Sun-tracking imaging system for intra-hour DNI forecasts

Yinghao Chu, Mengying Li, Carlos F.M. Coimbra[*]

Department of Mechanical and Aerospace Engineering, Jacobs School of Engineering, Center of Excellence in Renewable Resource Integration and Center for Energy Research, University of California in San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

ARTICLE INFO

ABSTRACT

A Sun-tracking imaging system is implemented for minimizing circumsolar image distortion for improved short-term solar irradiance forecasts. This sky-imaging system consists of a fisheye digital camera mounted on an automatic solar tracker that follows the diurnal pattern of the Sun. The Sun is located at the geometric center of the sky images where the fisheye distortion is minimized. Images from this new system provide more information about the circumsolar sky cover, which provides critical information for intra-hour solar forecasts, particularly for direct normal irradiance. An automatic masking algorithm has been developed to separate the sky area from ground obstacles and the image edges for each image that is collected. Then numerical image features are extracted from the segmented sky area and are used as exogenous inputs to MultiLayer Perceptron (MLP) models for direct normal irradiance forecasts. Sixty-seven days of irradiance and image measurements are used to train, optimize, and assess the MLP-based forecast models for solar irradiance. The results show that the MLP forecasts based on the newly proposed sky-imaging system significantly outperform the reference models in terms of statistical metrics and forecast skill, particularly for shorter horizons, achieving forecast skills 18%–50% higher than the skills of a reference MLP-based model that is based on a zenith-pointed, stationary sky-imaging system.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The market penetration of solar power generation is growing rapidly and this growth imposes challenges to the operations of electric power grids [1,2], which need to be balanced in real time [3]. Ground-level solar irradiance is highly variable due to atmospheric process, especially cloud cover and aerosol content. The variability of irradiance at ground level affects the reliability of solar power generation [4], which in turn compromises the stability and integration costs for high solar penetration grids [5]. Accurate solar irradiance and power forecasts are enabling technologies that have the potential to mitigate the uncertainty of solar power generation and to optimize demand and storage solutions [1,6–10].

Effective solar forecasting methods have been developed for various temporal horizons, ranging from several minutes to a few days. Commonly employed methods include regressive or stochastic learning models [4,11–21] and physical models based on remote-sensing or local-sensing techniques [7,19,20,22–27]. For intra-hour forecasts, advanced hybrid models that integrate stochastic learning and local sensing techniques have been developed in the recent years [3,19,20,28]. When assessed in real time, the hybrid models achieve forecast skills ranging from 6% to 32% over reference persistence models [1,28,29].

To date, local-sensing systems are mostly based on sky imagers or fisheye cameras [27]. The lenses of these imagers are stationary and typically zenith-oriented. In this work, a sky-imaging system consisting of a low-cost fisheye camera mounted on an automatic solar tracker is used. The lens of this proposed system tracks the trajectory of the Sun and provides sky images centered at the apparent position of the Sun in the sky. In comparison to whole sky images from stationary imagers, Sun-centered sky images provide more information about the circumsolar sky-cover with substantially less distortion. Therefore, this new sky-imaging system has high potential to further enhance the performance of intra-hour Direct Normal Irradiance (DNI) forecasts.

In general, sky images capture not only the sky area but also the ground obstacles and darken image edges (shown in Figs. 2–7). For stationary imaging systems, manually-annotated masks are commonly used to obtain the sky area and to discard the other areas that are not useful to solar forecasts [27,30]. However, in the

---

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

new sky-imaging system presented here, each image captured requires a specific mask because the camera is non-stationary. Generating manually-annotated mask for each sky image is not practical for automatic real-time forecasts. Therefore, a smart masking algorithm has been developed to automatically analyze the color gradients of an image and to segment the sky area from the other, less-informative areas. Once the sky area is identified using the automatic masking algorithm, image features are extracted from the sky area and are used as exogenous inputs to a stochastic-learning model (MultiLayer Perception (MLP) in this work) to forecast intra-hour DNI.

We describe the new sky-imaging system in Section 2. The automatic masking algorithm and DNI forecasting model are presented in Section 3. The statistic metrics used to assess the DNI forecasts are also presented in this section. Operational results are presented in Section 4, and conclusions are presented in Section 5.

2. Data

A Multi-Filter Rotating Shadowband Radiometer (MFR-7, manufactured by Yankee Environmental Systems) has been installed at University of California San Diego (latitude = 32.881°, longitude = −117.238°) to measure the DNI components of broadband solar irradiance. The DNI data are sampled every minute and are logged using a Campbell Scientific (CR1000) data logger. Two Vivotek fisheye cameras (model FE8171V) have been installed close to the MFR-7. These cameras collect 8-bit RGB sky images (1536 × 1536 pixels) using 3.1 MP CMOS sensors and a 360° panoramic-view lens. One camera has been installed in a stationary position with its lens pointing to the zenith. This stationary camera (named SkyCam) captures whole-sky images centered at the zenith. The other camera (named SunCam) is mounted on an Eppley automatic solar tracker that points the camera lens toward the apparent position of the Sun. A photo of the employed Sun-centered devices is presented in Fig. 1. The captured images (see sample images in Figs. 6 and 7) are transferred via FTP to a UCSD server once per minute. The DNI data and the sky images are stored in a MySQL database. DNI measurements and sky images with the same time label are grouped as data instances.

The analysis of this work uses 44,393 data instances (from June 20, 2013 to August 25, 2013; nighttime measurements have been

discarded). The first 30,000 data instances (from June 20, 2013 to August 03, 2013) are assigned as a training set for model training and optimization. The remaining 14,393 data instances (Aug 03, 2013 to Aug 25, 2013) are assigned as a testing set to assess the performance of the forecasting models. Both the training and testing sets include the diverse conditions of weather and cloud content.

The MFR-7 is a first-class radiometer that meets the accuracy requirements of this work. The fisheye lenses of both the SkyCam and the SunCam are regularly cleaned to maintain satisfactory image quality. Images with excessive amounts of dust are manually discarded. In addition, the analysis of this work uses data instances when the solar elevation angle is higher than 15° to reduce the effect of ground obstacles (e.g. trees, buildings).

