IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 5, 2020, accepted September 20, 2020, date of publication September 22, 2020, date of current version October 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3025967

Wind Speed Forecasting Based on Extreme Gradient Boosting

REN CAI[1], SEN XIE[1,2], BOZHONG WANG[3], RUIJIANG YANG[4], DAOSEN XU[4], AND YANG HE[4]

[1]College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060, China
[2]College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, China
[3]State Grid Hunan Electric Power Corporation Maintenance Company, Changsha 410082, China
[4]College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

Corresponding author: Sen Xie (senxie@szu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61903257.

ABSTRACT As the integration of wind power into electrical energy network increasing, accurate forecast of wind speed becomes highly important in the case of large-scale wind power connected into the grid. In order to improve the accuracy of wind speed forecast and the generalization ability of the model, Extreme Gradient Boosting (XGBoost) as an improvement from gradient boosting decision tree (GBDT) is trained and deployed in the cheaper central processing unit (CPU) devices instead of graphics processing unit (GPU) devices, thus, a wind speed forecast model based on Extreme Gradient Boosting is proposed in this paper. Firstly, the historical data is taken as a part of the input vectors for the model. Moreover, considering the monthly change of wind speed characteristics, the dataset of wind power is divided into four parts by month so that the models are constructed in different complexity by month. Finally, compared with back propagation neural networks (BPNN) and linear regression (LR) models, the experimental results show that the improved XGBoost model can promote the forecast accuracy effectively.

INDEX TERMS Short-term wind speed forecasting, XGBoost, time series, historical characteristics, power grid.

I. INTRODUCTION

A. MOTIVATION

As the serious concerns of environmental protection and energy crisis growing, the development of renewable energy, such as solar energy, hydro energy, wind energy, tidal energy, geothermal energy, etc., is becoming an important part of the energy portfolio worldwide [1]–[5]. However, wind power always exhibits high variability and volatility due to the fact that wind speed is influenced by synoptic weather patterns. The volatility and indirectness of wind speed and low energy density will reduce the reliability of power system operation [6], [7]. Therefore, it poses a challenge to the balance of supply and demand for electrical power and energy systems, especially when the wind power systems have low storage capacity. Wind speed forecast, which predicts the variability of wind power, plays an important role in the operational, management and real-time control of electrical energy systems. For example, wind speed forecast can help arrange the charging schedule of the energy storage systems and optimize energy transmission, thereby, avoiding the loss of available energy. Also, wind speed forecast mitigates the disturbance in the electrical energy system, reducing the reserve capacity and the generation cost. In recent years, wind speed forecast has attracted extensive attention in academia and industry.

So far, the literature related to wind speed forecast is divided into four typical classifications, containing physical model, statistical approach, machine learning and deep learning. It can readily conclude that neural networks (NNs), such as deep learning, is widely used due to their strong capabilities for feature extraction and data-mining.

Nevertheless, all of the methods above have obvious defects, and we will discuss them in the following.

B. LITERATURE REVIEW

1) PHYSICAL METHOD

This type of method is constantly employed for forecasting weather data, such as wind speed, wind direction, temperature, and so on [8]. Basically, numerical weather prediction (NWP) models [9] are used for predicting future data through

The associate editor coordinating the review of this manuscript and approving it for publication was Junjie Hu.

applying current physical data and topographical information [10]. Dong *et al.* proposed a comprehensive review of NWP models and put forward an effective hybrid model to predict day-ahead wind power [11]. However, these models have high requirements of computing resources for model simulation and need a great mass of information for predicting [12]. Besides, the physical model cannot be used for short-term forecasting because of their extensive training time [13].

2) STATISTICAL METHOD

The statistical method is appropriate for the time-series forecasting of wind speed. The statistical model mainly combines the characteristics of time-series data to forecast wind speed and optimize the model parameters by using the forecast errors [5]. Common statistical approaches include Kalman filter and Markov chain [14], auto regressive model, autoregressive moving average [15]–[17], autoregressive conditional heteroscedasticity model [18], exponential smoothing [19] and autoregressive integrated moving average [20], [21]. Nevertheless, among these models, assumptions need to be made with regard to the data distribution before modeling, and the obtained results are not satisfactory for nonlinear time series forecasting [22]. The spatial correlation model considers the spatial relationship of wind speed data at different location coordinates and enhances the forecasting stability.

