# *A Satellite Image Data based Ultra-short-term Solar PV Power Forecasting Method Considering Cloud Information from Neighboring Plant*

Fei Wang[a, b], Xiaoxing Lu[a], Shengwei Mei[c], Ying Su[d], Zhao Zhen[a, c], Zubing Zou[d], Xuemin Zhang[c], Rui Yin[e], Neven Duić[f], Miadreza Shafie-khah[g], João P. S. Catalão[h]

[a] Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China

[b] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China

[c] State Key Lab of Power System, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

[d] Institute of Science and Technology, China Three Gorges Corporation, Beijing 100038, China

[e] Dispatch and Control Center, State Grid Hebei Electric Power Co., Ltd, Shijiazhuang 050022, China

[f] Department of Energy, Power and Environmental Engineering, Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Ivana Lucica 5, HR-10000 Zagreb, Croatia

[g] School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland

[h] Faculty of Engineering of University of Porto and INESC TEC, 4200-465 Porto, Portugal

*Corresponding author at: Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China

E-mail address: georgiazhz@foxmail.com (Z. Zhen)

**Abstract-Accurate ultra-short-term PV power forecasting is essential for the power system with a high proportion of renewable energy integration, which can provide power fluctuation information hours ahead and help to mitigate the interference of the random PV power output. Most of the PV power forecasting methods mainly focus on employing local ground-based observation data, ignoring the spatial and temporal distribution and correlation characteristics of solar energy and meteorological impact factors. Therefore, a novel ultra-short-term PV power forecasting method based on the satellite image data is proposed in this paper, which combines the spatio-temporal correlation between multiple plants with power and cloud information. The associated neighboring plant is first selected by spatial-temporal cross-correlation analysis. Then the global distribution information of the cloud is extracted from satellite images as additional inputs with other general meteorological and power inputs to train the forecasting model. The proposed method is compared with several benchmark methods without considering the information of neighboring plants. Results show that the proposed method outperforms the benchmark methods and achieves a higher accuracy at 4.73%, 10.54%, and 4.88%, 11.04% for two target PV plants on a four-month validation dataset, in terms of root mean squared error and mean absolute error value, respectively.**

***Keywords—Ultra-short-term; PV power forecasting; Spatio-temporal; Satellite image***

## NOMENCLATURES

***Abbreviations***

| | | | |
|---|---|---|---|
| ANN | Artificial neural networks | IDFT | Inverse discrete Fourier transform |
| AR | Auto-regressive | IPSI | Image phase shift invariance |
| ARMA | Auto-regressive moving average model | MA | Moving average |
| BPNN | Back-propagation neural network | MAE | Mean absolute error |
| CART | Classification and regression trees | NSMC | National satellite meteorological center |
| CPS | The cross-power spectrum | PV | Photovoltaic |
| CMD | Cloud motion displacement | RBF | Radial basis function |
| CQ | Cloud quality | RMSE | Root mean squared error |
| DFT | Discrete Fourier transform | SCCF | Sample cross-correlation function |
| DNI | Direct normal irradiance | SVM | Support vector machines |
| ESSS | Exponential smoothing state space | SVR | Support vector regression |
| FPCT | Fourier phase correlation theory | UCSD | University of California, San Diego |
| GBDT | Gradient boosting decision trees | | |

***Nomenclatures***

***Nomenclatures for 2.1 and 2.2***

| | | | |
|---|---|---|---|
| $C_{xy}(\tau_i)$ | Correlation degree | $P$ | Pre-processed data of PV power |
| $C_{xx}(0)$ | The value of correlation degree when *x*=*y* and $\tau_i = 0$ | $P_{clear}$ | PV power in clear sky condition |
| $C_{yy}(0)$ | The value of correlation degree when *y*=*x* and $\tau_i = 0$ | $r$ | Radius of PV plant |
| $G_{CQ}$ | Cloud quality index | $r_{xy}$ | The correlation degree between two PV power output time series |
| $M$ | The number of all sky pixels in the circular | $X_t$ , $Y_t$ | Two PV power output time series |

| | | | |
|---|---|---|---|
| | region | | |
| $N$ | Number of calculated pixels in a circular region | $X_M, Y_M$ | The mean values of $X_t$ and $Y_t$ |
| $n$ | The length of $X_t$ and $Y_t$ | $\tau$ | Time lag |
| $p$ | Real data of PV power | $\varepsilon$ | Threshold of SCCF |
| ***Nomenclatures for 2.3.1 SVM*** | | | |
| $C$ | The tradeoff between empirical risk and model complexity | $T(x,y)$ | A series of given samples |
| $d$ | Degree of the polynomial kernel | $\lVert w \rVert^2$ | The describing function |
| $f$ | The complexity term | $x, y$ | The input and output of SVM |
| $F$ | The regression function | $\xi$ | Slack variable |
| $K(x_i, x_j)$ | Kernel functions | $\gamma$ | The width of the RBF kernel |
| $R_{reg}$ | The risk function in SVM | $\phi(x_i)$ | The inner products of SVM |
| ***Nomenclatures for 2.3.2 GBDT*** | | | |
| $F_0(x)$ | The loss function | $x$ | The vectors of features |
| $F_M(x)$ | Final model after *M* times of iteration | $y$ | The target of GBDT |
| $h_m(x)$ | Decision tree | $\gamma$ | A constant |
| $J$ | The number of leaf nodes of the decision tree | $\gamma_m$ | Multiplier in GBDT |
| $L$ | Differentiable loss function | $r_m$ | Pseudo-residuals |
| $M$ | Times of iteration | $\upsilon$ | Learning rate/ Shrinkage factor |
| $(x_i, y_i)$ | The training set of GBDT | | |
| ***Nomenclatures for 2.3.3 ARMA*** | | | |
| $p$ | The order of the AR process | $\beta_j$ | The coefficient of MA |
| $q$ | The order of MA error term | $e(t)$ | The coefficient of white noise |
| $\alpha_i$ | The order of AR coefficient | $S(t)$ | Forecasted value at time *t* |
| ***Nomenclatures for 2.4 and 2.5*** | | | |
| $A, B$ | The dimensions of a grayscale matrix | $G(x_i, y_i)$ | Gray value of $(x_i, y_i)$ in satellite image |
| $C, D$ | The dimensions of the convolution kernel matrix | $G(u,v)$ | The convolution form |
| $b_1$ | The mean value of the fitting curve | $h(x,y)$ | Convolution kernel matrix |
| $C(u,v)$ | The cross-power spectrum | $I(x,y)$ | Discrete convolution transform |
| $D_1$ | Cloud motion direction | $M, N$ | The grayscale matrix resolution of an image |
| $D_2$ | Physical distance direction of 2 plants | $P_T(t)$ | Power output of target plant at forecasting time |
| $f(x,y)$ | Satellite image | $u, v$ | The pixel's Frequency domain coordinates |
| $\lvert F(u,v) \rvert$ | The amplitude of $F(u,v)$ | $x, y$ | Cartesian coordinates of a pixel |
| $F^*(u,v)$ | The complex conjugate | $(x_0, y_0)$ | A displacement vector |
| $F(u,v)$ | The transformed form of $f(x,y)$ processed with DFT | $\tau'$ | Time lag between the target and neighboring plant pair |

# I. INTRODUCTION

## *1.1 Background and literature review*

In recent years, renewable energy application is becoming a rapidly evolving field to mitigate negative environmental issues, such as pollution, climate change, etc. for the sake of fossil energy utilization [1–3]. As one of the most promising renewable sources, solar energy can be converted into electricity by the means of photovoltaic (PV) power generation. However, due to the variability and uncertainty of ground-level irradiance, various technical challenges are posed for grid scheduling, dispatching, and regulation, thus increasing the cost to balance power generation and demand in real-time [4,5]. To reduce the adverse effects from grid-connected PV power plants, it is essential to apply several effective measures, including power flow optimization [6,7], demand response [8,9], backup generators, battery reserves [10], peaking units, multi-time scale PV power forecasting [11,12], etc. However, limitations are still existing in all these solutions. The power ascend/decline rate of backup generators is restricted by the unit ramp rate, which may result in difficulties to meet the incremental power generation need [13]. As for battery reserves, massive-scale energy storage is still difficult to realize for the sake of production costs and storage capacity restrictions [14]. Concerning the lack of information on the electricity consumption behavior of residential users, it's also hard to achieve demand response technologies [15]. As a low-cost strategy, high-fidelity PV power forecasting is extensively applied to mitigate the intermittency of PV power generation. Meanwhile, it can also provide effective support for other solutions [16].

Solar power forecasts with different time scales have been developed to meet various demands for the power industry. As two of the most popular forecasting fields in the last few decades, short-term PV power forecasting is widely utilized in the formulation of day-ahead generation plans [17], while ultra-short-term PV power forecasting is capable of offering guidance to real-time dispatching of the grid [18,19]. For ultra-short-term PV power forecasting, cloud cover is the main factor that affects the amount of irradiance reaching the ground surface, thus resulting in a power volatility effect. To be aware of the effect of cloud clusters in the atmosphere, local-sensing and remote-sensing devices are applied to track clouds, offering support to PV power forecasts [20,21]. For the first type, the ground-based sky imaging system is applied in the local-sensing-based research field, which can capture local clouds over the installed plant within sub-kilometers in real-time. Consequently, sky images are widely used for cloud characteristic extraction in minute timescale irradiance/PV forecasting, and in most cases focus on a single PV power plant due to the limits of its observation range. To improve the performance of irradiance/PV power forecasting, tracking cloud motion with high accuracy and robustness is a ground work to be done. In [22], a cloud motion vector calculation method-based image phase shift invariance (IPSI) was proposed to reduce outliers which generated by Fourier phase correlation theory (FPCT) method. Based on the work mentioned in [23], another two transforms such as wavelet transform and convolution transform were added to further decrease error rate of calculated cloud motion vector. After accomplishing cloud motion vector calculation, irradiance/PV power forecasts can be achieved. In [24], with the application of sky images, the accuracy of direct normal irradiance (DNI) forecasts are evaluated by various means of cloud transmittance and velocity calculation. In [25], digital image levels are transformed into irradiances, then maximum cross-correlation calculation is applied to achieve future predictions. Forecasting beam irradiance, diffuse irradiance, and global irradiance are evaluated and tested by using different statistical parameters, which shows a better performance of this proposed method. For the latter one, remote-sensing technique-based satellite images usually provide forecasts ranging from half an hour to six hours, and offers a wider view of observation. Therefore, it is conducive to achieve forecasts in the environment of growing number of PV power plant clusters [26]. As depicted in [27], hourly solar irradiance time series was able to be predicted by using satellite image analysis and a hybrid exponential smoothing state space (ESSS) model with artificial neural networks (ANN). The effectiveness of proposed method is shown better than those of other traditional forecasting models. In [28], with the application of physical method, solar radiation was estimated on the basis of the relation between clear-sky index and cloud cover index. The physical method required no ground data and was more suitable for the cases where the distances between PV plants were quite large.

At present, in the field of solar irradiance/PV power forecasting, major studies only focus training samples on historical data of the target forecasting plant. Only a few kinds of research are exploring spatial similarity and temporal correlation to achieve irradiance/PV power forecasting. Usually, when multiple PV plants are located in different regions with the same geographical locale, due to the short distance between every two plants, from a few to a couple of kilometers, it is common to observe that when there is an instantaneous large drop occurring in one PV plant output, after a while, another drop with the same variable shape will take place in another plant. Based on this phenomenon, forecasting on target plant with data from neighboring plants is practical. Several works of literature have investigated spatio-temporal correlation among multiple plants [29–31]. As depicted in [32], a framework is established to quantify spatial similarity and temporal correlation, then PV output forecasting in a certain geographical region has been achieved. In [33], intra-hour cloud locations and irradiance are forecasted for a network of six pyranometer ground plants in a microgrid at the University of California, San Diego (UCSD). Another multi-scale spatio-temporal PV power forecasting model by using autoregressive with exogenous input is established in [34], which delivers better accuracy than conventional temporal-only autoregressive models. As for [35], a multi-time scale forecast of PV generation was introduced based on spatio-temporal correlations among neighboring solar sites. The performance of the proposed method was compared with the conventional persistence model, and the improved forecast quality was studied by using historical data acquired from PV sites located in California and Colorado.

