Short-term wind speed forecasting based on signal decomposition algorithm

Xu Duan[1]

Shanghai Jiao Tong University

(Dated: Sep 8th 2021)

This paper is mainly about wind speed prediction. To reduce the randomness in origin wind speed data. a variety of signal decomposition techniques are applied. Including EMD (Empirical Mode Decomposition), EEMD (Ensemble Empirical Mode Decomposition), CEEMD (Complementary Ensemble Mode Decomposition), CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise), ICEEMDAN (Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise), VMD (Variation Mode Decomposition) and EWT (Empirical Wavelet Transform). After decomposition, different modes are fed to LSTM model respectively. Prediction results are used to reconstruct wind speed. These techniques can greatly improve wind prediction performance.

I. KNOWLEDGE BACKGROUND

A. Introduction

With the rapid growth of today's energy consumption, excessive consumption of traditional resources and the pollution of the environment problem increasingly serious. To deal with global energy problem, the development of new energy has become a consensus around the world. Wind power as a kind of widely distributed and abundant renewable energy, is gaining more and more attention in the world gradually. The proportion of wind power in power system is also increasing around the world. But the wind is uncertain and non-stationary, which implies incorporating wind farm may have serious impact on the stability of power system and its operation. Accurate prediction of wind speed, can not only reduce the cost of power network operation[1,2], but also is beneficial to the safety of wind power grid.

B. Literature review

Wind speed starts from physical prediction, which is similar to numerical weather prediction system. It gets humidity, pressure to predict the wind speed of the following day or more days. However, considering its prediction accuracy, its scope of application is limited.[3]

Then wind speed prediction turns to statistical prediction models which are chiefly reflected in linear models, such as AR (Auto Regressive), MA (Moving Average), ARMA (Auto Regressive Moving Average) and ARIMA (Auto Regressive Integrated Moving Average). However, due to the randomness and uncertainty of wind speed, the accuracy of traditional statistical prediction models is often not up to the required height.

With the great popularity of artificial intelligence , models such as ANN (Artificial Neural Network) and SVM (Support Vector Machine) has been widely accepted and adopted in dealing with time series prediction. The insufficiency of these models has greatly constrained their application in weed speed prediction. For example SVM is incapable in handling large amount of data.

In recent years, deep learning has been widely applied in various fields and has obvious advantages, which has won the attention and affection of many scholars. In particular, RNNs with the capability of maintaining states between different inputs show advantages in handling time sequences, but suffers from problems like vanishing and exploding gradients. LSTM, a variant of RNN (Recursive Neural Betwork), solves the problem by introducing controlling gates to select main useful information and abandon useless information in the process of propagating. LSTM has obvious advantages over other prediction models in time series prediction. traditional learning methods in many fields such as handwriting recognition[4] and stock analysis[5].

To tackle with nonlinear and non-stationary time series, Huang[6] proposed the EMD (Empirical Mode Decomposition) aiming to turn original non-stationary time series into a set of stationary time series. Later proposed EEMD (Ensemble Empirical Mode Decomposition) was proposed to deal with "mode mixing" problem of EMD[7]. To eliminate the error caused by random noise, Huang et al. revised EEMD and proposed CEEMD[8]. To ensure the completeness of signal decomposition, a novel CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) was proposed, which applies a iterative method[9]. ICEEMDAN (Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) was proposed to avoid the overlapping in the scales in the first few modes in CEEMDAN[10]. EMD and EMD-based methods are widely used today to recursively decompose a signal into different modes of unknown but separate spectral bands. Chen et al. utilize CEEMDAN and BP neural network to predict short-term wind speed[11].

However, EMD and EMD-based methods is known for limitations like sensitivity to noise and sampling. K. Dragomiretskiy and D. Zosso proposed an entirely non-recursive VMD (Variational Mode Decomposition) model, where the modes are extracted concurrently, showing promising practical decomposition results[12].

