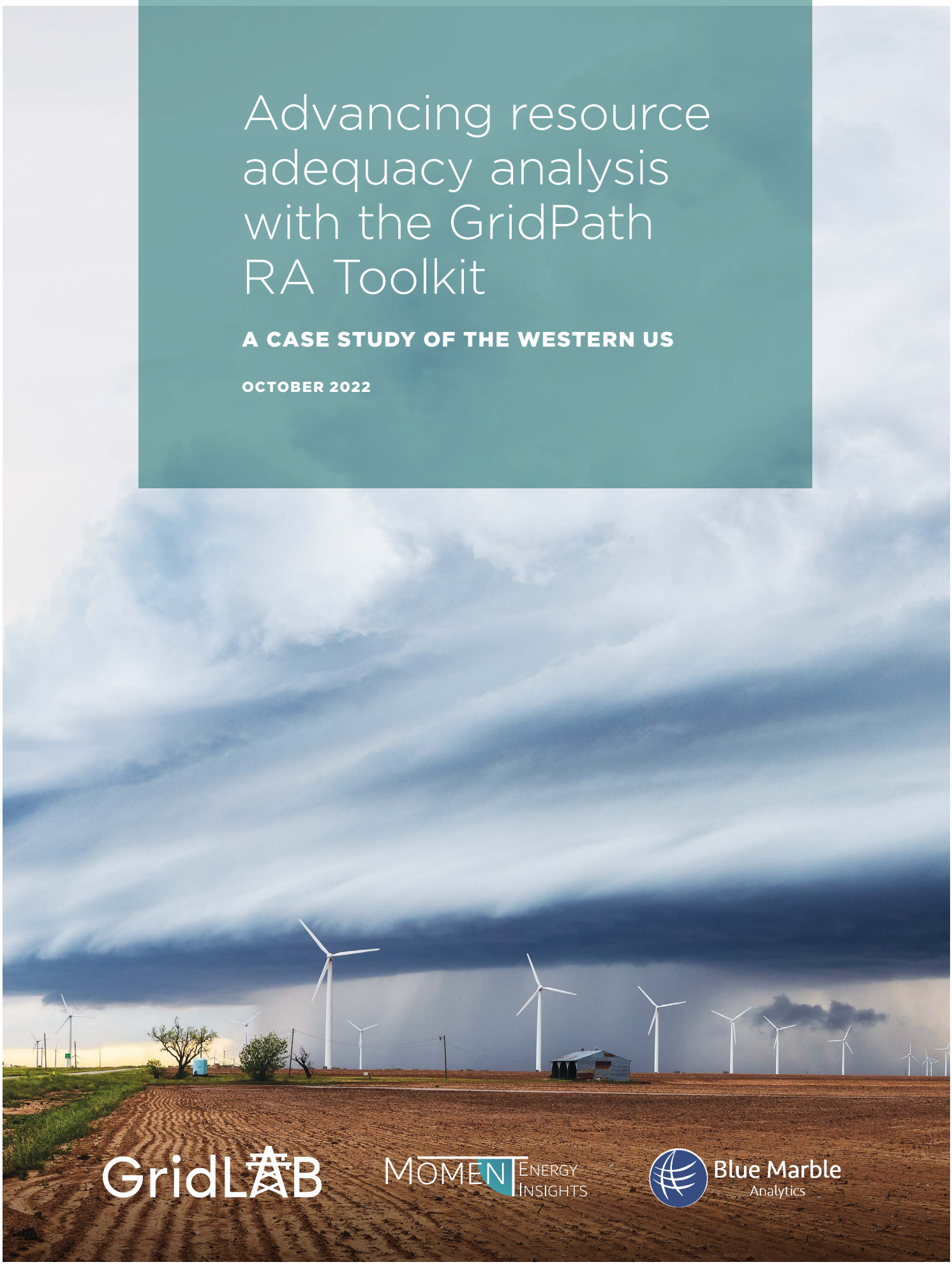
Advancing resource adequacy analysis with the GridPath RA Toolkit

A CASE STUDY OF THE WESTERN US

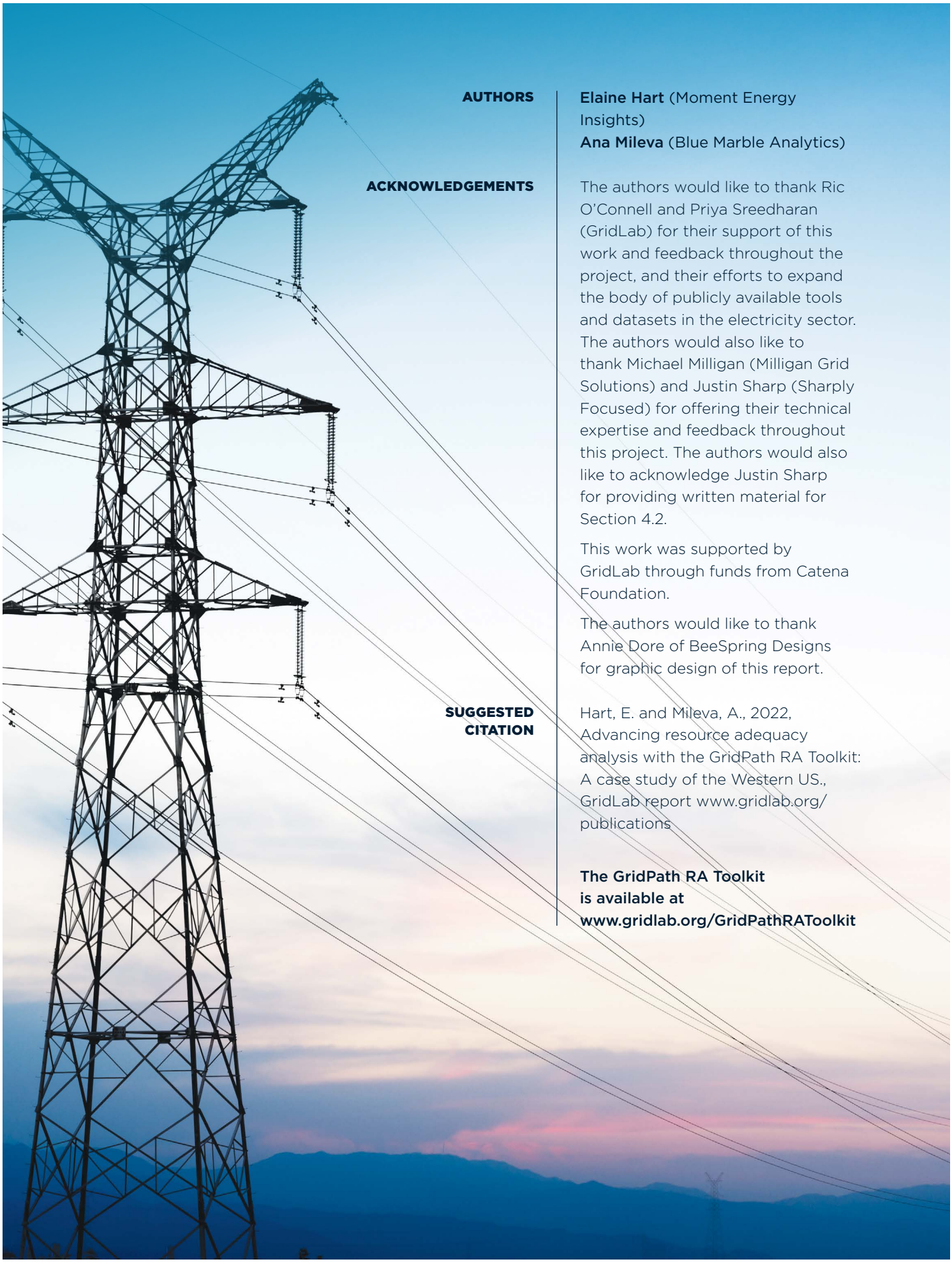
OCTOBER 2022



GridLAB

MOMEN ENERGY INSIGHTS

Blue Marble Analytics



AUTHORS

Elaine Hart (Moment Energy Insights)
Ana Mileva (Blue Marble Analytics)

ACKNOWLEDGEMENTS

The authors would like to thank Ric O'Connell and Priya Sreedharan (GridLab) for their support of this work and feedback throughout the project, and their efforts to expand the body of publicly available tools and datasets in the electricity sector. The authors would also like to thank Michael Milligan (Milligan Grid Solutions) and Justin Sharp (Sharply Focused) for offering their technical expertise and feedback throughout this project. The authors would also like to acknowledge Justin Sharp for providing written material for Section 4.2.

This work was supported by GridLab through funds from Catena Foundation.

The authors would like to thank Annie Dore of BeeSpring Designs for graphic design of this report.

SUGGESTED CITATION

Hart, E. and Mileva, A., 2022, Advancing resource adequacy analysis with the GridPath RA Toolkit: A case study of the Western US., GridLab report www.gridlab.org/publications

The GridPath RA Toolkit is available at www.gridlab.org/GridPathRAToolkit

TABLE OF CONTENTS

ABSTRACT

Robust and thoroughly vetted RA analyses will be critical to maintaining reliable supply as the power sector undergoes the clean energy transition. Current analytical methods used in RA analysis may be insufficient for analyzing RA of power systems that rely heavily on renewable energy and energy storage. Current practices are also characterized by limited transparency and inconsistency across jurisdictions. This work seeks to advance RA analysis by better characterizing three phenomena that are critical to understanding RA: weather-driven relationships between load and resource availability; capabilities and constraints of energy limited resources; and transmission flows and regional coordination. To improve transparency and accessibility, our approach leverages publicly available data and an open-source power system model (GridPath). The data and algorithms in this study are referred to as the GridPath RA Toolkit ("Toolkit") and are shared publicly.

The Toolkit offers two simulation modes to estimate resource adequacy challenges and their probabilities: a Monte Carlo Simulation mode, which samples from historical weather conditions to generate plausible combinations of load conditions and weather-driven resource availability (wind, solar, and thermal derates); and a Weather-Synchronized Simulation mode, which tests a more limited set of coherent weather conditions from the historical record. Both simulation modes test a wide range of hydro conditions and use Monte Carlo analysis to test unit forced outages.

We illustrate the Toolkit's features with a near term Western US case study that tests the physical capabilities of the Western Interconnect, excluding Canada and Mexico. We use the Monte Carlo Simulation mode to test three scenarios, each with a different hypothetical resource portfolio across the West in 2026. We find that if current retirement plans come to fruition and utilities take no incremental action between now and 2026 (the No Additions Scenario), the system would need 8.8-9.9 GW of additional capacity with at least 3 hours of duration to achieve a one-day-in-10-year standard. We also find that if California utilities add resources in line with the California Public Utility Commission's Preferred System Plan through 2026 (the California Additions Scenario), these additions would meet all of the identified West-wide capacity needs. We

also show that the retirement of an additional 11 GW of coal (the Less Coal Scenario) would not create an insurmountable resource adequacy challenge if utilities continue with current plans. Subregional analysis of these scenarios is provided to demonstrate how import policies can account for coherent weather across the broader West and to illustrate how sensitive RA findings are to import assumptions.

Finally, we investigate the limitations of Monte Carlo analysis for estimating weather-dependent system conditions in RA analysis by comparing the Monte Carlo Simulation results to Weather-Synchronized Simulation results. We find that Weather-Synchronized Simulation better captures correlations during the most extreme events, but is highly sensitive to the years over which data is available. We flag expansion of publicly available hourly datasets as critical to sound RA modeling and we introduce a statistical method for extending Weather-Synchronized Simulation results to a much longer historical or simulated weather record.

This work demonstrates the importance of adequately capturing weather-driven correlations in load and resource availability for interconnected systems over large geographic areas. It explores two methods of capturing these correlations and it illustrates how import policies can be designed to account for weather impacts outside of a particular subregion, utility, or RA program. It also shows how the expansion of public datasets, the development of open source tools, and the use of transparent and more granular RA metrics can be applied to improve resource adequacy decision making.

EXECUTIVE SUMMARY

MOTIVATION

Recent years have seen a renewed focus on resource adequacy (RA) in the Western United States, as aging coal plants begin to retire and new technologies and policies are quickly changing the composition of the Western grid. Clean technologies are increasingly cost competitive, but also create new technical challenges for grid planners and operators tasked with maintaining system reliability. Amidst these changes, the West has also seen extreme weather in recent years, leading to both unusually high demand for electricity and exacerbation of risks to grid infrastructure. Robust and thoroughly vetted RA analyses will be critical to maintaining reliable supply as the power sector continues to evolve in the coming years. Current approaches to RA analysis face many challenges, notably substantial data requirements and lack of transparency as the complexities of the power sector and power system models grow. Modern RA approaches must grapple with the impact of weather on both demand and resource availability, the increasing reliance on energy-limited resources such as batteries, and the potential for regional coordination to promote reliability through load and resource diversity across large interconnected systems. At the same time, they must offer transparency into the methodologies, assumptions, and data used to allow for meaningful vetting by regulatory agencies and other oversight bodies.

Modern RA approaches must grapple with the impact of weather on both demand and resource availability, the increasing reliance on energy-limited resources such as batteries, and the potential for regional coordination to promote reliability through load and resource diversity across large interconnected systems.

GRIDPATH RA TOOLKIT

With this motivation, GridLab sponsored this effort to build a toolkit, leveraging publicly available data sets and open-source power system analytical tools, that can assess resource adequacy for emerging power systems. The effort was conducted jointly by Moment Energy Insights and Blue Marble Analytics and leveraged GridPath, Blue Marble's open-source platform for power system planning and optimization. In addition to promoting transparency, using an open-source tool in this application allows for continued development of model capabilities and customization to specific systems over time by the study team and by other organizations—these features are important as the nature of the resource adequacy challenge and the assumptions and simplifications necessary to assess resource adequacy are system-dependent and will change over time.

In this phase, we focused on the following methodological priorities:

- **Weather-driven relationships between loads and resource availability across a large geographical region:** Appropriately valuing load and resource diversity as well as improved regional coordination requires RA models to carefully account for spatial and temporal correlations between electric loads and weather-driven resource availability constraints (e.g., wind availability, solar availability, and thermal derates). This is especially challenging when interactions can occur over geographical areas that are similar in size or larger than the spatial extent of critical weather phenomena.
- **Dispatch simulation capturing the capabilities and constraints of energy-limited resources:** Understanding the dynamic capabilities and limitations of energy-limited resources, such as hydropower and storage, is critical to evaluating their contributions toward resource adequacy. In particular, the ability to co-optimize energy-limited resources with variable renewables will help to unlock the potential of these resources.
- **Transmission flows and regional coordination:** Dynamic transmission flow optimization ensures that RA analysis takes into account physical transmission limits and transmission path constraints. It also ensures that the analysis recognizes the ability to co-optimize operations over large interconnected areas to unlock load and resource diversity benefits as well as the benefits of energy-limited dispatchable resources.

To test the capabilities of the system across a wide range of system conditions and to identify resource adequacy challenges, we relied on simplified optimization-based dispatch modeling using GridPath. For this RA analysis, GridPath minimized unserved energy subject to load, resource availability, contingency reserves, energy-related, and transmission constraints. The dispatch of all resources was co-optimized in order to account for the dynamic capabilities and limitations of energy-limited resources, like hydropower, energy storage, and hybrid renewable & storage systems, as well as the benefits of the transmission system.

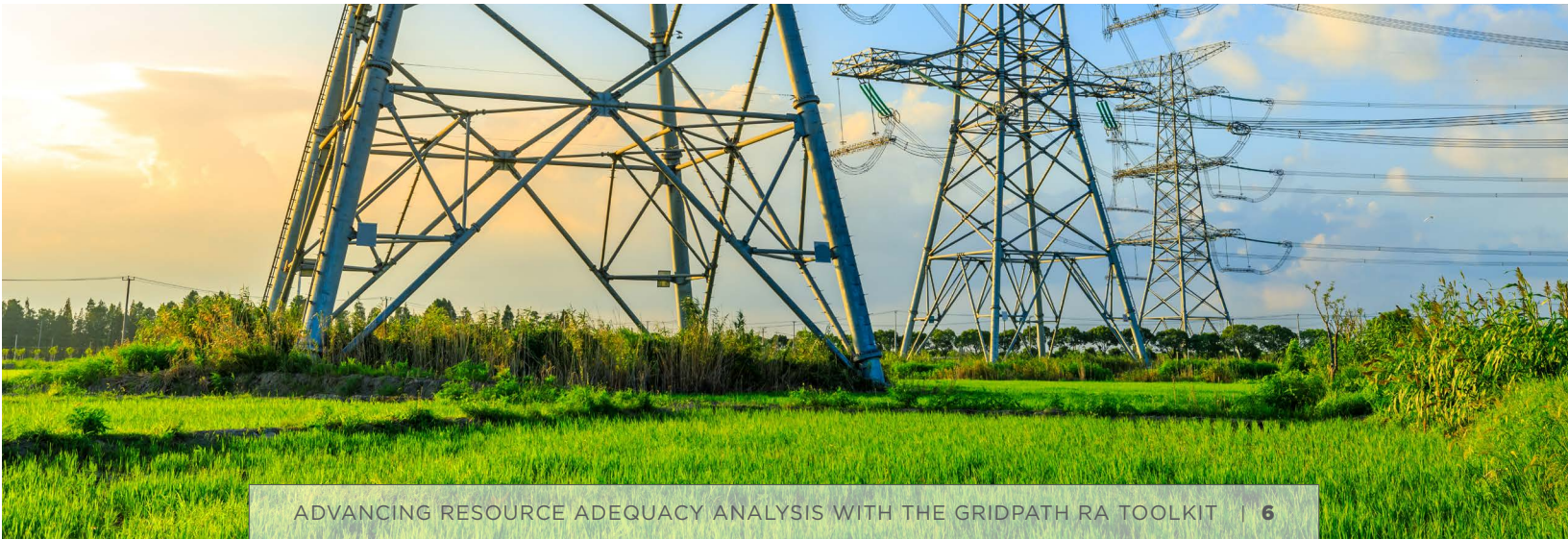
The GridPath RA Toolkit provides a *Monte Carlo Simulation* option that uses a Markov Chain approach to sample historical weather conditions and generate plausible load, renewable availability, and thermal derates. The sampling method aims to strike a reasonable balance between the number of possible combinations that can be tested and the physical plausibility of the resulting conditions. In this mode, Monte Carlo sampling is also used to randomly draw the hydro conditions for each simulated year and to simulate random forced outages.

While it is fairly common to use Monte Carlo analysis to sample weather conditions for RA analysis, this approach also has significant drawbacks. First, it depends on a subjective process of binning historical data and/or deciding

how to capture important geographical and temporal correlations. These decisions and their implications are often opaque, can be difficult to vet, and may result in conditions that are not physically realistic. Generating conditions in this manner also makes it difficult to identify the specific coherent weather conditions that drive RA risk over very large areas. Finally, while Monte Carlo analysis can provide very precise estimates by sampling very large numbers of conditions, such precision does not ensure accuracy and may lead to overconfidence by decision makers. To address some of the shortcomings of Monte Carlo analysis, the GridPath RA Toolkit also provides an option to simulate the system across a synchronized weather record. In this *Weather-Synchronized Simulation* mode, Monte Carlo is applied in a more limited fashion to simulate random thermal unit forced outages across all combinations of synchronized weather years and hydro years.

The GridPath RA Toolkit (Toolkit), which can be found at gridlab.org/GridPathRAToolkit, consists of GridPath, accompanying code that generates the datasets for use with GridPath via Monte Carlo or Weather-Synchronized Simulation, and the data inputs and settings used throughout the study. This initiative included a significant effort to develop renewable energy datasets based on publicly available sources and the resulting hourly renewables shapes are being made publicly available. A primary intention of sharing the datasets as part of the GridPath RA Toolkit is that researchers and analysts need not duplicate this step and may instead devote analytical efforts to asking and analyzing the important resource adequacy questions. The Toolkit provides information regarding system requirements for running a full simulation.

The GridPath RA Toolkit (Toolkit), which can be found at gridlab.org/GridPathRAToolkit, consists of GridPath, accompanying code that generates the datasets for use with GridPath via Monte Carlo or Weather-Synchronized Simulation, and the data inputs and settings used throughout the study.



WESTERN US CASE STUDY

In this report, we describe an application of the GridPath RA Toolkit to near-term (2026) resource adequacy challenges in the Western United States. We used this case study to test alternative approaches to RA modeling and to demonstrate how RA modeling might take into account regional coordination over a large and diverse footprint. The study area included all balancing area authority areas (BAAs) in the Western Interconnection, excluding those in Canada and Mexico. Transmission was represented by zonal constraints on flows between BAAs. We developed load and resource availability assumptions for each modeled BAA in 2026. The analysis reflected a physical model of the Western power grid and did not take into account resource ownership or contractual agreements for serving load. This means that the study results are broadly indicative of regional resource adequacy challenges, but cannot be used to attribute responsibility for shortages to individual load serving entities or RA programs.

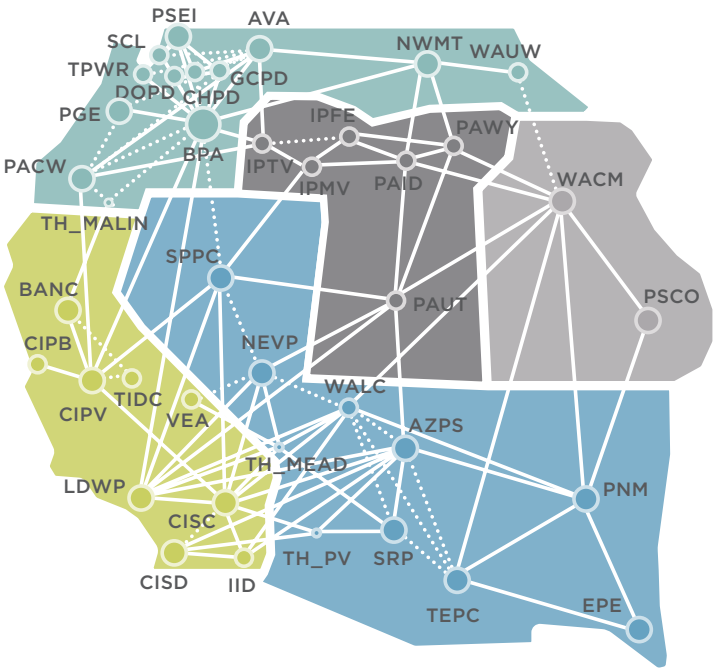


FIGURE ES.1.

Western US Case Study zonal topology.

WESTERN US DATASET

The availability of complete and transparent datasets for RA analysis can be an impediment to rigorous oversight. To further the state of RA modeling in the West, we provide a model-ready dataset for the Western US in the 2026 study year as part of the GridPath RA Toolkit. This data was developed

from publicly available datasets and publicly available tools, including EIA Forms 860M[1] and 923/906;[2] FERC Form 714;[3] the WECC 2026 Common Case;[4] NREL's National Solar Radiation Database (NSRDB),[5] Wind Integration National Dataset Toolkit (WIND Toolkit)[6] and System Advisor Model (SAM);[7] BPA's Total Load and Wind Generation Report;[8] CAISO's Daily Renewables Watch;[9] and the NCEI Global Surface Summary of the Day.[10] Some new tools were developed to efficiently collect this data and to generate the datasets required for power system analysis, including an empirical wind power curve estimator, a thermal derate estimator, and a tool for estimating hourly load in the study year based on historical data and economic forecasts.[11] To improve computational performance, much of the data provided as part of the GridPath RA Toolkit is aggregated to the WECC BAA level. Additional data

1 This study was based on the February 2021 EIA Form 860M, available at: https://www.eia.gov/electricity/data/eia860m/archive/xls/february_generator2021.xlsx

2 Available at: https://www.eia.gov/electricity/data/eia923

3 This study utilized the 2006-2020 Form 714 database, available at: https://www.ferc.gov/sites/default/files/2021-06/Form-714-csv-files-June-2021.zip

4 Available at: https://www.wecc.org/SystemAdequacyPlanning/Pages/Datasets.aspx

5 Sengupta, M., Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby. 2018. "The National Solar Radiation Data Base (NSRDB)." Renewable and Sustainable Energy Reviews 89 (June): 51-60.

6 Draxl, C., B.M. Hodge, A. Clifton, and J. McCaa. 2015. Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit (Technical Report, NREL/TP-5000-61740). Golden, CO: National Renewable Energy Laboratory. Draxl, C., B.M. Hodge, A. Clifton, and J. McCaa. 2015. "The Wind Integration National Dataset (WIND) Toolkit." *Applied Energy* 151: 355366. King, J., A. Clifton, and B.M. Hodge. 2014. Validation of Power Output for the WIND Toolkit (Technical Report, NREL/TP-5D00-61714). Golden, CO: National Renewable Energy Laboratory.

7 PySAM Version 2.2.2. National Renewable Energy Laboratory. Golden, CO. https://github.com/nrel/pysam.

8 Available at: https://transmission.bpa.gov/business/operations/wind/ (item 5)

9 Available at: http://www.caiso.com/market/Pages/ReportsBulletins/RenewablesReporting.aspx

10 Global Surface Summary of the Day - GSOD, National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce, available at: https://www.ncei.noaa.gov/access/search/data-search/global-summary-of-the-day

11 The code for these new tools can be requested from the study team and shared under an open-source license, but it will not be maintained over time.

used in the study and higher resolution datasets, including those listed in Table ES.1, can be obtained by contacting the study team.

TABLE ES.1.

Western US Case Study key data sources.

DATA	YEARS	SOURCES	GEOGRAPHICAL RESOLUTION
Resource Stack	2026	EIA Form 860M	EIA Generator
Hourly Load	2006-2020	Historical load: FERC Form 714 Load zone topology: 2026 WECC Common Case	WECC BAA
Hourly Wind	2007-2014	Hourly wind speed: NREL WIND Toolkit, Historical generation: EIA Form 923/906	EIA Plant
Hourly Solar	1998-2019	Hourly weather: NSRDB, Technological specifications: NREL System Advisor Model, LBNL's Utility-Scale Solar Data Update: 2020 Edition[12] and EIA Form 860M	EIA Plant
Hourly Thermal Derates	1998-2019	Forced outage rates: 2026 WECC Common Case Hourly temperature: NSRDB Winter and summer capacities: EIA Form 860M	EIA Generator
Monthly Hydro Energy, Pmin, and Pmax	2001-2020	Historical generation: EIA Form 923/906, Hourly data used to derive Pmin and Pmax: BPA Total Load and Wind Generation Report, CAISO Daily Renewables Watch, and WECC 2026 WECC Common Case	WECC BAA
Transmission & Interfaces	2026	2026 WECC Common Case	WECC BAA
Weather conditions	1949-2019	NCEI Global Surface Summary of the Day	16 sites across the West
Economic conditions	2006-2026	US Bureau of Economic Analysis,[13] 2021 EIA Annual Energy Outlook[14]	Western US

SCENARIO ANALYSIS

To better understand the nature of the near-term RA challenge in the West, we tested three scenarios using Monte Carlo Simulation: the No Additions Scenario, which reflects planned retirements and no planned resource additions; the California Additions Scenario, which layers on additional resources in CAISO based on the California Public Utility Commission's (CPUC) Preferred System Plan; and the Less Coal Scenario, which incorporates the additional resources in California while also retiring about 11 GW of coal elsewhere in the West. All three scenarios were modeled as a physical system

12 Bolinger, M., Seel, J., Robson, D., Warner, C., "Utility-Scale Solar Data Update: 2020 Edition,"Lawrence Berkeley National Laboratory, November 2020. Available at: https://emp.lbl.gov/publications/utility-scale-solar-data-update-2020

13 Annual GDP by State, SAGDP tables, available at: https://apps.bea.gov/regional/downloadzip.cfm

14 U.S. Energy Information Administration, Annual Energy Outlook 2021. Macroeconomic Indicators table, available at: https://www.eia.gov/outlooks/aeo/tables_side.php

that is unencumbered by institutional barriers to coordination and where short-term operational decisions can be made optimally to avoid lost load.

West-wide Findings

High-level results for the three scenarios are presented in Table ES.2.[15] We estimated that, without taking any action, the West could be physically short by approximately 8.8 to 9.9 GW in 2026 if planning to a traditional one-day-in-10-year *LOLE* standard. However, the identified shortage was much smaller than the amount of capacity additions in current utility plans in the West. Procurement authorized in California would effectively eliminate this capacity shortfall. The 28 GW of capacity added in the California Additions Scenario resulted in only seven reliability events in 1,000 years of simulated conditions, well below the one-day-in-10-year standard and meeting all of the perfect capacity needs identified in the No Additions Scenario. The Less Coal Scenario did not achieve a one-day-in-10-year standard, but well outperformed the No Additions Scenario, indicating that the resource adequacy contribution of the added resources in California exceeded that of the 11 GW of additionally retired coal units. The remaining capacity need identified for the Less Coal Scenario was about 3.3 to 4.2 GW to meet the one-day-in-10-year standard.

TABLE ES.2.

Western US Case Study key scenario RA metrics.

METRIC	NO ADDITIONS SCENARIO	CA ADDITIONS SCENARIO	LESS COAL SCENARIO
LOLE (days/10yrs)	17.3 - 19.0	0.02 - 0.12	3.73 - 7.86
Perfect Capacity Need *LOLE* = One day in 10 years	8.8 - 9.9 GW	0 GW	3.3 - 4.2 GW

In addition to the high-level reliability metrics above, GridPath provided detailed information on the nature of loss of load events encountered in the simulation. In the No Additions Scenario, loss of load events were identified only in the summer afternoon and evening hours. The highest probability of lost load was in the hour ending (HE) 18 (6-7 pm in Pacific Daylight Time) in August. Only 7 percent of the identified events were longer than 4 hours and all were

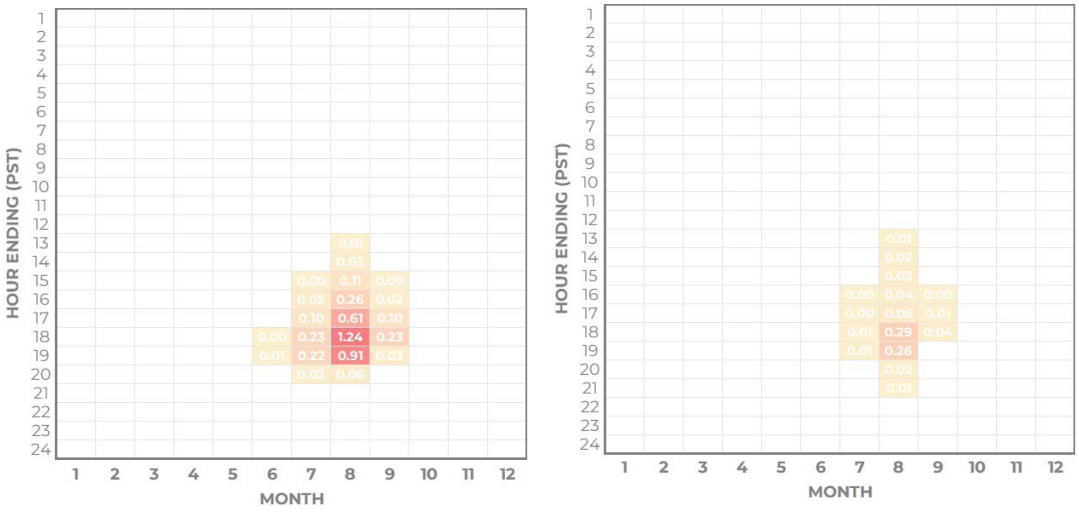
Our analysis of the maximum shortages and total energy shortages of the loss-of-load events for the No Additions Scenario suggested that the economically optimal solution to meet the one-day-in-10-year standard was unlikely to require beyond 4 hours of sustained duration, suggesting that battery storage may be well suited to alleviate near term RA challenges.

15 Ranges represent a 95% confidence interval.

less than 8 hours in duration. Our analysis of the maximum shortages and total energy shortages of the loss-of-load events for the No Additions Scenario suggested that the economically optimal solution to meet the one-day-in-10-year standard was unlikely to require beyond 4 hours of sustained duration, suggesting that battery storage may be well suited to alleviate near term RA challenges. Relative to the No Additions Scenario, the Less Coal Scenario eliminated most loss of load events outside of August afternoons and evenings, and 91% of the remaining events were less than 4 hours in duration.

FIGURE ES.2.

Loss of load hours per year in the (a) No Additions Scenario and (b) Less Coal Scenario.



Despite the retirement of almost 11 GW of coal resources, the Less Coal Scenario relied on the additional batteries to eliminate many of the capacity shortages from the No Additions Scenario. Resource availability on an example day from the two scenarios is shown in Figure ES.3. Panel (a) shows this day in the No Additions Scenario and panel (b) shows it for the Less Coal Scenario. On this day, storage charged early in the day and shifted the energy to the evening reducing the amount of unserved energy despite the reduced coal availability.

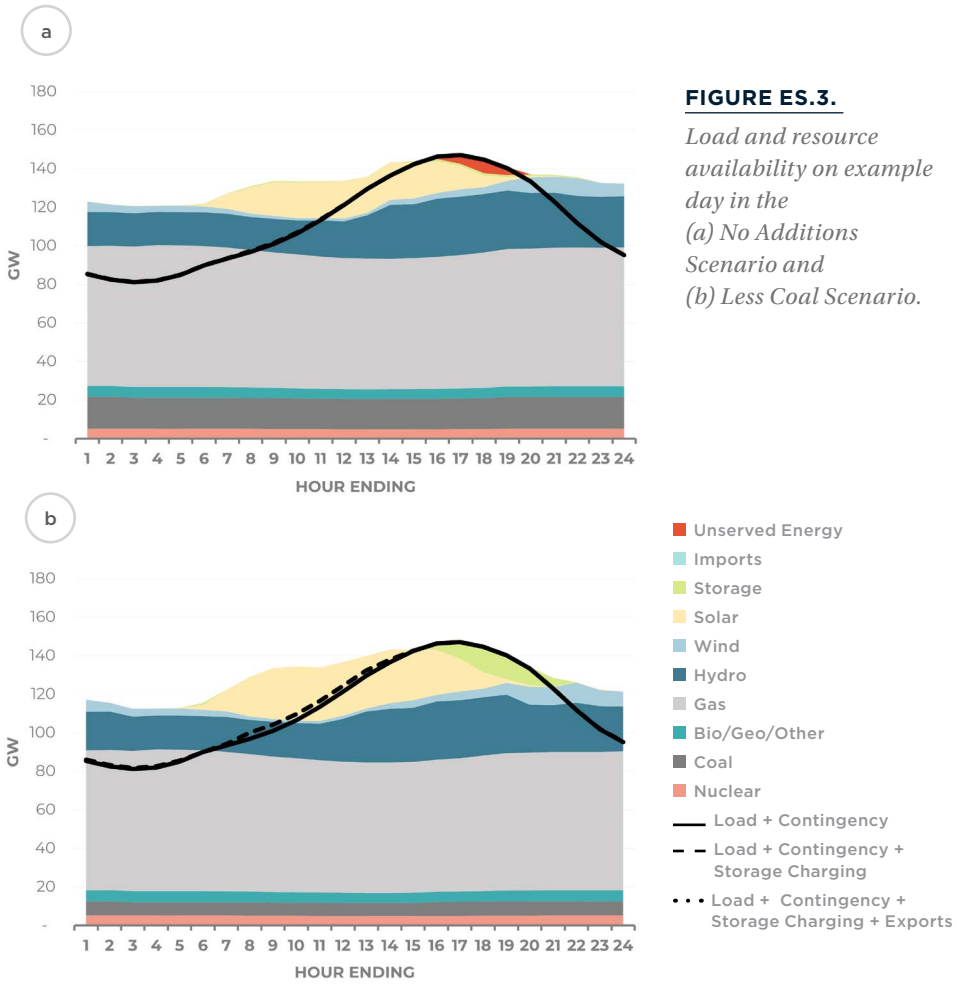


FIGURE ES.3.

Load and resource availability on example day in the (a) No Additions Scenario and (b) Less Coal Scenario.

Subregional Analysis

In addition to simulating each scenario on a West-wide basis, we also tested the No Additions Scenario and the Less Coal Scenario over two subareas that approximate the CAISO and Western Resource Adequacy Program (WRAP) footprints.[16] We tested these subareas with and without physically coherent import availability to understand how import policies may affect RA analysis findings. To examine the subareas, we first simulated each as a physical island with no import capabilities. We then compared the islanded simulation to the West-wide simulation to understand whether any lost load in the islanded simulation could be avoided by accounting for the rest of the West. In the No Additions Scenario, simulating CAISO as an island resulted in a high probability of lost load and a capacity need of 11.2 GW. Allowing for imports reduced the identified capacity need in the CAISO footprint to 8.2 GW, highlighting the significance of imports to CAISO's RA position.

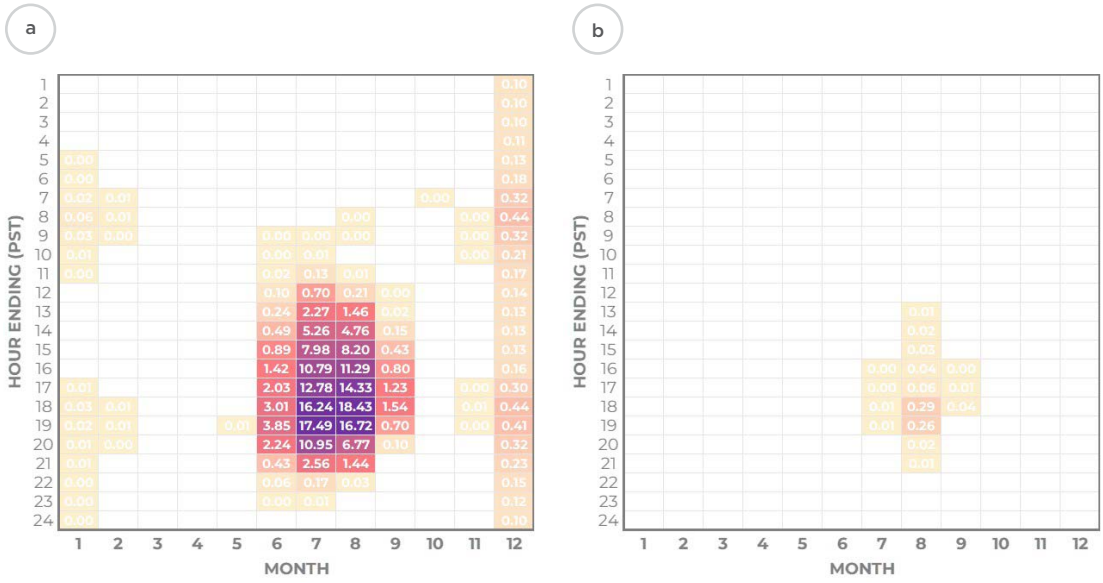
16 Note that the subareas do not exactly align with real resource adequacy programs because they consider loads and resources that are associated with each balancing area within a physical model of the West, rather than allocating resources to load serving entities (LSEs) based on ownership and contractual information.

We also found that import assumptions were a major driver of the identified needs for the WRAP subarea in the Less Coal Scenario, in which all of the coal in the WRAP footprint was retired. When WRAP was treated as an island in this scenario, shortages were observed in all months, except March and April, and were long in duration in both the summer and winter (See Figure ES.4). Accounting for imports alleviated most shortages in the WRAP subarea, with the remaining shortages limited to summer evenings, and similar in timing and duration to those identified in the West-wide analysis of the Less Coal Scenario.

Due to the highly interconnected nature of the West, we find that resource adequacy analysis that treats subareas or RA programs as islands can distort the observed RA challenges and may lead to suboptimal RA solutions, including potentially significant overbuild and employing solutions that are not well suited to the nature of the RA challenge.

FIGURE ES.4.

Loss of load hours per year for the WRAP subarea in the Less Coal Scenario when (a) the subarea is modeled as an island and (b) the subarea has access to imports.



Due to the highly interconnected nature of the West, we find that resource adequacy analysis that treats subareas or RA programs as islands can distort the observed RA challenges and may lead to suboptimal RA solutions, including potentially significant overbuild and employing solutions that are not well suited to the nature of the RA challenge.

WEATHER-SYNCHRONIZED SIMULATION INSIGHTS

We further explored the No Additions Scenario using Weather-Synchronized Simulation to better understand the limitations of Monte Carlo-based approaches to characterizing weather conditions in RA analysis. Weather-Synchronized Simulation tests fewer potential weather variations than Monte Carlo analysis, but it provides confidence that the findings reflect actual physical weather phenomena and all relevant spatial and temporal correlations. The Weather-Synchronized approach requires several years of time-synchronized load, wind, solar, and thermal derate data in order to be meaningful. Fully synchronized historical hourly load, wind, solar, and temperature data was available for the period 2007-2014, with load data also available through 2020, and solar and temperature data available through 2019. We developed synthesized wind, solar, and thermal derate hourly data for the respective missing years to lengthen the synchronized period to 2007-2020. We tested the full synchronized record across each available hydro year (2001-2020) and across 30 forced outage iterations, resulting in a total of 8,400 years of potential conditions.

Comparison to Monte Carlo Simulation Results

Using Weather-Synchronized Simulation to examine the No Additions Scenario resulted in higher loss of load expectation than Monte Carlo Simulation as well as larger shortages and slightly longer event durations, although all events were still 8 hours or less with similar timing in the summer evenings. Perfect capacity needs were about 1.8 GW higher in the Weather-Synchronized Simulation relative to the Monte Carlo Simulation. Notably, if we limited the analysis to the 8 years without synthesized data (2007-2014), the loss of load expectation and capacity needs were much lower because of the relatively high frequency of loss of load events in recent weather years. This finding highlights the sensitivity of RA analysis to the specific years of weather conditions that are included and underscores the importance of continued efforts to expand publicly available datasets like NREL's Wind ToolKit to recent weather years.

TABLE ES.3.

Key RA metrics in the No Additions Scenario using Weather-Synchronized Simulation.

METRIC	NO ADDITIONS SCENARIO (WEATHER-SYNCHRONIZED 2007-2020)	NO ADDITIONS SCENARIO (WEATHER-SYNCHRONIZED 2007-2014)	NO ADDITIONS SCENARIO (MONTE CARLO)
LOLE (days/10yrs)	24.6 - 25.3	10.7 - 11.3	17.3 - 19.0
LOLE = One day in 10 years	10.9 - 11.4 GW	6.3 - 6.6 GW	8.8 - 9.9 GW

Weather Insights

Weather-Synchronized Simulation allows for the identification of the specific weather conditions across the West that pose the greatest risk to resource adequacy in the 2026 timeframe. At a high level, we find that temperature remains the key weather driver of loss of load risk in this system. Notably, the days with non-zero loss of load probability see abnormally high temperatures across coastal load centers in California and the days with the greatest loss of load probability also see abnormally high temperatures across most of the West. The events in August 2020 serve as an example of the most challenging type of weather phenomenon for near-term RA in the Western United States. The August 2020 heat event was uncharacteristically hot across most of the West and the coincidence of unusually hot conditions at most load centers across the West resulted in very high loss of load risk in the simulation. This event was driven by a weather phenomenon known as the West Coast Thermal Trough, a self-reinforcing cycle that pushes air from the desert southwest northward between the Sierra/Cascades and the coast, which can bring coincident well-above-average temperatures to California and Western Oregon/Washington.[17] The weather conditions on two of the most challenging days during this event are shown in Figure ES.5.[18]

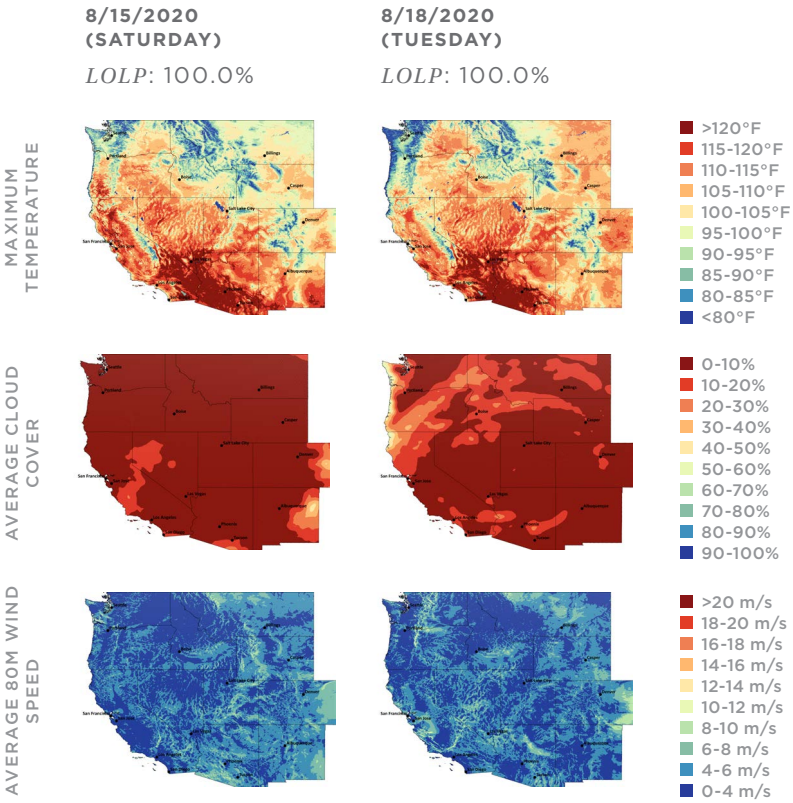
Weather-Synchronized Simulation allows for the identification of the specific weather conditions across the West that pose the greatest risk to resource adequacy in the 2026 timeframe.

17 Brewer, Matthew & Mass, Clifford & Potter, Brian. (2013). The West Coast Thermal Trough: Mesoscale Evolution and Sensitivity to Terrain and Surface Fluxes. Monthly Weather Review. 141. 2869-2896. 10.1175/MWR-D-12-00305.1.

18 All weather maps in this report were created with data from the NOAA High-Resolution Rapid Refresh (HRRR) analysis dataset, accessed via the HRRR Data Archive: AWS Open Data Program (https://mesowest.utah.edu/html/hrrr/)

FIGURE ES.5.

