Solar Energy Advances 2 (2022) 100024

Contents lists available at ScienceDirect

Solar Energy Advances

journal homepage: www.elsevier.com/locate/seja

Using probabilistic solar power forecasts to inform flexible ramp product procurement for the California ISO

Benjamin F. Hobbs[a,e,*], Jie Zhang[b], Hendrik F. Hamann[c], Carlo Siebenschuh[c], Rui Zhang[c], Binghui Li[b], Ibrahim Krad[d], Venkat Krishnan[d], Evangelia Spyrou[d], Yijiao Wang[a], Qingyu Xu[a], Shu Zhang[a]

[a] Robert O'Connor Sustainable Energy Institute, Johns Hopkins University, 3400 North Charles St., Baltimore, MD 21218 USA
[b] Department of Mechanical Engineering, The University of Texas at Dallas, 800 West Campbell Rd, Richardson, TX 75080 USA
[c] IBM Thomas J. Watson Research Center, 1101 Kitchawan Rd, Yorktown Heights, NY 10598 USA
[d] Energy Systems Integration Directorate, National Renewable Energy Laboratory, Golden, CO 80401 USA
[e] Market Surveillance Committee of the California Independent System Operator, 250 Outcropping Way, Folsom, CA 95630 USA

ARTICLE INFO

Keywords:
Probabilistic solar power forecasts
Flexible ramp product requirements
California
Production cost
Reliability
Machine learning

ABSTRACT

How can independent system operators (ISOs) take advantage of probabilistic solar forecasts to lower generation costs and improve reliability of power systems? We discuss one three-step approach for doing so, focusing on how such forecasts might help the California Independent System Operator (CAISO) prepare unexpected net load ramps, where net load equals gross demand minus wind and solar production. First, we enhance an existing solar forecasting system to provide well-calibrated hours-ahead probabilistic forecasts. We then relate the degree of uncertainty reflected in the forecasted prediction intervals (independent variables) to error distributions for net load ramp forecasts for the CAISO real-time market (dependent variable) using machine learning and quantile regression. Projected ramp forecast 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. Detailed descriptions are provided on the quantile regression and kth-nearest neighbor categorization methods for accomplishing that translation. Finally, a multiple time-scale look-ahead market simulation model is applied to a 118-bus IEEE Reliability Test System, modified to represent the CAISO generation mix and demand distributions. The model runs quantify how solar-conditioned ramp requirements can, first, decrease operating costs by reducing requirements compared to often conservative unconditional methods and, second, decrease generation scarcity events and consequently improve reliability by increasing flexibility requirements at times when unconditional forecast-based requirements understate actual ramp uncertainty. Solar-conditioned ramp requirements are found to reduce generation operating costs by about 2% for the test system (which would be equivalent to over $100 million per year for a CAISO-size system).

1. Introduction

Variable renewable generation is rapidly expanding in California and elsewhere, both in the form of behind-the-meter rooftop solar, and as grid-scale development of wind and solar resources [1]. As a result, the short-term variability in system net load that must be met by dispatchable thermal, hydro, and storage resources is also growing quickly [2]. It is widely recognized that improvements in solar forecasts can significantly reduce operating costs, improve system reliability, and even save capital costs in the long run [3]. Here, we focus on the value of a particular enhancement to solar forecasting technology: the development of probabilistic forecasts that not only predict a median or expected amount of solar insolation or power production, but also characterize the uncertainty around that central value. In particular, we describe a new approach for using probabilistic solar forecasts to improve the definition of requirements for operating reserves, and illustrate it with an application to procurement of flexible ramp product for the CAISO system.

The CAISO and other North American ISOs buy operating reserves of various types to accommodate this variability for the California market on various time scales [4]. Among these reserves are 30-minute replacement reserves, 10-minute spinning and non-spinning reserves, regulation-up, and regulation-down, with the last two handling unexpected variations on a sub-5-minute scale. These are acquired in the day-ahead CAISO market, with adjustments possible in its real-time market

* Corresponding author.
E-mail address: bhobbs@jhu.edu (B.F. Hobbs).

https://doi.org/10.1016/j.seja.2022.100024
Received 2 May 2022; Received in revised form 10 September 2022; Accepted 12 September 2022
Available online 15 September 2022

(which is part of the western US- and Canada-wide Energy Imbalance Market (EIM), whose other entities only procure or sell real-time energy in the EIM). The real-time market includes two separate but linked markets, one with 15-minute intervals that can allow commitment of short-start generation units, and a subsequent dispatch-only market with 5-minute intervals.

In addition to the operating reserves mentioned above, the CAISO has developed a new real-time product in 2016 called the *flexible ramp product* (FRP). FRP, which is the focus of our application, is designed to position supply and storage resources in each time interval of the EIM to feasibly accommodate unexpected deviations in net load ramp from that interval to subsequent intervals either in the downward or upward direction (designated "FRD" and "FRU," respectively), and to compensate resources for any resulting foregone energy revenues [5–7]. Finally, the CAISO is proposing a day-ahead product called the imbalance reserve product that will not only procure resources that could be needed to meet real-time FRP needs, but also accommodate deviations between day-ahead and real-time net load forecasts with a predetermined reliability of 95% [8].

Operating reserves can be expensive to procure, amounting to almost $200 million for the CAISO in 2020 (out of the total energy and ancillary services cost of $8 billion), or about $1/MWh [9]. Consequently, ISO engineers and the research community have paid significant attention to the general issues of what types, amounts, and locations of reserves to procure. Optimizing those procurement decisions must strike a balance between the costs of reserving capacity (e.g., costs of contracting fuel ahead of time in case it is needed, and the wear-and-tear and fuel cost of keeping extra capacity on stand-by), and the benefits in terms of improved reliability (avoiding voluntary or involuntary curtailment of load) and reduced fuel and other variable costs of expensive peaking generation that would otherwise be called upon during periods of scarcity [10].

A crucial input to calculating the desired amount of reserves is the degree of uncertainty in net load. This uncertainty depends upon time of day, season, and especially meteorological uncertainty that in turn translates into uncertainty in solar insolation, wind, stream flows, and temperature-dependent loads. Probabilistic forecasts of these quantities on appropriate timescales (e.g., 12-36 hours ahead for day-ahead power markets and minutes to hours ahead for real-time markets) can inform estimates of the magnitude of net load uncertainty that needs to be covered by reserves (e.g., [11]). Li and Zhang [12] summarize several approaches to using probabilistic forecasts to inform operating reserve procurement. The most common one is for "situational awareness", in which system operators informally consult these forecasts when adjusting reserve requirements. More formal "dynamic reserve" approaches that are under consideration use statistical models, such as quantile regression, to condition net load uncertainty on weather conditions or the amounts of variable renewables, and then to set the reserve requirements based on a maximum tolerable probability of net load exceeding the reserves [13]. Due to the exceptionally rapid growth in solar generation, increasing attention is being paid to using probabilistic solar forecasts to inform operating reserve procurement, but there is a lack of systematic methods to integrate such forecasts into operations and scheduling routines [12].

Here we focus on the potential for using probabilistic solar forecasts to set ramp product requirements for the CAISO system, which has the largest solar penetration of any U.S. ISO-based market.[1] Presently, in order to decide on how much of the two types of ramp products (FRD and FRU) to procure, the CAISO considers the distribution of ramp forecast errors in the relevant time interval over the previous few weeks. By creating a histogram of those errors, the 2.5% and 97.5% percentiles can be estimated; the CAISO uses these percentiles to define the MW of "uncertainty component" in the down- and up-directions to be acquired. The net down- and up-requirements for ramp in a given period are then defined as the expected ramp forecast adjusted down and up to account for those two uncertainty components, respectively.

Fig. 1, adapted from CAISO documentation, illustrates the definition of FRP for the 15-minute EIM market. The *y*-axis shows the MW quantity of energy and flexible capacity to be provided by the market, and the *x*-axis is time. Supply, storage, and demand-side resources submit bids (MW quantity and $/MWh prices) to a given market run 52.5 minutes prior to the start of the so-called "binding" interval (covering time t to t+15), for which financially binding schedules for those resources will be optimized and settled at the prices calculated by the market process. The optimization simultaneously optimizes short-start unit commitment and dispatch for the binding interval along with several subsequent "advisory" 15-minute intervals. Advisory schedules are not financially binding, but serve a "look ahead" purpose to ensure that resources are positioned in the binding interval so that net loads in later intervals are efficiently met. The solid line in the figure shows the forecast made at t-52.5 of energy demand in each interval (here, the binding interval plus the subsequent three advisory intervals). Finally, the figure shows, as dashed lines, how much resource capacity needs to be scheduled to meet FRU and FRD requirements in the advisory intervals. For example, the type and amount of resources scheduled in the binding interval [t,t+15] have to be able to feasibly move over 15 minutes to meet the forecast net load in the first advisory interval [t+15,t+30] plus the 97.5^{th} percentile error (the upper dashed line) in the case of FRU, and minus the 2.5^{th} percentile error (the lower dashed line) in the case of FRD. This constraint is also imposed in each advisory interval, with that interval's resource schedule having sufficient flexibility to meet the range of possible ramps defined by those errors in the next advisory interval, as illustrated.