Another novel method called EWT (Empirical Wavelet Transform) also show its usefulness compared to the classic EMD. EWT is a new approach to build adaptive wavelets to make prediction.[13]

C. Main contribution and chapter arrangement

The main contributions of this paper are the application of various signal decomposition techniques and compare their performance. Specifically, this paper creatively apply VMD and EWT to wind prediction and get good performance in contrast to other techniques.
The rest of this paper is arranged as follows: Section 2 mainly introduce the basic theory

of different signal decomposition algorithm, including EMD, EEMD, CEEMD, ICEEMD, VMD and EWT. Section 3 introduce the prediction model we applied in wind speed prediction. In section 4, we do experiment to compare the performance between different signal decomposition algorithm and make explanation. Finally, Section 5 gieves the conclusion of this study.

II. SIGNAL DECOMPOSITION ALGORITHM

Signal decomposition techniques are able to decompose original wind speed data into different modes with different features, thus reducing the randomness. Signals with less randomness and high stabilization can improve the performance of subsequent models.

A. EMD

EMD (Empirical Mode Decomposition) is an adaptive method to analyze non-stationary signals. It produces a local and fully data-driven separation of a signal in fast and slow oscillations. At the end, the original signal can be expressed as a sum of amplitude and frequency modulated (AM-FM) functions called Intrinsic Mode Functions (IMFs) and residual.

To be considered as an IMF, a signal must fulfill two conditions: (i) the number of extrema (maxima and minima) and the number of zero-crossings must be equal or differ at most by one; and (ii) the local mean, defined as the mean of upper and lower envelopes, must be zero.

Let x be the original signal. The algorithm can be described as follows:

- Step 1.Set $k = 0$ and find all extrema of $r_0 = x$.
- Step 2.Interpolate between minima (maxima) of r_k to obtain the lower (upper) envelope e_{min} (e_{max}).
- Step 3.Compute the IMF candidate $d_{k+1} = r_k - m$.
- Step 4.Is d_{k+1} an IMF?
  - Yes. Save d_{k+1}, compute the residue $r_{k+1} = x - \Sigma_{i=1}^{k} d_i$, do $k = k + 1$, and treat r_k as input data in step 2.

- No. Treat d_{k+1} as input data in step 2.

- Step 6.Continue until the final residue r_K satisfies some predefined stopping criterion.

The local nature of the EMD may produce oscillations with very disparate scales in one mode, or oscillations with similar scales in different modes. When this phenomenon is undesirable, and similar scales for each mode are preferred, this consequence of the method becomes a problem, named as ”mode mixing”.

B. EEMD

EEMD (Ensemble Empirical Mode Decomposition) is the ensemble version of EMD, IMFs are obtained from an ensemble of the original signal plus different realizations of finite variance white noise. By populating whole time-frequency space, EEMD reduces the mode mixing. Thus, modes with similar scales are obtained. Let x be the original signal. The algorithm can be described as follows:

- Step 1.Generate $x^{(i)} = x + \beta w^{(i)}$, where $w^{(i)}$ $(i = 1, ..., I)$ is a standard normal distribution white noise. $\beta > 0$ is a hyperparameter to adjust the variance of the white noise.

- Step 2.Decompose completely each $x^{(i)}$ $(i = 1, ..., I)$ by EMD, obtaining the modes $d_k^{(i)}$, where $k = 1, ..., K$ indicates the mode.

- Step 3.Assign $\bar{d}_k = \frac{1}{I}\Sigma_{i=1}^{I} d_k^{(i)}$.

It can be noticed that in EEMD, every $x^{(i)}$ is composed independently from other realizations and for every one of them a residue $r_k^{(i)} =_{k-1}^{(i)} -d_k^{(i)}$ is obtained at each stage, with no connection between the different realizations. This situation is the cause of some EEMD disadvantages: (i) Due to the randomness of white Gaussian noise, there will be reconstruction error. (ii) the decomposition is not complete and (iii) different realizations of signal plus noise might produce different number of modes.

C. CEEMD

CEEMD (Complementary Empirical Mode Decomposition) is proposed in order to deal with the reconstruction error. Noise is added in pairs to the original data (one positive and one negative) to generate two ensembles. However, there is still no guarantee that original signal plus white noise produce the same number of modes and the completeness in mode decomposition.

