Fuzzy Systems and Data Mining IX
A.J. Tallón-Ballesteros and R. Beltrán-Barba (Eds.)

doi:10.3233/FAIA231038

Application of a Hybrid Model of HPO and LSTM Neural Network Based on ICEEMDAN for Wind Speed Forecasting

Qi Bi, Yu-long Bai[1], Yu-ting Huang, Jun-tao Hong, Zhi-ru Chen and Zai-hong Hou
College of Physics and Electrical Engineering, Northwest Normal University, Lanzhou, China

Abstract. Wind power has the benefits of low cost, low emission, abundant resources, and renewability. The inherent randomness, intermittency, and fluctuation of wind power bring about the volatility of wind power generation. Ameliorating the precision of wind speed prediction has great significance. This study aimed to put forward a hybrid model of the hunter–prey optimization (HPO) and the long short-term memory (LSTM) neural network based on improved complementary ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) to acquire the exact wind speed forecast. First, the ICEEMDAN–HPO–LSTM model used ICEEMDAN to preprocess the raw wind speed sequence and then used the HPO–LSTM model to forecast each decomposed subsequence. Ultimately, the last predicted outcomes of the original wind speed sequence were attained by synthesizing all prediction subseries. Five comparison models were established based on three sets of data with different sequence lengths in Inner Mongolia, China, to test the dependability and utility of the model, and the advantages of the model were proved. The findings displayed that (1) the constitution of ICEEMDAN decomposition and HPO–LSTM could ameliorate the behavior of wind velocity forecast; and (2) the average values of the mean absolute error, mean absolute percentage error, root mean square error, and determination coefficient (R^2) of the three datasets were 0.22411, 4.60277%, 0.27590, and 0.99719, respectively. The proposed prediction model can be used for wind speed forecasts.

Keywords. Hunter–prey optimization, hybrid models, improved complementary ensemble empirical mode decomposition with adaptive noise, long short-term memory neural network, wind speed prediction

1. Introduction

In the wake of the worsening of the environment, sustainable energy, such as solar energy, water energy, and wind energy, accounts for an increasing proportion of energy generation [1]. Among the numerous nonfossil fuels, wind energy is unpolluted, resourceful, economic, and essential. Wind power generation shows a great development trend in the domain of renewable energy sources. The inherent randomness, intermittency, and fluctuation of wind power generation bring about the nondeterminacy of wind power generation [2], of which wind speed is the principal consideration. Previous studies have

[1] Corresponding Author: Yu-long Bai, Northwest Normal University, Lanzhou, China; E-mail: baiyulong@nwnu.edu.cn.

proposed a large number of wind speed forecast models to exactly forecast the wind speed, which can be roughly divided into four classes: physical models, statistical models, artificial intelligence models, and hybrid models [3].

Physical methods use weather data from a meteorological observatory or remote sensing data to model and forecast. This process requires many resources, and the modeling is complex. The statistical model predicts the wind speed through past records, which reflect the relevance between the input and the output [4]. Artificial intelligence methods emerged quickly in the domain of wind speed forecast in the past to seize these nonlinear characteristics of wind speed changes, greatly improving the prediction accuracy. Recently, hybrid models have been increasingly favorably received and valued by researchers. Many basic models are combined to generate a composite model with strong forecast capacity by taking advantage of the goodness of each model to restrain the poor efficiency of a single model. As the main current forecasting model, the structure of hybrid models has become increasingly complicated and varied. The composition of recent hybrid methods mainly includes data preprocessing, optimization algorithms, determining the weight of the hybrid model, and error postprocessing [5].

2. Methodologies

This section introduces the basic concepts of ICEEMDAN and long short-term memory (LSTM), as well as the detailed theory of the hunter–prey optimization (HPO) algorithm.