3) MACHINE LEARNING METHOD

Fuzzy logic models [23], support vector machines [24], and Gaussian mixture model are applied to analyze and predict the wind power [25]. These models can optimally fit the nonlinear time series data and are suitable for forecasting without presupposition [26]. Nevertheless, these models sometimes fall into local optimal, resulting in overfitting. A preliminary investigation of initial wind speed dataset shows that it is unstable and noisy. Thus, preprocessing methods and model optimization algorithms are adopted to abate the random disturbance of initial data [27], including wavelet packet decomposition [28], empirical mode decomposition [29], singular spectrum analysis [30], and others [31], [32]. The particle swarm optimization [33], genetic algorithms [34], cuckoo search [35] are used for optimization.

4) DEEP LEARNING METHOD

To further improve the statistical method and traditional machine learning method, deep learning algorithm are widely used [25], [36]–[40]. Deep neural network can simulate the hierarchical structure of human brain and make it have deeper emotional semantic expression ability, so it has been widely used in forecasting task. Compared with machine learning, deep learning has got rid of the constraints of complex feature engineering, but supervised deep learning still needs a large number of annotated data set training models, while unsupervised deep learning requires strict correlation between data, so this method needs to be further developed. At present, semi-supervised learning, which attracts more attention, only needs to use a small amount of labeled data to train the model, which is superior to the above methods in terms of forecasting. Nevertheless, despite the existing NNs-based forecasting model has their advantages, parts of engineers are unwilling to take use of it in practical application. This is because these models are generally built and deployed on graphics processing unit (GPU) devices, which results in a great deal of hardware cost and electric power cost. It is unaffordable in some application scenarios.

C. INNOVATIONS

Based on the literature review, the shortcomings of the above methods are summarized. On the one hand, the traditional methods, including physical method, statistical method and machine learning method, hardly learn the time-series features and get good performances though the GPU sources are not consumed. On the other hand, the NNs models always require the more expensive GPU devices instead of the cheaper CPU devices. Therefore, it is necessary to explore a method which could get good performances devices and get rid of the dependence on the GPU in the meantime.

Therefore, the Extreme Gradient Boosting (XGBoost) is adopted to forecast the wind speed, and the two innovation points are proposed to promote the model performance of XGBoost, as follows:

1) MERGING HISTORICAL DATA TO INPUT VECTOR

Firstly, the historical data is adopted as a part of the input vectors, that means the wind speed data at several points before the forecasting point and other environment variables are selected as the input dimensions of the model.

In this way, XGBoost is used to learn the historical characteristics and get rid of the recurrent structure of NNs, in other words, it does not need to depend on NNs but possess the potential alike NNs.

2) MODELING BY MONTH

To control the complexity of the model flexibly, the models are constructed month by month, and every model does not need to be considered to cover all the time characteristics all over the years. It means that a series of simpler models are get according to the months.

By the means, the cost of computing resource is saved when training and deploying models.

The paper is organized as follows. In Section II, the XGBoost algorithm for wind speed forecasting is introduced. In Section III, a wind speed forecasting model based on the XGBoost is established. Case study is given in Section IV. We conclude this work in Section V.

II. XGBOOST FOR WIND SPEED FORECASTING

In this paper, XGBoost is adopted to construct the models and forecast the wind speed, and the back propagation neural network (BPNN) [41], [42] model and the linear regression (LR) [43]–[47] model, which represent the NNs and traditional method, respectively, as the comparison models are used. We will introduce XGBoost algorithm emphatically in this section.

XGBoost, as multifactor forecasting model, is proposed by Chen and Guestrin [48], XGBoost is an efficient system implementation of Gradient Boosting, which is an improvement from Gradient Boosting Decision Tree (GBDT). Due to the fact that the XGBoost is competent for parallel computation, approximate tree matching, handling sparse data effectively as well as improvement for central processing unit (CPU) and memory, it has a good performance in machine learning. Its objective function can be calculated as follows:

$$Obj(\varphi) = \sum\nolimits_{i=1}^{n} l(\hat{y}_i, y_i) + \sum\nolimits_{k=1}^{g} \Omega(f_k) \quad (1)$$

$$(f_k) = \gamma T + \frac{1}{2}\lambda\omega^2 \quad (2)$$

where $\hat{y}_i$ and y_i are the predicted value and predicted value, respectively; l denotes the loss function; $\Omega(f_k)$ represents the regular term; f_k expresses a decision tree. Besides, XGBoost is applied a second order Tailor expansion. The tth loss function is regarded as follows:

$$L^{(t)} = \sum\nolimits_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i) + \Omega(f_t) \quad (3)$$

Then the second order Tailor expansion is applied for $L^{(t)}$:

$$L^{(t)} \cong \sum\nolimits_{i=1}^{n} [l(y_i, \hat{y}_i^{(t-1)} + g_i f_t(x_i) + \frac{1}{2} g_i f_t^2(x_i)] + \Omega(f_t) \quad (4)$$

$$g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \quad (5)$$

$$h_i = \partial^2_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \quad (6)$$

The structure of the XGBoost is revealed as Fig. 1.

