Paper ID APEN-MIT-2022_4885
Applied Energy Symposium: MIT A+B
July 5-8, 2022 • Cambridge, USA

Applying Optimized Preprocessing Technique and Extreme Learning Machine for Hybrid Wind Speed Forecasting System

Leyi Yu
State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources
North China Electric Power University
Beijing, China
yuleyi@ncepu.edu.cn

Zhiya Pan
State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources
North China Electric Power University
Beijing, China
panzhiya@ncepu.edu.cn

Xue Kong
State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources
North China Electric Power University
Beijing, China
kongxue@ncepu.edu.cn

Ying Zhang
State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources
North China Electric Power University
Beijing, China
zhangying@ncepu.edu.cn

Tao Wang
State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources
North China Electric Power University
Beijing, China
wt@ncepu.edu.cn

Yagang Zhang*
State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources
North China Electric Power University
Beijing, China
yagangzhang@ncepu.edu.cn

Abstract—In order to cope with the environmental problems of climate deterioration, reduce carbon emissions and develop environmentally friendly energy sources without delay. The global energy system is undergoing tremendous changes, accelerating the transformation of the energy structure, and promoting the development of the energy structure to a low-carbon or even carbon-free direction. Renewable energy has received attention due to its clean and green characteristics. Wind power generation, as an important way to develop renewable energy, faces grid security challenges due to the volatility and uncontrollability of wind speed. In this paper, a systematic prediction method is proposed for wind speed: the wind speed sequence is decomposed by the variational modal decomposition method according to the principle of envelope entropy, and the extreme learning machine network optimized by the bat algorithm is used to predict the data after decomposing. Besides, error sequence correction method based on kernel density estimation is proposed to predict residual. The availability of the model proposed in this paper is proved by experiments, and a good prediction effect is obtained. In order to heighten the utilization of wind energy, the wind power dispatching of relevant departments provides a reference method.

Keywords—wind speed prediction, variational mode decomposition, extreme learning machine, error correction

I. INTRODUCTION

In recent years, with the continuous increase of carbon dioxide emissions, it is urgent to solve the problem of global warming, which is not only related to the survival of global animals and plants, but also directly related to the future development of all mankind. Seeking a green and low-carbon development route has become an important issue for us to discuss now. The traditional power generation system is the main way to generate carbon dioxide. Reducing thermal power generation and increasing the proportion of clean energy power generation are the keys to achieving carbon emission reduction goals [1]. Wind energy is one of the most potential renewable energy sources, its intermittent characteristics pose a certain threat to grid connection and stable operation of power grid. Failure to address this issue will affect the scale and speed of wind farm deployment. Therefore, accurate prediction of wind speed can not only improve the efficiency of wind power generation, solve the problem of safe operation of the power grid, but also improve the possibility of wind farm development [2].

Research on wind speed prediction can generally be divided into these directions: physical model, statistical model, machine learning and hybrid model. Physical models mainly include numerical weather forecasting systems, which use external environmental factors such as air pressure, temperature, and altitude to infer changes in atmospheric conditions, thereby predicting changes in wind speed [3]. The statistical model mainly uses traditional statistical methods to fit and predict the wind speed. The commonly used methods mainly predict the stationary series, such as auto regressive (AR) [4], moving average (MA) [5] and auto-regressive moving average (ARMA) [6]. For prediction of non-stationary series, the auto-regressive integrated moving average (ARIMA) [7] model combined with the difference algorithm is often used for prediction. Commonly used machine learning includes support vector machine (SVM) [8] and various neural networks [9-11]. Neural network is widely used in wind speed prediction due to its good training effect. Common neural networks include: extreme learning machine (ELM) [9], back propagation neural network (BP) [10] and Elman neural network [11]. Nowadays, in order to better improve the prediction accuracy, a hybrid model combining multiple models is used to predict the wind speed. First, preprocess the original data, and then use the machine learning optimized by the optimization algorithm to train and predict the processed data, so that a better prediction effect can often be obtained [12-14].

