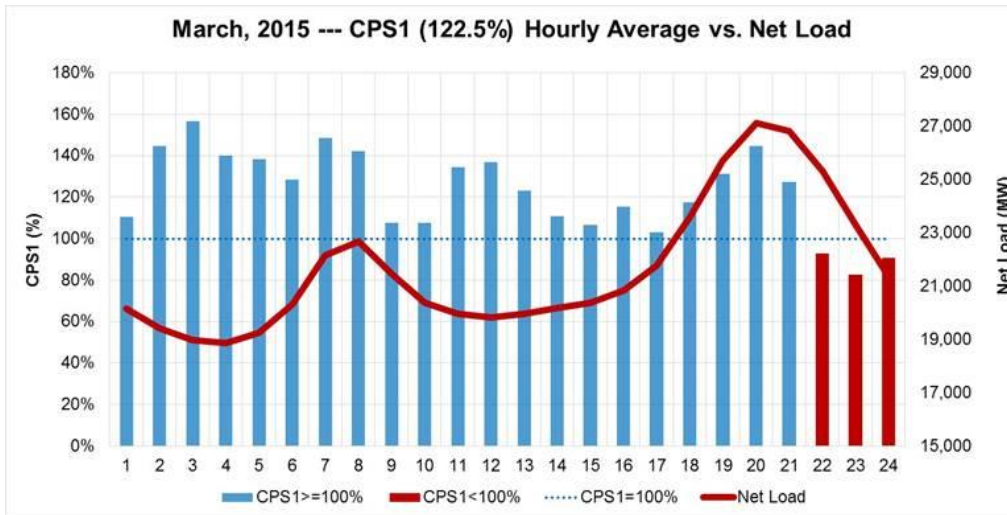

Figure 2.3: CPS1 Hourly Average (March 2015)

Another example of the evaluation, the plot in Figure 2.4 for a BA, shows that the average CPS1 score for December was 120.1 percent, which is good performance; however, the CPS1 scores for Hours 9 and 10 and Hours 16, 17, and 18 (red bars) were consistently below 100 percent for the entire month. This is an indication of downward ramping insufficiency for two hours and an upward ramping insufficiency for three hours, during which times the BA was leaning on the interconnection.
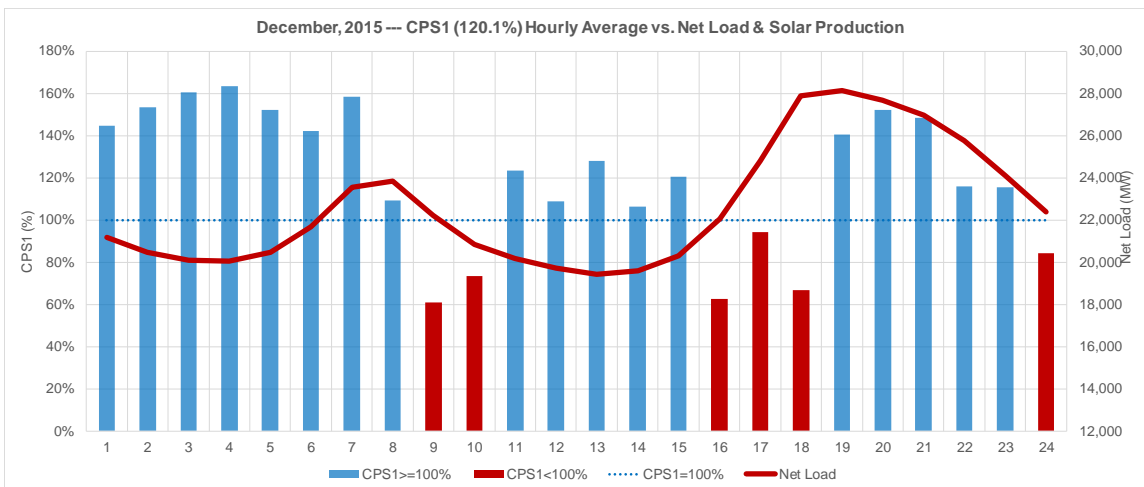

Figure 2.4: CPS1 Hourly Average (December 2015)

Chapter 3: Voltage Support

Background and Recommendations

Voltage regulation and reactive resource management are critical parts of planning and operating the BPS. Maintaining adequate voltage profiles across the BPS both pre- and post-contingency is a function of the reactive resources available and their location and utilization. Low-voltage events can cascade (creating a widespread event) and, while high-voltage events do not necessarily cascade, they can result in serious equipment damage and potentially in the loss of life. The ability to control the production and absorption of reactive power often becomes the driving force behind studying and operating the BPS over a wide range of conditions. This is especially so in areas where weak transmission systems supply load and generation or the transmission network can be subjected to heavy power transfers.

While voltage and reactive performance over a wide range of system conditions must be considered, there is a significant justification for focusing the analysis on much smaller sub-areas of the BPS. Due to the inability to efficiently or effectively transport reactive power on the BPS, planners and operators should consider defining sub-areas of the BPS within their study areas that have their own unique set of voltage/reactive performance issues that can only be rectified by local resources and remedial actions. The BPS varies widely from study area to study area based on the specific topology and electrical system characteristics. There has been a long held belief in many utility sectors that reactive resource reserves are the key to maintaining a robust system. This is true at a high level, but if those abundant reactive reserves are not electrically close to where they are needed, they are likely to be ineffective in managing voltages throughout the system.

Planning reactive capability can be viewed as an optimization problem. For example, consideration must be given to:

- **Dynamic Versus Static:** While static compensation devices are less costly, static devices are much less flexible than dynamic ones in dealing with the full range of potential operational issues. Use of static devices can result in operational issues and be unacceptable without due consideration of things like discreet switching bandwidths, the lack of timely response to contingencies, protection coordination, switching timing, and coordination with reclosing, etc.

- **BPS Versus Distribution Compensation:** While reactive compensation on the distribution system is generally cost effective, the BPS should not be dependent on the distribution system for support. First, the operation of the two systems is likely performed in different control rooms. Second, there is a limit to the ability of distribution compensation to correct/mitigate reactive concerns on the BPS.

- **Determining the Appropriate Split between Dynamic Versus Static Reactive Reserves:** This can be accomplished through managing the installed resources by determining the amount of operational dynamic reserves needed and maintaining that dynamic reserve by adjusting on-line static reserves. For example, switching in capacitor banks to increase the dynamic lagging reactive reserve in an area to prepare to respond to a post-contingent low-voltage condition.

- **Including Load Management Programs:** These can include things like voltage reduction and/or load shedding capabilities

- **Detailed Analyses Examining the Interaction and Use between Static and Dynamic Devices:** The choice of size, magnitude, and number of static and dynamic reactive devices should include an evaluation of expected use. These analyses can optimize the type of devices chosen and their size. For example, shunt capacitors can be installed to re-center a dynamic reactive device to respond to high and low voltage issues. However, the number of switch operations for a device is limited before maintenance must be performed. Therefore, the switching lifetime should be considered for local voltage control. If the switching occurs too frequently, there is an increased risk to switch failure due to overuse. Therefore, the

dynamic device should have a wider range of voltage setpoints to accomplish the local voltage control design goal.

Therefore, it is recommended that the Planning Coordinator and Transmission Planner, based on their knowledge of the unique characteristics of their system, should collaboratively develop a criterion for defining "reactive power sufficiency" sub-areas within their study areas. They should also involve the Reliability Coordinator and Transmission Operator who may have useful insight into sub-area characteristics based on operational experience in the selection process. Logical sub-areas can and are likely to cross company and jurisdictional boundaries. They should be developed based on electrical characteristics and reactive performance of the BPS.

Consider the inherent difference between a typical large urban area and that of a typical large rural area in such variables as: load level and load power factor (LPF), various overhead and underground transmission network configurations, dynamic and static reactive resources and the appropriate minimum and maximum voltage limits. While these two areas may be within the same Reliability Coordinator/BA/Planning Coordinator footprint, low voltage at heavy load periods, high voltage at light load periods, and the reactive performance of each area can vary significantly. That difference will drive both the criteria and the type of planning studies required to meet the objective of developing a robust reliable system. Similarly, those same differences may impact the way the real-time operations are managed. The determination of appropriate sub-areas within the larger footprint becomes the primary first step in planning and operating the BPS from a reactive power sufficiency standpoint. Considerations in defining sub-areas of the system are:

- Reactive performance within the footprint both pre- and post-contingency

  - Insufficient reactive compensation in a single area can impact or cascade to neighboring areas and affect overall BPS operation

  - The system can reach a state where even though voltages appear to be within the normal range, most available reactive resources are exhausted and the next contingency can degrade voltages and reactive performance pushing the BPS quickly into unacceptable performance

  - The loss of high voltage BPS facilities can load remaining facilities more heavily and result in significantly increased losses, which will negatively impact the voltage profile and the need for additional reactive resources

  - Outages of major reactive resources not only remove the reactive capability but can also result in large MW swings and increased losses, which will negatively impact the voltage profile and reactive resources

- Real power import, export, and flow-through characteristics (e.g., large power transfers within or between sub-sets can significantly increase reactive power losses, which will negatively impact the voltage profile and the need for additional reactive resources)

- Transmission topology and characteristics (e.g., high surge impedance loading, where real power transfers can reach a point that reactive consumption of the transmission system exceeds available reactive supply, negatively impacting the voltage profile)

- Charging from cables or long overhead lines during light load periods, where these facilities may produce voltages so high that leading reactive capability is exhausted and circuits must be opened, as a last alternative, to reduce voltages

- The types of reactive resources available which will have different lead/lag characteristics:

  - Synchronous generators and condensers

  - Nonsynchronous/inverter-based resources

  - Static devices (i.e., shunt capacitors, reactors, etc.)

- ▪ Nonsynchronous dynamic devices, (i.e., SVCs, STATCOMs, DVARS, etc.)

- ▪ HVDC terminals, such as voltage source converters that can supply reactive capability

- ▪ Line compensation devices that can be switched in and out such as for series compensation

- Real and reactive load distribution (while this is mainly in the distribution area of the system, real load and LPF can have both a positive and negative impact on the BPS and that contribution must be accounted for): transmission and distribution system planning and operations must be coordinated when managing reactive power transfer into and out of the distributions system. This will result in an effective optimization of installed reactive resources on each system to improve reliability and reduce cost.

Once appropriate sub-areas have been defined, the planners and operators must ensure compliance with all applicable NERC standards. More stringent regional or local criteria may also be used. The planners and operators will then need to develop appropriate sufficiency measures applicable to each unique sub-area. Sufficiency measures can, and most likely will, differ by sub-area based on the sub-area's reactive power characteristics.

As an example, a large urban area that has limited reactive resources and routinely imports large amounts of real power may have more stringent min/max voltage limits. It may also have certain on-line dynamic reactive resources and load power factor requirements.

A rural sub-area with relatively light real-power loads and long high-impedance overhead lines may have more relaxed min/max voltage limits and may need to maintain a specific reactive reserve. Reactive power sufficiency requirements by sub-area need to fit the reactive characteristics of the specific sub-area in order to ensure reliability is maintained and a certain degree of reactive power performance robustness is built into the system. Reactive power sufficiency measures therefore are not one size fits all. Potential sufficiency measures for sub-areas can be seen in Table 2.1.

