2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)
Melbourne, Australia, Nov 28 - Dec 1, 2016

Short-Term Wind Power Prediction Based on Intrinsic Time-Scale Decomposition and LS-SVM

L. L. Zhang*, M. S. Li*, *Member, IEEE*, T. Y. Ji*, *Member, IEEE*, Q. H. Wu*†, *Fellow, IEEE*
*School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, Guangdong, China.
†Department of Electrical Engineering and Electronics, The University of Liverpool, Liverpool, L69 3GJ, UK.
Corresponding author: Prof. Q. H. Wu; Email: wuqh@scut.edu.cn.

***Abstract*—This paper proposes a wind power prediction method based on intrinsic time-scale decomposition (ITD) and least square support vector machine (LS-SVM) to improve the accuracy of wind power forecast. The proposed method employs ITD as a preprocessing method to decompose wind power data into a set of proper rotation components and a monotonous baseline signal. Afterwards, the backward difference of each component is used as input of the LS-SVM model for training and prediction. Simulation studies are carried out on wind power data to evaluate the performance of the proposed method, and the results have shown that, by introducing the ITD, the proposed method outperforms the original LS-SVM method.**

***Index Terms*—intrinsic time-scale decomposition; wind power forecast; LS-SVM.**

I. Introduction

The capacity of wind power generation rapidly increased in many countries in the past decades and wind power will play a more important role in the future energy supply [1]. However, wind generation is highly dependent on wind speed at the wind turbine sites that results in high variability and uncertainty in the generated wind power [2]. Due to the intermittency and expected high penetration rate of wind power generation, the integration of wind power to electricity system brought some important challenges [3]. In order to operate wind farms as controllable plants, accurate and reliable techniques for wind farms output power prediction is much desired [4].

Different methods have been presented to forecast wind speed and the power produced by wind farms, which can be classified into two categories: physical methods and statistical methods [5]. Physical methods extrapolate the meteorological forecasts at the desired location and at turbine hub height, and use the manufacturer's power curve for estimating the wind power output [6]. Statistical methods require only the historical and current data to perform timely wind power prediction. For longer-term prediction, physical methods usually work better than statistical methods, however, they require more information and are much more complicated in computing. This paper focus on the statistical methods as they are more suitable for short-term prediction.

The commonly used statistical wind power prediction models include persistence models, autoregressive moving average (ARMA) model [7], kalman filters [8], grey predictors [9], [10], artificial neural networks (ANNs) [11], [12] and support vector machines (SVMs) [13], [14], etc. Each of these methods has its own advantages and shortcomings. In order to improve the accuracy of wind power prediction, different combinations of above methods have also been proposed [15]–[17]. Signal decomposition method like wavelet transform (WT) and empirical mode decomposition (EMD), as a processing tool, are also introduced to wind power prediction because of the ability for nonstationary signal approximation and frequency analysis. In [18]–[20], WT is applied to process wind data firstly and then respectively uses ARMA , ANN and SVM to predict wind power. The results demonstrated that the combined methods outperform the methods using original data directly to predict wind power. In [21], EMD is used to decompose wind power sequence into several intrinsic mode functions (IMFs) and a residual component, and then the SVM is used to build a forecast model for each component. The simulation results indicate that the hybrid model is superior to the model using SVM only. In [22], a forecasting model based on a mean trend detector and a local predictor is proposed to undertake short-term forecast using least square support vector machine (LS-SVM).

The intrinsic time-scale decomposition (ITD) [23] is an alternative method specifically formulated for nonlinear or non-stationary signal analysis. It overcomes the disadvantages of EMD such as boundary distortions and unidentified low-frequency components. In this paper, a statistical wind power prediction model which combines ITD and LS-SVM is presented. The backward differences of the proper rotation components and baseline signal obtained by ITD are used to train LS-SVM forecast model, respectively. This paper is organised as follows. In Section II, the foundations of ITD and LS-SVM are introduced. In Section III, the proposed wind power prediction method is detailed. Section IV describes the data used and conducts simulation studies on the data to evaluate the forecast performance of the proposed method. Finally, the conclusion is drawn in Section V.

II. Methodology Background

A. Intrinsic Time-Scale Decomposition

The ITD decomposes a complex signal into a set of proper rotation components (PRCs) and one monotonous trend. It constructs the piece-wise linear baseline signal between suc-

This project was supported by the China Postdoctoral Science Foundation (No. 2015M582383) and the Guangdong Innovative Research Team Program (No. 201001N0104744201).

cessive extrema and computes the instantaneous amplitude and frequency based on the piece-wise wave of each PRC [24].

