Inter-hour direct normal irradiance forecast with multiple data types and time-series

Tingting ZHU[1], Hai ZHOU[2], Haikun WEI[1] (iD)**, Xin ZHAO[1],**
Kanjian ZHANG[1], Jinxia ZHANG[1]

Abstract Boosted by a strong solar power market, the electricity grid is exposed to risk under an increasing share of fluctuant solar power. To increase the stability of the electricity grid, an accurate solar power forecast is needed to evaluate such fluctuations. In terms of forecast, solar irradiance is the key factor of solar power generation, which is affected by atmospheric conditions, including surface meteorological variables and column integrated variables. These variables involve multiple numerical time-series and images. However, few studies have focused on the processing method of multiple data types in an inter-hour direct normal irradiance (DNI) forecast. In this study, a framework for predicting the DNI for a 10-min time horizon was developed, which included the nondimensionalization of multiple data types and time-series, development of a forecast model, and transformation of the outputs. Several atmospheric variables were considered in the forecast framework, including the historical DNI, wind speed and direction, relative humidity time-series, and ground-based cloud images. Experiments were conducted to evaluate the performance of the forecast framework. The experimental results demonstrate that the proposed method performs well with a normalized mean bias error of 0.41% and a normalized root mean square error (nRMSE) of 20.53%, and outperforms the persistent model with an improvement of 34% in the nRMSE.

Keywords Inter-hour forecast, Direct normal irradiance, Ground-based cloud images, Multiple data types, Multiple time-series

✉ Haikun WEI
  hkwei@seu.edu.cn

  Tingting ZHU
  tingting_zhu2018@163.com

  Hai ZHOU
  zhouhai@epri.sgcc.com.cn

  Xin ZHAO
  zhaoxin0504@163.com

  Kanjian ZHANG
  kjzhang@seu.edu.cn

  Jinxia ZHANG
  jxzhang@seu.edu.cn

[1] Key Laboratory of Measurement and Control of CSE, Ministry of Education, School of Automation, Southeast University, Nanjing 210096, China

[2] State Key Laboratory of Operation and Control of Renewable Energy and Storage Systems (China Electric Power Research Institute), Nanjing 210003, China

1 Introduction

Solar energy is an important renewable energy resource, but is intermittent in the short period owing to the fluctuations of solar radiation. Dramatic fluctuations cause the energy output of a solar power plant to rapidly decrease from hundreds of megawatts to zero output within a few minutes, and bring about huge risk to the stability of the electrical grid [1]. Therefore, an accurate forecast of solar power is the premise and key technology of the grid-connection for photovoltaic (PV) or concentrated solar thermal (CST) plants [2, 3]. For various power system operations, the geographical and temporal requirements differ in a solar power forecast [4, 5].

Solar power significantly depends on the solar irradiance, such as the global horizontal irradiance (GHI) for PV plants and the direct normal irradiance (DNI) for CST plants [6], and thus the solar power output forecast can be transformed into a solar irradiance (GHI or DNI) forecast [7–9]. In this study, we focus on a DNI forecast. The DNI is mainly affected by surface meteorological variables and column integrated variables, such as clouds and water vapor. Except for clouds, these atmospheric variables have been considered in most clear-sky models for a DNI estimation [10] or numerical weather prediction for a day-ahead solar irradiance forecast [11, 12] and long-term solar irradiance estimation [13]. However, for an inter-hour DNI forecast in all sky conditions, clouds have been the only factor considered in certain forecast models, and other atmospheric variables have been ignored [14–19]. Recently, the majority of all-sky DNI forecast models have employed two main parts: clear-sky DNI forecast and cloud fraction forecast [20–22]. In this way, except for clouds, the effectiveness of other atmospheric variables is considered in a clear-sky DNI estimation. However, the accuracy of clear-sky forecast is limited by the immeasurability of most clear-sky atmospheric variables under all-sky conditions, and thus most clear-sky models applied in all-sky forecast are empirical with few atmospheric variables [23–25].

To achieve a higher accuracy of an inter-hour DNI forecast, [26] modified a clear sky model by considering the variation of the atmospheric components. Reference [27] also developed a forecast method using adaptive clear-sky models. However, the improvement in accuracy was limited by adjusting these models. In this study, a framework of an inter-hour DNI forecast was developed by directly considering the atmospheric variables, including the historically measured relative humidity, wind speed and direction, DNI, and clouds, presented through a numerical time-series or images. To fuse different types of data and avoid some inputs overwhelmed by the large magnitude of other variables, all inputs were transformed into dimensionless variables. Firstly, a ground-based cloud image was in-painted and corrected to obtain more accurate all-sky information, and the cloud covers in six key areas were then extracted separately, each of which was set as the input of the forecast model. Secondly, the measured historical DNI was transformed into a clear-sky index using a clear-sky model, and was set as an input of the forecast model. Finally, a support vector regression (SVR) model was employed to predict the DNI for the following 10 minutes, with multiple inputs including the cloud cover, clear-sky index, wind speed, and relative humidity (instead of water vapor).

The remainder of this paper is constructed as follows. The instrument and data collection are introduced in Section 2. A forecast framework based on multiple data types

and time-series is described in Section 3. Experiments carried out to evaluate the performance of the proposed method, along with the results and a discussion, are detailed in Section 4. Finally, some concluding remarks are provided in Section 5.

