1 Introduction

Solar forecasting is one of the enabling technologies for integration of intermittent renewable resources into the electric grid. Research on the topic is typically categorized by forecast horizon: intra-hour (<1 h ahead), intra-day (1–6 h ahead), and day-ahead (≥1–3 days ahead). Intra-hour forecasting generally relies on local telemetry and ground-based sky imagery (see, e.g., see Refs. [1–4]), while day-ahead forecasts are usually based on numerical weather prediction models and model output statistics (see, e.g., Refs. [5–10]). For the interval ranging from 1 to 6 h ahead, pure [11] or hybrid [12] remote sensing methods are competitive. The interplay of several forecasting techniques for solar energy integration is reviewed in Refs. [13–15].

1.1 Prior Work. A majority of previous solar forecasting research has focused on the solar resource rather than power output (PO). In particular, the existing literature is strongly biased toward forecasting of global horizontal irradiance (GHI) forecasts, the component of solar and atmospheric radiation that is the most relevant for nontracking photovoltaic (PV) power systems. There are fewer studies on predicting direct normal irradiance [4,16–18], and even fewer on the prediction of PO of operational solar power plants [6–8,19–22]. Relevant intra-day forecasting studies in the past have focused on satellite-to-irradiance models coupled with cloud motion vector (CMV) fields. Satellite-to-irradiance modeling is a mature, but still active area of research (see, e.g., Refs. [22–25]). Perez et al. [11] predicted hourly GHI for 1–6 h ahead at SURFRAD stations throughout the continental United States, while Coimbra and coworkers [12] developed a hybrid forecasting methodology based on CMV fields and artificial neural networks (ANNs) to predict GHI in Merced, CA, and Davis, CA, over horizons of 30–120 min ahead. Nonnenmacher et al. [16] forecasted GHI 1–3 h ahead in San Diego, CA, using streamlines determined from CMV fields. Zagouras et al. [26] investigated the use of lagged exogenous variables and spatiotemporal correlations to improve 1–3 h ahead GHI forecasts at seven CIMIS sites in California. Most recently, Mozorra Aguiar et al. [27] used ANNs to forecast GHI 1–6 h ahead for sites in the Canary Islands, off the northwest coast of Africa.

For a recent review of PO forecasting of solar power plants, readers are referred to Ref. [15]. In this work, we highlight the PO forecasting literature that focused on sites in California, the area of interest in this paper. Chu et al. [28] evaluated intra-hour forecasts of PO of a 48 MWp PV plant near Las Vegas, NV, while Lipperheide et al. [29] forecasted the PO of the same 48 MWp plant, but focused on horizons of less than 180 s ahead. Lonij et al. [30] used a network of residential PV systems in Tucson, AZ to forecast PO with horizons of 15–90 min. Pedro and Coimbra [19] evaluated 1–2 h ahead PO forecasts, generated without exogenous inputs, for a 1 MWp PV farm in Merced, CA. Larson et al. [7] recently forecasted day-ahead PO of two 1 MWp PV plants in San Diego County, CA.

1.2 Contributions. The goal of this work is to accurately and directly predict PO of PV power plants 1–6 h ahead. Rather than forecasting the solar resource and then developing a secondary resource-to-power (RTP) model, we focus on directly predicting the PO from the input variables. Removing the need for the RTP model minimizes data cross-dependencies and simplifies the implementation of the method by utilities and solar power plant operational managers. The proposed methodology only requires historical PO and satellite imagery for training, whereas other methods may require meteorological data streams (humidity, temperature, etc.) as well as diagnostic information related to power plant operations (e.g., PV panel temperature) for the RTP model, in addition to the inputs for the solar resource forecast. Such data streams may not be available due to (1) costs of instrument installation and maintenance, or (2) information security requirements at the power plant.

