California Energy System for the 21st Century

FINAL REPORT

Project	**Flexibility Metrics and Standards Grid Integration**
Title	***Role of Operating Flexibility in Planning Studies***
Project Team	
Pacific Gas and Electric Company	Antonio Alvarez, Will Dong, Ben Moradzadeh, Carl Nolen
San Diego Gas and Electric Company	Rob Anderson
Lawrence Livermore National Lab	Thomas Edmunds, John Grosh, Deepak Rajan
Astrape Consulting	Kevin Carden, Nick Wintermantel, Parth Patel , Alex Krasny
Electric Power Research Institute	Aidan Tuohy, Erik Ela, Eamonn Lannoye, Qin Wang
Date	September 12, 2017
Version Type	Final

.

Abstract

This project was conceived to examine the flexibility needs of the future California power grid. The analysis explores the need to establish generation planning metrics and standards that explicitly consider the operating flexibility needs of the system as the State pursues its aggressive renewable power generation goals. New methods and tools have been developed to use high resolution models of the grid that take into account uncertainties regarding renewable generation, load, equipment reliability, and economic growth. These models leverage high performance computational resources to fully explore the range of possible grid conditions that may lead to loss of load. The cost effectiveness of operating policies and hardware configurations that increase grid flexibility are examined with the tools to provide actionable information to grid planners and stakeholders.

Acknowledgement

The project team would like to thank the following Advisory Group members for their support and guidance throughout this project:

CPUC Energy Division (ED)
Maria Sotero
Patrick Young
Donald Brooks
David Miller
Forest Kaser

CPUC Office of Ratepayer Advocates (ORA)
Christopher Myers
Radu Ciupagea
Cindy Li

The California Independent System Operator (CAISO)
Shucheng Liu

The California Energy Commission (CEC)
Mike Jaske

The Utility Reform Network (TURN)
Kevin Woodruff

Southern California Edison Company (SCE)
Mark Nelson

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

1.1 Background

The California Energy Systems for the 21st Century (CES-21) Program is a collaborative research and development effort between the three California investor owned utilities – Pacific Gas and Electric Company (PG&E), Southern California Edison Company (SCE), and San Diego Gas and Electric Company (SDG&E) – and the Lawrence Livermore National Laboratory (LLNL). The objective of the CES-21 program, through two separate projects, is to explore the emerging challenges of cybersecurity and grid integration. The CES-21 program was approved by the California Public Utilities Commission (CPUC or Commission) on October 2, 2014 by Resolution E-4677.[1]

The Grid Integration Flexibility Metrics and Standards project ("Project") was conceived to examine the flexibility needs of the California Independent System Operator (CAISO) system, and to recommend, if appropriate, generation planning metrics and standards that explicitly consider the operating flexibility needs of the electric system. This report details the project's objectives, methods, results and recommendations, as well as requirements for the project managers set forth by the Commission.[2]

1.2 Project Requirements and Deliverables

In approving the CES-21 program, the Commission ordered specific requirements be met for successful completion of the project.

Table 1.1 below lists each requirement and demonstrates how the project delivered on these requirements.

Table 1.1 Project Requirements and Results

No.	Requirement	Delivered Results
1	Form a collaborative Advisory Group and meet at least once every six months to review and connect project results with relevant CPUC proceedings.	Formed an Advisory Group of CAISO, CEC, Energy Division (ED), ORA, SCE, and TURN. In total, the group held four meetings, and the project team provided several email updates.
2	Leverage learnings from PG&E's earlier collaborative review of planning model work.[3]	Based on the findings from the 2014 collaborative model review effort, selected the SERVM resource adequacy / production cost modeling tool to perform analysis for the project.
3	Present preliminary results and recommendations in a public workshop	Preliminary analysis and results were completed in 2015, and presented at a public workshop held on

[1] http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M116/K104/116104291.PDF
[2] *See*, Res. E-4677, OP 2-5.
[3] http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M116/K104/116104291.PDF

	using input assumptions from the 2014 Long-Term Procurement Plan (LTPP)[4]	1/6/2016.[5]
4	Demonstrate recommended metrics/standards in 2016 LTPP using at least one of the 2016 LTPP scenarios (Trajectory or expected scenario)	Final analysis was completed using 2016 LTPP assumptions. Project results and recommendations were presented to LTPP/IRP parties at a CPUC Integrated Resource Planning (IRP) proceeding on 8/15/2017. [6]
5	Provide 2016 LTPP parties opportunity to comment	Following the release of this CES-21 final report, the 2016 LTPP/IRP parties will be given the opportunity to provide written comments on the project's final results and recommendations.
6	Make database of detailed modeling input assumptions available	The entire set of input data used for the study will be made publicly available by the Energy Division.
7	Ensured ability of LTPP parties to license and use new or improved tools (if any)	Updated SERVM software is available for license by LTPP/IRP parties[7]
8	Offer one informal training session for Commission staff on new tools and models	The Project team held several calls during the project and met with CPUC staff on 8/16/2017 to provide training and updates on the SERVM tool and the CES-21 analytical framework.

1.3 Project Purpose

As a national energy leader, California has adopted aggressive goals to increase renewable generation to at least 50 percent of energy deliveries to customers by 2030, doubled efforts for cost-effective incremental energy efficiency, and invested in other alternatives such as transportation electrification. These efforts are key contributions to the wider goal of reducing GHG emissions economy-wide by 40 percent from 1990 levels by 2030. The electric grid needs to be operationally flexible to accommodate the diurnal patterns and hourly variability and forecast uncertainty of increased solar and wind generation needed to achieve the GHG emissions reductions. As a result, resource planners must gain a deeper understanding of the emerging flexibility needs of the system.

This need for deeper understanding is evident in recent LTPP proceedings. During the 2012 and 2014 LTPP proceedings, ED and CAISO staff facilitated a number of stakeholder workshops and working group meetings to discuss the flexibility needs of the CAISO system, with a particular focus on reliability. While significant progress was made through these discussions, a few important and challenging questions were not addressed fully, and provided an opportunity to be explored in this CES-21 research project.

[4] The project was implemented in two phases. Phase 1 was completed using 2014 LTPP assumptions with a simplified representation of the WECC. Phase 2 was completed using 2016 LTPP assumptions with a detailed representation of the WECC (see Technical Appendix for details). Unless otherwise noted, discussions in this report are based on results from Phase 2 of the study.
[5] http://www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=9281
[6] http://www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=6442454487
[7] www.astrape.com

Specifically, in an effort to enable resource planners to gain a deeper understanding of the emerging flexibility needs of the California electrical system, the project set out to answer the following questions:

1. **Reliability Impact – Did the range of projected CAISO system scenarios have sufficient capacity and operating flexibility to meet the 1 day in 10 years reliability standard in 2026?** Reliability is the primary binding constraint of all resource planning processes. In order to better understand the interaction between operating flexibility and reliability, the project examined a range of CAISO system scenarios (some with more flexibility, some with less flexibility), and then measured each scenario's results against a specific reliability standard.

2. **Other Impacts – How did operating flexibility, or the lack of it, impact costs and emissions (i.e., system operations)? What are the main drivers?** Given that flexibility is a multifaceted system characteristic that impacts system operations in different ways at different times of the year, this project was designed to analyze flexibility needs in a range of weather conditions, economic forecast scenarios and unit performance scenarios. Furthermore, the project sought to analyze and explore the relationship between different flexibility solutions and their effectiveness, including some system level solutions that had not been modeled in previous reliability and operating flexibility studies.

3. **New Standards – Are new planning standards needed to maintain operational flexibility; and if so, what would those standards be?** This is an explorative, research question that looks to the future needs of the planning community. Properly used, planning standards can provide for easily measured threshold tests, thus avoiding the need for detailed reliability studies. For example, instead of conducting the more time intensive Loss of Load studies, the Planning Reserve Margin (PRM) metric has often been used by planners as an estimate of a system's surplus or deficiency for peak capacity needs. It was thought that if the project were able to detect systematic flexibility deficiencies, it could potentially help quantify such flexibility needs with new easy-to-measure metrics and standards that could be used in planning, similar to how planners use PRM for peak capacity planning.

1.4 Project Scope

To support its stated objectives, the research project focused on the following areas.

First, the project focused on the ability of the generating system to provide adequate capacity at all times of the year while respecting generating unit constraints and considering forecast uncertainty. This is broader than typical resource adequacy analysis which generally assumes all available capacity can be used to serve load. However, the simulations were performed using a transportation model of the electric system, so more granular transmission reliability concerns were not addressed. For example, topics such as frequency response and voltage control were beyond the project scope.

Secondly, the project adopted a resource planning perspective and assumed the physical characteristics of the system – such as curtailment and net imports – could be fully accessed by system operations. In other words, the project did not take a position on policy issues, such as how much renewable curtailment is appropriate. Similarly, the project did not address some operational issues such as how much net import the system operator could actually rely upon, nor market design issues such as how much of the available physical flexibility would be economically provided to the system based on market compensations. Instead, the project simply assumed a range of clearly stated values, and focused on understanding their impact on resource planning.

Finally, the goal of the project was to provide directionally useful information, not precise results. Rather than drawing precise conclusions based upon static input assumptions, this project was designed to gain broader understanding through various sensitivities aimed at identifying key drivers and testing the magnitude of their impacts. The project's goal was to develop a robust framework and a set of insightful findings to help policy makers and other stakeholders to further explore and better understand the topic of operational flexibility.

1.4.1 Relationship with Other Resource Planning Studies / Analyses

Project results here should not be compared against those from any specific capacity expansion modeling exercises, as the assumptions developed for this analysis were only vetted to the level of providing directional information. Instead, insights drawn from this project can inform inputs into other resource planning analyses, and the analytical framework designed for this project can be used to examine particular scenarios in other modeling exercises.

1.4.2 Connection with Other Studies

Informed by the collective knowledge of the project team and the Advisory Group, the project built upon the latest knowledge in the resource planning space.[8] At the same time, this project was uniquely designed to answer its own objectives, which led to the detailed modeling of the entire WECC region, the inclusion of uncertainties, and the testing of the CAISO system under a unique set of system scenarios.

Although the project was focused on the operating flexibility needs and performance of the California electric grid, the project's analytical approach, as well as its results and recommendations, can inform other systems that have, or anticipate having, large amounts of renewable generation.

1.5 Summary of Findings and Recommendations

- Under the assumed resource mix studied, up to 50% RPS, the CAISO system has sufficient operating flexibility to meet demand in a reliable manner, subject to the assumption that the system operator can fully access the flexibility available including curtailments and net imports
- In terms of new planning standards, the CES-21 results suggest there is no need, at this time, to add additional flexibility-related standards for addressing reliability-related issues
- Planning Reserve Margin (PRM) is still a useful metric to assess adequacy, but the Effective Load Carrying Capability (ELCC) of all resources needs to be accurately calculated and used in the PRM calculation
- Sufficient load following capability must be carried in order to ensure intra-hour flexibility sufficiency – and there is a potential tradeoff between reliability and economics in calculating requirements
- Use of new Loss of Load Expectation (LOLE) metrics – $LOLE_{INTRA\text{-}HOUR}$ and $LOLE_{MULTI\text{-}HOUR}$ allow for greater understanding of the flexibility needs and resources. How these relate to $LOLE_{CAPACITY}$ needs to be further considered.

---

[8] For example, the CPUC's recent Effective Load Carrying Capabilities (ELCC) studies in the RA proceeding, CAISO's 2014 LTPP studies, and E3's 2016 WECC Flexibility Assessment

2 ANALYTICAL FRAMEWORK

2.1 Overview

In order to answer the key questions, this project needed to develop an analytical framework that can characterize the flexibility needs of a system and also capture its full impact on system operations.

For model inputs, the project need to consider a range of wind and solar profiles in order to reasonably cover generation patterns from renewable resources under different weather conditions. This is analogous to the need to simulate multiple forced outage patterns for conventional generators. The load corresponding to these weather patterns also had to be represented. To do this, the project leveraged Astrape Consulting's expertise to develop 35 sets of correlated wind, solar, and load hourly profiles based on historical weather patterns observed during 1980 – 2014.[9] Similarly, intra-hour volatility (e.g., forecast errors for wind, solar and load) also needed to be modeled in order to create a realistic representation of system conditions with high renewable penetration. Finally, other uncertainties such as economic load growth forecast errors and generation forced outages, commonly modeled in resource adequacy studies, are also included.

In selecting a modeling tool, the project needed to simulate system behavior at sub-hourly intervals over the entire year. This required an enhancement relative to previous resource adequacy tools that only focused on evaluating system needs during peak demand hours. The framework also needed to produce probabilistic results, a feature common in resource adequacy tools, in order to measure reliability. To provide statistically meaningful results in the presence of all of these sources of uncertainty, the test year would need to be simulated thousands of times. Hence, production cost modeling software with fast execution times was needed. With these features in mind, the Strategic Energy Risk Valuation Model (SERVM) modeling tool was selected for this project.

Finally, the analytical framework also needed metrics to capture the flexibility requirements and deficiencies of the system. Accordingly, the project developed new metrics to explicitly detect loss of load events due to the inability to meet multi-hour, or intra-hour ramping needs[10] rather than insufficient capacity. To help us understand the holistic impact of operational flexibility challenges, the framework also provided standard system performance measures such as production costs (including net market purchases), emissions, and renewable curtailment.