2 Data collection

All ground-based cloud (GBC) images and measured data employed were downloaded from the open database of the Solar Radiation Research Laboratory (39.74°N, 105.18°W, at 1828.8 m above sea level), which was provided by the National Renewable Energy Laboratory (NREL) [28].
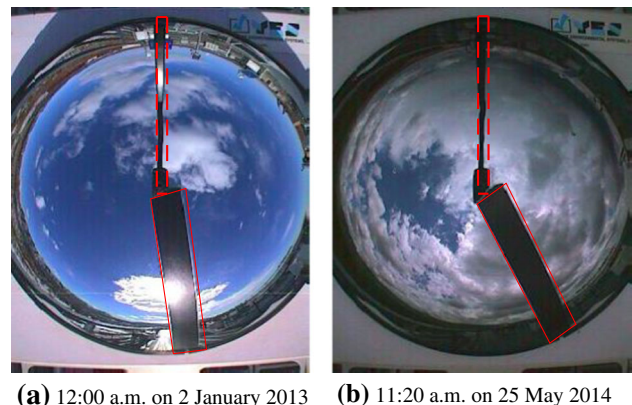
2.1 GBC images

A Total Sky Imager (TSI-880, Yankee Co.) was used to automatically capture 24-bit color sky images every 30 s. Figure 1 shows two raw GBC images in which a valid circular sky region is distorted by a fisheye lens, and a part of the sky is obstructed by both the support arm and a shadow band, which are marked with dashed-line and solid-line rectangles, respectively.

A total of 73 clear-sky GBC images were selected from 2013 to construct a clear-sky image dataset (CSID). All clear-sky images correspond to a solar zenith ranging from 89° to 17° (step size of 1°). The CSID set was constructed for the following preprocessing of the image inpainting to detect error pixels around the sun, as discussed in Section 3.1.

2.2 Measured numerical data

Some measured data were employed in the following proposed forecast framework, including the DNI, related humidity, wind speed, and wind direction. The DNI was

(a) 12:00 a.m. on 2 January 2013   **(b)** 11:20 a.m. on 25 May 2014

Fig. 1 Raw GBC images taken by TSI-880

measured using a pyrheliometer (Kipp & Zonen, CH1), mounted on a sun-following tracker and pointing precisely to the sun center within a 5° field of view, with a sampling interval of 1 min. An average of ten measured 1-min DNI values, as well as the wind speed and wind direction, were organized as the experimental data for the forecast model, contributing to the 10-min resolution dataset applied in this study. It is worth noting that the wind direction is from the north and ranges clockwise from 0° to 360°. The average of two wind directions is not always one-half of their arithmetic sum, but should be at the angular bisector of the smaller angle between the two directions. In other words, it is the opposite of their arithmetical average if their bias is greater than 180°. Considering the wind speed, the vector average method (VAM) [29] was used, and the average wind direction (D_{ave}) during a 10-min period was calculated as follows:

$$D_{ave} = \begin{cases} D'_{ave} & C_y > 0, C_x > 0 \\ D'_{ave} + 180° & C_y > 0, C_x < 0 \\ D'_{ave} + 180° & C_y < 0, C_x < 0 \\ D'_{ave} + 360° & C_y < 0, C_x > 0 \end{cases} \quad (1)$$

$$D'_{ave} = \arctan \frac{C_y}{C_x} \quad (2)$$

where $C_y = \sum_{i=1}^{10} S_i \sin D_i$ is the sum of y-components, S_i is the wind speed, D_i is the wind direction; and $C_x = \sum_{i=1}^{10} S_i \cos D_i$ is the sum of x-components.

3 DNI forecast based on multiple data types and time-series

A new framework was developed for predicting the inter-hour DNI by combining GBC images and multiple numerical time-series. Figure 2 shows the structure of the forecast framework. To make full use of all inputs, the GBC images should be pre-processed to obtain more realistic sky conditions, and the inputs of the forecast model should be non-dimensionalized to avoid some inputs with a large magnitude overwhelming the others.

3.1 GBC image processing

As shown in Fig. 3, the valid sky of a GBC image was first segmented into mathematical coordinates. The location of the sun (x_s, y_s) in the GBC image could then be calculated as follows:

$$r_s = r_e \cos z \quad (3)$$

$$x_s = x_0 + r_s \sin \alpha \quad (4)$$

$$y_s = y_0 + r_s \cos \alpha \quad (5)$$

where (x_0, y_0) is the center of the GBC image in Fig. 3a; r_s is the amplitude of the sun in a GBC image; r_e is the amplitude of the sun at the 90° solar zenith; z is the solar zenith angle; and α is the solar azimuth angle. The sun was then recognized as a red point, as shown in Fig. 3d, and the shadow band was then recognized as a rectangle (white region) whose symmetry axis uses the line determined based on the center of the GBC image and the sun point. Finally, the support arm is fixed in a GBC image such that it could be artificially designated (green region).

Because of strong sunlight, the pixels around the sun are whiter and brighter in a clear sky and it is easy to detect these pixels as clouds without other specific information. Hence, the clear-sky index (k) is used to determine whether the sun is obscured by clouds, and is defined as follows:

$$k = \frac{I_m}{I_{clr}} \quad (6)$$

where I_m is the measured DNI value; and I_{clr} is the clear-sky DNI value calculated using the clear-sky model developed in [26]. The clear-sky DNI indicates the theoretical maximum DNI under cloud-free conditions, and the clear-sky index (k) describes the attenuation of the DNI caused by clouds. Therefore, if k is bigger than the threshold, e.g. $T_k = 0.9$, the sun is viewed as unobstructed, and the white pixels around the sun are possible error pixels caused by an overexposure.

