Electric Power Systems Research 214 (2023) 108817

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Very short-term probabilistic prediction of PV based on multi-period error distribution

Sen Wang, Yonghui Sun *, Shanming Zhang, Yan Zhou, Dongchen Hou, Jianxi Wang

College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China

ARTICLE INFO

Keywords:
Multi-period error analysis
Very short-term
Probability prediction
EEMD
LSTM

ABSTRACT

As the penetration rate of photovoltaic (PV) in the grid increases, enormous challenges have been brought into power grid dispatcher's operation. Efficient and accurate PV power prediction is the key to solve this problem. Considering multi-period error distribution (MPED), a novel probabilistic prediction approach via ensemble empirical mode decomposition based on long short-term memory and backpropagation neural network (EEMD-LSTM-BP) is proposed. EEMD is utilized to study the characteristic of wave behaviors in different frequency domains. LSTM and BP are used to determine intrinsic mode functions (IMFs) and remaining components, respectively. Afterward, based on the prediction errors, PV power output fluctuation in different periods is analysed. The segment points are determined by Nadaraya-Watson (N-W) kernel regression. The bounds of prediction intervals (PIs) are quantified based on the error probability distribution. Based on the dataset of PV stations in Ningxia Province, the case studies verify the method's feasibility.

1. Introduction

Due to the depletion of fossil energy, renewable energy is becoming progressively critical. Therefore, the installed capacity of renewable energy in the world keeps increasing, especially photovoltaic (PV). However, the output of PV power generation has strongly unstable and random. At the same time, distributed generation is another feature of PV power generation. As the penetration rate of PV in the grid increases, it brings enormous challenges to power grid dispatchers' operation [1]. Therefore, high-precision PV power prediction technology can encourage scheduling optimization and optimize energy consumption. It is worth noting that probabilistic prediction can better reflect PV prediction uncertainty and improve more reliable information support for power system operation, optimization, and protection.

In recent years, advanced prediction methods including auto-regressive and moving average (ARMA), support vector machines (SVM), extreme learning machine (ELM), convolutional neural network (CNN), deep belief network (DBN) were studied in [2–4]. A CNN is a form of artificial neural network that is especially built to analyze pixel input in pattern recognition systems. A CNN employs a technology similar to a deep network that has been optimized for limited computational needs. Support-vector machines (SVMs, also known as support-vector networks) are convolutional networks learning techniques that examine information for machine learning assessment in computer vision. In complement to normal categorization, SVMs can do non-linear categorization quickly by employing the kernels technique, which involves indirectly translating their signals into high-dimensional subspaces. In [5], a long short-term memory (LSTM) network was used to forecast Santiago's solar irradiance. It took into account the correlation between consecutive hours. Compared with the conventional backpropagation neural network (BPNN), the prediction accuracy was verified. In [6], a novel two-way LSTM constituted by backpropagation through time (BPTT) was proposed. The numerical studies revealed that the prediction accuracy was improved on rainy days and sunny days. In [7], a novel method of PV power prediction based on ship micro-grid is proposed. Combine multiple machine learning algorithms to form a hybrid prediction method model.

Since deterministic prediction results inevitably bring errors and cannot reflect PV's uncertainty, probabilistic prediction has become a hot topic in recent years. In [8], a PV probabilistic prediction model based on a robust self-attention mechanism was proposed. In order to analyze data, quantile regression is a sort of logistic regression. Whenever the requirements of predictor variables are not satisfied, it is employed as an augmentation. Based on the certainty prediction results, quantile regression (QR) was used to analyze the model's robustness. In [9], after determining deterministic prediction errors, a t-position scale

* Corresponding author.
E-mail address: yhsun@hhu.edu.cn (Y. Sun).

https://doi.org/10.1016/j.epsr.2022.108817
Received 19 August 2021; Received in revised form 9 December 2021; Accepted 14 September 2022
Available online 23 September 2022

distribution that significantly improved the prediction accuracy was proposed. In [10], a probabilistic approach based on bootstrap and QR for PV generation was studied. For a PV power station, the conditions, including weather, irradiance, and cloud layer in the same period and season, have a highly similar characteristic [11–13]. Thus, it is significant to analyze the prediction error in a comparable period within each day [14]. Statistical analysis for the PV time series of high volatility requires a broader range of coverage. For PV time series of low volatility, the deterministic prediction has better performance. According to the PV generation, high prediction errors usually occur at noon [10]. Generalized cross validation (GCV) is a popular method for estimating variables in the framework of nonlinear systems and regularization methods. The calculation of the smoothing factor in curvatures is an ideal illustration. The prediction errors can be processed for data smooth based on generalized cross-validation (GCV) [15]. Meanwhile, considering the error characteristics in different periods, the uncertainty analysis can be further enhanced. Thus, the uncertainty analysis for the multi-period proposed in this paper is meaningful.

For the problem that high-frequency data cannot be efficiently predicted, ensemble empirical mode decomposition (EEMD) was used to decompose the original sequence [16]. Based on the above analysis, to enhance the quality of prediction intervals (PIs), a combination of EEMD, LSTM, and BPNN considering multi-period error distribution (MPED) is proposed. Firstly, EEMD is carried out for the data processing. Secondly, the intrinsic mode functions (IMFs), which LSTM trains represent fluctuation characteristic. And the remainder is predicted by BPNN. The uncertainty analysis then quantifies the multi-period segment points, and different prediction error probability distributions are selected. Finally, PIs can be quantified given the effect of deterministic prediction and uncertainty analysis.

This paper remains organized as follows: The deterministic prediction model based on EEMD and LSTM for PV power is described in Section II. In Section III, the proposed very short-term probability prediction method considering the MPED is analysed. Case studies are given for the verification of the raise method in Section Ⅳ. Section V concludes this paper.