3. Methods

3.1. Automatic masking algorithm

SunCam moves to track the diurnal pattern of the Sun. Therefore, automatic masking algorithm has been developed to segment the sky area from obstacles and image edge for each SunCam image. The algorithm is suitable for images from both SkyCam and SunCam, making it a universal algorithm. The automatic masking algorithm initiates with a Sun locating algorithm.

3.1.1. Sun locating algorithm

The Sun locating algorithm considers seven features of a sky image: Red (R), Green (G), Blue (B), Hue (H), Saturation (S), Value (V) and Intensity (I). The features are all normalized to $0-1$ range. As shown in Fig. 2, the Hue and Saturation of Sun area are relatively small while the Red, Green, Blue, Value and Intensity are relatively large. Therefore, image features F_8 is introduced, which is defined as:

$$F_8 = \text{logical}\left(\frac{R+G+B+V+I}{5} - \frac{H+S}{2} > \theta_e\right), \qquad (1)$$

where θ_e is a threshold which is set to 0.85 empirically. As shown in Fig. 3(b), only pixels in the circumsolar region have non-zero F_8.

The following process is used to eliminate the outsiders shown in Fig. 3(b): First define [#rows, #cols] as the locations of non-zero elements of F_8; Then define [#rows*, #cols*] as the locations of non-zero elements within one standard derivation of all [#rows, #cols]; finally the Sun location in an image is calculated as:

$$\text{SunL} = [\text{mean}(\#rows*), \ \text{mean}(\#cols*)]. \qquad (2)$$

Fig. 3(c) shows the algorithm finds the location of the Sun successfully.

3.1.2. Masking algorithm

As shown in Fig. 2, the Red, Blue, Green, Value and Intensity features of obstacles are smaller than that of sky. Therefore, a feature vector is defined to differentiate sky area from obstacles:

$$\overrightarrow{F} = [R, G, B, V, I]. \qquad (3)$$

The gradient of feature vector $\overrightarrow{F}$ is calculated to find the edges of obstacles (Fig. 3(d)),

$$\nabla \overrightarrow{F} = \frac{d\overrightarrow{F}}{d\overrightarrow{x}} = \sum_{i=1}^{5} \sqrt{\left(\frac{dF_i}{dx}\right)^2 + \left(\frac{dF_i}{dy}\right)^2}. \qquad (4)$$
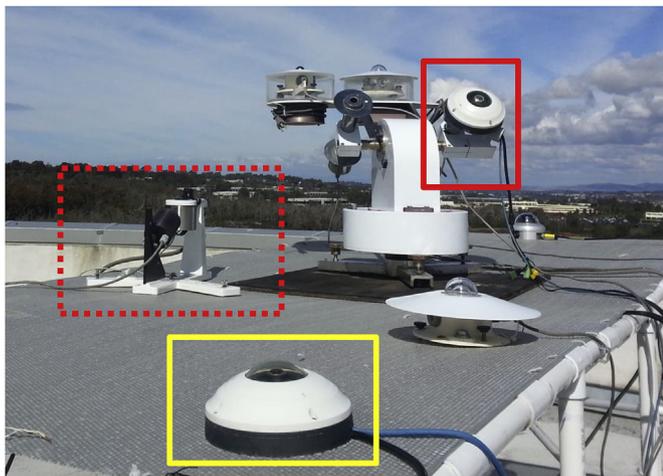
The edges are,

Fig. 1. Photo of the employed MFR-7 (quoted with red dash square), SkyCam (quoted with yellow square), and the SunCam (quoted with red square). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
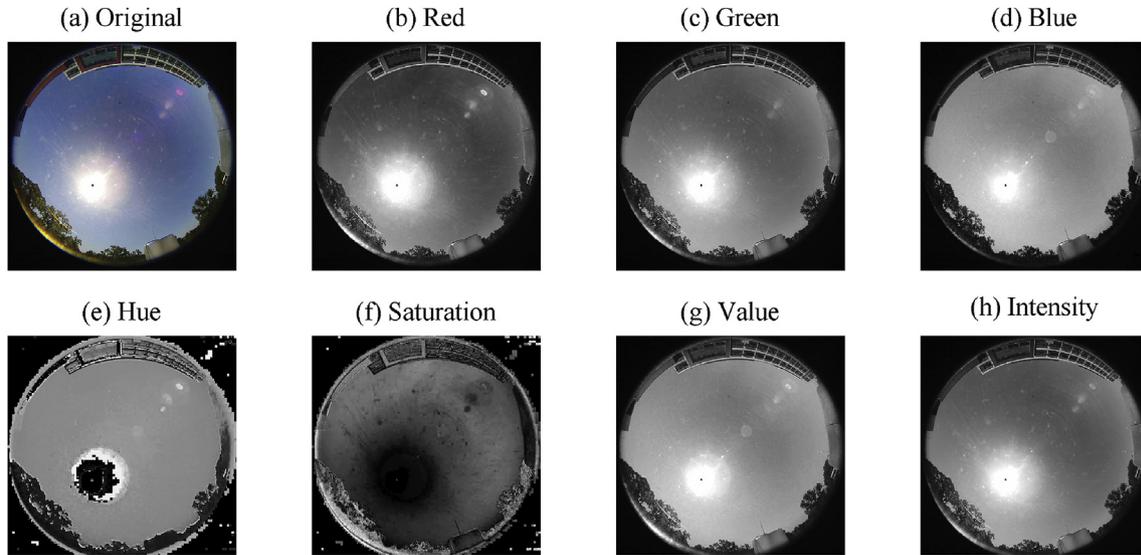
(a) Original   (b) Red   (c) Green   (d) Blue

(e) Hue   (f) Saturation   (g) Value   (h) Intensity

Fig. 2. Seven features of sample sky image.

