A Novel Multi-Step Cross-Decomposition Method Based on Wavelet Transform for Wind Power Prediction

Jianxun Lang[1]

[1]School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, China

Abstract. One of the main approaches to improve wind power prediction accuracy is to decompose wind-speed into different frequency-band components and use them as inputs of prediction model. Among the decomposition methods, wavelet transform is widely used due to its flexibility. However, the decomposition level and wavelet function need to be selected through trail-and-error, which is also called empirical decomposition method, because the effectiveness of a certain selection depends on the characteristic of wind farm and the prediction model. Therefore, it is difficult to find a general decomposition method that can be effective on different prediction models and wind farms. Aiming at this problem, a novel multi-step cross-decomposition method is proposed in this paper. The proposed method decomposes the wind-speed and power alternatively in each step, and after three steps of decomposition, the wind-speed can be decomposed to four different frequency-band components which will be used as the input of the prediction model. The prediction errors of proposed method and several empirical decomposition methods are compared on BPNN and SVM models. The results show that the proposed method is the only effective method on two prediction models for four wind farms.

1 Introduction

One of the main approaches to improve the wind power prediction (WPP) accuracy is to decompose the wind-speed provided by numerical weather prediction (NWP) to obtain several components with different frequency-band, and then apply them as the input of the prediction model. The main decomposition methods contain variational mode decomposition(VMD)[1-2] ,empirical mode decomposition (EMD) [3-4] and wavelet transform (WT) [5-6], among which WT is widely used due to its flexibility. When using WT to decompose wind-speed, there is no definite standard for the selection of decomposition level and wavelet function. Since not all WT decomposition methods can improve the prediction accuracy, researchers can only determine the decomposition level and wavelet function according to the prediction errors on a certain prediction model for a certain wind farm through trial-and-error, which is also called empirical decomposition method. Considering the diversity of selection of decomposition level and wavelet function , it is difficult to find a general decomposition method that works well on different wind farms and on different prediction models. Therefore, to find a deterministic decomposition method that can performs well on different prediction models for wind farms with different characteristics will greatly reduce the workload of researchers.

In this paper, a novel multi-step cross-decomposition method based on WT is proposed for wind-speed decomposition. First, an index called interpretation-ratio

(IR) is defined. Based on this index, a multi-step cross-decomposition between wind-speed and power is carried out. In each step, wind-speed and power will be decomposed alternatively, and the decomposition level and wavelet function which make the IR index maximum are selected as the optimal decomposition level(ODL) and the optimal wavelet function(OWF). After three steps decomposition, components with four different frequency-band of wind-speed can be obtained and will be used as the input of the prediction models. Finally, the prediction errors of proposed decomposition method and several empirical decomposition methods are compared on SVM and BPNN models to verify the effectiveness of the proposed method.

2 Definition of IR

Let $W=[w_1,w_2,...w_n]$ denotes a single wind-speed prediction series with n samples, the difference sequence of W can be calculated by formula (1):

$$dW = [w_2 - w_1, w_3 - w_2,..., w_n - w_{n-1}] \quad (1)$$

Let $P=\{p_1,p_2,...,p_n\}$ denotes a power observation series with n samples, the difference sequence of P can be calculated by formula (2):

$$dP = [p_2 - p_1, p_3 - p_2,..., p_n - p_{n-1}] \quad (2)$$

The index IR is defined as:

$$IR = \frac{sum(sign(dW) == sign(dP))}{n-1} \quad (3)$$

Where *sign* is the the symbolic function defined as

---

* Corresponding author: langjianxun2012@126.com

$$sign(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \qquad (4)$$

The NWP used in this paper provides 4 wind-speed prediction value at altitude 170m,100m,30m and 10m, which are denoted by W170, W100, W30, and W10 respectively, their difference sequence are denoted by dW170, dW100, dW30 and dW10 respectively. For the sake of brevity, a set W={W170,W100,W30,W10} is used to represent the 4 winds.