2. Theoretical background

2.1. Ensemble empirical mode decomposition

PV power has apparent periodicity, so the frequency domain decomposition method is significant. EEMD is a special algorithm that improves its own algorithm steps for modal aliasing as a new type of signal decomposition method. The EEMD is a noise-assisted approach frequently used for detecting roller bearing degradation. It is demonstrated that the suggested technique (performance enhanced EEMD) obtains a greater achievement in damage identification and generates a wider lead as compared to the existing version. It is primarily utilised decompose the underlying time series analysis into a limited and small number of oscillating phases based on regional distinctive timeframe. The technique adds Gaussian white noise to the empirical mode decomposition (EMD) [17]. Gaussian white noise (GWN) is a stable and harmonic chaotic system with normally distributed that is described by the following leadership asset: every two points of GWN are uncorrelated, regardless of their temporal proximity. This characteristic has the clear conclusion that the oscillation functions of a GWN. White relates to the notion that the enterprise has constant performance throughout the radio spectrum. It is analogous to the color white, that emits uniformly across all wavelengths in the spectral region. Gaussian since it has a temporal dimension data is normally distributed with an expected time realm mean of 0. The Gaussian distribution is a statistical distribution function in analytics since it accurately describes numerous physical events such as age, length, standardized tests, Intellectual capacity, total of two rolling dice, etc.

After multiple steps, the average is taken as the ultimate

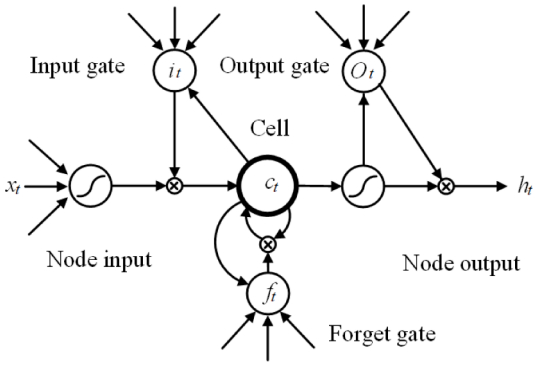


Fig. 1. LSTM structure.

experimental outcome. Gaussian white noise itself has a uniform frequency distribution, and the function is to make the signals at different scales have continuity and no longer alias signals [18,21]. There are three primary processes of EEMD detailed as follow:

1) The amplitude coefficient h of the white noise signal, e denotes the number of decompositions, and the maximum number of iterations of decomposition is E.
2) Run the EMD e times.

Firstly, a randomly generated white noise sequence $n_m(t)$ is added to change the original time series to a new sequence.

$$p_m(t) = p(t) + hn_m(t) \tag{1}$$

Secondly, the decomposition is performed to decompose the new sequence into IMF components, namely $c_{i,e}(t)$, and a residual component $r_{n,e}(t)$.

Finally, different white noise sequences are selected for each iteration, but still the same square root is required, then the previous steps are repeated, and finally different IMF components and residual components are combined.

1) The IMF components and the remaining components are uniformly processed, and the average value is calculated. The average value of each combination is the result of the final decomposition of EEMD. The intrinsic mode component is the oscillating component of a transmission that is acquired after the signal has been Hilbert-Huang transformed. This fundamental functionality must be retrieved correctly from the message in attempt to preserve the transmitter qualities and physiological significance.

$$\overline{c}_i(t) = \frac{\sum_{e=1}^{E} c_{i,e}(t)}{E} \tag{2}$$

$$\overline{r}_n(t) = \frac{\sum_{e=1}^{E} r_{n,e}(t)}{E} \tag{3}$$

2.2. Long short-term memory

LSTM, due to the creative input gate, forgotten gate, output gate structure, and time-series processing, has a strong predictive ability [19–20,22–24]. Its structure is shown in Fig. 1. The cell contains the information of long-term memory. In training algorithm, LSTM principally handles the overfitting problem. The memorizing operation is governed by a restriction in LSTMs. LSTMs use gateways that enter and shut to record, write, and access data. Training is a program that leads to skilful behavior by teaching people the fundamental skills required to

execute their professions. It is a minimal learning method that entails attaining enlightenment, honing skills, ideas, norms, or altering behavioural patterns in order to achieve organizational objectives. In the training process, the forgetting gate is the output of the above layer. And the sequence data to be input by this layer as input, through an activation function sigmoid, the output is f_t. The output value is in the interval [0,1], indicating the probability that the upper layer of the cell state is forgotten, 1 means complete reservation, and 0 means to abandon completely.

$$f_t = \sigma\left(W_h{}^f h_{t-1} + W_x{}^f x_t + b_i{}^f\right) \tag{4}$$

where f_t is the output of the forgetting gate, controlling the state of the upper layers in the cell. $W_*{}^*$ represents the weight. $b_*{}^*$ represents the bias.

In reality, tanh refers to a wide range of sigmoid values, particularly nonlinear activation features. Tanh and sigmoid are both S-shaped arcs; the biggest distinction is that sigmoid has a range from 0 and 1. Tanh, on the other hand, is around 1 and −1. However, as contrasted to sigmoid, the average of the tanh function is indeed near to 0. The *tanh* and *sigmoid* function are used as activation functions in each part of input gate.

$$i_t = \sigma\left(W_h^i h_{t-1} + W_x^i x_t + b_i^i\right) \tag{5}$$

$$\widetilde{C}_t = tanh(W_h{}^c h_{t-1} + W_x{}^c x_t + b_i{}^c) \tag{6}$$

where i_t and $\widetilde{C}_t$ are used as outputs in each part.

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \widetilde{C}_t \tag{7}$$

The output gate is used to control the number of filtered layers. The sigmoid function is used primarily since it occurs among (0 to 1). As a result, it is particularly useful for methods that allow us to anticipate the likelihood as an outcome. Because the likelihood of everything occurs only about 0 & 1, the sigmoid is the best option. The sigmoid function is used to get a value of [0,1] to calculate O_t. After processing with the *tanh* function multiply, output of this layer is obtained.

$$O_t = \sigma\left(W_h{}^o h_{t-1} + W_x{}^h x_t + b_i{}^o\right) \tag{8}$$

$$h_t = O_t \otimes tanh(C_t) \tag{9}$$

The LSTM here is used for deterministic prediction of IMFs. EMD is used to derive intrinsic mode functions (IMFs), and different sensitivity metrics also including Shannon's averaged entropy, approximation, sampling, phases, and Renyi are computed. Firstly, the IMFs are represented as $\mathbf{F} = \{f_1, f_2, \cdots, f_n\}$, the training dataset and testing dataset are represented as $\mathbf{F}^{'} = \{f_1, f_2, \cdots, f_m\}$ and $\mathbf{F}^{''} = \{f_{m+1}, f_{m+2}, \cdots, f_n\}$, respectively. After generating the training dataset, zero averaging of the elements f_i can be expressed as:

$$\mathbf{F}^{'} = \{f^{'}{}_1, f^{'}{}_2, \cdots, f^{'}{}_m\} \tag{10}$$

$$f_o^{'} = \left(f_o - \frac{\sum_{o=1}^{m} f_o}{m}\right) \sqrt{\frac{\sum_{o=1}^{m} \left(f_o - \sum_{o=1}^{m} f_o \Big/ m\right)^2}{m}} \tag{11}$$

where F is processed by data segmentation.