D. CEEMDAN

CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) is proposed aiming to deal with the drawbacks of CEEMD. CEEMDAN apply iterative method to complete signal decomposition. The general idea is as following: First, average the first mode $\bar{d}_1$ from $x^{(i)} = x + \beta w^{(i)}$, which is exactly the same as EEMD. Then, the first residue is obtained:$r_1 = x - \bar{d}_1$. The second mode is obtained from averaging the first EMD mode from the first residue plus first mode of white noise. The procedure continues until a stopping criterion is reached.
Let x be the original signal. And denote $E_k(\cdot)$ be the operator which produces the kth mode obtained by EMD. The algorithm of CEEMDAN is as following:

- Step 1.For every $i = 1, ..., I$ decompose each $x^{(i)} = x + \beta_0 w^{(i)}$ by EMD, obtaining the first mode of each $x^{(i)}$ and average them to get the first mode of original signal
$$\bar{d}_1 = \frac{1}{I}\Sigma_{i=1}^{I} E_1(x^{(i)}).$$
- Step 2.At the first stage ($k = 1$) calculate the first residue
$$r_1 = x - \bar{d}_1.$$
- Step 3.Obain the first mode of $r_1 + \beta_1 E_1(w^{(i)}),\ i = 1, ..., I,$ by EMD and define the second CEEMDAN mode as:
$$\bar{d}_2 = \frac{1}{I}\Sigma_{i=1}^{I} E_1(r_1 + \beta_1 E_1(w^{(i)})$$
- Step 4.For $k = 2, ..., K$ calculate the kth residue:
$$r_k = r_{(k-1)} - \bar{d}_k.$$
- Step 5.Obtain the first mode of $r_k + \beta_k E_K(w^{(i)}),\ i = 1, ..., I,$by EMD until define the $(K + 1)$th CEEMDAN mode as:
$$\bar{d}_{(k+1)} = \frac{1}{I}\Sigma_{i=1}^{I} E_1(r_k + \beta_k E_k(w^{(i)})).$$

- Step 6.Go to step 4 for the next k.

Iterate the steps 4 to 6 until the obtained residue cannot be further decomposed by EMD, either because it satisfies IMF conditions or because it has less than three local extrema. Observe that, by construction of CEEMDAN, the final residue satisfies:

$$x = \Sigma_{k=1}^{K} \bar{d}_k + r_K$$

ensuring the completeness property of the proposed decomposition and thus providing an exact reconstruction of the original data. Despite this, CEEMDAN still has some aspects in which it deserves to be improved:

- Recall the operator $E_K(\cdot)$ and define new operator $M(\cdot)$ which produces the local mean of the signal that is applied to. It can be noticed that $E_1(x) = x - M(x)$. Let $< \cdot >$ be the action of averaging throughout the realizations. For the first EEMD and original CEEMDAN modes we have:

  $$\bar{d}_1 =< E_1(x^{(i)}) >=< x^{(i)} - M(x^{(i)}) >=< x^{(i)} > - < M(x^{(i)}) >$$

  The difference between $< x^{(i)} >$ and original signal contributes to the amount of noise. So if we define:

  $$\bar{d}_1 = x- < M(x^{(i)}) >$$

  In this way, the amount of noise present in the modes are reduced.

- There is a strong overlapping in the scales in the first few modes (Take first two modes for example. The first mode are extracted adding white noise and the second one adding $E_1(w^{(i)})$)

E. ICEEMDAN

Taking into account the two previous aspects, here we propose ICEEMDAN. Recall the operator $E_K(\cdot)$ and $M(\cdot)$ and action $< \cdot >$. The algorithm is as following:

- Step 1. Calculate by EMD the local means of $x^{(i)} = x + \beta_0 E_1(w^{(i)})$ to obtain the first residue

  $$r_1 =< M(x^{(i)}) >$$

- Step 2. At the first stage ($k = 1$) calculate the first mode:

$$\bar{d}_1 = x - r_1$$

- Step 3. Estimate the second residue as the average of local means of $r_1 + \beta_1 E_2(w^{(i)})$ and define the second mode:

$$\bar{d}_2 = r_1 - r_2 = r_1 - < M(r_1 + \beta_1 E_2(w^{(i)})) >$$

- Step 4. For $k = 3, ..., K$ calculate the kth residue

$$r_k = < M(r_{k-1} + \beta_{k-1} E_k(w^{(i)})) >$$

- Step 5. Compute the kth mode

$$\bar{d}_k = r_{k-1} - r_k$$

- Step 6. Go to step 4 for next k.