After determining the error pixels, distinguishing them from the surrounding clouds is the greatest difficulty. Various cloud detection algorithms have been proposed, each of which has shown a satisfactory performance under a specific situation. Among them, the clear-sky background different (CSBD) algorithm outperforms other methods in a circumsolar area when the sun is not obstructed by clouds, but it fails to detect cloud which blocked the sun or is optically thick and dark [30]. The red-blue ratio (R/B) algorithm achieves better results when the clouds are optically thick and dark, but it fails to distinguish clouds from clear sky when cirrus or stratus clouds are present [31]. Combining the advantages of the CSBD and R/B algorithms in cloud detection, the error pixels in a GBC image (I_t) are marked using the following steps.

Step 1: search a clear-sky GBC image I_c based on the closest solar zenith from the constructed CSID introduced in Section 2.1.

Step 2: rotate the image I_c to make its azimuth angle the same as the target image I_t.

Step 3: calculate the green-channel histograms of the two images (I_c and I_t).

Step 4: adjust the green channel of the clear-sky image (I_c) by multiplying a ratio to make its green distribution the same as that of I_t.
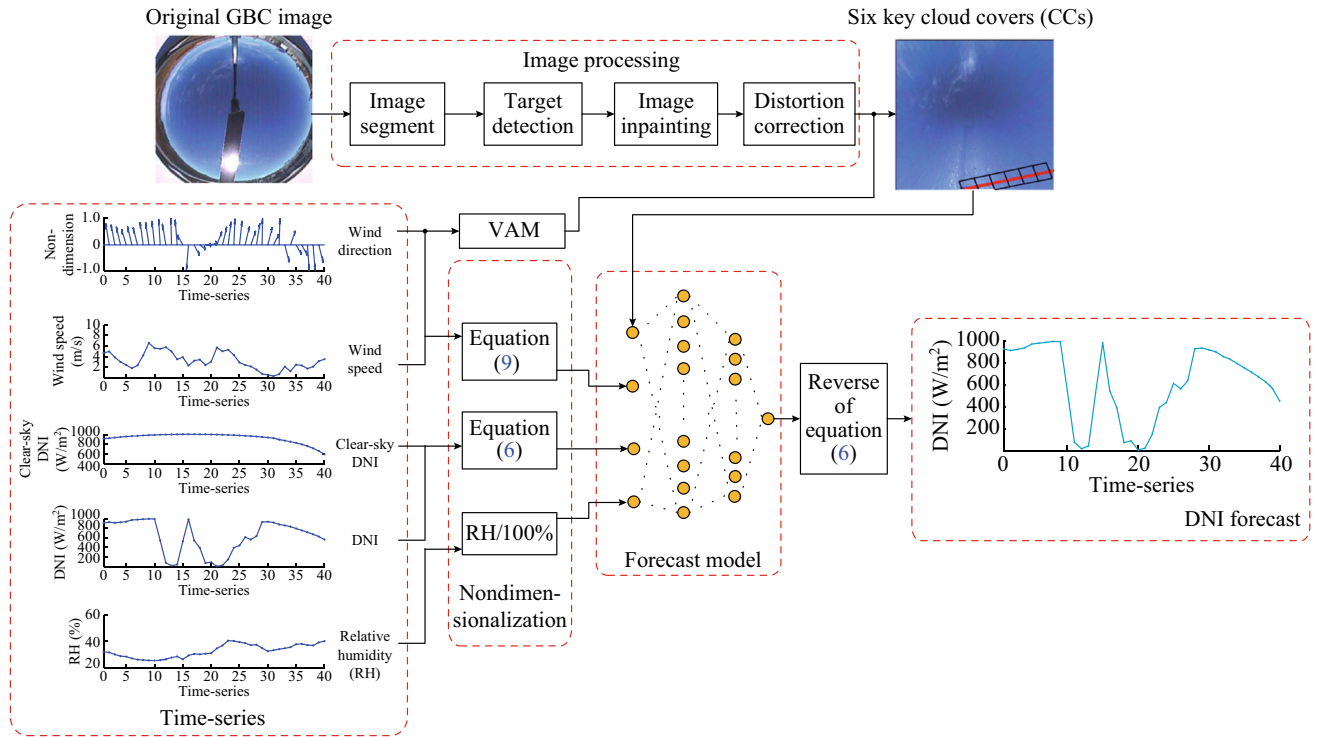
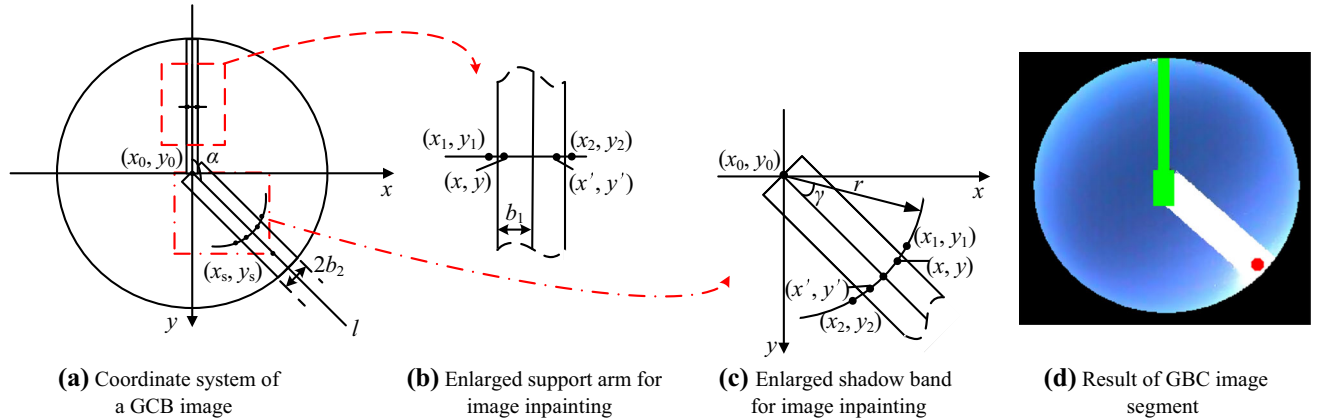**Fig. 2** Structure chart of DNI forecast framework based on multiple time-series and GBC images