The serial port segmentation length is L. Then the input of the segmented model can be expressed as:

$$D = \{D_1, D_2, \cdots, D_L\} \tag{12}$$

$$D_p = \{f^{'}{}_p, f^{'}{}_{p+1}, \cdots, f^{'}{}_{m-L+p-1}\} \tag{13}$$

The last row of the theoretical output $\boldsymbol{Y}$ can be expressed as:

$$G_f = \{f_{m-L+1}^{'}, f_{m-L+2}^{'}, \cdots, f_m^{'}\} \tag{14}$$

$$Y_f = \mathrm{LSTM}_{\mathrm{net}}^{*}(G_f) = \{y_{m-L+2}, y_{m-L+3}, \cdots, y_{m+1}\} \tag{15}$$

where G_f is the input of $\mathrm{LSTM}_{\mathrm{net}}^{*}$.

The predicted value is $\mathrm{y_{m+1}}$ when the time is $m + 1$. Combine the L-1 data segments in G_f with $\mathrm{p_{m+1}}$ to form the following formula:

$$G_{f+1} = \{f_{m-L+2}^{'}, f_{m-L+3}, \cdots, y_{m+1}\} \tag{16}$$

The predicted data segment is finally formed as:

$$Y_{\mathrm{o}} = \{y_{m+1}, y_{m+2}, \cdots, y_n\} \tag{17}$$

where F_{te} is the final predicted value.

$$Y_{\mathrm{te}} = \mathrm{de\_zscore}(Y_{\mathrm{o}}) = \{y_{m+1}^{*}, y_{m+2}^{*}, \cdots, y_n^{*}\} \tag{18}$$

$$y_k^{*} = y_k \sqrt{\frac{\sum_{o=1}^{m} \left(f_o - \sum_{o=1}^{m} f_o \Big/ n\right)^2}{m}} + \frac{\sum_{o=1}^{m} f_o}{m} \tag{19}$$

The loss function is calculated as follows:

$$\mathscr{L} = \sum_{i=1}^{L(m-L)} (y_i - g_i)^2 \Big/ [L(m-L)] \tag{20}$$

In the process of model training, the minimum loss is utilized as the optimization objective. The parameters including training steps step length, learning rate, sample number and feature dimension are initialized before model training.

3. Probabilistic prediction model considering multi-period error distribution

3.1. Deterministic prediction and error analysis

A deterministic prediction model combining EEMD, LSTM, and BPNN is utilized [23,24]. The PV time series is decomposed by EEMD firstly. The component of IMFs represents the fluctuation characteristic. LSTM is used to train and predict components of IMFs, and BPNN is used for the residual component.

Backpropagation in neural networks is an abbreviation for backward error propagation. It is a common technique for developing artificial neural networks. This approach aids in computing the distance of a losses system with regard to some of the channel's parameters. The backpropagation technique computes the derivative to reference for every value using the chain rule, each piece at a time, processing data downstream since the last protective layer unnecessary computations of intervening components in the chain rule. There are 4 main steps for the deterministic prediction as follows:

Step 1: The insufficient data are eliminated and corrected based on the average power values at the preceding and the following observations.

Step 2: The data were decomposed using EEMD to obtain multiple IMFs and the residual component. The components are classified based on high or low frequency.

Step 3: LSTM is utilized for deterministic prediction of high-frequency components, while select BPNN prediction for low-frequency components, considering the prediction performance and computational efficiency.

Step 4: According to the results of deterministic prediction, the prediction error can be carried out as the basis of uncertainty analysis.

The conventional methods for PIs are based on the prediction error, which approximately satisfies Gaussian distribution in the morning or all-day [25–27]. The distribution function of prediction error is as follows:

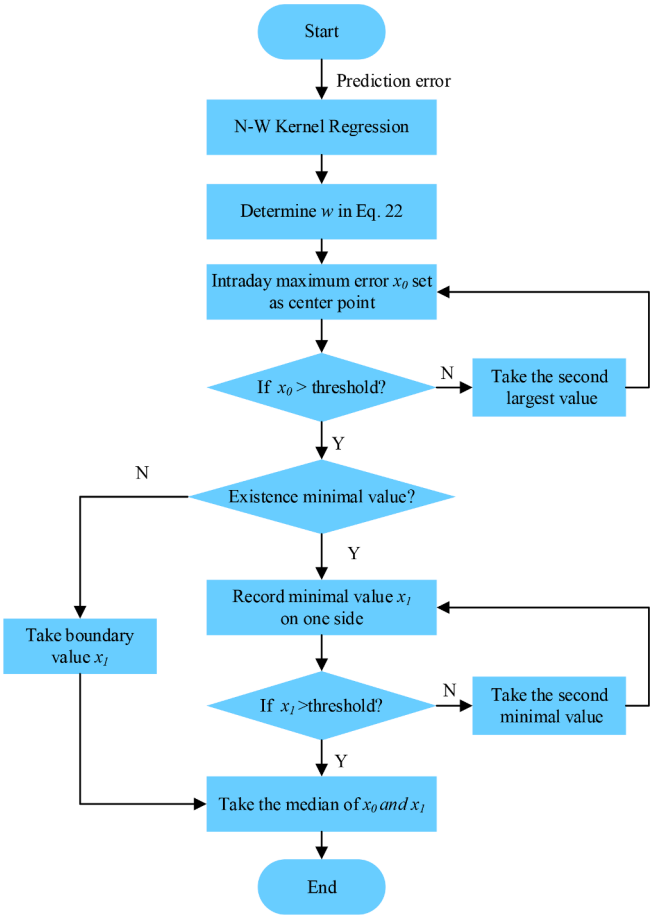


Fig. 2. Error distribution over different time periods.