(a) Original   (b) F_8 (c) Sun Location   (d) feature gradient

(e) Edge$_{raw}$ (f) Edge   (g) Edge points   (h) Mask
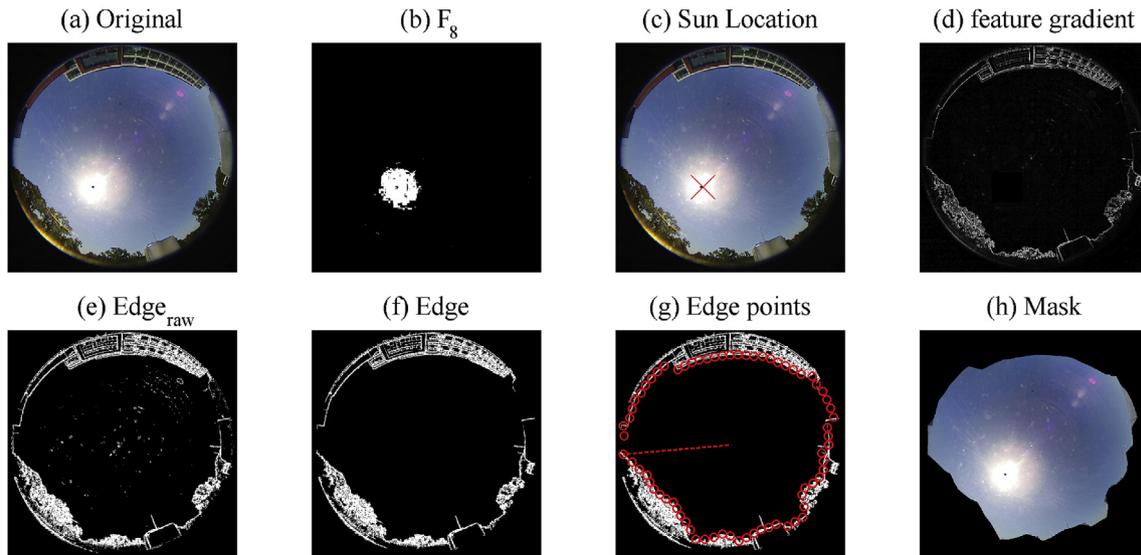
Fig. 3. The masking result of the sample sky image.

$$\text{Edge}_{raw} = \text{logical}\left(\nabla \vec{F} > \nabla \vec{F}_c \right). \tag{5}$$

where $\nabla \vec{F}_c$ is such when 10% pixels have $\nabla \vec{F} > \nabla \vec{F}_c$. The Edge$_{raw}$ of sample sky image is shown in Fig. 3(e). The center of the Sun appears as a black dot in some of the sky images because the Sun is too bright for the digital camera. The black dot results in a large feature gradient, which is similar to the edge between sky area and obstacles. Therefore, the Sun locating algorithm is used to identify and eliminate the black dot. The large feature gradient caused by dusts on cameras are also eliminated using Matlab 2013 *bwareaopen* function (Fig. 3(f)). The edge points are located and connected to generate the obstacle mask as shown in Fig. 3(g) and (h).

During partly-cloudy periods, the edges of clouds and sky can be mistakenly identified as edges of obstacles by the masking algorithm (shown in Fig. 4(b)). Because the location of the obstacles are nearly unchanged at the same time in several consecutive days

while the location of the clouds are different, the common edges from five images taken at the same time from five consecutive days are considered as the 'true' edges of obstacles. After the correction using five images, the algorithm works successfully for partly-cloudy images as shown in Fig. 4(c).

3.1.3. Performance of the masking algorithm on SunCam

The Sun always locates at the centers of the SunCam images. Therefore, distance d_e of the algorithm-derived Sun locations and the center of the images are calculated to assess the performance of the Sun locating algorithm. In general, the Sun-locating algorithm achieve excellent performance during clear period (average $d_e < 10$ pixels). However, the Sun-locating algorithm may have relatively large errors during overcast period when the Sun is obscured by the stratiform clouds. During partly-cloudy period, the cloud enhancement [31,32] may slightly degenerate the accuracy of the Sun-locating algorithm. The d_e calculated from days with different weather are plotted in Fig. 5. The average distance is 57-pixel that

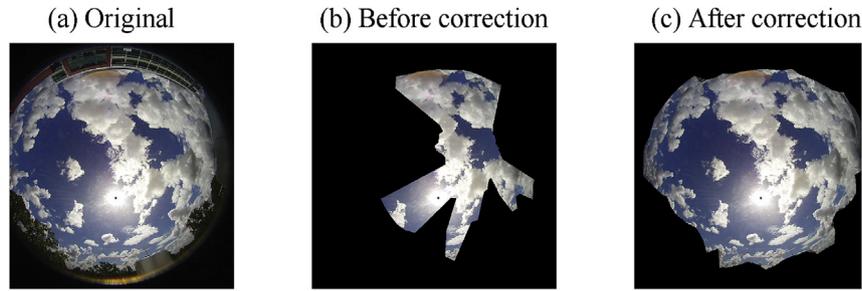(a) Original   (b) Before correction   (c) After correction

Fig. 4. The masking correction for cloudy time.

corresponds to 3.71% of the size of the image.

The automatic masking algorithm is applied to 15 randomly picked SunCam images, which consist of different weather conditions, and achieves satisfied performance. The example results of automatic masking algorithm for SunCam images are shown in Fig. 6. For both overcast time (Fig. 6(a)), partly-cloudy time (Fig. 6(b) and (c)) and clear time (Fig. 6(d)), the algorithm masks out obstacles correctly.