After training, the proposed methodology does not depend on ground telemetry or PO data to generate forecasts, i.e., the subsequent forecasts are not dependent on real-time data from the power plant. This means our forecast models are independent of data connection with the power plant, and therefore are not susceptible to failures due to sensor malfunctions or data transfer latency. While remote sensing is not fully dependable at all times, the reliability of remote sensing data transfer is typically much higher than local ground telemetry. Also, the two solar power plants used for testing and validation in this work are typical of hundreds of installations in California.

2 Forecast Methodology

We seek to develop a learnable mapping $f(\cdot)$ between a set of real-valued satellite-derived features $x^{(t)} \in \mathbb{R}^n$ at the current time $t$ and ground conditions $y^{(t)} \in \mathbb{R}$ at some future time $t + \Delta t$. More
specifically, we are interested in finding \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) for in-day forecasting of PO of PV solar farms (see Fig. 1). To accomplish these goals, we consider a linear model based on least squares (LS) and a nonlinear model based on support vector regression (SVR).

Our model selections are influenced primarily by two goals. First, both researchers and nonresearchers, e.g., power plant operators, should be able to implement the models. Least squares and support vector regression are standard regression methods, and are supported by most modern scientific programming environments, e.g., MATLAB, PYTHON, and R. This increases the likelihood that our forecast methodology will be adopted and that our results can be reproduced. Second, there should be prior literature supporting the use of the models for solar forecasting tasks. Indeed, both methods have been used successfully in published solar forecasting literature [7,31–33].

### 2.1 Least Squares

Least squares is a common linear regression method and can be solved efficiently. We use a regularized version of LS known as ridge regression, which solves the objective

\[
\minimize \sum_{i=1}^{m} \left( f(x_i; \theta, b) - y_i \right)^2 + \alpha ||\theta||_2
\]  

where \( \{\theta \in \mathbb{R}^n, b \in \mathbb{R}\} \) are the model parameters, and \( f(x_i; \theta, b) = \theta^T x_i + b \) is the forecast function. The \( \ell_2 \) regularization term \( (\alpha ||\theta||_2) \) is used to discourage overfitting, where \( \alpha \) is the regularization constant.

### 2.2 Support Vector Regression

Support vector regression is a form of support vector machines used for nonlinear regression tasks. In this study, we use \( \nu \)-SVR [34,35], as implemented in LIBSVM [36]. The objective function is

\[
\minimize \frac{1}{2} ||\theta||_2^2 + C \left( \nu e + \frac{1}{m} \sum_{i=1}^{m} (\xi_i + \xi_i^*) \right)
\]

subject to

\[
\begin{align*}
\nu & \leq \frac{1}{2} f(x_i; \theta, b) - y_i \\
\frac{1}{2} |\xi_i| & \leq \epsilon \\
\xi_i, \xi_i^* & \geq 0, i = 1, \ldots, m
\end{align*}
\]

where \( \{\theta \in \mathbb{R}^n, b \in \mathbb{R}, \xi \in \mathbb{R}^m, \xi^* \in \mathbb{R}^m, e \in \mathbb{R}\} \) are the model parameters, \( f(x_i; \theta, b) = \theta^T \phi(x_i) + b \) is the forecast function, and \( \phi : \mathbb{R}^n \rightarrow \mathbb{R}^m \) maps the input \( x_i \) to a higher dimensional feature space. The two remaining variables, \( \nu \in (0, 1) \) and \( C \in \mathbb{R} \), are hyperparameters. To improve computational performance, we use the radial basis function kernel: \( K(x, x') = \exp \left( \frac{-||x - x'||^2}{2\sigma^2} \right) \).

### 2.3 Smart Persistence

Following prior literature, we include a smart persistence model as a benchmark for evaluating our forecast methodology.

\[
\hat{y}(t + \Delta t) = \frac{y(t)}{y_{\text{clear}}(t + \Delta t)} y_{\text{clear}}(t + \Delta t)
\]

where \( y(t) \) is the target value at time \( t \), \( \Delta t \) is the forecast horizon, e.g., 1 h, \( \hat{y}(t + \Delta t) \) is the forecasted value, and \( y_{\text{clear}} \) is the clear sky value. When the target \( y \) is GHI, the fraction in Eq. (3) is commonly referred to as the clear sky index \( k_c \). We use the air mass-independent model described in Refs. [37–39] to estimate clear sky GHI. For clear sky PO, we used approximately 160 manually identified clear days, covering all four seasons, to fit a linear mapping between clear sky GHI and PO for each PV plant. As discussed in Ref. [7], the assumption of a linear relationship between GHI and PO is valid for the two sites.

