

Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest



David P. Larson, Lukas Nonnenmacher, Carlos F.M. Coimbra*

Department of Mechanical and Aerospace Engineering, Jacobs School of Engineering, Center of Excellence in Renewable Energy Integration and Center for Energy Research, University of California San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:

Received 2 July 2015

Received in revised form

30 November 2015

Accepted 8 January 2016

Available online xxx

Keywords:

Forecasting

Solar power output

Photovoltaics

Numerical weather prediction

Interannual performance variability

ABSTRACT

A forecasting method for hourly-averaged, day-ahead power output (PO) from photovoltaic (PV) power plants based on least-squares optimization of Numerical Weather Prediction (NWP) is presented. Three variations of the forecasting method are evaluated against PO data from two non-tracking, 1 MWp PV plants in California for 2011–2014. The method performance, including the inter-annual performance variability and the spatial smoothing of pairing the two plants, is evaluated in terms of standard error metrics, as well as in terms of the occurrence of severe forecasting error events. Results validate the performance of the proposed methodology as compared with previous studies. We also show that the bias errors in the irradiance inputs only have a limited impact on the PO forecast performance, since the method corrects for systematic errors in the irradiance forecast. The relative root mean square error (RMSE) for PO is in the range of 10.3%–14.0% of the nameplate capacity, and the forecast skill ranges from 13% to 23% over a persistence model. Over three years, an over-prediction of the daily PO exceeding 40% only occurs twice at one of the two plants under study, while the spatially averaged PO of the paired plants never exceeds this threshold.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

High renewable energy penetration grids are challenging to balance due to inherently variable generation weather-dependent energy resources. Solar and wind forecasting are proven methods for mitigating resource uncertainty and reducing the need for scheduling of ancillary generation. Several forecasting methodologies have been developed to target different forecast time horizons [1].

Short-term forecasting ($t < 1$ h) is mainly based on sky imaging techniques and time-series models, while satellite-based forecasts achieve usable results for time horizons of 1–6 h [2–4]. Usually, horizons larger than 6 h require numerical weather prediction models (NWP) to generate accurate results, although there are exceptions such as [5]. Recent advances in solar forecasting have mainly covered intra-day resource forecasts, driven by advances in sky image techniques and time-series modeling [6–16].

While intra-day forecasts are important for grid stability, day-ahead forecasts are critical for market participation and unit

commitment. Current market regulations in many world regions require day-ahead forecasts [17–19] and there are financial incentives to produce accurate forecasts [18]. Besides market participation, day-ahead forecasts can also be useful for energy storage dispatch [20,21]. In this paper, we focus on day-ahead forecasting of power output from photovoltaic power plants in the American Southwest.

1.1. Previous work

Most NWP models generate global horizontal irradiance (GHI) at the ground level as one of the outputs, with some newer generation models now including direct normal irradiance (DNI). Previous studies evaluated the accuracy of this GHI variable and suggested ways to improve it, which have mostly focused on spatial averaging and bias removal [22–28]. [22,27,29] showed that forecast errors for all sky conditions can be reduced by averaging the GHI forecasts from all NWP grid points within a set distance of the target site. In addition [22], showed that the forecast performance could be further improved through bias removal using a polynomial regression based on the solar zenith angle and clear sky index. [27] showed that a similar improvement can be achieved

* Corresponding author.

E-mail address: ccoimbra@ucsd.edu (C.F.M. Coimbra).

through Kalman filtering.

In contrast to our knowledge of the performance of day-ahead GHI predictions, the application of these forecasts directly to prediction of day-ahead power output (PO) of PV plants is poorly understood. There are currently less than a dozen published studies that cover the subject of day-ahead PV PO forecasts, the majority of which were published in the last 5 years. We summarize these papers below and in Table 1. It is important to note that this lack of knowledge is partly due to the difficulty of obtaining data from operational PV plants, due to security restrictions and lack of data infrastructure. However, data access should improve in the coming years due to energy policies that require PO forecasting and therefore necessitate the collection of PO data.

[30] applied the Danish Meteorological Institute's High Resolution Limited Area Model (HIRLAM) to forecast PO of 21 PV systems in Denmark. The PV systems had a total rated power capacity of 1–4 kWp each and one year of data was used to evaluate the forecasts. HIRLAM GHI forecasts were used as the main input to autoregressive-based PO models. [31] also used HIRLAM to forecast PO of a 6.2 kWp test site in Spain.

[29] and [32] used the European Center for Medium-Range Weather Forecasts Model (ECMWF) to generate regional, PO forecasts in Germany. [29] only evaluated two months (April and July 2006) of forecasts for 11 PV systems in Southern Germany, while [32] tested on 10 months (July 2009–April 2010) of PO data from approximately 400 representative PV systems.

[33] used an artificial neural network (ANN) to predict GHI in Trieste, Italy. The predicted GHI was then mapped directly to PO using the efficiency data of the studied 20 kWp PV system. Unfortunately, PO forecast results were only reported for four consecutive clear sky days in May 2009. Similarly [34], evaluated an autoregressive-moving-average model with exogenous inputs (ARMAX) forecast model that did not use NWP data as input. However [34], only forecasted the day-ahead mean daily PO, rather than day-ahead hourly values, for a 2.1 kWp PV system in China.

[27] forecasted PO of three small PV systems (6.72, 19.8 and 45.6 kWp) in mainland Canada, using the Canadian Meteorological Centre's Global Environmental Multiscale Model (GEM). The GEM's GHI forecasts were validated against ground measurements from the SURFRAD Network in the United States. Spatial averaging of 300 km–600 km and bias removal via Kalman filtering were used to improve the GEM forecast performance. Reported RMSE values were in the range of 6.4%–9.2% of the rated power of the PV systems for a 1 year testing set (April 2008–March 2009).

[35] and [36] generated regional forecasts for Japan using Support Vector Regression (SVR) together with inputs from the Japan Meteorological Agency's Grid Point Value-Mesoscale Model

(GPV-MSM). [35] used PO data from approximately 450 PV plants from four regions (Kanto, Chubu, Kansai, Kyushu) that had a net rated power of approximately 15 MWp while [36] used data from 273 PV plants spread over two regions (Kanto and Chubu), with an approximate total power of 8 MWp.

[37] and [38] forecasted PO of 28 PV plants in mainland France using Meteo France's Action de Recherche Petite Echelle Grande Echelle (ARPEGE) model. [37] presented a deterministic forecast that achieved a RMSE of 8–12% of the plant capacity over two years of testing data (2009–2010) while [38] presented a probabilistic forecast. Both methods used 31 variables from the ARPEGE, including GHI as well as environmental conditions, e.g., temperature, humidity and precipitation. Unfortunately, no actual values were reported for the PV plant power ratings or other technical details, which limits analysis into the applicability of the results to other PV plants and regions.

Most recently [39], forecasted the PO of five vertical-axis tracking PV plants in Spain, using a nonparametric model based on the Weather Research and Forecasting Model (WRF). The PV plants ranged in size from 775 to 2000 kWp and the forecasts were evaluated with data from 2009 to 2010. Quantile Regression Forests, a variation of random forests, was used to generate the PO forecasts, with WRF variables such as GHI, temperature and cloud cover as the inputs.

