A Combined Approach for Very Short Term Wind Power Probability Forecast

Yubo Tao, Yafei Zhang, Bo Chen and Jing Shi, *Member, IEEE*

***Abstract*—Wind power has been promoted to mitigate the energy crisis worldwide. It has beneficial impacts on economic and environment. However, owing to the chaotic nature of atmospheric movement, wind power generation always exhibits nonlinear and non-stationary uncertainties, which brings great challenges for power system operation and planning. To meet the challenges, a novel deep learning based combined approach is proposed for uncertainties. In this approach, a start-or-art point forecast approach is proposed based on wavelet transform and RNN-RBM deep learning. Raw wind power is decomposed into different frequencies. The nonlinear patterns are used to improve the forecast accuracy by this approach. Consequently, the probabilistic distribution of wind power data can be statistically formulated via the non-parametric approach. The experimental results demonstrate that this combined approach reduces the forecast error, which has a better performance than the other two comparison forecast approaches. The average accuracy of MAPE has increased by about 3.2%.**

***Index Terms*—Wind power forecast, probabilistic forecast, deep learning, combined approach**

I. INTRODUCTION

WITH the global energy interconnection 'power substitution' and 'clean alternative' put forward, wind power is becoming one of the most rapidly growing renewable energy sources, and regarded as an appealing alternative to conventional power generated from fossil fuel, which plays a very important role in national energy policies all around the world[1]-[3]. Because of its advantages of clean, low carbon and renewable energy, wind power has provided strong support for energy saving, emission reduction and environmental protection in various countries. According to the data released by the National Energy Administration in 2017, new installed capacity of grid-connected wind power is 15.03 million kilowatts; the accumulative installed capacity of grid-connected wind power reached 164 million kilowatts, accounting for 9.2% of the total installed capacity. The annual total wind power generation capacity is 305.7 billion kilowatt-hours, accounting for 4.8% of the total generation capacity, which is 0.7% higher than that in 2016[4].

Power system dispatching mode ensures the dynamic balance of power generation and load. Through the unit commitment, economic dispatch and primary frequency modulation strategy of output power stable generation units, power generation can track the load forecasting accurately and keep the dynamic balance of the generation with load. With the increase of the wind power capacity, the wind power spatial scale dispersion and the strong random fluctuation of the time scale can reduce the wind turbine controllability and affect the capacity of the power generation in the power system to track the load, which brings new challenges to the security, stability and economy. At present, one of the important approaches to solve this contradiction is to improve the wind power forecast. Accurate forecasting results can reduce costs and improve the reliability of wind power integration[5]-[6].

Considerable attention has been paid to wind power forecast these years. A large sum of research has been toward the development of accurate and reliable wind power forecasting models and many different approaches have been reported in the literature. Generally speaking, one of the most easily and widely used very short-term and short-term wind power forecasting approaches is data-driven modeling. Wind power is a typical time series with strong nonlinear features due to many factors affected[7]-[9]. A large number of nonlinear forecasting models have been introduced into the wind power forecast, including decision trees, decision forests, logistic regression, support vector machines and neural network. Many nonlinear forecasting models that usually perform well such as GARCH model, support vector regression and neural network have been focused, which mostly aimed at very short-team wind power

Manuscript received July 31, 2018; revised September 17, 2018.

Dr. Yubo Tao is with the State Grid Jiangsu Electric Power Co., Ltd Maintenance Branch Company, Jiangsu 211100 China (e-mail: rgtyb@foxmail.com).

Yafei Zhang is with the State Grid Jiangsu Electric Power Co., Ltd Maintenance Branch Company, Jiangsu 211100 China.

Bo Chen is with the State Grid Jiangsu Electric Power Co., Ltd Maintenance Branch Company, Jiangsu 211100 China.

Dr. Jing Shi is with the School of Electrical Engineering, Wuhan University, Wuhan 430072 China. (phone:027-68773898 e-mail: shi.jing@whu.edu.cn).

978-1-5386-6461-2/18/$31.00 ©2018 IEEE

forecast[10]-[13].

