Nonparametric Probabilistic Prediction of Regional PV Outputs Based on Granule-based Clustering and Direct Optimization Programming

Yonghui Sun, Yan Zhou, Sen Wang, Rabea Jamil Mahfoud, Hassan Haes Alhelou, *Senior Member*, *IEEE*, George Sideratos, Nikos Hatziargyriou, *Fellow Member*, *IEEE*, and Pierluigi Siano, *Senior Member*, *IEEE*

***Abstract*—Regional photovoltaic (PV) power prediction plays an important role in power system planning and operation. To effectively improve the performance of prediction intervals (PIs) for very short-term regional PV outputs, an efficient nonparametric probabilistic prediction method based on granule-based clustering (GC) and direct optimization programming (DOP) is proposed. First, GC is proposed to formulate and cluster the sample granules consisting of numerical weather prediction (NWP) and historical regional output data, for the enhanced hierarchical clustering performance. Then, to improve the accuracy of samples' utilization, an unbalanced extension is used to reconstruct the training samples consisting of power time series. After that, DOP is applied to quantify the output weights based on the optimal overall performance. Meanwhile, a balance coefficient is studied for the enhanced reliability of PIs. Finally, the proposed method is validated through multi-step PIs based on the numerical comparison of real PV generation data.**

***Index Terms*—Regional photovoltaic outputs, prediction intervals, granule-based clustering, direct optimization programming, nonparametric probabilistic prediction.**

I. Introduction

NOWADAYS, the high penetration and inherent variability of renewable energy generation introduce significant challenges to the power system operation [1]-[5]. Accurate prediction of renewable energy generation variations allows timely adjustment of the dispatching schedules [6]-[8], so as to reduce system reserves and consequent operational costs [9].

Several studies concerning probabilistic predictions of renewable generation have been done in the past years [10], [11]. These probabilistic predictions are very important for the quantitative prediction of photovoltaic (PV) system generation, similar to wind power generation [12], [13]. In [14] and [15], parametric prediction intervals (PIs) based on the error assumption of normal distribution and deterministic prediction were studied. In [16], a multi-model approach was studied via a combination of parametric and nonparametric PIs. Even though PV generation has high randomness and uncertainty, which brings difficulties in accurately assuming the error distribution, certain periodicity helps improve the utilization of training samples according to the specific type of weather [17]. Based on weather information such as cloud coverage, humidity, and solar irradiance, typical days of samples were classified as sunny, cloudy, and rainy days in [18]. Using the specific types of weather modeling, the performance of PIs was greatly improved. In [19], the prediction errors of PV generation were proven to be unable to completely satisfy assumed probability distributions such as Beta and Gaussian. Therefore, nonparametric methods are of great significance for the probabilistic prediction of PV generation. In [20], an efficient nonparametric PI approach was proposed based on extreme learning machine (ELM) and quantile regression (QR) for PV generation, achieving high reliability and efficient computation. In [21], a novel machine learning based linear programming (MLLP) approach was proposed, which considered both reliability and sharpness. The performance of PIs was evaluated by both reliability and overall performance considering sharpness [15]. Among them, the overall performance is a decisive criterion, the reliability is an important observation criterion, and the sharpness is an auxiliary criterion. The PI constructions of the

Manuscript received: September 13, 2022; revised: November 28, 2022; accepted: January 20, 2023. Date of CrossCheck: January 20, 2023. Date of online publication: March 16, 2023.

This work was supported by the National Natural Science Foundation of China (No. 62073121), the National Key R&D Program of China "Technology and application of wind power/photovoltaic power prediction for promoting renewable energy consumption" (No. 2018YFB0904200), and eponymous Complement S&T Program of State Grid Corporation of China (No. SGLNDKOOKJJS1800266).

Y. Sun (corresponding author), Y. Zhou, and S. Wang are with the College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China (e-mail: sunyonghui168@gmail.com; zhouyanxyzz@126.com; senwang@hhu.edu.cn).

R. J. Mahfoud is with the College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China (e-mail: rabea7mahfoud@hotmail.com).

H. H. Alhelou is with the Department of Electrical Engineering, Tishreen University, Latakia 2230, Syria (e-mail: h.haesalhelou@gmail.com).

G. Sideratos and N. Hatziargyriou are with the National Technical University of Athens, Athens 15773, Greece (e-mail: joesider@power.ece.ntua.gr; nh@power.ece.ntua.gr).

P. Siano is with the Department of Management & Innovation Systems, University of Salerno, Salerno 84084, Italy (e-mail: psiano@unisa.it).

DOI: 10.35833/MPCE.2022.000577

conventional nonparametric methods mainly focus on reliability or the combination of reliability and sharpness. It is worth mentioning that the overall performance of PIs depends not only on reliability and sharpness but also on the offsets of points outside PIs. Thus, to improve the forecasting performance considering overall performance and reliability, direct optimization programming (DOP) is proposed in this paper by directly optimizing the cost function of the overall performance criterion of PIs using a simple linear programming (LP) method.

Recently, lots of clustering techniques have been studied to improve the accuracy of samples' utilization. In [22], input data were clustered based on weather stability, the uncertainty obtained by a deterministic prediction model, and the uncertainty defined by several numerical weather prediction (NWP) updates using self-organized maps (SOMs). Using the above clustering technique, radial basis function neural networks (RBFNNs) could provide quantile predictions with high reliability. In [23], a novel RBFNN clustered the input samples based on their variable importance and significantly improved the forecasting accuracy of PV power. In [24], SOM divided the data into three nonlinear parts of the wind power curve. In [25], an improved fuzzy C-means (FCM) clustering algorithm was proposed to improve the accuracy. In [26], a clustering-based prediction method using weather forecast and historical power data was studied for regional PV power. In [27], the hierarchical clustering-based prediction method was studied. In this application, after each iterative calculation of hierarchical clustering, the centers of clusters, quantified by calculating the average of inputs in each cluster, were shifted. Besides, the quantification of cluster centers is affected by the outliers. These, however, may cause inaccurate clustering. Moreover, the hierarchical clustering only considered the distance of samples, neglecting analysis of variance. In [28] and [29], information granule-based neural networks (IGNNs) were designed to study the deterministic prediction of time-series data with the clustering of training samples. The experimental results showed that neural networks constructed on a basis of information granules (IGs) improved the forecasting performance in an efficient manner and produced meaningful estimates. In [30], based on IGNN, a novel approach of optimal granule-based PIs (OGPIs) was applied for enhanced forecasting performance by segmenting the power time series into granules to capture the variability. The numerical comparison revealed the effectiveness of granular computation to reduce the adverse effect caused by the volatility of high-resolution (1-min resolution) power time series. However, the IGs directly segmenting the power time series into granules will reduce the number of training samples, which may be infeasible due to the potential insufficiency of training samples under a typical season and weather condition with a resolution of 15 min or 1 hour. Inspired by the PI construction of IG, a granule-based clustering (GC) approach and the construction of related sample granule are proposed in this paper for the improvement of hierarchical clustering to consider the variance of samples and reduce the adverse impact of cluster center shifting and outliers.

The proposed sample granule represented by a matrix composed of several clustering samples is studied to improve the clustering performance, and is constructed based on the iterative quantifications of mergence and division processes for the training samples. Each row vector in the matrix represents a clustering sample consisting of PV power time series as prediction model input and solar irradiance from NWP. After the granulation process of clustering samples, the sample granules are clustered. Then, the sample granules are restored to original samples, of which the PV power data are used as prediction model input for training with unbalanced extension of training samples rather than the conventional clustering methods that only study the high correlation samples or directly segment the time series into granules. The objective of the proposed GC and unbalanced extension is to improve the training performance and avoid the potential lack of training samples for probabilistic prediction.

Moreover, in recent years, several research works have been devoted to regional power prediction. In [31], the statistical upscaling method was utilized to predict the very short-term PIs of regional PV power generation. The performance of the prediction model was affected by the accuracy of subregion division and the selection of representative PV stations. In [32], a prediction model for regional power output was proposed, considering the smoothing effect. The effectiveness of the smoothing method by weighting the historical regional wind power output was proved in [33]. From the perspective of regional power output, the influence of output fluctuations of local wind farms is ameliorated and the aggregated regional power output tends to be smooth, compared with the power output of a single wind farm. Similarly, the regional PV generation based analysis can reduce sudden changes of outputs and the impact of scattered clouds on the uncertainty of PV generation.

So far, with the increasing demands for effective prediction technologies, the significance of probabilistic predictions for regional PV power generation has also increased. In this paper, a novel nonparametric PI method for very short-term regional PV power generation is proposed. The main contributions of this paper are as follows.