Given a signal, X_t, we define an operator Λ, which extracts a baseline signal from X_t in a manner that causes the residual to be a proper rotation. More specifically, X_t can be decomposed as

$$X_t = \Lambda X_t + (1-\Lambda)X_t = L_t + H_t, \tag{1}$$

where $L_t = \Lambda X_t$ is the baseline signal and $H_t = (1-\Lambda)X_t$ is a proper rotation.

Suppose $\{X_t, t \geq 0\}$ is a real-valued signal, and denote the local extrema of X_t by $\{\tau_k, k = 1, 2, \dots\}$, and for convenience define $\tau_0 = 0$. In the case of intervals on which X_t is constant, but which contain extrema due to neighbouring signal fluctuations, τ_k is chosen as the right endpoint of the interval. To simplify notation, let X_k and L_k denote $X(\tau_k)$ and $L(\tau_k)$, respectively.

Suppose that H_t and L_t have been defined on $[0, \tau_k]$ and that X_t is available for $t \in [0, \tau_{k+2}]$. We can then define a (piece-wise linear) baseline-extracting operator, L, on the interval $(\tau_k, \tau_{k+1}]$ between successive extrema as follows:

$$\begin{aligned} L_t = \Lambda X_t &= L_k + (\frac{L_{k+1} - L_k}{X_{k+1} - X_k})(X_t - X_k) \\ &= L_t + H_t, t \in (\tau_k, \tau_{k+1},) \end{aligned} \tag{2}$$

where

$$\begin{aligned} L_{k+1} = &\alpha[X_k + (\frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k})(X_{k+2} - X_k)] \\ &+(1-\alpha)X_{k+1}, \end{aligned} \tag{3}$$

and $0 < \alpha < 1$ is typically fixed with $\alpha = 1/2$.

After defining the baseline signal according to equations (2) and (3), the proper-rotation-extracting operator, H, can be defined as:

$$HX_t \equiv (1-\Lambda)X_t = H_t = X_t - L_t. \tag{4}$$

Note that L_t and H_t respectively correspond to the components with higher frequencies and lower frequencies in the input signal. Therefore, a monotonic baseline signal and series of PRCs can be finally obtained by using the baseline signal as the input signal repeatedly. The instantaneous frequencies of the PRCs are successively decreasing at each subsequent level of the decomposition. More precisely,

$$\begin{aligned} X_t &= HX_t + \Lambda X_t = HX_t + (H+\Lambda)\Lambda X_t \\ &= (H + H\Lambda + \Lambda^2)X_t \\ &= (H + H\Lambda + (H+\Lambda)\Lambda^2)X_t \\ &= (H \textstyle\sum_{k=0}^{p-1} \Lambda^k + \Lambda^p)X_t \end{aligned} \tag{5}$$

B. Least Square Support Vector Machine

The LS-SVM model [25], is proposed for solving problems where priori knowledge is not available for the studied system and which is typically nonlinear and non-stionary.

For a given training dataset $\{\mathbf{x}_i, y_i\}_{i=1}^{I}$ with $\mathbf{x}_i$ as a d-dimensional input vector and y_i as the corresponding output, the regression model in the primal weight space for LS-SVM can be expressed in the form:

$$y = \omega^{\mathbf{T}} \varphi(\mathbf{x}) + b, \tag{6}$$

where ω is the weight vector, b is the bias, and $\varphi(\cdot)$ is the mapping to the higher dimensional feature space. In order to obtain ω and b, the following optimization problem to be solved is as:

$$\min \mathcal{J}(\omega, b, e) = \frac{1}{2}\omega^{\mathbf{T}}\omega + \frac{1}{2}\gamma \sum_{i=1}^{I} e_i^2, \tag{7}$$

subject to the equality constraint

$$y_i[\omega^{\mathbf{T}}\varphi(\mathbf{x}_i) + b] = 1 - e_i, i = 1, 2, \dots, I,$$

where γ is the regularization parameter, e_i is the regression error variable and $\mathcal{J}$ is the cost function which minimizes the error.