### 3 Data

We analyze our forecast methodology using backward-averaged hourly AC PO from two PV plants in Southern California: canyon crest academy in San Diego and La Costa Canyon (LCC) in Carlsbad. Both plants have nontracking PV panels at a fixed 5 deg incline, with AC nameplate capacities of 1 MWp each. Additionally, both CCA (32,959 deg. –117,190 deg) and LCC (33,074 deg. –117,230 deg) are less than 10 km from the Pacific Ocean, with temperate climates that result in minimal temperature effects on PV panel efficiency. However, the PV panels at LCC are aligned 30 deg southwest due to the site’s characteristics, leading to a difference in peak output of the two sites under identical conditions. For further details on the two sites, readers are referred to Ref. [7].

#### 3.1 Satellite Images

The geostationary operational environmental satellite (GOES) system is comprised of geosynchronous satellites operated by the National Oceanic and Atmospheric Agency through the National Environmental Satellite, Data, and Information Service. For the purposes of this study, we will be using visible wavelength images from GOES-15, which is currently designated as GOES West. The visible image channel is centered on 0.63 \( \mu \)m, with a spectral range of 0.53–0.75 \( \mu \)m, a spatial resolution of 1 km, and a temporal resolution of one image per 30 min.

A \( w \times w \) square region, centered on the target site, is extracted from each image and then flattened into a vector \( x^{(i)} \in \mathbb{R}^w \), where \( n = w \times w \). Each image is then normalized

\[
x^{(i)} = \frac{x^{(i)} - \text{avg}(x^{(i)})}{\sqrt{\text{var}(x^{(i)})} + 10}
\]

where \( x^{(i)} \in [0, 255]^m \) is the unprocessed 8 bit grayscale image vector, \( \text{avg}(\cdot) \) is the arithmetic mean value, and \( \text{var}(\cdot) \) is the variance. In Eq. (4), the numerator normalizes the brightness and the denominator normalizes the contrast. After normalization, the image vectors are stacked to create a matrix \( X \in \mathbb{R}^{n \times m} \), where \( m \) is the number of images.

![A diagram of the overall forecast methodology. First, a 480 × 680 pixel satellite image at time \( t \) is cropped down to a \( w \times w \) region of interest, centered around the target site. The \( w \times w \) cropped image is then transformed into a \( n \)-length feature vector \( x(t) \) and fed into a forecast model, e.g., SVR, to produce a prediction of the target output \( \hat{y}(t + \Delta t) \).](image-url)
Four years (2012–2015) of hourly PO measurements and satellite-derive features are merged to create data sets for training and testing our intra-day PO forecast methodology (see Fig. 2). Although the images are available every 30 min, we only use images from the top of each hour to generate the forecasts. Additionally, night values are removed from the data set, i.e., only data points with $\theta_p < 85^\circ$ deg are included. No other exogenous inputs, e.g., GHI and ambient air temperature, were considered. In particular, our choice to ignore temperature as a forecast input is justified as both sites lie in coastal areas with temperate climates, where temperature fluctuations have less than a 0.5% effect on panel efficiency.

The data sets used in our analysis exceed the overall time span of any known intra-day PO forecast studies. Pedro and Coimbra’s [19] study is the most analogous one to our paper due to its focus on PO from a 1 MWp PV plant. However, their forecast testing set consisted of eight months, compared to the 24 months in our testing sets. And while they studied one 1 MWp PV plant, our work evaluates PO forecasts for two separate 1 MWp PV plants. Note that this comparison is not meant to lessen the scientific value of Ref. [19]. Rather, the comparison is meant only to illustrate the value provided by the data sets in our paper.