Although these studies all presented day-ahead PO forecasting for PV plants, further research is still required, especially for sites in the United States. In this study we seek to provide the following contributions: (1) Introduction and evaluation of a PV PO forecast model for the American Southwest, a region with both high solar energy generation potential and a favorable political environment for solar, especially in California. (2) Investigation of the interannual forecast performance variability and (3) the occurrence of severe forecasting errors, as they relate to large-scale renewable energy integration. (4) Spatial smoothing through pairing of PV plants in proximity. To achieve these goals, we systematically test approaches and inputs to generate PO forecasts based on two operational PV plants in Southern California.

2. Data

2.1. Ground data

Two sites are used to evaluate day-ahead forecasts of GHI and PO: Canyon Crest Academy (CCA) and La Costa Canyon High School (LCC). CCA (32.959°, –117.190°) and LCC (33.074°, –117.230°) are both located in San Diego County, USA, and each feature 1 MW peak (MWp) of non-tracking photovoltaic (PV) panels. The PV panels for

Table 1
Summary of the current literature on day-ahead PO forecasting. "Type" refers to whether the forecast was for a single PV plant site or for an entire region while "Testing data" is how much data was used in the forecast model evaluation.

| | Region | NWP model | Type | PV plants | | |
|------|---------|-----------|----------|--------------|----------------|--------------|
| | | | | No. of sites | Total capacity | Testing data |
| [30] | Denmark | HIRLAM | Point | 21 | ~100 kWp | 1 year |
| [31] | Spain | HIRLAM | Point | 1 | 6 kWp | 10 months |
| [29] | Germany | ECMWF | Regional | 11 | Unknown | 2 months |
| [32] | Germany | ECMWF | Regional | 383 | ~100 MWp | 10 months |
| [35] | Japan | GPV-MSM | Regional | 454 | ~10 MWp | 1 year |
| [36] | Japan | GPV-MSM | Regional | 273 | ~10 MWp | 1 year |
| [27] | Canada | GEM | Point | 3 | ~100 kWp | 2 years |
| [37] | France | ARPEGE | Point | 28 | Unknown | 2 years |
| [38] | France | ARPEGE | Point | 28 | Unknown | 2 years |
| [39] | Spain | WRF | Point | 5 | ~10 MWp | 2 years |
| [33] | Italy | N/A | Point | 1 | 20 kWp | 4 days |
| [34] | China | N/A | Point | 1 | 2 kWp | 6 months |

both sites were installed in 2011 as part of a power purchase agreement (PPA) between the San Dieguito Union High School District and Chevron Energy Solutions.

Each site has two inverters, with the PV panels mounted above the parking lots at a 5° incline. At CCA, the panels are split over two parking lots, approximately 200 m apart and each with 500 kWp. Due to the site's characteristics, the north parking lot panels face (~20°) southwest, while the south parking lot panels face directly south. Meanwhile, the panels at LCC cover a single parking lot and face ~30° southwest.

Total PO from the PV panels at each site is available as 15-min averages. Additionally, 15-min averaged GHI data is available from co-located pyranometers. Fig. 1 illustrates the data availability for both sites. For this study, the GHI and PO data are backwards-averaged to hourly values and night values are removed, where we define night as zenith angles (θ_z) greater than 85°. Note that the maximum observed PO of CCA matches the nameplate capacity (1 MWp), while the maximum observed PO of LCC peaks at ≈0.8 MW, despite both sites using identical PV technologies. This difference is due to the greater (~30°) southwest orientation of the PV panels at LCC.

2.2. NWP data

We use forecasted variables from two publicly-available NWP models: the North American Mesoscale Forecast System (NAM) and the Regional Deterministic Prediction System (RDPS).

2.2.1. NAM

The NAM is a NWP model provided by the National Oceanic and Atmospheric Administration (NOAA) on a ≈12 km × 12 km spatial grid that covers the continental United States. Forecasts are generated four times daily at 00Z, 06Z, 12Z and 18Z, with hourly temporal resolution for 1–36 h horizons and 3 h resolution for 39–84 h horizons. Downward shortwave radiative flux (DSWRF) [W/m^2] at the surface, a synonym for GHI, is forecasted using the Geophysical Fluid Dynamics Laboratory Shortwave radiative transfer model (GFDL-SW) [40]. The NAM also forecasts total cloud cover (TCDC) [%], where the entire atmosphere is treated as a single layer.

In this study, NAM forecasts generated at 00Z were downloaded from the NOAA servers and degribbed for September 2013 to November 2014, for all nodes within 200 km of CCA and LCC. Fig. 2 shows the distribution of NAM forecast nodes over Southern California.

2.2.2. RDPS

The Canadian Meteorological Centre generates the RDPS model on a 10 km × 10 km spatial grid covering North America. As with the NAM, RDPS forecasts total cloud cover (TCDC) as a percentage for each grid element, not distinguishing between the layers of the atmosphere. The operational RDPS model generates forecasts daily at 00Z, 06Z, 12Z and 18Z with hourly temporal resolution. Only the 00Z forecasts are used in this study, to ensure fair comparison with the NAM forecasts. RDPS GHI forecasts were not available.

3. Forecasting models

3.1. Persistence

We include a day-ahead persistence forecast of both GHI and PO as a baseline. The persistence forecast assumes that the current conditions will repeat for the specified time horizon. In other words:

$$\hat{y}_{t+\tau} = y_t, \quad (1)$$

where $\hat{y}_{t+\tau}$ is the forecasted value of some variable y , at a time horizon of τ from some initial time t . For day-ahead persistence, $\tau = 24 \text{ h}$.

3.2. Deterministic cloud cover to GHI model

We use a modified version of the deterministic model from Refs. [19,41,42] to derive GHI from cloud cover:

$$\text{GHI} = \text{GHI}_{\text{CS}}[0.35 + 0.65(1 - \text{CC})], \quad (2)$$

where CC is the TCDC (0 = clear, 1 = overcast) from the NWP model and GHI_{CS} is the clear sky GHI from the airmass independent model described in Refs. [43–45].

3.3. Effects of spatial averaging

Previous studies have shown that NAM GHI forecast error can be reduced by spatially averaging the forecasts from all nodes within a set distance of the target site [22,27]. The distance (d) in km between a site and a NAM node can be calculated using the spherical law of cosines:

$$d = \arccos[\sin(\phi_1)\sin(\phi_2) + \cos(\phi_1)\cos(\phi_2)\cos(\lambda_2 - \lambda_1)]R, \quad (3)$$

where (θ_i, λ_i) is the latitude–longitude for a point i and R is the

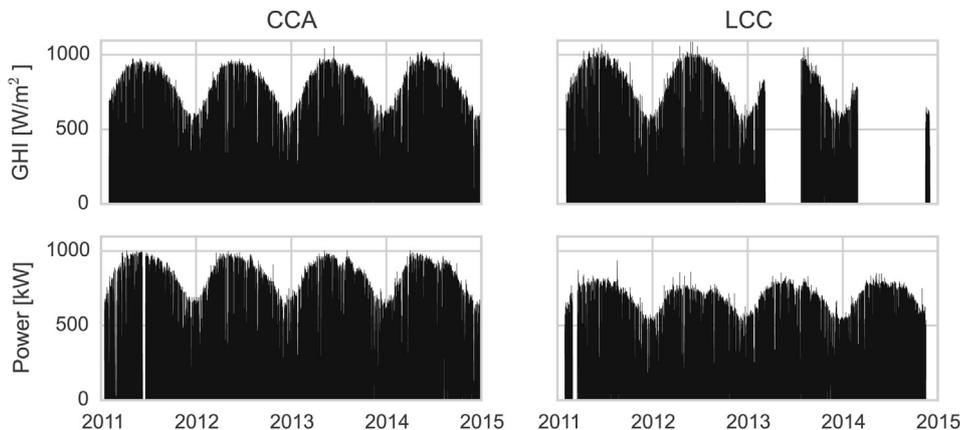


Fig. 1. Availability of ground measurements of GHI and PO for CCA (left column) and LCC (right column). The missing GHI data from LCC is due to a malfunctioning pyranometer.

