Ultra-Short-Term Solar PV Power Forecasting Method Based on Frequency-Domain Decomposition and Deep Learning

Lin Hu, Zhao Zhen, Fei Wang
Department of Electrical Engineering
North China Electric Power University
Baoding 071003, China
feiwang@ncepu.edu.cn

Gang Qiu, Yu Li
Dispatch and Control Center
State Grid Xinjiang Electric Power Co., Ltd
Urumqi 830018, China
qiugang412@163.com, liyu@xj.sgcc.com.cn

Miadreza Shafie-khah
School of Technology and Innovations
University of Vaasa
Vaasa 65200, Finland
mshafiek@univaasa.fi

João P. S. Catalão
Faculty of Engineering of the
University of Porto and INESC TEC
Porto4200-465, Portugal
catalao@fe.up.pt

2020 IEEE Industry Applications Society Annual Meeting | 978-1-7281-7192-0/20/$31.00 ©2020 IEEE | DOI: 10.1109/IAS44978.2020.9334889

***Abstract*—Ultra-short-term photovoltaic (PV) power forecasting can support the real-time dispatching of power grid and the optimal operation of PV power station itself. However, due to various meteorological factors, the photovoltaic power has great fluctuations. To improve the refined ultra-short-term forecasting technology of PV power, this paper proposes an ultra-short-term forecasting model of PV power based on optimal frequency-domain decomposition and deep learning. First, the amplitude and phase of each frequency sine wave is obtained by fast Fourier decomposition. As the frequency demarcation point is different, the correlation between the decomposition component and the original data is analyzed. By minimizing the square of the difference that the correlation between low-frequency components and raw data is subtracted from the correlation between high-frequency components and raw data, the optimal frequency demarcation points for decomposition components are obtained. Then convolutional neural network is used to predict low-frequency component and high-frequency component, and final forecasting result is obtained by addition reconstruction. Finally, the paper compares forecasting results of the proposed model and the non-spectrum analysis model in the case of predicting the 1 hour, 2 hours, 3 hours, and 4 hours. The results fully show that the proposed model improves forecasting accuracy.**

Keywords—PV power forecasting, ultra-short term, spectrum analysis, deep learning, frequency-domain decomposition

I. INTRODUCTION

1.1 Background and motivation

Facing the shortage of fossil energy and the deterioration of climate such as greenhouse effect, ozone hole and melting of polar glaciers, sustainable development of energy and environment has attracted worldwide attention. Due to the advantages of sustainability, clean and pollution-free, high flexibility, etc., PV power generation technology and related industries have experienced tremendous growth in the past few years [1-3]. PV power affected by various meteorological factors is highly uncertain, which will affect stable operation of the power grid [4-6]. PV power generation system is typically connected to the power grid to compensate for losses in conventional power generation systems. In a grid, some power generation systems that output constant power are called conventional power plants. Others convert their power in response to changes of PV generation and demand, which balances total power consumption with total power generation. These power plants that convert output power are called load-tracking power plants. When the output power of PV generation system is markedly increased or decreased, the load-tracking power plants must respond promptly. PV power forecast for the 15 min in advance enables the load-tracking power plant to react to sudden output changes timely. In order to stabilize the operation of the power grid, ultra-short-term PV power forecasting is particularly significant [7, 8].

1.2 literature review

Due to various meteorological factors, the photovoltaic power has great fluctuations [9-11]. It is difficult to achieve satisfactory results with traditional forecasting methods. In recent years, the method of first decomposing and then predicting photovoltaic power data has become a research hotspot. Among them, frequency-domain decomposition method excavates and extract PV power characteristics from the perspective of the frequency domain, which has become a current research hotspot. Literature [12] uses the variational mode decomposition (VMD) to decompose different frequency components from the historical PV power time series. Literature [13] decomposed the time series of solar photovoltaic power generation by wavelet decomposition. Literature [14] proposed a new forecasting model based on Hilbert Huang Transform (HHT) and integrating improved empirical mode decomposition (IEMD) with feature selection and forecasting engine. IEMD is used to decompose data. In the above literature, the data are only decomposed using existing models, but few literatures can support the rationality of the decomposition results through an effective method.

1.3 Contribution

To overcome the shortcomings that there is currently no effective method to support the rationality of the frequency domain decomposition results, this paper proposes an ultra-short-term forecasting model of PV power based on optimal frequency-domain decomposition and deep learning. First, the amplitude and phase of each frequency sine wave is obtained by fast Fourier decomposition. As the frequency demarcation point is different, the correlation between the decomposition component and the original data is analyzed. By analyzing the forecasting results of the decomposition components at different frequency demarcation points, the optimal frequency

demarcation points for decomposing low-frequency components and high-frequency components are obtained. Then convolutional neural network is used to predict low-frequency component and high-frequency component, and final forecasting result is obtained by addition reconstruction.

The main contributions of this paper include:

(1) By minimizing the square of the difference that the correlation between low-frequency components and raw data is subtracted from the correlation between high-frequency components and raw data, the optimal frequency demarcation points for decomposition components are obtained, which supports the rationality of the frequency-domain decomposition result.

(2) The method proposed in this paper is verified by the data of PV power station in China Ningxia. The results show that the proposed model improves the forecasting accuracy.

The rest of this article is organized as follows. Section 2 introduces the theory of FFT and deep learning. Section 3 introduces how to choose the optimal frequency demarcation point of frequency-domain decomposition. Section 4 introduces the model parameters and test results. In Section 6, conclusions are drawn.