### *1.2 Motivation and contribution*

High-fidelity PV power forecast in ultra-short-term time scale can provide guidance and detailed correction to grid dispatching and scheduling plan [36]. Meanwhile, it is conducive to reduce the charge frequency of the storage battery, thus prolong its service time. Therefore, related researches on PV power forecasting are essential to carry out. To sum up, in the field of ultra-short-term forecasts, satellite image-based methods address the problem of tracking cloud motions in a large region, thus promote forecasting accuracy after taking cloud characteristic input into consideration. For neighboring plant data-based methods, they achieve forecasts in another way by utilizing spatio-temporal information among regional plants. However, in the present work, primary spatio-temporal-based methods mainly focus on historical text data, such as power, irradiance, humidity, wind speed, wind direction, and other meteorological factors to train forecast models. Until now, there is still less or even no relevant literature presenting ultra-short-term PV power forecasting utilizing combining spatio-temporal correlation with satellite image information captured above the neighboring plant.

To fill this forecast gap, in this paper, we propose a method that integrates satellite image, PV power data collected from the neighboring plant, with the target plant's historical data, and other meteorological factors, to achieve PV power forecasting with a time horizon varying from 15 min to 4 h. First, the neighboring plant which has the strongest correlation with the target plant should be determined by calculating the maximum value of the sample cross-correlation function (SCCF). If the value of SCCF is over the threshold $\varepsilon$ and the cloud motion direction is from the neighboring plant to the target plant, then the neighboring plant can be applied, as well as the value of time lag $\tau$ corresponding to the maximum value of SCCF can be utilized with credibility. If not, a traditional method which only considers the data of the target plant is employed to achieve the forecast.

After ensuring the neighboring plant, neighboring cloud characteristic indexes are extracted as additional inputs with other general meteorological and power inputs to train the foresing models by using Support Vector Machines (SVM) and Gradient Boosting Decision Trees (GBDT), to verify the effectiveness of the proposed technical route. However, the method proposed with the spatio-temporal correlation between two plants can only achieve the forecasting within a time horizon as short as $\tau$ min. For other forecasting range between $\tau+15$ min and 4 h, a submethod should be applied. Based on historical cloud motion calculation between consecutive satellite image pairs, cloud motion above the neighboring plant in the future can be obtained by using linear extrapolation, then the regions which will cover the neighboring plant enable to be acquired. After cloud characteristic extraction and other inputs calculation, it is feasible to achieve the output forecasting of the target plant for the remaining time which the method with spatio-temporal correlation can not achieve. The comparisons with various benchmark methods show the performance of the proposed method is better over ultra-short-term PV power forecasts.

The main contributions of this paper include:

(1) Spatio-temporal correlation between plants in a PV power plant cluster is analyzed. Then the selection of an appropriate neighboring plant is introduced, which can provide prior knowledge of the power fluctuation in the target power plant.

(2) Verify the mapping relationship between cloud characteristics of the neighboring plant and solar PV power output of the target plant theoretically based on satellite remote sensing data.

(3) An ultra-short-term solar PV power forecasting method based on power data of neighboring plants and cloud information from satellite images is proposed, which can improve the forecasting accuracy.

(4) Actual data from two target plants are applied to evaluate the effectiveness of the proposed method.

## II. METHODOLOGY

### *2.1 Spatio-temporal correlation between two adjacent PV plants*

In a certain region, due to the similar geographical locale, PV outputs of two adjacent PV plants located at different places may exhibit a similar time-varying pattern, which can be defined as spatial similarity. In order to explore the correlations among several PV power output time series produced at different plants, the temporal correlation which is seemed as the pairwise similarity on spatial dimension needs to be evaluated. In other words, for the sake of cloud motion, there will be a lagging or leading effect between two PV output time series corresponding to two adjacent plants, which is defined as temporal correlation. To facilitate understanding, PV power output data collected from plant A and plant B at the same period of time are selected and presented in Fig.1, with a distance of up to 19.6 kilometers.

Since the variable-size surface area of the PV module has an impact on power output through influencing the amount of reaching irradiance, raw PV power data should be pre-processed by (1):

$$P(i) = p(i) / P_{clear} \tag{1}$$

where $P$ , $p$ , $P_{clear}$ mean pre-processed data, real data, and power in clear sky condition at each moment for data-processed power plant, respectively. For two PV output time series collected from two PV plants shown in Fig.1, we can observe that the two plants have similar time-varying patterns. The highly correlated events enable to facilitate determination of a relationship between the temporal difference in PV power outputs corresponding to the two plants and their geographical distance. As shown in Fig.1, it can be seen that the changes in PV power output time series of plant B consistently lag that of plant A with a time interval of around 50 min, thus present spatio-temporal correlation in a visual way.

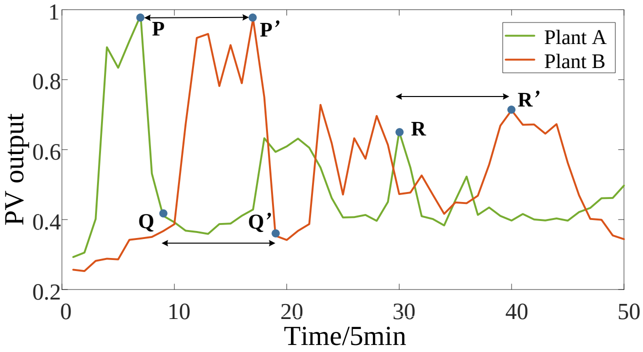


**Fig. 1.** Spatio-temporal correlation of PV power output time series between plant A and plant B with 19.6 kilometers of distance.

From the analysis mentioned above, it is not difficult to realize the significance that the data in plant A are beneficial to the PV power output forecast of plant B, especially for forecast time horizon within 1 h. Therefore, the calculation of time lag $\tau$ is necessary.

In this article, SCCF is taken into consideration to describe the correlation degree between two PV power output time series $X_t$ and $Y_t$ [37].

$$r_{xy} = \frac{C_{xy}(\tau_i)}{\sqrt{C_{xx}(0)C_{yy}(0)}} \ , \quad \tau_i = -k, -k+1, \cdots, 0, 1, 2, \cdots, k \tag{2}$$

$$C_{xy}(\tau_i) = \frac{1}{n}\sum_{t=1}^{n-\tau_i}(X_t - X_M)(Y_{(t+\tau_i)} - Y_M) \ , \quad \tau_i = -k, -k+1, \cdots, 0, 1, 2, \cdots k \tag{3}$$

In (2)-(3), $C_{xx}(0), C_{yy}(0)$ mean when $\tau_i = 0$ , the value of correlation degree $C_{xy}(\tau_i)$ in the case of $x = y$ or $y = x$ ; $n$ is the length of $X_t$ and $Y_t$ ; $X_M, Y_M$ are mean values of $X_t$ and $Y_t$ respectively; and $\tau_i$ represents time lag which set as a certain integer. After calculating $C_{xy}(\tau_i)$ with various $\tau_i$ from $-k$ to $k$, the maximum value which shows the strongest correlation for $X_t$ and $Y_t$ can be seemed as final time lag $\tau$. This would be a crucial factor that the two time series are necessary to obtain for the model training phase. However, a considerable quantity of paired plants has no significant correlation shown in the data series in practice. Therefore, it is necessary to set a threshold to examine whether the maximum value of time lag can be accepted, so as to verify whether the novel method proposed in this paper is appropriate for the predicted time horizon aiming at paired plants.

The figure of SCCF value in PV power time series indicates the time lag in which the correlation is strongest, and is constructed by a great number of time lag values which can describe the trend of the cross-correlation coefficient. In this way, it would be desirable that the correlation check can bring about the selection of neighboring plants having different distances. An appropriate neighboring plant can be chosen, which has the strongest correlation with the target plant, by extracting the maximum SCCF value from various pairs of PV power data measuring plants at the same region. The calculation of SCCF values for target plant A and neighboring plant B is shown in Fig.2 (a). Likewise, the SCCF for target plant A and neighboring plant C is shown in Fig.2 (b). The maximum SCCF values are marked by black points. From this figure, we can observe that the trend and maximum value of SCCF in Fig.2 (a) are both better than Fig.2 (b), which is decided on the distance from the target plant and cloud condition over the plants. In general, the higher the SCCF values between two plants, the better correlation is achieved. Also, in order to obtain the optimal performance of power forecasting, when the maximum SCCF values of two pairs of plants are nearly the same, the neighboring plant must be the one with a higher time lag value that contains more power information.

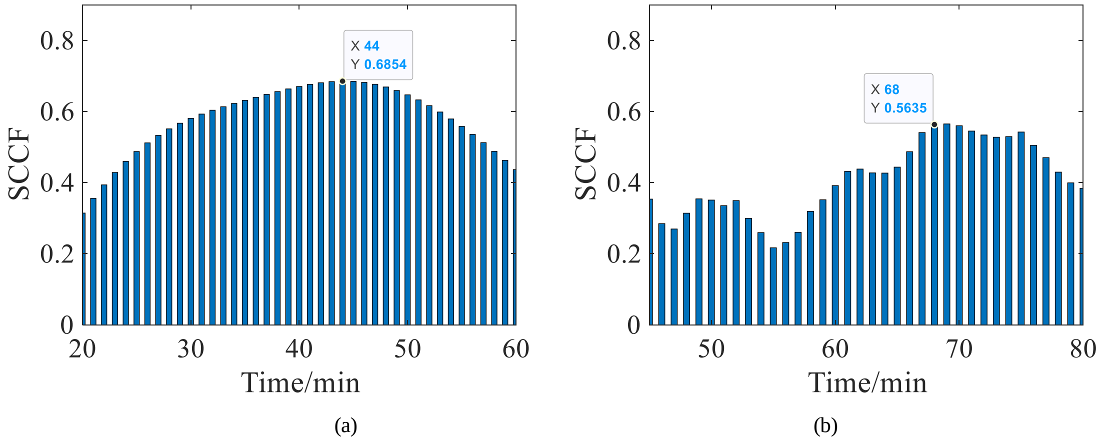


**Fig. 2.** (a) SCCF of solar power in the pair plant A-plant B; (b) SCCF of solar power in the pair plant A-plant C.

After selecting the neighboring plant which has the best correlation with the target plant, historical data of the neighboring plant can be applied to the target plant's power forecasting. With the calculated time lag $\tau$, the technical route of traditional solar PV power forecasting by using power data acquired from target and neighboring plants can be listed as follows.

1) Select a neighboring plant that has the strongest correlation with the target plant by calculating SCCF.

2) Extract time lag value $\tau$.

3) Add power data of neighboring plant at the current time with historical power data of target plant as model inputs.

4) Preliminary predicted value of target plant can be acquired by transporting these inputs into the forecasting model.

5) Prediction rectification on preliminary predicted value.

6) PV power forecasting for the target plant at the target time ($\tau$ min later) can be achieved.

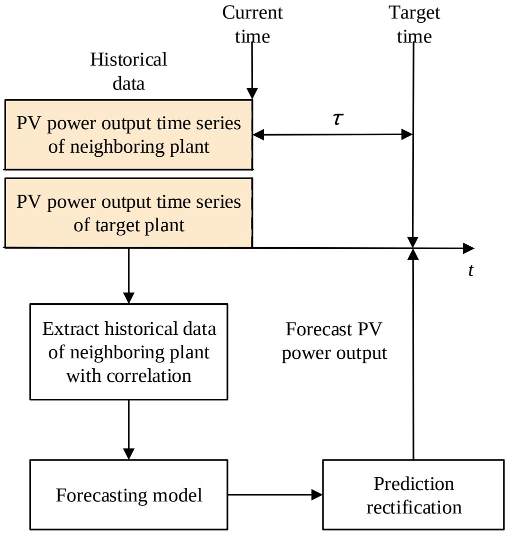

Current time
Target time
Historical data
PV power output time series of neighboring plant
$\tau$
PV power output time series of target plant
$t$
Extract historical data of neighboring plant with correlation
Forecast PV power output
Forecasting model
Prediction rectification

**Fig. 3.** A brief flow chart for the process of traditional PV power forecasting with historical data from the neighboring plant.

