Final Technical Report (FTR)

Agency/Office/Program	DOE/EERE/Solar Energy Technology Office	
Award Number	DE-EE0008215	
Project Title	Coordinated Ramping Product and Regulation Reserve Procurements in CAISO and MISO using Multi-Scale Probabilistic Solar Power Forecasts (Pro2R)	
Principal Investigator	Benjamin F. Hobbs, Theodore K. and Kay W. Schad Professor of Environmental Management bhobbs@jhu.edu 443-927-3376	
Business Contact	Lisa Walborn Senior Administrator lwalbor1@jhu.edu 410-955-2452	
Submission Date	June 13, 2022	
DUNS Number	001910777	
Recipient Organization	The Johns Hopkins University	
Project Period	**Start:** July 1, 2018	**End:** Dec. 31, 2021
Project Budget	Total $2,173,281 (DOE: $1,738,613; C/S: $434,668)	
Submitting Official Signature	B.F. Hobbs	

1. Acknowledgement: This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Forecasting II DE-FOA-0001649, Topic Area 3: Power Forecasts and Operational Integration Agreement Number EE0008215. Research performed by JHU (Ben Hobbs, Yijiao Wang, Qingyu Xu, Shu Zhang. Liang Chen); IBM (Hendrik Hamann (CoPI), Rui Zhang, Carlo Siebenschuh); Univ. Texas-Dallas (Jie Zhang (CoPI), Binghui Li, Li He, Mucun Sun); and NREL (Venkat Krishnan (CoPI), Xin Fang, Evangelia Spyrou, Ibrahim Kran).

2. Disclaimer: This report was prepared as an account of work sponsored by an agency of the US Government. Neither the US Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the US Government or any agency thereof. The opinions of authors expressed herein do not necessarily state or reflect those of the US Government or any agency thereof.

3. Executive Summary

How can probabilistic solar forecasts lower costs and improve reliability for independent system operator (ISO) markets? We tackle this question in three steps (Fig. 1). First, we enhance an existing solar forecasting system to provide well-calibrated hours-ahead probabilistic forecasts. We then relate the degree of uncertainty in those forecasts to error distributions for net load ramps for the California ISO (CAISO) using statistical and machine learning methods. Projected net load errors conditioned on solar uncertainty are translated into flexible ramp requirements that therefore reflect real-time meteorological and solar conditions, improving on typical ISO procedures. Finally, a multi-period look-ahead production cost model quantifies how conditional ramp requirements can a) decrease operating costs by lowering requirements compared to often conservative unconditional methods, and b) reduce generation scarcity events and consequently improve reliability by increasing flexibility requirements at times when unconditional forecast-based requirements understate actual ramp uncertainty.

Fig. 1. Organization of analysis of the cost savings and reliability improvements resulting from use of probabilistic solar forecasts to define ramp product requirements

In addition to the products just described (quantification of solar uncertainty, its translation into requirements for ramp capability product, and quantification of the benefits of more accurate ramp requirements), this project also developed a visualization system that alerts system operators of ramp and uncertainty conditions within the network based on solar forecasts. The system is called Resource Forecast and Ramp Visualization for Situational Awareness (RaVIS).

These four products represent significant advances in the state-of-the-art of probabilistic solar forecasting, development of weather-informed reserve requirements, production costing methods for estimating the benefits of more accurate reserve requirements, and visualization of system status, respectively. Yet the products are also practical and can be immediately implemented, potentially enabling system operators to save millions of dollars in ramp product procurement costs per year.

4. Table of Contents

5. Background

5.1 Task 1 Background: Solar Forecasting

The field of big data-driven probabilistic solar forecasting is evolving fast, driven by the rapid growth in cleaner energy sources, adoption of new decision-making processes by grid operators, and forecasting advances [1]. Solar forecasting research can be divided into (i) improved physics in numerical weather prediction models, (ii) developing non-physical (data-driven) forecasting approaches leveraging statistics, machine-learning and AI, and (iii) fusing physical and data-driven approaches [2].

What models may provide the best skills is a very complex question and depends on multiple factors including forecast horizons, forecast locations, and weather situations. In general, physics-based models perform well between 1-10 days and often cover large ranges, while data-driven approaches can have competitive or superior skills in short-term and longer-term forecasts but are often limited to point locations. The fundamental challenge with physics-driven approaches is not only that they are computationally very intensive as they require solving the full Navier-Stokes Equations on a “grid”, but perhaps more fundamentally that the physics is simply too complex and not fully understood, especially as it relates to turbulence, cloud formation and dissipation mechanisms. On the other hand, the trouble with purely data-driven approaches is that they not only perform poorly if there is not sufficient clean “training” data available but on a more fundamental level, they often fail in extreme weather situations, where there is no prior in the training data while at the same time good forecasting skill is most crucial.

Advances of physics-based numerical weather predictions have been described and summarized by Bauer et al [3]. An example of improving the forecast skill for solar radiation from a physics-based model is the development of WRF-Solar, a customization of the Weather Research and Forecasting (WRF) Model for solar forecasting applications. The customization included improved representation of aerosol–radiation feedback, the incorporation of cloud–aerosol interactions, and improved cloud–radiation feedback [4].

There is no shortage of research on data-driven methods, which include many regression, machine learning (ML), and AI models. Overviews can be found in [5-7]

It became evident from our review of prior work that we not only must integrate physical and data-driven approaches but also must make data-driven approaches more robust. Hence, more scalable approaches that leverage big data technologies are required. With such technologies models will have more training data, enhancing their robustness. Also, tens to hundreds of different models can be leveraged for model blending and selection.

Towards that end, the approach in this research project is unique and ahead of its time by pioneering the limits of data-intensive solar forecasting approaches. While scaling such approaches to larger and more complex situations remains a challenge, we have begun to make significant contributions by applying big geospatial data technologies to probabilistic solar energy forecasting, which for example enabled the scalable integration of multiple models and exploitation of massive data sets such as GOES-R.

5.2 Task 2 Background: Informing Reserve Requirements with Forecasts

The increased contributions of uncertain and variable resources has prompted changes in power market design and operator practices [8]. Operators assess solutions that help manage the uncertainty and variability of load and renewable generation. Among these are reserve products, especially the novel ramp product, which is an important feature in several U.S. power markets (MISO [9], CAISO [10], Southwest Power Pool [11]) to mitigate rapid net load changes caused by increased penetration of variable resources.

A key step in the market scheduling process is to define ramping product requirements. This involves trade-offs between market efficiency and system reliability. Too low or too high requirements can pose a risk to system reliability or system costs, respectively.

Historically, studies that define requirements for reserves and co-optimize energy and ancillary services can be divided into two categories based on how the requirements are determined: endogenous and exogenous [12,13]. Endogenous methods usually model the renewable and load uncertainties explicitly using a variety of techniques, such as stochastic optimization, robust optimization, and chance constraints, and simultaneously optimize where, how much, and when reserves are most economically provided to meet possible realizations of net load. However, the computational burden of endogenous methods is at least one of the reasons why no U.S. power system operators have adopted such a method. Instead, in practice, power system operators rely on exogenous methods. The exogenous methods estimate system-level ancillary service requirements *ex ante* with lead times specific to particular markets. Traditionally, system operators use deterministic methods, where certain percentages of peak load are used as reserve margins. This method, despite its simplicity, usually fails to satisfy system reliability standards efficiently owing to growing uncertainties caused by increasing renewable penetration.

Consequently, probabilistic methods [14] are sometimes adopted to estimate ancillary service needs based on the uncertainty levels of net load. Probabilistic forecasts explicitly account for uncertainty information, usually in the form of parametric probability distributions, or non-parametric forms such as quantiles, uncertainty intervals, and kernel density estimates. In the last decade, the value of probabilistic forecasts became increasingly recognized in power system applications [5]. For example, in [16], probabilistic wind and load forecasts were used to determine operating reserve requirements based on trade-offs between reliability and procurement costs. In a SF-II project supported by SETO, Ref. [17] also used probabilistic forecasts as input for generation of scenarios that were later used to determine operating reserve requirements. Refs. [18], [19] showed how to reduce the regulation requirements by employing probabilistic net load forecasts, which were derived from probability distributions of multiple net load components by convolution.

Although probabilistic forecasting is adopted in some markets, it is used primarily to improve situational awareness and rarely plays any important role in the decision-making process. Our work in this project represents a practical method that is easily implementable and extensible by any power market operators. This section features a particular method for relating solar probabilistic forecasts to reserve requirements, the k-nearest neighbors (kNN) method, which can also be used to define requirements for other market products that depend on historical data, such as regulation and non-spinning reserves.

5.3 Task 3 Background: Existing Visualization Capabilities & Forecast Integration

Control centers typically host several screens for visualizing network situational awareness, such as network one-line diagrams, circuit breaker/switch status, and graphs summarizing available reserves. Information about impending high-impact contingencies, as well as their impact on transmission line overloads, and in some cases system voltage and transient stability are displayed [20]. The idea of situational awareness in the form of real-time visualization, updates on event alerts, and plausible control locations/parameters is not new; however, enhancements are desirable because there are higher penetration levels of variable renewable generation and associated uncertainties.

With increasing uncertainty and emergent dynamic events, the awareness that is conventionally reported at aggregated levels is insufficient. The following capabilities are needed:

- Innovative ways to integrate variable renewable forecast data—including at higher spatial resolutions and for distributed site-level resources—apart from conventional aggregated data. Given the role of distributed variable renewable resources to provide system flexibility in the future, individual resource's forecasts will become key to improved visualization and observability to the system operators.
- Probabilistic forecast data integration that quantifies uncertainties related to variable renewables, which will be key to assessing risk associated with system dispatch.
- Timely alerts of excessive ramping events for net load and for each component of the net load—namely, different variable renewable resources.
- Zonal-, regional-, and even nodal-level generation flexibility information, which will ensure the timely deliverability of flexible resources in to offset unanticipated resource or net load ramps and to ensure reliability and resilience.

This section discusses some salient work performed in this space during the last decade—specifically, visualization capabilities that were built in response to the drastic increase in wind and solar generation penetration across the world, as reported in [21,22].

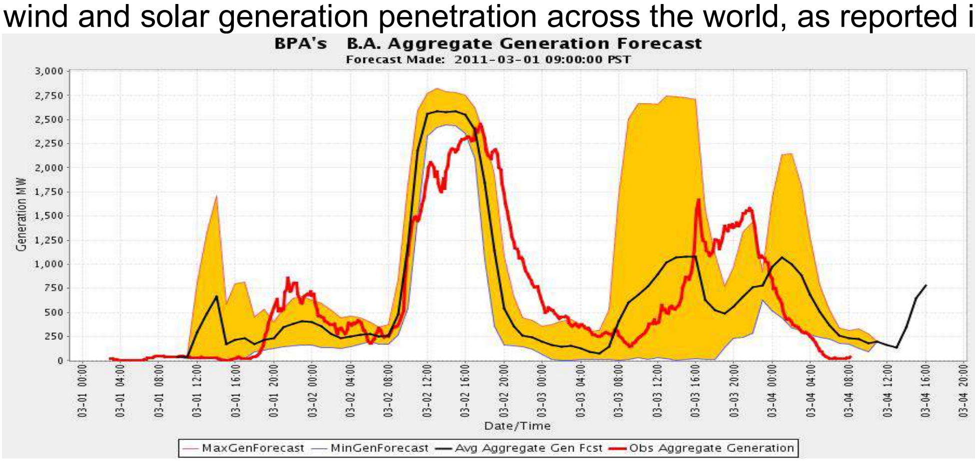


Fig. 2. BPA prototype of forecast display in the control center [21]

The simplest control center visualization tool for variable renewable forecasts is a time series of an aggregated average forecast (typically at 5 or 60 min intervals), bounded by upper and lower limits [21], as well as past observed generation. Fig. 2 [21, Fig. 57] shows an example from the Bonneville Power Administration. Worst-case scenarios define bounds in most cases, such as power loss from icing.

Some operators like to use these data to gain further information on the needed reserves, as shown in Fig. 3 from the Organization of the Nordic Transmission System Operators [21]. This is an example of using the variable renewable generation forecasts to estimate the needed reserves; it displays the available reserves across various regions.

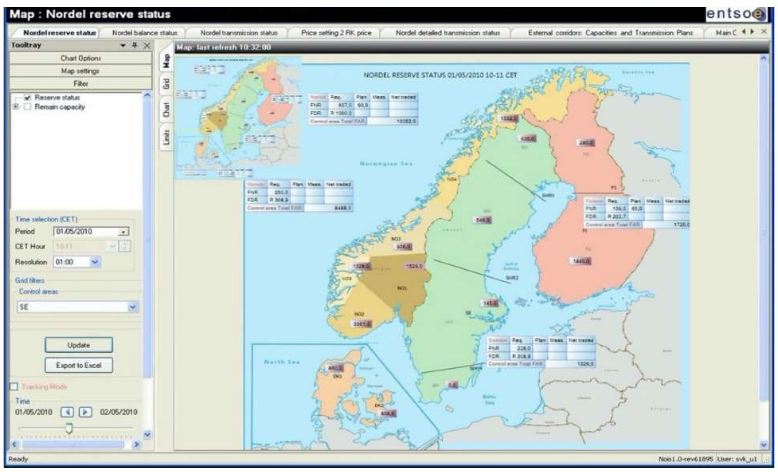


Fig. 3. Nordic Transmission System Operators reserve status situational awareness [21]

Fig. 4. Alberta Electric System Operator display on net load ramp [21]

Fig. 4 [21] presents another prototype decision support tool, developed by the Alberta Electric System Operator for their generation dispatch operations. This visualization and decision support tool facilitated the integration of variable renewable generation forecasts (specifically wind power in the Alberta, Canada, region) and system load forecasts into system operations. The visualization screen shows forecasts and available reserves (capacity & ramp) in the 60-minute time interval.

The illustrations shown so far include aggregated displays of forecasts and needed reserves. Fig. 5 [21, Fig. 78] shows regional (heat map) wind speed forecasts and expected ramp events and alerts in the Texas system. Fig. 6 (from [21]) shows another example from RTE France, where site-level wind power forecasts are related to real-time power flows on the transmission system and resulting bottlenecks.

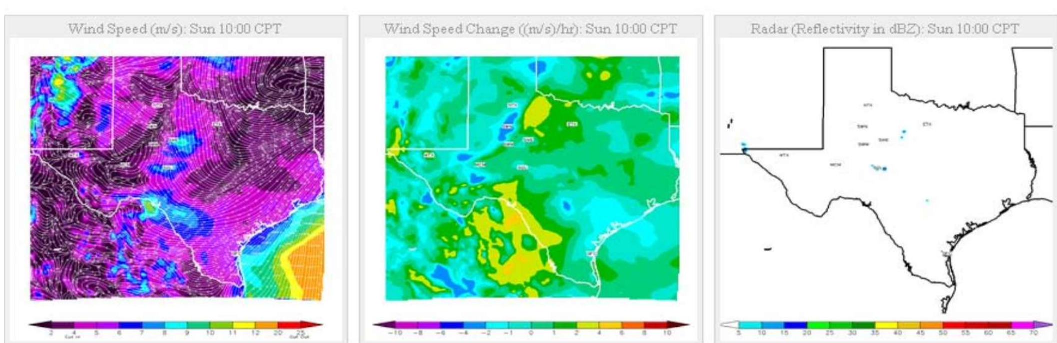


Fig. 5. Electric Reliability Council of Texas large ramp alert: (L) wind speed, (Center) wind speed change (hourly), (R) radar reflectivity [21, Fig. 78]

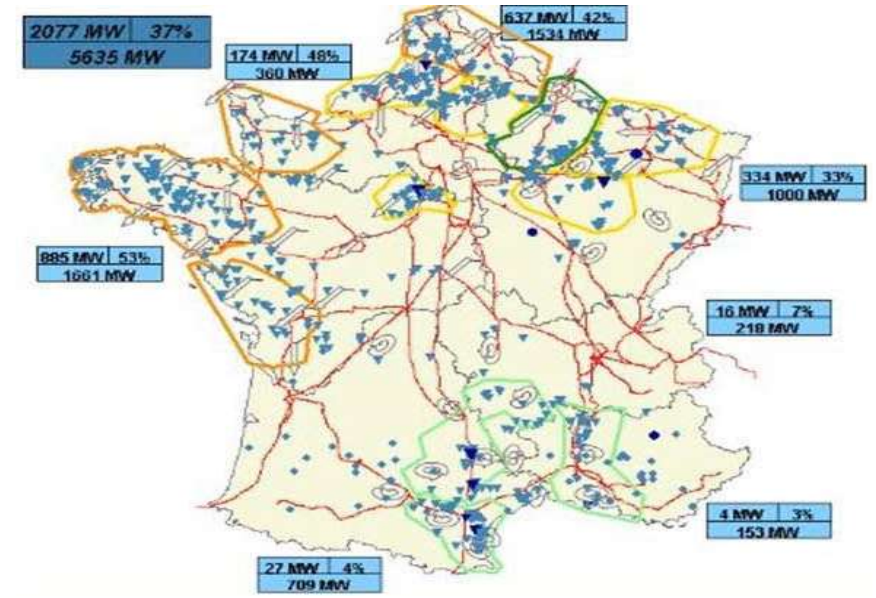


Fig. 6. RTE France main operator display for regional forecasts and congestion [21]

Most previous efforts used deterministic forecasts, and there were facilities to include additional "what if" scenarios and the consequent events and alerts. Fig. 7 shows a collaborative effort from the Pacific Northwest National Laboratory (PNNL) and CAISO to integrate probabilistic forecasts specifically to understand if there was enough real-time ramping capability in CAISO [23]. Fig. 7 shows the available generation capacity (the gray area) overlaid with the probabilistic net load forecasts, indicating the points in time where generation availability might not be enough to meet the net load changes. Fig. 7 shows—apart from the available generation and probabilistic forecast—the up/down ramping capability that is needed from the generation.

5.4 Task 4 Background: System Simulation to Assess Cost and Reliability Impacts of Improved Reserve Requirements Based on Solar Forecasts

Past projects have demonstrated the ability of ramp products to manage uncertainty and variability. Ref. [24] showed that ramping constraints can help operators meet expected variability; however, ramping constraints might not be cost-efficient when the deployment costs of ramping are not considered in the objective function that the market software optimizes. Refs. [25,26] demonstrated how ramping constraints can assist system operators in managing uncertainty, but ramping constraints might be less cost-efficient than an "ideal" stochastic unit commitment method. Ref. [27] discussed how the occurrence of variability and uncertainty can lead to energy imbalances and undesirable outcomes, such as power balance violations, real-time price spikes reflecting administrative penalties for violating constraints, leaning on regulation or interconnection, and out-of-market corrections. That paper concluded that a ramping constraint and product are more practical for managing uncertainty and variability than complex, time-consuming models and market formulations that capture stochastic processes and multiple possible futures.

However, the economic efficiency of ramping products depends on effective estimates of ramping requirements [25,28]. In practice, ISOs use samples of past forecast errors and calendar information, such as hour and type of day, in order to estimate a parametric or empirical probability distribution function (PDF) of net load (load minus renewable generation) [9,10]. Then, operators estimate either moments of parametric distributions [29] or percentiles [30] that they later use to determine ramping requirements. As systems change, ISOs assess the impact of additional factors on the PDF of net load. In particular, with the increase in variable energy resources (i.e., solar and wind generators), uncertainty related to forecasted generation by those resources might also increase. Weather conditions affect uncertainty related to forecasted generation by variable energy resources [31]. Therefore, state-of-the-art research is investigating whether real-time weather forecasts and measurements can be leveraged to estimate the net load PDF and associated balancing needs with increased accuracy.

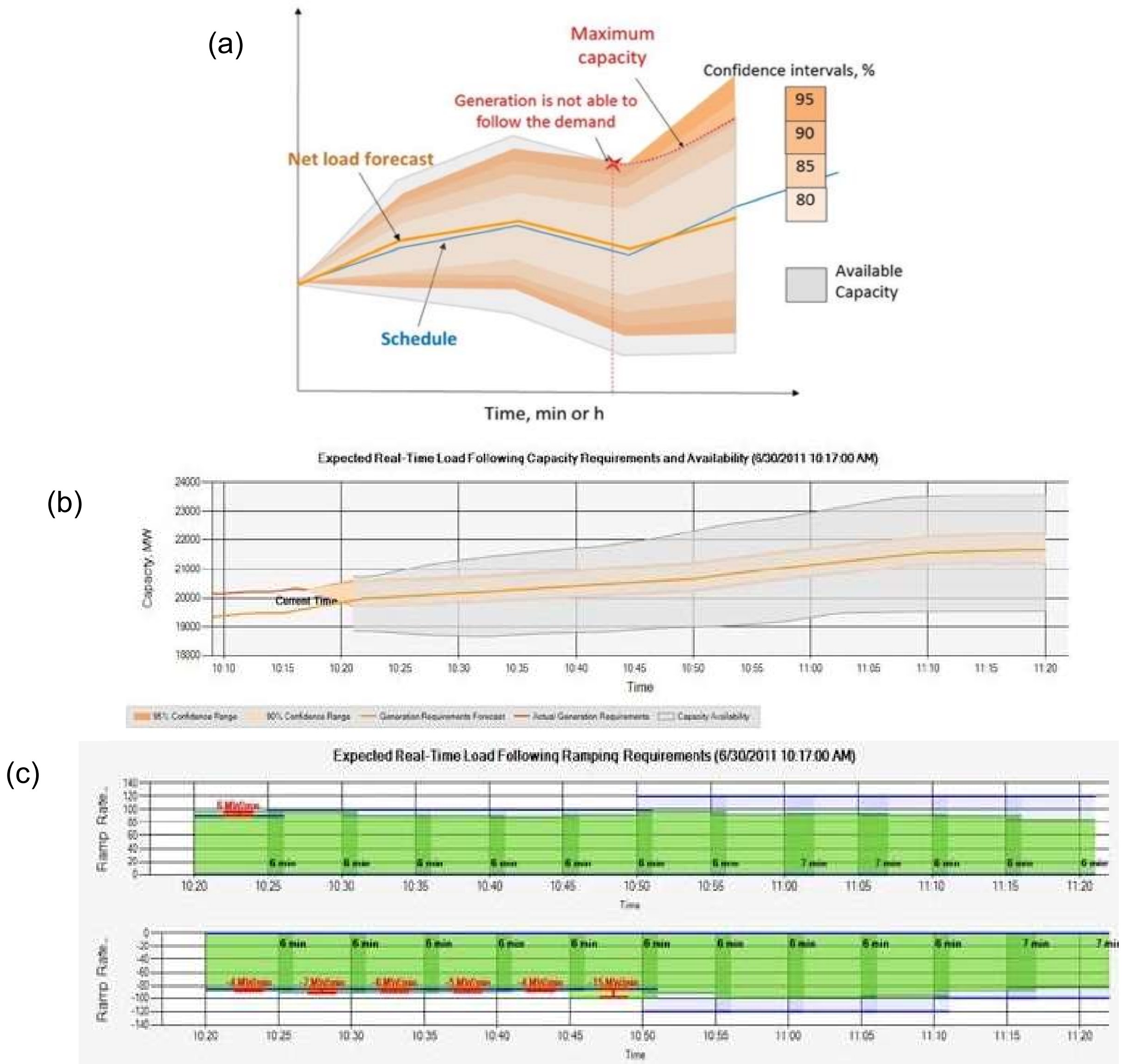


Fig. 7. CAISO ramping feasibility visualization developed by PNNL: (a) net load ramp, (b) load-following capacity requirements and availability, and (c) load-following ramping requirements [23, Used with permission]

For instance, CAISO is currently exploring the impact of weather variables on historical uncertainty of net load. Preliminary quantile regression analysis suggests that generation forecasts for variable energy resources and potentially temperature are statistically significant for the historical uncertainty of net load [32]. Ref. [33] used forecasts by numerical weather prediction models to develop probabilistic forecasts of net load and concluded that weather- informed probabilistic forecasts of net load can yield different (on average lower) balancing and regulation requirements than status-quo methods. Ref. [34] reached a similar conclusion by integrating forecasts of renewables as well as weather (irradiance, wind speed, temperature) into a dynamic reserves sizing methodology in Belgium.