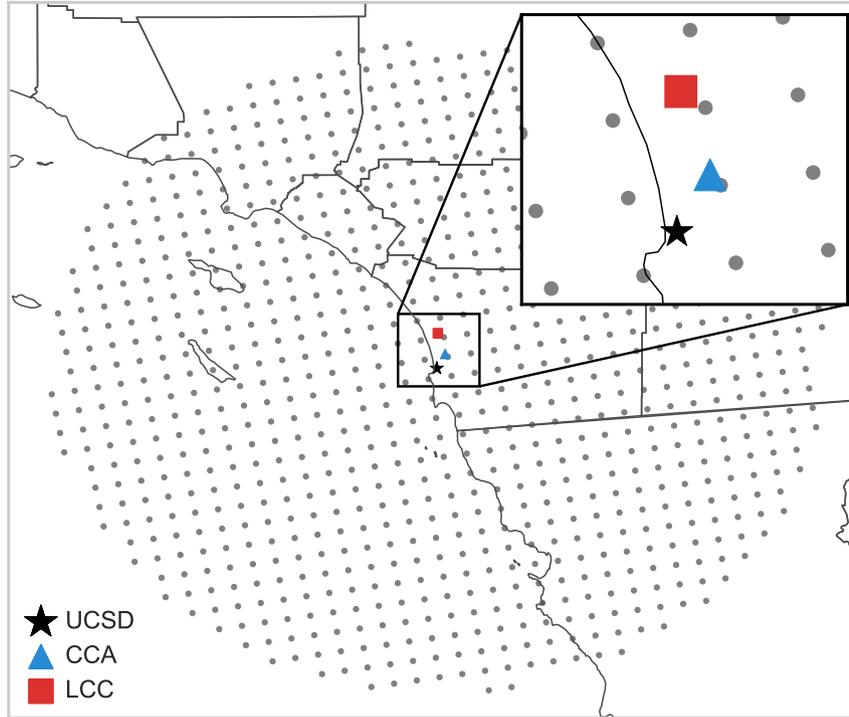


Fig. 2. Map of Southern California, with the NAM grid points marked with gray dots. UC San Diego (UCSD; star), Canyon Crest Academy (CCA; triangle) and La Costa Canyon (LCC; square) are also marked on the map. The area to the lower left is the Pacific Ocean, with Mexico shown in the bottom right.

radius of the Earth (6371 km). For both CCA and LCC, there are approximately 50 NAM nodes with 50 km, 200 nodes within 100 km, and 800 nodes within 200 km.

Fig. 3 shows the effect spatial averaging of the NAM_{GHI} GHI forecasts for CCA compared to using only the NAM node closest to the site, i.e., the naive node choice. The NAM forecasts from the available data set (September 2013–November 2014) are grouped into bins based on the clear sky index (k_t),

$$k_t = \frac{GHI}{GHI_{CS}}, \quad (4)$$

which provides a measure of the sky conditions during the forecasted time period ($k_t \rightarrow 0$: overcast, $k_t \rightarrow ->1$: clear). Then, the root mean square error (RMSE) is calculated between the forecasts in each k_t bin and the ground truth, and averaged to produce one RMSE value per k_t bin. The result is that spatial averaging reduces the RMSE of the forecasts for all sky conditions, as compared to the naive node choice. This is consistent with results presented in previous work, e.g. Refs. [22,27].

Hereafter, we will report results for the naive NAM forecasts (NAM_{GHI} , NAM_{CC}) as well as the 100 km spatially averaged versions (NAM_{GHI}^* , NAM_{CC}^*). 100 km is chosen over 200 km as there is a negligible difference in error between the two. Also, the 100 km case requires approximately 25% the NAM grid points as the 200 km case, and therefore has lower storage and computational requirements.

3.4. GHI to PO model

As the PV panels at both sites are non-tracking, we assume that there exists a linear function $f(\cdot)$ that can map GHI to PO. Fig. 4 illustrates the relationship between GHI and PO for CCA over a full year (September 2013–November 2014), with a reference linear fit included ($R^2 = 0.97$). The high R^2 value (>0.95) indicates that a linear mapping between GHI and PO is not an unreasonable assumption.

Formally, we define a linear function $f(x;w)$ as:

$$f(x;w) = \sum_{j=0}^n w_j x_j = w^T x, \quad (5)$$

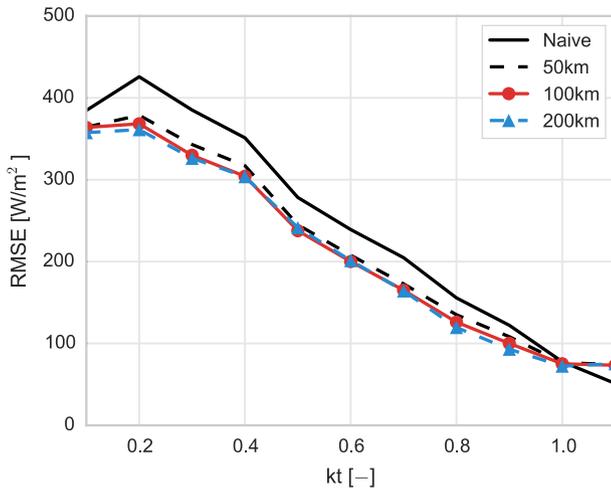


Fig. 3. Effect of spatial averaging the NAM_{GHI} GHI forecasts for CCA September 2013–November 2014. Naive refers to the NAM node physically closest to the site, while 50 km, 100 km, and 200 km are the unweighted average forecasts of all nodes within each search radius. Each marker represents the average RMSE of all forecasts generated for each k_t bin, i.e., each group of sky conditions.

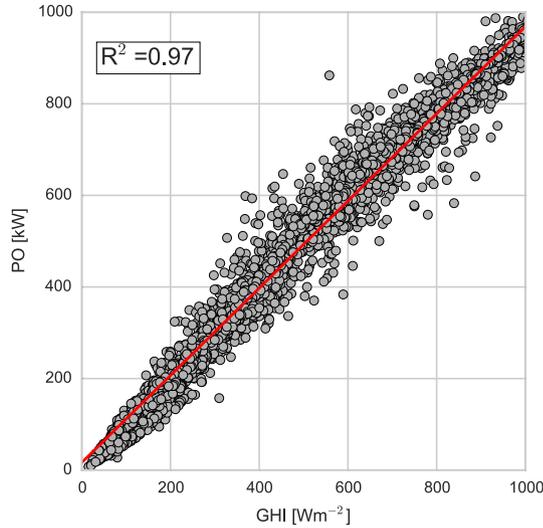


Fig. 4. GHI vs. PO for CCA over September 2013–November 2014, ignoring night values ($\theta_z > 85^\circ$). Each marker is a single data point, and the line is a least-squares fit between GHI and PO with a R^2 value of 0.97.

where w is the weight vector, x is the input vector, $x_0 = 1$ by convention, and n is the number of model parameters. Then we use ordinary least-squares to find the optimal weights w^* , i.e., the weights that minimize the error between the predictor $f(x;w)$ and the true value y .

