

# Multi-Step Wind Speed Forecasting Using Signal Decomposing Algorithms, Bat Optimization Algorithm and Least Squares Support Vector Machine

Zhou Li <sup>a)</sup> and Chunxiang Li <sup>b)</sup>

*Department of Civil Engineering, Shanghai University, Shanghai 200444, China.*

<sup>a)</sup> Corresponding author: lizhouxs@163.com

<sup>b)</sup> li-chunxiang@vip.sina.com

**Abstract.** Accurate forecasting of short term wind speed has been widely applied in the disaster early warning of civil engineering. Considering the characteristics of non-stationary and nonlinear of wind speed, the actual wind speed time series need to be decomposed first and then predicted. In this paper, a set of actual wind speed time series of typhoon is decomposed by four signal decomposing algorithms. (e.g., Wavelet Packet Decomposition/Ensemble Empirical Mode Decomposition/Variational Mode Decomposition/Empirical Wavelet Transform) And the features of intrinsic mode functions from these four methods are fully evaluated. Finally, multi-step wind speed forecasting experiment based on least squares support vector machine optimized by bat algorithm is carried out to testify the effectiveness of the signal decomposing algorithm. The results of experiences indicate that the Empirical Wavelet Transform is effective in the wind speed accurate forecasting.

**Key words:** non-stationary; wind speed; four signal decomposing algorithms.; bat algorithm; decomposing algorithm.

## INTRODUCTION

In history, between June and October each year, the typhoon will attack the coastal cities of Southeast China. With the construction of super high-rise buildings, large span bridges, stadiums, wind farms and other urban facilities in recent years, a series of wind speed prediction models are proposed by the scholars who want to reduce the loss in the typhoon [1-3].

Due to the nonlinear and non-stationary characteristics of the typhoon wind speed time series, the traditional prediction method is difficult to obtain accurate forecasting results. Fortunately, with the currently popular used concept of “decomposition and ensemble”, complex signal can use such process: decomposition-prediction-reconstruction, to obtain satisfactory prediction results. On the basis of the existing research, this paper discusses the effect of different signal decomposition algorithms in the multi-step prediction of wind speed.

The remaining section of this paper is organized as follows: Section 2 introduces the characteristics of four signal decomposition algorithms. Section 3 discusses the experiences and results of multi-step wind speed forecasting. Finally, section 4 outlines the conclusions and future works.