1) GC is proposed in this paper, which can construct sample granules to improve the clustering performance. It reduces the adverse impacts caused by the iterative calculation, including shifted cluster centers and outliers, and provides reasonable analysis of sample variance for the enhanced clustering performance.

2) After unbalanced extension for accurate utilization of all training samples, DOP is proposed to improve the overall performance of PIs based on the criterion of interval score by efficient LP. Moreover, a balance coefficient is proposed to enhance the reliability of PIs for better robustness.

The rest of the paper is organized as follows. In Section II, the methodology of the proposed method is described. The construction of PIs is presented in Section III. Case studies are performed to verify the performance of the proposed method in Section IV. Finally, conclusions are drawn in Section V.

II. Methodology of Proposed Method

In this section, the proposed GC is presented first. Then, the training samples are extended with unbalanced multiples by different weights to enhance the utilization of samples. Finally, DOP is proposed to obtain the output weights for optimal overall performance and reliability of PIs, considering the balance coefficient.

A. GC Theory

In the proposed GC, training samples are divided into sample granules based on their significant differences, which can reduce the deviation in the clustering process. Then, hierarchical clustering is utilized for the sample granules. The diagram of the proposed GC consisting of three stages is shown in Fig. 1, where g_G denotes the number of the existing sample granules. First, the data are processed in stage 1. Then, in stage 2, the sample granulation process, including the division process and the mergence process, is quantified. Finally, in stage 3, hierarchical clustering is utilized for the sample granules.

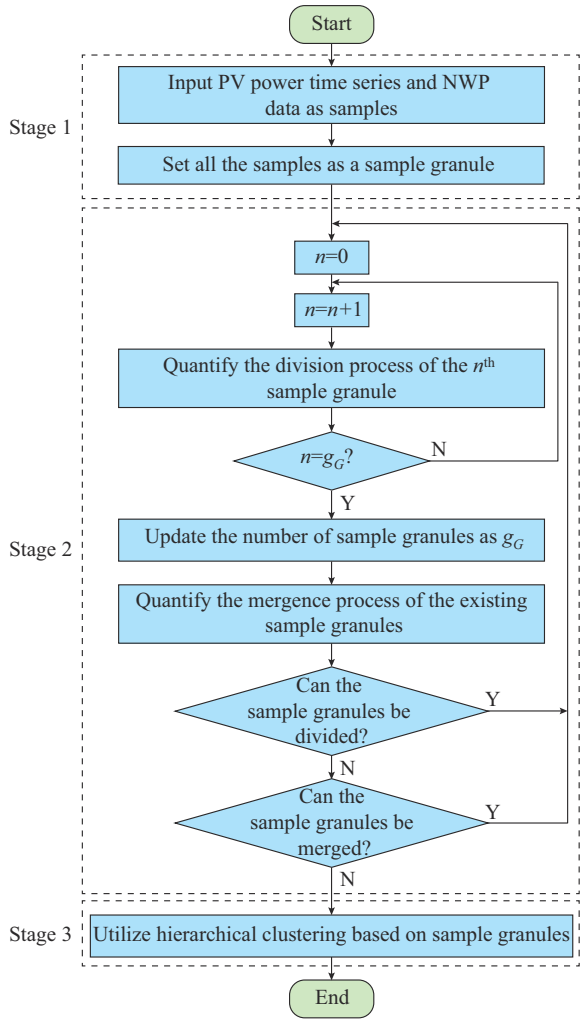

Fig. 1. Diagram of proposed GC consisting of three stages.

1) Stage 1: Data Preprocessing

Based on the time series of historical PV power generation and solar irradiance from NWP, the set of training samples is formed. Besides, considering that irradiance has a direct correlation with PV generation [18], its forecasting data are helpful in enhancing the clustering performance. Based on the capacity and solar irradiance forecast of each PV station, the regional composite irradiance is defined as:

$$R_c = \frac{\sum_{i=1}^{M} C_i I_i}{\sum_{i=1}^{M} C_i} \tag{1}$$

where R_c is the value of regional composite irradiance; M is the number of PV stations; and C_i and I_i are the capacity and the irradiance of the i^{th} PV station, respectively.

2) Stage 2: Sample Granulation Process

Sample granules consisting of regional PV power observations and regional composite irradiances are studied, and their input variables are in turn utilized as analysis variables. At each iteration, sample granules are divided or merged based on the Wikis likelihood criterion for optimal critical values of variables [34]. Each sample granule with a significant difference is divided into two new granules, while the sample granules with no significant difference are merged. The values of F statistic are utilized to determine whether the sample granules need to be divided or merged. This process goes on until all sample granules can no longer be divided or merged.

The F statistic of two sample granules $\boldsymbol{e}$ and $\boldsymbol{h}$ is quantified by:

$$F(d, n_e + n_h - d - 1) = \frac{1-\Lambda}{\Lambda} \frac{n_e + n_h - d - 1}{d} \tag{2}$$

$$\Lambda = \frac{|A|}{|A+B|} \tag{3}$$

$$A = \sum_{i=1}^{n_e} (\boldsymbol{e}_i - \bar{\boldsymbol{e}})^{\mathrm{T}} (\boldsymbol{e}_i - \bar{\boldsymbol{e}}) + \sum_{h=1}^{n_h} (\boldsymbol{h}_j - \bar{\boldsymbol{h}})^{\mathrm{T}} (\boldsymbol{h}_j - \bar{\boldsymbol{h}}) \tag{4}$$

$$B = \frac{n_e n_h}{n_e + n_h} (\bar{\boldsymbol{e}} - \bar{\boldsymbol{h}})^{\mathrm{T}} (\bar{\boldsymbol{e}} - \bar{\boldsymbol{h}}) \tag{5}$$

$$\bar{\boldsymbol{e}} = \frac{1}{n_e} \sum_{i=1}^{n_e} \boldsymbol{e}_i \tag{6}$$

$$\bar{\boldsymbol{h}} = \frac{1}{n_h} \sum_{j=1}^{n_h} \boldsymbol{h}_j \tag{7}$$

where F is the value of F statistic; Λ is the auxiliary variable, and the smaller the value of Λ, the larger the difference between the sample granules; A is the sum of squares of differences within groups; B is the sum of cross-product matrix; $\boldsymbol{e}_i$ and $\boldsymbol{h}_j$ are the i^{th} and j^{th} row vectors of $\boldsymbol{e}$ and $\boldsymbol{h}$, respectively; $\bar{\boldsymbol{e}}$ and $\bar{\boldsymbol{h}}$ are the mean vectors, which represent the centers of sample granules $\boldsymbol{e}$ and $\boldsymbol{h}$, respectively; n_e and n_h are the numbers of samples of $\boldsymbol{e}$ and $\boldsymbol{h}$, respectively; and d is the number of variables in each sample. $F_{1-\alpha_F}$ is the critical value of F statistic that determines if the difference of two sample granules is significant, which represents the value of $F_{1-\alpha_F}(d, n_e + n_h - d - 1)$, and $1-\alpha_F$ is the predetermined confidence. If $F \geq F_{1-\alpha_F}$, there is a significant difference be-

tween the sample granules. In the division process, the sample granules with submatrices of significant differences are divided, while the sample granules with insignificant differences are merged.

Based on (2)-(7), the division process is as follows.

Step 1: for each variable, sort the clustering samples according to their values. For n_e varying from 1 to $n_{e+h}-1$, where n_{e+h} is the number of samples of the original sample granule, quantify the values of Λ.

Step 2: obtain the minimum Λ and the corresponding critical values $F_{1-\alpha_F}$ and F. If $F \geq F_{1-\alpha_F}$, the granule needs to be divided into two new granules.

The detailed division process of sample granules is shown in Fig. 2, which aims to determine if the submatrices with the maximum F statistic value $F_{\max}$ need to be divided.