In general, the defined sub-areas should be somewhat autonomous relative to their reactive and voltage performance for N-1-1 conditions. This construct negates one sub-area relying too heavily on adjacent sub-areas for reactive support, thereby decreasing the chance of a cascading event. This also inherently builds in reactive margin so that in real-time, where multiple events may occur, there is enough robustness to prevent a widespread event from occurring. The end result is to plan and operate each sub-area so that the reliability of the broader BPS is maintained. There are numerous verified methods of studying, planning, building, and operating the BPS. At a high level, those aspects are addressed in the NERC Reliability Guideline: Reactive Power Planning and Operations (*reference when available*) that is summarized below.

Table 2.1: Potential Sub-area Sufficiency Measures				
Potential Sufficiency Measures	**Sub-area A** Typical large urban area with high MW imports	**Sub-area B** Typical rural area with matched amounts of generation and load	**Sub-area C** Typical large urban area with significant cable system	**Sub-area D** Typical rural area with large exports of MW
Pre-contingency Min/max voltage	1.01/1.04	0.98/1.04	0.98/1.045	1.01/1.04
Post-contingency Min/max voltage	.95/1.05	.92/1.05	.90/1.06	.95/1.05
On-line reactive reserves	Minimum of 3 generators in the sub-area, 100 MVA or greater, on-line above 80% load level	n/a	1 generator on line at all times, 2 of 3 SVC on-line at all times, within the sub-area	n/a
MW dispatch	Commit 1/3 of fast start resources at 95% load level	n/a	n/a	n/a
Min LPF required at heavy load periods	0.98 lagging	0.96 lagging	0.94 lagging	0.80 lagging
Max LPF allowed at light load periods	0.98 leading	0.97 leading	0.99 leading	0.98 leading
MW Imports	Limit pre contingency import into the sub-area to 90% of maximum capability	n/a	n/a	n/a
Line/Cable switching	n/a	n/a	Remove 2 of 8 cables in the sub-area at load levels below 40%	Open 1 of 4 possible overhead lines within the sub-area at load levels below 30%
Static reactive resources	All BPS capacitors on-line above 85% load	n/a	All BPS shunt reactors on-line below 50% load	n/a

Summary of NERC Reactive Power Planning and Operations Guideline

The NERC Reactive Power Planning and Operations Guideline (reactive guideline) provides strategies and recommended practices for reactive power operations, planning, and voltage control. Where applicable, it links them to NERC reliability standards.

Voltage must be controlled to protect system reliability and move power where it is needed in both normal and post-disturbance operations. Voltage is supported through the supply of reactive power; hence, reactive power is required to transfer large amounts of real power across the grid to serve the loads.

Reactive power must be supplied locally, mainly because of its dependence on voltage difference. It is usually necessary to site reactive support very near or at the location that is deficient. This reactive guideline covers specific areas such as:

- Basic discussion of voltage requirements and reactive supply characteristics

- Applicable NERC reliability standards

- Industry-accepted response and analysis time frames

- Various analysis techniques

- Planning and operational considerations

- Transmission and distribution dependencies

- Industry Practices:

    - Reactive planning practices and procedures

    - Transient voltage response criteria

Reactive power planning and reactive needs in the operating horizon vary significantly between Transmission Operators across the NERC footprint. Thus, in the operating horizon, sufficient reactive resources need to be available to ensure that voltage levels and reactive flows are controlled and maintained within limits in real-time or near real-time to protect equipment and the reliable operation of the interconnection. This includes ensuring reliable post contingency system conditions.

In the planning horizon, sufficient reactive resources need to be planned for so that the transmission system can meet planning performance requirements under a wide range of probable contingencies and result in a system that can be operated reliably over a broad spectrum of system conditions. These conditions range from the peak period of a heavy load day through an extremely light load period, such as the early morning hours during a spring or fall holiday weekend.

In addition to a myriad of changing conditions, there is an inherent difference in the reactive behavior between a typical large urban area and that of a typical large rural area in variables like load level and load power factor, various overhead and underground transmission network configurations, dynamic and static reactive resources, and the appropriate minimum and maximum voltage limits. In some areas, steady state voltage control and/or transient voltage response are of greater concern than in other areas. Due to these inherent differences the reactive performance of a system is often computed based on a more granular sub-system basis.

While reactive power planning and operational needs vary significantly across North America and Canada based on local system characteristics and practices, NERC Reliability Standards define a set of requirements to ensure reliable planning and operation of the bulk power system. The primary standards codifying reactive power

requirements in the operations and planning time frames include VAR-001-4, VAR-002-4, TPL-001-4, and TOP-004-2.

This reactive guideline provides detailed background discussion regarding the reactive characteristics of system components, system reactive behavior, and analytical techniques. This reactive guideline also provides recommended practices for reactive power operations, planning, and voltage control. Appendices A-G included in the reactive guideline provides the reactive planning practices, procedures, and requirements for an array of entities across North America. The entities have provided this information as typical of their practices and philosophies at the time that this reactive guideline was developed. These practices are expected to evolve over time.

Summary and Conclusions

This whitepaper explores directional measures to help the industry understand, evaluate, and prepare for the changing resource mix and other changes to the traditional characteristics of generation and load resources. This transition can and does have a profound impact on the transmission and distribution infrastructure. While manageable, these impacts have to be accounted for in energy policymaking, system planning, and system operations.

Below are the interim findings of the ERS Working Group. The working group will provide a report in 2017 on the trending analysis based on these recommendations.

Frequency

While Measure 4 is the more comprehensive perspective on frequency support as it looks at the minimum frequency following a disturbance and the timing of recovery, this interim report proposes a way for interconnections to identify minimum synchronous inertia while assuming current frequency response capabilities. If the interconnection is trending toward minimum synchronous inertia, there are alternatives (e.g., faster frequency response) that can be considered to supplement current frequency control practices. The NERC Resources Subcommittee (RS) started data collection for the frequency measures in June 2016.

The working group will also review alternative approaches for frequency support in very low-inertia conditions. The working group will monitor studies being conducted in other areas and by other parties (e.g., efforts in Ireland and the EU-funded MIGRATE project, IEEE, CIGRE, universities, national labs) that are investigating these issues and may be looking at future grid scenarios, such as for hours when most or all generation is from inverter-based resources. However, it is still important to forecast minimum inertia hours that will be encountered in future years as this drives the pace of developing adjustments that are necessary and appropriate.

Ramping and Balancing

The guideline for Measure 6 (Net Demand Ramping Variability) provides BAs with a screening approach to evaluate ramping concerns given their dispatchable and nondispatchable resource mix. Once the BA starts nearing ramping concerns, there are alternatives, such as use of more flexible resources, load shifting, demand management and changes to nondispatchable resources (both conventional and renewable).

The screening approach includes a CPS1 analysis and was developed in collaboration with the NERC RS and the NERC Reliability Assessment Subcommittee (NERC RAS) to provide BAs with methods to identify trends and indications of potential ramping concerns within their footprint. The guideline encourages coordination between the BA and the NERC subcommittees, and the subcommittees will provide ongoing support and monitoring assistance.

Voltage

Voltage control is a localized issue and is best evaluated on a sub-area basis. This may become more important with changes in the generation mix that may alter local voltage and short-circuit characteristics. The recommended approach has been documented in the NERC *Reliability Guideline on Reactive Power Planning and Operations.*

The NERC Performance Analysis Subcommittee (PAS) is collecting and the System Analysis and Modeling Subcommittee (SAMS) is analyzing data at the BA level concerning the monitoring/trending of voltage characteristics and will determine if this has ongoing value.

Appendix A: Summary Table of Recommendations

The complete list of ERSTF Measures and Industry Practices as defined in the 2015 ERSTF Final Framework Report are listed below in Table A.1:

Reference Number	Title	Brief Description	BA or Interconnection Level	ERSTF Recommendation	Ongoing Responsibility
1	Synchronous Inertial Response(SIR) at an Interconnection Level	Measure of kinetic energy at the interconnection level. It provides both a historical and future (3-years-out) view.	Interconnection	Measure	Resource Subcommittee and Frequency Working Group
2	Initial Frequency Deviation Following Largest Contingency	At minimum SIR conditions from Measure 1, determine the frequency deviation within the first 0.5 seconds following the largest contingency (as defined by the Resource Contingency Criteria (RCC) in BAL-003-1 for each interconnection).	Interconnection	Measure	Resource Subcommittee and Frequency Working Group
3	Synchronous Inertial Response at a BA Level	Measure 3 is exactly the same as Measure 1 but performed at the BA level. It provides both a historical and future (3 years out) view and will help a BA identify SIR-related issues as its generation mix changes.	BA	Measure	Resource Subcommittee and Frequency Working Group
4	Frequency Response at Interconnection Level	Measure 4 is a comprehensive set of frequency response measures at all relevant time frames: Point A to C frequency response in MW/0.1 Hz, Point A to B frequency response in MW/0.1 Hz (similar to ALR1-12), C:B Ratio, C:C' Ratio as well as three time-based measures (t_0 to t_C, t_C to $t_{C'}$, t_0 to $t_{C'}$), capturing speed of frequency response and response withdrawal.	Interconnection	Measure	Resource Subcommittee and Frequency Working Group
5	Real Time Inertial Model	Develop a real-time model of inertia including voltage stability limits and transmission overloads as criteria. This is an operator tool for situational awareness and alerts them if the system is nearing a limit and any corrective action is required.	BA	Industry Practice	BA

Table A.1: Summary of Measures and Industry Practices Recommendations

6	Net Demand Ramping Variability	Measure of net demand ramping variability at the BA level. It provides both a historical and future view.	BA	Measure	Reliability Assessment Subcommittee
7	Reactive Capability on the System	At critical load levels, measure static & dynamic reactive capability per total MW on the transmission system and track load power factor for distribution at low side of transmission buses.	TOP	Measure	Performance Analysis Subcommittee and the System Analysis and Modeling Subcommittee
8	Voltage Performance of the System	Measure to track the number of voltage exceedances that were incurred in real-time operations. This should include both pre-contingency exceedances and post-contingency exceedances. Planners should consider ways to identify critical fault-induced delayed voltage recovery (FIDVR) buses and buses with low short-circuit levels.	No Further Action	No Further Action	No Further Action
9	Overall System Reactive Performance	When an event occurs on the system related to reactive capability and voltage performance, measure to determine if the overall system strength poses a reliability risk. Adequate reactive margin and voltage performance should be evaluated across all horizons (planning, seasonal, real time). This type of post-mortem analysis comports with various requirements in existing and proposed NERC standards.	BA	Industry Practice	Event Analysis Subcommittee
10	System Strength	Based on short circuit contribution considerations, determine if low system strength poses a potential reliability risk. When necessary, calculate short circuit ratios to identify areas that may require monitoring or additional study.	Planning Coordinator	Industry Practice	Planning Coordinator

Appendix B: Inertia Data Collection for ERSTF Measures 1, 2 and 3

The *Essential Reliability Services Task Force Measures Framework Report* (December 2015) previously proposed the following measures for trending synchronous inertia:

- Measures 1 and 3: Kinetic energy (synchronous inertia) at the interconnection level (Measure 1) or at the BA level (Measure 3).