If the 4 winds are used together as the input of prediction model, then the IR is defined as:

$$IR = \frac{sum(m170 \,|\, m100 \,|\, m30 \,|\, m10)}{n - 1} \qquad (5)$$

Where the '|' represents the logical 'or' operation. m170, m100, m30 and m10 are calculated by formula (6)~(9).

$$m170 = sign(dW170) == sign(dP) \qquad (6)$$

$$m100 = sign(dW100) == sign(dP) \qquad (7)$$

$$m30 = sign(dW30) == sign(dP) \qquad (8)$$

$$m10 = sign(dW10) == sign(dP) \qquad (9)$$

3 Multi-step cross-decomposition based on maximum IR

3.1. Scope of decomposition level and wavelet

The power and wind-speed will be decomposed by experimental method, therefore the scopes of decomposition level and wavelet function should be determined first. In this paper, the range of decomposition level is set to 1~9，and the wavelet function is selected from the following wavelet series: {db$N$, bior$N_r.N_d$, coif$N$, sym$N$, fk$N$}, which have 60 kinds of wavelet functions.

3.2. 1th-step cross-decomposition process

The decomposition order of power and wind-speed should be determined first. The fluctuation of power observation series is usually higher than that of wind-speed prediction series, so the power series is selected to be decomposed first.

Choosing a wavelet function from aforementioned wavelet series and decomposing the power series to level 1~9 produces 60×9=540 results, with each result containing a scale component AP(1), where '(1)' denotes the '1th-step' decomposition. For each decomposition result, calculate the difference series dAP(1) by formula (2) then replace dP with dAP(1) in formula (5) will get a specific IR value. For 540 decomposition results, 540 IRs are produced. Finding out the maximum IR among the 540 IRs, the corresponding wavelet function is called optimal wavelet function for power decomposition in 1th-step，denoted by OWF_P(1). The corresponding decomposition level is called the optimal decomposition level for power decomposition in 1th-step, denoted by ODL_P(1). The corresponding AP(1) is denoted by AP(1)_opt.

Replaces dP in formula (3) with dAP(1) and starts to decompose the 4 winds. For each wind, e.g., W170, choosing a wavelet function from aforementioned wavelet series and decomposing the W170 to level 1~9 produces 60×9=540 results, with each result containing a scale component AW170(1). For each decomposition result, calculates the difference series dAW170(1) by formula (1) then replaces dW with dAW170(1) in formula (3) will get a specific IR value. For 540 decomposition results, 540 IRs are produced. Finding out the maximum IR among the 540 IRs, the corresponding wavelet function is called optimal wavelet function for W170 in 1th-step, denoted by OWF_AW170(1). The corresponding decomposition level is called the optimal decomposition level for W170 in 1th-step, denoted by ODL_AW170(1). The corresponding AW170(1) is denoted by AW170(1)_opt.

For W100,W30 and W10, repeat the above process independently to get their own optimal wavelet function, optimal decomposition level and optimal scale component.

The above process is called an one step cross-decomposition process. There are 4 wind farms numbered 1#~4# that will be used in this paper. Taking 1# wind farm as an example, the ODL and OWF for power and 4 winds in 1th-step cross-decomposition are listed in Table 1.

**Table 1.** ODL and OWF of 4 winds and power for wind farm 1# in 1th-step cross-decomposition

		ODL	OWF
W	W170	5	db16
	W100	5	sym9
	W30	5	db9
	W10	6	db9
P		5	sym9

For the sake of brevity, the four optimal scale components of 4 winds are represented by a set

$$AW(1)\_opt = \{AW170(1)\_opt, AW100(1)\_opt, \\ AW30(1)\_opt, AW10(1)\_opt\} \qquad (10)$$

3.3. 2th-step and 3 th step cross-decomposition process

After the 1th-step cross-decomposition process, the scale components,i.e.,components occupied the low-frequency-band are separated from the original wind series and power series. But it can be seen from Table 1 that the AW170(1)_opt, AW100(1)_opt and AW30(1)_opt components are both scale components obtained by 5-level decomposition, which indicates that these 3 components merely occupied the (0~π/32) scope of their normalized frequency-band (0~π); the AW10(1)_opt is obtained by 6-level decomposition, which indicate that it merely occupied the (0~π/64) scope of its normalized frequency-band (0~π),as shown in Fig.1.
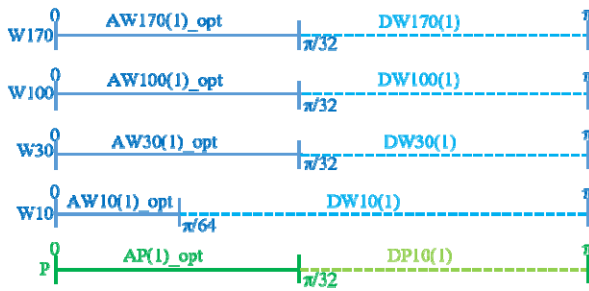
Fig.1. Demonstration of frequency-bands occupied by components of 4 winds and power