4 Results and Discussion

We evaluate our methodology by predicting PO at CCA and LCC. For each site, the forecasts are generated at the top of the hour using the most recent satellite image. In addition to the satellite-derived features, the LS and SVR models are supplied clear sky values of PO at the forecast target time $t + \Delta t$. The clear sky values are included to provide the models with both diurnal and seasonal information. Ground geometry is not included as input to the models in order to minimize data dependencies and improve the robustness of the models to ground-level operating conditions, e.g., delay or loss of data streams.

The forecasting models are trained on two years of data (2012–2013) and then tested on a separate set of two years (2014–2015) to ensure unbiased performance evaluations. Cross-validation is used to select the hyperparameters of the LS and SVR models for each horizon and site. Figure 2 illustrates the training and testing sets used for the four scenarios. The data sets shown are complete, i.e., they show the union of the output (PO) and input variables (satellite images). In Sec. 4.1, we discuss the error metrics used in evaluating the forecasts. Section 4.2 evaluates the effect of input image size choice on the forecast performance. In Sec. 4.3, we analyze the impact of the model choice, i.e., LS versus SVR. And in Sec. 4.4, we compare the performance of our forecast methodology to prior intra-day forecasting literature.

4.1 Error Metrics. We evaluate the models using standard error metrics: the root-mean-square error (RMSE) and the mean bias error (MBE) [2,13], as well as the forecast skill ($s$) proposed by Marquez and Coimbra [40]

$$s = 1 - \frac{\text{RMSE}}{\text{RMSE}_p} = 1 - \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (y(t) - \hat{y}(t))^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (y(t) - \bar{y}_p(t))^2}}$$

where $y$ is the measured value, $\hat{y}$ is the forecasted value, and $\bar{y}_p$ is the persistence forecasted value.

4.2 Effect of Image Size. The input image size may impact forecast performance. To investigate the effects, we compare the RMSE of LS and SVR intra-day forecast models trained on square image regions with $w \in \{4, 8, 16, 32, 64, 128\}$. As shown in Fig. 3, increasing the input image size decreases the RMSE for both the LS and SVR models in general for CCA. However, there are diminishing returns for both models as the number of features ($n = w \times w$) grows beyond $w = 64$. Doubling the image width from $w = 64$ to $w = 128$ increases the mean RMSE of the LS models by 5.6% as well as the minimum and maximum RMSE values by 2% and 5%, respectively. For the SVR models, increasing from $w = 64$ to $w = 128$ decreases the maximum RMSE by 8% while causing a negligible increase in mean RMSE (0.4%) and a non-negligible increase in the minimum RMSE, from 35 kW to 49 kW. A similar trend is observed for the LCC site. Hereafter, all forecast results are reported for models trained on $w = 64 \times 64$ image regions.

The optimal image size ($w = 64$) can be explained by two complementary mechanisms. First, both LS and SVR are well suited for data sets where the number of samples equals or exceeds the number of features, i.e., $m = n$ or $m > n$. Images with $w = 64$ produce $n \sim 4k$ features, which is the same order of magnitude as the number of samples in the training sets for each forecast horizon ($m \sim 3k - 6k$). In comparison, $w = 128$ images produce $n \sim 16k$ features, an order of magnitude larger than the number of samples, i.e., $m \ll n$. Second, larger image sizes cover larger spatial regions, and therefore can provide the models with information on cloud systems further away from the target site. In other words, while the first mechanism provides pressure to decrease $n$ relative to $m$, the second mechanism drives an increase in $n$. The balance between these two factors led to the optimal image width of $w = 64$.

4.3 Effect of Model Choice. Table 1 provides bulk statistics on the performance of the forecast models on the two years of testing data (2014–2015). For both sites, the SVR forecasts achieve the lowest RMSE and highest skill values for all horizons. For 1–2 h ahead, the difference in skill between the LS and SVR models is non-negligible ($\Delta s > 1\%$). However, beyond 3 h, the two models achieve similar skill values ($\Delta s < 1\%$). This result implies that while SVR is the best of the models evaluated, both LS and SVR are expressive enough to predict the intra-day PO behavior from satellite imagery. Additionally, for applications
where SVR is too complex to implement, users of our forecast methodology can instead employ LS and achieve comparable results.