The Lagrangian function for the optimization problem can be defined as

$$\mathcal{L}(\omega, b, e, a) = \mathcal{J}(\omega, e) - \sum_{i=1}^{N} a_i[\omega^{\mathbf{T}}\varphi(\mathbf{x}_i) + b + e_i - y_i], \tag{8}$$

where $a = [a_1, a_2, \dots, a_N]$ represents the Lagrangian multipliers. Solving (8), the regression model can be obtained as

$$\hat{y} = \sum_{i=1}^{I} a_i K(\mathbf{x}, \mathbf{x}_i) + b, \tag{9}$$

where $K(\cdot,\cdot)$ is the Kernel function. In this paper, RBF Kernel is employed.

III. Wind Power Prediction Method

A. Prediction Model

This paper proposed a wind power prediction method based on the ITD and LS-SVM, the flowchart of which is shown in Fig. 1. The steps are as follows:

1) Implement the ITD decomposition on the original wind power sequence and obtain the baseline signal L and proper rotation components $H_i, i = 1, 2, \cdots, m$.

2) Calculate the backward difference of the baseline signal, denoted by d_L, as

$$d_L(k) = L(k) - L(k-1) \tag{10}$$

and the backward difference of the proper rotation component H_i, denoted by d_{H_i}.

$$d_{H_i}(k) = H_i(k) - H_i(k-1) \tag{11}$$

3) Build a training data set for d_L and d_{H_i}, respectively, and set up the corresponding SVM prediction model with the selection of the optimal kernel functions and parameters to get the predicted value of each component, $\hat{d}_L$ and $\hat{d}_{H_i}$.

4) Calculate the p steps ahead predicted wind power as follows

$$\begin{aligned} \hat{x}(k+p) = &L(k) + \textstyle\sum_{j=1}^{p} \hat{d}_L(k+j) \\ &+ \textstyle\sum_{i=1}^{m}(H_i(k) + \sum_{j=1}^{p} \hat{d}_{H_i}(k+j)) \end{aligned} \tag{12}$$

5) Calculate and analyze the forecast error.

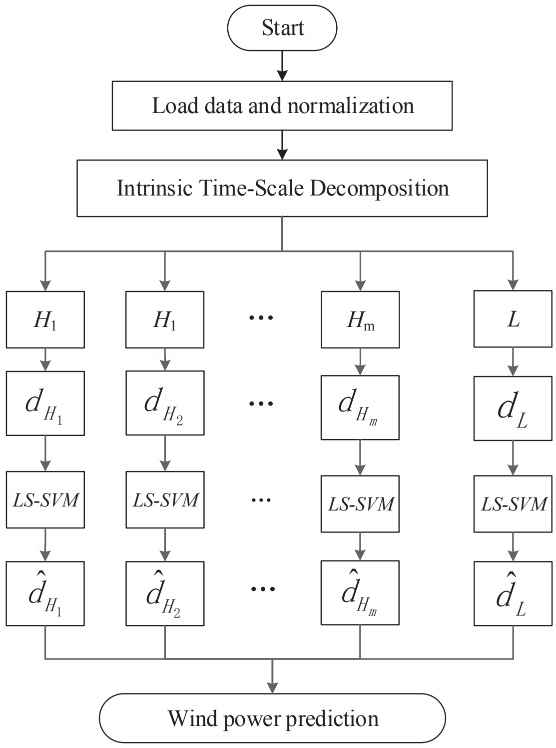

Fig. 1. The flowchart of the proposed model.

B. Performance Evaluation

Three indices, mean absolute percentage error (MAPE), the normalized root mean squared error (NRMSE) and normalized maximum absolute error (NMAE), are together applied to evaluated the performance of the forecast model, of which the definitions are as follows:

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} \times 100, \tag{13}$$

$$\text{NRMSE} = \frac{1}{P_{\text{inst}}}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \times 100}, \tag{14}$$

$$\text{NMAE} = \max\{\frac{|y_i - \hat{y}_i|}{P_{\text{inst}}}, i = 1, 2, \cdots, N\} \times 100 \tag{15}$$

where P_{inst} is the installed capacity of the wind farm at the forecasting moment, N is the size of dataset, $\hat{y}_i$ is the predicted value, y_i is the actual value and $\bar{y}$ is the mean of the actual values.

MAPE is a measure of accuracy in trend estimation. NRMSE assesses the quality of prediction in terms of variation and degree of bias. NMAE represents the margin of the error. The smaller the there indices are the better the model performs.

