Short-Term Local Prediction of Wind Power Based on Singular Spectrum Analysis and Self-Organizing Maps

To cite this article: T Y Ji et al 2018 IOP Conf. Ser.: Mater. Sci. Eng. 366 012015

View the article online for updates and enhancements.

Related content
- Segmentation of Natural Gas Customers in Industrial Sector Using Self-Organizing Map (SOM) Method
  A M Masbar Rus, R Pramudita and I Surjandari
- Russian Character Recognition using Self-Organizing Map
  D Gunawan, D’ Arisandi, F M Ginting et al.
- Applying self-organizing map and modified radial based neural network for clustering and routing optimal path in wireless networks
  Haider K Hoomod and Tuka Kareem Jebur
Short-Term Local Prediction of Wind Power Based on Singular Spectrum Analysis and Self-Organizing Maps

T Y Ji¹, L Liu¹, D Y Hong¹, M S Li¹ and Q H Wu¹

¹ School of Electric Power Engineering, South China University of Technology, Guangzhou, 510640, China

Corresponding author: Dr. M. S. Li. Email: mengshili@scut.edu.cn.

Abstract. Along with the increasing penetration of wind power into power systems, more accurate forecast of wind power becomes more and more important for real-time scheduling and operation. This paper proposes a novel model for short-term wind power forecast based on singular spectrum analysis (SSA) and self-organizing maps (SOM). In order to deal with the impact of high volatility of the original time series, SSA is utilized to extract the mean trend from the original time series. After that, SOM is applied to select the similar segments from mean trend, which are then employed in local prediction by support vector regression (SVR). Simulation studies are conducted on real wind power time series, and the final results indicate that the proposed model is more accurate and stable than other models.

1. Introduction
In recent years, wind energy as a non-pollution, promising type of energy has been the fastest growing renewable energy technology [1]. Obviously, with the increasing popularity of wind power inputting into power systems, the intermittency and randomness of wind have brought great challenge to the safe and stable operation of power systems [2]. Therefore, more accurate and stable forecast is required to guarantee the stability and security of power system operation without increasing the operating cost [3].

Many research works have been focusing on improving the accuracy of wind power forecast, and several forecast models have been proposed, which can be approximately classified into two categories: physical models and statistical models [4]. Physical models are usually referred to as meteorological prediction of wind power, which is related to the numerical approximation of the models that describe the state of atmosphere [5]. Unlike physical models, statistical models are established by discovering the relationship of historical data, whose advantages for wind power forecast are the simplicity of model construction, reduced computational requirement, and better performance than physical models in short-term forecast [6]. Therefore, in order to provide a more accurate guidance for power system real-time operation, a short-term forecast model belonging to the statistical models is proposed in this paper.

In order to gain more accurate forecast results, it is crucial to reduce the impact of strong volatility of wind power. To achieve this, a number of methods have been applied to extract the mean trend from the original time series. Empirical mode decomposition (EMD) [7], which is a typical approach among them, decomposes the original time series into multiple intrinsic mode functions (IMFs) and considers the first IMF or the first several IMFs as noise [8]. However, this practice is based on experience without theoretical guidance. In this paper, SSA, which is a powerful technique for
nonlinear time series analysis [9] [10], is applied to extract the mean trend from the original time series. The main advantage of SSA compared to other methods is that each component obtained by SSA has its own contribution rate, which makes the decomposition more meaningful.

According to the previous research, local prediction outperforms global prediction [11]. Thus, the SOM [12] [13] is used to select the similar segments of the forecast segment. The SOM is an unsupervised competitive learning algorithm which has been generally applied to vector quantization, data visualization tasks and clustering, thus, it is adopted in this research.

In this paper, a novel forecast model based on SSA and SOM for short-term wind power prediction is proposed. SSA is utilized to extract the mean trend from the original time series. After that, SOM is used to select the similar segments from the mean trend to participate in local prediction. Finally, the similar segments are trained by SVR [14].

2. Methodologies

2.1. Singular Spectrum Analysis (SSA)

SSA is an effective way to study the meaningful features of nonlinear time series. It extracts the mean trend from the original time series. The standard SSA is detailed in the following steps.

