Baseline and Target Values for Regional and Point PV Power Forecasts: Toward Improved Solar Forecasting

Jie Zhang[a,*], Bri-Mathias Hodge[a], Siyuan Lu[b], Hendrik F. Hamann[b], Brad Lehman[c], Joseph Simmons[d], Edwin Campos[e], Venkat Banunarayanan[f], Jon Black[g], John Tedesco[h]

[a]*National Renewable Energy Laboratory, Golden, CO 80401, USA*
[b]*IBM TJ Watson Research Center, Yorktown Heights, NY 10598, USA*
[c]*Northeastern University, Boston, MA 02115, USA*
[d]*University of Arizona, Tucson, AZ 85721, USA*
[e]*Argonne National Laboratory, Lemont, IL 60439, USA*
[f]*U.S. Department of Energy, Washington, D.C. 20585, USA*
[g]*ISO New England, Holyoke, MA 01040, USA*
[f]*Green Mountain Power, Colchester, VT 05446, USA*

Abstract

Accurate **solar photovoltaic (PV)** power forecasting allows utilities to reliably utilize solar resources on their systems. However, to truly measure the improvements that any new solar forecasting methods provide, it is important to develop a methodology for determining baseline and target values for the accuracy of solar forecasting at different spatial and temporal scales. This paper aims at developing a framework to derive baseline and target values for a suite of generally applicable, value-based, and custom-designed solar forecasting metrics. The work was informed by close collaboration with utility and independent system operator partners. The baseline values are established based on state-of-the-art numerical weather prediction models and persistence models in combination with a radiative transfer model. The target values are determined based on the reduction in the amount of reserves that must be held to accommodate the uncertainty of **PV** power output. The proposed reserve-based methodology is a reasonable and practical approach that can be used to assess the economic benefits gained from improvements in accuracy of solar forecasting. The financial baseline and targets can be translated back to forecasting accuracy metrics and requirements, which will guide research on solar forecasting improvements towards the areas that are most beneficial to power systems operations.

*Keywords: numerical weather prediction; operating reserve; ramp forecasting; **PV power forecasting***

1. Introduction

The penetration of solar power in the electric grid is steadily rising. For instance, the USA SunShot Vision Study reported that solar power could provide as much as 14% of U.S. electricity demand by 2030 and 27% by 2050 (Margolis et al., 2012). The increasing penetration of solar power has presented new challenges for the reliable and economic operation of the power grid because of the high variability and uncertainty of solar PV power production (Lave and Kleissl, 2010; Lew et al., 2010). At high levels of solar energy penetration, solar forecasting will become imperative for reliable electricity system operation because it helps to reduce the uncertainty associated with matching power generation to load demands.

1.1. Baseline of Solar Forecasting

To properly measure improvements that any new solar forecasting method can provide, it is important to first assess the state of the art in solar forecasting. A number of papers in the literature present global horizontal irradiance (GHI) forecast models for day-ahead and other similar timescale forecasts. Generally, a baseline model is used for comparison, which is selected from: (i) persistence models (Diagne et al., 2013; IEA, 2013;

Inman et al., 2013); (ii) numerical weather prediction (NWP) models without bias correction (Perez et al., 2010; 2013); and (iii) NWP models with bias correction (Lorenz et al., 2009; Mathiesen and Kleissl, 2011). Most of the literature includes a comparison to persistence models in which the forecast 24 hours (or 48 hours, etc.) ahead is set to the measurement of irradiance from the day (or two, etc.) before. Even when comparing multiple models for different geographic locations to additional baselines, the persistence model is generally included as a reference. It is expected that certain models have higher accuracy at different geographic locations. Therefore, multiple locations have to be considered in order to provide a fair overview of the accuracy of a proposed model.

Relatively fewer papers in the literature present day-ahead baseline forecasts (or similar timescales) for solar power predictions; however, the approaches to solar power forecasting baselines seem to be similar to those for irradiance forecasts. Baseline models include (i) persistence models that set the day-ahead forecasted PV power equal to the measured PV power 24 hours before that time (Lorenz et al., 2011; Lorenz et al., 2012; Paulescu et al., 2013); (ii) NWP + plane of array (POA) irradiance calculation + PV models (Lorenz et al., 2011; Lorenz et al., 2012; Paulescu et al., 2013); and (iii) NWP with bias correction + POA irradiance calculation + PV models (Lorenz et al., 2011; Lorenz et al., 2012; Paulescu et al., 2013). Baselines in the different PV power predictions for day-ahead forecasts almost always include a persistence model. They may also utilize one or more of the NWP forecast models. NWP forecasts (with or without bias correction) are used to calculate POA irradiance using standard formulae (Duffie and Beckman, 2006; Erbs et al., 1982). The POA irradiance is the input to PV models that are able to compute the forecasted DC power and the AC forecasted power when multiplied by derating efficiency inverter curves. In addition, approaches such as PVWatts (Marion et al., 2001) (which linearly relates the irradiance to PV power) or PV-Lib (Stein, 2012) (which uses the King models (King et al., 2004)) are used to calculate the PV power. The baseline PV models normally include temperature corrections, but often do not include other advanced corrections, such as for wind, elevation, or low incident angles.

1.2. Research Motivation and Objectives

The development of baseline and target values for solar forecasting is closely related to the objective of quantifying the economic benefit of solar forecasting, around which currently the industry has no consensus. This is not only because of the complicated hierarchy and structure of the electrical energy market, but also because of the lack of in-depth understanding about how forecast information may fit into the specific utility or ISO operational practices, which are often complex and unique to each organization. Our development of baseline values and target economic metrics for quantifying the benefits of the solar forecast system has been based on close collaboration with utility and ISO partners.

Target values for the solar forecasting technology will help establish the goals for improvements in solar PV power forecasting. Different design objectives and strategies are used in power system operations, depending on the various timescales of the forecast to ensure economic operations and reliability; thus, it is important to characterize solar forecasting at all timescales of interest. Major customers of solar forecasting technologies include utility companies, independent system operators (ISOs), distribution system operators, etc. As solar penetration levels increase, solar forecasting will become more important to solar energy producers and solar power plant developers. Therefore, baseline and target values should be made for different geographic and energy-market regions to evaluate the versatility of the technology.

The objective of this paper is to present a procedure to determine the baseline and target values for **solar PV** power forecasting metrics. A suite of generally applicable, value-based, and custom-designed metrics were adopted for evaluating solar forecasting for different scenarios. **It is important to note this study focuses on the forecasting of GHI and solar PV power.** Section 2 presents the methodology for determining baseline and target values. Section 3 discusses the results of point and region solar forecasting case studies. Concluding remarks and future work are given in the final section.

2. Methodology for Determining Baseline and Target Metrics Values of Solar Forecasting

Operations of power systems occur at different timescales that can be summarized, from longest to shortest, as unit commitment, load-following, economic dispatch, and regulation (Ela et al. 2011). Figure 1 shows a general power system load pattern for a single day. Unit commitment and scheduling are performed throughout the day to economically commit the units in the system to meet the general system load pattern of the day. During shorter periods of time (minutes to hours), the system redispatches its units (and sometimes starts and stops quick-start units) to counteract deviations from the schedule through load following. Regulation is the balancing of fast second-to-second and minute-to-minute random variations in load or generation (Ela et al.

2011). **These strategies represent the balancing during normal conditions of the power system. The forecasts of load and renewable are never 100% accurate and different types of reserves are used to help mitigate the effects of forecast errors. Typical ancillary service markets are spinning contingency reserve, non-spinning contingency reserve, and regulating reserve. Variable generation such as wind and solar can increase the requirements for normal balancing reserve, such as regulating reserve, which can increase the prices for those services. Improved solar forecasting is expected to help reduce the reserve requirement. In general, nuclear, biomass, and coal power plants are committed in the day-ahead (DA) market, while combined and steam turbines are committed in the 4-hour-ahead (4HA) market. Improved short-term solar forecasts allow operators make better DA market operation and unit commitment decisions, help real-time operations in the hour ahead, and inform operators about extreme events.** To understand the impact of solar forecasts on solar power integration, it is important to characterize solar forecast errors at all timescales of interest. One of the objectives of this study is to determine the baseline and target solar forecasting metrics over a number of different timescales. Four solar forecast horizons are investigated: DA forecasts, 4HA forecasts, 1-hour-ahead (1HA) forecasts, and 15-minute-ahead (15MA) forecasts.
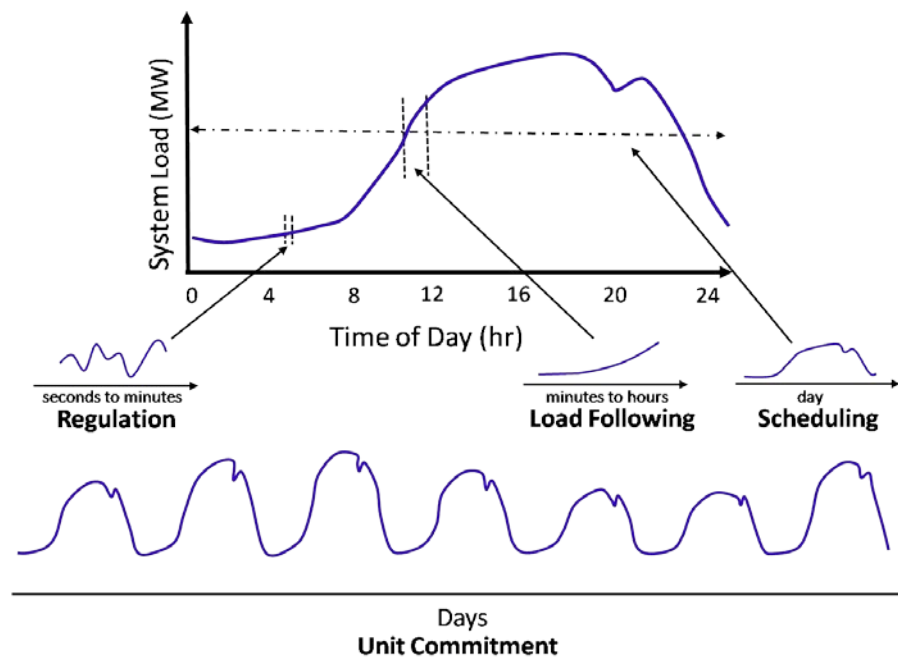

Figure 1. Power system operation timescales (Modified from Ela et al., 2011)

2.1. Metrics for Assessing Solar Forecasting

A suite of generally applicable, value-based, and custom-designed metrics for solar forecasting for a comprehensive set of scenarios (different time horizons, geographic locations, applications, etc.) were developed in previous work by the authors (Zhang et al., 2015). The proposed solar forecasting metrics can be broadly divided into four categories: (i) statistical metrics for different time and geographic scales; (ii) uncertainty quantification and propagation metrics, (iii) ramp characterization metrics; and (iv) economic metrics. A brief description of the metrics is given in Table 1, and detailed information about each metric can be found in (Zhang et al., 2015). A smaller value indicates a better forecast for most of the metrics, except for Pearson's correlation coefficient, skewness, kurtosis, distribution of forecast errors, and swinging door algorithm.