In addition to the indices above, a visual index is introduced to provide an intuitive comparison. For an ideal forecast model, there exists an ideal linear relationship between the predicted values and the actual values as follows:

$$\hat{y}_i = y_i. \tag{16}$$

This means that the best forecast model generates the predicted values, which are exactly equal to the actual measured values. However, it is a fact that the ideal linear relationship is beyond realization. Therefore, the mean value $\overline{d}$ of the distances between the scattered points and the line $\hat{y} = y$ is also introduced to qualify the forecast performance, together with the standard deviation of the distances σ.

IV. Simulation Studies

A. Data Description

The data of wind power generation used for simulation studies were collected from the public database of AESO [26], which were sampled once every 10 minutes. A three days segment of the wind power sequence, which consists of a total of 432 data samples was used to validate the performance. The first 144 samples, which are the data of the former day, are called the training data and are used for building the forecast model, and the rest are used to validate the accuracy of the prediction.

B. Wind Power Prediction by the Proposed Method

The training data sequence is firstly decomposed by ITD and the results is depicted in Fig. 2. It can be seen that the volatilities of obtained components are significantly lower than the original data. The PRCs become more and more stable from H_1 to H_4, and the last residual component is monotonous. The 1-step ahead wind power prediction results using the proposed model are shown in Fig. 3, where the normalized error is defined as: $error/P_{\text{inst}} \times 100\%$. For all the 288 testing samples in two days, the predicted wind power is very close to the actual wind power. The normalized errors of most samples fall between -2% and 2%.

C. Comparison Studies

For the purpose of comparison, the multi-step prediction performance of the proposed method and original LS-SVM model were also studied using the same data. The results are listed in Table I. Compared to the original LS-SVM model, the MAPE, NRMSE and NMAE of the proposed model are much smaller, which means that the proposed model provides more accurate and stable forecast. The results indicate that the performance of the original LS-SVM based forecast model can be improved more than 10% by introduced ITD as a processing method.

Furthermore, the visual index is investigated as well and the results of the 2-step ahead prediction are shown in Fig. 4. The line $\hat{y} = y$ means that the predicted power is exactly the same as the measured power. The scattered points depicted in green star represent the predicted power by LS-SVM model against the measured power while the scattered points depicted in red circle represent the predicted power by the proposed model against the measured power. As can be seen, there are several green

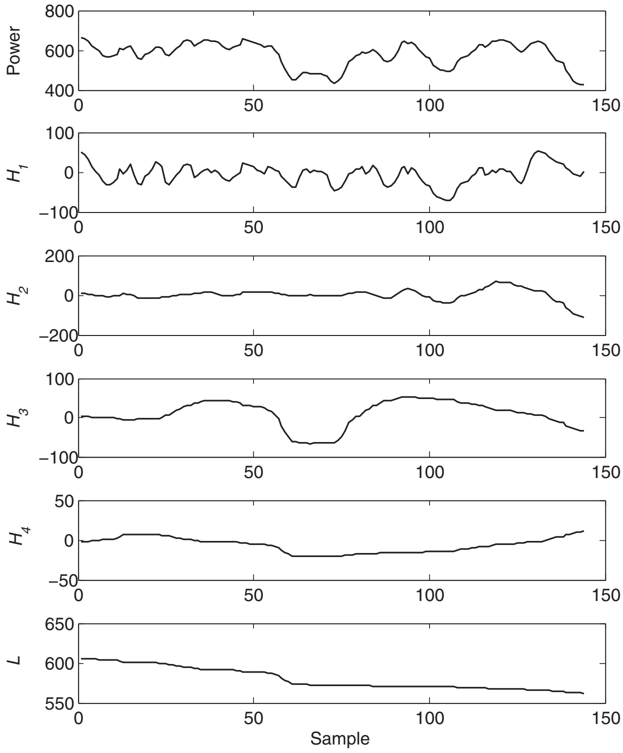

Fig. 2. The ITD decomposition results of wind power sequence.