- Measure 2: System frequency at minimum inertia conditions in the first half second following the largest contingency (as defined by the Resource Contingency Criteria in BAL-003-1)

Following up on this report, the ERSWG developed a spreadsheet for Measure 1 and 3 with the data elements, shown in Table B.1. This supported a data collection effort to inform trending and analysis.

Table B.1: Data Elements for Measure 1 and 3 Analysis				
Timestamp in UTC	Total Interconnect inertia contribution, MVA*s	Interconnect Load, MW	Production from Nonsynch Gen, MW	DC Exports/Imports, MW

Data collection for this measure started in June 2016 by the NERC Resources Subcommittee (RS). The aim is to set up inertia calculators and harmonize data collection for all four interconnections and then telemeter the data to NERC in real time with the following data resolution shown in Table B.2 (based on data availability):

Table B.2: Resolution of Data for Measure 1 and 3 Analysis			
ERCOT	EI	WECC	HQ
4 s	15 min	1 min	4 s

Peak Reliability Council is calculating inertia for the WECC system with one-minute resolution, based on real-time telemetry of generator production. Generator inertia constants and MVA ratings are available from WECC planning dynamic models.

Hydro Quebec and ERCOT, being single BAs, are collecting data for their respective areas with four-second resolution, based on real-time telemetry of generator production and inertia constants and MVA ratings from their respective dynamic models.

For the Eastern Interconnection, real-time on-line status of all synchronous generators is being acquired. Currently, nonsynchronous generator data is not available for the EI, but efforts are underway to begin capturing this data.

At the NERC RS meeting in July, 2016 all four interconnections were able to demonstrate inertia data availability for the period from June 1 through July 15. From collected synchronous inertia data, Measure 2 can be calculated for each interconnection at minimum inertia conditions as described in the *ERSTF Framework Report.*

Appendix C: Case Study—Calculation of Minimum Inertia in ERCOT

The example calculation below follows the sufficiency guideline process using actual data from ERCOT.

1. Calculate system inertia for every hour in a year in MVA*seconds. Figure C.1 shows box plots for system inertia in ERCOT.[14] Minimum inertia in each year and supporting data for the wind penetration record in each year are shown in Tables C.1 and C.2.
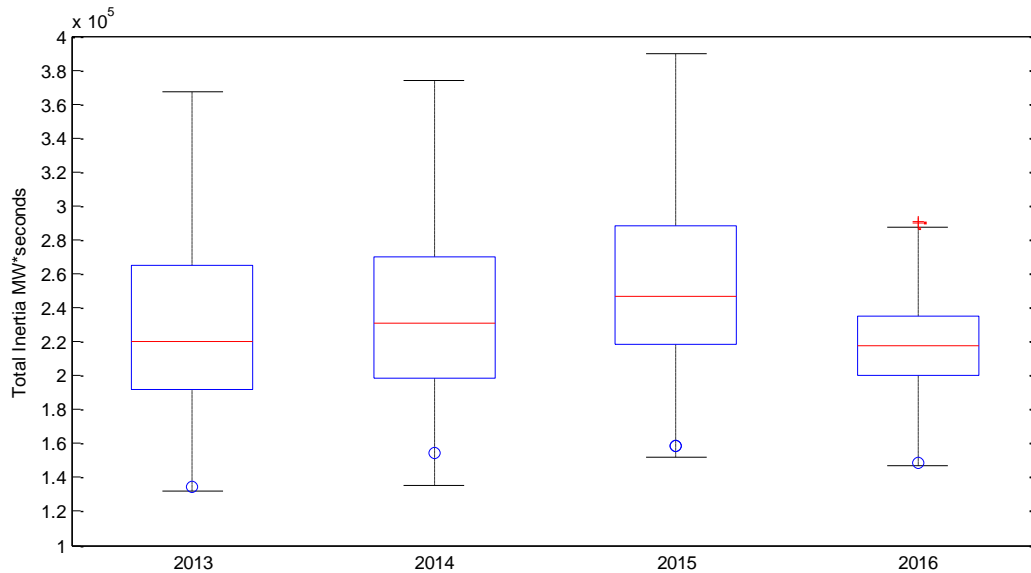

Figure C.1: Boxplot of the System Inertia from 01/01/2013 to 3/31/2016

Table C.1: Lowest Inertia in Different Year (GWs)			
2013	**2014**	**2015**	**2016**
132	135	152	147

---

[14] Note these values are somewhat higher than shown in the NERC *ERSTF Measures Framework Report*. This is due to different accounting of inertia contribution from private use networks (PUNs). A PUN is an electric network connected to the ERCOT Transmission Grid that contains load that is not directly metered by ERCOT (i.e., load that is typically netted with internal generation). Previously, PUN generation was only considered to be online if PUN net production was above 5 MW. However, in reality, PUN generation can be online and producing power even while exporting 0 MW into ERCOT. In this situation, PUN generation is still synchronously interconnected with the ERCOT grid and will provide inertia during contingency events. After recognizing this, these PUN generators with gross production above 5 MW are now being included in the total system inertia calculation.

Table C.2: Wind Generation Power and Load Demand at the Time of Record of Each Year				
	3/9 3:15 am	**11/3 2:28:56 am**	**12/20 3:05 am**	**3/23 1:10 am**
Installed Capacity ($P_{wind\_installed}$), MW	10,570	12,527	16,170	16,547
P_{wind}/P_{load}	35.8%	39.93%	44.71%	48.28%
P_{wind}, MW	8,773	9,882	13,058	13,154
$P_{wind}/P_{wind\_installed}$	83%	78%	81%	79%
Inertia, MW*seconds	134,196	154,599	158,970	148,798

2. As shown in Table 6, ERCOT's minimum system inertia for the past three years was 132 GWs in 2013. Using the expression below, it is possible to calculate the rate of change of frequency based on ERCOT's RCC, which is loss of two nuclear units with a total capacity of 2750 MW.

$$\text{RoCoF} = \Delta P_{MW}/(2* (KE_{min}-KE_{RCC}))*60 \qquad [\text{Hz/s}]$$

At 132 GWs of inertia, the RoCoF will be 0.69 Hz/s and it will take slightly over a second to get to the first stage of UFLS at 59.3 Hz.

3. ERCOT has load resources with under frequency relays providing up to 50 percent of the responsive reserve service requirement. These load resources will trip in 0.5 seconds after frequency reaches 59.7 Hz. With RoCoF as calculated above, 59.7 Hz will be reached in 0.43 seconds after the event, which means that load resources will trip in about 0.93 seconds after an event, and system frequency will be at about 59.35 Hz at that time and therefore will be arrested before involuntarily UFLS.

4. Gradually scaling down system inertia and performing the same analysis as in the previous step shows that at about 105 GWs of inertia, 59.7 Hz is reached at 0.35 seconds and 59.3 Hz is reached at 0.85 second, therefore this inertia value can be considered to be the theoretical minimum inertia for ERCOT, considering current frequency control practices.

5. Following the "bottom up" approach described in the previous section, establish if system inertia can fall below the minimum value found in Step 4 or if there is there sufficient inertia on-line from the generation units that are always on-line for other reasons:

    a. Currently there are no reliability-must-run units in ERCOT.

    b. Private use networks (PUNs) must be considered. Table 8 shows minimum and maximum of their inertia contribution for the past 3 years. The minimum in the past 3 years was 31.5 GWs.

Table 8: Inertia Contribution from PUNs		
Year	**Min inertia from PUN, GWs**	**Max inertia from PUNs, GWs**
2013	31.5	53.7
2014	48	52
2015	36	58

c. At least three nuclear units are normally on-line with total inertia of at least 18.4 GWs.

d. The NERC IFRO requirement is currently -381 MW/0.1 Hz. To ensure that the IFRO is fulfilled, ERCOT will procure sufficient amounts of responsive reserve service (RRS). The minimum "RRS from generation" requirement during low-inertia hours is 1348 MW. According to ERCOT Nodal Protocols, there is a 20% limit on how much capacity a single resource can offer towards RRS. Thus, 1348 MW will be distributed between generators as 6740/0.9=7488 MVA. In ERCOT various generation types are qualified to provide RRS:

   i. If all RRS is provided by combined-cycle (CC) units with H=4.97 seconds[15], the inertia contribution for generation resources providing RRS will be at least about 37 GWs.

   ii. If 1348 MW is provided by the smallest coal or gas-steam units that are qualified for providing RSS, then the inertia contribution for generation resources providing RRS will be about 22 GWs.

   iii. If RRS is provided by qualified hydro units in synchronous condenser mode (to their full capacity) and the rest is provided by qualified and participating gas steam units with the lowest inertia, then the inertial contribution from RRS resources would be 20 GWs.

e. Note that neither PUNs nor nuclear units are qualified to provide RRS in ERCOT. Therefore, there is no double counting of these units in "must run" inertia calculations.

f. Assume that, at the worst case, regulation reserves are provided by the same units that are providing RRS, so no additional unit commitment will be necessary to provide regulation.

6. If the ERCOT system has sufficient nonsynchronous, renewable generation to serve system load, then nuclear units, PUNs and units providing RRS will be the only ones supplying synchronous inertia to the system. Based on the considerations above, this total synchronous inertia will always be 70-87 GWs, unless PUNs significantly change their operating strategies (e.g., due to frequent negative energy prices at nighttime during winter/spring).

7. Theoretical minimum inertia obtained in step 4 is 105 GWs,

8. "Must run" units will only provide about 70-87 GWs of inertia, which is less than minimum sufficient inertia of 105 GWs.

9. Currently, due to low gas prices, Combine Cycle units are displacing Coal units in unit commitment. To provide 25 GWs of inertial response, an additional 9 CC units (600 MVA, H=4.97s) would be required to be on-line.

10. Dynamic studies conducted in 2014 also showed that when the inertia of the ERCOT system is less than 100 GWs, loss of the two largest units will cause voltage oscillations and voltage control issues in the Texas Panhandle area.

The example shown in Figure C.2 also shows how the minimum inertia needs may be reduced by increasing the frequency trigger or by shortening response time (from 30 cycles to 20 cycles) for fast frequency response in ERCOT. Note that at low-inertia values, the rate of change of frequency is faster and inclusion of load damping (equation (3) in the *ERSTF Measures Framework Report*) in the frequency calculation becomes important. The results in Figure C.2 are based on the analytical calculation using equation (3) in the *ERSTF Measures Framework Report* and are not yet verified by dynamic simulations. As discussed in step 9 above, the dynamic studies are showing voltage oscillations when operating below 100 GWs inertia. These issues would need to be addressed before moving on to lower inertia operation.