The requirements for ramp product are enforced by soft constraints in the CAISO real-time market's resource optimization software, in the form of demand curves for MW of FRD and FRU in each interval. These demand curves reflect the understanding that the incremental value of the product declines as more is procured, but that there is not a fixed requirement or threshold below which a product has a high value (reflected in a large violation penalty in the market software) and above which it has no value. Further, the CAISO recognizes [5], as does most of the industry [13], that defining requirements independent of information on weather and renewable energy production conditions will result, under some meteorological conditions, in overly conservative requirements well in excess of the actual need in a given day, thus inflating costs. Meanwhile, under other conditions when uncertainty is greater, procurement will be too little, exposing the system to risks of inadequate flexibility reserves and undesirably high probabilities of load balance violations. Since solar generation variability is a major source of net load uncertainty, it is logical to expect that forecasts of the net load ramp uncertainty components shown in Fig. 1 could be usefully conditioned on weather and renewable conditions [5], especially solar uncertainty.

In this paper, we summarize the procedures and some of the results of a research project directed at translating probabilistic solar forecasts into weather-conditioned projections of FRP needs for the CAISO system (for details, see [14]). Among these results are an updated Watt-Sun solar forecasting system, models linking probabilistic solar forecasts to system netload ramp uncertainty, and a comparison of system costs resulting from the revised requirements with the costs under the present CAISO unconditioned histogram approach for a IEEE 118-bus system modified to include a generation mix similar to the CAISO system (Fig. 2). The rest of the paper is organized into three parts: development of probabilistic solar forecasts from the Watt-Sun forecasting sys-

[1] See also [13], who propose alternative approaches for using probabilistic forecasts for reserves sizing for the CAISO, emphasizing approaches based on generating sets of scenarios of potential solar generation and then calculating implied reserve requirements. Our methodology also differs from the CAISO's proposed use of quantile regression to set FRP requirements [5] in our use of independent variables that represent the degree of solar uncertainty (e.g., the width of the central 50% prediction interval between the 25th and 75th percentiles of solar insolation), rather than just expected solar and wind output.

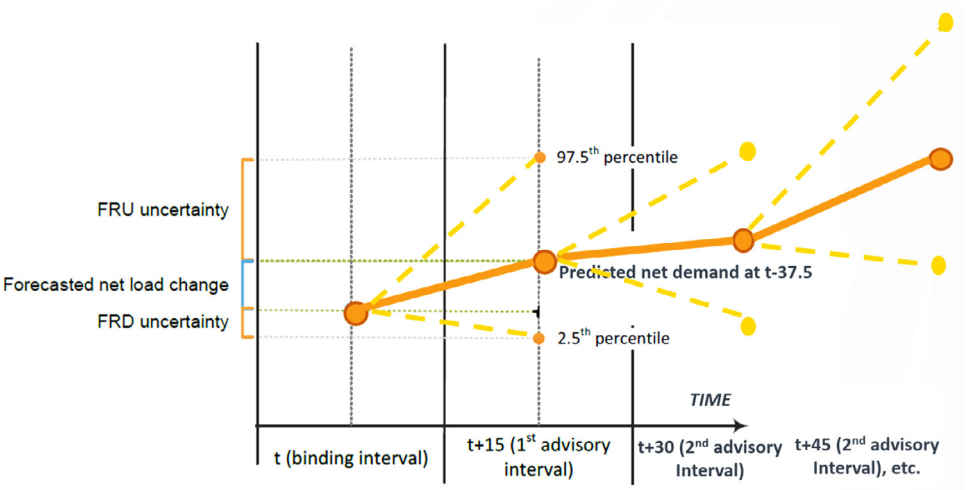

Fig. 1. Schematic of definition of the two components (uncertainty and forecast net load change) of CAISO flexible ramp product requirements (FRU = flexible ramp up, FRD = flexible ramp down), based on 2.5th and 97.5th percentiles of recent ramp forecast errors in the same time interval. Note that the requirements are defined for all intervals in the 15-minute real-time market (binding interval, and subsequent advisory intervals in the multi-interval optimization). (Source: Adapted by authors from [6]).

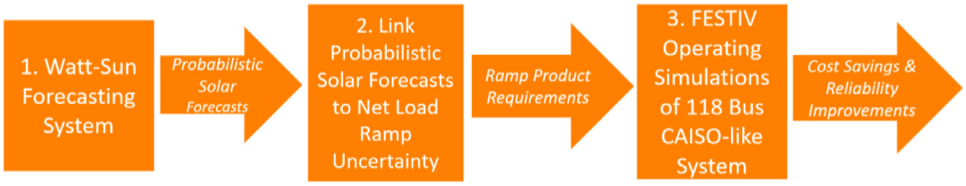

Fig. 2. Organization of analysis of solar forecast-based FRP requirements.

tem (Section 2); quantile regression- and machine learning-based modeling of the relationship of solar forecast uncertainty to forecast errors for CAISO real-time load ramps, yielding solar uncertainty-conditioned ramp requirements (Section 3); and production cost-based assessment of cost savings and reliability improvements for the test system (Section 4). That section carefully identifies the sources of cost savings that result from using solar uncertainty-informed requirements for ramp product. Finally, Section 5 concludes the paper.

2. Probabilistic solar forecasting

We have taken advantage of IBM's big-data platform Physical Analytics Integrated Repository and Services (PAIRS) [15,16] and quantile regression of forecast errors (one for each class of weather conditions identified by artificial intelligence [17]) to generalize the Watt-Sun solar forecasting system [15] to generate probabilistic forecasts. PAIRS curates terabytes of numerical weather prediction (NWP) models and historical data, including the high-resolution GOES-R imagery (NOAA L2 product's cloud optical depth). By blending multiple NWPs and imagery data using deep learning methods and quantile regressions to obtain a set of critical percentiles, we developed a big data-driven probabilistic forecasting system, whose flow chart is shown in Fig. 3. The system has been implemented for 10 observation sites each in the CAISO and Mid-Continent ISO footprints for global horizontal irradiance (GHI) forecasts. GHI represents the overall amount of shortwave solar radiation received at ground level. A prototype raster-based system using the GOES-R imagery has also been developed to create GHI irradiance forecasts on a 3 km pixel grid covering locations without direct observations.

Fig. 4 shows an example of Watt-Sun probabilistic forecasts of GHI, illustrating (from left to right) a sunny day, a partly cloudy day, and two very cloudy days. The graphs show that, in this case, the sunnier days not only have a higher median (green line), but also less uncertainty (measured by the width of the prediction interval between, for instance, the 25th and 75th percentiles, shown in orange and red, respectively).

Based on P-P plot scores, the Watt-Sun probabilistic forecasts are better calibrated than baseline persistence and High-Resolution Rapid Refresh (HRRR) bias-corrected forecasts (Fig. 5). A P-P score measures the total deviation (mean absolute value) of a plot of predicted versus empirical error cumulative distributions from the perfect (45°) calibration line (the total area between the two lines in each figure).

Figs. 6 and 7 provide further evidence of the quality of calibration of probabilistic Watt-Sun. Fig. 6 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 (Chi-squared) 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 a strong assumption.) The plot shows how the prediction intervals change from day-to-day; for instance, Dec. 1-8 were cloudy days with more solar uncertainty (lower medians and wider intervals) while Dec. 9-20 were sunnier (higher medians, narrower intervals). Plots like Fig. 7, which is an example binning of the observed GHI for all daylight hours in one month (red) compared to the 5%/20%/25%/25%/20%/5% expected frequencies, provide visual evidence of a good calibration, as its χ^2 value that indicates that the two distributions are not significantly different.