The effectiveness of such new sizing methods for balancing products [33,34] has been assessed in two ways. First, researchers record how frequently estimated requirements exceed the actual needs during a historical period, thereby estimating the reliability level of the proposed methods [33]. Second, researchers [33,34] compare the estimated requirements to the status quo requirements. They hypothesize that higher (than the status quo) requirements will improve system reliability, whereas lower requirements will reduce costs; however, there are no reports in the literature of estimates of the value of alternative balancing product sizing methods in terms of quantitative reliability and economic performance. Therefore, the validity of the hypothesis is yet to be tested.

6. Project Objectives

This project's objective is to assess the impacts of adopting improved flexible ramp product (FRP) requirements on power market operation from both economic and reliability perspectives. FRP is the procurement and possible deployment of spare capacity in operating markets in order to accommodate forecast or unpredicted changes, or "ramps," in net load (gross power demand minus wind and solar generation). This product is an increasingly important tool for managing net load variability and uncertainty in day-ahead and/or real-time markets run by three major ISOs (CAISO, MISO, and SPP). Solar generation is large and growing contributor to this variability and uncertainty, and its effective management will be critical to economic and reliable integration of the large amounts of solar energy that is anticipated to come on-line in coming decades. Achieving the nation's energy transition goals depends critically on the success of this integration.

Since net load variability and uncertainty depends strongly on weather conditions and instantaneous amounts of wind and solar output, so too will the amount of FRP needed to manage them. Too much FRP, and the economic efficiency of the grid will degrade because of unnecessary procurement costs are incurred. Too little FRP, and system reliability will be at risk or excessive costs will be incurred due to the need to quickly start up high-cost generators. Recent developments in forecasting tools that enable operators to characterize uncertainty in the components of net load—gross load, solar, and wind—have the potential to help define FRP needs more accurately since uncertainty in those components likely contribute significantly to uncertainty in net load ramps.

The objective of this project is to test this hypothesis by developing a state-of-the-art solar probabilistic forecasting tool (Task 1); developing and testing statistical and machine learning models that relate solar prediction width forecasts (a measure of solar uncertainty) to net load uncertainty, based on ISO data (Task 2); creation of visualization tools that communicate ramp and uncertainty forecasts to operators (Task 3); and production simulation analyses for an ISO to quantify the reliability and operating cost benefits of more accurate forecasts of FRP needs (Task 4). Detailed goals by task are summarized next. These tasks have been conducted in close consultation with two ISOs (MISO and CAISO), and have yielded practical and demonstrably beneficial tools for forecasting solar uncertainty and defining FRP requirements. In particular, as described later in this final report, the tools have met the forecast improvement and cost savings targets set forth in the project's milestones and go/no-go decisions.

Task 1 Goal: Advanced big data-driven probabilistic solar forecasting platform. Task 1 of this project includes subtasks to enhance the previous deterministic IBM Watt-Sun forecasting system to provide high fidelity and accurate probabilistic forecasts for the CAISO and MISO balancing areas. Because the Watt-Sun technology is based on machine-learned blending of multiple forecast models, it works best with a large amount of historical data (TB per day). Towards that end, the team integrated the Watt-Sun forecasting system into IBM's PAIRS big data technology, based on a Hadoop/Hbase cluster, allowing distributed and scalable processing [35,36]. PAIRS provides a generic and highly scalable platform to train Watt-Sun on hundreds of terabytes of big data from many forecasting and historical sources. Based on PAIRS big data-based error characterization, multi-expert machine learning methods were to be employed to enhance Watt-Sun's model blending, and to generate probabilistic solar power forecasts. This allows systematic and thorough characterizing of forecast errors of current models as a function of forecast horizon, weather situation and locations, and to learn a blended "super" model for many sites and regions in the respective ISOs. The resulting probabilistic forecasts have also been shared in mutually agreed upon formats to the validation team in Topic Area 1.

There are four innovations we planned to add to the original Watt-Sun forecasting system from Solar Forecasting I.

Innovation 1. The big data bus of the initial Watt-Sun system reached its limitation in terms of scalability and how much data can be injected. The reason is that the data size relevant to solar forecasting has been exploding, for example:

- Global Forecast System (GFS): 140GB/day, increasing to 1.5 TB/day
- Global Ensemble Forecast System: 302GB/day, increasing to 3 TB/day
- GOES: 80GB/day for CONUS, increasing to 900 GB/day

Thus, the Watt-Sun data management system has been replaced by PAIRS (*P*hysical *A*nalytics *I*ntegrated *D*ata *R*epository and *S*ervices), a scalable platform for geospatial-temporal data. PAIRS improved speed, maturity, and data processing throughput more than 50-fold compared to the previous system. It enables "automatic" fusing of satellite, weather and sensor data; support tens of PB; and can inject data much faster than the current system (eventually up to hundreds of TB/day). Crucially, PAIRS can distribute forecast data in a scalable matter. PAIRS supports a state-of-art REST API and SDK, enabling project partners and others to build applications on top of the system. In addition, PAIRS supports WMS mapping services as inputs to visualization applications.

Innovation 2. This goal was to develop a new short-term solar forecasting module, leveraging initial work from Solar Forecasting I, where we enhanced "convection-based" forecasting-based GOES satellite observations with a 2D Navier-Stokes equation. This module needed to be more thoroughly tested and then put into operation for the continental US. We leveraged the new GOES-R data in this task; that data has significantly better spatial, temporal, and spectral resolution than the previous GOES (-13/-14) satellite data.

Innovation 3. A powerful innovation that resulted from the Solar Forecasting I project is situation categorization for the machine-learning. This allowed separate modelling of specific weather situations, thereby enhancing overall accuracy. In Solar Forecasting I, the situations were identified using FANOVA (functional analysis of variance), which has its limitations because in essence it only identifies categories based on "point" validation data (such as from a single solar plant). A goal of Task 1 of this project was to extend this work by using deep learning techniques, which can identify situations based on full images or raster observations, such as from the GOES satellite.

Innovation 4. A primary goal of Task 1 of this project was to extend Watt-Sun's capabilities to include probabilistic estimates for irradiance for points and regions.

Task 2 Goal: Coordinated reserves procurement in UC/ED with probabilistic forecasts. The goal of this task's activities was to develop the modeling framework for usinig the probabilistic solar power forecasts to estimate requirements for flexible ramping product and dynamic regulation reserves in the ISO unit commitment and energy dispatch processes. Short-term (0-6 hr) as well as day-ahead forecasts were to be used to estimate reserves requirements for day-ahead and real-time markets. The amounts and reliability of requirements were then to be compared with baseline ISO methods.

Task 3 Goal: Visualization of probabilistic ramp forecasts for situational awareness. This task had the goal of developing visualization tools for presenting the probabilistic solar and net-load ramp forecasts to ISO control room operators, and update the visualization, as new forecasts become available. This task also involved performance of simulations to mimic real-time control center decisions to evaluate feasibility and impacts of advanced visualization on operational decisions. The activities included working with the ISOs, demonstrating early prototypes, identifying ISO-tailored functionalities, tool specifications and software requirements. In addition, it encompassed developing the back-end probabilistic ramp forecast database for front-end visualizations.

Task 4: Co-ordination with ISO for testing probabilistic solar forecast integration. This task began with efforts to understand the requirements and process of integrating probabilistic solar power forecasts-based products into ISO operations and control center visualizations. Then Level-2 integration testing was performed with ISO collaboration. The task focused on working closely with the ISO partners' forecasting, market operations, and control center teams throughout the course of the project to enable effective integration of forecasting products into their development environment and assess the cost-benefits. The impact of those products integration on system economics and reliability was first assessed using IEEE test systems and later with CAISO scale system models. Testing in BP2 took place on ISO-scale systems at NREL; it was also intended that BP3 testing would be conducted at the CAISO using off-line ISO testing software and databases, but COVID restrictions shifted the location of those simulations to NREL's high performance computing facility. This task also assessed the barriers for market adoption and develop mitigation or promotional strategies for market transformation.

Table 1 on the next two pages summarizes the milestones associated with each task, and the status of each milestone. Please note that Tasks 1,5, and 9 are the activities of Task 1 (forecasting system) occurring in BP1, BP2, and BP3, respectively. Similarly, Tasks 2, 6, and 10 are the year-by-year activities of Task 2, and so forth.

Table 1. Project Tasks, Milestone Completion Dates, and Milestone Status

Milestone Schedule		**Milestone Completion Dates**		**Milestone/ Target Status**
Task#/ Milestone#	**Milestone Description and Target (if applicable)**	**Original Planned**	**Actual**	
1.1	**Probabilistic Watt-Sun models for two regions**			
M1.1.1	*Develop live preliminary Demo #1 Version 0.0 of Watt-Sun. Target: Average Brier Score < 0.5 for 28 consecutive days for one point site using 4-hour forecasts to show that probabilistic scores are produced.*	*Q1(10/31/18)*	*Q2 (12/31/18)*	*Met, see separate Milestone Report.*
M1.1.2	*Identify data requirements from PAIRS for blending data sources in Watt-Sun V. 1.0.*	*Q4 (6/30/19)*	*Q4 (6/30/19)*	*Met*
M1.1.3	*Demo #2: iterative update; evaluate forecast performance. Target: Improved Watt-Sun (Version 1.0) probabilistic solar power forecasts, with point forecast accuracy > 10% compared to the best persistence-based baseline models,*	*Q3 (3/30/19)*	*Q3 (03/31/19)*	*Met*
1.2	**PAIRS platform advancement, forecasting error characterization**			
M1.2.1	*Revised data management plan; ascertain relevant data sources for solar forecasting. Parallel computing to curate/harminoze all data.*	*Q4 (6/30/19)*	*Q4 (6/30/19)*	*Met*
2.1	**Procedures to estimate FRP, regulation requirements using probabilistic forecasts**			
M2.1.1	*Extract ramp information from forecasts from the OpSDA method using Watt-Sun V. 0.0 forecasts. Target: Predictive intervals coverage probabilty (PICP) to measure the reliability of ramp predictions. For key ramp features in solar power and netload, the final method is expected to achieve PICP>90% for short-term forecasts, and PICP>80% for day-ahead forecasts for three point locations in a high solar system.*	*Q3 (3/30/19)*	*Q3 (3/30/19)*	*Met*
M2.1.2	*Estimate FRP requirements with net-load forecasts using Versions 0.0 and 1.0 of Watt-Sun. probabilistic forecasts. Target: Ramping requirements (the part caused by solar forecasting errors) with probabilistic solar forecasts will be 10% less than those with deterministic solar forecasts..*	*Q4 (6/30/19)*	*Q4 (6/30/19)*	*Met, Separate Milestone Report Submitted*
2.2	**Impacts & cost-benefit analysis in UC/ED**			
M2.2.1	*Impact/C-B analysis with IEEE 118 bus test system, with cost savings targets of >10-25% yearly cost savings based on simulations of flexible ramping product procurement.*	*Q4 (6/30/19)*	*Q4 (6/30/19)*	*Met, see Quarterly Report.*
M2.2.2	*Documentation of impact/C-B analysis using IEEE 118 bus test system*	*Q4 (6/30/19)*	*Q4 (6/30/19)*	*Met, see separate Milestone Report.*
3.1	**Functional & software requirements; tool development cycles**			
M3.1.1	*Prototype #1- RAVIS tool demonstration (to ISO and SETO staff) and functional specification with Version 0.0 Watt-Sun outputs. Tool demo will include solar and net-load ramp forecasts, along with their uncertainty levels*	*Q2 (12/31/18)*	*Q2 (12/31/18)*	*Met*
4.1	**Integration testing specifications from ISOs (See Milestone 8.1.1 below)**			
4.2	**Level 1-2 integration: cost-benefit analysis in UC,ED (See Milestone 8.2 below)**			
5.1	**Probabilistic Watt-Sun models for two regions**			
M5.1.1	*Develop Vers. 2.0 (short- and mid-term probabilistic solar power forecasts at 1-4 km resolutions). Version 2.0-Advanced Watt-Sun probabilistic solar power forecasts for 4 hour and 24 hour ahead for 6 points, shared with Topic Area 1 validation team. Target: Average improvement in point forecast accuracy (Brier Score) > 15% for 28 consecutive days for >6 points sites in the CAISO and MISO regions*	*Q7 (3/31/20)*	*Q7 (3/31/20)*	*Met: implemented 3 km resolution grid for CAISO region, with a 10 min resolution and 1 hr horizon. (Not shared with Topic Area 1 validation because team was not ready). See separate Milestone Report*
M5.1.2	*Provide accompanying report on PAIRS data used, multi-expert machine learning models and region-specific model blending performed report*	*Q7 (3/31/20)*	*Q7 (3/31/20)*	*Met*
5.2	**PAIRS platform advancement, forecasting error characterization**			
M5.2.1	*Identify data sources relevant for solar forecasting , docuument in website.*	*Q5 (9/30/19)*	*Q5 (9/30/19)*	*Met, see separate Milestone Report.*
M5.2.2	*Parallel computing to curate & harmonize all data with respect to common spatial & temporal grids. PAIRS platform advancement complete for continental US regions, with scalable data curation (rate >10 TB/day for at leastt 5 consecutive days) and multiple data sources integrated, verified via statistics published in IBM PAIRS website*	*Q5 (9/30/19)*	*Q5 (9/30/19)*	*Met, see separate Milestone Report.*
6.1	**Procedures to estimate FRP, regulation requirements using probabilistic forecasts**			
M6.1.1	*Ramp product requirements based on net-load forecasts from Watt-Sun V. 2.0.Target: Ramping requirements (the part caused by solar forecasting errors) with probabilistic solar forecasts will be 15% less those with deterministic solar forecasts*	*Q8 (6/30/20)*	*Q8 (6/30/20)*	*Met, see separate Milestone Report.*
7.1	**Functional & software requirements; tool development cycles**			
M7.1	*Prototype #2- RAVIS tool demonstration to ISO & SETO staff, and functional specification with Demo #2 (Version 1.0) Watt-Sun outputs based on staff feedback. Tool demo to include solar and net-load ramp forecasts, along with uncertainty levels . Obtain feedback from ISO operators on visualization requirements, integration specifications*	*Q6 (12/31/19)*	*Q6 (12/31/19)*	*Met, see separate Milestone Report.*
7.2	**Develop Visualization tool's alert & curtailment assessment capabilities**			
M7.2.1	*Develop RAVIS tool's ability to provide excessive down ramp alerts, to identify the least cost curtailment options. Qualitative comparisons of ramp visualization tool functionality based on Prototype #2 .with typical control center approaches,*	*Q7 (3/31/20)*	*Q7 (3/31/20)*	*Met, see separate Milestone Report.*
8.1	**Integration testing specifications from ISOs**			
M8.1.1	*Identify requirements for 1) forecast integration; 2) testing of improved ramping products. A report will detail the solar forecast formats, frequency of forecast updates for day-ahead and real-time integration, and reserve requirements formats*	*Q6 (12/31/19)*	*Q6 (12/31/19)*	*Met, see separate Milestone Report.*
M8.1.2	*Report assessing logistical feasiblity of integration with ISO energy management systems (dispatch and unit commitment).*	*Q6 (12/31/19)*	*Q6 (12/31/19)*	*Met, see separate Milestone Report.*

Table 1. Project Tasks, Milestone Completion Dates, and Milestone Status (Cont.)

Milestone Schedule		Milestone Completion Dates		Milestone/ Target Status
Task#/ Milestone#	Milestone Description and Target (if applicable)	Original Planned	Actual	
8.2	Level 1-2 integration: cost-benefit analysis in UC,ED for flexible ramping product procurement			
M8.2	*ISO system simulations and cost-benefits of forecast integration under all scenarios. Identify conditions which >10-25% yearly cost savings in flexible ramping product procurement can be ensured compared to the baseline*	Q8 (6/30/20)	Q9 (9/30/20)	*See separate Milestone8.2 Report. After COVID-related delay, Met as M12.2.0, below*
8.3	Level 2 integration: ISO development environment test (See Milestone 12.2.1)			
9.1	Probabilistic Watt-Sun models for two regions			
M9.1.1a	*Version 3.0-Advanced Watt-Sun probabilistic solar power forecasts for 4 hr and 24 hr ahead for 6 points, accuracy goals met. Target: Improvement in Point forecast accuracy > 20% compared to persistence baseline models as measured by P-P score (= the integral of the absolute value of the deviation from a 45 degree line of the plot of empirical distribution of solar GHI forecast errors versus their forecast cumulative probability)*	Q11 (3/31/21)	Q11 (3/31/21)	*Met, see separate Milestone Report.*
M9.1.1b	*Report on Version 3.0-Advanced Watt-Sun probabilistic solar power coordination efforts with TA1 team for forecast validation. Coordination efforts will include our sharing forecasts with the TA1 team via the PAIRS platform*	Q11 (3/31/21)	Q14 (9/30/21)	*Met*
M9.1.2	*A report (>5 pp.) providing guidelines for updating probabilistic Watt-Sun forecasting system based on new forecast data and updated forecasting error information*	Q9 (9/30/20)	Q9 (9/30/20)	*Met, see separate Milestone Report.*
10.1	Procedures to estimate FRP, regulation requirements using probabilistic forecasts			
M10.1.1	*Preparation of technical documentation of formulation and performance of probabilistic ramp forecasts, including journal article(s)*	Q11 (3/31/21)	Q11 (3/31/21)	*Met, see separate Milestone Report.*
11.1	Functional & software requirements; tool development cycles			
M11.1.1	*Final Version - RAVIS tool demonstration and functional specification with Version 3.0 Watt-Sun outputs. Demonstration will show capability of giving ramp alerts, regional ramps, and identify possible economic solar curtailment under ramp deficit conditions.*	Q12 (6/30/21)	Q12 (6/30/21)	*Met, see separate Milestone Report / RAVIS Technical Documentation*
12.1	Level 1-2 integration: impact & cost-benefit for regulation procurement in UC, ED			
M12.1	*Estimate regulation impacts, cost savings under baseline scenario using exogenous approach. Identify conditions which >10-25% yearly cost savings in regulation procurement can be ensured compared to the baseline.*	Q12 (6/30/21)	Q14 (12/31/21)	*Met, see Sect. 7.2.2. Day-ahead regulation method did not improve on ISO method; but real-time method lowered requirements & maintained reliability*
12.2	Level 1-2 integration: ISO development environment			
M12.2.0	*ISO integration testing of Watt-Sun probabilistic forecasts and revised FRP requirements, small system. For selected days from IBM's probabilistic forecasts, performing simulations using IEEE 118 bus test system (with CAISO generation mix), and estimating cost reduction*	Q8 (6/30/20)	Q9 (9/30/20)	*Met*
M12.2.1	*Simulation of large 1820-bus system and cost-benefit analysis for improved probabilistic forecasts based FRP procurments. First testing for a selected number of days and different times in the month, and observe the generation dispatches, clearing prices and system costs. This milestone will re-evaluate Milestone 12.2.0 using larger CAISO system. Target: 10-25% cost savings (due to production cost and reliability) using larger CAISO system*	Q10 (12/31/20)	Q12 (6/30/21)	*Met, see separate Milestone Report.*
M12.2.2	*A final report on feasibility and economic impacts (with a target of at least 12% savings in regulation and flexible ramping product procurement costs for high penetration solar systems) from testing the integration of probabilistic solar power forecasts for coordinated ramp product and dynamic regulation procurement under ISO development environment using CAISO scale system.*	Q12 (6/30/21)	Q14 (12/31/21)	*Met for FRP; no savings identified for solar-informed regulation*

7. Project Results and Discussion

Space limitations prevent us from present full descriptions of all results. The reader is referred to the 18 Milestone reports described in Table 1, above. The file name of each report is in the form "8215_QQQ_Johns Hopkins_MilestoneNNN_DATE," where QQQ is the project quarter, NNN is the milestone number, and DATE is the date of preparation.