---

[15] In 2014 about 68% of generation RRS was provided from Combined Cycle (CC) generation, and in ERCOT the average CC has an inertia constant of 4.97 seconds on 600 MVA base.
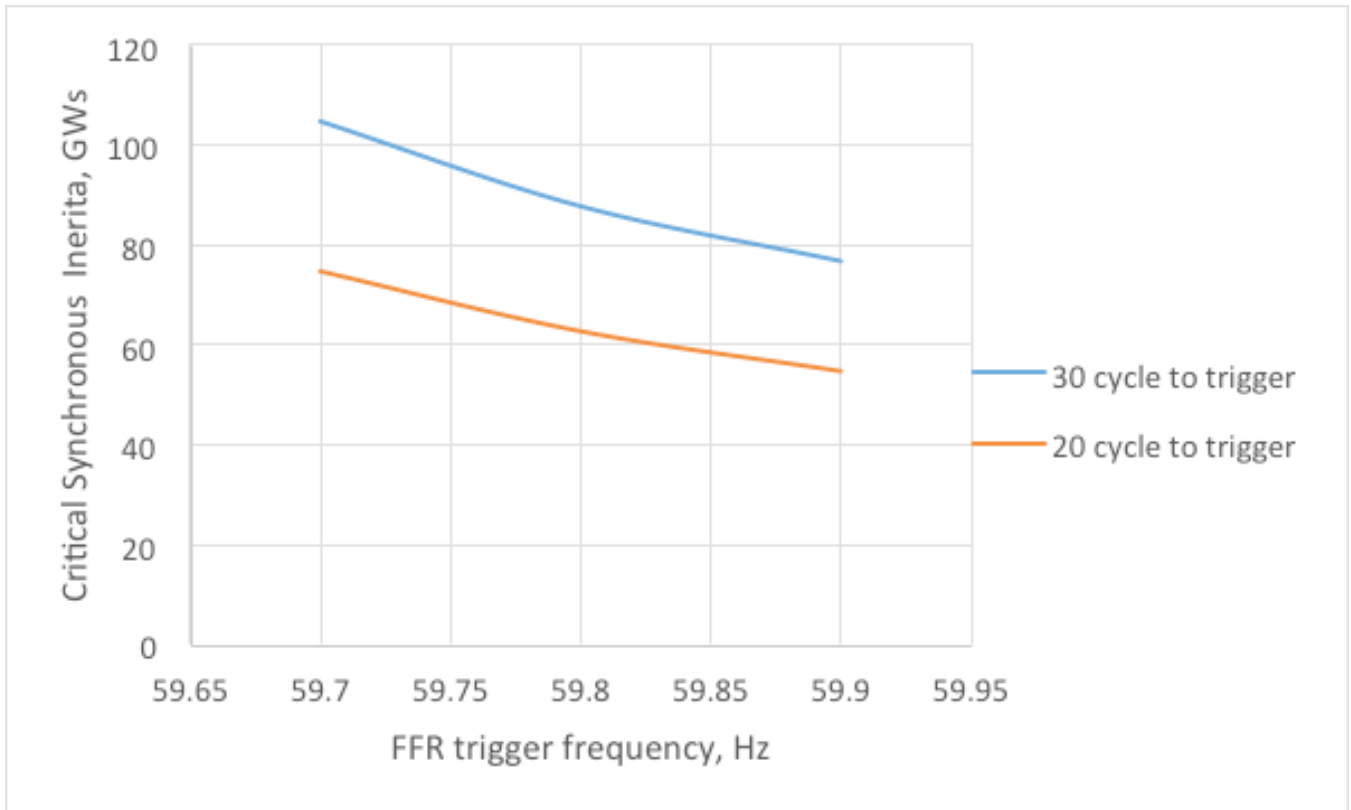
Figure C.2: Effect of FFR Trigger Frequency on SIR

Appendix D: Operational Monitoring and Forecasting of Synchronous Inertia

As changes to the operation of conventional generation resources and the continuous growth of wind and solar generation bring more uncertainties to how the grid is operated, there emerges a need to monitor system inertia in real time and help operators by predicting inertia for the near future. The system inertia will be added to voltage and frequency as another key indicator for monitoring the system operating conditions. Monitoring of synchronous inertia and frequency deviation based on Resource Contingency Criteria was also recommended as Measure 5 (a recommended industry practice) by the ERSTF.

ERCOT's Inertia Monitoring Tool and Dashboard

In order to streamline monitoring and analysis of the system inertia, as well as the contribution by individual generation types, ERCOT staff set up various data points to calculate synchronous inertia once a minute[16] by resource type and system total.

Additionally, a real-time inertia dashboard was experimentally set up to monitor inertia as shown in Figure D.1. The dashboard also shows the last 24 hours of inertia contributions by generator type. Monitoring by type is done to enable a more granular analysis of inertia trends.
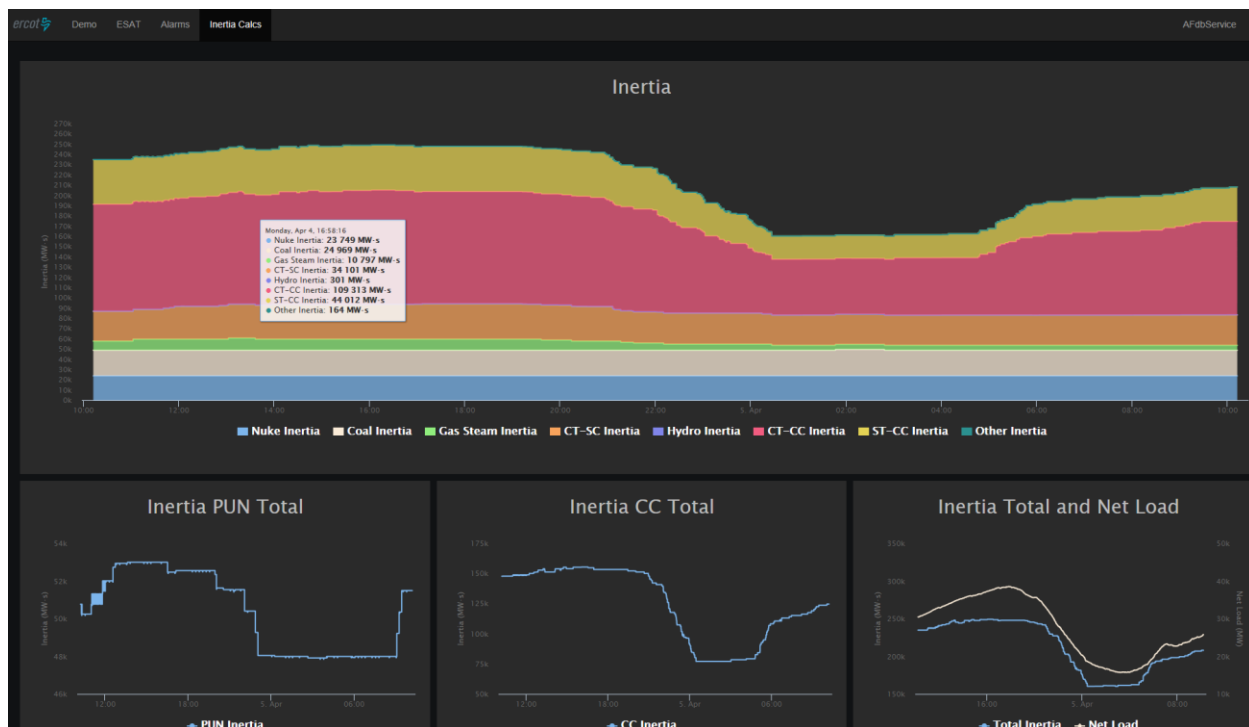
Figure D.1: A Dashboard to Monitor Inertia in Real Time

## ERCOT's Inertia Prediction Tool

The system inertia can be predicted using the current operation plan (COP) information submitted by the generators to provide some foresight into what the actual condition will be. The probability distribution function of three-hour-ahead inertia estimation error using COP data in 2015 is shown in Figure D.1 (a). This prediction will provide an opportunity for operators to evaluate the sufficiency of procured RRS as well as recognize the risk for extremely low-inertia conditions ahead of time and prepare a mitigation plan if needed. Figure D.1(b) shows the performance of three-hour-ahead RRS estimation, evaluated using 2015 data, with 3.78% of time showing an

---

[16] Currently being updated to calculate every four seconds to align with the Measure 1 and 2 data collection process recommended by ERSWG.

under estimation (three-hour-ahead RRS estimation is less than the actual RRS requirement) and 10.1% of time showing an over estimation. One example of under-estimation of the RRS requirement is depicted in Figure D.3. On November 4, 2015, the system lambda (the dashed line) dropped below $10/MWh in the earlier morning. In response to this, some generating units, which submitted "on-line" status in their COP ahead of time were eventually off-line in real time, which resulted in an over-estimation of the system inertia. When the energy price recovered, the generation units came back on-line so that three-hour-ahead estimation of the system inertia matched well with the actual system inertia after 6:00 am.
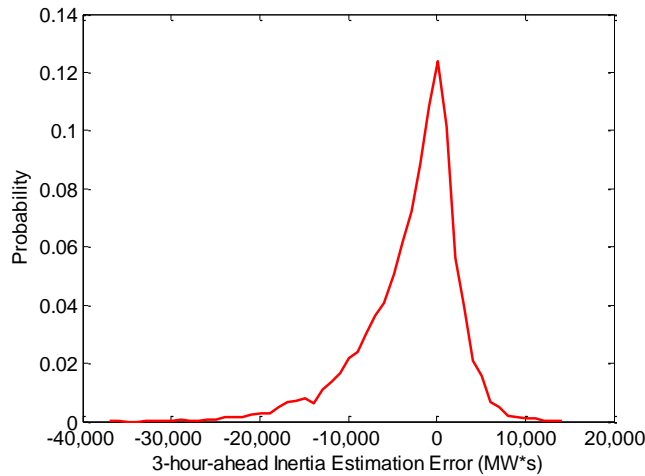
(a)   error in three-hour-ahead inertia estimation
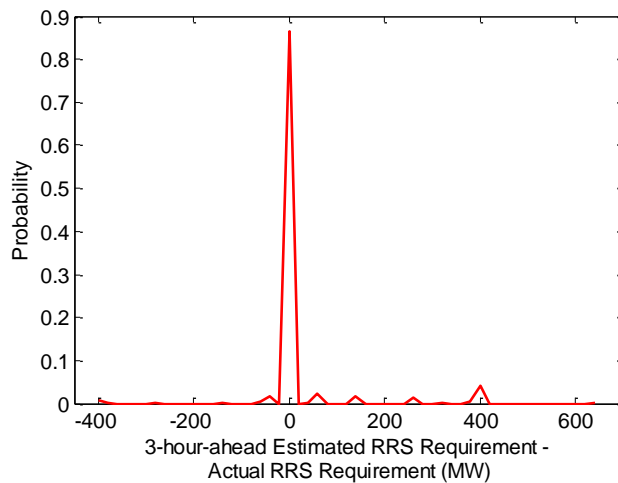
(b)   error in three-hour RRS requirement estimation