Table 1. Proposed metrics for solar power forecasting (Zhang et al., 2015)

Type	Metric	Description/Comment
Statistical Metrics	Distribution of forecast errors	Provides a visualization of the full range of forecast errors and variability of solar forecasts at multiple temporal and spatial scales
	Pearson's Correlation coefficient	Linear correlation between forecasted and actual solar power
	Root mean square error (RMSE) and normalized root mean square error (NRMSE)	Suitable for evaluating the overall accuracy of the forecasts while penalizing large forecast errors in a square order

	Root mean quartic error (RMQE) and normalized root mean quartic error (NRMQE)	Suitable for evaluating the overall accuracy of the forecasts while penalizing large forecast errors in a quartic order
	Maximum absolute error (MaxAE)	Suitable for evaluating the largest forecast error
	Mean absolute error (MAE) and mean absolute percentage error (MAPE)	Suitable for evaluating uniform forecast errors
	Mean bias error (MBE)	Suitable for assessing forecast bias
	Kolmogorov–Smirnov test integral (KSI) or KSIPer	Evaluates the statistical similarity between the forecasted and actual solar power
	OVER or OVERPer	Characterizes the statistical similarity between the forecasted and actual solar power on large forecast errors
	Skewness	Measures the asymmetry of the distribution of forecast errors; a positive (or negative) skewness leads to an over-forecasting (or under-forecasting) tail
	Excess kurtosis	Measures the magnitude of the peak of the distribution of forecast errors; a positive (or negative) kurtosis value indicates a peaked (or flat) distribution, greater or less than that of the normal distribution
Uncertainty Quantification Metrics	Rényi entropy	Quantifies the uncertainty of a forecast; it can utilize all of the information present in the forecast error distributions
	Standard deviation	Quantifies the uncertainty of a forecast
Ramp Characterization	Swinging door algorithm	Extracts ramps in solar power output by identifying the start and end points of each ramp
Economic Metrics	95th percentile of forecast errors	Represents the amount of non-spinning reserves service held to compensate for solar power forecast errors

2.2. Methodology for Determining Baseline Metrics Values

From the temporal point of view, the simplest approach to estimate forecasting baselines is that of climatology. The climatology approach consists of using a constant long-term average value throughout the entire forecasting period; the average value is often used as a benchmark for the forecasting skill with minimal effort. However, we consider that the solar forecasting sector has surpassed this benchmark, and a better baseline approach is needed.

Instead, we decided to use two other fundamental forecasting approaches: the model and the persistence approaches. The model approach corresponds to the use of products from NWP models, which rarely achieve useful skill at lead times smaller than a few hours because of the (spin-up) period they require to achieve numerical stability. The persistence approach (more specifically, Eulerian persistence) corresponds to using the persistence of the recent observations. This shows superior skill in the shorter forecasting periods and when atmospheric variability is smaller (e.g., dry climates, few clouds). Specifically, the persistence forecasts show better skill than the model forecasts in the short term, whereas the model forecasts show better skill (than the persistence) after a few hours in the forecasting period (Germann et al., 2006). **Other than NWPs, sky camera and geostationary satellite image analyses are also employed for short term forecasting (Chu et al. 2014; Marquez et al. 2013; Lorenz et al. 2012; Lorenz et al. 2014; Perez et al. 2010). Though accurate, such methods are sophisticated and often require proprietary data sources, which make them difficult to reproduce independently and thus impractical for the purposes of establishing baselines.**

Given the aforementioned considerations, the overall methodology for establishing baseline values is to use the persistence approach for 0- to 4-hour lead time, and to use the model approach for day-ahead lead time as summarized in Table 2. The establishment of baseline values for day-ahead forecasting for each of the metrics is based on a state-of-the-art weather model, specifically the North American Mesoscale Forecast System (NAM) (Mesinger et al., 2006), in combination with a streamer radiative transfer model (RTM) and the PV-Lib toolbox (Stein, 2012) for irradiance-to-power modeling. **The NAM model was chosen as this work is focused on baseline and target values for solar forecasting within the US. Without losing generality, the methodology presented may be applied outside North America by replacing the NAM model with a local NWP model or a global model such as the Global Forecast System (GFS), whichever is the best available for the region of interest (Lara-Fanego et al. 2012; Perez et al. 2013).** To remove substantial bias errors of day-ahead forecasts, a first order machine learning (linear regression model) model is applied based on the 3-day historical data. A modified persistence model is adopted for the 15MA, 1HA, and 4HA forecasts. In the following, we describe in detail the protocols for deriving the day-ahead baseline from numerical weather

prediction and hours-ahead baseline from persistence.

Table 2. Overall approach to determining baseline forecasts at different forecast horizons

Forecast Horizon	Weather Information	Irradiance Forecasts	Power Forecasts
15-min-ahead (15MA), 1-hour-ahaed (1HA), and 4-hour-ahead (4HA)	Persistence	Streamer RTM	Persistence of cloudiness
Day-ahead (0-23 hours ahead and 24-47 hours ahead)	NAM	Streamer RTM	(1) PV-Lib + linear regression; or (2) linear least square fit (if no PV specifications available)

2.2.1. Numerical weather predictions for DA solar forecasting

This day-ahead forecast uses NAM weather forecasting and a streamer RTM. The 5-km grid NAM forecast that runs at 06z daily is employed. Two NAM forecast windows, 0 to 23 hours ahead and 24 to 47 hours ahead, are used to derive solar irradiance forecasts for the two types of day-ahead (0 to 23 hours and 24 to 47 hours ahead) baseline values.

The vertical profiles (39 vertical levels) of pressure, temperature, geopotential height, humidity, cloud liquid water content, cloud ice content, and surface albedo are taken from the NAM forecast. The climatological monthly average of ozone concentration and aerosol optical depth are taken from the MODIS (Moderate Resolution Imaging Spectroradiometer) data set (Justice et al., 1998). Together these form the input for the streamer RTM (Key and Schweiger, 1998), which solves the radiative transfer equation for the plane-parallel geometry using the spherical harmonic discrete ordinate method to calculate GHI and direct normal irradiance (DNI) at the Earth's surface. The GHI, DNI, ambient temperature (2 m above ground), and wind speed (10 m above ground) are then fed into an irradiance-to-power model to derive forecasted AC PV power. The irradiance-to-power model consists of two parts: (i) the PV modules are modeled using the California Energy Commission (CEC) model, and (ii) the inverters are modeled using the Sandia National Laboratories model, both of which are implemented in PV-Lib (Stein, 2012).

2.2.2. Persistence for 15MA, 1HA, and 4HA solar forecasting

The 15MA, 1HA, and 4HA forecasts were synthesized using a persistence of cloudiness approach, as shown in Fig. 2. In this method, the solar power index (SPI) is first calculated, which represents the ratio between actual power (P) and clear-sky power (P_{CS}). Then, the solar forecast power is estimated by modifying the current power output by the expected change in clear-sky output. For the 1HA persistence of cloudiness approach, the forecast solar power at time $t+1$ can be calculated as follows.

$$P(t+1) = P(t) + SPI(t) \times \left[P_{CS}(t+1) - P_{CS}(t) \right] \tag{1}$$

where $P_{CS}(t+1)$ and $P_{CS}(t)$ represent the clear sky solar power at time t+1 and t, respectively; $P(t)$ is the actual solar power output at time t; and $SPI(t)$ is solar power index at time t.
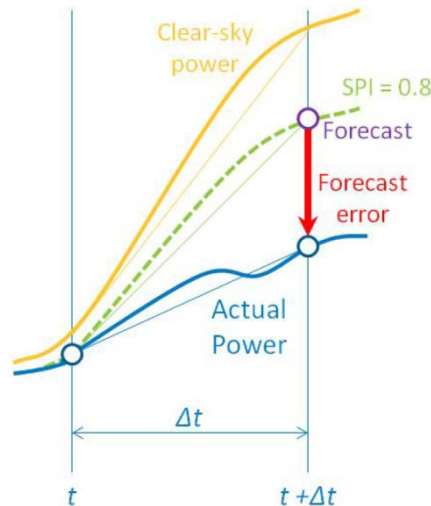

Figure 2. Persistence of cloudiness approach (Lew et al., 2013)

In this work, the clear-sky power $P_{CS}(t)$ is calculated for all test site locations following two steps. First, the standard summer atmospheric profile of temperature, pressure, and humidity without clouds is assumed. Streamer RTM (Key and Schweiger, 1998) is employed to calculate the Earth's surface-level GHI and DNI. Second, the CEC irradiance-to-power model, which is implemented in PV-Lib, is used to calculate AC PV

output from GHI and DNI. In the irradiance-to-power calculation, we assume an ambient temperature of 300 K and no wind.

2.3. Methodology for Determining Target Metrics

The baseline forecasting is divided into two parts: non-ramping period and ramping period. The target values of solar forecasting metrics are derived by (i) for the non-ramping period, applying uniform forecasting improvements by *x%* based on the baseline forecasting; (ii) for the ramping period, applying ramp forecasting improvements by *y%* based on the baseline forecasting; and (iii) deriving a complete set of target metrics. The values of *x%* and *y%* are determined based on the economic impacts of improved solar power forecasting (i.e., a reduction of 25% in reserve levels, which is based on a project partner utility consensus).

