Renewable Energy 62 (2014) 592–597

Contents lists available at ScienceDirect

Renewable Energy

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

Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm

Da Liu*, Dongxiao Niu, Hui Wang, Leilei Fan

School of Economics and Management, North China Electric Power University, Beijing 102206, China

ARTICLE INFO

Article history:
Received 29 January 2013
Accepted 15 August 2013
Available online 13 September 2013

Keywords:
Wind speed forecasting
SVM
GA
Wavelet transform
Input selection

ABSTRACT

Affected by various environment factors, wind speed presents characters of high fluctuations, autocorrelation and stochastic volatility; thereby it is hard to forecast with a single model. A hybrid model combining with input selected by deep quantitative analysis, Wavelet Transform (WT), Genetic Algorithm (GA) and Support Vector Machines (SVM) was proposed. WT was exploited to decompose the wind speed signal into two components, an approximation signal to maintain the major fluctuations and a detail signal to eliminate the stochastic volatility. SVM were built to model the approximation signal. Autocorrelation and partial correlation were applied to analyze the inner ARIMA Autoregressive Integrated Moving Average (ARIMA) relationship between the historical speeds thus to select the input of SVM from them, and Granger causality test was applied to select input from environment variables by checking the influence of temperature with different leading lengths. The parameters in SVM were fine-tuned by GA to ensure the generalization of SVM. A case study of a wind farm from North China demonstrates that this method outperforms the comparison models.

1. Introduction

In recent years, the wind power, a kind of clean renewable power resource, is widely used and plays more and more important role in energy industry around the world. The wind power is greatly affected by the wind speed, and there exits strong positive correlation between them. Forecasting the wind speed with higher accuracy helps to improve performance and reliability of wind turbines as well as power systems. Wind speed forecasting is getting more and more concerns [1–3].

The mostly used methods of wind speed forecasting can be classified as time series modeling and intelligent algorithm modeling. The former establishes linear or non-linear mapping relation of the historical wind speed to forecast the future speed by digging information contained in the historical signals, such as Auto-Regressive and Moving Average (ARMA) [4] and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) [5,6]. The latter builds a non-linear function of high dimension to fit the historical wind speed and its influence factors usually with minimization the training error, such as Artificial Neural Networks (ANN) [7,8]and Support Vector Machines (SVM) [9,10].

The time series models and intelligent models were improved to forecast wind speed with better performance by combining them with other data processing techniques. In Ref. [11] seasonal Autoregressive Integrated Moving Average (ARIMA) models were validated with a better sensitivity to the adjustment and prediction of the wind speed. In Ref. [12] a fractional-ARIMA was used to forecast wind speeds on the 24 h and 48 h horizons with significant improvements compared to a persistence model.

WT was used to eliminate the irregular fluctuation of the wind speed in Refs. [13,14]. Another method, Empirical Mode Decomposition (EMD) was also applied to decompose the wind speed into several intrinsic mode functions (IMFs) for modeling [15].

As it is well known that the input plays significant role in modeling, however in most cases input of the models was usually selected according to the researchers' experiences, deep quantitative analysis techniques were usually missed. In Ref. [15] input variables in FNN were selected by Partial AutoCorrelation function (PACF) to analyze the relationship between the candidate variables and the wind speed. Self-Organization Map (SOM) was used to cluster the 24 h data for modeling automatically according to their similarities in Ref. [16].

In Ref. [10] SVM with fine tuning parameters outperform the persistence model for one-step ahead wind speed forecasting. In Ref. [17] Particle Swarm Optimization (PSO) and differential evolution were applied to automated specification of ANN and the nearest neighbor search, suggesting an obvious reduction of the prediction

* Corresponding author.
E-mail address: liuda315@163.com (D. Liu).

0960-1481/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.renene.2013.08.011

error compared to the model with standard manually selected variables. SVM were applied to forecast the electricity price with GA adopted to optimize the model's parameters in Ref. [18].

Researches have been tried to combine the above techniques to improve the forecasting of wind speed. Literature [14] proposed a hybrid intelligent approach based on WT and a hybrid of ANN and fuzzy logic for short-term wind power forecasting in Portugal. In Ref. [19] a model using fuzzy logic and ANN was proposed to forecast wind speed. Literature [20] presented an adaptive very short-term wind power prediction using an ANN as a predictor along with adaptive Bayesian learning and Gaussian process approximation. In Ref. [21] ARIMA was used to forecast the wind speed signal and then ANN were applied to model the errors obtained by ARIMA to exploit the nonlinear tendencies that the ARIMA technique could not identify.

