Research Article

A Short-Term Wind Speed Forecasting Hybrid Model Based on Empirical Mode Decomposition and Multiple Kernel Learning

Yuanyuan Xu ^{1,2} and Genke Yang ^{1,3}

¹Department of Automation, Shanghai Jiaotong University, Shanghai 200240, China

²College of Information Science and Engineering, Huaqiao University, Xiamen 361021, China

³Ningbo Artificial Intelligence Institute, Shanghai Jiaotong University, Ningbo 315000, China

Correspondence should be addressed to Genke Yang; gkyang@sjtu.edu.cn

Received 10 August 2020; Revised 14 September 2020; Accepted 24 September 2020; Published 4 November 2020

Academic Editor: Jing Na

Copyright © 2020 Yuanyuan Xu and Genke Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Short-term wind speed forecasting plays an increasingly important role in the security, scheduling, and optimization of power systems. As wind speed signals are usually nonlinear and nonstationary, how to accurately forecast future states is a challenge for existing methods. In this paper, for highly complex wind speed signals, we propose a multiple kernel learning- (MKL-) based method to adaptively assign the weights of multiple prediction functions, which extends conventional wind speed forecasting methods using a support vector machine. First, empirical mode decomposition (EMD) is used to decompose complex signals into several intrinsic mode function component signals with different time scales. Then, for each channel, one multiple kernel model is constructed for forecasting the current sequence signal. Finally, several experiments are carried out on different New Zealand wind farm data, and the relevant prediction accuracy indexes and confidence intervals are evaluated. Extensive experimental results show that, compared with existing machine learning methods, the EMD-MKL model proposed in this paper has better performance in terms of the prediction accuracy evaluation indexes and confidence intervals and shows a better ability to generalize.

1. Introduction

Energy is crucial to social and economic development all over the world. With the acceleration of economic globalization, energy demands are growing rapidly [1, 2]. However, with the depletion of fossil fuels and the destruction of the environment, pollution-free and renewable energy sources must be sought [3]. As an inexhaustible and cleaner energy source, wind energy is regarded as an ideal substitute for conventional resources. Wind energy is the main force of this development and has become the main pillar of power supply in the world. According to a Global Wind Energy Council (GWEC) report, by the end of 2019, the total installed capacity of wind power reached 651 million kilowatts, which is 1.3 times higher than the 283.4 million kilowatts in 2012 and 5.93 times higher than the 94 million kilowatts in 2007. All installed wind turbines can cover 6% of global

electricity demand [4]. However, due to the intrinsic properties of wind such as instability, mutability, etc., wind energy sometimes seems to be uncontrollable.

Wind energy has the characteristics of volatility, intermittency, and low energy density, which makes the wind speed and output power of wind farms also uncertain [5]. This uncertainty brings huge challenges to the operation and dispatch of power systems. Short-term wind speed prediction is considered to be one of the most effective and economic means to improve the peak load regulation ability of power grids, enhance the ability of power grids to accept wind power generation, and improve the security and economy of power system operation [6, 7]. Constructing a model to accurately forecast wind speed is very difficult due to the intermittency and fluctuation of wind speed [8].

In past decades, various algorithms such as statistical approaches, machine learning, and deep learning have been

applied to the field of wind speed forecasting [6, 9,–13]. Each method employs different techniques and achieves good performance on different datasets and prediction horizons [14–16]. Statistical learning prediction models include autoregressive integrated moving average (ARIMA) [17, 18], exponential smoothing (ES) [19], grey forecasting method (GFM) [20], hidden Markov models (HMMs) [21], etc. In addition, artificial neural networks (ANNs) and machine learning models are also widely used in wind speed forecasting [22–29]. However, these models have some shortcomings, such as local minima or overfitting for ANNs and parameter sensitivity for genetic programming (GP) [30].

As one of the most popular machine learning methods, support vector machines (SVMs) are often used for regression with small samples and obtaining state-of-the-art performance in wind forecasting. In [31], a hybrid prediction model based on an SVM is proposed where an autoregressive model with time-delay coordinates is used for wind speed data feature selection and a genetic algorithm is used to optimize the parameters. A combined forecasting model based on Markov models and SVMs is proposed in [32]. A finite state Markov model is used to simulate wind fluctuation and seasonality. A prediction modeling method for shale shaker efficiency is proposed in [33], in which principal component analysis (PCA) is used to extract the nonlinear principal components of signals, and then LSSVMs (least squares support vector machines) are used to reconstruct the regression model. In [34], a novel model called IACO-LSSVM is proposed for wind speed forecasting. An improved ant colony optimization algorithm (IACO) is used to optimize the main parameters of LSSVM.

As wind speed data are usually intermittent and unsystematic with extreme nonlinear characteristics, an SVM is very sensitive to the selection of base kernels, especially on small sample data. To address this problem, the adaptive choice of optimal kernels is introduced according to the current testing case, which refers to linear combination techniques on base kernels, called multiple kernel learning (MKL) [35]. Given an input testing sample, the optimal linear combination of base kernels across different support samples is obtained using the training data to get the current testing value. This means, in different kernel domains, one testing sample may adaptively absorb those components that contributed to its intrinsic representation, which is naturally more robust than that with a fixed single kernel domain. The reason is that different kernel domains have different metric spaces, which may contribute to the final prediction score with different confidences. By using MKL, the metrics based on kernel domains can be adaptively combined such that a more robust predictor is obtained for wind forecasting.