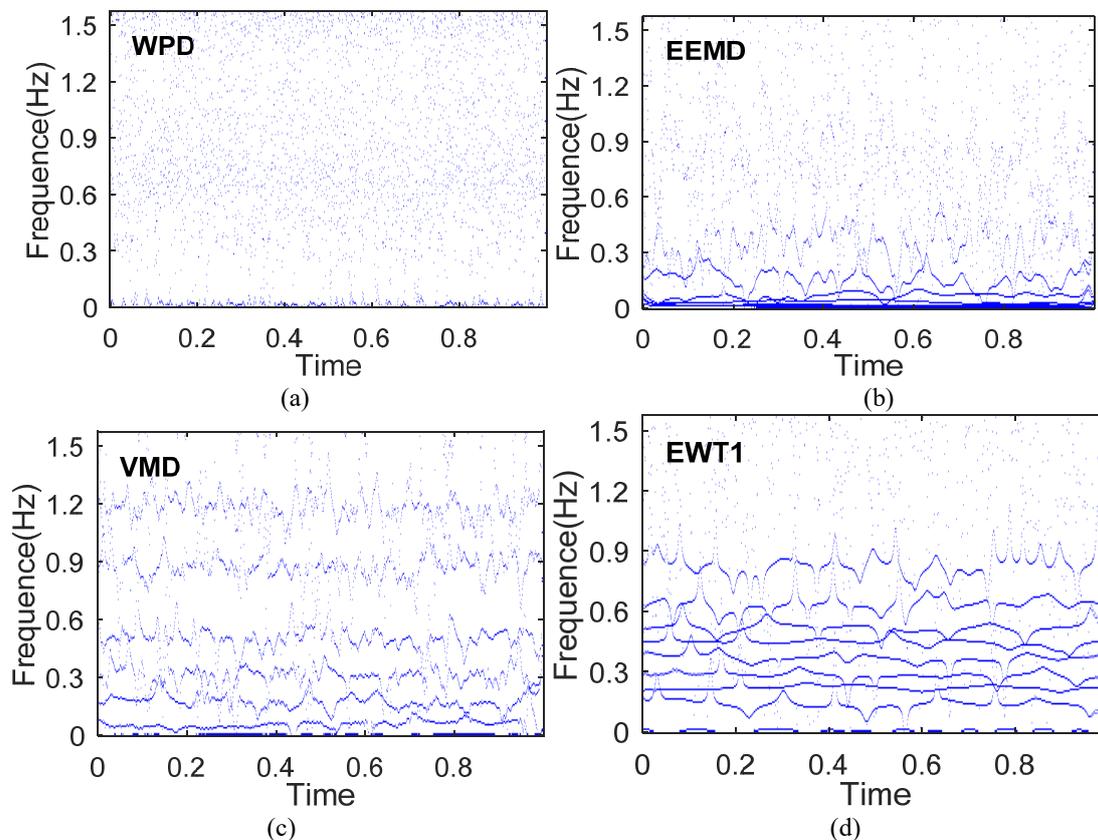
## WIND SPEED DECOMPOSITION

Various techniques for decomposing of wind speed time series have been proposed recently. This study summarizes the characteristics of the wavelet packet decomposition (WPD), ensemble empirical mode decomposition (EEMD), variational mode decomposition (VMD) and empirical wavelet transform (EWT) in detail.

Limited to the length of article, the theory of four decomposing methods can be found in reference [1-4], respectively.

The WPD method is a modified version of the wavelet decomposition (WD) method [1]. The main difference between them can be explained as: in the WD algorithm, only the detailed components will be decomposed. However, in the WPD algorithm, besides the detailed components, the approximate components will also be processed. So, for non-stationary signal, WPD can always have better decomposition results than WD. On the basis of empirical mode decomposition (EMD), Wu et al. [2] proposed the EEMD method to solve the mode mixing problem of EMD in a certain extent through adding white noise to the signal; simultaneously, the ensemble average is obtained by calculating all inherent modal functions (IMFs), which makes the decomposition more robust. However, the lack of mathematical theory and decomposition based on time domain of signal also restrict the improvement of EEMD method. VMD is a technology developed by Dragomiretskiy in 2014 [3], which can also be used to decompose signal into a limited number of IMFs. Each individual component has specific sparsity properties in the frequency domain. One of the main assumptions of this technique is that each IMF has a limited bandwidth in the frequency domain and is compact around a center frequency. Combining the advantages of wavelet transform and empirical mode decomposition, Gilles proposed EWT method in reference [4]. The EWT builds a set of band pass filters whose supports depend on the information of the spectrum of the analyzed signal. Thanks to many scholars' research, the EWT method contains a variety of Fourier spectral division methods, which are determined by parameters. In this paper, three typical classification methods are selected, their parameters are shown as follows: EWT1: `params.detect = 'locmaxmin'`; `params.reg = 'upenv'`; `params.N = 10`; EWT2: `params.detect = 'scalespace'`; `params.typeDetect = 'otsu'`; `params.reg = 'upenv'`; EWT3: `params.detect = 'scale space'`; `params.typeDetect = 'otsu'`; `params.reg = 'none'`; In comparison,, EWT1 focuses on the decomposition of low frequency parts, while EWT3 focuses on the decomposition of high frequency parts, and EWT2 is between the two.

Based on the method mentioned above, we can get a series of intrinsic modal functions of the original signal. By evaluating the Hilbert transform of each IMF and then plotting each curve in the time-frequency plane, Fig. 1 shows the time-frequency representation based on the WPD, EEMD, VMD, EWT1, EWT2 and EWT3, respectively. It could help us to see how the frequency of each IMF changes over time.