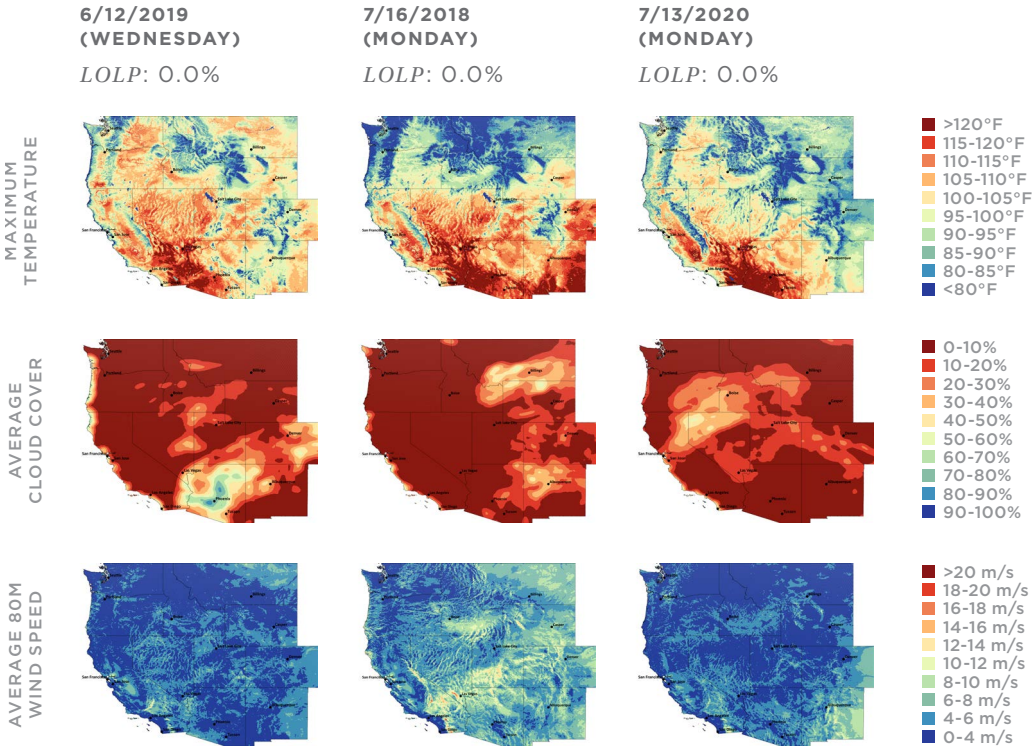
Regional weather on days during the August 2020 heat wave with 100% LOLP in the No Additions Scenario.



Days with more geographically isolated heat, which may result in very high load conditions in localized parts of the West, did not tend to pose loss of load risk in the modeled system due to geographical diversity. Three examples are shown in Figure ES.6. June 12th, 2019 was unusually hot in Portland and parts of California and the Southwest, but was relatively mild along coastal California load centers. July 16th, 2018 was unusually hot in Seattle, Portland, and areas East of the Cascades and Sierras, but was less extreme along the California coast and the Southwest. And July 13th, 2020 saw extreme heat in the Desert Southwest, but relatively mild conditions across the Northwest. Despite extreme heat in parts of the West on each of these days, none of them saw loss of load risk in the Weather-Synchronized Simulation thanks to geographical load diversity.

FIGURE ES.6.

Regional weather on days with localized heat but no simulated loss of load risk in the No Additions Scenario.



STATISTICAL ANALYSIS

Using the results from the Weather-Synchronized Simulation mode, we developed a statistical model based on logistic regression to estimate the probability of lost load as a function of daily weather conditions on each simulated day as well as several other parameters including the daylight hours, whether the day was weekday or weekend, and the West-wide hydro budget for the corresponding week. This statistical model can 1) provide important information about the drivers of RA challenges and 2) estimate the risk of loss of load events during times and weather conditions that were not directly simulated, both historical weather conditions and potentially under future climate scenarios. We note that because so few weather conditions result in RA risk, we relied on ensemble methods, random out-of-sample testing, and other measures to avoid overfitting.

Our analysis confirmed that temperature is the key driver of loss of load risk while higher wind speeds in key locations such as Wyoming and Montana may correlate with lower risk. Our final statistical model included temperature and

wind speed as the two weather variables; other weather variables such as dew point and pressure did not meaningfully improve the performance of the model.

We applied the logistic regression model to historical weather conditions for which the high resolution hourly data required for power system modeling was not available. As an example, we used the logistic regression model to estimate loss of load risk during the historic heat dome event in June 2021, which broke several high temperature records across the Northwest (116°F in Portland and 108°F in Seattle). Despite the historic heat in parts of the West on June 28, 2021 (see Figure ES.7), we estimated a loss of load probably of only 1.2% on this day because other parts of the West saw much milder conditions.

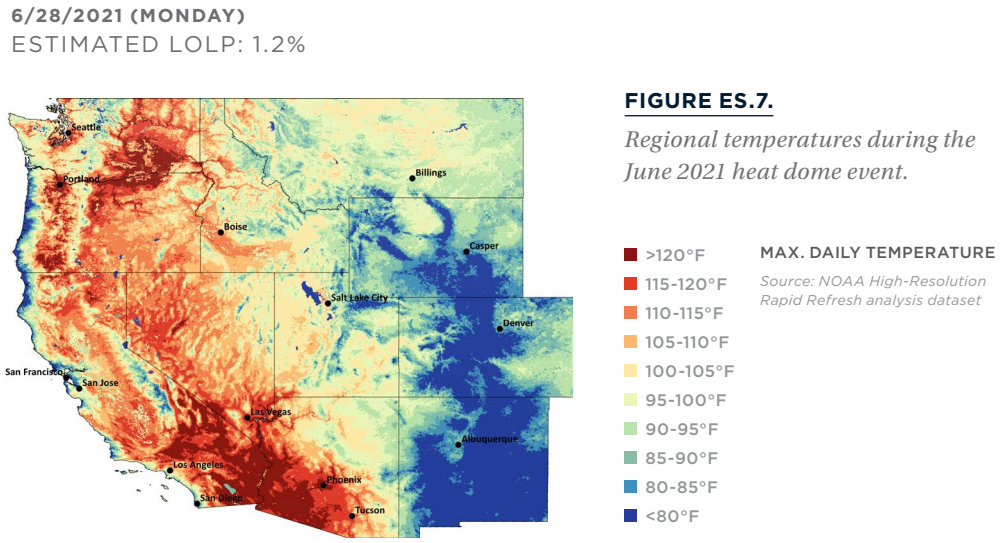


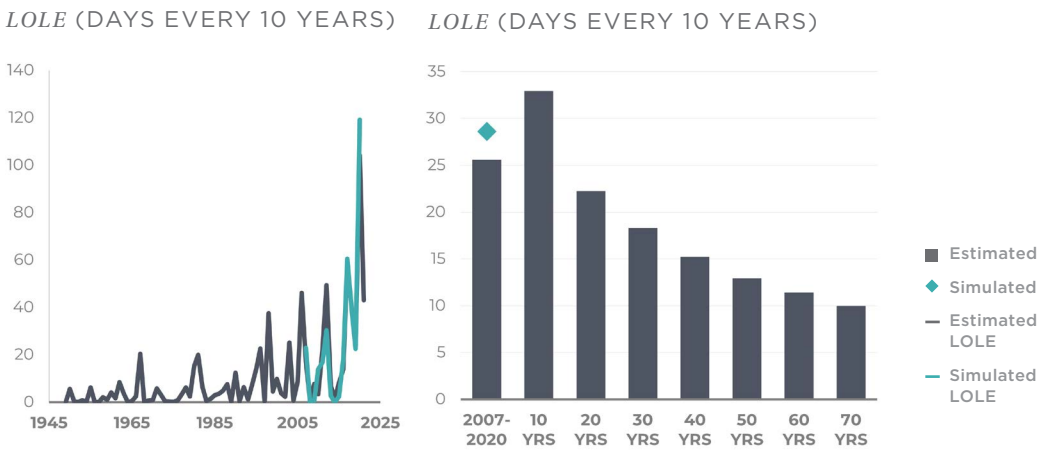
FIGURE ES.7.

Regional temperatures during the June 2021 heat dome event.

We also applied the logistic regression model to daily weather conditions going back to 1949 to investigate the loss of load risk across a longer historical record of weather conditions. Figure ES.8 shows the results of this analysis averaged across the hydro years. We found that weather patterns that drive the RA risk identified in the No Additions portfolio have increased in frequency since the 1990s. This trend has important implications for RA analysis more generally as historical weather data is often used to approximate future weather distributions.

FIGURE ES.8.

Statistical estimation of LOLE across historical weather conditions.



To explore the implications of this finding, we applied the statistical model across various historical weather records. We found that *LOLE* estimates trend lower as more distant historical years are used—and assumed equally likely—in the analysis. We note that without information about future weather and how it may compare to the historical data, decisions regarding the number of historical weather years to consider and/or the application of detrending methods to historical weather are effectively policy decisions based on risk tolerance. The development and use of future weather data, for example using climate simulations and downscaling, would greatly improve our understanding of the RA challenges future energy systems might face, and how to address those challenges.

The development and use of future weather data, for example using climate simulations and downscaling, would greatly improve our understanding of the RA challenges future energy systems might face, and how to address those challenges.

CONCLUSIONS AND NEXT STEPS

The Western US Case Study described in this report yields some key insights regarding the state of resource adequacy in the West and the relative strengths and weaknesses of RA methodologies, especially in the context of a changing climate.

INSIGHTS REGARDING THE WESTERN US CASE STUDY

- **Without any incremental resource additions, the Western United States power system could face a resource adequacy shortage in 2026. However, current California procurement plans include sufficient new capacity to eliminate this shortfall.**

  Without deploying new resources through 2026, we estimate that the West could be physically short by about 8.8 to 9.9 GW in 2026, if planning to a one-day-in-10-year *LOLE* standard. This shortage is much smaller than the amount of capacity additions in current utility plans in the West, including the procurement ordered California Public Utility Commission (CPUC) Decision D.21-06-035, which requires 11.5 GW of new net qualifying capacity through 2026. Incorporating capacity additions consistent with California's Preferred System Plan through 2026 eliminates all but seven RA events in 1,000 years of simulated dispatch.

- **Additional coal retirements do not seem to pose an insurmountable RA challenge in the near term.**

  The deployment of additional batteries and renewable resources in California appears to mitigate much of the needs associated with retiring a large portion of the West-wide coal fleet, even before considering capacity additions from utility plans in the rest of the West. We estimate perfect capacity needs of about 3.8 GW for this case and find that short-duration solutions are likely adequate as the majority of RA shortages were 4 hours or less and occured in the evenings on hot summer days. Energy-limited resources such as batteries or demand flexibility are well suited to this type of shortage. The system does not appear to be energy-limited, with plentiful resources available to charge storage outside of the high-risk hours and shift the needed energy to avoid shortages in the afternoon and evening.

Due to the highly interconnected nature of the West, resource adequacy analysis that treats a particular RA planning footprint as an island can distort the observed RA challenges and may lead to suboptimal RA solutions, including potentially significant overbuild.

- **Failure to account for regional dynamics can result in overbuild and a misunderstanding of the nature of the resource adequacy challenge.**

  Due to the highly interconnected nature of the West, resource adequacy analysis that treats a particular RA planning footprint as an island can distort the observed RA challenges and may lead to suboptimal RA solutions, including potentially significant overbuild. Even without full West-wide planning coordination, RA programs may benefit from adopting market access policies that are informed by West-wide analysis in order to properly account for interregional operational interactions.

METHODOLOGICAL DEVELOPMENTS AND INSIGHTS

- **The Weather-Synchronized Simulation approach to RA analysis has a number of benefits over Monte Carlo Simulation. However, data availability remains a key limitation.**

  The Weather-Synchronized Simulation approach developed for this report can provide a transparent analysis of resource adequacy challenges and metrics and reveal how weather and weather trends impact those metrics. Unlike Monte Carlo Simulation that mixes and matches load, wind, and solar conditions that may be unrealistic, physically inconsistent, and internally inconsistent in terms of relative likelihood, the Weather-Synchronized approach tests historically coherent weather conditions. It can therefore provide a much more transparent assessment of the drivers of RA risk and, importantly, statistical models derived from the results of Weather-Synchronized Simulations can be directly applied to test future weather conditions derived from climate modeling. The availability of more historical data, in particular, hourly wind datasets, would improve the performance of those statistical models.

- **Weather is the most important driver of RA challenges and the treatment of weather trends is a key determinant of RA risk level.**

  A key finding of this study is that weather conditions continue to be the most important factor influencing the magnitude of RA challenges in the West in the near term. Simulating forced outage conditions and testing multiple hydro years is important to reduce the uncertainty in that estimate. Our analysis found that the treatment of weather trends is a key factor affecting the magnitude of the RA challenge. Without information about future weather from climate simulations, the choice of how to take weather trends into account is a policy decision. However, any analysis that assumes that historical weather is reflective of future weather patterns is likely flawed.

- **The availability of more high-resolution historical power system data as well as information about likely future weather conditions would greatly improve our understanding of the RA challenge.**

  Regardless of the RA analysis approach, the availability of more historical data on load, wind, and solar output would improve the accuracy of the analysis. In particular, the lack of hourly wind data at hub height for recent historical years (after 2014) is a key limiting factor for capturing the weather-driven correlations of these variables with a higher level of confidence. Generating more wind data is a high priority. Since weather is a key determinant of RA risk, and recent weather trends pose questions about the validity of relying on the historical record for assessing future RA risk, the availability of data for future weather can improve our

understanding of how to plan for reliable future energy systems.

NEXT STEPS AND FUTURE WORK

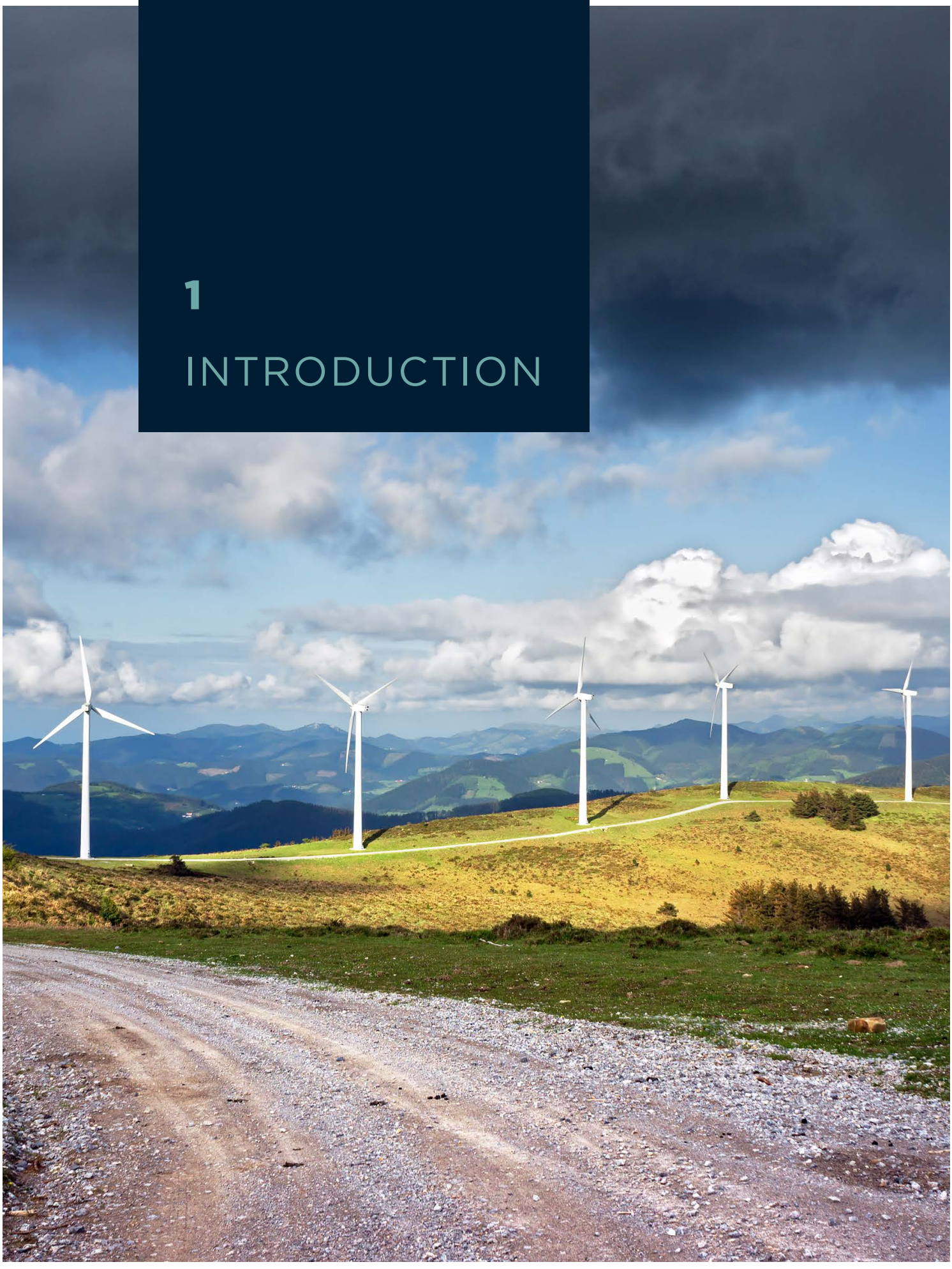
- **While the GridPath RA Toolkit provides a useful foundation for RA studies, climate sensitivities, electrification scenarios, and LSE or RA program modeling are key priorities for future RA analysis.**

As part of the GridPath RA Toolkit, we developed a methodological framework for RA modeling and a foundational RA dataset. We also identified several critical areas for future analysis. First, we highlight the importance of developing a deeper understanding of how weather patterns may change in the future, impacting loss-of-load risk. Second, in this study we focus on the near-term RA challenge, but longer term planning would require a more sophisticated treatment of the weather-sensitivity of loads in the context of increasing electrification. Third, layering ownership and contractual information onto the physical system dataset used in this study would make it possible to conduct LSE- or RA-program-specific analysis that is fully consistent with the broader regional dynamics in the West. Finally, a core purpose of this initiative was to develop an advanced, publicly available and transparent toolkit for resource adequacy analysis; our hope is that a broad set of users will leverage the GridPath RA Toolkit to further advance resource adequacy analysis for emerging power systems.

Finally, a core purpose of this initiative was to develop an advanced, publicly available and transparent toolkit for resource adequacy analysis; our hope is that a broad set of users will leverage the GridPath RA Toolkit to further advance resource adequacy analysis for emerging power systems.

1

INTRODUCTION



Recent years have seen a renewed focus on Resource Adequacy (RA) in the Western United States ("the West"), as aging coal plants begin to retire, and new technologies and policies are quickly changing the composition of the Western grid. Clean technologies are increasingly cost competitive, but also create new technical challenges for grid planners and operators tasked with maintaining system reliability. Amidst these changes, the West has also seen extreme weather in recent years, leading to both unusually high demand for electricity and exacerbation of risks to grid infrastructure. In response to recent challenges on the grid and in anticipation of continued change, multiple efforts across the region are being undertaken to ensure resource adequacy. In California, the California Public Utility Commission (CPUC) has authorized the procurement of 11.5 GW of new resources to address mid-term reliability needs arising from more extreme weather and the closing of the Diablo nuclear power plant.[19] Outside of CAISO, the Western Power Pool has convened a group of its members to design a resource adequacy program (the Western Resource Adequacy Program) that would allow participants to leverage diversity benefits of their loads and resources to meet resource adequacy objectives and to identify when regional resource adequacy challenges are afoot.[20]

The efforts that are underway to address resource adequacy in the West are critically important to maintaining reliability and continuing to transform the composition of the electric power sector, and their success will rely on sound RA analysis. Current approaches to RA analysis face many challenges, notably substantial data requirements, lack of transparency, and the potential for analytical inconsistency across the West, challenges which will grow as the complexities of the power sector—and the models used to simulate the power sector evolve. Modern RA approaches must grapple with the impact of weather on both demand and resource availability, the increasing reliance on energy-limited resources such as batteries, and the need to account for the ability of regional coordination to promote reliability through load and resource diversity in a highly interconnected system. At the same time, they must offer transparency into the methodologies, assumptions, and data used to allow for meaningful vetting by regulatory agencies and other oversight bodies. The Western United States, with a rapidly expanding fleet of variable renewables, a historical reliance on hydropower and a growing role for battery storage, and state clean energy policies that aim to continue to accelerate these trends, provides an instructive case study for modernized RA analysis.

The report describes a near term analysis of RA in the Western US using a new public tool for modern RA analysis—the GridPath RA Toolkit. Section 2 describes the data and methodologies employed by the Toolkit at a high level. Additional information can be found in the Technical Appendices. The Toolkit leverages GridPath, an open-source platform for power system planning and

19 CPUC Decision 21-06-035 in Rulemaking 20-05-003, 6/30/2021. Available at: https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M389/K603/389603637.PDF

20 More information can be found at: https://www.westernpowerpool.org/about/programs/western-resource-adequacy-program

optimization, and also includes the datasets and accompanying code required to run GridPath as an RA tool. In addition to promoting transparency, using an open-source tool in this application allows for continued development of model capabilities and customization to specific systems over time by the study team and by other organizations—as the nature of the resource adequacy challenge and the assumptions and simplifications necessary to assess resource adequacy are system-dependent and will change over time. In the initial phase of this work and the Western US case study described in this report, we focus on the following aspects of modern RA analysis:

- **Weather-driven relationships between loads and resource availability across the entire region.**

  Weather has always been a major driver of resource adequacy due largely to thermally driven electric loads, like air conditioning and electric heating, which tend to peak during the most extreme weather conditions. In systems with significant amounts of variable renewables, the weather is also a significant driver of resource availability.[21] In these systems, we are not only interested in understanding the hottest or the coldest days, but also days in which the weather conditions result in high loads and low resource availability. Probabilistic analysis for these systems requires not only an understanding of the distributions of the weather conditions that drive demand and resource availability (especially the tails of those distributions), but also the correlations between them.

  It can be challenging to appropriately account for these correlations even in systems with relatively limited geographic scope and a relatively straightforward relationship between weather and demand. However, it is even more challenging to account for all relevant correlations across a system as large as Western United States, where the weather experienced in one corner of the system may differ considerably from another and where the weather-sensitive electricity infrastructure, whether it be wind farms or baseboard heaters, varies widely across the region. It is not straightforward to identify the weather conditions that will pose the greatest challenges for resource adequacy in such a system nor is it straightforward to estimate their probabilities, but that is the central challenge of modern resource adequacy analyses.

21 Weather conditions can also affect the reliability of traditional thermal generation, a risk that was highlighted during the widespread outages in ERCOT in February of 2021. This study does not explore the specific failure modes that caused the outages in ERCOT, but we do address temperature-driven thermal derates and we note that weather-driven forced outage rates may be important for systems with a high reliance on thermal resources that are vulnerable to extreme weather.

- **Capabilities and constraints associated with dispatchable energy limited resources.**

  Hydropower and energy storage can make significant contributions to resource adequacy if their operations are optimized to account for load and the availability of renewables, and to prioritize avoiding lost load when the system is constrained. However, their contributions can also be limited due to energy constraints and losses. This study seeks to account for the dynamic capabilities of these resources in a manner that respects their key operating limitations. We employ a chronological dispatch approach to model not only the capacity adequacy of the system but its energy adequacy as well. This approach allows us to investigate the nature of potential RA events, including the magnitude and duration of shortages and the contributions of energy-limited solutions during these events. Note that changes to operational rules and incentive design may be required to ensure that these resources are best utilized to contribute to resource adequacy.

- **Transmission and regional coordination.**

  While much of the Western United States operates outside of a fully organized market, balancing areas across the West are highly interconnected and highly dependent on one another. During normal operations, entities regularly trade power across the West through bilateral agreements and, for some, the Western Energy Imbalance Market. During contingency events, like a sudden trip of a large generator, the stability of the grid is maintained by relying on generators in other parts of the region through formal pooling agreements, like the one managed by the Western Power Pool. And when an entity faces a potential energy and/or capacity emergency, a regional Reliability Coordinator notifies all market participants through the declaration of Energy Emergency Alerts, which can initiate additional short-term bilateral transactions to avoid load curtailments. In these ways, even though the West does not have a fully organized market, regional coordination is already critical to maintaining a reliable system.

  The extent to which resource adequacy analyses and planning studies acknowledge this interdependence across the West varies widely. Some entities attempt to estimate their reliance on the broader market during constrained conditions to avoid overbuilding resources, while others take a more conservative approach to ensure that their systems are adequate without reliance on the market. Because these assumptions are often driven by utility-specific considerations and subject to state

regulatory processes, there is no guarantee that they are internally consistent—that the same entities aren't assuming access to the same generation at the same time or that all available generation is being utilized where it's needed. This situation leads to inefficiencies and overbuild at best and reliability issues at worst.

This study provides an internally consistent analysis of the Western US to provide a holistic view of regional challenges and to inform more focused studies of individual utilities or RA programs. The study shows how transmission can be leveraged to mitigate regional resource adequacy challenges and how resource adequacy planners might take this into account when focusing on a particular part of the region with or without formal planning coordination.

To test the capabilities of the system across a wide range of system conditions, the GridPath RA Toolkit offers two simulation options: Monte Carlo Simulation or Weather-Synchronized Simulation. Monte Carlo Simulation allows the user to test more possible combinations of conditions, but may not preserve all relevant correlations and may not reflect physically realistic situations. Weather-Synchronized Simulation ensures that weather conditions are realistic and reflect all relevant correlations, but typically limits the number of weather conditions that can be tested due to limited data availability. In this report, we explore both methodologies.

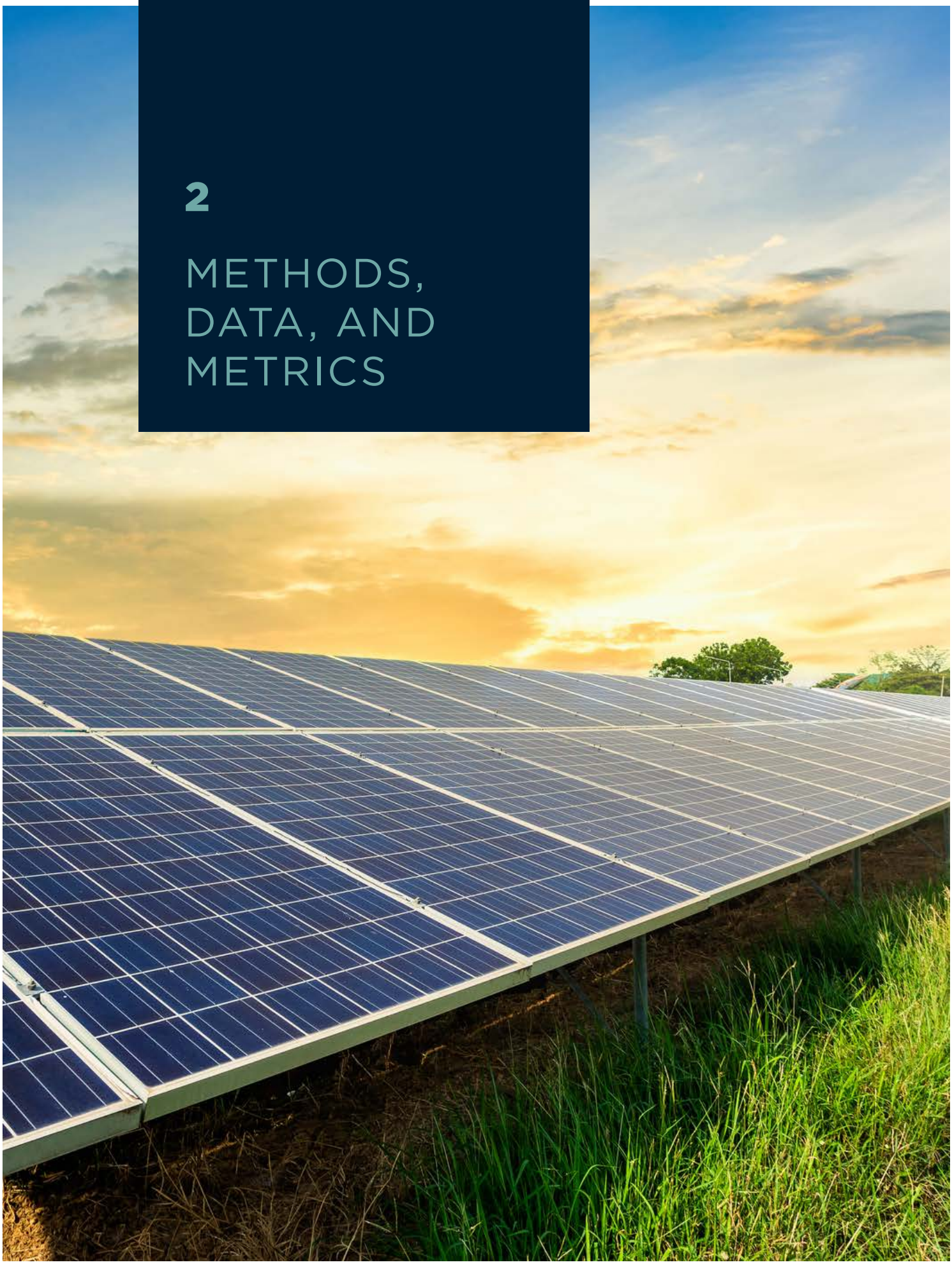
In **Section 3**, we use Monte Carlo Simulation to explore investigate multiple scenarios for the Western US in 2026: a No Additions Scenario representing the system if no action is taken; a CPUC Additions Scenario that includes procurement of new resources recently authorized in California; and a Less Coal scenario that includes those capacity additions but also retires a portion of the coal fleet in the rest of the West. In this part of the report, we seek to understand the nature of near term reliability challenges, their magnitude, timing, and duration as well as regional interactions when the system is highly constrained.

Section 4 of this report further explores the No Additions Scenario using Weather-Synchronized Simulation to better understand RA challenges based on historical weather conditions. In this section, we describe the historical weather events that pose the greatest RA risk and we present a novel statistical analysis relating key weather variables to loss of load probability and demonstrate how such analyses can be used to examine historical weather trends and associated RA risk.

Finally, in **Section 5**, we summarize key findings and offer recommendations for future work using the GridPath RA Toolkit or other RA modeling tools.

2

METHODS, DATA, AND METRICS



RA analyses typically consist of two primary components: an analysis of the relevant conditions that could pose RA challenges and an evaluation of the capabilities of the system during those conditions to avoid lost load. Depending on the system, approaches can range from highly simplified methods, like a traditional planning reserve margin (PRM) calculation, to highly complex and computationally challenging simulations. The complexities of the Western grid and modern power systems necessitate a rigorous approach that can account for complex weather patterns over large areas and operational decision-making that considers loads and resources over that entire area and across time. The GridPath RA Toolkit offers two options for modeling system conditions to address these complexities (Monte Carlo Simulation and Weather-Synchronized Simulation) and it leverages the dispatch optimization capabilities of GridPath to estimate system operational capabilities under those conditions. These tools require detailed information about loads, resources, and transmission on the system and multiple years of high resolution data. In this section, we describe: the methods that can be used to characterize a wide range of potential system conditions using the GridPath RA Toolkit; the constraints and objective function used in GridPath to simulate system capabilities under constrained conditions in this study; and the assumptions and methodologies used to transform public data into a usable format for RA analysis. More detailed information, including benchmarking analyses for the resource availability estimates, can be found in Appendix A.

The GridPath RA Toolkit offers two options for modeling system conditions to address these complexities (Monte Carlo Simulation and Weather-Synchronized Simulation) and it leverages the dispatch optimization capabilities of GridPath to estimate system operational capabilities under those conditions.

2.1 MODELING SYSTEM CONDITIONS

There are three primary sources of resource adequacy risk addressed in this study: weather risk, hydro risk, and forced outage risk. Accounting for these risks, in particular weather risk, is inherently challenging because there is limited historical data with the necessary spatial and temporal granularity to simulate grid operations and because the weather conditions that drive resource adequacy challenges tend to be rare. The GridPath RA Toolkit offers two options for modeling system conditions to address these risks, both of which are described in this section: Monte Carlo Simulation and Weather-Synchronized Simulation.

2.1.1 MONTE CARLO SIMULATION

As with many modern resource adequacy analyses, this study primarily relies on Monte Carlo Simulation to get around the fundamental issue of historical data availability. Figure 2.1 shows the data coverage for the timeseries data used to characterize weather and hydro risks in this study. While several years of hourly load, temperature, and solar data are publicly available from FERC and the National Renewable Energy Laboratory's (NREL) National Solar Radiation Database (NSRDB), hourly wind data at wind turbine hub heights is limited to the NREL Wind Integration National Dataset Toolkit (WIND ToolKit), which is available for 2007-2014. This severely limits the number of weather conditions that can be tested without using Monte Carlo Simulation, and importantly, it excludes the most recent years, which have seen extreme weather events.

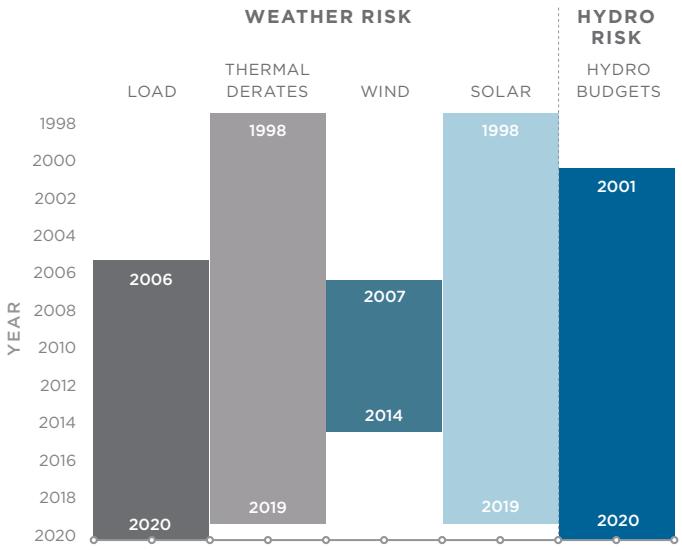


FIGURE 2.1.
Historical data availability.

When coherent data is limited in this manner, Monte Carlo Simulation, applied carefully, can be used to synthesize many years of random, but plausible, conditions using historical data. Figure 2.2 shows how RA analysis is conducted with the GridPath RA Toolkit in Monte Carlo Simulation mode.

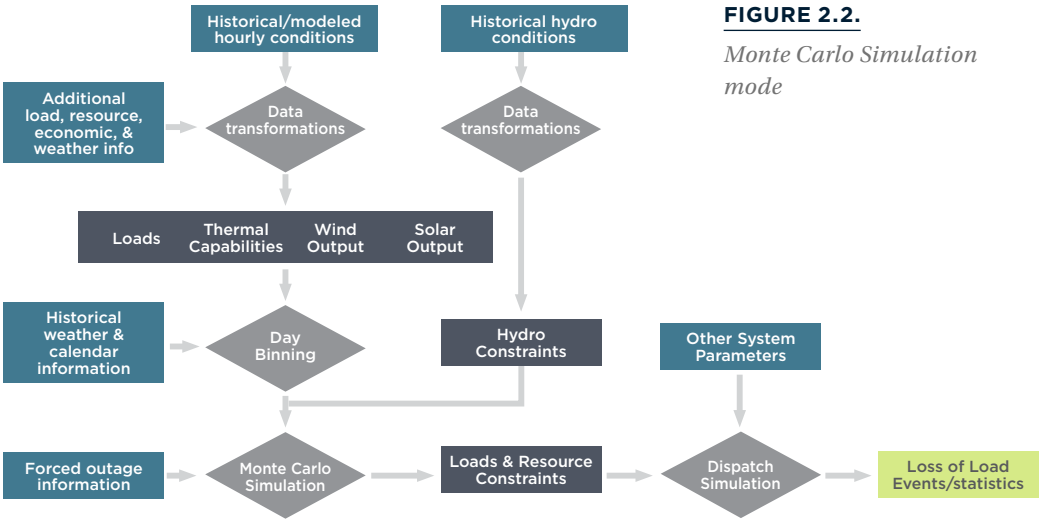


FIGURE 2.2.

Monte Carlo Simulation mode

The GridPath RA Toolkit offers two options for modeling system conditions to address these complexities (Monte Carlo Simulation and Weather-Synchronized Simulation) and it leverages the dispatch optimization capabilities of GridPath to estimate system operational capabilities under those conditions.

To represent each year, the Monte Carlo Simulation first randomly selects the hydro conditions for the year and then loops through the days of the year, randomly selecting weather-driven hourly load, wind, solar, and thermal shapes. To account for correlations between these variables due to weather, the model relies on daily weather binning. Days with similar weather conditions across the West are grouped together into bins[22] and the model uses a Markov Chain approach to randomly walk between weather bins based on historically-observed weather day transitions between 1991 and 2020. More information about the binning methodology can be found in Appendix B. After the weather bin is selected for each day, hourly load, wind, solar, and thermal shapes are randomly selected from within the bin. Mixing and matching these shapes from within the same weather bin allows the Monte Carlo method to synthesize many more potential system conditions than were actually recorded over the historical period from which the conditions are drawn. This approach captures some, but not all correlations between the variables—the weather binning attempts to group days in a manner that captures correlations, but any correlations between variables *within* each bin are neglected. Smaller bins tend to do a better job of capturing correlations, because the days are more similar to one another, but this reduces the number of potential combinations that can be investigated.

22 The bins also differentiate between months and load shape binning also takes into account weekends versus weekdays.

This is the key tradeoff for using Monte Carlo Simulation to generate weather conditions in RA analysis—the more weather conditions that Monte Carlo Simulation allows you to test (i.e., the greater the precision), the less confidence you have that the important correlations are preserved (i.e., the lower the accuracy).

As an alternative to Monte Carlo Simulation, the GridPath RA Toolkit includes an option to generate conditions using Weather-Synchronized Simulation, which considers only conditions that have been experienced in the historical record or that could be experienced coincidently across the study footprint based on high resolution weather simulation.

To preserve geographical correlations within each variable, daily shapes of each type are selected on the same day over the entire geographical footprint. For example, to synthesize a very hot August day, the Monte Carlo Simulation could randomly select August 30th, 2017 for solar and August 9th, 2012 for wind—two days with relatively similar weather conditions across the West. In this case, the Monte Carlo Simulation combines the solar shapes experienced across the entire West on August 30th, 2017 with the wind shapes experienced across the entire West on August 9th, 2012.

The Monte Carlo Simulation also randomly generates forced outages on a unit-specific basis for each hour using exponential failure and repair models. For this study, unit forced outages were independent and not weather-driven, however this capability could be developed for future studies.

The primary benefit of a Monte Carlo Simulation is the very large number of conditions that can be investigated, which can result in highly precise estimates of RA metrics. However, their accuracy can be questionable because the synthesized conditions may not be physically realistic due to the complexities of weather systems and their impacts on load and resource availability. As a result, Monte Carlo approaches may yield false precision and result in overconfidence by decision makers.

2.1.2 WEATHER-SYNCHRONIZED SIMULATION

As an alternative to Monte Carlo Simulation, the GridPath RA Toolkit includes an option to generate conditions using Weather-Synchronized Simulation, which considers only conditions that have been experienced in the historical record or that could be experienced coincidently across the study footprint based on high resolution weather simulation.[23] In this approach (see Figure 2.3), Monte Carlo analysis is only applied to simulate forced outages and not to generate plausible weather-driven variables. To capture hydro risk, each of the hydro years is tested across all of the synchronized weather years.

23 For this study, we rely on historical weather conditions, but we note that a similar approach could be used with the results of numerical weather modeling of alternative or potential future conditions.

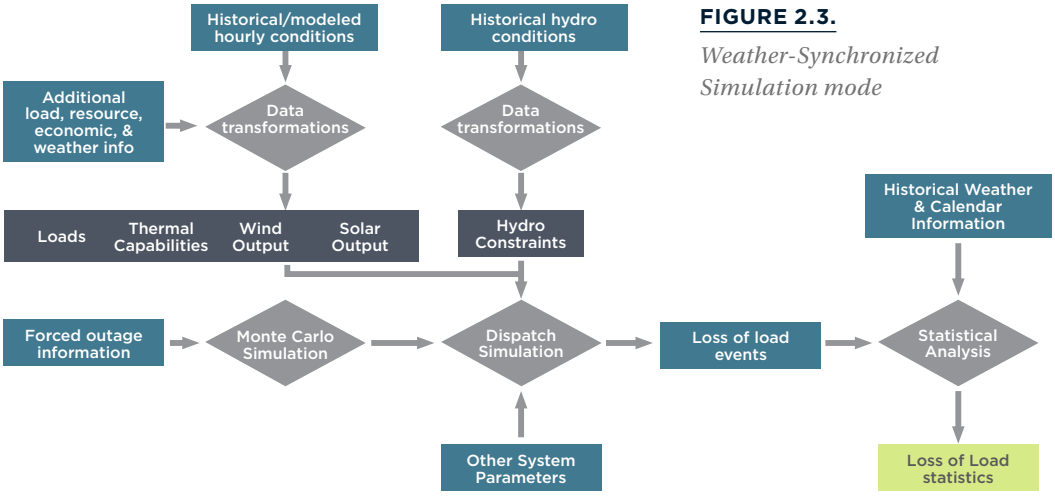


FIGURE 2.3.

Weather-Synchronized Simulation mode

Weather-Synchronized Simulation offers three primary benefits over Monte Carlo Simulation for characterizing weather-dependence:

1. It provides confidence that the conditions being modeled in the dispatch simulation reflect physically feasible weather conditions with all relevant temporal and geographical correlations over the study area, and preserves the relative frequency of different weather conditions;
2. It provides more transparency into the derivation of resource adequacy metrics; and
3. It allows for a relatively simple and transparent analysis into how weather and weather trends may impact resource adequacy.

The key disadvantage of this methodology is that publicly available coherent load, wind, thermal, and solar data with the granularity necessary to simulate operations across the Western grid is limited. Specifically, because the NREL Wind Toolkit data only spans 2007-2014, there are only 8 years of synchronized weather conditions available to test using public data, and the most recent years, which have experienced extreme weather events, cannot be tested. For the purposes of demonstrating the Weather-Synchronized Simulation approach, we chose to synthesize thermal, wind, and solar data using day matching to expand the coherent weather record to 2007-2020. More information about this approach can be found in Appendix C.

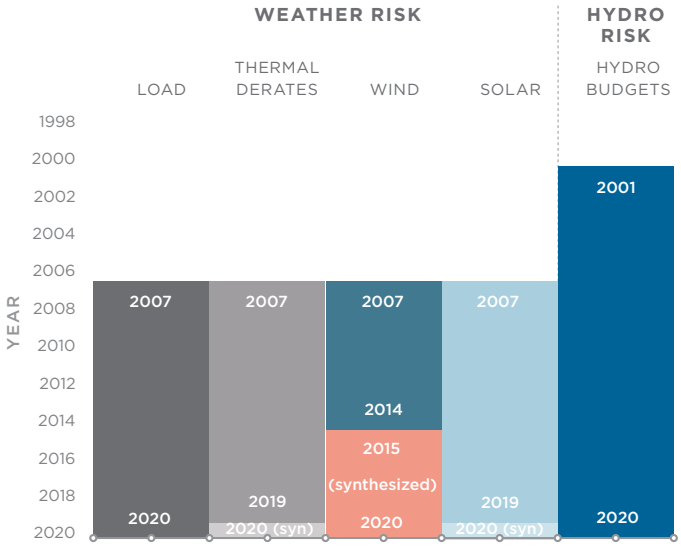


FIGURE 2.4.

Historical and synthetic data used in Weather-Synchronized Simulation mode.

For a near term study, synthesizing hourly wind, thermal, and solar data for some conditions may be reasonable, as load remains the primary driver of near term RA challenges and load data is available through 2020. However, for systems with a greater reliance on renewable resources, this approach may not be valid. We address this limitation in Section 4 and find that expansion of publicly available wind datasets in particular will be crucial for more robust RA modeling going forward.

2.2 DISPATCH MODELING

To test the capabilities of the system across a wide range of chronological conditions and to identify the conditions under which the system cannot reliably meet demand, the Toolkit leverages optimization-based dispatch modeling using GridPath, an open-source, versatile platform for power system planning and optimization. This approach accounts for the dynamic capabilities and limitations of energy-limited resources, like hydropower, energy storage, and hybrid renewable + storage systems, as well as the benefits of the transmission system.

To test the capabilities of the system across a wide range of chronological conditions and to identify the conditions under which the system cannot reliably meet demand, the Toolkit leverages optimization-based dispatch modeling using GridPath, an open-source, versatile platform for power system planning and optimization. This approach accounts for the dynamic capabilities and limitations of energy-limited resources, like hydropower, energy storage, and hybrid renewable + storage systems, as well as the benefits of the transmission system.

For the purposes of this study, we vastly simplify the dispatch problem within GridPath to prioritize avoiding lost load and to ensure relatively short runtimes. However, GridPath is highly flexible and can be used to layer additional complexities onto the dispatch problem, including resource economics and flexibility constraints, although this could increase computational resource requirements. For this study, GridPath is used to perform weekly optimizations of a zonal system with an hourly resolution over the set of generated system conditions. Each week is optimized independently and the model has perfect foresight for the whole week. This chronological dispatch approach allows us to capture the characteristics and capabilities of energy-limited resources such as energy storage and hydropower. The objective function used in this study minimizes total unserved energy over the week plus the largest hourly capacity shortage experienced during the week, thus taking into account both the energy and capacity components of resource adequacy. The dispatch of all resources is co-optimized in order to account for interactions between them, for example, between solar and storage.

Transmission is represented by constraining flows across branches that connect BAAs using a simple transport model and enforcing path limits across collections of branches. The objective function also includes a small penalty term on exports that ensures that each load zone prioritizes meeting its own load first.

In addition to meeting load, the model enforced contingency reserve obligations, which can be met by a subset of the modeled resources. Each modeled resource is also subject to various operational constraints, which are summarized in Table 2.1 and described in more detail in Appendix A.

TABLE 2.1.

Resource dispatch treatment.

RESOURCE TYPE	DISPATCH CONSIDERATIONS
Thermal Generators	• Always committed if available • Able to provide contingency reserves • Aggregated by technology and zone in optimization
Hydropower	• Output optimized to minimize unserved energy, subject to weekly hydro budget and corresponding min and max constraints • Able to provide contingency reserves • Aggregated (across dispatchable, run-of-river, and pumped storage) by zone in optimization
Wind Power	• Fixed shape • Aggregated by zone in optimization
Solar Power	• Fixed shape • Aggregated by zone in optimization
Battery Storage	• Charging and discharging optimized to minimize unserved energy, subject to battery constraints • Able to provide contingency reserves • Aggregated by zone in optimization
Hybrid Renewables	• Output optimized to minimize unserved energy, subject to renewable availability, battery constraints, and interconnection limits • Not aggregated in optimization (i.e., each hybrid project is explicitly optimized)

The GridPath RA Toolkit does not currently incorporate resource economics into the model formulation, however this could be explored in future work. The implicit assumptions behind neglecting economics are that resources will be committed in advance if they are available when operators see forecasted extreme weather conditions and that during emergency conditions, market and bilateral channels exist to ensure that all necessary transactions will occur and all available resources will contribute.