Moreover, since wind speed is usually collected as a one-dimensional signal varying with time, directly using raw signals to predict the next state is not trivial, especially when noise or irregular information accompanies the wind speed signal. To address this problem, we employ empirical mode decomposition (EMD) [36] to decompose the raw signal into several intrinsic mode functions (IMFs) with different frequency range and then feed these into the MKL predictor. Extensive experiments on public wind speed prediction data show that our

proposed MKL-based wind speed forecasting method is more effective and competitive than current methods.

The remainder of this paper is organized as follows. Section 2 discusses the proposed method and the hybrid model in detail. Section 3 describes the forecasting performance evaluation criterion. Experiments are discussed in Section 4. The last section concludes the paper.

2. The Proposed Method

2.1. Overview. Empirical mode decomposition is used to decompose complex signals into several intrinsic mode function component signals with different time scales.

Figure 1 shows the proposed wind speed forecasting process based on EMD and MKL, and the steps are as follows:

- (1) EMD is used to decompose the raw wind speed data into a number of IMFs. Since the raw wind speed signals are usually nonlinear and nonstationary, EMD can break down the signal without leaving the time domain, which is useful for analyzing the original signal in the next steps.
- (2) After EMD decomposition, multiple time-domain signals with different frequency bands are obtained. To reduce the scale variations of each signal, we normalize the EMD signals to the same scale space for the kernel learning that follows. The generated signals are scaled into the range $[-1, 1]$ using the translation and scaling transformations.
- (3) Construct one MKL model for each EMD decomposed signal. As prediction is highly complex, multiple kernel matrices on this signal data are adaptively chosen and combined to form the final prediction model.
- (4) Combine all prediction results for all EMD signals. To recover the predicted result of each MKL into the original scale space, the scale transform is inverted in contrast to step (2).

In engineering applications, single-step forecasting often fails to meet actual needs, and it is necessary to use multistep forecasting to predict more long-term points in the future. Therefore, it is necessary to select an appropriate prediction strategy to reduce the prediction error. For this, a recursive model is chosen. Three-step predictions are conducted as are often done in the wind speed prediction literature. The schematic diagram of the recursive multistep prediction model is given in Figure 2. Recursive multistep prediction uses the same prediction model and uses the predicted value of the previous step as a new input to the next step of prediction.

If x_i is the truth value, $\hat{x}_i$ is the predicted value, the size of the history window is 5, and the current time is k , then the predicted value of $\hat{x}_{k+1}$ at time $k+1$ is obtained by $\{x_{k-4}, x_{k-3}, x_{k-2}, x_{k-1}, x_k\}$. The predicted value of $\hat{x}_{k+2}$ at time $k+2$ is obtained by $\{x_{k-3}, x_{k-2}, x_{k-1}, x_k, \hat{x}_{k+1}\}$. The predicted value of $\hat{x}_{k+3}$ at time $k+3$ is obtained by $\{x_{k-2}, x_{k-1}, x_k, \hat{x}_{k+1}, \hat{x}_{k+2}\}$.

FIGURE 1: Wind speed prediction process based on EMD and MKL.

2.2. Wind Speed Signal Decomposition via EMD. EMD is a new method for time-frequency analysis. The basic idea is to decompose the original time series data into several subsequences with different frequency characteristics [36, 37]. EMD can decompose complex input signals into many intrinsic mode functions, which reflect the local characteristic signals of different time scales [38].

According to the explanation of IMF in [39], all functions involving time series data can be decomposed into the following:

$$y(t) = \sum_{i=1}^n f_i(t) + R_n(t), \quad (1)$$

where $f_i(t)$ ($i = 1, 2, \dots, n$) are the different IMFs and $R_n(t)$ is the residue.

Step 1: identify the local maxima and minima of the series $\{y(t)\}$, which can yield the corresponding upper envelope $\{y_{up}(t)\}$ and lower envelope $\{y_{down}(t)\}$ by means of a cubic spline interpolation of all the local extrema.

Step 2: compute the mean envelope $\{h(t)\}$:

$$h(t) = \frac{y_{up}(t) + y_{down}(t)}{2}. \quad (2)$$

Let $m(t) = y(t) - h(t)$.

Step 3: check whether $\{m(t)\}$ is an IMF. If $\{m(t)\}$ is an IMF, then set $f(t) = m(t)$ and replace $\{y(t)\}$ with the residual $R(t) = y(t) - f(t)$; if $\{m(t)\}$ is not an IMF, replace $\{y(t)\}$ with $\{m(t)\}$ and repeat steps 2 and 3 until the termination criterion is reached. The stopping condition of the iterative process is as follows:

$$\sum_{t=1}^l \frac{[m_{i-1}(t) - m_i(t)]^2}{[m_{i-1}(t)]^2} \leq \delta, \quad i = 1, 2, \dots, t = 1, 2, \dots, m, \quad (3)$$

where l is the length of the original series $\{y(t)\}$, i is the number of iterations, and δ is the termination threshold.