II. Methodology

2.1 Fast Fourier Transform

Discrete Fourier Transform (DFT) can discretize the frequency domain of a finite-length sequence, but its computational complexity is too large to process the problem in real time, thus leading to the Fast Fourier Transform (FFT) [15-17]. FFT decomposes the sequence of N into the two lines sequence of $N/2$. The subscript of one sequence is even, the subscript of the other sequence is odd. The DFT of two short sequences is the DFT of the original sequence. DFT. Repeat the above decomposition principle until it is decomposed into a sequence of n rows and columns. The FFT uses the periodicity and symmetry to improve the DFT algorithm, which greatly reduces the amount of computation. Compared with DFT, the FFT is uncomplicated and speedy.

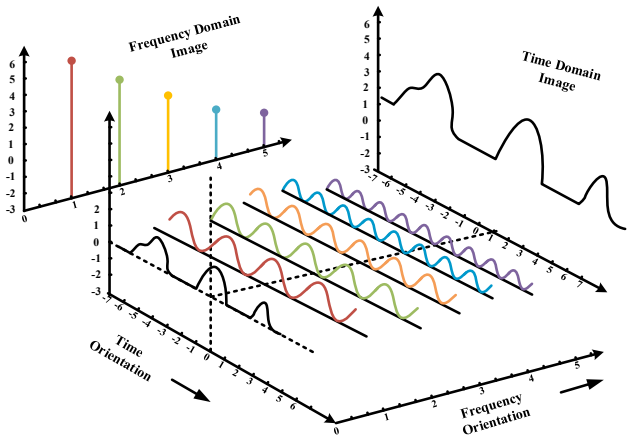


Fig. 1.Relationship between time domain signals and sine wave signals of different frequencies

Fig. 1 shows the relationship between the time domain signal and the sine wave signals of different frequencies. The black curve represents a piece of PV power and the color curve represents the sine wave signals of various frequencies that make up this time domain signal. The sine wave is the signal with the most single frequency component. Any complex signal can be seen as a composite of many sine waves with various frequencies and amplitudes. It can be considered that sine wave is the basis of all waveforms. Therefore, this paper uses FFT to decompose PV power in the frequency domain.

Discretized sequence decomposed into signals of multiple frequencies as shown in Eq. (1). According to Eq. (1), the sine wave and the cosine wave of the same frequency are superimposed by different coefficients to generate cosine waves of various phases of the same frequency. The modulus c_n of complex number in the frequency domain obtained by FFT represents the energy of the cosine wave corresponding frequency, and the angle of the complex number represents the phase θ_n of the cosine wave, thereby obtaining the amplitude spectrum and the phase spectrum. The actual amplitude a_k of the k-th point of the cosine is defined by Eq. (2). The physical frequency f_k of the k-th point is calculated as shown in Eq. (1).

$$\begin{aligned} f(n) &= a_0 + \sum_{n=1}^{N}\left[a_n \cos(2\pi f_n) + b_n \cos(2\pi f_n)\right] \\ &= c_0 + \sum_{n=1}^{N} c_n \cos(2\pi f_n + \theta_n) \end{aligned} \tag{1}$$

$$a_k = \frac{2c_k}{N} \tag{2}$$

$$f_k = k * \frac{f_s}{N} \tag{3}$$

In the formula, n is a positive integer. a_0 is constant component, a_n is the cosine component amplitude and b_n is the sinusoidal component amplitude. Where c_0 is the constant component, c_n is the amplitude, $f_n = \frac{2\pi n \Delta t}{N}$ is the frequency, θ_n is the phase, N is a positive integer, and Δt is the sampling interval.

2.2 Pearson correlation coefficient

Pearson's correlation coefficient (R) represents the curve fit of two sequences. Specifically, the critical criterion for R is a very strong correlation between 0.8 and 1.0, a strong correlation between 0.6 and 0.8, a medium correlation between 0.4 and 0.6, a weak correlation between 0.2 and 0.4, and very weak correlation or no correlation between 0 and 0.2[18]. R^2 is defined as shown in Eq. (4):

$$R_{X,Y} = \frac{N\sum XY - \sum X \sum Y}{\sqrt{N\sum X^2 - \left(\sum X\right)^2}\sqrt{N\sum Y^2 - \left(\sum Y\right)^2}} \tag{4}$$

2.3 Convolutional neural network

Convolutional Neural Network (CNN) is a common deep learning model with powerful feature extraction capabilities. CNN generally consists of five types of cell layers, namely the

input layer, the convolutional layer, the pooling layer, the fully connected layer, and the output layer [19]. The structure of CNN is shown in Fig. 2.

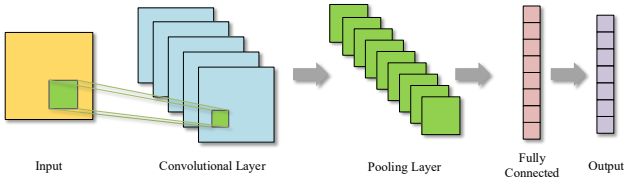


Fig. 2.The structure of CNN

2.3.1 *Convolution layer*

The role of the convolutional layer is to extract features from the input information. The convolutional layer is typically composed of multiple convolution kernels, each of which is used to compute a feature map. Each cell of the feature map is connected to the region of the adjacent cell in the previous layer. Convolution of the input by the convolution kernel and nonlinear processing of convolution results by activation function can acquire the new feature map. The formula for the convolutional layer is shown by Eq. (5)

$$y_{i,j,k}^{l} = \mathrm{F}((w_k^l)^T x_{i,j}^l + b_k^l) \tag{5}$$

Where w_k^l and b_k^l are the weights and bias of the *k*-th convolution kernel in the *l*-th convolutional layer, respectively. $x_{i,j}^l$ is the input information of (i,j) region in the *l*-th convolution layer. The weight w_k^l in the *l*-th convolution layer is shared by each region of the input information, which is weights sharing. $F(\bullet)$ is the activation function applied to the convolution layer, which can effectively improve the fitting ability of the model.