More specifically, for a given set of m example inputs $\{x^{(1)}, \dots, x^{(m)}\}$ and outputs $\{y^{(1)}, \dots, y^{(m)}\}$, the least-squares costs function is defined as:

$$J(w) = \sum_{i=1}^m (f(x^{(i)}; w) - y^{(i)})^2 = \|Xw - y\|_2^2. \quad (6)$$

The optimal weights w^* are then the w that minimizes the cost function:

$$w^* = \min_w \|Xw - y\|_2^2, \quad (7)$$

where $X \in \mathbb{R}^{m \times n}$, $w \in \mathbb{R}^n$, and $y \in \mathbb{R}^m$.

For the purposes of this study, we use the predicted GHI ($\widehat{\text{GHI}}$) from one of the NWP models and zenith angle (θ_z) as the model inputs. I.e. $n = 2$, $x_1 = \widehat{\text{GHI}}$ and $x_2 = \theta_z$, where θ_z is included so that temporal information is encoded into the model. Readers are referred to [46] and [47] for more detailed studies into modeling of PO of PV power plants.

4. Results and discussion

4.1. Error metrics

We use standard error metrics to evaluate the performance of both the GHI and PO forecasts: Mean Absolute Error (MAE), Mean Bias Error (MBE), Root Mean Square Error (RMSE), and skill (s) [7]:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |e_i|, \quad (8)$$

$$\text{MBE} = \frac{1}{N} \sum_{i=1}^N e_i, \quad (9)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N e_i^2}, \quad (10)$$

$$s = 1 - \frac{\text{RMSE}}{\text{RMSE}_p}, \quad (11)$$

where N is the number of forecasts, RMSE_p is the RMSE of the persistence forecast, and e is the error between the true value (y) and forecasted value (\hat{y}):

$$e_i = y_i - \hat{y}_i. \quad (12)$$

Additionally, for ease of comparison to previously published results, we also include normalized RMSE (nRMSE):

$$\text{nRMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{e_i}{\bar{y}}\right)^2}, \quad (13)$$

where \bar{y} is the mean true value over the considered period, e.g., September 2013–November 2014.

4.2. GHI forecast performance

Day-ahead GHI forecast performance and statistics for CCA for September 2013–November 2014 are shown in Table 2. As noted in Section 3.3, the spatially averaged forecasts ($\text{NAM}_{\text{GHI}}^*$ and NAM_{CC}^*) use all NAM nodes within 100 km of the site. The best GHI forecast is generated with RPDS, which reduces the RMSE 24% as compared to the persistence model. NAM_{CC} performs the worst, reducing RMSE only by 3% compared to persistence. However, the results presented are for GHI forecasts that have not undergone any post-processing or model output statistics (MOS), which have been shown to improve forecast performance (see Refs. [22,25]).

Without post-processing or MOS, the GHI forecasts underperform as compared to prior literature. Two of the most extensive studies on day-ahead GHI forecasting are [27] and [26]. In Ref. [27], the GEM NWP model is evaluated using seven SURFRAD Network sites in the US and achieves RMSE values in the range of ≈ 80 – 120 Wm^{-2} [26]. expands on [27] by providing results for a variety of common NWP models, including GEM, ECMWF and WRF. RMSE values reported in Ref. [26] are in the range of 70–200, 90–130, 80–115, and 95–135 Wm^{-2} for sites in the US, Central Europe, Spain and Canada respectively. Unfortunately, neither study directly evaluated RDPS or NAM.

4.3. Power output forecast performance

PO is forecasted for both CCA and LCC using the methods described in Section 3. For the 18 months of overlapping data

Table 2

Day-ahead GHI forecast performance for CCA, from September 2013 to November 2014. Night values have been removed (night $\equiv \theta_z \geq 85^\circ$). No post-processing or model output statistics (MOS) have been applied to the forecasts. The errors are relative to the mean GHI of the considered period (515 Wm^{-2}). Bold values indicate best performances.

| Method | MAE [%] | MBE [%] | RMSE [%] | s [–] |
|---------------------------------|-------------|------------|-------------|-------------|
| Persistence | 24.9 | 7.7 | 36.2 | – |
| RDPS _{CC} | 20.3 | 2.3 | 27.5 | 0.24 |
| NAM _{GHI} | 21.9 | –11.6 | 31.8 | 0.12 |
| NAM _{GHI} [*] | 20.5 | –7.6 | 28.3 | 0.22 |
| NAM _{CC} | 26.4 | 14.1 | 35.3 | 0.03 |
| NAM _{CC} [*] | 26.2 | 16.2 | 32.7 | 0.10 |

between the ground measurements and NWP models, odd months are used for the training set and even months for the testing set. Splitting the sets in this way, rather than in sequential partitions, ensures that the models are trained on data from all four seasons. Besides this monthly training, the effects of training on a full year of data and testing on another full year are discussed in Section 4.4.

Fig. 5 shows three exemplary days of forecasted and measured PO. The persistence model performs well on the first (clear) day, with no visible difference to the measured PO, while both NWP-based forecasts under-predict the PO. On the second exemplary day, the NWP-based models perform better than on the first day, but do not beat the persistence model. Since the persistence model is just a repetition of the previous day, persistence works well if the weather conditions between days do not change. Hence, persistence also outperforms the NWP models because it inherits a better representation of the temperature effects on the panels that are not an input to the GHI to PO model. However, the strength of the NWP models is to predict cloud cover and irradiance attenuation, and are hence performing better when conditions change between days. The third exemplary day shows this effect, where the persistence model under-predicts since the previous day (not shown) was overcast. These results are expected and consistent with previous findings on the performance of the persistence model [3,24].

Table 3 summarizes the PO forecast results and shows that all three NWP-based forecast models perform comparably for both sites ($s = 19\text{--}25\%$) and overall outperform the persistence model. However, the persistence model has a lower absolute error than the NWP-based models for 47% and 49% of the data points in the testing set for CCA and LCC respectively. This is likely a result of the temperate climate at both sites, i.e., that both sites have an abundance of periods where the weather conditions do not change and therefore the persistence model will excel.

Given the significant variations in skill for the GHI forecasts ($s = 10\text{--}24\%$), it might be surprising that all three PO forecasts perform almost equally well. However, the training of the GHI to PO model removes biases in the forecasts. This seems to be an advantage of our proposed model compared to [27], where the accuracy of the PV output prediction was strongly impacted by the accuracy of the GHI input.

4.4. Interannual performance variability

While all previous work has only focused on bulk error statistics, we also investigate the interannual performance variability. This is important since meteorological studies have shown that there are

relatively strong interannual variations in many weather phenomenon, e.g., DNI [48]. However, this variation is unknown for the evaluation of PV PO. Quantifying this variation will enable us to define what is a sufficient length of training and testing data sets for PO forecasts.

We compare the PO forecast performance per year for 2011–2014. RDPS_{CC} is used as the input source for the GHI to PO model and the model parameters are trained in the same way as the previous section. We use RDPS_{CC} rather than NAM_{GHI} and NAM_{CC} for two reasons. First, we have RDPS_{CC} forecasts for the entirety of 2011–2014, but NAM forecasts only for September 2013–November 2014. Second, as discussed in the previous section, the PO forecast performance is not strongly influenced by the input GHI forecast. Hence, we assume that the interannual RDPS_{CC} forecast results in this section are representative of all evaluated NWP-based models.