In this study, flexibility-related constraints and considerations such as ramp rates, minimum up and down times, and forecast errors are also neglected. However, given sufficient computational resources, these factors can be considered by GridPath. Enforcing flexibility-related constraints could result in additional unserved energy due to flexibility inadequacy. These various simplifications to our dispatch approach are what facilitate, computationally, a region-wide analysis that incorporates both probabilistic modeling and inter-regional interactions.

2.3 STUDY AREA AND TRANSMISSION TOPOLOGY

The study area includes all BAAs in WECC, excluding those in Canada and Mexico. The transmission topology, line ratings, and path limits are based on the 2026 WECC Common Case database,[24] which contains a nodal representation of the WECC transmission network, including individual branches, and groups of branches, called interfaces, across which path limits and other constraints are applied.

The solid lines in Figure 2.5 represent inter-zone transmission links for which some or all of the modeled flow is constrained by one or more interface limits. We also apply summer derates on inter-zone transmission capacities based on data from the 2026 WECC Common Case database. The final study results do not incorporate transmission outages. Early tests with significant transmission derates and large transmission path outages had minimal impacts on West-wide RA findings due to the high degree of redundancy on the high voltage transmission system. However, we flag transmission risk for subregional studies and especially for systems in which transmission outages could jeopardize access to large dispatchable resources as important areas of future work.

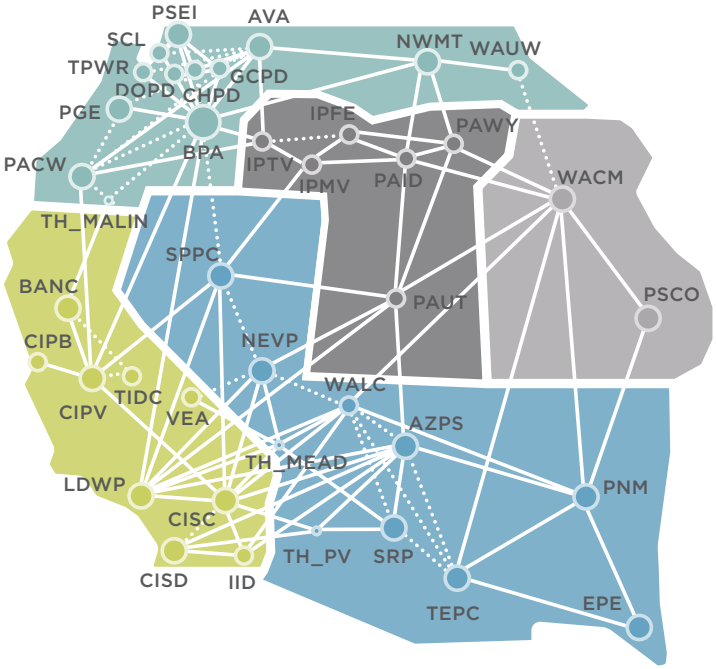


FIGURE 2.5.
Western US Case Study zonal topology.

We also note that reliance on publicly available data places some limitations on the types of conclusions that can be drawn with this study, particularly with respect to the zonal topology. This study is based on a physical representation

24 Available at: https://www.wecc.org/SystemAdequacyPlanning/Pages/Datasets.aspx

of the Western US grid and cannot provide insight regarding the resource adequacy position of an individual load serving entity (LSE) or collection of LSEs, because it does not contain contractual information. A shortage shown to be experienced in a particular BAA as modeled here may not be attributable or in actuality experienced in that BAA. The findings of this report and the associated data should therefore be interpreted as broadly indicative of regional resource adequacy positions, but not precise with respect to individual balancing area authorities (BAAs), Load Serving Entities (LSEs), or RA programs. Further insights could potentially be gained through the use of proprietary data and/or additional system-specific information. The information and methodologies provided in this study are intended as a foundation upon which such analyses could be pursued.

This dataset was developed for the Western US for the year 2026 and is made available as part of the GridPath RA Toolkit.

2.4 LOADS AND RESOURCE AVAILABILITY

This study relies on a large dataset that describes loads and resource availability across a wide range of potential conditions. This dataset was developed for the Western US for the year 2026 and is made available as part of the GridPath RA Toolkit. In the sections that follow, we provide high-level descriptions of the methodologies employed to transform the public data into datasets that can be used in RA analysis. Because the study footprint is so large and requires consistent data across numerous utilities and resources, we selected data sources and methodologies that could be applied universally across the entire footprint in a consistent manner. This necessitated some simplifications and approximations, which we describe in the following sections and in Appendix A. Appendix A also includes additional methodological details and benchmarking exercises.

2.4.1 LOADS

Load shapes were developed to estimate the hourly demand for electricity in the study year (2026) in each BAA across several years of historical weather conditions. The shapes were derived based on historical hourly loads from FERC Form 714,[25] historical weather conditions across the West, historical economic conditions, and forecasted economic conditions in the study year. This resulted in hourly loads for 2026 that corresponded to 15 years of weather conditions (2006-2020). Other RA studies have used more weather years to probe a potentially wider range of conditions. However, these studies must

25 This study utilized the 2006-2020 Form 714 database, available at: https://www.ferc.gov/sites/default/files/2021-06/Form-714-csv-files-June-2021.zip

grapple with the question of whether a longer historical record is necessarily more accurate given recent weather trends. In Section 4.3, we use weather statistics over a much longer historical record to delve into this topic in more depth and find that using a recent historical load record generally provides for more conservative estimates due to the higher frequency of extreme weather events in recent years.

Figure 2.6 shows the CAISO demand during two extreme weather events—late July of 2006 and mid-August in 2020. In each case, the figure compares the historical demand to the demand transformed to 2026, the study year.

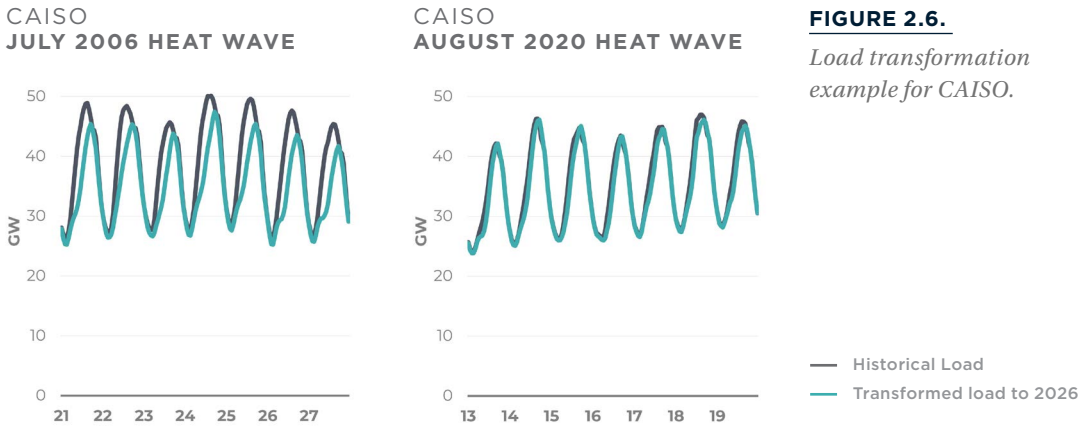


FIGURE 2.6.
Load transformation example for CAISO.

At a high level, historical trends suggest that CAISO demand may tend to be lower than it was in the historical record if the same weather conditions were to be experienced in the study year. This is due primarily to the expansion of behind-the-meter PV, which has significantly decreased demand during daylight hours in recent years. Intuitively, the transformation has a much larger impact on the demand associated with the July 2006 heat wave than that associated with the 2020 heat wave because there were significant changes in CAISO demand between 2006 and 2020, especially during daylight hours.

In contrast to the trends in CAISO, other parts of the West see exacerbated load conditions during historical weather events when transformed to the study year. Figure 2.7 compares the historical and transformed loads in Avista's BAA during the same weather events. In Avista, loads peak at higher levels and later in the day during heat events in more recent years, perhaps due to more widespread adoption of air conditioning. This trend results in increased demand during evening hours across both the 2006 and 2020 heat events. As with CAISO, the impact of the transformation on Avista's load is greater for the 2006 heat wave than the 2020 heat wave due to the changes in the load shape that were experienced between 2006 and 2020.

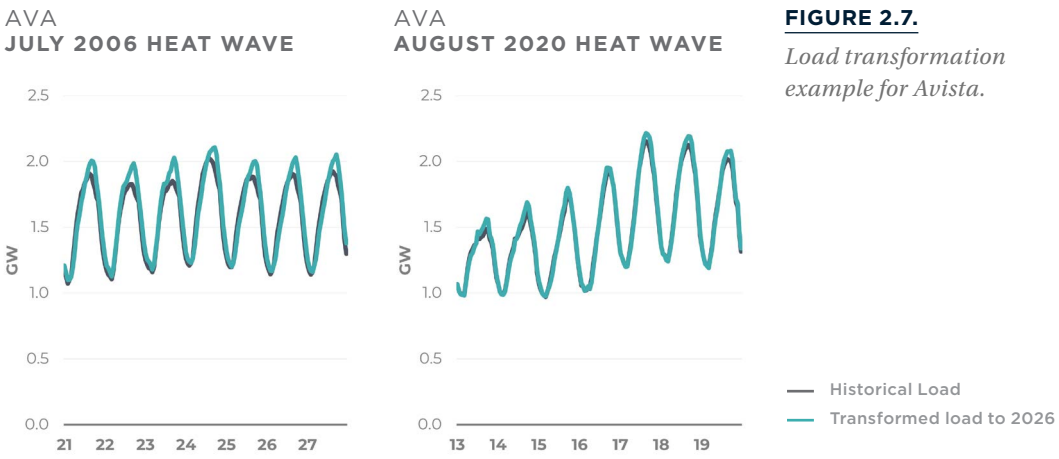


FIGURE 2.7.
Load transformation example for Avista.

We note that the load shapes used in this study diverge from utility forecasts, especially where historical trends in load shapes may not be indicative of the future. While the load shapes in this study may provide a reasonable characterization of load for a West-wide near-term RA analysis, more sophisticated methodologies that leverage additional information may be required for longer-term studies, where for example, electrification may significantly alter load shapes, or for studies that focus on an individual BAA or LSE.

2.4.2 DEMAND-SIDE RESOURCES

As described in the previous section, the impacts of behind-the-meter PV are embedded within the load shapes in this study and the load shape transformation estimates the impact of these resources continuing to grow over time based on their historical growth. To the extent that other demand-side resources, including demand response and other behind-the-meter generation and storage, have influenced the historical metered load, their estimated impacts in the study year are also embedded within the transformed load shape. In this way, demand-side resources are treated implicitly by the model. This is likely a conservative approach that captures the weather-sensitivity of demand-side resource availability, but may underestimate the impacts of demand-side resources in the future should they be adopted at rates that far exceed historical trends. More explicit modeling of weather-driven demand-side resource availability, especially for longer term scenarios, remains a priority for future studies.

2.4.3 GENERATION AND STORAGE

Several data sources and methods were employed to estimate the availability of generation and storage. In this section, each technology is summarized with a brief table describing the risks captured by the study, any necessary data transformations, and the data sources. More detailed information, including benchmarking information, can be found in Appendix A.

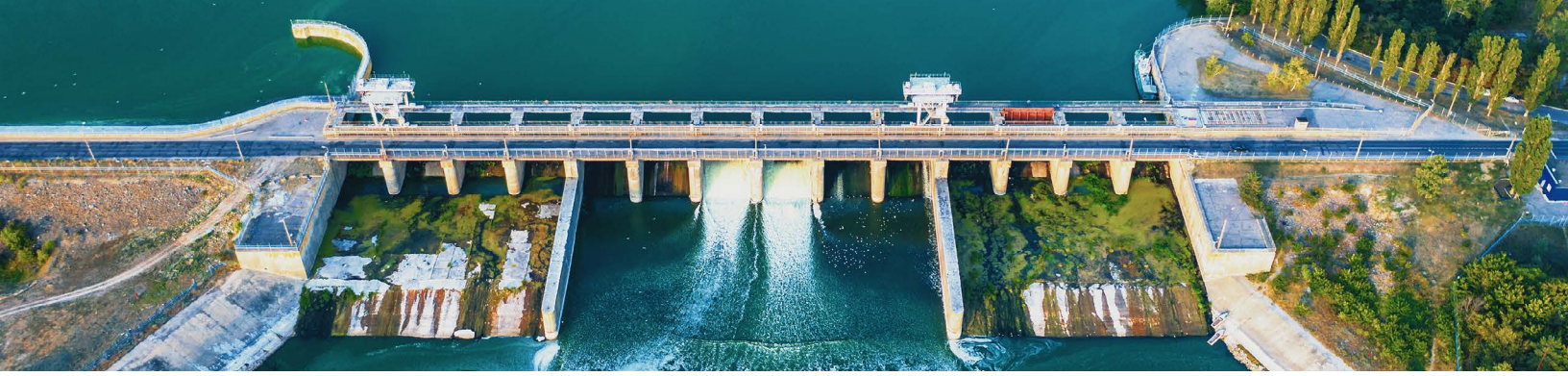


TABLE 2.2.

Thermal generator key inputs, assumptions and sources.

THERMAL GENERATORS	
Risks Analyzed	• Random forced outages (uncorrelated and weather-independent) • Temperature-based capacity derates
Data Transformation Methods	• Exponential failure and repair models tuned to forced outage rate (FOR) and mean time to repair (MTTR) information • Project-specific temperature derate functions derived from summer and winter capacities and site-specific historical temperature data
Data Sources	• Summer and winter capacities from EIA Form 860M (February 2021)[26] • Generator-specific FOR and MTTR information from 2026 WECC Common Case where available, averaged by technology type where not available • Site-specific hourly temperature data from the National Solar Radiation Database (NSRDB) (1998-2019)[27]

TABLE 2.3.

Hydropower key inputs, assumptions and sources.

HYDROPOWER	
Risks analyzed	• Interannual and intra-annual variation in hydro availability due to different water years and seasonal schedules • Forced outages (implicitly)
Data Transformation Methods	• Weekly aggregated hydro budgets estimated based on historical monthly hydro generation for each BAA • Aggregations include dispatchable, run-of-river, and pumped storage • Monthly hydro budgets for each BAA were adjusted to account for historical trends (detrended and extrapolated to the study year) • Min and max constraints were estimated as functions of the hydro budget based on historical aggregated hourly hydro generation data where available
Data Sources and Coverage	• Historical plant-specific monthly hydro generation from EIA Form 923/906 (2001-2020)[28] • Historical aggregated hourly hydro generation from the BPA Total Load and Wind Generation Report (2007-2020),[29] CAISO Daily Renewables Watch (April 2010-Jan 2021),[30] and WECC (2005) for estimation of min and max constraints

26 This study was based on the February 2021 EIA Form 860M, available at: https://www.eia.gov/electricity/data/eia860m/archive/xls/february_generator2021.xlsx

27 Sengupta, M., Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby. 2018. "The National Solar Radiation Data No Additions (NSRDB)." Renewable and Sustainable Energy Reviews 89 (June): 51-60.

28 Available at: https://www.eia.gov/electricity/data/eia923

29 Available at: https://transmission.bpa.gov/business/operations/wind/ (item 5)

30 Available at: http://www.caiso.com/market/Pages/ReportsBulletins/RenewablesReporting.aspx

TABLE 2.4.

Wind power key inputs, assumptions and sources.

WIND POWER	
Risks analyzed	• Hourly weather variability • Forced outages (implicitly)
Data Transformation Methods	• Plant-specific empirical logistic power curves derived based on historical performance where available, assumes 5% forced outage derate in all hours • Where historical generation data was not available, estimated performance based on similar projects
Data Sources and Coverage	• Modeled site-specific hourly 80m wind speed from NREL Wind ToolKit (2007-2014)[31] • Historical plant-specific monthly generation from EIA Form 923-906 (2007-2020)

TABLE 2.5.

Solar power key inputs, assumptions and sources.

SOLAR POWER	
Risks analyzed	• Hourly weather variability
Data Transformation Methods	• NREL System Advisor Model (SAM)[32] with project-specific parameters estimated based on technology (PV vs. solar thermal), location, and COD
Data Sources and Coverage	• Site-specific hourly weather data from NSRDB (1998-2019) • Technological specifications from LBNL's Utility-Scale Solar Data Update: 2020 Edition[33] and EIA Form 860M

TABLE 2.6.

Battery storage key inputs, assumptions and sources.

BATTERY STORAGE	
Risks analyzed	• Forced outages
Data Transformation Methods	• None
Data Sources and Coverage	• Duration information for individual battery projects was collected from various project websites, press releases, and news articles

31 Draxl, C., B.M. Hodge, A. Clifton, and J. McCaa. 2015. Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit (Technical Report, NREL/TP-5000-61740). Golden, CO: National Renewable Energy Laboratory. Draxl, C., B.M. Hodge, A. Clifton, and J. McCaa. 2015. "The Wind Integration National Dataset (WIND) Toolkit." *Applied Energy* 151: 355366. King, J., A. Clifton, and B.M. Hodge. 2014. Validation of Power Output for the WIND Toolkit (Technical Report, NREL/TP-5D00-61714). Golden, CO: National Renewable Energy Laboratory.

32 PySAM Version 2.2.2. National Renewable Energy Laboratory. Golden, CO. https://github.com/nrel/pysam.

33 Bolinger, M., Seel, J., Robson, D., Warner, C., "Utility-Scale Solar Data Update: 2020 Edition,"Lawrence Berkeley National Laboratory, November 2020. Available at: https://emp.lbl.gov/publications/utility-scale-solar-data-update-2020

TABLE 2.7.

Hybrid renewable key inputs, assumptions and sources.

HYBRID RENEWABLES	
Risks analyzed	• Hourly weather variability • Forced outages for battery systems
Data Transformation Methods	• Same as listed above for wind and solar availability • Hybrid solar availability assumes an inverter loading ratio (ILR) or 1 to avoid suboptimal clipping
Data Sources and Coverage	• Same as listed above for wind, solar, and battery systems

2.5 METRICS

RA metrics have been the subject of increasing discussion and scrutiny as the complexity of the RA problem has grown in recent years.[34] There are a number of different ways to describe the likelihood of lost load in an RA study and these different metrics and different variations on terminology can make it difficult to compare across studies. In this study, we do not propose new or different RA metrics or standards, but we report findings across a wide range of metrics in the interest of transparency and improved understanding. This section provides descriptions of those metrics that are referenced throughout the report.

In this study, we do not propose new or different RA metrics or standards, but we report findings across a wide range of metrics in the interest of transparency and improved understanding.

Loss of Load Probability (LOLP). An $LOLP$ is the probability of encountering lost load over a given period of time, or the probability of encountering a loss of load *event*, where an *event* can take on any number of definitions. The Northwest Power and Conservation Council has traditionally planned for a 5% $LOLP$, where an event is defined as a year in which lost load is encountered at some point in time, regardless of how many times it is encountered within that year. In this study, we denote this convention, which describes the probability that lost load will be encountered on any given year, as $LOLP_{year}$. Another common LOLP convention is to count events as days in which lost load occurs—regardless of how many times lost load occurs within that day. This convention, which we denote $LOLP_{day}$, represents the probability that lost load will be encountered on any given day. And finally, one can define an event as any hour in which lost load is observed and determine the corresponding $LOLP_{hour}$. This hourly $LOLP$ convention is more common for

34 For more background and discussion regarding RA metrics, we refer you to the ESIG Redefining Resource Adequacy for Modern Power Systems report, which can be found at https://www.esig.energy/resource-adequacy-for-modern-power-systems/.

studies that use convolution methods (versus chronological simulation, as in this study), because they cannot characterize multi-hour events.

Loss of Load Expectation (LOLE). Resource adequacy standards and study findings are often reported in terms of $LOLE$ instead of $LOLP$, but the two metrics are related. $LOLE$ metrics generally represent the expected number of events that are encountered over a given length of time. The $LOLE$ can be calculated as the corresponding $LOLP$ (depending on the event definition) times the length of time of interest. For example, the most common $LOLE$ standard is the *one-day-in-ten year* standard. To calculate the $LOLE$ in a manner consistent with this standard, you multiply the $LOLP_{day}$, the probability that any given day will see lost load, by the number of days in a ten-year period:

$$LOLE = LOLP_{day} \times 365 days/yr \times 10 yrs$$

Loss of Load Hours (LOLH). The $LOLH$ is a special case of an $LOLE$ metric—it is typically expressed as the expected number of hours per year that encounter lost load. The $LOLH$ is often used in studies where it is not possible to count multi-hour events because of the use of a convolution-based model. These studies sometimes adopt an $LOLH$ standard that reinterprets the one-day-in-ten-year standard as 24 hours in 10 years, or an $LOLH$ of 2.4 hours per year. The $LOLH$ can be calculated from the corresponding $LOLP$ in the following manner:

$$LOLH = LOLP_{hour} \times 8760 hrs/yr$$

Expected Unserved Energy (EUE). The expected unserved energy typically represents the average amount of energy shortages that will be experienced in a given year. It is calculated by summing all of the shortages identified in a simulation and dividing by the number of simulated years. We also report the normalized EUE, EUE_n, which is the EUE divided by the average annual load. The EUE_n is typically very small and is reported in parts per million (ppm) by multiplying by 10^6.

We also include metrics that provide more insight about the nature of lost load events. We denote the average energy shortage experienced on a day with lost load as and the average shortage experienced in an hour with lost load as $EUE^{\frac{day}{event}}$. These metrics provide a sense of the typical size (in terms of both energy and capacity, respectively) of shortages when they are experienced. They can also be calculated from the other loss of load metrics using the following formulas:

$$EUE^{\frac{day}{event}} = \frac{EUE}{LOLP_{day} \times day \times 365\ days/yr}$$

$$EUE^{\frac{hour}{event}} = \frac{EUE}{LOLH}$$

Average Event Duration. When multi-hour loss of load events can be described by an RA model, as is the case in this study, the average duration of such events can be derived from the metrics described above. If an event is defined as a day with lost load, then the average event duration is the average number of hours of lost load within a day that experiences lost load, which can be calculated as:

$$Average\ event\ duration = \frac{EUE^{\frac{day}{event}}}{EUE^{\frac{hour}{event}}}$$

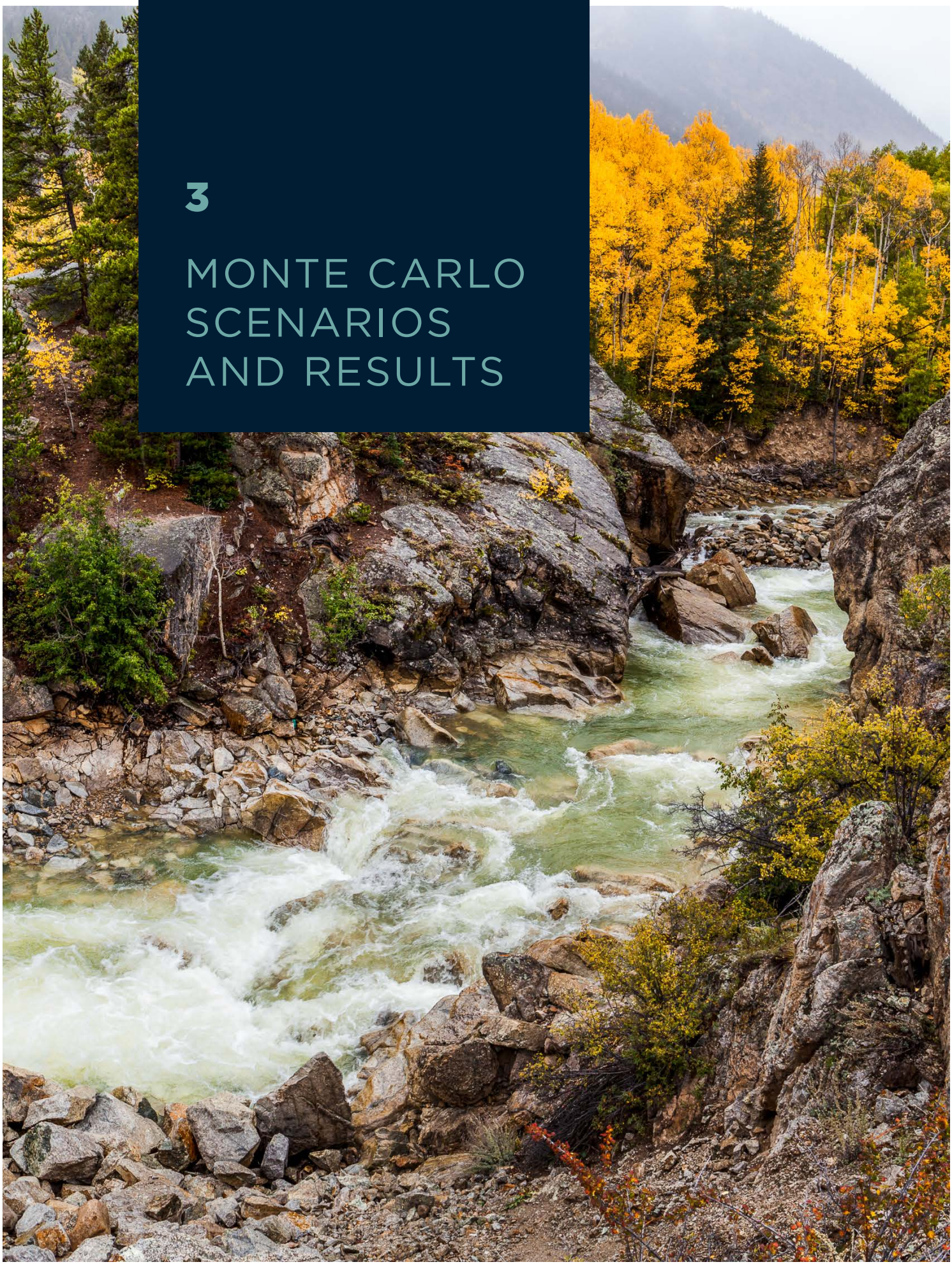
The resulting hours may be continuous or non-continuous throughout the day with this definition.[35]

In addition to these statistics, this report also includes statistical distributions and heatmaps to describe the timing, duration, and size of loss of load events.

35 While there may be interest in defining events as continuous periods of lost load (of which multiple can occur in one day), care must be taken to avoid over-interpreting simulation results in this manner, especially where degeneracies (i.e., different, but equally valid solutions) are possible. For example, an optimization model might not be able to differentiate between a solution where unserved energy is experienced across 12 continuous hours in a day (1 event with a 12-hour duration) versus a solution where it is experienced in alternating hours across the day (12 events, each with 1-hr duration). For this reason, this study focuses on the energy and capacity shortages experienced across a day, rather than individual continuous periods of lost load within the day.

3

MONTE CARLO SCENARIOS AND RESULTS



This study explores three scenarios in the 2026 time frame using Monte Carlo analysis, each using a different resource portfolio: a No Additions Scenario, which reflects planned retirement and no planned resource additions; a California Additions scenario, which layers on additional resources in CAISO based on the California Public Utility Commission's (CPUC) Preferred System Plan; and a Less Coal scenario, which incorporates the additional resources in California while also retiring about 11 GW of coal elsewhere in the West. All three scenarios are modeled as a physical system that is unencumbered by institutional barriers to coordination and where short-term operational decisions can be made optimally to avoid lost load. The following sections further describe each scenario and summarize the findings.

3.1 NO ADDITIONS SCENARIO

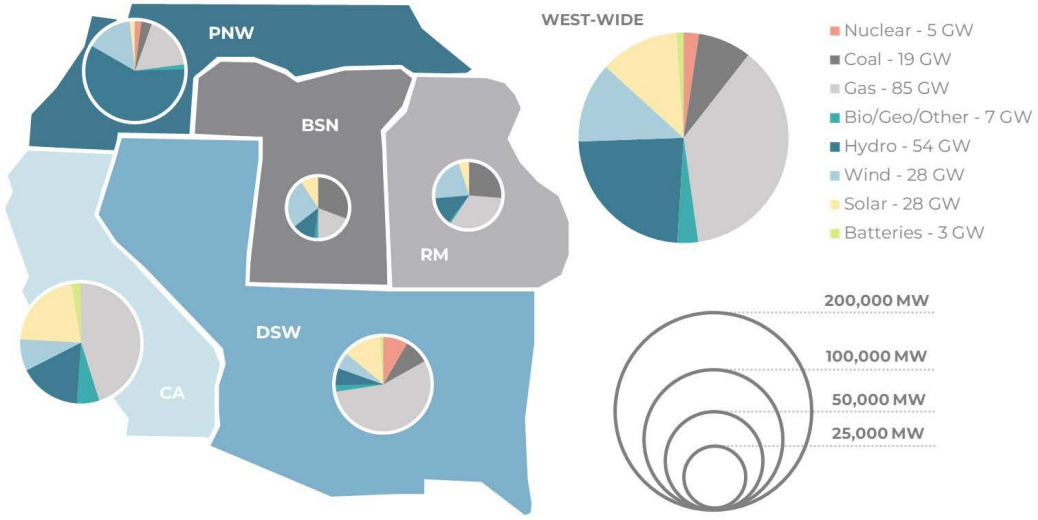
The No Additions Scenario estimates the underlying physical resource adequacy challenge in the Western United States without taking action. The scenario accounts for planned retirements, but not planned additions, unless they were already under construction as of February 2021. As such, the scenario describes the resource adequacy challenge for which current and forthcoming resource plans or procurement activities could be designed or modified to address.

The resource list is based on information in EIA Form 860M (February 2021) and includes resources with commercial online dates prior to January 1, 2026, retirement dates after December 31, 2026, and one of the following EIA status codes: operating; standby/backup; out of service but expected to return to service in next calendar year; under construction, less than or equal to 50% complete; under construction, more than 50% complete; or construction complete, but not yet in commercial operation. Projects that were planned or had regulatory approvals, but were not yet under construction as of February 2021, were excluded.

The resulting resource fleet in the No Additions Scenario is summarized in Figure 3.1.

FIGURE 3.1.

Resource composition in the No Additions Scenario.



3.1.1 LOSS OF LOAD METRICS

Table 3.1 summarizes key loss of load metrics for the No Additions Scenario (see Section 2.5 for additional information about each metric). Where listed, uncertainties and ranges reflect approximate 95% confidence intervals assuming a binomial distribution for loss of load events. Notably, this scenario did not achieve a one-day-in-10-year *LOLE,* as there were, on average, 18.2 days every 10 years that experienced lost load (or 1.82 days per year). The average loss of load event in this scenario lasted 2.3 hours and resulted in 7,597 MWh of unserved energy.

TABLE 3.1.

Loss of load metrics in the No Additions Scenario.

METRIC	VALUE	DESCRIPTION
$LOLP_{year}$	69% ± 3%	Probability of encountering lost load in a given year
$LOLP_{day}$	0.50% ± 0.02%	Probability of encountering lost load in a given day
$LOLP_{hour}$	0.048% ± 0.001%	Probability of encountering lost load in a given hour
LOLE (days/10yrs)	18.2 ± 0.8	Expected number of days with lost load in 10 years
LOLH (hrs/yr)	4.23 ± 0.14	Expected number of hours with lost load in a year
EUE (MWh/yr)	13,797	Expected amount of unserved energy in a given year
EUE_{norm} (ppm)	19.4	Average annual fraction of load not served x 10^6
$EUE^{\ day}_{\ event}$ (MWh/loss-of-load-day)	7,597	Average total energy shortage on a loss of load day
$EUE^{\ hour}_{\ event}$ (MW/loss-of-load-hour)	3,259	Average capacity shortage in a loss of load hour
Average Event Duration (hrs)	2.33	Average number of hours of lost load on a day with lost load

In addition to the average metrics described in Table 3.1, the GridPath simulation provides more descriptive information about the nature of lost load events experienced across the simulation. Figure 3.2 shows two ways of breaking out this information. Panel (a) breaks out the *LOLH* by month and hour of the day[36] to show when the system is most likely to encounter shortfalls. In the No Additions Scenario, loss of load events were identified solely in summer evening hours, with the highest probability of lost load occurring at HE 18 (6-7pm in Pacific Daylight Time) in August. Panel (b) breaks out loss of load events by their duration, or the number of hours of lost load experienced on each day with lost load. In the No Additions Scenario, 93% of events were four hours or less, 99% of events were 6 hours or less, and all events were 8 hours or less. Figure 3.3 breaks out loss of load events by the total energy shortage and the maximum capacity shortage experienced on each day. For each cell in this chart, a resource that can reliably provide the corresponding amount of energy and capacity when needed can eliminate all events in the cell as well as the events in those cells above or to the left of it.

36 Where not otherwise specified, hour of the day in this report refers to hour ending (HE) in Pacific Standard Time.

FIGURE 3.2.

Loss of load event timing and duration in the No Additions Scenario.

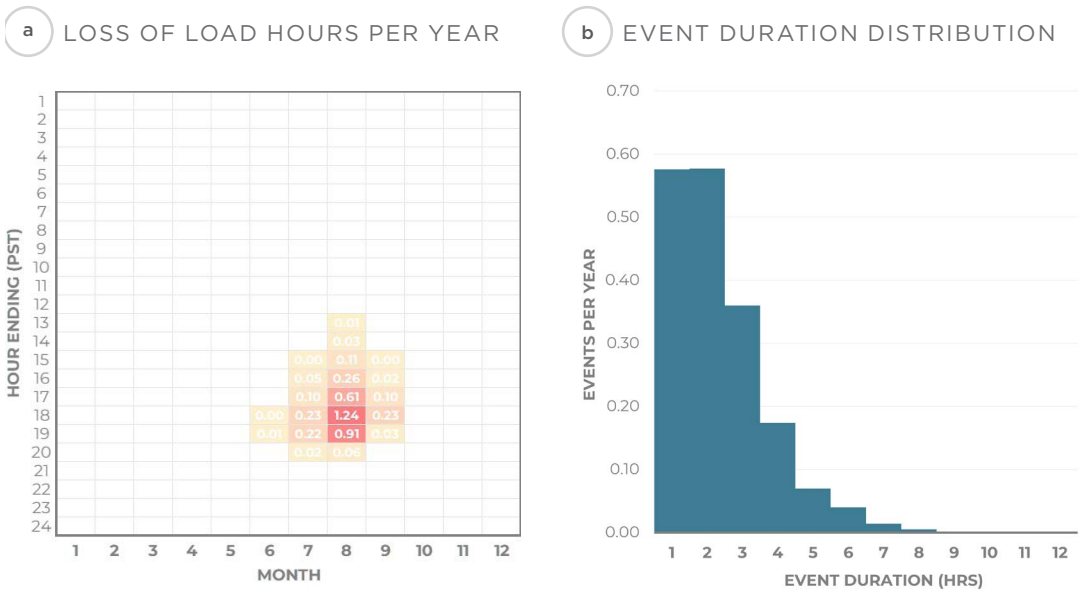
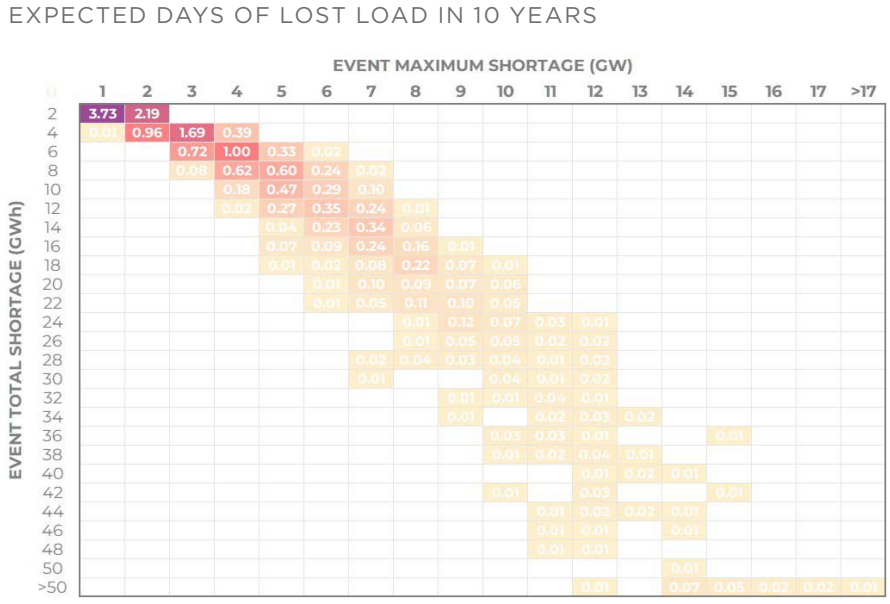


FIGURE 3.3.

Capacity and energy shortages in the No Additions Scenario.



Taken together, the loss of load statistics for the No Additions Scenario paint a picture of a resource inadequate system, with an *LOLE* that exceeds one day in 10 years. However, the observed shortages were fairly small relative to the size of the system, the durations of the shortages tended to be short, and the timing of the shortages was consolidated into a relatively predictable

seasonal and diurnal pattern. Both the short duration of the shortages and the timing of the shortages on hot summer evenings as the sun begins to set suggest that shorter duration energy-limited resources, like demand response and energy storage, may be well suited to meeting remaining needs and achieving resource adequacy. The next section delves into this more deeply.

Both the short duration of the shortages and the timing of the shortages on hot summer evenings as the sun begins to set suggest that shorter duration energy-limited resources, like demand response and energy storage, may be well suited to meeting remaining needs and achieving resource adequacy.

3.1.2 CAPACITY AND ENERGY NEEDS

To characterize the amount of capacity and energy needed to achieve resource adequacy on a system, one must first select a resource adequacy standard. To determine capacity and energy needs in this study, we primarily rely upon a traditional one-day-in-10-year $LOLE$ standard, however we also present results using alternative $LOLP$-based standards, including the 5% $LOLP$ standard that has traditionally been used by the Northwest Power and Conservation Council and the 2.4-hour-per-year $LOLH$ standard adopted by some utilities.

The simplest way to determine the capacity needs under an $LOLP$-based RA standard is to investigate the simulated shortages and their likelihoods using a duration curve. In a duration curve, the events are sorted from largest to smallest on the y-axis and the x-axis is calculated as the cumulative sum of the likelihoods of those events. The resource needs associated with a given $LOLP$-based RA standard can be found by reading the value along the duration curve that corresponds to the acceptable likelihood of an event. Figure 3.4 shows this for the one-day-in-10-year $LOLE$ standard. In this example, an event is any day with lost load and the capacity shortage associated with an event is the maximum amount of lost load observed during the event. The value of the duration curve at one day in 10 years is 9.3 GW. If this amount of capacity can be provided when needed, then the events to the right of the dotted line are eliminated and the events to the left of the dotted line are reduced in magnitude. This capacity addition effectively slides the duration curve downward until it hits the x-axis at one day in 10 years (as shown in the dashed line).

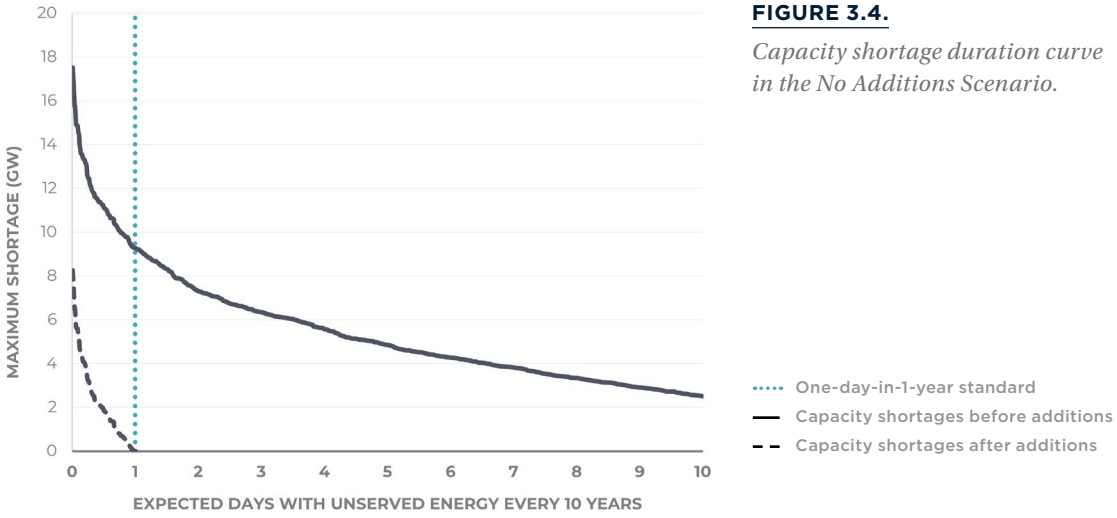


FIGURE 3.4.

Capacity shortage duration curve in the No Additions Scenario.

This same exercise can be undertaken to identify the needs associated with alternative resource adequacy standards (e.g., LOLP = 5% of years and *LOLH* = 2.4 hrs/year) using the same underlying hourly data from GridPath, but calculating events in different ways (e.g., an event is a year with unserved energy or an hour with unserved energy, respectively). Table 3.2 compares the capacity shortages identified using the one-day-in-10-year *LOLE* standard to those identified using a 5% *LOLP* standard or a 2.4 hour per year *LOLH* standard.

TABLE 3.2.

Perfect Capacity needs in the No Additions Scenario.

RA STANDARD	PERFECT CAPACITY NEED IN NO ADDITIONS SCENARIO
LOLE = One day in 10 years	8.8 - 9.9 GW
$LOLP_{year}$ = 5%	11.0 GW
LOLH = 2.4 hrs/yr	2.1 GW

For this particular system, the 5% *LOLP* metric was most stringent and a 2.4 hr/year *LOLH* standard was much less stringent than the one-day-in-10-year *LOLE* standard. The vast difference between the capacity shortages associated with the one-day-in-10-year *LOLE* standard and the 2.4 hr/year *LOLH* standard is not surprising given that the average event duration for this case was 2.3 hours. If the system hypothetically encountered 10 typical 2.3-hour events in a ten-year period, it would be far from achieving a one-day-in-10-year *LOLE* standard,

The vast difference between the capacity shortages associated with the one-day-in-10-year *LOLE* standard and the 2.4 hr/year *LOLH* standard is not surprising given that the average event duration for this case was 2.3 hours. If the system hypothetically encountered 10 typical 2.3-hour events in a ten-year period, it would be far from achieving a one-day-in-10-year *LOLE* standard, but would still meet a 2.4 hr/year *LOLH* standard.

but would still meet a 2.4 hr/year *LOLH* standard. This does not mean that an *LOLH* standard is any less valid than an *LOLE* standard—they simply represent different risk tolerances. We do not delve into the origins or justifications of various RA standards in this report, nor do we advocate for the adoption of one standard over another, but these findings make clear that identified resource needs may be highly sensitive to the adopted standard. In selecting a standard, policymakers and planners should be aware of this sensitivity and thoughtful about the potential cost impacts and impacts to health, human safety, and the economy of both potential RA shortages and any actions taken to avoid such shortages.

The simple exercises described above assume that additional capacity is available in all hours. This type of capacity, which has no energy limitations, likelihood of failure, or operational limitations, is often called "perfect capacity". Perfect capacity shortages can be a useful way to provide a simple high level description of the magnitude of RA challenges on a system and to make apples-to-apples comparisons between different scenarios, but it can also obscure important features of resource need.

To provide more visibility into the energy or duration requirements for achieving resource adequacy, a similar exercise can be undertaken in two dimensions: one that represents capacity and one that represents energy. The heatmap in Figure 3.3 is one way of describing this two-dimensional duration curve. To identify the amount of capacity and energy required to meet the one-day-in-ten-year standard, one must draw a two-dimensional box that achieves the standard, rather than a one-dimensional line. An example of this is shown in Figure 3.5. In this example, if a resource was available that could provide 26,000 MWh of energy on any day with lost load, with a maximum capacity of 11 GW, that resource would eliminate all of the events within the shaded box (which occur 17.2 times every 10 years), leaving only the events below and/or to the right of the box (which occur 0.97 days every 10 years).

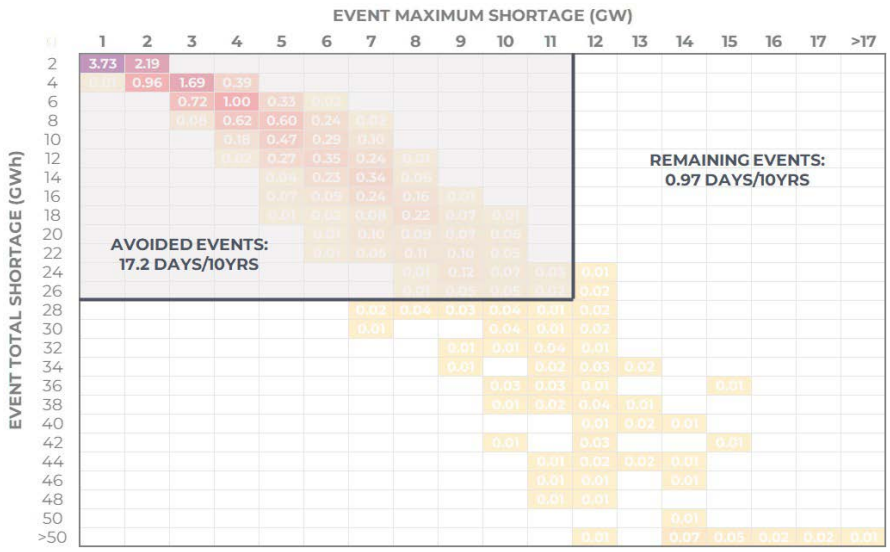


FIGURE 3.5.

Capacity and energy solution example in the No Additions Scenario.

While Figure 3.5 shows one potential strategy for meeting the one-day-in-10-year standard, multiple strategies exist (i.e., multiple boxes can be drawn) to meet the standard. Figure 3.6 shows several strategies for achieving the standard, represented as the resource capacity required as a function of the resource duration. Note that duration in this curve is shorthand for energy divided by capacity and does not limit the number of hours over which that energy and capacity could be provided. For example, a 100 MW 4-hr solution could be provided as 50 MW for 8 hours using this convention. This curve represents an efficient frontier for the capacity and energy additions that meet the standard at lowest cost. The curve makes clear the trade-off between capacity and duration for short-duration solutions (i.e., the longer duration solutions you have, the less total capacity you need), but also shows that increasing duration beyond three hours is not necessary to meet the reliability standard and provides little value in terms of avoiding capacity need for this particular system. The lowest cost point along this curve can be determined by examining the costs of solutions of different durations, but the rapid flattening of the slope suggests that the economically optimal duration is unlikely to be beyond four hours, and may be as low as two or three hours.

