Adaptive Image Features for Intra-Hour Solar Forecasts

Hugo T. C. Pedro,¹ Carlos F. M. Coimbra,¹ a) and Philippe Lauret²

¹)Department of Mechanical and Aerospace Engineering, Center for Energy Research
University of California San Diego, La Jolla, CA 92093, USA
²)Université de La Réunion, 15 Avenue René Cassin
97715 Saint-Denis Cedex 9, France

(Dated: 13 March 2019)

We introduce a simple and novel technique to extract dynamic features from sky images in order to increase the accuracy of intra-hour forecasts for both Global Horizontal Irradiance (GHI) and Direct Normal Irradiance (DNI) values. The proposed methodology is based on a block-matching algorithm that correctly identifies the bulk motion of clouds relative to the position of the Sun in the sky. Adaptive rectangular- and wedge-shaped Regions Of Interest (ROIs) are used to select the image pixels for the new features. Results show an average increase of 6.8% (6.7%) in forecast skill for GHI (DNI) across all horizons tested as measured against a model with global (non-adaptive) image features. Relative to clear-sky persistence the new model achieves skills ranging from 20% to 30% (22% to 35%) for GHI (DNI), among the highest ever reported for these time horizons. An analysis based on Mutual Information and Pearson correlation coefficients between the image features and the training data reveals overall improvements in all metrics. The proposed adaptive method also improves the predictability of ramp magnitude and direction.

I. INTRODUCTION

Increased market penetration of weather-dependent power resources results in unwanted fluctuations in the electric power grid. Successful grid integration of variable renewable resources requires coordinating efforts between solar producers, utility companies and independent system operators to minimize impacts on grid reliability and voltage or frequency oscillations. The success of these measures relies, to a large extent, on the use of forecasting models for solar resource variability.¹ Renewable forecasting is also an enabling technology for reducing energy storage needs, improve the lifetime of electrochemical storage systems, and for minimizing the usage of high levels of carbon-intensive spinning resources.

Various predictive models for solar irradiance and solar power generation have been developed in recent years with evolving improvements in forecast accuracy.²⁻⁵ In the work of Pedro and Coimbra⁶, the classical k-Nearest-Neighbor (kNN) model is optimized to forecast DNI and GHI intra-hourly. The optimization algorithm selects irradiance time series and sky image features that minimize the kNN forecast root mean square error (RMSE). In that work, and also here, image features are statistical metrics (average, standard deviation and entropy) of images collected with a sky camera collocated with the irradiance sensor. In a related paper (Pedro et al.⁷) the input and target data created in the previous work were used to train a random tree gradient boosting model (XGBoost), demonstrating that it is possible to improve the forecasting metrics for GHI and DNI through the use of properly trained XGBoost algorithms.

In the present work, we revisit this problem by introducing a new set of dynamic features derived from sky images. The motivation for this effort stems from the fact that in previous works the image features were derived from the whole set of pixels available after removing obstacles and saturated pixels. Here we exploit the question of whether more accurate forecasts could be obtained by taking into account image features sensitive to cloud motion dynamics in the vicinity of the region affected by short-term horizons. This question is addressed in this work by computing image features for Regions Of Interest (ROIs) that change with the computed average cloud direction. Cloud motion is

a)Electronic mail: ccoimbra@ucsd.edu.
evaluated using a simple Block-Matching Algorithm (BMA) on a sequence of images that precede the forecasting issuing time.

This work provides a simple but robust intra-hour forecasting model for GHI and DNI that takes into account region-specific cloud motion. The model bypasses some of the common tasks used in sky-imaging forecasting algorithms such as, the computation of cloud height and speed and resulting shadowing, pixel classification (cloud/clear-sky) and detailed cloud advection and estimation of cloud optical depth. By doing so, the model ignores some important physical properties but also avoids error accumulation from intermediate steps, which can be substantial as documented in Chow et al. for cloud identification, and in Wang et al. for cloud height and speed.