Most of the wind speed forecasting methodologies discussed above tried to use one of the techniques, such as signal decomposition and parameter selection to improve the modeling performance; however few considered these techniques compreh ensively. In addition reliable input selection for modeling with numerical investigation was usually missing in previous researches. This paper proposed to build an integrated forecasting model considering all these perspectives.

To evaluate impact factors and determine the lags of wind speed on these factors to choose proper input, not only ACF and PACF analysis was exploited to analyze their relationship, but also Granger Causality Test [22] which has been widely used in econometrics yet no applications shown in wind power forecasting field was proposed to analyze the causality between wind speed and its impact factors of different lags. Meanwhile WT was applied to decompose the original wind speed into approximation signal and detail signal to remove the irregular fluctuation of wind speed for further modeling. Then SVM served as the major part of the proposed modeling to simulate the wind speed. During the simulation the parameters were chosen automatically by GA to reduce the performance volatility of SVM with different parameters. The effectiveness of the hybrid method was examined by a case study from a wind farm located in northern China.

The rest of the paper is organized as follows: Section 2 gives a brief description of the principle of WT, SVM and GA; modeling approaches of the proposed technique are discussed in Section 3, an experiment study is put forward to validate the proposed method in Section 4, and discuss and conclusions are drawn in Section 5.

2. WT, SVM and GA

2.1. WT

WT is an effective method for signal processing. WTs can be divided into two categories: Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). DWT is a kind of WT for which the wavelets are discretely sampled. As with other WTs, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency and location information. In this study we take DWT to analysis the data.

WT usually decomposes a signal into an approximation component and many detail components, where approximation component contains the low-frequency information, the most important part to give the signal its identity; and the detail components to reveal flavor of the signal. Fig. 1 is a wavelet decomposition tree showing the decomposition process. Firstly, the signal $\mathbf{S}$ is decomposed into an approximation component $\mathbf{A}_1$ and a detail component $\mathbf{D}_1$; and then the $\mathbf{A}_1$ is further decomposed into another approximation component $\mathbf{A}_2$ and a detail component $\mathbf{D}_2$ if we

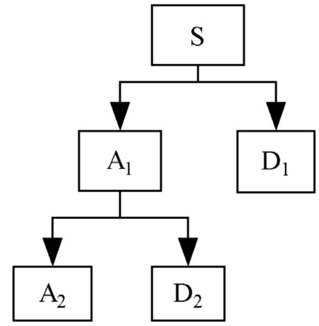

Fig. 1. Wavelet decomposition.

want to analyze the signal with higher level resolution. And so on until reaches a suitable number of levels.

The original wind speed signal is proposed to be decomposed into several components (one approximation component and multiple detail components) to reflect the features of the speed on different levels. The approximation is expected to present the main fluctuation of the wind speed and the details to contain the spikes and stochastic volatilities on different levels. A suitable number of levels can be decided by comparing the similarity between the approximation and the original signal.

2.2. GA

GA was proposed by John Holland et al. at the end of 1960s. It is an algorithm of nonlinear global optimization that inspired by the biological evolution mechanism (survival of the fittest, crossover, mutation, etc.). It is very suitable for the optimization of complex problems for simplicity and robustness, and it has been widely applied in various forecasting and optimization fields. The modeling approach of GA is listed as follows:

- Select a group of random candidate solutions
- Iterate the following steps until reach stop criterions
  - Calculating the fitness of the candidate solutions according to the adaptive condition
  - Producing the next generation according to the principle of proportionate (the one with higher fitness is more inclined to be chosen)
  - Perform a crossover and mutation operation to the candidate solutions and generate new ones
- Return the solutions.

2.3. SVM

SVM modeling is firstly developed by Vladimir N. Vapnik based on statistical learning theory. By mapping the sample space to a high dimensional and even infinite dimensional feature space (Hilbert space), SVM convert a nonlinear separable problem in the sample space to a linear separable problem in Hilbert space. It avoids the defect of overfitting that exists in traditional machine learning and curse of dimensionality. Similar to ANN and other intelligence algorithm, the performance of SVM seriously depends on the input and the parameters.

3. Approaches of W-SVM-GA modeling

In this section, the modeling approaches of the proposed method (W-SVM-GA) will be discussed in detail. The modeling flowchart contains three parts, as Fig. 2 shows.