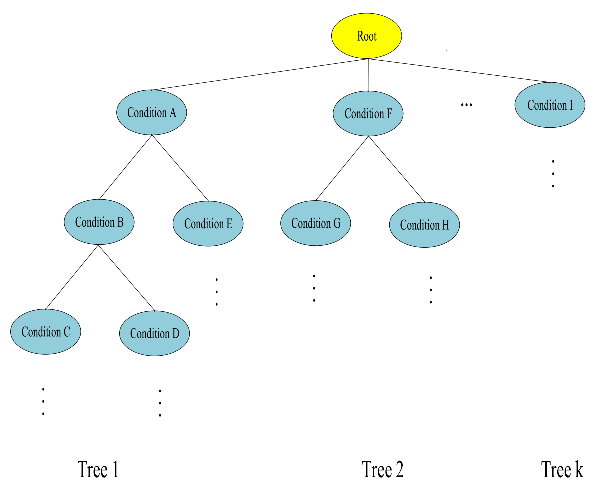


FIGURE 1. **The structure of XGBoost.**

III. WIND SPEED FORECASTING MODELING BASED ON XGBOOST

The abstract process of training for models is shown as Fig. 2. In addition, the specific training steps in this paper can be summarized as follows:

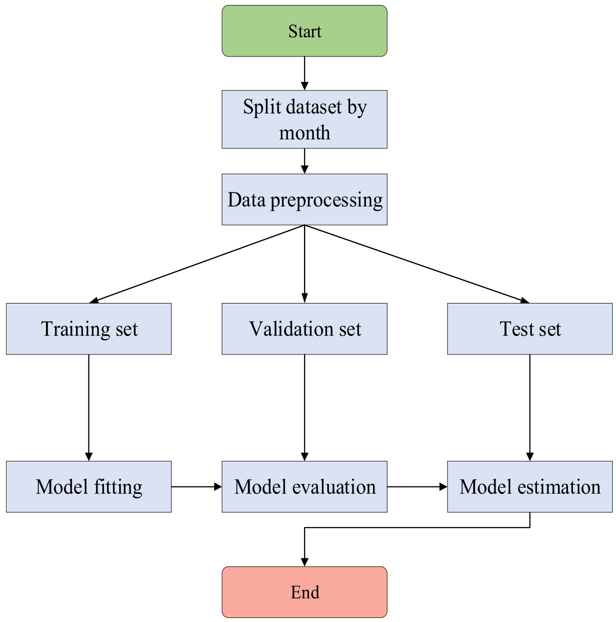


FIGURE 2. **The process of training of the models.**

(1) On account of the uncertain environment noise factors in the wind power station or the instability of the data collection system for wind speed, there exit a part of invalid data in the original data (such as empty values, values beyond common sense, etc.) that the data is preprocessed to eliminate invalid data.

(2) The dataset is split by month and then divided into training set, validation set and test set in the same rate because the wind speed in every month have potential different characteristics.

(3) The training set are sent to the models for training to learn the reasonable parameters about wind speed.

(4) The models are validated on the validation set so as to adjust the parameters.

(5) The models are verified in the test set to evaluate performance of the models, which ensures that the models are competent to forecast the wind speed.

IV. PERFORMANCE METRICS

In order to assess the overall forecasting performance of XGBoost comprehensively, four widely-used metrics in statistics are employed, namely mean absolute error (MAE), root mean square error (RMSE), coefficient of variation of RMSE (CV-RMSE) and coefficient of determination (R^2). MAE measures how close the forecasting results are to their actual values. RMSE is basically the standard deviation of the forecasting errors, which measures the concentration of data around the best fit line. CV-RMSE describes the fitness of the forecasting model. The lower the CV-RMSE, the smaller the residuals relative to the forecasted results, indicating that the model fits well. R^2 is the square of the correlation between the predicted value and their real value. The four indices are calculated as follows (7), (8), (9) and (10), as shown at the bottom of the next page.

V. CASE STUDY

In this section, 32686 wind speed samples were collected as the source dataset. Parts of the samples are shown as table 2. This dataset consists of various environmental and meteorological features.

TABLE 1. Parts of the samples.