2.1. Improved complementary ensemble empirical mode decomposition with adaptive noise

Improved complementary ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) is an improved signal handling method besides Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD), Complementary Ensemble Empirical Mode Decomposition (CEEMD), and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN). The improved CEEMDAN method effectively solved problems such as mode duplication, residual noise, and pseudo-Intrinsic Mode Function (IMF)components. Currently, ICEEMDAN decomposition is widely applied in the territory of metal price prediction [6], stock price prediction [7], and wind speed prediction [8]. The ICEEMDAN decomposition algorithm is as follows:

Step 1: Equation (1) indicates that white noise is added to the raw signal:

$$x^{(i)} = x + \beta_0 E_i[\omega^{(i)}] \tag{1}$$

where $\omega^{(i)}$ means the added white noise, and x is the signal used for disassembling.

Step 2: Equation (2) computes the numerical number of the IMF_1 modality component:

$$\tilde{d}_1 = x - r_1 = x - \frac{1}{I}\sum_{i=1}^{I} M[X^i] \tag{2}$$

Step 3: Equation (3) computes the numerical number of the IMF_2 modality component:

$$\tilde{d}_1 = r_1 - r_2 = r_1 - \frac{1}{I}\sum_{i=1}^{I} M[r_1 + \beta_1 E_2[\omega^{(i)}]] \tag{3}$$

Step 4: Equation (4) computes the numerical number of the kth modality weight IMF$_k$:

$$\tilde{d}_k = r_{k-1} - r_k = r_{k-1} - \frac{1}{I}\sum_{i=1}^{I} M[r_{k-1} + \beta_{k-1}E_k[\omega^{(i)}]] \tag{4}$$

2.2. Long short-term memory

Hochreiter et al. [9] proposed the LSTM neural networks to settle the prolonged dependence problem of conventional recurrent neural networks. LSTM can be better represented in a long sequence than the normal RNN. LSTM is widely used in practical applications and is suitable for tasks related to sequence learning, for example, voice recognition, time series analysis, and part of speech tagging. The unit structure and extension diagrams of LSTM are illustrated in Figure 1.

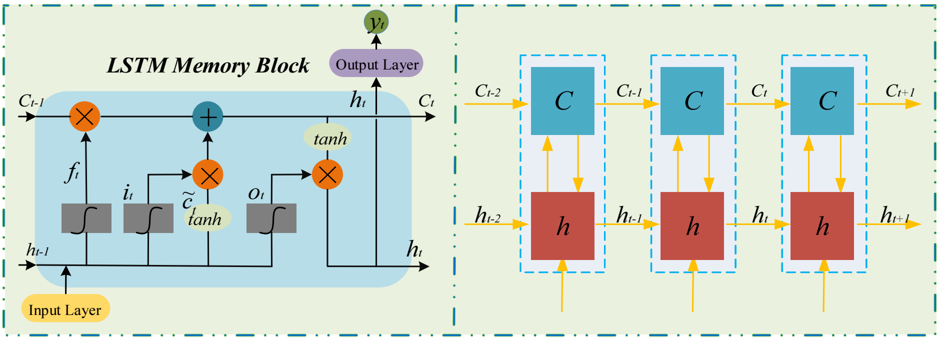

Figure 1. LSTM unit structure diagram and extension diagram.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5}$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{6}$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$
$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{7}$$

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{8}$$

$$o_t = \sigma(W_O \cdot [h_{t-1}, x_t] + b_O) \tag{9}$$

$$h_t = o_t \times tanh(C_t) \tag{10}$$

where i_t, f_t, and o_t are the input, forget, and output gate, respectively; C_t is the cell unit; W_* is the weight; b_* is the threshold value of each function; and σ represents the function of the sigmoid.

2.3. Hunter–prey optimization

Iraj Naruei et al. [10] proposed a new population-based optimization algorithm in 2022: the HPO. The algorithm simulates the conduct of predators such as panthers and lions that feed on bucks and antelope.

2.3.1. Initialization

The predator population is randomly initialized using Equation (11):

$$x_i = rand(1,d) * (u - l) + l \tag{11}$$

where x_i is the hunting position, l is the minimum (lower limit), u is the maximum (upper limit), and d is the amount of problem variables (dimension).