In Fig.1, the high-frequency-band components corresponding to the rest scope of W170, W100, W30, W10 and P are denoted by DW170(1), DW100(1), DW30(1), DW10(1) and DP(1) respectively, and can be calculated by formula (11):

$$DW170(1) = W170 - AW170(1)\_opt$$
$$DW100(1) = W100 - AW100(1)\_opt$$
$$DW30(1) = W30 - AW30(1)\_opt \qquad (11)$$
$$DW10(1) = W10 - AW10(1)\_opt$$
$$DP(1) = P - AW170(1)\_opt$$

For the sake of brevity, DW170(1), DW100(1), DW30(1), and DW10(1) are represented by a set DW(1), as shown in formula (12):

$$DW(1) = \{DW170(1), DW100(1), DW30(1), DW10(1)\} \quad (12)$$

DW(1) should be further decomposed, the decomposition process is illustrated below.

3.3.1 2th-step cross-decomposition process

Replacing W by DW(1) and P by DP(1) and repeating the decomposition process of 1th-step will produce AW(2)_opt, AP(2)_opt, DW(2) and DP(2). The corresponding optimal decomposition level and optimal wavelet function are listed in Table 2.

Table 2. ODL and OWF of 4 winds and power for wind farm 1# in 2th-step cross-decomposition

		ODL	OWF
DW(1)	DW170(1)	4	sym17
	DW100(1)	4	sym19
	DW30(1)	4	bior6.8
	DW10(1)	4	bior2.8
DP(1)		4	sym19

3.3.2 3th-step cross-decomposition process

Replacing W by DW(2) and P by DP(2) and repeating the process in 1th-step will produce AW(3)_opt, AP(3)_opt, DW(3) and DP(3). The corresponding optimal decomposition level and optimal wavelet function are listed in Table 3.

After 3 steps of decomposition, 4 different frequency-band components are separated from the original winds series, which can be written as a set of $\{AW(1)\_opt, AW(2)\_opt, AW(3)\_opt, DW(3)\}$, with a dimension of 16.

Table 3. ODL and OWF of 4 winds and power for wind farm 1# in 3th-step cross-decomposition

		ODL	OWF
DW(2)	DW170(2)	4	sym17
	DW100(2)	4	db10
	DW30(2)	5	fk4
	DW10(2)	5	fk4
DP(2)		4	db10

3.4 Input construction for prediction models

In addition to wind-speed, NWP also provides wind-direction predictions at altitude of 170m,100m,30m and 10m, which are be denoted by θ170, θ100, θ30 and θ10 respectively.

For each component in {AW(1)_opt, AW(2)_opt, AW(3)_opt, DW(3)}, e.g., AW170(1)_opt, use θ170 to further decompose the components to cosine and sine components, according to formula (13)~(14), which is also called triangle decomposition.

$$AW170\_opt\_cos = AW170\_opt \times \cos(\theta170) \quad (13)$$
$$AW170\_opt\_sin = AW170\_opt \times \sin(\theta170) \quad (14)$$

After triangle decomposition for each components in the set {AW(1)_opt,AW(2)_opt,AW(3)_opt,DW(3)}, the dimension is increased to 32 accordingly. The 32 components will together be used as the input of the prediction model later.

4 Compare and analysis

4.1 Construction of benchmark empirical decomposition methods

In order to verify the effectiveness of proposed decomposition method, several input sets need to be constructed with different empirical decomposition methods for comparison. In order to increase the diversity of benchmark decomposition methods, for each wind farm, suppose that its W100 were decomposed by wavelet A to level k in 1th-step decomposition, the following methods are used to construct the benchmark empirical decomposition methods, as shown in Table 4. The wavelet A and level k for a specific wind farm are listed in Table 5.

Table 4. Benchmark empirical decomposition methods for each wind farm

method	wavelet function	decomposition level
1	—	—
2	db3	3
3	A	3
4	A	k
5	db3	k

The '—' for method 1 means that method 1 does not use WT decomposition but just adopts the triangular decomposition to the 4 original winds.

Note that each component obtained through method 2~5 adopts the triangular decomposition as the proposed method.

Table 5. A and k for wind farm 1#~4#

wind farm	wavelet A	level k
1#	sym9	5
2#	db6	5
3#	sym17	5
4#	sym9	5

The components obtained by the proposed method and the methods listed in Table 4 will be used as the input of BPNN and SVM. The evaluation index of the prediction error is normalized-root-mean-square-error (NRMSE).