Rad	Tem	Pre	Hum	WD	TSR	TSS	WS
1.21	48	30.46	58	176.78	6:13:00	18:13:00	3.37
1.23	48	30.46	57	158.75	6:13:00	18:13:00	3.37
1.21	48	30.46	60	137.71	6:13:00	18:13:00	3.37
1.17	48	30.46	62	104.95	6:13:00	18:13:00	5.62
1.21	48	30.46	64	120.2	6:13:00	18:13:00	5.62
1.2	49	30.46	72	112.45	6:13:00	18:13:00	6.75
1.24	49	30.46	71	122.97	6:13:00	18:13:00	5.62
1.23	49	30.46	80	101.18	6:13:00	18:13:00	4.5
1.21	49	30.46	85	141.87	6:13:00	18:13:00	4.5
1.23	49	30.47	93	120.55	6:13:00	18:13:00	2.25

As table 1 shows, Rad, Tem, Pre, Hum, WD, TSR, TSS, WS represent radiation, temperature, pressure, humidity, wind direction, time of sunrise, time of sunset and wind speed, respectively.

A. EXPERIMENT SETTING

1) DATA DESCRIPTION

The wind speed dataset was collected from a wind power plant in the Northern Hemisphere. The dataset covers the whole September, October, November, December in the year of 2016 with a data resolution of 5 minute, and contains 7 features, including radiation, temperature, pressure, humidity, wind direction, time of sunrise, time of sunset. Negative values of wind speed in this dataset indicates that the corresponding feature cannot be observed due to some technical reasons. The wind speed observation of the dataset is shown in Fig. 3. From Fig. 3, the wind speed values vary greatly from month to month.

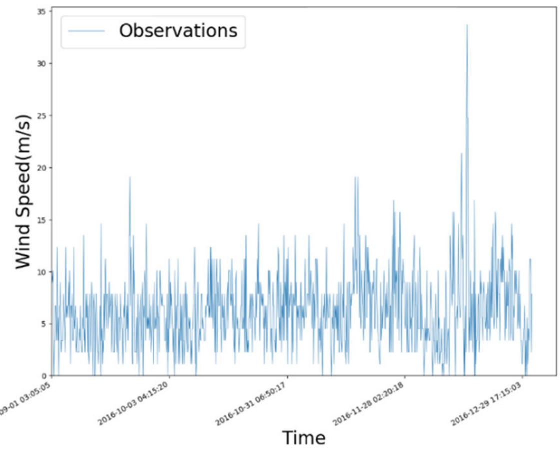


FIGURE 3. Wind speed observation.

2) SIMULATION SETTINGS

According to Section 2, an independent XGBoost forecast model is built to forecast wind speed. The forecasting model has 12 input parameters, including radiation, temperature, pressure, humidity, wind direction, and wind speed before 5/10/15/20/25/30/35 minutes. These input parameters are determined through trial and error. The output of the forecast model is the wind speed at the next time step. The learning rate of XGBoost is set to 0.005. The XGBoost-based forecast framework and two benchmark algorithms are all programmed in Python software. It is worth noting that we combinate the historical data, wind speed of 5/10/15/20/25/30/35 minutes before into the input vector, as shown as Fig. 4.

B. FORECASTING RESULT ANALYSIS

The entire wind speed dataset is divided into four independent parts, namely September, October, November, and December. Moreover, each part is split into a training set and a validation set.

The overall forecasting performance of XGBoost is comprehensively evaluated on test samples in four months. The evaluation metrics MAE, RMSE, CV-RMSE and R^2 are displayed in Table 2. The forecasting results of BPNN and

$$MAE = \frac{1}{N_S} \sum\nolimits_{i=1}^{N_S} |o_i - y_{fi}| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N_S} \sum\nolimits_{i=1}^{N_S} (o_i - y_{fi})^2} \tag{8}$$

$$CV - RMSE = \frac{\sqrt{\sum_{t=1}^{N_S} (o_i - y_{fi})^2/(N_S - 1)}}{\sum_{t=1}^{N_S} o_i/N_S} \tag{9}$$

$$R^2 = \frac{\left(N_S \sum_{t=1}^{N_S} o_i \times y_{fi} - \sum_{t=1}^{N_S} o_i \times \sum_{t=1}^{N_S} y_{fi}\right)^2}{\left[N_S \sum_{t=1}^{N_S} (y_{fi})^2 - \left(\sum_{t=1}^{N_S} y_{fi}\right)^2\right] \cdot \left[N_S \sum_{t=1}^{N_S} (o_i)^2 - \left(\sum_{t=1}^{N_S} o_i\right)^2\right]} \tag{10}$$

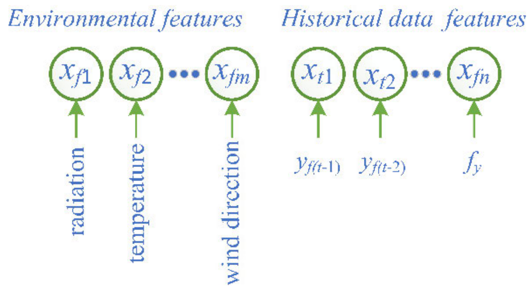


FIGURE 4. Input vector.