Part One (the top part): Data preprocessing and input variables selection. Decompose the wind speed **S** into an approximation signal **A** and a detail signal **D** by WT. As we discussed previously, a good decomposition should ensure the approximation signal **A** maintains the major fluctuations of the speed for later forecasting modeling while the detail signal **D** presents the local minor volatility to be eliminated.

We select temperatures on different prior orders as environment variable of wind speed by Granger causality test; and select historical wind speeds as endogenous variables by calculating autocorrelation and partial correlation.

After WT pretreatment and input selection, the original data set are divided into three sets: the training set, the validation set and the test set for model training, validation and test respectively.

Part Two (left part of the bottom): Model training and model validation. In this part, SVM are trained with the data pretreatment in Part One and initialization parameters. The training error and validation error are averaged and used as fitness function of GA, and then GA optimization is performed to search a group of parameters of SVM to minimize the fitness value.

Part Three (right part of the bottom): Modeling test. In this part, SVM with optimal parameters tuned by GA are used to forecast the future wind speed.

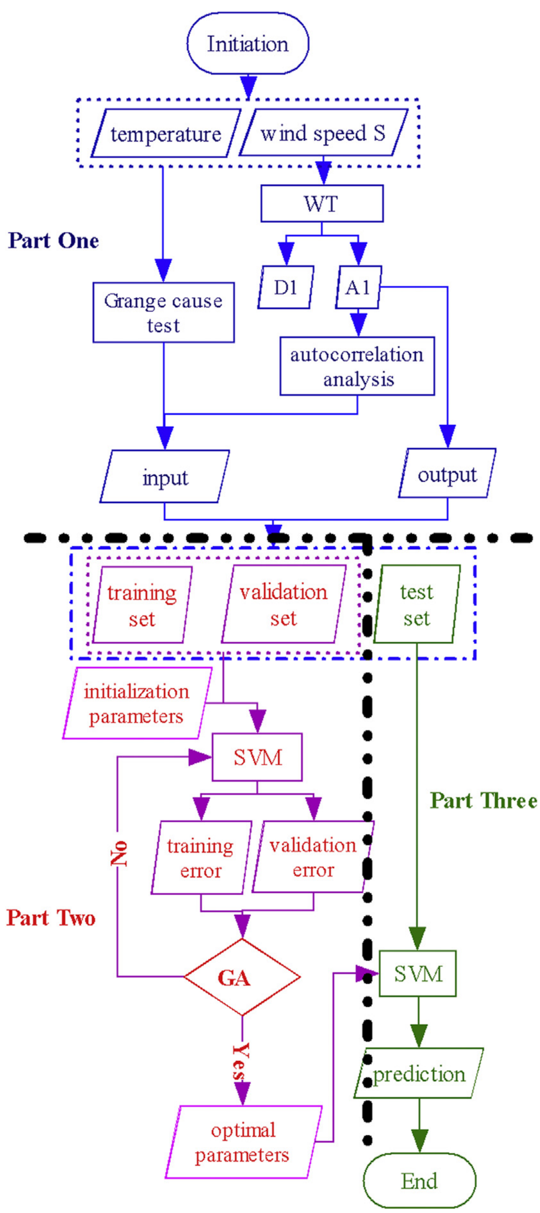

Fig. 2. Flowchart of the W-SVM-GA modeling.

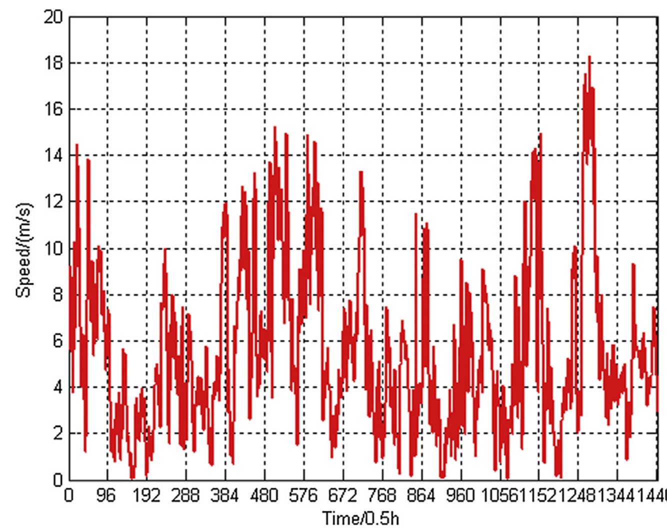

Fig. 3. Wind speed in September of a wind farm located in northern China.