Table 2.1 below summarizes the overall analytical framework developed for the project.

[9] See Technical Appendix for details

[10] Specifically, loss of load expectation (LOLE) due to multi-hour or intra-hour events.

Table 2.1 Analytical Framework Used for the Study

Inputs	Model	Results
Load and Resource Assumptions Each study case is a 2026 projected CAISO system with detailed WECC representation	**Strategic Energy Risk Valuation Model (SERVM)** A hybrid resource adequacy and production cost model	**System Performance** Reliability (capacity / flexibility), Cost, and Environmental Impact
Uncertainties considered for each study case • 35 weather years (correlated profiles for load / wind / solar) • 5 economic load growth uncertainty levels • 25 (or more) resource outage draws • Forecast errors for load / wind / solar (intra-day and intra-hour) • 20 study cases / scenarios[11]	Number of simulation iterations: 35 * 5 * 25 * 20 = 87,500 full years (8,760 hours each at 5 minute intervals) of simulated system operations	Key metrics captured: • Loss of load expectation (LOLE) due to lack of capacity • LOLE due to lack of flexibility (new metrics) • Production variable costs • CO_2 emissions • Renewable curtailment

The SERVM software was selected based on its unique set of features as reported in a recent collaborative review of planning models performed in 2014[12]. The features that are essential to this project include the ability to:

- Represent planning and operating uncertainties;
- Simulate system conditions within the hour and across all hours of the year;
- Calculate traditional reliability metrics and the ability to incorporate new operational flexibility metrics;
- Model a wide range of scenarios and sensitivities and complete the analysis within time available; and
- Calculate various system performance metrics, such as production costs, renewable curtailment, and GHG emissions, which are useful to assess the desirability of planning standards

[11] The list of study cases are described in the Data and Study Cases section of the report
[12] Pacific Gas and Electric Company, et al., Collaborative Review of Planning Models, (April 2014), available at www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=6626

In addition, parties to CPUC proceedings are already familiar with the model since the Energy Division is using it to estimate the effective load carrying capacity (ELCC) of wind and solar generation in the RA proceeding. SERVM is readily available and can be licensed by any party.

2.2 Key Metrics

This section lists and describes the key metrics produced by SERVM. Because SERVM is a stochastic modeling tool, each metric represents the expected annual value from a specific study case. However, if desired, iteration specific results (down to hourly levels) can be extracted from SERVM by re-running the desired study case.[13]

2.2.1 Reliability Metrics

LOLE is the main reliability metric used in the study. This is a generally accepted metric used in planning to measure the expected number of loss of load events over a given time period. The most commonly used time frame for LOLE is the number of events in 10 years, and that is the measure used in this study.[14]

In the past, the LOLE metric is solely used to measure loss of load events caused by capacity inadequacy (i.e., lack of available capacity to meet load during an hour of peak demand). In this study, loss of load events are further disaggregated by the type of resource deficiencies that caused them. The SERVM software logic that performs this disaggregation is detailed in the Technical Appendix.

- $LOLE_{CAPACITY}$ (events / 10 years) – Loss of load expectation due to generic capacity inadequacy to meet peak load
- $LOLE_{INTRA\text{-}HOUR}$ (events / 10 years) – Loss of load expectation due to flexible capacity inadequacy to meet intra-hour net load volatility
- $LOLE_{MULTI\text{-}HOUR}$ (events / 10 years) – Loss of load expectation due to flexible capacity inadequacy to meet multi-hour net load ramp
- $LOLE_{TOTAL}$ (events / 10 years) – Loss of load expectation due to capacity inadequacy of any kind[15]

2.2.2 Other CAISO System Level Metrics

- Renewables Curtailment (GWh) – Expected aggregate annual curtailment [16]
- Emissions (MMT) – Expected aggregate annual emissions calculated as the sum of all emissions from CAISO resources (using resource specific emissions rates) and the

[13] This can be accomplished by turning on detailed reporting features and re-running the case. By default, iteration specific results at the monthly, daily, or hourly level are not recorded in order to speed up simulation run-time.
[14] See CPUC's "Production Cost Modeling Requirements" ruling for additional discussion on reliability metrics http://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M167/K501/167501732.PDF
[15] Total LOLE represents the number of days with events of any LOLE type, and does not necessarily equal to the summation of LOLEs by type (e.g., two types of LOLE events can occur on a given day and only counts as one occurrence under $LOLE_{TOTAL}$).
[16] In this project, SERVM's economic commitment and dispatch algorithm attempted to minimize curtailment subject to economic and reliability constraints. No separate curtailment penalty was included in inputs to further limit curtailment beyond the economic commitment solution.

sum of all emissions from net imports (using an hourly import emissions rate based on a proxy heat rate of 8,000 Btu/kWh)

- Production Cost ($ Billions) – Expected aggregate annual CAISO cost to operate the system, including costs incurred by internal resources and also net purchase costs from external, non-CAISO regions
- Total Cost ($ Billions) – this is the sum of the production cost defined above and the expected approximation for the cost of curtailment (which is calculated by multiplying the Renewable Curtailment GWh metric by an assumed curtailment replacement cost of $50 / MWh)[17]

2.3 LLNL's High Performance Computing (HPC) Environment

The design of this project required a very large number of simulations. Specifically, to model the many uncertainties and range of study cases (discussed in the following section), more than 87,500 years of CAISO system operations had to be simulated. Moreover, LOLE values of one day in 10 years of operation at five minute intervals corresponds to detecting one event in over a million time intervals. Although SERVM includes many algorithm features and heuristics to speed execution time, the large scale of this computational campaign suggests the need for high performance computing (HPC) resources.

HPC resources and the expertise to utilize them were available at Lawrence Livermore National Laboratory (LLNL). The computer systems at LLNL contain over a million individual microprocessor cores, which allowed simultaneous execution of thousands of SERVM models in parallel. This enabled completion of the computational campaign thousands of times faster than execution on a single computer.

In order to access this computing power, however, the project team had to first reconfigure the SERVM software so that it could be deployed in an HPC environment and jobs executed in parallel. The Astrape Consulting and LLNL team was able to develop the software infrastructure for massively parallel deployment of the SERVM code. This research effort resulted in a new capability for industry. Now that the research effort has been completed, the capability could also be duplicated using commercial cloud computing services.

Deploying SERVM on LLNL's HPC system resulted in significant efficiency gains. For example, the ability to access 1,000+ cores of CPU and to process SERVM simulations in parallel resulted in program execution speeds hundreds of times greater than that of desktop computers.[18] Even compared to a cluster of dozens of desktop computers, the ability to access HPC systems effectively reduced run time from days to hours.

[17] Similar to past LTPP analysis, the project did not replace any curtailed energy with additional RPS resources (which is what a capacity expansion model would do); instead, a $50/MWh value is used to approximate the replacement cost for any curtailed energy.

[18] For example, a typical modern laptop computer has between 4 – 8 cores of CPU, so an HPC environment of 1,000 cores will achieve an efficiency gain of 1,000 / 4 to 8 = 250 to 125 fold.

3 DATA AND STUDY CASES

3.1 Projected 2026 CAISO System

Similar to the approach taken in previous LTPP cycles, the project first developed a system model for the 2026 planning year. The model captured both the CAISO system, using the approved 2016 LTPP assumptions and other systems in the WECC, using the latest available TEPPC 2026 Common Case.[19]

Detailed modeling of internal CAISO transmission and sub-regions of the WECC is especially important for this project as it allows for detection of flexibility issues caused by intra-regional transfer limitations that may otherwise be masked.[20] The Technical Appendix to this report provides additional details and lists all the sub-regions modeled.

Overall, the project's general modeling approach follows the guidelines provided by the Commission's September 23, 2016 ruling, and the attachment to this ruling titled "Production Cost Modeling Requirements."

3.2 Modeled Uncertainties

As discussed in the framework section, fundamental to a reliability study that examines portfolios with large amounts of variable generation is the need to model uncertainties. On top of the deterministic 2026 representation of the CAISO/WECC system, the project injected the following uncertainties:[21]

- 35 wind, solar, and load profiles and hydro inputs that correlate with historical weather patterns from 1980 – 2014
- 5 levels of economic load growth forecast errors
- Forecast errors for load, wind and solar (both hourly and within the hour)
- Generation forced outage patterns

With each modeling draw, a specific combination of uncertainties and hence a unique projection of a 2026 system is selected.

3.3 Study Cases / Scenarios

Whereas the range of uncertainties allows us to examine how random events affect a given study case, a set of carefully chosen study cases allows us to explore reliability challenges faced by different scenarios.[22]

---

[19] 2016 LTPP approved scenarios and assumptions http://www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=11673

[20] For instance, a coarser representation of the WECC may not reveal ramping limitations between sub-regions.

[21] See Technical Appendix for details on how each uncertainty is developed, including data sources and methods

[22] In this report, study cases and scenarios are used interchangeably to represent different 2026 CAISO systems.

Here, the Project tested the performance of the CAISO grid under 20 different scenarios with different amounts and types of renewable resources, and different amounts of flexibility being available to the grid, to determine at what point, and under what conditions, operating flexibility could become a reliability issue, and to quantify cost and emission impacts associated with higher or lower levels of flexibility.

At a high level, the scenarios are created by varying and testing the three aspects of reliability discussed under the Metric sections and grouped as such:

- **Capacity adequacy**: by varying the amounts and type of renewable resources (cases BC_01, BC_02, and BC_03);
- **Flexibility adequacy (intra-hour)**: by varying the amounts of flexible reserves, also known as load following reserves (cases SC_02 through SC_07); and
- **Flexibility adequacy (multi-hour)**: by varying the amounts of system flexibility in terms of
    - Ramping capability available from existing fossil fleet; (cases SC_08 through SC_11)
    - Ramping capability available through managing CAISO's net imports (cases SC_12 through SC_14)
    - Export capability to CAISO's neighboring balancing areas (cases SC_15 through SC_17)

Table 3.1 below provides a high level summary of the scenarios used for the final analysis. High level description of each group of cases is provided below, with additional details in the Technical Appendix.

Table 3.1 List of Study Cases

Case #	Type of Case	RPS % by 2026	Load Following	System P_{MIN}	Interchange 3-Hr Ramp	Net Exports Limit
BC_01	PRM Base Cases	33%	5% of Load	LTPP Default	Unlimited	2,000 MW
BC_02		43%	7% of Load			
BC_03		50%	9% of Load			
SC_01	**Reference Case**	**50%**	**9% of Load**	**LTPP Default**	**Unlimited**	**2,000 MW**
SC_02	Load Following (% of Load)		5% of Load			
SC_03			7% of Load			
SC_04			11% of Load			
SC_05	Load Following (Net Load Observed)		95th Pct			
SC_06			99th Pct			
SC_07			100th Pct			
SC_08	System P_{MIN} (+/- MW)			(-4,000)		
SC_09				(-2,000)		
SC_10				(+2,000)		
SC_11				(+4,000)		
SC_12	Interchange 3-Hour Ramp Limit				3,000 MW	
SC_13					6,000 MW	
SC_14					9,000 MW	
SC_15	Net Exports Limit					3,500 MW
SC_16						5,000 MW
SC_17						8,000 MW

3.3.1 *Capacity Adequacy Cases (BC_01, BC_02, BC_03)*

The Planning Reserve Margin (PRM) base cases represent 2026 systems with different levels of RPS penetration ranging from 33% to 50% (i.e., wind, solar) and also behind the meter PV; they are otherwise identical in load and generation assumptions. Specifically, the 43% RPS case (BC_02) is the "Reference Case" Scenario 2: the Default Scenario with the mid-level additional achievable energy efficiency sensitivity, described in the May 17, 2016 Assigned Commissioner's Ruling Adopting Assumptions and Scenarios for Use in the California Independent System Operator's 2016-17 Transmission Planning Process.

3.3.2 *CES-21 Reference Study Case (SC_01)*

This case is identical to the 50% RPS base case (BC_03), except in this reference case, instead of adding generic conventional resources, 600 MW of Energy Efficiency was added to achieve the $LOLE_{CAPACITY}$ standard of 1 day in 10 years.[23] All other study cases (SC_02 through SC_17) are built upon this reference case.

3.3.3 *Intra-Hour Flexibility Adequacy Cases (SC_02 through SC_07)*

These cases modeled different amounts of load following reserves available to mitigate intra-hour variability and forecast uncertainty of customer demand, and wind and solar generation. Two different methods were deployed to set LF requirements: one as a % of load, the other based on the amount of net load variation observed in the previous 60 days.[24]

3.3.4 *Multi-Hour Flexibility Adequacy Cases (SC_08 through SC_17)*

SC-08 through SC-11 quantified the impact of higher and lower ramping capability being available from the existing fossil fleet by adjusting their P_{MIN} levels, making these cases more or less flexible than the reference case.

SC-12 through SC-14 imposed maximum 3-hour ramping limits varying from 3,000 MW to 9,000 MW to CAISO imports and exports, making these cases less flexible than the reference case, which had no such limit.

SC15 through SC-17 examined the effect of expanding the net export limits to CAISO neighboring balancing areas from 3,500 MW to 8,000 MW in any given hour, making these cases more flexible than the reference case.

