IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 1, 2020, accepted July 19, 2020, date of publication July 22, 2020, date of current version July 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3011060

Short-Term Wind Power Forecasting Based on VMD Decomposition, ConvLSTM Networks and Error Analysis

ZEXIAN SUN AND MINGYU ZHAO
College of Electrical Engineering, North China University of Science and Technology, Tangshan 063009, China

Corresponding author: Zexian Sun (sunzx@ncst.edu.cn)

This work was supported in part by the Hebei Province Science and Technology Plan Project: Construction and Application of Wind Power Smart Capsule Cloud Management Platform Based on Big Data Technology under Grant 17214304D, and in part by the Funded Project of Hebei Natural Science Foundation: Data Scheduling Optimization and Demand Forecasting of Discrete Manufacturing Industry Based on Big Data under Grant F2018202206.

ABSTRACT Improving the predicted accuracy of wind power is beneficial to maintaining the secure operation and dispatching of the power system. Therefore, a combined model consisting of the variational mode decomposition(VMD), Convolutional Long short memory network(ConvLSTM) and error analysis is conducted for short-term wind power forecasting. Firstly, the VMD algorithm decomposes the wind power signal into an ensemble of components with different frequencies; A novel architecture embedding the convolution operation into LSTM network is procured as the preliminary forecasting engine, which is appropriate for extracting the spatial and temporal characteristics of each sub-series. Afterwards, all the predicted sub-signals would be aggregated to obtain the preliminary forecasting results; For the sake of further mining the unsteady features within the raw wind power series, LSTM modelling the trend of error sequence of the preliminary forecasting result is adopted. Eventually, the final forecasting results is obtained by integrating the forecasting error series and preliminary results. As a result, It can be easily demonstrated that by comparing with the contrastive models, the proposed model achieves the highest prediction performance for wind power series which is difficult to capture.

INDEX TERMS Wind power forecasting, variational mode decomposition, convolutional neural network, long short term memory network, error analysis.

I. INTRODUCTION

With in keeping with the fact that wind energy is strong intermittency, randomness and uncertainty, the large-scale share of wind energy has made great challenges to the stable operation of power system with an irreversible trend [1]. Wind power brings many more uncertainties than conventional generation. Efficient and reliable wind power forecasting becomes extremely important to optimize the operation cost and improve the reliability of the power system with increased wind penetration. Furthermore, accurate short-term wind power forecasting improves the utilisation of wind power, increases system reliability, reduces operating costs and allows efficient load management strategies. Therefore, accurately wind power forecasting is the basis of designing the optimal dispatching plan and efficient accommodation.

Based on the utilization of meteorological data and historical data, the forecasting models can be categorized into two types: Physical model and Statistical model [2], [3].

Physical models resolve a complex scenario considering the weather patterns to convert the predicted meteorological parameters into wind power curve to make an extrapolation of the trend of wind power series [4]. The numerical weather prediction(NWP) [5] model as the core of physical models has fully taken the complexity of the terrain and the resolution of the selected domain into account. A novel architecture integrating the fuzzy process and NWP model was proposed to perform the short-term wind speed forecasting, which significantly enhanced the forecasting precision [6]. However, due to the complex features of the computing process and the

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

atmospheric conditions, the applicability of Physical model is strongly limited.

In comparison, statistical models builds the mathematical functions to fit the mapping relationship between the historical data and output value. Autoregressive(AR) model [7], Autoregressive moving average(ARMA) model [8] and the autoregressive integrated moving average(ARIMA) model [9] have presented improved performance on disposing the inference problem by studying the statistical laws of wind power series. Besides, other methods like support vector machine(SVR) [10], [11] and time series approach [12] have been facilitated to perform wind power forecasting. With the assumption of linearization, statistical models would generate competitive predictions if the stationary condition is met. Therefore, there is still some room to be used for promoting due to the random and intermittent nature of the wind power series.

As the new branch of the statistical models, the artificial intelligence algorithms have been applied to capture the volatility characteristics of the wind speed and wind power series by mining the potential ability of modeling the nonlinear features. The core component in the artificial intelligence algorithms is artificial neural network(ANN) [13], which is extensively utilized in time series prediction domains. A novel methodology combining the adaptive neuron-fuzzy inference system(ANFIS), radial basis function neural network(RBFNN) and least squares support vector machines(LSSVM) is applied for short-term wind power forecasting [14]. An improved radial basis function neural network and an error feedback scheme for daily wind speed and wind power forecasting were employed and evaluated by comparing with four other artificial neural network-based forecasting models [15]. As we can see from the results, under the premise of maintaining the computational efficiency, the proposed model led to higher forecasting accuracy and stability compared with the benchmark models. Benefiting from data mining algorithm, the forecasting performance of wind power had been promoted by deeply searching the inner relation within it [16]. The recurrent neural network(RNN) integrating the bound estimation technology is proposed to build a powerful non-parametric method for deducing the wind speed signal [17]. It can be confirmed that the neural network(NN)-based model can surpass the physical and statistical models to a prominent degree referring to the forecasting performance. However, being prone to falling into the local minimum and overfitting phenomenons, ANN is in the failure of providing a stable predictor engine.

Motivated by the outperforming achievements in the field of computer version and image processing [18], the deep learning algorithms have been introduced to infer the time series. Originating from the conventional artificial neural network, deep learning algorithms overcome the drawbacks of ANN by discovering the distributed features of the raw data. Wang *et al.* [19] has proposed a methodology including the deep belief network(DBN), wavelet transform(WT) and spine quantile regression(QR) to learn the nonlinear and non-stationary features. RNN constitutes a very powerful model that can processes the dynamic sequences. Qin *et al.* [20] employed a multi-task method by training the convolution neural network(CNN) and a Long short-memory network(LSTM) for the wind signal forecasting. An innovative hybrid model including the VMD decomposition, Kullback-Leibler divergence, energy measure and LSTM prediction engine for wind speed causality processing was proposed [21].

Except the application of deep learning algorithms, decomposition technology had also been adopted to mine the volatility of the wind power series, which is in favour of getting rid of the adverse influence of the noise data and providing a more stable predict engine. Secondary decomposition technology consisting of WT algorithm and Singular Spectrum Analysis(SSA) was proposed to disaggregate the wind speed series into several sub-series [22]. Then, Elman neural network(ENN) was designed to forecast the wind speed series. Experimental results demonstrated that the proposed model was superior to the compared models. Peng *et al.* [23] combined the complementary ensemble empirical mode decomposition with adaptive noise(CEEMDAN) and VMD decomposition to build a two-stage decomposition scheme. Extreme learning machine(ELM) was exploited with Adaboost. RT technique to compare the forecasting capability with back propagation neural network(BPNN) and SVM model. Improved complete ensemble empirical mode decomposition adaptive noise(ICEEMDAN) combining the intelligent optimization algorithm was incurred for short-term wind speed forecasting [24].

Review the aforementioned literature, the following points can be summarized as:

(a) Physical models need some intricate environmental factors, of which a large amount of features for depicting the variation tendency of wind power series can be extracted. The prevailing difficulties in the process of capturing the topographic data make it a laborious task to guarantee the steerability of wind power forecasting.
(b) In consideration of the ascendency of seizing the linear characteristics, statistical approaches have privileges to eliminate the forecasting bias with respect to the steady series. However, as for the volatile wind power sequences, its theoretical supports cannot sufficiently dispose the non-linear distinctions.
(c) Being prone to falling into the local minimum and overfitting phenomenons, intelligence arithmetic is in the failure of providing a stable predictor engine.
(d) An idealized assumption is that the forecasting error series is a naturally white noise. Scientific assessments have confirmed that the previous articles promoted the forecasting performance to a limited degree for the lack of analyzing the error sequence.

In this paper, we propose a novel hybrid model based on VMD decomposition, a designed ConvLSTM network and error feedback scheme to proceed with the above intrinsic shortcomings, which greatly reduces the interference of the initial wind power data. The designed ConvLSTM network performs the convolution operation in the LSTM network, which replaces the matrix multiplication to capture the spatial features. The temporal dependency would be handled with the merits of LSTM network simultaneously, which can produce the efficient and reliable forecasting values. The main contributions of the paper are further explained as follows:

(a) VMD as a non-recursive signal decomposition technology is taken into consideration with the goal of furthest grasping the evolvement trend of the wind power series and then disaggregates the wind power series into a discrete set of components.