Figure D.2: Three-Hour-Ahead Estimation Error of Inertia and RRS Requirement
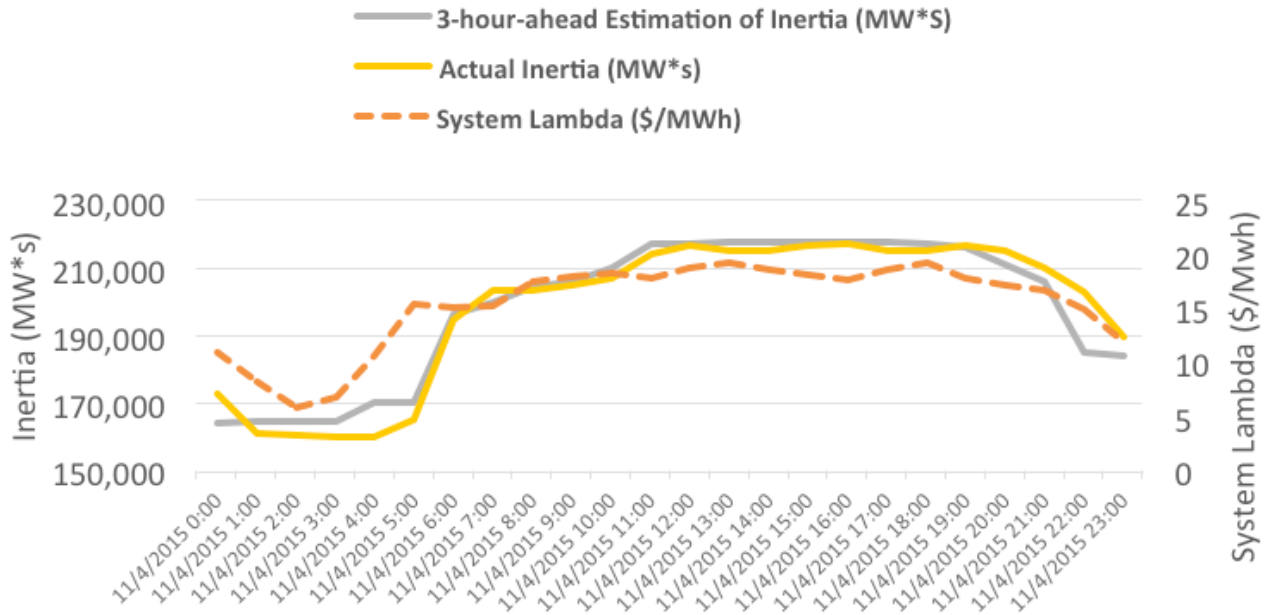
Figure D.3: System Lambda and Inertia on Nov. 4, 2015

Appendix E: NERC Resources Subcommittee Ramping Review

The NERC Resources Subcommittee (RS) will use the CPS1 Evaluation Methodology to evaluate CPS1 interconnection performance on a quarterly basis. The NERC RS will look at various levels of excursions (e.g., 100%, -100%, -200%, -300%, etc.) to identify BAs that may need assistance. This assistance may include aid in developing partnerships with other companies who have previously experienced the problem and aid in identifying potential solutions or assistance with additional problem evaluation.

The NERC RS will evaluate the interconnections by BA to identify areas that may have ramping deficiencies. The NERC RS will perform the analysis, coordinate with the NERC RAS on the results, and offer assistance to BAs as needed. Figures E.1–E.6 show examples of the various ways the data will be evaluated by the NERC RS.

While a BA is not required to provide hourly CPS1 data to the NERC RS, it is highly encouraged to support interconnection ramping evaluations.
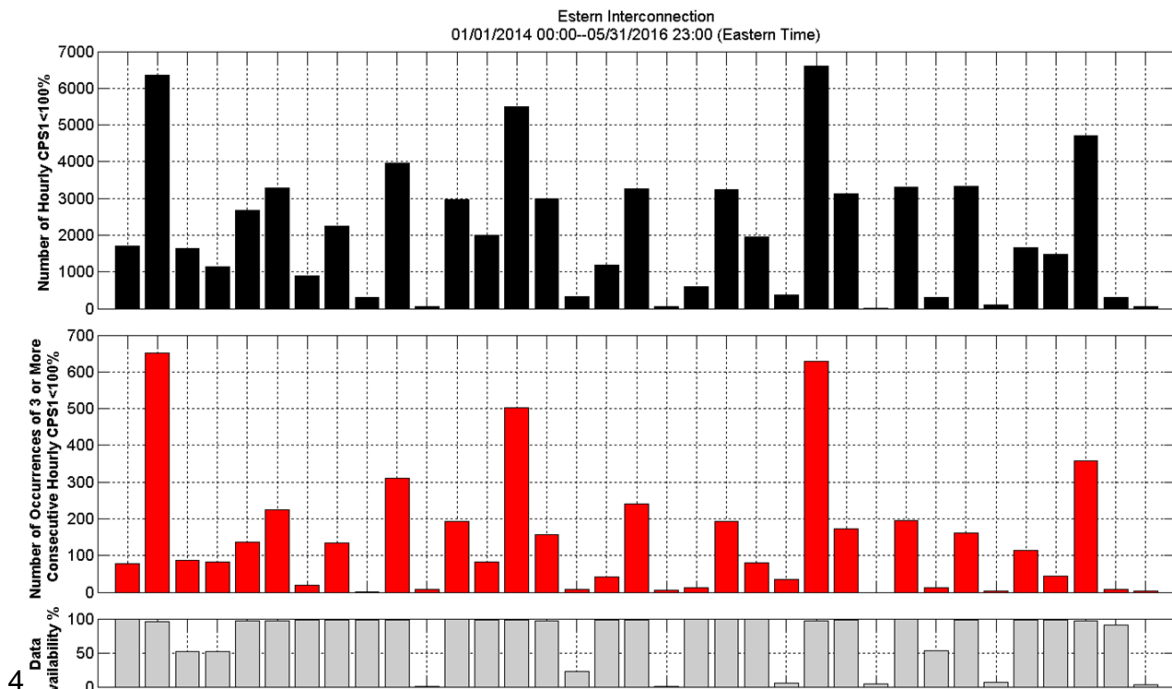

Figure E.1: Eastern Interconnection (EI) Data

CPS1 Violation Counts
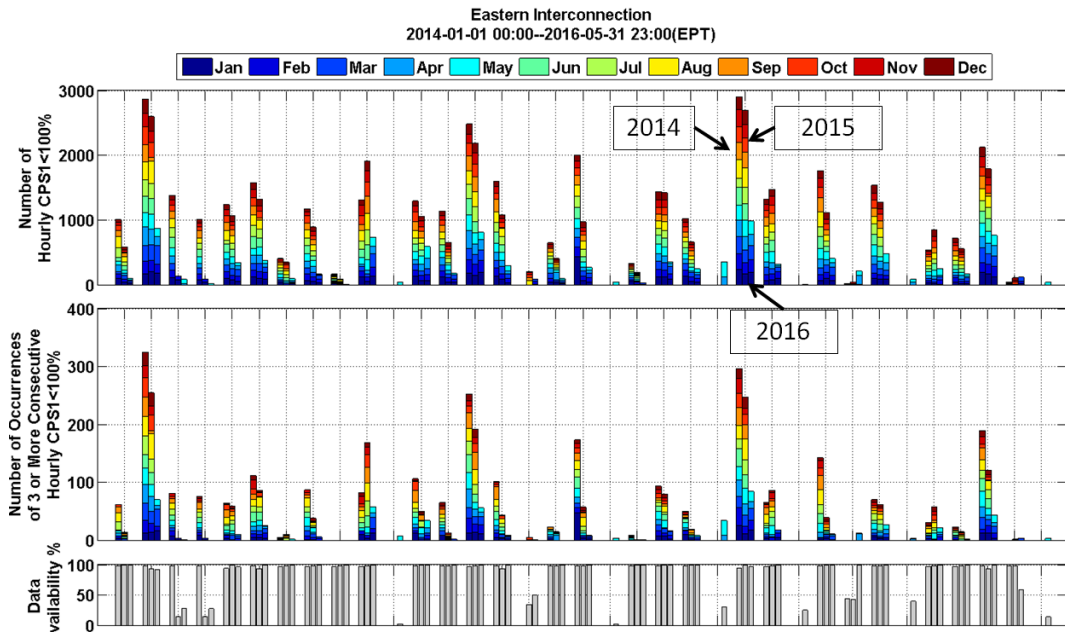## Group by BA, Year and Month of Year

Figure E.2: EI CPS1 Violation Counts by BA, Year, and Month of Year

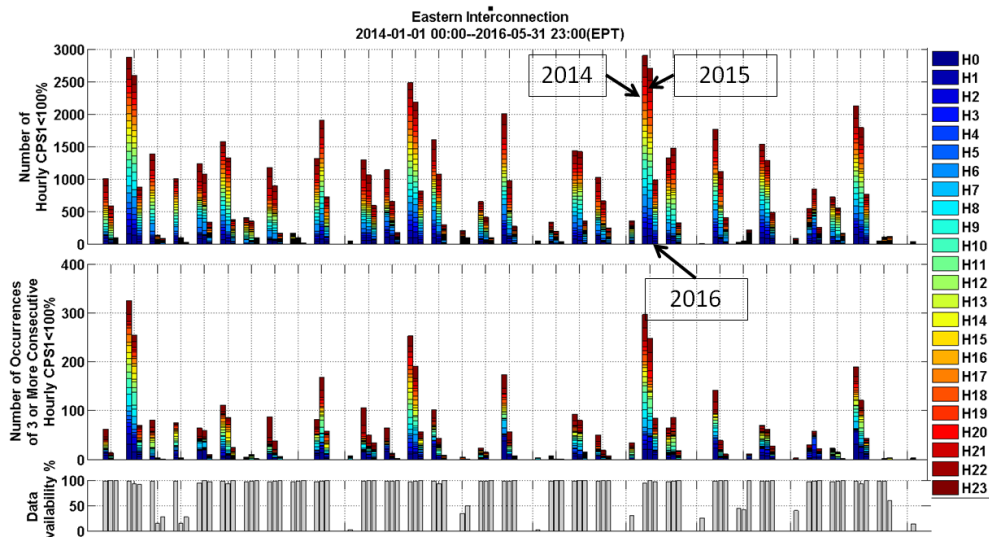## CPS1 Violation Counts
Group by BA, Year and Hour of Day