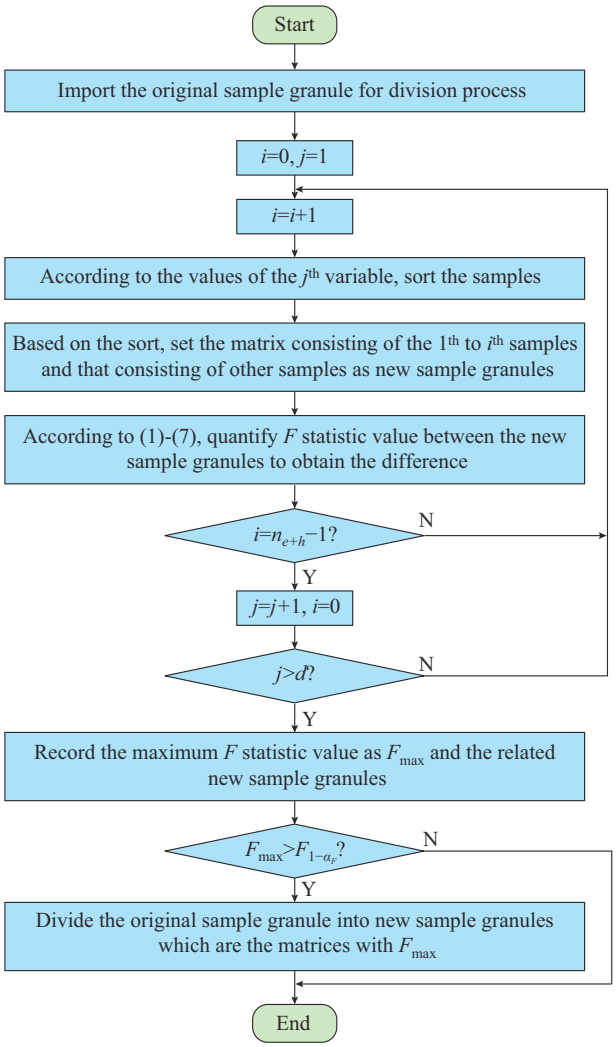

Fig. 2. Diagram of division process of sample granules.

Figure 3 shows the mergence process of sample granules, where $F_{\min}$ represents the minimum F statistic value. After each iteration of the division process, the existing sample granules are quantified in pairs with respect to the minimum Λ and the corresponding value of F, to determine whether the sample granules need to be merged.

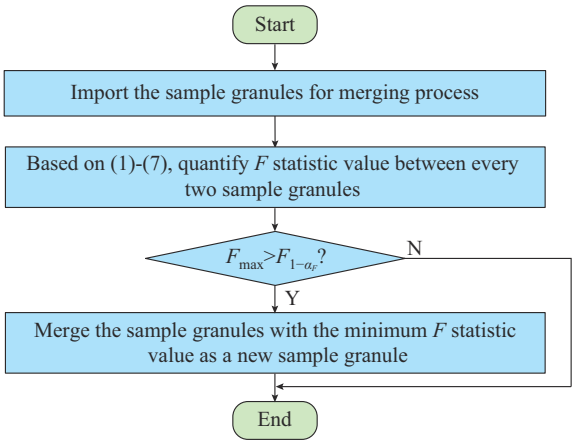

Fig. 3. Diagram of merging process of sample granules.

3) Stage 3: Hierarchical Clustering

Based on the result of granulation process, the centers of granules can be obtained by quantifying the mean value of the clustering samples. Then, the hierarchical clustering approach [27] is utilized to cluster the centers of sample granules. Different from the utilization of granules as inputs of the neural network in [30], the proposed granulation process is only used for enhanced clustering, not for prediction model training. After clustering, the sample granules are restored to the original samples, and then the PV output data are used for model training. In the proposed method, clustering samples are used for GC, while training samples are used for model training, given by:

$$\boldsymbol{St}_i = \{\boldsymbol{Sc}_i, R_i\} \tag{8}$$

where $\boldsymbol{St}_i$ and $\boldsymbol{Sc}_i$ denote the i^{th} clustering sample and training sample, respectively; and R_i is the corresponding regional composite irradiance.

B. Unbalanced Extension of Samples

The clustering method is usually utilized to select the samples with high correlation for training to improve the forecasting performance [8]. However, for the probabilistic prediction model, the removal of samples through screening will result in a certain loss of training sample information. In order to improve the accuracy of training samples' utilization, based on the clustering results, the unbalanced extension of samples is used to process the training samples. Based on the distances of clusters, the similarity of clusters, which is set as the coefficient of unbalanced extension, is quantified by:

$$Sim_i = \exp\left(-\sqrt{\sum_{j=1}^{d}(\bar{p}_j - \bar{c}_{i,j})^2}\right) \tag{9}$$

where Sim_i denotes the similarity between the cluster to be predicted and the i^{th} cluster of training samples; $\bar{p}_j$ denotes the j^{th} variable of the center of the cluster to be predicted; and $\bar{c}_{i,j}$ denotes the j^{th} variable of $\bar{c}_i$ which denotes the i^{th} cluster center.

In the proposed method, a cluster center is defined by:

$$\bar{g}_j = \frac{1}{T_g}\sum_{i=1}^{T_g} g_{i,j} \tag{10}$$

where $\bar{g}_j$ and $g_{i,j}$ denote the j^{th} variables of the cluster center and the i^{th} sample, respectively; and T_g denotes the number of samples.

The extension multiple to the samples from the cluster to be predicted is defined by:

$$E_i = \lfloor Sim_i/\varepsilon \rfloor + 1 \quad i = 1, 2, ..., N \tag{11}$$

where E_i denotes the extension multiple of samples for the i^{th} cluster; ε is the partition coefficient; $\lfloor\ \rfloor$ denotes the function of rounding down the number; and N denotes the number of clusters.

The unbalanced extension is applied to adjust the effects of sample clusters on the testing samples, rather than simply removing clusters of samples with low correlations. The training samples $\boldsymbol{S}_{tr}$ for a sample cluster to be predicted are reconstructed by:

$$\boldsymbol{S}_{tr} = \{\underbrace{\boldsymbol{S}_1, ..., \boldsymbol{S}_1}_{E_1}, \underbrace{\boldsymbol{S}_2, ..., \boldsymbol{S}_2}_{E_2}, ..., \underbrace{\boldsymbol{S}_N, ..., \boldsymbol{S}_N}_{E_N}\} \tag{12}$$

$$\boldsymbol{S}_i = \{\boldsymbol{S}_{i,1}, \boldsymbol{S}_{i,2}, ..., \boldsymbol{S}_{i,T}\} \quad i = 1, 2, ..., N \tag{13}$$

where $\boldsymbol{S}_i$ denotes all training samples of the i^{th} cluster, and its extension multiple is represented by E_i; $\boldsymbol{S}_{i,1}, \boldsymbol{S}_{i,2}, \ldots, \boldsymbol{S}_{i,T}$ denote the 1^{th} to T^{th} samples of the i^{th} cluster; and T is the number of training samples.

C. DOP

The reliability of PIs is evaluated by the average coverage error (ACE), which is the accuracy of PI coverage probability (PICP) according to PI nominal confidence (PINC) [35], given by:

$$|ACE| = |PICP - PINC| \tag{14}$$

The sharpness of PIs is given by:

$$\delta_t^\alpha(\boldsymbol{x}_i) = U_t^\alpha(\boldsymbol{x}_i) - L_t^\alpha(\boldsymbol{x}_i) \tag{15}$$

where $\boldsymbol{x}_i$ is the i^{th} input sample; $\delta_t^\alpha(\boldsymbol{x}_i)$ is the width of the i^{th} interval; $U_t^\alpha(\boldsymbol{x}_i)$ and $L_t^\alpha(\boldsymbol{x}_i)$ are the i^{th} upper and lower bounds as the prediction targets, respectively; and α denotes the PINC. To evaluate the overall performance of PIs, the interval score is utilized [14], [20], [30], [36]-[38] and formulated as:

$$Sc_t^\alpha(\boldsymbol{x}_i) = \begin{cases} -2(1-\alpha)\delta_t^\alpha(\boldsymbol{x}_i) - 4(L_t^\alpha(\boldsymbol{x}_i) - t_i) & t_i < L_t^\alpha(\boldsymbol{x}_i) \\ -2(1-\alpha)\delta_t^\alpha(\boldsymbol{x}_i) & t_i \in I_t^\alpha(\boldsymbol{x}_i) \\ -2(1-\alpha)\delta_t^\alpha(\boldsymbol{x}_i) - 4(t_i - U_t^\alpha(\boldsymbol{x}_i)) & t_i > U_t^\alpha(\boldsymbol{x}_i) \end{cases} \tag{16}$$

$$Sc_t^\alpha = \frac{1}{T_p}\sum_{i=1}^{T_p} Sc_t^\alpha(\boldsymbol{x}_i) \tag{17}$$

where $I_t^\alpha(\boldsymbol{x}_i)$ denotes the i^{th} PI; t_i denotes the i^{th} prediction target; T_p is the number of testing samples; $Sc_t^\alpha(\boldsymbol{x}_i)$ is the score of the i^{th} point; and Sc_t^α is the interval score.