This paper proposes a wind speed prediction scheme based on decomposition method, machine learning and error correction. The innovations are as follows:

- For the data preprocessing, for the problem that it is hard to determine the optimal parameters of the variational mode decomposition (VMD). The decomposition effect is measured by the envelope entropy value, and the optimal number of decomposition modes and penalty factor are selected.
- For machine learning models, in pursuit of better prediction results, the bat algorithm is used to optimize ELM models with stable performance and fast running speed.
- For the prediction of the error series, the kernel density estimation (KDE) method is used to sample the probability of the error series to obtain the error correction prediction value.

II. METHODOLOGY

A. Data pre-processing

This forecasting system uses improved VMD to processes the wind speed data. VMD is an adaptive data decomposition method by solving the variational problems. The method can effectively reduce the end effect and mode mixing. The non-stationarity of signal with high complexity is reduced by decomposing several intrinsic mode functions (IMF). The choice of VMD modal number K and penalty factor α is related to the effect of decomposition. If K is too small, the sequence information cannot be fully extracted and utilized. When K is too large, the sequence is excessively decomposed, resulting in a great impact from extra noise. Likewise, an inappropriate α also affects the accuracy of the decomposed IMF bandwidth. The selection of these two parameters depends on artificial settings. Formulating appropriate parameter selection principles is beneficial to improve the decomposition effect of VMD [6].

The minimum average envelope entropy is chosen as the index for selecting parameters of VMD. The envelope entropy reflects the complex characteristics of the signal through the fluctuation and sparsity of the signal [15-16]. A smaller entropy value means that the sequence contains more feature information and less noise. The formula of the envelope entropy E_p of the signal $y(t),(t=1,2,...,n)$ is as follow:

$$q(i)=a(i)/\sum_{i=1}^{n}a(i) \tag{1}$$

$$E_p=-\sum_{i=1}^{n}q(i)\log(q(i)) \tag{2}$$

In this formula, $a(t)$ is the Hilbert-transformed envelope signal of the signal $y(t)$, and $q(i)$ is the probability distribution sequence, n represents the number of signal points,.

The purpose of effective VMD decomposition results is to decompose nonlinear and complex data into a finite number of sequences with relatively simple structures. Therefore, it is considered that the signal decomposition effect is the best when the average value of the envelope entropy is the smallest.

B. Model buliding

Extreme learning machine (ELM) is used to build the main predictive model. ELM is different from traditional feedforward neural network [17]. The input weights and hidden biases are randomly generated, the optimal output weight is obtained according to the calculation, so as to achieve the training purpose.

Because the extreme learning machine does not need to repeat the iteration to reach the minimum training deviation, and the weight of the least square solution of its output is unique, it greatly reduces the calculation amount of the neural network. It effectively overcomes the problem that the traditional neural network is prone to local extreme values and overfitting. However, due to the randomness of weights and biases, for overly complex data, the prediction accuracy is limited by the hidden layer nodes [18]. To optimize model structure and improve the model prediction capability, it is essential to optimize the initial weights and biases.

C. Model optimization

Inspired by the ultrasonic echolocation of bats in nature, the bat optimization algorithm was proposed as a swarm intelligence search algorithm, which is widely used in various optimization scenarios [19].

The bat algorithm regards the individual position of each bat as the solution of the problem, and regards the optimization process as two stages: the global search stage and the local search stage [20]. In the first stage, individual bats enter the optimal solution candidate region by updating their positions. In the second stage, the pulse frequency and sound loudness of each individual bat are adjusted to achieve the purpose of fine search, and a new optimal solution is generated through comparison. The specific process of the bat algorithm is as follows:

Step1: Initialize the position $p_j(j=1,2,...,N)$ of the j-th bat, the corresponding initial velocity $v_j(j=1,2,...,N)$, and determine the range of the bat's frequency $[s_{min}, s_{max}]$, the individual initial sound intensity A_0, the initial pulse frequency r_0, and N is the population size. Set the fitness function $f(p)$.