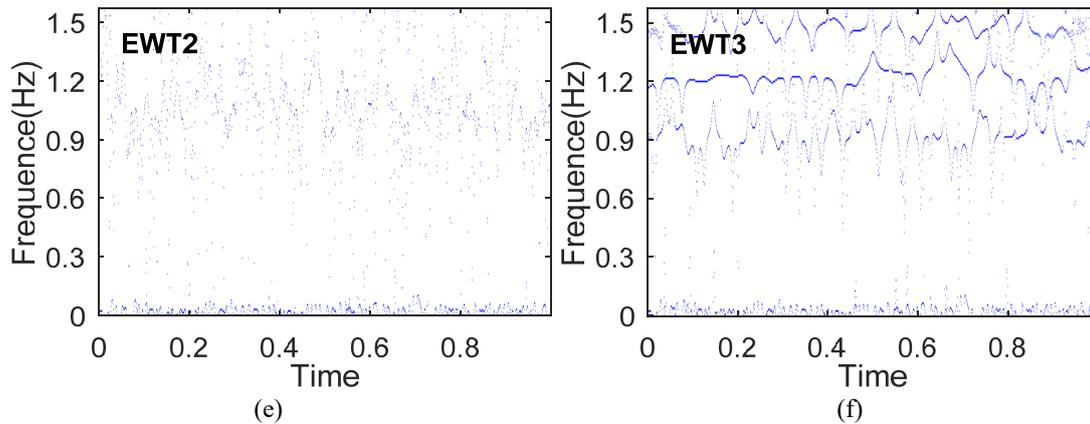


FIGURE 1. Comparison of Time-frequency representation

In order to quantify the complexity of the IMFs, Permutation entropy (PE) is introduced in this paper. For the convenience of application, we normalize PE by virtue of formula (1) to convert PE value into the range from 0 to 1.

$$h(p) = H(p) / \ln(N - d + 1) \quad (1)$$

Where  $H(p)$  is original PE value;  $N$  is the length of time series;  $d$  is embedding dimension. When the regularity of time series is better, the complexity is lower and PE is smaller [5].

What needs to be explained is that the actual wind speed time series in this paper comes from the 50-meter height of the Stonecutters Bridge, which is measured during the Typhoon Dujan. In this paper, the bat optimization is used to determine the penalty parameter and kernel parameter of the least squares support vector machine [6-7].

## WIND SPEED FORECASTING

### Hybrid Forecasting Strategy

Fig. 2 depicts the framework of multi-step wind speed forecasting experience. The whole process consists of the following four steps: step one decomposes actual data into several IMFs by WPD, EEMD, VMD and EWT method. Hilbert transform and permutation entropy are utilized to evaluate the frequency-domain characteristics and complexity of each IMF in step two. And in step three, the BA-LSSVM is employed to predict each IMF. Step four reconstruct all the forecasting results and calculate the error evaluation indexes.

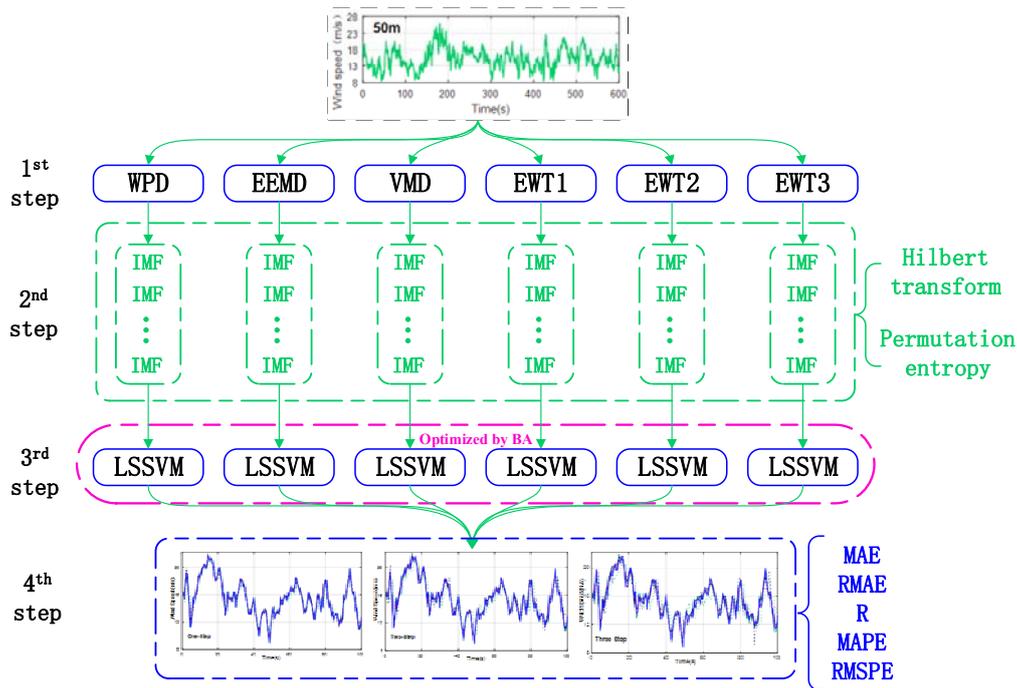


FIGURE 2. The flow chart for multi-step forecasting of wind speed

### Error Evaluation Indexes

To quantitatively determine the effectiveness of the involved models, the error evaluation indexes including mean absolute error (MAE), root-mean-square error (RMSE), correlation coefficient (R), mean absolute percentage error (MAPE) and root-mean-square proportional error (RMSPE). For these five indexes, the specific calculation formula is shown as follows:

$$MAE = \frac{1}{L} \sum_{n=0}^L |f(n) - h(n)| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{L} \sum_{n=0}^L (f(n) - h(n))^2} \quad (3)$$

$$R = \frac{cov(f(n), h(n))}{\sqrt{cov(f(n), f(n)) \times cov(h(n), h(n))}} \quad (4)$$

$$MAPE = \frac{1}{L} \sum_{n=0}^L \left| \frac{f(n) - h(n)}{f(n)} \times 100\% \right| \quad (5)$$

$$RMSPE = \sqrt{\frac{1}{L} \sum_{n=0}^L \left( \frac{f(n) - h(n)}{f(n)} \times 100\% \right)^2} \quad (6)$$