### *2.2 Cloud characteristic extraction*

Except for spatio-temporal correlation excavated in PV power output time series, cloud information should also be considered, because the fluctuation of PV power is subject to the rapid change of cloud cover proportion. In Fig.4, two satellite images are captured from plant A at time P and time Q. After time-varying pattern matching, time P' and time Q' can be acquired with a time lag $\tau$ , corresponding to time P and time Q. It is to be noted that plant A is covered with a few clouds in the former image, whereas, for the latter image, plant A is surrounded by thick white clouds. In other words, the PV power output value at time P is higher than the value at time Q, which is also appropriate for the case at time P' and time Q'. Therefore, two simple conclusions can be drawn: i) for the plant-blocked area, PV power output in the thin cloud or blue sky environment is much higher than that in a thick white cloud environment; ii) PV power output forecast aiming at target plant B by using satellite image information around neighboring plant A is workable theoretically.

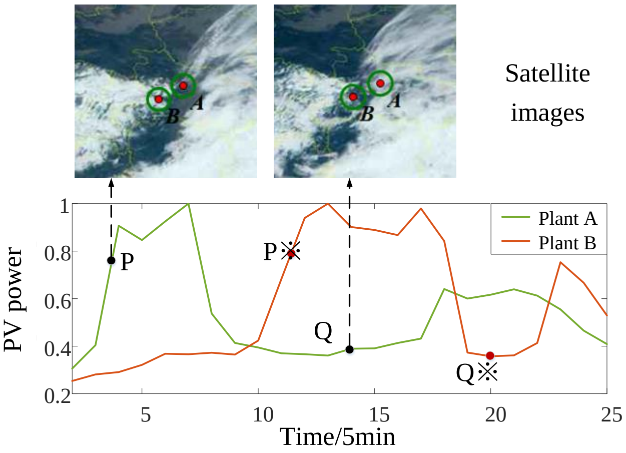


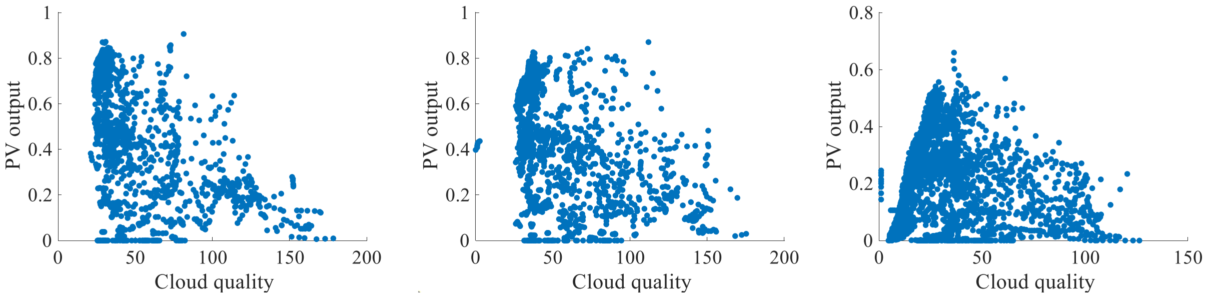
**Fig. 4.** Satellite image-PV power relation between plant A and plant B.

To achieve cloud characteristic extraction, in this paper, we define gray-related features of the satellite image, which are affected by irradiance, cloud thickness, steam dispersity, time, etc. as cloud quality (CQ). This type of index is capable of analyzing the shielding degree of cloud cluster to PV power output and is mainly influenced by cloud gray information. In this part, we extract gray values for each cloud pixel in a certain region, then add and average gray values to get the CQ index $G_{CQ}$ .

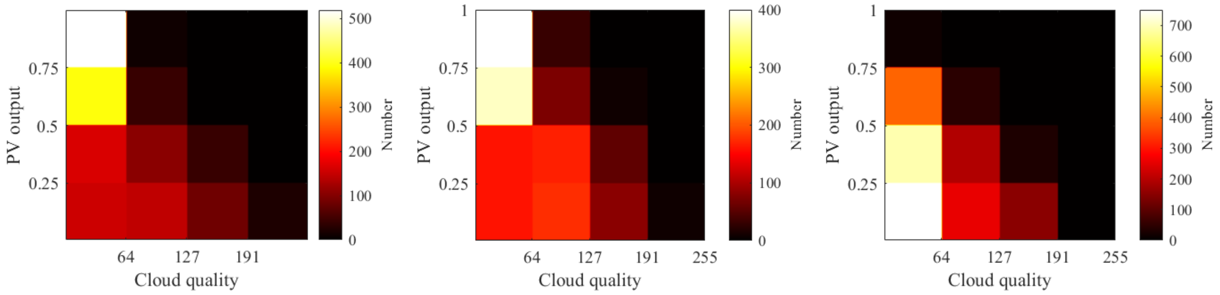
$$G_{CQ} = \frac{\sum_{i=1}^{N} G(x_i, y_i)}{M} \tag{4}$$

Here $G(x_i, y_i)$ means the gray value of coordinate $(x_i, y_i)$ in satellite image; $M$ means the number of all sky pixels in the circular region; *N* means the number of calculated pixels in a circular region that represents the target PV plant with radius *r* of the satellite image; The radius *r* of the selected circular subimage region can be set to the empirical value of 5 pixels, which can effectively cover the sky conditions of the target PV plant.

To verify the correlation between $G_{CQ}$ values and PV power theoretically, we extract $G_{CQ}$ values from circular regions covered multiple PV plants in each satellite image and PV power data at the same time, with 3-time horizons ranging from 9:00 to 10:00, 12:00 to 13:00, 16:00 to 17:00, respectively. The relation in each time horizon is presented by scattering and color block diagrams. The figures are shown as follows.



(a) (b) (c)



(d) (e) (f)

**Fig. 5.** Relationship between cloud quality values and PV output data. (a)(d) Time horizon ranges from 9:00 to 10:00. (b)(e) Time horizon ranges from 12:00 to 13:00. (c)(f) Time horizon ranges from 16:00 to 17:00.

It is shown that in the first two cases, most spots gather at the lower-left triangle, which indicates the mapping relation between cloud quality and PV output data. Generally, when the gray values are less than 64, a quite large proportion of PV output data are above 0.5. And when the gray values increase gradually, most points are located below. This conclusion obviously conforms to the actual situation that when the extracted region is in blue sky circumstance, the output usually rises high for the sake of a large amount of irradiance arrival; when cloud cluster is moving above the PV plant, the output value declines rapidly and maintains at a low level. Due to the diverse intensity of sunlight, the gray value of cloud pixel is hard to get above 200, resulting in few points located at the bottom right corner. For the third case depicted in Fig.(c), Fig.(f), the peak value of output can not arrive higher than 0.75 with the limitation of time. It is not hard to observe that in this period of time, the sun is gradually setting, which leads to a weak light environment. So it makes sense when the gray values keep low, in other words, at a darker circumstance, PV output extremely drops down. To sum up, we can draw a simple conclusion that cloud quality is one of the necessary factors to be considered in power forecasting. Besides, the values of PV power are also subject to the number of cloud pixels in the plant-blocked area by affecting the level of light penetration. Hence, taking account of the cloud pixel number and time information into the PV power forecast is highly beneficial.

### *2.3 Training model methods*

Information we select from the neighboring plant is explained and proved in detail, which can be considered as a part of inputs during model establishment. Except for model input extraction, the selection of machine learning methods is also non-negligible. Here we select three representative machine learning methods, including SVM, GBDT, Autoregressive moving average model (ARMA), to forecast the PV power of the target plant with a time horizon from 15 min to 4 h. In order to verify the effectiveness of the proposed method with model input based on spatio-temporal correlation and neighboring image information, we apply these models to make a contrast with other benchmarks without neighboring image information. The introduction of these models can be described as follows.

#### *2.3.1 Support Vector Machine*

SVM is one of the widely used machine learning methods developed from statistical learning theory and is well-received due to its better performance compared with other conventional methods. Its statistical learning theory provides effective theoretical support with a united frame, thus manage the problem of a limited learning sample [38]. The mathematical principle of SVM applied to regression forecast is denoted as follows.

With regard to a series of given samples: $T(x, y),\ (x_1, y_1),\ (x_2, y_2)\ , .... , (x_n, y_n) \in R^n \times R$ , assume the regression function as (5):

$$F = \left\{ f \,\middle|\, f(x) = w^T \cdot x + b, w \in R^n \right\} \tag{5}$$

Then the structure risk function in SVM can be formulated as (6):

$$R_{reg} = \frac{1}{2}\|w\|^2 + C \cdot R_{emp}[f] \tag{6}$$

where $\|w\|^2$ is the describing function; $f$ is the complexity term; $C$ is a constant value which means the tradeoff between empirical risk and model complexity. To further determine the optimal hyperplane in case of a linearly inseparable dataset, the main solution of the nonlinear Support Vector Regression (SVR) method is to map the input $x$ into higher dimensional feature space through the nonlinear mapping process. Then proper linear regression can be achieved in the feature space. Therefore, in this newly formed space, there will be a possibility that the data can be linearly separated. Then the regression problem can be denoted in another way:

$$min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\xi_i \tag{7}$$

which subject to (8).

$$y_i(w \cdot \phi(x_i) + b) \geq 1 - \xi_i, \quad \xi \geq 0,\ i = 1, ..., C > 0 \tag{8}$$

The inner products $\phi(x_i)$ in higher dimensional space can be substituted by various kinds of kernel functions $K(x_i, x_j)$. Appropriate selection of the calculation kernels is capable of performing necessary computations directly. Some frequently-used kernels are depicted as follows.

$$K(x, x_i) = \exp(-\gamma \left\| x - x_i \right\|^2) \tag{9}$$

$$K(x, x_i) = (1 + x \cdot x_i)^d \tag{10}$$

Equations (9)-(10) describe radial basis function (RBF) kernel and polynomial kernel respectively. Here $d$ represents the degree of the polynomial kernel, $\gamma$ represents a constant determining the width of RBF kernel.

In the actual situation, the kernel function selection would exert a great impact on the final realized effect. So it is crucial to choose a proper kernel function to optimize the kernel function solution. Generally, RBF kernel function, polynomial kernel function, and sigmoid function are the most commonly used kernel functions in SVR.

### *2.3.2 Gradient Boosting Decision Trees*

GBDT is essentially an ensemble machine learning technique where multiple decision trees are trained and used to predict unseen data, and it has proven to be one of the most powerful techniques for building predictive models because of its advantages over simplicity and effectiveness [39]. Here, the GBDT method is used to generate a prediction model for solar PV power forecasting, which is composed of Multiple Classification and Regression Trees (CART) by utilizing gradient boosting techniques [40]. To achieve a better understanding, here we present a well-articulated introduction to the theory of GBDT.

Assume $x$ is the vectors of features and $y$ is the target, $\{(x_i, y_i)\}_{i=1}^{n}$ is a training set. $F_M(x)$ is the final model after $M$ times of iteration and $L(y, F(x))$ is a differentiable loss function. Then details will be introduced to describe how the final model $F_M(x)$ is established by gradient boosting on the basis of decision trees.