Two types of forecasting improvements are implemented. The improvements are categorized by the appearance of large solar ramps, which are commonly one of the biggest concerns of high-penetration solar power scenarios. First, the start and end points of all significant ramps are extracted using the swinging door algorithm (see Section 2.3.1). The definition of significant ramps is based on the magnitude of solar power change. These improvements are generated through the following procedures:

- Uniform improvements of the time series excluding ramping periods: The forecast errors of the time series when there is not a significant ramp are uniformly decreased by a percentage (*x%*).
- Ramp forecasting magnitude improvements: Only significant ramps that are identified as a change greater than or equal to a threshold value (θ) are modified in the improved forecasts. The forecast errors of the time series with ramps are decreased by a percentage (*y%*).
- Ramp forecasting threshold: The ramp forecasting threshold (θ) is set as 10% of the solar power capacity in this paper.

Figure 3 illustrates the overall structure of the methodology to determine target metrics for solar forecasting.
(a) First, the reserve cost of the baseline solar forecasting (C_b) is calculated, and a 25% reduction is assumed for the target reserve level (C_t).
(b) Second, a set of (N) combinations of *x%* uniform forecasting improvement and *y%* ramp forecasting magnitude improvements are applied to the baseline forecasting. The reserve costs (C_i) from the N improved solar forecasting combinations are calculated. The N combinations are generated using a design of experiments method. In this paper, the Sobol's quasi-random sequence generator, a widely used method in design of experiments (Forrester et al., 2008; Zhang et al., 2013), was adopted to generate the N combinations. Sobol's sequences use a base of two to form successively finer uniform partitions of the unit interval and reorder the coordinates in each dimension. The Sobol sequence generator produces highly uniform samples of the unit hypercube. Note that other design of experiments methods can also be used, e.g., Latin hypercube sampling, fractional factorial, central composite, uniform designs, etc.
(c) Ideally when we reach target there would have a difference of zero, but because of the discrete set of N combinations, there will be small errors and the errors will be within a small bound. As N approaches infinity, the error would exactly equal zero. In this study, the set of *x%* and *y%* values, with the smallest difference between the ideal target reserve cost (C_t) and the reserve cost from the improved solar forecasting (C_i), is then selected for deriving the final target value for the solar forecasting. It is important to note that other sets of *x%* and *y%* combinations can also be used if the selection criterion is changed.
(d) Finally, the target solar forecasting is derived based on the selected set of *x%* and *y%* in the previous step. A complete set of target metrics is calculated based on the target solar forecasting value.
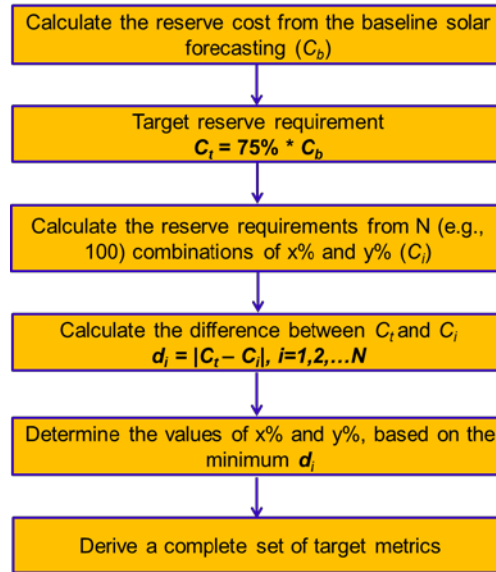
Figure 3. Overall structure to determine target metrics

It is important to note that proposed methodology is readily to be extended to arbitrary targets. In addition to the reserve cost used in this paper, the target can be extended to production cost, solar power curtailment, power system reliability, and other important power system metrics.

2.3.1. Swinging Door Algorithm

The swinging door algorithm extracts ramp periods in a time series of power signals, by identifying the start and end points of each ramp. The algorithm allows for consideration of a threshold parameter influencing its sensitivity to ramp variations. The only tunable parameter in the algorithm is the width of a "door", represented by ε in Fig. 4. The parameter ε directly characterizes the threshold sensitivity to noise and/or insignificant fluctuations to be specified. With a smaller ε value, many small ramps will be identified; with a larger ε value, only a few large ramps will be identified. A detailed description of the swinging door algorithm can be found in (Bristol, 1990; Florita et al., 2013).
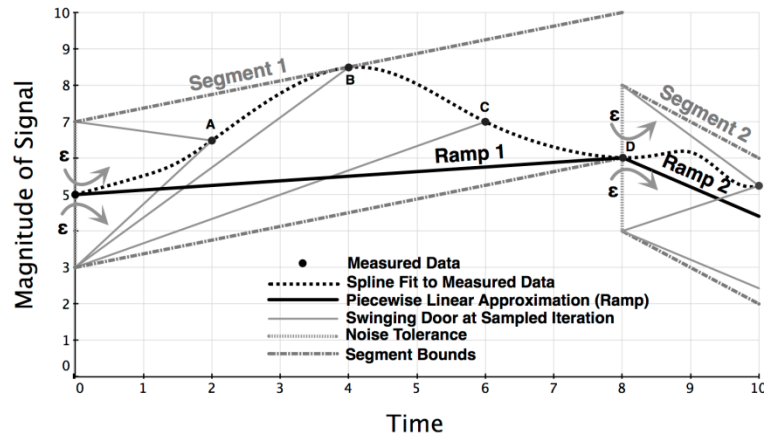

Figure 4. The swinging door algorithm for the extraction of ramps in power from the time series (Florita et al., 2013)

2.3.2. Flexibility Reserves for 15MA, 1HA, 4HA, and DA Forecasting

The reduction in the amount of reserves that must be carried to accommodate the uncertainty of solar power output is anticipated to be one of the significant cost savings associated with improved solar power forecasting. An advanced reserve calculation algorithm is thus applied to estimate the reserve reductions that various solar power forecasting improvements would allow. This methodology was originally developed during the Western Wind and Solar Integration Study Phase 2 (WWSIS-2) (Lew et al., 2013). Improved forecasting (on average) reduces the amount of reserves that must be held, and various types of flexibility reserves are defined by:

- For 15MA, 1HA, and 4HA solar power forecasting, spinning reserves are used to derive the target solar forecasting values. Spinning reserves represent the online capacity that can be deployed very quickly (seconds to minutes) to respond to variability. In this study, the spinning reserve for 0- to 4-hours-ahead

forecasting (R_s^{HA}) is defined as the 95% confidence interval ($\emptyset_{95}$) of solar power forecast errors (e^{HA}) at the 15MA, 1HA, or 4HA horizon.

$$R_s^{HA} = \emptyset_{95}(e^{HA}) \tag{2}$$

- For DA solar forecasting, both spinning and non-spinning reserves are used to derive the solar forecasting target. Non-spinning reserves represent the off-line or reserved capacity, or load resources (interruptible loads), capable of deploying within 30 minutes for at least 1 hour. In this paper, the spinning reserve for the DA forecasting (R_s^{DA}) is defined as the 70% confidence interval ($\emptyset_{70}$) of the DA solar power forecast errors (e^{DA}) (Hodge et al., 2015). The non-spinning reserve (R_{ns}^{DA}) is defined by the difference between the 95% confidence interval ($\emptyset_{95}$) and the 70% confidence interval ($\emptyset_{70}$) of the DA solar power forecast errors (e^{DA}) (Hodge et al., 2015).

$$R_s^{DA} = \emptyset_{70}(e^{DA}) \tag{3}$$

$$R_{ns}^{DA} = \emptyset_{95}(e^{DA}) - \emptyset_{70}(e^{DA}) \tag{4}$$

Considering the cost of holding and deploying reserves, this study assumes that the cost of non-spinning reserve per MW (C_{ns}^{MW}) is twice the cost of spinning reserve per MW (C_s^{MW}):

$$C_{ns}^{MW} = 2 \times C_s^{MW} \tag{5}$$

Equation (5) is derived based on the (i) start costs of two types of generators used for spinning and non-spinning reserves (gas turbine and oil turbine); and (ii) heat rates and fuel costs of four fuel types (biomass, nuclear, coal, and combined cycle). In this paper, the costs are selected according to the ISO-New England (ISO-NE) system.

3. Case Studies: System Operators, Utilities, and Energy Producers

Major customers for solar forecasting technologies are utility companies, independent system operators (ISOs), distribution system operators, etc. As solar penetration levels increase, solar forecasting will become more important to solar energy producers and solar power plant developers. The baseline and target values will be determined for different geographic and energy-market regions to evaluate the versatility of the technology.

3.1. Regional and Point **PV Power** Forecasts Scenarios

From the spatial point of view, there are two distinct types of test cases for the baseline and target values: point (single PV plant) and regional. In the point case, we are able to gather detailed information about the operational configuration of a particular PV plant (e.g., PV panel and inverter types). In conjunction with atmospheric estimates throughout the plant, we can perform physics-based numerical forecasts of the power production at the plant. We can also validate the point forecasts by using the observed power production. In the regional case, it may be impractical to gather the operational configurations for the multitude of solar PV producers in the region; however, an empirical relationship between the regional power production and the solar irradiance fields can be estimated.