7.1 Task 1: Developing Probabilistic Solar Forecasting using Watt-Sun/PAIRS

7.1.1. Year 1 Milestones: Probabilistic Watt-Sun 1.0. The original IBM Watt-Sun solar forecasting system is based on situation-dependent error analysis of multiple forecast models and subsequent multi-expert machine learning to blend such models to obtain the best possible forecast. This system was enhanced and integrated with a highly scalable big data platform, the IBM PAIRS Geoscope [37,38] (Fig. 8). PAIRS provides multiple benefits including its ability to scale, process literally unlimited data sizes and therefore to take advantage of truly big data to build more robust forecast models. PAIRS also allows disseminating the forecasts effectively, combining forecast data with other geo-spatial information and providing improved integration capabilities with other decision support systems such as the RAVIS (Resource Forecast and Ramp Visualization for Situational Awareness) visualization system discussed below.

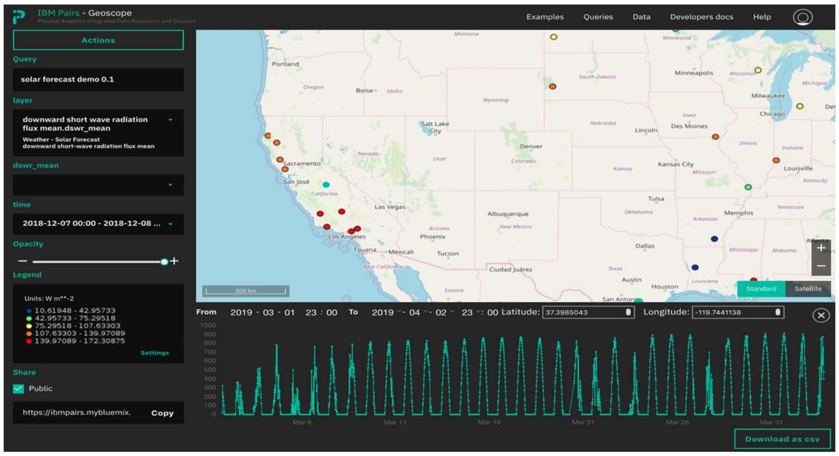


Fig. 8. The PAIRS platform-based Watt-Sun Solar Forecast

Subtask 1.1.1 provided, as a first demonstration, probabilistic forecasts (with a mean, 95% upper confidence interval, and 95% lower confidence interval), which were generated by this enhanced forecasting system (Version 0.0) for a single solar site in Topez, CA. It was shown that the Brier Score over three 28-day periods were below 0.4. The system provided the solar forecast at 15 minutes intervals with 4-hour forecast horizon and a refresh rate of every hour. This met **Milestone 1.1.1**. Along with the solar irradiance forecasts (i.e., Global Horizontal Irradiance (GHI), Direct Normal Irradiance (DNI), and Diffused Horizon Irradiance (DHI)), solar power forecasts were also provided, using the PV Lib [39] to convert the solar insolation to solar power production.

In **Milestone 1.1.2,** we reported on our techniques of model blending for solar forecasting and the data sources used for it. In this approach, the forecast errors of different numerical weather prediction models (NWPs) are analyzed and used to characterize weather situations that will be treated differently (i.e., trained upon individually). At the training stage, the multiple NWPs and measurement data are preprocessed to filter outliers, and the functional analysis of variance fANOVA method [40] is used for feature selection by using an empirical performance model (EPM). The EPM is based on random forests to analyze how much of the performance variance in the configuration space is explained by single parameters or combinations of few parameters. Then, forecast errors are predicted using a Random Forecast model, which are then used in the situation categorization. In the next step, for each situation in each category, we train a quantile regression model. For forecasting, first, forecast errors generated from a Random Forest model are used to categorize the situations, and then for each category, the corresponding trained quantile regression model is used to forecast GHI. Fig. 9 below depicts the approach.

The basic data sources required are outputs from NWPs and measurements. We used the PAIRS platform for these data sets and to be able to exploit very large data sets. The training and forecasting data needed for the machine learning were realized by querying PAIRS in a consistent and scalable manner. PAIRS takes care of the different spatial and temporal resolutions [35,36] of the input data. PAIRS has over 1000 different NWPs data sets and layers. For V. 1.0 we used the sub-hour High Resolution Rapid Refresh (HRRR) model [41], the North American Mesoscale Forecast System (NAM) [42], the US Climate Reference Network [43], and other data where noted.

A report on model blending and the data sources concluded **Milestone 1.1.2.**

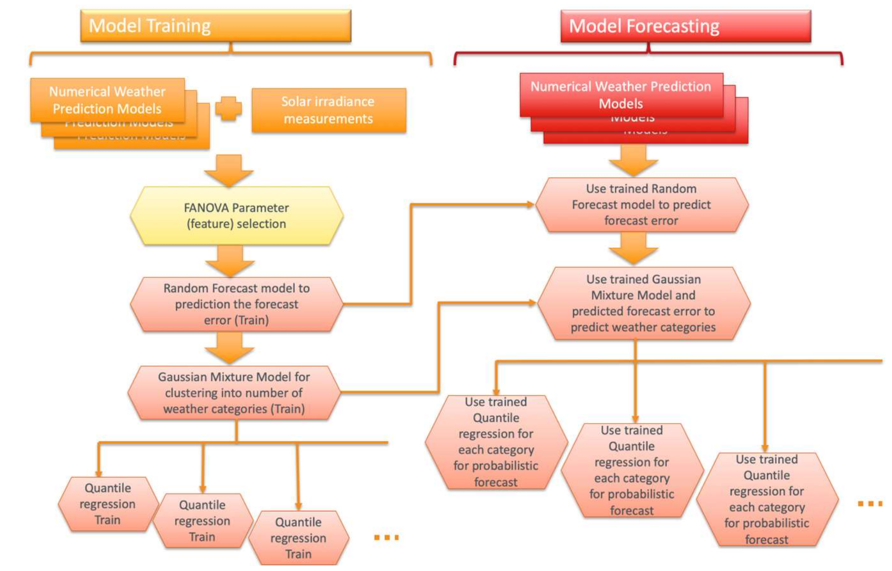


Fig. 9: Model blending Flow chart Overview

Milestone 1.1.3 concerned documentation and testing of an enhancement of the initial prototype of the PAIRS- integrated version of Watt-Sun, yielding Watt-Sun 1.0. We designed the original Watt-Sun system for deterministic solar forecasting instead of probabilistic. So we changed the last step of the situation-dependent error analysis to linear quantile regression because the forecast distribution is not normally distributed and therefore will be better captured by quantile regression. This contributed to improved probabilistic solar power forecasts. In addition, an Autoregressive Integrated Moving Average (ARIMA)-based smart persistence model was created for benchmark purposes. Such model is state-of-the-art at the ISOs and therefore useful to compare the Watt-Sun forecasts against it.

This second demonstration included probabilistic solar forecasts for 20 sites (10 sites in CAISO and 10 sites in MISO regions). Forecasts were operational with 5 quantiles (0.05, 0.25, 0.5, 0.75, 0.95), an extended forecast horizon of 24 hours at 15 min frequency and refresh rate. We evaluated these forecasts initially over two periods of 28 consecutive days. All Brier scores for each site were below 0.3, which is an improvement of 25% over Watt-Sun 0.0. Compared to the ISO standard ARIMA method 25% improvements for the 10 CAISO sites and 21% for the 10 MISO sites were achieved. These performances exceeded the Milestone 1.1.3 targets. Thus, **Milestone 1.1.3** was satisfied by submitting a report on Demo #2 (Version 1.0) of Watt-Sun probabilistic forecasts (Q3), and by Version 1.0 performing at least 10% better than a standard practice persistence-based method.

To meet Milestone **1.2.1,** a data management plan was developed and submitted. For all data management the PAIRS platform was used. This platform is maintained by IBM and has been growing by more 10 Terabytes/day since 2016. PAIRS standardizes the spatial and temporal resolutions of the ingested data [37,38] by employing a global reference systems of layered resolution layers, thereby linking data in space and time. The platform has both the input data for the model blending as well as the generated forecast data.

The platform does not only enable easy integration via OGC (open geographic consortium) compliant services (Web mapping services, Web processing services, Web coverage services, Web feature services) via an open source geoserver [44] with other applications but it also has also a very powerful, open-sourced SDK/API for querying, filtering, or more complex computational tasks without downloading the raw data. PAIRS is available through an academic license to everyone for research purpose.

7.1.2. Year 2 Milestones: Probabilistic Watt-Sun Version 2.0. Next, we developed version 2.0 of Watt-Sun probabilistic forecasts systems as part of the effort towards **Milestone 5.1.1**, which embodied several improvements. We concluded that the Brier score was less useful for assessing quality of probabilistic forecasts compared to the P-P-plot based metric, which we therefore used for forecast performance evaluations along with the Mean Absolute Percentage error. Watt-Sun 2.0 leverages level 2 (L2) data from CLAVR-x (Clouds from AVHRR (Advanced Very High-Resolution Radiometer) Extended System) [45], which provides real-time cloud information from a Geostationary Environmental Satellite (GOES-16) [46]. GOES-16 data can enable better solar forecasts for short-term forecast. A data pipeline was developed pushing GOES-16 data into PAIRS operationally. GOES-16 also opens the opportunity to provide gridded forecasts.

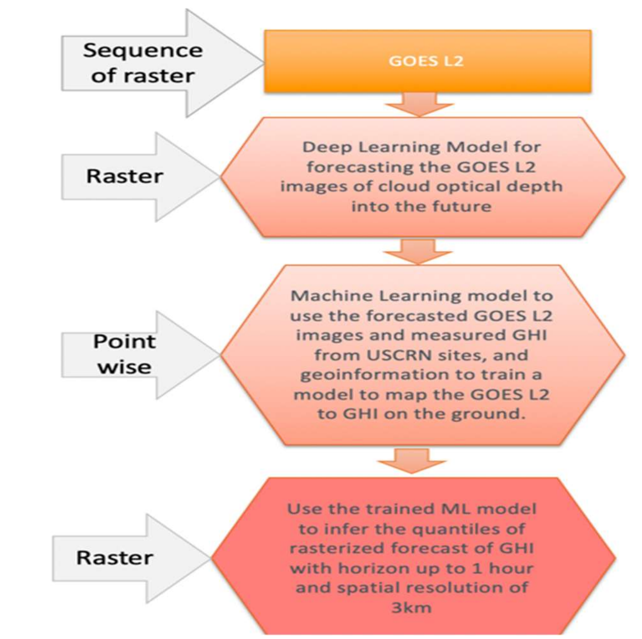


Fig. 10: Flowchart for our short-term solar irradiance forecast approach

Fig. 10 shows our approach for a measurement informed, gridded short-term solar irradiance forecasts with (i) a deep learning model to provide gridded forecasts from the GOES L2 data of cloud optical depth for a 1-hour horizon, (ii) a quantile regression-based ML model to map the forecasts to the measured ground-level GHI and (iii) trained ML model to forecast every grid point pixel with a 3km resolution of the defined region to generate the measurement-informed rasterized forecast.

We experimented with networks which are similar to the ImageNet challenge winning models AlexNet [47]. However, instead of a spatial convolution layer, we used the volumetric convolution neural network (VCNN) layer, which applies a 4D convolution over a 4D input tensor composed of several input planes (here, multiple timestamps and multiple bands from the GOES L2 product) (see Fig. 11). By using VCNN, the extracted features have access to multiple channels of weather information across both temporal and spatial dimensions. Therefore, the VCNN can learn the cloud movement patterns which is important for solar irradiance prediction.

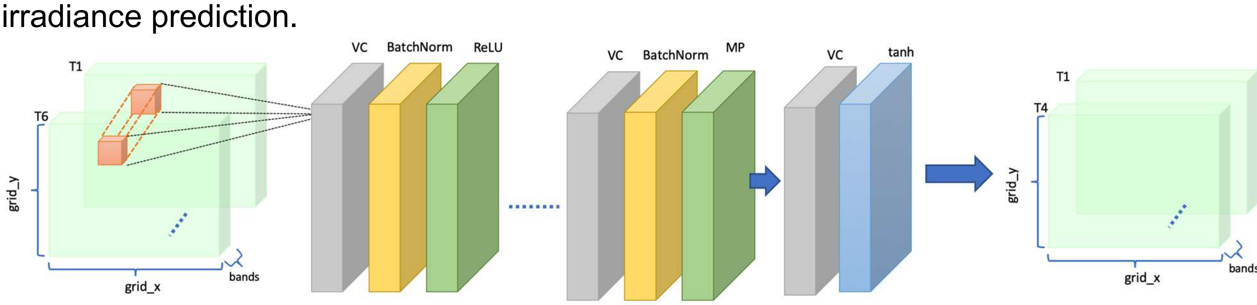


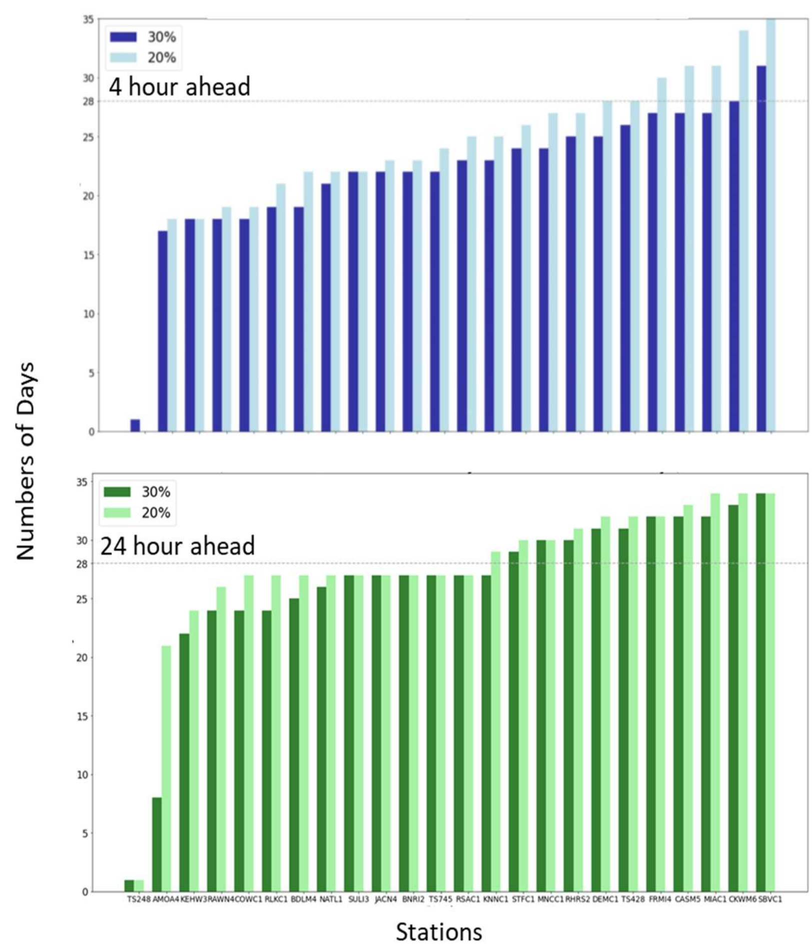
Fig. 11. Schematic diagram of a raster forecast Deep Neural Network (VCNN)

Milestone 5.1.1 includes development of Version 2.0-Advanced Watt-Sun probabilistic solar power forecasts for 4-hour and 24-hour ahead for 6 points, which are planned to be shared with Topic Area 1 validation team. The forecasts were to include an improvement in Point forecast accuracy > 15% compared to persistence baseline models as measured by Brier Score, especially for CAISO and MISO regions. Watt-Sun 2.0 performs better than the previous versions with an improvement in the P-P-plot score over bias-corrected HRRR GHI forecasts by 20%. Thus, this exceeded the target set by **Milestone 5.1.1**. In terms of rMAPE, Watt-Sun 2.0 performs better as well, with 17% improvement over HRRR bias-corrected forecast. In Watt-Sun 2.0 we also introduced an in-depth forecast calibration index. It enables evaluating the quality of calibration that allows for exploration of how well a probabilistic forecasting model performs over different intervals of the day, and over different months. A report on PAIRS data used, multi-expert machine learning models, and region-specific model blending performed, was generated and accepted by DoE, which satisfied **Milestone 5.1.2**

Milestone 5.2.1 addressed documentation of data sources for PAIRS updating and web-link documentation. There were three major enhancements to the PAIRS data. First, we significantly enhanced the data catalog in PAIRS, which is now completely standardized across different data sources and can be queried with elastic search. Second, in addition to the GOES-16 data, data pipelines for weather station data from the NOAA-ISD (Integrated Surface Database) [48] and RAWS [49] were developed. This data is now being curated, ingested and is available in PAIRS. Third, additional capabilities for batch data exports were added, especially for deep learning applications by implementing a SPARK connector on top the PAIRS HBASE data source. This was reported along with the list of available data in PAIRS, which was accepted by DoE. This concluded **Milestone 5.2.1**

Milestone 5.2.2 addressed documentation of US coverage of PAIRS, curation rate, and integration of multiple data sources. While the PAIRS catalog provides detailed documentation of data across space and time, fast sampling techniques are also available in the PAIRS systems to check interactively what regions, timestamps are available. This is documented in the PAIRS tutorials available on-line (https://pairs.res.ibm.com/tutorial/). In addition, in this subtask an improved real-time data curation monitoring system was developed and demonstrated. It is noted that PAIRS is a real-time system with data being updated and curated at a rate of 10 Terabytes and more each day. A report was written and submitting about the PAIRS data us coverage, curation rate and integration of multiple data sources, accomplishing **Milestone 5.2.2.**

7.1.3. Year 3 Milestones: Probabilistic Watt-Sun Version 3.0. **Milestone 9.1.1a** addressed development of Version 3.0-Advanced Watt-Sun probabilistic solar power and its forecast accuracy. The success value was a targeted point forecast accuracy improvement of 20% relative to the persistence baseline. The third version of the Watt-Sun forecast system was mainly enhanced by leveraging more training data. Initially, training data from HRRR was limited to a few months, while Watt-Sun 3.0 was able to take full advantage of almost three 3 years of historical data. The goal for Watt-Sun 3.0 was to achieve significantly improved calibration as measured by the P-P-plot metric as compared to the persistence baseline estimator, as shown in Figs. 12 and 13. We demonstrated that out of the 24 measurement stations for at least 6 of them and for at least 28 consecutive days across a period of 3 months, a relative improvement of 20% was achieved for forecast horizons of 4 and 24 hrs (Figs. 12, 13, respectively). This met the **Milestone 9.1.1a** target. Further, as a measure of the reliability of these improvements for 4-hr ahead forecasts, daily values of the P-P metric of forecast improvement were better by >30% for ≥ 3 weeks straight during that period for 17 of 24 CAISO & MISE sites tested, as shown in Fig.12.



Figs. 12,13: Forecasting performance: numbers of consecutive days of >20% & >30% improved P-P-plot performance at 24 locations for 4 & 24 hrs, respectively.

The Milestone 9.1.1b report documented V. 3.0-Advanced Watt-Sun probabilistic solar power coordination efforts with TA1 team for forecast validation. Coordination efforts included our sharing forecasts with the TA1 team via the PAIRS platform. To accommodate the validation efforts of the TA1 teams, changes to the Watt-Sun 3.0 systems were required, namely increasing the numbers of percentiles, retraining the forecasts for different locations, and changing the forecast frequencies. These activities were reported to the DoE, meeting **Milestone 9.1.1b.** The subsequent **Milestone 9.1.2** report provided guidelines for updating the probabilistic Watt-Sun forecasting system based on the most recent data.

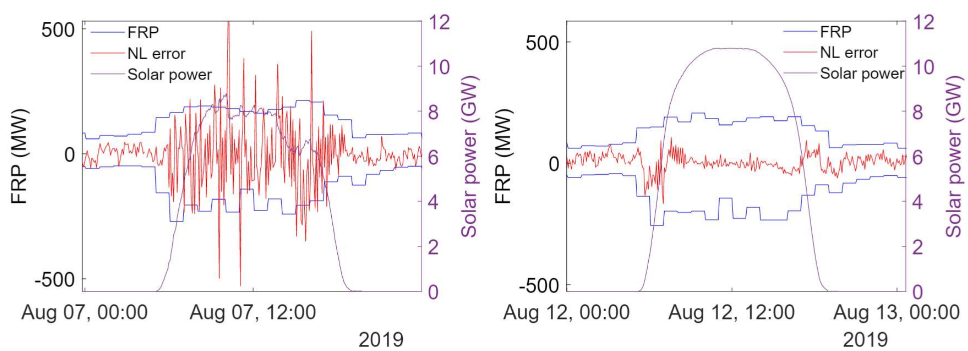
The final Task 1 (solar forecasting) effort concerned coordination with ARBITER. A real-time upload from Watt-Sun 3.0/PAIRS to the ARBITER platform was developed and realized. Except for scheduled PAIRS maintenance, data from the Watt-Sun 3.0 forecasting system was shared with the TA1 teams in real-time. For this we improved the reliability of the exogenous data pipeline (namely, NOAA's sub-hourly HRRR data ingestion into PAIRS Geoscope). Further, the software implementation of Watt-Sun 3.0 was implemented according to the object-oriented programming paradigm in which the real-time forecast ingestion functionality is realized through a set of class methods. This allows Watt-Sun 3.0 to be dockerized and run in IBM's cloud solution (the so-called Cloud Object Storage), which improves reliability and speed of the forecast availability for third parties.

7.2 Task 2 Results: Defining Reserve Needs with Probabilistic Solar Forecasts

This task addresses development, application, and testing of data-driven methods to improve estimates of FRP requirements by using probabilistic solar forecasts. Two approaches are used: machine learning and quantile regression (Sects. 6.2.1 and 6.2.2, respectively). In addition, we investigated use of those methods to improve projections of need for up-regulation (Sect. 6.2.3). Finally, we also investigated explicit convolution of net load components (gross load, wind, solar) to create requirements, but results were not immediately promising [15]. *Note:* All MW quantities here are transformed values to disguise the true values, consistent with non-disclosure agreements.