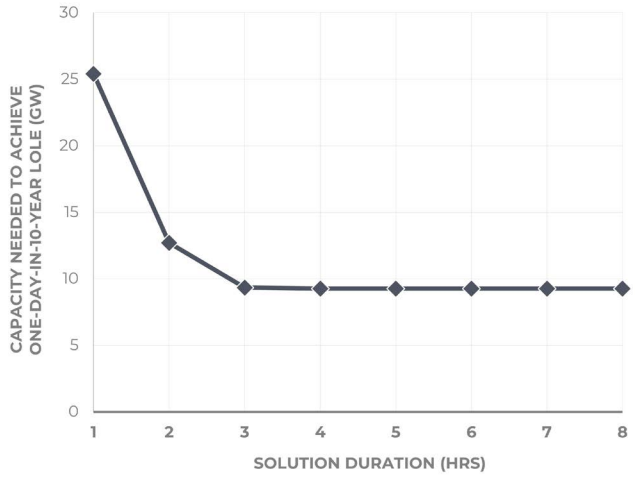


FIGURE 3.6.

Capacity and duration efficient frontier for the No Additions Scenario.

3.1.3 DRIVERS OF UNSERVED ENERGY

Net load (load minus wind and solar output) was a strong driver of unserved energy in the simulated years. However, as wind and solar are still not dominant in the simulated system, net load conditions in the No Additions Scenario were still primarily driven by load levels, with peak load and peak net load conditions occurring on the same day. The peak net load in the simulation was 142.1 GW, and unserved energy was not observed below WECC net load of 118.3 GW. The additional variation in unserved energy can be explained by the hydro and outage conditions in the simulation.

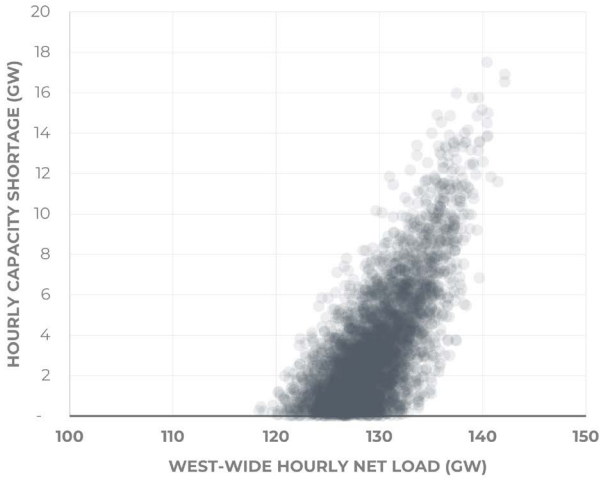


FIGURE 3.7.

Relationship between capacity shortages and net load in the No Additions Scenario.

Figures 3.8 and 3.9 show resource availability on some key days in the simulation. For thermal units (including coal, gas, nuclear, biomass, and biogas) the resource availability represents the resources' capacity adjusted for simulated outages and temperature-driven thermal derates. For hydro and storage, resource availability represents their output plus any contingency reserves provided. Wind and solar resource availability is based on their simulated output profiles.

On the peak load day in the Monte Carlo Simulation, load (including contingency reserves) reached 152.9 GW. The simulation included multiple instances of these load days, and the highest hourly capacity shortage observed on a day with these load conditions was 14.9 GW. The system resource availability on this day is shown in Figure 3.8(a). On this day, extreme load conditions combine with low wind and hydro availability as well as reduced gas availability (attributable largely to thermal derates) to produce a large shortage in the afternoon and early evening.

Net load (load minus wind and solar output) was a strong driver of unserved energy in the simulated years. However, as wind and solar are still not dominant in the simulated system, net load conditions in the No Additions Scenario were still primarily driven by load levels, with peak load and peak net load conditions occurring on the same day.

The peak net load day in the Monte Carlo Simulation (Figure 3.8(b)) had similar load conditions relative to the peak load day described above, peaking at 152.8 GW, but even lower wind and solar output. Peak net load on this day was 142.1 GW vs 140.1 GW on the day above. With similar hydro and gas availability on these two simulation days, the higher net load conditions resulted in a higher shortfall: the maximum capacity shortage observed on this peak net load day was 16.9 GW. The largest hourly capacity shortage in the simulation was 17.5 GW and occurred on a day with high peak net load conditions (140.0 GW) in addition to relatively low hydro and thermal availability. While the magnitude of the afternoon and evening shortages was large, sufficient resources were available during the night and the morning to produce energy that could be shifted to the times of grid stress using storage, a potential solution we test in the Less Coal Scenario.

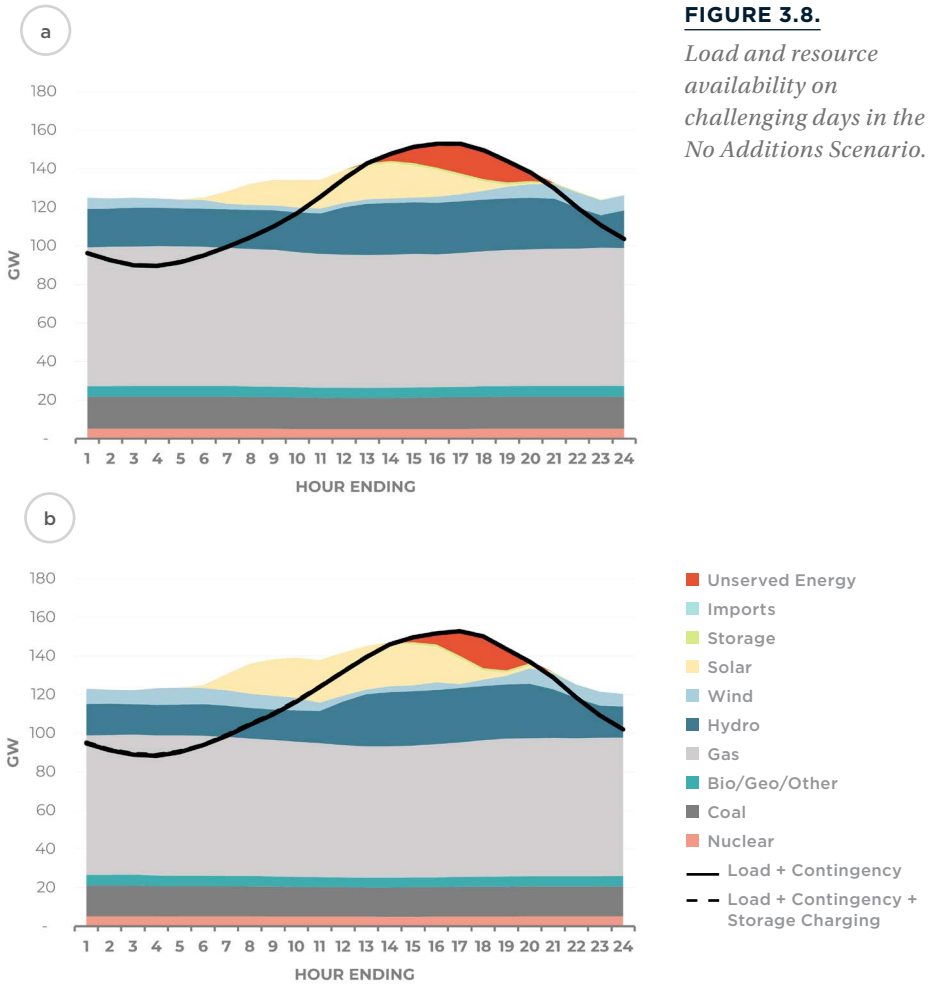


FIGURE 3.8.

Load and resource availability on challenging days in the No Additions Scenario.

The simulated system is still dominated by thermal resources such as gas, coal, and nuclear as well as hydropower, so in addition to load and net load conditions, thermal outages and hydro conditions can also drive unserved energy. Figure 3.9(a) shows an example of this. Nuclear, coal, and gas outages together with lower hydro availability resulted in a capacity shortage of almost 6 GW in HE 18 on this day. Higher thermal availability as well as more favorable wind conditions can result in no capacity shortages on days with the higher load and lower wind and solar output conditions (Figure 3.9(b)). In HE 18, the combined nuclear, coal, and gas availability on the former day was 88.9 GW and 93.3 GW on the latter. Hydro output levels were 30.2 GW and 33.6 GW respectively. Despite lower net load conditions, the former day (peak net load of 130.8 GW) experienced up to 5.7 GW of capacity shortage while the latter day (peak net load of 134.8 GW) did not encounter RA issues.

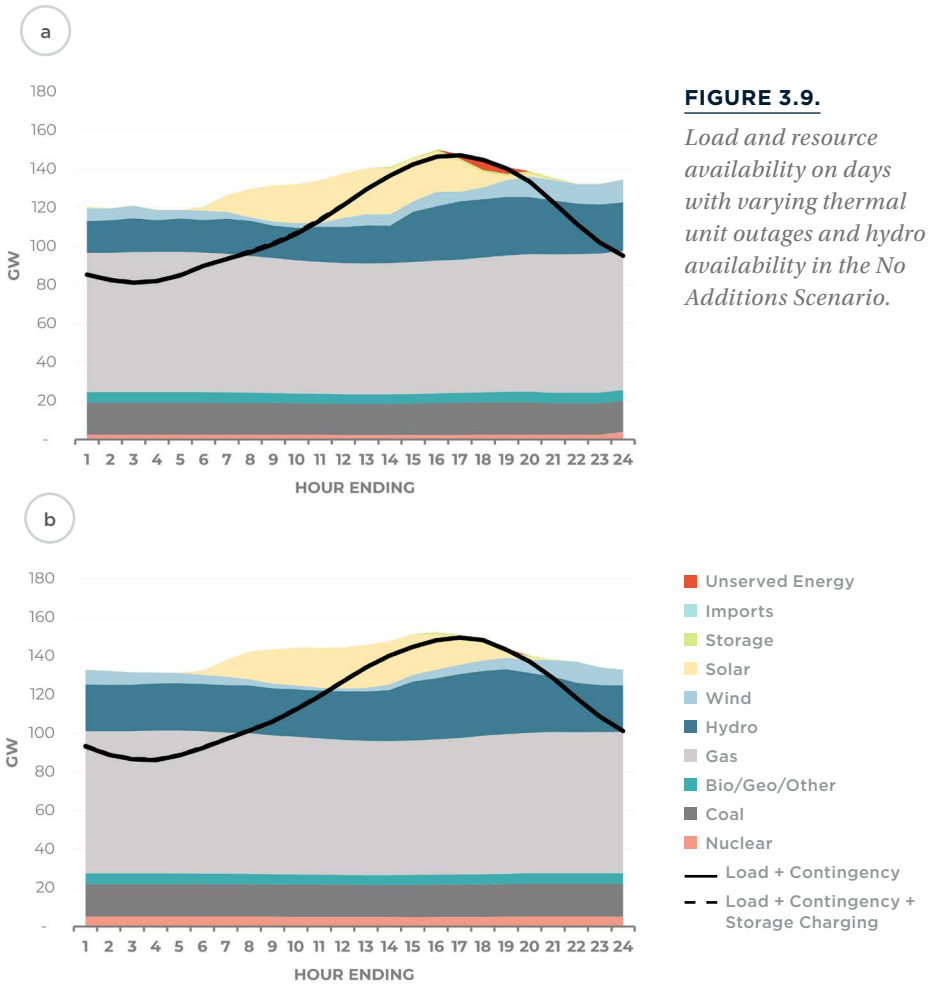


FIGURE 3.9.

Load and resource availability on days with varying thermal unit outages and hydro availability in the No Additions Scenario.

3.1.4 REGIONAL FLOWS

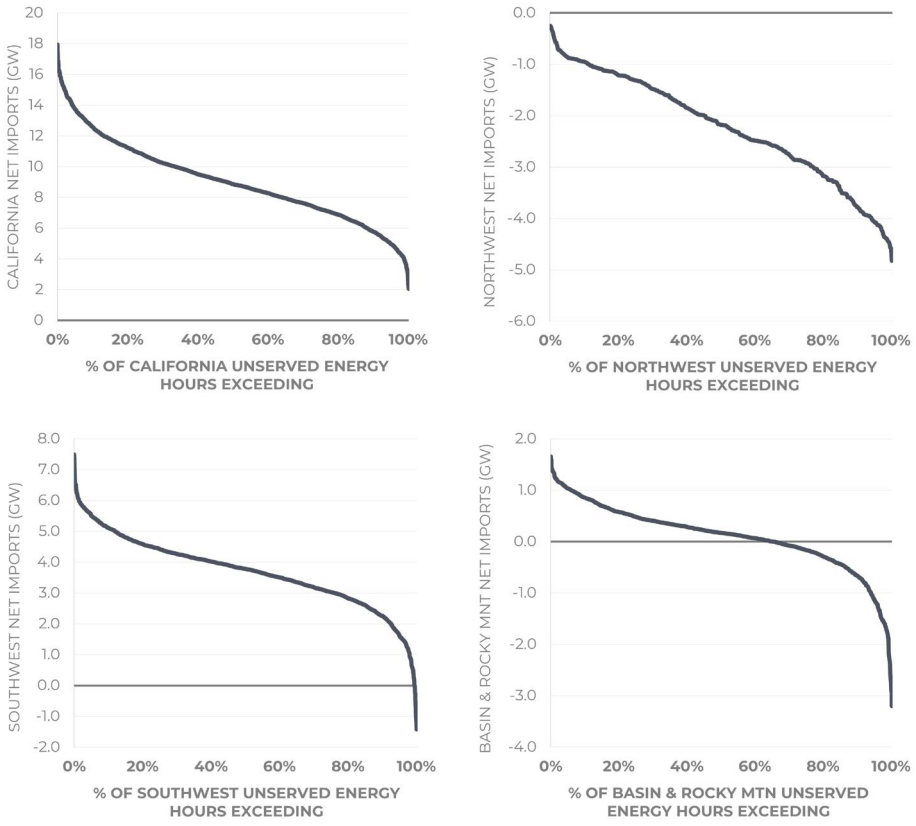
During the hours when the system experienced loss of load in the simulation, significant flows were observed between subregions within the Western US, in particular from the Pacific Northwest and Desert Southwest regions into California. On average, California imported 5.2 GW from the Desert Southwest and 3.3 GW from the Pacific Northwest during these hours.

Flows between subregions, however, varied widely across hours with unserved energy. Figure 3.10 shows duration curves of net imports into each subregion during the hours when the respective region experienced unserved energy. During the hours when it encountered loss of load, California imported between 2.1 GW and 17.9 GW, with at least 7 GW imported around 80 percent of the time and at least 6 GW imported around 90 percent of the time. When unserved energy occurred in the Desert Southwest, flows to the subregion varied between 1.4 GW of exports and 7.5 GW of imports. For the Basin-Rocky Mountain subregion, flows varied between 3.2 GW of exports and 1.7 GW of

imports during the subregion's hours of unserved energy. In the few hours when the Pacific Northwest experienced unserved energy, it remained a net exporter, with exports varying between 4.8 GW and 0.3 GW.

FIGURE 3.10.

Net imports during loss of load hours in the No Additions Scenario.



Higher shortages in California did appear to be correlated with lower import levels, although the relationship between the import level and the amount of unserved energy experienced was weak, with other factors such as net load level, hydro availability, and thermal outage conditions playing important roles.

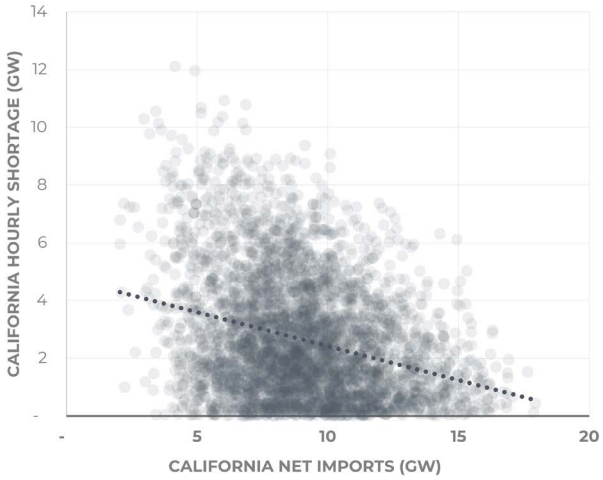


FIGURE 3.11.

Relationship between California imports and hourly shortages in the No Additions Scenario.

Loss-of-load events in California frequently coincided with high-risk conditions in the rest of the Western region. Unserved energy was experienced in the rest of the region in about 50 percent of the hours in which loss of load occurred in California. Conversely, California encountered a loss of load event in 91 percent of the hours in which unserved energy occurred in the rest of the West. An example is shown in Figure 3.12. In this week, California, the Northwest, the Southwest, and the Rocky Mountain—Basin regions all experienced high load levels on Monday and Tuesday, likely a high-temperature period across a wide area in the West. None of the regions were able to lean on neighbors to cover their shortages, with resources inadequate to meet the demand over a large geographic area.[37]

37 The formulation of the GridPath simulation ensures that each BA prioritizes meeting its own load before exporting energy. When it has surplus resources, a BA can export to other BAs without prioritizing those in its own region. This is why a region may experience unserved energy while exporting to other regions at the same time, as is the case for the Northwest and Southwest regions in Figure 3.12. Some BAs within those regions are experiencing lost load while others have surplus resources that they are exporting to other regions.

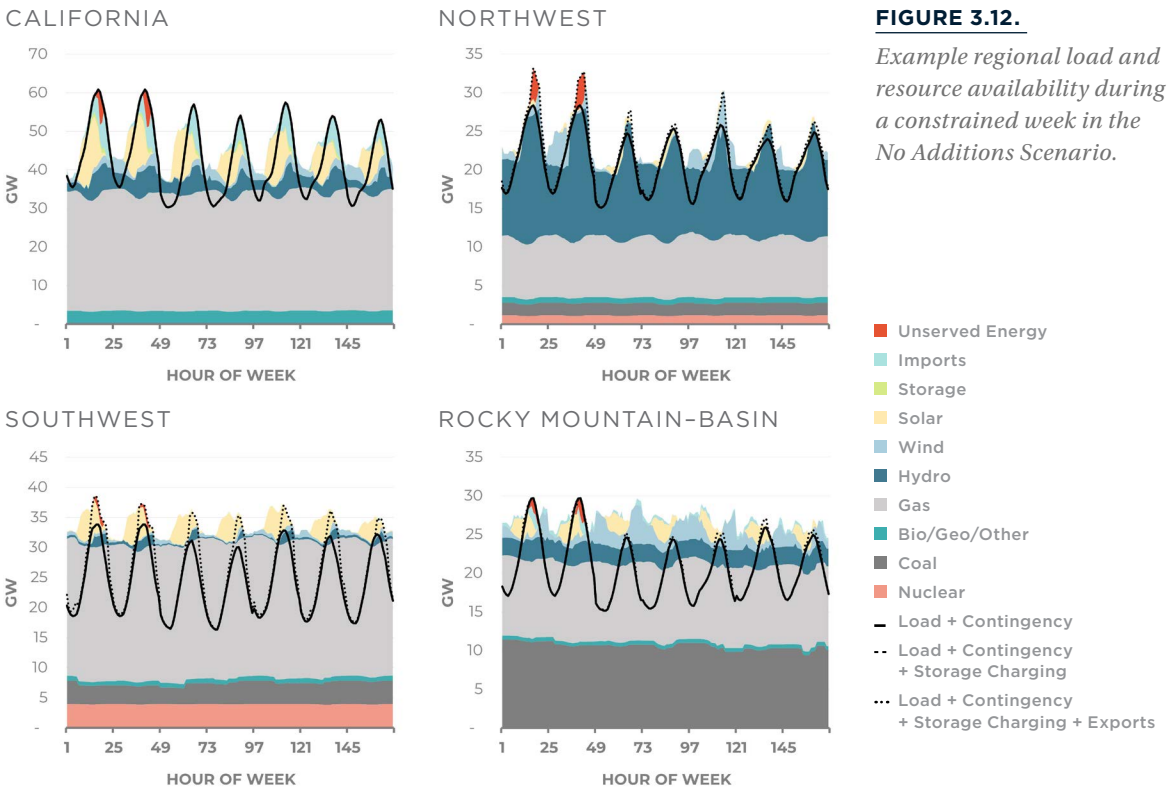


FIGURE 3.12.

Example regional load and resource availability during a constrained week in the No Additions Scenario.

Key Takeaways from the No Additions Scenario

- Without taking any action, we estimate that the West could be physically short by about 8.8 - 9.9 GW in 2026, if planning to a one-day-in-10-year LOLE standard. This shortage is much smaller than the amount of capacity additions in current utility plans in the West, including the procurement ordered with CPUC Decision D.21-06-035, which requires 11.5 GW of new net qualifying capacity through 2026.
- Shortages are expected to be relatively short in duration (mostly 4 hours or less) and occur in the evenings on hot summer days.
- Different combinations of capacity and energy could be secured to achieve a one-day-in-10-year LOLE standard, however we find that pursuing solutions beyond four hours in duration yields very small or negligible resource adequacy benefits for 2026.
- During loss of load events, California averaged approximately 5.2 GW of imports from the Desert Southwest and 3.3 GW of imports from the Pacific Northwest. Interregional flows were

key determinants of subregional resource adequacy positions and must be accounted for within RA analysis.

3.1.5 SUBREGIONAL ANALYSIS

In this section we explore how the RA challenges identified by different RA programs in the West might compare to those identified in a West-wide study. To investigate this, we broke the study footprint into subareas that roughly resemble the footprints of existing and proposed resource adequacy programs (CAISO and the Western Resource Adequacy Program, or WRAP). Table 3.3 lists the BAAs assigned to each subarea and Figure 3.13 shows the nodes and zonal transmission links in each subarea. The BAAs in the WRAP subarea were selected based on WRAP participants announced as of December 7, 2021. BANC was also included to simplify the transmission treatment for TIDC in the study. Trading hubs that sit adjacent to the two subareas (Mead and Palo Verde) were included in both subareas and any associated resources were allocated to the corresponding subarea based on publicly available information.

TABLE 3.3.

Subarea definitions.

SUBAREA	WECC BAAS/ZONES
CAISO	CIPB, CIPV, CISC, CISD, VEA, TH_Mead (partial), TH_PV (partial)
WRAP	AVA, AZPS, BANC, BPAT, CHPD, DOPD, GCPD, IPFE, IPMV, IPTV, NEVP, NWMT, PACW, PAID, PAUT, PAWY, PGE, PSEI, SCL, SPPC, SRP, TIDC, TPWR, TH_Malin, TH_Mead (partial), TH_PV (partial)
Excluded	EPE, IID, LDWP, PNM, PSCO, TEPC, WACM, WALC, WAUW

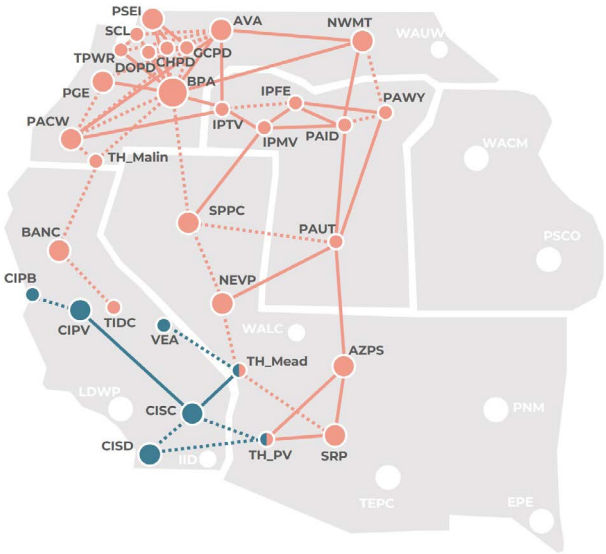


FIGURE 3.13.

Subarea zonal transmission topologies.

Note that the subareas do not exactly align with real resource adequacy programs because they consider loads and resources that are associated with each balancing area within a physical model of the West, rather than allocating resources to load serving entities (LSEs) based on ownership and contractual information. To the extent that there are resources in one subarea that are contractually obligated to serve loads in another subarea, actual resource adequacy positions will differ from this analysis. This analysis is therefore broadly indicative of regional dynamics, but not necessarily indicative of actual resource adequacy program positions. Further work to assign resources to LSE's participating in resource adequacy programs based on contractual information would offer more actionable contributions in this area, but may require confidential information.

The transmission topology in these scenarios was modified to remove transmission links between subareas. Therefore, only WECC paths and other interfaces that are fully within a subarea are enforced in these scenarios (these path constraints are enforced across the solid lines in Figure 3.13).

In the subregional analysis, we characterize RA needs under two very different policy constructs: one in which each footprint is treated as an island (the Island Policy); and one in which each footprint is able to rely, to some extent, on excess generation from neighboring systems, subject to transmission limits (the Import Policy). It is important to note that neither of these two policies is more physically accurate and in fact there are many equally valid alternative policies that could be adopted by a planner. Planning for resource adequacy in a particular system within the context of a broader market is fundamentally a policy decision about risk. In the most extreme example, a risk averse planner might assume no reliance on market availability when the system is constrained to eliminate the likelihood of an event where planners expect some assistance from outside the footprint that does not materialize. This policy (our Island Policy) will reduce resource adequacy risk, but will tend to lead to overbuild.

In contrast to the Island Policy, a highly cost-sensitive planner might instead look to the regional resource adequacy challenge and plan to address only the portion of that regional challenge that is attributable to their system. Attribution is subjective and can be a highly complex undertaking. Furthermore, without formal planning coordination (through for example, a resource adequacy program), this approach requires planners to make an assumption about how other entities will plan. This assumption could be made explicitly—for example, if a planner includes load and resource zones outside of their system in their resource adequacy analysis to account for imports, then the planner must decide whether to add resources outside of their footprint to ensure that external zones are resource adequate themselves. This assumption could also be made implicitly, by adopting an import capability assumption, which may not be specifically tied to resource additions outside the footprint, but certainly makes implications about the resource adequacy positions of neighboring entities when the system is constrained.

In this study, we adopt an import policy that allows the planner to account for excess generation from existing resources outside of their footprint, but does not assume additional resources will necessarily be built by other entities to help mitigate regional RA challenges. We call this a *Unilateral Imports* Policy because it accounts for imports, but also allows the planner to act unilaterally to achieve resource adequacy on their own system. In this policy, the planner plans for the minimum of the islanded subarea shortage and the West-wide shortage experienced in each hour.[38] For a given hour, the policy makes the following assumptions:

- If the West-wide shortage is smaller than the islanded subarea shortage, then there is some excess capability somewhere in the West, with adequate transmission, that could mitigate a portion or all of the shortage experienced in the subarea. In such an hour, the Unilateral Imports Policy assumes that, at worst, the subarea would experience all of the West-wide shortage, so the policy assumes that the subarea shortage is equal to the West-wide shortage, and the imports during that hour are equal to the amount of shortage avoided: the islanded subarea shortage minus the West-wide shortage.
- If the West-wide shortage is larger than the islanded subarea shortage, then there are factors outside of the subarea that are exacerbating the regional RA challenge beyond those experienced by the subarea. In the Unilateral Imports Policy, it is not the responsibility of the subarea planner to mitigate these external issues, so the planner considers only the shortage identified for the islanded subarea. In this hour, the subarea shortage is equal to the islanded subarea shortage and the imports are equal to zero, because the rest of the system is constrained at the same time.

While not as costly as an Island Policy, the Unilateral Imports Policy (or Imports Policy as it is referred to throughout this report) will still tend to result in overbuild because the West-wide shortages may be further reduced by actions taken by other entities. Further cost reductions can be achieved through formal resource adequacy coordination, which allocates a portion of the regional need to each entity and provides each planner with confidence that other entities will act to meet their allocated needs.

3.1.5.1 CAISO subarea

Table 3.4 lists the loss of load metrics for the CAISO subarea under both the Island and Import policies described above. When the CAISO subarea was treated as an island, the No Additions Scenario resulted in a very high likelihood of lost load—100% of the simulated years saw a shortage and the LOLE was 335 days every 10 years. However, adopting the import policy avoided nearly all of these shortages and resulted in loss of load risk in CAISO that closely aligned with the West-wide findings for the No Additions Scenario.

38 This requires the West-wide and subarea-specific simulations to consider the same underlying system conditions in each hour.

TABLE 3.4.

Loss of load metrics for the CAISO subarea in the No Additions Scenario, compared to West-wide loss of load metrics.

METRIC	NO ADDITIONS SCENARIO - CAISO ISLAND	NO ADDITIONS SCENARIO - CAISO IMPORTS	NO ADDITIONS SCENARIO - WEST-WIDE
$LOLP_{year}$	100%	69%	69%
$LOLP_{day}$	9.18%	0.50%	0.50%
$LOLP_{hour}$	0.991%	0.047%	0.048%
LOLE (days/10yrs)	335	18.2	18.2
LOLH (hrs/yr)	86.8	4.15	4.23
EUE (MWh/yr)	225,373	12,134	13,797
EUE_{norm} (ppm)	1,083	58	19.4
$EUE^{\ day}_{\ event}$ (MWh/loss-of-load-day)	6,724	6,685	7,597
$EUE^{\ hour}_{\ event}$ (MW/loss-of-load-hour)	2,595	2,922	3,259
Average Event Duration (hrs)	2.59	2.29	2.33

Figures 3.14 and 3.15 show the timing and duration of events under the Island and Import Policies. With the Island Policy, shortages were observed throughout May-October and sometimes spanned large swaths of the day—as early as HE 12 and as late as HE 24. HE 19 (8pm PDT) in August had the highest loss of load risk. With access to imports, the risk was significantly reduced across all periods and was concentrated primarily in HE 15 through HE 20 in July - September. The highest loss of load risk occurred slightly earlier at HE 18 (7pm PDT) in August, and the longest duration events observed with the Island Policy were mitigated (or shortened) by imports.

FIGURE 3.14.

Loss of load event timing and duration for the CAISO subarea without imports in the No Additions Scenario.

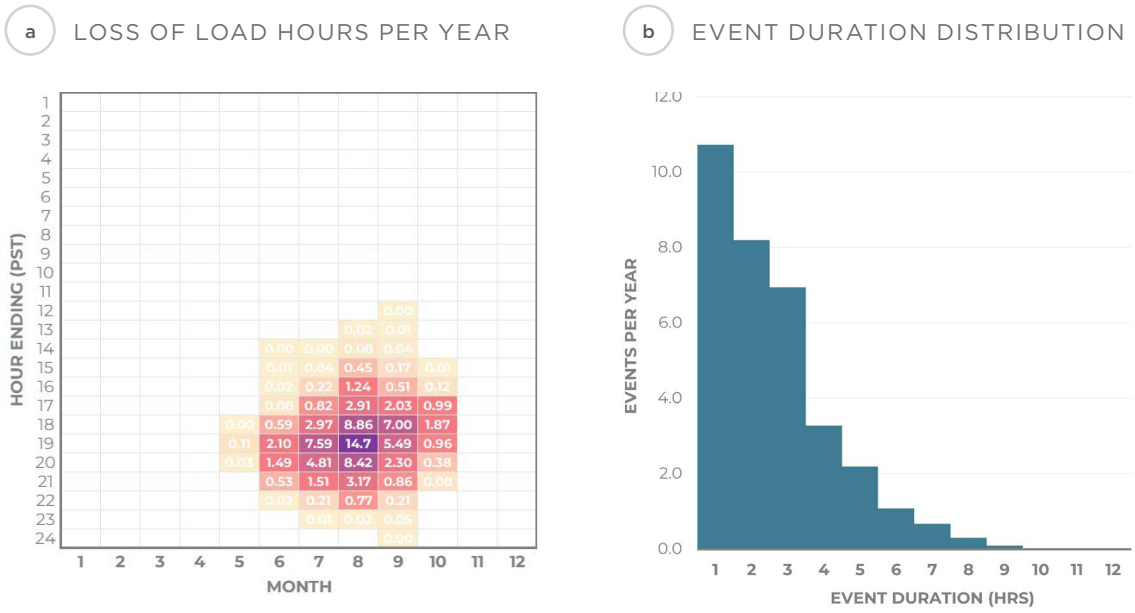


FIGURE 3.15.

Loss of load event timing and duration for the CAISO subarea with imports in the No Additions Scenario.

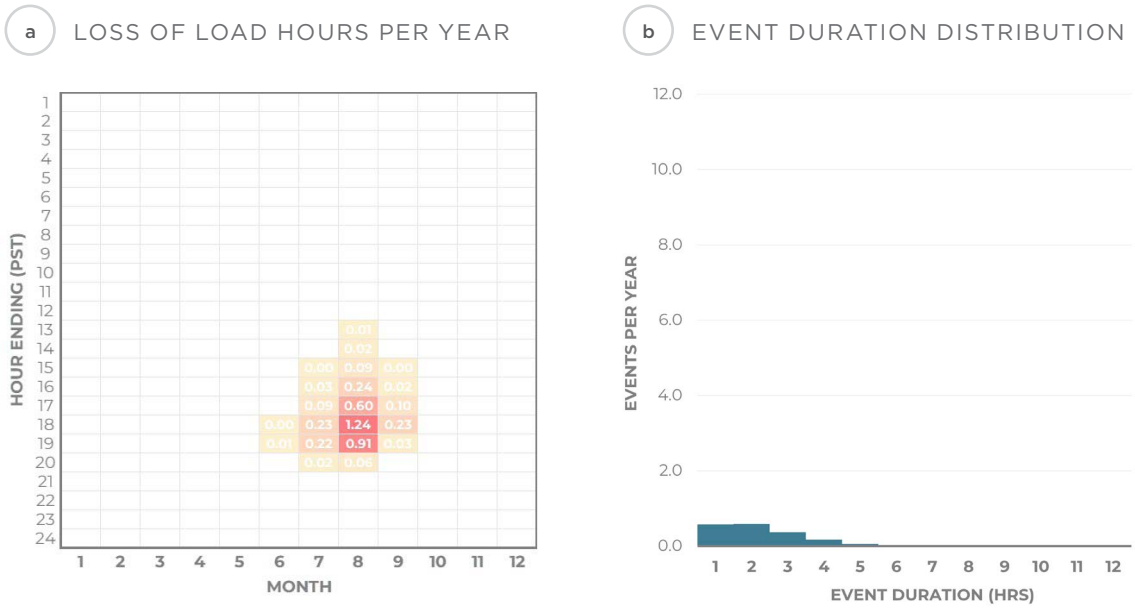


Table 3.5 summarizes the perfect capacity needs under various RA standards for the CAISO subarea under the Island and Import Policies. Across all three RA standards, the CAISO Island case yielded significant capacity needs. If using a one-day-in-10-year standard or a 5% LOLP standard, adopting the import policy reduced perfect capacity needs by about 3 GW. However, if using a 2.4 hour per year LOLH standard, adopting the import policy avoided about 6.1 GW of perfect capacity needs. In this example, the capacity avoided by adopting an import policy was highly sensitive to the RA standard. This example highlights the complexity of the relationships between different probability-based RA standards and the difficulty in translating between them without system-specific probabilistic analysis.

TABLE 3.5.

Perfect capacity needs for the CAISO subarea in the No Additions Scenario, compared to West-wide perfect capacity needs.

RA STANDARD	NO ADDITIONS SCENARIO - CAISO ISLAND	NO ADDITIONS SCENARIO - CAISO IMPORTS	NO ADDITIONS SCENARIO - WEST-WIDE
LOLE = One day in 10 years	11.2 GW	8.2 GW	9.3 GW
$LOLP_{year}$ = 5%	11.8 GW	8.8 GW	11.0 GW
LOLH = 2.4 hrs/yr	8.0 GW	1.9 GW	2.1 GW

The difference between the capacity needs in the Island and Import cases was even greater as energy limitations were applied, as shown in Figure 3.16 for the one-day-in-10-years RA standard. If limited to 2-hour resources, adopting the Import Policy avoided over 15.7 GW of capacity relative to the Island Policy. This is because accounting for interactions between neighboring systems tends to reduce the duration of shortages and the need for longer duration solutions.

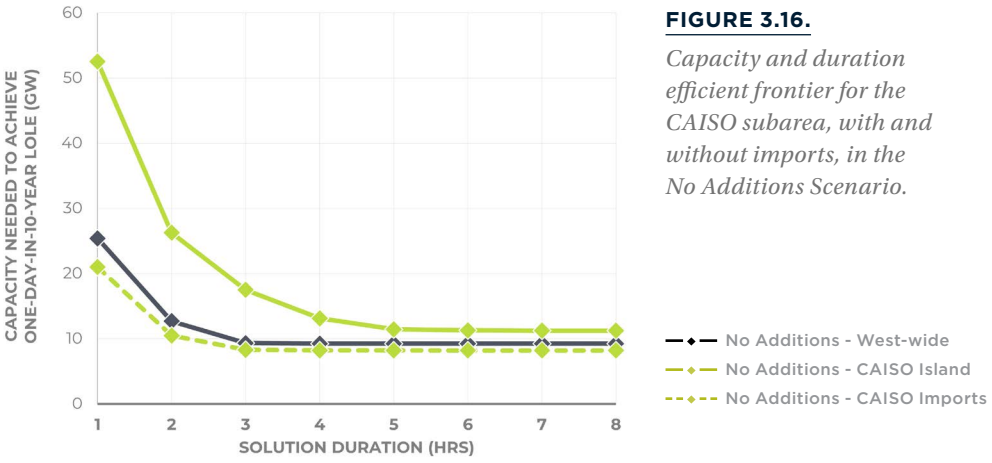


FIGURE 3.16.

Capacity and duration efficient frontier for the CAISO subarea, with and without imports, in the No Additions Scenario.

3.1.5.2 WRAP subarea

Table 3.6 lists the loss of load metrics for the WRAP subarea under both the Island and Import Policies. The WRAP Island case saw much less frequent lost load than the West-wide case, but still did not meet a one-day-in-10-year standard. While accounting for imports did avoid some lost load, the relative impact of imports was much smaller than was seen for the CAISO subarea.

TABLE 3.6.

Loss of load metrics for the WRAP subarea in the No Additions Scenario, compared to West-wide loss of load metrics.

METRIC	NO ADDITIONS SCENARIO - WRAP ISLAND	NO ADDITIONS SCENARIO - WRAP IMPORTS	NO ADDITIONS SCENARIO - WEST-WIDE
$LOLP_{year}$	20%	17%	69%
$LOLP_{day}$	0.08%	0.07%	0.50%
$LOLP_{hour}$	0.008%	0.006%	0.048%
LOLE (days/10yrs)	3.10	2.43	18.2
LOLH (hrs/yr)	0.74	0.55	4.23
EUE (MWh/yr)	824	703	13,797
EUE_{norm} (ppm)	2.4	2.1	19.4
$EUE^{\ day}_{\ event}$ (MWh/loss-of-load-day)	2,659	2,895	7,597
$EUE^{\ hour}_{\ event}$ (MW/loss-of-load-hour)	1,115	1,279	3,259
Average Event Duration (hrs)	2.38	2.26	2.33

As shown in Figure 3.17, the capacity needs identified for the WRAP subarea represented a small fraction of the West-wide need, and were not very sensitive to the import policy in this case. This finding is specific to this particular system and scenario. In Section 3.3, we explore a scenario where the WRAP subarea capacity needs are highly dependent on the import policy.

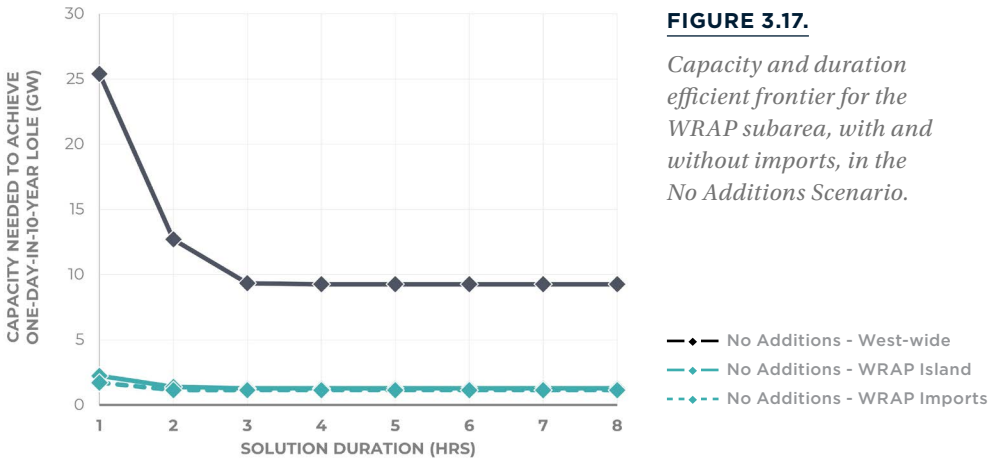


FIGURE 3.17.

Capacity and duration efficient frontier for the WRAP subarea, with and without imports, in the No Additions Scenario.

3.1.5.3 Benefits of planning coordination

The subarea analyses provide a way of quantifying the potential benefits of adopting import policies and coordinating on RA across subareas. Table 3.7 shows how different policies affect the total identified need across the subareas and compares this to the West-wide need. Because some BAs are not included in either subarea, the 4-hr capacity needs listed in the table represent a lower bound of the identified needs if all BAs were to be included. The difference between this value and the West-wide capacity need therefore represents a lower bound for the potential capacity savings of full coordination across the Western US. The actual overbuild will depend on the policies adopted in the remaining BAs.

TABLE 3.7.

4-hr capacity needs with various import policies adopted for the CAISO and WRAP subareas, compared to full coordination, in the No Additions Scenario.

CAISO POLICY	WRAP POLICY	4-HR CAPACITY NEED IN CAISO + WRAP SUBAREAS	OVERBUILD LOWER BOUND
Island	Island	14.4 GW	5.1 GW
Island	Imports	14.3 GW	5.0 GW
Imports	Island	9.5 GW	0.2 GW
Imports	Imports	9.4 GW	0.1 GW
Full West-wide coordination (Includes needs outside CAISO + WRAP)		9.3 GW	-

For this particular scenario, most of the benefits of full coordination between the CAISO and WRAP subareas could be achieved by adopting the Imports Policy for the CAISO footprint—this alone reduced the lower bound of the

overbuild from 5.1 GW to 0.2 GW. In interpreting these results, it is important to keep in mind that the subarea analyses each assume full coordination within their footprints, so the incremental benefit of full coordination across the footprints may be relatively small. If we were to calculate the potential overbuild without coordination within each subarea (i.e., each BA may be treated as an island or adopt an import policy), the benefits of full coordination would be much larger.

Key Takeaways from the Subregional Scenarios

- Due to the highly interconnected nature of the West, resource adequacy analysis that treats subareas or RA programs as islands distorts the observed RA challenges and may lead to suboptimal RA solutions, including potentially significant overbuild.
- To account for interregional operational interactions without full West-wide planning coordination, RA programs may adopt market access policies that are informed by West-wide analysis. This study describes one such policy, the Unilateral Imports Policy, which allows a planner to account for the ability of neighboring systems to help avoid shortages when they are not constrained, but does not assume that other entities will necessarily take planning steps to address RA.
- Applying the Unilateral Imports Policy to CAISO results in very similar RA findings for CAISO as the entire West in the No Additions Scenario. In this particular case, if CAISO were to adopt the Unilateral Imports Policy, CAISO would plan to address most of the identified resource adequacy needs in the West.

3.2 CALIFORNIA ADDITIONS SCENARIO

While the No Additions Scenario provides a useful reference point for ongoing planning and procurement activities that could influence the composition of the resource fleet in 2026, it does not provide a sense of how resource adequate the system might be 2026 given that some utility plans are likely to come to fruition between now and then. To offer additional insights into the nature of resource adequacy in 2026, we developed an alternative scenario (the California Additions Scenario) that incorporates additions in California

roughly consistent with the California's Public Utility Commission's Preferred System Plan through 2026. These additions are summarized in Table 3.8. Resource additions were allocated to CAISO BAs based on resource shares in the No Additions Scenario (for example, 54% of CAISO solar in the No Additions Scenario was in CISC, so 54% of the added solar was placed in CISC). For simplicity, we added these resources by scaling up BA-level aggregated resource shapes, which may tend to underestimate capacity factors and diversity benefits.

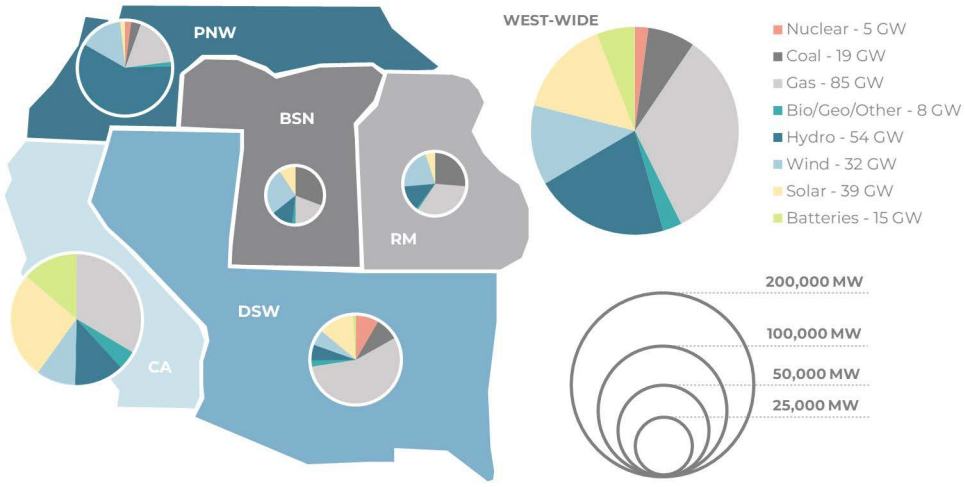
TABLE 3.8.

Resource additions in the California Additions Scenario, relative to the No Additions Scenario.

RESOURCE	TOTAL CAISO ADDITIONS
Biomass MW	+107
Geothermal MW	+184
Wind MW	+3,673
Utility-scale solar MW	+11,000
Storage MW	+12,749
Storage MWh	+51,780
Total MW	**+27,713**

FIGURE 3.18.

Resource composition in the California Additions Scenario.



In the West-wide simulation, the California Additions Scenario identified only seven events in 1,000 years of simulated conditions. This is well below the one-day-in-10-year standard, indicating that the 27.7 GW of additions was more than enough to meet the 8.8 - 9.9 GW of perfect capacity needs identified in the No Additions Scenario. When modeling CAISO as an island in this scenario, loss of load risk remained quite low and well below one day in 10 years.

Though not fully comparable, these results are similar to the findings of the SERVM analysis of the Preferred System Portfolio conducted by CPUC Staff, which identified an LOLE of 0.023 days every 10 years and EUE = 2.09 MWh/year. At a high level, this result demonstrates that portfolios of clean resources can make substantive contributions to resource adequacy—and that in the mid-2020s, resource adequacy in the West can likely be achieved without adding new fossil fuel resources.