3. Using probabilistic solar forecasts to create solar-conditioned requirements for flexible ramp product

3.1. Weather dependence of net load ramps

As stated above, the CAISO introduced ramp products in its real-time markets that procure generation "ramping" capacity so that potential net load ramp uncertainty can be managed feasibly. The purpose is to reduce the frequency of generation scarcity events and real-time price spikes. Specifically, the up- and down-FRP (FRU and FRD, respectively) address both expected ramps from one real-time interval to the next, plus an uncertainty component representing possible positive and negative errors, respectively, in net load forecasts. Presently, to determine the uncertainty component, the CAISO uses histograms of net load fore-

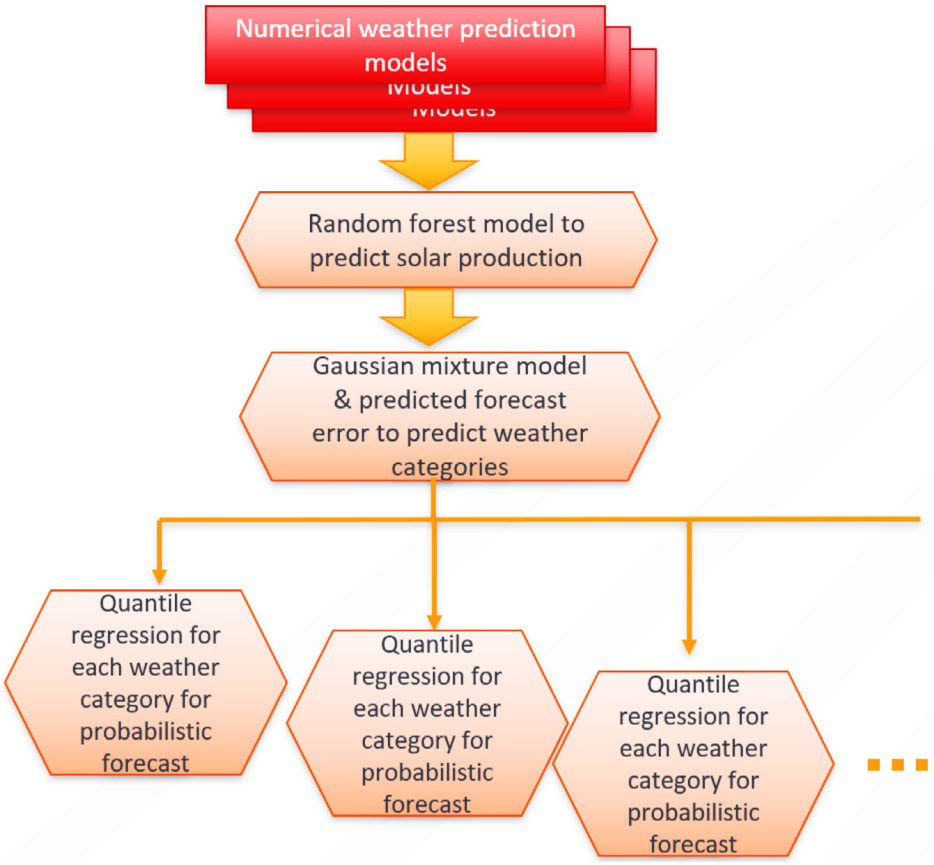

Fig. 3. Flow chart showing probabilistic Watt-Sun development of probabilistic forecasts of solar irradiance (GHI).

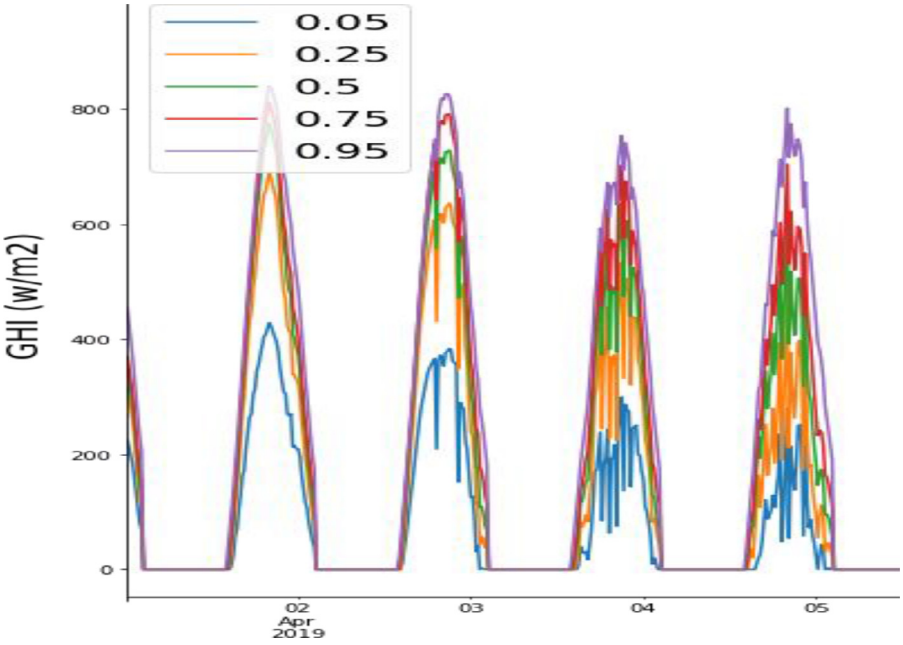

Fig. 4. Site-specific probabilistic global horizontal irradiance (GHI) forecasts (April 2-5, 2019, Topaz California solar site, 5th, 25th, 50th, 75th, and 95th percentiles), Probabilistic Watt-Sun 1.0 system.

cast errors on an hourly basis, considering all errors in that hour from the previous 40 days (if the day is a weekday) or 20 days (if the day is a weekend) [6]. After the construction of histograms, the upper and lower bounds of the 95% confidence interval are used as the up- and down-FRP requirements, respectively, for that hour. The 95% confidence interval for accounting for uncertainties is a well-accepted industry standard that strikes a balance between reliability and economics. We term these "unconditional" or "solar independent" FRP requirements.

A reasonable expectation is that the forecast error tends to be greater during a partially cloudy day, while being lower in a sunny or completely overcast day. However, by constructing histograms of forecast errors purely from historical data, the CAISO's unconditional method does not reflect the latest weather information and often leads to overestimation of ramp uncertainty under sunny weather conditions and underestimation under partly cloudy conditions. As clearly depicted in Fig. 8, below, CAISO's requirements in two close days (8/7/19 and 8/12/19) are

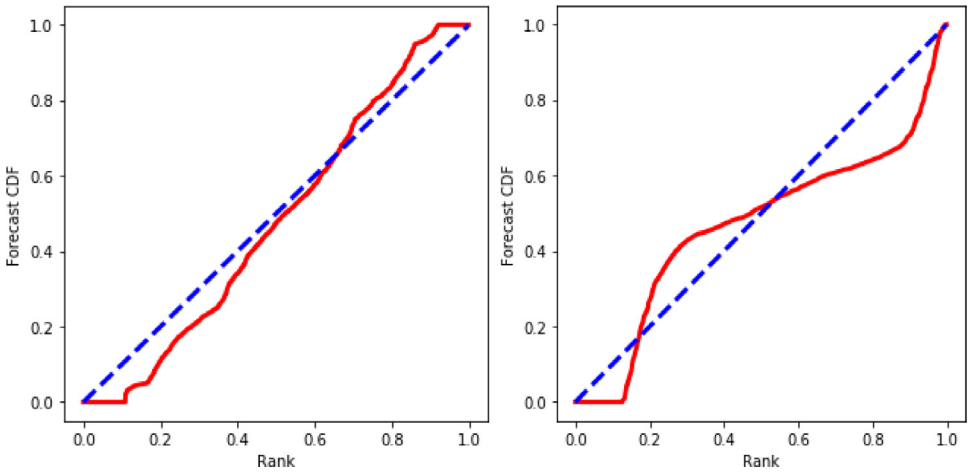

Fig. 5. Example P-P for Probabilistic Watt-Sun 1.0 forecasts (P-P Plot score=0.054) and HRRR bias-corrected forecasts from High-Resolution Rapid Refresh (HRRR) system [18] (P-P Plot score=0.086), May 2019, Topaz California site.