Step 4: repeat steps 1–4 to obtain all IMFs.

In short, we can use EMD to decompose the original wind speed data into different subsequences. These IMFs have different frequency characteristics, which is conducive to the implementation of the prediction model. In this paper, the MKL algorithm is used to model each subsequence.

2.3. Signal Prediction Model Based on MKL. MKL is used to learn the optimal kernel from a set of predefined base kernels and has been applied in many fields, such as image classification, face recognition, and so on. As far as we know, MKL is rarely used in wind speed prediction models.

Based on the definition of the kernel slack variable for each base kernel, we now build the MKL model for wind speed prediction. An L2-norm MKL is used in our approach because of its effectiveness in many applications, and it was originally used for classification as described in [40]. For wind speed forecasting, we modify it for the case of regression. Let $S = \{(x_i, y_{-i})\}$ ($i = 1, \dots, n$) denote the training pairs of the historic and the corresponding next prediction values; we consider a linear support vector machine of the form

$$h_{w,b}^{\sim}(X) = \langle \tilde{w}, \Psi(X) \rangle + b, \quad (4)$$

where the nonlinear mapping Ψ maps the original signals to a Hilbert space and $h_{w,b}^{\sim}(X)$ is the predicted result based on the input signals. To deal with the complexity of wind speed forecasting, we may employ M different feature mappings, i.e., $\Psi_m: \chi \rightarrow \mathcal{H}_m$, where $m = 1, \dots, M$. Then, the prediction model with multiple mappings is extended to the following equation:

$$h_{w,b,\theta}^{\sim}(X) = \sum_{m=1}^M \sqrt{\theta_m} \langle \tilde{w}_m, \Psi_m(X) \rangle \mathcal{H}_m + b, \quad (5)$$

where θ_m is the weight of the m -th submodel. Finally, this primal MKL optimization model can be written as follows:

FIGURE 2: The recursive multistep prediction model.

$$\begin{aligned}
 \inf_{w, b, \theta: \theta \geq 0} & C \sum_{i=1}^n V \left(\sum_{m=1}^M \langle W_m, \Psi_m(X_i) \rangle \mathcal{H}_m + b, y_i \right) \\
 & + \frac{1}{2} \sum_{m=1}^M \frac{\|W_m\|_{\mathcal{H}_m}^2}{\theta_m} \\
 \text{s.t.} \quad & \|\theta\|^2 \leq 1.
 \end{aligned} \tag{6}$$

In the above model, θ_m is used to regularize the weight of each model and V is a convex loss function. To solve this objective function, we may iteratively optimize θ and $\{w_1, \dots\}$. For the weight factor θ , by fixing $\{w_1, \dots\}$, we can derive the following solution according to the proposition in [41]:

$$\theta_m = \frac{\|W_m\|_{\mathcal{H}_m}^{2/(p+1)}}{\left(\sum_{m'=1}^M \|W_{m'}\|_{\mathcal{H}_{m'}}^{2p/(p+1)} \right)^{1/p}}, \quad \forall m = 1, \dots, M, \tag{7}$$

where $p=2$. If θ is fixed, we can derive the dual model by using dual theory, i.e.,

$$\begin{aligned}
 \sup_{\alpha} & 1^T \alpha - \frac{1}{2} \|\alpha^T Y K_m Y \alpha_{m=1}^M\|_{p^*} \\
 \text{s.t.} \quad & y^T \alpha = 0, \quad 0 \leq \alpha \leq C1,
 \end{aligned} \tag{8}$$

where K_m is the m -th kernel matrix. This model is regarded as an SVM model, and thus solutions are easily achievable via public tools, such as LibSVM [42]. Finally, the weighted kernel matrices $k_\theta = \sum \theta_m k_m$, $\theta_m \geq 0$ contribute to the prediction output, where θ_m is adaptively learned in the training data while keeping the max margin in the regression process.

It has been clearly observed that the structural risk function is well controlled by introducing the penalty parameter θ , and the solution of the MKL problem can be changed by varying this parameter, giving a novel perspective to such problems. For the kernel design, we may use different types of kernel, such as the RBF kernel, linear kernel, polynomial kernel, and so on. Moreover, different

parameters for these kernels can be employed to increase the quantity of kernel matrices.

3. Experiments

In this section, we verify our method using real wind speed data. First, we introduce the used dataset, then illustrate the evaluation criteria for wind speed forecasting, and finally conduct the experiments and compare the results with other methods.

3.1. Dataset. This study proposes a hybrid approach to predict short-term wind speed. Wind speed data from a wind farm location in New Zealand are used to demonstrate the effectiveness and reliability of the proposed hybrid EMD-MKL forecasting approach. In this paper, we use the open wind speed datasets of a wind farm in New Zealand to verify the validity and reliability of the novel EMD-MKL model. The data are from an organization called NIWA, published in the public report in [43].

We used data collected at the hub height of 85 meters in the STH1 wind farm and selected 5000 10-minute wind speed data samples in 2003 as the experimental data. Of these, 4500 samples are used to train the EMD-MKL model, and the remaining 500 samples are used for testing.