For readers familiar with the forecast skill metric, the values in Table 1 may, at first glance, appear unusually large. Skill values reported in the PO forecasting literature rarely exceed 40%, regardless of forecast horizon [15]. However, the definition of forecast skill means the reference persistence model can have a large impact, i.e., a poor performing reference model can inflate the performance of a forecast method, as measured by the skill. In this study, both the persistence and smart persistence models performed poorly (RMSE > 250 kW) over the entire intra-day horizon range, thereby resulting in abnormally large, but nonetheless valid forecast skill values.

Besides bulk statistics, it is also important to consider error distributions when evaluating forecast models. As shown in Larson et al. [7], violinplots present a compact solution to visualizing multiple forecast error distributions [41]. Figure 4 uses violinplots to compare the error distributions of the LS and SVR models across the studied forecast horizons. Together with the bulk statistics, the violinplots of Fig. 4 reinforce the point that the LS and SVR models are both expressive enough to predict the intra-day PO from the satellite imagery.

Table 2 summarizes forecast performance across all horizons (1–6 h ahead), but breaks down the results by season: Spring (March–May), Summer (June–August), Fall (September–November), and Winter (December–February). Forecast performance is reported as relative RMSE (rRMSE), which is the RMSE relative to the mean PO (PO) of the specified season. For all seasons and both sites, SVR achieves the lowest rRMSE values of the tested models. While the highest SVR rRMSE values occur during the Winter for both sites (0.294 for CCA and 0.295 for LCC), the lowest values are in the Fall for CCA (0.219) versus Summer for LCC (0.241). Additionally, while both sites employ identical PV technology and are separated by less than 15 km, there is a non-negligible difference in forecast performance between the sites for Spring, Summer, and Fall. One factor in this difference is the abundance of solar microclimates throughout the region. For more info, readers are referred to Ref. [42].

Figure 5 visualizes the distributions of forecast absolute errors (AE) as functions of solar zenith and azimuthal angles. As we do not consider nighttime forecasts, there are fewer data points for the longer horizons. For example, the CCA testing set contains ~6800 h ahead forecasts versus ~3900 h ahead forecasts. To prevent biasing the results, the plots only include 1 h ahead forecasts. By grouping the errors by quartile, we see that the largest errors, i.e., the fourth quartile of errors, occur primarily in the morning. This result can be explained by two factors: (1) the high albedo in the visible channel satellite imagery near sunrise and (2) the prevalence of a marine layer in the region. Recent studies have proposed the use of other imaging wavelengths, e.g., infrared, to reduce early morning forecast errors [43]. However, overlapping image datasets were not available at the time of this study.

4.4 Comparison to Prior Work. The performance of our forecast methodology is comparable to prior intra-day PO studies. The genetic algorithm enhanced ANN from Pedro and Coimbra [19] achieved skill values of 32.2% and 35.1%, relative to a smart persistence model, for horizons of 1 and 2 h, respectively. In comparison, our methodology achieves skill values of 30.6% (CCA) and 29.1% (LCC) for 1 h ahead, and 46.1% (CCA) and 45.1% (LCC) for 2 h ahead. The divergence in skill values for 2 h ahead is possibly due to the genetic algorithm-ANN model’s lack of exogenous inputs, which was a self-imposed constraint by the authors. Unfortunately, a lack of overlapping satellite imagery and PO data sets for the PV plant in Ref. [19] prevents direct comparison.

Table 1: Forecast performance on the testing set (2014–2015) for CCA and LCC for horizons of 1–6 h ahead. The column names (1 h, . . . , 6 h) indicate the forecast horizon in hours. Negative MBE values correspond to models biased toward overpredicting PO and vice versa. Skill is reported relative to a smart persistence model. Bolded values indicate the forecast with the best value for a specific error metric, site, and forecast horizon.