TABLE 2. Performance metrics of the models.

Months	Performance metrics	XGBoost	BPNN	LR
September	MAE	**1.5292**	1.8508	1.5487
	RMSE	**1.9796**	2.3512	1.9931
	CV-RMSE	**0.3796**	0.4509	0.3822
	R^2	**0.4183**	0.2295	0.4075
October	MAE	**1.8492**	2.0201	1.8502
	RMSE	2.3915	2.6280	**2.3910**
	CV-RMSE	0.4298	0.4723	**0.4297**
	R^2	0.2011	0.0866	**0.2049**
November	MAE	**2.0872**	2.5273	2.1143
	RMSE	**2.6753**	3.1884	2.7514
	CV-RMSE	**0.3882**	0.4626	0.3992
	R^2	**0.2192**	0.0188	0.2180
December	MAE	**1.9330**	2.4440	1.9337
	RMSE	2.4855	3.1757	**2.4826**
	CV-RMSE	0.3592	0.4590	**0.3588**
	R^2	0.6045	0.4276	**0.6102**
Mean	MAE	**1.8497**	2.2105	1.8618
	RMSE	**2.3830**	2.8359	2.4045
	CV-RMSE	**0.3892**	0.4612	0.3925
	R^2	**0.3608**	0.1906	0.3601

LR are also given for performance comparison. In order to present the forecasting performance of XGBoost, BPNN and LR visually, their real-time forecasting results over 200 minutes in four months are graphically given in Figs. 5-8. As shown in Table 1, the MAE index of XGBoost varies from 1.5292 m.s^{-1} to 2.0872 m.s^{-1} with an average of 1.8497 m.s^{-1}. The average values of the MAE index of BPNN and LR are 2.2105 and 1.8618, respectively.

Apparently, XGBoost has the best monthly MAE index compared to the other two benchmark algorithms, indicating that the forecasting results of XGBoost are closest to the real wind speed observations. In terms of RMSE and CV-RMSE, XGBoost performs the best in September and December. XGBoost performs slightly worse than LR in October and December, but much better than BPNN. Consequently, XGBoost has the best RMSE and CV-RMSE over the whole dataset among the benchmark algorithms, indicating that the forecasting curve of XGBoost is centered around the true wind speed curve. In addition, the average value of R^2 for XGBoost over the dataset performs the best. In summary, XGBoost has the best MAE, RMSE, CV-RMSE and R^2 throughout the whole dataset.

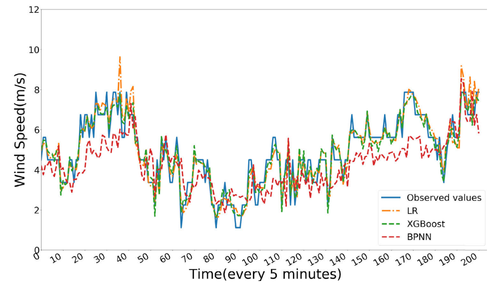


FIGURE 5. Forecasting results of the three algorithms in September.

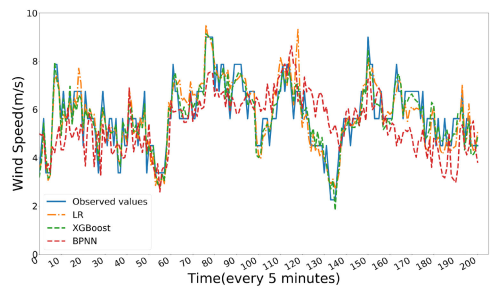


FIGURE 6. Forecasting results of the three algorithms in October.

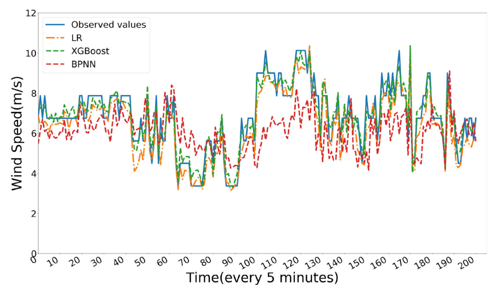


FIGURE 7. Forecasting results of the three algorithms in November.

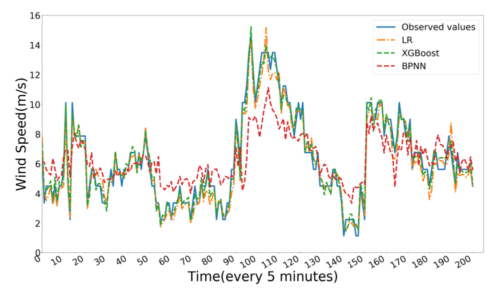


FIGURE 8. Forecasting results of the three algorithms in December.