$$f(x) = \frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma} \tag{21}$$

Nonparametric estimate is a quantitative approach for obtaining the linear model of a match to information in the presence of any proper instructions or limitations. Therefore, as consequence, the stochastic evaluation approaches have no significant related variables. Kernel regression is a non-parametric approach used in statistical to evaluate the parameter estimates of a stochastic process. The goal is to discover a non-linear relationship among different unknown parameters, X and Y. In order to improve the accuracy of uncertainty analysis, a nonparametric estimation technique can be used for error distribution. Here, Nadarava-Watson (N-W) kernel regression is used to estimate the error. The N-W kernel extrapolation estimation is a popular and versatile unbiased estimation of a linear regression that is frequently constructed by employing a predetermined frequency. The weight function is defined as [28,29]:

$$\aleph_i(x) = \frac{H\left(\frac{x-x_i}{w}\right)}{\sum_{j=1}^{n} H\left(\frac{x-x_j}{w}\right)} \tag{22}$$

where $H(\cdot)$ denotes the kernel function, and w is bandwidth which is positive. A kernel function is a way for taking information as input and transforming it into the needed format for execution. The term Kernel is chosen because the collection of arithmetic operations employed in SVM affords a window through which information may be manipulated. The n-th N-W kernel regression is:

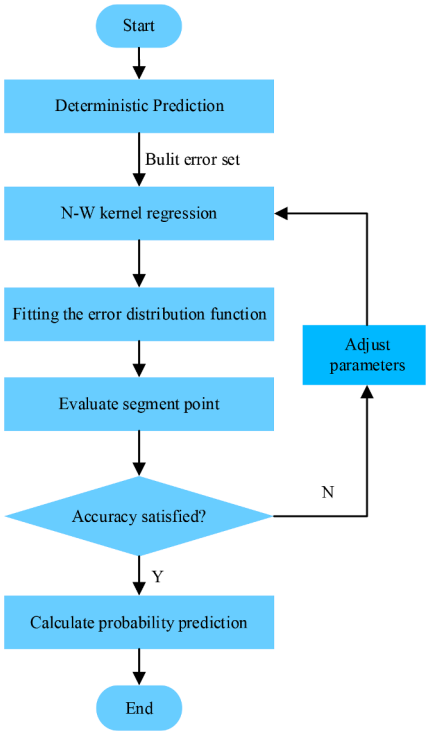


Fig. 3. Probability prediction model flowchart considering MPED.

$$\widehat{r}_n(x) = \sum_{i=1}^{n} \aleph_i(x) Y_i \tag{23}$$

3.2. Multi-period error distribution

To accurately quantify the cut-off points for periods, N-W kernel regression and the analysis of multi-period error are required, and the searching process of cut-off points is shown in Fig. 2. The attribution range verification is utilized based on the intraday error distribution scatter characteristic. According to the characteristic of prediction errors, the maximum point in each day which lies in the ownership scope, is set as the center point. Then, each side's cut-off point is quantified based on the existence validation, which is designed for typical days where the error distribution is relatively flat. The curve roughly conforms to the normal distribution or the bell type monotonic distribution; the algorithm takes the boundary value instead of the minimum value point. This category will lead to the phenomenon that the number of errors in the high error part is too large, which can be fitted through weights optimization. The threshold verification is designed to deal with the multi-peak phenomenon in the high error period. When the set threshold is less than the first minimum point's error frequency, the next extremum point is stored. The bimodal and multi-peak distributions reflect the error aggregation in the high stage, which should belong to the same kind of analysis.

3.3. Construction of probabilistic prediction

The model considering multi-period error distribution mainly based on the EEMD-LSTM-BP is demonstrated in Fig. 3. Firstly, the deterministic prediction model is utilized. Then, the optimization method is used to determine the segment points. The error distribution fitting is carried out separately at different time periods. Finally, the parameter estimation is calculated, and the corresponding analysis of probabilistic prediction is carried out. Mainly can be divided into three key steps:

Step 1: The error distribution of multiple data points is counted in the time segment, and the intraday error is analysed separately. The error

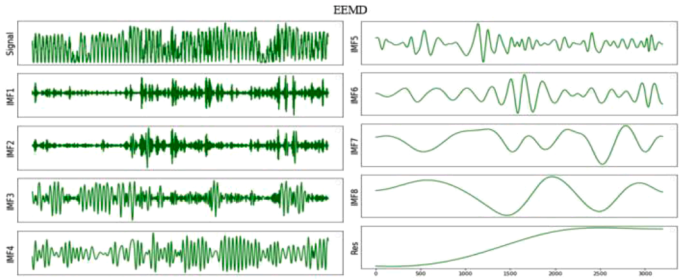


Fig. 4. The EEMD waveform for PV series in the training.

curve is fitted with N-W kernel regression, and smooth parameters are determined by generalized cross-validation. Then, the samples are weighted to determine the cut-off points.

Step 2: Fitting the error distribution function by time interval, based on N-W kernel regression. The confidence interval is set to calculate the distribution function and corresponding quantile of different periods, and the interval width is calculated.

Step 3: The PIs can be quantified based on the deterministic prediction and uncertainty analysis.

4. Case studies

4.1. Introduction of dataset

The dataset of PV stations in Ningxia Province is used for the numerical comparison in this paper. The prediction time scale is 4 h, and the resolution of the time series is 15 min. The datasets from March and May in 2018 are respectively utilized as the testing samples, while nearly 3000 PV power observations for training. And the prediction time is 8:00 to 18:00.

4.2. Evaluation criteria

For the numerical comparison, the accuracy of prediction interval coverage probability (PICP) according to PI nominal confidence (PINC), prediction interval nominal average width (PINAW), and interval score are utilized as the criteria [30–33].