2.3.2. Predator search

Equation (12) is the search mechanism of predators:

$$x_{i,j}(t+1) = x_{i,j}(t) + {}^{1}/_{2}\,[(2BZP_{pos(j)} - x_{i,j}(t)) + (2(1-B)Z\gamma(j) - x_{i,j}(t))] \tag{12}$$

where $x_{i,j}(t)$ is the present location of the predator, $x_{i,j(t+1)}$ is the next location of the predator, P_{pos} is the location of the game, γ is the mean numerical number of the whole locations, and Z is an adaptive parameter. B is the equilibrium parameter between exploration and development. The numerical number is reduced from 1 to 0.02 in the iterations, which is displayed by Equation (14):

$$P = Q_1 < B; IDX = (P == 0) \tag{13}$$

$$Z = Q_2 \otimes IDX + Q_3 \otimes (\sim IDX) \tag{14}$$

where Q_1 and Q_3 are stochastic vectors in the range [0, 1], P is the repertory figure of $Q_1 < B$, Q_2 is a random digit in [0, 1], IDX is the repertory figure of vector Q_1 satisfying the circumstance ($P == 0$), and B is the equilibrium parameter between exploration and development. The numerical number is reduced from 1 to 0.02 in the iterations. As displayed in Equation (15):

$$B = 1 - i\left(\frac{0.98}{M}\right) \tag{15}$$

where i is the numerical digit of current iterations and M is the maximum numerical digit of iterations.

In Equation (16), the expression of γ is as follows:

$$\gamma = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{16}$$

The Euclidean length is displayed as:

$$H_{euc(i)} = \left(\sum_{j=1}^{d}(x_{i,j} - \gamma_j)^2\right)^{\frac{1}{2}} \tag{17}$$

On the basis of the shooting scenario, when the predator seizes the game, the game is about to be killed and the predator moves to the position of the new game. To settle the problem, consider the mechanism of reduction or decrease as displayed in Equation (18):

$$k = round(B \times N) \tag{18}$$

where N is the number of search groups. At first, the value of k equals N of the algorithm. The last search target far away from the mean position a of the search thing is chosen as the quarry and seized by the hunter. The safest place is assumed to be a global optimal location because it gives a better survival opportunity for the prey, and the predator may pick another game. Equation (19) is used for interchanging the location of the game:

$$x_{i,j}(t+1) = T_{pos(j)} + BZ\cos(2\pi Q_4) \times [T_{pos(j)} - x_{i,j}(t)] \tag{19}$$

where $x_{i,j}(t)$ is the immediate location of the captured food, $x_{i,j}(t+1)$ is the next iteration location of the pillage, $T_{pos}(j)$ is the best location of the overall situation, Z is the adaptive parameter according to Equation (14), and Q_4 is a random digit within the range [0, 1]. The function of *cos* and its input arguments permit the next game location to become the global optimal position of different radii and angles. To select predators and prey, Equations (12) and (19) are combined to obtain the following equations:

$$x_i(t+1) = x_i(t) + {}^1/_2\,[(2BZP_{pos(j)} - x_i(t)) + (2(1-B)Z\gamma(j) - x_i(t))] \quad (20)$$

$$x_i(t+1) = T_{pos} + BZ\,cos(\,2\pi Q_4) \times [T_{pos} - x_i(t)] \quad (21)$$

where Q_4 is a random digit within [0,1] and β is a balance parameter, which is set to 0.1 in this study. If $Q_5 < \beta$, the hunt group is considered hunters, and the next location is updated with Equation (20); or else, the search group is considered as prey, and the next location is updated with Equation (21).

3. Framework of the proposed hybrid model

The hybrid model frame is displayed in Figure 2. The main process is as follows:

(1) ICEEMDAN is used for decomposing the primitive wind velocity data, and the decomposition result is to generate n intrinsic mode functions (IMF_1~IMF_n) and a residual R.

(2) HPO is used for optimizing LSTM model parameters.

(3) The ultimate forecast result is the combination of the IMF and the residual error forecast results.