2.3.2 *Pooling layer*

The role of the pooling layer is to reduce the size of the feature map generated by the convolution layer, and to effectively extract the feature information in the feature map. The formula for poor layer is shown by Eq. (6)

$$P_{i,j,k}^{l} = pool(y_{m,n,k}^{l}) \tag{6}$$

Where $(m,n) \in R_{i,j}$, and $R_{i,j}$ are the information at location (i,j).

2.3.3 *Fully connected layer*

The function of the fully connected layer is to summarize the distributed feature representations learned by the previous layer into the same space for subsequent applications. All the neurons in the previous layer are connected to each neuron in the current layer.

2.3.4 *Description of one-dimensional CNN*

One-dimensional(1D) CNN is a branch of CNN[20]. The convolution kernel window of CNN1D slides in a single direction (i.e. time step). Because photovoltaic power is time series data, this paper chooses CNN1D to build forecasting model. In the 1D convolution layer, the size of kernel is 3 and the corresponding stride is 1. In the 1D pooling layer, the size of kernel is 2 and the corresponding stride is 1. The specific flow chart is shown in Fig. 3.

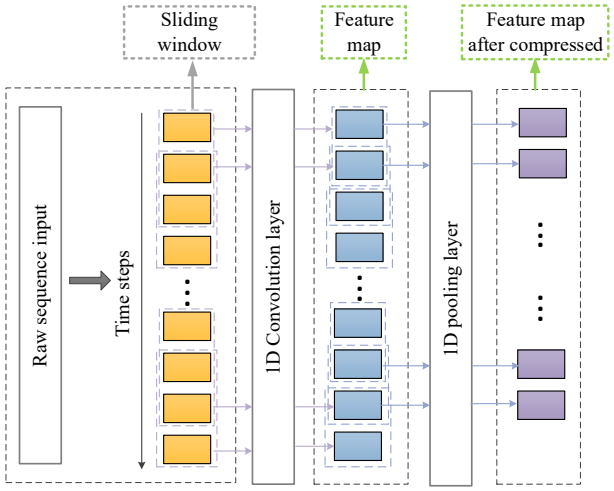


Fig. 3.Illustration of 1D convolution and 1D pooling

2.4 *Combined Model Framework*

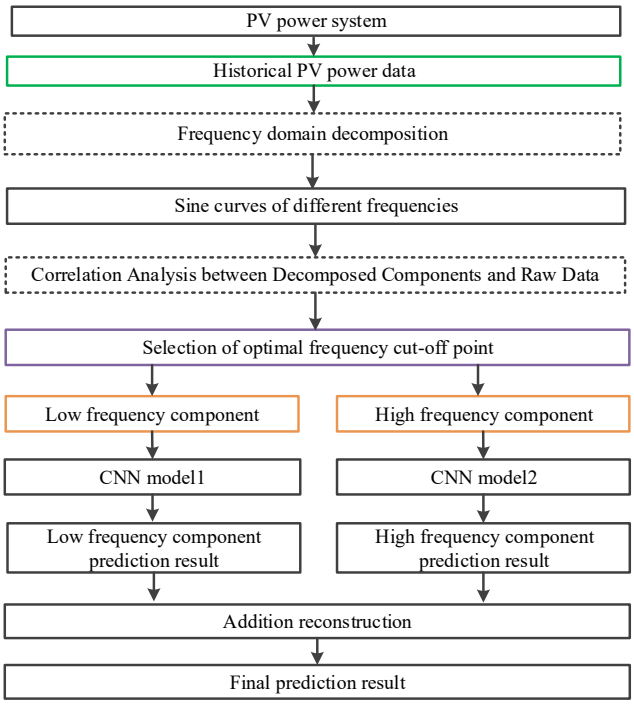


Fig. 4.Ultra-short-term forecasting model of PV power based on optimal frequency-domain decomposition and deep learning

The framework of ultra-short-term forecasting model of PV power based on optimal frequency-domain decomposition and deep learning proposed in this paper is shown in Fig. 4. First, the amplitude and phase of each frequency sine wave is obtained by fast Fourier decomposition. As the frequency demarcation point is different, the correlation between the decomposition component and the original data is analyzed. By minimizing the square of the difference that the correlation between low-frequency components and raw data is subtracted from the correlation between high-frequency components and raw data, the optimal frequency demarcation

points for decomposition components are obtained. Then convolutional neural network is used to predict low-frequency component and high-frequency component, and final forecasting result is obtained by addition reconstruction.

III. Spectrum Analysis

3.1 Frequency domain decomposition

This paper selects the PV data of Ningxia PV Power Station in China from 0:00 on January 1, 2017 to 23:45 on December 31, 2017 to train the model, and perform frequency domain decomposition through FFT to obtain the amplitude spectrum and a part of the phase spectrum, which are shown in Fig. 5 and Fig. 6, respectively.