Where  $f(n)$  is the original signal of testing dataset,  $h(n)$  is the forecasting results of testing dataset,  $n$  is the serial number of samples,  $L$  is the number of testing dataset samples, and  $COV$  refers to calculating the covariance matrix.

### Computational Results

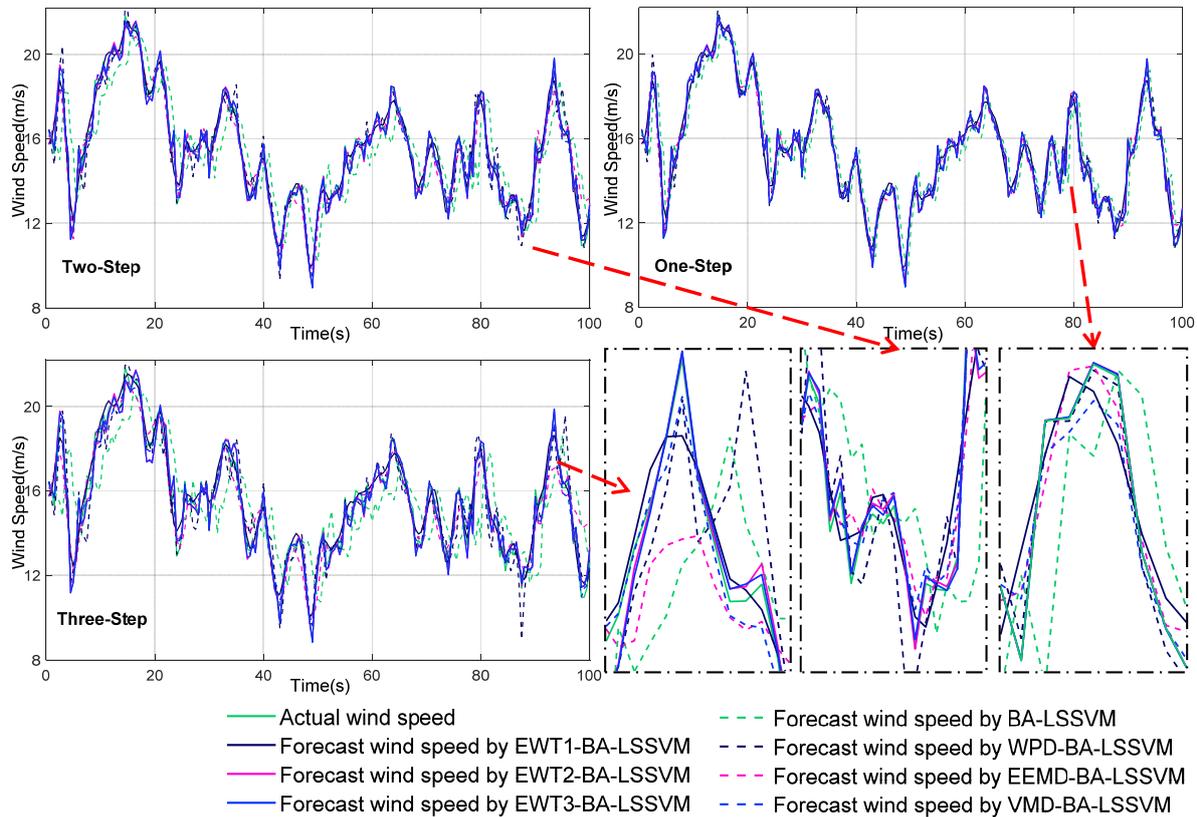


FIGURE 3. The results of the multi-step predictions for the actual wind speed series.