In our simulation, the size of the moving window is set to 5 samples for the 10-minute wind speed series. Then, we slide the windows containing the sample data to train the EMD-MKL model. 500 test samples can slide 495 times, which means that our model can obtain 495 prediction values.

3.2. Evaluation Criteria. In the evaluation of wind speed prediction models, the following error evaluation index functions are typically used: the mean absolute error (MAE), the root mean square error (RMSE), the mean absolute percentage error (MAPE), the mean squared prediction error (MSPE), and the maximum absolute percentage error (MaxAPE) [43–46]. These error evaluation index functions are defined as follows:

$$\begin{aligned}
\text{MAE} &= \frac{1}{n} \sum_{i=1}^n |e(i)|, \\
\text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n e(i)^2}, \\
\text{MAPE} &= \frac{1}{n} \sum_{i=1}^n \left| \frac{e(i)}{x(i)} \right|, \\
\text{MSPE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{e(i)}{x(i)} \right)^2}, \\
\text{MaxAPE} &= \max_{i=1}^n \left| \frac{e(i)}{x(i)} \right|,
\end{aligned} \tag{9}$$

where $e(i) = x(i) - y(i)$, n is the sample size, and $x(i)$ and $y(i)$ are the actual and forecasted values at time period i , respectively.

3.3. Confidence Interval. Based on the probability density function of the wind speed prediction error, the confidence interval under a certain confidence probability is calculated to establish the reliability of the forecasted value of the wind speed point.

The error of wind speed forecasting is defined as the deviation between the actual measured value V_{meas} of the wind speed at a certain point in time and the predicted value V_{pred} , i.e.,

$$e = V_{\text{meas}} - V_{\text{pred}}. \tag{10}$$

The fluctuation of the prediction error due to the different wind speed values is quite different. Therefore, this paper divides the wind speed prediction into multiple intervals and establishes the wind speed prediction error distribution of different wind speed intervals.

The maximum number of zones is V_{max} and V_{min} , respectively. The partition D_i is

$$D_i = [V_{\text{min}} + (i-1)\Delta V, V_{\text{max}} + i\Delta V], \quad i = 1, 2, \dots, n. \tag{11}$$

The aforementioned interval sample points may be insufficient to reflect the actual distribution of the prediction errors. So, a second division is needed, i.e., the division of a part of the distribution of adjacent samples and fewer points together, until the merger of the new interval after the number of sample points meets the requirements.

For a wind speed range of the prediction error samples, the probability density function is as follows:

$$f(e) = \frac{1}{Nh} \sum_{i=1}^1 K\left(\frac{e - e_i}{h}\right), \tag{12}$$

where e_i is the wind speed prediction error sample.

The probability density function of the prediction error is obtained by kernel density estimation and curve fitting, and then the cumulative probability distribution function is obtained by integral. Assuming that the cumulative probability distribution function of the prediction error is $F(\xi)$ (ξ is a random variable of the prediction error), the confidence interval of the true value of the wind power satisfying a confidence probability of $1 - \alpha$ is as follows:

$$[V_{\text{pred}} + \hat{F}(\alpha_1), V_{\text{pred}} + \hat{F}(\alpha_2)], \quad \alpha_2 - \alpha_1 = 1 - \alpha, \tag{13}$$

where $\hat{F}(\cdot)$ is the inverse of the cumulative probability distribution function $\Pr\{\xi \leq \hat{F}(q)\} = q$. In this paper, we take a symmetric probability interval, that is, $\alpha_1 = (\alpha/2)$ and $\alpha_2 = 1 - (\alpha/2)$.

According to the wind speed prediction error probability density curve and speed point prediction value, the confidence interval can be obtained at a certain confidence probability; the specific steps are as follows:

Step 1: for a certain predicted value, determine the wind speed interval to which it belongs and then find the predicted error probability density curve corresponding to the wind speed prediction interval

Step 2: apply cubic spline interpolation to the curve fitting to find the $\alpha/2$ and $((1 - \alpha)/2)$ subpoints of the wind speed prediction error

Step 3: use (13) to calculate the confidence interval of the wind speed forecasting

3.4. Simulation. In this section, in order to verify the effectiveness of the EMD-MKL model, we use two different hybrid models (EMD-MKL and EMD-SVM) to compare the results of multistep wind speed forecasting. Then, we implement 1-step to 3-step prediction for the 10-minute wind speed using the above two models. The CPU used in the experiment is Intel Core i7-6700k (8 m ache, 4.0 GHz), the memory is 64 G, and the operating system is Win 10 (64 bit).

First, we use EMD to decompose the selected wind speed data and decompose the complex wind speed signals into several subsequences with different frequency characteristics. We set $\delta = 0.3$ as the default value for EMD; there are 16 IMF subsequences and one residue in this experiment. The results of partial decomposition subsequence of wind speed data are shown in Figure 3.

Then, we use the SVM and MKL to train the eight subsequences. Through parameter optimization, the parameters of the SVM and MKL are set as follows:

SVM: cParam = 1 and gParam = 1.1953

MKL: cParam = 1 and gParam = [0.9, 0.925, 0.95, 0.975, 1, 1.025, 1.05, 1.075, 1.1]

Finally, the model reconstructs and combines the prediction results of subsequences to get the final prediction results. The prediction results of 1–3 steps of the two hybrid models are shown in Figures 4–6.