Table 4 shows the interannual performance variability of PO forecasts for CCA and LCC over 2011–2014. The PO model is trained separately on 2011, 2012, and 2013, and then tested on the subsequent years of data. For both CCA and LCC, the forecast performance on the test set depends more on the test set itself rather than the training set used. For example, the CCA forecast models trained on 2011, 2012 and 2013 achieved nRMSE values of 0.26, 0.30 and 0.28 respectively on the training set. But all three models achieved nRMSE of 0.24 on the 2014 testing set. A similar trend is seen in the 2013 testing sets and the LCC forecasts.

Although there is no prior literature that reports day-ahead PO forecast performance in our studied region of the American Southwest, it is important to consider our results in the context of other work. As noted in Table 4, our forecast methodology achieves nRMSE values in the range of 0.24–0.33. [30,36,39] reported nRMSE values in the range of 0.08–0.13, 0.26–0.35, and 0.10–0.30 respectively. Each study used at least 1 year of testing data, but the 21 PV systems in Ref. [30] had nameplate capacities of only 1–4 kWp. [36] and [39] evaluated forecasts with PV systems of total capacity ~10 MWp, and achieved nRMSE values closer in range to our results.

Error distributions are another important factor in evaluating forecast performance. Fig. 6 shows the forecast error distributions of three training–testing data sets and three sites: CCA, LCC and a third fictional site denoted as “Combined”. The “Combined” site allows us to consider the impact of pairing the two PV plants to simulate a single, spatially distributed 2 MWp power plant. All three sites and training–testing sets show similar distributions, with peaks centered at less than zero error. 2012 and 2013 have a

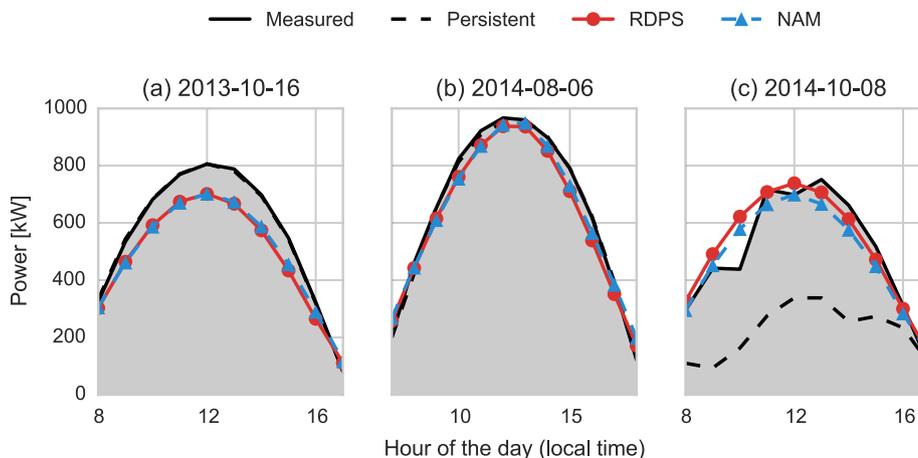


Fig. 5. Sample days where the NWP-based forecasts (RDPS_{CC} and NAM_{GHI}) for CCA perform (a) worse than; (b) as good as; and (c) better than day-ahead persistence.

Table 3

Day-ahead hourly PO forecasts for CCA and LCC, with MAE, MBE and RMSE reported as relative to the rated plant capacity of each site (1 MWp). The models were trained and tested using hourly data from September 2013–November 2014, with odd months used for training and even months for testing.

| Site | Method | Training | | | | Testing | | | |
|------|---------------------------------|----------|---------|----------|-------|------------|------------|-------------|-------------|
| | | MAE [%] | MBE [%] | RMSE [%] | s [–] | MAE [%] | MBE [%] | RMSE [%] | s [–] |
| CCA | Persistent | 9.0 | 0.4 | 15.6 | – | 8.8 | –0.5 | 14.7 | – |
| | RDPS _{CC} | 8.9 | 0.0 | 12.5 | 0.20 | 8.4 | 0.2 | 11.5 | 0.21 |
| | NAM _{GHI} | 9.6 | –0.0 | 13.2 | 0.15 | 8.5 | –0.3 | 11.3 | 0.23 |
| | NAM _{GHI} [*] | 9.3 | 0.0 | 12.7 | 0.18 | 8.4 | –0.2 | 11.0 | 0.25 |
| | NAM _{CC} | 9.2 | –0.0 | 12.8 | 0.18 | 8.8 | 0.1 | 11.6 | 0.21 |
| | NAM _{CC} [*] | 9.1 | –0.0 | 12.5 | 0.20 | 8.5 | 0.4 | 11.2 | 0.23 |
| LCC | Persistent | 8.0 | 0.3 | 14.0 | – | 7.2 | –0.3 | 12.0 | – |
| | RDPS _{CC} | 7.7 | 0.0 | 11.1 | 0.21 | 7.3 | 0.6 | 9.8 | 0.19 |
| | NAM _{GHI} | 8.1 | 0.0 | 11.6 | 0.17 | 7.3 | 0.2 | 9.5 | 0.21 |
| | NAM _{GHI} [*] | 8.0 | 0.0 | 11.2 | 0.20 | 7.2 | 0.3 | 9.3 | 0.23 |
| | NAM _{CC} | 8.1 | 0.0 | 11.4 | 0.18 | 7.7 | 0.4 | 9.8 | 0.18 |
| | NAM _{CC} [*] | 8.0 | 0.0 | 11.2 | 0.20 | 7.3 | 0.6 | 9.5 | 0.21 |

Table 4

Interannual variability of NWP-based PO forecasts for CCA and LCC. RDPS_{CC} is used as the NWP input for the forecast, with night values removed (night $\equiv \theta_z > 85^\circ$). The GHI-to-PO model is trained on one year of data and then tested on a separate year. RMSE is reported as relative to the PV plant capacity (1 MWp for CCA and LCC) while nRMSE is calculated using Equation (13) and the mean PO (\overline{PO}) of the considered period. Higher forecast skills (s) indicate better performance.

| Training | | | | | Testing | | | | |
|----------|----------------------|----------|-------|-----------|---------|----------------------|----------|-------|-----------|
| Year | \overline{PO} [kW] | RMSE [%] | s [–] | nRMSE [–] | Year | \overline{PO} [kW] | RMSE [%] | s [–] | nRMSE [–] |
| CCA: | | | | | 2012 | 466 | 14.0 | 0.17 | 0.30 |
| 2011 | 484 | 12.6 | 0.23 | 0.26 | 2013 | 476 | 13.7 | 0.15 | 0.29 |
| 2011 | 484 | 12.6 | 0.23 | 0.26 | 2014 | 481 | 11.7 | 0.23 | 0.24 |
| 2012 | 466 | 13.9 | 0.17 | 0.30 | 2013 | 476 | 13.6 | 0.16 | 0.29 |
| 2012 | 466 | 13.9 | 0.17 | 0.30 | 2014 | 481 | 11.7 | 0.22 | 0.24 |
| 2013 | 476 | 13.6 | 0.16 | 0.28 | 2014 | 481 | 11.7 | 0.23 | 0.24 |
| LCC: | | | | | 2012 | 374 | 12.4 | 0.13 | 0.33 |
| 2011 | 397 | 11.3 | 0.20 | 0.28 | 2013 | 392 | 11.7 | 0.14 | 0.30 |
| 2011 | 397 | 11.3 | 0.20 | 0.28 | 2014 | 410 | 10.2 | 0.20 | 0.25 |
| 2012 | 374 | 12.2 | 0.15 | 0.33 | 2013 | 392 | 11.8 | 0.13 | 0.30 |
| 2012 | 374 | 12.2 | 0.15 | 0.33 | 2014 | 410 | 10.5 | 0.17 | 0.26 |
| 2013 | 392 | 11.7 | 0.14 | 0.30 | 2014 | 410 | 10.3 | 0.20 | 0.25 |

higher occurrence of over-prediction (negative) errors than 2014.