(a) Coordinate system of a GCB image

(b) Enlarged support arm for image inpainting

(c) Enlarged shadow band for image inpainting

(d) Result of GBC image segment

Fig. 3 GBC image segment

Step 5: mark the pixel whose difference between the two green channels is greater than threshold T_g as a cloud (M_1).

Step 6: mark the pixel as a cloud (M_2) using the R/B algorithm.

Step 7: remove the error pixels ($M_2 \oplus M_1$) around the position of the sun.

The mirror gradient algorithm [32] was used to fill in the occlusions in a GBC image by combining the symmetry of the inpainted region and the texture features of the cloud. Because the width of a shadow band is too large to ignore the distortion, the polar coordinates were used to interpolate values from the pixels at the same image zenith angle of the target pixel [33], which differs from the inpainting algorithm in [32]. As Fig. 3c shows, the point (x, y) is viewed as the target to be inpainted, and its symmetrical point (x', y') in an arc with radius r is over the line of point (x_0, y_0) and the sun; the symmetric point (x_1, y_1) is then obtained with respect to the right border, and the point (x_2, y_2) symmetric to point (x', y') is obtained with respect to the left border. The polar angle (γ) is changed from 0° to 120° to search the inpainted pixels and their symmetric pixels, and the range of radius (r) is changed from 0 to r_e. Hence, the mirror gradient algorithm was modified as follows:

$$P(x,y) = \lambda_1(P(x_1,y_1) + \Delta P(x_1,y_1)) \\ + (1 - \lambda_1)(P(x_2,y_2) - \Delta P(x_2,y_2)) \tag{7}$$

$$P(x',y') = \lambda_1(P(x_2,y_2) - \Delta P(x_2,y_2)) \\ + (1 - \lambda_1)(P(x_1,y_1) + \Delta P(x_1,y_1)) \tag{8}$$

where λ_1 is the weight and ΔP is the gradient of the pixel P along the arc determined by the four pixels. For the support arm shown in Fig. 3b, a horizontal line was used instead of an arc because its width is narrow. If the point (x, y) is the target to be inpainted, then we obtain its symmetric point (x_1, y_1) with respect to the left border, point (x', y') with respect to the y-axis, and point (x_2, y_2) symmetric to point (x', y') with respect to the right border. The horizontal range was from $x = x_0 - b_1$ to $x = x_0$, and the vertical range was from $y = y_0$ to $y = y_0 - r_e$.

Finally, the distortion of the inpainted GBC image was corrected using the spherical orthogonal distortion correction method [34], and the pre-processing of the GBC image was completed.

3.2 Input nondimensionalization

The key point in predicting the inter-hour DNI is to predict the cloud motion, including the speed and direction. Because GBC images were sampled every 10 min, it is difficult to accurately identify the direction and speed of the cloud from two consecutive images. Considering the relationship between the cloud motion and wind, the wind speed and direction were used to estimate the cloud motion.

First, 30 groups of two consecutive GBC images were selected from January 2013, in which the cloud motion was very clear by the human eye. Then, the R/B algorithm was used to detect a cloud from the GBC image, and image enhancement and morphological methods were then used to find all cloud masses and their centroids. Finally, the cloud motion was calculated from two consecutive images comparing their centroids. It was found that a linear regression can be used to describe the relationship between the wind and cloud directions, as shown in Fig. 4 where the red line has a slope of 1, the blue points are the values of cloud direction calculated from the images and the corresponding wind direction, and their correlation coefficient is up to 0.74. Thus, the wind direction was defined as the cloud direction.