1) Prediction interval coverage probability

In commercial processes, predictions intervals are extensively utilized to forecast the dispersion of data sets. In the research, the precise comprehensive insurance likelihood as well as mean covering likelihood of the traditional specified range for a discrete probability distribution have not been correctly calculated.

$$\mathrm{PICP} = \frac{1}{n}\sum_{i=1}^{n}\lambda_i \tag{24}$$

where n represents the number of samples. If the i th actual observation lies in the PI, the value of λ_i is 1, otherwise the value of λ_i is 0.

2) Prediction interval nominal average width

$$\mathrm{PINAW} = \frac{1}{nR}\sum_{i=1}^{n}\left[U_t^{\alpha}(x_i) - L_t^{\alpha}(x_i)\right] \tag{25}$$

$$R = P_{pre}{}^{\max} - P_{pre}{}^{\min} \tag{26}$$

where P_{pre} represents the value of PV power prediction. It is also the deterministic predictive output. α is the significant level. $L_t^{\alpha}(x_i)$ and

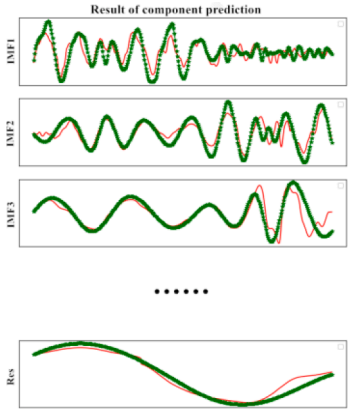


Fig. 5. Prediction of IMFs.

$U_t^{\alpha}(x_i)$ are the lower bounds and upper bounds are calculated by the method in this paper, respectively.

3) Prediction Interval score

$$\delta_t^{\alpha}(i) = U_t(x_i) - L_t(x_i) \tag{27}$$

$$S(i) = \begin{cases} -2\alpha\delta_t^{\alpha}(i) - 4[L_t(x_i) - t_i], t_i < L_t(x_i) \\ -2\alpha\delta_t^{\alpha}(i), t_i \in [L_t(x_i), U_t(x_i)] \\ -2\alpha\delta_t^{\alpha}(i) - 4[t_i - U_t(x_i)], t_i > U_t(x_i) \end{cases} \tag{28}$$

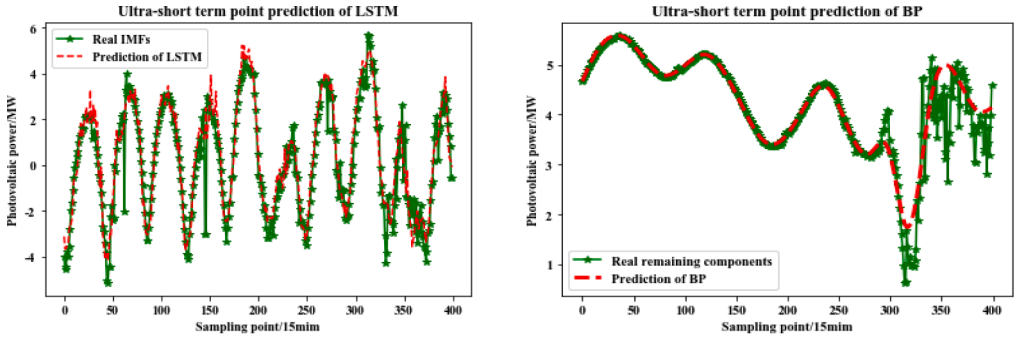
$$score = \frac{1}{n}\sum_{i=1}^{n} S(i) \tag{29}$$

where $\delta_t^{\alpha}(i)$ represents the i th interval width, and $S(i)$ represents the i th score.

4.3. Deterministic prediction

The PV power time series from March 22 to March 31 is used as testing samples. The EEMD algorithm decomposes the PV time series with undulating characteristics, and the result of decomposition is shown in Fig. 4. IMF1 to IMF7 mirror the fluctuation feature of the sequence in different frequencies, and the remainder represents the trend. As shown in Fig. 4, although the original signal is very noisy, the components have apparent periodicity and smoothness after the frequency domain decomposition, which lays a foundation for the prediction of the stationary components. Then, the deterministic prediction can be quantified.

Concerning the testing samples, the IMFs are respectively predicted based on the deterministic predictions. The actual observations and the prediction values are represented by green curves and red curves,



(a) Combination of prediction results of IMFs (b) Prediction result of residual curve

Fig. 6. The result of deterministic prediction.

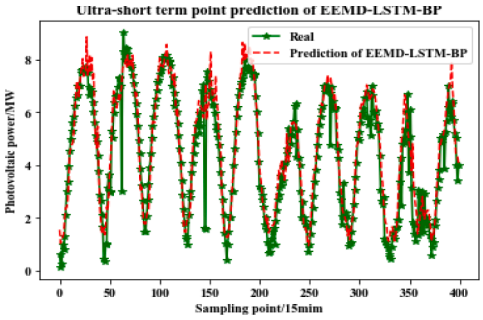


Fig. 7. The combination of deterministic prediction.

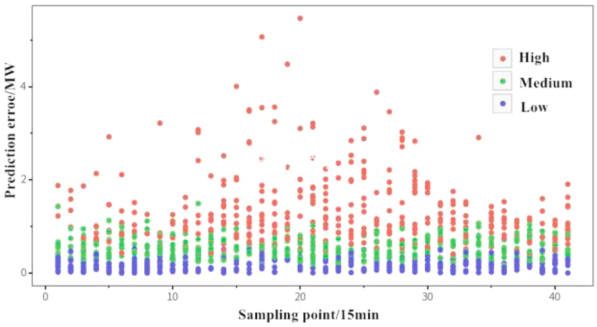


Fig. 8. Error distribution over different time periods.

respectively. The curve is reconstructed as IMFs with different frequencies. The prediction results are shown in Fig. 5.

Figs. 6 and 7 show the prediction results. Fig. 6 gives the prediction results of the IMFs and the residual reconstruction curve, respectively. Fig. 7 shows the overall prediction result according to the deterministic prediction.