FIGURE 3: The decomposition of the 10-minute wind speed series.

FIGURE 4: Results of 1-step wind speed prediction by EMD-SVM and EMD-MKL.

3.5. Comparisons. From these figures, it can be seen that in multistep prediction, compared with EMD-SVM, EMD-MKL has better results regarding the phase shift and amplitude fluctuation and has better performance in following trend change and predicting mutation points.

In Figures 7–11, we compare the recent state-of-the-art methods, including EMD-SVM, EMD-RBF (radial basis function), EMD-ELMAN (Elman neural network), EMD-

GMDH (group method of data handling), and EMD-MAD (multiple attribute decision) [33]. As observed from these figures, it can be seen that our proposed method is more competitive. We can see that EMD-MKL has worse results for 1-step forecasting but has high precision for 2-step and 3-step forecasting.

When comparing these models, it is seen that the performance of EMD-MKL is better than that of EMD-SVM.

FIGURE 5: Results of 2-step wind speed prediction by EMD-SVM and EMD-MKL.

FIGURE 6: Results of 3-step wind speed prediction by EMD-SVM and EMD-MKL.

FIGURE 7: MAE of 1-step to 3-step prediction results.

FIGURE 8: RMSE of 1-step to 3-step prediction results.

FIGURE 9: MSPE of 1-step to 3-step prediction results.

FIGURE 10: MAPE of 1-step to 3-step prediction results.

FIGURE 11: MaxAPE of 1-step to 3-step prediction results.

TABLE 1: Error indexes of 1-step to 3-step prediction results for CKS1.

Indexes	EMD-MKL			EMD-SVM		
	1-step prediction	2-step prediction	3-step prediction	1-step prediction	2-step prediction	3-step prediction
MAE (m/s)	0.4243	0.5557	0.6603	0.4580	0.6255	0.7681
RMSE (m/s)	0.5890	0.7463	0.8990	0.6031	0.8226	1.0233
MAPE	0.0636	0.0882	0.1046	0.0716	0.1004	0.1213
MSPE	0.0953	0.1339	0.2121	0.1002	0.1640	0.1892
MaxAPE	0.9507	1.4919	3.6608	0.7242	2.2323	2.4682

TABLE 2: Error indexes of 1-step to 3-step prediction results for STH3.

Indexes	EMD-MKL			EMD-SVM		
	1-step prediction	2-step prediction	3-step prediction	1-step prediction	2-step prediction	3-step prediction
MAE (m/s)	0.3078	0.4780	0.5633	0.3232	0.4980	0.6091
RMSE (m/s)	0.3863	0.5838	0.7206	0.4012	0.6252	0.8141
MAPE	0.0434	0.0641	0.0753	0.0445	0.0667	0.0841
MSPE	0.0646	0.0865	0.1039	0.0641	0.0915	0.1289
MaxAPE	0.4968	0.6480	0.5146	0.5176	0.6536	0.8827

TABLE 3: Error indexes of 1-step to 3-step prediction results for CKS2.

Indexes	EMD-MKL			EMD-SVM		
	1-step prediction	2-step prediction	3-step prediction	1-step prediction	2-step prediction	3-step prediction
MAE (m/s)	0.3714	0.5092	0.5882	0.4040	0.5551	0.6655
RMSE (m/s)	0.4621	0.6241	0.7392	0.4941	0.6839	0.8235
MAPE	2.3258	3.2074	3.7027	2.5151	3.4656	4.1333
MSPE	0.0233	0.0321	0.0370	0.0252	0.0347	0.0413
MaxAPE	0.1238	0.1708	0.1874	0.1092	0.1369	0.2234

Compared with EMD-SVM, EMD-MKL improved the RMSE index of 2-step and 3-step forecasting by 3.73% and 10.33%, respectively. EMD-MKL improved the MAE index of 2-step and 3-step forecasting by 5.44% and 11.56%, respectively. Thus, we can benefit from the use of multiple kernels, which enhance the model fitting capacity while keeping the

generalization ability with the max margin operation. The reason for these improvements is that for highly complex wind speed signals, the multiple kernel learning method can adaptively assign the weights of multiple prediction functions, which extends those of conventional wind speed forecasting methods using support vector machines.

TABLE 4: Error indexes of 1-step to 3-step prediction results for MWT2.

Indexes	EMD-MKL			EMD-SVM		
	1-step prediction	2-step prediction	3-step prediction	1-step prediction	2-step prediction	3-step prediction
MAE (m/s)	0.2524	0.3499	0.3905	0.2804	0.3728	0.4230
RMSE (m/s)	0.3318	0.4384	0.4815	0.3646	0.4664	0.5328
MAPE	0.0471	0.0637	0.0755	0.0532	0.0697	0.0835
MSPE	0.0694	0.0857	0.1144	0.0823	0.1002	0.1328
MaxAPE	0.5032	0.5570	1.0654	0.8320	0.6738	1.4598

FIGURE 12: Forecast interval estimation for EMD-SVM.