(b) Benefiting from the merits of fast learning and simple architecture, the ConvLSTM network is implemented as the predictor engine to remove the noise and extracts the temporal and spatial information of each sub-series. The preliminary results would be obtained by aggregating all the forecasting sub-series. Therefore, one ConvLSTM network is trained for each sub-series, which would result a component of ConvLSTM network with the number predefined by the VMD algorithm.

(c) The error feedback scheme making up for the preliminary results would also be built by LSTM network, which can be modeled by the error series generated from the validate set. The reason why we select LSTM algorithm to obtain the forecasting error series is that LSTM network can remember the irregular trend factor of the long-term periodic components. Final forecasting results would be aggregated by the preliminary results and predicted error series.

The rest of the paper is organized in the following way. Section II describes the specific methodology which consists of VMD decomposition, the designed ConvLSTM network and error analysis. Two cases are illustrated and analyzed to evaluate the proposed model in section III. Section IV offers the conclusion.

II. THE HYBRID VMD-CONVLSTM-LSTM MODEL

A. THE ARCHITECTURE OF THE PROPOSED MODEL

Fig.1 presents the overall framework of the VMD-ConvLSTM-LSTM model. The marching details are divided into four steps as follows:

(1) Decomposition by VMD: VMD algorithm, as the first module, decomposes the wind power signal into an ensemble of intrinsic mode functions(IMFs) owing specific sparsity properties in the bandwidth. The mathematical theory of the VMD algorithm are defined in Section II.B.

(2) Performing the forecasting task: Each IMF is transferred as the trajectory matrix by the sliding window algorithm. ConvLSTM network learns the corresponding matrix to extract the important attributes and bridge the long time information to carry out the issues of forecasting each IMFs. The preliminary forecasting results can be achieved by abbreviating the predictions of the sub-series. The mathematical theory of the CNN network, LSTM network and ConvLSTM network are defined in Section II.C, Section II.D and Section II.E, respectively.

(3) Applying the error scheme analysis: Error sequence would be gained by calculating the difference between real value and forecasting value. LSTM network which is on the heels of the preliminary model mines the latent information within it.

(4) Obtaining the final results and evaluating the proposed model: Integrating the preliminary forecasting results and error analysis to obtain the final results, which would be evaluated and compared with the benchmark models in the experimental section.

B. VARIATIONAL MODE DECOMPOSITION

As an adaptive and non-recursive signal processing method, VMD algorithm decomposes the original time series into several components u_k by determining the correlation frequency with limited bandwidth. Readers may refer to [21] for the detailed procedures about VMD.

The mathematical theory of the variational problem can be presented as

$$\begin{cases} \min\limits_{\{y_k\},\{w_k\}} \{\sum\limits_{k=1}^{K} \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) * y_k(t)] e^{-jw_k t} \right\|_2^2 \} \\ s.t. \sum_k y_k = z(t) \end{cases} \tag{1}$$

where $\{y_k\} = \{y_1, \ldots, y_K\}$ and $\{w_k\} = \{w_1, \ldots, w_K\}$ represent the set of all modes and respective centre frequencies. For converting the constrained variational problem to an unconstrained optimization one, the quadratic penalty α and Lagrange multipliers λ are introduced to encourage reconstruction fidelity and enforces constraints strictly, as the following formula shown

$$\begin{aligned} &L(\{y_k\}, \{w_k\}, \lambda) \\ &= \alpha\{\sum_{k=1}^{K} \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) y_k(t)] e^{-jw_k t} \right\|_2^2 \\ &\quad + \left\| z(t) - \sum_{k=1}^{K} y_k \right\|_2^2 \left\langle \lambda(t), z(t) - \sum_{k=1}^{K} y_k(t) \right\rangle \} \end{aligned} \tag{2}$$

By using the alternate direction method of multipliers (ADMM), each mode and its respective center frequency would be regulated to obtain the optimal y_k, w_k and λ, which stand for the saddle point of the unconstrained optimization problem from the mathematical perspective

$$\widehat{y}_k^{n+1} = \frac{\hat{z}(w) - \sum_{i \neq k} \widehat{y}_i(w) + \frac{\hat{\lambda}(w)}{2}}{1 + 2\alpha(w - w_k)^2} \tag{3}$$

$$\hat{w}_k^{n+1} = \frac{\int_0^\infty w \left| \widehat{y}_k(w) \right|^2 dw}{\int_0^\infty \left| \widehat{y}_k(w) \right|^2 dw} \tag{4}$$

$$\hat{\lambda}^{n+1}(w) = \hat{\lambda}^n(w) + \tau(\hat{z}(w) - \sum_k \widehat{y}_k^{n+1}) \tag{5}$$

where $\hat{z}(w)$, $\widehat{y}_i(w)$ and $\hat{\lambda}(w)$ represent the Fourier transforms of each variable, n is the number of iteration.

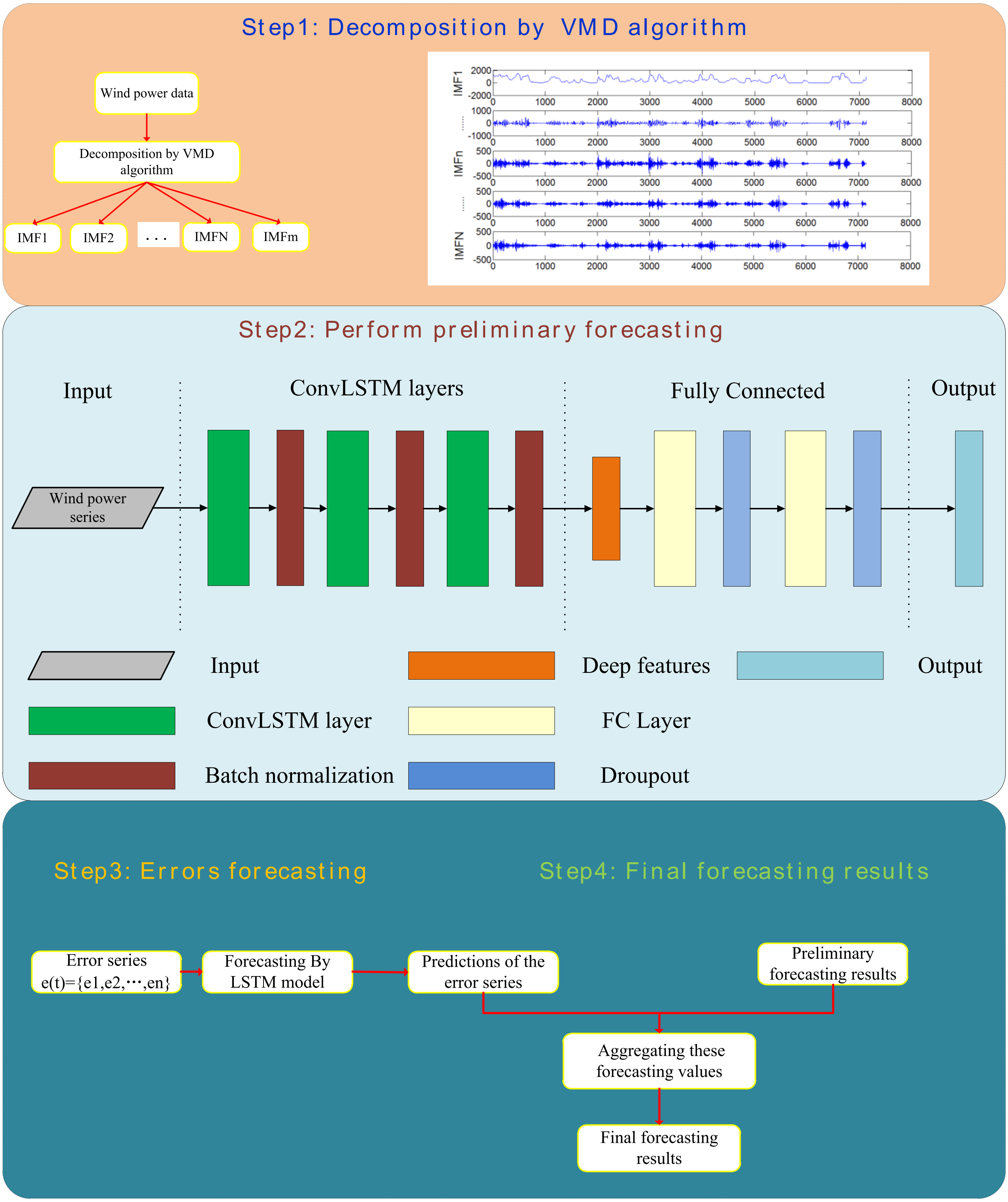

FIGURE 1. The framework of the VMD-ConvLSTM-LSTM model.