II. DATA AND INSTRUMENTATION

A. Data

The data sets used in this paper include time series for GHI and DNI obtained with a rotating shadowband radiometer (RSR2 instrument by Irradiance Inc.), plus sky images from a fisheye camera (model FE8171V by Vivotek In.), which is collocated with the RSR2. The RSR2 provides 1-minute averages for GHI and DNI every 1-minute. It should be noted that the RSR2 does not measure DNI directly, instead it measures GHI and diffuse horizontal irradiance (DHI) when the sensor is shaded. DNI is then computed using the fundamental relationship $GHI = DHI + DNI \cos(\theta_z)$, where $\theta_z$ is the zenith angle. Benchmarking of DNI RSR2 measurements shows acceptable uncertainty levels for applications in solar generation power production. The camera used to capture the sky images has a 3.1 MP CMOS sensor and is equipped with a 360° panoramic view lens, and provides 8-bit RGB images in JPEG format every minute. Data and images used in this work were collected from 2012 to 2013 in Folsom, CA. More details about this data set and the acquisition system can be found in Pedro and Coimbra.

B. Feature engineering

Following the work in Pedro and Coimbra several features are extracted from the irradiance data and the sky images. Features extracted from GHI and DNI time series include the backward average $B(t,w) = \langle I(t) \rangle_{t \in [t-w,t]}$ and the variability $V(t,w) = \sqrt{\langle \Delta I^2 \rangle_{t \in [t-w,t]}}$ for a window $w$ that varies from 5 minutes to 2 hours in 5-minute steps, where $\langle \cdot \rangle$ indicates the average. The third feature is simply the lagged irradiance values in 5-minute bins for the 2 hours that precede the forecasting issuing time.

The image features include simple statistical metrics for the image color data: average ($\mu$), standard deviation ($\sigma$) and entropy ($e$). These features are applied to the R, G, B color channels separately and also to the normalized red-to-blue ratio data (R/B). Image features in the previous study used all pixels available after removing obstacles and saturated pixels in the circumsolar region. Here we aim to verify whether or not the models can be improved by including the effect of cloud motion into the feature extraction algorithm. That is studied here by computing the image features listed above on a subset of pixels within the ROI that is a function of the average cloud motion. Popular algorithms to identify cloud motion from sky images usually fall under three distinct categories: block-matching, phase correlation and optical-flow. In this work, we use the Block-Matching Algorithm (BMA) due to its simplicity and robustness.

1. Block-Matching Algorithm

BMAs are used frequently for object tracking in a sequence of images or frames and are well suited when the inter-frame changes are small. In this work BMA determines the displacement
FIG. 1. Two consecutive images are annotated to obtain the manual cloud motion data used to validate the BMA. The same cloud structures are identified in both images (in this case 7 key points are highlighted). The sky image on the right also shows the location of key points (blue circles) from the previous image.

The displacement that minimizes Eq. 1 is determined using a grid search method and the cloud direction is calculated as: \[ \theta = \arctan\left(\frac{\delta_y}{\delta_x}\right) \]. Given that we are interested in obtaining a single value that characterizes the bulk cloud motion, we use the whole image in the BMA algorithm. This differs from applications that seek to determine the cloud motion vector field. In these cases the image is divided into small regions and the BMA is applied to each one separately.

The BMA is run for all images in the data set and validated against a set of cloud direction values obtained manually. These values were obtained by annotating cloud features that are easily identifiable in consecutive images as exemplified in Fig. 1. The validation included a large set of images selected from 15 \( \approx \) 1.5 hour periods that show distinct cloud directions. The left panel in Fig. 2 compares the BMA direction for five of these periods, obtained every 5-minutes, against the values obtained manually. The latter are then binned in 5-minute averages (shown as solid gray lines) that can be compared directly to the BMA output. That comparison is shown in the right panel of Fig. 2 that includes all the data (231 data points). The panel also indicates the bias and absolute error between these variables together with the correlation coefficient. Although the time series plot and the scatter plot shows instances of large error, on aggregated the errors are low and the correlation high. A possible approach to avoid large errors assumes that cloud direction changes slowly in sub-hour time frames. Thus, one cloud implement a simple filter using \( \theta \) from the last 5 to 30 minutes and replace outliers with the average value. Given the overall good accuracy of the BMA, the sensitivity of feature extraction and forecast accuracy relative to cloud direction perturbations is not explored in this work.

Finally, note that, with the displacement \((\delta_x, \delta_y)\) and the time elapse between frames it is possible to determine the apparent average velocity for the cloud field. This variable was not explored in this work given that, as shown below, the proposed adaptive features depend only on the cloud direction \( \theta \).