Step2: Global search stage: p_j approaches the current global optimal position p^*, and updates the velocity v_j and the current position p_j at the current time t by updating the frequency of the sound wave s_j emitted by the individual. The update formula is as follows:

$$s_j=s_{min}+\beta(s_{max}-s_{min}) \tag{3}$$

$$v_j(t)=v_j(t-1)+[p_j(t-1)-p^*]\cdot s_j \tag{4}$$

$$p_j(t)=v_j(t)+p_j(t-1) \tag{5}$$

Where β is a random vector in the range [0,1].

Step3: Approaching the prey stage: The individual bat reaches the vicinity of the optimal solution, strengthens its own search properties, and accurately locates the prey through sound loudness A_j and pulse frequency r_j.

$$A_j(t) = \eta \cdot A_j(t-1) \tag{6}$$

$$r_j(t) = r_0 \cdot [1 - e^{-\gamma(t-1)}] \tag{7}$$

Where η, γ is fixed constant. The specific update rules are as follow:

(1) Sample a random number $rand1$ in the interval [0,1], if $rand1 > r_j$, δ is a random number generated in the interval [-1,1], $\bar{A}$ is the mean value of loudness, then update the position according to the formula (8).

$$p_j(t) = p^* + \delta \bar{A} \tag{8}$$

Otherwise, skip to step 4.

(2) Generate a random number $rand2$ in [0,1], if $rand2 < A_j$, calculate the current fitness value f. if $f(p_j) < f(p^*)$, then let p_j be the new optimal solution, and update A_j and r_j according to formula (6)-(7).

Otherwise, go to step 4.

Step4: Iterative loop: Repeat step 2 and step 3, until the maximum number of iterations is reached. Output the final optimal position p^*, which is the required optimal solution.

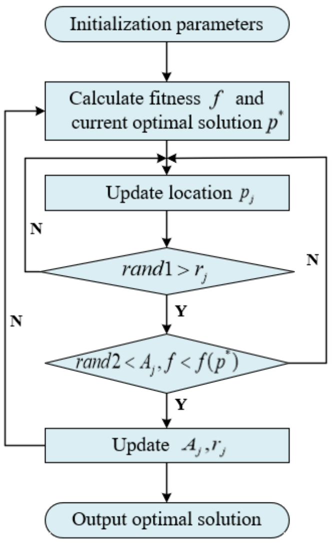


Fig. 1. Flow chart of bat algorithm

In this paper, the bat algorithm is used to determine ELM 's optimal weights and biases, and the fitness function is set as the mean square error of the training samples.

D. Error correction

In addition to the IMF, the VMD decomposition also includes a residual sequence that cannot be decomposed, and the residual sequence is a random sequence fluctuating around 0. In this paper, the kernel density estimation method is adopted to correct the residual series. Kernel density estimation (KDE) is nonparametric statistical technique that fits distribution sequence according to the characteristics of the sample data [2]. The implementation process is as follows: First, the kernel density is estimated for the residual sequence decomposed by VMD, and its probability distribution function is obtained. Second, the quantiles of the probability distribution are randomly sampled. Finally, the value corresponding to the quantile is obtained according to the residual distribution function, which is the corrected value of the obtained residual.

E. The proposed model

Based on the improved VMD, optimized ELM neural network and residual correction, this paper proposes a wind speed prediction method based on data preprocessing, machine learning model and error prediction. The forecasting process is as follows:

Step1: Based on the minimum average envelope entropy, the optimal VMD decomposition mode number and penalty factor are selected in turn.

Step2: Using the VMD with optimal parameters to decompose the wind speed original data, and obtain IMFs and a residual sequence.

Step3: Establish a BA-ELM model for each series obtained by decomposition, and output the prediction results of each component.

Step4: For the residual error after VMD decomposition, the KDE is used to fit the probability distribution of the residual, and sampling is performed according to distribution function, then residual predicted value is obtained.