C. CONVOLUTION NEURAL NETWORK

CNN [25] comprised of convolution layer and pooling layer is worldwide used to cope with complex characteristics from the multiple wind power series.

Convolution layer is a mathematical operation which merges the two sets of information to map the local features into global features. Convolution layer reduces the size of input vectors as the formula(6) shown:

$$y = \sigma(W^n \otimes x + b^n) \tag{6}$$

where n is the size of kernels. W, x, b are the connecting weights in kernels, the decomposed sub-series of wind power sequences which is taken as the input vector, the vector of bias, respectively. The symbol $\otimes$ presents the convolution operation. σ denotes the activated function, which is adopted by Rectified Linear Unit(ReLU) and the definition is

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \tag{7}$$

Pooling layer further reduces the size of representation of the convolution layer. The extent to which such potential can be developed is determined by the size of the pooling layer. Max pooling as a non-linear subsampling layer combines the output of neuron cluster from previous layer.

D. LONG SHORT TERM MEMORY NETWORK

Constituting an RNN variant, LSTM can preserve the long-term memory by conducting the constant error carousel(CEC) with the memory cell. The extracted vectors are taken as the input which flows to the adaptive units incorporated of the input gate, the forget gate and the output gate. Benefiting from the special mechanism, LSTM can bridge the long time range to carry out the issues of vanishing errors that traditional RNN would always occur. Figure.2 presents the schematic of LSTM.

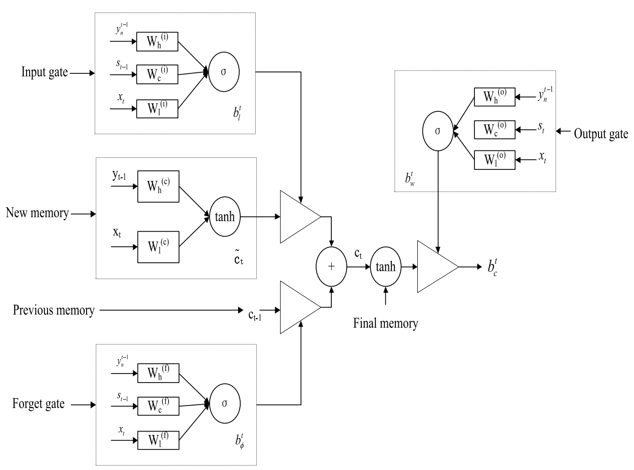

FIGURE 2. The schematic of LSTM.

For the input gate, assuming the input signal, the previous states of the hidden layer and the memory cells are x_i^t, y_n^{t-1} and s_c, respectively. The activation value b_l^t is calculated as the equation(8).

$$b_l^t = f(\sum_{i=1}^{I} w_{il}x_i^t + \sum_{h=1}^{H} w_{hl}y_n^{t-1} + \sum_{c=1}^{C} w_{cl}s_c^{t-1}) \tag{8}$$

where w_{il}, w_{hl}, w_{cl} denote the weight matrices connecting the three gate units and the hidden status, cell memory and input signal. I, H, C represent the dimension of the input gate, hidden gate and output gate, respectively. The activate function $f(\cdot)$ adopts the sigmoid function.

$$f(x) = \frac{1}{1+e^{-x}} \tag{9}$$

The forget gate determines the degree to which the last moment information are regulated to keep in the cell, the output vectors are described as

$$b_\phi^t = f(\sum_{i=1}^{I} w_{i\phi}x_i^t + \sum_{h=1}^{H} w_{h\phi}y_n^{t-1} + \sum_{c=1}^{C} w_{c\phi}s_c^{t-1}) \tag{10}$$

where $w_{i\phi}, w_{h\phi}, w_{c\phi}$ are the sets of weights of the forget gate.

The cell state would be modified by multiplying the previous memory cell state with the output value of forget gate, which is defined as

$$s_c^t = b_\phi^t s_c^{t-1} + b_h^t g(\sum_{i=1}^{I} w_{ic}x_i^t + \sum_{h=1}^{H} w_{hc}y_n^{t-1}) \tag{11}$$

where the activate function is the hyperbolic tangent function, as the formula(12) shows:

$$g(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{12}$$

The output gate filters the information obtained from the memory cell when performing the learning task, the mathematical processes are defined as the formula(13) shown:

$$b_w^t = f(\sum_{i=1}^{I} w_{iw}x_i^t + \sum_{h=1}^{H} w_{hw}y_n^{t-1} + \sum_{c=1}^{C} w_{cw}s_c^t) \tag{13}$$

As mentioned in the above context, the hidden state are defined as the product between the output of the output gate and the activation of the memory cell,that is

$$b_c^t = b_w^t g(s_c^t) \tag{14}$$

E. ConvLSTM NETWORK

ConvLSTM network involves the convolution operation in both input-to-state and state-to-state transitions, which gets rid of the matrix multiplication in LSTM. The temporal information can still determined by the current input vectors and historical cell states, which means that the ConvLSTM network not only take the merits of CNN network to mine the spatial features, but can capture the temporal features simultaneously. The mathematical theory of the ConvLSTM network is given as follows:

$$b_l^t = f(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1}) \tag{15}$$
$$b_\phi^t = f(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1}) \tag{16}$$
$$C_t = b_\phi^t \circ C_{t-1} + b_l^t \circ g(W_{xc} * X_t + W_{hc} * H_{t-1}) \tag{17}$$
$$b_w^t = f(W_{xo} * X_t + W_{ho} * H_{t-1} W_{co} \circ C_t) \tag{18}$$
$$b_c^t = b_w^t \circ g(C_t) \tag{19}$$

To differ from the full connected LSTM model, some mathematical symbols of the ConvLSTM network have been

TABLE 1. The detailed statistics of the four data sets.

Area	DataSet	Numbers	Statistical Indicator			
			Mean(KW)	Var(KW)	Max(KW)	Min(KW)
Wind farm1	All Samples	8929	11667.95	9.42*107	44138.63	-250.26
	Training	7142	13606.29	9.22*107	44138.63	-250.26
	Testing	1787	3921.13	2.71*107	27356.98	-215.68
Wind farm2	All Samples	7623	14087.54	9.21*107	41781.93	-215.68
	Training	6097	13751.80	8.79*107	41781.93	-215.68
	Testing	1526	15428.96	1.06*107	36130.55	-199.19
Wind turbine1	All Samples	8929	461.22	206883	1548.49	-16.99
	Training	7142	450.71	202193	1548.49	-16.99
	Testing	1787	503.22	223424	1548.25	-16.95
Wind turbine2	All Samples	8929	365.48	140902	1551.98	-43.1
	Training	7142	359.71	134995	1551.98	-43.1
	Testing	1787	388.56	163844	1537.02	-38.58

changed. $*$ denotes the convolution operation, $\circ$ denotes the dot operator. W is presented as the parameters of the network. H and C are the hidden state and cell state, respectively. In ConvLSTM network, the feature vectors generated by CNN activations are fed as input to the LSTM network. As a result, it performs both the spatial and temporal operations simultaneously to avoid the lost information.

III. CASE STUDY

A. MODELING DATA

The historical data deployed in this paper was collected from two wind farms in China between the September 1, 2012 and the March 1, 2013. To verify the generalized property of the VMD-ConvLSTM-LSTM model, we conducted four datasets involving two output series from wind turbines and the others from wind farms. The rated power of the wind turbine and wind farm is 1.5MW and 49.5 MW, respectively. Figure. 3 illustrates the trend of four training wind power series and the table1 presents the detailed statistics of the four data sets.