7.2.1 Using Machine Learning to Incorporate Solar Uncertainty into FRP Requirements. We developed a variety of numerical classifiers as labels of weather conditions based on the predicted uncertainty, variability, and cloudiness index of solar irradiance and power. In addition, we adopted principal components analysis (PCA) to reduce the dimension of multi-dimensional classifiers, which better reflects system-level uncertainties by including classifiers from multiple sites across CAISO. We applied a kNN-based method to identify historical days of similar weather condition, and used the realized FRP requirements from these days to construct predictive distributions of the FRP needs, which can be used to give weather-informed estimations of FRP requirements.

Fig. 14 compares the realized net load forecast errors with the published FRP requirements from 2 days in August 2019 in CAISO. Because the 2 days are close in time (5 days apart), the FRP requirements differ by less than 1% because of similar histograms in use, whereas the solar power profiles imply drastically different weather conditions, which potentially explain the greater uncertainty needs in the cloudy day than the sunny day. Therefore, if similar FRP amounts are procured, the system could experience a shortage of FRP in the cloudy day. This observation motivates the need for the latest probabilistic forecasts in estimating FRP requirements.



(a) A cloudy day. (b) A sunny day.

Fig. 14. Comparison of realized net load forecast errors with published FRP requirements for the CAISO.

The kNN-based method can be viewed as a direct extension of CAISO's original implementation (which chooses requirements based on the 2.5th and 97.5th percentiles of a histogram of ramp forecast errors from the previous several weeks) since both methods rely on historical data. However, in contrast to that CAISO baseline method, whose requirements are not conditioned on weather-of-the-day, the kNN-based method constructs weather-conditioned histograms by using probabilistic solar forecasts. Different numerical classifiers in total are used to feature the weather data, including solar irradiance, predictive intervals, cloudiness level, and clear-sky power index. In addition, in order to correctly characterize the solar power uncertainty over the entire CAISO region, multi-dimensional classifiers based on multiple sites at different locations are considered, and principal components analysis (PCA) is leveraged to reduce the dimension of multi-dimensional classifiers. We evaluate the FRP requirements from two perspectives: system reliability and market efficiency. Two metrics are investigated, including the frequency of FRP shortage and the amount of FRP oversupply.

As Fig. 15 shows, all 12 classifiers present similar trends as K (numbers of days selected for training) increases, i.e., the FRP oversupply increases sharply when K ≤ 20, which in turn results in a considerable drop of the frequency of the FRP shortage. When K > 20, however, both metrics remain relatively constant. This phenomenon suggests that a sufficient number of days (K > 20) are required to give a reliable estimation of FRP requirements across all classifiers. Note that in the baseline, the frequency of the FRP shortage increases slightly as K increases from 20 to 60, implying that 20 days could yield better performance than CAISO's business-as-usual (BAU) implementation, where K=30.

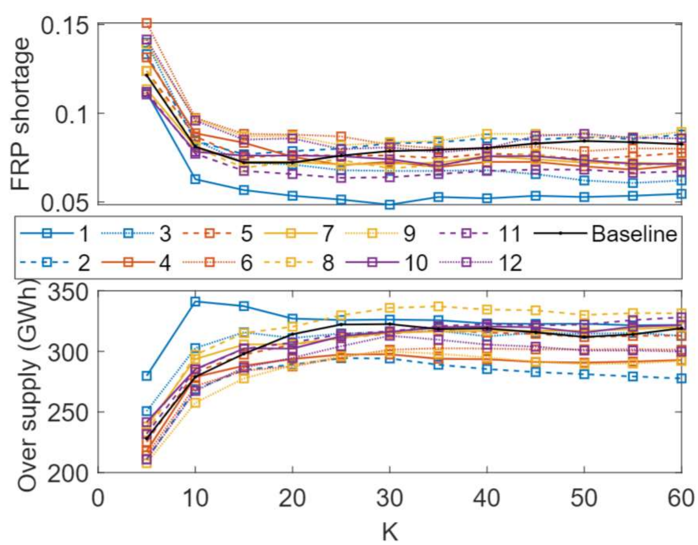


Fig. 15. The frequency of FRP shortage and FRP oversupply as a function of K in Feb. 2020.

Fig. 16 shows trade-offs between reliability and oversupply in the form of Pareto frontiers. The point at the intersection of two dashed lines represents CAISO's BAU implementation. The two dashed lines divide the plane into four quadrants (I, II, III, and IV), where points in Quadrant III indicate an improvement in both dimensions relative to the baseline, points in Quadrant I indicate a degradation in both metrics, and Quadrants II and IV represent tradeoffs improvement in one objective requires a degradation of the other. Note that a significant fraction of kNN points fall into Quadrant III and no point falls within quadrant I, suggesting that the kNN-based method can result in more economic and reliable solutions than the baseline. Besides, Fig. 16 shows the results when the kNN parameters are dynamically selected. A similar trade-off between reliability and oversupply is also observed when the size of the validation set (N) changes.

Fig. 17 shows the optimized Pareto frontiers of all 1-site cases and the multi-site case using PCA-kNN from February, August, and October 2020. All optimal Pareto frontiers present similar trends as in Fig. 16 when N varies. Although most 1-site cases present better performance than the baseline, their performance varies because of geographic differences. The variation of performance across sites implies the challenge of using one single site to characterize the weather condition of the whole CAISO region. By comparison, the optimized PCA-kNN frontier accounts for the weather conditions from all 5 sites, and results in less variation of performance. The optimized PCA-kNN frontier presents better performance in terms of both dimensions compared to the baseline and all 1-site cases in February and August.

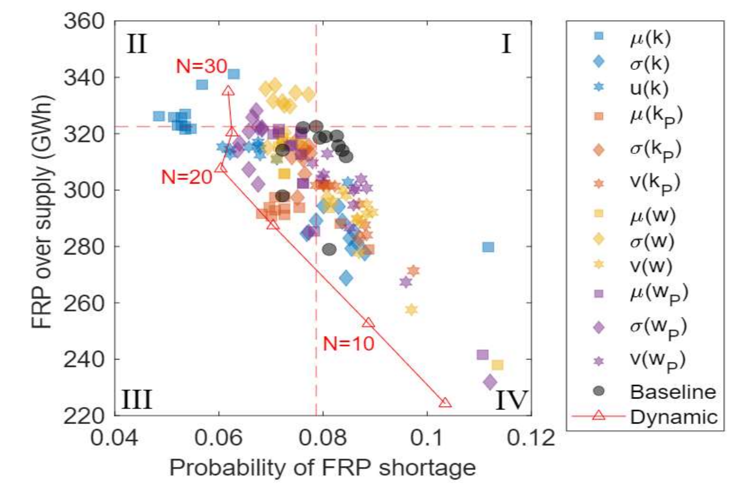


Fig. 16. Trade-offs between reliability and over-supply in February 2020. (One site)

To compare the results from the kNN-based methods with other data-driven methods, the evaluation metrics from the benchmark machine/deep learning (ML/DL) based models are calculated. Note that these models do not require the use of historical data to construct predictive distributions of FRP, therefore they are not affected by the number of neighbors K. As displayed in Fig. 17(d)-(f), the black dots that represent the original benchmark methods are all concentrated in Quadrant IV, indicating reduced oversupplies yet lower reliability

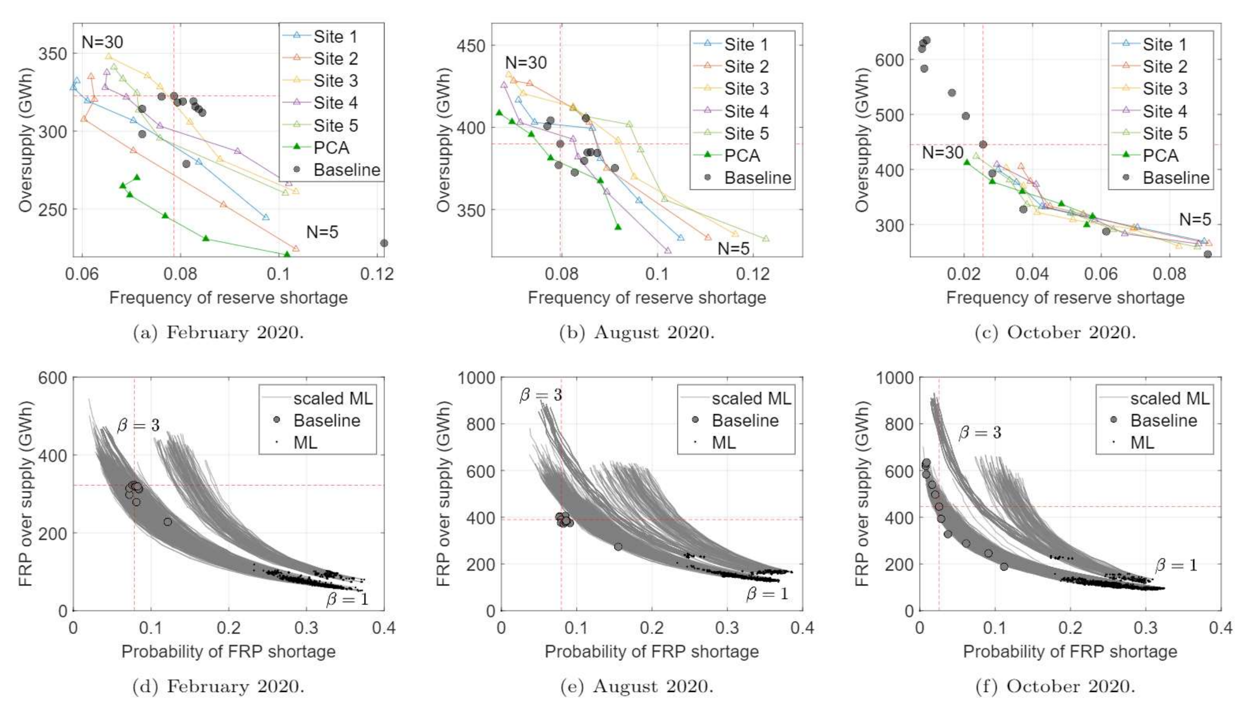


Fig. 17. (a)--(c) Pareto frontiers of all 1-site cases and the multi-site PCA-kNN case. (d)--(f) Frontiers from the ML/DL-based methods. The black dots represent the original results from ML/DL-based benchmarks---i.e., β=1---and the gray curves represent the Pareto frontiers when the factor β ranges from 1 to 3.

levels. The significantly greater chances of FRP shortage of the benchmark methods result in an unfair comparison, therefore, to make a comparison based on a more level playing field, we adjust the FRP requirements from the benchmark methods by multiplying them by a factor β, where β =1 indicates the original results. Owing to the trade-off between the oversupply and the chance of FRP shortage, a greater β results in more conservative FRP requirements. As Fig. 17 shows, the profiles of the adjusted benchmark results resemble the previous kNN profiles. Although several curves fall into Quadrant I, many of the adjusted benchmark results fall into Quadrant III, indicating potential improvement in both dimensions.

7.2.2 Machine-Learning Based Definition of Up-Regulation Requirements Using Solar Forecasts. In addition to FRP requirements, we also developed solar-informed regulation requirements determination methods. Although our attempt to improve **day-ahead** regulation requirements using solar-informed methods was unsuccessful in reducing those requirements and improving reliability, we did successfully reduce regulation-up requirements without compromising reliability using in **real-time**, as we now describe.

The historical ACE* signal in May 2020 from CAISO, which originally is at 1-min resolution, has been aggregated into a 5-min resolution (i.e., choosing maximum value in each 5 minutes) to be consistent with the OASIS solar MW production data [60] for analysis and the CAISO's present procedures for assessing the need for regulation. The regulation procurement baseline data is also selected from the same month. All the data are transformed consistent with non-disclosure agreements. This data, together with solar forecast data, is used to calibrate models that estimate ACE* two hours ahead of time, which can then be used to determine a real-time regulation amount that reflects information available at that time, with the goal of improving performance (maximize reliability and minimize procurement) relative to the present ISO system in which regulation requirements are determined day ahead based on an (unconditional) histogram of historical ACE* amounts over the previous weeks.

The hypothesis is that we can utilize historical ACE* and solar forecasts to forecast the future ACE* on a near real-time basis, then use the forecasted ACE* signal to determine regulation procurement, and that this procurement will perform better than the present ISO day-ahead method that doesn't explicitly consider weather. This forecast is done on a rolling basis with ACE* data for several weeks prior. We propose a simple approach of defining the regulation requirement as:

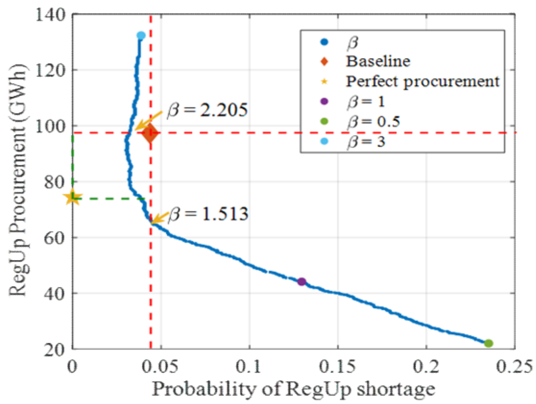
$$\mathrm{MAX}(R_{min},\ \beta\times\text{forecast ACE*})+z(t)\times R_{extra}$$

where R_{min} is a lower bound to the requirement, forecast ACE* is the two-hour ahead forecast area control error adjusted for the amount of regulation actually dispatched [8], β is a multiplier, and R_{extra} is an increase in procurement for t in day-time hours (6:00 to 20:00, when z(t) =1; otherwise z(t) =0). $R_{min,}$ β, and R_{extra} are tuned parameters to maximize performance. Three adaptive procurement strategies are examined:

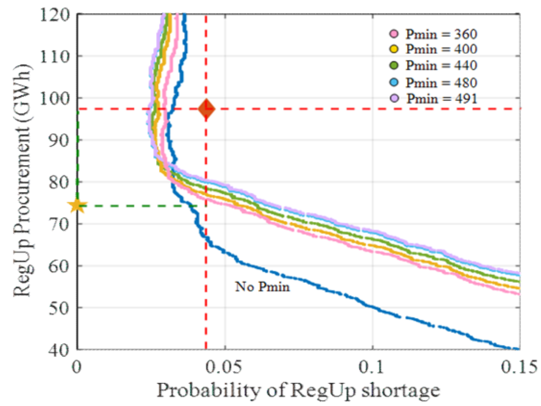
1. Multiplying the forecast ACE* value by a factor β ranging from 1 to 3.
2. Reducing the minimal regulation up procurement to R_{min} (lower bound for procurement), so the final requirement is MAX(R_{min}, $\beta\times$forecast ACE*). We note that the transformed minimal procurement for CAISO's baseline is 491 MW in each hour) in less fluctuating periods (i.e., 8 p.m. to 6 a.m.).

3. Adding an extra regulation up procurement R_{extra} to the forecasted ACE* value in more fluctuating periods, i.e., MAX(R_{min}, β×forecast ACE*)+R_{extra}*z(t).

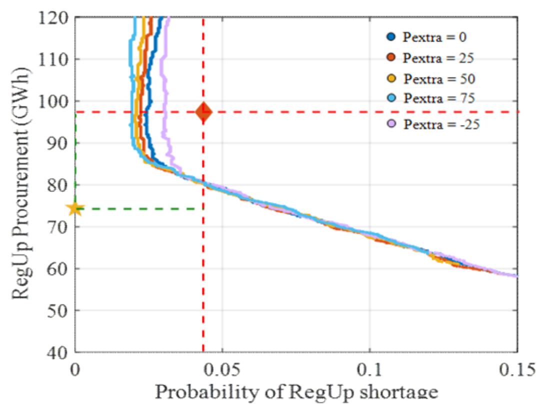
Figs. 18(a-c) show the Pareto analysis of regulation-up procurement for the last 7 days in May 2020. The figures show tradeoffs between reliability (probability of 5-min adjusted ACE* exceeding the requirement) and aggregated GWh supply over 7 days. The red diamond represents the CAISO baseline point based on reported OASIS regulation amounts, and the yellow star represents the ideal regulation procurement using perfect ACE* forecasts, which has 0 exceedance probability. Key insights include:



(a) sensitivity analysis of β



(b) sensitivity analysis of R_{min}



(c) sensitivity analysis of R_{extra} z

Fig. 18. (a)--(c) Pareto frontiers of regulation up procurement on the last 7 days of May 2020. (a) Multiplying factor β ranging from 0.5 to 3, R_{min} = 0. (b) Minimal procurement R_{min} for hours 8 p.m. to 6 a.m. (c) Extra procurement R_{extra} for hours 6 a.m. to 8 p.m

1. As observed from Fig. 18(a), a similar observation of β with the FRP requirement in Fig. 17(d-f) is obtained (i.e., a greater *β* results in more conservative regulation requirements), and no points fall into Quadrant I (where both reliability and GWh procured are worse). When β < 1.513, the curves fall into Quadrant IV, indicating a higher probability of regulation shortage than the baseline CAISO method, but less total GWh of procurement. A greater β results in more conservative regulation requirements, and the curves when β > 2.205 fall into Quadrant II, where procurement is greater than the CAISO baseline and the reliability is better.

2. As Ref. [8] explains, the regulation shortage always occurs during the sunshine periods. To evaluate the performance of adaptive procurement in different periods, we divide a day based on two time points: 6 a.m. and 8 p.m., and keep the β ranging from 0.5 to 3. As shown in Fig. 18(b), the regulation requirement could be improved by reducing the procurement in the less fluctuating periods. The curve of R_{min}=491 MW refers to the transformed CAISO baseline, where the regulation requirement is MAX(491MW, β×forecast ACE*), while the R_{min}=480 MW case (i.e., requirement is MAX(480MW, β×forecast ACE*) shows a very close proximity in Quadrants III and IV, indicating that slightly reduced procurement will not harm the reliability too much.
3. The results in Fig. 18(c) also shows that adding an extra procurement to the forecasted value during the 6 a.m. to 8 p.m. period also helps in promoting the reliability, however, with a higher amount of over-supply.

7.2.3 *Quantile Regression-based Estimation of Flexible Ramp Requirements.* Because the CAISO is already using quantile regression (QR) to assess flexiramp needs as a function of load and variable renewables, quantile regression has the advantage of familiarity as well as simplicity relative to the methods of Sect. 7.2.1. Therefore, we have also evaluated the potential of that method to improve requirement estimation (in terms of higher reliability and lower procurement amounts) relative to present practices. A variety of specifications were tested that included as independent variables various combinations of the following: deterministic load, wind, and solar forecasts as well as indicates of solar uncertainty (prediction interval widths for 2 hour ahead probabilistic forecasts, especially the difference between the 75th and 25th percentiles). As explained at the end of this section, in the course of this analysis, we developed an approach to assess and debias Watt-Sun forecasts in order to improve their calibration, which resulted in better predictions of FRP requirements for some months

The steps involved in creating QR-based flexiramp requirements were as follows. First, two separate QR estimations are performed for the 50th and 90th percentiles of the forecast error as a linear function of a set of independent variables related to weather and system conditions. (Four sets of such variables are considered here, as described in the Fig. 19's caption.) Second, the value of error for the desired reliability (say the 97.5th percentile, which would result in a 2.5% shortage rate) is obtained by fitting a normal distribution to the 50th and 90th percentiles and extrapolating. Out-of-sample validation found that this resulted in more stable estimates of extreme percentiles rather than using QR directly to estimate that percentile, due to small sample issues with the number of observations in the tail. This is done for using 30 days of data prior to the day of interest; the desired percentile is then estimated for that day given the value of the independent variables on that day. The FRP requirement is set equal to that value. The performance of the method is then assessed by comparing the realized forecast error against the requirement. This is repeated for each of the days in the month (March 2020 in Fig. 19), and four 15-minute intervals within each hour considered; this would give 30*4=120 observations to estimate the reliability and cost performance.

In Fig. 19, we did this for four levels of *a priori* reliability (10%, 5%, 3%, and 1.5% shortage probabilities) for each of four model specifications for the noon hour in March 2020. The best specification was one based on two independent variables: the average (across four sites) of the 25th-75th percentile prediction interval for GHI; and a nonlinear (sine wave) transformation of median GHI, again averaged over four sites. The transformation yields values of zero if GHI is at the minimum or maximum of the GHI observed over the last 30 days at that hour of day, and attains a maximum if GHI is halfway between those extremes; this reflects the fact that if solar is zero or if there is a clear sky, there is relatively less uncertainty than if GHI is somewhere between the extremes. As Fig. 19 shows, there is one version of that model (*a priori* reliability of 5%) that reduces oversupply by 20% (x-axis) and cuts the *ex post* frequency of FRP shortage by about half (from 7.5% to 4%, y-axis), relative the actual amount of FRP that the ISO procured for those intervals. Although that precise specification does not always result in improvements in each month and time interval we considered, it often did so, and therefore is worth considering as a relatively simple but effective way to make FRP requirements weather conditioned.