Step5: The predicted value of IMFs and the residual predicted value are added to obtain the deterministic forecasting results of wind speed.

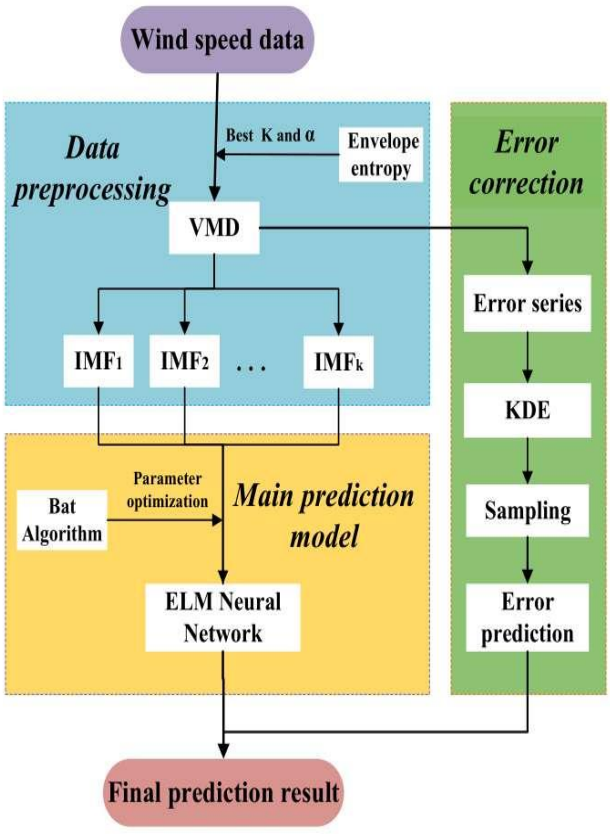


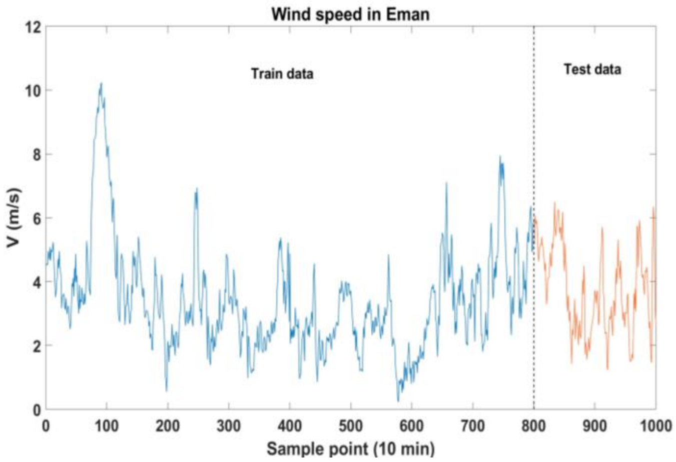
Fig. 2. Flow chart of proposed model

III. EXPERIMENT

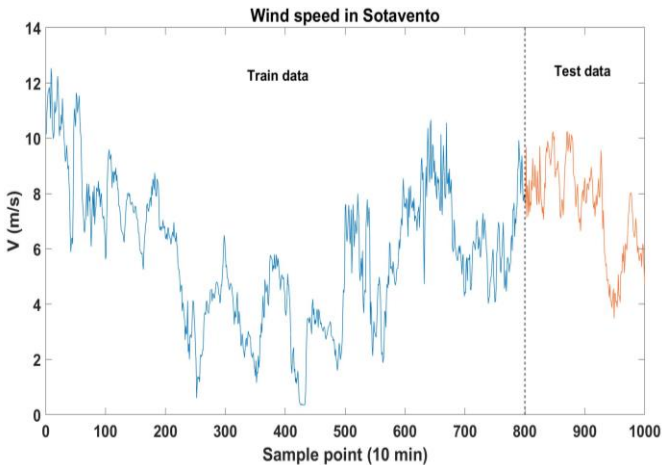
A. Data selection

In order to demonstrate that the prediction scheme proposed in this paper is capable to effectively predict the wind speed, we selected wind speed data from two different countries to conduct the experiment. Dataset 1 is the wind speed data of the Eman wind farm located in Hainan Province, China, in September 2020, and Dataset 2 is the wind speed data of the Sotavento wind farm in Spain, in October 2018. Data is sampled every ten minutes. This paper selects 1000 data points from the two data sets respectively, and the original sequence diagram of wind speed is presented in the Fig. 3. The two original series can be seen from Fig. 3(a) and Fig.3 (b) that the data of these two datasets show different fluctuation trends, which fully reflect the volatility of wind speed.

In this process, the first 80% of the data are divided into the training set, and the last 20% of the data are divided into the testing set.