However, the cloud speed does not have a clear relationship with the wind speed owing to the difficulty of cloud boundary detection and the lack of information regarding the cloud base height. Therefore, six cloud cover features were extracted from 20×20 pixel areas along the opposite direction of the wind (cloud), as shown in Fig. 2, to address the cloud effect on the DNI. The wind speed was
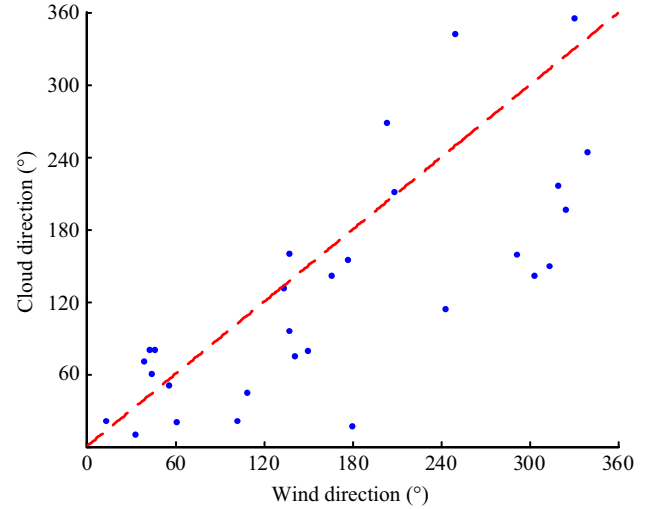

Fig. 4 Relationship between wind and cloud directions

set as one of the model inputs to adjust the weight of the cloud cover in the forecast model. Therefore, the GBC image was transformed to six cloud covers in 20×20 pixel areas dimensionlessly.

According to (6), the measured DNI can be transformed into clear-sky index k, thereby removing the influence of the sun's position on the DNI. Here, k is a dimensionless variable with a scale of [0, 1]. The relative humidity is the ratio of the partial pressure of water vapor to the equilibrium vapor pressure of water at a given temperature, and thus the relative humidity was selected as one input variable instead of the water vapor, and ranges from 0 to 1 instead of a percentage. To allow the wind speed (S) to be within the same scale, it was transformed as follows:

$$S' = \frac{S - S_{\min}}{S_{\max} - S_{\min}} \tag{9}$$

where $S_{\min}$ and $S_{\max}$ are the minimum and maximum wind speeds in the training set, respectively. Thus far, all inputs including the image and numerical time-series were transformed into dimensionless variables within the range of [0, 1].

3.3 DNI inter-hour forecast model

In view of the advantage of handling non-linear problems, SVR was applied as the DNI forecast model. The structure of the SVR model was introduced in detail in [35]. In this study, the kernel function of the SVR is the radial basis function. There are three model parameters: cost C, the tolerance of termination criterion ε and attenuation parameter v. The three model parameters are adjusted using the cross-validation method [36] with the training and validation sets. The extensive grid search method was used to determine the three hyperparameters,

where C was set in [1, 5] at a step size of 0.5, ε was set in [0.001, 0.1] at a step size of 0.001, and v was set in [1, 3] at a step size of 0.2, respectively. The range of these hyperparameters was set based on the experimental results and previous experience.

To determine the input order of the clear-sky index, the Bayesian information criterion (BIC) was used and the results are as shown in Fig. 5 where p is the number of inputs. Thus, it is reasonable to select $k(t - 3\Delta t)$, $k(t - 2\Delta t)$, $k(t - \Delta t)$ and $k(t)$ as the forecast model inputs, where Δt is the forecast interval of 10 min. The same dimensions of relative humidity were selected during the same period. For the wind speed and GBC image, only $S(t)$ and the cloud cover were selected concurrently. Finally, there were 15 inputs of the forecast model. The output of the SVR model was the next 10-min clear-sky index ($k(t+\Delta t)$), and the clear-sky index was then transformed into the DNI according to (6).

4　Results and discussion

80% of coupled data randomly selected from 2013 were set as the training set, and the remaining data (20%) were set as the validation set. Data from 2014 were set as the testing set. The accuracy of the forecast was assessed using three statistics, namely, the normalized mean bias error (nMBE), normalized mean absolute error (nMAE), and normalized root mean square error (nRMSE), which are defined as follows:

$$nMBE = \frac{1}{\bar{I}_m N}\left[\sum_{i=1}^{N}(I_{fi} - I_{mi})\right] \times 100\% \tag{10}$$

$$nMAE = \frac{1}{\bar{I}_m N}\left[\sum_{i=1}^{N}|I_{fi} - I_{mi}|\right] \times 100\% \tag{11}$$

$$nRMSE = \frac{1}{\bar{I}_m}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_{fi} - I_{mi})^2} \times 100\% \tag{12}$$

where N is the number of samples; I_{fi} is the predicted value of the forecast model; I_{mi} is the measured value; and $\bar{I}_m$ is the average of all measured DNIs.

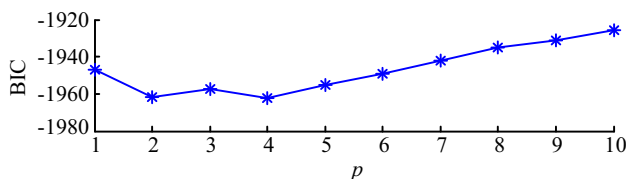To evaluate the performance of the proposed forecast framework at predicting the DNI in the next 10 min, two

Fig. 5 Results of BIC using clear-sky index in training set

Table 1 Performance of different models using only numerical DNI time-series as inputs for 10-min DNI forecast

Model	nMBE (%)	nMAE (%)	nRMSE (%)
Persistent	0.58	15.74	31.10
AR	0.39	17.29	30.12
ANN	0.48	17.48	29.23
SVR	0.02	16.47	29.04

groups of experiments were carried out: 1) using only the historical DNI time-series as inputs; 2) using both the numerical time-series and ground-based cloud images. Table 1 lists the four forecast models using only DNI time-series as inputs, including the persistent model (measured DNI as the presently predicted value), autoregressive (AR) model (with historical measured DNIs linearly fit as the predicted DNI), and artificial neural network (ANN) model (multilayer perceptron, with historical measured DNIs non-linearly fitted as the predicted DNI), which are introduced in Appendix A. Compared to the linear models (persistent and AR), the non-linear models (ANN and SVR) achieved better results in predicting the inter-hour DNI, and the SVR model performed slightly better than the ANN model, especially on nMAE.