TABLE 3.9.

Loss of load metrics in the California Additions Scenario, compared to the No Additions Scenario.

METRIC	CA ADDITIONS SCENARIO - WEST-WIDE	NO ADDITIONS SCENARIO - WRAP IMPORTS	NO ADDITIONS SCENARIO - WEST-WIDE
$LOLP_{year}$	0.6%	0.6%	69%
$LOLP_{day}$	0%	0%	0.50%
$LOLP_{hour}$	0%	0%	0.048%
LOLE (days/10yrs)	0.07	0.06	18.2
LOLH (hrs/yr)	0.01	0.02	4.23
EUE (MWh/yr)	8.96	19	13,797
EUE_{norm} (ppm)	0.0	0.09	19.4
$EUE^{\frac{day}{event}}$ (MWh/loss-of-load-day)	1,280	3,243	7,597
$EUE^{\frac{hour}{event}}$ (MW/loss-of-load-hour)	1,280	884	3,259

Though not fully comparable, these results are similar to the findings of the SERVM analysis of the Preferred System Portfolio conducted by CPUC Staff, which identified an *LOLE* of 0.023 days every 10 years and EUE = 2.09 MWh/year.[39] At a high level, this result demonstrates that portfolios of clean resources can make substantive contributions to resource adequacy—and that in the mid-2020s, resource adequacy in the West can likely be achieved

39 Available at: https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M451/K412/451412947.PDF

without adding new fossil fuel resources. While we found that the resource additions provided more capacity than was needed to achieve a one-day-in-10-year standard for this particular system, it is important to contextualize that finding with the broader set of uncertainties that could affect the grid between now and 2026, including load uncertainty, resource retirement uncertainty, and climate uncertainty.

3.3 LESS COAL SCENARIO

To explore the potential impacts of continued acceleration of coal retirements in the West, we developed a Less Coal Scenario that further accelerates coal retirements beyond those already announced through 2026. The Less Coal Scenario incorporated the additions in CAISO from the California Additions Scenario and retired an additional 10.9 GW of coal resources from the resource stack so that none of the BAs modeled within the WRAP subarea had remaining coal. This left 8.0 GW of coal remaining primarily in WACM, PSCO, and TEPC.

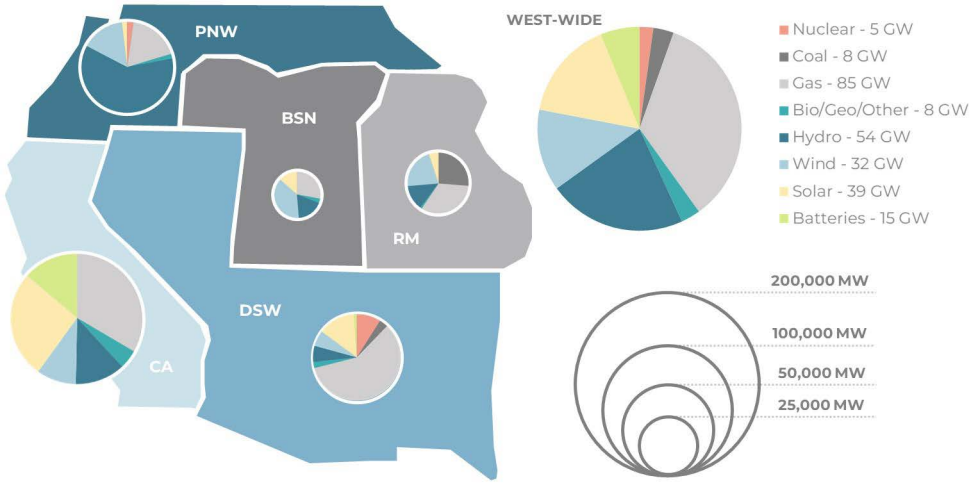
TABLE 3.10.

Resource additions and retirements in the Less Coal Scenario, relative to the No Additions Scenario.

RESOURCE	NET WEST-WIDE ADDITIONS
Biomass MW	+107
Geothermal MW	+184
Wind MW	+3,673
Utility-scale solar MW	+11,000
Storage MW	+12,749
Storage MWh	+51,780
Coal MW	-10,922
Total MW	**+16,791**

FIGURE 3.19.

Resource composition in the Less Coal Scenario.



3.3.1 WEST-WIDE ANALYSIS

Table 3.11 compares the loss of load statistics from the Less Coal Scenario to the No Additions Scenario. The Less Coal Scenario experienced fewer loss of load events than the No Additions Scenario, indicating that the resource adequacy contribution of the added resources in California exceeded that of the 10.9 GW of retired coal units. The system did not achieve a one-day-in-10-year standard, but had a relatively small *LOLE* of 4.13 days every 10 years.

TABLE 3.11.

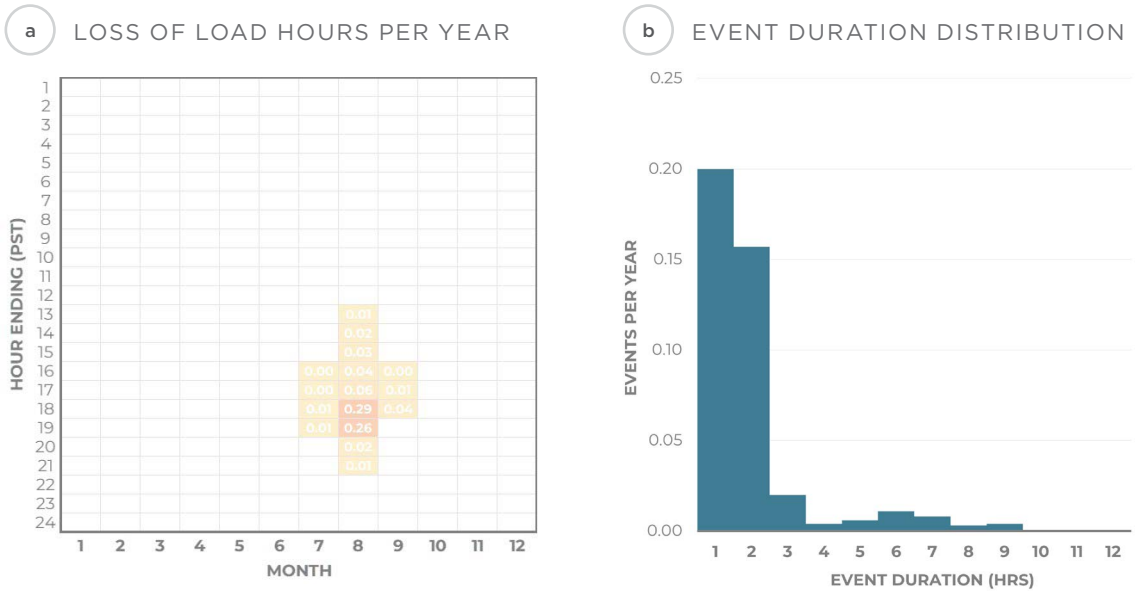
Loss of load metrics for the Less Coal Scenario, compared to the No Additions Scenario.

METRIC	LESS COAL SCENARIO	NO ADDITIONS SCENARIO
$LOLP_{year}$	29% ± 3%	69% ± 3%
$LOLP_{day}$	0.11% ± 0.01%	0.50% ± 0.02%
$LOLP_{hour}$	0.009% ± 0.001%	0.048% ± 0.001%
LOLE (days/10yrs)	4.13 ± 0.40	18.2 ± 0.8
LOLH (hrs/yr)	0.80 ± 0.06	4.23 ± 0.14
EUE (MWh/yr)	2,126	13,797
EUE_{norm} (ppm)	3.0	19.4
$EUE^{\frac{day}{event}}$ (MWh/loss-of-load-day)	5,147	7,597
$EUE^{\frac{hour}{event}}$ (MW/loss-of-load-hour)	2,650	3,259
Average Event Duration (hrs)	1.94	2.33

As shown in Figure 3.20, loss of load events were primarily experienced in August in this case, with a much smaller likelihood in July and September, and the vast majority of events (86%) were two hours or less. Unlike the No Additions Scenario, the Less Coal Scenario did see a small number of longer 9-hour events.

FIGURE 3.20.

Loss of load event timing and duration in the Less Coal Scenario.



Perfect capacity needs in the Less Coal Scenario were relatively small across all three of the RA standards listed in Table 3.12, and, as shown in Figure 3.21, were minimally affected by duration limitations.

TABLE 3.12.

Perfect capacity needs in the Less Coal Scenario.

RA STANDARD	PERFECT CAPACITY NEED IN LESS COAL SCENARIO
LOLE = One day in 10 years	3.3 - 4.2 GW
$LOLP_{year}$ = 5%	4.9 GW
LOLH = 2.4 hrs/yr	0 GW

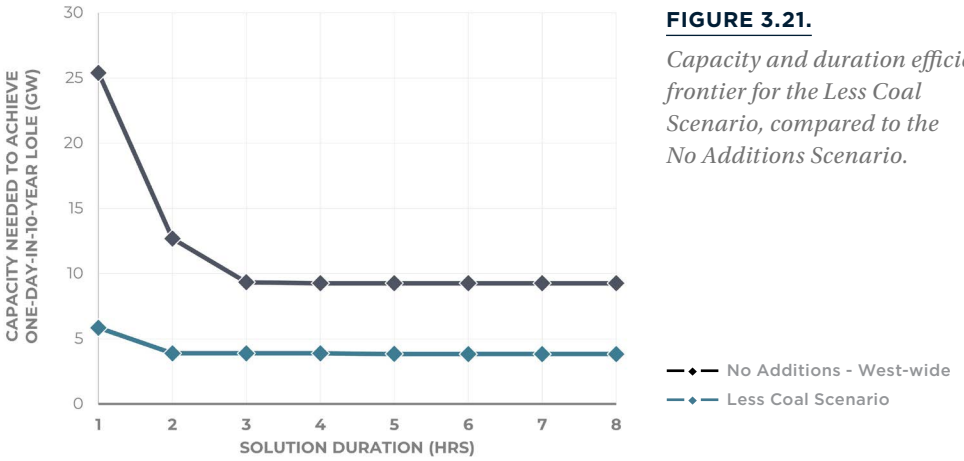


FIGURE 3.21.

Capacity and duration efficient frontier for the Less Coal Scenario, compared to the No Additions Scenario.

Despite the retirement of almost 11 GW of coal resources, the Less Coal Scenario relied on additional batteries to eliminate most of the capacity shortages from the No Additions Scenario. Resource availability on an example day from the two scenarios is shown in Figure 3.22. Panel A shows this day in the No Additions Scenario and panel B shows it for the Less Coal scenario. On this challenging day, storage charged early in the day and shifted energy to the evening, compensating for the reduced coal availability and reducing the amount of unserved energy. Notably, even after the coal retirements in the Less Coal scenario, the system did not appear to be energy-limited with plentiful resources available throughout the week to charge the storage and shift the needed energy to avoid shortages in the afternoon and evening.

Despite the retirement of almost 11 GW of coal resources, the Less Coal Scenario relied on additional batteries to eliminate most of the capacity shortages from the No Additions Scenario.

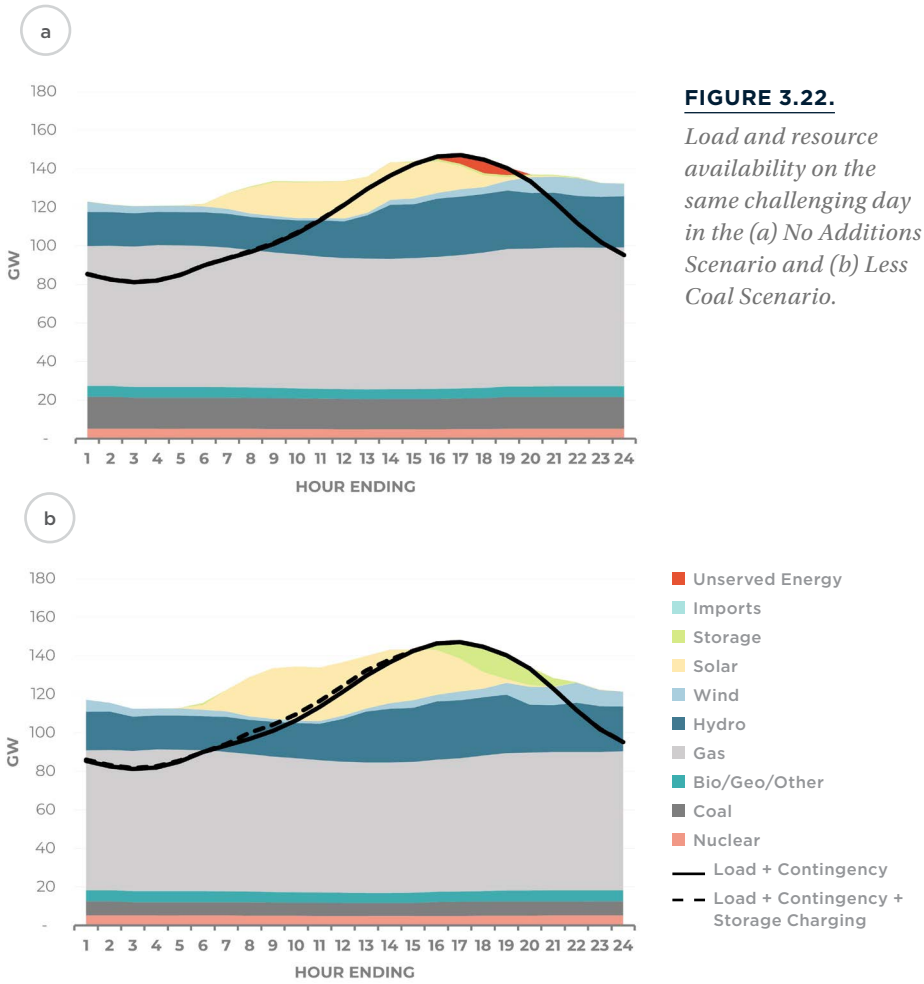


FIGURE 3.22.

Load and resource availability on the same challenging day in the (a) No Additions Scenario and (b) Less Coal Scenario.

Regional dynamics also changed: in the Less Coal scenario, California, with its added resources, eliminated the lost load and became a net exporter 51 percent of the hours in which it experienced unserved energy despite importing energy from the rest of the West in the No Additions Scenario. Import levels were much lower in the rest of the hours and RA events did not occur any other hours in California in the Less Coal scenario.

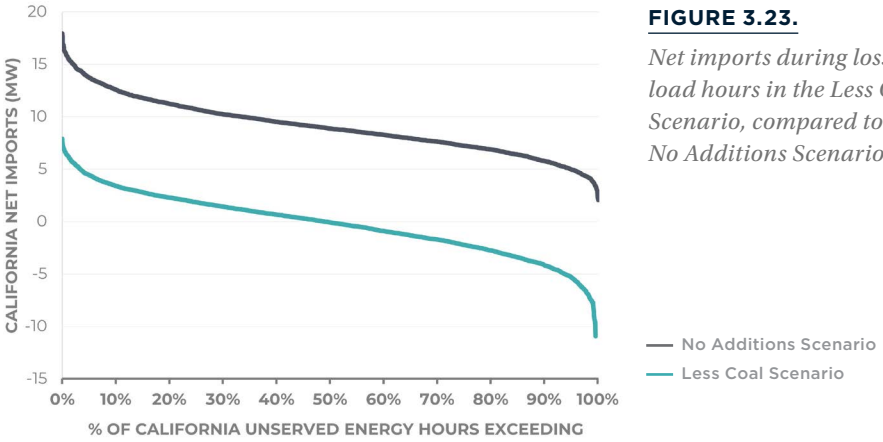


FIGURE 3.23.

Net imports during loss of load hours in the Less Coal Scenario, compared to the No Additions Scenario.

3.3.2 SUBAREA ANALYSIS

The subarea analysis described in Section 3.1.5 was repeated for the Less Coal Scenario and is described in the following sections.

3.3.2.1 WRAP subarea

Table 3.13 lists the loss of load metrics for the WRAP subarea in the Less Coal Scenario under both the Island and Import policies described in Section 3.1.5. When WRAP was treated as an island, the Less Coal Scenario resulted in a very high likelihood of lost load—100% of the simulated years saw a shortage and the LOLE was 451 days every 10 years. However, adopting the Import Policy avoided nearly all of these shortages and resulted in loss of load risk in WRAP that closely aligned with the West-wide findings.

TABLE 3.13.

Loss of load metrics for the WRAP subarea in the Less Coal Scenario, compared to West-wide loss of load metrics.

METRIC	LESS COAL - WRAP ISLAND	LESS COAL - WRAP IMPORTS	LESS COAL SCENARIO - WEST-WIDE
$LOLP_{year}$	100%	29%	29%
$LOLP_{day}$	12.4%	0.11%	0.11%
$LOLP_{hour}$	2.24%	0.009%	0.009%
LOLE (days/10yrs)	451	4.13	4.13
LOLH (hrs/yr)	196	0.80	0.80
EUE (MWh/yr)	275,929	2,118	2,126
EUE_{norm} (ppm)	808	6.2	3.0
$EUE^{\ day}_{\ event}$ (MWh/loss-of-load-day)	6,116	5,128	5,147
$EUE^{\ hour}_{\ event}$ (MW/loss-of-load-hour)	1,409	2,641	2,650
Average Event Duration (hrs)	4.34	1.94	1.94

Figures 3.24 and 3.25 show the timing and duration of events under the Island and Import policies for the WRAP subarea in the Less Coal Scenario. When WRAP was treated as an island, shortages were observed in all months, except March, and April, and could be very long in duration in both summer and winter seasons. However, many of these apparent challenges were alleviated by accounting for imports. In winter months and in many summer hours, there was enough excess generating capability in the rest of the West (and adequate transmission) to avoid most shortages in the WRAP subarea. The remaining shortages were very similar in timing and duration to those identified in the West-wide Less Coal Scenario (see Figure 3.20).

FIGURE 3.24.

Loss of load event timing and duration for the WRAP subarea without imports in the Less Coal Scenario.

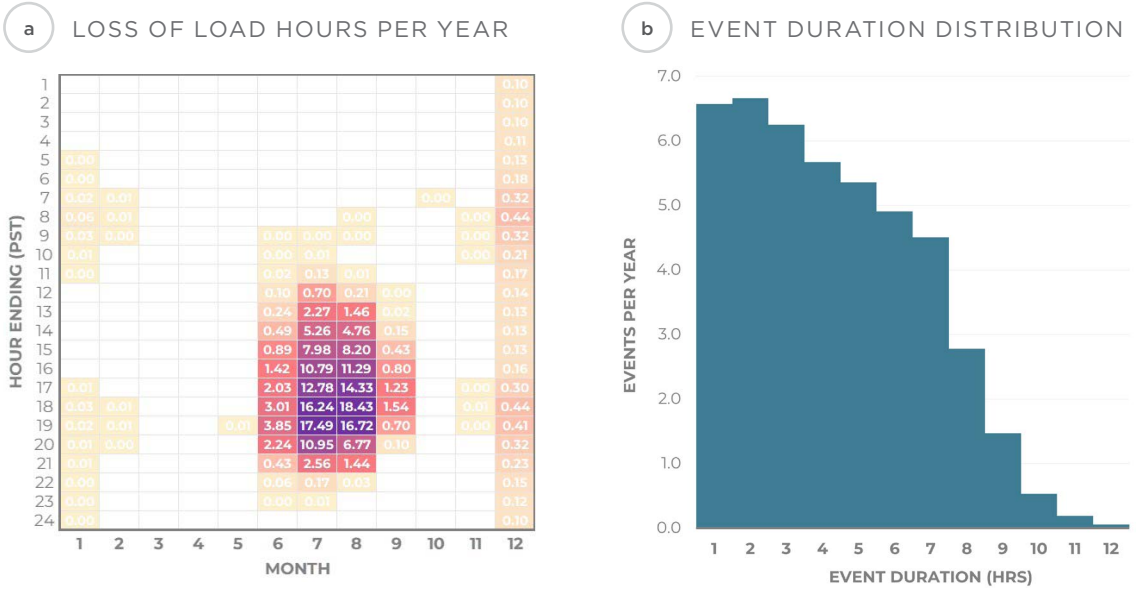


FIGURE 3.25.

Loss of load event timing and duration for the WRAP subarea with imports in the Less Coal Scenario.

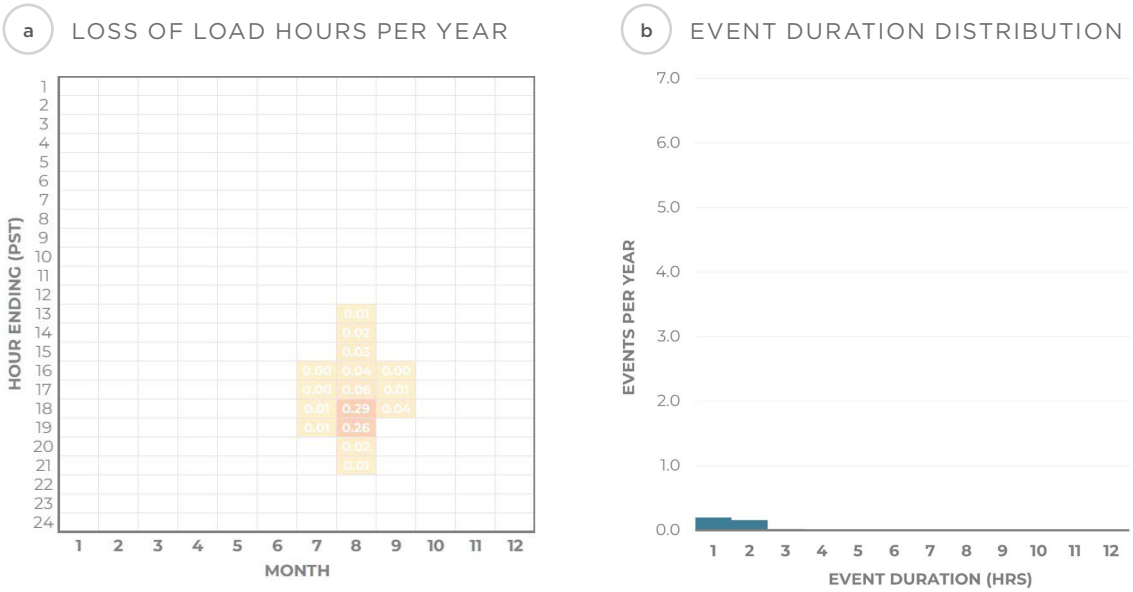


Table 3.14 summarizes the perfect capacity needs under various RA standards for the WRAP subarea under both the Island and Import Policies in the Less Coal Scenario. Across all three RA standards, the WRAP Island case yielded capacity needs that were similar in size to the amount of additional coal retired. However, accounting for imports eliminated the majority of these needs. Depending on the RA standard, accounting for imports avoided about 6-7 GW of perfect capacity needs.

Across all three RA standards, the WRAP Island case yielded capacity needs that were similar in size to the amount of additional coal retired. However, accounting for imports eliminated the majority of these needs.

TABLE 3.14.

Perfect capacity needs for the WRAP subarea in the Less Coal Scenario, compared to the West-wide perfect capacity needs.

RA STANDARD	LESS COAL SCENARIO - WRAP ISLAND	LESS COAL SCENARIO - WRAP IMPORTS	LESS COAL SCENARIO - WEST-WIDE
LOLE = One day in 10 years	10.1 GW	3.8 GW	3.8 GW
$LOLP_{year}$ = 5%	10.9 GW	4.9 GW	4.9 GW
LOLH = 2.4 hrs/yr	7.1 GW	0 GW	0 GW

Similar to the findings for the CAISO subarea in the No Additions Scenario, the difference between the capacity needs in the Island and Import cases was even greater as energy limitations were applied, as shown in Figure 3.26 for the one-day-in-10-years RA standard. In this case, accounting for imports not only reduced the capacity needs substantially, but also eliminated the sensitivity to solution duration. In this situation, imports offer a substitute for long duration solutions.

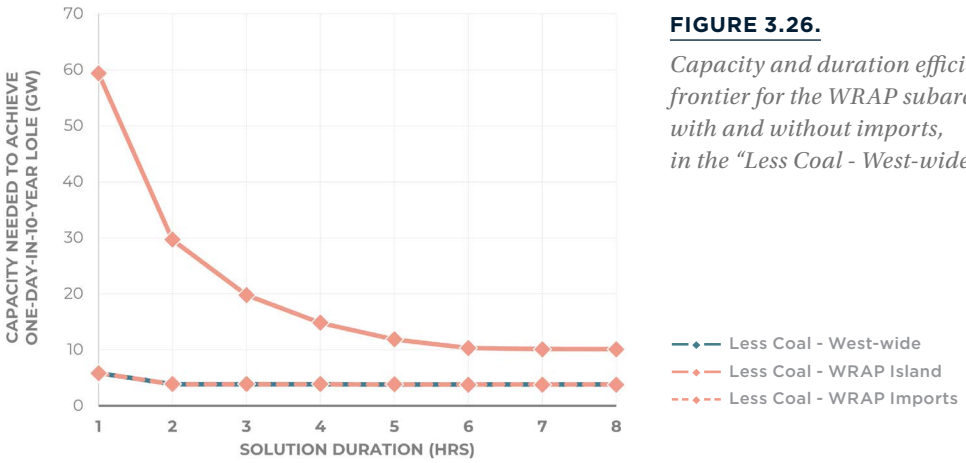


FIGURE 3.26.

Capacity and duration efficient frontier for the WRAP subarea, with and without imports, in the "Less Coal - West-wide.

3.3.2.2 CAISO subarea

Table 3.15 lists the loss of load metrics for the CAISO subarea under both the Island and Import Policies in the Less Coal Scenario. Because the resources in the CAISO subarea in this scenario were identical to those in the California Additions Scenario, the CAISO Island Policy loss of load risk was identical to that in the California Additions Scenario. In this case, accounting for imports had very little impact, because loss of load risk was so small before accounting for imports.

TABLE 3.15.

Loss of load metrics for the CAISO subarea in the Less Coal Scenario, compared to West-wide loss of load metrics.

METRIC	LESS COAL SCENARIO - CAISO ISLAND	LESS COAL SCENARIO - CAISO IMPORTS	LESS COAL SCENARIO - WEST-WIDE
$LOLP_{year}$	0.6%	0.6%	29%
$LOLP_{day}$	0%	0%	0.11%
$LOLP_{hour}$	0%	0%	0.009%
LOLE (days/10yrs)	0.06	0.06	4.13
LOLH (hrs/yr)	0.02	0.007	0.80
EUE (MWh/yr)	19	3.4	2,126
EUE_{norm} (ppm)	0.09	0.02	3.0
$EUE^{\ day}_{\ event}$ (MWh/loss-of-load-day)	3,243	559	5,147
$EUE^{\ hour}_{\ event}$ (MW/loss-of-load-hour)	884	479	2,650
Average Event Duration (hrs)	3.67	1.17	1.94

3.3.2.3 Benefits of planning coordination

In the No Additions Scenario, we found that most of the benefits of full coordination across the West could be achieved if the CAISO subarea were to adopt the Import policy described in Section 3.1.5. However, in the Less Coal Scenario, capacity needs were much more sensitive to the import policy adopted for the WRAP subarea. This was because the CAISO subarea was adequate without imports while the WRAP subarea relied much more on imports due to the additional coal retirements. The importance of the WRAP subarea import policy is demonstrated for the 4-hr capacity needs in Table 3.16. In this case, all of the incremental benefits of full coordination between WRAP and CAISO could be achieved if the WRAP subarea adopted the Import Policy.

TABLE 3.16.

4-hr capacity needs with various import policies adopted for the CAISO and WRAP subareas, compared to full coordination, in the Less Coal Scenario.

CAISO POLICY	WRAP POLICY	4-HR CAPACITY NEED IN CAISO + WRAP SUBAREAS	OVERBUILD LOWER BOUND
Island	Island	14.8 GW	10.9 GW
Island	Imports	3.9 GW	0
Imports	Island	14.8 GW	10.9 GW
Imports	Imports	3.9 GW	0
Full Coordination		3.9 GW	-

Again, it is important to keep in mind that the subarea analyses each assume full coordination within their footprints, so the incremental benefit of full coordination across the footprints may be relatively small. If we were to calculate the potential overbuild without coordination within each subarea (i.e., each BA may be treated as an island or adopt an import policy), the benefits of full coordination would be much larger.

When taken together, the No Additions Scenario and Less Coal Scenario demonstrate that the benefits of adopting import policies can vary widely depending on the system. Our findings support the general intuition that the shorter a system is on its own, the more important it is for that system to account for imports. In the No Additions Scenario, the CAISO subarea is quite short and the WRAP subarea is nearly adequate on its own. To avoid overbuild, the priority is for the CAISO subarea to adopt a reasonable import policy. In the Less Coal Scenario, the relative positions have switched—the CAISO subarea is adequate on its own and the WRAP subarea is quite short—so the WRAP import policy is most critical for avoiding overbuild.

Key Takeaways from the Less Coal Scenario

- The addition of the CPUC renewable and storage resources more than compensated for the retirement of the WRAP coal resources in the Less Coal Scenario, resulting in lower loss of load probability and smaller capacity shortages than the No Additions Scenario.
- The remaining shortages were primarily concentrated in August afternoons and evenings, and event durations remained relatively short with 92% of events lasting for 4 hours or less.
- The Less Coal system did not appear to be energy-limited, with plentiful resources available to charge storage outside of the critical system hours. Storage dispatch helped to eliminate a large fraction of the unserved energy events from the No Additions Scenario.
- Regional dynamics were fundamentally different, with California becoming a net exporter in the Less Coal scenario in 51% of the hours in which it experienced loss of load events despite importing energy from the rest of the West in the No Additions Scenario.

4

WEATHER-SYNCHRONIZED SIMULATIONS

Despite the retirement of almost 11 GW of coal resources, the Less Coal Scenario relied on additional batteries to eliminate most of the capacity shortages from the No Additions Scenario.

To further explore the nature of RA challenges in the West and the limitations of Monte Carlo analysis, we also conducted analysis of the No Additions Scenario using Weather-Synchronized Simulation. In this simulation, we tested the full synchronized record of load, wind, solar, and temperature conditions (2007-2020)[40] across each hydro year for which data was available (2001-2020), resulting in 14 x 20 = 280 years of potential weather and hydro conditions. For each of the 280 sets of weather/hydro conditions, we tested 30 years of potential forced outage conditions, generated using Monte Carlo Simulation with exponential failure and repair models.[41] This resulted in 8,400 years of potential conditions. To reduce runtimes, we first tested the 280 years of weather/hydro conditions under extreme forced outage conditions, where total thermal capacity was derated by 10% relative to maximum output. We then ran the full 30 iterations only on those weeks in which lost load was observed under the extreme forced outage conditions. This reduced the number of weeks that had to be simulated in GridPath from 436,800 to 29,430.

4.1 NO ADDITIONS SCENARIO RESULTS

The Weather-Synchronized Simulation of the No Additions Scenario resulted in higher $LOLE$ and $LOLH$ than the Monte Carlo Simulation, larger shortages, and slightly longer event durations, although all events were still 8 hours or less (see Table 4.1). Interestingly, the Weather-Synchronized Simulation resulted in a lower $LOLP_{year}$ than the Monte Carlo Simulation, suggesting that the Monte Carlo Simulation tended to distribute critical weather conditions more evenly across years (as opposed to concentrating them more into some weather years than others). This is not surprising, as the Markov Chain model only considered weather day transitions from one day to another, without accounting for longer term weather phenomena that could concentrate risk into particularly challenging seasons or years.

To investigate how sensitive these findings were to the particular weather years that were simulated, we also calculated loss of load metrics considering only those years for which wind data was available from the Wind Toolkit (2007-2014) and excluded the years over which we relied on synthesized wind data (2015-2020). These results are summarized in Table 4.1. In this test, we saw significantly lower RA risk ($LOLE$ = 10.9 vs. 24.9 days every 10 years) because extreme weather events occurred less frequently between

40 This record includes some synthesized data, as described in Appendix C.

41 To investigate the relative importance of random forced outage modeling, the first of the 30 iterations used a simple capacity derate for each unit equal to its forced outage rate.

2007 and 2014 than they did between 2015 and 2020. This test confirms that Weather-Synchronized Simulations over limited historical time periods may not accurately reflect RA risks and it highlights the importance of expanding publicly available datasets to more years, especially hourly wind datasets. This test also suggests that RA analysis, using Weather-Synchronized or Monte Carlo analysis, may be highly dependent on the weather conditions experienced during the particular years for which high resolution data is available. We explore this sensitivity further with a statistical analysis across a much broader set of weather conditions in Section 4.3.

TABLE 4.1.

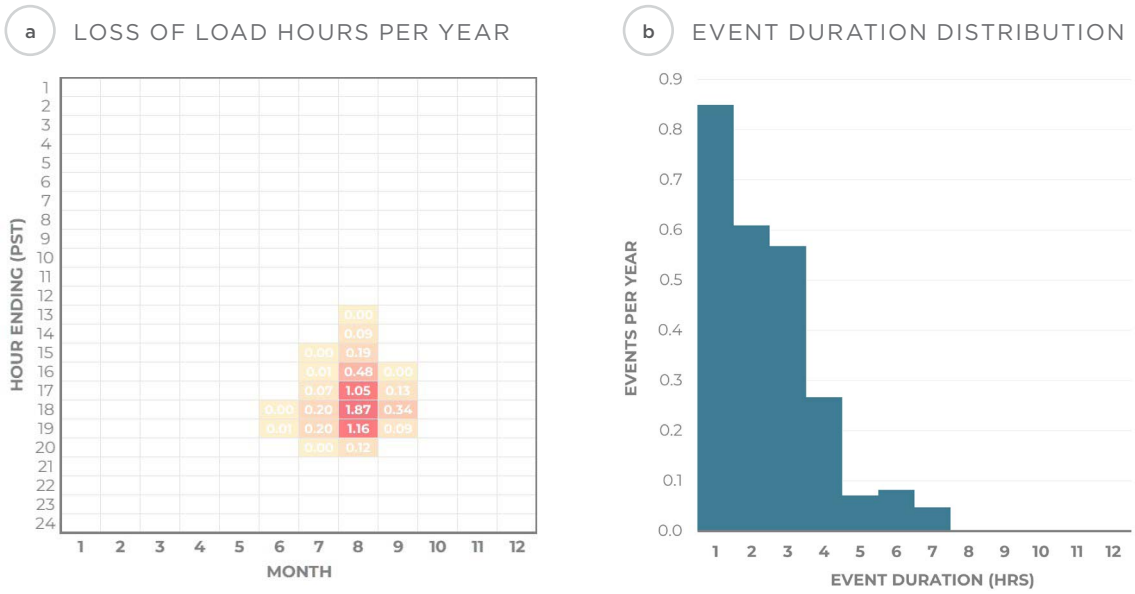
Loss of load metrics in the No Additions Scenario under Weather-Synchronized Simulation over different ranges of years, compared to Monte Carlo Simulation.

METRIC	NO ADDITIONS SCENARIO (WEATHER-SYNCHRONIZED 2007-2020)	NO ADDITIONS SCENARIO (WEATHER-SYNCHRONIZED 2007-2014)	NO ADDITIONS SCENARIO (MONTE CARLO)
$LOLP_{year}$	61.9%	49.4%	69%
$LOLP_{day}$	0.68%	0.30%	0.50%
$LOLP_{hour}$	0.069%	0.028%	0.048%
LOLE (days/10yrs)	24.9	10.9	18.2
LOLH (hrs/yr)	6.02	2.43	4.23
EUE (MWh/yr)	20,365	5,965	13,797
EUE_{norm} (ppm)	28.6	8.4	19.4
$EUE^{\frac{day}{event}}$ (MWh/loss-of-load-day)	8,171	5,449	7,597
$EUE^{\frac{hour}{event}}$ (MW/loss-of-load-hour)	3,382	2,452	3,259
Average Event Duration (hrs)	2.41	2.22	2.33

The timing of events in the Weather-Synchronized Simulation over 2007-2020 was similar to the Monte Carlo Simulation, except that more of the loss of load risk was concentrated in the most challenging hours of August.

FIGURE 4.1.

Loss of load event timing and duration in the No Additions Scenario under Weather-Synchronized Simulation.



Perfect capacity needs to meet a one-day-in-10-year standard in the Weather-Synchronized Simulation covering 2007-2020 weather years were 11.1 GW, about 1.8 GW higher than in the Monte Carlo Simulation (which covers load years 2006-2020). When the years with synthesized wind data (2015-2020) were excluded from the Weather-Synchronized Simulation, identified capacity needs were much smaller (6.4 GW).

TABLE 4.2.

Perfect Capacity needs in the No Additions Scenario under Weather-Synchronized Simulation over different ranges of years, compared to Monte Carlo Simulation.

RA STANDARD	NO ADDITIONS SCENARIO (WEATHER-SYNCHRONIZED 2007-2020)	NO ADDITIONS SCENARIO (WEATHER-SYNCHRONIZED 2007-2014)	NO ADDITIONS SCENARIO (MONTE CARLO)
LOLE = One day in 10 years	10.9 - 11.4 GW	6.3 - 6.6 GW	8.8 - 9.9 GW

To understand why the Weather-Synchronized Simulation identified greater RA risk than the Monte Carlo Simulation, we examined specific load circumstances in both the Weather-Synchronized Simulation and the Monte Carlo Simulation. Figure 4.2 shows the loss of load probability for each simulated load day in July - September 2020. Recall that in the Monte Carlo Simulation, these days

are randomly drawn many times throughout the simulation and each time they are paired with different wind, solar, and thermal conditions from similar weather days in the historical record. The *LOLP* corresponding to the Monte Carlo Simulation in the figure is equal to the number of times that unserved energy was observed when the specified load day was randomly drawn, divided by the total number of times that the load day was drawn. In the Weather-Synchronized Simulation, each day with synchronized weather conditions was tested across various hydro and forced outage conditions. The *LOLP* corresponding to the Weather-Synchronized Simulation in the figure is equal to the number of times that unserved energy was observed for the specified weather day, divided by the total number of hydro and forced outage iterations that were tested (20 hydro years x 30 forced outage iterations = 600 total iterations for each weather day).

This tendency of the Monte Carlo Simulation to underestimate risk on the most challenging load days was observed more generally across the simulated days. We found that mixing and matching load, wind, solar, and thermal conditions from days with similar, but not identical, weather in the Monte Carlo Simulation tended to overestimate resource availability on those days.

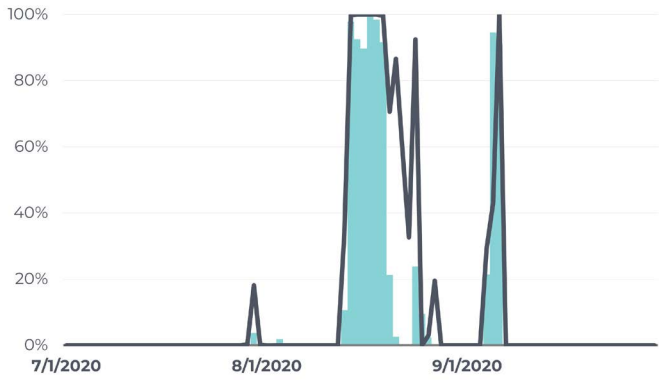


FIGURE 4.2.
Daily LOLP on example load days in Weather-Synchronized and Monte Carlo Simulations.

During the summer of 2020, there were three load days in which the Monte Carlo Simulation identified greater RA risk than the historical weather conditions suggested (8/3/2020, 8/25/2020, and 9/5/2020), but there were several more days in which the actual historical weather conditions yielded higher *LOLP* than the estimates provided by the Monte Carlo Simulation. This tendency of the Monte Carlo Simulation to underestimate risk on the most challenging load days was observed more generally across the simulated days. We found that mixing and matching load, wind, solar, and thermal conditions from days with similar, but not identical, weather in the Monte Carlo Simulation

tended to overestimate resource availability on those days. Figure 4.3 and Figure 4.4 show this overestimation, averaged across the load days with non-zero *LOLP* in the Weather-Synchronized Simulation (i.e., the "critical load days"). Notably, the resource availability overestimation was greatest in some of the hours of the day with the highest loss of load risk (HE 16 - HE 18) and the average overestimation in those hours was similar in size to the capacity shortage differences between the Weather-Synchronized and Monte Carlo Simulations.

FIGURE 4.3.

Resource availability on critical load days in Weather-Synchronized and Monte Carlo Simulations.

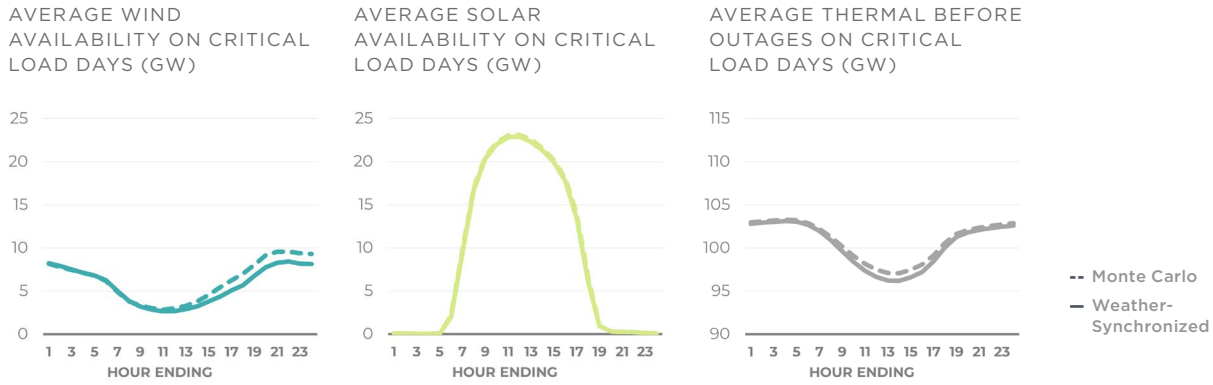
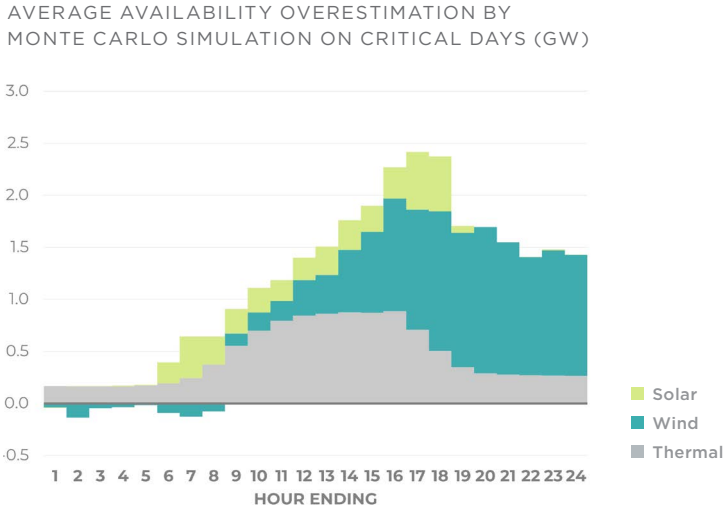


FIGURE 4.4.

Average difference in resource availability on critical load days between Weather-Synchronized and Monte Carlo Simulations.



How well a Monte Carlo Simulation captures the risk on these critical days will depend on how the historical days are binned and what other measures are taken to account for correlations. Providing this type of simple shape analysis with a Monte Carlo Simulation may help planners contextualize Monte Carlo findings and provide a high level sense of the extent to which the Monte Carlo approach may be underestimating RA risk.

This tendency of the Monte Carlo Simulation to underestimate risk on the most challenging load days was observed more generally across the simulated days. We found that mixing and matching load, wind, solar, and thermal conditions from days with similar, but not identical, weather in the Monte Carlo Simulation tended to overestimate resource availability on those days.

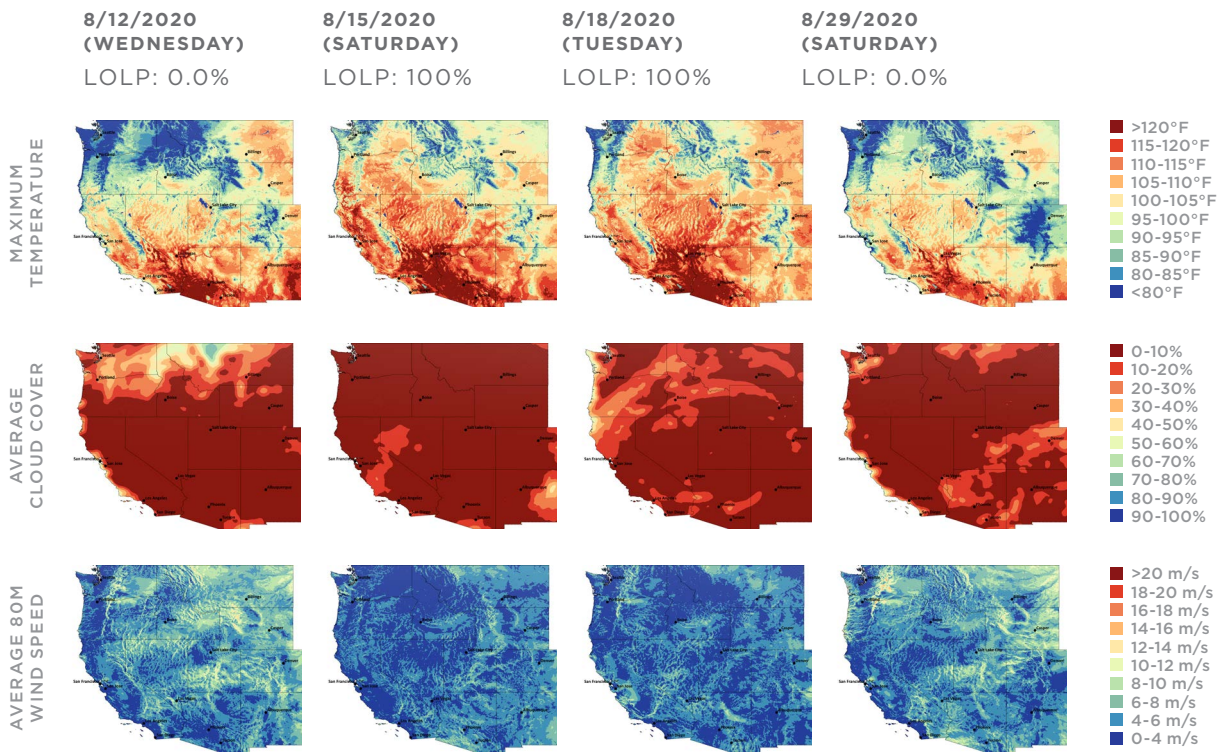
4.2 WEATHER INSIGHTS

One benefit of Weather-Synchronized Simulation is the ability to identify the coherent weather conditions, and the underlying physical weather phenomena that pose the greatest RA risk. At a high level, we find that temperature remains the key weather driver of loss of load risk in this system. Notably, the days with non-zero loss of load probability see abnormally high temperatures in California and the days with the greatest loss of load probability also see abnormally high temperatures across most of the West. The events in August 2020 serve as an example of the most challenging type of weather phenomenon for near-term RA in the Western United States. The August 2020 heat event was uncharacteristically hot across most of the West from August 14th through August 19th, with the highest region-wide temperatures at the major load centers occurring on August 15th and 16th (see Table 4.3). The coincidence of unusually hot conditions at most load centers across the West during this event resulted in very high loss of load risk in the simulation. This event was driven by a weather phenomenon known as the West Coast Thermal Trough (WCTT), a self-reinforcing cycle that pushes air from the desert southwest northward between the Sierra/Cascades and the coast, which can bring coincident well-above-average temperatures to California and Western Oregon/Washington. Figure 4.5 shows maximum daily temperatures, average daily cloud cover, and average daily 80-meter wind speeds across the West during four days in August 2020—a day before the arrival of the widespread heat event, two days during the heat event, and a day well after the event subsided.[42]

42 All weather maps in this report were created with data from the NOAA High-Resolution Rapid Refresh (HRRR) analysis dataset, accessed via the HRRR Data Archive: AWS Open Data Program (https://mesowest.utah.edu/html/hrrr/)

FIGURE 4.5.