3.3.5 *Additional Storage Sensitivity Cases*

Given the ongoing public interest in battery storage as a grid integration solution, the project tested two additional set of sensitivity cases to better understand storage's contribution in terms of:

1. Reliability contribution; and
2. Economic and curtailment benefits

[23] See Results section for analysis on the amount of capacity that is needed to reach the 1 day in 10 years standard for the PRM base cases.
[24] Net load is calculated as load net of wind and solar generation

To examine storage's reliability contributions, the project tested three cases by adding 3,000 MW, 6,000 MW, and 10,000 MW of 4 hour duration battery storage to the reference study case and measured the average capacity value (i.e., ELCC) for the entire class of 4-Hour battery storage.

For the economic and curtailment benefit runs, the project created four cases, each adding 1,000 MW of a different type of storage device – 2-hour, 4-hour, 6-hour, and 8-hour battery storage – to the reference case.

3.4 Access to Input Data

The SERVM model inputs used for the final analysis will be made publicly available by the Energy Division.

4 RESULTS

This section presents study case results along with some high level descriptions. Detailed interpretation and synthesis of results are captured in the next section on recommendations.

The results section is organized in the same three groupings of cases as presented earlier:

1. Capacity adequacy results
2. Intra-hour flexibility adequacy results
3. Multi-hour flexibility adequacy results

Table 4.1 below provides a summary of results for the 20 defined study cases. Key metrics shown in this table are defined in the Key Metrics subsection of this report.

Table 4.1 Summary of Results (CES-21 Study Cases)

Case #	Type of Case	Description	LOLE CAPACITY[25]	LOLE INTRA-HOUR	LOLE MULTI-HOUR	LOLE TOTAL	Curtailment		Emissions	Total Cost
			(Events / 10 Years)				(GWh)[26]	(%)[3]	(MMT)	($ Billion)[27]
BC_01	PRM Base Cases	33% RPS	1.0	0.1	0.0	1.0	242	0.2%	61	7.2
BC_02		43% RPS	1.0	0.1	0.0	1.0	2,652	2.1%	52	6.4
BC_03		50% RPS	1.0	0.1	0.1	1.0	6,129	4.9%	49	6.5
SC_01	**Study Case**		**1.0**	**0.1**	**0.1**	**1.0**	**6,466**	**5.2%**	**48**	**6.4**
SC_02	Load Following (% of Load)	5%	0.8	0.6	0.0	1.4	5,503	4.4%	47	6.1
SC_03		7%	0.9	0.1	0.0	0.9	5,961	4.8%	47	6.3
SC_04		11%	1.1	0.1	0.0	1.1	7,045	5.6%	49	6.7
SC_05	Load Following (NL Observed)	95 Pct	0.9	99.5	13.6	113.0	4,797	3.8%	46	5.9
SC_06		99 Pct	0.7	25.3	1.5	27.4	4,987	4.0%	46	6.0
SC_07		100 Pct	0.7	2.4	0.0	3.1	5,624	4.5%	47	6.2
SC_08	P_{MIN} (+/- MW)	(-4,000 MW)	1.0	0.2	0.1	1.2	3,751	3.0%	46	6.0
SC_09		(-2,000 MW)	1.0	0.1	0.1	1.0	4,802	3.8%	47	6.2
SC_10		(+2,000 MW)	0.9	0.1	0.1	0.9	9,940	8.0%	49	6.7
SC_11		(+4,000 MW)	1.0	0.1	0.0	1.0	15,447	12.4%	51	7.3
SC_12	Interchange 3-Hour Ramp Limit	3,000 MW	1.7	0.2	0.1	1.7	8,548	6.8%	49	8.5
SC_13		6,000 MW	0.9	0.1	0.0	0.9	6,835	5.5%	48	7.1
SC_14		9,000 MW	0.9	0.1	0.0	0.9	6,572	5.3%	48	6.7
SC_15	Net Exports	3,500 MW	1.0	0.1	0.0	1.0	5,259	4.2%	48	6.3
SC_16		5,000 MW	1.0	0.1	0.0	1.0	4,553	3.6%	47	6.3
SC_17		8,000 MW	1.1	0.1	0.0	1.1	4,113	3.3%	47	6.3

[25] See Framework section for definition of key metrics
[26] This study did not model resources needed to replace any curtailed energy in order to meet a given RPS %
[27] This includes the total system production cost (includes cost of net imports), plus an approximated cost of curtailment (by assuming a replace cost of $50 / MWh)

4.1 Capacity Results (PRM Cases)

As discussed in the Study Case section, the Planning Reserve Margin (PRM) cases tested three different systems, each carrying a different level of wind and solar generation with otherwise identical load and generation. **Table 4.2** below shows the generation portfolio by resource type for each of the three scenarios.

Table 4.2 Name Plate Capacity by Resource Type (Planning Reserve Margin Cases)

Resource Type (Name Plate MW)	33% RPS	43% RPS	50% RPS
Aggregated GHG Free Portfolio	**38,888**	**50,000**	**54,289**
Solar (IFM + BTM PV)[28]	13,075	23,897	27,495
IFM	8,035	12,764	16,362
BTM	5,040	11,133	11,133
Wind	6,027	6,317	7,008
Other Renewables[29]	4,522	4,522	4,522
Energy Efficiency (EE)[30]	4,491	4,491	4,491
Energy Storage	1,350	1,350	1,350
Demand Response	1,559	1,559	1,559
Hydro and PSH[31]	7,863	7,863	7,863
Conventional			
Fossil Resources (CAISO)	26,740	26,740	26,740
Imports	11,665	11,665	11,665

4.1.1 ELCC Results

As shown in **Table 4.2**, the 43% RPS scenario carries far more solar and wind resources than the 33% RPS case, by a combined 11,112 MW. This difference in name plate capacity, however, does not directly translate into difference in dependable capacity to mitigate loss of load events. As various other planning studies have shown, when it comes to reliability assessments, a meaningful comparison can only be made if resources are measured by their reliability contributions – not name plate capacity – via methods such as the Effective Load Carrying Capability (ELCC) calculation. This calibration is especially necessary at higher renewable levels and particularly important for non-dispatchable resources such as solar and wind, whose reliability contributions are significantly impacted by the particular portfolio mix, which affects the timing of the system reliability need (i.e., the hours of system peak net load).

In this project, with the exception of fossil resources and imports, ELCC calculation is performed for every resource type – including demand side resources such as Energy

[28] In front of the meter and behind the meter PV

[29] This includes geothermal and biomass resources, also includes certain small, non-dispatchable hydro resources

[30] Energy Efficiency values are based on IEPR Mid Base - Mid AAEE forecast (e.g., 1xAAEE)

[31] Pumped storage hydro

Efficiency (EE) and BTMPV – in order to capture any changes in reliability contribution as more renewables were added to the system.[32]

Figure 4.1 below shows the average ELCC value for the aggregated GHG free portfolio and selected resource types, for the three cases. Overall, the aggregate ELCC decreases as RPS levels rise, largely driven by the diminishing ELCC value of solar.[33] The effect on other resource types are also visible, though less dramatic. For example, there is a slight decrease in EE due to the shift in timing of the peak net load.[34] Conversely, there is a slight increase in the value of storage due to shortening of the duration of the system peak net load as more solar is added to the system, an effect further discussed in the storage sensitivities section.

[32] For this project, ELCCs are calculated relative to a generic fossil resource. First, each projected system is calibrated to a given reliability level by adding or removing generic fossil resources to achieve an $LOLE_{CAPACITY}$ of 1 day in 10 years. Then, for each resource type evaluated (e.g., wind), the entire portfolio of this specific resource type is removed from the system (e.g., 5,000 MW of wind). Following this, generic fossil resources were added to the system until an $LOLE_{CAPACITY}$ of 1 day in 10 years is regained (e.g., 1,000 MW of generic resource). The amount of generic capacity added divided by the name plate of all the specific resource type removed is the ELCC % shown here. Such an ELCC calculation implicitly considers the reliability needs of the system across all hours of the day, thus obviating the need to focus on specific hours and can accurately reflect system changes (e.g., when the system peak net load is pushed further into the evening)

[33] The average ELCC of solar resources did not decline as rapidly as expected between 43% and 50%. The primary reason is that the mix of solar resources between 33% and 40% RPS was heavily weighted toward BTMPV while between 43% and 50% RPS was heavily weighted toward fixed and tracking utility scale PV. The solar profiles for BTMPV reflect suboptimal orientation and tilt and thus provide limited output in late afternoon hours, while the utility-scale solar configurations are more optimized and show higher output in these hours. So while the net load peak was later in the day in the 50% RPS cases, the more optimized solar shapes partially offset the impact of the net load shift.

[34] The EE data is limited to a static 8,760 hourly profile published by the CEC, which was assumed as constant and used across all the weather years in this analysis. Depending on the EE programs, this assumption may have underestimated or overestimated EE's ELCC, and is an area that can benefit from future research.

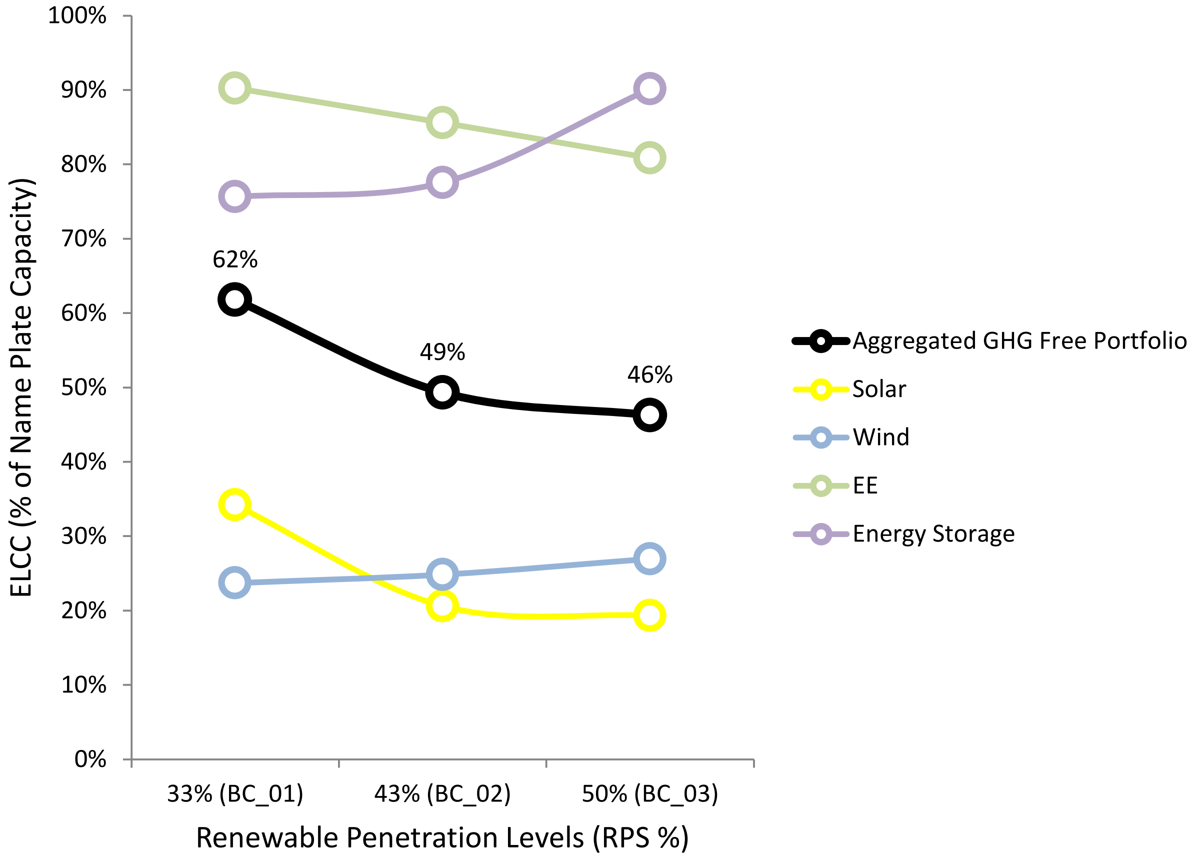


Figure 4.1 Average Effective Load Carrying Capability (ELCC) at Various RPS % Levels

4.1.2 Capacity Adequacy Results

For the PRM cases, the project team performed a capacity adequacy analysis for each of the three systems. **Table 4.3** below shows the amount of capacity that is needed for each system to reach the reliability standard of $LOLE_{CAPACITY}$ of 1 day in 10 years.[35]

Table 4.3 Reliability Results for As-Is PRM Base Cases

	33% RPS	43% RPS	50% RPS
Reliability Results ("as is")			
$LOLE_{CAPACITY}$ (days / 10 years)	2.9	1.9	1.4
Deficiency / (Surplus) to reach 1 day in 10 years standard			
Generic Resource Additions (MW)	1,348	730	393

As these results show, all three systems, as is, are less reliable than the standard.

4.1.3 Calculating Planning Reserve Margin

Having calibrated all three cases to a common reliability standard, the corresponding Planning Reserve Margins are calculated and shown in **Table 4.4** below.