**Step 1:** Embedding. Given an original time series \( y_1, y_2, \ldots, y_N \) of length \( N \) and a parameter \( L \) as the embedding dimension of SSA, the original time series \( y \) is transformed to \( L \)-lagged vectors \( x_1, x_2, \ldots, x_K \), and each \( x_i = [y_{i}, y_{i+1}, \ldots, y_{i+L-1}]^T \), \( i = 1, 2, \ldots, K \) is the expression of each \( L \)-lagged. The trajectory matrix \( X \in \mathbb{R}^{L \times K} \) is expressed as

\[
X = [x_1, x_2, \ldots, x_K] = \begin{bmatrix}
y_1 & y_2 & \cdots & y_K \\
y_2 & y_3 & \cdots & y_{K+1} \\
\vdots & \vdots & \ddots & \vdots \\
y_L & y_{L+1} & \cdots & y_N
\end{bmatrix}. \tag{2}
\]

**Step 2:** Singular values decomposition (SVD). \( d \) eigenvalues which meet the condition \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0 \) are produced by computing the eigenvalues of the covariance matrix \( S = XX^T \). \( u_1, u_2, \ldots, u_d \) are the corresponding orthogonal eigenvectors. Then the trajectory matrix \( X \) is decomposed by SVD into

\[
X = e_1 + e_2 + \cdots + e_d
\]

\[
e_i = \sqrt{\lambda_i} u_i v_i^T, v_i = X^T u_i \sqrt{\lambda_i}
\]

\[
d = \min\{L, K\}, i = 1, 2, \ldots, d
\]

where the vectors \( v_1, v_2, \ldots, v_d \) are regarded as the principle components of trajectory matrix \( X \). Therefore, the collection \((\sqrt{\lambda_i}, u_i, v_i)\) is defined as the \( i \)th triple feature vector of matrix \( X \).

**Step 3:** Grouping. The indices \( \{1, 2, \ldots, d\} \) are divided into \( m \) independent groups \( I_1, I_2, \ldots, I_m \) and each group \( I_j = \{I_{j,1}, \ldots, I_{j,r_j}\}, j = 1, 2, \ldots, m \) is denoted as a group of indices. The trajectory matrix \( X \) is expressed as

\[
X = E_{I_1} + E_{I_2} + \cdots + E_{I_m}
\]

\[
E_{I_j} = e_{I_{j,1}} + e_{I_{j,2}} + \cdots + e_{I_{j,r_j}}
\]

where the contribution rate of \( E_{I_j} \) is \( \sum_{i \in I_j} \lambda_i / \sum_{i = 1}^d \lambda_i \). In this paper, the original time series is decomposed into two components. Hence, we have \( X = E_{I_1} + E_{I_2} \).
Step 4: Diagonal averaging. Each matrix $E_{ij} \in \mathbb{R}^{L \times K}$ is converted to a time series by the following ways. Let $z_{il}, i=1,2,\cdots,L; l=1,2,\cdots,K$ be the elements of the matrix with $E_{ij}$. Set $L^* = \min\{L,K\}$, $K^* = \max\{L,K\}$, and $N = K + L - 1$. $z_{il}^* = z_{il}$ when $L < K$; otherwise, $z_{il}^* = z_{il}$. The matrix $E_{ij}$ is transformed into a time series $\left[ y_1^{(j)} y_2^{(j)} \cdots y_N^{(j)} \right]$ via the following equation

$$
\begin{align*}
 y_k^{(j)} = \begin{cases} 
 \frac{1}{k} \sum_{q=1}^{k+1} z_{k+q-k-2,1}^*, & 1 \leq k \leq L^* \\
 \frac{1}{L^*} \sum_{q=1}^{L^*} z_{k+q-k-2,L^*}^*, & 1 \leq k \leq K^* \\
 \frac{1}{N-K+2} \sum_{q=k-K^*+2}^{N-k+1} z_{k+q-k-2,K^*}^*, & K^* < k \leq N, 
\end{cases}
\end{align*}
$$

(5)

In this paper, the diagonal averaging is applied to $E_i$ and $E_j$, respectively, and the two corresponding time series are $y^{(1)} = \left[ y_1^{(1)} y_2^{(1)} \cdots y_N^{(1)} \right]$ and $y^{(2)} = \left[ y_1^{(2)} y_2^{(2)} \cdots y_N^{(2)} \right]$ according to (5). Thus, the original time series $y$ is expressed by $y = y^{(1)} + y^{(2)}$, where $y^{(1)}$ is the mean trend and $y^{(2)}$ is treated as noise.

An example is given in Figure 1, where the mean trend is extracted from a wind power time series collected from Elia, the transmission system operator (TSO) of Belgium [15], and the contribution rate of each eigenvector is demonstrated in Table 1.

![Figure 1](image.png)

**Figure 1.** A time series of wind speed from the Elia database and the mean trend extracted by SSA

**Table 1.** The contribution rate of each eigenvector in embedding dimension $L=20$

| Contribution rate (%) | (1) | (2) | (3) | (4) | (5)-(20) |
|-----------------------|-----|-----|-----|-----|---------|
| 98.438                | 1.153 | 0.194 | 0.074 | 0.141 |

2.2. Self-Organizing Maps (SOM)

Considering the continuity characteristic of meteorology, wind power at a certain time instant is closely related to data collected in a short period of time before. Thus, instead of using all the sample segments to train the model, it is more effective to use only the segments that bear the same trend of the forecast segment. The process of SOM is explained in 4 steps as follows.