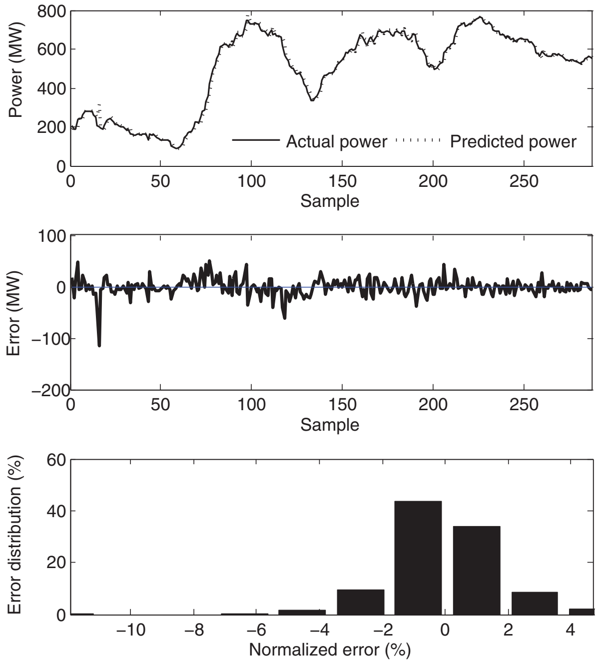

Fig. 3. Wind prediction using the proposed method (1-step adead).

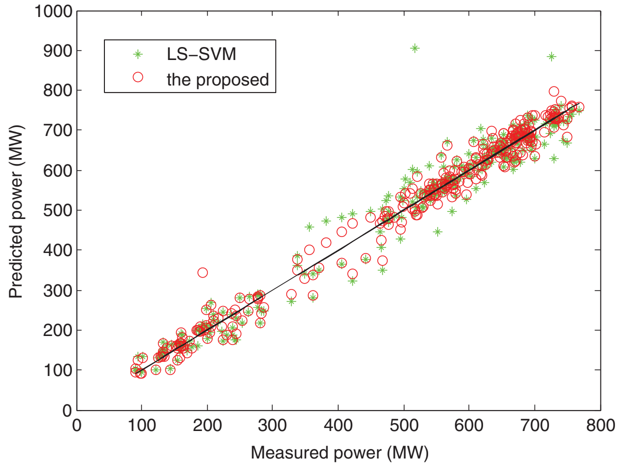

Fig. 4. Camparision of the proposed method and original LS-SVM model (2-steps ahead).

TABLE I
MAPE, NRMSE AND NMAE OF DIFFERENT MODELS

	p-step	LS-SVM	Proposed	Improvement
MAPE	2	6.7451	5.4919	18.58%
	4	10.9296	9.3231	14.70%
	6	14.5666	12.6635	13.06%
NRMSE	2	4.6758	3.1482	32.67%
	4	6.4381	5.1607	19.84%
	6	8.3693	7.1023	15.14%
NMAE	2	42.7040	16.7013	60.89%
	4	36.6125	20.0549	45.22%
	6	31.6801	25.4273	19.73%

star points deviating far from the line, which means that the LS-SVM is unstable. In contrast, for the proposed model, the red circle points lean closer to the line. The results of the other two cases, 4-steps and 6-steps ahead prediction, are together shown in Table II. From the table we can see that the mean value $\overline{d}$ and the standard deviation σ of the distances between the scattered points by the proposed method are much smaller. This means that the predicted values of the proposed method track the actual values better.

V. Conclusion

In this paper, a hybrid ITD-SVM wind power prediction method has been established. The introduction of ITD method and the backward difference operation ensure the inputs of LS-SVM are stationary, which benefits the prediction. Simulation studies have been carried out on wind power generation data from a public database. The results have demonstrated that the proposed method outperforms the original LS-SVM method in terms of accuracy and stability.

TABLE II
$\overline{d}$ AND σ OF DIFFERENT MODELS

	p-step	LS-SVM	Proposed	Improvement
$\overline{d}$	2	26.8184	20.6890	22.85%
	4	41.8913	35.2991	15.74%
	6	56.1809	49.3751	12.11%
σ	2	32.7352	19.5815	40.18%
	4	40.4826	30.5644	24.50%
	6	50.7759	41.1233	19.01%

REFERENCES

[1] B. Ernst, B. Oakleaf, M. L. Ahlstrom, M. Lange, C. Moehrlen, B. Lange, U. Focken, and K. Rohrig. Predicting the wind. *IEEE Power and Energy Magazine*, 5(6):78–89, Nov 2007.

[2] A. Ouammi, H. Dagdougui, and R. Sacile. Short term forecast of wind power by an artificial neural network approach. In *Systems Conference*, pages 1 – 5, 2012.

[3] M. Gholami, G. B. Gharehpetian, and M. Davari. One day ahead prediction of wind speed based on power and exponential. In *Power Engineering and Optimization Conference*, pages 449 – 452, 2010.