4. Case-studies

4.1. Data preprocessing

The wind speed data every 0.5 h in a wind farm of North China in September 2012 are taken for experiments to validate our proposed modeling technique. Fig. 3 shows the wind speed of 1440 samples in September. It can be seen that the wind speed fluctuates severely, ranging from around 0 m/s to near 20 m/s. From Fig. 3 we cannot see any apparent regularity of the wind speed.

WT is adopted to decompose the original wind speed to eliminate its stochastic volatility for further modeling. The wind speed S, the approximation component $\mathbf{A}_1$ and the detail component $\mathbf{D}_1$ decomposed from S by WT are shown in Fig. 4.

From Fig. 4, we can clearly see that $\mathbf{A}_1$ shows high similarity to **S**, presenting major fluctuation of wind speed; at the same time the

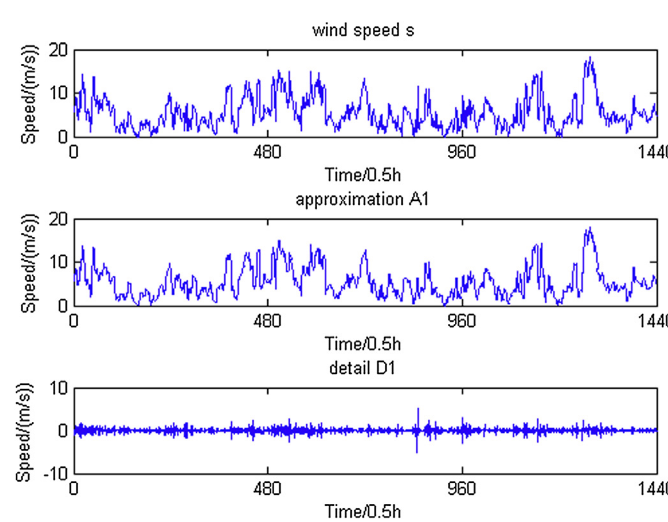

Fig. 4. Original wind speed signal and its approximation component and detail component decomposed by WT.

other minor irregularity that neglected by $\mathbf{A}_1$ appears in $\mathbf{D}_1$. So we just take $\mathbf{A}_1$ as the wind speed to model for efficiencies.

4.2. Selection of input

We will build a SVM model to forecast the wind speed in the following. As discussed previously, the performance of SVM depends on their input and parameters, so here we labor to select the input by the techniques of correlation analysis and Granger causality test that have been widely adopted in econometrics to analyze the relation between the variables.

The correlation analysis and partial correlation analysis are applied to select the historical wind speeds that have highest correlation on the target speed as the input of SVM model. Fig. 5 is the plot of correlation analysis and partial correlation analysis of the wind speed.

It can be seen from Fig. 5 that there exits high autocorrelation in $\mathbf{A}_1$ (AutoCorrelation Function, ACF of $\mathbf{A}_1$, the top part of the figure). By analyzing the partial correlation of $\mathbf{A}_1$ (PACF of $\mathbf{A}_1$, the middle part of the figure) and the partial correlation of squared $\mathbf{A}_1$ (PACF of squared $\mathbf{A}_1$, the bottom part of the figure), we can know that only Lag 1 and Lag 3 of ACF and PACF are significant after the elimination of internal correlation. So we select a_{t-1} and a_{t-3} from the approximation signal $\mathbf{A}$ as parts of the input.

In order to mine the information of the environment except the inner information contained in historical wind speed, the correlation between temperature and wind speed is analyzed. Fig. 6 reveals the fluctuations of the wind speed and the temperature. It can be seen that the fluctuation of the wind speed lags behind the temperature.

Granger causality test is applied to quantitatively analyze the exact dependence of the two variables on different lags. Table 1 shows the Granger causality test of the wind speed and the temperature from Lag 1 to Lag 4.

As can be seen obviously from Table 1, for Lag 2 and Lag 3 the supposition that "temperature does not cause wind speed" is rejected on the significance level of 0.005, indicating that the temperature will affect the wind speed lagged two and three (1 h and 1.5 h). Here we choose the temperatures 1.5 h ahead (Lag 3) as an environment variable for SVM input besides the historical speeds. It also can be seen from Table 1 that for Lag 1 the supposition that "wind speed does not cause temperature" is rejected on the significance level of 0.05, illustrating that the wind speed has great impact on temperature 0.5 h later. However this information contributes less to improving the wind speed forecasting, so it is out of our consideration.