Regional weather conditions before, during, and after the August 2020 heat event.



August 2020 Weather Event Analysis

AUTHOR | *Justin Sharp, Sharply Focused*

The August 2020 heat event was a manifestation of a phenomenon known as the West Coast Thermal Trough (WCTT),[43] a self-reinforcing cycle that has significant impacts on temperature, humidity and wind, most typically in the regions between the Sierra Nevada/Cascade Mountains and the West Coast.[44] During these events, large scale weather conditions result in high pressure building to the east of the mountain barriers resulting in easterly flow across the mountains. This easterly flowing air is compressed as it descends off the mountains causing warming and an extension of the low pressure area that typically exists over the Desert Southwest due to the strong heating that is prevalent there; this further enhances the offshore flow and moves the trough northwards. Eventually, a trigger, such as a

43 A trough is a local minimum in air pressure. A thermal trough is a pressure minimum created by locally relatively warm air.

44 Brewer, Matthew & Mass, Clifford & Potter, Brian. (2013). The West Coast Thermal Trough: Mesoscale Evolution and Sensitivity to Terrain and Surface Fluxes. Monthly Weather Review. 141. 2869-2896. 10.1175/MWR-D-12-00305.1.

weak weather system approaching the West Coast, pushes the trough inland and cool air floods inland from the ocean. The trough will then either migrate across the mountains and merge with the Great Basin thermal low, or recede back to California and dissipate. The process often results in the final warm day in any given location exhibiting the highest temperature of the event,[45] often followed by a sharp transition to cooler temperatures. Almost all heat waves in West Coast states are associated with the WCTT. However, the orientation and extent of the WCTT is dictated by both large scale and local scale conditions, which produce significantly different outcomes because of the sharp contrast in marine and continental air mass characteristics. Some configurations allow for offshore flow to extend along the entire West Coast producing coincident heat at many or all major western load centers, especially in summer and early autumn. In other orientations, some combination of the San Diego area, LA Basin or the San Francisco Bay may experience much cooler onshore flow as the trough brings high temperatures to regions further north and/or inland, or the trough may never extend beyond southern Oregon before dissipating so that conditions at Pacific Northwest load centers remain mild.

TABLE 4.3.

Maximum daily temperatures during August 2020 heat event.

	SEA	PDX	SFO	SMF	LAX	SAN	PHX
8/12/2020	71°F	74°F	64°F	97°F	73°F	76°F	111°F
8/13/2020	75°F	78°F	83°F	100°F	83°F	84°F	115°F
8/14/2020	81°F	88°F	84°F	106°F	81°F	87°F	117°F
8/15/2020	88°F	99°F	86°F	109°F	87°F	85°F	114°F
8/16/2020	98°F	95°F	81°F	112°F	93°F	78°F	115°F
8/17/2020	88°F	93°F	81°F	105°F	78°F	83°F	115°F
8/18/2020	81°F	84°F	77°F	106°F	85°F	83°F	115°F
8/19/2020	83°F	87°F	77°F	101°F	85°F	82°F	115°F
8/20/2020	80°F	84°F	72°F	88°F	85°F	85°F	112°F
8/21/2020	71°F	78°F	75°F	92°F	81°F	85°F	100°F

The August 2020 event was complex in its details; major load centers in California, Oregon and Washington experienced

45 The period leading up to this transition day is often also characterized by dry, windy, easterly flow conditions that bring greatly elevated wildfire risk, but provide little resource for generation at typical wind farm locations that are sited to capitalize on the more prevalent westerly flow regime.

interspersed periods of cooler flow from the oceans, and hot flow descending from the mountains. However, coincident heat impacted most major load centers on August 15th and 16th. The event began on August 13th and the hottest day across California and Oregon load centers was August 15th, after which some coastal regions transitioned to onshore flow. The greatest coverage of the heat throughout California, Oregon, and Washington was August 16th, when Seattle reached 98°F (which is the second highest August day on record), and Olympia reached 99°F. The worst of the heat only lasted for one day in Washington, with temperatures moderating, though still remaining above normal, on August 17th. While some of the coastal California cities had cooled by a few degrees, the 16th was also the warmest day in several Central Valley locations including Sacramento and Fresno. August 18th and 19th saw further moderation with cooler air infiltrating further inland from the Bay Area northwards as the WCTT migrated inland and weakened. However, conditions became conducive to offshore flow again in southern California, with a larger portion of the LA Basin experiencing well above normal temperatures. This highlights the complexity that is often present during these events, where slight changes in the weather pattern can lead to profound differences in loads in major coastal and near coastal cities. By August 20th, the WCTT event had ended, along with the coincident heat from San Diego to Seattle. This is consistent with a reduction in the LOLP observed in both the Weather-Synchronized and Monte Carlo Simulations, despite continued well-above normal conditions in Southern California coastal cities.

There are significant and consistent correlations between this type of heat event and deviations from average wind and solar output. Most wind projects along the West Coast are sited to take advantage of the more prevalent onshore conditions (i.e., flow from ocean to land) in order to maximize energy production during more typical conditions. In these events, however, the flow shifts to offshore (i.e., from land to ocean), which is unfavorable for generation at most wind project locations, as can be seen in Figure 4.5. Offshore flow is also generally associated with clearing skies and strong solar resource performance. However, as the WCTT moves northwards, coastal load centers in California, where behind-the-meter solar is prevalent, typically transition to onshore flow, resulting in localized cooling and low clouds. This transition can also bring stronger winds to localized areas, as can be seen on August 18th and August 29th in Figure 4.5. This phenomenon underscores the importance of examining physically coherent weather across the West within RA analysis, as local conditions along the coast could have large and quickly shifting impacts on

load, distributed solar output, and wind production and may not be indicative of conditions elsewhere in the West.

The August 2020 event illustrates complex relationships between power system variables across time and geography and reveals how challenging it can be to capture key weather correlations over large areas in RA analysis. Recall that the wind and solar data for 2020 in the Weather-Synchronized Simulation is synthesized based on historical generation in BPA and CAISO on these days (see Appendix C), so the simulated conditions do not exactly match the weather conditions described in this section. However, the synthetic shapes generally reflect the same weather regime along the West Coast as was experienced historically on these days. More complete publicly available datasets for energy system modeling would increase confidence that Weather-Synchronized Simulation captures all of these dynamics, especially as the system relies more heavily on wind and solar.

Days with more geographically isolated heat, which may result in very high load conditions in localized parts of the West, did not tend to pose loss of load risk in the modeled system due to geographical diversity. In general, most of these hot days exhibit many of the same meteorological features just discussed and a WCTT is often present. However, details in the antecedent conditions and the evolving pressure pattern can profoundly change where the dividing line between onshore and offshore flow is found, and the initial temperature distribution of the air mass also dictates how extreme the temperature evolution will become and where. Three examples are shown in Figure 4.6.

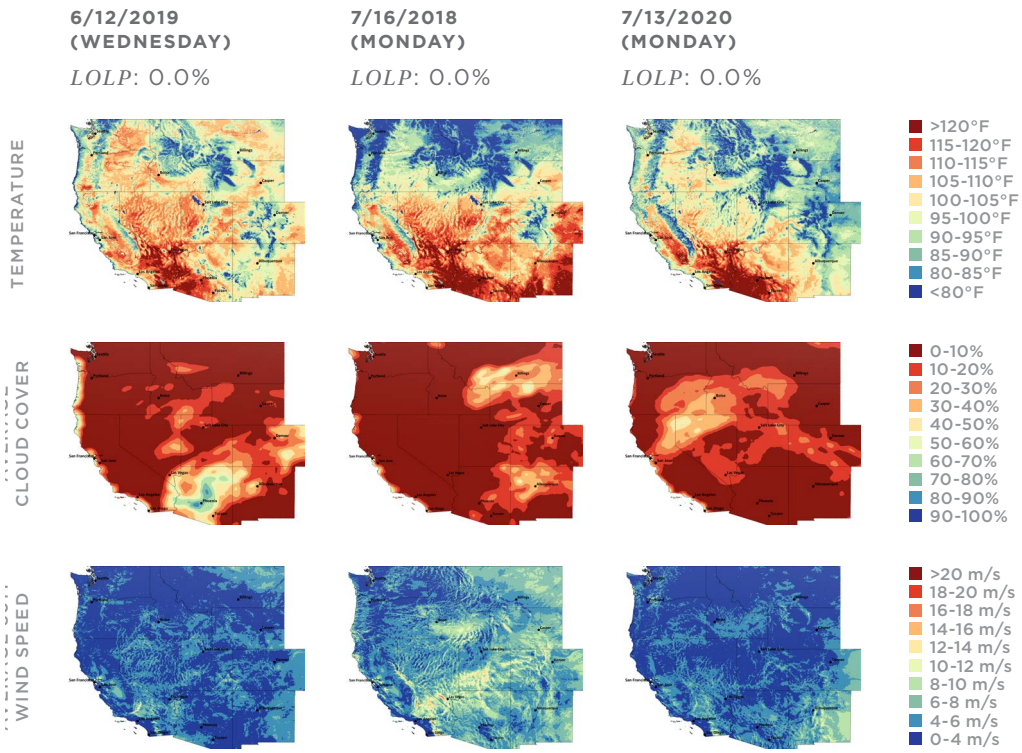
June 12th, 2019 was unusually hot in Portland (98°F) and Seattle (95°F) and interior California (Sacramento was 103°F) and the Southwest (112°F in Phoenix) were also warm. However it was relatively mild along coastal California load centers (72°F in Los Angeles, and 74°F in San Diego). During this event, Los Angeles and San Diego never transitioned to offshore flow, with temperatures peaking on the 10th at 80°F in LA and 77°F in San Diego. By the 12th, the flow was onshore along the entire California coastline and while the temperature on June 10th reached 97°F in San Francisco, on June 12th the high temperature only reached 79°F. July 16th, 2018 was also unusually hot in Seattle (92°F), Portland (98°F), and areas east of the Cascades and Sierras, but was less extreme along the California coast, with San Francisco experiencing the strongest onshore flow and a high of only 69°F while it was 79°F in Los Angeles and San Diego. The Southwest saw typical temperatures for this time of year (105°F in Phoenix), with cloud cover associated with monsoon rainfall. July 13th, 2020 saw extreme heat in the Desert Southwest (114°F in Phoenix) as well as the California Central Valley, but was marked by relatively mild

conditions across the Northwest (76°F in Seattle and 80°F in Portland) and cool and cloudy coasts.

Despite extreme heat occurring in parts of the West on each of these days, none of them saw loss of load risk in the Weather-Synchronized Simulation due to geographical load diversity. It is important to note that this load diversity is weather driven, and that the same weather drives wind and solar resource availability. In particular, the phenomenon that drives high demand also drives low wind generation. As penetration levels increase, it will become more crucial to fully account for these dynamics in resource adequacy studies, and to use dynamically consistent and coincident weather data for all fields.

FIGURE 4.6.

Regional weather on days with localized heat but no simulated loss of load risk in the No Additions Scenario.



4.3 STATISTICAL ANALYSIS

The synchronized weather case provides for more transparency into the relative drivers of RA risk via statistical analysis. To demonstrate this, we used logistic regression to estimate the loss of load probability on each simulated day as a function of the daily weather conditions (see Appendix E), number of daylight hours, hydro conditions, and whether the historical day fell on a weekend or weekday. Figure 4.7 compares the estimated *LOLE* using this approach to the *LOLE* based on Weather-Synchronized Simulation. The left panel shows each day in terms of a variable, w, which represents how strongly the day reflects the conditions that drive loss of load risk. The right panel compares the average *LOLE* between the Weather-Synchronized Simulation and the logistic regression estimate for each simulated year. More information about this analysis, including the steps taken to avoid overfitting, can be found in Appendix G.

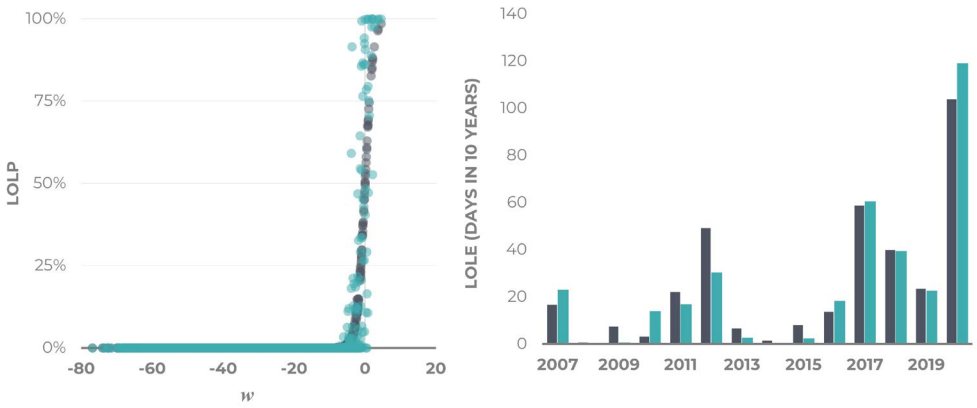


FIGURE 4.7

Logistic regression estimates for the No Additions Scenario

The logistic regression confirmed that weather conditions—and temperature in particular—are by far the biggest driver of loss of load risk.

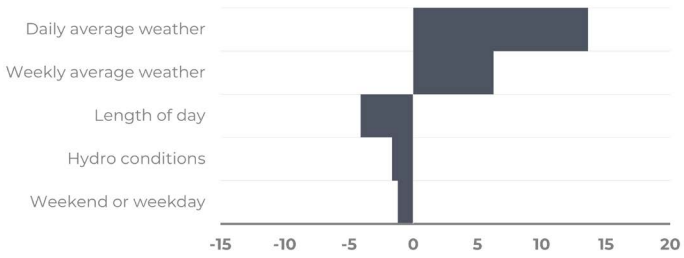


FIGURE 4.8.

Logistic regression coefficients for the No Additions Scenario.

We also found that higher wind speeds at key locations around the West, in particular Wyoming and Montana, tended to reduce loss of load risk in the No Additions Scenario, however it is difficult to discern the extent to which

this was due to wind generation or correlations between temperatures and wind speeds. We did not find that including dew point and sea level pressure information yielded materially improved statistical model performance for this particular system. These findings are specific to this particular system and portfolio of resources. Temperature may be a less dominant driver for systems with greater penetrations of renewables, in which other weather conditions may play a larger role.

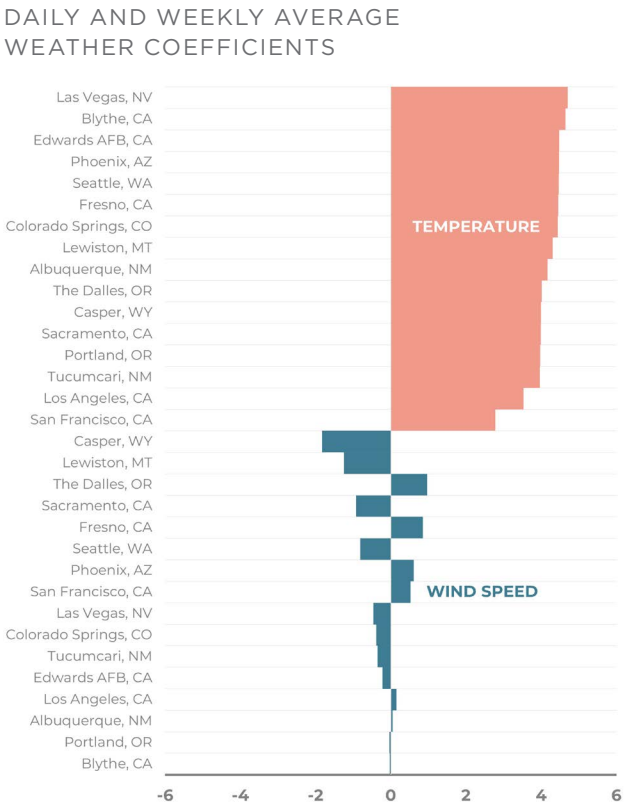


FIGURE 4.9.

Weather coefficients in the logistic regression for the No Additions Scenario.

4.3.1 HISTORICAL WEATHER PATTERNS

In addition to providing information about the drivers of RA risk for a given system, this analysis can be used to estimate loss of load risk during conditions that were not directly simulated, either historical weather conditions or simulated future conditions under climate scenarios. As an example, we used the logistic regression model to estimate loss of load risk during the historic heat dome event in June 2021, which broke several high temperature records across the Northwest (116°F in Portland and 108°F in Seattle). Despite the historic heat in parts of the West on June 28, 2021 (see Figure 4.10), we estimated a loss of load probably of only 1.2% on this day because other parts of the West saw much milder conditions.

6/28/2021 (MONDAY)
ESTIMATED LOLP: 1.2%

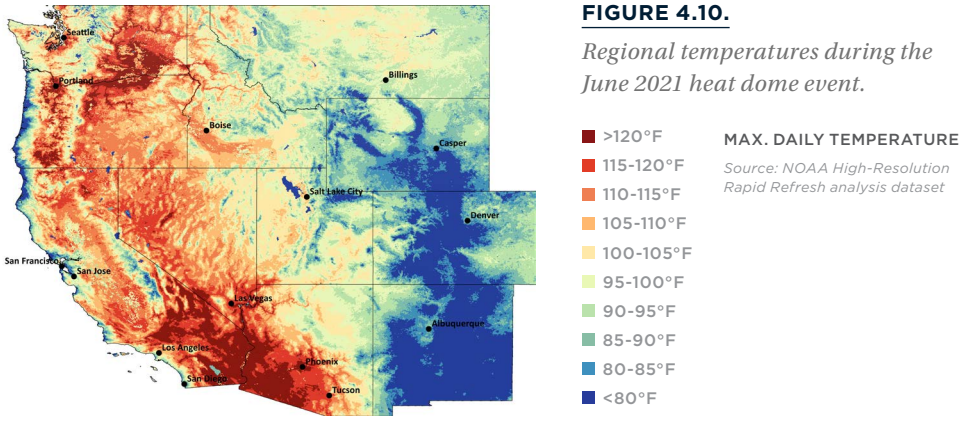


FIGURE 4.10.
Regional temperatures during the June 2021 heat dome event.

Because most RA analyses rely on historical weather to approximate future weather, we used the logistic regression model to examine potential RA risks over a much longer historical record, going back to 1949, to examine historical weather trends and their potential implications for RA analysis. Figure 4.11 shows the *LOLE* calculated based on each historical weather year, averaged across the 20 hydro years.

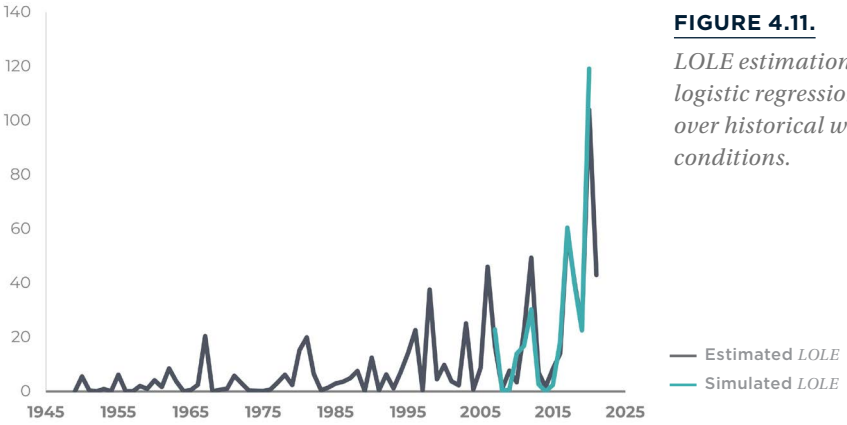


FIGURE 4.11.
LOLE estimation using logistic regression model over historical weather conditions.

As shown in Figure 4.11, the statistical analysis revealed some critical weather conditions in the historical record that were not simulated, but were likely to have caused RA challenges. The figure also shows that prior to the 1990s, the statistical model suggested that the weather conditions resulting in RA risk occurred less frequently. This can also be seen in Figure 4.12, which shows w (an indicator of how extreme each historical day was with respect to those factors that drive loss of load risk) for every day in the historical record. Figure 4.13 shows the same information, but zoomed in on the highest values in order to highlight the frequency of extreme events. More information about w can be found in Appendix G.

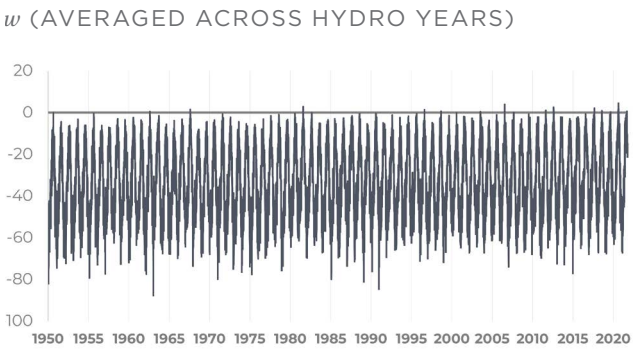


FIGURE 4.12.

Critical weather indicator, w, over the historical weather record.

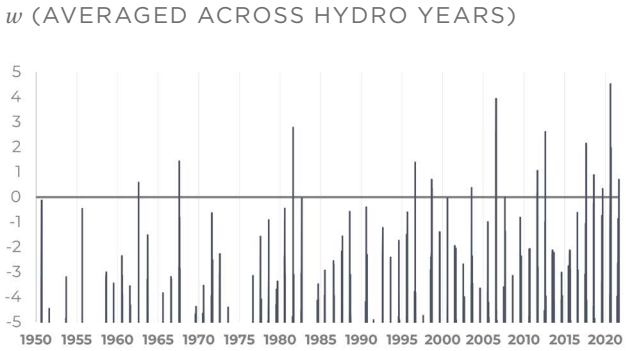


FIGURE 4.13.

Critical weather indicator, w, over the historical weather record (focusing on extreme events).

This trend in the frequency of critical weather has implications for RA analysis that uses historical weather data as a proxy for or indication of future weather conditions. Many RA analyses, including the Monte Carlo analysis used in this report, assume that future weather follows the distribution of weather conditions experienced over some historical period of time. Using the statistical model for the No Additions Scenario, we estimated the implications of this assumption across various historical weather records (see Figure 4.14). As more historical years were included in the analysis, and assumed to be equally likely as recent years, the lower the LOLE estimation became. While the observed trend suggests that it would be overly risky to assume that future weather conditions will reflect the distribution of weather across the last 70 years, it is not immediately clear that a particular historical period is necessarily superior for this type of analysis. Without information about potential future weather, via climate modeling and downscaling, the decision to rely

Without information about potential future weather, via climate modeling and downscaling, the decision to rely on historical weather and a particular number of years is effectively a policy decision. A risk averse planner might choose to plan as if future weather will look like the last 10 years and a planner with more risk tolerance might choose to plan as if future weather will look like the last 30 years.

on historical weather and a particular number of years is effectively a policy decision. A risk averse planner might choose to plan as if future weather will look like the last 10 years and a planner with more risk tolerance might choose to plan as if future weather will look like the last 30 years.

Some RA analyses attempt to account for historical trends by detrending the weather data underpinning the analysis. We explore this strategy in Appendix F.5 and find that while weather detrending may yield more stable results (i.e., the RA risk is not as sensitive to the number of historical years considered), this strategy further abstracts from actual weather conditions and the physical phenomena that drive them and may not accurately reflect the types of weather events that could actually be experienced on the system and their probabilities.

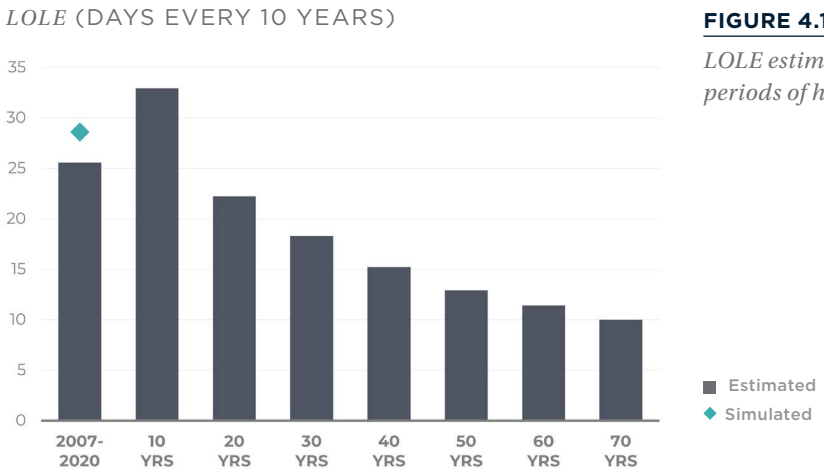


FIGURE 4.14.

LOLE estimation over various periods of historical weather.

4.4 COMPARISON TO MONTE CARLO APPROACH

At a high level, we found that the nature of the identified RA challenge was similar between Weather-Synchronized Simulation and Monte Carlo Simulation, as long as data from recent weather years were included in the Weather-Synchronized Simulation. The primary benefits of Weather-Synchronized Simulation over Monte Carlo were:

- Increased transparency into the physical weather conditions that drive RA risk
- Increased confidence in the accuracy of relevant weather correlations, both temporal and spatial
- The ability to derive statistical models based on lower resolution weather data to estimate RA risk during events for which high resolution weather data may not be available, including historical events or simulated future events under climate scenarios.

The primary drawback of Weather-Synchronized Simulation was its sensitivity to the specific weather years being simulated, which was exacerbated when the analysis was limited to publicly available high resolution wind data (2007-2014).

In Appendix F, we compare convergence behavior between the Monte Carlo and Weather-Synchronized Simulations and we test other modeling decisions to gain additional methodological insights. In this analysis, we found:

- The Monte Carlo analysis converged a bit more quickly than the Weather-Synchronized Simulation (in terms of the number of years simulated and the uncertainty in the resulting RA metrics), but this improved computational performance came at the cost of accuracy for this particular system. See Appendix F.1 and Appendix F.2.
- Modeling several iterations of forced outages across the 280 weather/hydro year combinations in the Weather-Synchronized Simulation did not materially impact the RA metrics, but did reduce the uncertainties in those metrics. For this system, estimating forced outages using derates equal to unit forced outage rates yielded reasonable estimates for LOLE and capacity need. This finding is specific to this particular system and may not hold for systems that are smaller relative to thermal unit sizes. See Appendix F.2.
- Modeling a single hydro year rather than 20 hydro years may not accurately capture hydro-related risk and did not offer material computational benefits for this system (e.g., improved convergence or precision). See Appendix F.3.
- Testing different combinations of weather conditions with weekend or weekday load patterns did not materially impact annual RA metrics. Using the statistical model, we estimated that increases in identified risk for challenging weather days that historically fell on weekends would be roughly offset by decreases in identified risk for challenging weather days that historically fell on weekdays. See Appendix F.4.

Key Takeaways from the Weather-Synchronized Simulation

- A Weather-Synchronized approach to RA analysis has a number of benefits over Monte Carlo Simulation to provide confidence that physically accurate weather conditions with all relevant correlations are captured. However, data availability remains a key limitation.
- Weather-Synchronized Simulation yielded similar RA metrics to Monte Carlo Simulation, provided that weather conditions from recent years (2015-2020) were tested.

- Weather is the most important driver of RA challenges and the treatment of weather trends is a key determinant of RA risk level.
- The availability of more high-resolution historical power system data as well as information about likely future weather conditions would greatly improve our understanding of the RA challenge.
- Statistical analysis can be used to estimate RA metrics associated with weather conditions for which high resolution data may not be available, including historical weather or simulated future weather under climate scenarios.

5

CONCLUSIONS AND NEXT STEPS

This initiative introduces the GridPath RA Toolkit—a publicly available resource adequacy analysis toolkit for the Western United States that leverages public data sets and the open-source model GridPath—and applies the Toolkit towards an independent and transparent evaluation of the state of resource adequacy in the Western United States in 2026. This report documents our insights regarding the relative strengths and weaknesses of RA modeling methodologies in the context of a modern, highly interconnected power system.

5.1 WESTERN US CASE STUDY FINDINGS

In the No Additions Scenario, which incorporates planned retirements, but excludes planned resources through 2026, we estimate that the West could be physically short by about 8.8 - 9.9 GW in 2026, if planning to a one-day-in-10-year LOLE standard. We find that event durations are relatively modest and the economically optimal duration for RA solutions is unlikely to exceed 3 or 4 hours. The identified shortage is much smaller than the amount of capacity additions in current utility plans in the West, including the procurement ordered in CPUC Decision D.21-06-035, which requires 11.5 GW of new net qualifying capacity through 2026. Incorporating capacity additions consistent with California's Preferred System Plan through 2026 eliminates all but seven RA events in 1,000 years of simulated dispatch.

Additional coal retirements do not pose an insurmountable RA challenge in the near term if California utilities procure the quantities of clean resources set forth in California's Preferred System Plan. In the Less Coal Scenario, the deployment of additional batteries and renewable resources in California appears to mitigate much of the needs associated with the retirement of an additional 11 GW of coal elsewhere in the West, even before considering capacity additions from utility plans outside of California. We estimate perfect capacity needs of about 3.8 GW for this case and find that short-duration solutions are likely adequate as the majority of RA shortages were 4 hours or less and occured in the evenings on hot summer days. Energy-limited resources such as batteries or demand flexibility are well suited to this type of shortage. The system does not appear to be energy-limited, with plentiful resources available to charge storage outside of the high-risk hours and shift the needed energy to avoid shortages in the afternoon and evening.

Due to the highly interconnected nature of the West, resource adequacy analysis that treats a particular RA planning footprint as an island can distort the observed RA challenges and may lead to suboptimal RA solutions, including potentially significant overbuild. Even without full West-wide planning coordination, RA programs may benefit from adopting market access policies that are informed by West-wide analysis in order to properly account for interregional operational interactions. In the No Additions Scenario, the

capacity needs associated with the CAISO footprint were reduced by about 3 GW by accounting for market access. And in the Less Coal Scenario, the capacity needs associated with the WRAP footprint were reduced by over 6 GW by accounting for market access. In both cases, accounting for market access also significantly reduced event durations, highlighting the potential for load and resource diversity across large areas to help mitigate energy shortages in the near-term.

5.2 METHODOLOGICAL FINDINGS

The Weather-Synchronized RA methodology developed for this report can provide a transparent analysis of resource adequacy challenges and metrics and reveal how weather and weather trends impact those metrics. Unlike Monte Carlo Simulation, which mixes and matches load, wind, and solar conditions in a manner that may not be physically realistic, the Weather-Synchronized approach tests historical weather conditions. It can therefore provide a much more transparent assessment of the drivers of RA risk. The Weather-Synchronized simulation revealed that actual extreme weather days tended to carry more risk than extreme days arising from Monte Carlo simulation. How well a Monte Carlo-based analysis captures the risk on the most challenging days will depend on subjective simulation design decisions, such as weather binning methodologies or other approaches to estimating the effects of correlated weather phenomena. Despite the benefits of Weather-Synchronized Simulation, this approach can be highly sensitive to the years for which synchronized high resolution load, wind, solar, and thermal derate data is available and these datasets are limited.

Statistical models can also be derived from the results of Weather-Synchronized Simulations in order to estimate loss of load risk during conditions for which high resolution synchronized data may not be available, including historical conditions or future weather conditions derived from climate modeling. The statistical analysis in this study suggests that weather conditions—temperature in particular—are the most important factors influencing loss of load risk in the West in the near term. It also suggests that weather trends in recent years may have implications for RA analyses that leverage historical data to estimate future weather conditions. Without information about potential future weather, via climate modeling and downscaling, the decision to rely on historical weather and a particular number of years is effectively a policy decision.

Regardless of the RA analysis approach, the availability of more historical data on load, wind, solar output, and thermal derates would improve the accuracy of the analysis. In particular, the lack of wind data in the NREL WIND Toolkit for recent historical years (after 2014) is a key limiting factor for capturing the weather-driven correlations of these variables with a higher level of confidence.

Generating more wind data is a high priority. As weather is a key determinant of RA risk and weather trends pose questions about the reliance on the historical record for assessing future RA risk, the availability of data for future weather can further improve our understanding of how to plan for reliable future energy systems.

5.3 FUTURE WORK

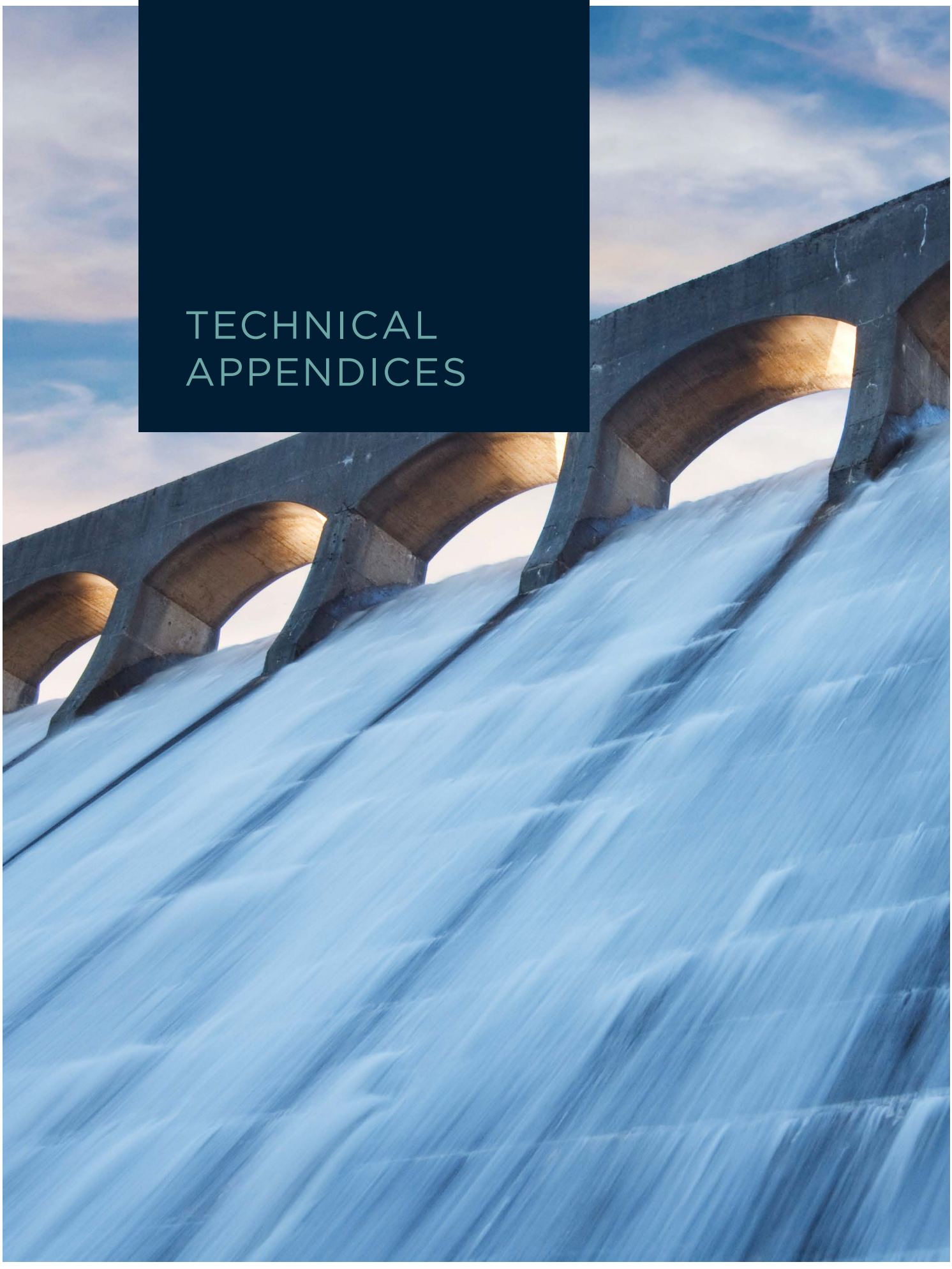
This study provides a useful foundation for future RA analysis. We've identified the following priorities for future RA analysis using the methodologies described in this report or other tools:

- **Climate sensitivities.** This study demonstrates that the historical weather record does not offer a stable distribution of conditions for use in RA studies. We do not have a strong understanding of how future weather patterns may impact loss of load risk. However, the statistical analysis approach described in this study may offer a relatively simple approach to probing this question. The statistical model for LOLP described in Section 4.3 could be applied to daily average weather conditions from downscaling climate models, rather than the historical record, to estimate LOLP in the study year. This would require simulated weather conditions at the same locations as were used to develop the statistical model and would require several years of potential weather conditions for the year 2026.
- **Electrification scenarios.** Because this study focuses on near-term RA, it relies on relatively simple estimates of future weather-sensitive load shapes based on historical trends. For longer-term studies, these estimates are less valid, especially in the context of electrification. To conduct longer-term studies, additional focus must be placed on the weather-sensitivity of electrified loads and the associated impacts on load shapes.
- **Weather-driven outages.** While temperature-dependent thermal derates are considered in this study, correlated outages driven by weather patterns are not. Future work to assess the importance of these common mode generator failures is dependent on the availability of data to describe the relationship between outage probability distributions and weather.
- **LSE or RA program modeling.** This study uses a physical representation of the Western United States to determine the physical capabilities of the system. This approach can be used to quantify physical RA challenges, but it cannot be used to allocate responsibility for RA to individual load serving entities or RA programs. Such an exercise would require ownership and contractual information for all of the resources contributing to meeting load in the LSE or RA program footprint, some of which may be confidential. However, with this ownership and contractual information, the dataset and model developed for this study can be used to conduct LSE, or RA program-specific analysis in a manner that is fully consistent with

the broader dynamics and constraints across the West during the periods that are most challenging for RA.

Finally, a core purpose of this initiative was to develop an advanced, publicly available and transparent toolkit for resource adequacy analysis that will contribute to the broader resource adequacy dialogue across the US. Our hope is that a broad set of users will leverage the GridPath RA Toolkit to further advance resource adequacy analysis for emerging power systems.

TECHNICAL APPENDICES



APPENDIX A.
LOAD AND RESOURCE AVAILABILITY MODELING DETAILS

A.1 LOAD SHAPES

Hourly historical load data was obtained from FERC Form 714 for respondents in WECC over the years 2006-2020 and was mapped to the WECC BAAs based on load shapes in the 2026 Common Case. All data was adjusted to Pacific Standard Time, which in some cases required manual adjustments. Manual adjustments were also applied to remove outliers via visual inspection and in specific cases where we could find documentation of a load shedding event in the historical record. In these circumstances (listed in Table A.1), the shed load was added back to the historical load to more closely reflect demand during the period.

TABLE A.1.

Historical load adjustments due to load shed events.

DATE	BAA	MAX OUTAGE (MW)	CAUSE
9/20/2015	CISO	150	Plant outage + Local Import Limit[46]
9/12/2016	PNM	150	Transmission outage[47]
2/2/2011	SRP	300	Many - cold weather event[48]
2/2/2011	EPE	523	Many - cold weather event
2/3/2011	EPE	250	Many - cold weather event
2/4/2011	EPE	250	Many - cold weather event
8/14/2020	CISO	1,072	Many - heat event[49]
8/15/2020	CISO	698	Many - heat event

Some BAA load shapes were approximated as linear combinations of the FERC respondent load shapes. This approach was used to better align load shapes with the 2026 Common Case load zone topology (especially in cases where multiple LSEs are in the same Common Case load zone) or to estimate loads in BAAs where there were large and/or abrupt changes in load in the historical record that could not be readily adjusted to reflect conditions in the study year. This could have been caused by a change in the footprint for a given respondent or the loss or gain of a very large load within a relatively small

46 https://www.nerc.com/pa/rrm/ea/energyemergency/EEA3_Report_20150920_SDGE.pdf
47 https://www.nerc.com/pa/rrm/ea/energyemergency/EEA3_Report__PNM_20160912_Disturbance.pdf
48 https://www.nerc.com/pa/rrm/ea/February%202011%20Southwest%20Cold%20Weather%
49 http://www.caiso.com/Documents/Final-Root-Cause-Analysis-Mid-August-2020-Extreme-Heat-Wave.pdf

system sometime during the historical record.[50] For these BAAs, we estimated the hourly load shape as a linear combination of the loads in other nearby BAAs. We selected the linear coefficients that resulted in load shapes that yielded the closest alignment between the estimated historical load in 2009 and the 2009 load shapes available in the 2026 Common Case. These load approximations are summarized in Table A.2.

TABLE A.2.

Load shape calibration to the WECC 2026 Common Case 2009 load shapes.

BAA	RESPONDENT LOADS IN LINEAR COMBINATION	REASON
BANC	SMUD, MID, constant	Combining multiple respondents into single load zone
BPAT	BPA, constant	Adjusting for a persistent difference between 2009 BPA load in FERC Form 714 and 2009 BPAT load in 2026 Common Case.
LDWP	LDWP, constant	Adjusting for a persistent difference between 2009 LDWP load in FERC Form 714 and 2009 LDWP load in 2026 Common Case.
NEVP	NEVP	Adjusting for a persistent difference between 2009 NEVP load in FERC Form 714 and 2009 NEVP load in 2026 Common Case, potentially due to VEA.
PNM	PNM, TSGT, PSC, constant	Combining multiple respondents into single load zone
TEPC	APS, PNM, SRP, TEP, constant	Adjusting for a persistent difference between 2009 TEPC load in FERC Form 714 and 2009 TEP load in 2026 Common Case.
VEA	NEVP	VEA data not available from FERC Form 714
WACM	PSC, constant	Abrupt change in load trend in recent years
WALC	APS, constant	Multiple abrupt changes in load in the historical record
WAUW	NWE, constant	Abrupt change in load early in the historical record

Figure A.1 shows two examples of the linear load estimations summarized in Table A.2. In both cases, the 2009 historical load from FERC Form 714 had a similar shape, but was out of alignment with the 2009 load shapes in the Common Case, suggesting a difference in the boundaries between the Common Case load zones and the reporting areas of the FERC Form 714 respondents. In both cases, using a linear combination resulted in load shapes that were based on the FERC Form 714 data (and could therefore be developed for the other historical years), but also aligned with Common Case load data that was available.

50 Manual adjustments were made to the CHPD load shape to remove demand associated with the Wenatchee Works Aluminum Smelter in December 2015.

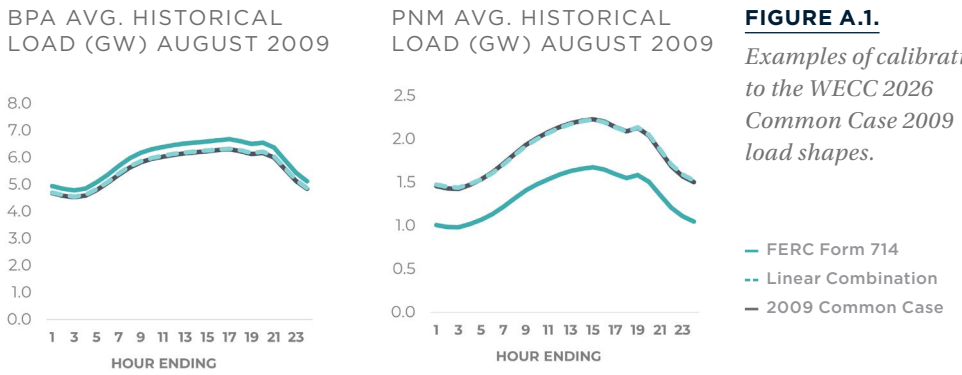


FIGURE A.1.

Examples of calibration to the WECC 2026 Common Case 2009 load shapes.

Some load zones in the Common Case represent only part of the footprint for the associated FERC respondent, specifically those in CAISO, Idaho Power, and PacifiCorp's BAAs. For these load zones, we used the 2009 load shapes in the 2026 Common Case to derive month-hour specific allocation factors for each load zone and applied these allocation factors to the FERC Form 714 data.

To transform the historical hourly load shapes to the study year, we first binned the historical load data by month, hour ending (HE), and weekdays versus weekends. Within each bin, j, we express the load on each day, i, as a function of various daily weather indicators (X_i), the historical year that the day falls within (y_i), a normalized measure of the total annual real GDP of the Western states in that year (z_{yi}), a constant (k_j), and a residual term (ε):

$$l_{ij} = A_j X_i^2 + B_j X_i + c_j y_i + d_j z_{yi} + k_j + \varepsilon_{ij}$$

The daily weather indicators (X_i) were derived for each bin by performing Y-aware principal component analysis (PCA) on the weather data described in Appendix E for each of the days in the bin and selecting enough components to explain 85% of the variance. We used linear regression to solve for the coefficients in each bin, A_j, B_j, c_j, d_j, and k_j, and we calculate the residual for each day associated with the linear model:

$$\varepsilon_{ij} = l_{ij} - A_j X_i^2 - B_j X_i - c_j y_i - d_j z_{yi} - k_j$$

To estimate the load in the study year, l_{ij}^{2026}, we apply this linear model to the same weather conditions, but we update the year to 2026 and the GDP to a forecasted value for 2026:

$$l_{ij}^{2026} = A_j X_i^2 + B_j X_i + 2026 \times c_j + d_j z_{2026} + k_j + \varepsilon_{ij}$$

$$l_{ij}^{2026} = l_{ij} + (2026 - y_i) \times c_j + (z_{2026} - z_{yi}) \times d_j$$

The historical real GDP across the Western United States in 2006-2020 was obtained from the Bureau of Economic Analysis[51] and the real GDP across the Western United States in 2026 was forecasted by applying the national percent GDP growth from 2020 to 2026 from the 2021 EIA Annual Energy Outlook.[52] The resulting values of z are shown in Figure A.2.