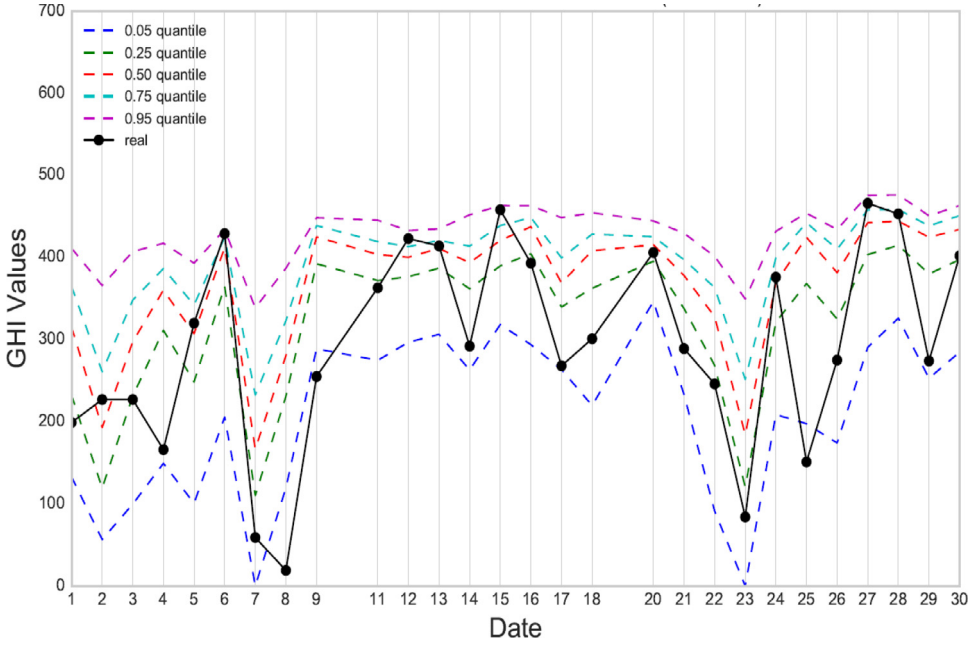

Fig. 6. Data for χ^2 calibration test. Dec. 2019 actual GHI and Watt-Sun probabilistic forecast quantiles, 1:30 local time, Topaz site (Note, Dec. 10, 19, 31 values missing).

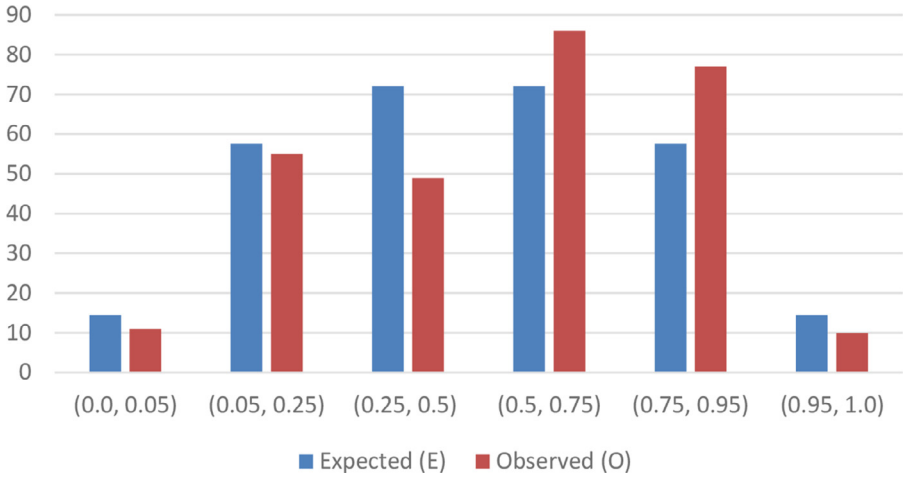

Fig. 7. One month's expected number of expected GHI observations by bin (fractile ranges) versus observed (daylight hours only), illustrating quality of calibration.

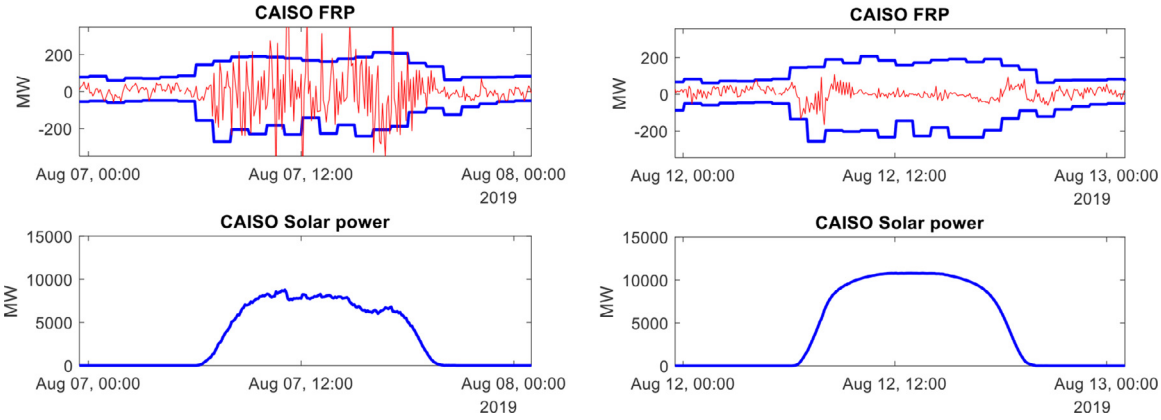

Fig. 8. Top: realized net load forecast errors (red) compared with the uncertainty components of CAISO's estimated FRP requirements (upper blue line is FRU, lower blue line is FRD) from two days in August 2019 (cloudy (left) and sunny (right)). The bottom curves show the CAISO's real-time forecasts of system-wide solar power production, consistent with those types of days.

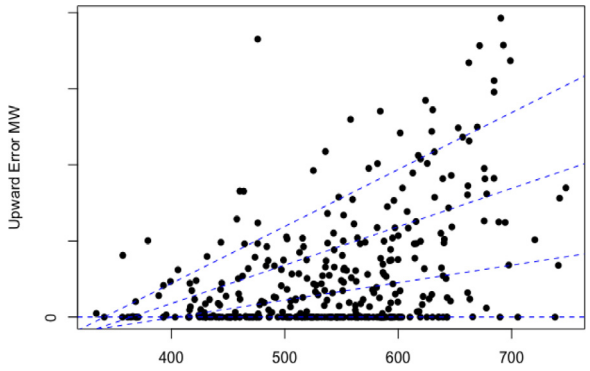

Fig. 9. Quantile regression results (11 a.m.-2 p.m. May 2019) for upward (over-forecast) CAISO errors in real-time load forecasts (*y*-axis, normalized scale), as function of GHI 50% confidence interval width (*x* axis). From top, the blue dashed lines are the results for linear quantile regression estimates of the 90th, 75th, 50th, and 25th percentiles of load forecast errors.

almost identical since the histograms are similar. However, as the blue lines in the figure show, the weather conditions in these two days are totally different: the bottom of the left figure presents a lower and fluctuating profile of system-wide solar power, suggesting relatively cloudy weather, while the lower right profile shows a smooth and stable output, consistent with sunny conditions. Consequently, the need for FRP is very different on those two days: greater uncertainties in cloudy days should contribute to higher FRP requirements, while sunny weather should be a reason for lower FRP requirements due to reduced uncertainty. If the same amount of FRP is procured in both days, the cloudy day might have a high probability of FRP shortage, and hence a risk of capacity shortages, while the market efficiency would be reduced in the sunny day due to over-procurement.

In order to improve upon current industry practice of weather-unconditioned reserve requirements, we used quantile regression and a machine learning method (kth nearest neighbor (kNN) classification) to relate the 95% confidence intervals in net load forecast errors to uncertainty in forecasted GHI in order to create solar-conditioned FRP requirements. These methods are summarized in the next two subsections. Fig. 9 illustrates that there is a strong dependence of up-ramp uncertainty (positive errors, shown on the y-axis)[2] on the width of the 25th-75th percentile prediction interval (shown on the x-axis), the latter derived from the Watt-Sun probabilistic forecasting system (from solar forecast data such as Figs. 4 and 6).

3.2. Using quantile regression (QR) to link solar forecasts to FRP needs

The QR method [19], which we used to obtain the relationships shown as dashed lines in Fig. 9, can be viewed as an advanced version of histogram method used by the CAISO. QR boosts the prediction performance of the baseline histogram method by integrating the latest weather information, including but not limited to probabilistic solar forecasts, into estimates of flexible ramping requirements that the power grid needs to address net load uncertainty. Specifically, in contrast to the baseline method where a histogram of unconditional historical net load forecast errors are referred to when calculating ramp needs, the weather-informed QR method relies on an estimated posterior distribution of net load forecast errors given the inputs of meteorological and other variables (especially the forecast probability distribution of solar irradiance) at a future interval. The posterior distribution of net load errors is thus conditioned on the information input (for instance, if the value of the x-axis in Fig. 9 is relatively high, i.e., a relatively large value of the width of the prediction interval defined by the 25th-75th percentiles of GHI), as the blue dashed lines in Fig. 9 show.