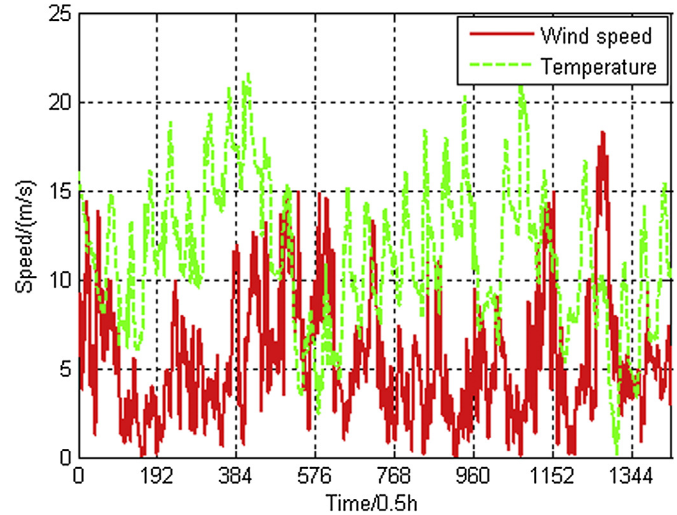

Fig. 6. Relationship between the wind speed and temperature.

Under the above considering, the input for SVM we choose are the historical wind speed 0.5 h and 1.5 h ahead, the temperature 1.5 h ahead. Next, we divide the 1437 wind speeds in September (1440 samples with three samples removed for three lag lengths) into training set, validation set and test set. The number of samples in training set is 765, and both 336 in validation set and test set.

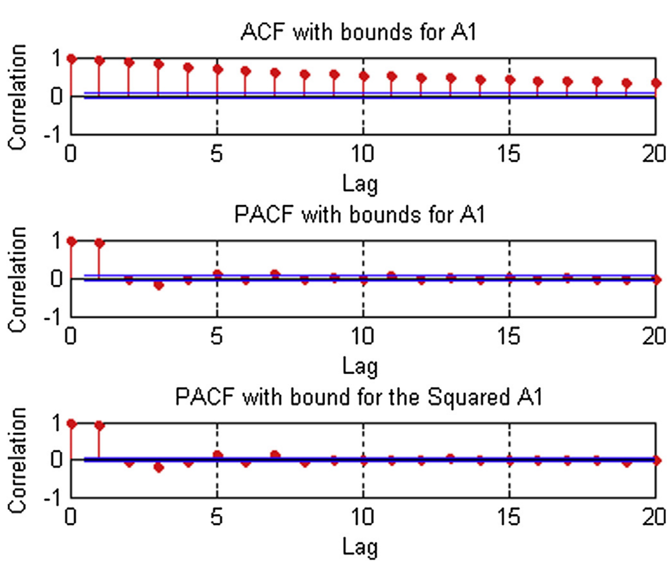

Fig. 5. ACF and PACF between $\mathbf{A}_1$ and squared $\mathbf{A}_1$.

4.3. Model performance evaluation

To examine the performance of model, MAE (see formula (1)), MAPE (see formula (4)) and RMSE (see formula (5)) are selected to measure the forecast accuracy.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|I_a| \times 100\% \tag{1}$$

where n is the number of the speeds to be forecasted, and

$$I_a = \hat{y}_i - y_i \tag{2}$$

where $\hat{y}_i$ is the ith forecast wind speed and y_i is the actual speed of the same point.

$$I_p = \frac{\hat{y}_i - y_i}{y_i} \times 100\% \tag{3}$$

and

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}|I_p| \times 100\% \tag{4}$$

Table 1
Probability to reject in Granger causality test between the wind speed and temperature.

Supposition	Lag 1	Lag 2	Lag 3	Lag 4
Temperature does not cause wind speed	0.1739	0.0023	0.0022	0.0069
Wind speed does not cause Temperature	0.0481	0.3539	0.3195	0.6577

Table 2
Parameters of GA.