[35] In all three scenarios, the project assumed the same Energy Efficiency level of 1xAAEE

Table 4.4 PRM Calculation – Method 1 (Treating all resources as supply side measured by ELCC)

Line #	PRM Calculation	33% RPS	43% RPS	50% RPS
	Demand			
1	Gross Consumption (MW)	54,727	54,727	54,727
	Supply			
2	Aggregate GHG Free Portfolio (excluding EE)	19,971	20,817	21,490
3	Fossil Resources	26,740	26,740	26,740
4	Imports	11,665	11,665	11,665
	Demand Side Resources Modeled as Supply			
5	Energy Efficiency	4,053	3,844	3,631
	Deficiency / (Surplus) to reach LOLE standard			
6	Generic Resource Additions	1,348	730	393
	PRM to satisfy LOLE standard (%)[36]	**116.5%**	**116.6%**	**116.8%**

It is critically important to understand that there are multiple methods to calculating PRM used throughout the electric industry. Each PRM is derived based on a specific method, and meaningful comparison between PRMs can only occur if the methods match; otherwise, comparisons are meaningless. For the PRMs shown above, two unique features define its method: 1) all resources are calculated with their ELCC (even demand side resources); 2) demand side resources are treated as supply resources and not netted out of the gross consumption data.

Using a different method, the resulting PRM that correspond to the same reliability standard would look significantly different. The following example illustrates that point. **Table 4.5** below shows the same three systems, except in how EE is treated in the PRM calculation. Under this method, EE is treated as a load modifier, where it is netted out of load based on EE's contribution at the time of system coincident, gross peak.

Table 4.5 PRM Calculation – Method 2 (Treating EE as load modifier)

Line #	PRM Calculation	33% RPS	43% RPS	50% RPS
	Demand			
1	Gross Consumption (MW)	54,727	54,727	54,727
2	Energy Efficiency	4,491	4,491	4,491
	Supply			
3	Aggregate GHG Free Portfolio (excluding EE)	19,971	20,817	21,490
4	Fossil Resources	26,740	26,740	26,740
5	Imports	11,665	11,665	11,665
	Deficiency / (Surplus) to reach LOLE standard			
6	Generic Resource Additions	1,348	730	393
	PRM to satisfy LOLE standard (%)[37]	**118.9%**	**119.3%**	**120.0%**

[36] Calculated as the ratio between a) sum of lines 2 through 6 and b) line 1

[37] Calculated as the ratio between a) sum of lines 3 through 6 and b) difference between lines 1 and 2

This set of PRMs is higher than the previous set for two reasons First, using EE's output at the gross peak load in this set assumes a higher reliability contribution for EE than it would actually provide during time of the net load peak (this effect is shown using an illustrative example in **Figure 4.2** below). Second, netting EE from gross demand in the PRM calculation essentially credits EE with the full PRM (i.e., PRM is calculated here by dividing capacity needed against the gross load net of EE), thus produces a higher PRM.

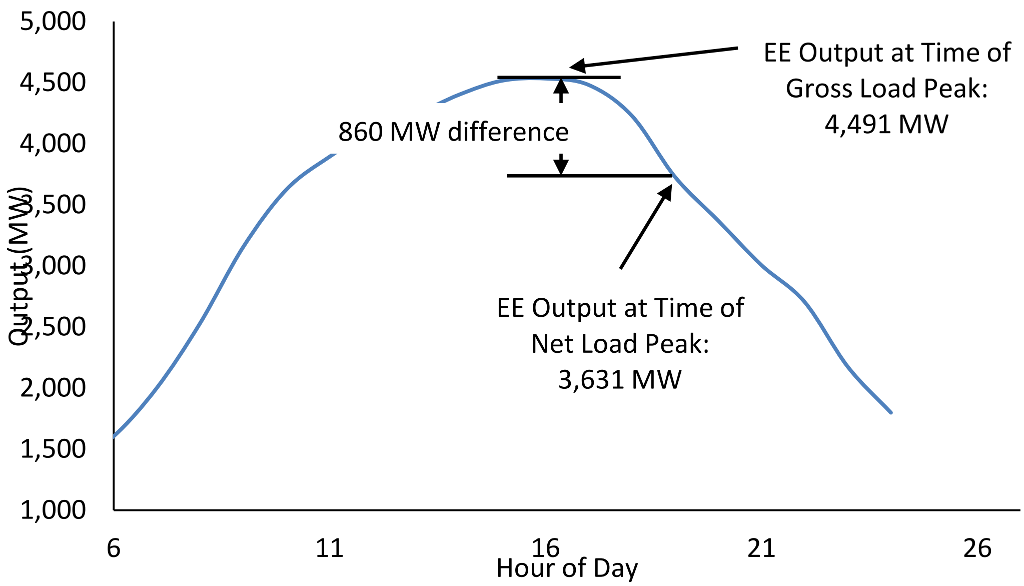


Figure 4.2 Energy Efficiency Output at Time of Gross and Net Load Peaks

However, these two calculations are done on the same system. Re-calculating the PRM using a different formula doesn't change the MW of generic resource additions needed to satisfy the LOLE standard.

These PRM tables show the importance of applying and using PRM correctly, a topic that is further discussed in the recommendations section.

Other than capacity adequacy, these PRM cases also needed different levels of LF to maintain the same $LOLE_{INTRA\text{-}HOUR}$.[38] Very little $LOLE_{MULTI\text{-}HOUR}$ events were detected in all three cases. Instead, system flexibility challenges showed up under economic metrics such as curtailment. Specifically, the 33%, 43%, 50% RPS cases resulted in annual curtailments of 0.2, 2.6, and 6.1 TWh respectively (corresponding to 0.2%, 2.1%, and 4.9% of annual output from all RPS eligible resources).

[38] The relationship between LF and $LOLE_{INTRA\text{-}HOUR}$ is presented in the Intra-Hour Flexibility results section, and the recommendations section. For the 33%, 43%, and 50% RPS PRM base cases, hourly load following requirements were set to 5%, 7%, and 9% of hourly load in order to maintain a similar level of $LOLE_{INTRA\text{-}HOUR}$ at roughly 0.1 events / 10 years.

4.2 Intra-Hour Flexibility Results (Load Following Cases)

As discussed in the study case section, the project tested seven different levels of load following reserves. **Figure 4.3** below shows the hourly amount of load following that is carried in each case. It illustrates the wide range of load following covered among these cases (e.g., the maximum hourly load following reserves difference between the 11% of hourly load and the NL 95th percentile cases is over 5,000 MW).

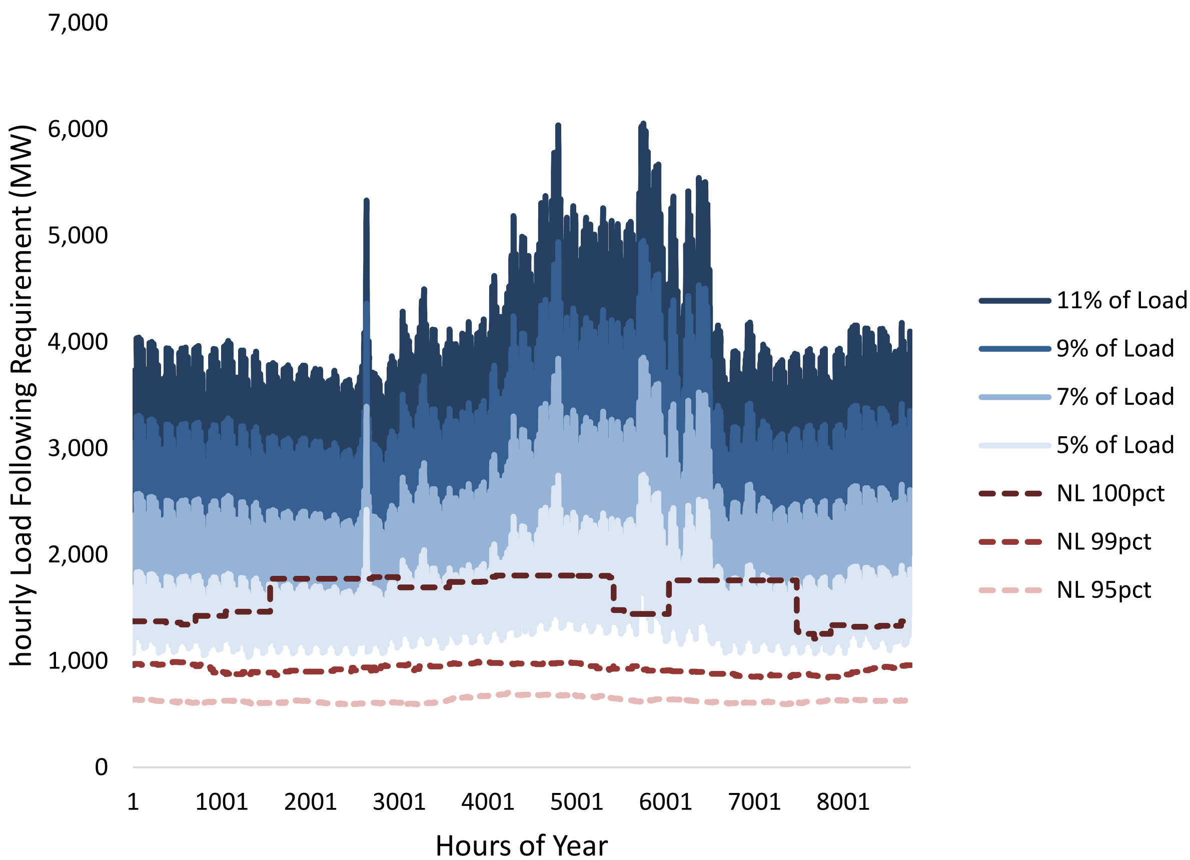


Figure 4.3 Hourly Load Following Requirements (Load Following Cases)

For each of these cases, reliability results were measured by the $LOLE_{INTRA\text{-}HOUR}$ metric. Results in **Table 4.6** shows a clear relationship between the amount of load following reserved and the number of $LOLE_{INTRA\text{-}HOUR}$ events detected.[39]

Table 4.6 Load Following Requirement vs. $LOLE_{INTRA\text{-}HOUR}$

Case #	LF Method	Description	Annual LF Amount (TWh)	$LOLE_{INTRA\text{-}HOUR}$ (Events / 10 Years)
SC_05	NL Observed	95 Pct	6	99.5
SC_06		99 Pct	8	25.3

[39] As described in the Study Case section, the Net Load (NL) observed method sets LF reserves based on the volatility observed in the previous 60 days. For example, the NL 100th percentile (case SC_07) uses the largest volatility observed in the previous 60 days. But even that does not eliminate $LOLE_{INTRA\text{-}HOUR}$, as what is observed on day 61 may be higher than any of those observed in prior days.

SC_07		100 Pct	14	2.4
SC_02	% of Gross Load	5%	14	0.6
SC_03		7%	19	0.1
SC_01		9%	25	0.1
SC_04		11%	31	0.1

While the timing of the addition of reserves was not optimized in any case, the difference in LOLE between cases SC_07 and SC_02 highlights the impact that timing has on results. Both cases supplied 14 TWh of annual load following, but $LOLE_{INTRA\text{-}HOUR}$ ranged from 0.6 to 2.4. Case SC_07 utilized a rolling 60-day window for setting reserve requirements. When large volatility events drop from the window, load following requirements drop, and the likelihood of events rises. But the addition of load following as a function of load (Cases SC_01-SC_04) is likewise not optimized for cost or reliability either.

It is worth noting that in the majority of the cases studied (all but the 95th percentile case, which carried far less reserves), $LOLE_{INTRA\text{-}HOUR}$ events occurred mostly during the low load, high renewable seasons, where less resources is committed to serve load yet a large amount of intra-hour volatility existed on the system due to large output from variable generations.[40]

Other than reliability, results from these cases show a converse trend in system cost. That is, while carrying additional reserves help mitigate $LOLE_{INTRA\text{-}HOUR}$, it comes at a higher cost as more resources are committed. For instance, the difference in total system costs between the cases with the least amount of load following reserves (SC_05) and the most (SC_04) was nearly $800 million a year.

The Recommendations section further interprets these results and provides a discussion on the level of load following reserves to be considered in planning studies.

4.3 Multi-Hour Flexibility Results

4.3.1 System P_{MIN} Cases

In these study cases, relative to the reference scenario, more or less flexible systems were created by decreasing or increasing the overall P_{MIN} level of the system.

Reliability Results

Figure 4.4 below shows the reliability results measured by the $LOLE_{MULTI\text{-}HOUR}$ metric for all the cases.[41]

[40] Also in these cases, a subtle relationship was observed between $LOLE_{INTRA\text{-}HOUR}$ and $LOLE_{CAPACITY}$, where because LF is set at a higher level, the increased use of energy limited resources and commitment of fossil generators for reserves resulted in slightly higher $LOLE_{CAPACITY}$

[41] Recall the $LOLE_{MULT\text{-}HOUR}$ metric detects any multi-hour ramping insufficiency, which is the renewable integration challenge most commonly illustrated by CAISO's "duck chart."