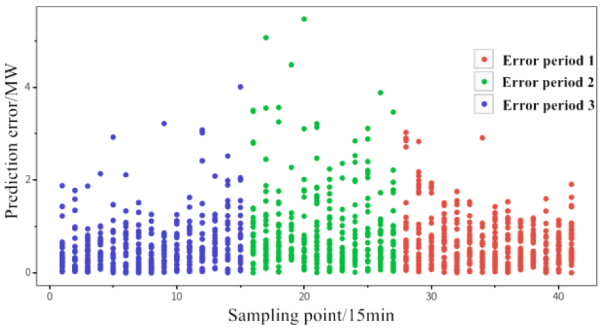


Fig. 10. Segment point partition diagram of error set.

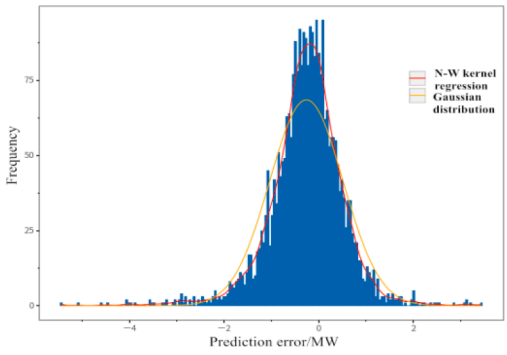
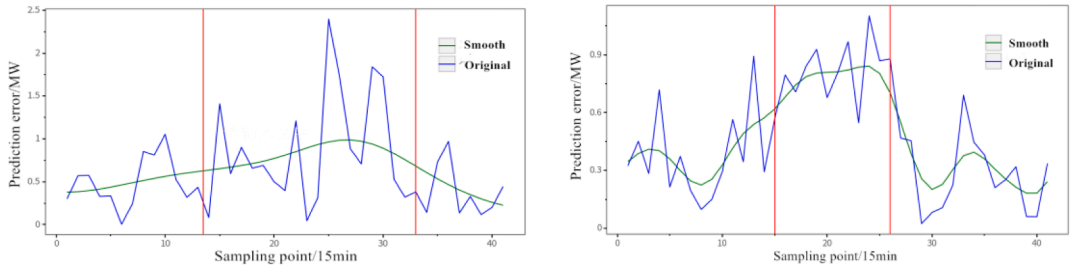


Fig. 11. The fitting curve of error distributions.

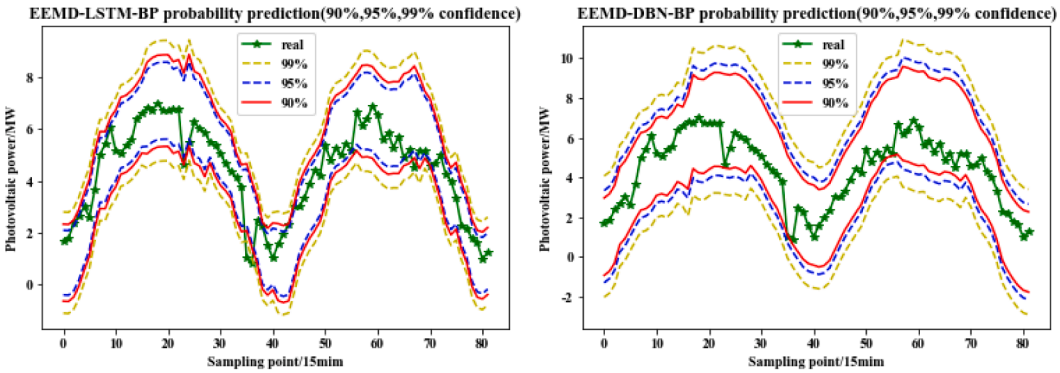
4.4. *Uncertainty analysis for multi-period*

By analysing the regional historical output and the samples of prediction errors on May 5 and 6, the central trend of power fluctuation in a certain period can be quantified. Thus, a time period error probability



(a) May 5 (b) May 6

Fig. 9. Multi-period point division for typical days.



(a) PIs based on EEMD-LSTM-BP (b) PIs based on EEMD-DBN-BP

Fig. 12. EEMD-LSTM-BP and EEMD-DBN-BP probability prediction.

Table 1
EEMD-LSTM-BP typical daily prediction error.

Confidence	PINAW/MW	PICP	Score
90%	2.5267	91.46%	-0.6296
95%	3.0108	95.12%	-0.3518
99%	3.9568	97.56%	-0.0791

Table 2
EEMD-DBN-BP typical daily prediction error.

Confidence	PINAW/MW	PICP	Score
90%	4.1847	96.34%	-0.8801
95%	4.9863	97.56%	-0.5171
99%	6.5531	98.78%	-0.1311

prediction model can be established.

Fig. 8 shows the distribution of error range identification for daily values from 8:00 to 18:00. According to the magnitude, prediction errors are equally classified into high, medium, and low values. The first 33% of the red high-error values exist broadly at noon. The maximum point of the trend spreads in the middle to both sides. Considering the size of daily PV observations, the daily values are divided into three time periods. The division of time periods for some typical days based on MPED approach is shown in Fig. 9.

For the optimal segment points of error periods, the kernel regression is utilized. The error periods are shown in Fig. 10, which are 8:00–11:45, 11:45–14:45, and 14:45–18:00, respectively, through the training dataset. For the testing days, the curve of error distributions fitted by N-W kernel regression and Gaussian distribution are shown in Fig. 11. From the result, the curve based on N-W kernel regression has better accuracy.

After fitting probability distribution, probability prediction is carried out from May 8 to May 9 on different confidences, including 90%, 95%, and 99%. In this paper, the proposed PIs are quantified based on EEMD-LSTM-BP, which is the key to enhancing probabilistic prediction performance. The accuracy of deterministic prediction affects the PIs coverage performance and sharpness. The PIs based on EEMD-DBN-BP is utilized as a benchmark for numerical comparison to reveal the proposed approach's effectiveness. According to the results given in Fig. 12, the prediction capability of EEMD-LSTM-BP at the abrupt fluctuation point in the noontime is better, and the PIs of EEMD-LSTM-BP have better sharpness, which reveals the necessity of LSTM in the prediction model.

To compare the PIs performances, Tables 1 and 2 show the numerical results. Obviously, the LSTM-based PIs have better reliability, sharpness, and overall performance compared to the DBN-based PIs.