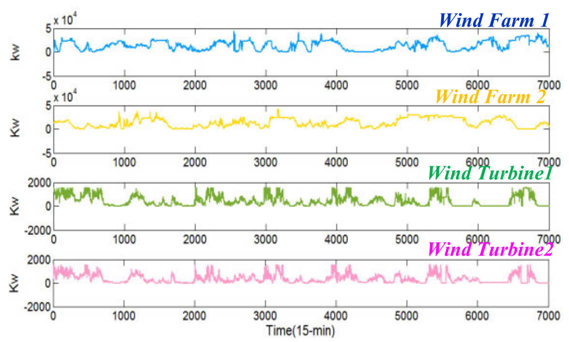

FIGURE 3. The training wind power series.

B. PERFORMANCE EVALUATION INDEXES

Four evaluated metrics, Root Mean Square Error(RMSE), Mean Relative Error(MRE), Mean Absolute Error(MAE) and Mean Squared Error(MSE) are regarded as the forecasting accuracy indexes to provide comprehensive understandings of the fitting effects of the involved approaches, which can be calculated by formulas (20-23):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}[x(t)-\hat{x}(t)]^2}{N}} \tag{20}$$

$$MRE = \frac{\sum_{i=1}^{N}\left(\left|x(t)-\hat{x}(t)\right| \big/ x_r\right)}{N} \tag{21}$$

$$MAE = \frac{\sum_{i=1}^{N}\left|x(t)-\hat{x}(t)\right|}{N} \tag{22}$$

$$MSE = \frac{\sum_{i=1}^{N}[x(t)-\hat{x}(t)]^2}{N} \tag{23}$$

where $x(t)$ is the actual data, x_r is the rated value, $\hat{x}(t)$ is the predicted data, N is the number of forecasting samples.

C. EXPERIMENT 1: ANALYSIS FOR THE FORECASTING RESULTS WITH TWO WIND TURBINES

To verify the generalized capacity of the proposed VMD-ConvLSTM-LSTM model, seven benchmark models including BPNN model, Elman mdoel, LSTM model, the VMD-BP model, the VMD-Elman model, the VMD-LSTM model and VMD-ConvLSTM model are employed to conduct the 15-min, 30-min and 1-h power forecasting of two wind turbines. Figures 4-9 give the comparison results between the proposed model and benchmark models in a visible manner. The evaluated metrics of two wind turbines on different time scales are illustrated in Tables2-3.

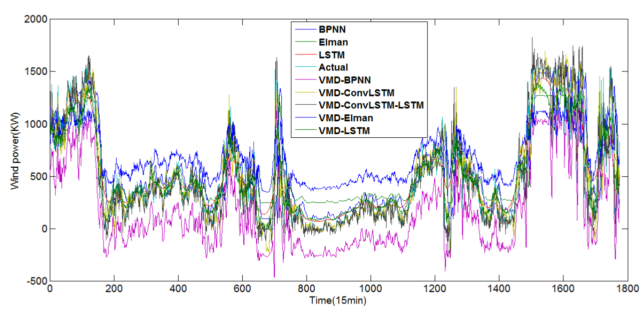

FIGURE 4. 15-min wind power forecasting for the wind turbine#1.

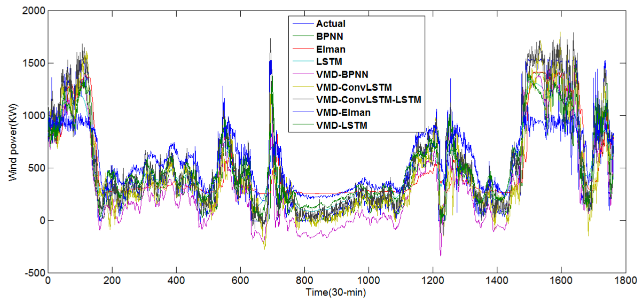

FIGURE 5. 30-min wind power forecasting for the wind turbine#1.

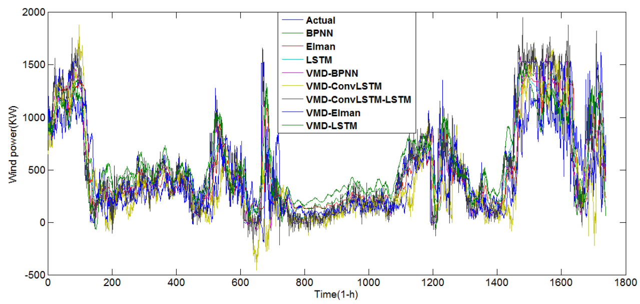

FIGURE 6. 1-h wind power forecasting for the wind turbine#1.

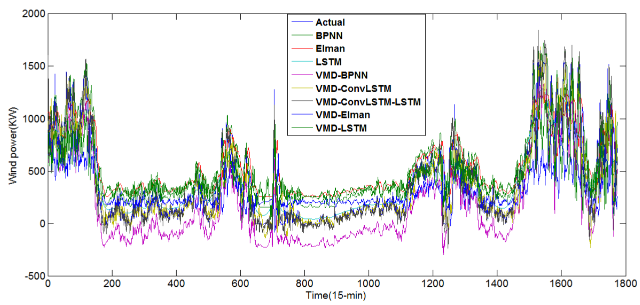

FIGURE 7. 15-min wind power forecasting for the wind turbine#2.

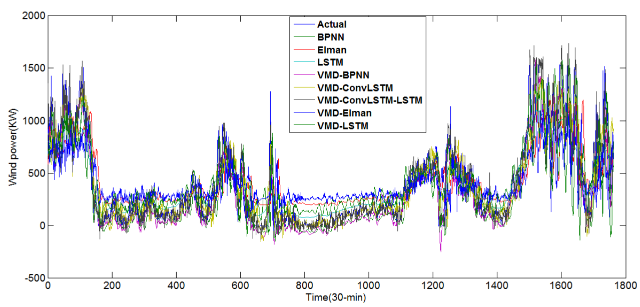

FIGURE 8. 30-min wind power forecasting for the wind turbine#2.

From the results given in Tables 2-3 and Figures 4-9, the general survey of the comparisons in this experiment can be summed up below:

(1)The single LSTM model deeply seizes the dynamic fluctuation of the wind power sequences and exhibits slightly superior forecasting precision and stability to the other two single models. For the wind turbine #1, in terms of the RMSE

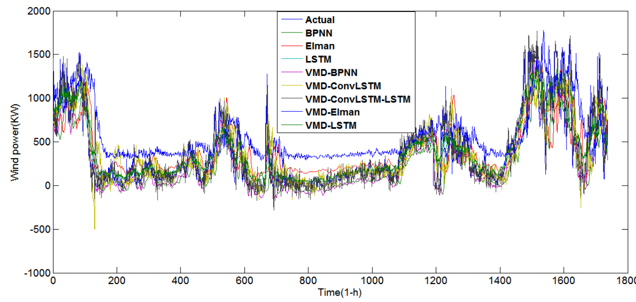

FIGURE 9. 1-h wind power forecasting for the wind turbine#2.

metric, LSTM approach in different time scales produces the values as 111.35KW, 197.26KW and 293.56KW. The corresponding values of MRE are 0.056KW, 0.091KW and 0.132KW, respectively. The MAE and MSE share the analogical performance, in which the enhancements of MAE are 249KW, 39KW and -13.5KW by comparing the LSTM model with BPNN model and the relevant promotion by comparing the LSTM model with Elman model of MAE are 165KW, 106.5KW and 22.5KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the LSTM model can improve the MRE by 0.041KW, 0KW, 0.001KW, the MAE by 61.5KW,0KW, 1.5KW, the MSE by 21767.59KW, 2099.06KW, 3730.91KW and the RMSE by 55.77KW, 6.08KW, 7.91KW compared to the BPNN model; the LSTM model can improve the MRE by 0.044KW, 0.106KW, 0.015KW, the MAE by 66KW, 159KW, 22.5KW, the MSE by 65803.69KW, 88514.62KW, 66835.07KW and the RMSE by 138.97KW, 172.87KW, 115.4KW compared to the Elman model; This consideration indicates that traditional single model limited by the local minimum and over-fitting problems cannot produce the forecasting performance generated from the deep learning methods, in which there are still some room to be used for promoting to some certain extent.