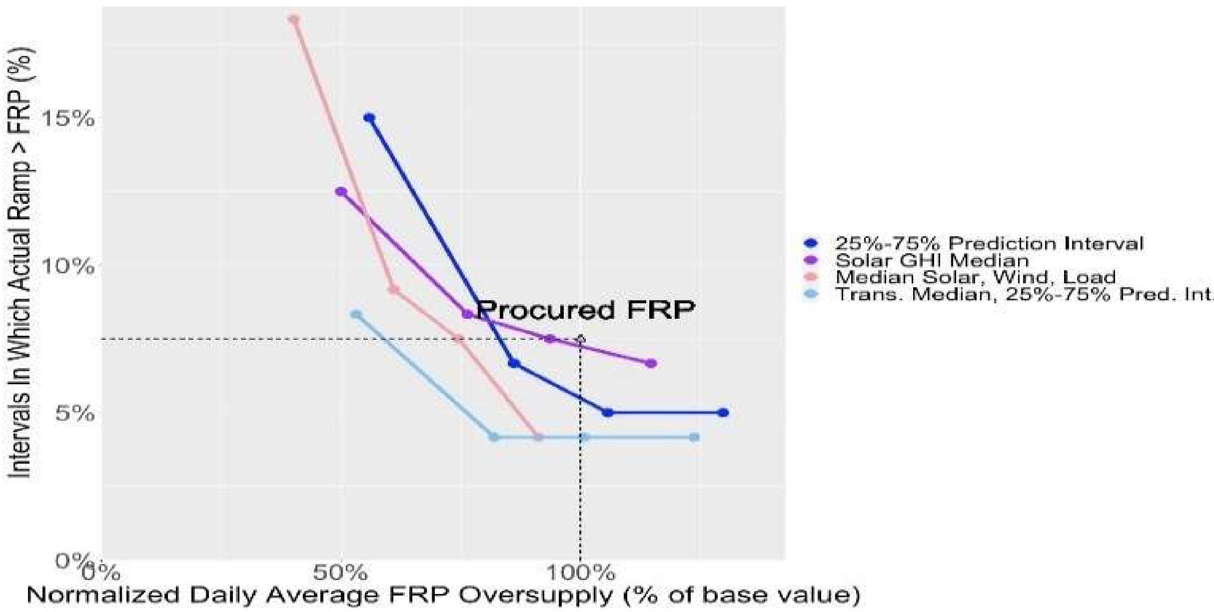


Fig. 19. Pareto plot (reliability (FRP shortage frequency) vs cost (excess FRP)) showing performance of four QR specifications (for 12:00-13:00 local time, March 2020), compared to performance of CAISO-procured FRP. Specifications include: linear using width of 25th-75th prediction interval for solar GHI (blue); linear using median solar GHI (purple); linear using median GHI plus CAISO wind and gross load forecast (pink); and linear with GHI interval and sine transformation of median GHI (light blue). Each set of solar variables is averaged across four CAISO solar sites. The points on each curve (upper to lower) correspond to 10%, 5%, 3%, and 1.5% a priori frequencies of FRP shortage.

Preliminary analyses of calibration of Watt-Sun forecasts indicated that cloudy and sunny days had different quality of calibration, so we developed a method for classifying day types and then adjusting forecasts that, based on out-of-sample tests, improved forecasts. We used χ^2 (chi-squared) tests in which the expected frequencies of observations within each of the 6 bins (0-0.05, 0.05-0.25, 0.25-0.5, 0.5-0.75, 0.75-0.95, and 0.05-1.0) are compared to observed frequencies. Fig. 20(a) shows actual GHI and Watt-Sun's reported probability distributions for the fifteen-minute interval centered on 1:30 p.m. local time for Dec. 2019 for the Topaz, CA site. There are 28 days of data, of which 9 days have values falling above the 50th percentile. Whether this could happen by chance can be assessed by a, e.g., χ^2 test; if the test does not reject the hypothesis that the observations were drawn from the shown distribution, then it would be concluded that the model is well-calibrated for that period. (Such a test for a set of daily observations for one particular time is reasonable if it is assumed that errors from day to day are independent, which is admittedly a strong assumption.) Plots like Fig. 20(b) below, which is an example binning of the observed GHI for all daylight hours in one month, provide visual evidence of a good calibration. Analyses of results of the last version of Watt-Sun indicated that the calibration differences between cloudy and sunny days had become less pronounced, and adjustment of forecasts was no longer required.

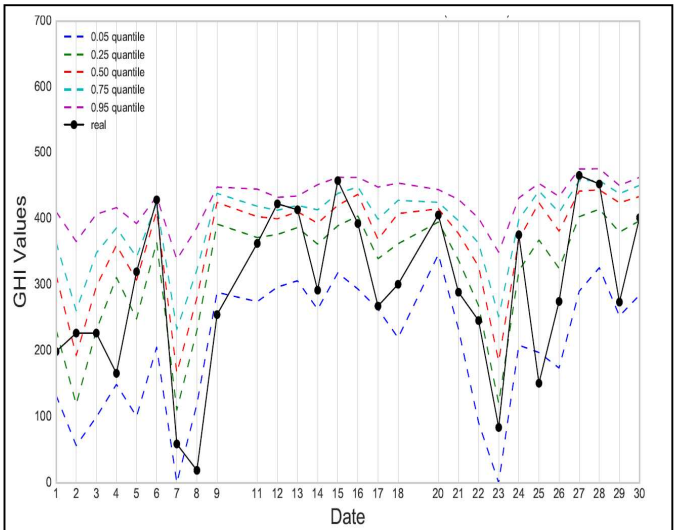


Fig.20(a) Data for χ^2 calibration test. Dec. 2019 actual GHI and Watt-Sun probabilistic forecast quantiles, 1:30 local time, Topaz site (Note, Dec. 10, 19, 31 values missing)

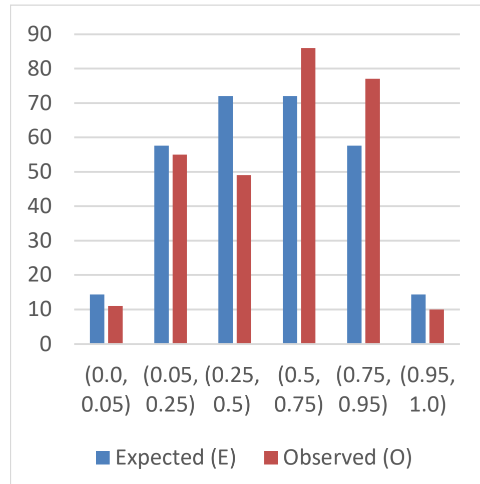


Fig. 20(b). One month's expected number of expected GHI observations by bin (fractile ranges) versus observed (daylight hours only), illustrating quality of calibration

In sum, for most intervals tested between March and Aug. 2020 (10-11 am, 12-1 pm, 2-3 pm, 4-5 pm), QR formulations were found that improved upon the CAISO base method and to actually procured FRP (from OASIS). In our application of the base method, a histogram was created of the last 30 days of ramp forecast errors for the relevant interval; the 2.5th and 97.5th percentiles then defined the down and up FRP uncertainty components, respectively, which when added to the forecast ramp yielded the requirements.

7.3. Results: Ramp and Solar Uncertainty Visualization for Situational Awareness

7.3.1 Overview. This section describes a flexible, open-source visualization tool for situational awareness related to operating a power system with high shares of variable renewables. The tool is named Resource Forecast and Ramp Visualization for Situational Awareness (RAVIS), and it is built at NREL under DOE EERE SETO funding. RAVIS is a research-grade tool intended to help researchers, forecast vendors, and system operators (such as utilities, ISOs, and balancing authorities) who use variable renewable forecasts for efficient operation of their systems. This tool provides a way to integrate advanced forecasts for variable renewable generation, including probabilistic forecasts, and helps operational control centers and forecasting teams at utilities to develop situational awareness and timely mitigation strategies. The tool is flexible enough for end users to tailor the data integration, visualization, and alerts to their needs and use cases.

As mentioned in Section 5.3, although control room visualizations are relatively mature, much development is needed to prepare for a future with high shares of variable renewable generation and for integrating advanced probabilistic forecasts, including ingesting site-specific distributed forecasts in addition to the conventional regionally aggregated forecasts. Additionally, RAVIS provides researchers around the world with an open-source alternative to demonstrate their developments. The source codes of RAVIS are publicly available for anyone to download and use at https://github.com/ravis-nrel/ravis. The link also provides guidelines and instructions to install and modify the code, including answers to some frequently asked questions.

We now describe the function and architecture of RAViS in detail. As a prototype of the tool and demonstrating use cases of variable renewable integration, RAVIS currently integrates site-specific solar power forecasts in the CAISO and MISO footprints from the IBM Watt-Sun forecasting platform, and superimposes market simulation data for the CAISO footprint from an in-house NREL market clearing tool (FESTIV).

RAVIS uses a technology suite that is assembled to provide optimum visualization facility while maintaining a wide pool of potential deployment and client environments. The tool is designed to take advantage of web application technologies and open-source visualization libraries and tooling. This will enable deployment in any environment, using any operating system, and it is easily scalable high spatial and temporal levels of visualization.

However, RAVIS is not a turnkey system. It is the product of a research endeavor, and it is not intended as a commercially viable product. To successfully deploy and operate RAVIS, the user must have a minimum basic understanding of web application software development and operations support knowledge. Some experience with NodeJS development and a working understanding of web-based mapping, including serving vector tile data, are also highly recommended.

7.3.2 Interactions with Users. The modular dashboard of RAVIS contains configurable panes for viewing probabilistic time-series forecasts; ramp event alerts on the look-ahead timeline; spatially resolved resource sites and forecasts; and system simulation and market clearing data, such as transmission line utilization, nodal prices, and available generation flexibility. The tool has the ability to alert the viewer to significant up or down ramps for both individual variable renewable sites as well as regionally aggregated net load ramps, and alerts can also be qualified with respect to available flexible generation.

Fig. 21 shows the RAVIS user interface. RAVIS contains four customizable panes: (1) site-specific and regional event alerts at various look-ahead times; (2) a regional overview of aggregated renewable resources; (3) a site-specific zoom-in view of distributed resources along with GIS information (not shown in Fig. 21 but viewable when a user clicks on or selects a region); and (4) a regional and site-specific forecast time-series viewer.

To demonstrate how users interact with RAVIS, forecast data for 10 solar PV sites each from the CAISO and MISO footprints were downloaded from the IBM Watt-Sun PAIRS Geospatial Analytics forecasting and data platform. Table 2 shows site information.

We now describe two modes of interaction:

1. ***Dynamic metadata lookup mode**:* As the user moves the cursor over a node in the

detailed regional pane, metadata for that station pops up, with information on the size of the plant (or aggregated size of several renewable generation plants if cursor moved in the regional view), and whether a significant ramp event is detected. A mock-up of this is shown in Fig. 21. If there is no significant ramping detected (based on the definition of a significant ramp set by the ramp configuration parameter), “Ramping nominal” will be shown. If a significant ramp is detected, the size of the ramp will be displayed. For instance, the site Medora, in North Dakota, and the aggregated forecasts in the eastern region have nominal ramping, whereas at the aggregated regional view, the western and central regions see significant ramping—59.85 MW and 65.84 MW, respectively.

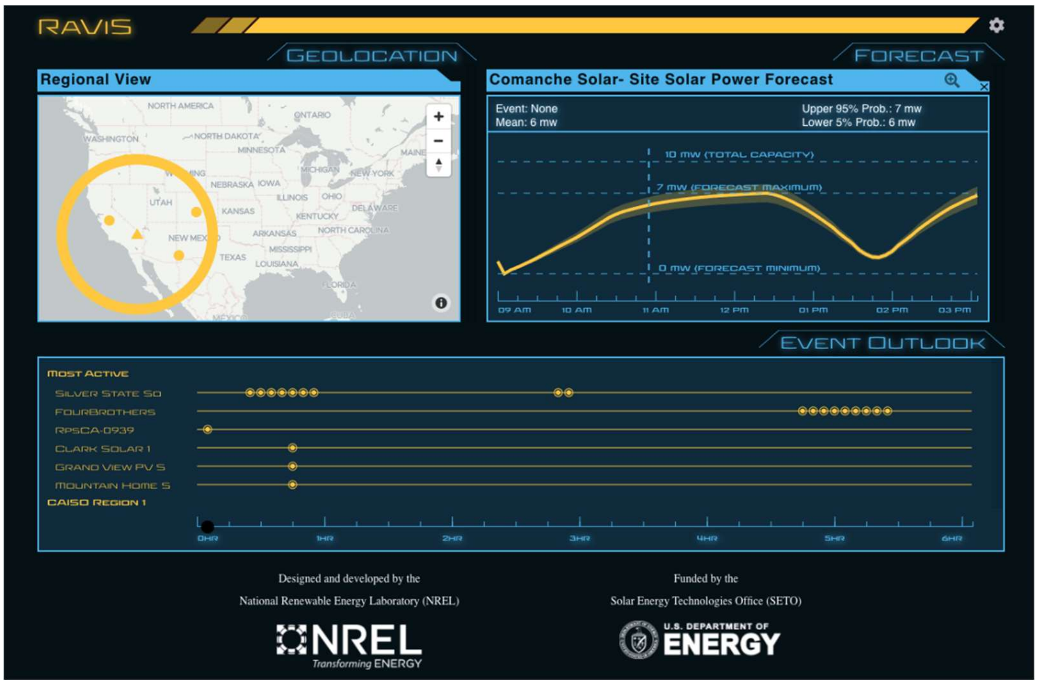


Fig. 21. RAVIS use interface: event alerts and spatiotemporal probabilistic forecasts

Table 2. CAISO and MISO Sites from IBM Watt-Sun Forecasting System

	CAISO			MISO		
Sites	**Station ID**	**Latitude**	**Longitude**	**Station ID**	**Latitude**	**Longitude**
1	CA_Topaz	35.38	-120.18	AMOA4	33.58	-91.8
2	RSAC1	38.47	-122.71	FRMI4	40.64	-91.72
3	RLKC1	40.25	-123.31	BNRI2	37.24	-89.37
4	SBVC1	34.45	-119.7	SULI3	39.07	-87.35
5	KNNC1	40.71	-123.92	NATL1	31.49	-93.19
6	MIAC1	37.41	-119.74	BDLM4	42.62	-85.65
7	MNCC1	34.31	-117.5	CASM5	47.37	-94.61
8	STFC1	34.12	-117.94	CKWM6	30.52	-88.98
9	DEMC1	35.53	-118.63	TS428	46.89	-103.37
10	COWC1	39.12	-123.07	RHRS2	43.87	-103.44

2. ***Viewing forecast time series mode***: When a user clicks on a particular node, either an individual site or the aggregated region, the time-series forecasts will be displayed in the right-hand pane. Fib. 16b shows the time series for the western region as well

as a single site in California.

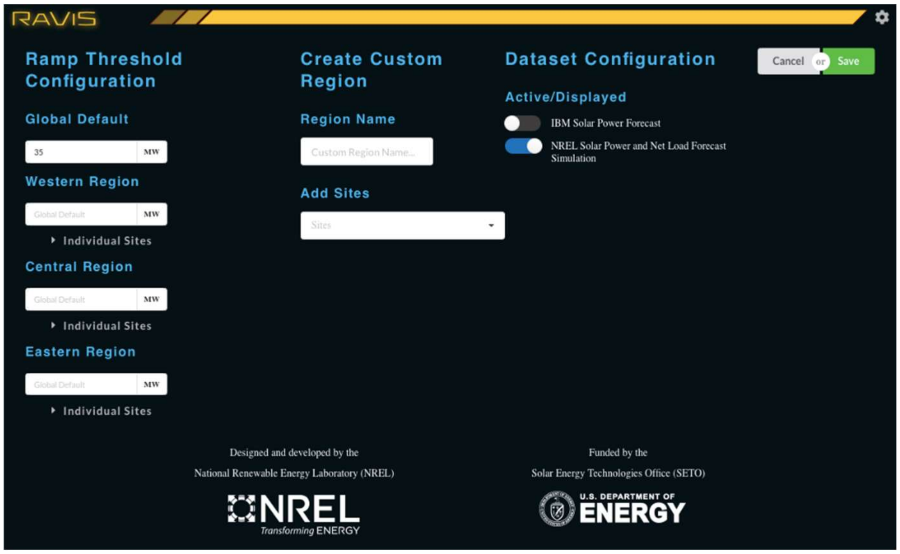


Fig.22. Configuration features in RAVIS

The tool is endowed with design flexibility and customization features useful for potential end users, including on-the-fly configuration and viewer updating features, as shown in Fig. 22. These can be accessed in the tool through the "gear" symbol shown in the top right corner. Currently implemented features include:

1. ***Ramp definition:*** Ramp definitions (MW change per minute) can be done at the global, regional, and site levels. This is shown in Fig. 22 (left).
2. ***Forecast zones or plant aggregation:*** Custom regions with selected plants can be created if users want to closely monitor them. This is shown in Fig. 22 (middle).
3. ***Time-series pane customization:*** The forecast pane includes a "+/-" symbol in the right corner that allows viewers to adjust the y-axis to the size of the ramp event. By default, the y-axis shows the renewable plant size or the total regional capacity.
4. ***Comprehensive data assimilation:*** Each end user has their own needs, and so the data ingestion in this tool is highly flexible for integrating visualization widgets of interest. For example, the tool can ingest additional data layers from various forecast vendors, electricity generation scheduling and market clearing data, and network topology and transmission data for comprehensive situational awareness. This is shown in Fig. 22 (right), with toggles selected to add more layers of data. This report introduces a use case that integrates additional electricity market-related data for understanding the interrelationships among forecasts, ramp uncertainties, and various system operating metrics.

7.3.3 Use Cases. One major use case discussed in this section includes the integration of detailed, site-specific probabilistic solar power forecasts and the detection of ramp events at both individual solar power plant and aggregated regional levels. Fig. 23 illustrates visualizing ramp alerts at different time instances, based on the input forecast data from the IBM Watt-Sun solar power forecasting platform, and the ramp definition parameters set in the configuration window. Fig. 23 shows several site-level solar power ramp alerts that are being detected in the central and eastern regions. In this example, note all the detected ramps in the site and at the aggregated regions are down-ramps and are shown in the visualization by the direction of the arrows; and the summary statistics in the time-series pane. This section discusses several such examples and use cases in detail.

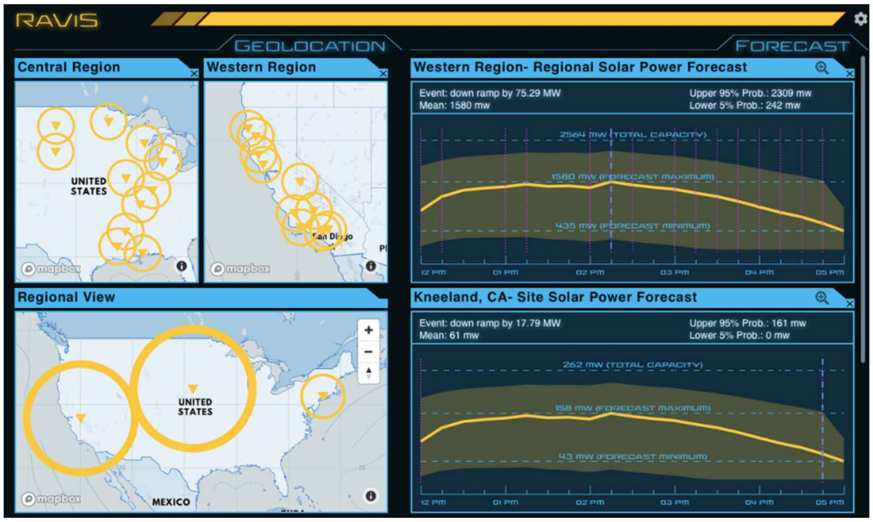


Fig. 23. Site-specific and regional renewable power ramp alerts and time-series forecasts for California and U.S.

Fig. 24. Comprehensive situational awareness with forecasts and data from the market clearing process. The market clearing process typically consumes forecasts as one input to ascertain system operation and generation dispatch decisions that result in anticipated nodal prices, transmission utilization, and available generation flexibility.

Another use case illustrates how the RAVIS tool can ingest multiple data layers in addition to variable renewable power forecasts for a comprehensive visualization capability. For instance, Fig. 24 shows a use case where RAVIS integrates 5-minute-resolution solar power forecasts and net load forecasts (load minus wind and solar) developed by NREL for the CAISO system for March 2020 (7 a.m.–12 p.m.). Additionally, RAVIS integrates market simulation results available from the NREL in-house simulation tool (FESTIV) for the modeled independent system operator system. These market results include:

- Network nodes and transmission topologies (the black dots and orange lines respectively in Fig. 24 (left), where lines >75% utilization are only shown.)
- Nodal clearing prices (appearing when the cursor is moved over the black dots.)
- Aggregated net load forecasts (Fig. 24 (right) shows the time-series probabilistic net load forecast data for the San Diego region.)
- Available generation flexibility in the upward and downward directions at the nodal

and aggregated system level (shown by the orange lines overlaid with the net load time-series forecasts in San Diego, thereby enabling operators to see whether there is sufficient generation flexibility to meet the net load uncertainties.)

The full report on RAVIS (Milestone 11.1.1's report) discusses these use cases in detail and the types of insights a system operator could gain and use. All these additional visualization features can be added by toggling the customization parameter discussed in Fig. 24 provided such data are available and fed to RAVIS appropriately.

7.4. Task 4: ISO Interaction and FESTIV Simulation of Solar Forecast-Informed Ramp Requirements

7.4.1 ISO Staff Interactions and Feedback. We held a total of 26 meetings with ISO staff and others to request data; learn about ISO market rules, forecasting, regulation, and FRP requitement methods; and to present and receive feedback on project results; and to demonstrate and obtain feedback on the RaVIS visualization system. The dates, topics, and attendees are listed below:

Meetings with ISO staff on solar and load data, ISO forecasting and requirements methods, and project results:

- 2018: July 12 (MISO), Aug. 3 (CAISO), Aug. 13 (MISO), Aug. 31, Sept. 28, Dec. 6, Dec. 18 (CAISO), Dec 21 (MISO)
- 2019: Jan. 7 (CAISO), Jan. 23 (MISO), Mar. 11, Apr. 2, May 15, July 2 (CAISO), Oct 4; Oct 8-9 (MISO), Dec. 5, Dec. 19 (CAISO)
- 2020: April 8, May 21, Sept. 4, Nov. 18 (CAISO), Dec. 12 (MISO)
- 2021: March 22 (CAISO)

RAVIS Demonstrations: Jan. 17, 2020 (MISO, CAISO, SETO Staff), May 20, 2021 (MISO, CAISO, Excel Energy, SETO Staff)

7.4.2 Overview of Simulations. We compared two methods for estimating uncertainty-related ramping needs in terms of system performance using two approaches. In the first approach (***theoretical comparison***), we develop a framework that classifies market intervals into five types with different anticipated performance in terms of reliability and economics. We use that framework to choose a few simulation days with different profiles (frequency of different types of intervals) for preliminary testing of new methods.