FIGURE 13: Forecast interval estimation for EMD-MKL.

TABLE 5: Multistep forecasting results with confidence intervals.

	EMD-SVM			EMD-MKL		
	1-step prediction	2-step prediction	3-step prediction	1-step prediction	2-step prediction	3-step prediction
$\hat{F}(\alpha_1)$	-2.8694	-2.8616	-3.5215	-1.2349	-1.2083	-1.3172
	-2.9508	-4.0046	-4.4386	-1.6884	-2.0611	-2.0446
	-1.8067	-2.8289	-3.6129	-1.4967	-1.9216	-2.4128
	-2.2733	-3.0428	-3.8066	-1.7741	-2.0374	-2.1500
	-2.7841	-2.5692	-3.1499	-1.4236	-1.6013	-1.8308
	-2.8305	-2.9841	-3.4719	-1.3805	-1.4735	-1.4825
	-1.3578	-1.8935	-2.3102	-1.0058	-0.9542	-1.2542
$\hat{F}(\alpha_2)$	0.4018	0.8493	0.9181	0.3781	0.6035	0.6765
	0.7219	1.2186	1.9246	0.6073	0.7331	0.9268
	2.6543	2.7081	3.0932	0.9786	1.0913	1.1415
	3.1104	2.9839	3.4246	1.3237	1.2719	1.4514
	1.9904	2.8054	3.3526	1.3216	1.3711	1.6346
	2.8094	2.8293	3.3477	1.4083	1.5131	1.7677
	2.8896	3.3032	4.2679	1.3966	1.9544	2.1963

TABLE 6: Multistep forecasting results with confidence intervals.

	EMD-SVM			EMD-MKL		
	1-step prediction	2-step prediction	3-step prediction	1-step prediction	2-step prediction	3-step prediction
$\hat{F}(\alpha_2) - \hat{F}(\alpha_1)$	3.2712	3.7109	4.4396	1.613	1.8118	1.9937
	3.6727	5.2232	6.3632	2.2957	2.7942	2.9714
	4.4610	5.537	6.7061	2.4753	3.0129	3.5543
	5.3837	6.0267	7.2312	3.0978	3.3093	3.6014
	4.7745	5.3746	6.5025	2.7452	2.9724	3.4654
	5.6399	5.8134	6.8196	2.7888	2.9866	3.2502
	4.2474	5.1967	6.5781	2.4024	2.9086	3.4505

3.6. *More Results.* To further illustrate the advantages of the proposed models, this study provides four different wind speed series CKS1, STH3, CKS2, and MWT2. The above examples of the forecasted results are listed in Tables 1–4. From these tables, the same conclusion can be drawn as in Section 3.5.

3.7. *Result Analysis with Confidence Intervals.* Using equations (10)–(13), the confidence intervals can be calculated. The results are shown in Figures 12 and 13. As can be seen, most of the truth values fall within the confidence intervals and only individual points are outside the interval, indicating that the confidence of most of the predictions is 90%. The larger the confidence level, the longer the confidence interval. The envelope of the confidence interval can reflect the range of the true value of the short-term wind speed from the probability point of view and better reflect the wind speed.

The training values are divided into seven intervals; each prediction value V_{pred} belongs to one of these seven intervals. The predicted values in the same interval have the same $\hat{F}(\alpha_1)$ and $\hat{F}(\alpha_2)$; thus, the length of the confidence intervals of the predicted values in the same interval is equal, that is, the value of $\hat{F}(\alpha_2) - \hat{F}(\alpha_1)$ is equal.

From the confidence interval, Tables 5 and 6 show that the confidence interval distance of the predicted value of the

MKL method is smaller than the confidence interval distance of the predicted value of the SVM method when we have a confidence level of 90%. The range of the obtained MKL is smaller; we know that if the confidence interval is smaller, then the accuracy of the forecast is higher. So, the MKL method has better accuracy than the SVM method.

4. Conclusion

This paper presented a hybrid short-term wind speed prediction method using EMD and MKL. For complex wind speed signals, multiple kernel learning methods can adaptively allocate the weights of multiple prediction functions, while EMD generates more subsequences with different frequency characteristics and enhances the stationarity of signal features, using multichannel signals to increase the robustness of the signal features. Through multistep wind speed prediction, it was demonstrated that, compared with EMD-SVM, EMD-MKL has better results in terms of phase offset and amplitude fluctuation and has better performance in following trend changes and forecasting mutation points, having a higher prediction accuracy and stability. The results showed that the prediction accuracy of 10-minute wind speed of EMD-MKL under different error criteria was higher than that of EMD-SVM, and the EMD-MKL method also showed high stability. Therefore, the EMD-MKL model has high potential for wind speed prediction.

This paper selected data from several wind farms in New Zealand for testing and verified that EMD-MKL can generalize well. In addition, based on the probability density curve of wind speed prediction error and the prediction value of mutation points, we conclude that the fluctuation width of EMD-MKL in the confidence interval is small and has good prediction performance. Therefore, the EMD-MKL hybrid model has great potential in wind speed prediction. In recent years, with the extensive application of deep learning in various fields, we expect to combine deep learning with wind power forecasting to produce more research results.