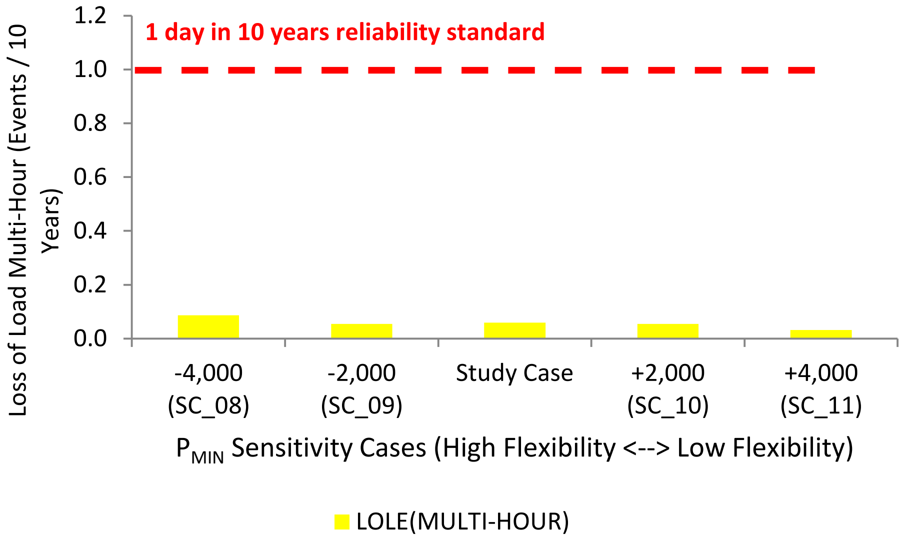


Figure 4.4 Impact of System P_{MIN} on Multi-Hour LOLE

These reliability results show very few $LOLE_{MULTI-HOUR}$ events were detected in all of the cases, indicating the system is physically able to manage the large ramping needs presented in a 50% RPS scenario even under highly challenging situations (e.g., the most inflexible +4,000 P_{MIN} case). Clearly, other sources of flexibility were available to help the system maintain balance. As it turns out, two of the primary drivers are curtailment and net imports.

Curtailment Results

Figure 4.5 below shows the level of curtailment by case. Results show a sharp increase in curtailment as system P_{MIN} is increased; and conversely, a drop in curtailment as system P_{MIN} is decreased.

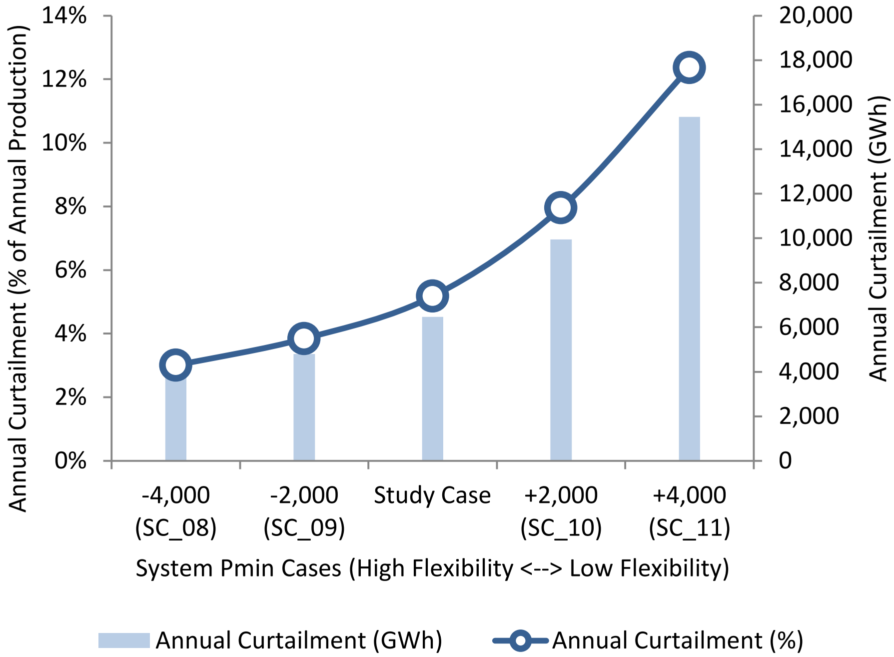


Figure 4.5 Impact of System P_{MIN} on Curtailment

These results revealed a relationship between two flexibility solutions: curtailment and system P_{MIN}. This relationship – incremental reduction in curtailment with reduction in system P_{MIN} – is shown in **Table 4.7** below. These results suggest the marginal curtailment benefit may be a function of the flexibility or inflexibility of the underlying system. In these cases, an additional MW of decrease in system P_{MIN} resulted in much higher curtailment benefit for an inflexible system than a flexible one. For instance, the marginal curtailment benefit of 1 MW of P_{MIN} reduction is 2.8 GWh at the highest level of P_{MIN} studied. This indicates that having lower P_{MIN} would benefit the system 2,800 hours per year when the system is at such a high P_{MIN} baseline. However at lower P_{MIN} baselines, the benefit is much lower. Between -4,000 MW and -2,000 MW P_{MIN}, the marginal curtailment reduction is only 0.5 GWh per MW of P_{MIN} reduction, suggesting only 500 hours per year of benefit.

Table 4.7 Curtailment Benefits from decreasing System P_{MIN}

Cases	Annual Curtailment (GWh)	Incremental Curtailment Reduction between cases (GWh)	Marginal Curtailment Reduction (GWh per incremental MW of P_{MIN} Reduction)
P_{MIN} +4,000 MW (SC_11)	15,447	**5,507**[42]	**2.8**
P_{MIN} +2,000 MW (SC_10)	9,940	**3,474**	**1.7**
Reference Case (SC_01)	6,466	**1,664**	**0.8**
P_{MIN} -2,000 MW	4,802	**1,051**	**0.5**

[42] Here, the incremental reduction between cases SC_10 and SC_11 is 15,447 – 9,940 = 5,507 GWh; and the marginal curtailment reduction is 5,507 GWh / 2,000 MW of P_{MIN} = 2.8

(SC_09)			
P_{MIN} -4,000 MW (SC_08)	3,751	**N/A**	N/A

Net Import Results

In addition to curtailment, net import is another source of flexibility available to the system. **Figure 4.6** below shows the hourly mileage for the +4,000 P_{MIN} case.[43] These results show that the the projected CAISO system is consistently using net import to help balance the daily morning and evening net load ramps, across all seasons of the year.

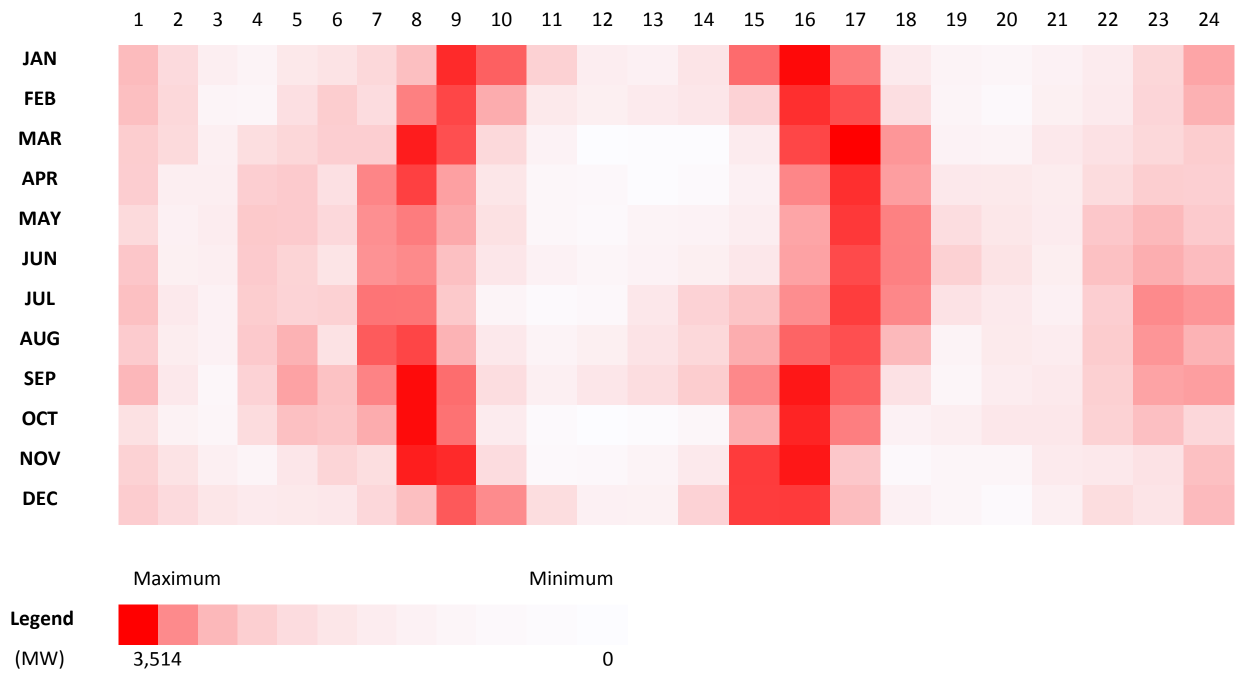


Figure 4.6 Average Hourly Net Import Mileage by Season and Hour (System P_{MIN} Cases)

Whereas curtailment and net import data reflects specific aspects of system operations, total production costs and emissions captures holistic, system level impacts. These results are shown in **Figure 4.7** below.

[43] Hourly mileage is calculated as the absolute hourly difference between CAISO net import levels

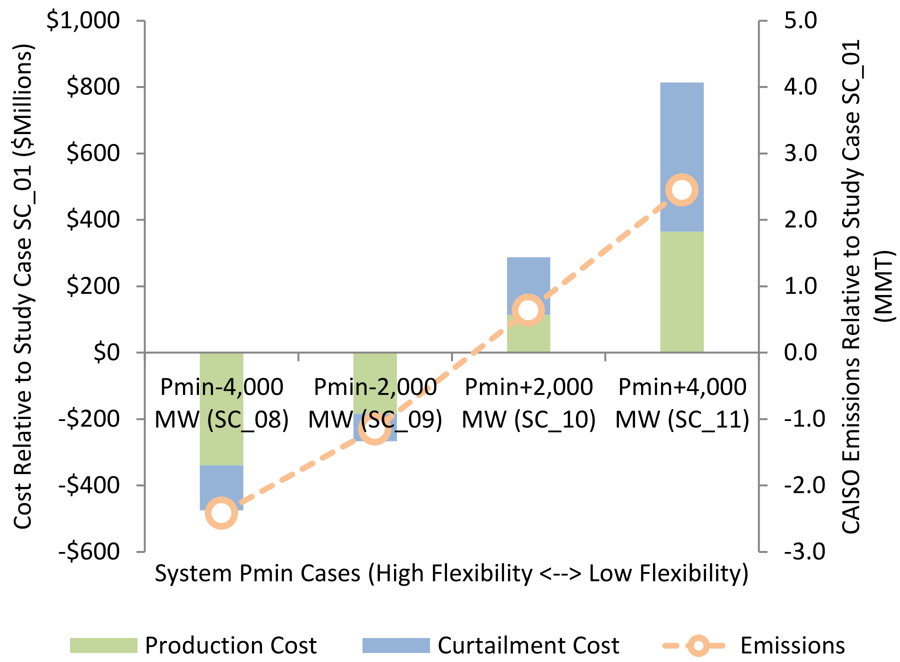


Figure 4.7 Impact of System P_{MIN} on Costs and Emissions

Here, the results confirm that in addition to more curtailments, inflexible systems also incur more production costs (due to inefficient dispatch and commitment of resources) and also produce more emissions.

4.3.2 Interchange 3-Hour Ramp Cases

Relative to the reference study case, these study cases limited the system's 3-hour ramping capability from net imports; thus, decreasing the flexibility of the system.
Similar to the P_{MIN} cases, these results showed a consistent trend between system flexibility, curtailment and system costs. Specifically, as system flexibility is decreased by further reducing the 3-hour ramping capability in each case, curtailments and production costs increased.

For these cases, results also showed the amount of 3-hour ramp from net imports that the CAISO could benefit from under a 50% RPS world. **Figure 4.8** below shows the largest one thousand instances of CAISO 3-hour ramp modeled for the projected 2026 year, for each of the study cases. Also shown in the chart, for comparison purposes, is the historical actual data for the years 2014 through 2016.[44]

As shown by the 9,000 MW and 6,000 MW study cases, these results indicate that under a 50% RPS scenario, the CAISO can benefit from having access to more than 4,000 MW of 3-Hour net import capability for approximately a thousand hours of a year. Restricting that access, as shown by the 3,000 MW study case, severely limits a source of flexibility that the CAISO is already and increasingly relying upon as more renewables are integrated onto the grid.