Population Size	20	Crossover Fraction	0.8	Selection function	Stochastic uniform
Stall generations	100	Termination tolerance	1e-6	Crossover function	Crossover heuristic
Elite Count	2	Initial Population	Random	Fitness Scaling	Shift linear
Migration Fraction	0.2	Population Type	Double Vector	Mutation function	Gaussian

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(I_a)^2} \tag{5}$$

4.4. Selection of SVM parameters

As discussed previously, the performance of SVM modeling lies on its parameters. GA is developed to tune crucial parameters of SVM by minimizing errors generated in the training set and validation set. Here we choose *nu*-SVM as a basic model, RBF as the kernel function. The other parameters, cost (c), *nu* of SVM, tolerance of termination criterion (e) and gamma of kernel function RBF (g) are tuned by GA automatically. Some parameters in GA modeling are listed in Table 2.

The fitness function of GA is the MAPE generated in the training set and the validation set.

Fig. 7 reveals the best fitness and the mean fitness of GA for different generations.

From Fig. 7 we can see that the fitness drops quickly along with generation iterations. At beginning phase, the fitness is around 1.7, which means that MAPE much bigger than 1, denoting the poor performance of SVM with random initial parameters. However with the iteration increasing, the fitness falls quickly, illuminating that GA has found better parameters for SVM in a short period. After 20 generations, there is no obvious change of the fitness, denoting that the best parameters have been obtained by GA. The parameters of SVM optimized by GA are listed in Table 3.

The test set is used to exam the SVM model with the parameters obtained by GA. Fig. 8 shows the comparison of actual speed and the forecasting residual series by SVM. It can be seen that forecasting by SVM is acceptable for the forecast error (residual) varies around 4 m/s even during the pike speed periods.

The autocorrelation and partial correlation of the residual by SVM is further studied to exam the wind speed forecasting, as Fig. 9 shows. From Fig. 9, we can see that there is no significant autocorrelation and partial correlation of the residual series and neither is the partial correlation of the squared residuals. Then we can conclude that the information that contained in the historical speed signal is mined well by SVM modeling.

A persistent model and SVM with parameters tuned by GA (SVM-GA) are developed to compare with the proposed model W-SVM-GA based on three criteria, MAE, MAPE and RMSE, shown in Table 4. It can be seen from Table 4 that W-SVM-GA outperforms the SVM-GA (based on the MAE and MAPE criterions) suggesting that WT helps to improve the performance of SVM. The latter two

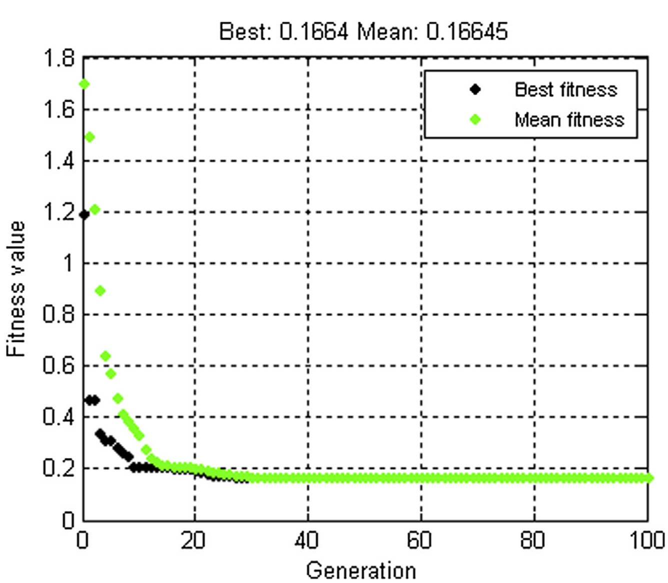

Fig. 7. Fitness of GA along with generation iterations.

Table 3
Optimal parameters of SVM.

Parameter	Value of the parameter
c, nu, e, g	3.9932, 0.6863, 0.0004, 0.0286

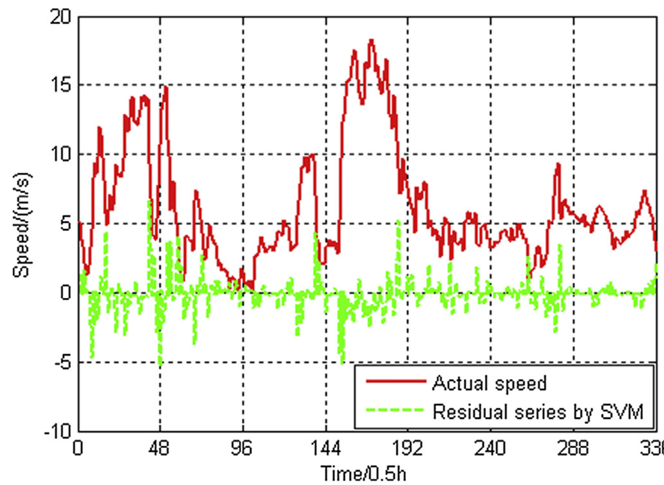

Fig. 8. Actual speed and the residual series (I_a) by SVM.