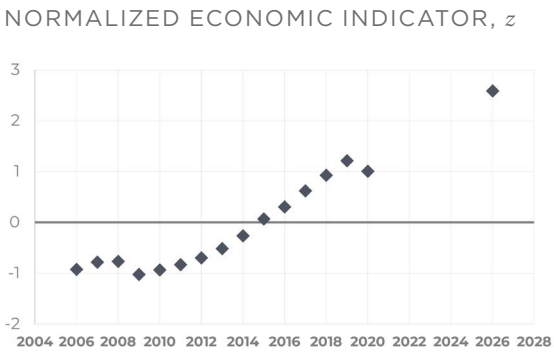


FIGURE A.2.

Normalized economic indicator used in load shape adjustment.

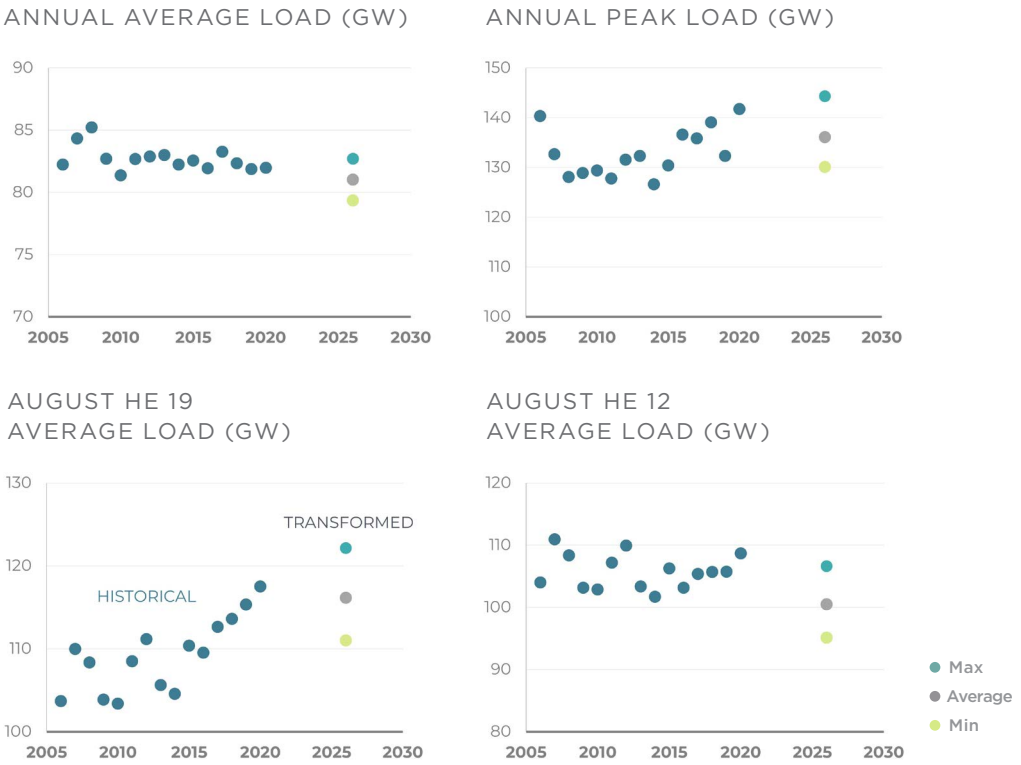
Figure A.3 summarizes the West-wide loads resulting from the load transformation using some high level metrics, and compares them to the historical West-wide loads. In each plot, the 2026 dots represent the min, maximum, and average annual metrics associated with the load across the 15 weather years transformed to 2026 economic conditions. The blue dots represent the actual historical values. The transformed data generally represent a continuation of the trends observed in the historical record, but also reflect the wide variation in load across the weather years. In some cases, for example the annual peak load and the average load during August HE 19, the transformed data appears to be less extreme than the trend might suggest. In these cases, the trend in the historical data may be significantly influenced by a trend in weather conditions. Because the load transformation assumes the same weather conditions as were experienced historically, it does reflect changes in load that may be associated with weather trends. More discussion of weather trends and their implications for resource adequacy analysis can be found in Section 4.3. In other metrics, specifically the average load during August HE 12, the transformed load seems to accelerate the historical trend. In this case, we are seeing the effects of continued rooftop PV buildout, which offset potential increases in load due to the weather and other factors.

51 Annual GDP by State, SAGDP tables, available at: https://apps.bea.gov/regional/downloadzip.cfm

52 U.S. Energy Information Administration, Annual Energy Outlook 2021. Macroeconomic Indicators table, available at: https://www.eia.gov/outlooks/aeo/tables_side.php

FIGURE A.3.

Comparison of 2026 transformed load statistics to historical data.



A.2 THERMAL GENERATOR AVAILABILITY

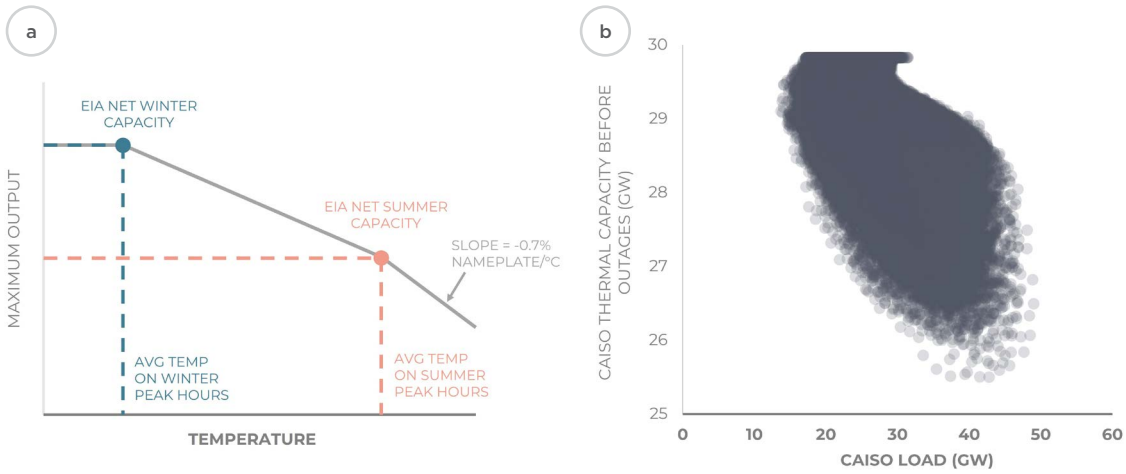
The maximum hourly output from all thermal units (including coal, gas, nuclear, biomass, and biogas) is estimated as a function of temperature using the following methodology.

For each unit, the maximum output in each hour was estimated by a unit-specific piecewise linear function of temperature, as shown in Figure A.4(a). To derive this piecewise linear function, we examined the hourly load shape of the BA to which the unit was assigned to identify the "peak hours" in the winter and summer, the four hours of the day in each season with the highest average load, over which net capacity testing may have occurred. We then assigned the net winter and summer capacities from EIA Form 860M to the average temperatures over these hours in 2019 at the project site and assumed linear changes in maximum output at temperatures between these two points. At temperatures above the average summer peak hours temperature at the site, we derated capacity by 0.7% of nameplate per degree Celsius, a simple

approximation based on information from GE[53] and Wartsila[54]. At temperatures below the average winter peak hours temperature, we held the maximum output constant.

FIGURE A.4.

Thermal derate curve approach and resulting availability before outages.



This approach results in correlations between temperature-driven loads and thermal resource availability, as illustrated in Figure A.4(b), which shows the total hourly availability of all thermal units in CAISO in a test simulation, before accounting for forced outages, as a function of the CAISO load.

Forced outages were applied to each unit after calculating the hourly maximum output. Forced outages were applied randomly within Monte Carlo Simulation using exponential failure and repair models with parameters tuned to achieve a unit-specific forced outage rate (FOR) and mean time to repair (MTTR). Random unit forced outages were assumed to be independent and uncorrelated. For units that could be mapped to the WECC 2026 Common Case, forced outage parameters from the Common Case were used. For all remaining units, weighted-average forced outage parameters by technology type from the Common Case were used, which are listed in Table A.3.

53 Brooks, Frank J., GE Power Systems, "GE Gas Turbine Performance Characteristics," https://www.ge.com/content/dam/gepower-new/global/en_US/downloads/gas-new-site/resources/reference/ger-3567h-ge-gas-turbine-performance-characteristics.pdf

54 Wartsila, "Combustion Engine vs Gas Turbine: Derating due to Ambient Temperature," https://www.wartsila.com/energy/learn-more/technical-comparisons/combustion-engine-vs-gas-turbine-derating-due-to-ambient-temperature

TABLE A.3.

Weighted average outage parameters applied to unmapped units.

TECHNOLOGY	WEIGHTED AVERAGE FORCED OUTAGE RATE	WEIGHTED AVERAGE MTTR (HRS)
Biomass	3.4%	38
Coal	4.7%	38
GasCC	3.3%	34
GasCT	3.0%	51
GasIC	3.3%	37
GasST	3.4%	40
Geothermal	3.1%	24
Nuclear	3.1%	298
Other	4.4%	41

In the dispatch simulation, thermal generators are assumed to be committed in all time steps in which they are available (i.e., not experiencing a forced outage) under the assumption that system operators have some advance notice when a capacity shortage may be imminent. As with other resource types, scheduled maintenance is not reflected in the simulation as it is assumed that maintenance will be scheduled for periods with a very low risk of encountering a shortfall. We also do not apply energy limitations to thermal units. The scenarios investigated in this study do not exhibit significant energy shortages, but we flag this as an important area of further investigation in future studies, especially where fuel constraints or emissions regulations may materially affect the availability of thermal resources.

A.3 HYDROPOWER

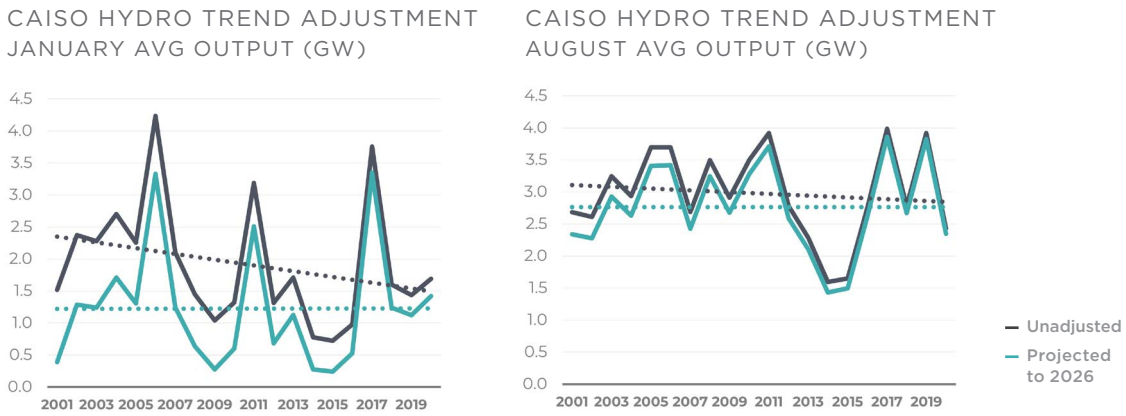
Hydro resources across the West vary widely in their operational characteristics, availability, storage, and non-power-related constraints. Because of the complexity and diversity of hydro resources as well as interactions between hydro resources that exist on the same river system, hydropower availability was estimated in aggregate within each BAA based on historical operations. The aggregated dispatch of hydro resources within each BAA was simulated within the dispatch optimization based on minimum output, maximum output, and weekly energy budget constraints.

Aggregated weekly energy budgets were derived from historical monthly hydro generation data 2001 to 2020 (including hydro with storage, run-of-river hydro, and pumped storage hydro) in each BAA from EIA Form 923/906.

The historical record showed indications of trends in hydro generation over time. To account for these trends, we detrended the hydro data for each BAA and month and projected to the study year of 2026 using a simple linear regression. We also bounded the adjusted energy by the historical maximum observed energy across all months to avoid non-physical hydro conditions. The resulting projected energy budgets (expressed as average output) are shown for January and August in CAISO and BPA in Figure A.5 and Figure A.6, respectively.

FIGURE A.5.

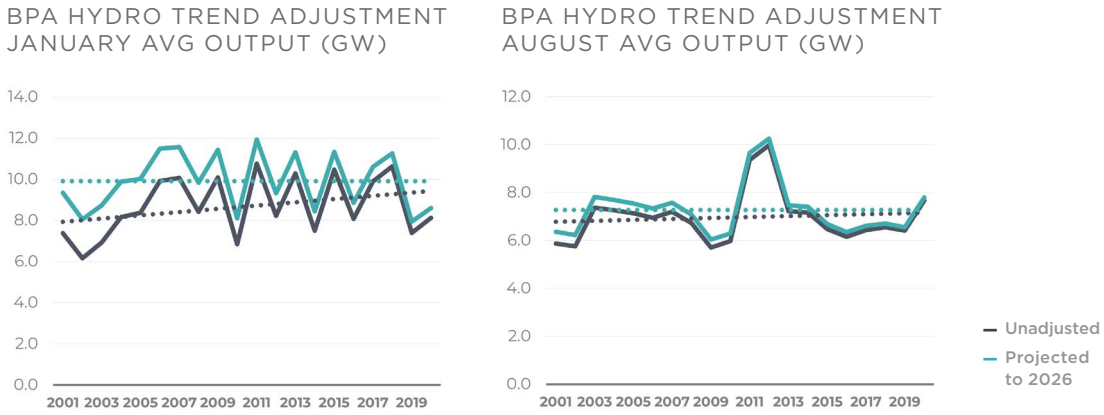
CAISO hydro trend adjustments.



In the figures, the observed and adjusted linear trends are represented by dotted lines. Figure A.5 shows that the average hydro generation in January months in CAISO tended to decline over the period from 2001 to 2020. The hydro adjustment therefore reduces the January hydro budgets corresponding to the earlier years quite significantly so that the slope of the hydro budget is flat over time (dotted green line). The unadjusted and adjusted trend lines meet at the study year (2026) so that the mean hydro budget across the entire adjusted dataset is equal to the projected mean hydro budget in 2026 based on the linear trend in the historical dataset. The same type of adjustment can be seen for August, although historical generation in August had a much less dramatic trend over the years.

FIGURE A.6.

BPA hydro trend adjustments.



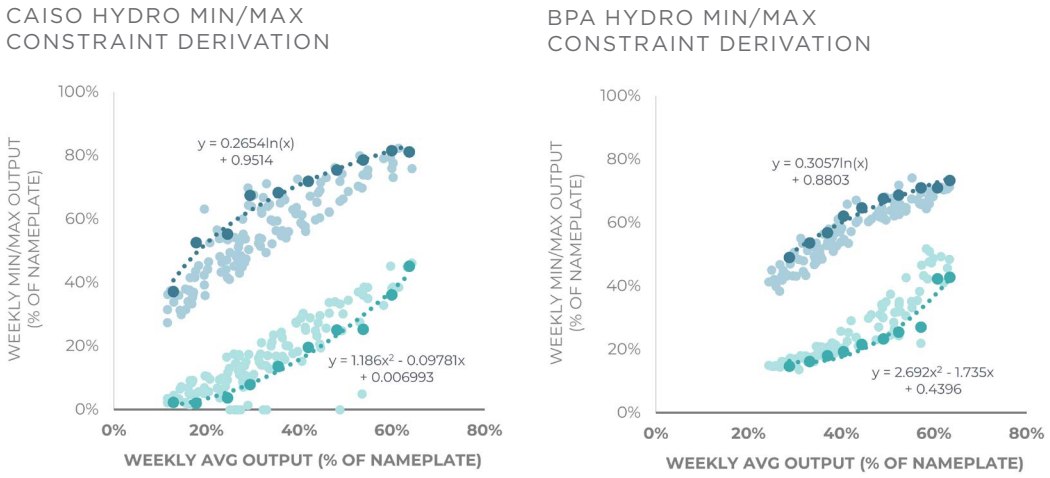
BPA hydro trends differ significantly from CAISO hydro trends, with generation in both January and August tending to increase over 2001-2020. In this case, the hydro detrending process increases hydro budgets for the earlier years, more dramatically in January than in August.

To derive minimum and maximum output constraints for the aggregated hydro in each BAA, we used historical hourly hydro data to investigate the relationship between the average generation over a given week and the minimum and maximum output over the course of the week. Because hourly historical hydro generation was not available for every plant or BAA, we adopted a number of approximations in this process, in particular outside of CAISO and BPA.

For CAISO, we used historical hourly hydro data from the CAISO Daily Renewables Watch between April 2010 and January 2021. For each week, we calculated the minimum, maximum, and average hydro output. These are plotted in Figure A.7. We then binned each week based on its average hydro output and for each bin, we calculated the 15th percentile of the minimums and the 95% of the maximums (shown in dark green and dark blue, respectively). We then fit curves to these points. The minimum points were fit to a quadratic function and the maximum points were fit to a logarithmic function. We also identified a range over which application of these functions seemed reasonable. For CAISO, this range was 12.9-63.8% of nameplate for both the minimum and maximum functions. To ensure feasibility below this range, maximum output constraints were assumed to be the maximum of the energy budget and the maximum output associated with the 12.9% threshold and minimum output constraints were assumed to be the minimum of the energy budget and the minimum output associated with the 12.9% threshold. The same approach was applied for energy budgets above the 63.8% threshold.

FIGURE A.7.

Hydro minimum and maximum output constraint derivation.



For BPA, we used historical hourly hydro generation from the BPA Total Load and Wind Generation Report from 2007-2020. We used the same methodology described above for CAISO, except the 5th percentile of the binned data was used to estimate the minimum values and the 90th percentile of the binned data was used to estimate the maximum values. For BPA, the resulting functions were applied over the range 32.2-63.4% of nameplate capacity for the minimum output and 28.8-63.4% for the maximum output. Outside of this range, we used the same approach described above to ensure feasibility.

To benchmark this approach, we compared the estimated minimum and maximum constraints to historical generation when the West was constrained. Figure A.8 shows this comparison for hydro resources in CAISO over the month of August 2020. The green shaded area highlights August 14th and 15th, when CAISO experienced rolling blackouts due to supply shortages. This exercise suggests that the hydro constraints may slightly underestimate the capabilities of the CAISO hydro fleet when the system is constrained.

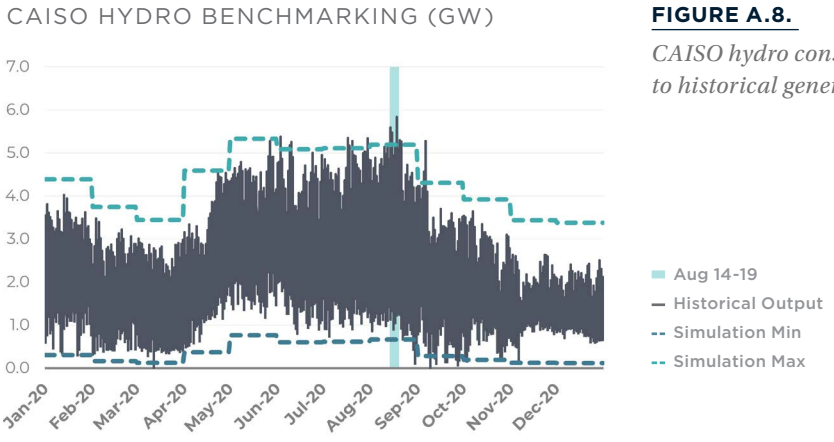


FIGURE A.8.

CAISO hydro constraint benchmarking to historical generation.

The same period is shown for BPA hydro in August 2020 in Figure A.9. While the maximum output constraint aligns well with the maximum output during the week of the heat wave, the historical output ramped up to that level over the course of about four days, rather than being available during the most critical periods on August 14th and 15th. This could have been due to a combination of factors, including forecast errors, scheduling constraints (August 15th and 16th fell on the weekend, which can complicate scheduling), and transmission congestion on COI. The difference between the capability of the hydro fleet and its actual dispatch during this critical period is an important reminder that operational modernization must accompany resource development to achieve reliability of supply in an efficient manner.

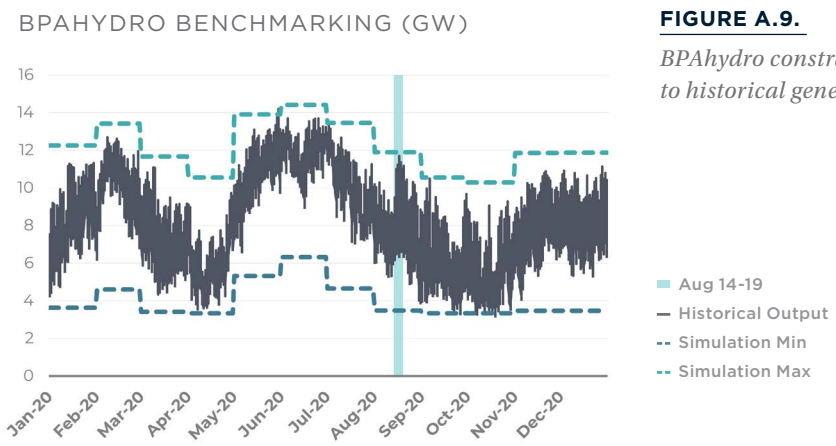


FIGURE A.9.

BPAhydro constraint benchmarking to historical generation.

Outside of CAISO and BPA, we relied on hourly data for 46 hydro projects in the year 2005, which was available through WECC and represented 44% of the remaining hydro in the No Additions Scenario, to derive the minimum and maximum functions for the BAs listed in Table A.4.

TABLE A.4.

Hydro minimum and maximum constraints.

BAA	MINIMUM FUNCTION	APPLICABLE RANGE	MAXIMUM FUNCTION	APPLICABLE RANGE	% OF HYDRO MW WITH HOURLY DATA
AVA	0.035	0-100%	0.50ln(x)+1.40	12-41%	74%
BANC	$1.59x^2$-1.11x+0.28	45-82%	0.59ln(x)+1.33	21-82%	68%
CHPD	0.02	0-100%	0.34ln(x)+1.10	30-57%	30%
GCPD	$3.13x^2$-2.72x+0.67	43-61%	0.83	0-100%	99%
PACW	$2.11x^2$-0.73x+0.12	17-36%	0.34ln(x)+1.09	13-58%	67%
TH_Mead (Hoover)	0.01	0-100%	0.27ln(x)+1.17	5-23%	100%
TIDC	$1.19x^2$-0.80x+0.13	39-113%	0.27ln(x)+1.16	4-113%	91%
WACM	$3.98x^2$-1.64x+0.26	21-40%	0.35ln(x)+0.91	18-40%	89%
WALC	$0.43x^2$-0.06x+0.04	5-67%	0.39ln(x)+1.26	5-50%	60%

We then grouped all of the hydro projects in the No Additions Scenario based on whether they were assigned to CAISO, assigned to BPA, or had hourly generation information in 2005 from WECC. For these projects, we aggregated the energy budgets and applied the associated minimum and maximum functions so that the dispatch optimization could solve for the hourly dispatch within the specified ranges. For all other projects, we made the conservative assumption of fixing hydro output based on the energy budget — effectively treating these projects as run-of-river. This approximation will tend to underestimate the capabilities of the hydro fleet. Table A.5 shows the nameplate capacity that fell into each of these categories by BAA. In total, 73% of the hydro fleet (in terms of nameplate capacity) was represented by the minimum and maximum functions, while 27% of the fleet was represented by flat dispatch based on the energy budget. Note that some of the hydro incorporated into the minimum and maximum functions may be run-of-river. In that case, the limited flexibility of those plants is reflected in the historical data and will influence the minimum and maximum constraints accordingly (i.e., will tend to tighten them or bring them closer to the average output). Similarly, forced outages are implicitly captured to the extent that they have reduced the historical output that is used to derive the maximum output function.

TABLE A.5.

Breakdown of hydro modeling approach by aggregated hydro resource.

BAA	MW WITH MIN/MAX FUNCTION	MW MODELED AS FIXED	TOTAL MW
AVA	834.7	297	1131.7
AZPS	0	40	40
BANC	1743.8	817.2	2561
BPAT	20464.4	202.2	20666.6
CHPD	588	1358.8	1946.8
CIPV	6600.2	0	6600.2
CISC	1696.4	0	1696.4
CISD	0	56.6	56.6
DOPD	0	696.6	696.6
GCPD	2175.6	16	2191.6
IID	0	82.6	82.6
IPFE	0	92.4	92.4
IPMV	0	479.9	479.9
IPTV	0	1541.6	1541.6
LDWP	0	1928.6	1928.6
NWMT	0	735.7	735.7
PACW	561.1	276.7	837.8
PAID	0	342.2	342.2
PAUT	1.5	104.5	106
PAWY	0	2.9	2.9
PGE	0	694.2	694.2
PNM	0	23.6	23.6
PSCO	0	382.5	382.5
PSEI	0	406.6	406.6
SCL	0	1958.3	1958.3
SPPC	0	17.8	17.8
SRP	0	243.4	243.4
TH_Mead	2078.8	0	2078.8
TIDC	125.3	12.2	137.5
TPWR	0	835.4	835.4
WACM	2521.5	322	2843.5
WALC	254.8	169.5	424.3
WAUW	0	173.1	173.1
Total	**39,646 (73%)**	**14,310 (27%)**	**53,956**

A.4 WIND POWER

The hourly availability of wind resources was estimated on a site-specific basis using 80-m wind speed data from the NREL WIND Toolkit over 2007-2014 and empirically-derived logistic function power curves that were tuned to best match project-specific historical monthly generation data from EIA Form 923/906.

Three different methodologies were applied based on the amount of available historical generation data. For projects that came online prior to 2014 and that had overlapping periods with historical monthly generation and hourly simulated wind speed data, we calibrated the power curve to minimize the root mean square error of the associated monthly generation (Method 1). For projects that came online after 2014 and had adequate historical data to estimate an average historical capacity factor, we calibrated the power curve to achieve that average capacity factor across the 2007-2014 simulation period (Method 2). For projects that did not have adequate historical data to estimate an average capacity factor (very recent projects or projects under construction as of February 2021), we estimated the average capacity factor based on the zone and commercial online date (COD) and calibrated the power curve to achieve this capacity factor (Method 3). Some projects were not represented well by the empirical power curve or did not have enough data to fit a curve. For these projects, we assumed that the hourly shape was the same as the aggregated shape for the BAA to which they were assigned.

Table A.6 summarizes the amount of wind capacity over which each method was applied and compares the average historical capacity factors over the various project-specific calibration periods to the simulated capacity factor over 2007-2014 for each BAA for the projects that utilized any of the three methods.

TABLE A.6.

Breakdown of wind simulation methodologies applied by aggregated wind resource.

BAA	METHOD 1 (MW)	METHOD 2 (MW)	METHOD 3 (MW)	NONE (MW)	SIMULATED CAP FACTOR (2007-2014)	HISTORICAL CAP FACTOR (VARIOUS PERIODS)
AVA	105.3	0	144	0	38.5%	37.6%
AZPS	227.3	0	0	0	23.4%	22.3%
BPAT	4573.6	466.8	361.3	6	28.2%	28.0%
CIPB	116.2	131.9	57.5	0	36.7%	36.1%
CIPV	1135.9	10.9	164.6	33.9	30.3%	29.5%
CISC	3181.6	524.5	186.9	241.2	26.2%	26.1%
CISD	315.4	144	0	0	25.1%	24.1%
EPE	0	0	249.8	0	38.9%	38.9%
IPFE	79.2	0	0	0	35.6%	35.4%
IPMV	326.9	0	0	115	30.0%	29.5%
IPTV	45	50	0	100.7	31.1%	30.9%
LDWP	440.5	0	0	0	23.0%	22.5%
NWMT	630.7	139.4	160	1.8	36.5%	35.8%
PACW	508	10	40	0	28.1%	27.4%
PAID	411.3	0	0	0	33.5%	32.4%
PAUT	18.9	65.3	0	0	28.2%	28.2%
PAWY	1336.1	80	2787.1	58.4	33.6%	33.5%
PNM	496.4	842.8	306.2	0	38.3%	38.2%
PSCO	2141.1	1133.4	970.1	6	37.3%	36.6%
PSEI	705.3	0	136	4.3	29.4%	28.8%
SPPC	150	0	0	0	23.5%	21.0%
TEPC	60.4	30	0	0	26.9%	26.7%
WACM	218.5	325	382.5	0	40.5%	40.2%
WALC	0	0	350	0	22.0%	22.0%
Total	**17,224 (61%)**	**3,954 (14%)**	**6,296 (22%)**	**567 (2%)**	**31.7%**	**31.3%**

To further benchmark this approach, we used it to estimate the total hourly wind generation in CAISO and BPA in 2013 and compared it to reported actuals in the Daily Renewables Watch and the BPA Total Load and Wind Generation Report, respectively. We found that the approach tended to underestimate wind power output in both systems, and had an hourly rms error of 10.2% of nameplate for CAISO and 11.2% of nameplate for BPA. As can be seen in Figure A.10 and Figure A.11, which show the comparisons for August 2013, the approach generally captures the timing and magnitude of wind events and diurnal trends. Differences between the historical and the simulated data could arise due to a number of factors, including: differences between the resources included in the Daily Renewables Watch and BPA Report and the resources listed in EIA Form 860M that are assigned to CAISO and BPA, differences between the beginning of metered operations and the commercial operation dates listed in EIA Form 860M, numerical weather simulation uncertainties in NREL's WIND Toolkit, and uncertainties in the power curve estimation approach. Given all of these potential sources of error and the reliance on publicly available information, we determined that the observed error was acceptable for the purposes of this study.

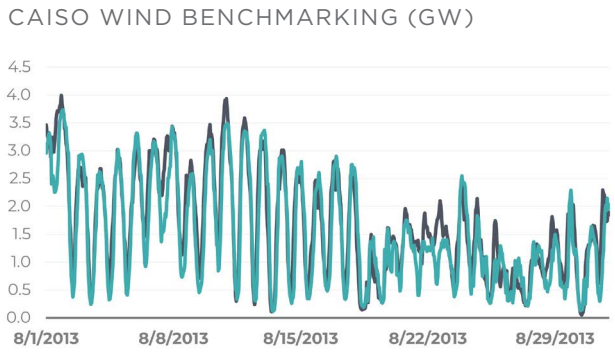


FIGURE A.10.

CAISO wind benchmarking to historical generation.

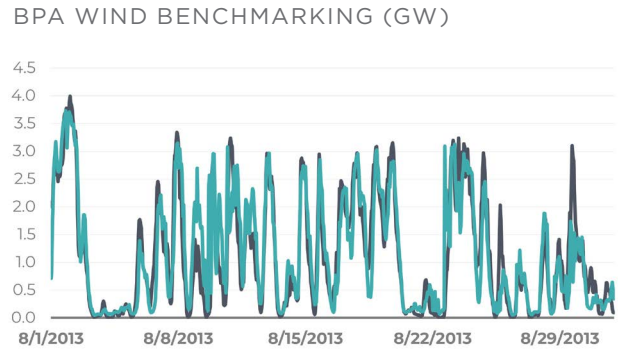


FIGURE A.11.

BPA wind benchmarking to historical generation.

A.5 SOLAR POWER

The hourly availability of solar resources was estimated on a site-specific basis using weather data over 1998-2019 from the National Solar Radiation Database (NSRDB). Output from each solar resource was simulated using NREL's System Advisor Model (SAM). For each project, the inverter loading ratio was estimated based on the project commercial operation date (COD) and market trends reported in LBNL's Utility-Scale Solar Data Update: 2020 Edition (see Figure A.12). Where fixed versus tracking information was not available, we simulated output under both configurations and took a weighted average of the resulting shapes with weights based on the project COD and market trends reported in LBNL's Utility-Scale Solar Data Update: 2020 Edition (see Figure A.12).

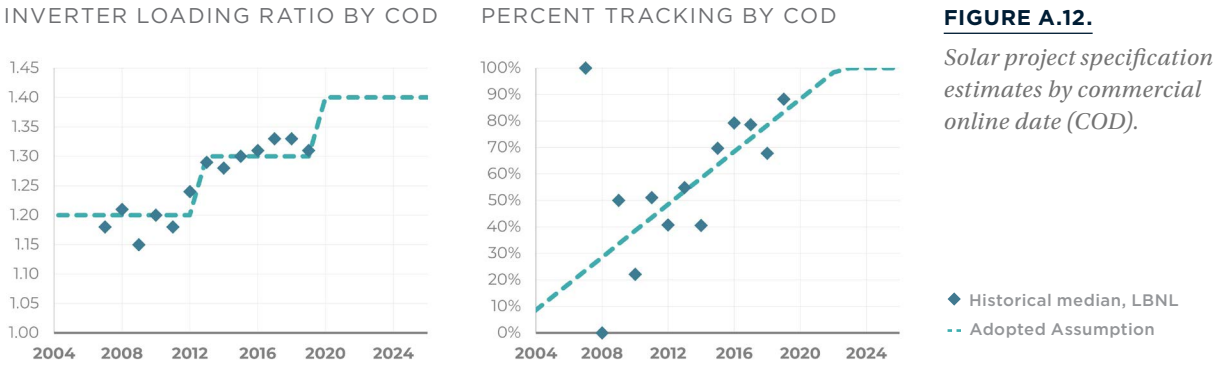


FIGURE A.12.
Solar project specification estimates by commercial online date (COD).

For fixed tilt systems, we estimated the tilt angle based on the historical weighted average tilt angle as a function of latitude as reported by EIA.[55]

55 Available at: https://www.eia.gov/todayinenergy/detail.php?id=37372#:~:text=Most%20utility%2Dscale%20fixed%2Dtilt,U.S.%20Energy%20Information%20Administration%20(EIA)

TABLE A.7.

Fixed tilt solar tilt angle estimates by latitude.

LATITUDE	ESTIMATED TILT ANGLE
20-25	16
25-30	18
30-35	20
35-40	22
40-45	23
45-50	27

For hybrid resources with solar PV components, solar power availability was simulated in SAM with an inverter loading ratio of 1 to avoid suboptimal clipping. Solar thermal projects were modeled in SAM using the default parabolic trough and power tower configurations with thermal storage where specified in publicly available project information.

To benchmark this approach, we used it to estimate the hourly generation from solar resources in CAISO between 2017 and 2019, based on resource CODs in EIA Form 860M, and we compared it to reported solar generation from the CAISO's Daily Renewable Watch. We found that the performance varied by month. Average solar capacity factors were slightly underestimated in June–August and slightly overestimated in other months. The root-mean-square errors of the hourly simulated capacity factors for each month were smallest during summer months and largest during the spring. Figure A.13 shows the simulated and historical data across the month of September 2017. This month began with a strong heat wave, leading to peak demand conditions on September 1. On this day, the simulated shapes slightly underestimate solar availability relative to historical actuals. Throughout the rest of the month, the simulated shapes generally reflect the day-to-day-variations observed in historical solar output.

TABLE A.8.

Solar estimation error in 2017 - 2019.

MONTH	MONTHLY AVERAGE CAPACITY FACTOR ERROR	ROOT-MEAN-SQUARE ERROR OF HOURLY CAPACITY FACTOR
1	+0.5%	5.1%
2	+0.9%	6.7%
3	+1.2%	7.2%
4	+1.2%	8.0%
5	+0.6%	6.7%
6	-0.9%	4.2%
7	-1.2%	3.7%
8	-1.0%	3.8%
9	+2.2%	3.5%
10	+1.4%	4.5%
11	+1.3%	6.0%
12	+1.2%	5.8%
2017-2019	+0.6%	5.6%

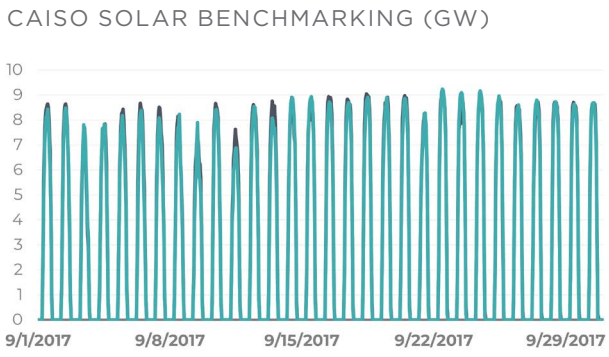


FIGURE A.13.

Solar benchmarking to historical generation.

Differences between the historical and the simulated data could arise due to a number of factors, including: differences between the resources included in the Daily Renewables Watch and the resources listed in EIA Form 860M that are assigned to CAISO, differences between the beginning of metered operations and the commercial operation dates listed in EIA Form 860M, numerical weather simulation uncertainties in the National Solar Radiation Database, solar output simulation uncertainties in SAM, and differences between actual project specifications and the approximations that we employed in SAM. Given all of these potential sources of error and the reliance on publicly available

information, we determined that the observed error was acceptable for the purposes of this study.

A.6 BATTERY STORAGE

Battery dispatch is modeled endogenously by GridPath. Batteries are associated with three variables in each hour of the simulated week: the charging level, the discharging level, and the energy available in storage. The first two are constrained to be less than or equal to the battery's power capacity, and the third is constrained based on the battery's duration. The model tracks the state of charge in each hour based on the charging and discharging decisions in the previous hour, with adjustments for charging and discharging efficiencies. Battery system availability is modeled with a fixed 3.5% derate in all hours to approximate forced outages. We did not simulate random battery unit forced outages because we did not have enough information about the nature of battery forced outages (effective unit sizes and mean time to repair) and we expected unit sizes to be very small relative to the size of the system.

A.7 HYBRID RENEWABLES

The dispatch of hybrid projects, e.g., solar and battery hybrids, is modeled endogenously in GridPath. We model the batteries in these systems as being able to charge from the solar or wind component only and not from the grid. We track power availability from the renewable resource via an hourly capacity factor parameter and the model determines the amount of power that goes directly to the grid and the amount that is stored. The former is limited by the power capacity of the renewable component and the latter by the charging power capacity of the battery. The model tracks the state of charge of the battery. Total grid power output from the project (from the renewable component and from storage discharging) is constrained by an interconnection limit.

APPENDIX B.
MONTE CARLO SIMULATION DETAILS

The Monte Carlo approach employed in this study synthesizes several years of plausible hourly load, wind availability, solar availability, and temperature-driven thermal derate data over which the system operations can be simulated. Synthetic days are built by combining load, wind, solar, and temperature derate shapes from different but similar days in the historical record. The first step in this process is to bin historical days by month and their weather conditions. Each bin contains multiple days of data with respect to each variable and those days share similar characteristics. To synthesize a plausible day of conditions, the model randomly selects load, wind, solar, and thermal derate days from within the same bin and combines the conditions across the variables. To preserve geographical correlations, the same day is selected for all sites or locations associated with each variable. For example, the model might combine the load from August 13th, 2012 with the solar availability from August 27th 2011, the wind availability from August 31st, 2007, and the thermal derates from August 16th, 2015 because these days all fall within the same August weather bin. In this example, the loads across the entire system reflect conditions on August 13th, 2012 and the availability across all solar projects reflects the conditions on August 27th, 2011. Load days are also binned by weekend and weekday.

B.1 WEATHER BINS

For this study, we designed weather bins to capture the weather phenomena across the West that drive net load in the study year. This is a complicated exercise because weather patterns across such a large geographic area are complex and diverse. To characterize the weather in the West, we collected daily average weather data (LIST) at 16 locations across the West, each chosen for its proximity to a load center and/or a strong wind or solar resource. More information about the historical weather data can be found in Appendix E. To reduce this complex system down to relatively simple parameters that could be used to categorize each day, we used Y-aware principal component analysis (PCA) of the historical weather data and partitioned days based on their percentiles with respect to the first two principal components. This process is described in more detail below.

1. First, we calculated the weather-based hourly net load (load minus wind minus solar minus temperature-driven thermal availability) for all historical weekdays for which hourly data was available across all the variables (2007-2014).
2. For each historical day with hourly net load data, we calculated a daily net load score intended to reflect the potential to experience energy and capacity shortages on that day: the average of the daily average net load and the daily maximum net load.

3. We then pulled daily weather variables representing conditions across the West (see Appendix E for more information) for those same days.
4. We partitioned the days by month and within each month, we performed Y-aware PCA to reduce the N-dimensional weather data down to two dimensions:
   a. We normalized the daily weather variables so they had mean 0 and standard deviation 1 to prepare for PCA.
   b. We then conducted Y-aware PCA on the normalized daily weather variables with the daily net load scores. This process involved scaling each normalized weather variable by the slope of the daily net load score with respect to that normalized variable and then performing PCA on the scaled data. This approach places more emphasis on weather variables that may impact unserved energy and it results in orthogonal "weather modes" that can be used to partition the days based on weather. We also adjusted the signs of the principal components so that they pointed in the direction of increasing net load score.
   c. We then transformed each day into the coordinate system defined by the first two principal components.

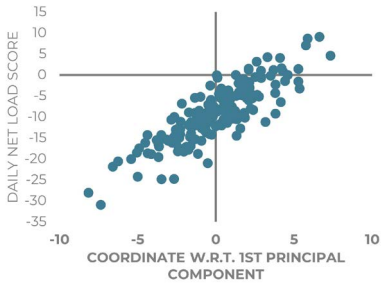
Figure B.1 shows the first principal component for August and the relationship between the coordinate with respect to this component and the daily net load scores. The same is shown for January in Figure B.2. The analysis suggests that the most challenging days in August from a net load perspective are likely to have very high temperatures in the Central Valley California, relatively high temperatures in the Pacific Northwest, and relatively still wind conditions at key locations in the Pacific Northwest and California.

AUGUST FIRST PRINCIPAL COMPONENT

	TEMPERATURE	WIND SPEED	PRESSURE	DEW POINT
Seattle, WA	0.204	-0.003	0.010	0.012
Portland, OR	0.228	0.000	-0.003	0.043
The Dalles, OR	0.120	-0.176	0.015	0.003
Lewiston, MT	0.025	-0.003	0.163	-0.041
Casper, WY	-0.003	-0.108	0.240	-0.003
Colorado Springs, CO	-0.005	-0.001	0.190	-0.003
Albuquerque, NM	-0.001	0.037	0.089	-0.001
Tucumcari, NM	-0.005	-0.045	0.138	-0.002
Phoenix, AZ	0.072	0.022	0.003	0.011
Blythe, CA	0.140	-0.036	0.002	0.084
Las Vegas, NV	0.102	-0.187	0.042	0.098
Edwards Airforce Base, CA	0.378	-0.261	0.000	0.048
Los Angeles, CA	0.126	-0.021	-0.087	0.117
Fresno, CA	0.405	-0.150	-0.034	0.046
Sacramento, CA	0.364	-0.148	-0.043	-0.003
San Francisco, CA	0.044	0.001	-0.055	0.003

FIGURE B.1.

First weather principal component for August and relationship to daily net load score.



The most challenging January days appear to be colder than usual across the West, especially in the Pacific Northwest, with relatively high pressure conditions everywhere but California and the Desert Southwest. However,

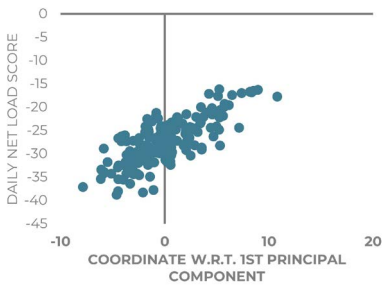
unlike August, none of the January days have a daily net load score exceeding zero, suggesting that January may not see material West-wide resource adequacy challenges in this study.

JANUARY FIRST PRINCIPAL COMPONENT

	TEMPERATURE	WIND SPEED	PRESSURE	DEW POINT
Seattle, WA	-0.241	-0.083	0.057	-0.303
Portland, OR	-0.319	0.005	0.034	-0.354
The Dalles, OR	-0.306	-0.032	0.097	-0.344
Lewiston, MT	-0.100	-0.059	0.148	-0.130
Casper, WY	-0.219	-0.014	0.125	-0.268
Colorado Springs, CO	-0.186	-0.026	0.069	-0.103
Albuquerque, NM	-0.061	0.000	0.003	-0.033
Tucumcari, NM	-0.121	-0.020	0.052	-0.024
Phoenix, AZ	-0.057	0.009	0.001	-0.015
Blythe, CA	-0.085	0.013	0.001	-0.022
Las Vegas, NV	-0.106	0.002	0.001	-0.063
Edwards Airforce Base, CA	-0.105	0.003	0.000	-0.048
Los Angeles, CA	-0.052	0.016	0.002	-0.048
Fresno, CA	-0.068	0.000	0.005	-0.076
Sacramento, CA	-0.125	0.001	0.006	-0.116
San Francisco, CA	-0.144	-0.001	0.006	-0.147

FIGURE B.2.

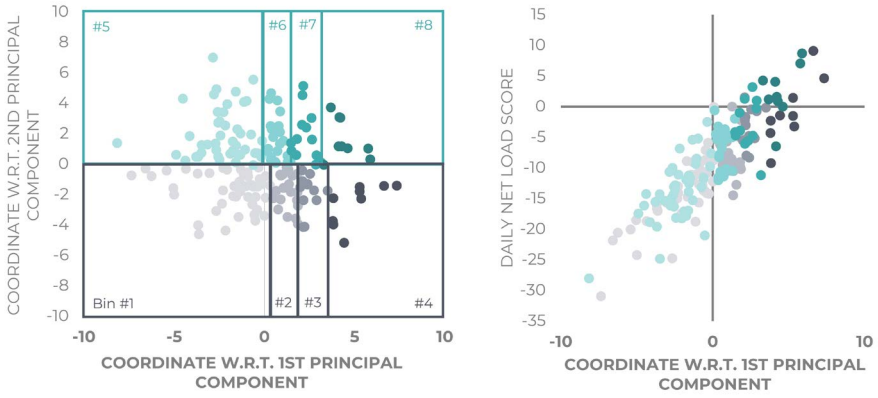
First weather principal component for January and relationship to daily net load score.