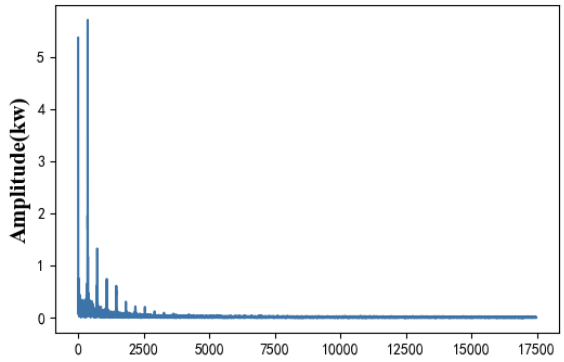


Fig. 5. Amplitude spectrum of PV power

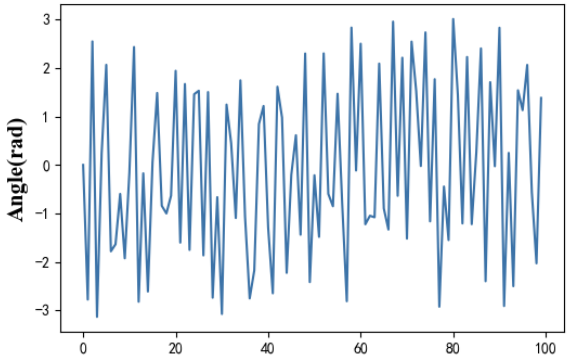


Fig. 6. Phase spectrum of PV power

3.2 Selection of optimal frequency demarcation point

Through FFT spectrum analysis, PV power can be decomposed into low-frequency components and high-frequency components. The low-frequency component represents the regular part of PV power, which indicates its trend characteristics, while the high frequency component represents the randomness of PV power, which indicates its fluctuation characteristics affected by other factors such as weather. How to accurately decompose low-frequency components and high-frequency frequencies has always been a difficult problem. In order to solve this problem, this article chooses the optimal frequency demarcation point from the perspective of the correlation between decomposed data and raw data. Fig. 7 shows the correlation between decomposed components and raw data. When frequency-domain decomposition is performed on PV power data, the more frequencies selected, the stronger the correlation between low-frequency components and raw data, and the weaker the correlation between high-frequency components and raw data.

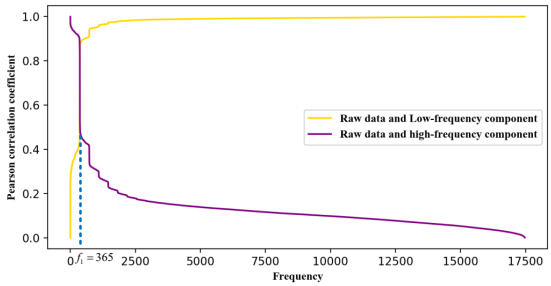


Fig. 7. Correlation between decomposition components and raw data before weighting

To illustrate the effect of the correlation between the decomposition component and raw data on the forecasting results, this paper uses the CNN model to predict low-frequency components and high-frequency components from 0:00 on January 1, 2017 to 23:45 on December 31, 2017. Table I compares forecasting results of low-frequency components at different frequencies. Table II compares the forecasting results of high-frequency components at different frequencies. The frequency demarcation point selected in Tables I and Table II are frequency nodes with relatively large amplitudes in amplitude spectrum of Fig. 5. As the correlation between the low-frequency component and raw data increases, the curve fitting effect becomes better. As the correlation between the high-frequency component and raw data decreases, the curve fitting effect becomes worse. The relationship between the selected frequencies and the correlation of decomposition component and raw data shows that as the frequency number increases, the better forecasting result of the low-frequency component and the worse forecasting result of the high-frequency component, so the two are contradictory. Low-frequency components represent the regular part of PV power which can be accurately predicted, while high-frequency components are relatively random, which are difficult to predict. If the proportion of low-frequency components in raw data can be increased, accurate forecasting of low-frequency components to balance forecasting errors of high-frequency components will effectively improve the overall forecasting accuracy. Therefore, the idea of optimal frequency demarcation point selection proposed in this article is that low-frequency component consider to select the highest frequency demarcation point as possible, and accounts for a large proportion, and then the second-level is that high-frequency component, which hope the frequency selection is not too high. Otherwise, forecasting is too difficult to get a certain level of balance.

From Fig. 7, it can be seen that when the frequency reaches a certain value, as the frequency band increases, the correlation between low-frequency components and raw data does not increase significantly. It can be considered that there is no significant increase in the correlation between low-frequency components and raw data, while the correlation between high-frequency component and raw data continues to decline. This optimal frequency demarcation point that balances both low-frequency components and high-frequency components not only ensures sufficient extraction of low-frequency components, but also does not cause the difficulty

of predicting high-frequency components to an unacceptable level due to excessive extraction of low-frequency components. In this paper, the optimal frequency is solved by minimizing the square of the difference that the correlation between low-frequency components and raw data is subtracted from the correlation between high-frequency components and raw data, which can be regarded as an optimization problem. The specific formula is given in Eq. (7).

$$\begin{aligned} &\min F(f)=\left[L(f)-H(f)\right]^2 \\ &s.t. \quad f\in\left[0,\frac{N}{2}\right] \end{aligned} \tag{7}$$

Among them, $F(f)$ is an objective function. When the objective function reaches the minimum value, the correlation between low-frequency components and raw data and the correlation between high-frequency components and raw data reach an optimal balance. $L(f)$ is the correlation between low-frequency components and raw data, $H(f)$ is the correlation between high frequency components and raw data, and N is the total number of sampling frequencies.