VI. CONCLUSION

In this paper, the model of XGBoost are applied to forecast the wind speed, and compared with the models of BPNN and LR. Among these models, XGBoost has the best MAE, RMSE, CV-RMSE and R^2 throughout the whole dataset, which represents the model fitting ability on the overall dataset.

The result shows the advantages of the improved XGBoost model on the forecasting for wind speed. It can forecast the wind speed in the practical application scenario accurately to provide technical support for the device safety and the protection of wind speed system.

Additionally, the improved XGBoost combinate the methods that adopting historical data and simplifying the models by month, which successfully saves computing resources and promote the performance of the model.

However, it is a pity that more data cannot be obtained in this paper, which may influence on model accuracy. Therefore, a larger dataset will be gotten for a more excellent performance. In addition, the sudden noise samples also cause much forecasting error and we should solve this problem in the future research.

REFERENCES

[1] H. Wang, Y. Liu, B. Zhou, C. Li, G. Cao, N. Voropai, and E. Barakhtenko, "Taxonomy research of artificial intelligence for deterministic solar power forecasting," *Energy Convers. Manage.*, vol. 214, Jun. 2020, Art. no. 112909.

[2] Z. Wang, J. Zhang, Y. Zhang, C. Huang, and L. Wang, "Short-term wind speed forecasting based on information of neighboring wind farms," *IEEE Access*, vol. 8, pp. 16760–16770, 2020.

[3] D. Xu, Q. Wu, B. Zhou, C. Li, L. Bai, and S. Huang, "Distributed multi-energy operation of coupled electricity, heating, and natural gas networks," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2457–2469, Oct. 2020, doi: 10.1109/TSTE.2019.2961432.

[4] C. Huang, L. Wang, and L. L. Lai, "Data-driven short-term solar irradiance forecasting based on information of neighboring sites," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9918–9927, Dec. 2019.

[5] V. V. S. N. Murty and A. Kumar, "Multi-objective energy management in microgrids with hybrid energy sources and battery energy storage systems," *Protection Control Mod. Power Syst.*, vol. 5, no. 1, pp. 1–20, Dec. 2020.

[6] Y. Nan, H. Yu, and Y. Di, "Study on probability distribution of wind power fluctuation based on nacemd and improved nonparametric kernel density estimation," *Power Syst. Technol.*, vol. 43, no. 3, pp. 910–917, 2019.

[7] X. Yindi, L. Kaipei, Q. Liang, O. Tinghui, and H. E. Jiayi, "Short-term wind power prediction method based on dynamic wind power weather division of time sequence data," *Power Syst. Technol.*, vol. 43, no. 9, pp. 3353–3359, 2019.

[8] J. Zhao, Z.-H. Guo, Z.-Y. Su, Z.-Y. Zhao, X. Xiao, and F. Liu, "An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed," *Appl. Energy*, vol. 162, pp. 808–826, Jan. 2016.

[9] J. Yang, M. Astitha, L. D. Monache, and S. Alessandrini, "An analog technique to improve storm wind speed prediction using a dual NWP model approach," *Monthly Weather Rev.*, vol. 146, no. 12, pp. 4057–4077, Dec. 2018.

[10] B. Wu, M. Song, K. Chen, Z. He, and X. Zhang, "Wind power prediction system for wind farm based on auto regressive statistical model and physical model," *J. Renew. Sustain. Energy*, vol. 6, no. 1, Jan. 2014, Art. no. 013101.

[11] A. U. Haque, M. H. Nehrir, and P. Mandal, "A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1663–1672, Jul. 2014.

[12] M. Lydia, S. S. Kumar, A. I. Selvakumar, and G. E. P. Kumar, "Linear and non-linear autoregressive models for short-term wind speed forecasting," *Energy Convers. Manage.*, vol. 112, pp. 115–124, Mar. 2016.

[13] E. Erdem and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," *Appl. Energy*, vol. 88, no. 4, pp. 1405–1414, Apr. 2011.

[14] M. Louzazni, H. Mosalam, A. Khouya, and K. Amechnoue, "A non-linear auto-regressive exogenous method to forecast the photovoltaic power output," *Sustain. Energy Technol. Assessments*, vol. 38, Apr. 2020, Art. no. 100670, doi: 10.1016/j.seta.2020.100670.

[15] F. O. Hocaoglu and F. Serttas, "A novel hybrid (Mycielski-Markov) model for hourly solar radiation forecasting," *Renew. Energy*, vol. 108, pp. 635–643, Aug. 2017.