In addition to the annual statistics and error distributions shown in Table 4 and Fig. 6, it is useful to evaluate the error distributions on shorter time scales. Fig. 7 examines the error distribution of both sites broken down by the hour of the day and by the month of the year for the 2014 testing set. Violin plots are used instead of box plots to enable visualization of not only the error magnitudes, but also their distribution densities [49]. From the hour of day perspective, forecasts for both sites over-predict at sunrise and sunset, and under-predict during the majority of the day. On the monthly time scales, the forecasts tend to under-predict during the rainy season (December–March), when the coastal areas are more impacted by mobile mid-latitude weather systems and their associated cloud patterns than by the diurnal pattern of marine stratus coverage.

For system operators, over-prediction of power can be more harmful than under-prediction due to the technical complexity required for up-ramping of load following and backup units in the power system, which is greater than for curtailing excess output from intermittent producers. Meanwhile, for power producers, curtailing solar PO is unfavorable as the power potential and profits of the plant are reduced. To better understand the occurrence of such errors, we analyze the severe over-prediction errors per year, both on hourly and daily timescales.

Table 5 summarizes the occurrence of over-prediction errors that exceed a range of thresholds, relative to the rated plant

capacity. We define over-prediction errors greater than 20% as severe. Although both sites are approximately 7 km from the Pacific Ocean and 13 km from each other, LCC has a lower frequency of over-prediction events than CCA on both hourly and daily timescales. And this trend is consistent for all three years of testing data (2012, 2013 and 2014).

The pairing of the two sites (“Combined”) has an impact on the occurrence of severe over-predictions. On the hourly timescale, the occurrence of severe over-prediction events for the “Combined” site is reduced for all thresholds by 0.2–1.9% compared to CCA. Similarly, daily over-prediction events are reduced by up to 2.5%. These results indicate that the pairing of PV plants can enable spatial smoothing of severe over-prediction events.

5. Conclusions

This study presented a methodology to generate day-ahead power output forecasts for two PV plants in the American Southwest. The forecasts are based on publicly available numerical weather prediction models from the National Oceanic and Atmospheric Administration, and the Canadian Meteorological Centre. Four years of ground measurements (2011–2014) from two 1 MWp, non-tracking PV plants in San Diego County, USA were used in this study. Forecasts of the two sites achieved annual RMSE of 11.7–14.0% and 10.2–12.4% relative to the rated plant capacities (1 MWp), as well as annual forecast skills of 15–23% and 13–20%.

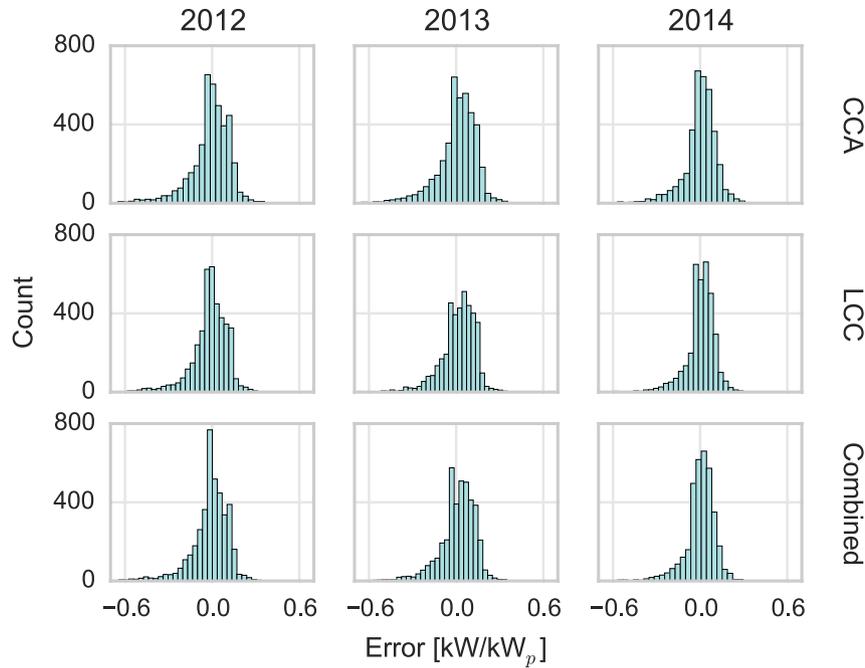


Fig. 6. Power forecast error distributions for CCA (top row), LCC (middle row) and treating the two sites as one combined power plant (bottom row). The columns are for three separate testing sets (2012, 2013, and 2014), with the PO forecast trained on the year prior to the testing set (e.g. 2012 was trained on 2011). As with Fig. 7, the error is calculated using Equation (12) relative to the plant capacity, where negative errors are over-predictions (i.e. the forecasted PO was greater than the measured PO) while positive values are under-predictions.

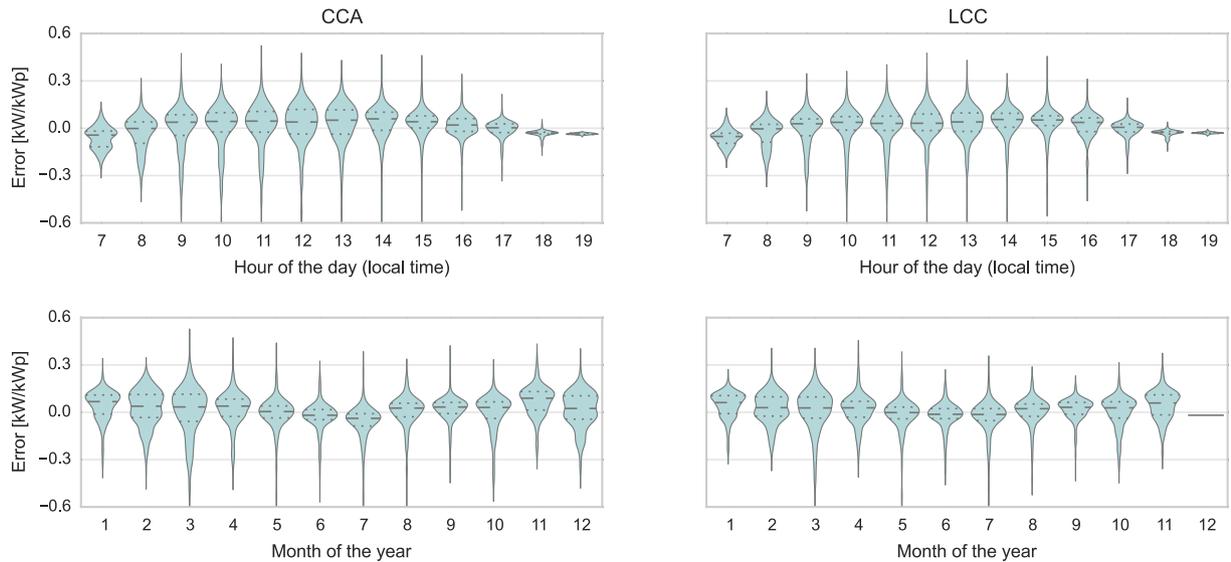


Fig. 7. Day-ahead PO forecast errors for CCA (left column) and LCC (right column), grouped by hour of the day (top row) and by month of the year (bottom row). The results here are for the PO model trained on 2013 and then tested on 2014. Error is as defined in Equation (12) relative to the PV plant capacity. Therefore, negative error values are over-predictions by the PO model and positive values are under-predictions. The dashed lines show the 25th, 50th, and 75th percentiles, while the width represents the density, which is approximated using kernel density estimation (KDE).