(2)For comparison II, as for the VMD-LSTM model versus the single LSTM model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-LSTM model can improve the MRE by 0.009KW, 0.025KW, 0.026KW, the MAE by 13.5KW, 37.5KW, 39KW, the MSE by 1277.01KW, 19520.95KW, 51510.75KW and the RMSE by 5.89KW, 58.01KW, 107.37KW compared to the LSTM model. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-LSTM model can improve the MRE by 0.041KW, -0.002KW, 0.041KW, the MAE by 61.5KW, -3KW, 61.5KW, the MSE by 10438.9KW, 1753.03KW, 30539.52KW and the RMSE by 34.83KW, 5.25KW, 79.47KW compared to the LSTM model. This peroration suggests that the VMD algorithm can pose positive impacts on the uncertainty analysis of the wind power series.

(3)When comparing the VMD-ConvLSTM-LSTM model with the BPNN model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.02KW, 0.077KW,

TABLE 2. The evaluated results between the proposed model and seven benchmark models for the wind turbine #1.

Forecasting approaches	Time scale	MRE(KW)	MAE(KW)	MSE(KW)	RMSE(KW)
BPNN	15-min	0.222	333	122234.14	349.62
	30-min	0.117	175.5	42791.06	206.86
	1-h	0.123	184.5	66002.75	256.91
ELMAN	15-min	0.166	249	90324.29	300.54
	30-min	0.162	243	118384.16	344.07
	1-h	0.147	220.5	93305.81	305.46
LSTM	15-min	0.056	84	12398.82	111.35
	30-min	0.091	136.5	38911.51	197.26
	1-h	0.132	198	86177.47	293.56
VMD-BPNN	15-min	0.109	163.5	110529.65	332.46
	30-min	0.075	112.5	32112.64	179.20
	1-h	0.087	130.5	16708.15	129.26
VMD-ELMAN	15-min	0.124	186	57941.30	240.71
	30-min	0.124	186	59005.27	242.91
	1-h	0.112	168	59467.70	243.86
VMD-LSTM	15-min	0.047	70.5	11121.81	105.46
	30-min	0.066	99	19390.56	139.25
	1-h	0.106	159	34666.72	186.19
VMD-ConvLSTM	15-min	0.036	54	4550.85	67.46
	30-min	0.066	99	13409.64	115.80
	1-h	0.069	103.5	23189.20	152.28
VMD-ConvLSTM-LSTM	15-min	**0.022**	**33**	**1402.50**	**37.45**
	30-min	**0.040**	**60**	**5389.03**	**73.41**
	1-h	**0.037**	**55.5**	**5530.90**	**74.37**

TABLE 3. The evaluated results between the proposed model and seven benchmark models for the wind turbine #2.

Forecasting approaches	Time scale	MRE(KW)	MAE(KW)	MSE(KW)	RMSE(KW)
BPNN	15-min	0.142	213	49746.84	223.04
	30-min	0.084	126	30856.44	175.66
	1-h	0.113	169.5	57499.24	239.79
ELMAN	15-min	0.145	217.5	93782.94	306.24
	30-min	0.190	285	117272.00	342.45
	1-h	0.141	211.5	120603.40	347.28
LSTM	15-min	0.101	151.5	27979.25	167.27
	30-min	0.084	126	28757.38	169.58
	1-h	0.112	168	53768.33	231.88
VMD-BPNN	15-min	0.107	160.5	30377.00	174.29
	30-min	0.069	103.5	18944.77	137.64
	1-h	0.085	127.5	26370.51	162.39
VMD-ELMAN	15-min	0.141	211.5	57686.43	240.18
	30-min	0.127	190.5	58167.79	241.18
	1-h	0.121	181.5	59609.22	244.15
VMD-LSTM	15-min	0.060	90	17540.35	132.44
	30-min	0.086	129	27004.35	164.33
	1-h	0.071	106.5	23228.81	152.41
VMD-ConvLSTM	15-min	0.044	66	9735.77	98.67
	30-min	0.059	88.5	17838.27	133.56
	1-h	0.044	66	9245.78	96.155
VMD-ConvLSTM-LSTM	15-min	**0.017**	**25.5**	**827.14**	**28.76**
	30-min	**0.042**	**63**	**6229.94**	**78.93**
	1-h	**0.027**	**40.5**	**3533.11**	**59.44**

0.086KW, the MAE by 300KW, 115.5KW, 129KW, the MSE by 120831.64KW, 37402.03KW, 60471.85KW and the RMSE by 312.17KW, 133.45KW, 182.54KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.125KW, 0.042KW, 0.086KW, the MAE by 187.5KW, 63KW, 129KW, the MSE by 48919.7KW, 24626.5KW, 53966.13KW and the RMSE by 194.28KW, 96.73KW, 180.35KW, respectively.

(4)When comparing the VMD-ConvLSTM-LSTM model with the Elman model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.144KW, 0.122KW, 0.11KW, the MAE by 216KW, 183KW, 165KW, the MSE by

88921.79KW, 112995.13KW, 87774.91KW and the RMSE by 312.17KW, 133.45KW, 182.54KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.128KW, 0.148KW, 0.114KW, the MAE by 192KW, 222KW, 171KW, the MSE by 92955.8KW, 111042.06KW, 117070.29KW and the RMSE by277.48KW, 263.52KW, 287.84KW, respectively.

(5)When comparing the VMD-ConvLSTM-LSTM model with the LSTM model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.034KW, 0.051KW, 0.095KW, the MAE by 51KW, 76.5KW, 142.5KW, the MSE by 10996.32KW, 33522.48KW, 80646.57KW and the RMSE by 73.9KW, 123.85KW, 219.19KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.084KW, 0.042KW, 0.085KW, the MAE by 126KW, 63KW, 127.5KW, the MSE by 27152.11KW, 22527.44KW, 50235.22KW and the RMSE by 138.51KW, 90.65KW, 172.44KW, respectively.

(6)When comparing the VMD-ConvLSTM-LSTM model with the VMD-BPNN model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.087KW, 0.035KW, 0.05KW, the MAE by 130.5KW, 52.5KW, 75KW, the MSE by 109127.15KW, 26723.61KW, 11177.25KW and the RMSE by 295.01KW, 105.79KW, 54.89KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.09KW, 0.027KW, 0.058KW, the MAE by 135KW, 40.5KW, 87KW, the MSE by 29549.86KW, 12714.83KW, 22837.4KW and the RMSE by 145.53KW, 58.71KW, 102.95KW, respectively.

(7)When comparing the VMD-ConvLSTM-LSTM model with the VMD-Elman model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.102KW, 0.084KW, 0.075KW, the MAE by 126KW, 130.5KW, 168KW, the MSE by 56538.8KW, 53616.24KW, 53936.8KW and the RMSE by 203.26KW, 169.5KW, 169.49KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.124KW, 0.085KW, 0.094KW, the MAE by 186KW, 127.5KW, 141KW, the MSE by 56859.25KW, 51937.85KW, 56076.11KW and the RMSE by 211.42KW,162.25KW, 184.71KW, respectively.

(8)When comparing the VMD-ConvLSTM-LSTM model with the VMD-LSTM model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.025KW, 0.026KW, 0.069KW, the MAE by 37.5KW, 39KW, 103.5KW, the MSE by 9719.31KW, 14001.53KW, 29135.82KW and the RMSE by 68.01KW, 65.84KW, 111.82KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.043KW, 0.044KW, 0.044KW, the MAE by 64.5KW, 66KW, 66KW, the MSE by 16713.21KW, 20774.41KW, 19695.7KW and the RMSE by 103.68KW, 85.4KW, 92.97KW, respectively.