Three PV plants were chosen among hundreds of sites available by the solar utilities in the Watt-sun (Lu et al., 2015a; 2015b) research consortium as point test cases: Smyrna, Green Mountain Power (GMP), and Tucson Electric Power (TEP). The selection was based on the best quality, continuity, and variety of power production observations at the sites. In addition, two regional test cases, ISO-New England (ISO-NE) and California-ISO (CAISO) were chosen to cover two distinct atmospheric conditions: a cloudier and more humid climate for the ISO-NE region, in contrast to relatively drier climate in the CAISO region. A few assumptions were made: (i) data points at nighttime were removed when the actual or forecasted power was zero; (ii) hourly point forecasts were used for DA, 4HA, and 1HA forecasts; (iii) 15-minute average forecasts were used for 15MA forecasts; and (iv) a set of 100 combinations of x% and y% were evaluated for 5 case studies (CAISO, ISO-NE, TEP, GMP, and Smyrna). The locations of the three point and two regional test cases are shown in Fig. 5. **Table 4 lists the details of the five cases, including the forecast horizons, irradiance and power forecasts methods, validation methods and periods. Due to data limitations, the evaluation period is mainly during the summer and fall seasons. Solar forecasting during the winter season may be slightly different based on the geographic locations, such as the tendency for over-forecasting or under-forecasting.**
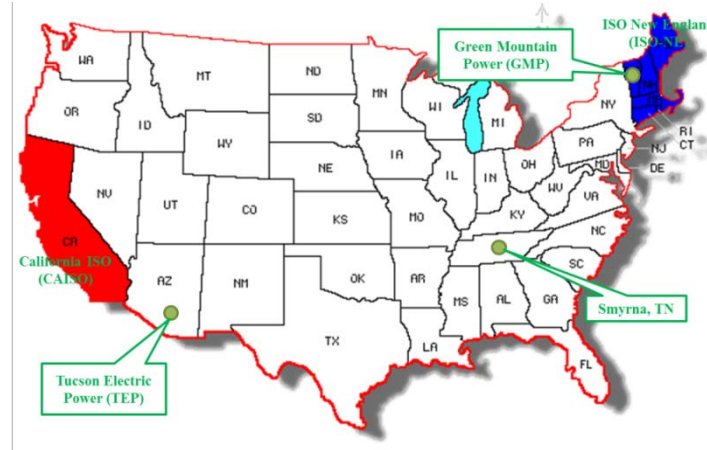
Figure 5. Locations of the three point and two regional test cases

Table 4. Test cases of system operators, utilities, and energy producers

| Test Case | Role | Forecast Horizon | | Irradiance Forecasts | Power Forecasts | Validation | Evaluation Period |
		15MA, 1HA, and 4HA	Day-ahead				
ISO-NE	System operator	Persistence	NAM	Streamer RTM	N/A	GHI –12 MesoWest sites	**03-May-2013 to 30-Oct-13**
CAISO	System operator	Persistence	NAM	Streamer RTM	Linear least square fit	Aggregated Power	**04-May-2013 to 30-Oct-13**
GMP	Utility	Persistence	NAM	Streamer RTM	PV-Lib	Direct Power measurements	**03-May-2013 to 30-Oct-13**
TEP	Utility	Persistence	NAM	Streamer RTM	PV-Lib	Direct Power measurements	**02-June-2013 to 30-Oct-13**
Smyrna	Energy producer	Persistence	NAM	Streamer RTM	PV-Lib	Direct Power measurements	**03-May-2013 to 30-Oct-13**

3.2. Baseline and Target Metrics Values for Multiple Scenarios

3.2.1. CAISO Baseline and Target Metrics Values

For CAISO, data is available in 1-hour intervals in energy production summed throughout the CAISO region (OASIS Website, http://oasis.caiso.com/). The actual and forecasted data used in this case are between 04-May-2013 and 30-Oct-13. **Solar power is derived by assuming a constant power generation within one hour. The total capacity considered in this study is approximately 4,100 MW. For CAISO, the forecasts of aggregated hourly solar power output are derived from the NAM model without PV specification using a persistence-weighted-by-irradiance method. Namely, one starts with the measured solar power output of the current hour in a day (P_c) as well as the NAM GHI forecast (averaged over the entire CAISO service territory) of the current hour (GHI_c) and GHI forecast of the same hour one or two days ahead (GHI_{da}). The day-ahead solar power output (P_{da}) is then computed as: $P_{da} = P_c \times GHI_{da}/GHI_c$. Moreover, in order to remove systematic bias-error, the day-ahead solar energy forecast, derived using either PV-lib or persistence-weighted-by-irradiance as noted above, are corrected using linear regression with historical data from three days immediately preceding the forecast.**

Figure 6 shows the amounts of the baseline and target reserves for different forecast horizons. The baseline reserves at different timescales are shown on the top the figure (e.g., 227.75 MW and 448.25 MW in Fig. 6(a)); the target reserves are listed on the bottom of the figure (e.g., 168.17 MW and 335.48 MW in Fig. 8(a)); the percentages of the uniform and ramp forecasting improvement are listed in the middle (e.g., 25.13% and 30.88% in Fig. 6(a)). For DA forecasts in Fig. 6(a), the uniform forecasting improvement and ramp forecasting improvement were determined based on the combined reduction (25%) in spinning and non-spinning reserves costs. For the 15MA, 1HA, and 4HA forecasts, the improvements were determined based only on spinning reserves.

(a) Day-ahead forecasts at CAISO   (b) 15MA, 1HA, and 4HA forecasts at CAISO
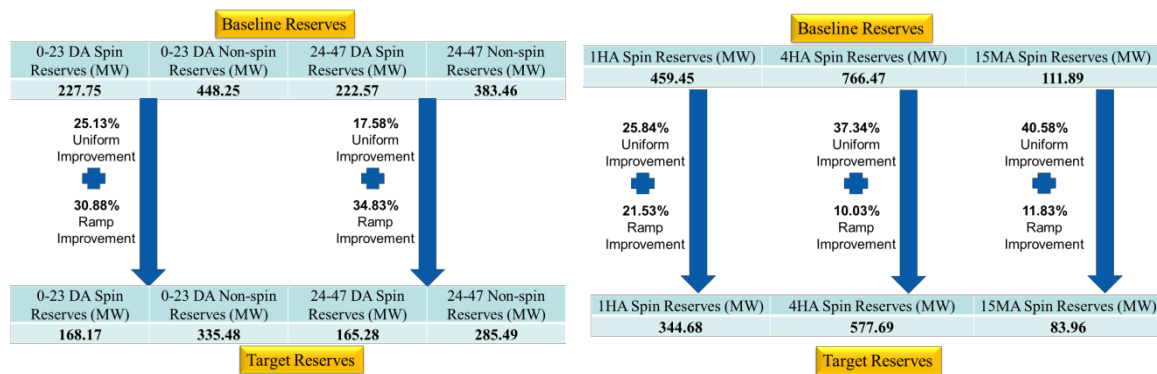
Figure 6. Baseline and target reserves values; uniform and ramp forecasting improvements were determined based on 25% reduction in reserve costs

Figures 7(a) and 7(b) illustrate the distributions of solar power forecast errors for the baseline and target forecasting, respectively. It shows that (i) the 15MA forecasts perform the best among all forecast horizons, as shown by the peak of the distribution, and (ii) the distribution of the target forecast errors is relatively skinnier than the corresponding distribution of the baseline forecast errors.

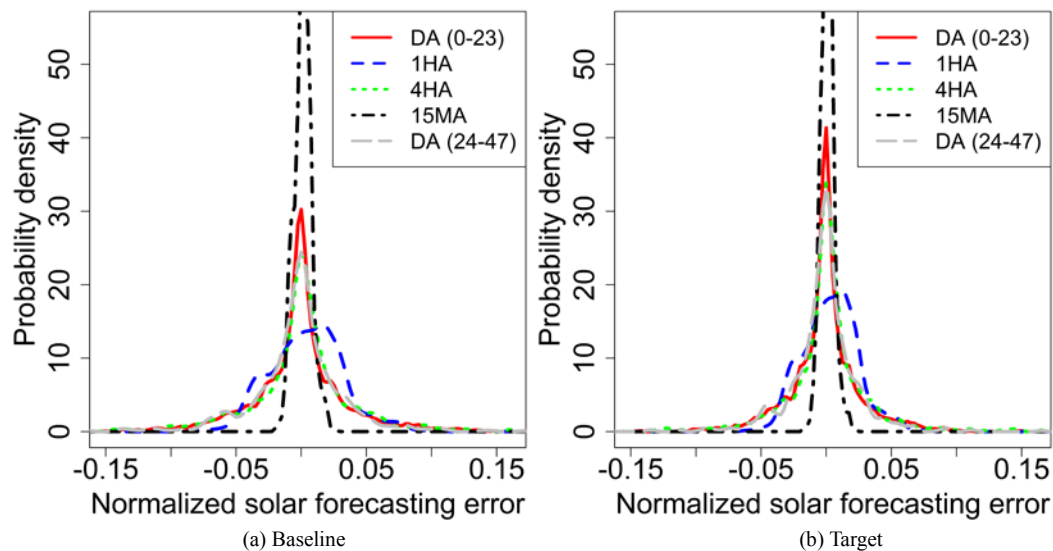(a) Baseline   (b) Target

Figure 7. Distribution of baseline and target solar power forecast errors at DA, 4HA, 1HA, and 15MA forecast horizons (CAISO)

The baseline values and target metrics are summarized in Table 5. The NAM model for the DA forecasts and the persistence model for the 4HA, 1HA, and 15MA forecasts are accurate with very high correlation coefficients and small RMSE and MAPE values. The relatively larger RMSE and MAE values in the 4HA forecasts are partially attributed to the inherent assumption that indirect light and panel temperature changes of more than 4 hours are not accounted for. The 15MA baseline forecast has a high accuracy at 1% NRMSE by capacity, which indicates that the baseline persistence model is a strong predictor of PV output. Such a small error is likely an indication that geographic averaging across a large region cancels the errors. It is interesting to notice that reducing the spinning reserve costs (by 25%) using the Fig. 3 algorithm, may lead to a specific metric being worsened. Target values are better than baseline values for most metrics. However, the target renyi entropy values are slightly increased compared to the baseline for DA forecast horizons. This is because the renyi entropy utilizes all of the information present in the forecast error distribution. MBE and standard deviation reflect the improvement based on the proposed reserve methodology; other information present in the forecast error distribution (e.g., skewness and kurtosis) may not necessarily reflect the improvement.
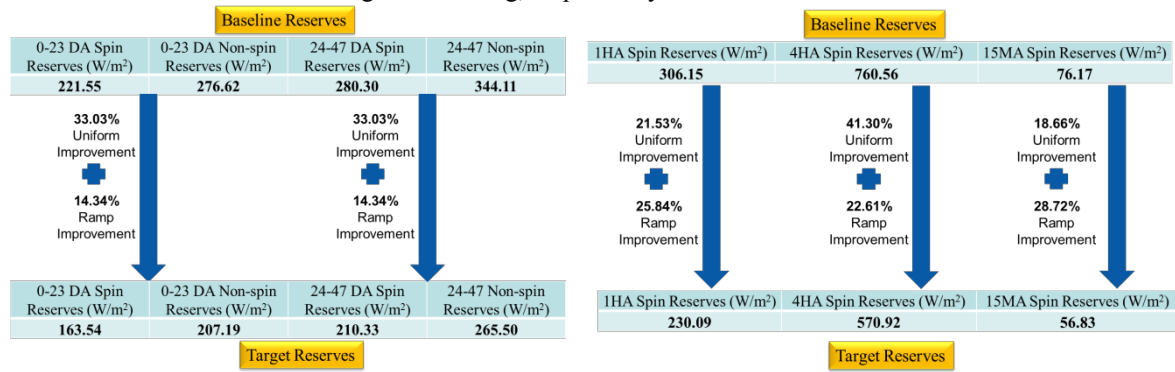
Table 5. Baseline and target metrics values for CAISO at different forecast horizons