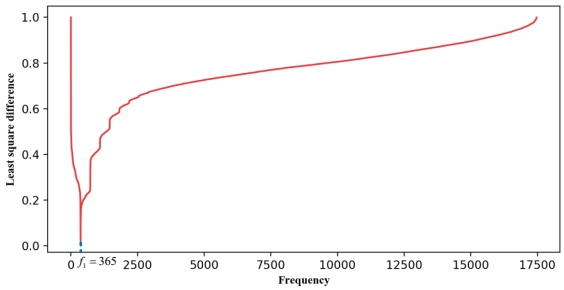


Fig. 8. The graph of the squared difference before weighting

Fig. 8 is a graph of the obtained square of the difference that the correlation between low-frequency components and raw data is subtracted from the correlation between high-frequency components and raw data. It can be seen from the figure that when the frequency demarcation point f_1 is 365, the correlation between low-frequency components and raw data and the correlation between high-frequency components and raw data reach a balance. However, due to the strong regularity and high forecasting accuracy of low-frequency components, this article first hopes that the proportion of low-frequency components is high. When the frequency value continues to increase, the correlation between the low-frequency components and raw data can be improved. However, when the correlation between the low-frequency components and raw data is basically unchanged, the forecasting accuracy of low-frequency components basically reaches the maximum value. At this time, the accuracy of final forecasting result mainly depends on high-frequency components. Therefore, in order to give priority to low-frequency components, the frequency value is continuously increased. This article adds weight values to the objective function in Eq. (7). By multiplying the correlation between low-frequency components and raw data by a low weight value and multiplying the correlation between high-frequency components and raw data by a high weight value. In this way, the design idea of giving priority to low-frequency components is achieved. The new objective function is shown by Eq. (8). This paper chooses $\alpha = 0.173$, $\beta = 0.827$, and obtains the frequency demarcation point $f_2 = 1825$.

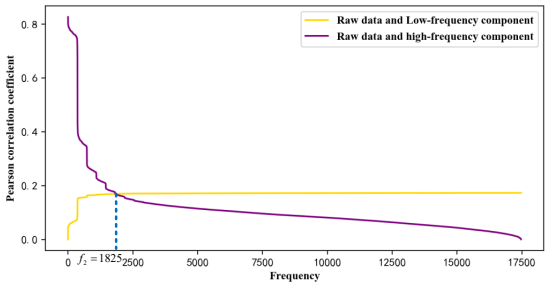


Fig. 9. Correlation between decomposition components and raw data after weighting

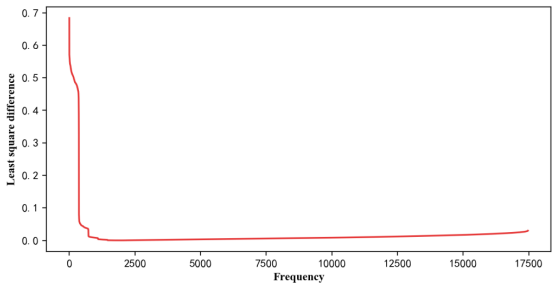


Fig. 10. The graph of the squared difference after weighting

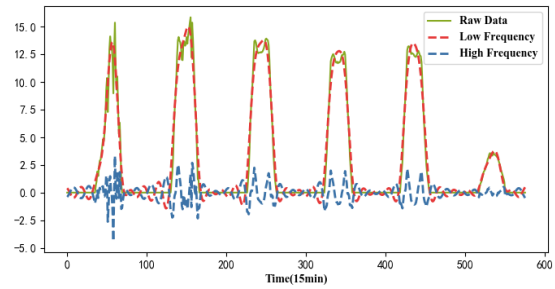


Fig. 11. Optimal frequency decomposition result

Fig. 10 shows the square of the difference that the correlation between low-frequency components and raw data is subtracted from the correlation between high-frequency components and raw data after weighting. It can be seen that after f_2, as the frequency band increases, the correlation between low-frequency components and raw data does not increase significantly, while the correlation between high-frequency components still maintains a downward trend. Therefore, f_2 is the optimal frequency demarcation point f_o. Fig. 11 shows the decomposition results at frequency demarcation point f_o.

$$\begin{aligned} &\min F(f)=\left[\alpha L(f)-\beta H(f)\right]^2 \\ &s.t. \quad f\in\left[0,\frac{N}{2}\right] \end{aligned} \tag{8}$$

Where α is the weight of the correlation between the low frequency component and the original data, and β is the weight of the correlation between the high frequency component and the original data.

TABLE. I. COMPARISON OF R OF LOW FREQUENCY COMPONENTS AT DIFFERENT FREQUENCIES

Time Scale of forecasting	Evaluation indexes					
	365	***730***	***1092***	***1458***	***1825***	***2185***
1h	0.9243	0.9654	0.9792	0.9994	0.9997	0.9998
2h	0.9067	0.9534	0.9635	0.9947	0.9980	0.9984
3h	0.8286	0.8438	0.8912	0.9032	0.9132	0.9245
4h	0.7395	0.7627	0.7821	0.7932	0.8060	0.8125

TABLE. II. COMPARISON OF FORECASTING RESULTS R2 OF HIGH FREQUENCY COMPONENTS AT DIFFERENT FREQUENCIES

Time Scale of forecasting	Evaluation indexes					
	365	***730***	***1092***	***1458***	***1826***	***2186***
1h	0.6722	0.6584	0.4783	0.2947	-0.1446	-0.1876
2h	0.6612	0.6457	0.4378	0.2653	-0.3992	-0.5743
3h	0.6576	0.6369	0.3989	0.2375	-0.7690	-1.0467
4h	0.6378	0.6128	0.3648	0.2014	-1.3688	-1.5674