DOP is proposed to directly quantify the output weights of the optimal overall score. In (16), the interval scores for three cases, including the actual points above, below, and inside the intervals, are expressed. The interval score is negative, and the closer to zero its value is, the better the interval overall performance is. Equations (16) and (17) can be modified as:

$$\left|S_Q\right| = \frac{1}{T}\sum_{i=1}^{T}(-2\alpha\delta_t^\alpha(\boldsymbol{x}_i) + 2D_i) \tag{18}$$

$$D_i = \begin{cases} \left|U_t^\alpha(\boldsymbol{x}_i) - t_i\right| + \left|L_t^\alpha(\boldsymbol{x}_i) - t_i\right| & t_i \notin I_i \\ U_t^\alpha(\boldsymbol{x}_i) - L_t^\alpha(\boldsymbol{x}_i) & t_i \in I_i \end{cases} \tag{19}$$

where $\left|S_Q\right|$ denotes the absolute value of interval score S_Q, and the smaller value of $\left|S_Q\right|$ is, the better the overall performance is; and D_i denotes the sum of distances between the i^{th} observation and the i^{th} bound of PI. When the actual value of power is within the PI, D_i is the width of the interval. Otherwise, the value of D_i is quantified based on the width of the i^{th} interval and the offset of the point outside PI. The overall performance of PIs is determined not only by the accuracy of coverage and the width of PIs, but also by the offsets of points outside PIs. Thus, the average offset (AO) of points outside PIs, considered by the interval score, is utilized to observe the performance, which is defined by:

$$AO = \frac{1}{N_o}\sum_{i=1}^{N_o} O_i \tag{20}$$

$$O_i = \begin{cases} t_i - U_t^\alpha(\boldsymbol{x}_i) & t_i > U_t^\alpha(\boldsymbol{x}_i), i = 1, 2, ..., N_o \\ L_t^\alpha(\boldsymbol{x}_i) - t_i & t_i < L_t^\alpha(\boldsymbol{x}_i), i = 1, 2, ..., N_o \end{cases} \tag{21}$$

where O_i denotes the offset of the i^{th} interval; and N_o denotes the number of points outside PIs.

Based on (18), the cost function of LP for optimal overall performance with PINC of α is given by:

$$\min_{\overline{\xi}_i, \underline{\xi}_i, \boldsymbol{\beta}_U, \boldsymbol{\beta}_L} \sum_{i=1}^{T}\left[-2\alpha k(f(\boldsymbol{x}_i, \boldsymbol{\beta}_U) - f(\boldsymbol{x}_i, \boldsymbol{\beta}_L)) + 2\overline{\xi}_i + 2\underline{\xi}_i\right] \tag{22}$$

s.t.

$$f(\boldsymbol{x}_i, \boldsymbol{\beta}_U) - f(\boldsymbol{x}_i, \boldsymbol{\beta}_L) \geq 0 \tag{23}$$

$$\begin{cases} 0 \leq f(\boldsymbol{x}_i, \boldsymbol{\beta}_U) \leq 1 \\ 0 \leq f(\boldsymbol{x}_i, \boldsymbol{\beta}_L) \leq 1 \end{cases} \tag{24}$$

$$-\overline{\xi}_i \leq t_i - f(\boldsymbol{x}_i, \boldsymbol{\beta}_U) \leq \overline{\xi}_i \tag{25}$$

$$-\underline{\xi}_i \leq t_i - f(\boldsymbol{x}_i, \boldsymbol{\beta}_L) \leq \underline{\xi}_i \tag{26}$$

$$\begin{cases} \overline{\xi}_i \geq 0 \\ \underline{\xi}_i \geq 0 \end{cases} \tag{27}$$

where $f(\cdot)$ denotes the output function of ELM; $\boldsymbol{\beta}_U$ and $\boldsymbol{\beta}_L$ denote the output weights of upper bound $f(\boldsymbol{x}_i, \boldsymbol{\beta}_U)$ and lower bound $f(\boldsymbol{x}_i, \boldsymbol{\beta}_L)$, respectively; $\overline{\xi}_i$ denotes the distance between the upper bound and the i^{th} actual value; $\underline{\xi}_i$ denotes the distance between the lower bound and the i^{th} actual value; and k is the balance coefficient for the reliability and overall performance, which is selected to avoid cases, where the interval score is optimal but the reliability is low. Similar to direct quantile regression (DQR) [34], the proposed DOP formulates the model input and output based on ELM, which has an extremely fast learning speed. Besides, the ELM-based method overcomes the limitations of conventional neural networks, such as overtraining, local minimum,

and high computational burden. ELM becomes a linear system after the hidden layer, which motivates its effective integration with the LP-based approaches, and compared with other neural network methods including deep learning methods, it is not affected by the iterative calculation of the training process.

III. Construction of PIs

In this section, the construction of PIs is presented. Since GC, unbalanced extension, and DOP have been analyzed, the steps of the proposed method for construction of PIs are briefly summarized as follows.

Step 1: the parameters are initialized, and the dataset is imported after normalization.

Step 2: the clustering samples consisting of regional PV output observations and solar irradiance data from NWP are constructed and clustered based on the proposed GC.

Step 3: after GC, the unbalanced extension of samples consisting of regional power time series is formulated according to the similarity coefficients between the centers of the specified cluster and other clusters.

Step 4: the output weights of PIs are quantified based on DOP, by successively assuming that the testing samples belong to each cluster.

Step 5: according to the cluster labels of testing samples and the output weights, the bounds of PIs can be quantified.

In the quantification of sample granules, considering solar irradiance data further improves the clustering performance of the future trend of power time series. It should be noted that the solar irradiance data will not be used in model inputs while training and testing. That is, the input of a training or testing sample only consists of PV power time series due to the inadequate prediction accuracy of NWP for very short-term prediction. In the proposed method, the hyperparameters of GC and unbalanced extension, including the partition coefficient ε, the confidence a_F, and the cluster number of GC, are selected mainly based on the prior tests of training samples, considering the computational efficiency and forecasting performance. Here, the above hyperparameters can be obtained according to the forecasting performance of deterministic or probabilistic prediction. To ensure the consistency of hyperparameter selection in the numerical comparisons of deterministic and probabilistic predictions, the hyperparameters of GC and unbalanced extension are optimized according to the deterministic prediction accuracy.

Remark 1: the output prediction performance of a PV station is easily affected by scattered clouds which can block solar radiation [26]. However, the overall analysis of regional PV power generation and composite irradiance is not sensitive to a small number of floating clouds, so it is of significance to study.

Remark 2: in conventional hierarchical clustering [27], Euclidean distance which only quantifies the samples' difference is the criterion for clustering. Besides, after each iterative calculation of hierarchical clustering, the centers of clusters will be shifted. Those shifts with high frequencies and outliers reduce the accuracy. Therefore, in the proposed GC, a granulation process is used to classify the samples before their hierarchical clustering, considering the variance of samples and reducing the frequency of cluster centers' movement and the adverse effects of outliers.

Remark 3: in the preprocessing of training samples, clustering is usually applied to remove the samples with low correlation [14] and improve the accuracy of deterministic prediction. Different from deterministic prediction, probabilistic prediction requires all samples to be considered to exploit the potential information of probability. Removing some samples can result in a shortage of training sample information for probabilistic analysis. Hence, the unbalanced extension of samples is important to improve the accuracy of samples' utilization.

Remark 4: the LP of nonparametric PIs is applied to quantify the output weights based on reliability and sharpness [21]. In the proposed method, DOP considering the optimal interval score including AO is utilized to quantify the output coefficients. In addition, the balance coefficient is used to improve the robustness of the proposed method.

IV. Case Studies

A. Introduction of Dataset

To verify the effectiveness of the proposed method, two datasets are considered. In each dataset, the data on the days with high irradiance, humidity, or cloud coverage are selected to form a sample group with the weather condition of sunny days, rainy days, or cloudy days, respectively. The datasets are given as follows.

1) Dataset 1: this dataset consists of the output and NWP data of 20 PV stations with 15-min resolution in the northeast of China in March-June 2019. And each sample group consists of 30-day data, of which 9-day data are used for testing and the rest data are used for training.

2) Dataset 2: this dataset consists of the output and NWP data of 45 PV stations with 1-hour resolution in October-February of the Global Energy Forecasting Competition 2014 (GEFCom2014) [39]. And each sample group consists of 90-day data, of which 30-day data are used for testing and the rest data are used for training.

The time series of regional PV power and irradiance data are used as clustering samples, and considering the prediction accuracy and correlation of NWP, only the time series of PV power are utilized as sample inputs for PIs of very short-term forecasting [15]. The time horizon to be predicted for Dataset 1 is 07:00-17:00, and the daytime data are used for Dataset 2. The generalization performance of ELM is stable on a wide range of numbers of hidden nodes [40]. Furthermore, the lengths of input data are 8 and 4 for Datasets 1 and 2, respectively. Datasets 1 and 2 are used after normalization, according to the installed capacities and the maximum historical observations, respectively.