(a) Wind speed data in Eman



(b) Wind speed data in Sotavento

Fig. 3. Original wind speed data

B. Experimental results

Establish models for the two datasets respectively. To prove the validity and superiority of the prediction scheme proposed in this paper, the BP, Elman, long short-term memory neural network (LSTM), and ELM are selected. At the same time, in order to verify the improvement effect of the optimization method, the VMD-ELM, IVMD-ELM, IVMD-BA-ELM combined model and the model proposed in this paper are used for comparative experiments. To facilitate the presentation, 100 data points are intercepted, and the prediction results are shown in Fig.4 - Fig.7.

Fig.4 and Fig.5 show that the combined model has a higher prediction effect than the single model, and among the four neural networks, the single ELM neural network model has relatively higher accuracy. In the simulation process, the ELM network is the fastest, and in the ultra-short-term forecast of wind speed, there is no doubt that ELM model has computational advantages.

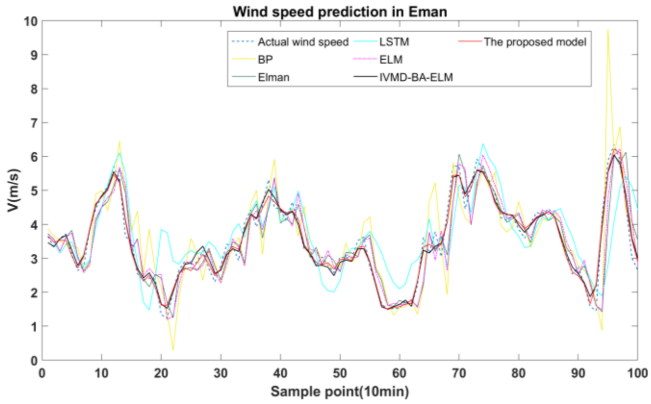


Fig. 4. Prediction results of dataset1 (a)

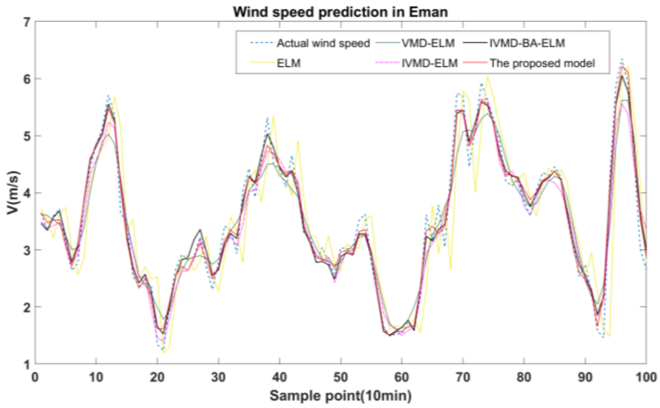


Fig. 5. Prediction results of dataset2 (a)