[16] J. L. Torres, A. García, M. De Blas, and A. De Francisco, "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)," *Energy*, vol. 79, no. 2005, pp. 65–77, 2004, doi: 10.1016/J.SOLENER.2004.09.013.

[17] H. Liu, H.-Q. Tian, and Y.-F. Li, "An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system," *J. Wind Eng. Ind. Aerodyn.*, vol. 141, pp. 27–38, Jun. 2015, doi: 10.1016/J.JWEIA.2015.02.004.

[18] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica J. Econ. Soc.*, vol. 50, pp. 987–1007, 1982.

[19] M. U. Yousuf, I. Al-Bahadly, and E. Avci, "Current perspective on the accuracy of deterministic wind speed and power forecasting," *IEEE Access*, vol. 7, pp. 159547–159564, 2019.

[20] The Global Wind Energy Council. *Annual Global Wind Report*. Accessed: Sep. 3, 2019. [Online]. Available: https://gwec.net/wp-content/uploads/2019/04/GWEC-GlobalWind-Report-2018.pdf

[21] B. Kanna and S. N. Singh, "Long term wind power forecast using adaptive wavelet neural network," in *Proc. IEEE Uttar Pradesh Sect. Int. Conf. Electr., Comput. Electron. Eng. (UPCON)*, Varanasi, India, Dec. 2016, pp. 671–676.

[22] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, "A review of deep learning for renewable energy forecasting," *Energy Convers. Manage.*, vol. 198, Oct. 2019, Art. no. 111799.

[23] S. Baran, "Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components," *Comput. Statist. Data Anal.*, vol. 75, pp. 227–238, Jul. 2014.

[24] L. Li, Y.-Q. Liu, Y.-P. Yang, S. Han, and Y.-M. Wang, "A physical approach of the short-term wind power prediction based on CFD pre-calculated flow fields," *J. Hydrodyn.*, vol. 25, no. 1, pp. 56–61, Feb. 2013.

[25] H. B. Azad, S. Mekhilef, and V. G. Ganapathy, "Long-term wind speed forecasting and general pattern recognition using neural networks," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 54–553, Apr. 2014.

[26] X. Kong, X. Liu, R. Shi, and K. Y. Lee, "Wind speed prediction using reduced support vector machines with feature selection," *Neurocomputing*, vol. 169, pp. 449–456, Dec. 2015.

[27] X. Luo, J. Sun, L. Wang, W. Wang, W. Zhao, J. Wu, J.-H. Wang, and Z. Zhang, "Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4963–4971, Nov. 2018.

[28] C. Ren, N. An, J. Wang, L. Li, B. Hu, and D. Shang, "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting," *Knowl.-Based Syst.*, vol. 56, pp. 226–239, Jan. 2014.

[29] Y. Liu, L. Guan, C. Hou, H. Han, Z. Liu, Y. Sun, and M. Zheng, "Wind power short-term prediction based on LSTM and discrete wavelet transform," *Appl. Sci.-Basel*, vol. 9, no. 6, p. 17, 2019.

[30] K.-P. Lin, P.-F. Pai, and Y.-J. Ting, "Deep belief networks with genetic algorithms in forecasting wind speed," *IEEE Access*, vol. 7, pp. 99244–99253, 2019.

[31] Q. Cao, B. T. Ewing, and M. A. Thompson, "Forecasting wind speed with recurrent neural networks," *Eur. J. Oper. Res.*, vol. 221, no. 1, pp. 148–154, Aug. 2012.

[32] X. Yuan, C. Chen, Y. Yuan, Y. Huang, and Q. Tan, "Short-term wind power prediction based on LSSVM–GSA model," *Energy Convers. Manage.*, vol. 101, pp. 393–401, Sep. 2015.

[33] X. Zhao, C. Wang, J. Su, and J. Wang, "Research and application based on the swarm intelligence algorithm and artificial intelligence for wind farm decision system," *Renew. Energy*, vol. 134, pp. 681–697, Apr. 2019.

[34] K. M. Nor, M. Shaaban, and H. A. Rahman, "Feasibility assessment of wind energy resources in malaysia based on NWP models," *Renew. Energy*, vol. 62, pp. 147–154, Feb. 2014.

[35] G. Li, J. Shi, and J. Zhou, "Bayesian adaptive combination of short term wind speed forecasts from neural network models," *Renew. Energy*, vol. 36, no. 1, pp. 352–359, 2011.