3.2. Multilayer preceptron

MultiLayer Preceptrons networks (MLPs) are widely used as stochastic learning tools for solar modeling and forecasting [33–35]. MLPs are one of the most established ANN architectures and are capable of performing non-linear classification and regression [36]. The basic processing elements of MLP are called neurons, which are placed in layers. The layers between the first input layer and the last output layer are called hidden layers [37]. In this work, the employed MLPs have 1 hidden layer with 7 neurons as suggested by Chu et al. [27,29]. Each neuron of the hidden layer sums the weighted inputs, adds a bias to the sum, and then process the sum using an activation function to generate outputs. The signal processing of neurons can be mathematically expressed as:

$$Y_i = f\left(\sum_{j=1}^{M}\left(w_{ij}X_j + \beta_{ij}\right)\right), \tag{6}$$

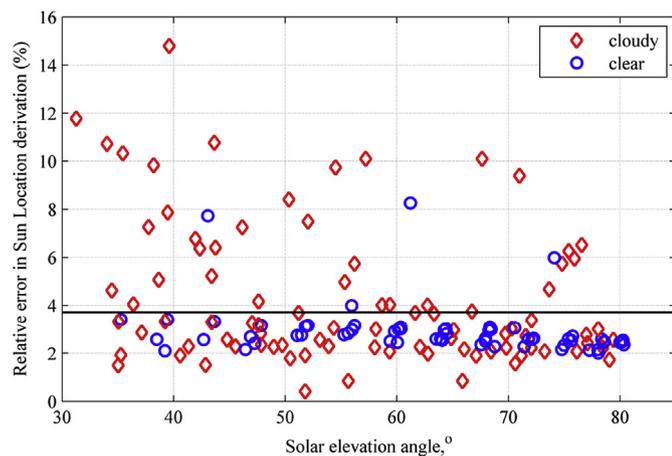where Y_i is the output of the i-th neuron on current layer, w_{ij} and β_{ij}

Fig. 5. Relative Sun location derivation error with respect to solar zenith angle is plotted. The relative error is defined as d_e/M. Results of six different days are plotted, including both clear periods and cloudy periods. The six days are June 21st, June 28th, July 8th, July 15th, July 23rd and August 14th.
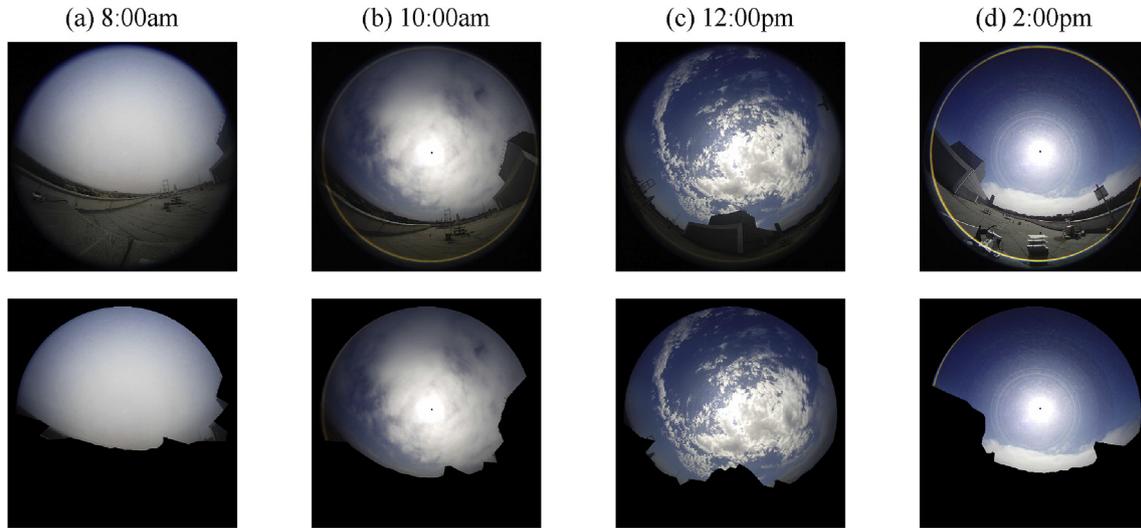
are the weight and bias of the j-th input on the i-th neuron, M is the number of inputs, X_j is j-th output from previous layer, and $f(\cdot)$ is the activation function, which is a sigmoidal function in this work [38].

$$f(y) = \frac{1}{1 + e^{-y}}. \tag{7}$$

A supervised learning process is applied to estimate the weight w_{ij} and bias β_{ij} using the training data. The learning process is the Bayesian regularization process with Levenberg-Marquardt optimization [3]. This is an iterative process that stops once the discrepancy with respect to target is lower than a preset value. The procedure of the training can be summarized as [39]:

1. Set the number of layers, number of neurons in each layer, and a tolerance parameter $\varepsilon > 0$
2. Initialize the values of MLP parameters weight vector (w_j^i, $i = 0,1,2,…,L$) randomly.
3. Calculate the neuron outputs $X_j^i = f((w_j^i)^T X^{(i-1)})$ and final output errors $\varepsilon_j^L = Y_j - X_j^L$ using the initial MLP weights and the training data.
4. Calculate the output delta $\delta_j^L = \varepsilon_j^L f'((w_j^L)^T X^{L-1})$.
5. Calculate the propagation errors for hidden neurons $\varepsilon_j^i = \sum_{i=1}^{N+1}\delta_j^{i+1}w_j^{i+1}$ and the hidden delta $\delta_j^i = \varepsilon_j^i f'((w_j^i)^T X^{(i-1)})$.
6. Update the weight vector $w_j^{i+1} = w_j^i - \eta\delta_j^i X^{i-1}$, where η is the learning rate coefficient.
7. Repeat step 3–6 recursively until the difference in error changes $\Delta\varepsilon = \varepsilon^i - \varepsilon^{i-1}$ less than the tolerance parameter $\Delta\varepsilon \le e_T$.
8. Output the weights determined in the final round.

Once the training is finished, the MLP-based model is able to perform real-time forecasts with new inputs $X_{new}(t)$.

$$Y(t + \Delta t) = \text{MLP}(X_{new}(t)). \tag{8}$$

3.3. Image processing and numerical inputs

Stochastic learning methods require numerical inputs. Therefore, images need to be processed and translated into numerical information. The numerical information is then used as exogenous inputs to the proposed stochastic MLPs. In this work, we calculate three Normalized Red Blue Ratio (NRBR) parameters [27,40] for an image: mean, standard deviation, and entropy. Mean

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\text{NRBR}_i, \tag{9}$$

where n is the number of pixels in the sky area of an image. Standard deviation

(a) 8:00am   (b) 10:00am   (c) 12:00pm   (d) 2:00pm

Fig. 6. Masking results of SunCam images.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{NRBR}_i - \mu)^2}. \tag{10}$$

Entropy

$$e = -\sum_{j=1}^{n_B} p_j log_2\left(p_j\right), \tag{11}$$

where p_j is the relative frequency for the jth bin (out of $n_B = 256$ evenly spaced bins).