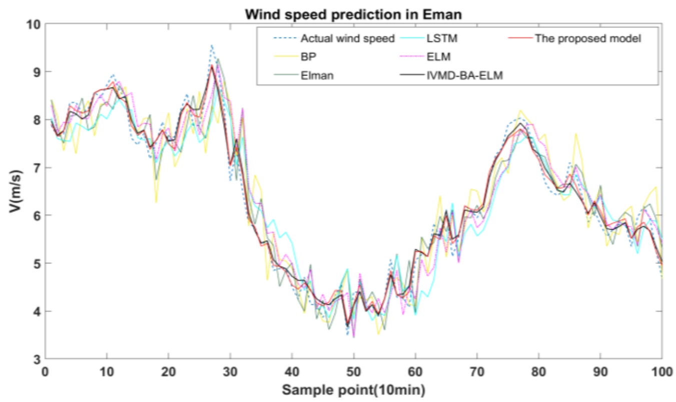


Fig. 6. Prediction results of dataset1 (b)

Fig.6 and Fig.7 show that the improved methods for VMD and ELM and the error correction methods used in this paper are effective. The consequences show that the proposed model not only simulates the fluctuation trend of wind speed well, but also solves the hysteresis problem of most of the comparative models, and also has a certain ability to predict wind speed climbing

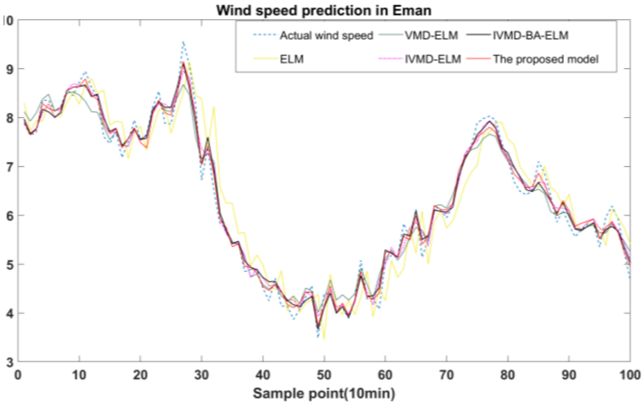


Fig. 7. Prediction results of dataset2 (b)

C. Model evaluation

For the propose of evaluating the forecasting effect of the model more comprehensively, the mean absolute error (MAE), mean squared error (MSE) and mean absolute percentage error (MAPE) are used to evaluate the prediction accuracy of the model. The calculation results are shown in Table Ⅰ and Table Ⅱ.

TABLE I. MODEL ERROR METRICS COMPARISON 1

Model	Eman		
	MSE	***MAE***	***MAPE***
BP	0.6944	0.6126	18.41%
LSTM	0.7072	0.6515	21.19%
Elman	0.4290	0.4837	15.03%
ELM	0.3534	0.4542	13.85%
VMD-ELM	0.1202	0.2626	8.14%
IVMD-ELM	0.0756	0.2101	6.46%
IVMD-BA-ELM	0.0440	0.1651	5.21%
Proposed model	0.0508	0.1562	4.97%

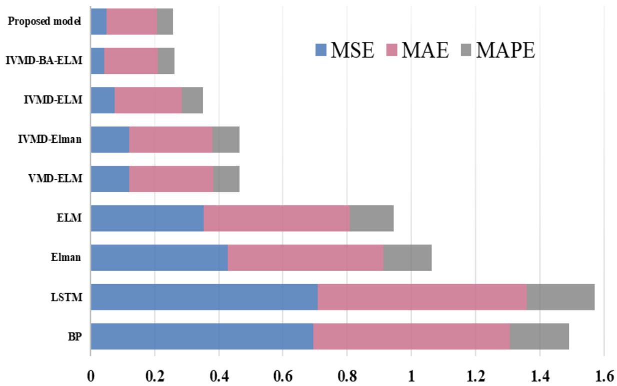


Fig. 8. Model evaluation of Eman

TABLE II. MODEL ERROR METRICS COMPARISON 2

Model	Sotavento		
	MSE	***MAE***	***MAPE***
BP	0.5259	0.5396	7.50%
LSTM	0.405	0.4821	6.89%
Elman	0.4341	0.5072	7.17%
ELM	0.3881	0.4685	6.63%
VMD-ELM	0.1175	0.2653	3.68%
IVMD-ELM	0.1015	0.2508	3.48%
IVMD-BA-ELM	0.0903	0.2268	3.17%
Proposed model	0.0581	0.1847	2.58%