2. Adaptive ROIs for feature extraction

We augment the set of image features by considering only pixels in ROIs upstream of the sun’s position in the direction identified by the BMA. In this work we consider two shapes for the ROIs as
FIG. 2. BMA validation. Left: The time series plot compares the cloud direction obtained using BMA (solid lines with open markers) against the values determined from manual annotation in the sky images (solid dots) for five ≈1.5-hour periods with distinct cloud direction. The solid gray lines show the average of the values obtained manually in 5-minute bins that are directly comparable with the BMA results. Right: The scatter plot shows the BMA direction against the averaged direction from the images annotated manually for a larger dataset with 15 ≈1.5-hour periods. The colored dots indicate the correspondence with the time series in the left panel. Black dots indicate data not shown in the left panel. The annotation in the top left corner indicates the bias error, absolute error and correlation coefficient between the two directions using all data.

FIG. 3. ROIs for feature extraction. Panels A and C indicate how the rectangular and wedge ROIs are created. The ROIs are a function of the sun’s position (orange marker) and the cloud direction (blue arrow) identified by the BMA. Panels B and D show three ROIs for each case. The ROIs in B and D are obtained with \( w = \{50, 100, 200\} \) pixels and \( \alpha = \{45^\circ, 90^\circ, 180^\circ\} \), respectively. Features are computed with the ROIs pixels. The remaining pixels are ignored.

The rectangle width and wedge angle were not subject to optimization. The values were chosen as a compromise between covering the range of possible values for \( w \) and \( \alpha \) and the total number of features generated since that, for each ROI, there are 12 new features (4 color maps and 3 statistical metrics). Thus, this feature engineering technique augments the set of predictors with \( 12 \times 6 = 72 \) new features.
FIG. 4. Illustration of the forecasting timeline with data and images acquired on 2013-4-14 (PST). In this example forecasts for DNI and GHI for horizons $\delta = 5$ and 30 minutes are issued at time $t = 15:00$. The forecasting models use features derived from the irradiance time series and images collected up to $t$. Note that, as shown in the figure, the forecast models produce predictions for the average value of irradiance in the forecasting window $[t,t+\delta]$.

C. Forecasting algorithm

In order to provide a complete comparison, the methodology presented in two previously published papers $^6,^7$ is repeated here, i.e., forecast models are trained separately for DNI and GHI for horizons ($\delta$) from 5 to 30 minutes. Forecasts are issued every 5 minutes during daytime (solar elevation $\geq 5^\circ$ or $\theta_z \leq 85^\circ$). Figure 4 illustrates the timing for the forecasts, the data available for extracting forecasting inputs and the forecasting targets (for 5 and 30 minutes).

The forecasting algorithm used here aims to determine the relationship between the inputs and the outputs for a training dataset (22,203 samples). The results presented below are then computed using an independent testing set (15,807 samples). The splitting of the data set and the forecasting model are explained in detail in Pedro and Coimbra $^6$ and Pedro et al. $^7$, respectively. For the sake of completeness and introduce some nomenclature the forecasting framework is summarized here.

Mathematically, we seek to obtain models of the form

$$\hat{I}(t+\delta) = f(x_1(t), x_2(t), \ldots, x_N(t)) \times I_{cs}(t+\delta)$$

where $I$ denotes irradiance (GHI or DNI), $\hat{I}(t+\delta)$ denotes the predicted average irradiance value for the time window $[t,t+\delta]$, $I_{cs}$ is a clear-sky model $^{17}$ for the irradiance $I$ and $f(\cdot)$ is the mapping function between the features $x_1(t), x_2(t), \ldots, x_N(t)$ and the corresponding clear-sky index $k_{cf}(t+\delta) = I(t+\delta)/I_{cs}(t+\delta)$. As illustrated in Fig. 4 the features are extracted from irradiance data and sky images for the time window that precedes the forecasting issuing time $t$.

The mapping $f(\cdot)$ is determined using the XGBoost algorithm. $^{18}$ XGBoost is a popular and efficient implementation of the gradient boosting technique that uses a greedy algorithm for splitting the tree and regularized learning to prevent overfitting. Implementation details for this particular case are available in Pedro et al. $^7$.