In order to deal with strongly nonlinear features of wind power, the first combination strategy for forecasting model was reported in 1999 that the improved the performance compared with a single model[14]. The basic idea of the combination approach is to combine different approaches, retaining advantages of each approach. Reference[15] proposed a forecasting model combining neuro-fuzzy and artificial neural network. Reference[16] proposed a wind power forecasting models based on many algorithms. Decomposition of the original data was proposed from 2009, which can improve the regression models[17], [18]. In addition, deep learning is one focus of the research in recent years. Deep learning networks allow computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction[19].

Not only need accurate point forecasts but also the uncertainty of the forecast is essential for determining the size of the operating reserves necessary to balance the generation with load. This paper proposes a novel combined forecasting approach, support vector regression is used to forecast the stationary components which are obtained from wavelet decomposition. For the strong non-stationary uncertainty component, wind power forecasting approach based on deep learning has better results. A non-parametric approach based on kernel density estimation is proposed to provide the complete wind power forecast distribution, so that the point forecasts can be extended to the probability forecasts, which can provide more decision information.

II. Data Statistical Analysis

A. Characteristics of wind power

The uncertainty of wind power is difficult to describe. Wind uncertainty is mainly formed by turbulence and its fluctuation. Atmospheric motion includes a variety of time scales. Movement in different scales has different roles in the transmission process of matter and energy, the component separation of different scales can be separately studied the characteristics of different scales.

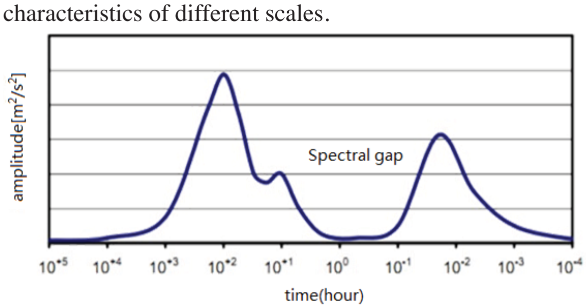

Fig. 1. Horizontal wind speed spectrum at 100m height

Figure 1 is the horizontal wind speed spectrum at 100m height. From this figure, there appears to be two major energy peaks in the spectrum; one peak occurs at the period of about 100 hours, and the second peak occurs at the period of close to 0.01 hour (i.e. close to 1 minute). Between the two peaks, a broad spectral gap lasts about one hour. The spectral gap seems to exist under varying terrain and synoptic conditions[20]. It is worth noting that weather movement process can be decomposed into two different time scales process because of the spectral gap, which is the stationary component and the non-stationary uncertainty component. The time interval between the non-stationary component and stationary component is close to 1 hours. Wind power is calculated by the wind speed at the hub height of the wind turbines plugged into the corresponding wind power curve, so that wind power and wind speed have similar characteristics. This analysis is based on the basic atmospheric dynamics' principles.

B. Analysis of wind power components

Wavelet transform is a mathematical tool for nonlinear and non-stationary signal analysis, which has the capability to separate the signal energy among different frequency bands[21]. Wavelet decomposition can convert a raw wind power series in a set of constitutive sub-series. The decomposed sub-series are separately forecasted and then the aggregate calculation of the sub-series to form the final forecast value of wind power.

In this paper, according to the literature[22], db4 function is selected to be the mother wavelet because it provides a proper balance between the wavelength and smoothness, so that Mallat algorithm is adopted.

III. Forecasting Approaches

A. Deep learning approach based RNN-RBM

In 2006, Hinton used the Restricted Boltzmann Machine (RBM) to build the Deep Belief Network (DBN). The unsupervised pre-training process caused extensive attention, making the training of the deep neural network possible. RBM consists of visible units and hidden units. Each visible unit is connected to all the hidden units and this connection is undirected. The bias unit is connected to all the visible units and hidden units. This restriction allows for more efficient training algorithms that are available for the general class of Boltzmann Machines, such as the gradient-based contrastive divergence algorithm.