Mathematically, QR estimates the vector of coefficients β_q for each line in Fig. 9 (one line for each quantile q considered) by choosing β_q that minimizes Eqn. (1), below. Define y_i as the value of the independent variable for the i-th observation (the forecast error for net ramp, as explained in Footnote 2, above, or just its positive component, as in the case of upward errors as in Fig. 9), and let x_i be the vector of independent variables corresponding to observation i. For instance, in a multivariate QR with two input variables, we could have $x_{i0} = 1$ to allow for an intercept, x_{i1} = median predicted solar GHI, and x_{i2} = width of the 25th-75th percentiles-based prediction interval for GHI. Linear programming is then used to accomplish the minimization shown below:

[2] The upward ramp forecast error for a given interval is calculated by the following procedure which is unique to the CAISO. Consider the ramp forecast between intervals [t,t+15] and [t+15,t+30] in Fig. 1. The upward ramp error is the difference between 1) the highest binding interval forecast from the 5-minute market among the three 5-minute intervals in period [t+15,t+30] (not shown in Fig. 1) and 2) the first advisory interval forecast made for the 15-minute interval [t+15,t+30] in the previous 15-minute market run (i.e., whose binding interval was [t,t+15]), shown as the second orange dot on the solid orange line in Fig. 1. Thus, this is the difference between two forecasts, and not between a forecast and actual net load. A positive error indicates that the forecast closer in time to actual operations (the binding interval forecast) is higher than was anticipated in the earlier market run, and so the system has to meet a steeper ramp in net load than anticipated.

$$\underset{\beta_q}{MIN} \sum_i q\, MAX\left(0, y_i - \beta_q x_i\right) + \sum_i (1-q) MAX\left(0, -y_i + \beta_q x_i\right) \quad (1)$$

In words, one starts with an assumed value of a quantile q between 0 and 1 (e.g., q=0.9 would yield the coefficients $\beta_{0.9}$ defining the relationship between the 90th percentile and the input variables, which is the top blue dashed line in Fig. 9). The first term in (1) is the sum of the differences (weighted by q) between the observed value of y_i and the predicted value $\beta_q\, x_i$ only for i in which the observed value exceeds what is predicted. The second term is instead 1-q times the sum of those differences for i where instead the observed value is below the predicted value. By weighing the first term $q/(1-q)$ times as much as the second term, QR attempts to find the relationship (a line in the case of Fig. 9, or, more generally, the hyperplane) that results in fraction q of the observations lying below the relationship.

In this study, load and weather forecast variables (also known as classifiers) are used as regressors in our QR models to predict distributions of upward and downward net load errors at future intervals. Since, the CAISO is interested in setting the FRP down and up requirements at the 2.5th and 97.5th percentiles, respectively, in theory we could do two QR regressions, one for $q = 0.025$ and one for $q = 0.975$. However, when a relatively small sample is used, such as the last 30 days, the resulting relationships are unstable because few observations occur in these tails. For this reason, we use more stable QR relationships obtained for intermediate rather than extreme values of q, and then assume normality to extrapolate tail percentiles. Thus, in the case of upward errors our QR relationships are estimated for q =0.5 and 0.9. Then a normal distribution is calibrated to those values (by obtaining the implied mean and standard deviation), and then the desired 97.5th percentile is extrapolated using that distribution. For downward errors, instead q=0.1 and 0.5 are used in the same fashion, with the 2.5th percentile estimated by calibrating a normal distribution to those percentiles.

The independent variables we used in the QR are constructed as follows. First, two solar variables are considered. For a selection of up to 10 solar sites across the state of California, 2 hour-ahead probabilistic GHI forecasts over selected sites with 15-min resolution are used to calculate GHI_m and GHI_w. The former denotes the average (across sites) of the 50th percentile (median) of probabilistic forecasts of GHI. Its value is of particular interest to the procurement of FRP, as an inverse-U relationship between net load uncertainty and median GHI_m is revealed by our analysis of the data: it is observed that net load forecasts errors of a particular time interval are smaller when the value of GHI_m of the forecast interval is near its high or low extremes. Meanwhile, net load uncertainty tends to be highest when GHI_m is intermediate. The second input, GHI_w is defined as the average (over sites considered) of the 25th-75th prediction interval width, and measures the 2-hour-ahead uncertainty in ground-level solar radiation within a given 15-min interval. As Figs. 8 and 9 indicate, solar uncertainty is a useful predictor to determine the amount of FRP procurement. In addition to probabilistic solar forecasts, 15-min wind forecast and load forecasts from the real-time CAISO energy market are other two independent variables we considered for to estimate FRP needs, as those variables are linearly related to net load (i.e., net load is total load minus wind and solar power).

To summarize, the following steps are used to estimate and apply QR models predict FRP-up requirements. (Analogous steps are used for FRP-down requirements.)

1 *Training Data.* For a given day, training data (net load forecast errors y_i and independent variables x_i, as defined above) are pooled by hour of the day for the previous N days of the same type (e.g., we considered $N = 30$ in the case of weekdays). Thus, the models are updated daily in a rolling fashion.
2 *Training.* QR relationships are estimated by solving Eqn. (1) for q = 0.5 and 0.9, yielding two relationships for each daylight hour.
3 *Determine the FRP Requirements.* FRP-up requirements for a given hour are estimated by inserted the forecast values of x_i for that hour in the two estimated QR relationships, yielding estimated 50th and 90th percentiles for net load error in the upward direction. From those values, the 97.5th percentile is extrapolated by assuming a normal distribution for the net load forecast error. The latter percentile is the amount of FRP that can then be used to define the FRP requirement in the fifteen-minute market optimizations for that hour. For simplicity, our FESTIV-based simulations in Section 4 assume fixed requirements that can be relaxed at a fixed per MW penalty; in reality, the CAISO defines a "demand curve" for FRP that has an increasing marginal penalty for larger deviations from the requirement.

3.3. Using kth-nearest neighbor classification to link solar forecasts to FRP needs

In addition to the QR method illustrated in Fig. 9, various specifications of a machine learning approach (based on the kth-nearest neighbor (kNN) classification method [20]) were also tested with different combinations of GHI variables at various solar generation sites in the CAISO. Examples of GHI variables considered include median, 50% confidence interval width, and volatility in 50% prediction intervals from 15-minute to 15-minute market interval. The kNN-based method is a non-parametric classification approach, and it can be viewed as a direct extension of the CAISO's original implementation, since both methods rely on historical data. In both methods, the upward and downward FRP requirements at a certain time interval are given by the predictive distributions of errors in upward and downward net load forecasts, respectively. The difference, however, is that the kNN-based method constructs posterior (weather-conditioned) histograms by using probabilistic solar forecasts.

A brief summary of our application of the kNN-based method to FRP requirements estimation is given below, and more details can be found in [21].

1 *Characterize the weather conditions*: Given probabilistic solar irradiance forecasts during a time interval, the kNN-based method uses a set of numerical classifiers to characterize the weather conditions during that interval.
2 *Find similar historical days and construct weather-conditioned distributions of net load forecast errors*: We then find historical intervals with similar weather conditions to the target interval, where the similarity of a pair of time intervals is represented by Euclidean distances between two numerical classifiers. After sufficient similar historical days are identified, we construct histograms using net load forecast errors from these days and then fit cumulative density functions for upward and downward errors.
3 *Determine the FRP requirements*: The kNN-based method then determines the up- and down-FRP requirements based on the 95% confidence intervals of the cumulative density functions.

The performance of the kNN-based method relies heavily on the numerical classifier and the number of closest neighbors. Our study suggests that performance is also sensitive to the size of validation set, which consists of N previous days [21]. Fig. 10 compares the out-of-sample performance of kNN-based models that estimate ramp requirements relative to the unconditional method for February 2020, which are trained with data starting in January 2020. We use the probability of FRP shortage to assess the reliability levels of the FRP requirements, which measures the frequency of actual net load forecast errors exceeding the FRP requirements during the evaluated period. The amounts of oversupply are calculated by summing up excess FRP requirements over actual net load forecast errors over the evaluated period, and are used to measure the economic performance.[3] The figure displays trade-offs between reliability levels and oversupplies in the form of Pareto frontiers

[3] Specifically, "oversupply" for any particular 15 minute real-time market interval is measured by the amount that the FRU requirement procured in interval