The high volatility of PV may result in the low coverage of PIs, which reduces uncertainty analysis reliability, particularly under high PINC circumstances. To test the feasibility of MPED in this research paper, PIs based on multi-period analysis and PIs with uniform error algorithms instead of multi-period analysis are used for the numerical analysis. The EEMD-LSTM-BP and N-W kernel regression are utilized for both the PIs methods. The outputs from 12:00 to 14:45 on May 7 utilized as prediction target and the local PIs on PINC 95% are shown in Fig. 13. From the figure, the curve of output has high fluctuation. Thus, the width of PIs in this period is more comprehensive so that the PIs have high reliability. The numerical comparison is shown in Table 3, to reveal upper, lower bounds and whether PIs cover each actual value. From the

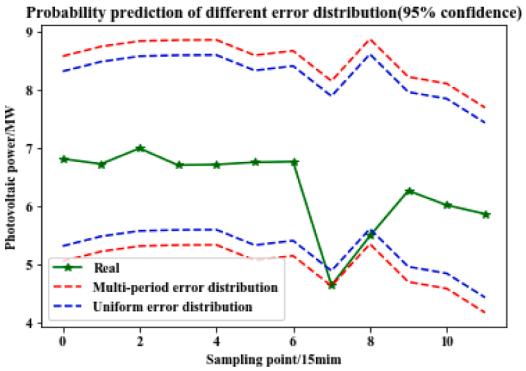


Fig. 13. PIs between 12:00–14:45 on May 7.

Table 3
The upper and lower bounds of PIs on PINC 95%.

Actual value (MW)	Multi-period analysis based upper bound (MW)	Multi-period analysis based lower bound (MW)	Uniform error based upper bound (MW)	Uniform error based lower bound (MW)	Covered?
6.82	8.5875	5.0543	8.3263	5.3155	Y/Y
6.73	8.7544	5.2212	8.4932	5.4824	Y/Y
7.00	8.8489	5.3156	8.5877	5.5769	Y/Y
6.71	8.8655	5.3323	8.6043	5.5935	Y/Y
6.72	8.8699	5.3367	8.6087	5.5979	Y/Y
6.76	8.6027	5.0694	8.3414	5.3307	Y/Y
6.77	8.6808	5.1475	8.4195	5.4087	Y/Y
4.64	8.1603	4.6271	7.8991	4.8883	Y/N
5.49	8.8849	5.3516	8.6236	5.6128	Y/N
6.27	8.2308	4.6975	7.9696	4.9588	Y/Y
6.02	8.1162	4.5830	7.8550	4.8442	Y/Y
5.87	7.6997	4.1665	7.4385	4.4277	Y/Y

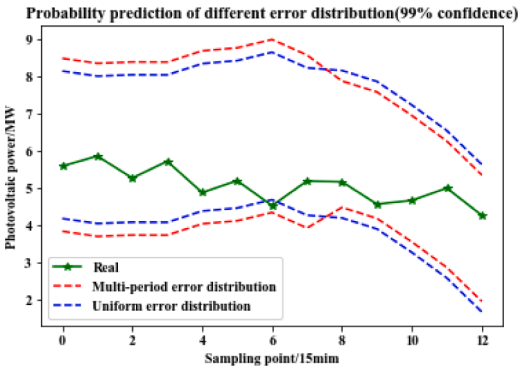


Fig. 14. PIs between 13:00–16:00 on May 8.

numerical result, in this period, the proposed PIs have better sharpness and reliability.

To further reveal the effectiveness of the PIs, the local PIs from 13:00 to 16:00 on May 8 are utilized for the comparison of probability prediction approaches. Fig. 14 shows the prediction results under PINC 99%. It is clearly demonstrated that the proposed PIs have better performance of coverage on PINC 99%, based on the advanced multi-time analysis.

5. Conclusion

A novel very short-term probability prediction approach has been proposed for PV based on MPED. By analysing the characteristic of PV output, the proposed method has reduced the width of local PIs and obtained good reliability, which has been analysed based on the numerical results. Besides, the numerical result has also revealed the high accuracy of deterministic prediction based on the methods proposed in this paper, which can provide a reasonable baseline for the PIs, which the numerical comparison has verified. Thus, on the high PINCs, it is valuable to make full use of reliable coverage characteristics of the proposed model.

Future research could pay close attention to the following two aspects. On the one hand, the effective screening of data input factors. On the other hand, the consideration of advanced clustering approaches.

CRediT authorship contribution statement

Sen Wang: Conceptualization, Investigation, Writing – original draft. **Yonghui Sun:** Conceptualization, Methodology, Writing – review & editing. **Shanming Zhang:** Investigation, Visualization. **Yan Zhou:** Supervision, Validation, Data curation. **Dongchen Hou:** Investigation, Software, Visualization. **Jianxi Wang:** Writing – review & editing.

Declaration of Competing Interest

The author(s) declared no potential conflicts of interest with respect to the research, author- ship, and/or publication of this article.

Acknowledgments

This work was partially supported by the National Key R&D Program of China (Technology and application of wind power/photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200), partially supported by the eponymous Complement S&T Program of State Grid Corporation of China (SGLNDKOOKJJS1800266) and partially supported by the National Natural Science Foundation of China under Grant 62073121.