III. RESULTS

In the training stage we developed 12 models: two irradiance components (GHI and DNI) and six forecast horizons. After obtaining the trained models with XGBoost we computed GHI and DNI predictions for the testing dataset. In this section we assess whether or not the adaptive image features improve the forecast accuracy based on those results. We analyze the forecasting error for the proposed model, which is denoted in this section as XGB+CM where “CM” denotes cloud motion. The model’s performance is compared to that of the model presented in Pedro et al. $^7$ which is denoted as XGB and the kNN model in Pedro and Coimbra $^6$. 
FIG. 5. GHI RMSE and RMSE forecast skill as a function of the forecast horizon. The figure shows the results for the new model “XGB+CM” and compares it against results from Pedro et al.\(^7\) (XGB) and Pedro and Coimbra\(^6\) (kNN) for 15807 samples in the testing set. The annotations next to the markers quantify the skill improvement relative to the model below.

### A. Bulk Error Metrics

Here we assess the forecasting error for a given horizon \((\epsilon(t, \delta) = I(t + \delta) - \hat{I}(t + \delta))\) in terms of bulk error metrics: Mean Absolute Error, \(\text{MAE} = \sum_{i=1}^{M} |\epsilon_i| / M\); Root Mean Square Error \(\text{RMSE} = (\sum_{i=1}^{M} \epsilon_i^2 / M)^{1/2}\); and forecast skill \(s = 1 - E / E_p\), where \(M\) is the number of test samples and \(E\) denotes one of the previous metrics. The subscript \(p\) denotes the clear-sky index persistence model, which is explained in detail in Pedro and Coimbra.\(^6\)

Figure 5 plots the forecast skill and RMSE for the GHI forecasts as a function of the horizon. The figure allows to compare results for the model introduced here (in dark blue) against results from the previous two papers. For each forecast horizon the improvements from kNN to XGB, and from XGB to XGB+CM are annotated in the figure. Figure 6 shows the same for the DNI forecasts.

The results presented in these figures indicate that, indeed, the adaptive image features explored in this work improve the forecast accuracy. The figures show that the XGB+CM model adds between 5.6% and 8.0% to the forecast skill relative to the XGB model. As reported in Pedro et al.\(^7\) XGB improves skill between 1.8% and 4.9% relative to the kNN model. The pronounced improvement between XGB+CM and XGB (larger than that obtained between XGB and kNN) can be attributed to the new adaptive image features since everything else is the same in these two models. In the following section we try to establish the reason behind the higher predictive power of the new features.

### B. Relevance of adaptive features

Here we use Pearson correlation (PC) coefficient and Mutual Information (MI) to assess the relevance of the image features to the prediction of irradiance. These two metrics are often used in...
feature selection for machine learning models. In this work feature selection was not applied for two reasons. Firstly, to maintain conditions as similar as possible to the two previous works. Secondly, XGBoost finds the most important features during its iterative process. Instead, we use these tools to understand the improvements seen in the previous section.

In this case, we followed a comprehensive approach in which the PC and MI are computed between each image feature and the target data for all forecast horizons. Results for the 5-minute horizon are compiled in Figs. 7 and 8 for the PC and MI values, respectively. In each figure the top panel corresponds to GHI data and the bottom panel to DNI data. For every panel, rows correspond to a different operator, and columns correspond to pairs of (ROI, color data). The pairs are identified by the labels in the x-axis (color data) and the labels in the top-left corner of the figure (ROIs). For each feature operator (each row) the largest absolute value is identified by the cross marker and the respective annotation.

These figures allow to draw some conclusions:

- In general, the red-to-blue ratio data shows the largest values of PC and MI regardless of the statistical operator used to compute the feature.
- The statistical operator that results in the largest values of PC and MI is entropy.
- In the case of MI, the largest values are obtained when using the ROI \( r_1 \). In the case of PC that role is played by \( r_3 \).

The analysis for the other forecast horizons yields similar results. In order to summarize these results, the largest values for MI and PI for each horizon and the corresponding feature are plotted in Fig. 9 and Fig. 10, respectively. These figures show results for two cases: i) considering all features (solid lines) and ii) considering only the non-adaptive features (dashed lines) used in the previous works. The annotation that accompanies each point in the figure \((\text{Operator(Color data,Mask)})\) indicates how the corresponding feature was obtained.