[44] Actuals are based on CAISO's daily Renewables Watch data

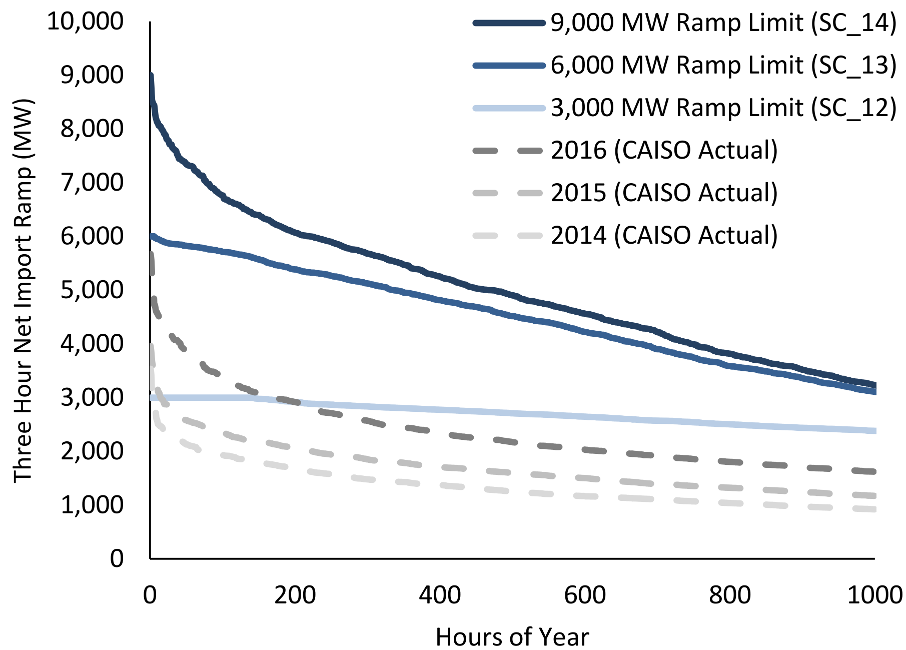


Figure 4.8 CAISO 3-Hour Net Import Ramp (Modeled Results vs. Historical Actuals)

4.3.3 Net Export Cases

Relative to the reference study case, the Net Export cases expanded the CAISO system's ability to export power in over-supply conditions, hence, increasing the flexibility of the system.

Similar to the P_{MIN} cases, results in **Figure 4.9** showed a clear relationship between flexibility and curtailment.

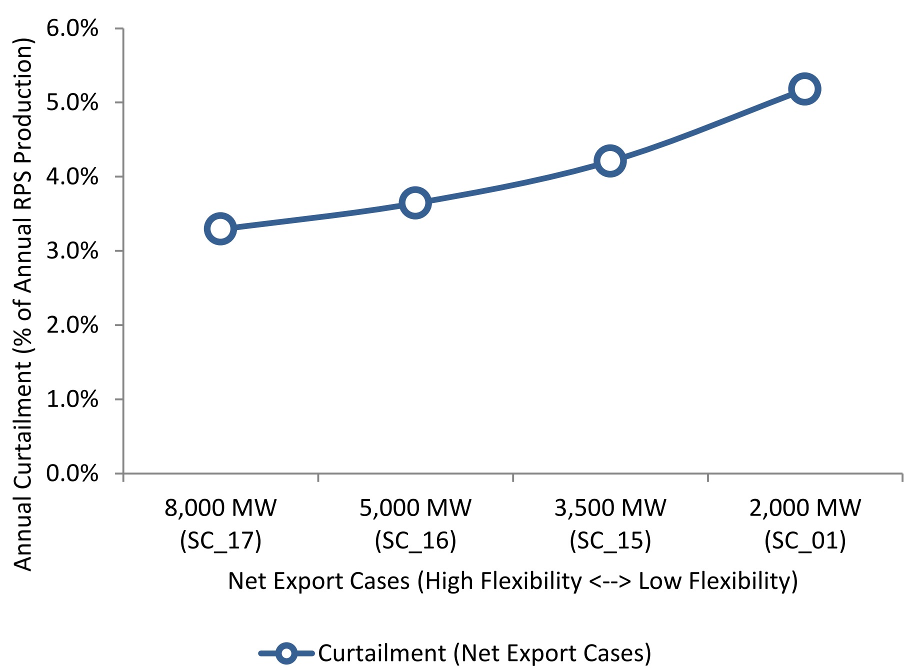


Figure 4.9 Impact of Net Export on Curtailment

Furthermore, **Table 4.8** below shows the magnitude of curtailment reduction as a function of the net export capability.

Table 4.8 Curtailment Benefits from Increasing Net Export

Net Export Cases	Annual Curtailment (GWh)	Incremental Curtailment Reduction between cases (GWh)	Marginal Curtailment Reduction (GWh per incremental MW of Net Export)
2,000 MW (SC_01)	6,470	**1,211**	**0.6**
3,500 MW (SC_15)	5,259	**706**	**0.4**
5,000 MW (SC_16)	4,553	**440**	**0.2**
8,000 MW (SC_17)	4,113	**N/A**	**N/A**

Similar to the P_{MIN} cases, these results indicate a diminishing gain in curtailment reduction as the system becomes more flexible (i.e., further expanding its net export capability). Part of this is due to the observation that the hours when the CAISO is experiencing extreme over-supply conditions at least partially coincides with similar situations in neighboring areas, thus limiting the CAISO's ability to export, regardless of modeled net export limit setting. Again, the marginal curtailment column indicates the utilization of the increased net export capability. Between the highest levels of net export capabilities, the marginal curtailment benefit is only 0.2 GWh per incremental MW of net export capability. This indicates the increased capability is only utilized approximately 200 hours per year.

4.3.4 *Additional Storage Sensitivities*

As described in the study case section, two sets of storage sensitivity cases were run.

Capacity Value

The first set of cases focused on understanding the capacity value of storage as more storage is added to the system.

Figure 4.10 below shows the average ELCC value for the entire class of 4-Hour storage as the project team added 3,000 MW, 6,000 MW and 10,000 MW of 4-Hour storage to the CES-21 Reference Study Case (SC_01).

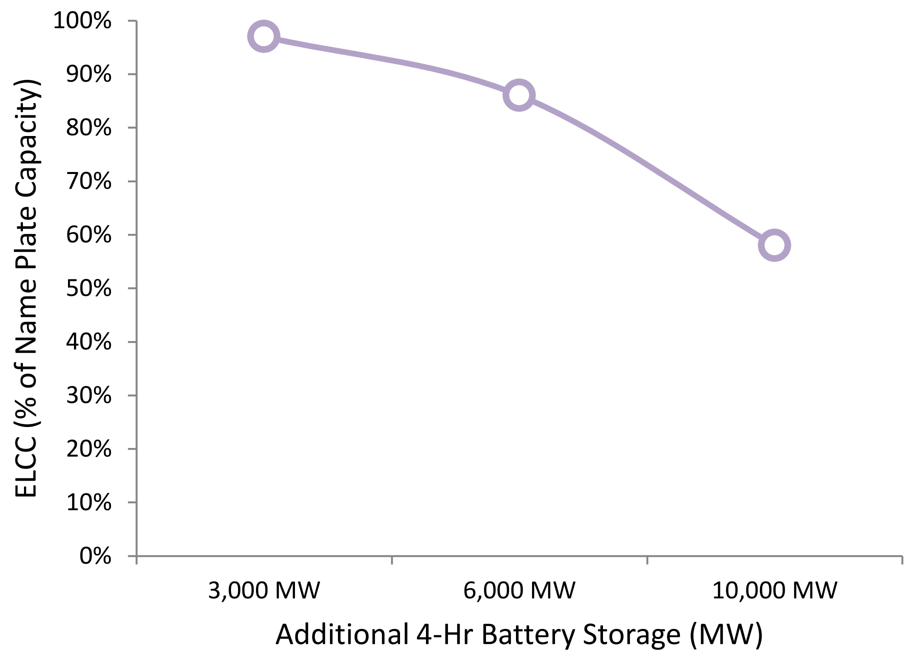


Figure 4.10 Average ELCC vs. Amount of Storage

These ELCC results indicate that adding 3,000 MWM of 4-Hour storage does not affect storage's capacity value. However, as more storage is added to the system, ELCC value for the entire class of storage resources decreases, and in the case of adding 10,000 MW of storage, this class average ELCC can drop below 60%. This is mainly driven by changes in the shape of the peak net load. Storage charges during off-peak hours in order to discharges during peak hours, flattens the peak net load in the process. These results show that if 10,000 MW of storage is added to a 50% RPS system, they will flatten the peak net load so much that a 4-hour battery device can only cover a portion of the peak.

Economic and Curtailment Benefits

The second set of storage cases focuses on understanding the economic and curtailment benefits from storage products of different durations. The economic and curtailment benefits of adding 1,000 MW of storage devices (of various durations) to the 50% RPS reference case are shown below in **Table 4.9** and **Table 4.10**, respectively. These results show a similar trend: for the studied 50% RPS reference system, the marginal economic and curtailment benefit is highest for a shorter duration storage device.

These cases presume that the storage resources can be used to serve ancillary services which do not require significant shifts in the real-time storage level of the resource. This means that a 2-hour storage product can likely serve ancillary service requirements all 24 hours of the day. In comparison, these results show the incremental load-shifting value of increasing the storage capability to 4-hours per day is relatively small (as shown by the more limited opportunity to provide curtailment benefits)

Table 4.9 Economic Benefits of Energy Storage

Cases	Total Benefit ($/kW-year) [45]	Incremental Benefit between cases ($/kW-year) [46]	Marginal Benefit ($/kW-year per incremental hour of storage capacity) [47]
Add 1,000 MW of 2 HR Storage	48	48	**24**
Add 1,000 MW of 4 HR Storage	80	32	**16**
Add 1,000 MW of 6 HR Storage	96	16	**8**
Add 1,000 MW of 8 HR Storage	100	4	**2**

Table 4.10 Curtailment Benefits of Energy Storage

Cases	Total Benefit (MWh of Curtailment Reduction / MW of Storage)	Incremental Benefit between cases ((MWh of Curtailment Reduction / MW of Storage)	Marginal Benefit (MWh of Curtailment / MW of Storage per incremental hour of storage capacity)
Add 1,000 MW of 2 HR Storage	266	266	**133**
Add 1,000 MW of 4 HR Storage	472	206	**103**
Add 1,000 MW of 6 HR Storage	588	116	**58**
Add 1,000 MW of 8 HR Storage	601	12	**6**

5 FINDINGS AND RECOMMENDATIONS

5.1 Overview

As discussed earlier in the report the main goal of this project was to determine whether there is a need to revise planning standards to reflect the changing conditions of the electric grid. Due to high levels of variable energy resources, there are concerns about whether there is sufficient operational

[45] Includes CAISO production cost benefits, net purchase cost benefits, and the economic scarcity rent; does not capture any resource costs.
[46] Represents the incremental benefit from an additional two hour chunk of storage capacity (e.g., the incremental benefit going from a 2-Hour to a 4-Hour storage device is 80 – 48 = 32 ($/kW-year)
[47] Taking the incremental value and dividing it by the two hour block of storage capacity

flexibility in the system to manage the increased variability and uncertainty associated with wind and solar power. A well-designed standard is expected to provide a relatively easy to calculate metric that shows whether a system has a sufficient resource mix to reliably meet load, within certain tolerances. While the relative economics are considered when setting criteria for the standard (e.g. whether to set desired reserve margin at 15% or 20%), the calculation on how resources contribute to reliability does not consider the economics of doing so. The economics related to operating the system, and other operational practices, are not usually robustly considered when assessing the system's ability to meet reliability standards since dispatch decisions to ensure reliability will trump normal operational practices or economic concerns. However, this ignores any potential interaction between economic concerns, operational concerns, and reliability. In many conventional systems with predominantly dispatchable resources this approach may be reasonable, but the significant projected changes to the resource mix in California compel a realistic simulation of system operations to determine whether standards should consider operational flexibility explicitly.

Resource adequacy has traditionally been assessed by calculating the risk of not meeting demand, using metrics such as Loss of Load Expectation (LOLE), which determines the expected number of intervals during a given time horizon in which load will not be met. While this provides a detailed risk-based calculation, this method can be very time consuming and is not easily done especially if multiple load serving entities are making simultaneous planning decisions that impact the overall reliability of an electrical system. To address this other methods have developed around resource needs, with planning reserve margin (PRM) as a common approach. This is calculated as the additional capacity above expected peak demand divided by peak demand, and is required to cover the uncertainty in peak demand forecasting and generator availability in the future year. This can be done using a simple, transparent calculation. Calculated properly, PRM allows for easy comparison across different candidate portfolios, while also being easier to allocate procurement responsibility across different entities, such as the various load serving entities in California. This allocation would be challenging using a LOLE-type approach, so planners need to have a simpler metric to use for this. As computational power increases, solution algorithms improve and data availability improves, even more detailed studies are now possible, as were carried out using the SERVM tool in this project, where LOLE can be calculated explicitly while considering operational flexibility issues. Therefore, when and how PRM can still provide value, especially in light of increased renewable penetration, was a key consideration in this project.

The SERVM results described earlier clearly show that the assumed resource mix studied, up to 50% RPS, has sufficient capacity and flexibility to meet demand in a reliable manner. This finding is subject to several important assumptions as discussed later, including the assumption that operating practices, represented in detail here, will be adjusted to reflect the increased uncertainty in the system at higher RPS penetrations. In terms of planning standards, this suggests there is no need to add additional flexibility-related standards for addressing reliability-related issues. This is not to say that additional metrics cannot be useful indicators, or that economic or market related issues may not result in the need for new metrics; it also does not mean that planning processes used in the past always guarantee sufficient flexibility. Indeed, the introduction of $LOLE_{INTRA\text{-}HOUR}$ and $LOLE_{MULTI\text{-}HOUR}$ show that operational assumptions can affect the calculation of typical reliability metrics and thus there is a need to better consider flexibility issues in reliability studies. However, the study did demonstrate how the continued use of the PRM requires robust calculations of the Equivalent Load Carrying Capability (ELCC) of resources to indicate a reliable system.