Metrics	DA (24-47) Baseline	DA (24-47) Target	DA (0-23) Baseline	DA (0-23) Target	4HA Baseline	4HA Target	1HA Baseline	1HA Target	15MA Baseline	15MA Target
Correlation coefficient	0.97	0.98	0.98	0.99	0.96	0.97	0.98	0.99	1.00	1.00
RMSE (MW)	168.39	120.05	150.54	110.82	184.62	149.17	119.91	90.75	29.01	21.42
NRMSE by capacity	0.04	0.03	0.04	0.03	0.04	0.04	0.03	0.02	0.01	0.01
MaxAE (MW)	2728.00	1777.89	860.02	619.10	1736.00	1561.86	1252.63	982.93	313.16	276.12
MAE (MW)	98.56	71.74	98.91	72.68	111.97	85.35	93.98	70.95	22.24	15.45
MAPE by capacity	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.01	0.00
MBE (MW)	-8.25	-6.55	-5.72	-4.46	4.45	4.38	16.74	12.42	4.43	3.35
KSIPer (%)	16.61	14.12	16.36	14.71	31.02	22.68	38.91	31.88	16.93	13.82
OVERPer (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00
Standard dev. (MW)	168.25	119.92	150.47	110.76	184.63	149.16	118.77	89.91	28.68	21.16
4RMQE (MW)	472.71	311.54	237.20	174.47	371.25	326.40	204.29	158.20	50.08	42.77
N4RMQE by capacity	0.11	0.07	0.06	0.04	0.09	0.08	0.05	0.04	0.01	0.01
95th percentile (MW)	304.10	227.02	339.69	251.51	386.30	298.80	229.08	175.52	55.85	42.29
Renyi entropy	3.09	3.24	4.21	4.23	3.44	3.13	4.54	4.51	4.42	4.07
NRMSE by clear sky power	0.31	0.22	0.26	0.19	0.27	0.22	0.19	0.14	0.02	0.02
MAPE by clear sky power	0.18	0.13	0.17	0.13	0.16	0.12	0.15	0.11	0.02	0.01

3.2.2. ISO-NE Baseline and Target Metrics Values

For ISO-NE, solar generation is mostly behind the meter and interconnected to the distribution system, so the value of an improved forecast technology will lead to improved net load (i.e., the load minus the PV) forecasting – especially for the day-ahead unit commitment process. Currently, in this case study, the metrics were calculated based on solar irradiance instead of power due to the data access limitation. Note that the spin and non-spin reserves are not the best metric for irradiance, but we kept the methodology the same for the sake of consistency. The actual and forecasted data used in this case are between 03-May-2013 and 30-Oct-13. The baseline and target metrics for ISO-NE are summarized in Table 6. The capacity used for normalization is 1000 W/m^2. It is observed that DA (both 0-23 and 24-47 hours ahead) baseline forecasts performed better than the 4HA baseline forecasts. This can be partially attributed to the cloudy weather of ISO-NE region; the persistence forecast is significantly affected by the cloud movement. It is important to note that for the ISO-NE case, the irradiance is calculated by averaging a set of sites. However, since the available sites are closely located in a small geographical region and their irradiance values are highly correlated. In the CAISO case much of the short term variability is smoothed out by geographic diversity. The MBE values of all forecast horizons are positive, indicating an over-forecasting tendency at all forecast horizons; the MBE values of 4HA forecasts (both baseline and target) are larger than those of other forecast horizons.

Figure 8 shows the baseline and target reserves values (in terms of irradiance) at different forecast horizons. To achieve the target reserves, there is more ramp forecasting improvement than uniform improvement for DA and 4HA forecasts, and there is more uniform improvement than ramp forecasting improvement for shorter timescale forecasts (1HA and 15MA). Figures 9(a) and 9(b) illustrate the distributions of solar power forecast errors for ISO-NE baseline and target forecasting, respectively.

(a) Day-ahead forecasts at ISO-NE (b) 15MA, 1HA, and 4HA forecasts at ISO-NE

Figure 8. Target reserves values based on uniform and ramp forecasting improvement (ISO-NE)

(a) Baseline   (b) Target
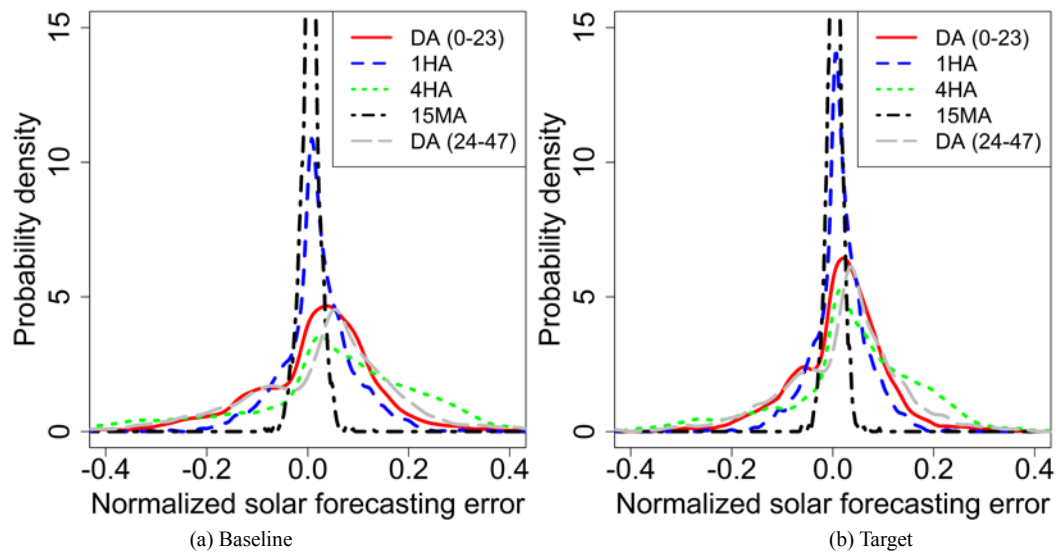
Figure 9. Distribution of baseline and target solar power forecast errors at DA, 4HA, 1HA, and 15MA forecast horizons (ISO-NE)
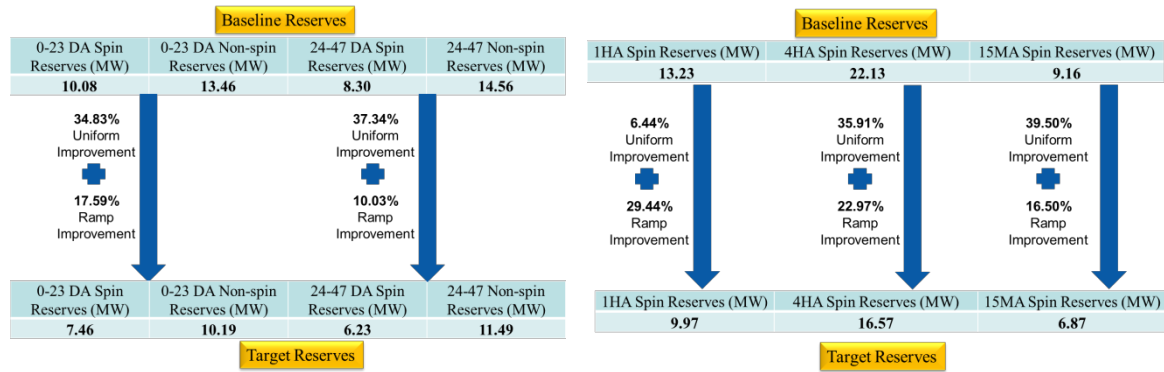
Table 6. Baseline and target metrics values for ISO-NE at different forecast horizons

Metrics	DA (24-47) Baseline	DA (24-47) Target	DA (0-23) Baseline	DA (0-23) Target	4HA Baseline	4HA Target	1HA Baseline	1HA Target	15MA Baseline	15MA Target
Correlation coefficient	0.80	0.90	0.86	0.93	0.73	0.85	0.96	0.97	1.00	1.00
RMSE (W/m^2)	152.55	115.75	122.27	89.73	192.21	143.12	73.25	55.19	18.32	13.68
NRMSE by capacity	0.15	0.12	0.12	0.09	0.19	0.14	0.07	0.06	0.02	0.01
MaxAE (W/m^2)	617.54	528.96	513.11	369.03	715.12	553.44	357.36	265.00	129.05	91.99
MAE (W/m^2)	119.13	88.80	92.58	67.83	147.58	107.52	52.99	40.12	13.02	9.81
MAPE by capacity	0.12	0.09	0.09	0.07	0.15	0.11	0.05	0.04	0.01	0.01
MBE (W/m^2)	24.05	19.62	15.98	14.80	52.87	39.39	18.57	14.12	4.60	3.51
KSIPer (%)	170.83	148.47	147.98	121.29	144.25	131.54	60.92	47.50	20.74	17.21
OVERPer (%)	89.12	67.55	68.99	47.07	75.54	62.44	7.85	4.48	0.00	0.00
Standard dev. (W/m^2)	150.69	114.11	121.25	88.52	184.85	137.63	70.87	53.36	17.73	13.23
4RMQE (W/m^2)	212.81	165.39	175.62	128.06	261.63	198.49	107.36	80.52	28.17	20.91
N4RMQE by capacity	0.21	0.17	0.18	0.13	0.26	0.20	0.11	0.08	0.03	0.02
95th percentile (W/m^2)	315.68	234.01	254.17	184.18	380.34	286.38	154.16	116.37	38.41	28.55
Renyi entropy	5.29	5.11	5.18	5.16	5.22	5.10	4.74	4.80	4.42	4.48
NRMSE by clear sky irradiance	0.28	0.22	0.22	0.16	0.30	0.23	0.12	0.09	0.03	0.02
MAPE by clear sky irradiance	0.22	0.17	0.17	0.12	0.23	0.17	0.09	0.07	0.02	0.02

3.2.3. TEP Baseline and Target Metrics Values

The **PV power** data collected from Tucson Electric Power (TEP) is with 15-minute interval and the capacity is 25 MW. The data shows a large number of clear-sky days which enable us to develop an improved clear-sky model for single-axis tracking. The actual and forecasted data used in this case are between 02-June-2013 and 30-Oct-13. Table 7 lists the baseline values and target metrics at different forecast horizons. Baseline values of 15 MA and 1HA forecasts present large correlation coefficients, because the persistence model benefits from many consecutive clear-sky days at TEP. The TEP data shows a strong effect of panel temperature especially in summer. Though a number of consecutive clear-sky days exist at the TEP region, the persistence method, which is generally more accurate in clear-sky days, is significantly affected by the large panel temperature difference over 4 hours.