4. Case study

4.1. Data description

The study data came from wind power plants in Inner Mongolia, China. The time span per step of the primitive wind speed data was 15 min, with three groups of diverse datasets. The statistical data of three wind speed datasets are displayed in Table 1. The sizes of these three datasets were diverse: 5 days, 7 days, and 7 days. In Table 1, T, T_a, and T_b, respectively, represent the size of total samples, training set samples, and test set samples. Min, mean, max, std, var, skewness, and kurtosis are the abbreviations of the minimum, mean, maximum, standard deviation, variance, skewness, and kurtosis of the total sample, respectively.

The top 80% of each wind speed series was a training set, and the remaining was a test set. According to skewness, the first and the third datasets were biased to the left, and the value was less than zero. This indicated more abnormal values on the left side of the data. The skewness values of the second dataset were all greater than zero, indicating abnormal digits on the right side of the distribution. Kurtosis is usually used for identifying abnormal values in a preset dataset. The higher the kurtosis, the higher the peak in the data sequence.

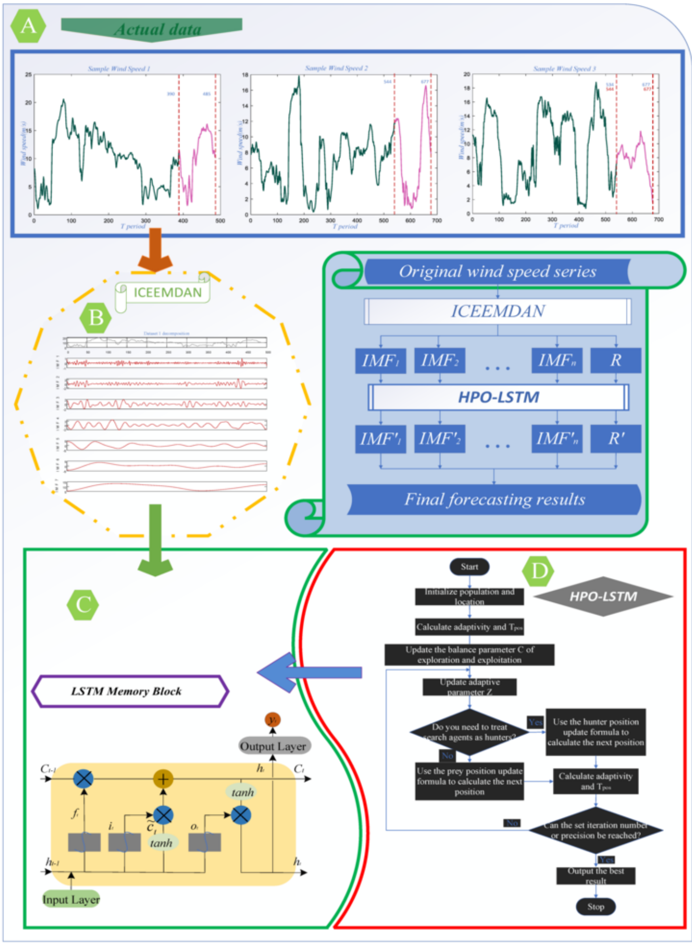

Figure 2. Framework of the hybrid forecasting model.

Table 1. Statistics of the three datasets

Dataset	***T*/period**	T_a	T_b	**Min (m/s)**	**Mean (m/s)**	**Max (m/s)**	**Std**	**Var**	**Skewness**	**Kurtosis**
Dataset 1	485	390	95	1.09	9.9225	20.6	4.7064	22.1498	-0.0468	-1.0239
Dataset 2	677	544	133	0.33	7.1295	17.85	4.0561	16.4521	0.4233	-0.3569
Dataset 3	677	544	133	0.74	9.0723	18.83	4.7239	22.3151	-0.0542	-1.1370

4.2. Model parameter setting

In this study, the hyperparameters of the LSTM neural network were optimized.

All simulations were conducted using the MATLAB R2021a and Python 3.9 platform operating on a Windows 10, with a 3.40-GHz Intel Core i5-11300H CPU and a 64-bit 16-GB RAM.

The hyperparameters in this study are shown in Table 2.