B. Numerical Comparison of Clustering Methods

To verify the effectiveness of the proposed GC, *K*-means-based method [41], hierarchical clustering-based method (CM) [27], and IGNN [28] are utilized as benchmarks for the numerical comparison of deterministic prediction perfor-

mances based on the efficient ELM. The root mean squared errors (RMSEs) and mean absolute errors (MAEs) of deterministic predictions [15] on sunny days, rainy days, and cloudy days are shown respectively from different datasets to reveal the effectiveness of the proposed GC for enhanced performance.

Tables I and II reveal the numerical comparisons covering March to June from Dataset 1 and October to February from Dataset 2, respectively, of which the look-ahead time for deterministic prediction is 1-hour. The hyperparameters of all the methods are selected according to the prediction accuracy while training. Meanwhile, the irradiance data from NWP are considered with PV output observations for clustering, and only the PV output observations are utilized for model training of the neural network.

TABLE I
PREDICTION ERRORS BASED ON DATASET 1

Method	Sunny days		Rainy days		Cloudy days	
	MAE (%)	RMSE (%)	MAE (%)	RMSE (%)	MAE (%)	RMSE (%)
K-means	3.11	4.28	6.31	7.93	7.48	9.53
Hierarchical CM	4.10	5.44	6.15	7.51	7.76	9.36
IGNN	3.70	4.76	5.79	7.17	6.71	9.24
GC	2.58	3.44	5.19	6.57	5.78	7.31

TABLE II
PREDICTION ERRORS BASED ON DATASET 2

Method	Sunny days		Rainy days		Cloudy days	
	MAE (%)	RMSE (%)	MAE (%)	RMSE (%)	MAE (%)	RMSE (%)
K-means	5.43	9.36	8.19	11.13	7.85	12.36
Hierarchical CM	5.70	7.71	8.13	11.22	7.88	12.21
IGNN	5.41	7.42	8.08	10.07	7.22	9.55
GC	4.71	6.83	7.27	9.42	6.71	8.17

According to the comparison result, in most cases, the prediction accuracy of IGNN is better than that of *K*-means based method and hierarchical CM. However, segmenting the power time series into IGs tends to result in potential insufficient training data. For very short-term forecast of PV output, the output observations need to be classified and used according to season and weather conditions, and each classification roughly covers 1 to 3 months to ensure the accurate use of training data. If the dataset spans too many days, it will lead to confusion about different seasonal characteristics which reduces the performance of model training, while if the dataset spans only few days, it will lead to poor training effect of the prediction model. The above dataset application of PV output for very short-term forecast reveals the need for efficient and comprehensive training sample construction. Hence, to ensure the sufficient and precise training data, the proposed sample granules are restored to the original samples with reasonable multiples rather than directly studying the sample granules, similar to the conventional utilization of IGs. Besides, with iterative quantifications of mergence and division processes, the proposed GC tends to find out the differences between the samples or sample granules, which can further improve the clustering performance. According to the numerical comparison between the deterministic prediction methods shown in Tables I and II, the proposed GC has the highest deterministic prediction accuracy.

C. Numerical Analysis of PIs

To reveal the prediction performance of DOP, the data from Datasets 1 and 2 on rainy days, during which the PV generation has strong volatility and uncertainty and the prediction performance is the worst with PINC of 90% and look-ahead time of 1-hour, are used to display the influence of balance coefficient k in (22).

Figure 4 reveals the interval scores and $|ACE|$ in the cross-validation of training with different values of balance coefficients. When $k=1$, the interval score is directly applied as the cost function of DOP without the balance of reliability and overall performance. As shown in Fig. 4(a) and (b), DOPs with $k=1.029$ and $k=0.976$ have the best performances in terms of reliability and overall performance of PIs, respectively. The interval score, as the cost function of LP ($k=1$), is sometimes less sensitive to $|ACE|$, resulting in low reliability. Thus, the balance coefficient should be optimized. Here, for numerical comparison, the methods based on DQR [20], [34], MLLP [21], bootstrap-based ELM (BELM) [15], and the conditional probability-based PIs (CPPI) [14] are utilized as benchmarks. The least absolute shrinkage and selection operator (LASSO) penalty coefficient in MLLP and the selection of divided intervals for conditional probability in CPPI are both optimized by particle swarm optimization (PSO).

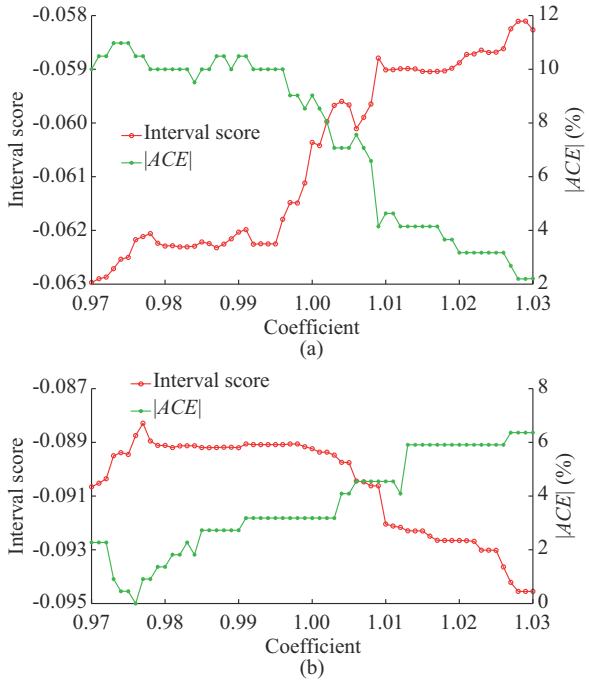

Fig. 4. PI performances with different values of balance coefficient. (a) Dataset 1. (b) Dataset 2.

For a comprehensive comparison, the forecasting perfor-

mances of the benchmarks and DOP are shown in Tables III-V. In Table III, the PI performances with the look-ahead time of 1-hour are given to reveal the numerical comparison of all the criteria analyzed in Section II-C. Meanwhile, considering the data resolutions of Datasets 1 and 2 are 15-min and 1-hour, respectively, the numerical comparisons with the look-ahead time of 30-min and 90-min for Dataset 1 are given in Table IV, while the numerical comparisons with the look-ahead time of 1-hour and 2-hour for Dataset 2 are given in Table V. As analyzed in Section II-C, interval score is the decisive criterion, reliability is an important observation criterion, and sharpness and AO are the auxiliary criteria. That is, the reasonable PIs require excellent overall performance and low coverage deviation. Among all the methods, DOP with the optimal balance coefficient has the best forecasting performance, since the reliability and overall performance are directly used as the optimization target. BELM and CPPI are methods of parametric PIs, which are affected by the accuracy of point prediction and error assumption. BELM, DQR, and CPPI quantify PIs mainly based on the reliability, while for MLLP, both the reliability and sharpness are considered. However, based on the quantification of interval score, the offsets of points outside PIs should be considered, which means that the PIs should not deviate too much from the points outside PIs to improve the rationality of probabilistic prediction.

TABLE III
COMPARISON RESULTS WITH LOOK-AHEAD TIME OF 1-HOUR BASED ON DATASET 1

Method	Sunny days				Rainy days				Cloudy days			
	PICP (%)	AW	AO	Score	PICP (%)	AW	AO	Score	PICP (%)	AW	AO	Score
DQR	93.77	0.1622	0.0202	–0.0375	88.62	0.2372	0.0284	–0.0604	82.11	0.2297	0.0311	–0.0682
MLLP	93.77	0.1725	0.0094	–0.0367	91.33	0.2351	0.0323	–0.0582	88.08	0.2737	0.0260	–0.0672
BELM	93.50	0.1524	0.0280	–0.0402	87.80	0.2238	0.0339	–0.0613	83.47	0.2423	0.0440	–0.0776
CPPI	94.31	0.1622	0.0205	–0.0371	87.80	0.2271	0.0320	–0.0611	87.80	0.2466	0.0385	–0.0681
DOP ($k=1$)	92.14	0.1428	0.0197	–0.0348	87.53	0.2209	0.0237	–0.0560	86.72	0.2294	0.0318	–0.0628
DOP (optimal k)	91.60	0.1397	0.0195	–0.0345	89.43	0.2301	0.0231	–0.0558	88.35	0.2347	0.0317	–0.0623