4.2 Comparison on BPNN model

The NRMSE on BPNN model for proposed methods and empirical decomposition methods are shown in Table 6. For each method, the average value of NRMSE on 4 wind farms(i.e.,average error) is calculated and listed in the last column of Table 6. If the average error of a certain method is larger than that of method 1, this decomposition method is ineffective, otherwise,effective.

Table 6. Prediction error on BPNN model (NRMSE,%)

	1#	2#	3#	4#	average
method 1	16.01	15.65	15.59	17.07	16.08
method 2	16.39	16.16	16.5	17.12	16.54
method 3	15.89	16.19	16.6	17.01	16.42
method 4	14.59	14.73	14.24	15.55	14.78
method 5	14.88	14.74	15.36	16.28	15.32
proposed	13.62	13.11	13.73	13.72	13.55

The following conclusions can be drawn from Table 6.

(1) The average error of method 1 is 16.08%. The average errors of method 2 and method 3 are 16.54% and 16.42% respectively, both larger than that of method 1, hence the method 2 and method 3 are ineffective. The average error of proposed method is 13.55%, which is less than that of method 1, so that the proposed method is effective. With a same input dimension of 32, the proposed decomposition method is effective while method 2 and method 3 are ineffective.

(2) The average errors of method 4 and method 5 are 14.78% and 15.32% respectively, both less than that of method 1, hence the method 4 and method 5 are effective. But the two average errors are still larger than that of the proposed method. Considering that method 4~method 5 both have a input dimension of 48 which is higher than that of the proposed method, the proposed method can achieve a higher prediction accuracy with a lower input dimension.

4.3 Comparison on SVM model

The NRMSE on SVM model for proposed methods and empirical decomposition methods are shown in Table 7. For each method, the average error on 4 wind farms is calculated and listed in the last column of Table 7.

It can be seen from Table 7 that the average errors of method 2~5 are 17.11%, 16.92%, 14.83% and 16.34% respectively, all larger than the average error 14.44% of method 1, which indicates that method 2~5 are all ineffective on SVM model. However, the average error of the proposed method, 12.68%, is less than that of method 1, which indicates that the proposed method is the only effective methods on SVM model.

Table 7. Prediction error on SVM model (NRMSE,%)

	1#	2#	3#	4#	average
method 1	14.44	14.49	14.62	15.15	14.44
method 2	17.11	18.02	17.68	18.41	17.11
method 3	16.92	17.41	17.49	18.41	16.92
method 4	14.83	16.41	15.09	16.47	14.83
method 5	16.34	16.34	15.15	16.42	16.34
proposed	12.68	12.07	13.93	12.9	12.68

5 Conclusion

When using the the empirical decomposition method, the effectiveness of a certain decomposition method depends on the decomposition level, wavelet function, prediction model and the characteristic of wind farm. It needs a lot of work to find an effective method through trail-and-error. The multi-step cross-decomposition method is a deterministic method and can be effective on both BPNN and SVM models. Moreover, it can achieve a higher prediction accuracy with a lower input dimension, which will facilitate the the follow-up design.

References

1. L. Han, Z. Rongchang, W. Xuesong, J. Huitian, Multi-step wind power forecast based on VMD-LSTM. J. IET.Renew.Power.Gener **13**, 1690-1700 (2019)
2. S. Zexian, Z. Shenshen, Z. Jingxuan, Short-term wind power forecasting on multiple scales using VMD decomposition, K-Means clustering and LSTM principal computing. J. IEEE.Access **7**, 166917-166929 (2019)
3. Y. Zhang, K. Liu, L. Qin, Wind power multi-step interval forecast based on ensemble empirical mode decomposition-sample entropy and optimized extreme learning machine. J. Power.Syst.Technol **40**, 2045-2051 (2016)
4. N. Bokdea, A. Feijóob, K. Kulataa, Analysis of differencing and decomposition preprocessing methods for wind speed prediction. J. Appl.Soft.Comput **71**, 926-938 (2018)
5. De. Giorgi, S. Campilongo, A. Ficarella, Comparison between wind power prediction models based on wavelet decomposition with least-squares support

vector machine (ls-svm) and artificial neural network (ann). J. Energies **7**, 5251-5272 (2014)

6. A. Tascikaraoglu, B. M. Sanandaji, K. Poolla, Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using wavelet transform. J. Appl. Energy **165**, 735-747 (2016)