Initiate $F_0(x)$ with a constant $\gamma$, then the loss function can be minimized by (11).

$$F_0(x) = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, \gamma) \tag{11}$$

After that, gradient boosting iteration will be continued. For the $m$ th iteration, pseudo-residuals $r_m$ can be calculated by (12).

$$r_{im} = -[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}]_{F(x)=F_{m-1}(x)} , \quad i = 1, \ldots, n \tag{12}$$

Then decision tree $h_m(x)$ is trained with a fixed $J$ depth by using $\{(x_i, r_{im})\}_{i=1}^{n}$. After finishing the calculation of $h_m(x)$, the multiplier $\gamma_m$ can be derived as (13).

$$\gamma_m = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma \cdot h_m(x_i)) \tag{13}$$

Update the model $F_m(x)$ by (14).

$$F_m(x) = F_{m-1}(x) + \upsilon \cdot \gamma_m \cdot h_m(x) \tag{14}$$

In this function, $\upsilon$ is the learning rate, which has another name called “shrinkage factor”. Usually the greater the $\upsilon$ is, the shorter computation time and worse performance during the learning process. Then after $M$ times of iteration, we have the final model $F_M(x)$.

### *2.3.3 Auto-Regressive and Moving Average Model*

In stationary random sequence analysis, the ARMA model is one of the most widely used methods which have been in-depth studied and investigated by scholars. After years of development and implementation, the ARMA model has been achieved as a progressive, integral, systematic modeling method. Meanwhile, it owns a statistical sense of perfection and a substantial theoretical basis.

$$S(t) = \sum_{i=1}^{p} \alpha_i S(t-i) + \sum_{j=1}^{q} \beta_j e(t-j) \tag{15}$$

From (15), it is easy to observe that the ARMA model is mainly composed of two parts: auto-regressive (AR) part and moving average (MA) part. In this function, $S(t)$ is forecasted value at time $t$. For the AR part, $p$, $\alpha_i$ represent the order of AR process and AR coefficient. For the MA part, $q$ is the order of MA error term, $\beta_j$, $e(t)$ represent the coefficient of MA and white noise respectively. Here the white noise can generate random uncorrelated variables which contain zero-mean value and constant variance [41]. In general, ARMA requires a lot of historical data to establish the model. In this paper, ARMA (*p*, *q*) is applied as one of the benchmark methods to forecast PV power value within 4 h ahead.

### *2.4 Cloud displacement vector calculation*

After establishing a forecasting model by applying power and cloud information of neighboring and target plants as inputs, PV power data of target plant at the current time as output, model inputs in the next 4 h need to be required according to real-time satellite image and historical power data. As the output power of solar PV plants is mainly affected by the amount of irradiance reaching the ground's surface, and related to the cloud distribution over the plants at the corresponding moment in time, the regions in which clouds are located cover the target plant are necessary to be acquired. In previous researches, different kinds of digital image processing techniques were applied to achieve cloud motion tracking. In order to simplify the process of cloud motion displacement (CMD) calculation and improve its performance, in this article, an improved FPCT method based on convolution transform is applied. This transform method is verified to meet the IPSI property [42], further opens up new possibilities in cloud motion results and reduces the probability of final CMD with few credibilities.

Traditional FPCT method is capable of acquiring an object's motion information, especially rigid parallel motion, by means of transforming image information from the time domain into the frequency domain. After the image processing with Fourier translation-based FPCT method, it is easy to find that the image displacement information is mainly saved in phase spectrum, unrelated to amplitude spectrum, which further brings convenience to calculate CMD by dealing with information in phase spectrum. The detailed deduction of FPCT is elaborated by (16)-(19).

The grayscale matrix resolution of an image $f(x,y)$ is assumed as $M\times N$. After processed with discrete Fourier transform (DFT), the transformed form $F(u,v)$ which is corresponding to $f(x,y)$ can be defined as follows:

$$F(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi(\frac{ux}{M}+\frac{vy}{N})}=\left|F(u,v)\right|e^{-j\phi(u,v)} \quad (u=0,1,...,M-1;\ v=0,1,...,N-1) \tag{16}$$

where $x,\ y$ represent the cartesian coordinates of a pixel, and $u,\ v$ represent the pixel's Frequency domain coordinates.

In the ideal case, if two satellite images $f_1(x,y)$ and $f_2(x,y)$ only differ with a displacement vector $(x_0,y_0)$, then the formula can be expressed as (17) to represent the relation between $f_1(x,y)$ and $f_2(x,y)$.

$$f_2(x,y)=f_1(x-x_0,y-y_0) \tag{17}$$

Then the cross-power spectrum (CPS) can be denoted as (18):

$$C(u,v)=\frac{F_1(u,v)F_2^*(u,v)}{\left|F_1(u,v)F_2^*(u,v)\right|}=\mathrm{e}^{j2\pi(\frac{ux_0}{M}+\frac{vy_0}{N})} \tag{18}$$

where $F^*(u,v)$ is complex conjugate, $\left|F(u,v)\right|$ is the amplitude of $F(u,v)$. After processing CPS $C(u,v)$ by using inverse discrete Fourier transform (IDFT), the result of CMD $(x_0,y_0)$ can be finally denoted as (19).

$$F^{-1}\{C(u,v)\}=\delta(x-x_0,y-y_0) \tag{19}$$

However, considering the generation, dissipation, and deformation of clouds, the ideal condition of rigid cloud motion would not happen frequently, which results in a considerable amount of noise during image registration and makes the true displacement value $(x_0,y_0)$ submerged by other noise pulses. Hence, in order to solve this problem, we choose the same satellite image pair for multiple times using convolution transform with different kinds of convolution kernel matrix $h(x,y)$, then FPCT calculation is proceeded, thus generating plenty of CMD calculation results.

Before image pre-processing, the convolution transform we select should be verified corresponding to the IPSI characteristic, which means that the information in the phase spectrum remains the same either before or after the transformation. In this part, we assume that a grayscale matrix $f(x,y)$ has the dimensions $A\times B$, convolution kernel matrix $h(x,y)$ has the dimensions $C\times D$. Then the discrete convolution transform $I(x,y)$ is denoted as follows:

$$\begin{aligned}I(x,y)&=f(x,y)*h(x,y)\\&=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}h(m,n)\cdot f(x-m,y-n)\\&(x=0,1,2,...,M-1;\ y=0,1,2,...,N-1;\\&M=A+C-1;\ N=B+D-1)\end{aligned} \tag{20}$$

where $f(x,y)*h(x,y)$ means the convolution of $f(x,y)$. The convolution theorem is represented as (21):

$$G(u,v)=F(f(x,y)*h(x,y))=F(u,v)H(u,v) \tag{21}$$

Hence, according to (18) and (21), the CPS matrix can be acquired from (22):

$$C(u,v)=\frac{G_1(u,v)G_2^*(u,v)}{\left|G_1(u,v)G_2^*(u,v)\right|}=\mathrm{e}^{j2\pi(\frac{ux_0}{M}+\frac{vy_0}{N})} \tag{22}$$

which is equal to (18), so as to verify the fact that convolution transform satisfies the IPSI theory.

After being processed with various convolution kernels, multiple CMD results can be obtained. Assume that the coordinates of CMDs are:

$$D=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\} \tag{23}$$

In order to determine the desired final displacement, it is necessary to extract the most credible CMD value from the result dataset, which was generated by the improved FPCT method mentioned above. Here we utilize the Gaussian distribution fitted curve to achieve final displacement vector extraction.

In Fig. 6, CMD values of a pair of satellite images are shown in the X-coordinate, the values in Y-coordinate represent the number of points locating at different coordinates counting as 121 in total. From the view of the CMD distribution diagram, we can draw a simple conclusion explicitly: most result points converge in a reliable region that may be the correct CMD in great probability. During observation of the coordinate points based on the density and distance distribution in Fig. 6, the Gaussian distribution curve enables to fit the displacement point distribution. By using the automatic curve-fitting function in MATLAB, here $b_1$ denotes the mean value of the fitting curve which represents the final CMD value after rounding off.

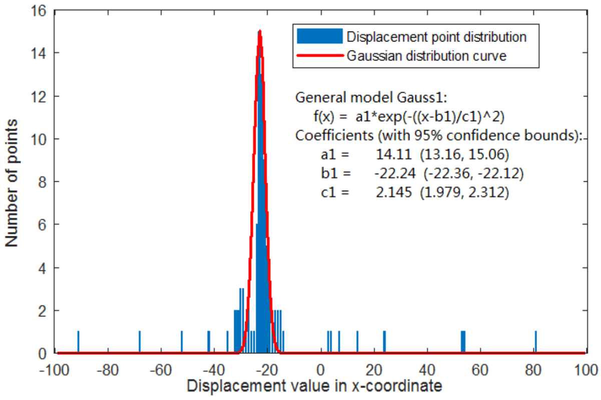


**Fig. 6.** Distribution of displacement points with Gaussian distribution fitted curve.

### *2.5 Ultra-short-term PV power forecasting based on spatio-temporal correlation and neighboring information*

In this section, the framework of the proposed method using spatio-temporal correlation and neighboring information for PV power forecasting is introduced in detail. The implementation of this framework is depicted as Fig. 7, which denotes that the method mainly consists of 3 major parts.

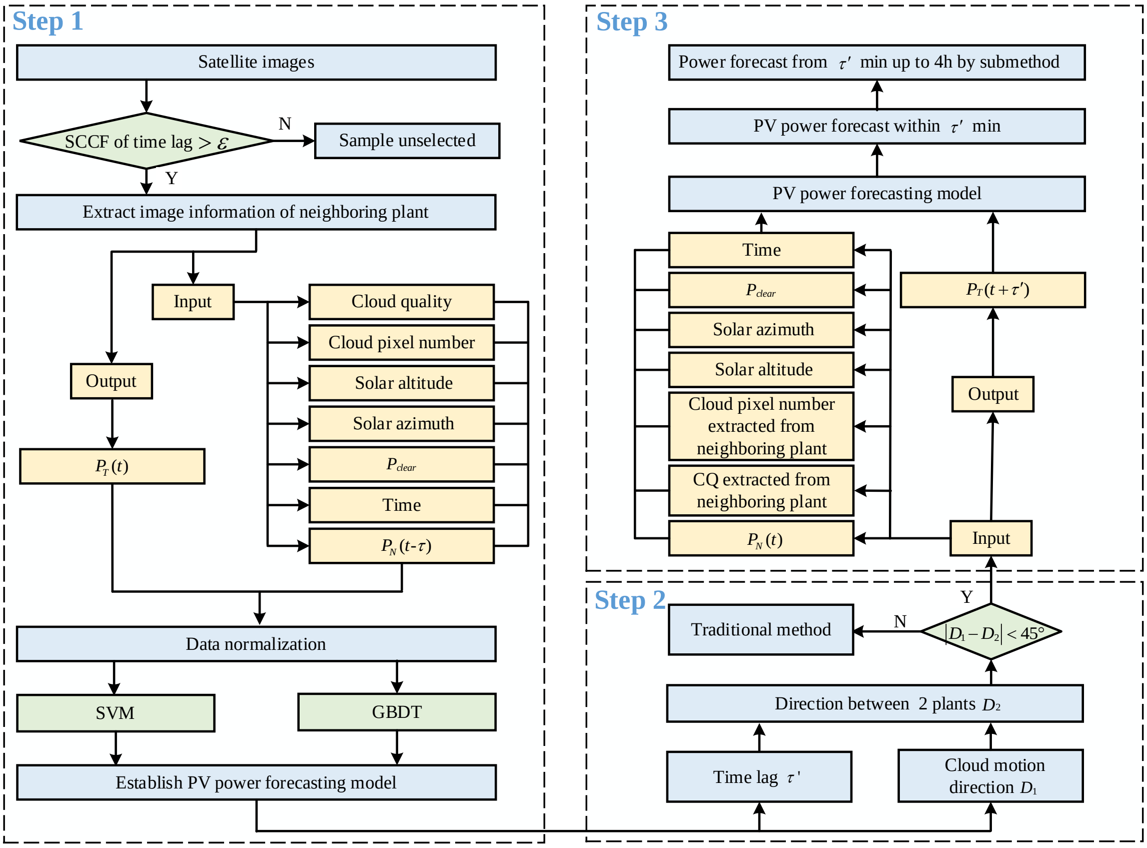

Step 1
Satellite images
SCCF of time lag > $\varepsilon$
N
Sample unselected
Y
Extract image information of neighboring plant
Input
Output
Cloud quality
Cloud pixel number
Solar altitude
Solar azimuth
$P_{clear}$
Time
$P_N(t-\tau)$
$P_T(t)$
Data normalization
SVM
GBDT
Establish PV power forecasting model
Step 3
Power forecast from $\tau'$ min up to 4h by submethod
PV power forecast within $\tau'$ min
PV power forecasting model
Time
$P_{clear}$
Solar azimuth
Solar altitude
Cloud pixel number extracted from neighboring plant
CQ extracted from neighboring plant
$P_N(t)$
$P_T(t+\tau')$
Output
Input
Step 2
Traditional method
N
Y
$|D_1-D_2|<45°$
Direction between 2 plants $D_2$
Time lag $\tau'$
Cloud motion direction $D_1$