Data Availability

The data used to support the findings of this study have been deposited in the generating synthetic wind data of NIWA (<https://www.niwa.co.nz/environmental-information/research-projects/synthetic-wind-data>).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Key R&D Program of China (2017YFA60700602 and 2018YFC0809200) and the Ministry of Industry and Information Technology of China under project “Industrial Internet Platform Test Bed for Optimizing the Operation of Motor and Driving Equipment” (2018-06-46).

References

- [1] Y. Jiang, X. Chen, K. Yu, and Y. Liao, “Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm,” *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 1, pp. 126–133, 2017.
- [2] J. Hu, J. Wang, and G. Zeng, “A hybrid forecasting approach applied to wind speed time series,” *Renewable Energy*, vol. 60, pp. 185–194, 2013.
- [3] Z. Qu, K. W. Mao, and Z. W. LiZhang, “Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network,” *Renewable Energy*, vol. 133, pp. 919–929, 2019.
- [4] WWEA—World Wind Energy Association, *Wind Power Capacity Worldwide Reaches 597 GW, 50,1 GW Added in 2018*, World Wind Energy Association, Bonn, Germany, 2019, <http://www.wwindea.org>.
- [5] J. Wang and J. Hu, “A robust combination approach for short-term wind speed forecasting and analysis - combination of the ARIMA (autoregressive integrated moving average), ELM (extreme learning machine), SVM (support vector machine) and LSSVM (least square SVM) forecasts using a GPR (Gaussian process regression) model,” *Energy*, vol. 93, no. 1, pp. 41–56, 2015.
- [6] G. Santamaría-Bonfil, A. Reyes-Ballesteros, and C. Gershenson, “Wind speed forecasting for wind farms: a method based on support vector regression,” *Renewable Energy*, vol. 85, pp. 790–809, 2016.
- [7] A. Heydari, D. Astiaso Garcia, F. Keynia, F. Bisegna, and L. De Santoli, “A novel composite neural network based method for wind and solar power forecasting in microgrids,” *Applied Energy*, vol. 251, Article ID 113353, 2019.
- [8] M. Majidi Nezhad, A. Heydari, D. Groppi, F. Cumo, and D. Astiaso Garcia, “Wind source potential assessment using sentinel 1 satellite and a new forecasting model based on machine learning: a case study Sardinia islands,” *Renewable Energy*, vol. 155, pp. 212–224, 2020.
- [9] H. Liu and C. Chen, “Data processing strategies in wind energy forecasting models and applications: a comprehensive review,” *Applied Energy*, vol. 249, pp. 392–408, 2019.
- [10] Y. Wang, J. Wang, and X. Wei, “A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: a case study of wind farms in northwest China,” *Energy*, vol. 91, pp. 556–572, 2015.
- [11] Y. Zhu and W. X. Zheng, “Multiple lyapunov functions analysis approach for discrete-time-switched piecewise-affine systems under dwell-time constraints,” *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 2177–2184, 2020.
- [12] Y. Xu, W. Yan, G. Yang, J. Luo, T. Li, and J. He, “CenterFace: joint face detection and alignment using face as point,” *Scientific Programming*, vol. 2020, Article ID 7845384, 8 pages, 2020.
- [13] Y. Zhu, W. X. Zheng, and D. Zhou, “Quasi-synchronization of discrete-time lur’e-type switched systems with parameter mismatches and relaxed PDT constraints,” *IEEE Transactions on Cybernetics*, vol. 50, no. 5, pp. 2026–2037, 2020.
- [14] Y. Zhu and W. X. Zheng, “Observer-based control for cyber-physical systems with periodic DoS attacks via a cyclic switching strategy,” *IEEE Transactions on Automatic Control*, vol. 65, no. 8, pp. 3714–3721, 2020.
- [15] C. Wang, H. Zhang, W. Fan, and P. Ma, “A new chaotic time series hybrid prediction method of wind power based on EEMD-SE and full-parameters continued fraction,” *Energy*, vol. 138, p. 977, 2017.
- [16] S. Samadianfard, K. S. Hashemi, A. M. MostafaeipourIzadyar, N. Nabipour, and S. Shamshirband, “Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm,” *Energy Reports*, vol. 6, pp. 1147–1159, 2020.
- [17] R. G. Kavasseri and K. Seetharaman, “Day-ahead wind speed forecasting using f-ARIMA models,” *Renewable Energy*, vol. 34, no. 5, pp. 1388–1393, 2009.
- [18] J. L. Torres, A. Garcia, M. De Blas, and A. De Francisco, “Forecast of hourly average wind speed with ARMA models in Navarre (Spain),” *Solar Energy*, vol. 79, no. 1, pp. 65–77, 2005.
- [19] Y.-h. Liao, S.-p. Su, J.-s. Kang, L. Li, F. Fan, and Y.-m. Su, “Wind speed forecasting based on improved exponential smoothing method and Markov model,” *Journal of Electric Power and Technology*, vol. 31, no. 1, pp. 85–89, 2016, in Chinese.
- [20] C.-Y. Huang, Y.-W. Liu, W.-C. Tzeng, and P.-Y. Wang, “Short term wind speed predictions by using the grey prediction model based forecast method,” in *Proceedings of the 2011 IEEE Green Technologies Conference (IEEE-Green)*, Baton Rouge, LA, USA, 2011.
- [21] L. S. Johnson, S. R. Eddy, and E. Portugaly, “Hidden Markov model speed heuristic and iterative HMM search procedure,” *Bmc Bioinformatics*, vol. 11, no. 1, p. 431, 2010.
- [22] L. Kamal and Y. Z. Jafri, “Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan,” *Solar Energy*, vol. 61, no. 1, pp. 23–32, 1997.