5. With the two-dimensional weather data, we then partitioned the data based on percentiles with respect to each dimension. This exercise was subjective, but based on a couple of key observations and principles. We found that the first principal component generally explained about 2 times more of the variance in the weather conditions than the second component, so we decided to break the data into more bins with respect to the first component (4 bins) than the second (2 bins). We also wanted to ensure that the bins representing the most extreme conditions contained enough days to ensure a variety of potential combinations, but not so many days that dramatically different weather conditions were being lumped together into the same bin. For this study, we sought to ensure that the most extreme bin included at least about 10 days across the synchronous record (2007-2014). We also observed that the historical data contained about 160-180 weekdays per month. Given these observations and priorities, we decided to first split the data in half based on whether each day was above or below the median coordinate with respect to the second principal component and then within each set of days, further partition the data based the percentile with respect to the first principal component with cutoffs at 90%, 75%, and 50%. The resulting bins are shown for August weekdays between 2007 and 2014 in Figure B.3. In the left panel, each day is shown in the coordinate system of the first two principal components to show how the partitioning into bins occurs. The right panel shows the same information that is shown in the right panel of Figure B.1, but with each day colored based on its bin assignment.

FIGURE B.3.

Weather binning across first two weather principal components.



B.2 SIMULATION DETAILS

After binning the historical data, each simulated year was developed via the following process:

1. We randomly drew the hydro conditions for the year from the historical or trended hydro record
2. We initialized the weather conditions with a randomly selected December day. The weather bin corresponding to the selected day was considered the prior weather bin for the first day of the year.

   We then looped through 364 days (i.e., 52 weeks) for each simulated year and for each day, we:

   a. Randomly selected a historical day in the month for which the previous day's weather conditions were in the prior weather bin. The weather bin corresponding to this selected day became the current weather bin. This Markov chain approach helped to ensure that the day-to-day weather transitions were consistent with historical observations.

   b. Determined whether the day was a weekday or weekend day (the day type) based on the study year calendar.

   c. Randomly selected solar, wind, and thermal derate shapes from within the current weather bin.

   d. Randomly selected a load shape from the subset of the days in the current weather bin that corresponded to the day type (weekend or weekday); and

   e. Set the prior weather bin equal to the current weather bin before proceeding to the next day.

In this study, forced outages and weather-based availability were modeled independently. However, in future work, the forced outage rate could itself be weather-driven and synthesized for each day using the methodology described above. This would require information that relates the probability of forced outages to weather conditions.

APPENDIX C.
DATA SYNTHESIS FOR THE WEATHER-SYNCHRONIZED SIMULATION

For the purposes of demonstrating the Weather-Synchronized Simulation approach, we chose to synthesize thermal, wind, and solar data using day matching to expand the coherent weather record to 2007-2020.

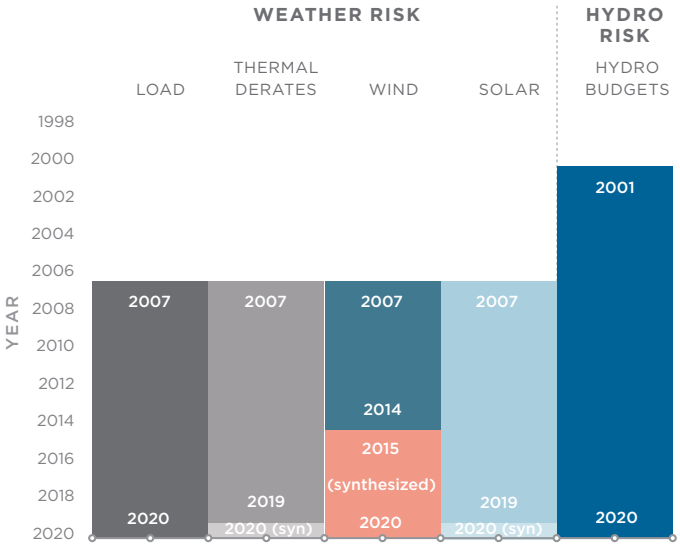


FIGURE C.1.
Historical and synthetic data used in Weather-Synchronized Simulation mode.

For each day in 2015-2020, we estimated hourly wind conditions by selecting the day in 2012-2014 that provided the closest match based on hourly historical wind generation in BPA and CAISO. Candidate day matches were limited to days in the same month or the month before or after the day to be matched. And the best match was selected by minimizing the root mean square error of the historical hourly wind capacity factors in BPA and in CAISO between the matching days. Historical hourly capacity factors in BPA and CAISO were estimated based on hourly generation data from the BPA Total Load and Wind Generation Report and the CAISO Daily Renewables Watch, respectively, as well as commercial online date information for projects included in each report. Even though wind generation data was available before 2012, we limited the candidate days to 2012-2014 because the rapid development of wind projects prior to 2012 led to less stable historical capacity factor estimates in those years.

An example of this day matching approach is shown in Figure C.2 for August 14, 2020. The closest matching wind day to August 14, 2020 across CAISO and BPA was September 13, 2012. On both days, CAISO wind output fell during the day to very low levels and increased somewhat during the evening hours, while BPA wind stayed very low throughout the day. We therefore use the wind shapes from September 13, 2012 to estimate the wind shapes for August 14, 2020 in the synthesized dataset.

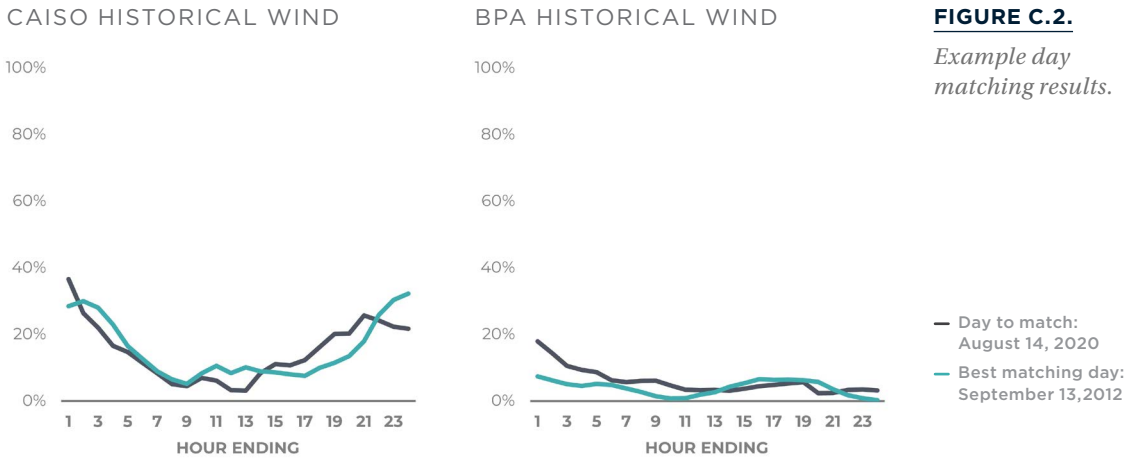


FIGURE C.2.

Example day matching results.

We applied similar approaches to synthesize solar and thermal derate data for 2020. We estimated the hourly solar shapes by selecting the day in 2018-2019 that provided the closest match based on hourly historical solar generation in CAISO. And we estimated the hourly temperature by selecting the day in 1998-2019 that provided the closest match based on daily average temperatures at the 16 locations across the West for which we had daily average weather data.

APPENDIX D.
DISPATCH MODELING DETAILS

To perform the dispatch modeling, we use GridPath, an open-source, versatile platform for power system planning and optimization developed by Blue Marble Analytics. We use GridPath to perform weekly optimizations with an hourly resolution over a wide range of system conditions. Each week is optimized independently and the model has perfect foresight for the whole week. The objective function used in this study is to minimize total unserved energy over the week plus the largest hourly capacity shortage experienced during the week.

GridPath has additional functionality not utilized in this study but that could be part of future RA studies including:

- Economic commitment and dispatch (this study minimizes unserved energy)
- Multi-stage commitment and scheduling (this study assumes perfect foresight)
- Flexibility constraints such as ramp rates and min up & down times (this study does not enforce flexibility constraints)
- Capacity expansion, including policy constraints such as renewable portfolios standards or carbon targets (this study does not optimize resource selections)
- RA contractual layer (this study models physical system only)

More information about gridpath is available at https://www.gridpath.io. The codebase can be found on GitHub at https://github.com/blue-marble/gridpath. Extensive documentation is available at https://gridpath.readthedocs.io/en/latest/.

D.1 CONTINGENCY OBLIGATIONS

Contingency obligations are modeled across four reserve-sharing groups — the Western Power Pool (WPP), the Southwest Reserve Sharing Group (SWSG), the California Independent System Operation (CAISO), and the Los Angeles Department of Water and Power (LADWP)—each with a contingency obligation equal to 6% of its load in each hour. The reserve sharing groups and BAAs included in each group are listed in Table D.1.

TABLE D.1.

Reserve sharing groups.

RESERVE-SHARING GROUP	BAAS
WPP	AVA, BANC, BPAT, CHPD, DOPD, GCPD, IPFE, IPMV, IPTV, NEVP, NWMT, PACW, PAID, PAUT, PAWY, PGE, PSCO, PSEI, SCL, SPPC, TIDC, TPWR, WACM, WAUW
SRSG	AZPS, EPE, IID, PNM, SRP, TEPC, WALC
CAISO	CIPB, CIPV, CISC, CISD, VEA
LADWP	LDWP

We allow coal and gas as well as hydro and battery resources located within each reserve-sharing group's BAAs to contribute toward the respective contingency obligation requirement.

D.2 TRANSMISSION REPRESENTATION

To convert from a nodal to a zonal representation, we made the following approximations:

- There are no transmission constraints within each Balancing Area. Flows are only modeled between Balancing Areas.
- Branches that connect Balancing Areas are aggregated into single lines where possible. Branches that are part of an interface (i.e., subject to path limits) are modeled individually when necessary to track interface flows and enforce interface constraints.
- To account for intra-Balancing Area branches that are part of an interface that crosses a Balancing Area boundary, we apply the following approximation: we use a recursive search algorithm to identify the subset of adjacent branches within 10 nodes of the intra-BA interface branch. Any branches within that subset that cross a BA-BA boundary determined to be part of the interface are modeled explicitly and are included in the associated interface constraints. A few remaining branches were manually allocated to interfaces.

We explicitly model transmission losses by applying a constant loss factor of 2 percent on transmission line flows.

APPENDIX E.
HISTORICAL WEATHER DATA

In both the Monte Carlo weather binning process and the statistical analysis that accompanies the Weather-Synchronized Simulation, we rely on historical weather data at various locations across the Western US. We selected the weather data (both the weather variables and sites) to maximize the number of years for which data could be obtained while also taking into consideration weather factors that may be related to regionally resource adequacy. The weather sites, which are listed in Table E.1, were selected because of their proximity to load centers and/or strong wind or solar resources as well as the availability of a long and consistent historical weather record.

Daily Global Surface Summary of the Day files were downloaded from the National Centers for Environmental Information (NCEI) as far back as 1936.[56] Most sites required pulling data from multiple station IDs, as station ID conventions have changed over time and some sites required multiple nearby stations to produce a complete record. For each site, we combined data from the station IDs listed in Table E.2. If overlapping data was available from multiple stations on the same day, we prioritized the station listed first in the table. We then inspected the combined data and kept only days for which at least 85% of the four weather variables across the sixteen sites had records. This yielded 26,579 days of weather data between December 1948 and October 2021. Table E.1 lists the percentage of days with missing data for each weather variable across this record. For these days, we filled in missing data using weighted k-Nearest Neighbor imputation.

56 Global Surface Summary of the Day - GSOD, National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce, available at: https://www.ncei.noaa.gov/access/search/data-search/global-summary-of-the-day

TABLE E.1.

Key findings using 30 forced outage iterations, compared to a forced outage derate approximation.

		DATA COVERAGE (12/19/1948 - 10/4/2021)			
SITE	STATION ID(S)	TEMPERATURE	DEW POINT	SEA-LEVEL PRESSURE	WIND SPEED
Albuquerque International Airport, NM	72365023050, 99999923050	100.0%	100.0%	99.9%	100.0%
Blythe, CA	74718823158, 99999923158	79.1%	79.1%	79.0%	79.1%
Casper Natrona County Airport, WY	72569024089, 99999924089	98.3%	98.3%	98.2%	98.3%
Colorado Springs Municipal Airport, CO	72466093037, 99999993037	100.0%	100.0%	99.9%	100.0%
Dallesport Airport, WA	72698824219, 99999924219	89.0%	88.9%	88.9%	89.0%
Edwards Air Force No Additions, CA	72381023114	91.6%	91.6%	91.5%	91.6%
Fresno Yosemite International Airport, CA	72389093193, 99999993193	99.0%	99.0%	99.0%	99.0%
Harry Reid International Airport, NV	72386023169, 99999923169	98.6%	98.6%	98.6%	98.6%
Los Angeles International Airport, CA	72295023174, 99999923174	100.0%	100.0%	100.0%	100.0%
Lewiston Airport, MT	72677624036, 99999924149	99.9%	99.9%	94.7%	99.9%
Portland International Airport, OR	72698024229, 99999924229	100.0%	100.0%	99.9%	100.0%
Phoenix, AZ (Phoenix Airport and Phoenix Luke Air Force No Additions)	72278023183, 99999923183, 72278523111, 99999923111	100.0%	100.0%	99.9%	100.0%
Sacramento Airport, CA	72483023232, 99999923232	100.0%	99.8%	97.4%	100.0%
San Francisco International Airport, CA	99999923234, 72494023234	100.0%	100.0%	100.0%	100.0%
Seattle, WA (Seattle Tacoma International Airport, Seattle Boeing Field)	72793024233, 99999924233, 72793524234, 99999924234	100.0%	100.0%	100.0%	100.0%
Tucumcari Municipal Airport, NM	72367623048, 99999923048	93.4%	93.3%	70.6%	93.3%

APPENDIX F.
ADDITIONAL MODELING INSIGHTS

F.1 MONTE CARLO SIMULATION CONVERGENCE BEHAVIOR

Figure F.1 and Figure F.2 show the convergence behavior of the Monte Carlo Simulation of the No Additions Scenario. The number of years to achieve reasonable convergence in an RA analysis will depend on the system under study and the reliability criteria. For this specific system, simulating 1,000 years was adequate to provide *LOLE* and capacity need estimates that were stable relative to the amount of uncertainty.[57]

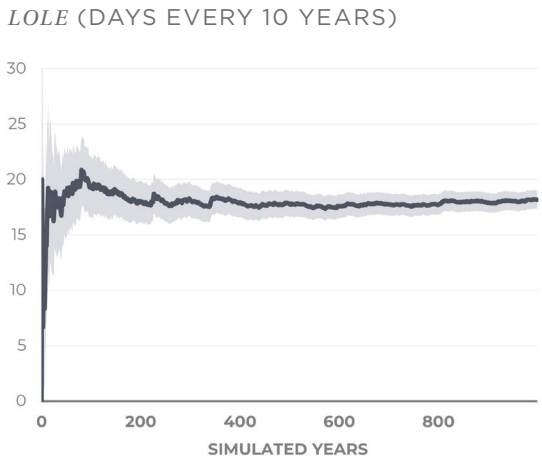


FIGURE F.1.
LOLE convergence for the No Additions Scenario using Monte Carlo Simulation.

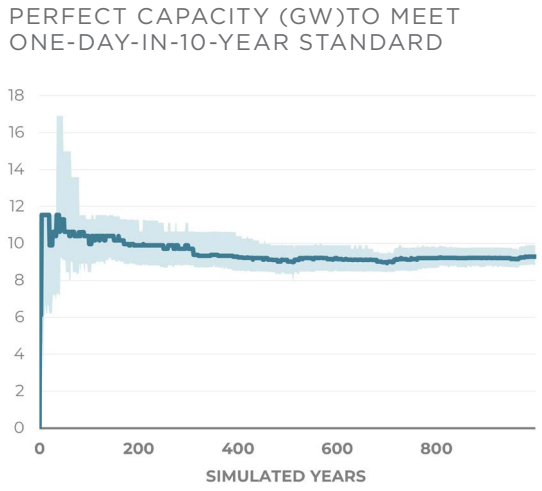


FIGURE F.2.
Capacity need convergence for the No Additions Scenario using Monte Carlo Simulation.

While the perfect capacity needs were estimated to be 9,267 MW for this system, the uncertainty in the *LOLE* implies that the perfect capacity

57 The error bands in both figures represent approximate 95% confidence intervals, assuming a binomial distribution for loss of load days.

needs could fall between about 8,800 MW and 9,900 MW, with about 95% confidence. Given the size of the system and the size of the shortage, this amount of uncertainty seems reasonable to inform decision making. More precise estimates could be achieved by simulating more years, but this would come at a computational cost and may lead to false precision given the scale of uncertainty regarding the state of the system in the study year.

F.2 WEATHER-SYNCHRONIZED SIMULATION CONVERGENCE BEHAVIOR

Figure F.3 and Figure F.4 show the convergence behavior of the Weather-Synchronized Simulation of the No Additions Scenario. Recall that each iteration consists of 14 weather years and 20 hydro years and those same 280 combinations of weather and hydro years are tested across 30 unique forced outage iterations via Monte Carlo Simulation. The first forced outage iteration in each simulation estimates forced outages using flat derates equal to the unit forced outage rates, while the other iterations use hourly exponential failure and repair models to simulate random forced outages throughout each year.

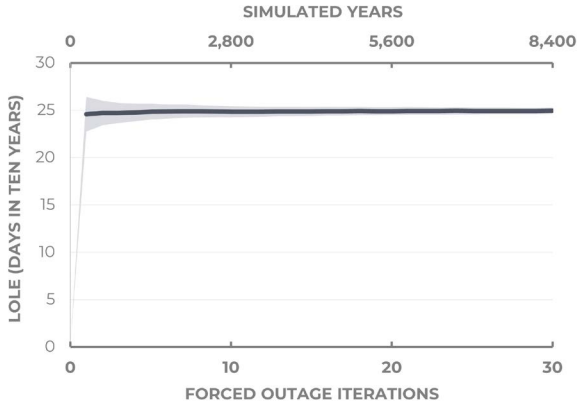


FIGURE F.3.

LOLE convergence for the No Additions Scenario using Weather-Synchronized Simulation.

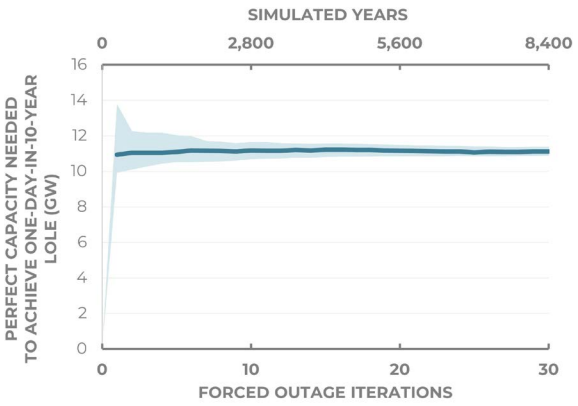


FIGURE F.4.

Capacity need convergence for the No Additions Scenario using Weather-Synchronized Simulation.

Figure F.3 and Figure F.4 show that increasing the number of forced outage iterations did not have a significant impact on the LOLE or capacity need estimates, but did reduce the uncertainty in those estimates. Estimating the effects of forced outages using simple derates, rather than Monte Carlo modeling, yielded values that were relatively close to the converged values.

However, using a derate approximation resulted in much higher uncertainties simply because fewer conditions were tested. We have less confidence in how well we understand the frequency of lost load when we only test 280 possible years compared to 8,400 possible years.

TABLE F.1.

Key findings using 30 forced outage iterations, compared to a forced outage derate approximation.

METRIC	30 FORCED OUTAGE ITERATIONS	FORCED OUTAGE DERATE APPROXIMATION
LOLE **(Days every 10 years)**	24.9 ± 0.3	24.6 ± 1.8
Perfect Capacity to meet *LOLE* **= 1 day in 10 years**	11,111 ± 285 MW	10,928 ± 2,866 MW

How much confidence should a planner expect to have in this type of exercise? It will depend on the system under study, the types of decisions that need to be made, and their timelines. For a system as large as the Western United States, precision to 100-200 MW is probably not necessary and could give policymakers and planners false confidence in their understanding of future risks. However, uncertainty in the range of several thousand MW may make it difficult to make decisions or weigh the prudency of plans and actions. For this system, about 1,000 MW of uncertainty across the Western United States may be reasonable and this was achieved with about five forced outage iterations in the Weather-Synchronized Simulation, or 1,400 simulated years. This level of precision was achieved a bit more quickly in the Monte Carlo Simulation, which required about 400-500 simulated years to estimate capacity needs to within 1,000 MW (see Figure F.5). For this system, the same degree of precision was achieved more quickly using Monte Carlo analysis, but potentially at the expense of accuracy and transparency.

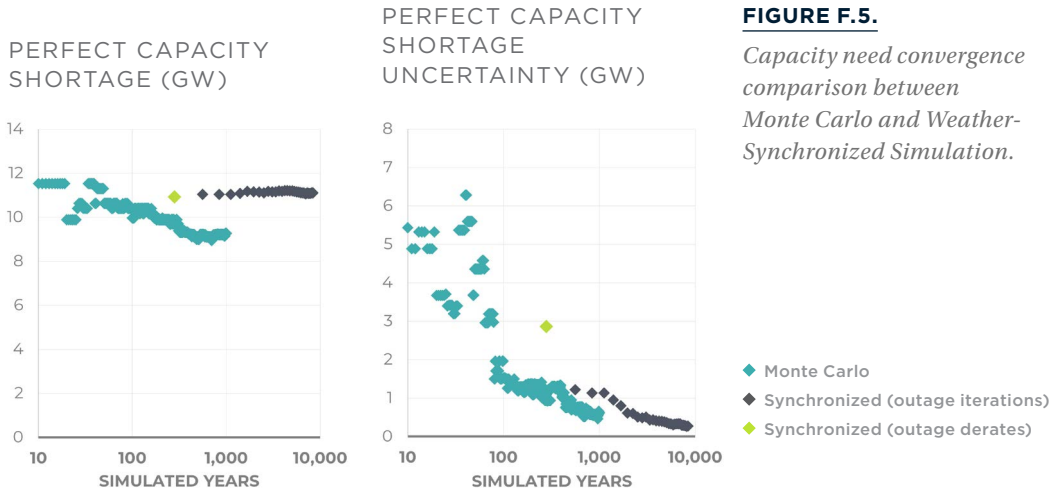


FIGURE F.5.

Capacity need convergence comparison between Monte Carlo and Weather-Synchronized Simulation.

F.3 HYDRO MODELING INSIGHTS

The level of attention given to hydro variability in RA analysis tends to depend on the extent to which the system relies on hydro resources. In the Pacific Northwest, for example, there is a tradition of testing the system under critical hydro conditions to ensure resource adequacy and the Northwest Power and Conservation Council has developed a more sophisticated approach that tests several decades of water years when examining regional resource adequacy. This study shows that hydropower plays a key role in supporting resource adequacy in the broader West, and so the level of rigor required to characterize hydropower in such studies is of interest.

One way of simplifying the analysis with respect to hydropower is to test a single representative hydro year, instead of many possible hydro years. We conducted two tests to examine whether this approach could be used to achieve reasonable results with fewer iterations for the modeled system. In the first test, we calculated the loss of load metrics based on a single year, 2003, which closely aligned with average monthly hydro conditions across the West (based on the West-wide hydro budgets). In the second, we calculated the loss of load metrics based on the year 2015, which had the lowest West-wide hydro budgets during June-September, in which loss of load risk was highest. The resulting perfect capacity needs and capacity need uncertainties are shown in Figure F.6, compared to Weather-Synchronized Simulation that considered all 20 hydro years.

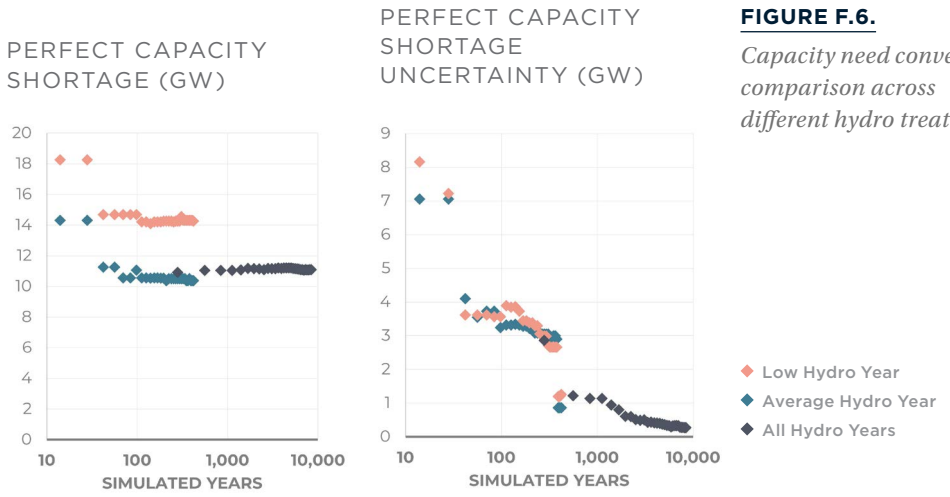


FIGURE F.6.
Capacity need convergence comparison across different hydro treatments.

We found that relying on a low hydro year to estimate hydro risk for this system overestimated capacity needs by about 3,000 MW. Relying on a single relatively typical hydro year only slightly underestimated capacity needs, but did not yield significantly improved convergence behavior or precision relative to modeling 20 years of hydro conditions.

F.4 COINCIDENCE OF WEATHER PATTERNS WITH DAYS OF THE WEEK

The Weather-Synchronized Simulation tests only days that have actually occurred in the historical record. Because there are so few days that experience weather conditions extreme enough to create potential RA risks, the results may meaningfully depend on whether those weather conditions were experienced on a weekend or weekday. For example, the most extreme weather conditions that were experienced during the heat wave in which CAISO experienced lost load in August 2020 occurred on August 16th, a Sunday. If those same conditions had been experienced on a weekday with the additional loads associated with weekday activities, the consequences could have been even greater. The statistical model for *LOLP* provides a simple way to test how important this risk may be for RA analysis because it allows for the estimation of *LOLP* under various combinations of weather conditions and weekends/weekdays that were not directly simulated. To test this, we applied the statistical model across the entire historical weather record twice — once assuming that each day was a weekday and once assuming that each day was a weekend day — and we calculated a weighted average of the resulting *LOLPs*.

This test tended to increase the estimated *LOLP* on weekends, but also tended to decrease the estimated *LOLP* on weekdays, which can be seen for the August 2020 heat wave in Figure F.7.

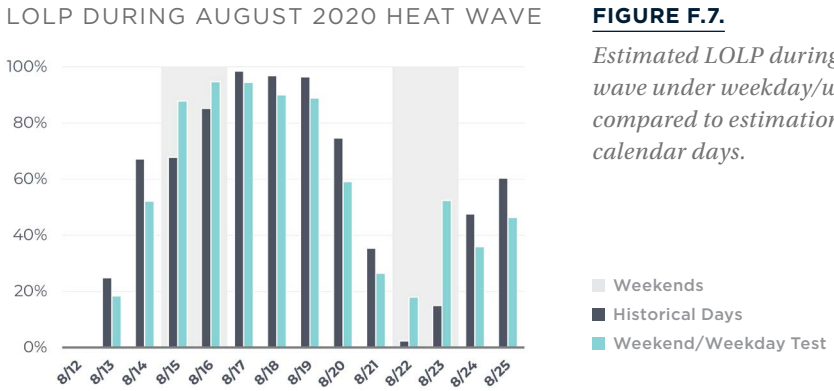


FIGURE F.7.

Estimated LOLP during August 2020 heat wave under weekday/weekend sensitivity, compared to estimation using historical calendar days.

As a result of these offsetting effects, we did not see significant differences between the resulting annual *LOLE* estimates and those that were estimated based on the calendars associated with the historical weather conditions (see Figure F.8). In this case, relying on historical calendar-based weekdays for RA analysis is likely sufficient.

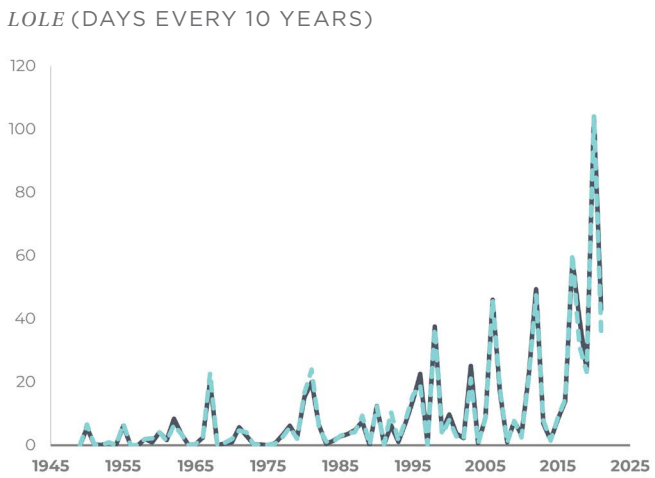


FIGURE F.8.

Estimated historical LOLE under weekend/weekday sensitivity, compared to estimation using historical calendar days.

F.5 WEATHER DETRENDING

Some RA analyses attempt to address historical weather trends by detrending the historical weather data underpinning their models—effectively adjusting the weather data so that the distribution is stable across the historical record. There are many ways to detrend data. In the example below, we simply translated the weather data in each year using a linear function so that the slope of the detrended data over the years was equal to zero and the mean was equal to the mean value in 2021.[58] This theoretically adjusted for historical trends, but did not assume that those trends would continue to the study year. We made this adjustment independently for each month of data to account for different seasonal trends. Figure F.9 shows the detrended indicator, w, when the slopes were calculated based on the entire historical record (over 70 years). Figure F.10 shows the same information, but zoomed in to focus on the most extreme events. The detrended data saw very high values of w more frequently throughout the historical record, which influenced the *LOLE* (Figure F.11).

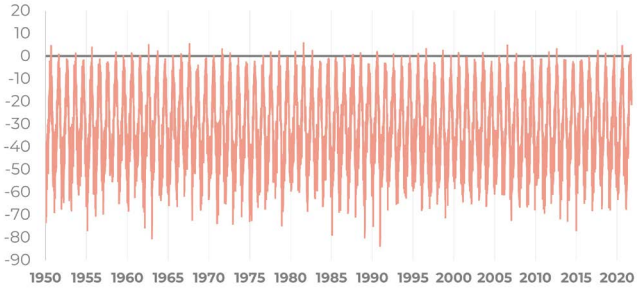


FIGURE F.9.

Critical weather indicator, w, over detrended weather record.

58 Note that the statistical analysis relies on the indicator, w, which has a linear relationship with each of the weather variables. Rather than transforming each weather variable individually and re-calculating w, the linearity of the transformations allowed us to simply transform w to capture the combined effects of transforming each weather variable.

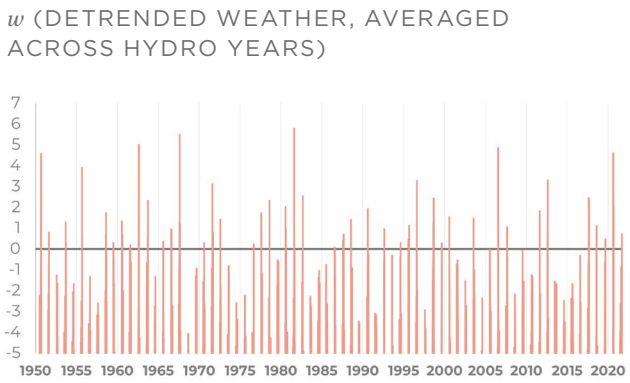


FIGURE F.10.

Critical weather indicator, w, over detrended weather record (focusing on extreme events).

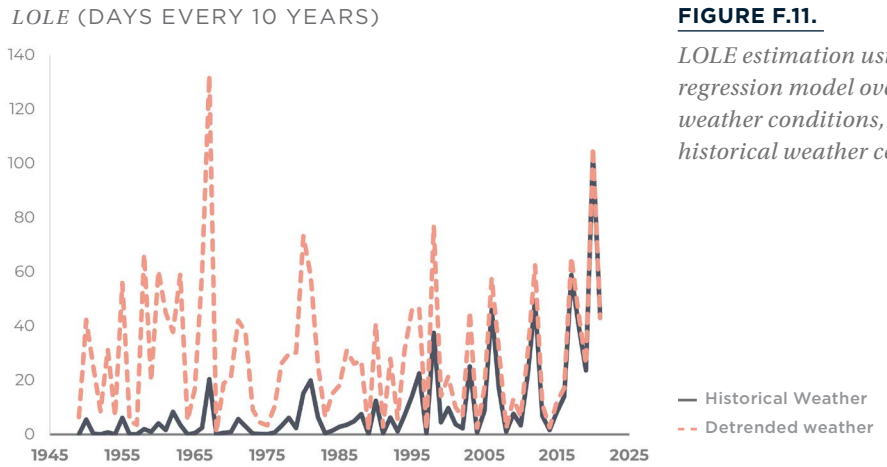


FIGURE F.11.

LOLE estimation using logistic regression model over detrended weather conditions, compared to historical weather conditions.

Figure F.11 shows the resulting *LOLE* when considering various historical periods of detrended weather. Note that for each historical period, the weather was detrended over the full 70+ year record, not just the years over which the *LOLE* was calculated. In this particular case, the *LOLE* estimates using detrended data were relatively stable (*LOLE* = 27 to 29 days every 10 years) when considering between 20 and 70 years of detrended weather conditions.

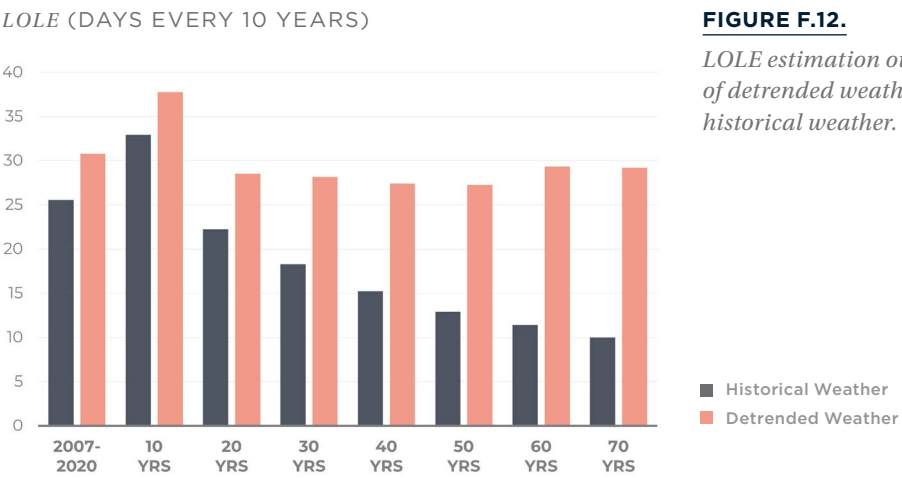


FIGURE F.12.

LOLE estimation over various periods of detrended weather, compared to historical weather.

While detrending offers a simple solution to adjust for past weather trends and achieve more stable estimates, it is important to remember that the detrended weather data does not represent real physical weather conditions and their potential frequencies in the future. As with averaging across historical weather conditions, without more information about future weather conditions, the decision to detrend a historical weather dataset for use in RA analysis is a policy decision based on risk tolerance. In this case, using detrended historical weather data provided a more conservative estimate than using historically observed weather conditions.

APPENDIX G.
STATISTICAL ANALYSIS DETAILS

This section describes a statistical analysis developed from the Weather-Synchronized Simulation results and used to estimate the relative contributions of various factors to the loss of load probabilities. In addition to providing information about the drivers of RA risk, this analysis can also be used to estimate loss of load during conditions that were not directly simulated. In this report, we use this functionality to test historical weather conditions for which high resolution data was not available and feasible combinations of weather and day of the week that were not experienced in the historical record. In the future, this type of approach could also be used to explore RA risk for future climate scenarios. It is worth noting that the statistical analysis draws from the results of the Weather-Synchronized results, which are specific to a particular power system configuration and year. To apply the statistical analysis to a future power system would entail first developing a physical model that represents the future power system (including the load and resources) and then applying the Weather-Synchronized Simulation approach before a statistical model could be developed.

The statistical model uses logistic regression to estimate loss of load probability on each day as a function of: the daily weather conditions; a parameter representing the variation in daylight hours across the year (to, for example, help the model differentiate between a hot day in June versus a hot day in August); a parameter differentiating between weekends (1) and weekdays (0); and a parameter representing the West-wide weekly hydro budget applied to the day (and the given hydro year) in the GridPath simulation.

Because there are 64 weather variables in the historical dataset (see Appendix E) and the GridPath simulation yielded only 105 unique weather days with lost load (in at least one iteration of at least one hydro year), care had to be taken to avoid overfitting with respect to weather. In this following section, we describe how we examined the effects of weather variables in order to significantly reduce the number of variables representing weather in the model.

G.1 IDENTIFYING KEY WEATHER DRIVERS

To isolate the impacts of weather, we first conducted a statistical analysis on the synchronized case with extreme forced outage conditions. Using an extreme forced outage case provided more days with lost load for the model to consider and avoided noise associated with random forced outages. The forced outage case resulted in lost load (in at least one of the 20 hydro years) in only 202 unique weather days across the simulation. To determine the key

weather drivers of RA risk, we fit a logistic regression model to estimate the *LOLP* as a function of weather and non-weather parameters:

$$LOLP_{ij} = 1/(1 - e^{-w_{ij}})$$

$$w_{ij} = aX_i + bY_i + cz_{ij} + d$$

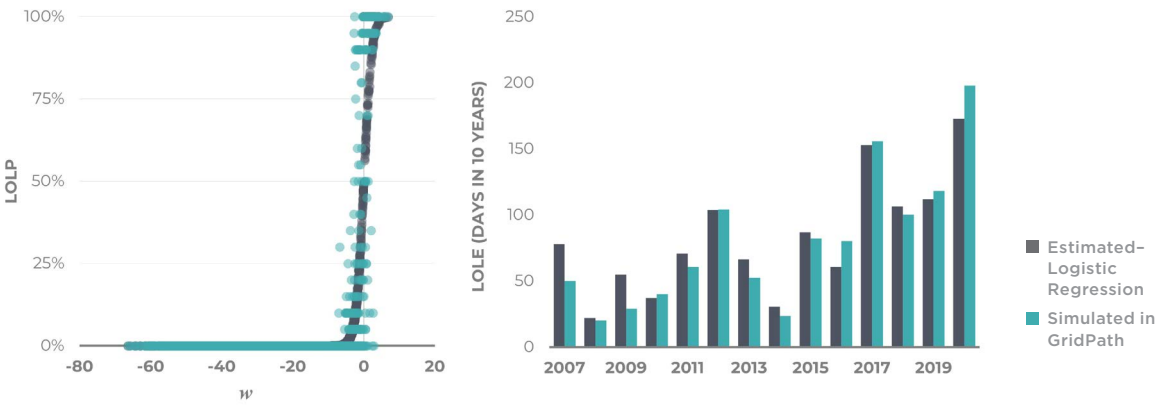
where X_i contains weather information for day i, Y_i contains the information for the day that does not depend on the weather or hydro conditions (daylight hours and weekend versus weekday), and z_{ij} contains the hydro conditions for the day on hydro year j. To reduce the dimensionality of the weather dataset and to help prevent overfitting, we used only the temperature and wind speed weather variable across the West.[59]

To avoid overfitting, we trained the model on approximately 80% of the data, taking care to ensure that about 80% of the unique weather days with lost load in one or more hydro years were in the training set so that weather conditions leading to lost load in the training and testing sets were materially different. We tested performance across 30 randomly generated training and testing sets to select a regularization parameter (C = 5). Performance was measured by the sensitivity and precision of the estimated day classifications as well as the root-mean-square error of the estimated *LOLP* in each simulated year.

After selecting the regularization parameter, we applied Bootstrap Aggregation across 100 randomly-generated training datasets. Figure G.1 compares the estimated *LOLE* using this approach to the *LOLE* based on the GridPath simulation of the synchronized extreme derate case. The left panel shows each day in terms of the vector w (averaged across the hydro years) and the right panel compares the average *LOLE* for each simulated year.

FIGURE G.1.

Logistic regression estimates for the No Additions Scenario under extreme forced outage assumption.



59 We separately tested a model that used principal component analysis across all of the weather variables to reduce the dimensionality and found a very small difference in model performance.

To determine the weather coefficients for the *LOLP* estimation, we examined *a*, the vector of weather coefficients in each of the 100 logistic regression fits. Each set of coefficients were slightly different because they were trained on different randomly-generated subsets of the data. We took the average across the 100 iterations to derive a single vector representing the contribution of each normalized weather variable to the likelihood of lost load. These are shown in Figure G.2.

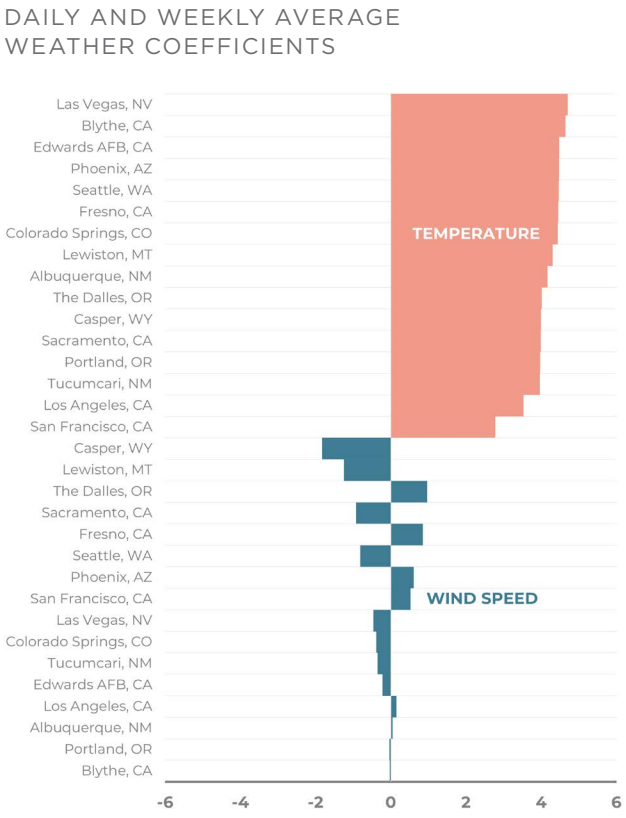


FIGURE G.2.

Weather coefficients in the logistic regression for the No Additions Scenario.

The analysis confirmed that temperature was a key indicator of loss of load risk in the Weather-Synchronized Simulation under the extreme derate assumption. Wind speed appeared to be less important, although high wind speeds in Wyoming and Montana did seem to correlate with lower loss of load probabilities.

G.2 TRAINING THE MODEL ON SIMULATED DAYS

We used a similar logistic regression approach to fit a model to the full set of simulated days across the forced outage iterations in the Weather-Synchronized Simulation. Because forced outages were less extreme in this case than the case described in the previous section, there were even fewer unique weather days with lost load — only 105 unique weather days between

2007 and 2020. To avoid overfitting, we used the coefficients derived in the prior section to reduce the daily weather conditions down to a single variable that represented how extreme the weather was on each day.

In addition to the daily weather variable, we also included a weekly weather variable equal to the average of the daily weather variables across the week in which each day was optimized in the dispatch simulation. This helped to account for GridPath's ability to trade off unserved energy between days within the same week and to account for limitations that may exacerbate RA challenges on a given day if it is surrounded by other challenging days (e.g., energy constraints). We then fit a logistic regression model to estimate the loss of load probability for each day based on the two weather parameters as well as the non-weather parameters listed in Section G.1 (each normalized to have mean 0 and standard deviation 1). The formulation was identical to the model described in the prior section, except that X contained only the daily and weekly weather variables rather than the 32 temperature and wind speed variables. We used the same approach to select the regularization parameter (C=0.1), and also applied Bootstrap Aggregation across 100 randomly-generated training datasets to arrive at the final estimates. Figure G.3 shows the resulting logistic regression coefficients.

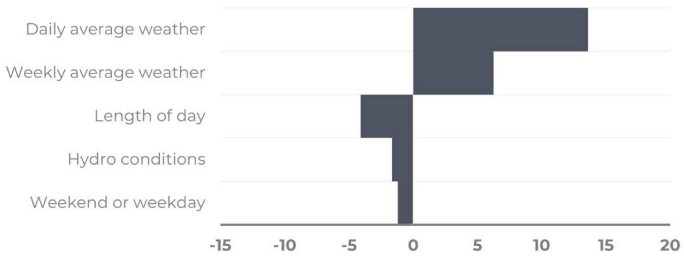


FIGURE G.3.
Logistic regression coefficients for the No Additions Scenario.

Figure G.3 compares the estimated $LOLE$ using this approach to the $LOLE$ based on the GridPath simulation of the No Additions Scenario under the synchronized weather approach. The left panel shows each day in terms of the vector w (averaged across the hydro years) and the right panel compares the average $LOLE$ for each simulated year.

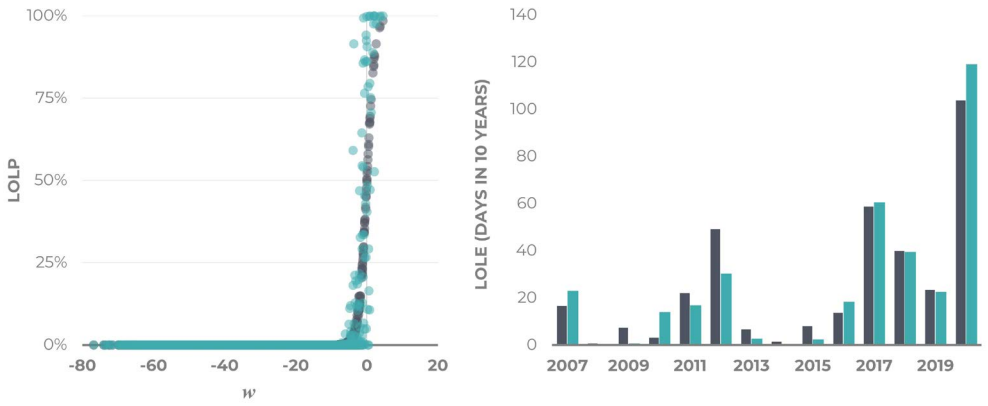


FIGURE G.4.

Logistic regression estimates for the No Additions Scenario

Note that the statistical analysis employed in this study was greatly simplified because all lost load was observed on hot summer days. Because this weather mode dominated the risk of lost load, we could estimate the classes (days with lost load versus days without lost load) as linearly separable and apply a linear classifier (i.e., logistic regression). If very different weather modes, for example cold snaps or more moderate temperature low wind days, resulted in lost load in the Weather-Synchronized Simulation, then the methodology described above could potentially be applied after an initial clustering step, or a more sophisticated neural network modeling approach could be applied. Generally, when using more sophisticated models, even more care must be taken to avoid overfitting.