Our analysis leads to the following main conclusions:

- 1 Our proposed methodology can reduce the error in PV power output forecasts for day-ahead market participation in San Diego County, a key region for solar installations in the American Southwest.
- 2 Our methodology can be applied without any prior knowledge of the NWP model that produced the input GHI forecast.

- 3 The interannual forecast performance variability depends more heavily on the data set used in the performance evaluation rather than the data set used in training the forecast model.
- 4 Evaluation of severe over-prediction errors should be considered when evaluating day-ahead PO forecasts. Our methodology results in severe (>20% of the nameplate capacity) hourly over-prediction errors less than 9% of the time over three years for both sites. However, the occurrence of such severe over-predictions may be larger for other sites.

Table 5

Occurrence of over-prediction errors that exceed 20%, 30%, 40%, and 50% of the plant capacity. The year denotes the year used for testing, with the previous year used for training. “Hourly” are the number of hourly data points, “Daily” are the number of days where the average of the hourly forecasting errors within each day exceed the threshold, and the values in parentheses are occurrences relative to the total number of data points [%]. 2012 and 2013 had approximately 4000 hourly and 365 daily data points per site, while 2014 had approximately 3600 hourly and 320 daily data points per site. “Combined” represents the scenario in which CCA and LCC are treated as a single, spatially distributed 2 MWp power plant.

| Error | Hourly | | | Daily | | |
|-------|-----------|-----------|-----------|----------|----------|----------|
| | CCA | LCC | Combined | CCA | LCC | Combined |
| 2012: | | | | | | |
| >20% | 347 (8.6) | 272 (6.7) | 288 (7.1) | 22 (6.0) | 18 (4.9) | 21 (5.7) |
| >30% | 182 (4.5) | 146 (3.6) | 154 (3.8) | 11 (3.0) | 7 (1.9) | 8 (2.2) |
| >40% | 100 (2.5) | 73 (1.8) | 83 (2.1) | 1 (0.3) | 0 (0.0) | 0 (0.0) |
| >50% | 53 (1.3) | 22 (0.5) | 35 (0.9) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| 2013: | | | | | | |
| >20% | 302 (7.6) | 208 (5.2) | 228 (5.7) | 17 (4.7) | 7 (1.9) | 8 (2.2) |
| >30% | 143 (3.6) | 85 (2.1) | 100 (2.5) | 4 (1.1) | 0 (0.0) | 2 (0.5) |
| >40% | 66 (1.7) | 28 (0.7) | 39 (1.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| >50% | 24 (0.6) | 6 (0.2) | 14 (0.4) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| 2014: | | | | | | |
| >20% | 212 (5.9) | 162 (4.5) | 168 (4.7) | 10 (3.1) | 7 (2.2) | 8 (2.5) |
| >30% | 85 (2.4) | 61 (1.7) | 69 (1.9) | 2 (0.6) | 2 (0.6) | 2 (0.6) |
| >40% | 38 (1.1) | 27 (0.8) | 30 (0.8) | 1 (0.3) | 0 (0.0) | 0 (0.0) |
| >50% | 21 (0.6) | 13 (0.4) | 16 (0.4) | 0 (0.0) | 0 (0.0) | 0 (0.0) |

5 Spatial smoothing of over-prediction errors is possible through pairing of two spatially-distributed PV plants.

Additionally, in order to reduce the obstacles to safe and economic operation of solar power, research should shift focus towards power output forecasting. While improving irradiance prediction accuracy is important for many applications, our results suggest that for the day-ahead power forecasts, statistical methods can compensate for systematic errors in the irradiance predictions and are suitable for modeling the complexity of power plant operations.

Acknowledgments

The authors gratefully acknowledge the technical support provided by the administration and staff of Canyon Crest Academy, particularly Dr. E. Gerstin.