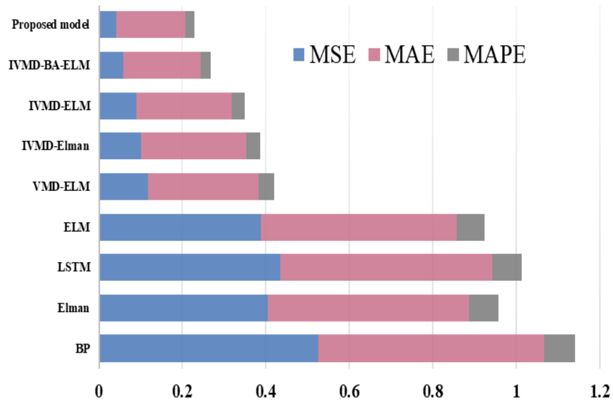


Fig. 9. Model evaluation of Sotavento

According to the above tables and pictures, it can be seen that the error indicators of the proposed model have decreased compared with other models, which further proves the superiority of the model.

TABLE III. IMPROVE PERFORMANCE

Compared Model	Reduction rate of MAE (%)	
	Eman	***Sotavento***
VMD-ELM	42.18%	43.37%%
IVMD-ELM	53.74%	46.47%
IVMD-BA-ELM	63.65%	51.60%
Proposed model	65.61%	60.58%

Table Ⅲ takes the MAE value as an example to show the decline degree of the combined model error indicators compared to the ELM single prediction model.

The following conclusions can be drawn:

- Compared with BP, LSTM and Elman, the ELM model has better prediction accuracy and prediction rate.
- The prediction consequent of the combined model is better than that of a single neural network, especially the data decomposed by VMD can greatly improve the prediction accuracy. From the original wind speed series characteristics of the two datasets, for more complex data, the model proposed in this paper improves to a greater extent.

- Compared with the unoptimized model, the improved VMD and the ELM network optimized by the bat algorithm have a certain degree of accuracy improvement. Moreover, the effect of the model is further improved after error sequence prediction. In terms of MAE value, the accuracy of the proposed models in dataset1 and dataset2 is improved by 65.61% and 60.58% respectively compared with the ELM model, which verifies the validity of the model.

IV. Conclusion and Discussion

This paper proposes a hybrid model system for ultra-short-term wind speed forecasting including series decomposition, intelligent algorithm model and error correction. Firstly, in order to determine the optimal parameters of the variational modal decomposition, the minimum average envelope entropy is used as the selection criterion to reduce the data complexity. Secondly, in order to resolve the randomization of ELM model parameters, choose the bat algorithm optimize the weight and bias of ELM, and good preliminary prediction results are obtained. Finally, for the residual error of VMD decomposition, KDE is used to fit the error sequence distribution, and the error is corrected by sampling to realize the purpose of measuring the error sequence and improving the prediction accuracy.

Combining the above aspects, the Eman wind farm and Sotavento historical wind farm wind speed data are selected to verify the performance of composite model. The experimental results show that compared with other neural network models, the prediction results of the proposed model are more in line with the initial wind speed series. Simultaneously, the error metrics has decreased significantly, and the prediction accuracy has been improved. In conclusion, the model proposed in this paper has achieved satisfactory results and provided a technical basis for the operation of wind farms.

Acknowledgment

This research was supported partly by the National Natural Science Foundation of China (51637005) and the S&T Program of Hebei (G2020502001).

References

[1] Y. Han, L. Mi, L. Shen, C. Cai, Y. Liu, K. Li and G. Xu, ”A short-term wind speed prediction method utilizing novel hybrid deep learning algorithms to correct numerical weather forecasting,” Applied Energy, vol.312, pp.118777, 2022

[2] Y. Zhang, Y. Zhao, X. Shen and J. Zhang, “A comprehensive wind speed prediction system based on Monte Carlo and artificial intelligence algorithms,” Applied Energy, vol.305, pp.1-19, 2022.