Table 2. LSTM Parameters

Model	Symbol	Meaning	Value
LSTM	n_i	Number of input layer nodes	10
	n_h	Number of hidden layer nodes	Interval optimization with lower and upper limits of [10, 300]
	n	Number of hidden layers	2
	n_o	Number of output layer nodes	1
	η	Fixed learning rate	Interval optimization with lower and upper limits of [0. 005, 0. 05]
	T	Size of batch	One tenth of the training set
	E_p	Epochs of training	Interval optimization with lower and upper limits of [100, 800]

4.3. Experiment: comparison with other models

In this study, the wind speed series in the prediction dataset was delimited into a training set and a test set in the ratio of 8:2. Then, single models, such as Autoregressive integrated moving average (ARIMA), Back propagation neural network (BPNN), Long short-term memory (LSTM), VMD (Variational Mode Decomposition)–LSTM, EMD (Empirical Mode Decomposition)–HPO–LSTM, and ICEEMDAN–HPO–LSTM, were used to forecast the dataset. The test results displayed that the ICEEMDAN–HPO–LSTM model had the optimum forecast effect and could efficaciously increase the forecast precision. Table 3 shows the test results, and the bold values signify that the prediction results are better.

We compared four performance indicators, MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Square Error), and R^2, to display the forecast endings of various kinds of models more distinctly, as displayed in Table 3.

Figure 3 displays the forecast results of six diverse models for three groups of wind speed record. The single model forecast results were unstable and needed further improvement. Each single model was run 10 times to test the dependability of the experimentation.

5. Discussions and Outlook

Increasing the precision of wind speed forecasting is important for wind energy transformation. In this study, we proposed and analyzed an ICEEMDAN–HPO–LSTM hybrid dynamic prediction model based on ICEEMDAN decomposition and HPO. The model was validated using three wind speed datasets and contrasted with the outcomes of five other models. The undermentioned conclusions were based on the point prediction evaluation index, line chart, scatter chart, and histogram of the model.

Table 3. Data of Model Comparison Experiment (MATLAB R2021a, Python 3.9, Windows 10, 3.40-GHz Intel 11300H, 16-GB RAM)

Dataset	Models	MAE	MAPE	RMSE	R^2
Dataset 1	ARIMA	1.0876	17.2565	1.3031	0.91976
	BPNN	0.34717	0.05104	0.51003	0.98818
	LSTM	0.65687	11.2856%	0.91751	0.95905
	VMD-LSTM	0.36054	5.7927%	0.50174	0.98931
	EMD-HPO-LSTM	0.5753	8.5384%	0.70799	0.98423
	ICEEMDAN-HPO-LSTM	**0.33749**	**4.9487%**	**0.39986**	**0.9962**
Dataset 2	ARIMA	1.7561	75.0865	2.0755	0.96636
	BPNN	0.24501	0.082264	0.31824	0.99616
	LSTM	0.55392	18.9988%	0.67019	0.98545
	VMD-LSTM	0.44384	16.5308%	0.51864	0.99332
	EMD-HPO-LSTM	0.26948	9.5948%	0.34023	0.99644
	ICEEMDAN-HPO-LSTM	**0.22282**	**6.6486%**	**0.27119**	**0.99847**
Dataset 3	ARIMA	0.5686	9.6747	0.70735	0.92882
	BPNN	0.12992	0.017991	0.16005	0.99431
	LSTM	0.24185	3.9137%	0.3078	0.97938
	VMD-LSTM	0.16646	2.4528%	0.21519	0.98919
	EMD-HPO-LSTM	0.16104	2.5178%	0.20979	0.99069
	ICEEMDAN-HPO-LSTM	**0.11203**	**2.211%**	**0.15666**	**0.99685**

(1) The ICEEMDAN mean decomposed into a few subsequences without transforming the original data, thus improving the overall accuracy.

(2) Verifying the proposed model using three datasets and five comparison models showed that the proposed hybrid model had great forecast precision and robustness.