Results also investigated the main drivers of the overall reliability of the system, as discussed in the various sensitivities shown. These showed that minimum stable levels of dispatchable generation can have a significant impact on results, while assumptions about how much flexibility can be obtained from the rest of the interconnection can also have a significant impact on the ability to meet load and manage variability and uncertainty in California. The ability to curtail renewable resources was shown to be crucial, while the assumed load following requirements, which drive commitment of generation, can be very important.

These results are mainly focused on the long term needs of the system. However, similar concerns will occur when looking several months to years in advance, as shown by the recently developed Flexible Resource Adequacy Criteria Must Offer Obligation (FRAC-MOO) construct in the CAISO market.[48] Investigating the need for such a construct was somewhat outside the scope of this project, as FRAC-MOO is not just focused on whether the resource mix can provide sufficient reliability, but also on procuring those resources and ensuring they offer flexibility into the CAISO market. However, results here do show that, if the market can access the flexibility available, there is no shortfall in ability to provide flexibility up to 50% renewable penetration.

In general, it was shown that existing planning standards can still ensure reliability, assuming the relevant components are calculated sufficiently. The project's analysis does show a need to consider intra-hour and multi-hour ramps more explicitly in long term planning, but the Planning Reserve Margin techniques developed and used previously can still provide useful indicators of resource adequacy.

5.2 Capacity Adequacy

In the past, PRM was calculated based on adding up nameplate capacity of the dispatchable generation on the system plus the dependable output of variable energy resources, and ensuring they have a specific margin over the expected demand. This study showed that for PRM to continue to serve as a reasonable reliability standard, the process for calculating PRM will need to account for the Effective Load Carrying Capability (ELCC) of all resources with any dispatchability constraint. Therefore, it is recommended that ELCC is calculated based on studies like the ones completed in this project, and revisited when significant changes to the resource mix occur. Results can then be used to inform the PRM, which allows for quick comparison of different plant portfolios. Results would need to be revisited if the portfolio of wind and solar resources changes sufficiently from the portfolio assumptions used in the calculations; this may also be true for demand side resources and energy storage. If the ELCC is calculated as such, then this project found a PRM of 17% appears to provide sufficient reliability to meet the LOLE standard as described above (subject to operational assumptions described later). This study did not explicitly look at how much the wind and solar resource penetration (or any other resource mix changes) would need to change before revisiting the ELCC calculations. Given the rate of change in California's power system, a two to three-year cycle for calculating ELCC contributions of new resources seems reasonable.

In the IRP process, PRM could therefore still be used as a metric to assess resource adequacy, with the ELCC being based on outcomes of detailed studies. The results section shows the results calculated for this system including sensitivities on issues like energy storage; this type of analysis would need to be repeated for any future analysis with different underlying system assumptions. There may also be value in the IRP process to determine an ELCC for both existing resources and marginal or new resources since the IRP process is likely to be looking for the best resource to add to

[48] https://www.caiso.com/informed/Pages/StakeholderProcesses/FlexibleResourceAdequacyCriteria-MustOfferObligations.aspx

the existing system and these ELCC values can vary based on the underlying system and assumptions about operations.

Recommendation #1: PRM is still a useful metric to assess adequacy, but the ELCC of all resources needs to be accurately calculated and used in the calculation.

5.3 Intra-Hour Ramping

As shown in the results, assumptions about the amount of load following reserves required can have a significant impact on the likelihood of having sufficient flexibility available to meet intra-hour load changes. As such, load following requirements need to be carefully understood for future studies. Traditionally, sufficient load following was made available through the day ahead hourly market, and then economic dispatch at time resolutions closer to real time (fifteen-minute market and real time dispatch). Aspects like adding a look ahead in the dispatch and forecasting wind and solar in the real time and fifteen-minute markets also helps increase the amount of available capacity. However, with increasing levels of renewable penetration, explicit load following is often carried in planning studies focused on renewable integration.

With increasing renewable penetration, **Figure 5.1** below shows that there is a significant increase in ramping, particularly when moving from 33% to 43% penetration. Here, the solid lines show the absolute ramping on a daily basis for four different time series – load only, and net load for the three cases. This is calculated by adding up the absolute value of 5 minute ramps in a given day (e.g. if the net load ramped up 20 MW in one period and down 5 MW in the next, the absolute ramping mileage would be 25 MW). The numbers here are less important than the relative changes, as absolute ramping does not impact operations, but does show how much overall additional ramping is required. As shown, renewables add ramping throughout the year. The dotted lines show cumulative ramping – compared to load only ramping, 33% renewables increases ramping by 172%, 43% renewables increases it by 237%, and 50% renewables increases ramping by 259%.

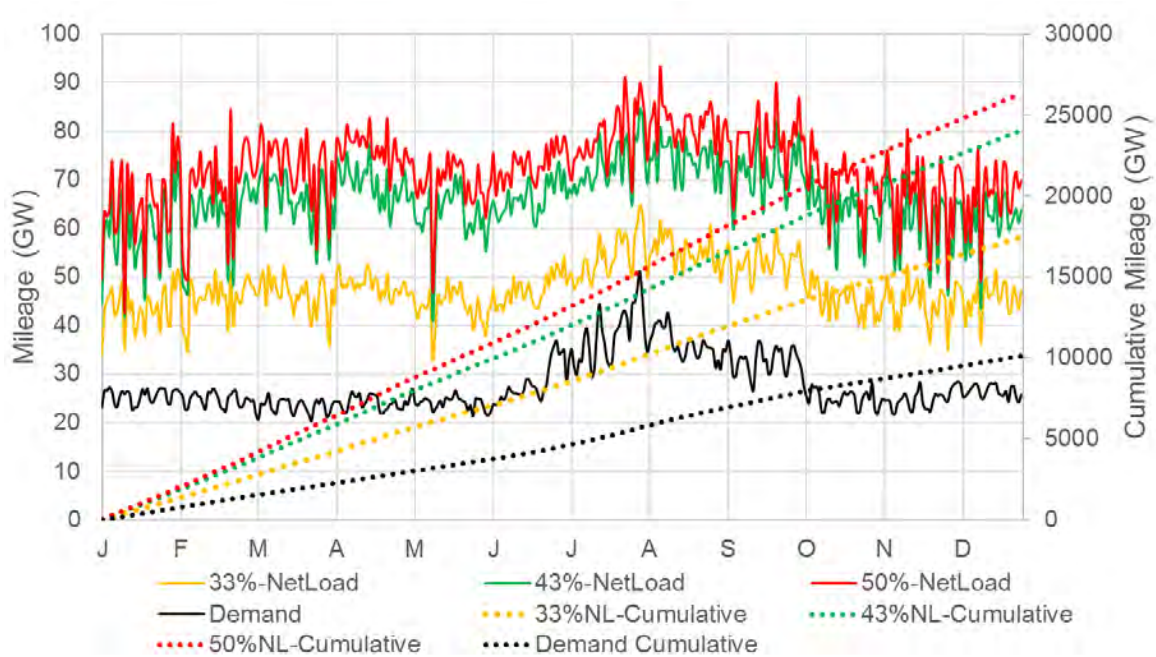


Figure 5.1 Daily and Cumulative Absolute Ramping Mileage for Different RPS Cases

This increase in ramping would be expected to put increased strain on intra-hour load following, and a noteworthy outcome of the studies performed here was that the $LOLE_{INTRA\text{-}HOUR}$ metric was shown to be heavily influenced by the amount of load following carried. This type of LOLE hasn't traditionally been included in reliability studies in the past – as such, it can be thought of as a somewhat new metric, or at least subset of existing metrics, that have been introduced here. The purpose of this, as well as the $LOLE_{MULTI\text{-}HOUR}$ is to look at the ability of the system to meet intra-hour (or multi-hour) ramps. Here, it was decided to continue to use a 1 day in 10 years total LOLE as the standard; however, there may be a need to further consider what the appropriate requirements for the new LOLE standard, or at least these new indicators, should be. Hence, when presenting results, the project team has tended to show both capacity and intra-hour/multi-hour ramping, and then ensure that total LOLE is less than 1 day in 10 years. One could also consider separating out these metrics from $LOLE_{CAPACITY}$, where that metric is still focused on the 1 day in 10 years standard, whereas the new metrics may be assessed against a different standard. Due to the close relationship between $LOLE_{INTRA\text{-}HOUR}$ and operational decisions, this metric can be significantly altered by changing operational decisions, and thus may not be as important as a planning reliability standard. For the purposes of this study, though, it was decided to use the total LOLE as the standard – the assumption being that there is a need to ensure reliability metrics can be met using the planned resource mix.

In actual operations, the operator may decide on some relaxation of the intra-hour flexibility requirements as a trade-off between small Area Control Error deviations (while still maintaining NERC and WECC standards) and lower costs. For example, the Net Load Observed method used here, while showing a very large $LOLE_{INTRA\text{-}HOUR}$ in comparison to the percentage of load-based load following requirement, is still relatively low in terms of how it compares with NERC and WECC operational standards, where frequency deviation due to small supply-demand imbalances is allowed on a relatively frequent basis. Here, the project team wanted to determine whether the system could meet all variability and uncertainty as well as capacity requirements within the 1-day-in-10 standard. In operations, it will be up to the policy of system operators to determine the appropriate requirements. A final point to note on the introduction of these new indicators is that the current baseline is not known; for example, it may be that while the $LOLE_{CAPACITY}$ of the current system is significantly lower than 1 day in 10 years, when the other LOLE indicators are also included, the current system may already be significantly greater than the 1-in-10 standard. The project team has taken the conservative assumption here that, even with new ways to consider loss of load, the system should still be planned for a 0.1 total LOLE.

The results of the study related to Load Following are shown in **Figure 5.2** below for the 50% penetration cases. The blue dots, and right hand axis show the $LOLE_{INTRA\text{-}HOUR}$, while the orange dots and left hand axis show the costs in billions of dollars. Based on the study results shown in **Table 4.6**, at 50% penetration, calculating load following using the NL Observed method – where short term variability of wind, solar and load variability is considered in calculating the required amounts – is insufficient to ensure that intra-hour variability and uncertainty can be met. Using a percentage of load, and significantly increasing beyond current requirements to 9% of load, can ensure sufficient intra-hour ramp capability is always available. Clearly, there is a cost to this, which needs to be understood more before it comes to operating the system, but these results at least show the future resources on the system (including interchanges) can be operated to nearly always meet load, as well as ramps in load (or net load) levels. The system operators may determine a more optimal method to determine reserve requirements, but results here show the potential cost and reliability implications of varying this target reserve level. Based on the figure, it would appear that there is a "sweet" spot for the modeled system where the $LOLE_{INTRA\text{-}HOUR}$ is sufficiently low, but costs are still kept relatively low, with 25 GWh, corresponding to 9% of load, the amount chosen for the CES-21 reference study case.

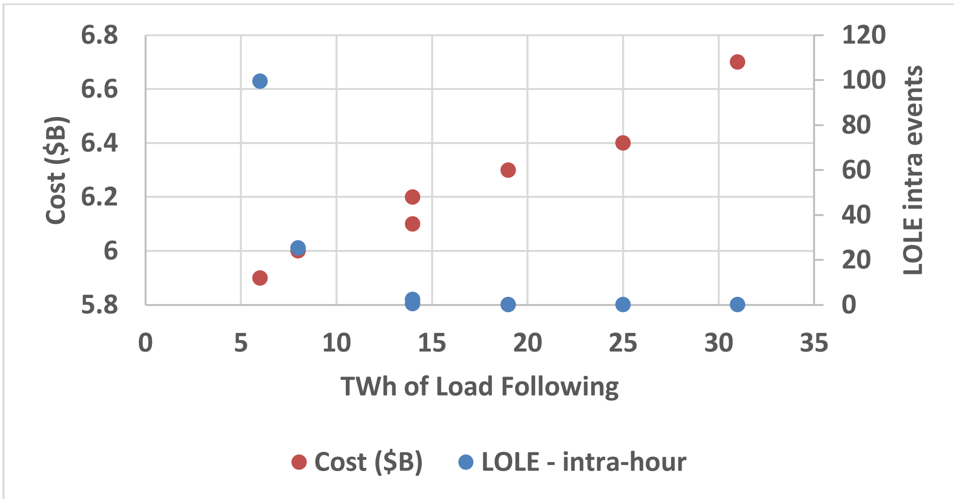


Figure 5.2 Cost and Load Following Impacts of Different Load Following Levels

For example, moving from 14 GWh (corresponding to 5% of load) to 25 GWh (9% of load) increases system operating costs by approximately $300m, but also reduces the intra-hour shortfalls and helps bring total Loss of Load Expectation under 1 event in 10 years. In comparison, using the 99th percentile of net load ramping observed in previous two months increases the intra-hour shortfalls to 25 events in 10 years, but saves an additional $100m beyond the 5% of load case. 25 events in 10 years appears to be a lot; however, the duration and magnitude may be short and small enough respectively to be acceptable as operators determine the actual requirements. As stated earlier, the conservative approach is taken here to ensure the system can meet load at (practically) all times. Another approach would be to study the LOLE of the current system, including intra-hour shortages, and then assume that the future systems being studied will have the same LOLE. Depending on what the actual LOLE may be, this could be either a more conservative or optimistic approach.