Step 1: Initialization. The mean trend $y^{(1)}$ is reconstructed into a higher-dimensional phase space with the embedding dimension of $s$ and the time constant $\tau$. It can be reconstructed into $N - (s - 1) \tau$ segments and each segment is expressed as

$$ y_i^{(1)} = \left[ y_1^{(1)} y_2^{(1)} \cdots y_{i+(s-1)\tau}^{(1)} \right] $$

(6)
where $i = 1, 2, \cdots, N - (s - 1)\tau$.

Initialize the parameters connection weights vector $\mathbf{w}_j(0) = [w_{j1}(0) \ w_{j2}(0) \ \cdots \ w_{jD}(0)], \ j = 1, \cdots, D$ of $D$ neurons, learning rate $\eta_0$ and radius of the neighborhood $\sigma_0$ with small random values.

**Step 2**: Competition. For each input pattern, the neurons compute their respective values of a discriminant function, which provides the basis for competition. The particular neuron with the smallest value of the discriminant function is the winner. In this paper, the discriminant function is the squared Euclidean distance between the mean trend $y_i^{(i)}$ and the weight vector $\mathbf{w}_j(k), k = 1, 2, \cdots, T, \ \text{for each neuron } j$

$$d_j(y_i^{(i)}) = \|y_i^{(i)} - \mathbf{w}_j(k)\|$$

(7)

**Step 3**: Cooperation. The winner neuron determines the spatial location of a topological neighborhood neurons, thereby providing the basis for cooperation among neighboring neurons. The topological neighborhood $T_{j,j'(y_i^{(i)})}$ is defined as

$$T_{j,j'(y_i^{(i)})} = \exp\left(-\frac{S_j^{(y_i^{(i)})} / 2}{2\sigma^2}\right)$$

(8)

$$\sigma(k) = \sigma_0 \exp(-k / T), \ k = 1, 2, \cdots T$$

(9)

where $I(y_i^{(i)})$ is the index of the winner neuron, $S_{j,j'(y_i^{(i)})}$ is the distance between the connection $\mathbf{w}_j(k)$ and $\mathbf{w}_{j'(y_i^{(i)})}(k)$, $\sigma(k)$ is radius of the neighbourhood at step iteration $k$.

**Step 4**: Adaptation. The weight vector $\mathbf{w}_j(k)$ of the current winner neuron and the learning rate $\eta(k)$ are simultaneously adjusted according to the following learning rule

$$\Delta \mathbf{w}_j = \eta(k) \cdot T_{j,j'(y_i^{(i)})}(k) \cdot (y_i^{(i)} - w_j(k))$$

(10)

$$\eta(k) = \eta_0 \exp(-k / T), \ k = 1, 2, \cdots T$$

(11)

An instruction on how to implement the proposed model is shown in the framework of Figure 2.

3. Simulation Studies

3.1. Performance qualification

In order to assess the performance of the proposed model, two commonly used criteria are employed to measure the errors: normalized mean absolute error (NMAE) and normalized root mean squared error (NRMSE) [16]. The NMAE indicates how close the forecast results are to the actual values, and the NRMSE evaluates the standard deviation between the forecast results and the actual values, which reflects the stability of the forecast model. The smaller the NMAE/NRMSE is, the more accurate and stable the forecast results are. They are defined by:

$$\text{NMAE} = \frac{1}{N} \sum_{i=1}^{N} \left| y_i - \hat{y}_i \right| \times 100\%$$

(12)

$$\text{NRMSE} = \frac{1}{Y} \left( \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \right)^{1/2} \times 100\%$$

(13)

where $N$ is the size of the validation data set, $\hat{y}_i$ is the forecast value, $y_i$ is the actual value, $Y$ is the installed capacity of the wind farm at the forecasting moment.
3.2. Forecast results and discussion
To evaluate the performance of the SSA-SOM-SVR model, comparison studies are carried out on the persistence (Per.), SVR and SSA-SVR models for 1~20 look-ahead steps, and the performances have been illustrated in Figure 3. It is obvious that the SSA-SOM-SVR model achieves the smallest NMAEs and NRMSEs. Furthermore, NMAE and NRMSE are improved by 12.38% and 13.90% in 20 look-ahead steps, respectively, by the SSA-SOM-SVR model over the Per. model. To make a more comprehensive study, the numerical results of 4, 12, 20 look-ahead steps are listed in Table 2. As wind may have various natural characteristics in different seasons, the forecast is carried out in March, June, September and December, respectively. It is obvious that the proposed model gains the smallest NMAEs and NRMSEs for all situations, which indicates that the proposed model outperforms the other ones in both accuracy and stability.