For the second approach (***practical comparison***), we focused on ramping requirements estimated a few hours in advance and contrasted two estimation methods: (a) a baseline (industry-inspired) method that considers calendar information and past errors and (b) an alternative (research-inspired) method that uses probabilistic solar forecasts and other weather information. Results on a small 118-bus system suggest that weather-informed estimation methods could yield different ramping requirements than existing calendar-based methods. A 0.5M$ production cost savings from the new requirements for a three week period in March 2020, when extrapolated to annual savings, amounts to about $8M/yr. This value is one-third of the total CAISO FRP procurement costs of $25M in 2018, and is of the same order of magnitude of FRP costs in 2019 and 2020. Moreover, preliminary results indicate that system conditions and ramping product design strongly influence the benefits. Due to space limitations, details are not provided here on the 118 bus simulations, but are available in our quarterly reports and papers [50].

We then compared results for an approximation of the western US power system for a sample of eight days in March 2020. We found that the solar forecast-informed approach decreased costs of power production (fuel and non-fuel O&M) and FRP procurement. These results are described in detail below. Although the small sample of days means that the results are not definitive, they do illustrate the potential value of using probabilistic solar forecasts to inform reserves procurement. Details are provided below.

7.4.3 Theoretical Comparison of Baseline and Solar Forecast-Informed FRP Requirements [61]. Day-ahead power markets or scheduling algorithms provide generator schedules usually with a lead time of approximately 14–38 hours, given net load forecasts and the technical constraints of thermal and hydro generators with respect to capacity, startup, shutdown, minimum on/off-time, and energy limits. However, uncertainty and variability in load and variable renewable output might lead to real-time deviations from the day-ahead schedules. Balancing markets, such as real-time markets in CAISO [10] and intraday markets in Europe [53], as well as balancing products such as ramping products and regulation, aim to address deviations and prevent energy imbalances by procuring flexibility and enabling the system to respond at lower cost to deviations from expected conditions. In this context, flexibility is defined as the load-following [54] ramping capability during a market interval starting from the last financially binding schedule. Therefore, a system needs flexibility or balancing capability equal to the system-wide net imbalance.

Sizing methods aim to estimate ramping requirements that will satisfy system flexibility needs within a target reliability level. Here, we compare system performance under an alternative set of ramping requirements (possibly provided by a weather-informed method using solar forecasts) to the system performance under the baseline (status quo) method. The “alternative” method can estimate: (a) higher, (b) lower, or (c) identical requirements compared to the “baseline” method for each market interval, as shown in Fig. 25.

Requirements under case (c) (identical) will cause no differences in system performance between the two methods. But under cases (a) and (b), different levels of requirements between the two methods might lead to differences in system performance. We use “might” because differences in ramp requirements are necessary but not sufficient for differences in system performance. In particular, the market software includes constraints that indirectly procure flexibility greater than or equal to the ramping requirements; so available flexibility could exceed the requirements. Moreover, the market software allows for deficits, i.e., procured flexibility lower than the requirements. In sum, the level of available flexibility, which depends on both overall system conditions and the requirements, affects system performance. For example, a system could be overly flexible during a market interval and its available flexibility could be higher than both sets of requirements, resulting in the same system performance under cases (a) and (b).

Fig. 25 further categorizes possible outcomes of the comparison of system performance under the alternative (solar forecast-informed) vs. the baseline estimation method by dividing the market intervals (a), (b), and (c) into five categories: a1, a2, b1, b2, and c. A “1” (i.e., cases (a1), (b1) indicate that the realized ramping needs are more than the ramping requirements estimated by a method, and the system does not necessarily have enough resources ready to ramp up, meaning that system reliability might be at risk. A “2” indicates that the realized ramping needs are less than the requirements of both methods (see cases (a2) and (b2) of the framework shown in the figure, so that both methods estimate adequate requirements for that market interval; hence, we do not anticipate any power balance violations caused by flexibility shortages and the method with the lowest requirements might incur lower production cost

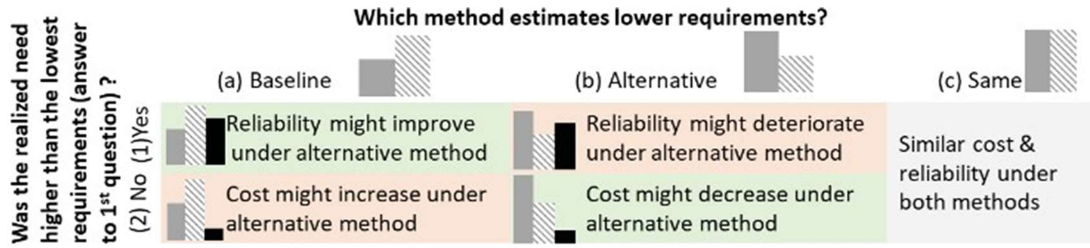


Fig. 25. Theoretical framework classifying market intervals by comparing ramping requirements from 2 methods (baseline/alternative, shown as solid/pattern grey bars, respectively) and the realized ramping needs (illustrated with solid black bars). Green, pink, and grey boxes indicate potentially improved, deteriorated, and similar system performance, respectively, under the alternative ramping requirement.

***7.4.4 Practical Comparison: Overview of System Simulation Model*.** Analysts usually quantify system performance through production cost simulations, which minimize aggregate costs, i.e., the sum of production cost (economic) and penalties for unserved energy (reliability). Analysts can formulate ramping products in tools such as the Flexible Energy Scheduling Tool for Integrating Variable Generation (FESTIV) [55] and conduct simulations with different sets of ramping requirements. FESTIV consists of a day-ahead scheduling algorithm run once per day and two rolling algorithms: real-time unit commitment (RTUC) and real-time dispatch (RTD). Following CAISO's conceptual design [30], we simulate a ramping product in FESTIV RTUC by adding constraints (1)–(6) for all time intervals $t \in \{1,2,\dots \text{HRTC-1}\}$:

$$\mathrm{costru_t} \geq \mathrm{WTP} * \mathrm{frus_t};\;\; \mathrm{costrd_t} \geq \mathrm{WTP} * \mathrm{frds_t} \tag{1a,b}$$

$$\alpha * \textstyle\sum_g \mathrm{fru_{g,t}} + \mathrm{frus_t} \geq \mathrm{FRUR_t};\;\; \alpha * \textstyle\sum_g \mathrm{frd_{g,t}} + \mathrm{frds_t} \geq \mathrm{FRDR_t} \tag{2a,b}$$

$$\mathrm{gen_{g,t}} - \alpha * \mathrm{frd_{g,t}} - \mathrm{rs_{g,RD,t+1}} \geq \mathrm{PMIN_g} * \mathrm{u_{g,t+1}} \tag{3}$$

$$\mathrm{gen_{g,t}} + \alpha * \mathrm{fru_{g,t}} + \textstyle\sum_{\mathrm{res} \neq \mathrm{RD}} \mathrm{rs_{g,res,t+1}} \leq \mathrm{PMAX_{g,t+1}} * \mathrm{u_{g,t+1}} \tag{4}$$

$$\mathrm{gen_{g,t+1}} - \mathrm{gen_{g,t}} \leq \alpha * \mathrm{fru_{g,t}};\; \mathrm{gen_{g,t}} - \mathrm{gen_{g,t+1}} \leq \alpha * \mathrm{frd_{g,t}} \tag{5a,b}$$

$$\{\alpha \mathrm{fru^{-}_{g,t}} + \tfrac{\beta}{2}(\mathrm{rs_{g,RU,t}} + \mathrm{rs_{g,RU,t+1}});\; \alpha \mathrm{frd_{g,t}} + \tfrac{\beta}{2}(\mathrm{rs_{g,RD,t}} + \mathrm{rs_{g,RD,t+1}})\} \leq \mathrm{RR_g} * \mathrm{D_{RTUC}} \tag{6a,b}$$

In words, each constraint *m* (*m* =1,2,5,6) is divided into (*m*a) for upward ramping and (*m*b) for downward ramping. Eqs. 1 estimate the penalties (costru, costrd) for ramping product deficits (frus, frds), which are added to the objective function of the scheduling optimization. Eq. (2) procures ramp to meet requirements (FRUR, FRDR), whereas eqs. (3) and (4) guarantee that the unit's generation (gen) and ramp will not violate its minimum and maximum operating limits, respectively, depending on its commitment (u). Eqs. (5) count expected change in generation schedule toward the ramping product. Eqs. (6) ensure that the ramping capability for the units is shared between ramp and reserve (rs) products. To ensure that flexibility produced in the real-time (short-start) unit commitment RTUC will be available in the real-time (5 min) dispatch RTD, we record RTUC schedules for the second interval (t=2) and force them in the subsequent run. Thus, we model $gen_{g,t}$, rs_g, $res_{,t}$, $u_{g,t}$ as parameters at t=1 and as decision variables for t>1.

A comparison of the aggregate cost under different uncertainty-related ramping requirements in absolute terms is informative; however, a comparison in relative (percentage) terms can be misleading because, by far, the bulk of aggregate cost is not caused by the uncertainty of net load or the procurement of uncertainty-related ramping products. That is why we propose and employ a metric called **"uncertainty-induced costs."** The "uncertainty-induced costs" are equal to (a) the aggregate cost of a simulation with net load uncertainty and ramping product minus (b) the aggregate cost of another simulation without uncertainty (i.e., perfect net load forecast). By subtracting (b), this metric omits the portion of the aggregate cost that would be incurred if net load were known with certainty.

We numerically illustrate the value of ramping sizing methods using a modified IEEE 118-bus system that mimics the annual generation mix of CAISO [56] with ~10% solar penetration (in terms of annual energy) in line with 2017 CAISO levels [57]. The case study closely follows the CAISO market structure in FESTIV [55] by simulating three markets: day-ahead with hourly resolution and a 24-hour horizon, RTUC with 15-min resolution and a 3-hour horizon, and RTD with 5-min resolution and a 1-hour horizon. Moreover, to reflect the balancing nature of RTUC and RTD, we include CAISO must-run rules [10].

ISOs have multiple balancing products to address uncertainty and variability. Here, we focus on a single uncertainty—forecast errors of solar generation with a lead time of 2–3 hours—because novel methods aim to quantify uncertainty of solar irradiance [58,59]. In practice, ISOs account for uncertainty induced by gross load and wind forecast errors as well. We also estimate requirements for one operating reserve product: the flexible ramping product (FRP) in RTUC using the baseline and alternative methods. The 1–4 hr lead time of RTUC in our simulation (as opposed to the actual 1–5 hrs in CAISO) facilitates integration of weather-informed probabilistic forecasts with a lead time of a few hours.

We assume that net load can take one of three values in each market interval t (i.e., lower, mean, upper). We estimate RTUC FRP requirements in the down and up directions that address both forecasted movement from t to t+1 and uncertainty at t+1 using (7) and (8) as follows. We set regulation and spinning reserves at 1% and 3% of gross load, respectively, and the value of lost load at $6,500/MWh, which is the administrative penalty for power balance violations in the scheduling runs of CAISO [10]. For simplicity, we do not procure flexible ramping product in the RTD market because we assume that solar generation is known with certainty in the CAISO RTD market (which closes at t-7.5 minutes, as opposed to the time when RTUC is run, which is at t-52.5 minutes).

$$\mathrm{FRDR}_{t} = \max\left(0, \mathrm{NL}_{t,\mathrm{mean}} - \mathrm{NL}_{t+1,\mathrm{lower}}\right) \tag{7}$$

$$\mathrm{FRUR}_{t} = \max\left(0, \mathrm{NL}_{t+1,\mathrm{upper}} - \mathrm{NL}_{t,\mathrm{mean}}\right) \tag{8}$$

7.4.5. Practical Simulation: System Performance for Large System: System Description. The large-scale CAISO system (with 1820 buses) is simulated in FESTIV, and we performed a cost-benefit analysis of the improved flexible ramping product. Two overarching tasks were performed to work towards this milestone:

1. Porting over FESTIV simulation to the high-performance computing (HPC) resource for gaining speed-up, and for running several scenarios in parallel.

2. Performing CAISO 1820-bus simulations for various scenarios: 1) perfect forecast without uncertainty, 2) with CAISO baseline FRP, and 3) with new improved FRP proposed by our team. Similar to the small system (IEEE 118 bus simulation analysis) analysis (summarized in our quarterly reports and [50]), our goal was to perform such an analysis using the large-scale CAISO system for ascertaining improved cost-benefit metrics related to economic savings and reliability benefits.

We simulated FESTIV for the WECC-sized power system model, summarized below.

Table 3. Summary of WECC scale simulation model

Number of buses	1,820
Number of transmission lines	3,053
Number of generators	3,787
Non-renewable units/installed capacity [MW]	1,742/158848
Hydro units/installed capacity [MW]	1,570/64008
Wind units /installed capacity [MW]	85/20004
Solar units /installed capacity [MW]	390/18540
Number of regions captured	40

FESTIV has been developed and used at NREL over the last 10-odd years, and this system represents the largest system ever simulated within FESTIV. Previous efforts mostly used IEEE 118 bus system, and therefore this marks ~15 fold increase in system size, which posed computational challenges. In order to find a tractable solution to the optimization of such a large system, several modeling assumptions were made.

1. First, a detailed representation of power system is only maintained for the California system operator's footprint (CAISO), and remaining regions outside the jurisdiction of the CAISO are modeled as single nodes interacting with CAISO.
2. Second, in order to capture real-world transaction costs, a flat $/MWh hurdle rate is assigned to exchanges of power between regions.
3. Third, hydro is not optimized but rather follows historical profiles as published by CAISO and BPA. The BPA profiles are applied to BPA hydro assets. The CAISO profiles are applied to all remaining hydro assets in the model.
4. Fourth, in order to simplify the solution space of the optimizations, start-up and shut-down trajectories are not explicitly modeled.
5. Fifth, flat start-up costs are used rather than dynamic start-up costs based on generator usage.

7.4.6. Practical Simulation: Illustrative Large System Dispatch. With these and the previously discussed assumptions in place, to illustrate the use of FESTIV for the large system and our input assumptions, we present details on the FESTIV simulation day-ahead market for one day, March 9th 2020 below. The overall system load is shown in Fig. 26. Note, the figure shows 48-hour load forecast, as the 1st 24-hour period is used by the day-ahead unit commitment as financially binding, and the 2nd 24-hour period is advisory only to prevent "end effect" distortions at the close of the 1st 24 hours.

The system-wide peak demand is just above 100 GW and exhibits a daily evening peak with a smaller peak during the morning hours. The output of the renewable generation assets is shown in Fig. 27.

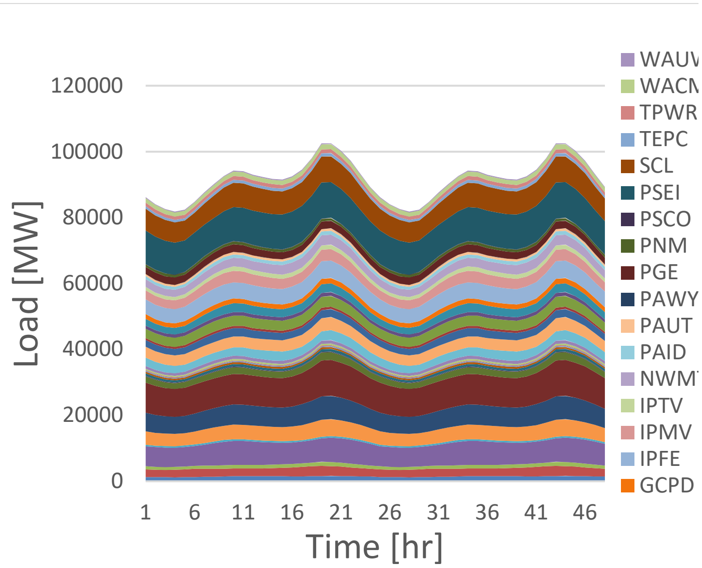


Fig. 26. System load in for regions in WECC

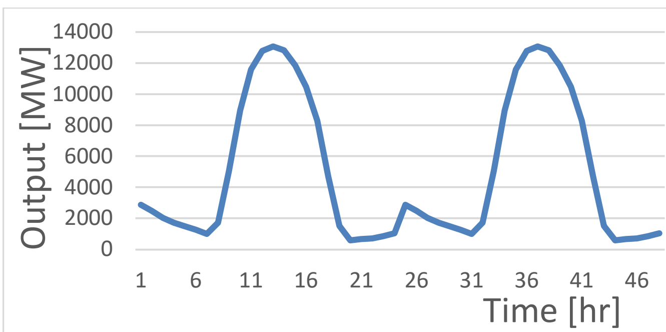


Fig. 27. Aggregated wind and solar output in the WECC model

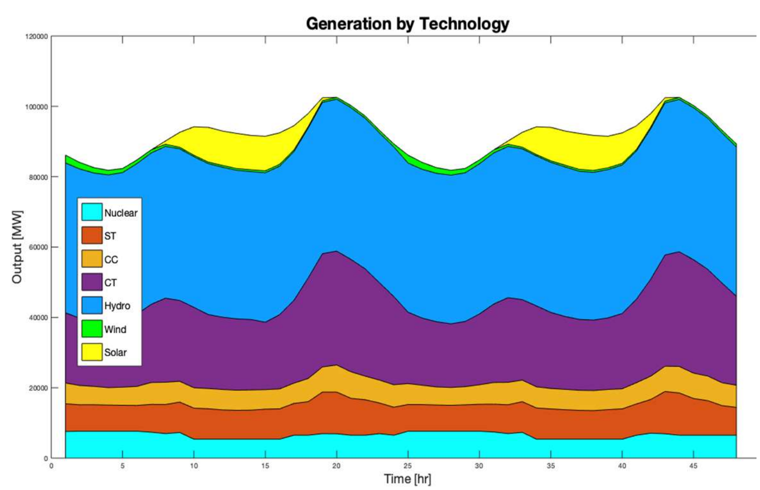


Fig. 28. FESTIV dispatch of generation by technology

The system shows large penetration of solar generation, topping ~12 GW of power during the day. The model is able to co-optimize energy and ancillary service requirements, and the calculated generation stack is presented in Fig. 28 below.

The day-ahead SCUC takes about 6-7 hours to finish the whole process (reading data, solving, writing output data). Note, the simulation results shown above are for the basecase without FRP (i.e., perfect forecast case). We discuss our comparison of two other scenarios below, namely the baseline CAISO FRP and new FRP for comparison.

7.4.7. Practical Simulation: Comparisons of Large System Baseline and Solar-Informed FRP Requirements. Scenarios with transmission constraints in this CAISO system were simulated to ascertain the benefits from weather informed FRP procurements. The production cost results are used to compare the results in this section. Three different simulation scenarios are considered and summarized below in Table 4.

Scenario 1 in that table establishes a baseline for the analysis with ramp uncertainties covered by current CAISO practice. Scenario 2 captures operational impacts due to updated FRP requirements calculated using probabilistic forecasts, following our proposed enhanced solar forecast-conditioned approach. Scenario 3 establishes a reference cost of the system, without net-load uncertainties. Any costs incurred above this value can be attributed to provision of grid services to counter uncertainties experienced by the system.

Table 4. Scenarios Simulation for Operational Impact Assessment

Scenario 1	Business as usual using published FRP requirements from OASIS system
Scenario 2	New FRP requirements calculated using probabilistic forecasts
Scenario 3	Operator has perfect visibility of the system, i.e., zero netload forecast errors

Due to HPC resource bottlenecks and data availability, the team performed a simulation of the CAISO system without transmission constraints for a three-week period in March, 2020, but only performed simulations of the CAISO system with transmission constraints for selected 8 days during this three-week period. These days included 3/16-3/20 and 3/23-3/25 in that year. Simulation results with network constraints are summarized below.

Figs. 29 and 30 show the FRP requirements using the baseline method (CAISO method) and proposed new FRP method (conditioned on weather, including information from the Watt-Sun probabilistic solar forecasts summarized in Section 9.1 above, using the BP2 models created by UT-Dallas, described in Section 9.2). Apart from being conditioned on solar power forecasting, we also see that the new solar forecast-informed FRP matches the pattern of expected morning and evening ramps well. The FRP-down requirement from the new FRP method is greater during morning hours when solar power is expected to ramp up and net-load is expected to ramp-down, thereby requiring more FRP-down. On the other hand, the new method's FRP-up is greater during evening hours, when solar power is expected to reduce, consequently causing net-load ramp to increase. Apart from these trends, other intermediate and diurnal patterns are due to changes in expected diurnal and hourly solar power forecasts and associated uncertainties. The proposed new solar-conditioned FRP method adapts itself well to the changing weather and forecasts; while the baseline method has highly similar FRP procurements for several contiguous days (due to the baseline method using historical 20-40 days as the basis for procurement, without considering latest weather changes).

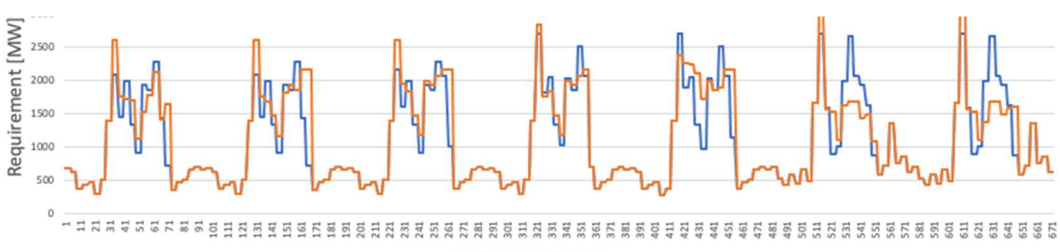


Fig. 29. Flexible Down Ramp Requirement: Baseline Vs. New FRP

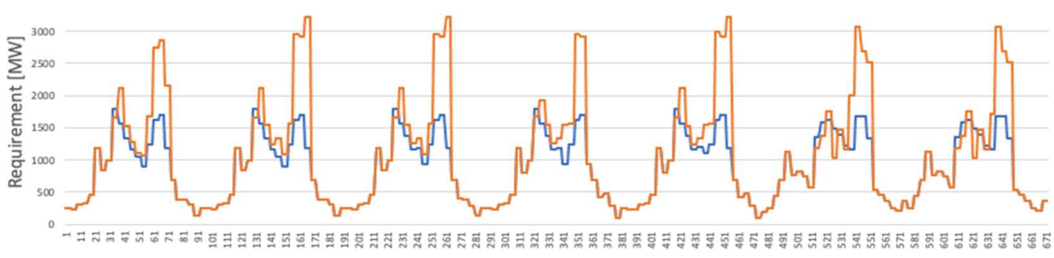


Fig. 30. Flexible Up Ramp Requirement: Baseline Vs. New FRP

The below tables and figures summarize cost impacts, including system production costs and uncertainty costs (i.e., the difference between experienced costs and the costs under the ideal circumstance of perfect forecasts). With the full network constraints considered, \$0.43 million is saved during the 8 days simulated with the proposed FRP method.