In conclusion, the model offered a more trustworthy and exact wind speed forecasting method using wind power generation systems. The proposed model can solve not only the wind speed prediction issue but also other time sequence forecasting problems. For instance, the model can be used to forecast metal prices, carbon prices, and solar radiation intensity of solar power generation. In the future, a metaheuristic algorithm should be used to optimize model parameters, correct the predict errors, and adopt double decomposition or quadratic decomposition to improve prediction accuracy and stability.

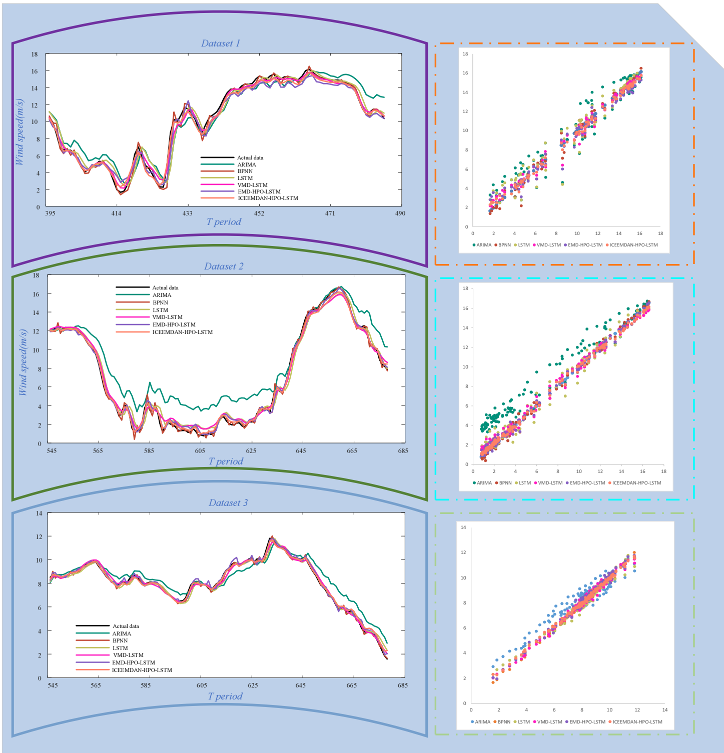

Figure 3. Comparison diagram of single model forecast experiments.

References

1. Wang, L., X. Li, and Y. Bai, Short-term wind speed prediction using an extreme learning machine model with error correction. Energy Conversion and Management, 2018. 162: p. 239-250.
2. Bai, Y., et al., Double-layer staged training echo-state networks for wind speed prediction using variational mode decomposition. Applied Energy, 2021. 301(OCT.15): p. 117461.
3. Ding, L., et al., Predicting short wind speed with a hybrid model based on a piecewise error correction method and Elman neural network. Energy, 2022. 244: p. 122630.
4. Ma, Z., et al., Application of hybrid model based on double decomposition, error correction and deep learning in short-term wind speed prediction. Energy Conversion and Management, 2020. 205: p. 112345.
5. Ding, L., et al., Serial-parallel dynamic echo state network: A hybrid dynamic model based on a chaotic coyote optimization algorithm for wind speed prediction. Expert Systems with Applications, 2023. 212: p. 118789.
6. Huang, Y.-t., et al., Application of a hybrid model based on the Prophet model, ICEEMDAN and multi-model optimization error correction in metal price prediction. Resources Policy, 2022. 79: p. 102969.
7. Wang, J., et al., Asian stock markets closing index forecast based on secondary decomposition, multi-factor analysis and attention-based LSTM model. Engineering Applications of Artificial Intelligence, 2022. 113: p. 104908.
8. Emeksiz, C. and M. Tan, Wind speed estimation using novelty hybrid adaptive estimation model based on decomposition and deep learning methods (ICEEMDAN-CNN). Energy, 2022. 249: p. 123785.
9. Alessandrini, S., et al., An analog ensemble for short-term probabilistic solar power forecast. Applied Energy, 2015. 157: p. 95-110.
10. Naruei, I., F. Keynia, and A.S. Molahosseini, Hunter-prey optimization: algorithm and applications. SOFT COMPUTING, 2022. 26(3): p. 1279-1314.