![Figure 3. The performance of the four models on the Elia data set in June (a) NMAEs (b) NRMSEs](image)

| Table 2. Performance evaluation of the four forecast models on the Elia data set (%) |
|-----------------------------------------------|-----------|-----------|-----------|-----------|
| month | Per | SVR | SSA-SVR | SSA-SOM-SVR |
| 4-step | | | | |
| NMAE | 2.6018 | 2.5433 | 1.3322 | 1.0318 |
| NRMSE | 3.6154 | 3.5836 | 1.8750 | 1.4994 |
| 12-step | | | | |
| NMAE | 5.1307 | 5.1524 | 4.2830 | 3.8897 |
| NRMSE | 6.9956 | 6.9042 | 5.8747 | 5.2941 |
| 20-step | | | | |
| NMAE | 7.0272 | 6.9700 | 6.2935 | 6.1695 |
| NRMSE | 9.5660 | 9.1924 | 8.3764 | 8.2779 |
4. Conclusion
This paper proposes a forecast model based on SSA, SOM and SVR for short-term wind power prediction. The SSA is employed to extract the mean trend from the original time series. Afterwards, SOM is used to select similar segments to perform local prediction. In order to evaluate the accuracy of the proposed model and highlight the effect of SSA and SOM methods, simulation studies are conducted on the Per., SVR and SSA-SVR models for the purpose of comparison, and the final results prove that the application of SSA and SOM improve the accuracy and stability.

Acknowledgments
This work was supported by Guangdong Innovative Research Team Program (No. 201001N0104744201) and the State Key Program of National Natural Science of China (No. 51437006).

References
[1] Zhang J, Mingjian Cui, Bri-Mathias Hodge, et al. Ramp forecasting performance from improved short-term wind power forecasting over multiple spatial and temporal scales[J]. Energy, 2017, 122:528-541.
[2] Wu J L, Ji T Y, Li M S, et al. Multi-step wind power forecast based on similar segments extracted by mathematical morphology[C].Power and Energy Engineering Conference. IEEE, 2015:1-6.
[3] Zhang C, Wei H, Zhao X, et al. A Gaussian process regression based hybrid approach for short-term wind speed prediction[J]. Energy Conversion & Management, 2016, 126:1084-1092.
[4] Foley A M, Leahy P G, Marvuglia A, et al. Current methods and advances in forecasting of wind power generation[J]. Renewable Energy, 2012, 37(1):1-8.
[5] Liu H, Shi J, Erdem E. Prediction of wind speed time series using modified Taylor Kriging method[J]. Energy, 2010, 35(12):4870-4879.
[6] Akçay H, Filik T, Yan J. Short-term wind speed forecasting by spectral analysis from long-term observations with missing values[J]. Applied Energy, 2017, 191:653-662.
[7] Guo Z, Zhao W, Lu H, et al. Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model[J]. Renewable Energy, 2012, 37(1):241-249.
[8] Boudraa A O, Cexus J C. EMD-Based Signal Filtering[J]. IEEE Transactions on Instrumentation & Measurement, 2007, 56(6):2196-2202.
[9] Zhang X, Wang J, Zhang K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by cuckoo search algorithm[J]. Electric Power Systems Research, 2017, 146:270-285.
[10] Zhang W, Su Z, Zhang H, et al. Hybrid Wind Speed Forecasting Model Study Based on SSA and Intelligent Optimized Algorithm[J]. Abstract and Applied Analysis,2014,(2014-4-3), 2014, (2014):1-14.
[11] Lau K W, Wu Q H. Local prediction of non-linear time series using support vector regression[J]. Pattern Recognition, 2008, 41(5):1539-1547.
[12] Liu J, Djurdjanovic D. Topology Preservation and Cooperative Learning in Identification of Multiple Model Systems[J]. IEEE Transactions on Neural Networks, 2008, 19(12):2065-72.
[13] Júnior A H S, Barreto G A, Corona F. Regional models: A new approach for nonlinear system identification via clustering of the self-organizing map[J]. Neurocomputing, 2015, 147:31-46.
[14] Zhu L, Wu Q H, Li M S, et al. Support vector regression-based short-term wind power prediction with false neighbours filtered[C]. International Conference on Renewable Energy Research and Applications. IEEE, 2014:740-744.
[15] Elia. Wind power data [online]. Available:http://www.elia.be/en/griddata/power-generation/wind-power.
[16] Bludszuweit H, Dominguez-Navarro J A, Llombart A. Statistical Analysis of Wind Power Forecast Error[J]. IEEE Transactions on Power Systems, 2008, 23(3):983-991.