Table 2 lists the test results of different forecast models considering GBC images in predicting the DNI, where the AR-I, ANN-I and SVR-I models maintained the same structure as the AR, ANN and SVR models in Table 1, while adding six cloud covers (CCs) extracted from a GBC image as inputs. The ANN-I model was introduced in [19], and the forecast model based on digital image processing (FM-DIP) was developed in [15]. The forecast skill (Fs) is a criterion used to assess the performance of a forecast model over the persistent model, which is defined as follows:

$$Fs = \frac{nRMSE_{per} - nRMSE_f}{nRMSE_{per}} \times 100\% \tag{13}$$

where $nRMSE_{per}$ is the nRMSE of the persistent model; and $nRMSE_f$ is the nRMSE of the forecast model. Figure 6 shows the forecast skills of all models listed in Tables 1 and 2. Among these forecast models, the non-linear models

Table 2 Performance of different models with different inputs for 10-min DNI forecast

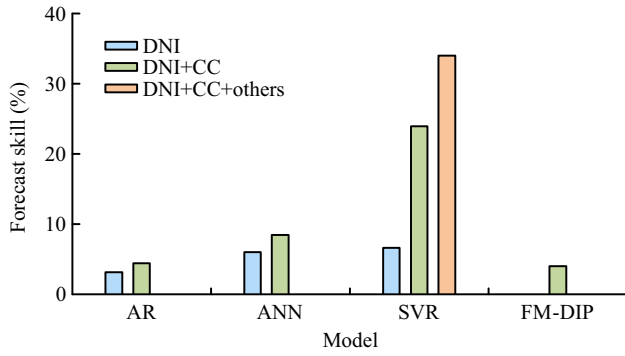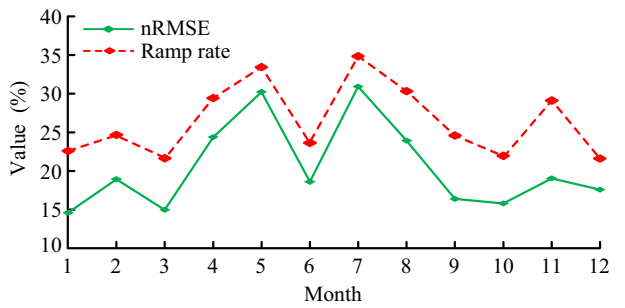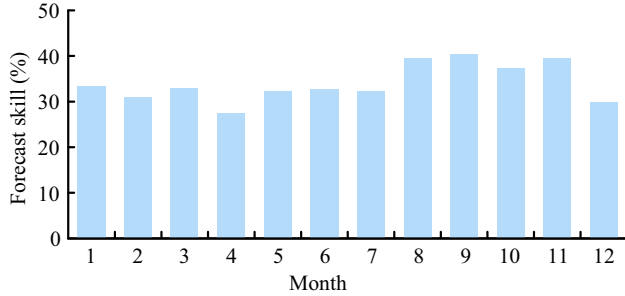| Model	nMBE (%)	nMAE (%)	nRMSE (%)
AR-I	− 0.04	17.85	29.73
ANN-I	0.43	18.04	28.47
SVR-I	0.98	16.03	23.66
FM-DIP	− 3.78	20.01	29.85
Proposed	0.41	14.13	20.53

Fig. 6 Forecast skills of different models compared to persistent model

(a) nRMSE of proposed method and ramp rate of predicted DNI for each month

(b) Monthly improvement of proposed method on nRMSE over persistent model

Fig. 7 Monthly distribution of the ramp rate of DNI time-series and nRMSE of proposed method for testing set (2014)

performed better than the linear ones; in terms of input, the model with images used as inputs achieved better results with the same structure. For the three SVR models with different inputs, the proposed model achieved an optimal performance, with 34% improvement over the persistent model in terms of the nRMSE, which is mainly owing to the supplementary consideration of the water vapor effects, in addition to the historical DNI and cloud. It is therefore reasonable that the consideration of various atmospheric variables (including clouds and water vapor) in the DNI forecast model can improve the accuracy.

To investigate the performance and fluctuations of the proposed method, the ramp rate (RR) was used to evaluate the fluctuations of the DNI time-series, and is defined as follows:

$$RR = \frac{N_{ramp}}{N_{total}} \times 100\% \tag{14}$$

$$N_{ramp} = \sum_{t=1}^{N_{total}} \text{sgn}\left(\frac{|I_m(t) - I_m(t-1)|}{\max(I_m(t), I_m(t-1))} > 0.5\right) \tag{15}$$

where N_{total} is the number of total testing sample couples and N_{ramp} is the number of fluctuations defined as the bias between two consecutive DNIs greater than half of the larger DNI. In addition, Fig. 7a shows the nRMSE of the proposed method and the ramp rate of the predicted DNI for each month. The nRMSE agrees well with the ramp rate of the predicted DNI time-series, with a constantly increasing or decreasing trend. The nRMSE reached up to 31% in July, whereas the ramp rate was 35%. Throughout the year of the testing data (2014), the proposed method reached a steady improvement in the nRMSE of 34% on average over the persistent model, as shown in Fig. 7b.