The two figures show that the largest MI and PC increase considerably when the adaptive features are considered. The most notable case occurs for the DNI’s PC (Fig. 10(right)) in which the maximum PC values increase from \( \approx 0.45 \) (without CM) to \( \approx 0.67 \) (with CM).
FIG. 7. Correlation coefficient between the image features and the target data for the 5-minute forecast. In the top panel the target data is GHI and in the bottom panel it is DNI. Labels in the x-axis indicate the image data used to compute the features listed in the y-axis. The annotation in the top-left corner of the figure indicates the ROIs for feature engineering: * = whole image, w = wedge ROIs, and r = rectangular ROIs.

FIG. 8. Same as Fig. 7 but for the Mutual Information metric.

Focusing only on the results when the adaptive features are considered we see that in the case of MI (Fig. 9) the largest value for all cases is obtained with entropy computed on the red-to-blue ratio color data with the $r_1$, the smallest rectangular ROI. In the case of PC, the largest values are again obtained with entropy applied to the red-to-blue ratio color data, but the selected ROI varies. In the GHI case, the largest PC values are obtained with $r_3$ and in the DNI case they are obtained with the $w_1$ ROI. The preference of larger ROIs for GHI and smaller ROIs for DNI is not surprising since the first is a global measure of the solar irradiance and the second is a directional measure.

Note that, the increase of PC and MI values with forecast horizon may seem counterintuitive at first sight. It is easily explained: it results from the fact the forecast horizon is coupled with the averaging of the measured data. For instance, the 5-minute forecast is compared against measured data averaged in 5-minute bins, whereas the 30-minute forecast is compared against 30-minute data. The longer averaging window for longer horizons results in smoother data and larger PC and MI.

185 MI (Fig. 9) the largest value for all cases is obtained with entropy computed on the red-to-blue ratio color data with the $r_1$, the smallest rectangular ROI. In the case of PC, the largest values are again obtained with entropy applied to the red-to-blue ratio color data, but the selected ROI varies. In the GHI case, the largest PC values are obtained with $r_3$ and in the DNI case they are obtained with the $w_1$ ROI. The preference of larger ROIs for GHI and smaller ROIs for DNI is not surprising since the first is a global measure of the solar irradiance and the second is a directional measure.

Note that, the increase of PC and MI values with forecast horizon may seem counterintuitive at first sight. It is easily explained: it results from the fact the forecast horizon is coupled with the averaging of the measured data. For instance, the 5-minute forecast is compared against measured data averaged in 5-minute bins, whereas the 30-minute forecast is compared against 30-minute data. The longer averaging window for longer horizons results in smoother data and larger PC and MI.

185 MI (Fig. 9) the largest value for all cases is obtained with entropy computed on the red-to-blue ratio color data with the $r_1$, the smallest rectangular ROI. In the case of PC, the largest values are again obtained with entropy applied to the red-to-blue ratio color data, but the selected ROI varies. In the GHI case, the largest PC values are obtained with $r_3$ and in the DNI case they are obtained with the $w_1$ ROI. The preference of larger ROIs for GHI and smaller ROIs for DNI is not surprising since the first is a global measure of the solar irradiance and the second is a directional measure.

Note that, the increase of PC and MI values with forecast horizon may seem counterintuitive at first sight. It is easily explained: it results from the fact the forecast horizon is coupled with the averaging of the measured data. For instance, the 5-minute forecast is compared against measured data averaged in 5-minute bins, whereas the 30-minute forecast is compared against 30-minute data. The longer averaging window for longer horizons results in smoother data and larger PC and MI.
FIG. 9. For each horizon the figure shows the largest MI value and the corresponding feature indicated by the (Operator(Color data,Mask)) annotation. In each panel two cases are shown: the largest MI when considering the new adaptive features (solid line) and the largest MI when considering non-adaptive features (dashed line). Results for GHI and DNI are shown in the left and right panels, respectively.

FIG. 10. Same as Fig. 9 but for the absolute value of the Pearson correlation coefficient.

This analysis demonstrates that the features introduced in this work are better predictors of the irradiance, according to the MI and PC relevance metrics, than the non-adaptive metrics considered in previous works. However, as mentioned before, the exact effect that the adaptive ROI pixel selection has on the forecast accuracy cannot be assessed due to the non-linear nature of XGBoost. For instance, removing lower importance features from the set of predictors based on this analysis, may actually reduce forecast accuracy. Nevertheless, the conclusion is clear: The use of directional ROIs yields image features with larger PC and MI values which are explored by the XGBoost algorithm to create more accurate irradiance predictions.