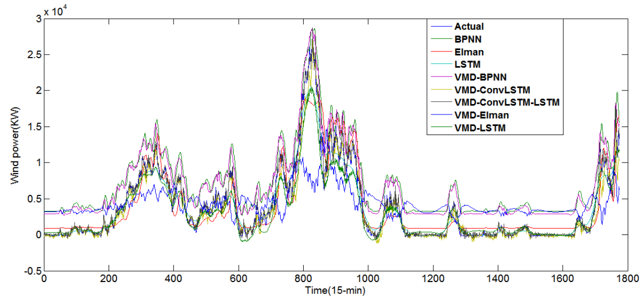

FIGURE 10. 15-min wind power forecasting results of the wind farm #1.

(9)When comparing the VMD-ConvLSTM-LSTM model with the VMD-CNN-LSTM model, it can be observed that in wind turbine #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.014KW, 0.026KW, 0.032KW, the MAE by 21KW, 39KW, 48KW, the MSE by 3148.35KW, 8020.61KW, 17658.3KW and the RMSE by 30.01KW, 42.39KW, 77.91KW, respectively. In terms of the wind turbine #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.027KW, 0.017KW, 0.017KW, the MAE by 40.5KW, 25.5KW, 25.5KW, the MSE by 8908.63KW, 11608.33KW, 5712.67KW and the RMSE by 69.91KW, 54.63KW, 36.715KW, respectively.

(10)The assessments achieved on the above mentioned results have indicated that the proposed model offers more attractive results than the contrast models relative to the forecasting accuracy and stability. Furthermore, the wind power series of two wind farms would be implemented to evaluate the universality of the proposed model in the next section

D. EXPERIMENT 2: ANALYSIS FOR THE FORECASTING RESULTS WITH TWO WIND FARMS

In this section, the proposed VMD-ConvLSTM-LSTM model which has the same parameters space with the above mentioned case is employed to yield the volatile output of two wind farms. The forecasting results which are illustrated in Figures 10-15 and tables4-5 have verified that the proposed model strengthens the forecasting stability and accuracy in a outstanding degree.

The wind power series that were forecast by the proposed model and benchmark models are compared to the actual measurements in Figures10-15 and tables 4-5. It can be observed that the forecasting results of wind farm share the analogical behaviors with the wind turbine. From the results given in Tables 4-5 and Figures 10-15, the general survey of the comparisons in this experiment can be summed up below:

(1)The values in tables 4-5 show that the LSTM model outperforms conventional single models for short-term wind power forecasting. In two cases, the values of MRE, MAE,

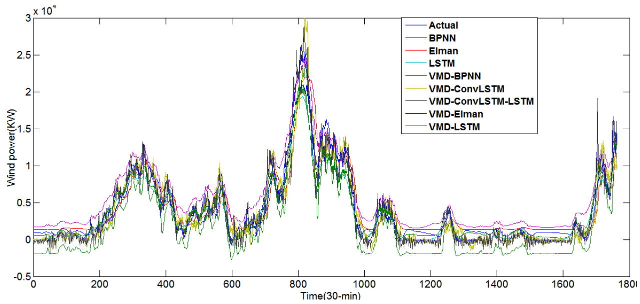

FIGURE 11. 30-min wind power forecasting results of the wind farm #1.

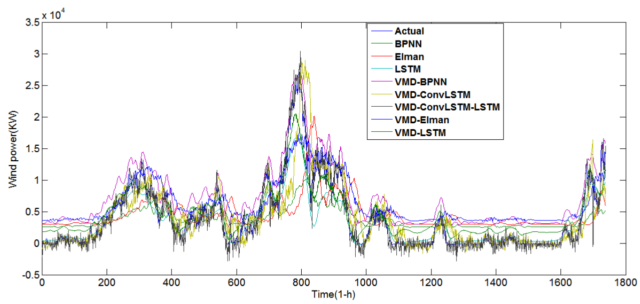

FIGURE 12. 1-h wind power forecasting results of the wind farm #1.

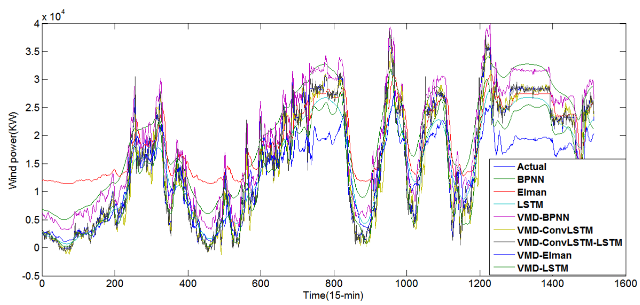

FIGURE 13. 15-min wind power forecasting results of the wind farm #2.

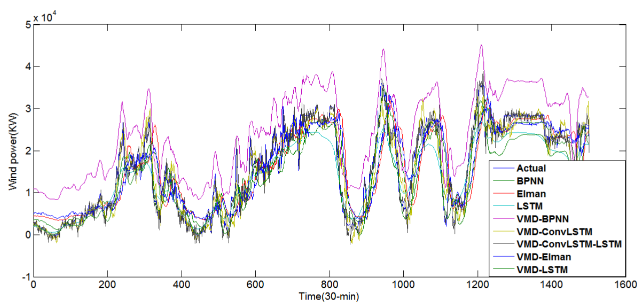

FIGURE 14. 30-min wind power forecasting results of the wind farm #2.

MSE, RMSE of the LSTM model are lower than the two single forecasting models. For the wind farm #1, in terms of the RMSE metric, LSTM approach in different time scales produces the values as 1384.01KW, 1542.07KW and 2077.34KW. The corresponding values of MRE are 0.0194KW, 0.021KW and 0.026KW, respectively. The MAE and MSE share the analogical performance, in which the

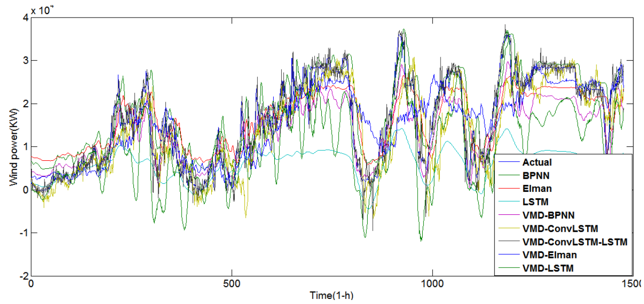

FIGURE 15. 1-h wind power forecasting results of the wind farm #2.

enhancements of MAE are 2059.2KW, 1039.5KW and 1831.5KW by comparing the LSTM model with BPNN model and the relevant promotion by comparing the LSTM model with Elman model of MAE are 1118.7KW, 2376KW and 6534KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the LSTM model can improve the MRE by 0.4039KW, 0.0313KW, 0.037KW, the MAE by 19993.05KW, 1549.35KW, 1831.5KW, the MSE by 658143782KW, 27719586.31KW, 32977032.09KW and the RMSE by 23456.07KW, 2991.46KW, 2437.28KW compared to the BPNN model; the LSTM model can improve the MRE by 0.1699KW, 0.2783KW, 0.187KW, the MAE by 8410.05KW, 13775.85KW, 9256.5KW, the MSE by 177639070.5KW, 393224242.35KW, 270026156.7KW and the RMSE by 11224.09KW, 16939.15KW, 11796.8KW compared to the Elman model.

(2)For the forecasting results of the two farms, the MRE metric obtained from the VMD-LSTM model varies from 0.0148KW to 0.085KW, with the average of 0.035KW. As for the single LSTM model, the range of the MRE metric is from 0.0194KW to 0.091KW, with the average of 0.0396KW. For the wind farm #1, from 15-min to 1-h predictions, VMD-LSTM model can produce the improvements of MRE as 0KW, 0.0062KW and 0.0034KW compared with the LSTM model. In terms of wind farm #2, from 15-min to 1-h predictions, VMD-LSTM model can produce the improvements of MRE as 0.0177KW, 0.0135KW and 0.006KW compared with the LSTM model. The MAE, MSE and RMSE metrics present the similar phenomenon, which gives a strong proof that the VMD algorithm can pose positive impacts on the more broader applicability of the wind power series.