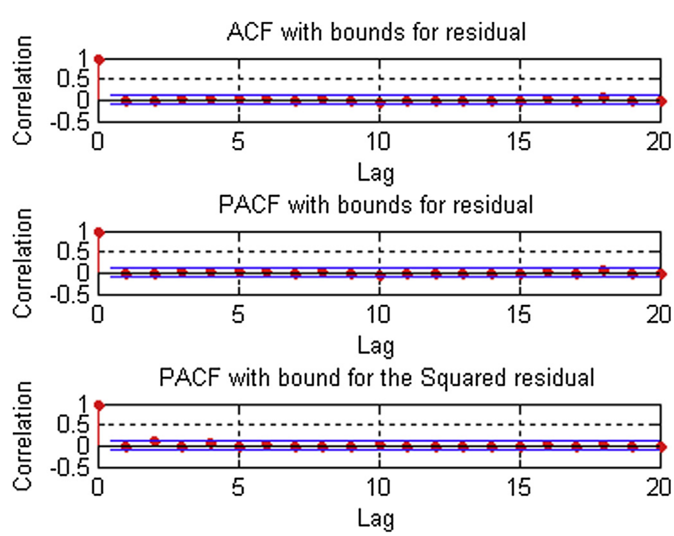

Fig. 9. Correlation of the residual series by SVM.

Table 4
Performance evaluations.

Model	MAE (m/s)	MAPE (%)	RMSE (m/s)
Persistent	0.8356	22.64	1.2182
SVM-GA	0.7843	17.80	1.2125
W-SVM-GA	0.6169	14.79	1.2234

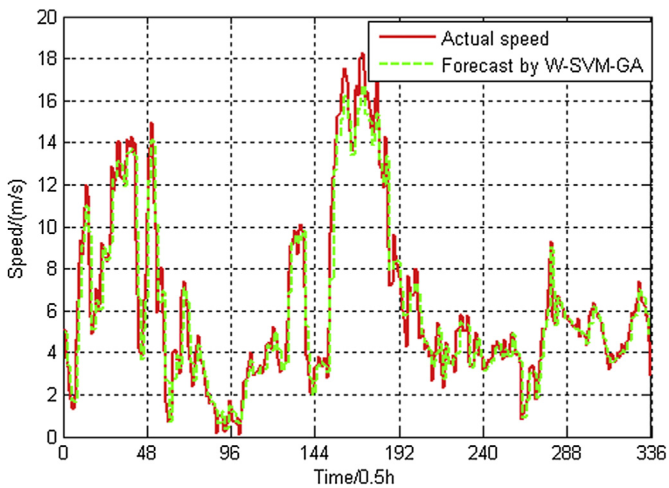

Fig. 10. Comparison between the actual speed and the forecast by W-SVM-GA.

models perform much better than the persistent model on the criterion of MAPE because we take the MAPE as the fitness function of GA; and it denotes the powerful modeling of SVM-GA. However RMSEs of the latter two models are slightly bigger than their counterparts, suggesting that the GA modeling pays more attention on the MAPE and ignores the other indices if we take the MAPE as the only criterion for the fitness.

Maybe more criterions should be involved to develop a comprehensive index as the GA fitness to find more robust parameters for SVM in wind speed forecasting.

Fig. 10 is a comparison between the actual wind speed and the forecasting. We can see that W-SVM-GA gives acceptable predictions most of time, while it inclines to underestimate the speed for those extremums.

5. Discuss and conclusions

This paper proposed a hybrid intelligent algorithm to forecast the short-term wind speed. Firstly the WT was used to preprocess the original wind speed signal to eliminate the random fluctuation of the wind speed. Then the SVM model was established to forecast the approximation of the wind speed signal obtained by WT. GA was proposed to search the optimum parameters to avoid the risk of instability caused by improper parameters selected. When selecting the input variables of SVM, we used Granger causality test to choose proper lags of the environment variable temperature, and used autocorrelation and partial correlation to choose proper lags of historical speeds. The proposed method is more efficient than a persistent model and a SVM-GA model without WT.