TABLE. III. ERROR INDEXES OF LOW FREQUENCY COMPONENT FORECASTING AT DIFFERENT TIME SCALES

MODEL	Evaluation indexes		
	MAE	***RMSE***	***R***
1h	0.0683	0.0860	0.9997
2h	0.3532	0.4520	0.9980
3h	1.0509	1.3302	0.9232
4h	1.4887	2.0172	0.8060

TABLE. IV. ERROR INDEXES OF HIGH FREQUENCY COMPONENT FORECASTING AT DIFFERENT TIME SCALES

MODEL	Evaluation indexes		
	MAE	***RMSE***	***R***
1h	0.4618	0.8336	-0.1446
2h	0.4864	0.8782	-0.3992
3h	0.5467	0.9453	-0.7690
4h	0.5712	0.9867	-1.3688

TABLE. V. FORECAST 1 HOUR RESULT ERROR

MODEL	Evaluation indexes		
	MAE	***RMSE***	***R***
No frequency decomposition	1.1001	2.1467	0.7654
Frequency decomposition	0.4011	0.8198	0.9695

TABLE. VI. FORECAST 2 HOUR RESULT ERROR

MODEL	Evaluation indexes		
	MAE	***RMSE***	***R***
No frequency decomposition	1.6755	2.8731	0.5864
Frequency decomposition	0.4875	0.9297	0.9610

V. CASE STUDY

4.1 Simulation data and simulation platform

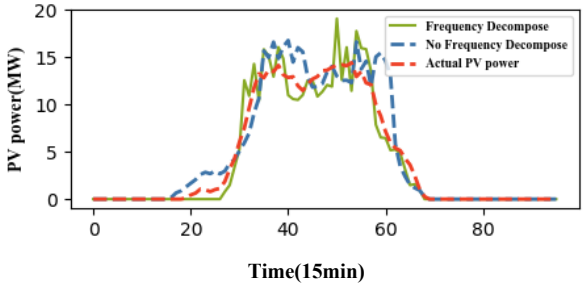
The data used for the simulation in this paper are historical weather data and historical power with 15 min time resolution from 0:00 on January 1, 2017 to 23:45 on January 31, 2018 in Ningxia PV Power Station in China. This article uses the data from 0:00 on January 1, 2017 to 23:45 on December 31, 2017 as training set, and the data from January 1, 2018 0:00 to January 31, 2018 23:45 as testing set. The time scale of input data is 24 hours, and the time scale of output data is 1 hour, 2 hours, 3 hours, and 4 hours.

We use Python 3.6.1 with Tensorflow and scikit-learn to perform all the simulations.

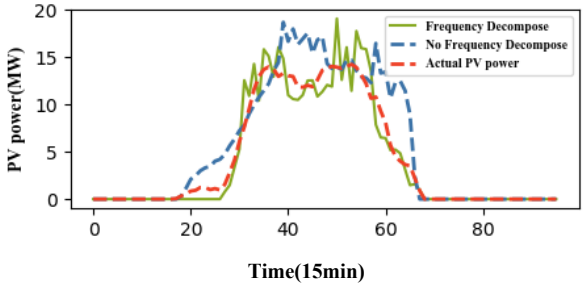
4.2 Data processing

During model training, to eliminate the difference in magnitude between each dimension data, the samples of input data will be normalized. The range of values of all samples is converted to [0, 1], which avoids large differences in sample magnitude leads to large network forecasting errors. Its formula is shown in (9):

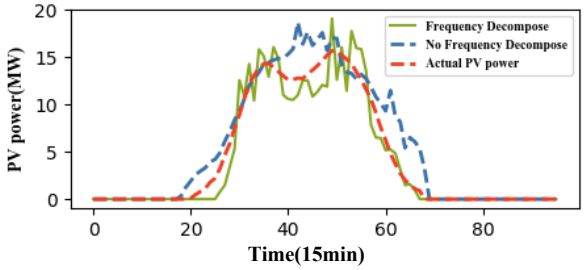
$$M^{*} = \frac{M - M_{\min}}{M_{\max} - M_{\min}} \tag{9}$$



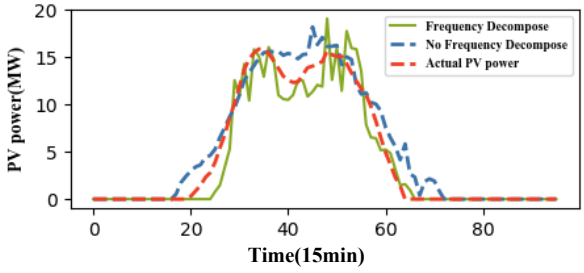
(1) 1hour



(2) 2 hour



(3) 3 hour



(4) 4 hour

Fig. 12. Final forecasting results at different scales

4.3 Performance criterion

In order to evaluate the performance of the forecasting model, we employ two effective error indexes that are Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Under the same set of training data, the smaller MAE and RMSE, together with the higher R, and the better the forecasting model effect. The mathematical formulas for the two error indexes are shown in (10) and (11):

$$MAE = \frac{\sum_1^N |y_t - \hat{y}_t|}{N} \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{t-1}^N (y_t - \hat{y}_t)^2}{N}} \tag{11}$$

4.4 Test results

4.4.1 Decomposition component forecasting results

Table III and Table IV show forecasting error indexes of low-frequency and high-frequency components at different time scales. From the results, it can be seen that as the number of hours increases, MAE and RMSE also increase, and forecasting accuracy decreases. This shows that as the time scale increases, the difficulty of forecasting increases.

