Energy Conversion and Management 148 (2017) 895–904

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

An improved Wavelet Transform using Singular Spectrum Analysis for wind speed forecasting based on Elman Neural Network

Chuanjin Yu, Yongle Li *, Mingjin Zhang

Department of Bridge Engineering, Southwest Jiaotong University, 610031 Chengdu, Sichuan, PR China

ARTICLE INFO

Article history:
Received 13 February 2017
Received in revised form 15 May 2017
Accepted 26 May 2017

Keywords:
Wavelet transform
Singular Spectrum Analysis
Elman Neural Network
Wind speed prediction

ABSTRACT

To raise the wind speed prediction accuracy, Wavelet Transform (WT) is widely employed to disaggregate an original wind speed series into several sub series before forecasting. However, the highest frequency sub series usually has a great disturbance on the final prediction. In the study, for raising the forecasting accuracy, Singular Spectrum Analysis (SSA) is applied to make further processing on the highest frequency sub series, instead of making no modification on or getting rid of it. So a hybrid decomposition technology called Improved WT (IWT) is proposed. Meanwhile, a new hybrid model IWT-ENN combined with IWT and Elman Neural Network (ENN) is also designed. The procedure of IWT is systematically investigated. Experimental results show that: (1) the performance of the hybrid model IWT-ENN has a great improvement compared to that of others including the persistence method, ENN, Auto-Regressive (AR) model, Back Propagation Neural Network (BPNN) and Empirical Mode decomposition (EMD)-ENN; (2) compared to the two general strategies where the highest frequency sub series is without retreatment or eliminated, the new proposed hybrid model IWT-ENN has the best prediction performance.

1. Introduction

Wind power, as one of the plentiful and renewable energies, has becoming increasingly competitive and universally used [1–4]. Owing to the wind inherent characteristics of intermittence and fluctuation, to protect the safety of the power utilization and integration, it is essential and significant to make high precision forecasting of wind speed [5,6].

In recent years, there have been many researches about hybrid models merged with decomposition technologies, including Empirical Mode decomposition (EMD) [7,8] and Wavelet Transform (WT) etc., to make predictions on wind speed and power. EMD is a data driven method, without any prior knowledge. But WT is still one of the most popular method, whose process has a more clear physical meaning.

Catalão et al. [9,10] proposed two hybrid models based on WT and ANN, including Back Propagation Neural Network (BPNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) Neural Network optimized by Particle Swarm Optimization (PSO) to make short-term wind power. The time series was fell part into several sub series by WT and then were individually modeled and predicted. Liu et al. [11] applied WT to dissolve the short-term wind power time series into a sequence of sub series, which were respectively forecasted by Support Vector Machines (SVMs) with the wavelet kernel function and radial basis kernel function for two different hybrid models. Liu et al. [12–14] designed some hybrid models based on Wavelet Transform combined with Multilayer Perception (MLP) Neural Network, ANFIS Neural Network and Extreme Learning Machines respectively to make wind speed prediction. Osório et al. [15] proposed a method involved WT and the ANFIS Neural Network, which was optimized by Evolutionary Particle Swarm Optimization (EPSO) to make wind power prediction. Kiplangat et al. [16] came up with a hybrid model with Auto-Regressive (AR) models combined with WT to make wind speed prediction, whose performance was greater than that of AR and Fractionally Autoregressive Integrated Moving Average Model (f-ARIMA) models proved by some experiments. Tascikaraoglu et al. [17] utilized WT for decomposing the wind speed data into more stationary components and then applied spatio-temporal models on each sub series for prediction. The basic idea of these hybrid models mentioned above is that firstly to decompose a time series into a sequence of sub series by WT, then suitable regression models are adopted to make predictions, all of which are added up as the final prediction. More similarity works can be found in

* Corresponding author.
E-mail addresses: 7818534@qq.com (C.J. Yu), lele@swjtu.edu.cn (Y.L. Li), zhang108119@163.com (M.J. Zhang).

http://dx.doi.org/10.1016/j.enconman.2017.05.063
0196-8904/

Refs. [18–21]. It can be seen that the performance of these hybrid models combined with WT are more great than those without WT, since the subseries decomposed by WT behave more stationary with less fluctuations.

However, the subseries with the highest frequency band always causes a big disturbance on the final prediction. Resulting from that, it is eliminated to improve prediction accuracy by some researchers. Tascikaraoglu et al. [22] used WT to cut up the original wind speed series into a number of better-behaved subseries which were separately forecasted by BPNN except for the highest frequency band. Liu et al. [23] came up with a model based on WT and SVM. The original wind series was decomposed into an approximation signal and a detail signal by WT. The detail signal was eliminated while the approximation signal was forecasted by SVM, whose parameters were optimized by Genetic Algorithm (GA). Wang et al. [24] proposed a model combined with WT, Seasonal Autoregressive Integrated Moving Average Model (SARIMA) and BPNN. There were two parts decomposed by WT from the short-term load series. Only the approximation signal was forecasted by SARIMA and BPNN respectively at first, which then are combined by weighting to get the final short-term load predictions.