Figure 10 shows the baseline and target reserve requirements. To achieve the target reserves, there is generally more uniform improvement than ramp forecasting improvement (for the DA, 4HA, and 15 MA forecasts). Figures 11(a) and 11(b) illustrate the distributions of solar power forecast errors for baseline and target forecasting, respectively.

(a) Day-ahead forecasts at TEP (b) 15MA, 1HA, and 4HA forecasts at TEP

Figure 10. Target reserves values based on uniform and ramp forecasting improvement (TEP)
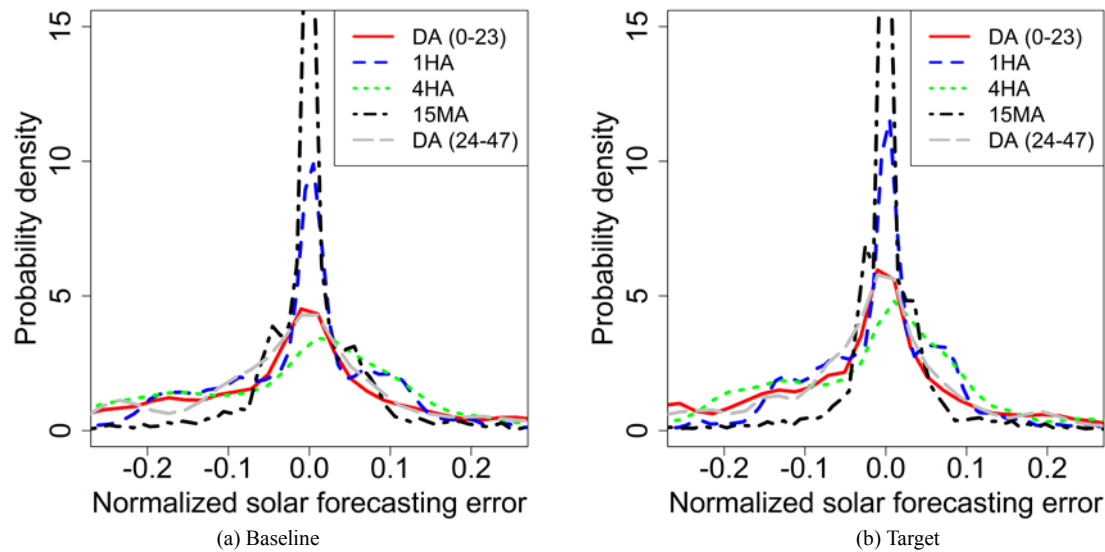

(a) Baseline (b) Target

Figure 11. Distribution of baseline and target solar power forecast errors at DA, 4HA, 1HA, and 15MA forecast horizons (TEP)

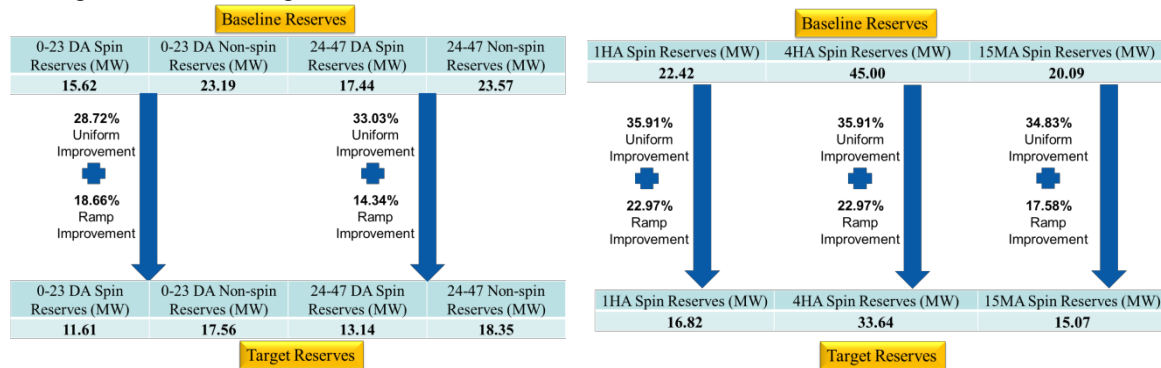Table 7. Baseline and target metrics values for TEP at different forecast horizons

Metrics	DA (24-47) Baseline	DA (24-47) Target	DA (0-23) Baseline	DA (0-23) Target	4HA Baseline	4HA Target	1HA Baseline	1HA Target	15MA Baseline	15MA Target
Correlation coefficient	0.63	0.78	0.59	0.77	0.65	0.78	0.85	0.91	0.94	0.97
RMSE (MW)	5.30	3.99	5.82	4.21	5.00	3.68	3.12	2.34	2.04	1.55
NRMSE by capacity	0.21	0.16	0.23	0.17	0.20	0.15	0.12	0.09	0.08	0.06
MaxAE (MW)	18.09	15.64	22.57	18.60	17.86	13.75	15.66	13.00	16.60	13.86
MAE (MW)	3.56	2.67	4.03	2.94	3.62	2.67	2.09	1.58	1.11	0.78
MAPE by capacity	0.14	0.11	0.16	0.12	0.14	0.11	0.08	0.06	0.04	0.03
MBE (MW)	-1.23	-0.85	-1.60	-1.07	-0.34	-0.27	-0.21	-0.15	0.04	0.03
KSIPer (%)	180.43	160.49	232.49	193.20	112.23	80.25	75.36	49.03	40.29	23.13
OVERPer (%)	99.06	81.51	148.15	112.13	36.26	17.93	14.28	6.33	12.31	1.52
Standard dev. (MW)	5.16	3.90	5.59	4.08	5.00	3.68	3.12	2.33	2.04	1.55
4RMQE (MW)	7.80	5.90	8.20	5.95	7.16	5.24	4.92	3.73	3.94	3.15
N4RMQE by capacity	0.31	0.24	0.33	0.24	0.29	0.21	0.20	0.15	0.16	0.13
95th percentile (MW)	12.62	9.44	12.58	9.21	11.09	8.28	6.59	5.06	4.60	3.46
Renyi entropy	4.99	4.79	4.92	4.75	5.18	5.12	4.19	4.28	3.02	2.69
NRMSE by clear sky power	0.34	0.25	0.39	0.28	0.30	0.22	0.19	0.14	0.13	0.10
MAPE by clear sky power	0.23	0.17	0.27	0.20	0.22	0.16	0.13	0.10	0.07	0.05

3.2.4. GMP Baseline and Target Metrics Values

GMP has relatively high solar penetration compared to CAISO and ISO-NE. Currently there is approximately 47 MW PV installed behind the meter, which represents about 5% of peak load in the GMP region. GMP has goals to make renewables 20% of total annual energy by 2017 and 90% by 2050. The actual and forecasted data used in this case are between 03-May-2013 and 30-Oct-13. The day-ahead forecasts without linear regression at GMP have large errors (23% NRMSE by capacity for both 1- and 2-day ahead forecasts). This is likely due to the highly variable clouds, changing weather, and mountains in the region. The day-ahead forecasts were significantly improved after the first order machine learning (linear regression). Table 8 shows

the baseline and target values at different forecast horizons. It is shown that most target metrics are improved from baselines.

Figure 12 shows the baseline and target reserves. For all forecast horizons, there are more uniform improvements than the ramp forecasting improvements. Figures 13(a) and 13(b) illustrate the distributions of solar power forecast errors for baseline and target forecasting, respectively. The 4HA forecast tends to underforecast the power generation compared to other forecast horizons, which might be due to morning clouds in the region and the shading from mountains.

(a) Day-ahead forecasts at GMP

(b) 15MA, 1HA, and 4HA forecasts at GMP

Figure 12. Target reserves values based on uniform and ramp forecasting improvement (GMP)
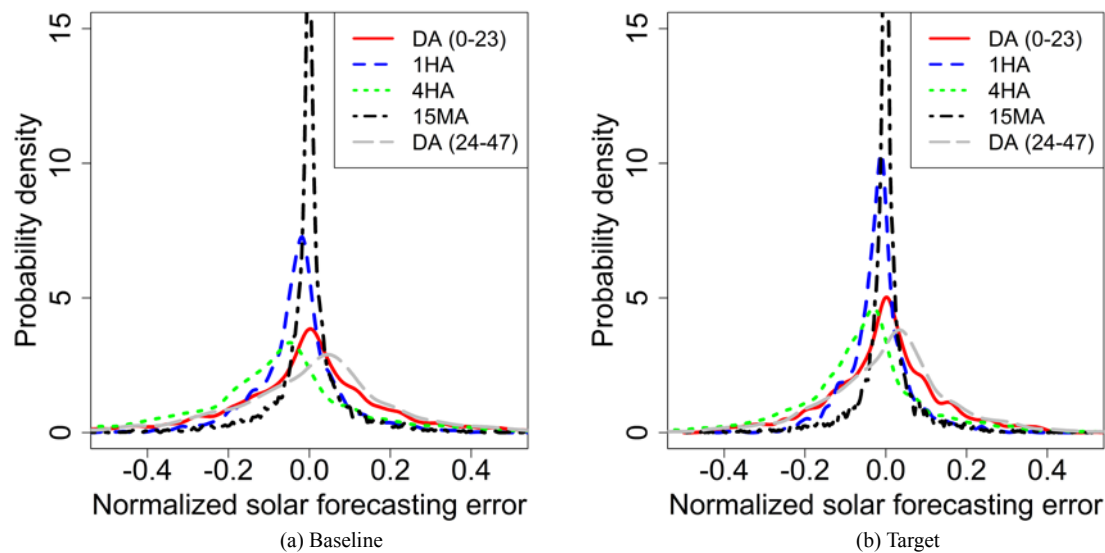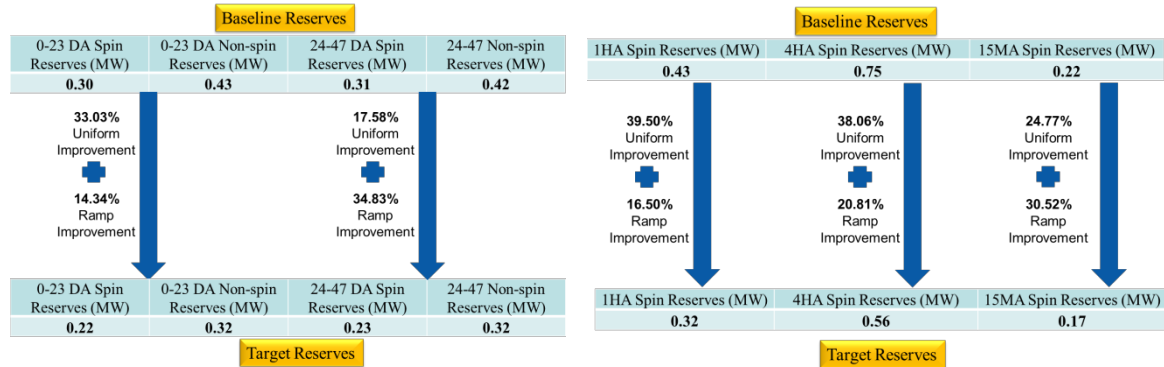