5 Conclusion

A framework for predicting the inter-hour DNI was proposed, including the nondimensionalization of the input variables, employment of a forecast model based on the SVR, and transformation of outputs. Multiple input variables, involving the historical measured DNI, relative humidity, ground-based cloud image, and wind speed, were employed in the developed forecast model. The data from the NREL dataset for the entire year of 2013 were used to train the forecast model using a cross-validation method. The experimental results demonstrated that the proposed model was comparable to other models in terms of the forecast accuracy, and achieved a 34% improvement in the nRMSE over the persistent model.

The performance of the proposed forecast model is affected by the ramp rate of the DNI time-series, with a higher accuracy when the monthly DNI changes more gently, and vice versa. Therefore, considering the ramp rate and thereby constructing different forecast models will further improve the forecast accuracy, and should be investigated in the future studies.

Appendix A

For the persistent model, the predicted value I_f is expressed as follows:

$$I_f(t + \Delta t) = I_m(t) \tag{A1}$$

where Δt is the forecast step (set to 10 min in this study).

For the AR model, the target I_f is expressed as follows:

$$I_f = a_0 + a_1 x_1 + \cdots + a_p x_p \tag{A2}$$

where a_i ($i = 0, 1, \ldots, p$) are the model parameters; x_i ($i = 0, 1, \ldots, p$) are the inputs (historical DNIs); and p is the number of inputs ($p = 4$).

For the ANN model, the target I_f is expressed as follows:

$$g_j = \sum_{i=1}^{n_1} w_{ij} x_i + b_j \tag{A3}$$

$$u_j = \frac{2}{1 + e^{-2g_j}} - 1 \tag{A4}$$

$$I_f = \sum_{j=1}^{n_2} w_j u_j + c \tag{A5}$$

where w_{ij}, b_j, w_j and c are the model parameters, and are determined using a learning algorithm with the training set; and n_1 and n_2 are the number of input nodes and hidden nodes in an ANN model, respectively.

References

[1] Senatla M, Bansal R (2018) Review of planning methodologies used for determination of optimal generation capacity mix: the cases of high shares of PV and wind. IET Renew Power Gener 12(11):1222–1233

[2] Lamsal D, Sreeram V, Mishra Y et al (2018) Kalman filter approach for dispatching and attenuating the power fluctuation of wind and photovoltaic power generating systems. IET Gener Transm Distrib 12(7):1501–1508

[3] Prada J, Dorronsoro JR (2018) General noise support vector regression with non-constant uncertainty intervals for solar radiation prediction. J Mod Power Syst Clean Energy 6(10):1–13

[4] Yang D, Quan H, Disfani VR et al (2017) Reconciling solar forecasts: geographical hierarchy. Sol Energy 146:276–286

[5] Yang D, Quan H, Disfani VR et al (2017) Reconciling solar forecasts: temporal hierarchy. Sol Energy 146:332–346

[6] Law EW, Prasad AA, Kay M et al (2014) Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting – a review. Sol Energy 108:287–307

[7] Hong T, Pinson P, Fan S et al (2016) Probabilistic energy forecasting: global energy forecasting competition 2014 and beyond. Int J Forecast 32(3):896–913

[8] Durrani SP, Balluff S, Wurzer L et al (2018) Photovoltaic yield prediction using an irradiance forecast model based on multiple neural networks. J Mod Power Syst Clean Energy 6(2):255–267

[9] Gafurov T, Usaola J, Prodanovic M (2014) Modelling of concentrating solar power plant for power system reliability studies. IET Renew Power Gener 9(2):120–130

[10] Engerer NA, Mills FP (2015) Validating nine clear sky radiation models in Australia. Sol Energy 120:9–24

[11] Mathiesen P, Collier C, Kleissl J (2013) A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting. Sol Energy 92(4):47–61

[12] Lima FJL, Martins FR, Pereira EB et al (2016) Forecast for surface solar irradiance at the Brazilian northeastern region using NWP model and artificial neural networks. Renew Energy 87:807–818

[13] Ozoegwu CG (2018) New temperature-based models for reliable prediction of monthly mean daily global solar radiation. J Renew Sustain Energy 10(2):023706

[14] Fu CL, Cheng HY (2013) Predicting solar irradiance with all-sky image features via regression. Sol Energy 97(5):537–550

[15] Alonso-Montesinos J, Batlles FJ, Portillo C (2015) Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images. Energy Convers Manag 105:1166–1177

[16] Kurtz B, Kleissl J (2017) Measuring diffuse, direct, and global irradiance using a sky imager. Sol Energy 114:268–277

[17] Cheng HY, Yu CC (2015) Multi-model solar irradiance prediction based on automatic cloud classification. Energy 91:579–587

[18] Cheng HY (2017) Cloud tracking using clusters of feature points for accurate solar irradiance nowcasting. Renew Energy 104:281–289

[19] Chu Y, Pedro HTC, Coimbra CFM (2013) Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning. Sol Energy 98(4):592–603

[20] Chu Y, Li M, Pedro HTC et al (2015) Real-time prediction intervals for intra-hour DNI forecasts. Renew Energy 83:234–244