From the literature reviewed, it can be seen that WT can actually assist on improving the hybrid models' performance for wind speed prediction. However, the decomposed sub series with the highest frequency usually bother the final prediction accuracy. In order to enhance the forecasting capacity, it is crucial and required to minimize the impact of the highest frequency sub series. To further raising the forecasting accuracy, different from the two common strategies including making no modification on or getting rid of the highest frequency sub series mentioned above, a new hybrid decomposition technology called Improved WT (IWT) is proposed, where SSA is combined with WT to make further operation on the highest frequency sub series. What's more, a new hybrid model IWT-ENN is came up with to make wind speed prediction, where ENN is extensively adopted and proven to do well in wind speed prediction [25]. In the text that follows, the procedure of IWT is systematically investigated. Meanwhile, the forecasting performance of IWT-ENN is compared to that of others including the persistence method, ENN, AR, BPNN and EMD-ENN.

The rest of paper is organized as follows: Section 2 reviews decomposition technologies and proposes the new hybrid decomposition method; Section 3 introduces the forecasting method; Section 4 brings up the new hybrid model; Section 5 presents some experimental results and makes comparative studies.

2. Decomposition technology

2.1. Wavelet transform

Using wavelet transform, the original wind speed series is decomposed into a sequence of sub series with better behavior and more predictable. The Continuous Wavelet Transform (CWT) of a signal $X(t)$ can be illustrated as follows:

$$CWT(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} X(t)\psi^*\left(\frac{t-b}{a}\right)dt \tag{1}$$

where a, b is the scale factor and translation parameter respectively, $\psi(t)$ is the mother wavelet and $*$ denotes the complex conjugate. There is a digitally counterpart of CWT, Discrete Wavelet Transform (DWT), with less computation and approximately accuracy. It is defined as follows:

$$DWT(m,n) = 2^{-(m/2)}\sum_{x=0}^{N-1} X(t)\psi^*\left(\frac{t-n2^m}{2^m}\right) \tag{2}$$

where m and n is the scale factor and translation parameter respectively ($a = 2^m$, $b = n2^m$), and N is the length of the signal $X(t)$.

Based on the Mallat's algorithm [26], the multi-resolution process is achieved where the signal $X(t)$ is disaggregated into some sub series involved with " approximation " and "detail " information respectively. A typical DWT process of the signal $X(t)$ with 3 decomposed level is shown in Fig. 1.

In this study, a wavelet function of type Daubechies of order 6 (abbreviated as Db6) is adopted as the mother wavelet $\psi(t)$ [27,28]. Another is worth mentioning, the characteristic of wavelet decomposition is that the more scales the original signal be decomposed, the better performance the decomposed signals have, but great errors will be generated at the same time [21]. The decomposed level of wavelet transform is 3 in the study by default, which is widely used in many studies [9,10,12].

2.2. Singular Spectrum Analysis

SSA is a popular approach for time series analysis, including trend of quasi-periodic component detection and extraction and denoising [29,30]. The process of SSA is combined with two main stages decomposition and reconstruction. It is detailed described as follows.

Stage I. Decomposition
Step a. Embedding

For a signal $X(t) = (X_1, \ldots, X_N)$, make $Y_i(t) = (X_i, \ldots, X_{i+k-1})$ for L dimensions to make up a trajectory matrix Y:

$$Y = \begin{bmatrix} X_1 & \cdots & X_k \\ \vdots & \ddots & \vdots \\ X_L & \cdots & X_N \end{bmatrix} \tag{3}$$

where $k = N - L + 1$ and Y is Hankel matrix with equal elements along the diagonals ($i + j$ = const)

Step b. Singular Value Decomposition (SVD)

Apply SVD on the matrix XX^T to get its eigentirples (λ_i, U_i, V_i) in descending order by λ_i, in which λ_i is the ith sigular value, U_i and V_i is the ith left and right eigenvector respectively. The trajectory matrix Y can be rewritten as:

$$Y = Y_1 + \ldots + Y_d, Y_i = \sqrt{\lambda_i}U_iV_i^T \tag{4}$$

where $d = \mathrm{rank}(Y)$.

Stage II. Reconstruction
Step c. Grouping

Select r out of d eigentirples. Denote $I = \{I1, \ldots, Ir\}$, which is a sequence of r selected eigentirples and designated as: $Y_I = Y_{I1} + \ldots + Y_{Ir}$. Y_I represents the original data Y, while the others (d-r) eigentirples are supposed as the noise term ϵ.

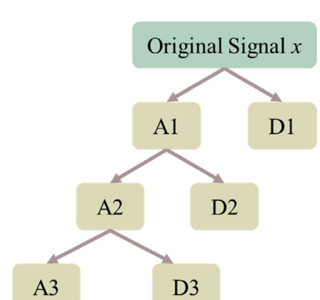

Fig. 1. Wavelet Transform process with 3 decomposed level: A represents the approximate component and D represents the detail component.