TABLE IV
COMPARISON RESULTS WITH DIFFERENT LOOK-AHEAD TIME BASED ON DATASET 1

Method	Sunny days				Rainy days				Cloudy days			
	30-min		90-min		30-min		90-min		30-min		90-min	
	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score
DQR	94.31	–0.0274	94.04	–0.0492	87.26	–0.0456	85.37	–0.0787	81.30	–0.0467	82.93	–0.0898
MLLP	92.14	–0.0267	94.31	–0.0473	88.62	–0.0422	91.60	–0.0721	87.53	–0.0450	85.91	–0.0890
BELM	92.14	–0.0289	87.26	–0.0526	88.08	–0.0441	85.91	–0.0774	81.03	–0.0540	85.37	–0.0984
CPPI	94.04	–0.0263	85.45	–0.0482	88.35	–0.0414	88.89	–0.0746	84.28	–0.0491	85.91	–0.0949
DOP ($k=1$)	93.50	–0.0250	92.95	–0.0439	88.08	–0.0391	88.08	–0.0695	84.12	–0.0412	84.55	–0.0786
DOP (optimal k)	88.89	–0.0242	89.16	–0.0401	90.24	–0.0387	90.51	–0.0693	88.35	–0.0411	89.16	–0.0760

TABLE V
COMPARISON RESULTS WITH DIFFERENT LOOK-AHEAD TIME BASED ON DATASET 2

Method	Sunny days				Rainy days				Cloudy days			
	1-hour		2-hour		1-hour		2-hour		1-hour		2-hour	
	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score	PICP (%)	Score
DQR	91.82	–0.0684	92.42	–0.0908	87.88	–0.0931	91.21	–0.1305	85.15	–0.0965	87.27	–0.1273
MLLP	88.48	–0.0657	92.73	–0.0857	92.73	–0.0894	92.42	–0.1222	84.55	–0.0906	87.56	–0.1179
BELM	86.36	–0.0741	86.97	–0.0995	88.18	–0.0896	94.26	–0.1235	88.18	–0.1017	85.15	–0.1362
CPPI	83.94	–0.0738	86.06	–0.0966	91.82	–0.0923	91.82	–0.1223	84.55	–0.0939	88.48	–0.1247
DOP ($k=1$)	87.27	–0.0600	95.45	–0.0837	92.12	–0.0889	92.73	–0.1200	88.18	–0.0830	88.79	–0.1082
DOP (optimal k)	89.09	–0.0588	91.21	–0.0819	89.39	–0.0885	89.70	–0.1192	89.09	–0.0827	90.61	–0.1067

In Tables VI and VII, the results of upscaling method (UM) [31], CM [27], IGNN [28], OGPI [30], hierarchical clustering-based DOP (HDOP) with the removal of low correlation samples, and DOP are used as benchmarks for the proposed method. Datasets 1 and 2 are utilized for numerical comparisons, respectively. In UM, the PV stations with both MAE and RMSE less than 10% and the correlation with regional output of more than 0.9, are selected as representative PV stations. UM and CM quantify the bounds of PIs based on the results of deterministic prediction and error assumption. The cluster numbers of all the CMs are obtained according to the prior test of training samples. In the numeri-

cal comparison of probabilistic predictions, IGNN quantifies parametric PIs based on the deterministic prediction and prediction errors of assumed Gaussian distribution, and even though its deterministic prediction is accurate, the probabilistic prediction performance is still affected by the accuracy of error assumption. OGPI is a novel granule computing-based framework for PI construction without the adverse effect of deterministic prediction and prior error assumption, and the parameters are optimized by PSO.

TABLE VI
COMPARISON RESULTS OF PIS IN MARCH-JUNE BASED ON DATASET 1

PINC (%)	Method	Sunny days				Rainy days				Cloudy days			
		PICP (%)	AW	AO	Score	PICP (%)	AW	AO	Score	PICP (%)	AW	AO	Score
90	UM	93.22	0.1885	0.0220	–0.0436	88.35	0.3432	0.0543	–0.0939	92.91	0.3159	0.0509	–0.0784
	CM	94.04	0.1650	0.0264	–0.0393	85.09	0.2057	0.0325	–0.0605	83.47	0.2559	0.0387	–0.0768
	IGNN	92.41	0.1659	0.0262	–0.0422	89.43	0.2398	0.0252	–0.0586	88.89	0.2963	0.0387	–0.0764
	OGPI	94.63	0.1925	0.0180	–0.0424	92.68	0.2591	0.0290	–0.0603	88.54	0.3082	0.0228	–0.0721
	HDOP	92.41	0.1596	0.0124	–0.0357	89.43	0.2264	0.0236	–0.0552	86.99	0.2439	0.0285	–0.0636
	DOP	91.60	0.1397	0.0195	–0.0345	89.43	0.2301	0.0231	–0.0558	88.35	0.2347	0.0317	–0.0623
	Proposed	91.33	0.1471	0.0119	–0.0335	90.24	0.2217	0.0270	–0.0549	89.97	0.2423	0.0312	–0.0610
95	UM	96.48	0.2213	0.0236	–0.0255	92.68	0.4087	0.0464	–0.0544	96.75	0.3820	0.0548	–0.0453
	CM	96.48	0.2154	0.0309	–0.0259	89.43	0.2430	0.0311	–0.0375	88.35	0.2873	0.0402	–0.0475
	IGNN	95.93	0.1983	0.0351	–0.0255	95.93	0.3037	0.0397	–0.0368	94.04	0.3530	0.0341	–0.0434
	OGPI	96.59	0.2359	0.0143	–0.0255	95.85	0.3181	0.0249	–0.0359	96.10	0.3836	0.0157	–0.0408
	HDOP	96.75	0.2115	0.0127	–0.0228	93.22	0.2529	0.0239	–0.0318	92.95	0.2958	0.0265	–0.0371
	DOP	94.85	0.1720	0.0229	–0.0219	93.50	0.2533	0.0269	–0.0323	92.68	0.2733	0.0334	–0.0371
	Proposed	95.12	0.1882	0.0132	–0.0214	94.58	0.2574	0.0244	–0.0310	95.66	0.3101	0.0212	–0.0347

TABLE VII
COMPARISON RESULTS OF PIS IN OCTOBER-FEBRUARY BASED ON DATASET 2

PINC (%)	Method	Sunny days				Rainy days				Cloudy days			
		PICP (%)	AW	AO	Score	PICP (%)	AW	AO	Score	PICP (%)	AW	AO	Score
90	UM	91.54	0.2971	0.0685	–0.0826	93.10	0.4459	0.0375	–0.0995	92.48	0.4342	0.0543	–0.1032
	CM	91.21	0.2354	0.0425	–0.0620	92.42	0.3950	0.0498	–0.0941	83.44	0.3042	0.0665	–0.1044
	IGNN	89.09	0.2418	0.0412	–0.0663	88.48	0.3301	0.0463	–0.0873	89.09	0.3145	0.0516	–0.0854
	OGPI	89.09	0.3592	0.0227	–0.0818	90.91	0.4182	0.0314	–0.0951	83.94	0.3404	0.0317	–0.0885
	HDOP	87.58	0.2053	0.0365	–0.0588	89.39	0.3697	0.0337	–0.0883	88.48	0.2647	0.0587	–0.0800
	DOP	89.09	0.2409	0.0243	–0.0588	89.39	0.3759	0.0313	–0.0885	89.09	0.3255	0.0403	–0.0827
	Proposed	90.30	0.2186	0.0319	–0.0561	90.30	0.3840	0.0221	–0.0854	90.91	0.3209	0.0385	–0.0782
95	UM	94.04	0.3547	0.0673	–0.0515	96.06	0.4884	0.0529	–0.0572	96.54	0.7109	0.0492	–0.0785
	CM	93.03	0.3355	0.0605	–0.0504	96.36	0.4908	0.0311	–0.0536	86.69	0.3628	0.0762	–0.0767
	IGNN	93.03	0.3051	0.0480	–0.0439	94.85	0.4037	0.0504	–0.0507	93.64	0.3652	0.0582	–0.0513
	OGPI	93.33	0.4251	0.0184	–0.0474	94.55	0.4638	0.0205	–0.0508	91.21	0.4127	0.0286	–0.0513
	HDOP	95.76	0.2628	0.0437	–0.0337	93.94	0.3939	0.0353	–0.0480	94.24	0.3614	0.0440	–0.0463
	DOP	94.24	0.2784	0.0269	–0.0340	94.24	0.4000	0.0389	–0.0490	93.94	0.3679	0.0411	–0.0468
	Proposed	94.55	0.2746	0.0255	–0.0330	95.15	0.4289	0.0236	–0.0475	94.55	0.3784	0.0352	–0.0455

As analyzed in [30], by segmenting the power time series into granules, OGPI captures the variability of data with high resolution (1-min resolution) and has a good forecasting performance. However, for the resolutions of 15-min and 1-hour, the effectiveness of OGPI is not stable. In most cases, the methods based on granule-based PIs including IGNN and OGPI have better forecasting performances than the conventional UM and CM, but their forecasting performances are not as good as the DOP. In the proposed method, to ensure the sufficient and precise training data, after the proposed GC, the utilization of training samples is based on an unbalanced extension rather than the conventional method of removing the low correlation samples. From the comparison of DOP and HDOP, the removal of low correlation samples cannot obviously improve the forecasting performance of DOP which quantifies PIs without deterministic prediction, while the proposed unbalanced extension of samples and GC improve the forecasting performance, compared with DOP. The proposed method quantifies the upper and lower bounds of PIs directly with the optimal overall performance and reliability. This means that the objective of PI construction is to obtain the optimal forecasting performance considering the

reliability, sharpness, and offsets of points outside PIs, rather than the conventional methods, which only aim at reliability or consider both reliability and sharpness. When the value of PICP is 100%, all the actual power values are within PIs, then, "none" is defined to denote AO. Of all the methods, PIs of the proposed method have the best forecasting performance.