4.4.2 Ultimate forecasting results

Fig. 12 shows final forecasting results at different time scales, and Tables V-VIII shows the comparison of error indicators with and without frequency-domain decomposition model. As the number of hours increases, compared with the model without frequency-domain decomposition, MAE of the proposed model has improved accuracy by 63.54%, 70.90%, 43.23% and 32.91%, MAE of proposed model has improved accuracy by 61.81%, 67.64%, 49.61% and 33.70%, and the R of proposed model has improved accuracy by26.67%, 63.88%, 83.56% and 78.47% respectively. It can be seen that proposed frequency domain decomposition method improves forecasting accuracy. High-frequency component is more volatile after the absence of low-frequency components, which may lead to reduced forecasting accuracy of high-frequency component. However, the method of selecting the optimal frequency demarcation point proposed in this paper, which not only improves the proportion of low-frequency components in raw data, but also balances the difficulty of predicting high-frequency components to a certain extent. Therefore, after adding and reconstructing low frequency component and high frequency component, the accuracy of ultimate results obtained is generally improved.

VI. Conclusion

To overcome the shortcomings that there is currently no effective method to support the rationality of the frequency domain decomposition results, this paper proposes an ultra-short-term forecasting model of PV power based on optimal frequency-domain decomposition and deep learning. First, the amplitude and phase of each frequency sine wave is obtained by fast Fourier decomposition. As the frequency demarcation point is different, the correlation between the decomposition component and the original data is analyzed. By minimizing the square of the difference that the correlation between low-frequency components and raw data is subtracted from the correlation between high-frequency components and raw data, the optimal frequency demarcation points for decomposition components are obtained. Then convolutional neural network is used to predict low-frequency component and high-frequency component, and final forecasting result is obtained by addition reconstruction. The results fully show that the ultra-short-term forecasting model of photovoltaic power proposed by the optimal frequency domain decomposition and deep learning improves the forecasting accuracy.

In the future, the research on utilizing this proposed ultra-short-term solar PV power forecasting model to electric vehicle charging demand management [21, 22], demand response [23-25], energy storage management [26, 27], load pattern [28-30], aggregator aggregated capacity forecasting and energy trading in nanogrid [31-34] will be conducted.

ACKNOWLEDGMENT

This work was supported by the National Key R&D Program of China (Technology and application of wind power/photovoltaic power forecasting for promoting renewable energy consumption, 2018YFB0904200) and eponymous Complement S&T Program of State Grid Corporation of China (SGLNDKOOKJJS1800266). Also, João P. S. Catalão acknowledges the support by FEDER funds through COMPETE 2020 and by Portuguese funds through FCT, under POCI-01-0145-FEDER-029803 (02/SAICT/2017).

REFERENCES

[1] Y. Wu, S. Chang, and P. Mandal, "Grid-Connected Wind Power Plants: A Survey on the Integration Requirements in Modern Grid Codes," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 5584-5593, Nov.-Dec. 2019.

[2] Y. Wu, J. Lin, and H. Lin, "Standards and Guidelines for Grid-Connected Photovoltaic Generation Systems: A Review and Comparison," *IEEE Trans. Ind. Appl.*, vol. 53, no. 4, pp. 3205-3216, Jul.-Aug. 2019.

[3] Y. Wu, P. Sun, T. Wu, J. Hong, and M.Y. Hassan, "Probabilistic Wind-Power Forecasting Using Weather Ensemble Models," *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 5609-5620, Nov.-Dec. 2018.

[4] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, "A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern forecasting framework," *Energy Convers. Manag.*, vol. 212, Art. no. 112766, May. 2020.

[5] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duić, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Appl. Energy*, vol. 253, Art. no. 113595, Nov. 2019.

[6] Z. Zhen, J. Liu, Z. Zhang, F. Wang, H. Chai, Y. Yu, X. Lu, T. Wang, and Y. Lin, "Deep learning based surface irradiance mapping model for solar PV power forecasting using sky image," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3385-3396, Jul.-Aug. 2020.

[7] F. Wang, Z. Xuan, Z. Zhen, Y. Li, K. Li, L. Zhao, M. Shafie-khah, and J. P.S. Catalão, "A minutely solar irradiance forecasting method based on real-time sky image-irradiance mapping model," *Energy Convers. Manag.* vol. 220, Art. no. 113075, Sep. 2020.

[8] Y. Wu, W. Tan, S. Huang, Y. Chiang, C. Chiu, and C. Su, "Impact of Generation Flexibility on the Operating Costs of the Taiwan Power System Under a High Penetration of Renewable Power," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, pp. 2348-2359, May-Jun. 2020.

[9] H. Yang, C. Huang, Y. Huang, Y. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp: 917-926, Jul. 2014.

[10] J. Shi, W. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Trans. Ind. Appl.*, vol. 48, no. 3, pp: 1064-1069, May-Jun. 2012.

[11] P. Lin, Z. Peng, Y. Lai, S Cheng, Z. Chen, and L. Wu, "Short-term power forecasting for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets," *Energy Convers. Manag.*, vol. 177, pp. 704-717, Dec. 2018.

[12] H. Zhang, L. Cheng, T. Ding, K. Cheung, Z. Liang, Z. Wei, and G. Sun, "Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network," *IET Gener. Tran. & Distri.*, 2018, vol. 20, 4557-4567.