C. Ramps Events

In the previous two sections, we explored the forecast accuracy for the whole testing set and the importance of the new image features. These analyses already allow to answer positively to the question posed at the top of this paper. Nevertheless, it is useful and important to have a more granular picture of the forecasting error. This can be accomplished by binning the testing data in terms of ramps events and analyzing the error for each bin. Ramps are defined as step changes in the measured irradiance time series \( R(t) = (I(t + \delta) - I(t)) / \delta \). This definition takes into account both ramps caused by the deterministic diurnal irradiance cycle and ramps due to weather.
The frequency of ramps of a certain magnitude can be assessed from the histograms of $|R|$ shown in Fig. 11. The figure shows that the magnitude of most measured ramps (between 71 and 98% for DNI and 67 and 99% for GHI) are below the 0.05Wm$^{-2}$/s threshold, which corresponds to a small step change in irradiance of 15Wm$^{-2}$ in 5 minutes or 90Wm$^{-2}$ in 30 minutes. The histograms also show that larger ramps are much more frequent for DNI than GHI.

After assigning a ramp value to every forecasting issuing time it is easy to study the models’ performance as a function of the observed ramp magnitude. That much is showed in Fig. 12 that plots the MAE forecast skill for the XGB and XGB+CM models for all $|R|$ bins and forecast horizons. We use MAE in this analysis instead of RMSE since the sample sizes vary widely from case to case, and MAE represents a better comparison between models sampled at different rates.24

The figure shows that XGB+CM outperforms XGB in all cases and that there are appreciable increases in the forecast skill for $|R| \geq 0.05$Wm$^{-2}$/s. For the shortest forecast horizon $\delta = 5$ minutes and $|R| < 0.05$Wm$^{-2}$/s both models show negative skill since, by definition, the persistence model is very accurate when there are no or very small changes in the target variable. However, the error incurred in such cases is very small and has little impact in the overall error metrics.

With this analysis it is also possible to ascertain whether or not the ramp direction is predicted correctly. In this case up-ramps ($I(t+\delta) > I(t)$) and down-ramps ($I(t+\delta) < I(t)$) are flagged in the observed data, and the models’ performance is measured via the True Positive Rate (TPR) computed as TPR = TP/P $\times 100$ [%], where TP is the number of True Positives and P the number of events flagged.

Tables I and II show the TPR for the XGB and XGB+CM models for GHI and DNI, respectively. The results are discriminated by forecast horizon and binned by ramp magnitude as in the previous analysis. In order to summarize the results, the TPR values listed group up-ramps and down-ramps for a given model, horizon and ramp magnitude. The values between parenthesis indicate the difference in TPR between the two models. The table also lists for each horizon and ramp bin the number of samples identified in the testing set.
TABLE I. Number of samples and TPR values for the XGB and XGB+CM forecasts discriminated by ramp bin and forecast horizon for the GHI forecasts (testing set).

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>[0.0, 0.05]</th>
<th>[0.05, 0.1]</th>
<th>[0.1, 0.2]</th>
<th>[0.2, 0.3]</th>
<th>[0.3, 0.4]</th>
<th>$\geq 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 min</td>
<td>10545</td>
<td>3264</td>
<td>912</td>
<td>406</td>
<td>332</td>
<td>222</td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>88.6</td>
<td>89.7</td>
<td>75.4</td>
<td>74.6</td>
<td>81.0</td>
<td>85.1</td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>88.3 (-0.3)</td>
<td>90.0 (+0.3)</td>
<td>76.8 (+1.3)</td>
<td>80.3 (+5.7)</td>
<td>80.4 (+5.5)</td>
<td>92.3 (+7.2)</td>
</tr>
<tr>
<td>10 min</td>
<td>14044</td>
<td>893</td>
<td>477</td>
<td>105</td>
<td>34</td>
<td>6</td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>91.0</td>
<td>78.5</td>
<td>84.7</td>
<td>92.4</td>
<td>94.1</td>
<td>100.0</td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>91.2 (+0.1)</td>
<td>82.6 (+4.1)</td>
<td>89.7 (+5.0)</td>
<td>96.2 (+6.8)</td>
<td>97.1 (+2.9)</td>
<td>100.0 (+0.0)</td>
</tr>
<tr>
<td>15 min</td>
<td>14786</td>
<td>502</td>
<td>136</td>
<td>10</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>91.8</td>
<td>80.9</td>
<td>94.1</td>
<td>80.0</td>
<td>100.0</td>
<td>–</td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>92.1 (+0.2)</td>
<td>84.7 (+3.8)</td>
<td>96.3 (+2.2)</td>
<td>90.0 (+10.0)</td>
<td>100.0 (+0.0)</td>
<td>–</td>
</tr>
<tr>
<td>20 min</td>
<td>15057</td>
<td>227</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>92.5</td>
<td>85.5</td>
<td>77.8</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>92.7 (+0.2)</td>
<td>90.7 (+5.3)</td>
<td>85.2 (+7.4)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>25 min</td>
<td>15119</td>
<td>62</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>93.1</td>
<td>83.9</td>
<td>83.3</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>93.1 (+0.1)</td>
<td>90.3 (+6.5)</td>
<td>83.3 (+10.0)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>30 min</td>
<td>15046</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>93.7</td>
<td>82.4</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>93.6 (-0.2)</td>
<td>94.1 (+11.8)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