The NRBR parameters are employed instead of detecting actual cloud cover information for two reasons: (1) Available cloud detection methods are not specified for the SunCam images and their accuracy are compromised when applied to SunCam images. (2) Comparing to cloud detection methods, calculating the NRBR parameters significantly reduces the computational cost and achieves competitive performance [29]. Sample original and NRBR images for both SkyCam and SunCam are shown in Fig. 7.

In addition to the NRBR parameters, seven time-lagged DNI measurements (ranging from 0 to 30 min in steps of 5 min) are used as endogenous inputs because the latest measurements of DNI are highly informative for very short-term DNI forecasts [3]. To evaluate the benefits of SunCam-based forecasts, three MLP-based forecasting models are trained and tested using three different input sets respectively: MLP endogenous model, which uses endogenous inputs only; MLP with SkyCam, which uses endogenous inputs with SkyCam NRBR parameters; and the proposed MLP with SunCam, which uses endogenous inputs with SunCam NRBR parameters.

3.4. Model assessments

Four statistical metrics are used to assess the performance of the employed MLP-based models: mean biased error (MBE)

$$\text{MBE} = \frac{1}{n}\sum_{t=1}^{n}\left(\widehat{B}(t) - B(t)\right), \tag{12}$$

where the n is the number of testing instances, $\widehat{B}$ is the prediction of the MLP-based models and B is the measured irradiance mean

absolute error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left(\left|\widehat{B}(t) - B(t)\right|\right), \tag{13}$$

root mean square error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\widehat{B}(t) - B(t)\right)^2}, \tag{14}$$

and forecasting skill (s)

$$s = 1 - \frac{\text{RMSE}}{\text{RMSE}_p}. \tag{15}$$

The forecast skill measures the improvements of the MLP-based models over a reference model in terms of RMSE.

In this work, the reference model is the persistence model, which is the simplest forecasting model and is widely recommended as a reference model for solar forecasts in literature [1,9,21,28,41]. The persistence model assumes that the clear-sky index [1] persists within the forecast horizon. Therefore, persistence model achieves excellent performance under stationary weather conditions. The persistence model is mathematically expressed as:

$$\widehat{B}_p(t + FH) = \frac{B(t)}{B_{\text{clr}}(t)} \times B_{\text{clr}}(t + FH), \tag{16}$$

where $\widehat{B}_p$ is the prediction of the persistence model, subscript p represents persistence, t is the time when forecast is performed, FH is the forecast horizon, B is the forecast variable (DNI in this work), and B_{clr} is the clear-sky DNI, which is predicted using a clear sky model.

The clear-sky model used in this work is an empirical model [3,7]. This empirical model computes the dependence of the clear-sky DNI on the sine of the solar elevation angle by a seven-order polynomial expression. Least square method is applied to derive the parameters of this polynomial expression using data from six clear-sky days, which are manually selected from the training set. The mean relative error for this empirical model is 1.93% when tested using the training data. Compared to most available clear-sky
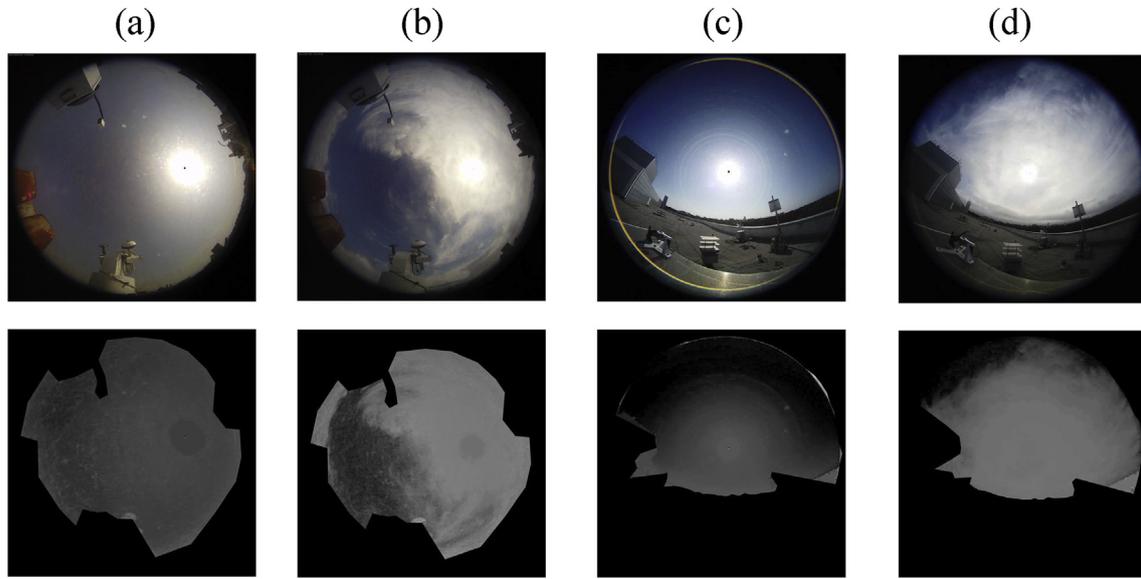
Fig. 7. Sample original images (top row) and NRBR images (bottom row). (a) and (b) SkyCam images (2013-06-30 and 2013-07-02), (c) and (d) SunCam images (2013-06-30 and 2013-07-02).

models, this empirical model is superior in terms of accuracy, simplicity, and ease of implementation, particularly for the available data period [29]. If Linke Turbidity coefficient is available, the Ineichen clear-sky model [42,43] is recommended for longer period at different locations.