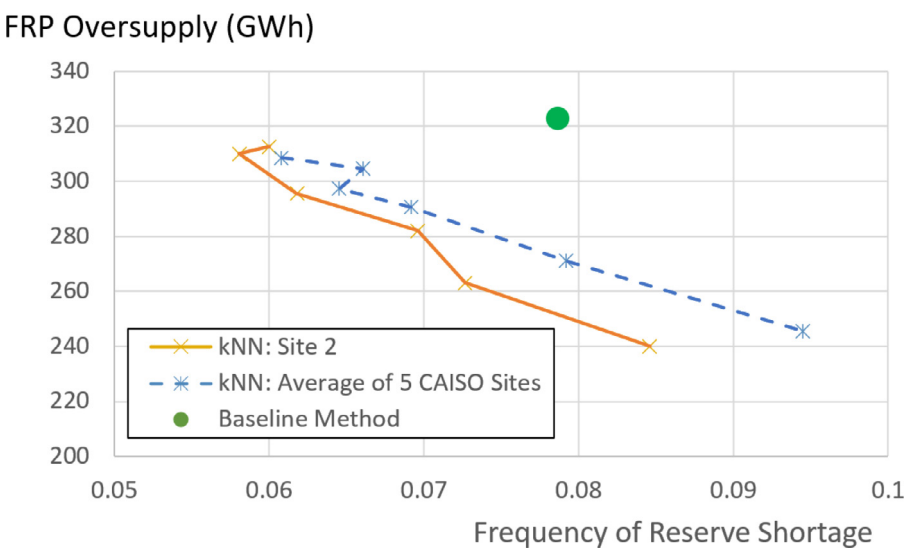

Fig. 10. Pareto diagram showing performance of two sets of kNN-based machine learning requirements for FRU for February 2020, based on probabilistic solar forecasting data from each of five solar production sites in the CAISO in January 2020. Performance is compared to the CAISO baseline histogram method, which is not conditioned on weather. The markets represent performance under alternative estimations from more conservative (upper left, showing more requirements and lower likelihood of reserve shortage, resulting from using larger validation sets, starting from N=30 previous days at the extreme upper left) to less conservative (lower right, resulting from using smaller validation sets, starting from N=5 at the extreme lower point).

for the case of 2-dimensional classifiers as we vary the size of validation sets.

The circular point in Fig. 10 represents the baseline set of requirements using the CAISO's unconditional histogram method. Points lying to the southwest of the baseline indicate improvements (reductions) in both shortage of reserves (an index of reliability, y axis) and oversupply of FRP (x axis), hence these are better solutions than the original implementation. The colored lines represent the performance of two machine-learning based models derived from data from either one or five solar production sites in the CAISO. A set of points results from adjusting the parameters of each model to yield more (upper left) to less (lower right) conservative requirements. By "conservative", we mean larger MW flexible ramp-up requirements and therefore lower probability of actual ramp errors lying outside the requirement, but higher ramp product procurement cost.

In Fig. 10, most of the kNN points fall southwest of the baseline method, suggesting that the solar forecasting-based method can both reduce average procurement requirements while improving average reliability levels. Therefore, the user can tune the method to emphasize either FRP oversupply reduction or reliability improvements. For instance, the figure shows that the best of the two Pareto curves (Site 2) can reduce FRP oversupply to meet the present reliability performance (just under 8%, as shown on the x axis) by a fifth (from 325 to 250 MW, the difference between the baseline and the orange curve, measured at 8% reliability). Alternatively, oversupply could be maintained at about 320 MW while reducing the frequency of reserve shortage from 8% to 6%.

Ultimately the solar forecast-based requirements used in the simulations of the next section consisted of the kNN results based on the latest weather information. To select the optimal parameters for the kNN-based method–i.e., classifiers and the number of nearest neighbours–we designed a multi-objective optimization algorithm to dynamically select the parameters [21].

Fig. 11 gives an example of a set of FRU (left) and FRD (up) requirements for the CAISO resulting from this method (blue) that were used in the simulation of the next section, contrasting them to the FRU and FRD requirements respectively based upon the CAISO baseline unconditional histogram method (red) (downloaded from [22]). By definition, the requirements differed only during daylight hours, because those are the only times with solar forecast information. The results show that for this particular day, the solar-informed method increases ramp up requirements relative to the baseline, especially during late afternoon hours, in order to enhance system reliability, while ramp down requirements are reduced through most of the day. This suggests that the probabilistic solar forecasts are consistent with greater risks of decreased solar production later in the day, but show smaller than typical risks of increased production throughout daylight hours. We next describe our analysis of the implications of these changed requirements for system production costs and FRP procurement costs.

4. Production cost assessment of benefits of solar forecast-based flexible ramp product requirements

Our simulation of the benefits of solar-conditioned FRP requirements considered the IEEE 118-bus reliability test system [23,24] with alterations to reflect California ISO generation mix and demand conditions. In this application, we used requirements developed using the kNN method of Section 3.3. The size of the modified 118-bus system is approximately one-tenth of the CAISO in terms of load, and was altered to reflect 2017 renewable penetration (solar penetration of 1.2 GW, providing about 10% of the system's annual energy). The test system also included steam units (0.9 GW), combustion turbines (1.8 GW), combined cycle plants (2.4 GW), hydro plants (1 GW), and wind (0.6 GW). The detailed generator and network data can be found at https://github.com/NREL/FESTIV_MODEL/tree/master/Input/SF2_JHUprime_input/. Locations of wind and solar facilities in the 118-bus system are shown in Fig. 12. Wind and solar capacity factors are based on CAISO data. Due to data and high-performance computing facility availability, we focused on operations during a few selected days (March 9-30, 2020).

We considered three scenarios in this analysis to identify the benefits of better FRP requirements. Scenario 1 is the 'baseline' simulation that is based the current unconditional FRP methodology, using available historical data on ramp requirements from the CAISO (e.g., red lines, Fig. 11). Scenario 2 is the 'conditional FRP' scenario in which FRP requirements are updated based on the kNN method summarized in Section 3.3 to develop more efficient ramping requirements based on probabilistic solar power forecasts (e.g., blue lines, Fig. 11). The differences in production costs between Scenarios 1 and 2 describe the impact of using solar-conditioned FRP. Scenario 3 is the 'perfect' forecast scenario which models system operations if the operator has perfect knowledge of the system (i.e., zero netload forecast errors). The difference between Scenario 3 and any of the other scenarios' production cost gives a quantitative measure of "*uncertainty induced*" costs. Those are the increases in costs that occur under uncertainty, which are the expenses of

[t,t+15] (dot at end of first upper dashed line, Fig. 1) is in excess of the amount actually needed to meet the binding interval net load forecast in the following interval [t+15,t+30].

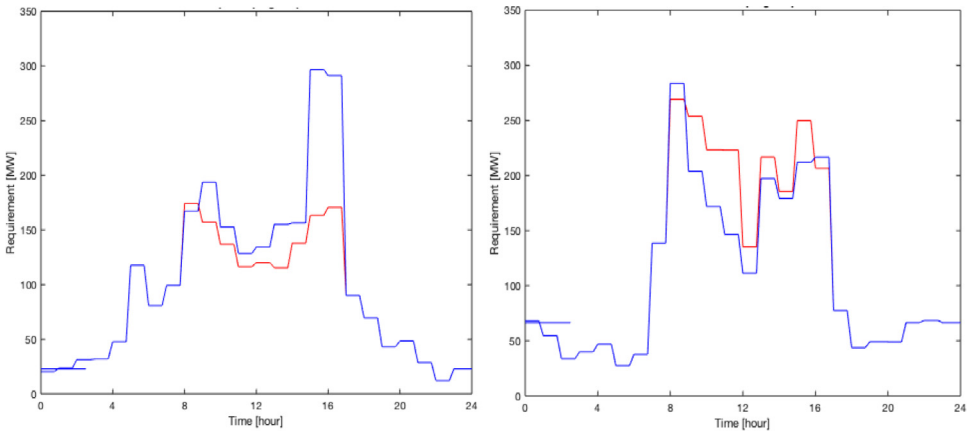

Fig. 11. Flexible ramp requirements for example day during March 2020 from CAISO baseline unconditional method (red) versus kNN method (blue) for including solar uncertainty information in requirements. Left: FRU requirements; right: FRD requirements.

Table 1
Summary of FESTIV simulation results for CAISO-like IEEE 118 Bus System, March 9-30, 2020.

Index\Scenario:	1 (Baseline)	2 (Solar-Informed FRP)	3 (Perfect Info)
Production Cost [$M]	23.445	23.049	23.002
Uncertainty Cost [$M]	0.443	0.048	0 (by definition)
Renewable Curtailment [GWh]	69.1	63.9	63.3

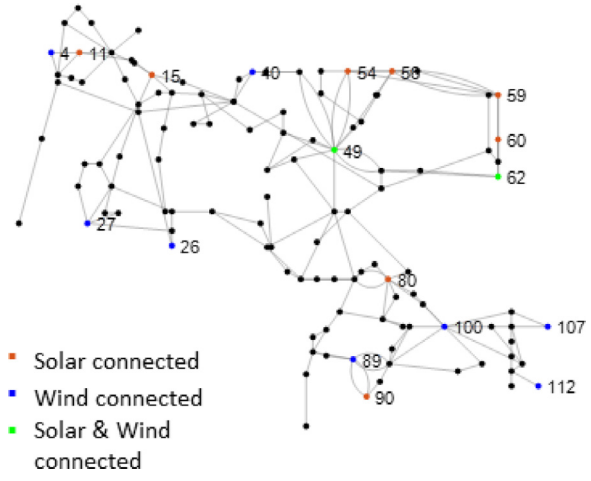

Fig. 12. Sites for solar and wind power plants in the modified 118-bus system. The orange/blue/green colors of nodes, respectively, indicate buses with solar PV power plants, wind power plants, or both wind and solar connected to them. The spatial representation of the 118-bus system is based on Fig. 1 of Ref. [25].