In this analysis one should focus on the values corresponding to larger ramps since, as mentioned above, the error accumulated when $|R(t)| < 0.05$ Wm$^{-2}$/s contributes little to the overall forecast accuracy. Thus, for $|R(t)| \geq 0.05$ Wm$^{-2}$/s XGB and XGB+CM average across all horizons TPR = 86.6% and 90.9%, respectively, for the GHI forecasts. For the DNI forecasts these values are 85.7% and 87.8%. These TPR improvements of XGB+CM over XGB are modest, specially in the case of DNI, however any small improvement in periods of large ramps will have a large effect in the overall forecast accuracy by reducing large mismatches with the measured data. It is also noteworthy that larger improvements in TPR are seen for larger ramps which strengthens the importance of the adaptive image features. The only exception occurs for the 15-minute DNI forecast for ramps $|R(t)| \in [0.3, 0.4]$ Wm$^{-2}$/s, however that involves a very small number of nine samples.

Finally, the effect of this improvement can be seen in Figs. 13 and 14. These figures show the time series plot for measured data and forecasted data with XGB and XGB+CM for GHI and DNI, respectively, for the 15-minute forecast (other forecast horizons show similar results). The figures show 5 days of increasingly larger daily average $|R|$ as annotated over each daily plot. The bottom panels show the instantaneous error and the daily MAE for each model. The figures show that the inclusion of the adaptive image features in the XGB+CM model reduces, in most cases, the instantaneous error during periods of large irradiance variability and does not degrade the forecast otherwise. That is also supported quantitatively by the daily MAE values, that show larger improvements for larger daily average $|R|$.

IV. CONCLUSIONS

In this paper we proposed a feature engineering procedure to extract information from sky images that can be used in machine learning algorithms. The procedure relies on a simple block-matching algorithm that accurately identifies the average cloud direction. Once this vector is known, an adaptive region of interest is used to select image data for the feature operators. The new adaptive image features, together with features from Pedro and Coimbra, are supplied to an XGBoost algorithm to produce intra-hour forecast for GHI and DNI. The model here proposed is a simple one: it maps irradiance and image features directly into the target forecasts avoiding common steps in solar forecasting from sky images (e.g. cloud identifications, estimation of cloud velocity field and height...
TABLE II. Same as Table I but for the DNI forecasts.