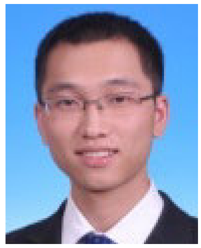
[36] P. Jiang, R. Li, and H. Li, "Multi-objective algorithm for the design of prediction intervals for wind power forecasting model," *Appl. Math. Model.*, vol. 67, pp. 101–122, Mar. 2019, doi: 10.1016/j.apm.2018.10.019.

[37] M. G. De Giorgi, A. Ficarella, and M. G. Russo, "Short-term wind forecasting using artificial neural networks (ANNs)," *WIT Trans. Ecol. Environ.*, vol. 121, no. 1, pp. 197–208, 2009, doi: 10.2495/ESUS090181.

[38] S. N. Singh, and A. Mohapatra, "Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting," *Renew. Energy*, vol. 136, pp. 758–768, Jun. 2019.
[39] R. G. Kavasseri and K. Seetharaman, "Day-ahead wind speed forecasting using f-ARIMA models," *Renew. Energy*, vol. 34, no. 5, pp. 1388–1393, May 2009.
[40] Z. Song, Y. Jiang, and Z. Zhang, "Short-term wind speed forecasting with Markov-switching model," *Appl. Energy*, vol. 130, pp. 103–112, Oct. 2014.
[41] Y. Luo, B. Lin, C.-B. Wen, and M. Luo, "Conducting a correlation model between TCM constitution and physical examination index based on BPNN algorithm," *Digit. Chin. Med.*, vol. 1, no. 1, pp. 84–89, Mar. 2018.
[42] D. K. Ghose and S. Samantaray, "Modelling sediment concentration using back propagation neural network and regression coupled with genetic algorithm," *Procedia Comput. Sci.*, vol. 125, pp. 85–92, 2018.
[43] H. Pombeiro, R. Santos, P. Carreira, C. Silva, and J. M. C. Sousa, "Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks," *Energy Buildings*, vol. 146, pp. 141–151, Jul. 2017.
[44] V. Bianco, O. Manca, and S. Nardini, "Electricity consumption forecasting in italy using linear regression models," *Energy*, vol. 34, no. 9, pp. 1413–1421, Sep. 2009.
[45] T. A. Reddy, *Applied Data Analysis and Modeling for Energy Engineers and Scientists*. New York, NY, USA: Springer, 2011.
[46] D. H. C. Toe and T. Kubota, "Development of an adaptive thermal comfort equation for naturally ventilated buildings in hot–humid climates using ASHRAE RP-884 database," *Frontiers Architectural Res.*, vol. 2, no. 3, pp. 278–291, 2013.
[47] G. Woodal, "Methodology for energy-efficiency measurements applicable to ICT in buildings (eeMeasure)," *Smart*, vol. 2011, no. 72, pp. 1–23, 2012.
[48] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785–794.
[49] G. Fu, "Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system," *Energy*, vol. 148, pp. 269–282, Apr. 2018.

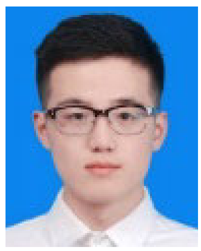
REN CAI received the B.S. degree in industrial design from South China Agricultural University, Guangzhou, China, in 2018. He is currently pursuing the master's degree in control science and engineering with Shenzhen University. His research interests include natural language processing, interpretable deep learning, data mining, and high-performance computing.

SEN XIE received the B.Eng. degree in electrical engineering and automation from the Hunan University of Technology, China, in 2011, the M.Eng. degree in control theory and control engineering from Liaoning Technical University, China, in 2014, and the Ph.D. degree in control science and engineering from Central South University, China, in 2018. She is currently a Postdoctoral Researcher with Shenzhen University, China. Her research interests include modeling and optimal control of complex industrial process, data processing, and intelligent analysis.



BOZHONG WANG received the B.S. and M.S. degrees in electrical engineering from North China Electric Power University, Beijing, China, in 2009 and 2012, respectively. He is currently a Power Grid Engineer with State Grid Hunan Electric Power Corporation Maintenance Company, Changsha, China. His research interests include power system operation, reliability, maintenance, and protection.

RUIJIANG YANG received the B.S. degree in agricultural electrification from Nanjing Agricultural University, Nanjing, China, in 2019. She is currently pursuing the M.S. degree with the College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China. Her research interests include power system planning and operation.



DAOSEN XU received the B.S. degree in electrical engineering and automation from Zhengzhou University, Zhengzhou, China, in 2019. He is currently pursuing the M.S. degree with the College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China. His research interest includes resilience of prosumer clusters.

YANG HE received the B.S. degree in electrical engineering and automation from China Three Gorges University, Yichang, China, in 2019. She is currently pursuing the M.S. degree with the College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China. Her research interests include power system planning and operation.