References

[1] D.W. Vandermeer, J. Munkhammar, J. Widén, Probabilistic forecasting of solar power, electricity consumption and net load: investigating the effect of seasons, aggregation and penetration on prediction intervals, Sol. Energy 171 (Sep. 2018) 397–413.
[2] M.J. Sanjari, H.B. Gooi, N.C. Nair, Power generation forecast of hybrid PV-Wind system, IEEE Trans. Sustain. Energy 11 (2) (Apr. 2020) 703–712.
[3] Y. Sun, S. Zhai, H. Cui, D. Nan, K. Wang, Frequency regulation strategy for private EVs participating in integrated power system of REs considering adaptive Markov transition probability, Electr. Power Syst. Res. 173 (Aug. 2019) 291–301.
[4] Z. Wang, W. Wang, C. Liu, B. Wang, Forecasted scenarios of regional wind farms based on regular vine copulas, J. Mod. Power Syst. Clean. Energy 8 (1) (Mar. 2020) 77–85.
[5] X. Qing, Y. Niu, Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM, Energy 148 (1) (Apr. 2018) 461–468.
[6] K. Chen, Q. Huo, Training deep bidirectional LSTM acoustic model for LVCSR by a context-sensitive-chunk BPTT approach, IEEE-ACM Trans. Audio Speech Lang. 24 (7) (Mar. 2016) 1185–1193.
[7] S. Wen, C. Zhang, Y. Xu, Y. Tang, Y. Huang, A hybrid ensemble model for interval prediction of solar power output in ship onboard power systems, IEEE Trans. Sustain. Energy 12 (1) (Jan. 2021) 14–24.
[8] S. Sharda, M. Singh, K. Sharma, RSAM: robust self-attention based multi-horizon model for solar irradiance forecasting, IEEE Trans. Sustain. Energy 12 (2) (Apr. 2021) 1394–1405.
[9] K.Y. Bae, H.S. Hang, D.K. Sung, Hourly solar irradiance prediction based on support vector machine and its error analysis, IEEE Trans. Power Syst. 32 (2) (Mar. 2017) 935–945.
[10] A. Altamimi, D. Jayaweera, Reliability of power systems with climate change impacts on hierarchical levels of PV systems, Electr. Power Syst. Res. 190 (Jan. 2021) 106830–106843.
[11] X.G. Agoua, R. Girard, G. Kariniotakis, Probabilistic models for spatio-temporal photovoltaic power forecasting, IEEE Trans. Sustain. Energy 10 (2) (Apr. 2019) 780–789.
[12] M.J. Sanjari, H.B. Gooi, Probabilistic forecast of PV power generation based on higher order Markov chain, IEEE Trans. Power Syst. 32 (4) (Jul. 2017) 2942–2952.
[13] A.T. Eseye, J. Zhang, D. Zheng, Short-term photovoltaic solar power forecasting using a hybrid wavelet-PSO-SVM model based on SCADA and meteorological information, Renew. Energy 118 (Apr. 2017) 357–367.
[14] G. Zhang, et al., Wind power prediction based on variational mode decomposition multi-frequency combinations, J. Mod. Power Syst. Clean Energy 7 (2) (Mar. 2019) 281–288.
[15] P. Craven, G. Wahba, Smoothing noisy data with spline functions, Numer. Math. 31 (Dec. 1978) 377–403.
[16] C. Geng, F. Wang, J. Zhang, Z. Jin, Modal parameters identification of power transformer winding based on improved empirical mode decomposition method, Electr. Power Syst. Res. 108 (Mar. 2014) 331–339.
[17] Y. Ren, P. Suganthan, N. Srikanth, A comparative study of empirical mode decomposition based short-term wind speed forecasting methods, IEEE Trans. Sustain. Energy 6 (Jan. 2015) 236–244, vol. 1.
[18] S. Wang, Y. Sun, Y. Zhou, R.J. Mahfoud, D. Hou, A new hybrid short-term interval forecasting of PV output power based on EEMD-SE-RVM, Energies 13 (1) (Jan. 2020) 87–104.
[19] G. Memarzadeh, F. Keynia, Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm, Electr. Power Syst. Res. 192 (Mar. 2021) 106995–107009.
[20] Y. Sun, P. Wang, S. Zhai, D. Hou, S. Wang, Y. Zhou, Ultra short-term probability prediction of wind power based on LSTM network and condition normal distribution, Wind Energy 23 (Sep. 2019) 63–76.
[21] K. Sekaran, P. Chandana, N.M. Krishna, S. Kadry, Deep learning convolutional neural network (CNN) With Gaussian mixture model for predicting pancreatic cancer, Multimed. Tools Appl. 79 (15–16) (2019) 10233–10247.
[22] J. Tulensalo, J. Seppänen, A. Ilin, An LSTM model for power grid loss prediction, Electr. Power Syst. Res. 189 (Dec. 2020) 106823–106827.
[23] C. Ma, G. Dai, J. Zhou, Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM-BILSTM method, IEEE Trans. Intell. Transp. Syst. (2021), https://doi.org/10.1109/TITS.2021.3055258.
[24] J. Ospina, A. Newaz, M.O. Faruque, Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model, IET Renew. Power Gener. 13 (7) (May 2019) 1087–1095.
[25] P. Singh, A. Shankar, A novel optical image denoising technique using convolutional neural network and anisotropic diffusion for real-time surveillance applications, J. Real Time Image Process. (2021).
[26] N. Zhang, C. Kang, Q. Xia, J. Liang, Modeling conditional forecast error for wind power in generation scheduling, IEEE Trans. Power Syst. 29 (3) (May 2014) 1316–1324.
[27] H. Sheng, J. Xiao, Y. Cheng, Q. Ni, S. Wang, Short-term solar power forecasting based on weighted gaussian process regression, IEEE Trans. Ind. Electron. 65 (1) (Jun. 2017) 300–308.
[28] Y. Kuang, Q. Wu, Y. Wang, N. Dey, F. Shi, R.G. Crespo, R.S. Sherratt, Simplified inverse filter tracked affective acoustic signals classification incorporating deep convolutional neural networks, Appl. Soft Comput. 97 (2020), 106775.
[29] Z. Gan, C. Li, J. Zhou, G. Tang, Temporal convolutional networks interval prediction model for wind speed forecasting, Electr. Power Syst. Res. 191 (Feb. 2021) 106865–106875.

[30] M. Yang, C. Shi, H. Liu, Day-ahead wind power forecasting based on the clustering of equivalent power curves, Energy 281 (Mar. 2021) 119515–1119525.
[31] C. Wan, J. Wang, J. Lin, Y. Song, Z. Dong, Nonparametric prediction intervals of wind power via linear programming, IEEE Trans. Power Syst. 33 (1) (Jan. 2018) 1074–1076.
[32] J. Gao, H. Wang, H. Shen, Machine learning based workload prediction in cloud computing, in: 2020 29th International Conference on Computer Communications and Networks (ICCCN), 2020.
[33] Y. Zhou, Y. Sun, S. Wang, L. Bai, D. Hou, R.J. Mahfoud, P. Wang, A very short-term probabilistic prediction method of wind speed based on ALASSO-nonlinear quantile regression and integrated criterion, CSEE J. Power Energy Syst. (2021), https://doi.org/10.17775/CSEEJPES.2020.05370.