<table>
<thead>
<tr>
<th>δ</th>
<th>No Samp</th>
<th>[0, 0.05]</th>
<th>[0.05, 0.1]</th>
<th>[0.1, 0.2]</th>
<th>[0.2, 0.3]</th>
<th>[0.3, 0.4]</th>
<th>≥ 0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Samp</td>
<td>11137</td>
<td>1650</td>
<td>1138</td>
<td>512</td>
<td>554</td>
<td>683</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>67.7</td>
<td>80.6</td>
<td>74.4</td>
<td>68.9</td>
<td>73.1</td>
<td>82.1</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>73.4 (+5.7)</td>
<td>80.2 (-0.4)</td>
<td>76.0 (+1.6)</td>
<td>72.9 (+3.9)</td>
<td>75.6 (+2.5)</td>
<td>90.2 (+8.1)</td>
<td></td>
</tr>
<tr>
<td>10 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Samp</td>
<td>12988</td>
<td>1141</td>
<td>810</td>
<td>352</td>
<td>223</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>76.8</td>
<td>78.9</td>
<td>77.4</td>
<td>89.5</td>
<td>91.9</td>
<td>97.8</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>79.9 (+3.1)</td>
<td>80.9 (+2.0)</td>
<td>80.5 (+3.1)</td>
<td>92.0 (+2.6)</td>
<td>95.1 (+3.1)</td>
<td>100.0 (+2.2)</td>
<td></td>
</tr>
<tr>
<td>15 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Samp</td>
<td>13880</td>
<td>878</td>
<td>546</td>
<td>122</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>79.6</td>
<td>77.9</td>
<td>87.4</td>
<td>94.3</td>
<td>100.0</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>82.2 (+2.6)</td>
<td>80.0 (+2.1)</td>
<td>90.1 (+2.7)</td>
<td>96.7 (+2.5)</td>
<td>88.0 (-11.1)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>20 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Samp</td>
<td>14359</td>
<td>686</td>
<td>266</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>81.9</td>
<td>84.0</td>
<td>88.0</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>83.6 (+1.7)</td>
<td>87.0 (+3.1)</td>
<td>92.5 (+4.5)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>25 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Samp</td>
<td>14649</td>
<td>494</td>
<td>44</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>83.1</td>
<td>87.4</td>
<td>88.6</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>84.1 (+1.1)</td>
<td>89.1 (+1.6)</td>
<td>88.6 (+0.0)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>30 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Samp</td>
<td>14791</td>
<td>272</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB)</td>
<td>83.5</td>
<td>89.3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>TPR(XGB+CM)</td>
<td>84.6 (+1.1)</td>
<td>90.8 (+1.5)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

FIG. 13. Time series plots for five days with increasing daily average |R|. The top panel shows the measured GHI time series and the forecasted values with XGB and XGB+CM for the 15-minute forecast horizon. The annotation indicates the average |R| for the corresponding day. The bottom panel shows the instantaneous error for each forecast model. The annotations indicate the daily MAE in Wm<sup>-2</sup> for each model.

FIG. 14. Same as Fig. 13 but for the DNI forecasts.
and advection) that introduce error into the final prediction.

The results presented in this paper allow to conclude that:

- The BMA accurately identifies the angle that characterizes the bulk cloud motion. It was also seen that, instantaneously the BMA can show large error relative to the cloud direction obtained from images annotated manually. In such cases, filtering outliers based on recent BMA outputs may reduce large mismatches.

- The forecast skill for model XGB+CM that uses the new features is substantially larger than that of XGB\(^7\) and kNN.\(^6\) Furthermore, the results also show that adding the new features to the set of predictors results in a greater skill increase than just replacing the kNN algorithm by XGBoost. Typical forecast skills using the proposed adaptive method range from 20% to 35% for both GHI and DNI, with slightly higher values observed for DNI predictions. The RMSE values for GHI predictions are quite stable around 30 W m\(^{-2}\) over forecast horizons ranging from 5 to 30 minutes, while RMSE values for DNI increases from about 55 W m\(^{-2}\) for 5-min horizon forecasts to above 65 W m\(^{-2}\) for 30-min horizon forecasts.

- The new adaptive features show substantially larger values for metrics used for feature selection such as Pearson correlation coefficient and Mutual Information. The analysis of feature importance also indicates that features obtained from large ROIs are more important to GHI since it is a global measure of the irradiance. Although the effect of cloud motion it still very relevant (e.g. features derived for the incoming hemisphere are more important than features from the whole image). In the case of DNI, larger importance is assigned to features from small ROIs due to the directional nature of this variable.

- The new features also improve the forecasting of large irradiance ramps, both in terms of ramp magnitude and ramp direction.

Finally, the results of this study allows us to conclude that irradiance forecasts do benefit from adaptive sky image features, i.e. features that depend on cloud motion in a region near the apparent position of the Sun in the sky. A possible extension to this work may include the BMA filtering mentioned above, and the modification of the feature extraction algorithm to account for cloud velocity. For instance, for slow moving clouds pixels near the sun have a larger weight than pixels towards the image’s edge during feature engineering, and vice-versa for fast moving clouds.

V. REFERENCES