4. Results and discussion

The MLP-based models, which are trained using different input sets, are evaluated using the independent testing data and their performances are compared to the reference persistence model. The testing process simulates the real-time forecast scenario on data collected from Aug 03, 2013 to Aug 25, 2013. The forecasting performance in terms of the assessment metrics (defined in Section 3.4) are presented in Table 1. All MLP-based models and the persistence models exhibit small MBEs. However, the MLP-based models significantly outperform the persistence models in terms of RMSE and forecast skill. MLP-based models that consider sky imaging information achieve significantly higher forecast skill than the MLP-based models that only consider endogenous inputs. The improvement on forecast accuracy of stochastic learning models

due to the integration of sky imaging system is also observed and extensively discussed in earlier studies [28,44]. The results show that the MLP-based models with SunCam achieve the highest forecast skills that are 18%—50% higher than the forecast skills of models with regular stationary SkyCam system. The SunCam images are focused on the Sun and therefore provide more detailed sky-cover information about the circumsolar region. The circumsolar sky-cover is particular important for very short-term solar forecasts. Therefore, the SunCam-based forecasts achieve higher forecast skills for shorter forecast horizons (e.g. 5-min) as presented in Table 1.

To further analyze the performance of the employed forecasts, the Cumulative Density Functions (CDF) of the absolute forecast errors for the 5-, 10-, 15- and 20-min horizons are plotted in Fig. 8. In general, the averaged cumulative probability at different error levels tends to decrease with the increase of forecast horizon since longer horizons are more difficult to forecast [1]. However, more than 60% of the errors from all employed models have magnitudes less than 0.1 kW/m^2 regardless of the forecast horizons. Most of these small errors are observed during stationary weather conditions such as clear or overcast periods. These results match the

Table 1
Testing results of the different forecasting models for 5, 10, 15, and 20 min horizons. The values in brackets are the relative errors.

| Metrics	Models	Forecast horizons			
		5-min	10-min	15-min	20-min
MBE (W/m^2)	Persistence	−2.5 (−0.003)	−5.5 (−0.007)	−8.0 (−0.01)	−11.3 (−0.014)
	MLP Endogenous	−3.9 (−0.005)	−5.8 (−0.007)	−7.7 (−0.01)	−10.9 (−0.013)
	MLP with SkyCam	−0.6 (−0.001)	−1.9 (−0.002)	−3.3 (−0.004)	−7.0 (−0.009)
	MLP with SunCam	7.1 (0.009)	6.6 (0.008)	5.1 (0.006)	2.2 (0.003)
MAE (W/m^2)	Persistence	71.5 (0.088)	88.4 (0.109)	104.8 (0.129)	117.0 (0.144)
	MLP Endogenous	71.6 (0.088)	86.3 (0.106)	101.4 (0.125)	112.8 (0.139)
	MLP with SkyCam	68.8 (0.085)	85.7 (0.106)	102.4 (0.126)	115.3 (0.142)
	MLP with SunCam	65.2 (0.08)	81.9 (0.101)	97.1 (0.12)	109.5 (0.135)
RMSE (W/m^2)	Persistence	139.8 (0.172)	169.6 (0.209)	195.2 (0.241)	215.8 (0.266)
	MLP Endogenous	129.7 (0.16)	156.3 (0.193)	178.1 (0.22)	194.7 (0.24)
	MLP with SkyCam	124.3 (0.153)	151.7 (0.187)	172.5 (0.213)	188.1 (0.232)
	MLP with SunCam	116.3 (0.143)	143.2 (0.177)	165.3 (0.204)	183.1 (0.226)
Forecast skill	MLP Endogenous	7.3%	7.8%	8.7%	9.8%
	MLP with SkyCam	11.1%	10.6%	11.6%	12.8%
	MLP with SunCam	16.9%	15.5%	15.3%	15.1%
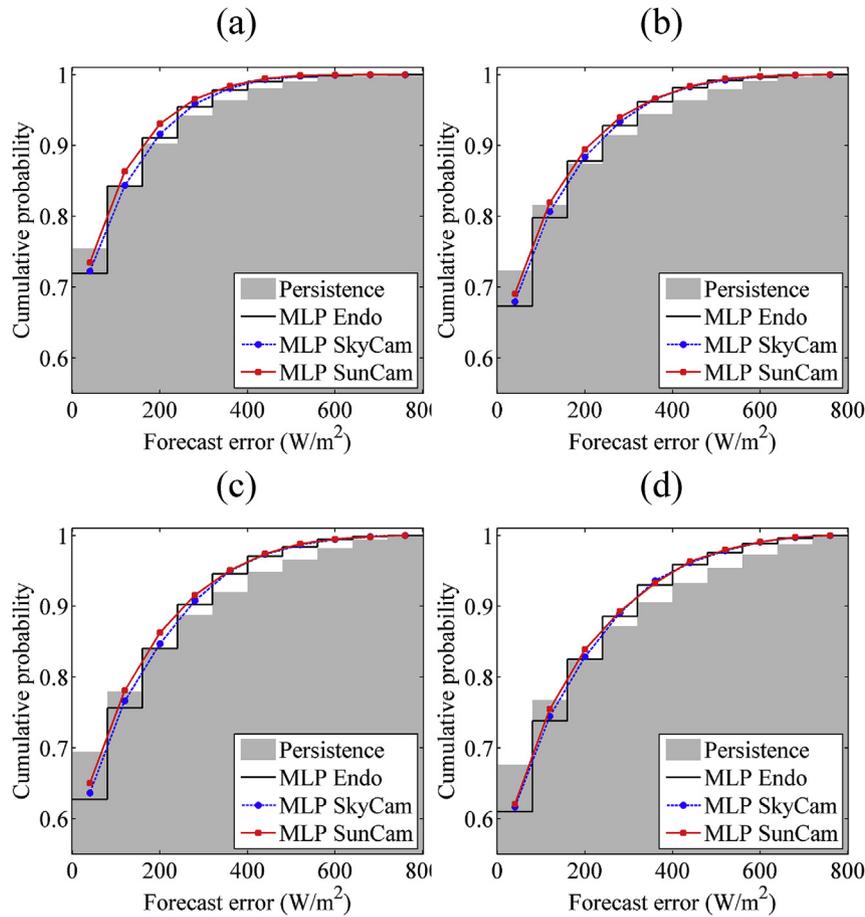
Fig. 8. Plots of Cumulative Density Functions (CDFs) of forecast errors. (a) 5-min forecasts, (b) 10-min forecasts, (c) 15-min forecasts, and (d) 20-min forecasts. The size of bins used to create the CDF is 80 W/m².