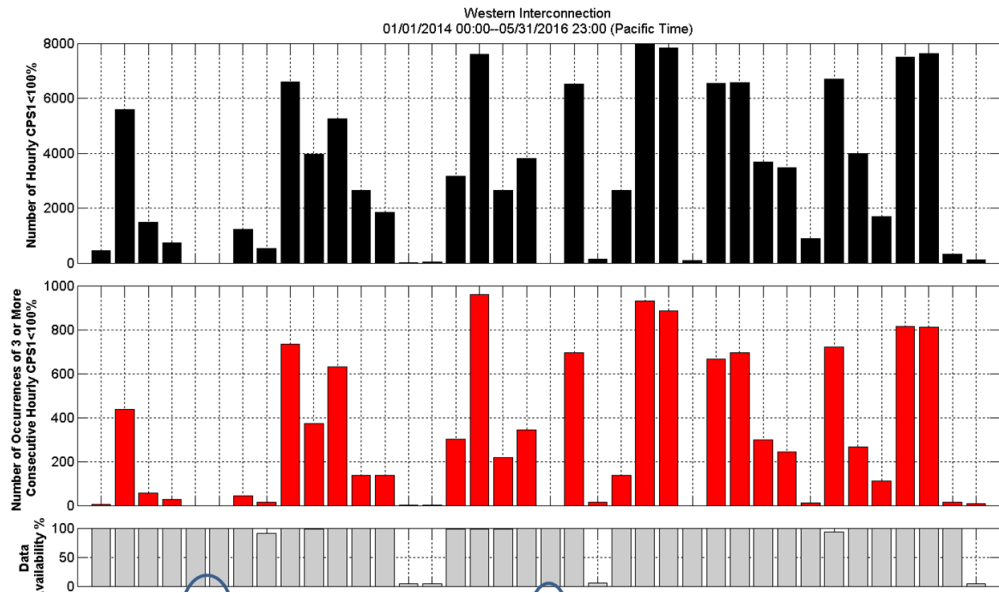

Figure E.3: EI CPS1 Violation Counts by BA, Year, and Hour of Day

Figure E.4: Western Interconnection (WI) Data
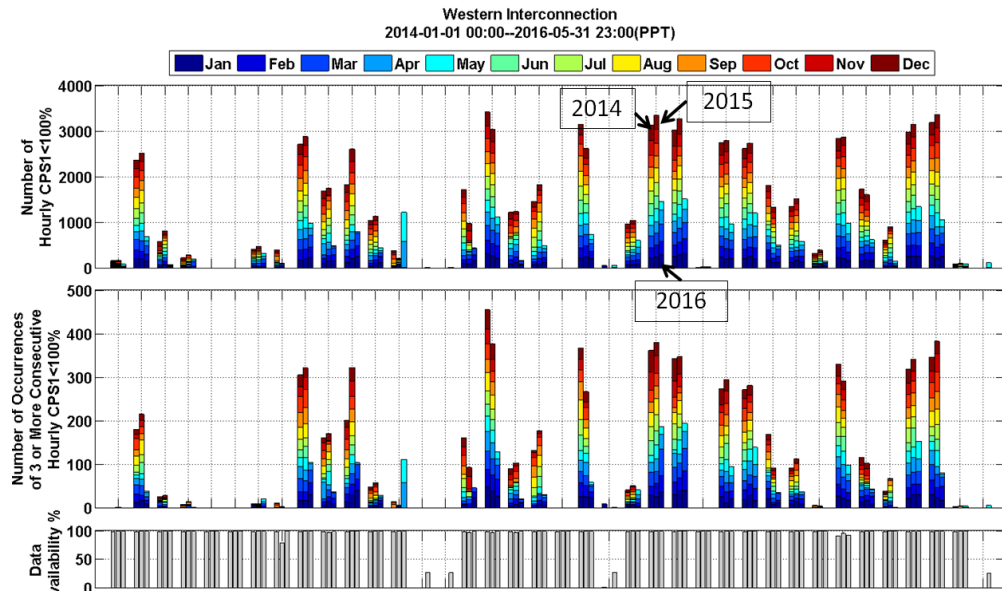
CPS1 Violation Counts
Group by BA, Year and Month of Year

Figure E.5: WI CPS1 Violation Counts by BA, Year, and Month of Year

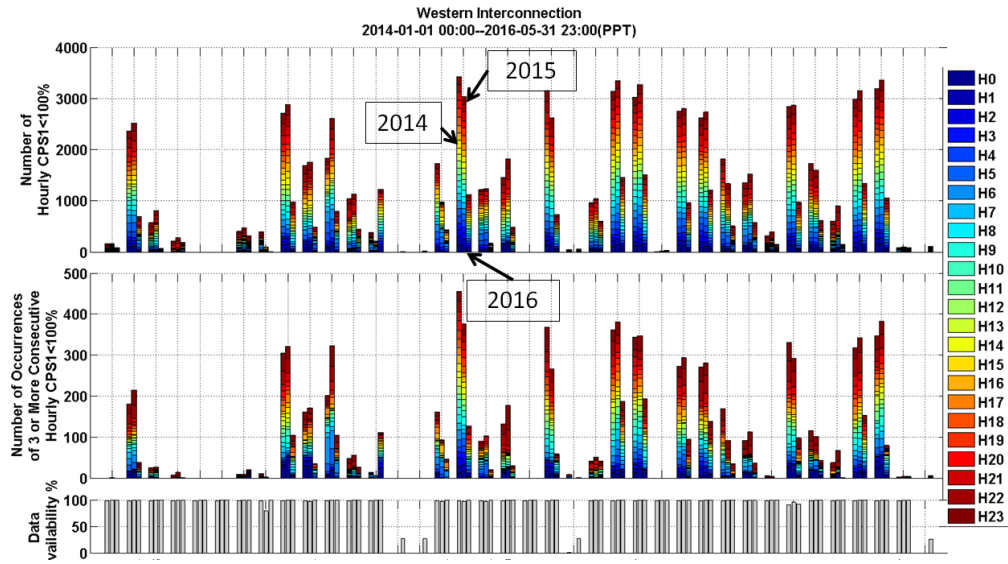# CPS1 Violation Counts
Group by BA, Year and Hour of Day

Figure E.6: WI CPS1 Violation Counts by BA, Year, and Hour of Day

Appendix F: Methodology for Analyzing Trends in Future Ramping Needs

Once a BA has identified that ramping and CPS1 concerns may be on the horizon, a deeper investigation may be required.

Load, Variable Generation, Net-Load ramping analysis

Overview
This is a basic level of ramping analysis that looks at load, variable generation, and net load (load minus variable generation) ramps. The BA can use the net demand ramping variability methodology, as defined in Appendix B of the 2015 *ERS Measures Framework Report*, for this evaluation process. This level of analysis is helpful in seeing what kind of ramps can be expected in the future years.

Methodology
The required data for this level of analysis is load, variable generation, and net load time series for the future study period. These could come from wind, solar, and load profiles/forecasts for the system. Planned wind and solar resources with signed interconnection agreements (and, possibly, other financial commitments) should be included in the analysis. It is also important to evaluate the added impact of Distributed Energy Resources (DERs) on system load. High levels of DERs can offset demand from the transmission system during sunrise, resulting in the need to back off dispatchable resources and increase production during sunset.

The recommended wind, solar and load data resolution is one minute.

One way to obtain high-resolution data for a future year is to use actual one-minute load, wind, and solar data for the most recent year. Using data from the same year maintains the correlation between the load demand and the solar/wind production.

Future Year Load: Scale the actual one-minute load data points by the load growth factor for the future year,

1. **Future$_{Year-Wind}$**: Scale the actual one-minute wind production data by the factor (Wind Capacity $_{Future\_Year}$ /Wind Capacity $_{Current\_Year}$)

2. **FutureYear-Solar**: Scale the actual one-minute solar production data by the factor (Solar Capacity Future_Year /Solar Capacity Current_Year)

3. Calculate the one-minute net-load data for the future year

4. **Future Year_Net_Load**: Future Year_Load - FutureYear_Wind - FutureYear_Solar

However, note that scaling wind, solar, and load time series is not ideal since this methodology will likely exaggerate variability. As wind and solar plants are built in different geographical locations, the resulting production time series is likely to be smoother than the scaled time series due to geographical dispersion of the resources.

ERCOT, for example, is using DNV GL's SFLEX tool for this, which can interpolate hourly wind, solar, and load data into as high as one-minute resolution. This tool takes geographical location of the planned wind and solar plants into account when producing high-resolution power production time series from hourly data.

Once the high-resolution data is obtained, the analysis itself can be performed by using dedicated tools such as Electric Power Research Institute's (EPRI) InFLEXion tool or Excel. Ramps are calculated by simply finding the difference between the load, variable generation, and net load at two points in time. A shifting window should be used in calculating ramps. For example, if looking at one-hour ramps while using one-minute resolution data, the

first two ramps of a day would be from 00:00 to 01:00 and from 00:01 to 01:01. ERSTF Measure 6 recommended calculating one-hour and three-hour net-load upward and downward ramping magnitude for a future year (e.g., three years out). This is required to ensure that dispatchable resources have the ramping capability to meet expected and unexpected changes within an operating hour. Multiple three-hour ramps are important for BAs with high penetrations of solar generation because of the higher ramping needs of the system during sunrise and sunset.

Once the ramps are calculated, they can be displayed in a number of different ways, such as a duration curve of all ramps encountered or the magnitude of ramps encountered at certain times of day during certain seasons.

Appendix G: Resource Flexibility Analysis Method

Overview

A more complex level of analysis looks at both the ramps a system could encounter in the future years as well as existing and planned resources' ability to serve the ramps the system may encounter. The amount of flexibility a fleet has available at a given time, referred to as "available flexibility," is a function of the unit capacities, minimum generation levels, ramp rates, start-up time, shut down times, and so on. Subtracting the net load from this available flexibility yields the system's net flexibility. Net flexibility can be calculated for specific times by comparing the available flexibility at a given time with the net load ramp encountered at that time, but it can also be calculated for hypothetical ramps. For example, the available flexibility at a specific date and time could be compared to the 98th percentile of all ramps encountered when beginning at the net load level seen at that time. This is helpful in planning for multiple risk levels.

EPRI has created several flexibility metrics that can be helpful in performing ramping analysis. One metric is periods of flexibility deficit (PFD), which is a count of the number of intervals in the study period where net flexibility is below zero. PFD is calculated for a specific time horizon as well as for ramp direction. Another metric is expected unserved ramp (EUR), which is the total magnitude of negative net flexibility.

Performing this level of analysis requires significantly more data and tools than the basic analysis described above. ERCOT is using the methodology described below for performing this analysis. The methodology assumes that the study period is a full year and the time resolution is five minutes, but it can easily be modified to accommodate different study periods and time resolutions.

Methodology

The first step in performing flexibility study for a future year is to prepare a generator database of all generators that will be active in the study year; this includes existing generators that are not scheduled to retire before the study year as well as planned generators that are scheduled to be on-line during the study year. Load resources and storage resources that are being used in energy dispatch or for providing ancillary services (AS) should also be included.

It is necessary to obtain hourly profiles for wind, solar, and load for the study year as well as interpolate these profiles into higher resolution (preferably one-min) as described in Appendix H above.

The hourly wind, solar, and load profiles are used as input into a production cost simulation tool, along with the generator database prepared earlier. Any resources with fixed schedules or must run units should be modeled as such in the production cost simulation. ERCOT uses Energy Exemplar's Plexos to perform 365 daily optimization runs with a one-day look ahead in order to obtain unit commitment for the entire study year. Unit commitment also considers AS requirements for each hour in a year. If any capacity is reserved on resources (in addition to or in lieu of AS markets) to meet NERC standards (e.g., BAL-001, BAL-002, and BAL-003), this should also be included. Qualification of resources for provision of various AS (or reserves) should be assigned correctly in the production cost simulation. Any other regulations and limits that may affect available flexibility should be modeled in detail.

Once each unit's commitment status is found for each hour of the year from the hourly production cost simulation run and the temporal resolution of the hourly wind, solar, and load profiles have been increased to five-minute resolution, a sequential production cost simulation run is performed in order to obtain each unit's five-minute dispatch for the study year to serve the given net load. Again, in this dispatch run, capacity reserved for provision of the AS is not available for dispatch, except for nonspinning reserves that are generally available in ERCOT in energy scarcity situations. This dispatch data, along with the generator database and the five-minute load data,

are the inputs to EPRI's InFLEXion tool. The InFLEXion tool can perform all of the analysis described in this document as well as additional ramping/flexibility analyses.