[21] Bernecker D, Riess C, Angelopoulou E et al (2014) Continuous short-term irradiance forecasts using sky images. Sol Energy 110:303–315

[22] Cheng HY, Yu CC (2016) Solar irradiance now-casting with ramp-down event prediction via enhanced cloud detection and tracking. In: Proceedings of IEEE international conference on multimedia and expo (ICME), Seattle, USA, 11–15 July 2016, pp 1–6

[23] Marquez R, Coimbra CF (2013) Intra-hour DNI forecasting based on cloud tracking image analysis. Sol Energy 91:327–336

[24] Quesada-Ruiz S, Chu Y, Tovar-Pescador J et al (2014) Cloud-tracking methodology for intra-hour DNI forecasting. Sol Energy 102(4):267–275

[25] Chu Y, Li M, Coimbra CF (2016) Sun-tracking imaging system for intra-hour DNI forecasts. Renew Energy 96:792–799

[26] Zhu T, Wei H, Zhao X et al (2017) Clear-sky model for wavelet forecast of direct normal irradiance. Renew Energy 104:1–8

[27] Bone V, Pidgeon J, Kearney M et al (2018) Intra-hour direct normal irradiance forecasting through adaptive clear-sky modelling and cloud tracking. Sol Energy 159:852–867

[28] Stoffel T, Andreas A (1981) NREL Solar Radiation Research Laboratory (SRRL): Baseline Measurement System (BMS);

Golden, Colorado (Data) (No. NREL/DA-5500-56488), United States

[29] Han S, Liu Y, Yang W et al (2012) An optimized vector average method for calculation of average wind direction. Power Syst Technol 36(5):68–72

[30] Yang J, Min Q, Lu W et al (2016) A total sky cloud detection method using real clear sky background. Atmos Meas Tech 8(12):13073–13098

[31] Long CN, Ackerman TP, Gaustad KL et al (2006) Estimation of fractional sky cover from broadband shortwave radiometer measurements. J Geophys Res Atmos 111(D11204):1–11

[32] Zhu X, Zhou H, Zhu T et al (2018) Pre-processing of ground-based cloud images in photovoltaic system. Autom Electr Power Syst 42(6):104–151

[33] Ghonima MS, Urquhart B, Chow CW et al (2012) A method for cloud detection and opacity classification based on ground based sky imagery. Atmos Meas Tech 5(11):2881–2892

[34] Zhu H (2006) Correction of radial distortion image based on straight line characteristic. J Shanghai Univ Eng Sci 20(2):153–155

[35] Suykens JAK, Brabanter JD, Lukas L et al (2002) Weighted least squares support vector machines: robustness and sparse approximation. Neurocomputing 48(1):85–105

[36] Lendasse A, Wertz V, Verleysen M (2003) Model selection with cross-validations and bootstraps—application to time series prediction with RBFN models. Artificial neural networks and neural information processing. Springer, Berlin, pp 573–580

Tingting ZHU is currently working toward the Ph.D. degree in pattern recognition and artificial intelligence at School of Automation, in Southeast University, Nanjing, China. She achieved a fellowship jointly awarded by the Fonds de Recherche du Québec—Nature et Technologies (FRQNT) and the China Scholarship Council and studied as a visiting student at department of Atmospheric and Oceanic Sciences, in McGill University, Canada, from 2017 to 2018. Her current research interests include solar power generation forecast, machine learning and climate feedback.

Hai ZHOU is a master, senior engineer and the director of Electric Meteorological Simulation and Applied Research Department, New Energy Research Center, China Electric Power Research Institute, and a backbone member of the scientific and technological research team on new energy power generation, power forecast and optimal dispatching in State Grid Corporation of China, mainly engaged in warning of meteorological disasters, monitoring and evaluation of meteorological resources in power grids. His research interests include numerical weather prediction, wind power, and photovoltaic power forecast.

Haikun WEI received the B.S. degree from the Department of Automation, North China University of Technology, China, in 1994, and the M.S. and Ph.D. degrees from the Research Institute of Automation, Southeast University, China, in 1997 and 2000, respectively. He was a Visiting Scholar with the RIKEN Brain Science Institute, Japan, from 2005 to 2007. He is currently a Professor with the School of Automation, Southeast University. His research interests include machine learning and renewable energy application.

Xin ZHAO received the B.S. degree and the M.S. degree from Jiangnan University, Wuxi, China, in 2011 and 2014. She is currently working toward the Ph.D. degree in control science and engineering in the Southeast University, Nanjing, China. Her current research interests include machine learning and renewable energy power generation forecast.

Kanjian ZHANG received the B.S. degree in mathematics from Nankai University, China, in 1994, and the M.S. and Ph.D. degrees in control theory and control engineering from Southeast University, China, in 1997 and 2000, respectively. He is currently a Professor with the School of Automation, Southeast University. His research interests include nonlinear control theory and its applications, especially robust output feedback design and optimization control.

Jinxia ZHANG received the B.S. and Ph.D. degrees from the Department of Computer Science and Engineering, Nanjing University of Science and Technology, China, in 2009 and 2015, respectively. She was a Visiting Scholar with the Visual Attention Lab, Brigham and Women's Hospital, and also with the Harvard Medical School from 2012 to 2014 in the USA. She is currently a Lecturer with the School of Automation, Southeast University. Her research interests include visual attention, visual saliency detection, computer vision and machine learning.