(a) Baseline

(b) Target

Figure 13. Distribution of baseline and target solar power forecast errors at DA, 4HA, 1HA, and 15MA forecast horizons (GMP)

Table 8. Baseline and target metrics values for GMP at different forecast horizons

Metrics	DA (24-47) Baseline	DA (24-47) Target	DA (0-23) Baseline	DA (0-23) Target	4HA Baseline	4HA Target	1HA Baseline	1HA Target	15MA Baseline	15MA Target
Correlation coefficient	0.67	0.82	0.72	0.85	0.66	0.80	0.91	0.95	0.94	0.97
RMSE (MW)	9.44	7.19	8.63	6.47	10.87	8.02	5.21	3.83	4.29	3.23
NRMSE by capacity	0.20	0.15	0.18	0.14	0.23	0.17	0.11	0.08	0.09	0.07
MaxAE (MW)	38.10	30.06	30.05	24.45	45.43	35.00	29.10	22.42	31.16	23.81
MAE (MW)	7.03	5.35	6.21	4.69	7.89	5.74	3.64	2.64	2.42	1.73
MAPE by capacity	0.15	0.11	0.13	0.10	0.17	0.12	0.08	0.06	0.05	0.04
MBE (MW)	0.07	0.21	-0.36	-0.21	-3.76	-2.68	-1.25	-0.90	-0.07	-0.04
KSIPer (%)	138.02	147.06	108.81	119.86	213.14	148.51	79.73	57.25	10.06	12.87
OVERPer (%)	63.21	67.18	32.46	42.56	126.45	62.94	13.99	1.19	0.00	0.00
Standard dev. (MW)	9.45	7.19	8.63	6.47	10.20	7.56	5.06	3.73	4.29	3.23
4RMQE (MW)	13.42	10.25	12.28	9.17	15.91	11.98	8.00	5.99	7.77	6.09
N4RMQE by capacity	0.28	0.22	0.26	0.19	0.34	0.25	0.17	0.13	0.16	0.13
95th percentile (MW)	20.38	15.31	19.51	14.51	23.70	17.32	11.38	8.55	10.04	7.53
Renyi entropy	5.33	5.24	5.34	5.31	4.95	4.86	4.56	4.45	3.40	3.17
NRMSE by clear sky power	0.34	0.26	0.35	0.27	0.41	0.30	0.21	0.15	0.18	0.13
MAPE by clear sky power	0.25	0.19	0.26	0.19	0.29	0.21	0.14	0.10	0.10	0.07

3.2.5. Smyrna Baseline and Target Values

The developed reserve-based method to determine target forecasts is more suitable for large regions (such as ISOs and utilities), since it is not common to hold reserves at individual solar plant level. However, the developed method is also applicable to individual plant, and this Smyrna case study is to further evaluate the effectiveness the developed method. Smyrna site has a 1 MW PV installation. The actual and forecasted data used in this case are between 03-May-2013 and 30-Oct-13. The baseline and target metrics are summarized in Table 9. It is observed that the 1- and 2-day ahead baseline forecasts have similar levels of errors, e.g., 17% NRMSE by capacity, 0.17 standard deviation, etc. The baseline and target reserves are shown in Fig. 14. Figures 15(a) and 15(b) illustrate the distributions of solar power forecast errors for baseline and target forecasting, respectively.

(a) Day-ahead forecasts at Smyrna   (b) 15MA, 1HA, and 4HA forecasts at Smyrna

Figure 14. Target reserves values based on uniform and ramp forecasting improvement (Smyrna)
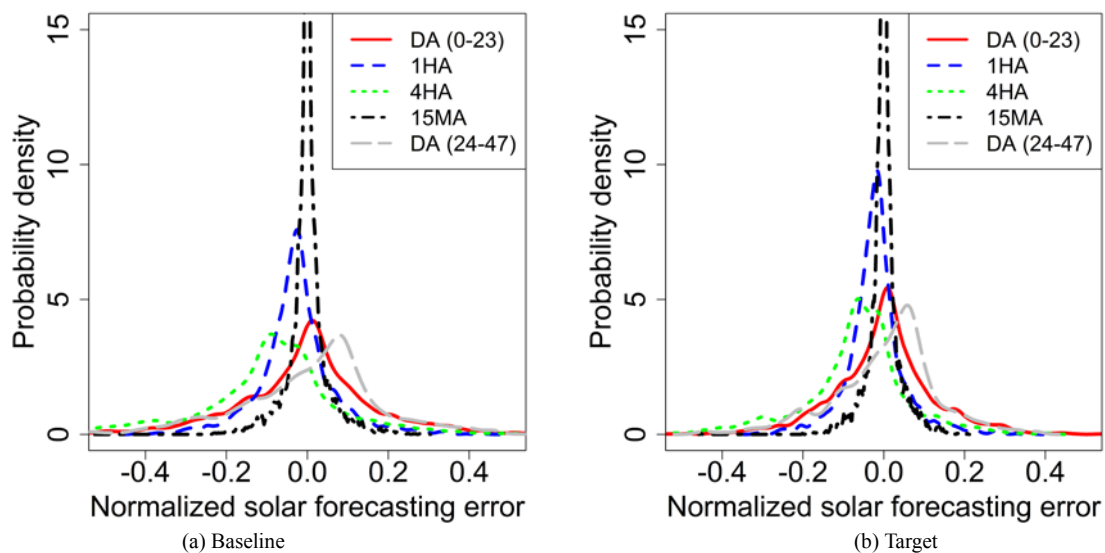

(a) Baseline   (b) Target

Figure 15. Distribution of baseline and target solar power forecast errors at DA, 4HA, 1HA, and 15MA forecast horizons (Smyrna)

Table 9. Baseline and target metrics values for Smyrna at different forecast horizons

Metrics	DA (24-47) Baseline	DA (24-47) Target	DA (0-23) Baseline	DA (0-23) Target	4HA Baseline	4HA Target	1HA Baseline	1HA Target	15MA Baseline	15MA Target
Correlation coefficient	0.71	0.85	0.72	0.85	0.70	0.83	0.91	0.95	0.92	0.96
RMSE (MW)	0.17	0.13	0.17	0.12	0.19	0.14	0.10	0.07	0.05	0.04
NRMSE by capacity	0.17	0.13	0.17	0.12	0.19	0.14	0.10	0.07	0.05	0.04
MaxAE (MW)	0.56	0.46	0.65	0.55	0.72	0.57	0.44	0.36	0.65	0.45
MAE (MW)	0.13	0.10	0.12	0.09	0.14	0.11	0.07	0.05	0.03	0.02
MAPE by capacity	0.13	0.10	0.12	0.09	0.14	0.11	0.07	0.05	0.03	0.02
MBE (MW)	0.01	0.00	0.00	0.00	-0.09	-0.07	-0.03	-0.02	0.00	0.00
KSIPer (%)	148.33	145.36	105.06	117.06	261.30	190.75	87.89	65.30	5.73	12.97
OVERPer (%)	68.59	64.43	29.19	38.65	172.13	105.64	15.65	3.20	0.00	0.00
Standard dev. (MW)	0.17	0.13	0.17	0.12	0.17	0.13	0.09	0.07	0.05	0.04
4RMQE (MW)	0.24	0.18	0.24	0.18	0.28	0.21	0.15	0.11	0.09	0.06
N4RMQE by capacity	0.24	0.18	0.24	0.18	0.28	0.21	0.15	0.11	0.09	0.06
95th percentile (MW)	0.37	0.27	0.36	0.27	0.43	0.32	0.22	0.17	0.11	0.08
Renyi entropy	5.48	5.43	5.18	5.00	5.19	5.07	4.81	4.73	3.13	3.24
NRMSE by clear sky power	0.32	0.24	0.33	0.25	0.35	0.26	0.19	0.14	0.10	0.07
MAPE by clear sky power	0.24	0.18	0.24	0.18	0.26	0.19	0.13	0.10	0.06	0.04

*3.3. Discussion on Baseline and Target **PV Power** Forecasting*

The baseline and target solar forecasting methodology were evaluated and applied to three photovoltaic plants (TEP, GMP, and Smyrna) and two ISOs (CAISO and ISO-NE). It is a complex task to determine the economic benefits of improved solar power forecasting. The proposed reserve-based methodology is arguably a reasonable methodology which can approximately quantify the economic benefits gained from improved solar forecasting. With the 25% reduction in reserves, other metrics were significantly improved. For example, at CAISO, the RMSE of day-ahead forecasts was reduced by 28.7%; the MAE was reduced by 27.2%; and the MBE was reduced by 20.6%. At ISO-NE, the RMSE of day-ahead forecasts was reduced by 24.1%; the MAE was reduced by 25.5%; and the MBE was reduced by 18.4%. These are significant to power system operators. Indeed one pathway to enable such needed improvement in forecasting is via machine learning technology. Towards this end, as part of the project work performed under the SunShot Initiative's Improving the Accuracy of Solar Forecasting program, a system for improving solar forecast, Watt-sun, is being developed by this team (Lu et al., 2015a; 2015b). Watt-sun uses big-data information processing technologies and applies machine-learnt, situation-dependent blending of multiple models to enhance system intelligence, adaptability and scalability. The baseline and target metrics developed in this paper are used to guide the development of the Watt-sun system.