[3] C. William, Y. Liu, Y. Liu, Y. Zhang, M. William and W. Thomas, “The impact of model physics on numerical wind forecasts,” Renewable Energy, vol.55, pp. 347-356, 2013.

[4] P. Poggi, M. Muselli, G. Notton, C. Cristofari and A. Louche, “Forecasting and simulating wind speed in Corsica by using an autoregressive model,” Energy Conversion and Management, vol.44 (20), pp.3177-3196, 2003.

[5] G. Riahy and M. Abedi, “Short term wind speed forecasting for wind turbine applications using linear prediction method,” Renewable Energy, vol.33 (1), pp.35-41. 2008.

[6] Y. Zhang, Y. Zhao, C. Kong and B. Chen, “A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic, ” Energy Conversion and Management, vol.203, pp.112254,2020.

[7] M. Liu, L. Ding, and Y. Bai, “Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction,” Energy Conversion and Management, vol. 233, pp.113917, 2021.

[8] M. Liu, Z. Cao, J. Zhang, L. Wang, C. Huang, and X. Luo. “Short-term wind speed forecasting based on the Jaya-SVM model,” International Journal of Electrical Power & Energy Systems, vol. 121, pp.106056, 2021.

[9] L. Hua, C. Zhang, T. Peng, C. Ji and M. Nazir. “Integrated framework of extreme learning machine (ELM) based on improved atom search optimization for short-term wind speed prediction,” Energy Conversion and Management, vol. 252, pp.115102, 2022.

[10] Y. Zhang, B. Chen, G. Pan and Y. Zhao. “A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting,” Energy Conversion and Management, vol. 195, pp. 180-197, 2019.

[11] L. Ding, Y. Bai, M. Liu, M. Fan and J. Yang. “Predicting short wind speed with a hybrid model based on a piecewise error correction method and Elman neural network,” Energy, vol. 244(A), pp.122630, 2022.

[12] A. Altan, S. Karasu and E. Zio. “A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer,” Applied Soft Computing, vol.100, pp.106996, 2021.

[13] S. Jalali, S. Ahmadian, M. Khodayar, A. Khosravi, M. Shafie-khah, S. Nahavandi, J. Catalão,” An advanced short-term wind power forecasting framework based on the optimized deep neural network models,” International Journal of Electrical Power & Energy Systems,vol.141,pp. 108143,2022.

[14] M. Shivaie, M. Mokhayeri, M. Kiani-Moghaddam, A. Ashouri-Zadeh,” A reliability-constrained cost-effective model for optimal sizing of an autonomous hybrid solar/wind/diesel/battery energy system by a modified discrete bat search algorithm,” Solar Energy, vol. 189, pp. 344-356, 2019.

[15] X. Wan, W. Sun, K. Chen, and X. Zhang,” State degradation evaluation and early fault identification of wind turbine bearings,” Fuel, vol. 311, pp. 122348, 2022.

[16] R. Goyal, K. Reddy,” Numerical investigation of entropy generation in a solar parabolic trough collector using supercritical carbon dioxide as heat transfer fluid,” Applied Thermal Engineering, vol. 208, pp. 118246, 2022.

[17] H. Liu, X. Mi, and Y. Li, ” An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm,” Renewable Energy, vol. 123, pp. 694-705, 2018.

[18] Y. Hu, L. Chen,” A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm,” Energy Conversion and Management, vol. 173, pp. 123-142, 2018.

[19] Q. Yang, N. Dong, J. Zhang,” An enhanced adaptive bat algorithm for microgrid energy scheduling,” Energy, vol. 232, pp. 121014, 2021.

[20] S. Ali, G. Yang, C. Huang,” Performance optimization of linear active disturbance rejection control approach by modified bat inspired algorithm for single area load frequency control concerning high wind power penetration,” ISA Transactions, vol. 81, pp. 163-176, 2018