RBM essentially perform a binary version of factor analysis and consists of a matrix of weights W_{jk} associated with the connection between hidden unit h_j and visible unit v_k, as well as bias weights a_k for the visible units and b_j for the hidden units. RBM is further restricted to abandon visible-visible and hidden-hidden connections. Given these, the energy of a configuration is defined as

$$E(v,h) = -h^T W v - a^T v - b^T h \\ = -\sum_k a_k v_k - \sum_j b_j h_j - \sum_{jk} w_{jk} v_k h_j \quad (1)$$

The joint distribution over the visible units and hidden units is defined by

$$p(v,h) = \frac{e^{-E(v,h)}}{Z} \quad (2)$$

Where Z is a normalization factor, which computation times is increasing by the number of hidden unit and visible unit number. It is difficult to get the actual distribution calculation.

When given a visible unit or hidden unit state, the implied probability of activation layer is independent, the probability of nodes of the hidden unit of the j or the visible unit k are respectively in (3) and (4):

$$p(h_j = 1|v) = sigmoid(b_j + \sum_k w_{jk} h_k) \quad (3)$$

$$p(v_k = 1|h) = sigmoid(c_k + \sum_j w_{jk} h_j) \quad (4)$$

For the hidden units and visible units are in (5) (6):

$$p(h|v) = \prod_j p(h_j|v) \quad (5)$$

$$p(v|h) = \prod_k p(v_k|h) \quad (6)$$

DBN is a l-layer neural network. The input can be written as a vector $x=h_0$, and ($h_1,\cdots\cdots, h_{l-1}$) is hidden unit, h_l is output layer. The 1:l-1 levels of the underlying network is constituted by RBM, using sigmoid function. However, the top-level activation function can use a pure linear function owing to regression problems. It defines a joint probability distribution for the hidden unit of input layer x and l. The joint probability density x and l is in (7).

$$p(x, h^1, \cdots\cdots, h^l) = (\prod_{i=1}^{l-1} p(h^{i-1}|h^i)) p(h^{l-1}, h^l) \quad (7)$$

The basic idea of DBN learning algorithm is to make the network hierarchical, from down to up, unsupervised training every layer, taking the underlying RBM hidden unit as the input of last RBM, and after the end of the unsupervised pre-training, supervised learning is used to precisely adjust network. The training process is shown in Figure 2, balancing the energy of the RBM, which is firstly formed between the visible unit and the first hidden unit, through the training methods on the last section; the second RBM is composed by h^1 as the input and h^2, adjusting the parameters such that the RBM training is completed; this process is repeated to train the third RBM[23], [24].

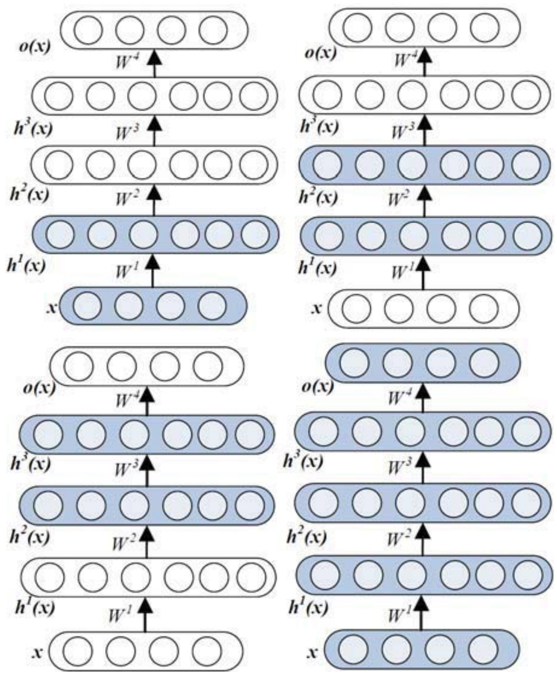

Fig. 2. Diagram of DBN training process

In order to forecast the time series more accurately, the RNN-RBM combined forecasting approach based on the energy model can be considered, which may have high dimension at a certain time.