Fig. 3 demonstrates the comparing results of multi-step wind speed forecasting based on different signal decomposing algorithm. Table 2 gives the error estimated results of the forecasting experiences. From Table 1, it can be found that: (a) The BA-LSSVMs with signal decomposing algorithms have better forecasting performance than the single BA-LSSVM algorithm. The reason of the phenomenon is that the signal decomposing algorithms can transform complex signals into a number of simple components; (b) In general, the three cases of EWT are better than the other three decomposing methods, this is mainly due to the good time frequency concentration of EWT; (c) Comparison of three cases of EWT shows that the decomposition of the decomposition focusing on the high frequency parts of signals obtained the best prediction performance. The reason is that the high frequency parts of wind speed signals are too complex to predict accurately. Additionally, we can draw the following conclusions from the time frequency representation and permutation entropy of IMFs: (d) Compared with the decomposition based on time-domain, the decomposition based on frequency-domain makes the mode have smaller permutation entropy, and the prediction result is better; (e) The modes with large bandwidth in frequency domain are more difficult to predict.

**TABLE 1.** Forecasting results of error evaluation indexes for different models.

	Index	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
One Step	MAE	0.1586	0.0285	0.1011	0.2633	0.0924	0.0151	0.0149
	RMSE	0.4755	0.0925	0.1847	0.3437	0.2309	0.0258	0.0225
	R	0.9889	0.9996	0.9983	0.9949	0.9974	0.9999	0.9999
	MAPE	0.0107	0.0019	0.0069	0.0179	0.0063	0.0010	0.0010
	RMSPE	0.0332	0.0062	0.0129	0.0246	0.0160	0.0017	0.0015
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Two Step	MAE	0.2387	0.0659	0.1357	0.2654	0.0936	0.0338	0.0320
	RMSE	0.6759	0.2029	0.2891	0.3451	0.2360	0.0777	0.0689
	R	0.9776	0.9980	0.9959	0.9949	0.9973	0.9997	0.9998
	MAPE	0.0161	0.0044	0.0093	0.0180	0.0064	0.0022	0.0021
	RMSPE	0.0474	0.0136	0.0210	0.0246	0.0165	0.0049	0.0044
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Three Step	MAE	0.2978	0.1013	0.1646	0.2730	0.1026	0.0636	0.0641
	RMSE	0.8293	0.3182	0.3855	0.3534	0.2573	0.1611	0.1570
	R	0.9661	0.9951	0.9928	0.9946	0.9968	0.9987	0.9988
	MAPE	0.0200	0.0068	0.0113	0.0185	0.0070	0.0042	0.0042
	RMSPE	0.0585	0.0216	0.0285	0.0252	0.0182	0.0104	0.0102

Notes: Model 1: BA-LSSVM; Model 2: WPD-BA-LSSVM; Model 3: EEMD-BA-LSSVM; Model 4: VMD-BA-LSSVM; Model 5: EWT1-BA-LSSVM; Model 6: EWT2-BA-LSSVM; Model 7: EWT3-BA-LSSVM.

## CONCLUSION

This paper compares the influence of different signal decomposition methods on the accuracy of multi-step wind speed forecasting and evaluates the characteristics of different modes. The experimental results show that the application of signal decomposing algorithms can obviously improve the prediction precision of multi-step wind speed forecasting, especially the EWT algorithm for high frequency parts decomposition. This paper provides a general conclusion for the study of signal decomposition algorithms apply to the field of multi-step wind speed forecasting. It also provides some guidance for the improvement of the original decomposing algorithm.

## ACKNOWLEDGMENTS

This study is supported by the National Natural Science Foundation of China (Grant No. 51778354 and Grant No. 51378304). The authors also would like to thank Prof. Youlin Xu for providing the actual wind speed data of this study generously.

## REFERENCES

1. H. Liu, H.Q. Tian, Y.F. Li. Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms. *Energy Conversion & Management*, 2015, 100:16-22.
2. Z.H. Wu, N.E. Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 2009, 1(01):1-41.
3. Dragomiretskiy K, Zosso D. Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, 2014, 62(3):531-544.
4. Gilles J. Empirical Wavelet Transform. *IEEE Transactions on Signal Processing*, 2013, 61(16):3999-4010.
5. Lian J, Liu Z, Wang H, et al. Adaptive variational mode decomposition method for signal processing based on mode characteristic[J]. *Mechanical Systems & Signal Processing*, 2018, 107:53-77.
6. T. Niu, J.Z. Wang, K.Q. Zhang, P. Du. Multi-step-ahead Wind Speed Forecasting Based on Optimal Feature Selection and a Modified Bat Algorithm with the Cognition Strategy. *Renewable Energy*, 2018, 118.
7. B.Z. Zhu, D. Han, P. Wang, Z.C. Wu, T. Zhang, Y.M. Wei. Forecasting carbon price with empirical mode decomposition and least squares support vector regression. *Pricing and Forecasting Carbon Markets*. Springer International Publishing, 2017:521-530.