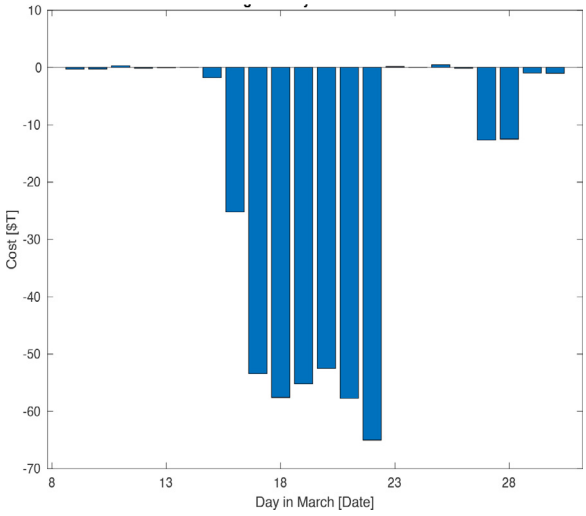

Fig. 13. Distribution of production cost reductions from using solar forecast-conditioned FRP requirements for 22 sample days in March 2020 compared to the CAISO baseline method, based upon 118-bus IEEE RTS 'CAISO-like" system

generation commitment and dispatch changes to procure ramp product and to adjust resource schedules to correct for forecast errors, including unexpected real-time generation costs that occur if not enough ramp is available to meet actual needs.

In addition, if system reliability (as measured, for instance, by expected unserved load) differs among the scenarios, this is also captured in the production costs, which include the value of lost load in the model's objective function. However, in all our simulations, there was no unserved demand, so the different FRP requirements in Scenarios 1 and 2 did not affect reliability.

All the analysis was performed using the FESTIV (Flexible Energy Scheduling Tool for Integrating Variable generation) [26] FESTIV is a production cost model that commits and dispatches day-ahead and real-time energy, procures reserves and ramp product, and simulates forecast errors between day-ahead and real-time, as well as real-time forecast errors that occur in the rolling real-time market processes.

Table 1 shows that the FRP requirements based on probabilistic solar forecasts (Scenario 2) are highly successful, reducing uncertainty-related costs by almost 90% (from $443K to $48k) in a 22-day period in March 2020. Most of these benefits occurred in a one-week period mid-month (Fig. 13). Use of those requirements decreased production costs by about 2% of total operating costs. Since actual energy and operating reserve costs of the CAISO system amount to about $8 billion/year, as noted above, if extrapolated to an entire system the size of the CAISO for one year, the savings would amount to over $100 million. However, such an extrapolation is only suggestive of the benefits of FRP, because results may be very different in other months and for a larger system.

We now examine the simulation results in detail to identify how the solar-informed FRP requirements of Scenario 2 improved system operations. A large share of the benefits of the solar-conditioned method (Scenario 2) arose in this case due to reduced FRP requirements for some intervals, leading to reduced commitment of conventional gener-

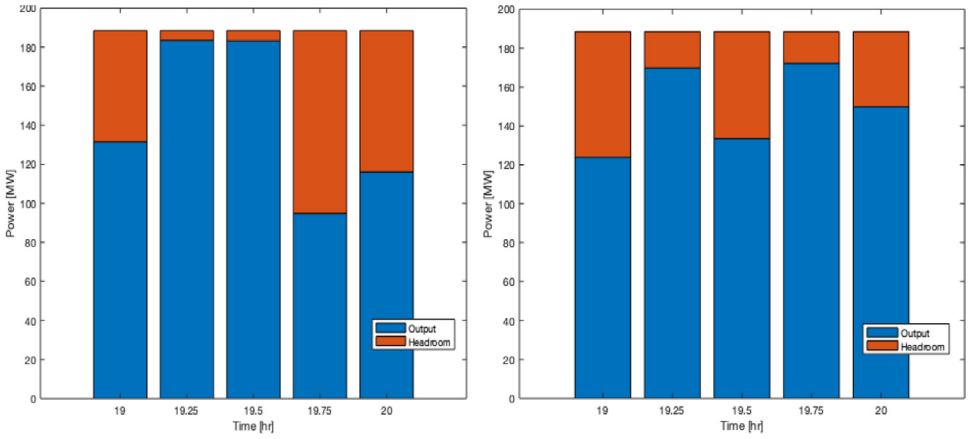

Fig. 14. Real-time dispatch of four impacted generators on March 16, 2020, under baseline FRP requirements (Scenario 1, left) and solar conditioned requirements (Scenario 2, right), illustrating that insufficient FRP-up in baseline case resulted in inadequate headroom in intervals two and three, necessitating real-time commitment of a costly fifth large generator in the fourth interval (time 19.75).

ation and turning off of large generators with spinning capacity. This enabled the system to accommodate more solar output and reduce curtailment.

Fig. 14 illustrates this impact by comparing the dispatch of four particular thermal generators in the evening of March 16, 2020, under Scenario 1 (baseline unconditional FRP requirements) and Scenario 2 (solar-uncertainty conditioned FRP requirements) (left and right sides of figure, respectively). The blue areas are the real-time energy production, and the orange areas are the "head room" that can accommodate unexpected upward ramps in the net load. The baseline scenario (left) has relatively low evening FRU requirements, so the day-ahead market believes that it has sufficient excess undispatched capacity on-line to meet the solar to thermal handoff in the evening. In real-time, however, the ramp-up turned out to be higher than expected (for instance, due to less solar output than expected), and this was reflected in the small amounts of "head room" (orange areas) in the second and third 15-minute intervals in the left side of Fig. 14. When this happened, the real-time market detected that available capacity was depleted in those intervals, and so a larger (fifth) generator was then committed in the real-time market's short-term unit commitment procedure. As a result, the four generators back down slightly in the fourth and fifth intervals to accommodate the minimum operating capability of the newly committed generator. By contrast, the larger FRU requirements in the solar-conditioned ramp scenario (Scenario 2) calculations (e.g., Fig. 11) force commitment in the day-ahead market of an additional smaller generator outside of the set of four generators considered in Fig. 14. This additional commitment leaves sufficient headroom to accommodate the larger than expected up-ramp in net load in real-time. The extra headroom shown as the orange areas in the second and third intervals in the right-side figure eliminate the need to commit the larger (fifth) generator we committed in the baseline scenario (Scenario 1). The result is significant cost-savings, as shown above for this day in Fig. 13.

The solar-conditioned FRP requirements also provided other benefits by providing more capacity in some other intervals by procuring more FRP and avoiding generation scarcity, particularly during the evening ramping intervals as solar production tapered off. In some days, smaller generators were committed by the day-ahead model, which turned out to yield a lower minimum thermal production (in particular, a lower sum of Pmin levels for on-line generators) (e.g., Fig. 15). This allowed the system to deploy rather than curtail solar production in some hours, saving production costs.

We note that systems of different sizes or with more or less generation flexibility could change the above quantitative conclusions. Also, the results for winter or summer conditions in the CAISO might be quantitatively different. However, we conjecture that the qualitative improvements seen by the use of solar conditioned FRP requirements are likely to still apply; this should be confirmed by future research.

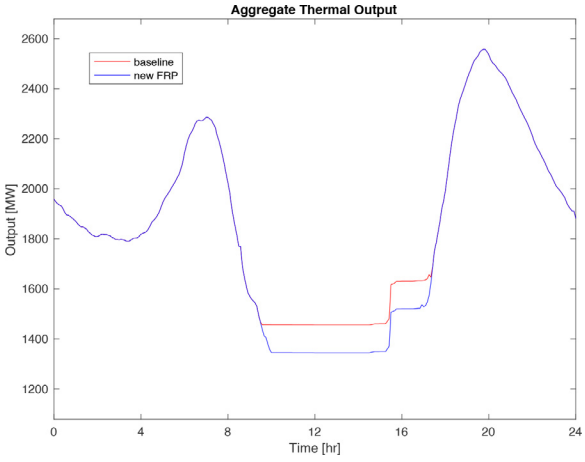

Fig. 15. Actual dispatch of all thermal generation in "CAISO-like" IEEE 118-bus system under baseline (unconditional) FRP requirements (red) and solar conditional requirements (blue). The reduced thermal production in the solar conditioned case indicates that the system can accommodate increased utilization (reduced curtailment) of solar output.