This work could be expended in several directions. More environment variables, such as the air pressure, precipitation, air humidity besides the temperature could be added into the input for SVM to augment external information. More criteria such as MAE and RMSE could be introduced into the fitness function in GA optimization to ensure a comprehensive promotion of the performance in further studies.

Acknowledgments

This work was supported by National Natural Science Foundation of China (NSFC) (70901025) and in part by the Fundamental Research Funds for the Central Universities (13MS32). Also we are grateful to the referees for their helpful suggestions.

Da Liu is with the School of Economics and Management, North China Electric Power University, Beijing 102206 China (E-mail: liuda315@163.com).

References

[1] Foley AM, Leahy PG, Marvuglia A, McKeogh EJ. Current methods and advances in forecasting of wind power generation. Renewable Energy 2012;37:1–8.
[2] Zheng XX, Fu Y. Research of wind speed and wind power forecasting. Renewable and Sustainable Energy 2012;347-353(1–7):611–4.
[3] Cao Qing, Ewing Bradley T, Thompson Mark A. Forecasting wind speed with recurrent neural networks. European Journal of Operational Research 2012;221:148–54.
[4] Erdem E, Shi J. Arma based approaches for forecasting the tuple of wind speed and direction. Applied Eenergy 2011;88:1405–14.
[5] Liu Heping, Erdem Ergin, Shi Jing. Comprehensive evaluation of arma–garch (-m) approaches for modeling the mean and volatility of wind speed. Applied Energy 2011;88:724–32.
[6] Jiang W, Yan Z, Feng DH, Hu Z. Wind speed forecasting using autoregressive moving average/generalized autoregressive conditional heteroscedasticity model. European Transactions On Electrical Power 2012;22:662–73.
[7] Li Gong, Shi Jing. On comparing three artificial neural networks for wind speed forecasting. Applied Energy 2010;87:2313–20.
[8] Cadenas E, Rivera W. Short term wind speed forecasting in la venta, oaxaca, mexico, using artificial neural networks. Renewable Energy 2009;34:274–8.
[9] Ortiz-Garcia EG, Salcedo-Sanz S, Perez-Bellido AM, Gascon-Moreno J, Portilla-Figueras JA, Prieto L. Short-term wind speed prediction in wind farms based on banks of support vector machines. Wind Energy 2011;14:193–207.
[10] Zhou JY, Shi J, Li G. Fine tuning support vector machines for short-term wind speed forecasting. Energy Conversion and Management 2011;52:1990–8.
[11] Cadenas Erasmo, Rivera Wilfrido. Wind speed forecasting in the south coast of oaxaca, méxico. Renewable Energy 2007;32:2116–28.
[12] Kavasseri Rajesh G, Seetharaman Krithika. Day-ahead wind speed forecasting using f-arima models. Renewable Energy 2009;34:1388–93.
[13] Catalao JPS, Pousinho HMI, Mendes VMF. Short-term wind power forecasting in portugal by neural networks and wavelet transform. Renewable Energy 2011;36:1245–51.
[14] Catalao JPS, Pousinho HMI, Mendes VMF. Hybrid intelligent approach for short-term wind power forecasting in portugal. IET Renewable Power Generation 2011;5:251–7.
[15] Guo ZH, Zhao WG, Lu HY, Wang JZ. Multi-step forecasting for wind speed using a modified emd-based artificial neural network model. Renewable Energy 2012;37:241–9.
[16] Niu Dongxiao, Liu Da, Wu Desheng Dash. A soft computing system for day-ahead electricity price forecasting. Applied Soft Computing 2010:868–75.
[17] Jursa René, Rohrig Kurt. Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models. International Journal of Forecasting 2008;24:694–709.
[18] Liu Da, Niu Dongxiao, Xing Mian, Nie Qiaoping. Day-ahead price forecast with genetic-algorithm-optimized support vector machines based on garch error calibration. Automation of Electric Power Systems 2007:31–4.
[19] Monfared M, Rastegar H, Kojabadi HM. A new strategy for wind speed forecasting using artificial intelligent methods. Renewable Energy 2009;34:845–8.
[20] Blonbou R. Very short-term wind power forecasting with neural networks and adaptive bayesian learning. Renewable Energy 2011;36:1118–24.
[21] Cadenas E, Rivera W. Wind speed forecasting in three different regions of mexico, using a hybrid arima-ann model. Renewable Energy 2010;35:2732–8.
[22] Granger CWJ. Some recent development in a concept of causality. Journal of Econometrics 1988;39:199–211.