**Fig. 7.** The framework of the proposed PV power forecasting method based on spatio-temporal correlation and neighboring information

1) Model establishment: In this module, historical data from several PV plants located at different places within a certain region are collected. Then the neighboring plant which has the strongest correlation with the target plant can be determined by calculating the maximum value of SCCF, as well as the value of time lag $\tau$. In order to forecast the PV power output of the target plant, the power datum of the neighboring plant, which is $\tau$ min before the forecasting moment, can be selected to guide and facilitate the process of forecasting. Besides, as analyzed in Section 2.2, cloud characteristics extracted from the sky area above PV plants are verified to have a direct impact on PV power output value at the corresponding time. Here cloud quality and cloud pixel number in a circular atmospheric region above the neighboring plant with a fixed radius are selected as two model inputs. Similarly, the selected moment of cloud characteristics of the neighboring plant is also $\tau$ min before the forecasting time. With the consideration of the effects of the sun's movement on power forecasting, solar altitude, solar azimuth, and corresponding time are also included. Except for cloud information, power value of the target plant under a clear sky environment in a whole day is also necessary. For the sake of the earth's revolution around the sun, the daily extraterrestrial irradiance of a certain region accords with adjacent similarity and annual periodicity. The adjacent similarity means that the changing rule of irradiance in one day is similar to the one of the neighboring days. As for annual periodicity, it presents the phenomenon that the irradiance data in the same ahargana of different years are nearly the same. Therefore, in order to obtain the value of PV power output under a clear sky environment, power data in a clear sky condition selected from neighboring days for the same predicted PV plant can be utilized to further achieve the power forecasting. Then after data normalization, the forecasting model can be established by applying SVM and GBDT respectively.

2) Correlation judgment: In order to determine whether the proposed method should be used to achieve PV power forecasting of the target plant during a period of time, it is necessary to detect whether there is a neighboring plant with a strongly spatio-temporal correlation to the target plant existing in target plant's perimeter zone. If the peak value of all SCCF curves of checked surrounding plants is greater than the threshold $\varepsilon$, and the cloud motion direction is from the neighboring plant with a maximum value of SCCF towards the target plant, then the time lag $\tau$ can be determined and the neighboring plant's historical data could be used in next forecasting part. If not, the traditional forecasting method without neighboring data is applied to generate ultra-short-term solar PV power forecasting.

3) Power forecasting: After determining the neighboring plant and the value of time lag $\tau$, model inputs could be obtained, such as cloud quality, cloud pixel number of neighboring plant, forecasting time $t$, solar azimuth, solar altitude, clear sky power output of target plant $P_{clear}$, the power output of neighboring plant $\tau$ min before forecasting time $P_N(t-\tau)$, to acquire power output of target plant at forecasting time $P_T(t)$. However, for the sake of wind speed and geographic distance between target plant and neighboring plant, generally, the time lag value $\tau$ is not high enough to reach 4 h, or even closer. Therefore, the method considering spatio-temporal correlation can only generate forecasting results from 15 min up to $\tau$ min ahead. For the horizon in the range of $\tau$+15 min to 4 h, the submethod should be applied which is depicted as Fig.8. As cloud motion is a reflection of the atmospheric physical motion process, CMDs calculated from a series of consecutive satellite image pairs should usually be similar due to inertia. Hence, based on historical cloud motion calculation, cloud motion above the neighboring plant in the future can be obtained by using linear extrapolation, then the regions which will cover the neighboring plant can also be acquired. After cloud

characteristic extraction and other inputs calculation, it is feasible to achieve the output forecasting of the target plant with a time horizon between $\tau$+15 min to 4 h.

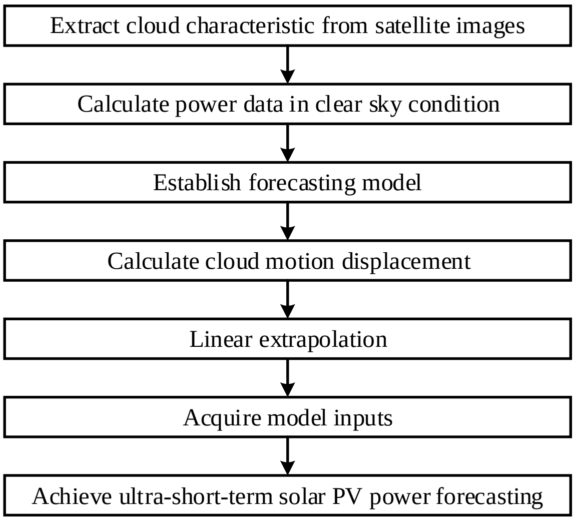

Extract cloud characteristic from satellite images
Calculate power data in clear sky condition
Establish forecasting model
Calculate cloud motion displacement
Linear extrapolation
Acquire model inputs
Achieve ultra-short-term solar PV power forecasting

**Fig. 8.** The sub-method of PV power forecasting with a time horizon between $\tau$+15 min to 4 h.

The main technology route of ultra-short-term solar PV power forecasting considering cloud information from the neighboring plant is described as follows:

1) Calculate all SCCF indexes of two PV power time series generated from each surrounding plant and target plant pair with a fixed-size observation window. If the maximum value of SCCF is less than the threshold $\varepsilon$, the group of samples should not be regarded as a part of the model training sample set. If not, the corresponding time of the maximum value of SCCF can seem as time lag $\tau$;
2) Extract neighboring plant's circular subimage region with radius $r$;
3) Acquire cloud quality, cloud pixel number from subimage, solar altitude, solar azimuth, power in clear sky condition $P_{clear}$, time, power output of neighboring plant $\tau$ min ago as model inputs, power output of target plant as model output;
4) Select one of these two machine learning methods to train the groups of samples: i) SVM; ii) GBDT.
5) Calculate time lag $\tau'$ between forecasted target plant and neighboring plant pair;
6) Calculate cloud motion direction $D_1$ and physical distance direction of 2 plants $D_2$;
7) If the absolute value of the difference between $D_1$ and $D_2$ is less than 45°, then skip to step 8); or else, apply traditional method without neighboring data to achieve the forecasting;
8) Acquire cloud quality, cloud pixel number from neighboring plant's subimage, solar altitude, solar azimuth, power in clear sky condition $P_{clear}$, predicted time, power output of neighboring plant $\tau'$ min before predicted time as model inputs,
9) Forecast power output of target plant within the range between 15 min and up to $\tau'$ min by using one of the established models mentioned in step 4);
10) Forecast power output of target plant within the range between $\tau'$ +15 min and up to 4 h by using sub method.

## III. CASE STUDY

### *3.1 Data*

The dataset used in this study contained actual solar PV power output with 15 min observation intervals, from January 2018 to June 2019, for 21 PV systems monitored in Jilin Province, China, which is shown in Fig. 9. The region selected lies in the latitude range 42.73-45.83° and the longitude range 122.83-125.43°. Satellite image datasets are acquired from a stationary meteorology satellite named Fengyun-4A, which is in charge of the National Satellite Meteorological Center (NSMC). The file size of each image is approximately 170 Mb and the time resolution per graph is about 5 min. In order to be consistent with power data, the time resolution of images should also be modified to 15 min. In this paper, we choose 2 adjacent PV plants pair as testees: plants 503 and 508, plants 519 and 520, to verify the performance of the proposed method. Detailed geographical information of these 4 PV plants are depicted in Table 1.

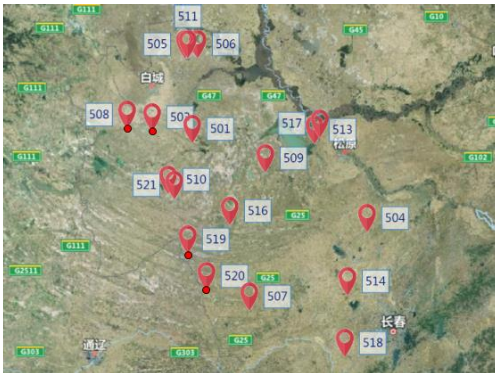


**Fig. 9.** Geographical distribution of 21 PV plants.

**Table 1**. Location information of 2 PV plants pair.

| | Plant | Lat | Lon | Distance |
|---|---|---|---|---|
| 1 | 503 | 45°30´ N | 122°83´ E | 19.6km |
| 2 | 508 | 45°32´ N | 122°58´ E | |
| 3 | 519 | 44°40´ N | 123°20´ E | 33.5km |
| 4 | 520 | 44°13´ N | 123°39´ E | |

### *3.2 Simulation process*

In this part, the key process of model establishment is analyzed and introduced in detail. Data averaged each 15 min from 2 pairs of target and neighboring plants: plant 503 and plant 508, plant 519 and plant 520 are utilized in the study. Plants 503 and 519 are target plants and plants 508 and 520 are neighboring plants. During the process of model establishment, solar data records from Jan. 01, 2018 to May 31, 2019 in Jilin province are considered as training and validation sets. To clearly observe the performance of forecasting methods, solar PV power data of each target plant are predicted by both 5 methods, from Jun. 01, 2019 to Jun. 30, 2019, counting up to 30 days in total. To realize the proposed method elaborately, we choose Jul. 30, 2019 at 9:00 a.m. as the forecasted start time, to further explain the forecasting process, which could also be shown as Fig. 10.

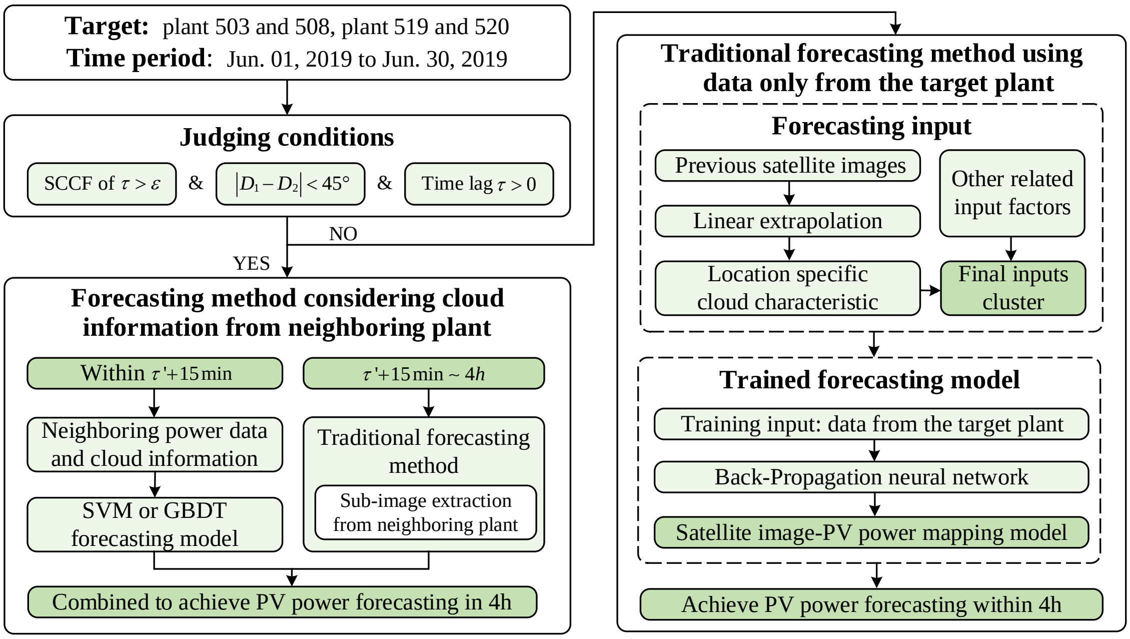

Target: plant 503 and 508, plant 519 and 520
Time period: Jun. 01, 2019 to Jun. 30, 2019
Judging conditions
SCCF of $\tau > \varepsilon$
&
$|D_1 - D_2| < 45°$
&
Time lag $\tau > 0$
NO
YES
Forecasting method considering cloud information from neighboring plant
Within $\tau'$+15min
$\tau'$+15min ~ 4$h$
Neighboring power data and cloud information
SVM or GBDT forecasting model
Traditional forecasting method
Sub-image extraction from neighboring plant
Combined to achieve PV power forecasting in 4h
Traditional forecasting method using data only from the target plant
Forecasting input
Previous satellite images
Linear extrapolation
Location specific cloud characteristic
Other related input factors
Final inputs cluster
Trained forecasting model
Training input: data from the target plant
Back-Propagation neural network
Satellite image-PV power mapping model
Achieve PV power forecasting within 4h