F. VMD

There is a slightly different definition of IMF in VMD. IMF are amplitude-modulated-frequency-modulated (AM-FM) signals, written as:

$$u_K(t) = A_k(t)\cos(\phi_k(t)),$$

where the phase $\phi_k(t)$ is a non decreasing function, $\phi_k'(t) \geq 0$, the envelope is non-negative $A_k(t) \geq 0$, and both the envelope $A_k(t) \geq 0$ and the instantaneous frequency $\omega_k(t) := \phi_k'(t) \geq 0$ vary much slower than the phase $\phi_k(t)$.

The goal of VMD is to decompose a real valued input signal f into a discrete number of IMFs. And each IMF should be mostly compact around a center pulsation ω_k.

In order to assess the bandwidth of a mode, we propose the following scheme: 1) for each mode u_k, compute the associate analytic signal by means of Hilbert transform in order to obtain a unilateral frequency spectrum. 2) for each mode, shift the mode's frequency spectrum to "baseband", by mixing with an exponential tuned to the respective estimated center frequency. 3) The bandwidth is now estimated through the H^1 Gaussian smoothness of the demodulated signal, i.e. the squared L^2-norm of the gradient. The result constrained variational problem is the following:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} u_k = f$$

where f is the original signal. $u_k := u_1, ..., u_K$ and $\omega_k := \omega_1, ..., \omega_K$ are notations for the set of all modes and their center frequencies, respectively.

We make use of both a quadratic term α and Lagrangian multipliers λ in order to render the problem unconstrained. Therefore, we introduce the augmented Lagrangian $\mathcal{L}$ as follows:

$$\mathcal{L}\left(\{u_k\},\{\omega_k\},\lambda\right) := \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle.$$

The solution to the original minimization problem is now found as the saddle point of the augmented Lagrangian $\mathcal{L}$ in a sequence of iterative sub-optimizations called alternate direction method of multipliers (ADMM). In the process of finding saddle point, u_k, ω_k, λ_k are updated.

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k(\omega) \right|^2 d\omega}$$

$$\lambda^{n+1} \leftarrow \lambda^n + \tau \left(f - \sum_k u_k^{n+1} \right)$$

until convergence:$\Sigma_k \|u_k^{n+1} - u_K^n\|_2^2 / \|u_k^n\|_2^2 < \epsilon$

G. EWT

EWT combines the strength of wavelet's formalism with the adaptability of EMD. The main idea is to extract the different modes of a signal by designing an appropriate wavelet filter bank. This construction leads us to a new wavelet transform, called the empirical wavelet transform. From the Fourier point of view, decomposition is equivalent to building a set of bandpass filters. Assume that the Fourier support $[0, \pi]$ is segmented into N contiguous segments. We denote ω_n to be the limits between each segments (where $\omega_0 = 0$ and $\omega_N = \pi$). Each segment is denoted $\Lambda_n = [\omega_{n-1}, \omega_n]$, then it is easy to see that $\bigcup_{n=1}^{N} = [0, \pi]$.Centered around each ω_n, we define a transition phase of width $2\tau_n$.

The empirical wavelets are defined as bandpass filters on each Λ_n. Utilizing the idea used in the construction of both Littlewood-Paley and Meyer's wavelets, we define the empirical scaling function and the empirical wavelets.

$$\hat{\phi}_n(\omega) = \begin{cases} 1 & \text{if } |\omega| \le (1-\gamma)\omega_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega| - (1-\gamma)\omega_n)\right)\right] & \text{if } (1-\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_n \\ 0 & \text{otherwise} \end{cases}$$

and

$$\hat{\psi}_n(\omega) = \begin{cases} 1 & \text{if } (1+\gamma)\omega_n \le |\omega| \le (1-\gamma)\omega_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_{n+1}}(|\omega| - (1-\gamma)\omega_{n+1})\right)\right] & \text{if } (1-\gamma)\omega_{n+1} \le |\omega| \le (1+\gamma)\omega_{n+1} \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega| - (1-\gamma)\omega_n)\right)\right] & \text{if } (1-\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_n \\ 0 & \text{otherwise} \end{cases}$$