A PC with Intel(R) Core(TM) i7-7700 CPU @ 2.8 GHz and 8 GB RAM is used for computations. The computational time of GC and DOPs of all clusters is 71 s and 49 s, respectively, while the prediction time of the proposed probabilistic prediction method is not more than 5 s. Since the training of the prediction model is usually carried out at least every few days, and the prediction time scale of very short-term is usually tens of minutes, the proposed method has good computational efficiency and forecasting performance, so it is suitable for very short-term probabilistic prediction.

Based on Datasets 1 and 2, Figs. 5 and 6 illustrate the PIs at different time points under different weather conditions with PINC of 90%. The sampling time of Figs. 5 and 6 is 15 min and 1 hour, respectively. It is clear from Fig. 5(a) and Fig. 6(a) that the PIs have high reliability and sharpness on sunny days, because the curve of power output changes smoothly. On rainy days and cloudy days, the curves of power outputs which change dramatically are more uncertain, compared with the outputs on sunny days.

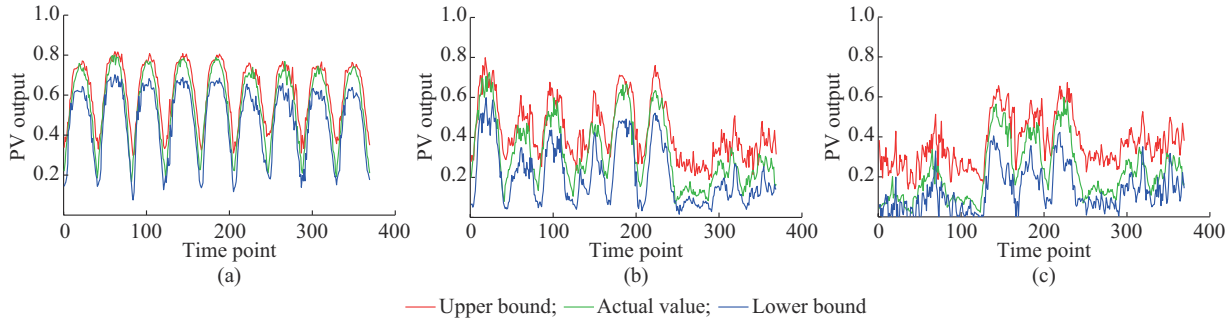

Fig. 5. PIs based on Dataset 1 under different weather conditions with PINC of 90%. (a) Sunny days. (b) Rainy days. (c) Cloudy days.

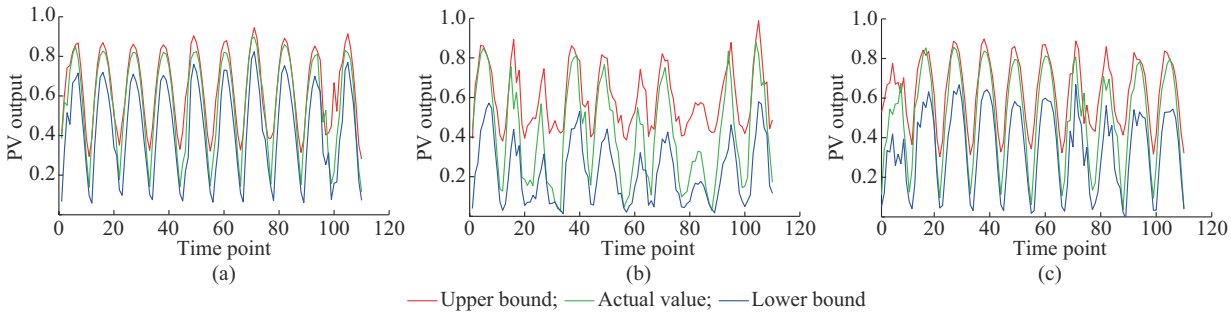

Fig. 6. PIs based on Dataset 2 under different weather conditions with PINC of 90%. (a) Sunny days. (b) Rainy days. (c) Cloudy days.

V. Conclusion

In this paper, a novel nonparametric method of very short-term probabilistic prediction for regional PV outputs based on GC and DOP is proposed. First, DOP is proposed to optimize the overall performance of nonparametric PIs. Compared with the conventional methods of PIs, the proposed DOP can effectively quantify the output weights of the optimal overall performance. The numerical comparison of calculation methods for PIs verifies the effectiveness of DOP under different weather conditions. Second, the balance coefficient is proposed to further improve the robustness of DOP, which considers the possibility that PIs with excellent overall performance have low reliability caused by the strong volatility and uncertainty of PV generation. The effect of balance coefficient is analyzed by numerical comparison. Third, the proposed GC is used to improve the clustering performance, and its effectiveness is verified by the numerical comparison of deterministic predictions with the prediction methods based on hierarchical clustering and K-means. Then, an unbalanced extension of samples is applied to enhance the samples' utilization, which is verified to be effective by numerical comparison among HDOP, DOP, and the proposed method. Finally, the effectiveness of the proposed method is verified by the numerical comparisons with the existing methods for regional PV power output.

Since the forecasting performance of PV power is obviously affected by the classification accuracy of daily weather types which usually have certain changes within a few hours, it is necessary to further consider the time scale of half a day or even several hours for optimized classification, which will be studied in future work.

References

[1] B. Mohandes, M. S. E. Moursi, N. Hatziargyriou *et al.*, "A review of power system flexibility with high penetration of renewables," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3140-3155, Jul. 2019.

[2] Y. Wang, Y. Sun, and V. Dinavhi, "Robust forecasting-aided state estimation for power system against uncertainties," *IEEE Transactions on*