Recommendation #2: Sufficient load following capability must be carried in order to ensure intra-hour flexibility sufficiency – and there is a potential tradeoff between reliability and economics in calculating requirements

Recommendation #3: Use of new metrics – $LOLE_{INTRA-HOUR}$ and $LOLE_{MULTI-HOUR}$ allow for greater understanding of the flexibility needs and resources. How these relate to $LOLE_{CAPACITY}$ needs to be further considered.

5.4 Multi-Hour Ramping and Flexibility Options

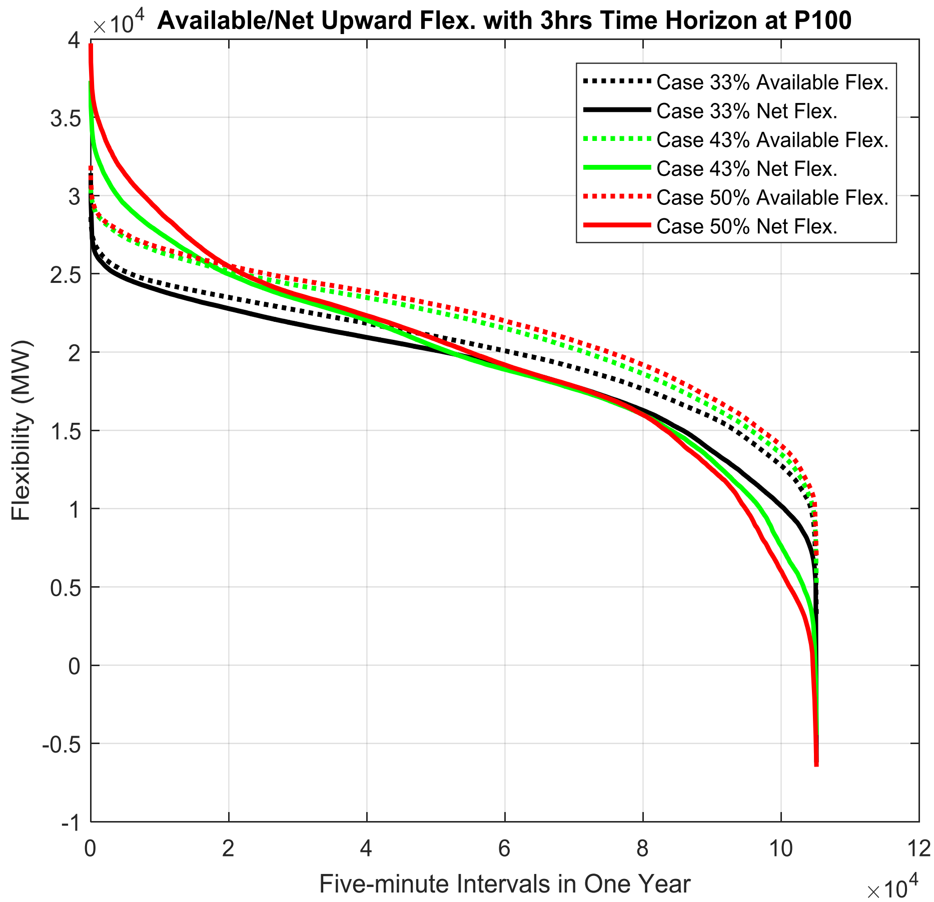
As shown in the results, multi-hour ramping constraints are not as frequently binding as capacity or intra-hour ramping constraints. It was shown that, if sufficient capacity exists, then intra-hour ramping is more frequently a binding constraint in terms of meeting load. No reliability deficiencies were found in most cases because the study assumed imports and exports can vary hourly, and renewable generation can be economically dispatched[49] or curtailed, to make up the multi-hourly ramping up/down of net load. It was also shown that, in general, reduced multi-hour ramping flexibility has little or no negative reliability impact but a significant cost impact. Reducing flexibility from minimum generation levels, interchange ramping or export limits all have cost and emissions implications. Therefore, careful study of these issues is warranted in any future activities, particularly those related to economic build-out of the resource mix.

Longer duration ramps, from one hour to several hours, have been identified as a potential challenge, and are demonstrated in the well-known CAISO 'duck curve'. With increased renewable penetration, particularly solar, there is a need to ensure that longer ramps can be managed in a reliable fashion. The results here show that, assuming the system has access to renewable curtailment, interchange, and the ability to commit and dispatch all resources in an operational context, that the system can meet longer duration ramps.

The EPRI InFLEXion tool was used to analyze single years for several of the cases in more detail. Below, the amount of ramping available in that interval was calculated based on resource characteristics, with flexibility also available from interchanges based on the same assumptions described earlier; this is shown as available ramping. That was then compared with the potential requirements, which were based on calculating a certain percentile of the ramps observed during similar net load conditions; this is shown as net ramping. (note the actual flexibility available was compared with the largest potential ramp, rather than what actually occurred, such that this is a very conservative answer, and that the net flexibility may be greater than available when no upwards ramping is expected in the period). Even making the very conservative 100th percentile assumption, the number of periods when there were potential shortfalls in the three-hour time horizon were extremely low, with less than 1% of all intervals showing insufficient ramp capability in the 50% case, and even lower amounts in 43% and 33% cases.

---

[49] Variable energy resources can be dispatched by curtailing in advance of upward ramps, then reducing curtailment to increase generation.



Since 2015, the CAISO has been procuring a flexible ramping capacity under the Flexible Resource Adequacy – Must Offer Obligation (FRAC-MOO) program. There, 3 hour ramps are analyzed to determine a requirement by time of day and year, and then resources are procured that must provide themselves as available to the market. It should be noted that the results in this study do not necessarily invalidate the need for that construct, but do show that it is not necessarily needed from a long-term planning perspective if by default all resources provide their respective flexibility to the market. So the construct could still be needed to ensure the market has access to the relevant resources, and that availability to ramp is adequately considered in the market and available to short term operations.

From a long-term planning perspective, the scenarios described in the Results section can provide insights into how different assumptions and potential flexibility options can impact costs, emissions and, in some cases, reliability. From a technical resource flexibility perspective, minimum stable generation levels are shown to be an important source of flexibility. Therefore, as more renewables are integrated into the system, there will be a need to ensure that conventional resources can operate over as wide a range as possible. It will be important not to lose any of the existing flexibility as a reduction in operating range can have a significant impact on results.

Other study cases related to multi-hour ramping provided greater insight into how operational practices and policies can impact outcomes. Curtailment of renewables is shown to be very important to provide flexibility, particularly when flexibility from other resources is reduced; for example, reducing operating range of the conventional resources by 4,000 MW (Case SC_11) from current assumed ranges results in increasing curtailment from under 5% to over 12 %. This system can still be operated reliably, but curtailment is extremely high. Therefore, planning studies should consider curtailment as an option, but need to ensure that it is not overly excessive.

Treatment of interchanges is also an important aspect. Here, the Energy Imbalance Market, and any expansion of the CAISO, are not considered explicitly, but flexibility from interchange is shown to be important. Net import is shown to help balance solar variability in particular; this is already happening in CAISO. As coordination between regions continues to increase in the Western Interconnection, planning studies should ensure that this is considered as a potential source of flexibility, but also that assumptions made are realistic. The results showed that 3-hour interchange ramps are typically higher than what has been observed in recent years, and that limiting ramping on the interchange does have significant cost (and small reliability) impacts. The most expensive system costs produced in this study are for the case where interchange ramp was limited to 3,000 MW in three hours (Case SC_12).

5.5 Use of the Analytical Framework for Further Studies

One important insight from this study was that it showed how detailed modeling, with high temporal resolution, representation of neighboring areas, modeling of the impact of uncertainty on operations and detailed representation of system operations, can provide additional insight beyond more simple models. Models with less of this type of operational detail can often be used to determine potential generation expansion activities, but cannot always provide the insight into both reliability and economics as shown here. One of the outcomes of this study was that it is clear that, on a regular basis, detailed studies informing planning decisions should be revisited. At the same time, there is likely no need to continually evaluate the large range of scenarios studied here, including things like load following, P_{MIN}, interchange ramps, etc. An obvious question is therefore when such studies should be performed, and when less detailed studies can be used.

The example of the storage studies performed here show how this detailed approach can be used to study resource options in the future. As shown there, capacity contribution of storage can be calculated using the modeling approach used, while also looking at aspects such as the economic and renewable curtailment benefits of different levels of energy storage. Duration of energy storage was shown to be important, particularly when moving from very short durations of less than two hours to several hours; beyond a few hours showed less benefit.

In terms of when the studies should be performed, the specific answer is subjective. However, the study provides several potential sufficient changes to require additional studies. An obvious reason would be that renewable penetration being considered has changed sufficiently from previous studies. For example, the difference between the 33% and 50% cases here were significant. Therefore, if moving from one of the scenarios studied here to a higher penetration, or a significantly different mix of wind and solar, one should redo the study. Similarly, if significant changes occur to CAISO operations, then there may be a need to revisit the study; in particular increasing the size of the ISO may be a reason to revisit the analysis as it can have an impact on flexibility available from interchanges. Other reasons to perform similar studies again would be if there are significant changes in underlying conventional resources, such as decreased flexibility from those resources, particularly in terms of operating range. However, with small changes to the resource mix, the need for studies such as this one would not be as great.

6 CONCLUSION

This project has shown a method to assess the costs and reliability impacts of increasing renewable penetrations. A large range of different sensitivities on operational assumptions and flexibility resources are used to understand future system operations. While no new standards are observed to be needed, the methods used here can inform future planning activities, including the IRP. The general recommendations and insights described in the previous section should be considered appropriately when moving forward with planning in California. By reporting on the studies done here, the project team hopes to show how one can study these issues using an appropriate set of tools and data. Further work is described next. It is clear that the detailed modeling performed here could be used in conjunction with the resource expansion models used in IRP. Selecting a subset of potential future cases and analyzing them using the detailed approach described here can provide significant insights into reliability, costs, emissions and what the system operator may need to do to operate the system for the future resource mixes in the IRP.

The study described here was intended as research, and therefore specific numbers for some inputs (e.g. on costs of DR or minimum stable level of generation resources) were based on imprecise data. However, this did not affect the ability of the study to meaningfully analyze, in a research framework, the need for flexibility metrics and standards. As such, the specific outcomes here are less important than the general directional findings. More specific studies on particular aspects would be required if making planning decisions; this project was about developing a framework that can be used to make such decisions in the future.

6.1 Future Work

A number of potential issues were identified that could be further examined, ranging from operational or market assumptions and policy, technical characteristics of the system and data used in these modeling exercises.

From an operational policy, the importance of curtailment to maintaining reliability was shown here, and should be further considered. For example, tranches of curtailment may need to be identified. In the results here, nearly all curtailment can be managed via wind and solar curtailment only, but a handful of hours indicate curtailment above the hourly output of all wind and solar resources. In those cases, there is a question of whether hydro/BTM-PV should be curtailed when RPS is exhausted. More generally, the costs of different levels of curtailment may need to be considered, e.g. curtailment up to a small amount may be relatively inexpensive, but will become progressively more expensive. The framework described here could be used.

Another operational issue is the need to ensure that merchant generation from other regions is available. In reality, long-term contracts and other potential markets may limit flexibility from other regions. While the model here analyzed the entire WECC system with significant detail, it may still need further analysis on the likelihood and potential impact if flexibility is not available.

As discussed earlier, load following requires further study in a number of aspects. The framework used here could be used to study more efficient methods to carry reserves while minimizing costs of doing so. Regardless of costs, there may be a need to consider the load following impact on LOLE, and therefore ELCC of new resources. For example, increasing load following can have an impact on capacity shortfalls; if this is the case, there may be a need to revisit ELCC of the resources causing the need for load following. On the other hand, more thought needs to be given to what the appropriate

$LOLE_{INTRA\text{-}HOUR}$ should be; this is a new metric so assuming it gets rolled into an overall LOLE needs to be considered carefully.

In terms of technical flexibility, further work is likely needed, mainly to ensure that the assumptions about flexibility resources are accurate; this includes both cost and flexibility attributes. For example, minimum generation level has been shown to be capable of reduction, but at potential capital and operating costs. Similarly energy limited Demand Response resources also have economic parameters that should be more carefully studied.

From a data perspective, the data for intra-hour variability and day ahead and hour ahead uncertainty may need further investigation. Relatively simple methods were used to scale data here, such that the diversity benefits associated with increased renewable penetration (where per-unit variability often decreases due to increased geographic diversity) were not captured in a detailed statistical fashion. The forecast errors assumed may also need further analysis, to ensure they are reflective of the uncertainty that would actually be seen in operations.

Through collaborative research between PG&E, SDG&E, LLNL, the project team, and the Advisory Group, this CES-21 project has successfully investigated the feasibility of maintaining operational flexibility as renewable generation increases, identified some economic tradeoffs for achieving that flexibility, and provided a valuable tool and framework for IRP stakeholders to quantitatively analyze new planning scenarios as California's electric grid continues to evolve.