**Fig. 10.** Flowchart of the simulation process.

First, we need to determine whether the proposed model is suitable for a specific forecasting time period. The time lag value $\tau$ should accord with the following terms:

1) the max value of SCCF $\tau$ must exceed a certain threshold level $\varepsilon$, in case PV power time series of two plants may have a weak correlation, especially on the occasion of undesired cloud motion direction;

2) calculated time lag $\tau$ should accord to the fact, which means the cloud motion direction should be the same with relative geographical direction of the selected target and neighboring plants;

3) the value of time lag $\tau$ must be positive.

If calculated $\tau$ fits the above-mentioned conditions, then the power output and cloud information of neighboring plant at time $t-\tau$ can be seemed as guidance to facilitate the forecasting of target plant output at time *t*.

If not, a traditional forecasting method based on cloud characteristic extraction by using linear extrapolation can be applied to achieve the forecast. As for this method, only target plant's data are used, which contains cloud quality, cloud pixel number, forecasting time *t*, solar azimuth, solar altitude, clear sky power output and real solar data records. A "satellite image – PV power" mapping model is established by using the Back-Propagation neural network (BPNN) with model output solar data records and other factors identified above as model inputs. Due to cloud motion is an inertia-based process and it is incredible to generate vigorous change in velocity within a short time, according to calculated CMDs between each pair of consecutive satellite images during the previous time period, the locations where clouds cover the plant can be acquired after calculated displacement values average and linear extrapolation. Then location-specific cloud characteristics with other factors are able to be put into the training model, thus forecasting PV power output within 4 h can be acquired.

When the calculated time lag $\tau$ is reasonable, which indicates that there is a strong spatio-temporal correlation between the target and neighboring plants, it is theoretically feasible to adopt neighboring information with regular forecasting factors such as time, meteorological information, and historical power output data as model inputs, to conduct the future forecasting process. In this paper, the inputs of the proposed forecasting model not merely contain the neighboring power data like traditional spatio-temporal correlation-based method, but also contain the neighboring cloud information with the consideration of a graphical standpoint, including cloud quality, cloud pixel number from neighboring plant's subimage, solar altitude, solar azimuth, power in clear sky condition $P_{clear}$, predicted time, the power output of neighboring plant $\tau'$ min before the predicted time $t$. Then the output of the target plant at time $t$ can be forecasted by using the SVM or GBDT forecasting model. However, with regard to the restriction of lagging duration, the forecasting method considering spatio-temporal correlation further increases the difficulties in the length of predictable time horizon. For the forecasting PV power output within the range between $\tau'$ +15 min and up to 4 h, the applied method is similar to the traditional forecasting method based on cloud characteristic extraction by using linear extrapolation, except for the starting region of subimage extraction, which is from the neighboring plant rather than target plant.

Here we apply a set of approaches to test both of these methods' performance of ultra-short-term PV power forecasting for two target plants. Five forecasting methods are exhibited as follows.

1) Method 1: Proposed ultra-short-term solar PV power forecasting with neighboring plant's cloud information and power data by using SVM.

2) Method 2: Proposed ultra-short-term solar PV power forecasting with neighboring plant's cloud information and power data by using GBDT.

3) Benchmark 1: Method 1 without cloud characteristic inputs.

4) Benchmark 2: Method 2 without cloud characteristic inputs.

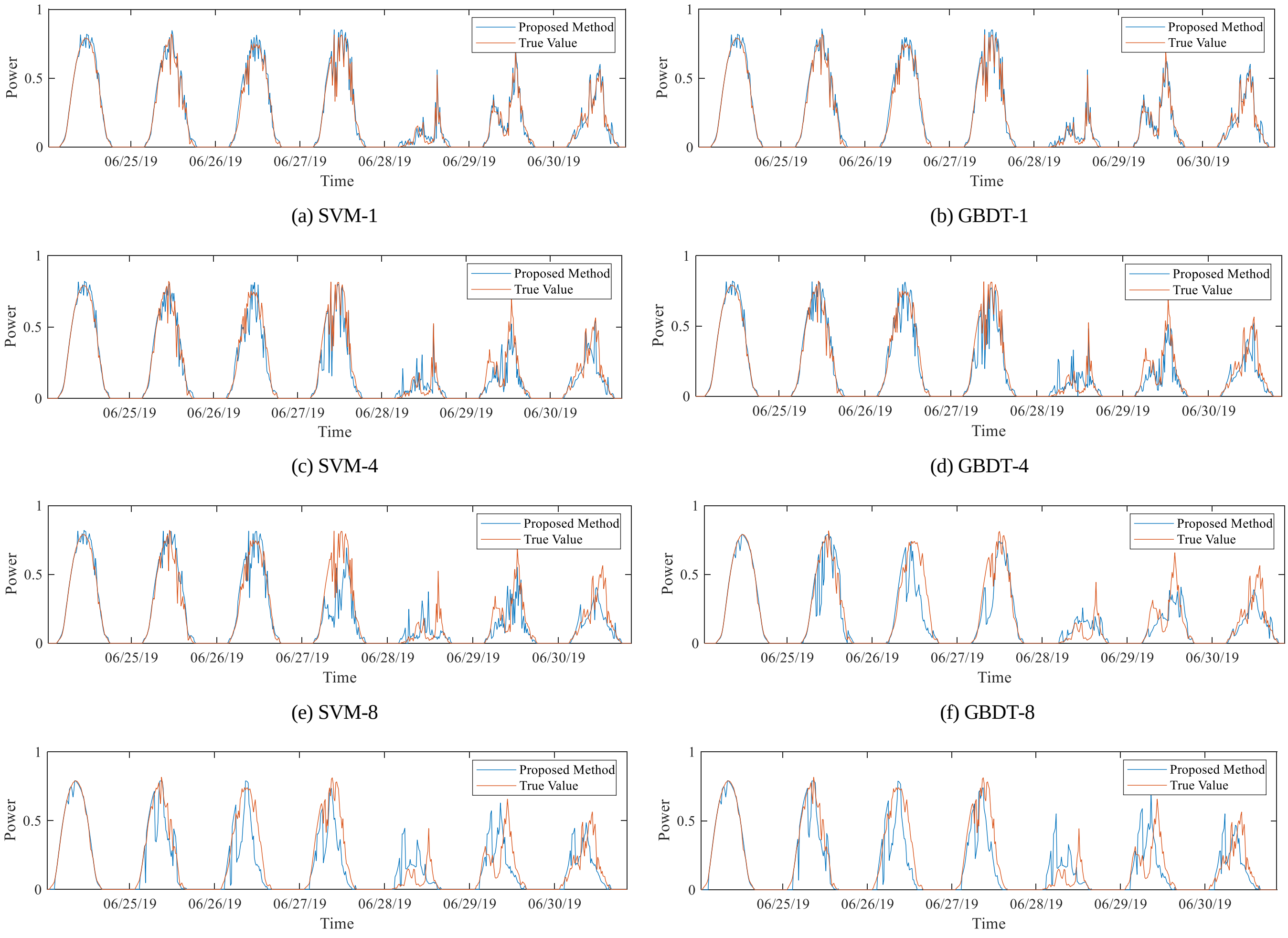
5) Benchmark 3: PV power forecasting by using ARMA.

### *3.3 Results and discussion*

A detailed evaluation of the accuracy of solar PV power forecasting was performed in this Section for the ensemble of two target PV plants: plant 503 and plant 519. The proposed algorithm-based spatio-temporal correlation and neighboring image information is applied as a primary forecasting approach to the considered datasets. As indicated before, 30 days of the forecasting stage from Jun. 01, 2019 to Jun. 30, 2019 are considered as test data.

To evaluate the effectiveness of the proposed method, we compare this model with traditional forecasting methods (Benchmark 1 and Benchmark 2) and ARMA (Benchmark 3), which apply the same artificial intelligence-based models with the proposed method, respectively SVM and GBDT models. To provide a convenient multi-viewpoints observation, for plant 503, forecasting PV power results from Jun. 24, 2019 to Jun. 30, 2019 are presented as Fig. 10, Fig. 11 and Fig. 12. The 1st, 4th, 8th, 16th forecasting points, corresponding to the results 15 min, 1 h, 2 h, 4 h after the predicted initial time respectively, are denoted in each figure, and further show the forecasting accuracy at different times explicitly.

It can be seen that the forecasting accuracy of ARMA is the lowest among all these artificial intelligence-based models. This method is widely used in the prediction of smooth time series by exploring the laws of data during the data mining process. For the time series with sharp irregular fluctuation, the forecasting results of this method are not ideal, thus deliver a relatively poor performance. As for other methods, forecasting accuracy are both higher than ARMA due to model input addition with power data from neighboring plant based on spatio-temporal correlation. It can also be observed that compared to Benchmark 1 and Benchmark 2, Method 1 and Method 2 have a more effective and robust capability to complete the forecast for the time horizon within $\tau$ min, while the performances of Benchmarks pair and Methods pair are practically similar with the time horizon between $\tau$+15 min and 4 h. To evaluate various forecasting methods comprehensively, statistical parameters such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) are applied to modify the forecasting results of the proposed method and benchmarks according to different evaluation criteria. In Table 2, the calculated statistics of all methods are listed in conditions of two target plants. In terms of forecasting accuracy, it can be seen that Method 1 and Method 2 outperform the other approaches with respect to all error metrics. Considering the MAE and RMSE values as an explanation, the proposed method achieves a higher accuracy at 4.12%, 9.48%, and 3.70%, 8.54% for plant 503 and plant 519, respectively.