<table>
<thead>
<tr>
<th>Site</th>
<th>RMSE (kW)</th>
<th>MBE (kW)</th>
<th>s (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 h</td>
<td>2 h</td>
<td>3 h</td>
</tr>
<tr>
<td>CCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smart Pers.</td>
<td>137.1</td>
<td>201.5</td>
<td>268.5</td>
</tr>
<tr>
<td>LS</td>
<td>114.4</td>
<td>126.1</td>
<td>134.6</td>
</tr>
<tr>
<td>SVR</td>
<td>104.8</td>
<td>121.8</td>
<td>132.0</td>
</tr>
<tr>
<td>LCC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smart Pers.</td>
<td>117.9</td>
<td>173.7</td>
<td>233.4</td>
</tr>
<tr>
<td>LS</td>
<td>96.6</td>
<td>106.5</td>
<td>114.5</td>
</tr>
</tbody>
</table>
Wolf et al. [33] forecasted PO for 15 min–5 h ahead for 921 PV systems of unspecified nameplate capacities. Although not reported directly, approximate forecast skills can be derived from Figs. 8 to 9 of Ref. [33]. The SVR-based models achieved skill scores in the range of 20–36% for horizons of 1–5 h ahead. Our methodology, which also uses SVR, achieves skill scores of 24–64% for the same forecast horizons. However, it is important to note that Wolf et al. used power data with a 15 min temporal resolution while we forecasted hourly averaged values.

Although our work focuses on forecasting PO, comparisons to intra-day solar resource forecasting literature can still be of value. GHI forecast results are particularly well suited for comparison as the PO of both CCA and LCC are linearly correlated with GHI due to their use of nontracking PV panels [7]. And fortunately, there have been multiple intra-day GHI forecasting studies at sites within the same region as both power plants. Nonnenmacher and Coimbra [16] achieved GHI forecast skill scores of ~8–16% for horizons of 1, 2, and 3 h at the University of California San Diego campus (32.88, –117.23). Zagouras et al. [26] obtained skills of ~13–23% (1–3 h ahead) for the CIMIS weather station at Torrey Pines (32.90, –117.25). Perez et al. [11] produced forecasts with skills of ~6–19% (1–6 h ahead) for the SURFRAD station at the Desert Rock Airport, NV, (36.62, –116.02). The skill scores (>23%) achieved by our methodology, compared to these three analogous studies, indicates that our results are competitive with other intra-day forecast methods.

Another important comparison is between the intra-day PO forecast results of this paper and the day-ahead PO results of Larson et al. [7]. Both papers use PO data from the same PV plants, i.e., CCA and LCC, and cover similar time ranges (2011–2014 for Ref. [7] versus 2012–2015 for this study). Larson et al. reported day-ahead PO forecasts with skills of ~15–23% for CCA and ~13–20% for LCC, relative to a persistence model. In comparison, the intra-day PO forecasts attain higher skill scores than the day-ahead scenarios (~24–68% for CCA and ~24–69% for LCC). Intra-day forecasts achieving higher skill scores than day-ahead forecasts are in agreement with results from the forecasting literature, e.g., Ref. [11].

### 5 Impact of Next-Generation Satellite Imagery

In November 2016, the first satellite in the next-generation GOES-R series was launched, and is currently undergoing testing. Although the transition from the current GOES-N Series will not begin until approximately Fall 2017, we can already consider the potential impacts of the next-generation imagery provided by GOES-R on the proposed forecasting methodology. The new ABI will provide three times the spectral bands, four times the spatial

---

**Table 2 Performance of forecasts across all studied horizons (1–6 h ahead) on the testing set (2014–2015). Results are broken down by site and season: Spring (3/1–5/31), Summer (6/1–8/31), Fall (9/1–11/30), and Winter (12/1–2/28). RMSE values are reported relative to the mean PO of the season (PO). Bold values indicate the forecast model with the lowest RMSE for a specific season.**