where we have chosen τ_n proportional to $\omega_n : \tau_n = \gamma\omega_n$, where $0 < \gamma < 1$. And $\beta(x)$ is most used as

$$\beta(x) = x^4(35 - 84x + 70x^2 - 20x^3)$$

Then the detail coefficients and the approximation coefficients are defined as:

$$\mathcal{W}_f^{\mathcal{E}}(n, t) = \langle f, \psi_n \rangle = \int f(\tau)\overline{\psi_n(\tau - t)}d\tau$$

$$\mathcal{W}_f^{\mathcal{E}}(0, t) = \langle f, \phi_1 \rangle = \int f(\tau)\overline{\phi_1(\tau - t)}d\tau$$

The empirical mode f_k is given by

$$f_0(t) = \mathcal{W}_f^{\mathcal{E}}(0, t) * \phi_1(t)$$

$$f_k(t) = \mathcal{W}_f^{\mathcal{E}}(k, t) * \psi_k(t)$$

III. PREDICTION MODEL

The prediction model we choose is LSTM model. The LSTM (Long Short Term Memory) recurrent network is a special RNN (recurrent neural network) model. As discussed by (Bengio et al.), the RNN approach has the weakness of long-term dependency in practical applications. To overcome the major constraint of RNN model, (Hochreiter and Schmidhuber) developed a LSTM algorithm which is more suitable for processing and predicting important events with relatively long intervals and delays in time series. The network has been widely used in image text recognition, image processing, video data recognition and other fields. As shown in FIG. 1, the main structure of LSTM is a memory block with three

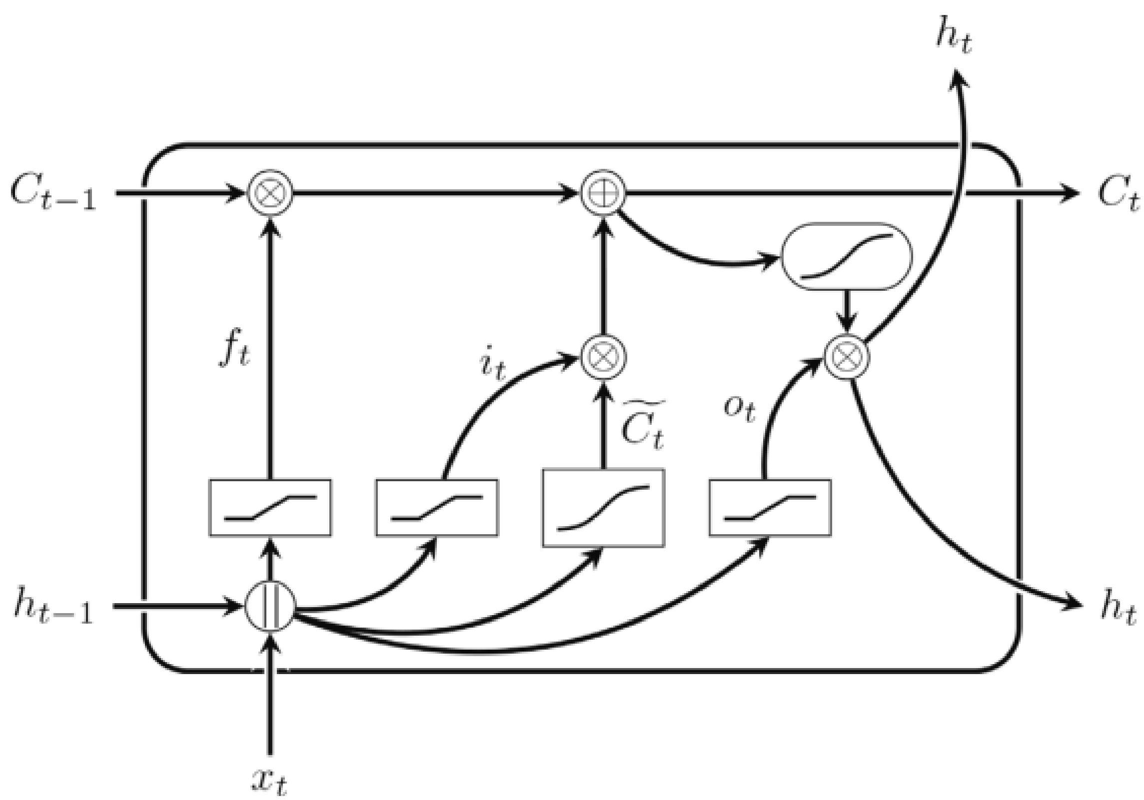

FIG. 1. EMD result

gates and a memory cell. Mathematically, the LSTM algorithm is composed of six core equations, governed by