Table 5. Operational Impact from the Large 1820-bus CAISO system (5 day period)

March 16th-20th 2020 (5 days)	Baseline	New FRP	Perfect Forecast
Production cost (M\$)	66.9	66.5	52.31
Uncertainty cost (M\$)	14.59	14.19	
Savings (M\$)		0.332	
Savings from prod cost (%)		0.5%	
Reduction in uncertainty cost (%) (Actual minus Perfect Forecast cost)		2.34%	

Table 6. Operational Impact from the large 1820-bus CAISO system (3 day period)

March 23th-25th 2020 (21 days)	Baseline	New FRP	Perfect Forecast
Production cost (M\$)	39.74	39.65	33.94
Uncertainty cost (M\$)	5.8	5.71	
Savings M\$)		0.097	
Savings from prod cost (%)		0.24%	
Reduction in uncertainty cost (%) (Actual minus Perfect Forecast cost)		1.6%	

While the reduction in production cost is more modest compared to the significant reduction we reported earlier with the smaller IEEE 118 test system (Hobbs et al. 2021), the reduction is in itself non-negligible in terms of magnitude and in line with expectations for large system simulations. There is a 0.4% reduction (~\$430,000 reduction in 8 days) in production costs saving which stem from the slight modification of generation utilization throughout the day (March 16-20).

Table 7 gives system FRP procurement costs (price paid times quantity procured). With the proposed FRP method, procurement is reduced over 11% in the first 5 days (March 16-20) simulation in Table 5, and over 50% in the 3 days simulation shown in Table 6. The size of this cost saving depends on the FRP profile, and the system operation conditions. High solar penetration and conventional generation flexibility-limited scenarios may further increase the benefits of using probabilistic forecast-based FRP procurements.

Table 7. Total FRP procurement cost (\$K) from the large 1820-bus CAISO system

Time Period	Baseline (Scenario 1)	New FRP (Scenario 2)
5 days: March 16-20	76.14	67.53 (**11.3%** savings ~\$8.6K)
3 days: March 23-25	60.38	28.9 (**52.15%** savings ~ \$31.5K)

The minor changes due to improved FRP produces about 40.1K\$ savings in FRP procurement costs for 8-day period. This value is about one-tenth of the production cost decreases noted above, likely because the marginal price of FRP does not reflect changes in lumpy start-up and minimum run costs associated with changes in unit commitment. Note that FRP procurement costs in CAISO in 2018 were about \$25M/yr, and costs in 2019 and 2020 were about a third of this level according to CAISO Department of Market Monitoring reports. If we annualize this 40.1K\$ per 8 days to an annual savings, they extrapolate to about 1.9M$/year in savings, which is 7% of the 2018 actual FRP procurement costs, and 20% of the per year costs in 2019 and 2020.

Thus, we have observed procurement cost **savings of 11.3% (March 16-20) and 52.15% (March 23-25)** with the large model including transmission constraints. (These results are consistent with more extensive 3-week simulations within March 2020 without the transmission network, where we observed procurement cost savings of up to **12.7%** in the large CAISO system **without the transmission network.)** Therefore, simulation results on the large CAISO system meet the project's success value of 10-25% cost savings.

The large system (with transmission) daily production costs for the baseline CAISO and new FRP requirements cases are listed in Table 8. They show considerable variation day-to-day, with the new requirements better on average, but actually worse (higher) in two of three days . The number of units in baseline CAISO requirement and new FRP requirements cases are shown in Fig. . The number of online units is close but not identical, with the greatest differences obvious in the last day-and-a-half, which is also when the new FRP requirements save the most money. More units are committed to meet the evening peak, and less are committed in the middle of the night.

Table 8. Daily Generation cost in the large system simulation

With transmission	Base	NewFRP	New-Old	%Change
16-Mar	15357910.7	15428729	70818	0.5%
17-Mar	12531502	12433532	-97970	-0.8%
18-Mar	12215683.6	12291530	75846	0.6%
19-Mar	11938095.9	11909754	-28342	-0.2%
20-Mar	12329596.7	12073389	-256208	-2.1%

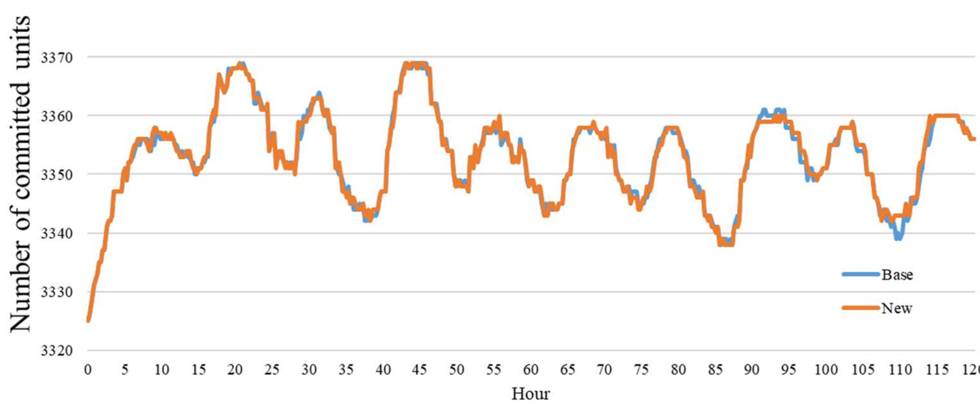


Fig. 31. Number of units committed (with-transmission constraints cases, March 16-20

We now examine the generation dispatch differences on the last two days (March 19-20), which are summarized in the Table below. Fig. 32 shows a time series of the differences in generation from various sources between the new FRP and base FRP cases. From Table 9, it can be seen that the main generation dispatch is that the new case obtains more energy from steam and combined cycle units, and less from costly combustion turbines and hydropower (which can have high opportunity costs). This is the source of the cost savings from using the new requirements.

Fig. 32. Generation dispatch difference between base case and new FRP case

Table 9. Generation dispatch difference (MWh) over March 19-20 between the new FRP requirements and the baseline FRP requirements case

	Nuclear	Steam	CC	CT	Hydro	Wind	PV
Sum	2276.4	12933.3	11367.9	-12294	-14587.7	-536	840

8. Significant Accomplishments and Conclusions

8.1 Task 1: Forecasting Accomplishments and Conclusions

8.1.1. Progress towards a scalable forecasting system. From our involvement in SETO's SF-I, we learned that by utilizing multiple NWPs, one can build a model to understand when, where, and how each NWP model performs, and by building a machine learning model to categorize different situations and treat each situation respectively, we can build a better solar forecast model. The key innovation of this previous research was that we found an effective way to enable customized modeling for each weather situation, which made the machine-learning model much more performant for extreme weather conditions. Competitive approaches would generally perform not as well and suffer from an overall averaging effect in extreme situations where little data was available.

However, while it is generally believed that machine-learning (ML) and AI approaches hold significant promise for improved solar forecasting, it is also clear that they will only perform as good as the input data used for training. This is true for ML in general but especially for situation-dependent and probabilistic ML approaches.

One of the main accomplishments of this research is that we have made significant progress towards building a much scalable solar forecasting system. However, scalability of solar forecasting will remain a huge challenge despite the progress we and others have accomplished here. It will therefore warrant much more research as the amount of data to process is growing exponentially and becomes more complex. Watt-Sun 3.0 is the only solar forecasting system, which can be easily scaled to larger data sets.

Most solar forecasting systems combine relational databases and object (or file-based) storage. Both approaches will not scale well for forecasting applications requiring data in hundreds of Terabytes. Towards that end, we have integrated a scalable backend technology, PAIRS Geoscope, which is key-value based and therefore scalable to hundreds of Petabytes. It avoids slow file-based operations and enables parallel compute via frameworks such as MapReduce or Spark. In addition, the PAIRS system deals automatically with different representations, reference systems, projections, spatial and temporal resolutions of different inputs of weather information or weather models automatically by operational industrial-scale ingestion pipelines from major content providers such as ECMWF, NOAA, etc.. Adding new weather information only requires activation of new pipelines, which will stream new data into the system as they become available. In this project we added numerous data sources important for solar forecasting such as HRRR and GOES-16. We implemented more ways by which PAIRS can support extracting "data frames" for subsequent ML tasks using an open-sourced SDK and Spark connectors. We develop techniques to deal with missing data and improved the reliability of the input data ingest, e.g., by developing a full monitoring and alerting tool for missing data. IBM has plans to open-source the PAIRS Software soon. The PAIRS SDK is already open-sourced and PAIRS provides OGC (Open geographic consortium) compliant web services. PAIRS is available for non-commercial purposes without a charge under an academic license.

This accomplishment is also significant because it will provide not only provide a path forward to improve forecast skills by improved efficiency and leveraging more data, which are exponentially increasing. The PAIRS platform has cross-industry use cases which means it is more economic than single industry-focused systems. PAIRS is already the backbone of the IBM's Weather Business Solutions. PAIRS has served up to 110M API requests per day and supports industry solutions/applications in energy, agriculture, insurance, and the public sector. For example, in the utility sector, PAIRS now supports commercial solutions including renewables forecasting, load forecasting, vegetation management, outage predictions, energy trading etc.. Towards that end, PAIRS includes many other non-energy related data which will become more important for decision making tools. Examples of such data include land use, population, and critical infrastructure.

8.1.2. Rasterized probabilistic solar irradiance forecasts. The scalability of the PAIRS system has been exemplified by developing rasterized, short-term (1 hour ahead) probabilistic solar forecasting data product based on GOES-16 and individual measurement stations. GOES-16 provides up to 3 Terabytes of new data per day, from which a significant fraction is being ingested and curated in PAIRS in near real-time. While ingesting all that data is a challenge, it is equally a challenge to compute forecasts in a timely manner, especially considering that short-term solar forecasts. PAIRS and the described improvements (e.g., batch exports for SPARK ML, etc.) enabled us to provide such a rasterized probabilistic short-term solar forecast which is believed to be uniquely differentiated from other forecasting systems.

A key innovation in this accomplishment is a deep learning model to forecast GOES L2 data from optical cloud depth. This model uses instead of a spatial convolution layer, a volumetric convolution neural network layer, which allowed conducting convolutions to automatically extract features that are spatial temporal correlated. Another important innovation is that this deep learning model was coupled to a quantile regression-based machine-learning model, which mapped the forecasted L2 data and the measured ground-level GHI. Fig. 33 shows an hour ahead GHI forecasts based on the approach for the 50% quantile on 02/15/2020 on 00:00 UTC for part of the Western United States.

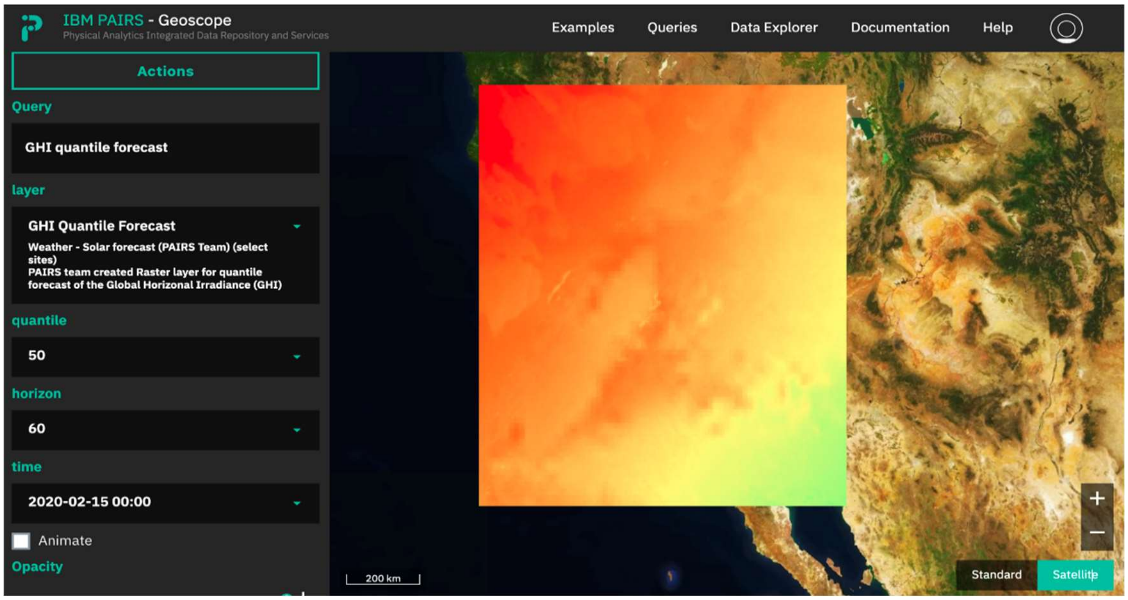


Fig. 33. Example of Rasterized 1-hour probabilistic ahead forecasts

8.1.3. Watt-Sun solar forecasting performance. The different versions of the Watt-Sun forecasting systems delivered consistently improved accuracies, consistent with the milestone targets. These continual improvements are the foundation of IBM's commercial offerings in this space.

- Watt-Sun 0.0, which was an initial prototype, delivered a Brier score of 0.4 over three 28-day periods for a single California site providing GHI forecasts at 15 minutes intervals with 4-hour forecast horizon and a refresh rate of every hour.
- Watt-Sun 1.0 delivered GHI solar forecasts for 20 sites (10 sites in CAISO and 10 sites in MISO regions). Forecasts were operational with 5 quantiles (0.05, 0.25, 0.5, 0.75, 0.95) with a forecast horizon of 24 hours at 15 min frequency and refresh rate. Brier scores for each site were below 0.3 over two periods of 28 consecutive days, which is an improvement of 25% over Watt-Sun 0.0. Compared to an ISO standard ARIMA method Watt-Sun 1.0 achieved improvement of larger than 20%.
- Watt-Sun 2.0 was evaluated for the same 20 sites. Forecasts were made for the same 5 quantiles with a forecast horizon of 24 hr at 15 min frequency and refresh rate. Forecasts were evaluated using the P-P plot score in comparison to bias-corrected HRRR GHI forecasts. Watt-Sun 2.0 not only performed better than Watt-Sun 1.0 based on Brier scores but also showed a 20% improvement in the P-P plot score over bias-corrected HRRR GHI forecasts. In terms of rMAPE. Watt-Sun 2.0 also performs better, with a 17% improvement over HRRR bias-corrected forecast.

Finally, Watt-Sun 3.0 was evaluated based against persistence for 24 sites, but the same forecast horizons, frequencies, and quantiles as Watt-Sun 1.0 and 2.0. Watt-Sun 3.0 showed that *79% of Watt-Sun 3.0 probabilistic forecasts are better calibrated than the (probabilistic) persistence estimator*. More specifically, for the forecast horizon of 4 hours and for 7 stations, a consistent improvement of at least 20% of the P-P-plot metric was achieved by Watt-Sun 3.0 as compared to a persistence estimator. Additionally, 2 stations achieve a consecutive run of at least 28 days for a 30% P-P-plot metric improvement. For the forecast horizon of 24 hours and for 11 stations, a consistent improvement of at least 20% of the P-P-plot metric was achieved as compared to a persistence estimator.

8.2 Task 2: Reserve Requirement Definition Accomplishments and Conclusions

Contributions of this task include: 1) development of a systematic way to prepare, process, and extract key features from probabilistic solar forecasts, which can be used by market operators in conjunction with data-driven methods to make weather-informed decisions; 2) demonstration of the trade-offs between system reliability and market efficiency as different parameters or procurement strategies are selected; and 3) modeling FRP and regulation as a two-objective (reliability and procurement) optimization problem.

By applying the proposed kNN-based FRP requirements method to the CAISO 15-min RT market, our results from three representative months in 2020 suggest that the proposed method can improve the performance compared to the baseline in terms of both system reliability and oversupply relative to need. Because of the performance sensitivity to the kNN parameters, we also proposed a selection method to dynamically identify the best kNN parameters per hour of the day, which allows power system operators to maximize the benefits of the kNN-based method. Key insights from our results include:

1. The performance of the kNN-based method varies across classifiers and the kNN parameters. Generally, a greater K results in more conservative estimations, i.e., greater FRP requirements and increased reliability levels.
2. Despite variations in performance, the solar/weather-informed decisions have better reliability and market efficiency most of the time, notably during early morning and late afternoon, when sunrise/sunset cause greater uncertainties.
3. Although different classifiers from a single site explain weather uncertainty from different perspectives, they also present considerable correlations. In addition, a comparison of results using forecasts from different sites suggests that geographical variations can affect performance. In response, the PCA-based method effectively characterizes the system-level weather uncertainty by using only a reduced set of principal components, and results in improved performance when compared with all single-site cases. The proposed approach is extendable to include more sites (e.g., the entire CAISO region) to further improve the performance, and therefore promises even greater benefits in real applications.
4. Different from the kNN-based method that is bounded by similar days, the machine learning-based methods learn to estimate reserves with greater variations. Moreover, the ML/DL-based methods tend to underestimate FRP requirements, which further result in increased frequencies of FRP shortages. This is because FRP requirements are subject to a lower bound of 0 MW but no upper bound, and because the ML/DL-based methods put the same weight on the two objectives of FRP shortages and oversupplies, so underestimating FRP requirements result in smaller prediction errors. Thus, a better objective function is required in the ML/DL-based methods to better account for real world trade-offs between reliability and economics.
5. Our solar forecast-conditioned requirements for FRP and regulation-up consistently showed better performance than the CAISO baseline methods which did not consider weather or solar conditions. For FRP, although the ML/DL-based benchmarks can potentially reduce FRP oversupply, at the same time our methods present better reliability. Further, the optimized PCA-kNN results outperform the best adjusted

CAISO benchmark results in Aug. 2020, indicating superior and robust performance. In general, our optimized PCA-kNN results present better performance than CAISO's baseline across all months. In addition, quadratic regression methods for estimating requirements also outperformed CAISO's baseline method in out-of-sample tests, although which particular specification depends on the month and time interval. Meanwhile, for regulation-up, short-term (2 hr-ahead) prediction of area control error (adjusted for reg-up deployment) using solar forecasts and time series data yields real-time reg-up requirements that outperform the present day-ahead method.

8.3 Task 3: Visualization Accomplishments and Conclusions

We have shown how RAVIS successfully integrates forecast data from the IBM Watt-Sun system, in addition to in-house NREL forecasts, that can be refreshed periodically. We have also provided information on the innovations and unique features of this tool, the basic architecture and open-source libraries used, and illustrations of use cases where site-specific solar power forecasts have been integrated and ramp alerts are visualized at the site and regional levels. One use case also includes the integration of market clearing data to provide a proof of concept for plug-and-play layered architecture of heterogeneous data integration and seamless integration of system operational data visualization in conjunction with forecast data. Example data includes network topology, nodal attributes (flexibility, prices), transmission congestion, and available aggregated generation flexibility (compared to regional aggregated net load probabilistic forecast and ramps).

The vision is for this tool to be used by end users who can tweak it as driven by their particular needs. End users can expand it to include alerts to ingest more sensor data and weather data into this common platform, including site-level cyber anomalies and regional stability alerts, and they can superimpose severe weather events to forecast their evolution and the affected regions. Additionally, given the importance of co-simulating bulk transmission and distribution systems to study the role of distributed energy resources (DERs) to mitigate stability issues under various IEEE 1547 standards, the visualization can be expanded to include high spatial resolution of distribution networks with customer-sited DERs that are connected to respective transmission substation nodes. It is acknowledged that each end user might have their own needs for features; therefore, although this final report summarizes the set of standard developments related to specific tasks involved in the SETO-funded project, the tool is flexible for further R&D.

8.4 Task 4: Simulation-Based Benefits Accomplishments and Conclusions

The baseline FRP requirement (using the CAISO's method) and the new FRP requirement with the proposed solar forecasting method have been compared for a large transmission network (WECC, including the CAISO). The multi-timescale power system simulation tool, FESTIV, has been used for this simulation. Three cases representing the baseline FRP requirement, the new FRP requirement, and perfect forecasting are analyzed. We find that using the proposed solar-informed FRP requirement estimation method, system generation cost and FRP procurement expense can be reduced in this large system. The technical challenge of large-scale transmission network simulation have been solved using NREL high performance computers, which can be leveraged for future studies.

9. Budget and Schedule

See the final RPPR2 and SF-425 reports for financial details. The federal project budget has been underspent slightly (4%), which has not impaired the ability of the project team to accomplish the goals of the project. The budgeted cost-match goal been met. Due to COVID-caused obstacles to use of the high-performance computing facility at NREL, there were delays in completing the Task 4 system simulations to assess the benefits of solar uncertainty-informed ramp requirements, and the simulations were less extensive than initially planned. To accommodate the adjustments and delays in the project tasks, a 6 month no cost extension was requested and granted.

10. Path Forward

10.1 Task 1: What's next for probabilistic solar forecasts? Algorithms to scaling

Despite the progress the technical community has made in advancing methods and techniques for solar forecasting, it is becoming increasingly evident that improved algorithms alone, whether physics-based, data-driven or a combination, will not bring about the required forecasting accuracies, especially for high impact situations of extreme weather conditions, where it matters the most. In fact, it is a perhaps painful realization of the "AI revolution" that such approaches and techniques are often not scalable and must leverage much bigger data sets to become more robust [51], to yield reliably the required accuracy, and to be able to inform high-value decisions for operations as the grid becomes more complex and as we are moving from renewable energy forecasting to integration. That is, research must shift from developing algorithms to a focus on scaling the underlying information architectures that support these approaches. On a practical level, this means that such AI-enhanced big data technologies (i) must not only be generalizable to make them economic. Further, (ii) they must be extendable as one seeks to manage more complex and larger energy systems as exponentially more data is becoming available.