<table>
<thead>
<tr>
<th></th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Winter</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CCA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PO (kW)</td>
<td>572.9</td>
<td>585.0</td>
<td>502.9</td>
<td>425.5</td>
<td>527.2</td>
</tr>
<tr>
<td>Smart Pers.</td>
<td>0.516</td>
<td>0.533</td>
<td>0.532</td>
<td>0.586</td>
<td>0.539</td>
</tr>
<tr>
<td>LS</td>
<td>0.227</td>
<td>0.239</td>
<td>0.226</td>
<td>0.300</td>
<td>0.244</td>
</tr>
<tr>
<td>SVR</td>
<td>0.223</td>
<td>0.232</td>
<td>0.219</td>
<td>0.294</td>
<td>0.238</td>
</tr>
<tr>
<td><strong>LCC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PO (kW)</td>
<td>456.1</td>
<td>477.9</td>
<td>407.4</td>
<td>346.1</td>
<td>425.9</td>
</tr>
<tr>
<td>Smart Pers.</td>
<td>0.541</td>
<td>0.594</td>
<td>0.592</td>
<td>0.641</td>
<td>0.590</td>
</tr>
<tr>
<td>LS</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.296</td>
<td>0.259</td>
</tr>
<tr>
<td>SVR</td>
<td>0.243</td>
<td>0.241</td>
<td>0.244</td>
<td>0.295</td>
<td>0.252</td>
</tr>
</tbody>
</table>
resolution, and up to five times the temporal resolution of the current GOES-N satellites.\(^2\)

A direct comparison of forecasts with GOES-N and GOES-R imagery for the two California sites is not possible at this time. However, the Advanced Himawari Imager (AHI) on the Japanese Meteorological Agency’s Himawari-8 geostationary weather satellite has comparable specifications to the GOES-R ABI: 16 spectral bands, 0.5 km per pixel resolution visible images, and six full disk images per hour. While the AHI full disk images do not overlap with California, the images are available from July 2015–present.

5.1 Experiment Setup. Here, we aim to provide an analysis of the potential impact of the GOES-R ABI images on the proposed methodology using images from the Himawari-8 AHI and ground telemetry from sites in Australia. The goal is to understand whether the proposed methodology will still be applicable once the transition to the next-generation satellite imagery is complete. Additionally, the following analysis provides insight into whether the methodology can be adapted to multiple satellite imagery products with differing qualities, and therefore whether or not it can be successfully applied to sites outside of North America.

We compare the forecast performance between AHI visible images at full resolution (0.5 km per pixel; comparable to the GOES-R ABI), and images that we have down-sampled to 1.0 km per pixel to reduce model complexity and storage requirements. All other variables are fixed; we evaluate the same intra-day forecast models described in Sec. 2, extract features from hourly grayscale images, and train on historical hourly PO data from a nontracking PV power system.

Full resolution, full disk images from Himawari-8 were downloaded from Japan’s National Institute of Information and Communications Technology (NICT). NICT provides a publicly accessible web portal of historical Himawari-8 images.\(^3\) Although available as red/green/blue composites, we convert all images to grayscale to minimize the influence of the additional spectral channels, which are not the focus of this analysis, and then perform the preprocessing detailed in Sec. 3.1.

For the purposes of this analysis, we use publicly available data from the Australian PV Institute Solar Map, funded by the Australian Renewable Energy Agency.\(^4\) The Australian PV Institute provides hourly power generation data from 53 PV installations across Australia, with nameplate capacities in the range of ~2–35 kWp. We select the 20 kWp PV install in Blacktown, NSW (~33.763 deg., 150.907 deg.), as it has one of the largest nameplate capacities among the AVPI sites with data that overlaps with the available Himawari-8 images, and has minimal quality control issues.

One year of overlapping Himawari-8 images and PO data from Blacktown is available for the analysis (January 2016–2017). The year of data is partition into a training set consisting of the odd months (January, March, etc.), and a testing set of the even months (February, April, etc.). As noted in Larson et al. [7], this partitioning ensures that the models are trained, and also tested, on data from all four seasons. Additionally, as with the two California sites, night values were removed from both the training and testing sets. Based on the training set results, a spatial image size of 128 km × 128 km was selected for full analysis.