(3)The forecasting accuracy of the VMD-ConvLSTM-LSTM model is higher than the BPNN model obviously, it can be observed that in wind farm #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.0552KW, 0.0311KW, 0.047KW, the MAE by 1831.5KW, 346.5KW, 2326.5KW, the MSE by 8285607.27KW, 767502.93KW, 14146893.4KW and the RMSE by 2952.65KW, 1373.6KW, 2711.92KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.431KW, 0.058KW, 0.097KW, the MAE by 21334.5KW, 2871KW, 4801.5KW, the MSE

TABLE 4. The evaluated results between the proposed model and seven benchmark models for the wind farm #1.

Forecasting approaches	Time scale	MRE(KW)	MAE(KW)	MSE(KW)	RMSE(KW)
	15-min	0.061	3019.5	11765723.21	3430.12
BPNN	30-min	0.042	2079	5015718.576	2239.58
	1-h	0.063	3118.5	15715198.78	3964.24
	15-min	0.042	2079	15041908.99	3878.39
ELMAN	30-min	0.069	3415.5	35741266.56	5978.4
	1-h	0.158	7821	91985170.99	9590.89
	15-min	0.0194	960.3	1915483.68	1384.01
LSTM	30-min	0.021	1039.5	2377887.362	1542.04
	1-h	0.026	1287	4315341.476	2077.34
	15-min	0.059	2920.5	9432453.713	3071.23
VMD-BPNN	30-min	0.037	1831.5	4598537.136	2144.42
	1-h	0.059	2920.5	11039404.95	3322.56
	15-min	0.057	2821.5	17872009.9	4227.53
VMD-ELMAN	30-min	0.046	2277	1891862.703	1375.45
	1-h	0.058	2871	12779338.03	3574.82
	15-min	0.0194	960.3	1915483.68	1384.01
VMD-LSTM	30-min	0.0148	732.6	1073316.72	1036.01
	1-h	0.0226	1118.7	3829222.786	1956.84
	15-min	0.009	445.5	483344.7529	695.23
VMD-ConvLSTM	30-min	0.014	693	855773.0064	925.08
	1-h	0.0216	1069.2	3376369.5	1837.49
	15-min	**0.0058**	**287.1**	**227977.6009**	**477.47**
VMD-ConvLSTM-LSTM	30-min	**0.0109**	**539.55**	**749921.3604**	**865.98**
	1-h	**0.016**	**792**	**1568305.382**	**1252.32**

TABLE 5. The evaluated results between the proposed model and seven benchmark models for the wind farm #2.

Forecasting approaches	Time scale	MRE(KW)	MAE(KW)	MSE(KW)	RMSE(KW)
	15-min	0.443	21928.5	663439533.6	25757.32
BPNN	30-min	0.082	4059	37562802.32	6128.85
	1-h	0.128	6336	63740583.41	7983.77
	15-min	0.209	10345.5	182934822.1	13525.34
ELMAN	30-min	0.329	16285.5	403067458.4	20076.54
	1-h	0.278	13761	300789708	17343.29
	15-min	0.0391	1935.45	5295751.563	2301.25
LSTM	30-min	0.0507	2509.65	9843216.012	3137.39
	1-h	0.091	4504.5	30763551.32	5546.49
	15-min	0.064	3168	11759275.47	3429.18
VMD-BPNN	30-min	0.035	1732.5	71141127.63	8434.52
	1-h	0.078	3861	23163621.38	4812.86
	15-min	0.134	6633	74636294.99	8639.23
VMD-ELMAN	30-min	0.244	12078	231424111.5	15212.63
	1-h	0.457	22621.5	792041394.8	28143.23
	15-min	0.0214	1059.3	1585509.089	1259.17
VMD-LSTM	30-min	0.0372	1841.4	5013971.856	2239.19
	1-h	0.085	4207.5	36833003.76	6069.02
	15-min	0.021	1039.5	2853802.062	1689.32
VMD-ConvLSTM	30-min	0.042	2079	8751657.222	2958.32
	1-h	0.079	3910.5	27627848.69	5256.22
	15-min	**0.012**	**594**	**1464221.003**	**1210.05**
VMD-ConvLSTM-LSTM	30-min	**0.024**	**1188**	**3569076.64**	**1889.2**
	1-h	**0.031**	**1534.5**	**5503199.892**	**2345.89**

by 661975312.6KW, 33993725.68KW, 58237383.52KW and the RMSE by 24547.27KW, 4239.65KW, 5637.88KW, respectively.

(4)The forecasting accuracy of the VMD-ConvLSTM-LSTM model is higher than the Elman model obviously, it can be observed that in wind farm #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.0362KW, 0.0581KW, 0.142KW, the MAE by 891KW, 1683KW, 7029KW, the MSE by 11561856.05KW, 31493050.91KW, 90416865.61KW and the RMSE by 3400.92KW, 5112.42KW, 8338.57KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.197KW, 0.305KW, 0.247KW, the

MAE by 9751.5KW, 15097.5KW, 12226.5KW, the MSE by 181470601.1KW, 399498381.7KW, 295286508.1KW and the RMSE by 12315.29KW, 18187.34KW, 14997.4KW, respectively.

(5)The forecasting accuracy of the VMD-ConvLSTM-LSTM model is higher than the LSTM model obviously, it can be observed that in wind farm #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.0136KW, 0.0101KW, 0.01KW, the MAE by 673.2KW, 499.95KW, 495KW, the MSE by 1687506.08KW,1627966.0012KW, 2747036.0932KW and the RMSE by 906.54KW, 676.06KW, 825.02KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.0271KW, 0.0267KW, 0.06KW, the MAE by 1341.45KW, 1321.65KW, 2970KW, the MSE by 3831530.56KW, 6274139.37KW, 25260351.42KW and the RMSE by 1091.2KW, 1248.19KW, 3200.6KW, respectively.

(6)The forecasting accuracy of the VMD-ConvLSTM-LSTM model is higher than the VMD-BPNN model obviously, it can be observed that in wind farm #1, from 15-min to 1-h predictions, theVMD-ConvLSTM-LSTM model can improve the MRE by 0.0532KW, 0.0261KW, 0.043KW, the MAE by 2633.4KW, 1291.95KW, 2128.5KW, the MSE by 9204476.11KW, 3848615.77KW, 9471099.57KW and the RMSE by 2593.76KW, 1278.44KW, 2070.24KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.052KW, 0.011KW, 0.047KW, the MAE by 2574KW, 544.5KW, 2326.5KW, the MSE by 10295054.47KW, 67572050.99KW, 17660429.49KW and the RMSE by 2219.13KW, 6545.32KW, 2466.97KW, respectively.

(7)The forecasting accuracy of the VMD-ConvLSTM-LSTM model is higher than the VMD-Elman model obviously, it can be observed that in wind farm #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.0512KW, 0.0351KW, 0.042KW, the MAE by 2534.4KW, 1737.45KW, 2079KW, the MSE by 17644032.3KW, 1141941.3421KW, 11211032.65KW and the RMSE by 3750.06KW, 509.47KW, 2322.5KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.122KW, 0.22KW, 0.426KW, the MAE by 6039KW, 10890KW, 21087KW, the MSE by 73172073.99KW, 227855034.9KW, 786538194.9KW and the RMSE by 7429.18KW, 13323.43KW, 25797.34KW, respectively.

(8)The forecasting accuracy of the VMD-ConvLSTM-LSTM model is higher than the VMD-LSTM model obviously, it can be observed that in wind farm #1, from 15-min to 1-h predictions, theVMD-ConvLSTM-LSTM model can improve the MRE by 0.013KW, 0.015KW, 0.028KW, the MAE by 673.2KW, 193.05KW, 326.7KW, the MSE by 1687506.07KW, 323395.35KW, 2260917.40KW and the RMSE by 906.54KW, 170.03KW, 704.52KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.0094KW, 0.0132KW, 0.054KW, the MAE by 465.3KW, 653.4KW, 2673KW, the MSE by 121288.08KW, 1444895.22KW, 31329803.86KW and the RMSE by 49.12KW, 349.99KW, 3723.13KW, respectively.