discussions in Refs. [7,28] that most solar forecasting models, particularly the persistence models, achieve excellent performance during stationary weather conditions when irradiance variability is low. However, the errors of persistence models increase rapidly with the irradiance variability during partly-cloudy periods [27] because the persistence forecasts are unable to predict the irradiance ramps caused by cloud covers [27]. Consequently, persistence forecasts have a relatively higher percentage of errors with high absolute values (>0.2 kW/m²).

The persistence errors represent the levels of variability in the DNI time series [28]. The CDFs of the MLP forecasts approach 1 faster than the persistence forecast, indicating that the MLP forecasts significantly reduce the forecasting uncertainty in DNI time series. Compared to the persistence forecasts, all MLP-based forecasts have less occurrence of large errors and therefore have higher cumulative probabilities for moderate error level (0.1 kW/m² <|ε| < 0.4 kW/m²). The SunCam-MLP models show significantly higher cumulative probability for the moderate error level, indicating that SunCam-based MLP forecasts have less proportion of mid- and large-size errors. For instance, in Fig. 8, the probability of SunCam-MLP errors greater than 0.3 kW/m² is 0.06 while the probability of persistence errors greater than 0.3 kW/m² is 0.086, which result in heavier tails in the CDFs of persistence error. Therefore, the SunCam-MLP forecasts achieve the lowest RMSE and highest forecast skills in Table 1.

To quantitatively analyze the differences among CDFs from different models, the areas above the CDF are calculated using the following mathematics expression:

Table 2
The results of A_ε for different models and forecast horizons.

Models	Forecast horizons			
	5-min	10-min	15-min	20-min
Persistence	50.2	63.7	78.9	88.6
MLP Endogenous	48.7	63.4	78.5	87.3
MLP with SkyCam	46.9	60.7	75.1	84.4
MLP with SunCam	42.3	57.1	70.7	82.2

$$A_\varepsilon = \int_0^\infty (1 - \text{CDF}(\varepsilon))d\varepsilon, \tag{17}$$

where ε is the absolute error in the x-axis of Fig. 8. A_ε for all investigated forecasts are presented in Table 2. Smaller value of A_ε indicates that the CDF approach 1 faster and the percentage of large errors is smaller. For instance, the results show that A_ε increases with forecast horizon for all models, which matches the previous discussion that longer horizon forecasts are more challenging and have larger forecasting uncertainties. The SunCam-based MLP forecasts achieves the lowest A_ε for all forecast horizons.

5. Conclusions

A novel sky-imaging system (SunCam) was developed using a low-cost, high-resolution fisheye camera mounted on an automatic

solar tracker. The camera is mounted on the solar tracker and is continuously orientated toward the Sun to provide Sun-centered sky images. An automatic masking algorithm is developed to calibrate the solar position and to separate the sky area from ground obstacles in the sky images. Numerical image features are extracted from the SunCam images to be used as exogenous inputs to MultiLayer-Perceptron-network (MLP) models, and these models are used to automatically forecast DNI for 5-, 10-, 15-, and 20-min horizons.

The performance of the SunCam-based MLP models is assessed using data from 23 continuous days, which include widely different weather conditions and cloud cover content. The performance of the SunCam-MLP model is compared to three reference models: a persistence model, an MLP-based model that considers only endogenous inputs, and an MLP-based model that considers image features extracted from a traditional stationary sky-imaging system (SkyCam).

The results show that the new MLP-based model with the proposed sky-imaging achieves robust improvements in forecasting skills, ranging from 15.1% to 16.9% depending on forecast horizons. There is a robust improvement over both the reference persistence model and the forecast obtained with a fixed camera (SkyCam) over all forecast horizons.

This work successfully demonstrates the potential of proposed imaging system to improve the short-term solar forecasts. The current imager has been implemented an expensive Eppley Sun tracker (USD 12,000). A new Sun tracker with light CCD camera is under development in our group that will substantially reduce the instrument cost below USD 500. The new tracking instrument will enable multiple tests of the proposed imaging system for longer time periods and at different locations.