- [23] A. Malmberg, U. Holst, and J. Holst, "Forecasting near-surface ocean winds with Kalman filter techniques," *Ocean Engineering*, vol. 32, no. 3-4, pp. 273-291, 2005.
- [24] P. Louka, G. Galanis, N. Siebert et al., "Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 12, pp. 2348-2362, 2008.
- [25] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renewable Energy*, vol. 29, no. 6, pp. 939-947, 2004.
- [26] M. A. Mohandes, S. Rehman, and T. O. Halawani, "A neural networks approach for wind speed prediction," *Renewable Energy*, vol. 13, no. 3, pp. 345-354, 1998.
- [27] P. Flores, A. Tapia, and G. Tapia, "Application of a control algorithm for wind speed prediction and active power generation," *Renewable Energy*, vol. 30, no. 4, pp. 523-536, 2005.
- [28] T. G. Barbounis and J. B. Theocharis, "A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation," *Neurocomputing*, vol. 70, no. 7-9, pp. 1525-1542, 2007.
- [29] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Economics*, vol. 30, no. 5, pp. 2623-2635, 2008.
- [30] M. Dätig and T. Schlurmann, "Performance and limitations of the Hilbert-Huang transformation (HHT) with an application to irregular water waves," *Ocean Engineering*, vol. 31, no. 14-15, pp. 1783-1834, 2004.
- [31] Y. Chen, C. R. Chen, and P. Tan, "Short-term wind speeds prediction of SVM based on simulated annealing algorithm with Gauss perturbation," *IOP Conference Series: Earth and Environmental Science*, vol. 267, Article ID 042032, 2019.
- [32] L. Yang, J. M. He, and V. Vittal, "Support-vector-machine-enhanced Markov model for short-term wind power forecast," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 791-799, 2015.
- [33] F. Wang, *Study on the Theory and Method of Short-Term Wind Speed Prediction and Simulation in Wind Farm*, North China Electric Power University, Beijing, China, 2013.
- [34] J. Yang, "A novel short-term multi-input-multi-output prediction model of wind speed and wind power with LSSVM based on improved ant colony algorithm optimization," *Cluster Computing*, vol. 22, no. S2, pp. 3293-3300, 2019.
- [35] G. R. G. Lanckriet, N. Cristianini, P. L. Bartlett et al., "Learning the kernel matrix with semidefinite programming," *Journal of Machine Learning Research*, vol. 5, no. 1, pp. 27-72, 2002.
- [36] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903-995, 1998.
- [37] J. Wang, W. Zhang, Y. Li, J. Wang, and Z. Dang, "Forecasting wind speed using empirical mode decomposition and Elman neural network," *Applied Soft Computing*, vol. 23, pp. 452-459, 2014.
- [38] Z. Guo, W. Zhao, H. Lu, and J. Wang, "Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model," *Renewable Energy*, vol. 37, no. 1, pp. 241-249, 2012.
- [39] H. Liu, H. Tian, X. Liang, and Y. Li, "New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, mind evolutionary algorithm and artificial neural networks," *Renewable Energy*, vol. 83, pp. 1066-1075, 2015.
- [40] W. M. Czarnecki and J. Tabor, "Two ellipsoid support vector machines," *Expert Systems with Applications*, vol. 41, no. 18, pp. 8211-8224, 2014.
- [41] C. M. Wilson, K. Li, X. Yu, P.-F. Kuan, and X. Wang, "Multiple-kernel learning for genomic data mining and prediction," *BMC Bioinformatics*, vol. 20, no. 1, 2019.
- [42] LIBSVM—A Library for Support Vector Machines, <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
- [43] R. Turner, X. Zheng, N. Gordon et al., "Multi-year ten-minute synthetic wind speed time-series for 15 actual or proposed New Zealand wind farm," NWA report no. WLG2009-43, National Institute of Water and Atmospheric Research, Auckland, New Zealand, 2009.
- [44] H. Liu, H.-Q. Tian, X.-F. Liang, and Y.-F. Li, "Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks," *Applied Energy*, vol. 157, pp. 183-194, 2015.
- [45] Y. Xu, G. Yang, and T. Yao, "An EMD-SVM model with error compensation for short-term wind speed forecasting," *International Journal of Information Technology and Management*, vol. 18, no. 2-3, pp. 171-181, 2019.
- [46] W. Zhang, F. Liu, X. Zheng, and Y. Li, "A hybrid EMD-SVM based short-term wind power forecasting model," in *Proceedings of the 2016 IEEE PES Asia-Pacific Power & Energy Engineering Conference*, Xi'an, China, 2016.