(a) SVM-1

(b) GBDT-1

(c) SVM-4

(d) GBDT-4

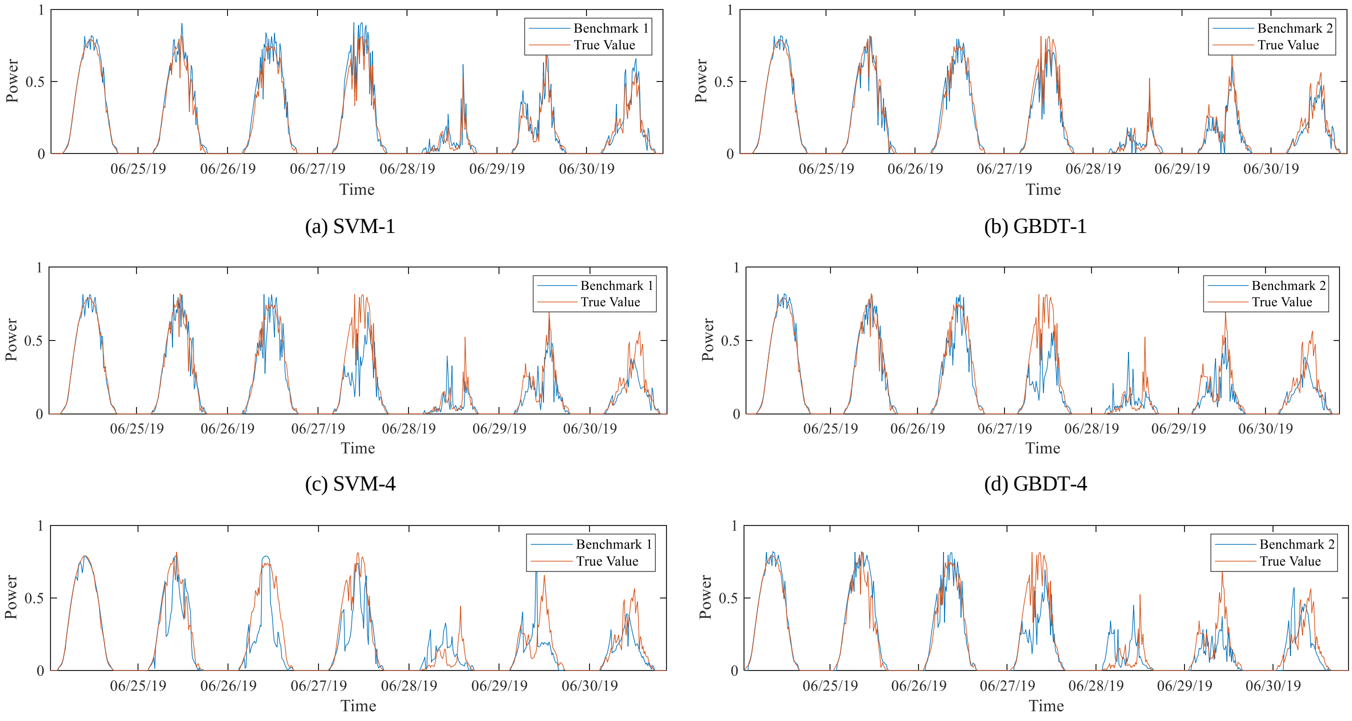
(e) SVM-8

(f) GBDT-8

(g) SVM-16

(h) GBDT-16

**Fig. 10.** Forecasting results of proposed method by using SVM and GBDT aiming at 503 PV plant.



(a) SVM-1

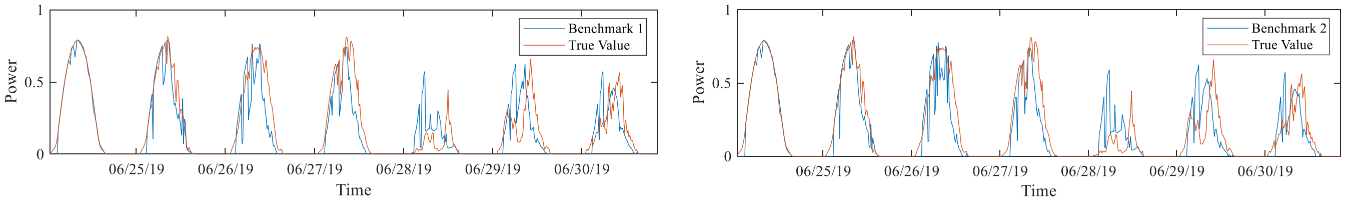
(b) GBDT-1

(c) SVM-4

(d) GBDT-4

(e) SVM-8

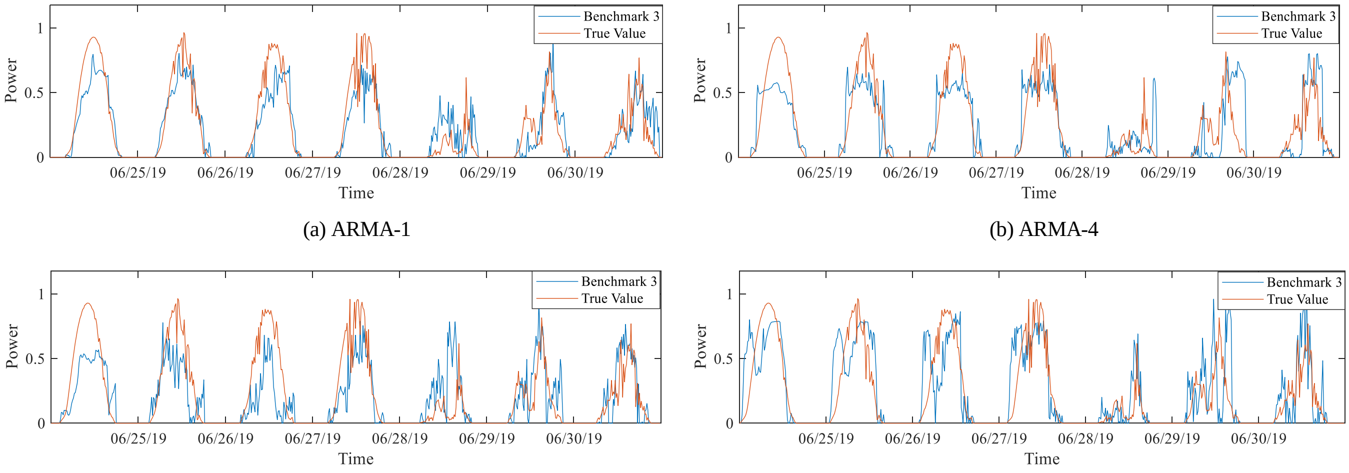
(f) GBDT-8



(g) SVM-16

(h) GBDT-16

**Fig. 11.** Forecasting results of Benchmark 1 and Benchmark 2 by using SVM and GBDT aiming at 503 PV plant.



(a) ARMA-1

(b) ARMA-4

(c) ARMA-8

(d) ARMA-16

**Fig. 12.** Forecasting results of Benchmark 3 by using ARMA aiming at 503 PV plant.

**Table 2**. 30 days' forecasting accuracy of different models.

| Plant | Indexes | Method 1 | Method 2 | Benchmark 1 | Benchmark 2 | Benchmark 3 |
|---|---|---|---|---|---|---|
| 503 | MAE | **4.12%** | 4.14% | 4.42% | 4.60% | 8.47% |
|  | RMSE | **9.48%** | 9.54% | 10.14% | 10.55% | 18.28% |
| 519 | MAE | 3.84% | **3.70%** | 4.22% | 4.28% | 8.57% |
|  | RMSE | 8.77% | **8.54%** | 9.85% | 9.95% | 18.45% |

In order to avoid that the good performance of the proposed method is time-dependent, a supplementary case is further carried out, which aims to forecast the solar PV power output from Feb. 01, 2019 to Jun. 30, 2019, counting up to 4 months in total. The training sets are from Jan. 01, 2018 to Jan. 31, 2019. The results indicate that among all the tested days for target plants 503 and 519, 14.24% and 15.72% of the data satisfy the judging conditions and use the neighboring plant's cloud information-based forecasting method while the others use the traditional method for ultra-short-term solar PV power forecasting. The forecasting accuracy is shown in Table 3, which presents that the proposed method 1 and method 2 are still superior to the benchmark methods in terms of forecasting accuracy. In this 4-month case, the proposed method achieves a higher accuracy at 4.73%, 10.54%, and 4.88%, 11.04% for plant 503 and plant 519, respectively. It verifies the effectiveness of the proposed method together with the one-month forecasting results.

**Table 3.** 4 months' forecasting accuracy of different models.

| Plant | Indexes | Method 1 | Method 2 | Benchmark 1 | Benchmark 2 | Benchmark 3 |
|---|---|---|---|---|---|---|
| 503 | MAE | **4.73%** | 4.84% | 5.02% | 4.94% | 8.39% |
|  | RMSE | **10.54%** | 10.71% | 11.14% | 10.95% | 18.08% |
| 519 | MAE | 4.92% | **4.88%** | 5.14% | 5.24% | 8.61% |
|  | RMSE | 11.13% | **11.04%** | 11.65% | 11.85% | 18.55% |

According to the above comparison of PV power forecasting methods, the advancement and effectiveness of the proposed methods could be verified. The primary advantageous property of the proposed PV power forecasting method is that it could achieve a better forecast precision, which can be attributed to its combining consideration of the spatio-temporal correlation between two adjacent plants. Since the fluctuation characteristic of PV power output is closely related to cloud movement conditions, which establishes the decisive role of cloud feature extraction in the accurate solar PV forecasting process. The consideration of satellite image data captured above both the target and neighboring plants offers the proposed method easy access to the accurate and real-time cloud features. In this way, these cloud features could be regarded as the inputs of the forecasting model for both training and testing, and realize the improvement of forecasting precision compared with Benchmark method 1-3, which does not consider the spatial correlation of neighboring plants. Benchmark method 3 only considers the target plant's historical data to train the forecasting model, hence, it presents the lowest accuracy; On the basis of Benchmark method 3, Benchmark methods 1 and 2 further combine the time-lag information of cloud movement in their forecasting process and thus better results are achieved. However, compared to the proposed method, it only considers the temporal correlation to achieve PV power forecasting, therefore, its performance is inferior to that of the proposed method.

In terms of the limitations, the proposed method shows inadequacy in two aspects: 1) Heavier computation burden. When it comes to computation burdens, Benchmark method 3 is the least time-consuming one and undertakes the minimum computational burden, and Benchmark methods 1,2 also outperform the proposed method. This is due to these methods have no need to extract cloud characteristics from the high-resolution satellite cloud images, which is a relatively comparatively complicated process. In addition, Benchmark method 3 further saves itself some computation time because the relatively complex linear extrapolation step in method 2 is not required in method 3; 2) More restrictions and more relevant information are needed. Before the application of the neighboring plant's cloud information-based forecasting method, it is necessary to detect whether there is a neighboring plant with a strongly spatio-temporal correlation to the target plant existing in the target plant's perimeter zone and whether the cloud motion direction is from the neighboring plant to the target plant. This places restrictions on the extensive applications of this method in comparison with Benchmark method 3 which could achieve forecasting as long as the PV output data of the target plant is available.

It should be noted that, although the proposed method could achieve a great prediction precision, many factors will cause an accuracy decline effect in the forecasting process. Even though two PV power time series is matching well in a certain observation window, there is no guarantee that the power output values during the forecasting time period still accord with the same condition. In other words, the lagging effect will not be constant all the time which reflecting as a changing time lag value. Besides, cloud generation, elimination, deformation are complex atmospheric physics processes. The motion of clouds is hard to track and highly related to wind speed and direction. During cloud motion between two plants, the thickness and shape of clouds will change seriously when arriving at the target plant. Under the circumstance that cloud distribution is highly dispersed, such as blocky clouds, the clouds may even disappear before arrival. All these cases will result in invalid extraction of cloud characteristic input in the proposed method, which fails to bring improvement on forecasting accuracy. What's more, it is a very complicated task to acquire accurate PV power output of target plant in clear sky condition, which is affected by the expansion and demolition of PV power plants, change of temperature, power generation limit by the grid, etc. These factors directly influence the generating capacity and will cause an increased error during forecasting.

### *3.4 Applications and economic benefits*

The economic benefits of the ultra-short-term PV power forecasting method can be categorized into three main aspects:

Firstly, for PV power plants, they attempt to improve the ultra-short-term PV power forecasting accuracy not only for the compliment with the operational requirements of the power system, but also for avoiding penalty due to inaccurate prediction. Take a 100 MW PV plant in Northeast China as an example, each 1% improvement in prediction accuracy could save about 180 thousand yuan of penalty throughout the year under the current assessment standard. That is, the forecasting accuracy is directly related to their economic benefits.

Secondly, with the deepening reform of the electricity market, accurate PV power forecasting lays a solid foundation for the market participants like PV power plants in formulating transaction strategies. Provided that 10% of the electricity consumption will be traded and settled in the spot market, the fluctuation of electricity bills caused by PV output forecasting error and non-optimal market bidding strategy will reach tens of billions of yuan per year. At that time, the commercial returns brought by improving the accuracy of power forecasting by 1% will far exceed the penalty savings of grid assessment.