The 25% reduction in spinning reserve would be significant and constitutes an ambitious goal. The annual hourly average price of spinning reserves is $10.11 per MW in 2006 (CAISO, 2007). Using the proposed reserve reduction methodology, the day-ahead (24-47 hours) spinning reserve reduction is 57.29 MW for CAISO. With the improved solar forecasting, the annual cost saving of day-ahead spinning reserves would be approximately 5 million dollars. It is important to note that previous studies (Hodge et al., 2015) have found that reserve savings are only 5-10% of total savings from improved unit commitment and economic dispatch. The financial baseline and targets can be translated back to forecasting accuracy metrics and requirements, which will guide research on solar forecasting improvements towards the areas that are most beneficial to power systems operations.

By comparing Tables 5–9, the forecast errors (e.g., NRMSE, MAPE, etc.) over large areas (CAISO and ISO-NE) are smaller than those for individual solar plant. Such small errors are likely an indication that geographic averaging across a large region cancels the errors, showing the smoothing effect from geographic diversity for solar. Additionally, geographic smoothing over large areas means that large errors are less common in balancing areas or interconnection size areas.

4. Conclusion

The development of baseline and target values for solar forecasting is closely related to the objective of quantifying the economic benefit of solar forecasting, around which currently the industry has no consensus. This is not only because of the complicated hierarchy and structure of the electrical energy market, but also because of the lack of in-depth understanding about how forecast information may fit into the specific utility or ISO operational practices, which are often complex and unique to each organization. Our development of baseline values and target economic metrics for quantifying the benefits of the solar forecast system has been based on close collaboration with utility and ISO partners.

Although solar forecasts are likely to have a multitude of economic benefits, the industry agrees that improved solar forecast accuracy will lead to a reduction in the amount of minimum reserves that must be carried to accommodate the uncertainty of solar power output. Toward this end, we have provided a methodology (Figure 3) to derive the required uniform and ramp forecast improvements to provide a certain percent reduction in spinning reserve costs. In this paper the example of 25% reduction is used, but this was for explanation purposes only. Further, the cost reduction need not be applied to spinning reserves only because other costs can be added. The step-by-step procedure in Figure 3 should still be followed, only with the new cost function. For the test cases in this paper, the actual amount of reduction in spinning and non-spinning reserves could have substantial financial impact: we note that even at present, the amount of reserve reduction for CAISO will be several hundred MW, which will correspond to an annual savings of millions of dollars. The savings will continue to grow in the next years as the level of solar power penetration increases in the region.

Acknowledgements

References

Bristol, E., 1990. Swinging door trending: adaptive trend recording? ISA National Conference Proceedings, 749–753.

CAISO, Ancillary Service Markets, Department of Market Monitoring – California ISO, April 2007.

Chu, Y., Nonnenmacher, L., Inman, R. H., Liao, Z., Pedro, H. T. C., Coimbra, C. F. M., 2014. A smart image-based cloud detection system for intra-hour solar irradiance forecasts. Journal of Atmospheric and Oceanic Technology, 31, 1995–2007.

Diagne, M., David, M., Lauret, P., Boland, J., Schmutz, N., 2013. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews, 27, 65–76.

Duffie, J. A., Beckman, W. A., 2006. Solar Engineering of Thermal Processes. 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc.

Ela, E., Milligan, M., Kirby, B., 2011. Operating Reserves and Variable Generation. NREL/TP-5500-51978. Golden, CO: National Renewable Energy Laboratory.

Erbs, D. G., Klein, S. A., Duffie, J. A., 1982. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. Solar Energy, 28, 293–302.

Germann, U., Zawadzki, I., Turner, B., 2006. Predictability of precipitation from continental radar images Part IV: limits to prediction. Journal of the Atmospheric Sciences, 63(8), 2092–2108.

Florita, A., Hodge, B.-M., Orwig, K., 2013. Identifying wind and solar ramping events. IEEE 5th Green Technologies Conference, Denver, Colorado.

Hodge, B.-M., Florita, A., Sharp, J., Margulis, M., and Mcreavy, D., 2015. The Value of Improved Short-Term Wind Power Forecasting. NREL/TP-5D00-63175. Golden, CO: National Renewable Energy Laboratory.

Forrester, A., Sobester, A., Keane, A., 2008. Engineering design via surrogate modelling: a practical guide. Wiley, New York.

Inman, R. H., Pedro, H. T. C., Coimbra, C. F. M., 2013. Solar forecasting methods for renewable energy integration. Progress in Energy and Combustion Science, 39(6), 535–576.

International Energy Agency (IEA), 2013. Photovoltaic and solar forecasting: State of the art. Tech. Rep. IEA PVPS T14-01:2013, Paris, France.

Justice, C. O. et al., 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. IEEE Transactions on Geoscience and Remote Sensing, 36(4), 1228–1249.

Key, J., Schweiger, A. J., 1998. Tools for atmospheric radiative transfer: Streamer and FluxNet. Computers and Geosciences, 24(5), 443–451.

King, D. L., Kratochvil, J. A., Boyson, W. E., 2004. Photovoltaic array performance model. SAND2004-3535, Albuquerque, NM: Sandia National Laboratories.

Lara-Fanego, V., Ruiz-Arias, J., Pozo-Vázquez, D., Santos-Alamillos, F., Tovar-Pescador, J., 2012. Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain). Solar Energy, 86 (8), 2200-2217.

Lave, M., Kleissl, J., 2010. Solar Variability of Four Sites across the State of Colorado. Renewable Energy, 35(12), 2867–2873.

Lew, D. et al., 2010. Western Wind and Solar Integration Study. NREL/SR-550-47781. Golden, CO: National Renewable Energy Laboratory.

Lew, D. et al., 2013. The Western Wind and Solar Integration Study Phase 2. NREL/TP-5500-55888. Golden, CO: National Renewable Energy Laboratory.

Lorenz, E. et al., 2009. Benchmarking of different approaches to forecast solar irradiance. in Proc. 24th European Photovoltaic Solar Energy Conference, 4199–4208.

Lorenz, E., Kühnert, J., Heinemann, D., 2012. Short Term Forecasting of Solar Irradiance by Combining Satellite Data and Numerical Weather Predictions. In Proceedings of 27th EUPVSEC Frankfurt.

Lorenz, E., Kühnert, J., Heinemann, D., 2014. Overview of Irradiance and Photovoltaic Power Prediction, In: Weather Matters for Energy, Editors: A. Troccoli, L. Dubus, S.E. Haupt (Springer 2014).

Lorenz, E., Scheidsteger, T., Hurka, J., Heinemann, D., Kurz, C., 2011. Regional PV power prediction for improved grid integration. Progress in Photovoltaics: Research and Applications, 19(7), 757–771.

Lorenzm E., Heinemann, D., Kurz, C., 2012. Local and regional photovoltaic power prediction for large scale grid integration: Assessment of a new algorithm for snow detection. Progress in Photovoltaics: Research and Applications, 20(6), 760–769.

Lu, S., Hwang, Y., Khabibrakhmanov, I., Marianno, F. J., Shao, X., Zhang, J., Hodge, B.-M., and Hamann, H. F., 2015a. Machine Learning Based Multi-Physical-Model Blending for Enhancing Renewable Energy Forecast Improvement via Situation Dependent Error Correction. European Control Conference, Linz, Austria.

Lu, S., Hwang, Y., Khabibrakhmanov, I., Dang, H., van Kessel, T., Marianno, F., and Hamann, H. F., 2015b. Machine Learning based Multi-Model Blending for Enhancing Renewable Energy Forecasting. 95th American Meteorological Society Annual Meeting, Phoenix, Arizona.

Marion, B., Anderberg, M., George, R., Gray-Hann, P., Heimiller, D., 2001. PVWATTS version 2 – enhanced spatial resolution for calculating grid-connected PV performance. Proceedings of the NCPV Program Review Meeting, Lakewood, CO.

Marquez, R., Coimbra, C. F. M., 2013. Intra-hour DNI forecasting based on cloud tracking image analysis. Solar Energy, 91, 327–336.

Mathiesen, P., Kleissl, J., 2011. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental united states. Solar Energy, 85(5), 967–977.

Margolis, R., Coggeshall, C., Zuboy, J., 2012. SunShot Vision Study. U.S. Department of Energy. Washington, D.C..

Mesinger, F. et al., 2006. North American regional reanalysis. Bulletin of the American Meteorological Society, 87(3), 343–360.

Paulescu, M., Paulescu, E., Gravila, P., Badescu, V., 2013. Weather Modeling and Forecasting of PV Systems Operation. Springer, London.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renne, D., Hoff, T. E., 2010. Validation of short and medium term operational solar radiation forecasts in the U.S.. Solar Energy, 84(12), 2161–2172.

Perez, R., Lorenz, E., Pelland, S., Beauharnois, M., Van Knowe, G., HemkerJr, K., Heinemann, D., Remund, J., Müller, S. C., Traunmüller, W., et al., 2013. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. Solar Energy 94, 305-326.

Stein, J. S., 2012. The Photovoltaic Performance Modeling Collaborative (PVPMC). 38th IEEE Photovoltaic Specialists Conference (PVSC), Austin, Texas.

Zhang, J., Chowdhury, S., Zhang, J., Messac, A., Castillo, L., 2013. Adaptive Hybrid Surrogate Modeling for Complex Systems. AIAA Journal, 51(3), 643–656.

Zhang, J., Florita, A., Hodge, B.-M., Lu, S., Hamann, H. F., Banunarayanan, V., Brockway, A. M., 2015. A Suite of Metrics for Assessing the Performance of Solar Power Forecasting. Solar Energy, 111, 157–175.