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

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

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_c)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \cdot \tanh(C_t)$$

where $\cdot$ is matrix multiplication, $*$ is element-wise multiplication. $f_t, i_t, \tilde{C}_t, C_t, o_t$ and h_t are process functions. W_f, W_i, W_C and W_o are wights, b_f, b_i, b_c and b_o are bias. Here Eq.(4) denotes information selection stage or forgetting stage. This stage is mainly to selectively forget the input information from the previous moment. Useless information is filtered by this gate. Eq.(5) - Eq.(7) is the processing stage. Input gate selectively remembers the current input. Therefore current state is updated based on the combination of past and current useful information. Eq.(8)-Eq.(9) is the information output stages, the updated information will be outputted after the appropriate evaluation by output gate.

IV. EXPERIMENTS AND RESULT ANALYSIS

The process of our experiments are as following:

- Step 1. Fill Missing values of original wind speed.
- Step 2. Decompose wind speed using different signal decomposition algorithms.
- Step 3. Normalize each mode respectively.
- Step 4. For each mode, train a LSTM model and predict respectively.
- Step 5. Reconstruct wind speed.

A. Superiority of Signal Decomposition

In this subsection, we compare prediction performance between emd signal decomposition and no decomposition process. Here is the result of EMD.

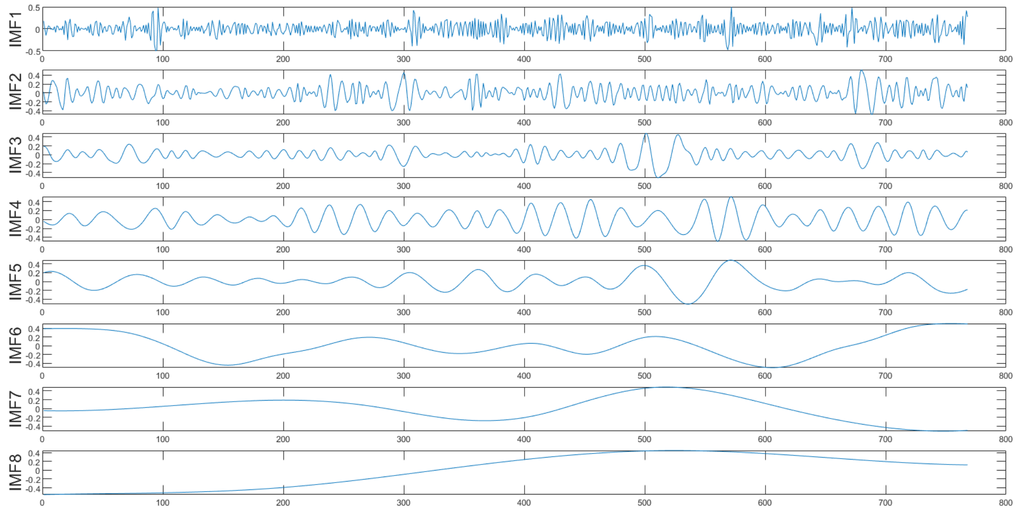

FIG. 2. EMD result

Here is the prediction result

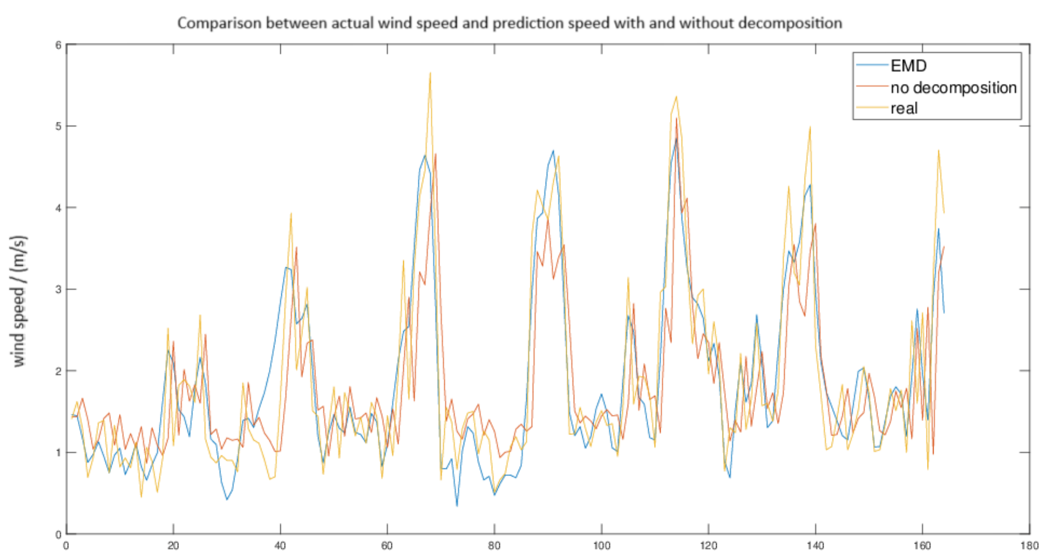

FIG. 3. Comparison between EMD and no signal decomposition

Here is the performance

	EMD	no signal decomposition
rmse	0.4772	0.8140
correlation coefficient	0.9109	0.7103

TABLE I. Compare between emd signal decomposition and no decomposition

It is obvious that EMD has greatly improved the stability of input wind speed signal, thus has better prediction performance.

B. Comparison Between Different Signal Decomposition Algorithms

Follow the steps above. We get the performance of different signal decomposition algorithms as follows:

	EMD	EEMD	CEEMD	CEEMDAD	ICEEMDAN	VMD	EWT
rmse	0.4772	0.4423	0.4213	0.4395	0.4376	0.3098	0.3388
correlation coefficient	0.9109	0.9242	0.9371	0.9267	0.9258	0.9634	0.9561

TABLE II. Comparison Between Different Signal Decomposition Algorithms

The following is the overall prediction performance:

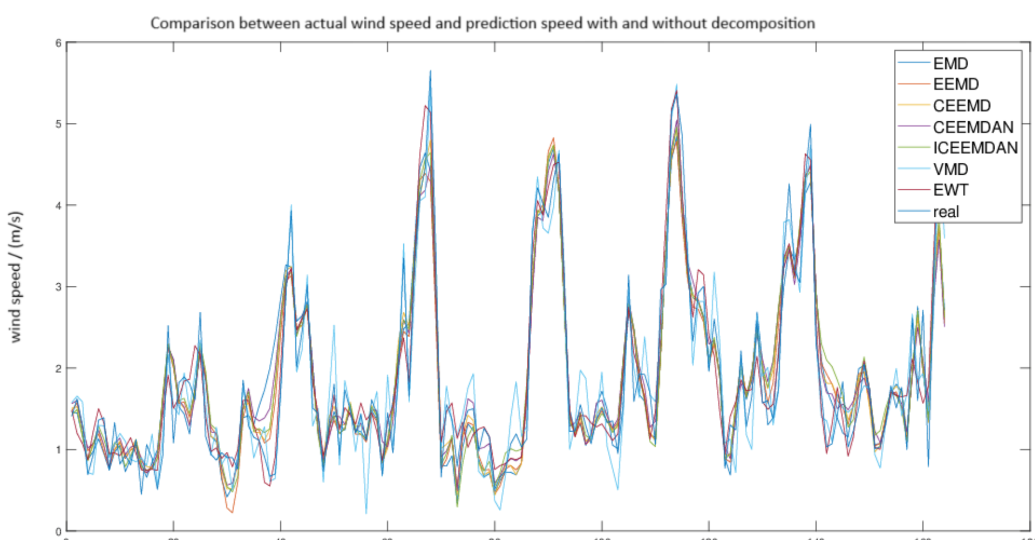

FIG. 4. Comparison Between Different Signal Decomposition Algorithms

The following is part of the previous graph:

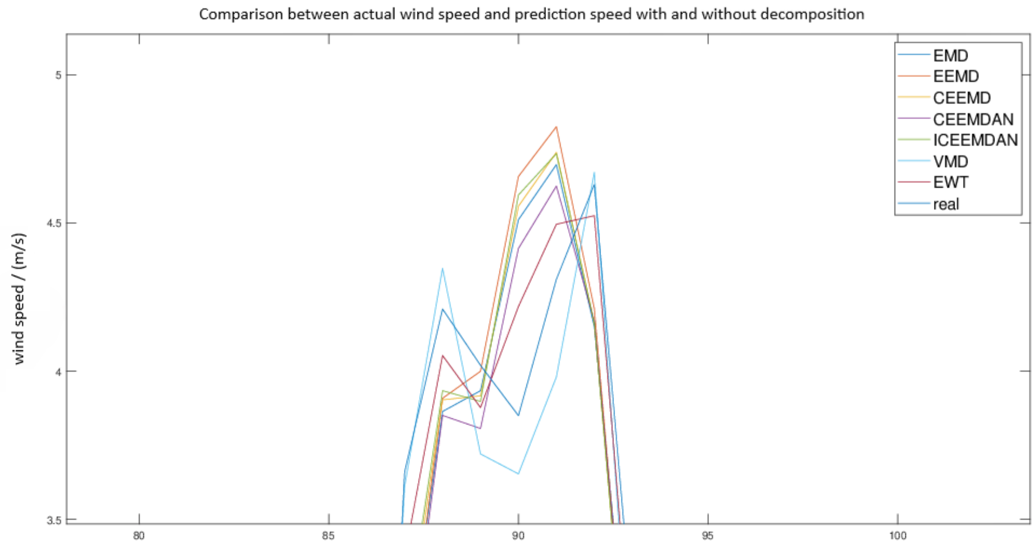

FIG. 5. Zoom in of the previous graph

V. CONCLUSION

Result shows that VMD has the best performance, and EWT is the second. And those modified EMD seems to have little improvement on EMD.

REFERENCES

[1] J. P. Carvallo, N. Zhang, S. P. Murphy, B. D. Leibowicz, and P. H. Larsen, "The economic value of a centralized approach to distributed resource investment and operation," Applied Energy **269**, 115071 (2020).

[2] T. Mao, B. Zhou, X. Zhang, W. Yao, and Z. Zhu, "Accommodation of clean energy: Challenges and practices in china southern region," IEEE Open Journal of Power Electronics **1**, 198–209 (2020).

[3] X. Zhao, J. Liu, D. Yu, and J. Chang, "One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data," Energy Conversion and Management **164**, 560–569 (2018).

[4] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, "A novel connectionist system for unconstrained handwriting recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence **31**, 855–868 (2009).

[5] K. Chen, Y. Zhou, and F. Dai, "A lstm-based method for stock returns prediction: A case study of china stock market," in *2015 IEEE International Conference on Big Data (Big Data)* (2015) pp. 2823–2824.

[6] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. Math. Phys. Eng. Sci. **454**, 903–995 (1998).

[7] Z. WU and N. E. HUANG, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," Advances in Adaptive Data Analysis **01**, 1–41 (2009), https://doi.org/10.1142/S1793536909000047.

[8] J. R. Yeh, J. S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method," Advances in Adaptive Data Analysis **2**, 135–156 (2010).

[9] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2011) pp. 4144–4147.

[10] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble emd: A suitable tool for biomedical signal processing," Biomedical Signal Processing and Control **14**, 19–29 (2014).

[11] G. Chen, B. Tang, X. Zeng, P. Zhou, P. Kang, and H. Long, "Short-term wind speed forecasting based on long short-term memory and improved bp neural network," International Journal of Electrical Power & Energy Systems **134**, 107365 (2022).

[12] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," IEEE Transactions on Signal Processing **62**, 531–544 (2014).

[13] J. Gilles, "Empirical wavelet transform," IEEE Transactions on Signal Processing **61**, 3999–4010 (2013).