Sample results

Figure G.1 shows the average one-hour net load ramps for each hour of each day in the same time period. Negative numbers in this figure are times where the average net load ramp was in the downward direction rather than upward.
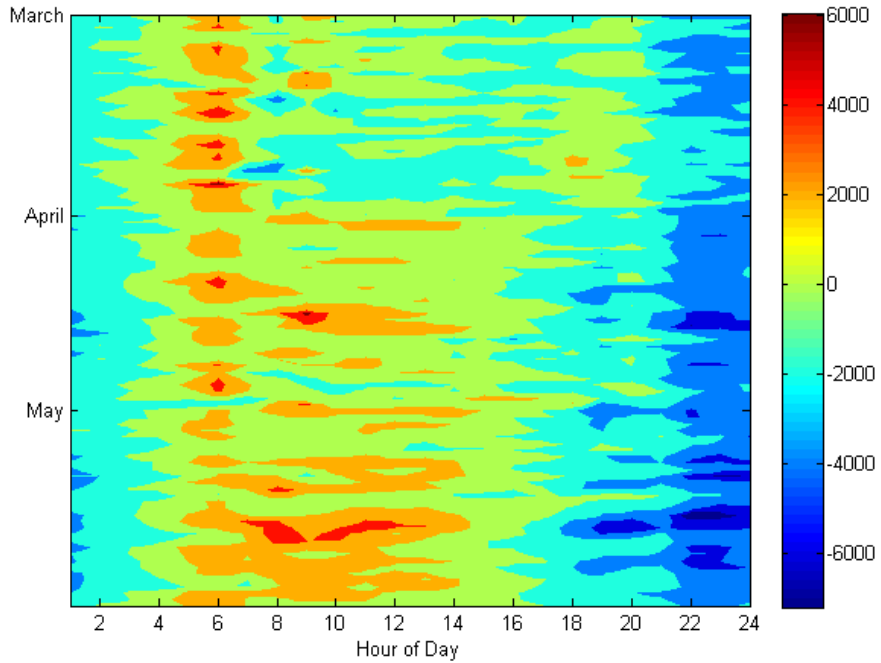

Figure G.1: Net load One-Hour Ramps in Spring 2017

Figure G.2 shows the average amount of available upward flexibility that can be provided by the resource fleet to serve one-hour net load ramps for each hour of each day in the spring of 2017. As mentioned above, the available flexibility is calculated based on what units are on-line and able to increase generation from their current operation point (as obtained from production cost simulation) as well as units that are off-line but could start up quickly enough to help serve the ramp if necessary, including off-line nonspinning reserves. Note that in this level of analysis, all reserves are considered available to follow expected net load ramps.
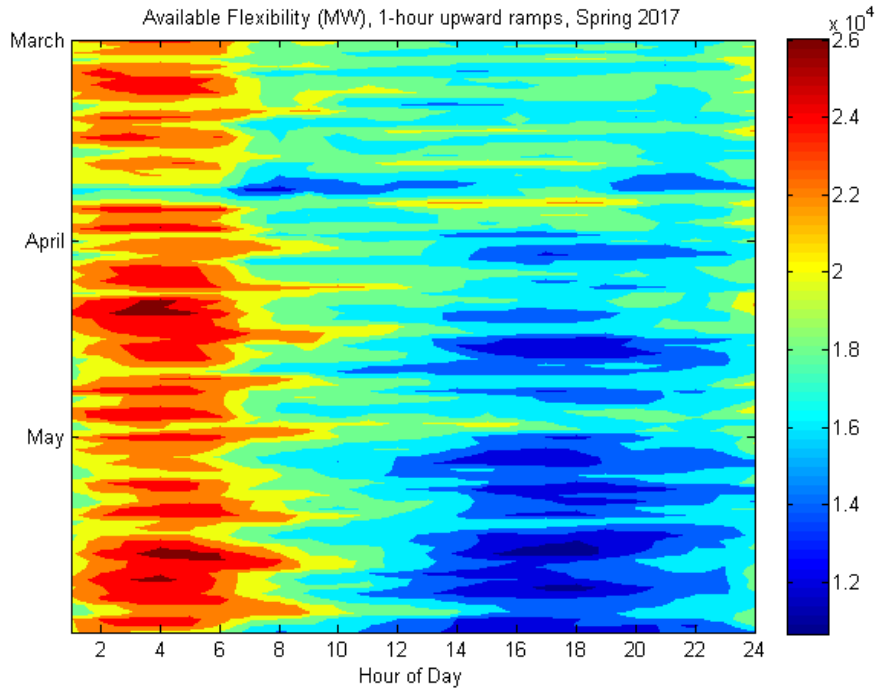

Figure G.2: Available flexibility for One-Hour Upward Ramps in Spring 2017

Figure.G.3 shows the average net flexibility at each hour of each day in the time period. The net flexibility is simply the difference between the available flexibility and the net load ramp at a given time. If there were times with negative net flexibility, this would indicate intervals where the system may face a ramping deficit.
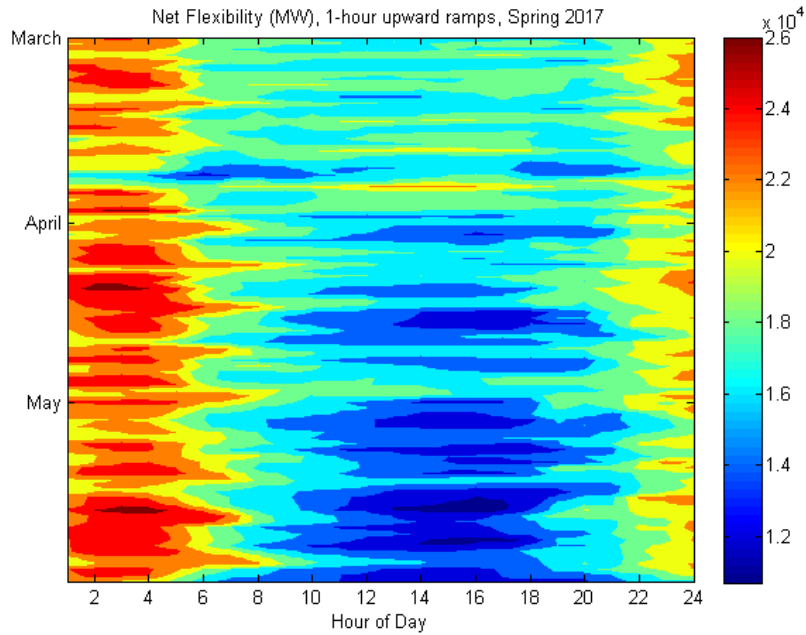

Figure G.3: Net flexibility for One-Hour Upward Ramps in Spring 2017

High-Level Work Flow and Data Requirements to Perform Future Ramping Analysis

Ramping Analysis

- Calculate load, VER generation, net load ramps

  - Net load = load - VER generation

  - Data needed: load time series, VER time series

    - Resolution: less than or equal to the time horizon of the ramps you are interested in

  - Ramps can be calculated in dedicated tools like InFLEXion or simply Excel

  - Shifting window (e.g. for one-hour ramps: 00:00-01:00, 00:05-01:05, etc.)

Flexibility Analysis

- Calculate available flexibility, net flexibility

  - Available flexibility based on units status, capacity, min gen, ramp rate, start time, stop time, etc.

  - Net flexibility = available flexibility - net load ramp

  - For net flexibility, can either compare available flexibility to actual ramps or potential ramps, e.g. maximum ramp encountered at given net load level or 98th percentile of all 05:00–05:05 ramps encountered in the time series

  - Available/net flexibility can be calculated in dedicated tools like InFLEXion or manually (more difficult)

- Flexibility metrics: Periods of Flexibility Deficit, Expected Unserved Ramp

  - Periods of Flexibility Deficit (PFD): number of intervals with negative net flexibility

  - Expected Unserved Ramp (EUR): total magnitude of negative net flexibility

Preparing the Data

1. Prepare generator database of existing units as well as units that are scheduled to come on-line by/during a study year

2. Run hourly production cost simulation of study year for unit commitment (e.g. ERCOT uses Energy Exemplar's tool Plexos)

   a. Daily optimization with 1 day look ahead

3. Increase resolution as necessary (e.g. ERCOT uses DNV GL's SFLEX tool to increase one-hour wind, solar, and load data to five-minute resolution)

   a. Hourly profiles that were used for Plexos run are also fed into SFLEX to obtain five-minute data resolution

4. Rerun production cost simulation at higher resolution for unit dispatch

   a. Sequential run

5. Feed results into InFLEXion

Appendix H: CAISO Examples of Ramping and Operational Tools

CAISO is experiencing rapid growth of VER resources, and their experience has influenced the development of these recommendations in the following ways:

As CAISO has seen increased VER generation penetration, the methodologies outlined in this guideline have been a valuable tool for CAISO. With over 12,000 MW of transmission- and distribution-connected solar resources and 5,800 MW of wind within its footprint, CAISO evaluates its net-load ramping needs on a monthly basis to evaluate the adequacy of its fleet make-up and recommend changes as necessary.

Additionally, CAISO monitors its control performance standards (CPS1) scores on an hourly basis to ensure it does not unnecessarily lean on the interconnection during hours of greatest ramping needs. By evaluating its CPS1 score on an hourly basis, CAISO was able to identify a correlation between significant intra-hour and multi-hour ramps and CPS1 excursions below 100 percent across the same time frame. As shown in, Figure H.1 CAISO's maximum three-hour upward ramping need was approximately 6,000 MW in 2013. However, in 2016, CAISO's three-hour upward ramping has already reached 10,800 MW and by 2020 its three-hour ramping requirements are expected to reach 13,000 MW.
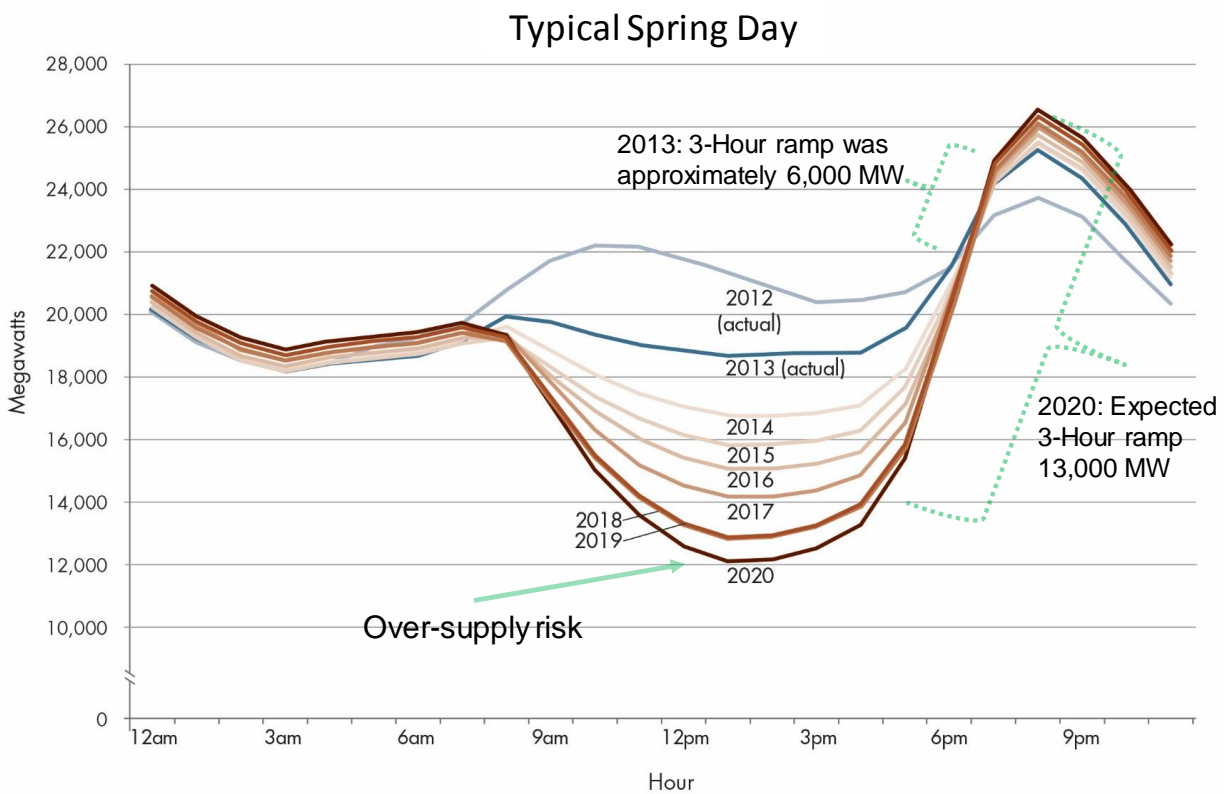