Acknowledgments

References

[1] R.H. Inman, H.T.C. Pedro, C.F.M. Coimbra, Solar forecasting methods for renewable energy integration, Prog. Energy Combust. Sci. 39 (6) (2013) 535–576.
[2] IEA, Technology Roadmap Solar Photovoltaic Energy, Tech. Rep., IEA, 2014.
[3] Y. Chu, H.T.C. Pedro, C.F.M. Coimbra, Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning, Sol. Energy 98 (2013) 592–603.
[4] M. Lave, J. Kleissl, Solar variability of four sites across the state of Colorado, Renew. Energy 35 (12) (2010) 2867–2873.
[5] G.K. Singh, Solar power generation by PV (photovoltaic) technology: a review, Energy 53 (2013) 1–13.
[6] S. Sayeef, C. Scientific, Solar Intermittency: Australia's Clean Energy Challenge: Characterising the Effect of High Penetration Solar Intermittency on Australian Electricity Networks, CSIRO, Sydney, Australia, 2012.
[7] S. Quesada-Ruiz, Y. Chu, J. Tovar-Pescador, H.T.C. Pedro, C.F.M. Coimbra, Cloud-tracking methodology for intra-hour DNI forecasting, Sol. Energy 102 (2014) 267–275.
[8] S.R. West, D. Rowe, S. Sayeef, A. Berry, Short-term irradiance forecasting using skycams: motivation and development, Sol. Energy 110 (2014) 188–207.
[9] Y. Chu, B. Urquhart, S.M.I. Gohari, H.T.C. Pedro, J. Kleissl, C.F.M. Coimbra, Short-term reforecasting of power output from a 48 MWe solar PV plant, Sol. Energy 112 (2015) 68–77.
[10] W.K. Yap, V. Karri, An off-grid hybrid PV/diesel model as a planning and design tool, incorporating dynamic and ANN modelling techniques, Renew. Energy 78 (0) (2015) 42–50.
[11] S. Sinha, S. Kumar, T. Matsumoto, T. Kojima, Application of system identification modeling to solar hybrid systems for predicting radiation, temperature and load, Renew. Energy 22 (1–3) (2001) 281–286.
[12] M. Chaabene, M. Ben Ammar, Neuro-fuzzy dynamic model with Kalman filter to forecast irradiance and temperature for solar energy systems, Renew. Energy 33 (7) (2008) 1435–1443.
[13] J. Cao, X. Lin, Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks, Energy Convers. Manag. 49 (6) (2008) 1396–1406.
[14] L. Martín, L.F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, M. Cony, Prediction of global solar irradiance based on time series analysis: application to solar thermal power plants energy production planning, Sol. Energy 84 (10) (2010) 1772–1781.
[15] 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 (5) (2010) 807–821.
[16] A. Mellit, H. Eleuch, M. Benghanem, C. Elaoun, A.M. Pavan, An adaptive model for predicting of global, direct and diffuse hourly solar irradiance, Energy Convers. Manag. 51 (4) (2010) 771–782.
[17] H.T.C. Pedro, C.F.M. Coimbra, Assessment of forecasting techniques for solar power production with no exogenous inputs, Sol. Energy 86 (7) (2012) 2017–2028.
[18] R. Marquez, C.F.M. Coimbra, Proposed metric for evaluation of solar forecasting models, ASME J. Sol. Energy Eng. 135 (2013) 0110161–0110169.
[19] R. Marquez, V.G. Gueorguiev, C.F.M. Coimbra, Forecasting of global horizontal irradiance using sky cover indices, ASME J. Sol. Energy Eng. 135 (2013) 0110171–0110175.
[20] 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.
[21] L. Nonnenmacher, A. Kaur, C.F.M. Coimbra, Verification of the SUNY direct normal irradiance model with ground measurements, Sol. Energy 99 (2014) 246–258.
[22] A. Hammer, D. Heinemann, E. Lorenz, B. Luckehe, Short-term forecasting of solar radiation: a statistical approach using satellite data, Sol. Energy 67 (1999) 139–150.
[23] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, D. Renne, T.E. Hoff, Validation of short and medium term operational solar radiation forecasts in the US, Sol. Energy 84 (5) (2010) 2161–2172.
[24] 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 testbed, Sol. Energy 85 (11) (2011) 2881–2893.
[25] R. Marquez, C.F.M. Coimbra, Intra-hour DNI forecasting methodology based on cloud tracking image analysis, Sol. Energy 91 (2013) 327–336.
[26] L. Nonnenmacher, C.F.M. Coimbra, Streamline-based method for intra-day solar forecasting through remote sensing, Sol. Energy 108 (2014) 447–459.
[27] Y. Chu, H.T.C. Pedro, L. Nonnenmacher, R.H. 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 (2014) 1995–2007.
[28] Y. Chu, H.T.C. Pedro, M. Li, C.F.M. Coimbra, Real-time forecasting of solar irradiance ramps with smart image processing, Sol. Energy 114 (2015) 91–104.
[29] Y. Chu, M. Li, H.T.C. Pedro, C.F.M. Coimbra, Real-time prediction intervals for intra-hour DNI forecasts, Renew. Energy 83 (2015c) 234–244.
[30] B. Urquhart, M. Ghonima, D. Nguyen, B. Kurtz, J. Kleissl, Solar Energy Forecasting and Resource Assessment, Elsevier, 2013.
[31] P. Emck, M. Richter, An upper threshold of enhancend global shortwave irradiance in the troposphere derived from field measurements in tropical mountains, J. Appl. Meteorol. Climatol. 47 (11) (2008) 2828–2845.
[32] R. Tapakis, A.G. Charalambides, Enhanced values of global irradiance due to the presence of clouds in Eastern Mediterranean, Renew. Energy 62 (2014) 459–467.
[33] C.M. Bishop, Neural networks and their applications, Rev. Sci. Instrum. 65 (6) (1994) 1803–1832.
[34] 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 (5) (2011) 746–756.
[35] S.X. Chen, H.B. Gooi, M.Q. Wang, Solar radiation forecast based on fuzzy logic and neural networks, Renew. Energy 60 (0) (2013) 195–201.
[36] A. Sözen, E. Arcaklıoğlu, M. Özalp, N. Çağlar, Forecasting based on neural network approach of solar potential in Turkey, Renew. Energy 30 (7) (2005) 1075–1090.
[37] B. Garca-Domingo, M. Piliougine, D. Elizondo, J. Aguilera, CPV module electric characterisation by artificial neural networks, Renew. Energy 78 (0) (2015) 173–181.
[38] A. Mellit, S.A. Kalogirou, Artificial intelligence techniques for photovoltaic applications: a review, Prog. Energy Combust. Sci. 34 (5) (2008) 574–632.
[39] D.J. Rumelhart, D. Zipser, Feature discovery by competitive learning, Cogn. Sci. 9 (1985) 75–112.
[40] Q. Li, W. Lu, J. Yang, A hybrid thresholding algorithm for cloud detection on ground-based color images, J. Atmos. Ocean. Technol. 28 (2011) 1286–1296.
[41] A. Kaur, H.T.C. Pedro, C.F.M. Coimbra, Ensemble re-forecasting methods for enhanced power load prediction, Energy Convers. Manag. 80 (2014) 582–590.
[42] P. Ineichen, R. Perez, A new airmass independent formulation for the Linke turbidity coefficient, Sol. Energy 73 (3) (2002) 151–157.
[43] P. Ineichen, A broadband simplified version of the Solis clear sky model, Sol. Energy 82 (2008) 758–762.
[44] 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.