References

- [1] R. Inman, H.T.C. Pedro, C.F.M. Coimbra, Solar forecasting methods for renewable energy integration, *Prog. Energy Combust. Sci.* 6 (2013) 535–576.
- [2] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker Jr., D. Renné, T.E. Hoff, Validation of short and medium term operation solar radiation forecasts in the US, *Sol. Energy* 84 (2010) 2161–2172.
- [3] L. Nonnenmacher, C.F.M. Coimbra, Streamline-based method for intra-day solar forecasting through remote sensing, *Sol. Energy* 108 (2014) 447–459.
- [4] A. Zaouras, H.T.C. Pedro, C.F.M. Coimbra, On the role of lagged exogenous variables and spatial-temporal correlations in improving the accuracy of solar forecasting methods, *Renew. Energy* 78 (2015) 203–218.
- [5] F. Wang, Z. Mi, S. Su, H. Zhao, Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters, *Energies* 5 (2012) 1355–1370.
- [6] C.W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields, B. Washom, Intra-hour forecasting with a total sky imager at the UC San Diego solar energy test, *Sol. Energy* 85 (11) (2011) 2881–2893.
- [7] R. Marquez, H.T.C. Pedro, C.F.M. Coimbra, Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs, *Sol. Energy* 92 (2013) 176–188.
- [8] M.I. Gohari, B. Urquhart, H. Yang, B. Kurtz, D. Nguyen, C. Chow, M. Ghonima, J. Kleissl, Comparison of solar power output forecasting performance of the total sky imager at the University of California, San Diego sky imager, *Energy Procedia* 49 (2014) 2340–2350. Proceedings of the SolarPACES 2013 International Conference.
- [9] Y. Chu, H.T.C. Pedro, L. Nonnenmacher, R. Inman, Z. Liao, C.F.M. Coimbra, A smart image-based cloud detection system for intra-hour solar irradiance forecasts, *J. Atmos. Ocean. Technol.* 31 (9) (2014) 1995–2007.
- [10] Y. Chu, M. Li, H.T.C. Pedro, C.F.M. Coimbra, Real-time prediction intervals for intra-hour DNI forecasts, *Renew. Energy* 83 (2015a) 234–244.
- [11] Y. Chu, M. Li, H.T.C. Pedro, C.F.M. Coimbra, Real-time forecasting of solar irradiance ramps with smart image processing, *Sol. Energy* 114 (2015b) 91–104.
- [12] D. Bernecker, C. Riess, E. Angelopoulou, J. Hornegger, Continuous short-term irradiance forecasts using sky images, *Sol. Energy*. ISSN: 0038-092X 110 (2014) 303–315.
- [13] H.T.C. Pedro, C.F.M. Coimbra, Nearest-neighbor methodology for prediction of intra-hour global horizontal and direct normal irradiances, *Renew. Energy* 80 (2015) 770–782.
- [14] H.Y. Cheng, C.C. Yu, S.J. Lin, Bi-model short-term solar irradiance prediction using support vector regressors, *Energy* 70 (2014) 121–127.
- [15] M. Chaabene, M.B. Ammar, Neuro-fuzzy dynamic model with Kalman filter to forecast irradiance and temperature for solar energy systems, *Renew. Energy* 33 (2008) 1435–1443.
- [16] L. Martin, L.F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, M. Cony, Prediction of global solar irradiance based on time series analysis: application of solar thermal power plants energy production planning, *Sol. Energy* 84 (2010) 1772–1781.
- [17] B. Kraas, M. Schroedter-Homscheidt, R. Madlener, Economic merits of a state-of-the-art concentrating solar power forecasting system for participation in the Spanish electricity market, *Sol. Energy* 93 (2013) 244–255.
- [18] J. Luoma, P. Mathiesen, J. Kleissl, Forecast value considering energy pricing in California, *Appl. Energy* 125 (2014) 230–237.
- [19] L. Nonnenmacher, A. Kaur, C.F.M. Coimbra, Day-ahead resource forecasting for concentrated solar power integration, *Renew. Energy* 86 (2015) 866–876.
- [20] R. Hanna, J. Kleissl, A. Nottrott, M. Ferry, Energy dispatch schedule optimization for demand charge reduction using a photovoltaic-battery storage system with solar forecasting, *Sol. Energy* 103 (2014) 269–287.
- [21] M. Delfanti, D. Falabretti, M. Merlo, Energy storage for PV power plant dispatching, *Renew. Energy* 80 (2015) 61–72.
- [22] P. Mathiesen, J. Kleissl, Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States, *Sol. Energy* 85 (5) (2011) 967–977.
- [23] C. Voyant, M. Muselli, C. Paoli, M. Nivet, Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation, *Energy* 39 (1) (2012) 341–355. Sustainable Energy and Environmental Protection 2010.
- [24] V. Lara-Fanego, J. Ruiz-Arias, D. Pozo-Vázquez, F. Santos-Alamillos, J. Tovar-Pescador, Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain), *Sol. Energy* 86 (8) (2012) 2200–2217.
- [25] P. Mathiesen, C. Collier, J. Kleissl, A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting, *Sol. Energy* 92 (2013) 47–61.
- [26] R. Perez, E. Lorenz, S. Pelland, M. Beauharnois, G. Van Knowe, K.J. Hemker, D. Heinemann, J. Remund, S. Miller, W. Traummiller, G. Steinmauer, D. Pozo, J.A. Ruiz-Arias, V. Lara-Fanego, L. Ramirez-Santigosa, M. Gaston-Romero, L. Pomares, Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe, *Sol. Energy* 94 (2013) 305–326.
- [27] S. Pelland, G. Galanis, G. Kallos, Solar and photovoltaic forecasting through post-processing of the global environmental multiscale numerical weather prediction model, *Prog. Photovolt. Res. Appl.* 21 (3) (2013) 284–296.
- [28] H. Ohtake, K. Shimose, J. Fonseca, T. Takashima, T. Oozeki, Y. Yamada, Accuracy of the solar irradiance forecasts of the Japan Meteorological Agency mesoscale model for the Kanto region, Japan, *Sol. Energy* 98 (2013) 138–152. Part B.
- [29] E. Lorenz, J. Hurka, D. Heinemann, H. Beyer, Irradiance forecasting for the power prediction of grid-connected photovoltaic systems, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2 (2) (2009) 2–10.
- [30] P. Bacher, H. Madsen, H. Nielsen, Online short-term solar power forecasting, *Sol. Energy* 83 (2009) 1772–1783.
- [31] D. Masa-Bote, M. Castillo-Cagigal, E. Matallanas, E. Caamano-Martin, A. Gutierrez, F. Monasterio-Huelin, J. Jimenez-Leube, Improving photovoltaics grid integration through short time forecasting and self-consumption, *Appl. Energy* 125 (2014) 103–113.
- [32] E. Lorenz, T. Scheidsteger, J. Hurka, D. Heinemann, C. Kurz, Regional PV power prediction for improved grid integration, *Prog. Photovolt. Res. Appl.* 19 (7) (2011) 757–771, <http://dx.doi.org/10.1002/pip.1033>.
- [33] A. Mellit, A.M. Pavan, A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy, *Sol. Energy* 84 (2010) 807–821.
- [34] Y. Li, Y. Su, L. Shu, An ARMAX model for forecasting the power output of a grid connected photovoltaic system, *Renew. Energy* 66 (2014) 78–89.
- [35] J.G. Fonseca, T. Oozeki, H. Ohtake, K. Shimose, T. Takashima, K. Ogimoto, Regional forecasts and smoothing effect of photovoltaic power generation in Japan: an approach with principal component analysis, *Renew. Energy* 68 (2014a) 403–413.
- [36] J. G. Fonseca, T. Oozeki, H. Ohtake, T. Takashima, K. Ogimoto, Regional forecasts of photovoltaic power generation according to different data availability scenarios: a study of four methods, *Prog. Photovolt. Res. Appl.*
- [37] M. Zamo, O. Mestre, P. Arbogast, O. Pannekoucke, A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, Part I: Deterministic forecast of hourly production, *Sol. Energy* 105 (0) (2014a) 792–803.

- [38] M. Zamo, O. Mestre, P. Arbogast, O. Pannekoucke, A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, Part II: Probabilistic forecast of daily production, *Sol. Energy* 105 (2014b) 804–816.
- [39] M.P. Almeida, O. Perpignan, L. Narvarte, PV power forecasting using a nonparametric PV model, *Sol. Energy* 115 (2015) 354–368.
- [40] A. Lacis, J. Hansen, A parameterization for the absorption of solar radiation in the Earth's atmosphere, *J. Atmos. Sci.* 31 (1) (1974) 118–133.
- [41] R. Marquez, C.F.M. Coimbra, Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database, *Sol. Energy* 85 (2011) 746–756.
- [42] R. Perez, K. Moore, S. Wilcox, D. Renne, A. Zelenka, Forecasting solar radiation—preliminary evaluation of an approach based upon the national forecast database, *Sol. Energy* 81 (6) (2007) 809–812.
- [43] P. Ineichen, R. Perez, A new airmass independent formulation for the Linke turbidity coefficient, *Sol. Energy* 73 (3) (2002) 151–157.
- [44] P. Ineichen, Comparison of eight clear sky broadband models against 16 independent data banks, *Sol. Energy* 80 (4) (2002) 468–478.
- [45] C.A. Gueymard, Clear-sky irradiance predictions for solar resource mapping and large-scale applications: improved validation methodology and detailed performance analysis of 18 broadband radiative models, *Sol. Energy* 86 (8) (2012) 2145–2169.
- [46] Y.M. Saint-Drenan, S. Bolfinger, R. Fritz, S. Vogt, G.H. Good, J. Dobschinski, An empirical approach to parameterizing photovoltaic plants for power forecasting and simulation, *Sol. Energy* 120 (2015) 479–493.
- [47] A. Dolara, S. Leva, G. Manzolini, Comparison of different physical models for PV power output prediction, *Sol. Energy* 119 (2015) 83–99.
- [48] C.A. Gueymard, S.M. Wilcox, Assessment of spatial and temporal variability in the US solar resource from radiometric measurements and predictions from models using ground-based or satellite data, *Sol. Energy* 85 (2011) 1068–1084.
- [49] J.L. Hintze, R.D. Nelson, Violin plots: a box plot-density trace synergism, *Am. Stat.* 52 (2) (1998) 181–184.