5. Conclusions

As solar and wind capacity is added to power systems, the need for—and cost of—operating reserves grows. Furthermore, it is reasonable to expect that the potential benefits of conditioning reserve requirements on weather and, especially, solar and wind operating conditions will increase as well. In this paper, we have used probabilistic solar forecasts from the probabilistic Watt-Sun forecasting system to create solar-conditioned requirements for the California Independent System Operator's newest type of operating reserves, called the flexible ramp product.

We provide details on two methods to relate solar forecast uncertainty (prediction interval width) to upward and downward uncertainty in 15-minute ramps: quantile regression and a k^{th}-nearest neighbor-based methodology. Based on the 2.5th and 97.5th percentiles of the resulting solar-adjusted conditional probability distributions of ramp uncertainty, we can define requirements for the flexible ramp product. Using the k^{th}-nearest neighbor-based requirements, we then estimate the economic and reliability benefits from using those requirements rather than the CAISO baseline (unconditional histogram system) method. This is done by production costing simulations of a modified 118-bus IEEE Reliability Test System, whose generator mix characteristics and load shapes resemble the California system. Those simulations indicate a potential production cost reduction of 2% for this system from using solar

uncertainty-conditioned forecasts, savings which arise from more efficient unit commitment and reduced solar curtailment.

Declaration of Interest

None.

Acknowledgments

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 Energy Technologies Office (SETO) Agreement Number EE0008215. We thank the following Independent System Operator staff for their advice and data provided over the course of the project: Amber Motley, Guillermo Bautista-Alderete, Clyde Loutan, Rebecca Webb, and Hong Zhou of the California ISO, and Stephen Rose and Blagoy Borissov of the Mid-Continent ISO. Guidance from DOE SETO staff, especially John Seuss and Kemal Celik, are gratefully acknowledged. Any opinions or errors expressed in the paper are solely the responsibility of the authors, and do not necessarily represent the position of any of the organizations employing the authors, the sponsoring agency, or the individuals mentioned in these acknowledgments.

References

[1] US Energy Information AgencyElectricity in the US, US Department of Energy, Washington, DC, 2022 www.eia.gov/energyexplained/electricity/electricity-in-the-us.php accessed 20 June.

[2] A.D. Mills, J. Seel, D. Millstein, H. Kim, M. Bolinger, W. Gorman, Y. Wang, S. Jeong, R.H. Wiser, Solar-to-Grid: Trends in System Impacts, Reliability, and Market Value in the United States with Data through 2019, Lawrence Berkeley National Laboratory, Berkeley, CA, 2021 https://emp.lbl.gov/sites/default/files/solar-to-grid_technical_report.pdf accessed 10 Oct. 2021.

[3] C.B. Martinez-Anido, B. Botor, A. Florita, S. Lu, H.F. Hamann, B.-M. Hodge, The value of day-ahead solar power forecasting improvement, Solar Energy 129 (2016) 192–203.

[4] U. Helman, B.F. Hobbs, R.P. O'Neill, The design of US wholesale energy and ancillary service auction markets: Theory and practice, in: F.P. Sioshansi (Ed.), Competitive Electricity Markets, Elsevier, 2008, pp. 179–243.

[5] California Independent System OperatorFlexible Ramping Product Refinements Initiative Appendix C - Quantile Regression Approach, Report, CAISO, 2020 http://www.caiso.com/InitiativeDocuments/AppendixC-QuantileRegressionApproach-FlexibleRampingProductRequirements.pdf (Accessed date: 10 October 2021).

[6] California Independent System Operator, Final Proposal: Flexible ramping product refinements, August 31, 2021, www.caiso.com/InitiativeDocuments/FinalProposal-FlexibleRampingProductRefinements.pdf (accessed 10 Oct. 2021).

[7] B. Wang, B.F. Hobbs, Real-time markets for flexiramp: A stochastic unit commitment-based analysis, IEEE Trans. Power Syst. 31 (2015) 846–860.

[8] G. Angelidis, Day-Ahead Market Enhancements, Appendix B: Draft Technical Description, California ISO, Folsom, CA, 2021 Vers. 6.1.Aug. 6 www.caiso.com/InitiativeDocuments/AppendixB-DraftTechnicalDescription-Day-AheadMarketEnhancements.pdf accessed 10 Oct. 2021.

[9] California Independent System OperatorAnnual Report on Market Issues and Performance 2020, Department of Market Monitoring, Folsom, CA, 2021 www.caiso.com/Documents/2020-Annual-Report-on-Market-Issues-and-Performance.pdf accessed 20 June 2022.

[10] M.A. Ortega-Vazquez, D.S. Kirschen, Optimizing the spinning reserve requirements using a cost/benefit analysis, IEEE Trans. Power Syst. 22 (2007) 24–33.

[11] Z. Zhou, A. Botterud, J. Wang, R.J. Bessa, H. Keko, J. Sumaili, V. Miranda, Application of probabilistic wind power forecasting in electricity markets, Wind Energy 16 (2013) 321–338.

[12] B. Li, J. Zhang, A review on the integration of probabilistic solar forecasting in power systems, Solar Energy 210 (2020) 68–86.

[13] N. Costilla-Enriquez, M. Ortega-Vasquez, A. Tuohy, A. Motley, R. Webb, Operating dynamic reserve dimensioning using probabilistic forecasts, IEEE Trans. Power Syst. (2022), doi:10.1109/TPWRS.2022.3163106.

[14] B.F. Hobbs, V. Krishnan, H. Hamann, Jie Zhang, R. Zhang, C. Siebenschuh, B. Li, L. He, X. Fang, E. Spyrou, I. Krad, Y. Wang, Q. Xu, S. Zhang, Coordinated Ramping Product and Regulation Reserve Procurements in CAISO and MISO using Multi-Scale Probabilistic Solar Power Forecasts (Pro2R), Final Report, Submitted to Solar Energy Technologies Office, US Department of Energy, 2022 June www.osti.gov/servlets/purl/1873393 Accessed 4 Sept. 2022.

[15] H.F. Hamann, A multi-scale, multi-model, machine-learning solar forecasting technology, IBM, Final Technical Report (2017) February www.osti.gov/servlets/purl/1395344/(accessed . 10 Oct. 2021).

[16] S. Lu, H. Hamann, IBM PAIRS: Scalable big geospatial-temporal data and analytics as-a-service, M. Werner, Y.-Y. Chiang (eds.), Handbook of Big Geospatial Data, Ch. 1, Springer (2021) 3-23.

[17] S. Lu, Y. Hwang, I. Khabibrakhmanov, F.J. Marianno, X. Shao, J. Zhang, B.M. Hodge, H.F. Hamann, Machine learning based multi-physical-model blending for enhancing renewable energy forecast-improvement via situation dependent error correction, 2015, Eur. Control Conf. (ECC) (2015) 283–290.

[18] National Oceanic and Atmospheric AdministrationHRRR Documentation, 2022 nomads.ncep.noaa.gov/txt_descriptions/HRRR_doc.shtml accessed 1 May.

[19] L. Hao, D.Q. Naiman, *Quantile regression*, Quantitative Applications in the Social Sciences No. 149, Sage Publishing, 2007.

[20] P. Cunningham, S.J. Delany, k-nearest neighbour classifiers - a tutorial, ACM Comput. Surv. 54 (2021) Article 128.

[21] 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, Appl. Energy 313 (2022) 118812.

[22] California Independent System Operator, Open Access Same Time Information System (OASIS), oasis.caiso.com/mrioasis/logon.do (accessed 20 June 2022).

[23] B.-M. Hodge, C. Brancucci Martinez-Anido, Q. Wang, E. Chartan, A. Florita, J. Kiviluoma, The combined value of wind and solar power forecasting improvements and electricity storage, Appl. Energy 214 (2018) 1–15.

[24] Illinois Institute of TechnologyThe IEEE 118-bus 54-unit 24-hour system, 2022 motor.ece.iit.edu/data/JEAS_IEEE118.doc accessed 20 June.

[25] I. Pena, C.B. Martinez-Anido, B.M. Hodge, An extended IEEE 118-bus test system with high renewable penetration, IEEE Trans. Power Syst. 33 (2017) 281–289.

[26] E. Ela, B. Palmintier, I. Krad, FESTIV (Flexible Energy Scheduling Tool for Integrating Variable generation, National Renewable Energy Laboratory, Golden, CO, 2019 www.osti.gov/biblio/1504740 accessed 10 Oct. 2021.