5.2 Comparison Results. Table 3 summarizes the intra-day forecasting performance for the 20 kWp Blacktown site using the Himawari-8 images. Both the LS and SVR models achieve forecast skill values that are similar to the CCA and LCC sites (~30–70%). Additionally, as with the two California sites, the SVR models have overall lower RMSE and higher skill values than the LS models. However, for certain applications, e.g.,

\(^2\)For a complete comparison between the ABI and the imager on the current GOES-N Series, please see: http://www.goes-r.gov/spacesegment/abi.html

\(^3\)http://himawari8.nict.go.jp/

\(^4\)http://pv-map.apvi.org.au
generating forecasts on low-power devices, the difference in performance may not be statistically significant enough to justify the added complexity and lower interpretability of the SVR models, as compared to LS.

Based on this analysis, down-sampling the image inputs does not have a statistically significant negative impact on the forecast performance. The absolute difference in RMSE between the low-resolution and high-resolution images is less than or equal to 0.4% for both the LS and SVR models. The lack of forecast performance degradation, coupled with the benefit of lower computation and storage requirements of the low-resolution image features, suggests that down-sampling the image inputs is a viable strategy.

Before moving on, we note a few caveats to these results. First, to enable a more direct comparison to the other intra-day results in this study, we did not consider the impact of the increased temporal resolution of the next-generation satellite images. Higher spatial and temporal resolution images may enable detection of small-scale and short-lived cloud systems, which current satellite images can miss, and therefore could outweigh the lower computation and storage benefits of down-sampling the images.

Second, our analysis only considers a year of data (six months of training and six months of testing) for one sub-1 MWp PV site. While we took steps to ensure the validity of our analysis (see Sec. 5.1), a longer data set would add increased confidence to the results. Additionally, we presented an analysis of the potential impact of down-sampling the image inputs to the intra-day forecast results for CCA and LCC. Data for a larger scale PV power plant in the area covered by Himawari-8 were not available at the time of this study.

6 Conclusions

We presented a methodology for forecasting PO of PV power plants for horizons of 1–6 h ahead. Our methodology enables direct prediction of PO of operational solar farms from satellite imagery. This has two major benefits: first, by directly forecasting PO, we remove the need for intermediate irradiance forecasts and resource-to-power modeling, which are typically site specific; second, by removing the need for additional data dependencies, e.g., meteorological telemetry, we ensure that the methodology is generalizable to power plants with varying levels of pre-existing instrumentation. Together, these two features make the proposed forecasting methodology well poised to be applied to a range of operational PV power plants.

We evaluated the proposed direct remote sensing forecast methodology using two nontracking, 1 MWp PV plants in California. Four years (2012–2015) of combined PO and satellite imagery data were used to analyze the forecast methodology’s performance. Over a testing set of two years (2014–2015), our methodology achieved RMSE values of ~90–136 kW over horizons of 1–6 h ahead and forecast skill scores of ~24–69%, relative to the baseline smart persistence model.

Additionally, we presented an analysis of the potential impact of next-generation satellite imagery from GOES-R on the intra-day forecasts. Using PO data from a PV power plant in Australia and images from the Himawari-8 geostationary satellite as a proxy for GOES-R, we evaluated the effects of the higher resolution imagery. The results showed that the methodology can easily be deployed for the imaging products of other satellites. The higher spatial resolution imagery does not improve or degrade the forecast performance. Down-sampling the images to reduce computation and storage costs was also shown to be compatible with the methodology. These results help quantify the impact that a new generation of satellite imagery systems, e.g., GOES-R, will have on remote sensing-to-power solar forecasting performance.

Acknowledgment

The authors gratefully acknowledge the technical support provided by the administration and staff of Canyon Crest Academy, particularly Dr. E. Gerstin. Additionally, the authors would like to thank Dr. K. T. Murata and the NICT Science Cloud project for technical support related to accessing historical Himawari-8 images.

References