Power Systems, vol. 35, no. 1, pp. 691-702, Jan. 2020.
[3] Y. Sun, X. Wu, J. Wang *et al.*, "Power compensation of network losses in a microgrid with BESS by distributed consensus algorithm," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 4, pp. 2091-2100, Apr. 2021.
[4] P. Siano, "Evaluating the impact of registered power zones incentive on wind systems integration in active distribution networks," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 2, pp. 523-530, Apr. 2015.
[5] J. Yang, M. Yang, M. Wang *et al.*, "A deep reinforcement learning method for managing wind farm uncertainties through energy storage system control and external reserve purchasing," *International Journal of Electrical Power and Energy Systems*, vol. 119, p. 105928, Jul. 2020.
[6] M. Yang, X. Chen, and B. Huang, "Ultra-short-term multi-step wind power prediction based on fractal scaling factor transformation," *Journal of Renewable and Sustainable Energy*, vol. 10, no. 5, pp. 1-10, Jun. 2018.
[7] P. Siano, C. Cecati, H. Yu *et al.*, "Real time operation of smart grids via FCN networks and optimal power flow," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 4, pp. 944-952, Nov. 2012.
[8] L. Ye, C. Zhang, Y. Tang *et al.*, "Hierarchical model predictive control strategy based on dynamic active power dispatch for wind power cluster integration," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4617-4629, Nov. 2019.
[9] A. Singla, K. Singh, and V. K. Yadav, "Optimization of distributed solar photovoltaic power generation in day-ahead electricity market incorporating irradiance uncertainty," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 3, pp. 545-560, May 2021.
[10] S. V. Medina and U. P. Ajenjo, "Performance improvement of artificial neural network model in short-term forecasting of wind farm power output," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 3, pp. 484-490, May 2020.
[11] Z. Wang, W. Wang, C. Liu *et al.*, "Forecasted scenarios of regional wind farms based on regular vine copulas," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 1, pp. 77-85, Sept. 2020.
[12] Y. Zhao, L. Ye, P. Pinson *et al.*, "Correlation-constrained and sparsity-controlled vector autoregressive model for spatio-temporal wind power forecasting," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5029-5040, Sept. 2018.
[13] L. Ge, Y. Xian, J. Yan *et al.*, "A hybrid model for short-term PV output forecasting based on PCA-GWO-GRNN," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1268-1275, Nov. 2020.
[14] Y. Sun, P. Wang, S. Zhai *et al.*, "Ultra short-term probability prediction of wind power based on LSTM network and condition normal distribution," *Wind Energy*, vol. 23, no. 2, pp. 63-76, Oct. 2019.
[15] C. Wan, Z. Xu, P. Pinson *et al.*, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1033-1044, May 2014.
[16] Y. Lin, M. Yang, C. Wan *et al.*, "A multi-model combination approach for probabilistic wind power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 226-237, Jan. 2019.
[17] J. R. Andrade and R. J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1571-1580, Oct. 2017.
[18] J. Wang, H. Zhong, X. Lai *et al.*, "Exploring key weather factors from analytical modeling toward improved solar power forecasting," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1417-1427, Mar. 2019.
[19] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation – with application to solar energy," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3850-3863, Sept. 2016.
[20] C. Wan, J. Lin, Y. Song *et al.*, "Probabilistic forecasting of photovoltaic generation: an efficient statistical approach," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2471-2472, Apr. 2016.
[21] C. Wan, J. Wang, J. Lin *et al.*, "Nonparametric prediction intervals of wind power via linear programming," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1074-1076, Jan. 2018.
[22] G. Sideratos and N. Hatziargyriou, "Probabilistic wind power forecasting using radial basis function neural networks," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1788-1796, Nov. 2012.
[23] G. Sideratos and N. Hatziargyriou, "A distributed memory RBF-based model for variable generation forecasting," *International Journal of Electrical Power and Energy Systems*, vol. 120, p. 106041, Sept. 2020.
[24] G. Sideratos and N. Hatziargyriou, "An advanced statistical method for wind power forecasting," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 258-265, Mar. 2007.
[25] M. Yang, C. Shi, and H. Liu, "Day-ahead wind power forecasting based on the clustering of equivalent power curves," *Energy*, vol. 218, p. 119515, Mar. 2021.
[26] X. Zhang, Y. Li, S. Lu *et al.*, "A solar time based analog ensemble method for regional solar power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 268-279, Jan. 2019.
[27] L. Han, T. Shang, J. Shu *et al.*, "Time series data intelligent clustering algorithm for landslide displacement prediction," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 4, pp. 4131-4140, Oct. 2018.
[28] X. Zhu, W. Pedrycz, and Z. Li, "Development and analysis of neural networks realized in the presence of granular data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 9, pp. 3606-3619, Sept. 2020.
[29] W. Pedrycz, A. Jastrzebska, and W. Homenda, "Design of fuzzy cognitive maps for modeling time series," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 1, pp. 120-130, Feb. 2016.
[30] S. Chai, Z. Xu, and W. K. Wong, "Optimal granule-based PIs construction for solar irradiance forecast," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3332-3333, Jul. 2016.
[31] M. Pierro, M. Felice, E. Maggioni *et al.*, "Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data," *Solar Energy*, vol. 158, pp. 1026-1038, Dec. 2017.
[32] J. G. Fonseca, T. Oozeki, H. Ohtake *et al.*, "Regional forecasts and smoothing effect of photovoltaic power generation in Japan: an approach with principal component analysis," *Renewable Energy*, vol. 68, pp. 403-413, Aug. 2014.
[33] M. Yang, L. Zhang, Y. Cui *et al.*, "Investigating the wind power smoothing effect using set pair analysis," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1161-1172, Jul. 2020.
[34] D. T. Phong, T. Pham-Gia, and D. N. Thanh, "Exact distributions of two non-central generalized Wilks's statistics," *Journal of Statistical Computation and Simulation*, vol. 89, no. 10, pp. 1798-1818, Apr. 2019.
[35] C. Wan, J. Lin, J. Wang *et al.*, "Direct quantile regression for nonparametric probabilistic forecasting of wind power generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2767-2778, Jul. 2016.
[36] C. Wan, Z. Xu, P. Pinson *et al.*, "Direct interval forecasting of wind power," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4877-4878, Nov. 2013.
[37] Y. Zhou, Y. Sun, S. Wang *et al.*, "A very short-term probabilistic prediction method of wind speed based on ALASSO-nonlinear quantile regression and integrated criterion," *CSEE Journal of Power and Energy Systems*. doi:10.17775/CSEEJPES.2020.05370
[38] Y. Zhou, Y. Sun, S. Wang *et al.*, "Performance improvement of very short-term prediction intervals for regional wind power based on composite conditional nonlinear quantile regression," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 1, pp. 60-70, Jan. 2022.
[39] T. Hong, P. Pinson, S. Fan, *et al.*, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896-913, Jul.-Sept. 2016.
[40] G. Huang, Q. Zhu, and C. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, Dec. 2006.
[41] X. Huang, Y. Ye, and H. Zhang, "Extensions of K-means-type algorithms: a new clustering framework by integrating intracluster compactness and intercluster separation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 8, pp. 1433-1446, Aug. 2014.

Yonghui Sun received the Ph. D. degree in electrical engineering from the City University of Hong Kong, Hong Kong, China, in 2010. He is currently a Professor with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China. His research interests include stability analysis and control of power systems, optimal planning and operation of the integrated energy system, optimization algorithms, and data analysis.

Yan Zhou received the Ph. D. degree in electrical engineering from Hohai University, Nanjing, China. His research interests include power big data analysis, machine learning, power system operation optimization, and renewable energy generation forecasting.

Sen Wang is currently pursuing the Ph. D. degree in electrical engineering

with Hohai University, Nanjing, China. His research interests include power big data analysis, renewable energy grid-connected control, power system operation optimization, and renewable energy generation forecasting.

Rabea Jamil Mahfoud received the Ph.D. degree in electrical engineering from Hohai University, Nanjing, China, in 2020. His current research interests include optimal operation and planning of power systems, optimization algorithms, and data analysis.

Hassan Haes Alhelou is currently a Faculty Member with Tishreen University, Lattakia, Syria. He was a recipient of the Outstanding Reviewer Award from Energy Conversion and Management Journal in 2016, ISA Transactions Journal in 2018, Applied Energy Journal in 2019, and many other awards. He is included in the year 2018 and 2019 Publons list of the Top 1% Best Reviewer and Researchers in the field of engineering. His major research interests include power systems, power system dynamics, power system operation and control, dynamic state estimation, frequency control, smart grids, microgrids, demand response, load shedding, and power system protection.

George Sideratos received the Electrical and Computer Engineering degree from the National Technical University of Athens (NTUA), Athens, Greece, in 2002 and the Ph. D. degree in electrical engineering from NTUA, in 2010. He is currently a Senior Researcher in the Power System Laboratory of NTUA. His research interests include wind power and load forecasting and artificial intelligence techniques.

Nikos Hatziargyriou is a Professor in power systems at the National Technical University of Athens, Athens, Greece. He has over 10-year industrial experience as Chairman and CEO of the Hellenic Distribution Network Operator (HEDNO), Athens, Greece. He was Chair and Vice-chair of the EU Technology and Innovation Platform on Smart Networks for Energy Transition (ETIP-SNET) representing E. DSO. He is Life Fellow Member of IEEE, past Chair of the Power System Dynamic Performance Committee (PSDPC) and currently Editor in Chief of the IEEE Transactions on Power Systems. He is included in the 2016, 2017, and 2019 Thomson Reuters lists of the top 1% most cited researchers and he is 2020 Globe Energy Prize laureate. His research interests include smart grids, microgrids, distributed and renewable energy sources, and power system security.

Pierluigi Siano received the M.Sc. degree in electronic engineering and the Ph.D. degree in information and electrical engineering from the University of Salerno, Salerno, Italy, in 2001 and 2006, respectively. He is a Professor and Scientific Director of the Smart Grids and Smart Cities Laboratory with the Department of Management & Innovation Systems, University of Salerno. Since 2021 he has been a Distinguished Visiting Professor in the Department of Electrical & Electronic Engineering Science, University of Johannesburg. He has been the Chair of the IES TC on Smart Grids. He is Editor for the Power & Energy Society Section of IEEE Access, IEEE Transactions on Power Systems, IEEE Transactions on Industrial Informatics, IEEE Transactions on Industrial Electronics, IEEE Systems. His research interests include demand response, energy management, the integration of distributed energy resources in smart grids, electricity markets, and planning and management of power systems. In these research fields, he has co-authored more than 700 articles including more than 410 international journals that received in Scopus more than 17000 citations with an H-index equal to 63. In the period 2019-2022 he has been awarded as a Highly Cited Researcher in Engineering by Web of Science Group.