(9)The forecasting accuracy of the VMD-ConvLSTM-LSTM model is higher than the VMD-ConvLSTM model obviously, it can be observed that in wind farm #1, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.0032KW, 0.0031KW, 0.0056KW, the MAE by 158.4KW, 153.45KW, 277.2KW, the MSE by 255367.15KW, 105851.65KW, 1808064.12KW and the RMSE by 217.76KW, 59.1KW, 585.17KW, respectively. In terms of the wind farm #2, from 15-min to 1-h predictions, the VMD-ConvLSTM-LSTM model can improve the MRE by 0.009KW, 0.018KW, 0.048KW, the MAE by 445.5KW, 891KW, 2376KW, the MSE by 1389581.06KW, 5182580.58KW, 22124648.79KW and the RMSE by 479.27KW, 1069.12KW, 2910.33KW, respectively.

IV. CONCLUSION

Accurate wind power forecasting is crucial in the electricity market. This paper proposes a short-term wind power forecasting model including VMD decomposition, ConvLSTM predictor and error series modelling. VMD decomposition technology is firstly applied to eliminate the non-stationary features of the raw wind power series. ConvLSTM model is implemented as the predictor engine for the preliminary results. Considering the valuable information within the error sequence, the LSTM network modeling the error sequence is built to further strengthen the stability and accuracy of the forecasting results. Seven different models are executed on two experiments, it is worth mentioning that the proposed scheme can significantly increase the forecasting accuracy and stability. For two experiments, the accuracy improvement in MRE, MAE, MSE and RMSE metrics of benchmark models by the proposed scheme are introduced in the experimental section. It demonstrates that forecasting of short wind power output is a promising research direction, and the proposed hybrid model has much potential for promoting the operational performance of wind power system.

However, there are still limitations in this paper. The initial parameters of the VMD algorithm,CNN model and LSTM network are selected by the expertise, wihch would result a large computational cost to obtain the optimal paramters. In the next research work, we will search other intelligent algorithms to optimize the proposed model.

REFERENCES

[1] L. Han, H. Jing, R. Zhang, and Z. Gao, "Wind power forecast based on improved long short term memory network," *Energy*, vol. 189, Dec. 2019, Art. no. 116300, doi: 10.1016/j.energy.2019.116300.

[2] A. Ahmed and M. Khalid, "A review on the selected applications of forecasting models in renewable power systems," *Renew. Sustain. Energy Rev.*, vol. 100, pp. 9–21, Feb. 2019, doi: 10.1016/j.rser.2018.09.046.

[3] E. B. Ssekulima, M. B. Anwar, A. Al Hinai, and M. S. El Moursi, "Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: A review," *IET Renew. Power Gener.*, vol. 10, no. 7, pp. 885–989, Jul. 2016, doi: 10.1049/iet-rpg.2015.0477.

[4] J. Zhang, J. Yan, D. Infield, Y. Liu, and F.-S. Lien, "Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model," *Appl. Energy*, vol. 241, pp. 229–244, May 2019, doi: 10.1016/j.apenergy.2019.03.044.

[5] C. Dou, Y. Zheng, D. Yue, Z. Zhang, and K. Ma, "Hybrid model for renewable energy and loads prediction based on data mining and variational mode decomposition," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 11, pp. 2642–2649, Jun. 2018.

[6] 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.

[7] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "ARIMA-based time series model of stochastic wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 667–676, May 2010.

[8] 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.

[9] M. Caporin and J. Preś, "Modelling and forecasting wind speed intensity for weather risk management," *Comput. Statist. Data Anal.*, vol. 56, no. 11, pp. 3459–3476, Nov. 2012.

[10] Y. Zhang, J. Le, X. Liao, F. Zheng, and Y. Li, "A novel combination forecasting model for wind power integrating least square support vector machine, deep belief network, singular spectrum analysis and locality-sensitive hashing," *Energy*, vol. 168, pp. 558–572, Feb. 2019, doi: 10.1016/j.energy.2018.11.128.

[11] A. Habib, R. Abbassi, A. J. Aristizabal, and A. Abbassi, "Forecasting model for wind power integrating least squares support vector machine, singular spectrum analysis, deep belief network, and locality-sensitive hashing," *Wind Energy*, vol. 23, no. 2, pp. 235–257, Feb. 2020, doi: 10.1002/we.2425.

[12] G. Sun, C. Jiang, P. Cheng, Y. Liu, X. Wang, Y. Fu, and Y. He, "Short-term wind power forecasts by a synthetical similar time series data mining method," *Renew. Energy*, vol. 115, pp. 575–584, Jan. 2018, doi: 10.1016/j.renene.2017.08.071.

[13] Y. Li, H. Wu, and H. Liu, "Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction," *Energy Convers. Manage.*, vol. 167, pp. 203–219, Jul. 2018, doi: 10.1016/j.enconman.2018.04.082.

[14] J. Liu, X. Wang, and Y. Lu, "A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system," *Renew. Energy*, vol. 103, pp. 620–629, Apr. 2017, doi: 10.1016/j.renene.2016.10.074.

[15] G. W. Chang, H. J. Lu, Y. R. Chang, and Y. D. Lee, "An improved neural network-based approach for short-term wind speed and power forecast," *Renew. Energy*, vol. 105, pp. 301–311, May 2017, doi: 10.1016/j.renene.2016.12.071.

[16] W. Wu and M. Peng, "A data mining approach combining K-means clustering with bagging neural network for short-term wind power forecasting," *IEEE Internet Things J.*, vol. 4, no. 4, pp. 979–986, Aug. 2017, doi: 10.1109/JIOT.2017.2677578.

[17] Z. Shi, H. Liang, and V. Dinavahi, "Direct interval forecast of uncertain wind power based on recurrent neural networks," *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1177–1187, Jul. 2018, doi: 10.1109/TSTE.2017.2774195.

[18] M. Längkvist, L. Karlsson, and A. Loutfi, "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern Recognit. Lett.*, vol. 42, pp. 11–24, Jun. 2014.

[19] H. Z. Wang, G. B. Wang, G. Q. Li, J. C. Peng, and Y. T. Liu, "Deep belief network based deterministic and probabilistic wind speed forecasting approach," *Appl. Energy*, vol. 182, pp. 80–93, Nov. 2016, doi: 10.1016/j.apenergy.2016.08.108.

[20] Y. Qin, K. Li, Z. Liang, B. Lee, F. Zhang, Y. Gu, L. Zhang, F. Wu, and D. Rodriguez, "Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal," *Appl. Energy*, vol. 236, pp. 262–272, Feb. 2019, doi: 10.1016/j.apenergy.2018.11.063.

[21] J. Wang and Y. Li, "Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy," *Appl. Energy*, vol. 230, pp. 429–443, Nov. 2018, doi: 10.1016/j.apenergy.2018.08.114.

[22] C. Yu, Y. Li, and M. Zhang, "An improved wavelet transform using singular spectrum analysis for wind speed forecasting based on elman neural network," *Energy Convers. Manage.*, vol. 148, pp. 895–904, Sep. 2017, doi: 10.1016/j.enconman.2017.05.063.

[23] T. Peng, J. Zhou, C. Zhang, and Y. Zheng, "Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and AdaBoost-extreme learning machine," *Energy Convers. Manage.*, vol. 153, pp. 589–602, Dec. 2017, doi: 10.1016/j.enconman.2017.10.021.

[24] J. Song, J. Wang, and H. Lu, "A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting," *Appl. Energy*, vol. 215, pp. 643–658, Apr. 2018, doi: 10.1016/j.apenergy.2018.02.070.

[25] H. J. Sadaei, P. C. de Lima e Silva, F. G. Guimarães, and M. H. Lee, "Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series," *Energy*, vol. 175, pp. 365–377, May 2019, doi: 10.1016/j.energy.2019.03.081.

ZEXIAN SUN was born in Tangshan, Hebei, China, in 1991. He received the Ph.D. degree in automation from the North China University of Science and Technology, in 2019.

He has been teaching with the College of Electrical Engineering, North China University of Science and Technology. His research interests include wind power and the application of deep learning in the wind energy.

MINGYU ZHAO was born in Tangshan, Hebei, China, in 1991.

She has been teaching with the College of Electrical Engineering, North China University of Science and Technology. Her research interest includes the application of deep learning in the wind energy.