Thirdly, there would be an increasing number of entities with the demand for power forecasting, which will create additional economic benefits for the forecasting method. For example, except for traditional centralized PV power plants and the independent system operator, who will continue to improve their forecasting accuracy, the deep penetration of distributed PV will also bring incremental necessities for renewable power forecasting.

## IV. CONCLUSION

An ultra-short-term PV power forecasting method combining spatio-temporal correlation with cloud information from the neighboring plant is proposed in this paper. The PV power output has both random and periodic fluctuation due to cloud motion and daily change of sunlight. Thus, more detailed information which can track the volatility should be applied to the forecasting model. The relationship between two PV power time series from the target and neighboring plants is firstly explored which can be transformed and indicated as time lag $\tau$ . Then the mapping relationship is proved between cloud characteristic indexes extracted from neighboring satellite images and the target plant's solar power data. In addition to neighboring cloud information and historical power data, other factors such as time, solar altitude, solar azimuth, power in clear sky conditions are taken into consideration to training the forecasting model by using SVM and GBDT. Then the proposed model can be well-established. To this end, only when the two PV power time series have a strong correlation, and the calculated time lag $\tau$ is corresponding to the geographical direction between the neighboring plant and target plant, the proposed method can be used. Or else, the traditional method is applied to achieve the forecast. To verify the accuracy of this novel method all-sidely, forecast models without cloud information, merely with historical text data of neighboring and target plants are considered to contrast. Data of two target plants named plant 503 and plant 519 are occupied to test this proposed method. Simulation results using actual data show that the proposed method can promote the accuracy of ultra-short-term solar PV power forecasting, and the neighboring cloud information is also suggested to help obtain more accurate forecasting results.

## ACKNOWLEDGEMENT

This work was supported by the Science & Technology project of China Three Gorges Corporation (Research on key technologies of intelligent joint control operation and maintenance of grid-connected-friendly wind and solar storage station group, WWKY-2021-0173), the Management Consulting Project of State Grid Hebei Electric Power Co., Ltd (Research on construction of comprehensive evaluation system and priority dispatch strategy for renewable energy station power forecasting, SGHE0000DKWT2000464), and the National Natural Science Foundation of China (52007092).

## REFERENCES

[1] Elum ZA, Momodu AS. Climate change mitigation and renewable energy for sustainable development in Nigeria: A discourse approach. Renew Sustain Energy Rev 2017;76:72–80. doi: 10.1016/j.rser.2017.03.040.

[2] Armeanu DS, Joldes CC, Gherghina SC, Andrei JV. Understanding the multidimensional linkages among renewable energy, pollution, economic growth and urbanization in contemporary economies: Quantitative assessments across different income countries' groups. Renew Sustain Energy Rev 2021;142:110818. doi: 10.1016/j.rser.2021.110818.

[3] Wang B, Wang Q, Wei YM, Li ZP. Role of renewable energy in China's energy security and climate change mitigation: An index decomposition analysis. Renew Sustain Energy Rev 2018;90:187–94. doi:10.1016/j.rser.2018.03.012.

[4] Li K, Zhang P, Li G, Wang F, Mi Z, Chen H. Day-Ahead Optimal Joint Scheduling Model of Electric and Natural Gas Appliances for Home Integrated Energy Management. IEEE Access 2019;7:133628–40. doi:10.1109/ACCESS.2019.2941238.

[5] Wen L, Zhou K, Yang S, Lu X. Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting. Energy 2019;171:1053–65. doi:10.1016/j.energy.2019.01.075.

[6] Li S, Gong W, Wang L, Yan X, Hu C. Optimal power flow by means of improved adaptive differential evolution. Energy 2020;198:117314. doi:10.1016/j.energy.2020.117314.

[7] Sureshkumar K, Ponnusamy V. Power flow management in microgrid through renewable energy sources using a hybrid modified dragonfly algorithm with bat search algorithm. Energy 2019;181:1166–78. doi:10.1016/j.energy.2019.06.029.

[8] Li K, Wang F, Mi Z, Fotuhi-firuzabad M, Duić N, Wang T. Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation. Appl Energy 2019;253:113595. doi:10.1016/j.apenergy.2019.113595.

[9] Li K, Liu L, Wang F, Wang T, Duić N, Shafie-khah M, et al. Impact factors analysis on the probability characterized effects of time of use demand response tariffs using association rule mining method. Energy Convers Manag 2019;197:111891. doi:10.1016/j.enconman.2019.111891.

[10] Angenendt G, Zurmühlen S, Figgener J, Kairies K, Sauer D. Providing frequency control reserve with photovoltaic battery energy storage systems and power-to-heat coupling. Energy 2020;194:116923. doi:10.1016/j.energy.2020.116923.

[11] Boland J, David M, Lauret P. Short term solar radiation forecasting : Island versus continental sites. Energy 2016;113:186–92. doi:10.1016/j.energy.2016.06.139.

[12] Wang F, Zhang Z, Liu C, Yu Y, Pang S, Duić N, et al. Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting. Energy Convers Manag 2019;181:443–62. doi:10.1016/j.enconman.2018.11.074.

[13] Yuan X, Ji B, Zhang S, Tian H, Chen Z. An improved artificial physical optimization algorithm for dynamic dispatch of generators with valve-point effects and wind power. Energy Convers Manag 2014;82:92–105. doi:10.1016/j.enconman.2014.03.009.

[14] Mohammed A, Pasupuleti J, Khatib T. Simplified performance models of photovoltaic/diesel generator/battery system considering typical control strategies. Energy Convers Manag 2015;99:313–25. doi:10.1016/j.enconman.2015.04.024.

[15] Wang F, Li K, Liu C, Mi Z, et al. Synchronous Pattern Matching Principle-Based Residential Demand Response Baseline Estimation : Mechanism Analysis and Approach Description. IEEE Trans Smart Grid 2018;9:6972–85. doi:10.1109/TSG.2018.2824842.

[16] Wang F, Xuan Z, Zhen Z, Li K, Wang T, Shi M. A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. Energy Convers Manag 2020;212:112766. doi:10.1016/j.enconman.2020.112766.

[17] Wang F, Mi Z, Su S, Zhao H. Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. Energies 2012;5:1355–70. doi:10.3390/en5051355.

[18] Trapero J, Kourentzes N, Martin A. Short-term solar irradiation forecasting based on Dynamic Harmonic Regression. Energy 2015;84:289–95. doi:10.1016/j.energy.2015.02.100.

[19] Gao M, Li J, Hong F, Long D. Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. Energy 2019;187:115838. doi:10.1016/j.energy.2019.07.168.

[20] Zhang C, Du Y, Chen X, et al. Cloud motion tracking system using low-cost sky imager for PV power ramp-rate control. In: 2018 IEEE International Conference on Industrial Electronics for Sustainable Energy Systems (IESES). Hamilton, New Zealand; Jan 2018. p. 493-498. doi:10.1109/ieses.2018.8349927.

[21] Marquez R, Coimbra CFM. Intra-hour DNI forecasting based on cloud tracking image analysis. Sol Energy 2013;91:327–36. doi:10.1016/j.solener.2012.09.018.

[22] Wang F, Zhen Z, Liu C, Mi Z, Hodge B, Shafie-khah M, et al. Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting. Energy Convers Manag 2018;157:123–35. doi:10.1016/j.enconman.2017.11.080.

[23] Zhen Z, Xuan Z, Wang F, Sun R, Dui N, Jin T. Image phase shift invariance based multi-transform-fusion method for cloud motion displacement calculation using sky images. Energy Convers Manag 2019;197:111853. doi:10.1016/j.enconman.2019.111853.

[24] Li M, Chu Y, Pedro HTC, Coimbra CFM. Quantitative evaluation of the impact of cloud transmittance and cloud velocity on the accuracy of short-

term DNI forecasts. Renew Energy 2016;86:1362–71. doi:10.1016/j.renene.2015.09.058.

[25] Alonso-montesinos J, Batlles F, Portillo C. Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images. Energy Convers Manag 2015;105:1166–77. doi:10.1016/j.enconman.2015.09.001.

[26] Marquez R, Pedro HTC, Coimbra CFM. Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs. Sol Energy 2013;92:176–88. doi:10.1016/j.solener.2013.02.023.

[27] Dong Z, Yang D, Reindl T, Walsh WM. Satellite image analysis and a hybrid ESSS/ANN model to forecast solar irradiance in the tropics. Energy Convers Manage 2014;79:66–73. doi:10.1016/j.enconman.2013.11.043.

[28] Boata RS, Gravila P. Functional fuzzy approach for forecasting daily global solar irradiation. Atmos Res 2012;112:79–88. doi:10.1016/j.atmosres.2012.04.011.

[29] David M, Ramahatana F, Liandrat O. Spatial and temporal variability of PV output in an insular grid : Case of Reunion Island. Energy Procedia 2014;57:1275–82. doi:10.1016/j.egypro.2014.10.117.

[30] Zhang B, Dehghanian P. Spatial-Temporal Solar Power Forecast through Use of Gaussian Conditional Random Fields. 2016 IEEE Power Energy Soc Gen Meet, vol. 2, 2016. p. 16–20. doi:10.1109/pesgm.2016.7741503.

[31] Jamaly M, Kleissl J. Spatiotemporal interpolation and forecast of irradiance data using Kriging. Sol Energy 2017;158:407–23. doi:10.1016/j.solener.2017.09.057.

[32] Zhang R, Ma H, Hua W, Kumar T. Data-Driven Photovoltaic Generation Forecasting based on Bayesian Network with Spatial-Temporal Correlation Analysis. IEEE Trans Ind Informatics 2019;16(3):1635-1644. doi:10.1109/TII.2019.2925018.

[33] Wai C, Urquhart B, Lave M, Dominguez A, Kleissl J, Shields J, et al. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. Sol Energy 2011;85:2881–93. doi:10.1016/j.solener.2011.08.025.

[34] Yang C, Xie L. A Novel ARX-based Multi-scale Spatio-temporal Solar Power Forecast Model. 2012 North American Power Symposium (NAPS); 2012. p. 1–6. doi:10.1109/naps.2012.6336383.

[35] Yang C, Thatte A, Xie L. Multitime-Scale Data-Driven Spatio-Temporal Forecast of Photovoltaic Generation. IEEE Trans Sustain Energy 2014;6(1): 104–112. doi:10.1109/TSTE.2014.2359974.

[36] Wang F, Zhen Z, Wang B, Mi Z. Comparative Study on KNN and SVM Based Weather Classification Models for Day Ahead Short Term Solar PV Power Forecasting. Appl Sci 2017;8(1):28. doi:10.3390/app8010028.

[37] Akinci TC, Nogay HS. Wind Speed Correlation Between Neighboring Measuring Stations. Arab J Sci Eng 2012;37(4):1007-1019. doi:10.1007/s13369-012-0223-4.

[38] Shi J, Lee W, Liu Y, Yang Y, Wang P. Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines. IEEE Trans Ind Appl 2012;48(3):1064–1069. doi:10.1109/TIA.2012.2190816.

[39] Wen Z, Shi J, He B, Chen J, Ramamohanarao K, Li Q. Exploiting GPUs for Efficient Gradient Boosting Decision Tree Training. IEEE Trans Parallel Distrib Syst 2019;30:2706–17.

[40] Ma J, Cheng J. Identification of the numerical patterns behind the leading counties in the U.S local green building markets using data mining. J Clean Prod 2017;151:406–418. doi:10.1016/j.jclepro.2017.03.083.

[41] Jachan M, Matz G, Hlawatsch F. Time-Frequency ARMA Models and Parameter Estimators for Underspread Nonstationary Random Processes. IEEE Trans Signal Process 2007;55:4366–81. doi:10.1109/tsp.2007.896265.

[42] Balci M, Foroosh H. Subpixel estimation of shifts directly in the Fourier domain. IEEE Trans Image Process 2006;15:1965–72. doi:10.1109/TIP.2006.873457.