[4] T. H. M. El-Fouly, E. F. El-Saadany, and M. M. A. Salama. One day ahead prediction of wind speed and direction. *IEEE Transactions on Energy Conversion*, 23(1):191–201, March 2008.

[5] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh. Current methods and advances in forecasting of wind power generation. *Renewable Energy*, 37(1):1–8, 2012.

[6] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, et al. Wind power forecasting: state-of-the-art 2009. Technical report, Argonne National Laboratory (ANL), 2009.

[7] J. Zhang and C. Wang. Application of arma model in ultra-short term prediction of wind power. In *Computer Sciences and Applications (CSA), 2013 International Conference on*, pages 361–364, Dec 2013.

[8] F. Cassola and M. Burlando. Wind speed and wind energy forecast through kalman filtering of numerical weather prediction model output. *Applied Energy*, 99(6):154–166, 2012.

[9] T. H. M. El-Fouly, E. F. El-Saadany, and M. M. A. Salama. Grey predictor for wind energy conversion systems output power prediction. *IEEE Transactions on Power Systems*, 21(3):1450–1452, 2006.

[10] T. H. M. El-Fouly, E. F. El-Saadany, and M. M. A. Salama. Improved grey predictor rolling models for wind power prediction. *IET Generation Transmission & Distribution*, 1(6):928–937, 2007.

[11] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret. Wind power forecasting using advanced neural networks models. *IEEE Transactions on Energy Conversion*, 11(4):762–767, 1996.

[12] M. A. Mohandes, S. Rehman, and T. O. Halawani. A neural networks approach for wind speed prediction. *Renewable Energy*, 13(3):345–354, 1998.

[13] J. Y Zhou, J. Shi, and G. Li. Fine tuning support vector machines for short-term wind speed forecasting. *Energy Conversion and Management*, 52(4):1990–1998, 2011.

[14] Y. Wang, D. L. Wu, C. X. Guo, Q. H. Wu, W. Z. Qian, and J. Yang. Short-term wind speed prediction using support vector regression. In *IEEE PES General Meeting*, pages 1–6, July 2010.

[15] H. W. Peng, F. R. Liu, and X. F. Yang. A hybrid strategy of short term wind power prediction. *Renewable Energy*, 50:590–595, 2013.

[16] J. P. S. Catalao, H. M. I. Pousinho, and V. M. F. Mendes. Hybrid intelligent approach for short-term wind power forecasting in portugal. *IET Renewable Power Generation*, 5(3):251–257, 2011.

[17] H. Liu, H. Q. Tian, and Y. F. Li. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Applied Energy*, 98(1):415–424, 2012.

[18] J. W. Zeng and W. Qiao. Short-term wind power prediction using a wavelet support vector machine. *IEEE Transactions on Sustainable Energy*, 3(2):255–264, 2012.

[19] L. Cao and R. Li. Short-term wind speed forecasting model for wind farm based on wavelet decomposition. In *International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, pages 2525–2529, 2008.

[20] R. R. B. De Aquino, M. M. S. Lira, J. B. De Oliveira, M. A. Carvalho, O. N. Neto, and G. J. De Almeida. Application of wavelet and neural network models for wind speed and power generation forecasting in a brazilian experimental wind park. In *Proceedings of the 2009 international joint conference on Neural Networks*, pages 172–178, 2009.

[21] W. Zhang, F. Liu, X. Zheng, and Y. Li. A hybrid EMD-SVM based short-term wind power forecasting model. In *Power and Energy Engineering Conference*, 2015.

[22] J. L. Wu, T. Y. Ji, M. S. Li, P. Z. Wu, and Q. H. Wu. Multistep wind power forecast using mean trend detector and mathematical morphology-based local predictor. *IEEE Transactions on Sustainable Energy*, 6(4):1216–1223, Oct 2015.

[23] M. G. Frei. Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals. *Royal Society of London Proceedings*, 463(463):321–342, 2007.

[24] Jinshan Lin. Comparison between intrinsic time-scale decomposition and hilbert-huang transform. In *International Conference on Information Science & Engineering*, pages 5005–5008, 2010.

[25] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, J. A. K. Suykens, and T. Van Gestel. *Least squares support vector machines*, volume 4. World Scientific, 2002.

[26] AESO. Wind power data. http://www.aeso.ca/gridoperations/20544.html.