Figure H.1: Increase in Ramping Needs as VER Production Increases

Net Ramping Variability Tool

CAISO is in the process of developing an operating tool to alert operators if the committed resources have adequate ramping capability to meet the uncertainty and variability associated with load, wind, and solar for any particular hour. The operator would have the option of viewing flexible capacity needs (or *Gen. Requirements Forecast*[17]) and availability of flexible capacity for various time frames in the future.

Net Load Ramping Capacity Requirement

The generation forecast screen refreshes every minute and can show the forecast trajectory for up to 24 hours in advance. The operator can also define a confidence range around the *Generation Requirements Forecast* curve by activating the confidence ranges in percent. A 90% confidence band means that there is a 90% confidence level that the *Generation Requirements Forecast* curve can meet the uncertainty and variability associated with load, wind and solar for any particular hour. Figure H.2 shows the generation scheduled through the market with two confidence bands (90% and 95%) that represents the net ramping needs of the system looking out for the next hour.
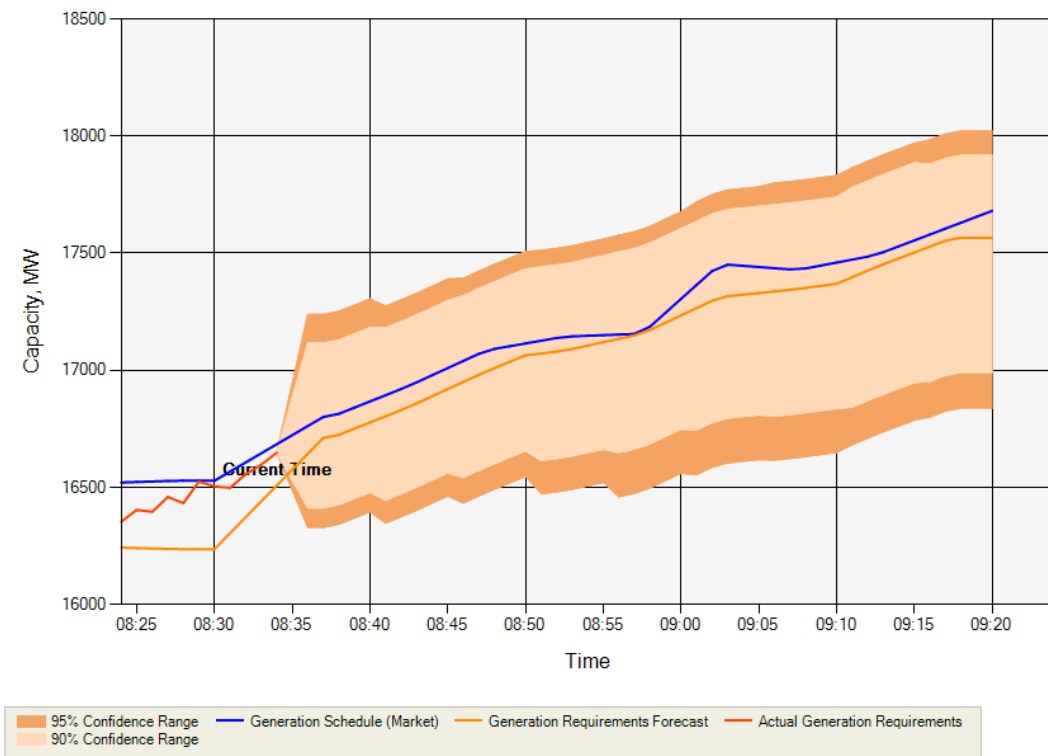

Figure H.2: Real-Time Net-Load Ramping Capacity Requirement

---

[17] Generation Requirements Forecast = Load Forecast - Wind Forecast - Solar Forecast - Interchange Schedule

Available Ramping Capability

As shown in Figure H.3, the operator has the option of overlaying the upward and downward generating capacity available for dispatch from conventional resources. The capacity availability shown is the actual ramping capability of the committed conventional resources and is based on a five-minute ramp rate and is calculated in each of the real-time dispatch (RTD) run. The look-ahead horizon is the RTD five-minute intervals for the advisory intervals. However, the operator can look ahead up to 24 hours. The available capacity is typically coned shaped because the available ramping capacity is calculated from the current time. For example, the capacity availability shown 20 minutes away from the "current time" is the actual upward and downward ramping capability of the committed conventional resources over a 20-minute period.

When this shaded area covers the *Generation Requirements Forecast* confidence bands, the committed resources have enough capacity and ramping capability to meet the expected wind/solar uncertainty and variability for the horizon displayed.
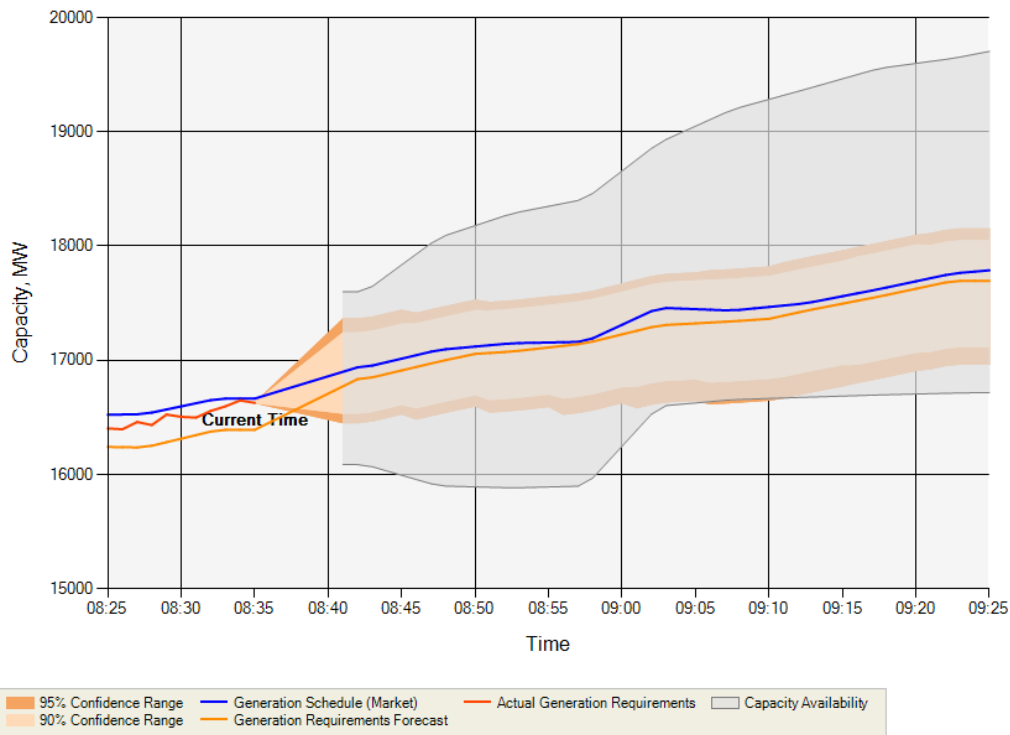

Figure H.3: Ramping Capacity availability as determined by RTD for the next hour

When any portion of the shaded area falls within the *Generation Requirements Forecast* confidence bands, the committed resources cannot meet the capacity and/or ramping capability to meet the expected wind/solar uncertainty and variability for the horizon displayed.

Working Group Membership

Name	Entity
Gerald Beckerle	Ameren
Dave Canter	American Electric Power
Jim Fletcher	American Electric Power
Richard Hydzik	Avista Corporation
Clyde Loutan	California Independent System Operator
Lorenzo Kristov	California Independent System Operator
Andrew Clarke	Duke Energy
John Hughes	Electricity Consumers Resource Council
Robert Entriken	Electric Power Research Institute
Aidan Tuohy	Electric Power Research Institute
Jens Boemer	Electric Power Research Institute
Jack Cashin	Electric Power Supply Association
Julia Matevosjana	Electric Reliability Council of Texas
Hassan Hamdar	Florida Reliability Coordinating Council
Jason McDowell	General Electric
Nicholas Miller	General Electric
Paul Turner	Georgia System Operations Corporation
Sasoon Assaturian	Independent Electricity System Operator
John Simonelli	Independent System Operator of New England
Mike McMuller	MISO
Barry Mather	National Renewable Energy Laboratory (NREL)
Paul McCurley	National Rural Electric Cooperative Association
Mark Ahlstrom	NextEra Energy
Robert Cummings	North American Electric Reliability Corporation
Michelle Marx	North American Electric Reliability Corporation
John Moura	North American Electric Reliability Corporation
Ryan Quint	North American Electric Reliability Corporation
Pooja Shah	North American Electric Reliability Corporation
Elliott Nethercutt	North American Electric Reliability Corporation
Levetra Pitts	North American Electric Reliability Corporation
Nicole Segal	North American Electric Relibaility Corporation
Thomas Coleman	North American Electric Relibaility Corporation
Lee Pedowicz	Northeast Power Coordinating Council
Quoc Le	Northeast Power Coordinating Council
Gary Keenan	Northwest Power Pool
Dariush Shirmohammadi	California Wind Energy Association
Troy Blalock	South Carolina Electric and Gas (SCANA)
Todd Lucas	Southern Company
Rodney O'Bryant	Southern Company
Cindy Hotchkiss	Southern Company
Thomas Siegrist	Stone, Mattheis, Xenopoulos & Brew, P.C.
Jagan Mandavilli	Texas Reliability Entity
Brian Evans-Mongeon	Utility Services
Anthony Jankowski	WE Energies
Layne Brown	Western Electricity Coordinating Council