[13] V. Kushwaha, N. Pindoriya, "A SARIMA-RVFL hybrid model assisted by wavelet decomposition for very short-term solar PV power generation forecast,". *Renewable Energy*, vol. 140, pp: 124-139, 2019.

[14] W. Zhang, H. Dang, R. Simoes, "A new solar power output forecasting based on hybrid forecast engine and decomposition model," *ISA Trans.*, vol. 81, pp. 105-120, 2018.

[15] M. Alam, S. Rehman, L. Al-Hadhrami, J. Meyer, "Extraction of the inherent nature of wind speed using wavelets and FFT," *Energy for Sustainable Development*, vol. 22, pp. 34-47, 2014.

[16] M. Yang, X. Huang, "Ultra-short-term forecasting of photovoltaic power based on periodic extraction of PV energy and LSH algorithm," *IEEE Access*, vol. 6, pp: 51200-51205, 2018.

[17] M. Yang, X. Huang, "An evaluation method of the photovoltaic power forecasting quality," *Mathematical Problems in Engineering*, 2018.

[18] T. Kim, S. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp: 72-81, 2019.

[19] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, K. Li, "Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting," *Appl. Sciences*, vol. 8, no.8, Art. no. 1286, 2018.

[20] F. Wang, Z. Zhang, C. Liu, Y. Yu, S. Pang, N. Duic, M. Shafie-Khah, and JPS. Catalao, "Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting," *Energy Convers. Manag.*, vol. 181, pp: 443–462, 2019.

[21] F. Teng, Z. Ding, Z. Hu, and P. Sarikprueck, "Technical Review on Advanced Approaches for Electric Vehicle Charging Demand Management, Part I: Applications in Electric Power Market and Renewable Energy Integration," *IEEE Trans. Ind. Appl.*, Early Access, 2020, DOI: 10.1109/TIA.2020.2993991.

[22] Z. Ding, Y. Lu, L. Zhang, W. Lee, and D. Chen,, "A Stochastic Resource-Planning Scheme for PHEV Charging Station Considering Energy Portfolio Optimization and Price-Responsive Demand," *IEEE Trans. Ind. Appl.*, vol.54, no.6, pp.5590-5598, Nov.-Dec. 2018.

[23] Z. Ding, F. Zhu, Y. Wang, Y. Lu, and L. Zhang, "Flexible Demand Resource Pricing Scheme: A Stochastic Benefit-Sharing Approach," *IEEE Trans. Ind. Appl.*, vol.55, no.6, pp.5554-5563, Nov.-Dec. 2019.

[24] K. Li, L. Liu, F. Wang, T. Wang, N. Dui, M. Shafie-khah and J. P. S. Catalão, "Impact factors analysis on the probability characterized effects of time of use demand response tariffs using association rule mining method," *Energy Convers. Manag.*, vol. 197, Art. no. 111891, Oct. 2019.

[25] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalao, "Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6972–6985, Nov. 2018.

[26] F. Hafiz, M.A. Awal, A.R. de Queiroz, and I. Husain, "Real-Time Stochastic Optimization of Energy Storage Management Using Deep Learning-Based Forecasts for Residential PV Applications," *IEEE Trans. Ind. Appl.*, vol.56, no.3, pp. 2216-2226, May.-Jun. 2020.

[27] T. Ku, C. Chen, C. Lin, C. Hsu, and H. Chuang, "Transformer Management System for Energy Control of Customer Demand Response and PV Systems," *IEEE Trans. Ind. Appl.*, vol.55, no.1, pp. 51-59, Jan.-Feb. 2019.

[28] K. Li, X. Cao, X. Ge, F. Wang, X. Lu, M. Shi, R. Yin, Z. Mi, and S. Chang, "Meta-heuristic optimization based two-stage residential load pattern clustering approach considering intra-cluster compactness and inter-cluster separation," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3375-3384, Jul.-Aug. 2020.

[29] F. Wang, K. Li, N. Duić, Z. Mi, B.M. Hodge, M. Shafie-khah, and J. P. S Catalão, "Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns," *Energy Convers. Manag.*, vol. 171, pp. 839-854, Sep. 2018.

[30] L. Zhao, Z. Yang, and W. Lee, "The Impact of Time-of-Use (TOU) Rate Structure on Consumption Patterns of the Residential Customers," *IEEE Trans. Ind. Appl.*, vol. 53, no. 6, , pp. 5130-5138, Nov.-Dec. 2017.

[31] S. Yan, K. Li, F. Wang, X. Ge, X. Lu, Z. Mi, H. Chen, and S. Chang, "Time-frequency features combination-based household characteristics identification approach using smart meter data," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, , pp. 2251-2262, May-Jun. 2020.

[32] F. Wang, B. Xiang, K. Li, X. Ge, H. Lu, J. Lai, and P. Dehghanian, "Smart households' aggregated capacity forecasting for load aggregators under incentive-based demand response programs," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1086-1097, Mar.-Apr. 2020.

[33] Z. Zhang, H. Tang, P. Wang, Q. Huang, and W. Lee, "Two-Stage Bidding Strategy for Peer-to-Peer Energy Trading of Nanogrid," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1000-1009, Mar.-Apr. 2020.

[34] H. Xu, Y. Lin, X. Zhang, and F. Wang, "Power system parameter attack for financial profits in electricity markets," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3438-3446, Jul. 2020, DOI: 10.1109/TSG.2020.2977088.