On a technical level the challenges can perhaps be best explained by the notion of *data gravity*, which results from two facts. *First,* the data sets required to fuel more complex and robust models for renewable energy generation have become very big. For example, take weather or climate data. Every day, hundreds of TBs of weather/climate-related information are generated (e.g., weather stations, radar or satellite observations, or model forecasts), which are growing >25%/y [51]. For example, 100 TBs daily takes more than a week to move from a storage device to the memory of processor for subsequent computation (at 100 MB/s read speed of a hard disk). *Second,* most applications require additional data sets (e.g., advanced metering infrastructure or electrical grid networks) that must be linked, for instance to enhance the fidelity of a machine-learning model.

Data gravity describes the fact that data sets are too big to be moved and thus big data tends to attract more data - in the same way a bigger mass attracts a smaller mass. Data gravity also means that data movement must be minimized. While in other domains (i.e., consumer choice prediction), specific and proprietary technologies have been invented to overcome some of the discussed challenges, one should keep in mind that in case of the energy sector we are dealing not only with even more data, but the information is also much more complex and heterogenous, compared to the consumer market where one might work with a common data set such as web pages. On a technology level, data gravity does not only mean that next generation scalable systems and platforms must heavily leverage (public and/or private) cloud computing but also one must invent new information architectures that will (i) drastically minimize the data movement and (ii) enable indexing and linking of highly heterogenous data so that they become query-able and search-able (in the same way we can search today 40B webpages in 0.5s). *The core of a new program could include research, development, demonstration and commercialization of such a novel information architecture and system, which will enable vastly superior exploitation of energy relevant data for the application of renewable energy integration.* Research questions and subtopics may include:

- Novel forms of indexing plus distributed data and computing architectures (for processing energy relevant data, e.g., time series, geospatial, weather, and AMI data),
- New architectures and designs to facilitate complex "in data" computation with minimal data movements,
- Hybrid computational approaches to facilitate distributed learning across multiple data repositories, and
- Development of benchmarks for testing new information architectures.

10.2 Task 2: Informing Ancillary Services Requirements

In general, there are at least three open research questions for the development of practical weather-informed estimation methods for ramping requirements. First, methods that provide probabilistic forecasts of solar irradiance are yet to be thoroughly validated during long periods using metrics such as P-P plots, sharpness indices, and Brier scores. Translations into solar power uncertainty also need to be validated. Second, uncertainty related to wind and gross load should also be considered. Third, multiple balancing products address uncertainty of net load. Future work will assess the potential of weather-informed methods for additional balancing products including ramping and regulation using the framework. We now discuss these directions in more detail.

It is estimated that CAISO's solar power uncertainty contributes at least half of overall net load uncertainty. Therefore, we have focused on the uncertainty component caused by solar power only, and we assume perfect foresight for wind and load. Future work should address quantification of economic and reliability benefits of using integrated solar, wind, and load probabilistic forecasts in system operations to inform ancillary service and flexible ramp requirements. Those requirements should be defined taking a multi-objective approach, as we have shown in our Pareto methods for defining solar uncertainty-informed requirements considering reliability and economics (Sect. 7.3). The appropriate balance of the two objectives would be user and situation dependent and would best be selected after studying the trade-offs embodied in the set of efficient (Pareto) alternatives.

In the FRP estimation work, a critical caveat in our analysis is that the solar power output is converted from solar irradiance by using simulation tools, which may not reflect the actual irradiance-to-power conversion because of imperfect knowledge of technical parameters of real-world solar power plants, such as inverter capacities, PV panel capacities, and degradation rates. Future work can calibrate the simulated results by comparing it with real-world data or conducting sensitivity analyses to address parametric uncertainties. Comparative studies can also be conducted to demonstrate the impact on the model and, ultimately, power system performance by using simulated PV power as classifiers.

Incorporation of solar forecasts in requirements definitions could be extended to other ancillary services, as we show for regulation. Due to limited data availability, only regulation-up is examined, and preliminary results are obtained. Future work can aggregate the merits of the classification methodology for intra-hour short-term forecasting, and adaptive procurement for regulation and other ancillary service requirement estimation in electricity markets. In addition, improved forecasts will likely result from training separate ML/DL models for different weather classifications using more years of data.

The algorithms and outcomes of this project could be leveraged for potential commercialization through future DOE SIBR/STTR projects. In addition to utility-scale solar generation, the methods could extend probabilistic solar forecasting to behind-the-meter (BTM) generation, and account for BTM solar in determining needed reserve products.

10.3 Task 3: Ramp and Uncertainty Visualization

Advanced real-time operation and control of renewable resources. We've illustrated how operators can be informed by visualization of ramps and solar uncertainty. Future work could emphasize helping end users to obtain a much more comprehensive situational awareness, allowing them to relate resource forecasts and their ramps to power system operational metrics, such as generation flexibility, transmission congestion, and nodal prices. Any of these system metrics could be used as additional alerts, and an operator could take appropriate corrective actions knowing the status of the grid in real time. For instance, while responding to nodal inflexibilities, it would be good to anticipate transmission congestion to ensure that the decision to add additional reserves can be effectively delivered to the required location. If such delivery is challenging because of congestion, an operator could resort to controlling the outputs of variable renewable resources in that location (again, by looking at the resource forecasts and anticipated ramps in those local resources) and curtail or adjust ramp rates to meet the reliability and avoid a ramp scarcity event and the consequent real-time price spikes. These are some advanced control room real-time operations that could be made feasible by an interactive and dynamic visualization platform that provides comprehensive situational awareness by ingesting heterogeneous data along with probabilistic solar, wind, and load forecasts.

Component-level forecasts, ramping, and innovative solutions for additional flexibility. The use cases presented in Section 8.3 illustrated the visualization of solar power forecasts and net load forecasts, but this tool is customizable to visualize every component of net load individually—i.e., utility-scale solar, distributed solar power, wind power and load power—and their aggregated form as net load. Therefore, a system operator could look at the relative contributions of each component of the net load variability and uncertainty at the nodal or regional level and accordingly devise mitigation strategies as well as compensation mechanisms. One could quantify the components that aid system reliability and the components that deleteriously impact reliability and identify locations where additional flexibility in the system might be beneficial. Such analysis could also encourage the owners of these variable renewable power plants to develop innovative solutions for enhancing flexibility and reliability, including investigating options to install co-located or AC-/DC-integrated storage and hybrid systems.

10.4 Task 4: Simulation-Based Benefits of Using Probabilistic Forecasts

Preliminary results are encouraging for the development of more accurate ramping requirement estimation methods (possibly weather-informed), based on simulations of a small test system as well as a large (Western US) system during a sample of days in March 2020, as described above. However, net cost and reliability simulations for a longer period (e.g., a year) are needed to provide definitive benefit estimates that will be most useful to operators. Considering longer periods will be useful for identifying (a) issues such as nondeliverability of reserves that prevent the realization of benefits and (b) periods during which more accurate quantification could be most valuable. Given the importance of system effects, future simulations should use realistic systems under a wide range of scenarios, including varying penetration levels of renewable resources. We will also analyze correlations of real-time power balance violations with traditional reliability metrics such as control performance standards [52] because violations provide information on both economic (out-of-market corrections) and reliability performance.

11. Inventions, Patents, Publications, and Other Results

See RPPR2 for workshops (3), websites (2), and other conference/workshop talks (29).

1. ARTICLE: B. Li and J. Zhang, A review on the integration of probabilistic solar forecasting in power systems, *Solar Energy*, 210, 68-86, 2020, doi.org/10.1016/j.solener.2020.07.066
2. ARTICLE: C. Albrecht, B. Elmgreen, O. Gunawan, H. Hamann, L. Kleein, S. Lu, F. Mariano, C. Siebenschuh, J. Schmude, Next-generation geospatial-temporal information technologies for disaster management, *IBM Journal of Research and Development*, 64(1/2), 12 May 2019, DOI: 10.1147/JRD.2020.2970903
3. ARTICLE: B. Li, C. Feng, C. Siebenschuh, R. Zhang, E. Spyrou, V. Krishnan, B.F.Hobbs, J. Zhang, Sizing Ramping Reserve Using Probabilistic Solar Forecasts: A Data-Driven Method, *Applied Energy*, 313 (2022): 118812, doi.org/10.1016/j.apenergy.2022.118812
4. ARTICLE: J. Fox, E. Ela, B.F. Hobbs, J. Sharp, J. Novacheck, A. Motley, R.J. Bessa, P. Pinson, and G. Kariniotakis, Forecasting and Market Design, Supporting an Increasing Share of Renewable Energy, *Power & Energy Magazine (IEEE)*, 19(6), Nov/Dec 2021, 77-85.
5. ARTICLE: B.F. Hobbs, V. Krishnan, et al., “Development of Flexible Ramp Product Procurement for the California ISO using Probabilistic Solar Power Forecasts”, *Solar Energy Advances*, in review.
6. ARTICLE: B.F. Hobbs, V. Krishnan, et al., How Can Probabilistic Solar Power Forecasts Be Used to Lower Costs and Improve Reliability in Power Spot Markets? A Review and Application to Flexiramp Requirements, *Open Access J. Power Engineering (IEEE)*, in review.
7. PROCEEDINGS: E. Spyrou, V. Krishnan, Q. Xu and B. F. Hobbs, "What Is the Value of Alternative Methods for Estimating Ramping Needs?," *2020 IEEE Green Technologies Conference*, 2020, pp. 159-164, doi: 10.1109/GreenTech46478.2020.9289786.
8. PROCEEDINGS: P. Edwards, H. Sky, V. Krishnan, “RAVIS: Resource Forecast and Ramp Visualization for Situational Awareness of Variable Renewable Generation”, *e-Energy '21: Proc. of the 12th ACM Intl. Conference on Future Energy Systems*, Torino, Italy, June 2021, pp. 345-354; www.nrel.gov/docs/fy21osti/79786.pdf, doi.org/10.1145/3447555.3466594
9. PROCEEDINGS: B. Li, J. Zhang, B. Hobbs, "A Copula Enhanced Convolution for Uncertainty Aggregation", *IEEE 11th Conf. on Innovative Smart Grid Tech*, Washington, DC, Feb. 2020, ieeexplore.ieee.org/abstract/document/9087644, DOI: 10.1109/ISGT45199.2020.9087644
10. BOOK CHAPTER: S. Lu, H. Hamann, "IBM PAIRS: Scalable big geospatial-temporal data and anlytics as-a-service" Chapter 1, in *Handbook of Big Geospatial Data*, M. Werner, Y.-Y. Chiang (eds.), Springer, pp. 3-23, 2021, doi.org/10.1007/978-3-030-55462-0_1, ISBN: 978-3-030-55462-0
11. PATENT: J.B Chang, H.F Hamann, L. Klein, S. Lu, Multimodal analyte sensor network, Applied 1/8/2019, Application US10871479B2, Awarded 12/22'2020
12. REPORT: P. Edwards, H. Sky, V. Krishnan, RaVIS: Ramp Visualization for Situational Awareness: An Introduction to the Open-Source Tool and Use Cases, NREL Technical Report NREL/TP-5D00-79746, www.nrel.gov/docs/fy21osti/79746.pdf, May 2021
13. GITHUB: https://github.com/ravis-nrel/ravis, RaVIS ramp visualization system, 3/2021

12. References

[1] Tuohy, A., et al. "Solar forecasting: methods, challenges, and performance." *IEEE Power and Energy Magazine* 13.6 (2015): 50-59.
[2] Antonanzas, J., et al., 2016. Review of photovoltaic power forecasting. Sol. Energy 136, 78–111.
[3] Bauer, P.,et al. "The quiet revolution of numerical weather prediction." *Nature* 525.7567 (2015): 47-55.
[4] Jimenez, P.A., et al. "WRF-Solar: Description and clear-sky assessment of an augmented NWP model for solar power prediction." *Bulletin of the American Meteorological Society* 97.7 (2016): 1249-1264.
[5] Li, B., and J. Zhang. "A review on the integration of probabilistic solar forecasting in power systems." *Solar Energy* 210 (2020): 68-86.
[6] Ghalehkhondabi, I., et al. "An overview of energy demand forecasting methods published in 2005–2015." *Energy Systems* 8.2 (2017): 411-447.
[7] Voyant, C., et al. "Machine learning methods for solar radiation forecasting: A review." *Renewable Energy* 105 (2017): 569-582.
[8] E. Ela et al., "Electricity markets and renewables: A Survey of potential design changes and their consequences," *IEEE Power Energy Mag*., vol. 15, no. 6, pp. 70–82, 2017.
[9] Midcontinent ISO, "Energy and Operating Reserve Markets," BPM-002-r18 Sep-23-2017, 2017.
[10] CAISO, "Business Practice Manual for Market Operations Version 60." 2019.
[11] Southwest Power Pool Market Monitoring Unit, "State of the Market 2018," 2019
[12] A.A. Thatte, L. Xie, A metric and market construct of inter-temporal flexibility in time-coupled economic dispatch, *IEEE Trans. Power Syst.* 31 (5) (2016) 3437–3446.
[13] G.G. Dranka, et al., A review of co-optimization approaches for operational and planning problems in the energy sector, *Appl. Energy* 304 (2021) 117703.
[14] H. Holttinen, et al., Methodologies to determine operating reserves due to increased wind power*, IEEE Trans. Sust.. Energy* 3(4) (2012) 713–723.
[15] B. Li, J. Zhang, "A Copula Enhanced Convolution for Uncertainty Aggregation", *IEEE 11th Conf. on Innovative Smart Grid Tech*, Washington, DC, Feb. 2020
[16] M. A. Matos, R. J. Bessa, Setting the operating reserve using probabilistic wind power forecasts, *IEEE Trans. Power Syst*. 26 (2) (2010) 594–603.
[17] N. Costilla-Enriquez, M. Ortega-Vazquez, A. Tuohy, et al., Operating dynamic reserve dimensioning using probabilistic forecasts, arXiv preprint arXiv:2108.09362 (2021).
[18] P. Etingov, L. Miller, Z. Hou, et al., Balancing needs assessment using advanced probabilistic forecasts, in: 2018 IEEE Intl. Conf. on Prob. Methods Applied to Power Systems (PMAPS), IEEE, 2018, pp. 1–6.
[19] P. Etingov, L. Miller, Z. Hou, et al., Improving BA control performance through advanced regulation requirements prediction, in: 2018 IEEE Power & Energy Society General Meeting (PESGM), 2018.
[20] Marinelli, M., et al. *Demonstration of Visualization Techniques for the Control Room Engineer in 2030.* ELECTRA Deliverable D8.1. WP8: Future Control Room Functionality. backend.orbit.dtu.dk/ws/files/130245204/D8_1_Demonstration_visualization_for_control_room_2030.pdf
[21] Jones, L. E. 2011. *Strategies and Decision Support Systems for Integrating Variable Energy Resources in Control Centers for Reliable Grid Operations*. Washington, D.C. Alstom Grid. https://www.osti.gov/servlets/purl/1032531
[22] Pennock, K., Beaucage, P., Makarov, Y.V., et al., 2016. *Comprehensive Solutions for Integration of Solar Resources into Grid Operations.* AWS Truepower, LLC. www.osti.gov/servlets/purl/1234682
[23] Etingov, P. V., et al.. 2012. *Online Analysis of Wind and Solar Part I: Ramping Tool*, PNNL-21112, PNNL, Richland, WA, digital.library.unt.edu/ark:/67531/metadc842084/m2/1/high_res_d/1060669.pdf
[24] E. Ela and M. O'Malley, "Scheduling and Pricing for Expected Ramp Capability in Real-Time Power Markets," IEEE Trans. Power Syst., vol. 31, no. 3, pp. 1681–1691, 2016.
[25] B. Wang and B. F. Hobbs, "Real-Time Markets for Flexiramp: A Stochastic Unit Commitment-Based Analysis," IEEE Trans. Power Syst., vol. 31, no. 2, pp. 846–860, 2016.
[26] A. Daraeepour, D. Patino-Echeverri, and A. J. Conejo, "Economic and environmental implications of different approaches to hedge against wind production uncertainty in two-settlement electricity markets: A PJM case study," Energy Econ., vol. 80, pp. 336–354, 2019.
[27] K. H. Abdul-Rahman, H. Alarian, M. Rothleder, et al., "Enhanced system reliability using flexible ramp constraint in CAISO market," in IEEE Power and Energy Society General Meeting, 2012.
[28] C. Wu, G. Hug, and S. Kar, "Risk-Limiting Economic Dispatch for Electricity Markets with Flexible Ramping Products," IEEE Trans. Power Syst., vol. 31, no. 3, pp. 1990–2003, 2016.

[29] N. Navid, G. Rosenwald, “Ramp Capability Product Design for MISO Markets,” MISO, DRAFT, 2013.
[30] California ISO, “Flexible Ramping Product Draft Final Technical Appendix,” CAISO, 2016.
[31] J. Dobschinski et al., “Uncertainty Forecasting in a Nutshell: Prediction Models Designed to Prevent Significant Errors,” IEEE Power Energy Mag., vol. 15, no. 6, pp. 40–49, 2017.
[32] M. Poage, D. Tretheway, “Day-Ahead Market Enhancements,” CAISO, 2019. www.caiso.com/Documents/PresentationDayAheadMarketEnhancementsWorkingGroupMeeting81319.pdf .
[33] P. Etingov et al., “Balancing needs assessment using advanced probabilistic forecasts,” in 2018 International Conference on PMAPS 2018 – Proceedings.
[34] K. De Vos, N. Stevens, O. Devolder, et al., “Dynamic dimensioning approach for operating reserves: Proof of concept in Belgium,” Energy Policy, vol. 124, no. June 2018, pp. 272–285, 2019.
[35] Klein LJ, et al. (2015) PAIRS: A scalable geo-spatial data analytics platform. *IEEE Intl. Conf. on Big Data*, 1290-1298
[36] Lu S, et al. (2016) IBM PAIRS: Curated Big Data Service for Accelerated Geospatial Data Analytics and Discovery, *IEEE Intl. Conf. on Big Data*.
[37] Lu, S., and H.F. Hamann. "IBM PAIRS: Scalable Big Geospatial-Temporal Data and Analytics As-a-Service." *Handbook of Big Geospatial Data*. Springer, Cham, 2021. 3-34.
[38] Klein, Levente J., et al. [35] op. cit..
[39] Andrews, R.W., et al. "Introduction to the open source PV LIB for python Photovoltaic system modelling package." *2014 IEEE 40th photovoltaic specialist conference (PVSC)*. IEEE, 2014.
[40] Lu, Siyuan, et al. "Machine learning based multi-physical-model blending for enhancing renewable energy forecast-improvement via situation dependent error correction." *2015 European control conference (ECC)*. IEEE, 2015.
[41] https://rapidrefresh.noaa.gov/hrrr/ - accessed on 2/21/2022
[42] www.ncei.noaa.gov/products/weather-climate-models/north-american-mesoscale - acc. 2/21/2022
[43] www.ncei.noaa.gov/access/crn - accessed 2/21/2022
[44] https://geoserver.org - accessed 2/21/2022
[45] Thomas, S. M., et al. "Comparison of NOAA's operational AVHRR-derived cloud amount to other satellite-derived cloud climatologies." *Journal of climate* 17.24 (2004): 4805-4822.
[46] www.goes-r.gov/multimedia/dataAndImageryImagesGoes-16.html
[47] Alom, M.Z., et al. "The history began from AlexNet: A comprehensive survey on deep learning approaches." *arXiv preprint arXiv:1803.01164* (2018).
[48] www.ncei.noaa.gov/products/land-based-station/integrated-surface-database - accessed 2/21/2022
[49] https://raws.dri.edu/ - accessed 2/21/20
[50] B.F. Hobbs, V. Krishnan, J. Zhang, et al., “Development of Flexible Ramp Product Procurement for the California ISO using Probabilistic Solar Power Forecasts”, *Solar Energy Advances*, in review.
[51] Overpeck, J.T., et al. "Climate data challenges in the 21st century." *Science* 331.6018 (2011): 700-702.
[52] NERC, “Reliability Standards for the Bulk Electric Systems of North America,” North American Electric Reliability Corporation, Feb. 29, 2016.
[53] EPEX SPOT, “XBID: Cross-Border Intraday Market Project.” www.epexspot.com/en/market-coupling/xbid_cross_border_intraday_market_project
[54] E. Hirst and B. Kirby, “Separating and measuring the regulation and load-following ancillary services,” Util. Policy, vol. 8, no. 2, pp. 75–81, 1999.
[55] E. Ela, M. Milligan, and M. O’Malley, “A flexible power system operations simulation model for assessing wind integration,” in IEEE Power and Energy Society General Meeting, 2011.
[56] B.-M. Hodge, C. Martinez-Anido, et al., “The combined value of wind and solar power forecasting improvements and electricity storage,” *Appl. Energy*, vol. 214, Mar. 2018, pp. 1–15, 2018.
[57] Dept. of Market Monitoring CAISO, “2017 Annual Report on Market Issues & Performance,” 2018.
[58] B.F. Hobbs et al., “Coordinated Ramping Product and Regulation Reserve Procurements in CAISO and MISO Using Multi-Scale Probabilistic Solar Power Forecasts Introduction,” 2018. www.energy.gov/sites/prod/files/2018/10/f56/Solar-Forecasting-2-Kickoff-JHU.pdf .
[59] B. F. Hobbs et al., “Coordinated Ramping Product and Regulation Reserve Procurements in CAISO and MISO using Multi-Scale Probabilistic Solar Power Forecasts (Pro2R),” 2019. www.energy.gov/sites/prod/files/2019/10/f67/9 Solar-Forecasting-2-Annual-Review_The-Johns-Hopkins-University.pdf
[60] “CAISO Open Access Same-time Information System (OASIS).” https://oasis.caiso.com
[61] E. Spyrou et al., "What Is the Value of Alternative Methods for Estimating Ramping Needs?," *2020 IEEE Green Tech. Conference*, 2020, pp. 159-164, doi: 10.1109/GreenTech46478.2020.9289786.