Recurrent Neural Networks (RNNs) are models where the idea is to make use of sequential information. An RNN takes advantage of the temporal dependence and uses previous calculations as input to each new calculation. The purpose was to further utilize the forecasting capability of the two models and to create a model that allowed more freedom to describe the temporal dependencies involved. The model extends the RNN model by adding an RBM at each time step. The output layer of the RNN is no longer a direct representation of the visible units intended to forecast, but instead lays ground to the parameters for the RBM model. This can be seen in Figure 3. The bottom layer constitutes the RNN model and the top-two layers constitutes the RBM model.

The bias vectors for the RBM model $b_v(t)$ and $b_h(t)$ are updated through the hidden units for the RNN layer $u(t$-1) in (8):

$$b_v(t) = b_v + W_{uv} u(t-1) \\ b_h(t) = b_h + W_{uh} u(t-1) \quad (8)$$

where b_v and b_h are the initial bias vectors for the visible units and the hidden units in the RBM layer. The vector $u(t)$ represents the hidden units for the RNN layer at time t and is calculated as (9)

$$u(t) = \tanh\left(b_u + W_{uu}u(t-1) + W_{vu}v(t-1)\right) \quad (9)$$

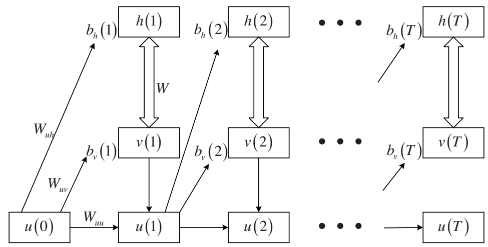

Fig. 3. A graphical illustration of the RNN-RBM model

B. *Probabilistic forecasting based on non-parametric approach*

Mainly due to the stochastic nature and intermittent nature, it is very difficult to accurately forecast wind power. The significant error is usually generated. The uncertainty of wind power forecasting is an extension of the point forecasts. Knowing the uncertainty of the forecast enables the system operator to assess the risk of the point forecast. The uncertain forecasting approach can give the probability that the wind power falls within a certain interval. The larger interval of forecasting, the higher probability that the forecasting results falls in the interval.

The non-parametric approach does not rely on any assumption about the distribution to model for the forecasting error and is suitable to estimate the uncertainty of the wind power forecast. The commonly approach is kernel density estimation. To adopt the approach of forecast bin, sorting data in a small to large order, data is divided into power bins for modeling the distribution. Then, according to the probability density function of wind power forecasting error and point forecasting results, an interval estimate of a confidence level can be obtained to reflect the fluctuation range of wind power. The calculation steps are as follows:

Step 1: the forecasting results of wind power are arranged from small to large and then divided into equal bins;

Step 2: in a bin, the forecasting error is analyzed, and the probability density function is estimated by the non-parametric estimation approach;

Step 3: the confidence interval of the forecasting results is calculated through probability density function.

C. *Support vector regression*

Vapnik developed a machine learning algorithm called support vector machines(SVM), which is based on statistical learning theory, to solve pattern recognition and classification problems. As a very specific type of learning algorithms characterized by the capacity control of the decision function, the use of the kernel functions and the sparsity of the solution. Established on the unique theory of the structural risk minimization principle to estimate a function by minimizing an upper bound of the generalization error, SVM are shown to be very resistant to the over-fitting problem, eventually achieving high generalization performance in solving various time series forecasting problems. PSO algorithm is used to optimize parameters. For more detailed descriptions see the reference[25], [26].

IV. EXPERIMENT VERIFICATION

A. *Forecasting framework*

The forecasting model is a combined of SVR and deep RNN-RBM, as shown in figure 4. The raw wind power data is first normalized and then decomposed into the stationary component and the non-stationary uncertainty component via wavelet toolbox. The stationary component has smaller degrees of uncertainty than the non-stationary component. A quick SVR model is used to forecast the behavior of it. The hybrid deep learning model is designed for the non-stationary component as accurately as possible. Then, synthesizing all of the forecasting frequencies via wavelet reconstruction, and anti-normalization produces the final point forecasting results of wind power. Finally point forecasting results is extended to confidence interval via the non-parametric approach.

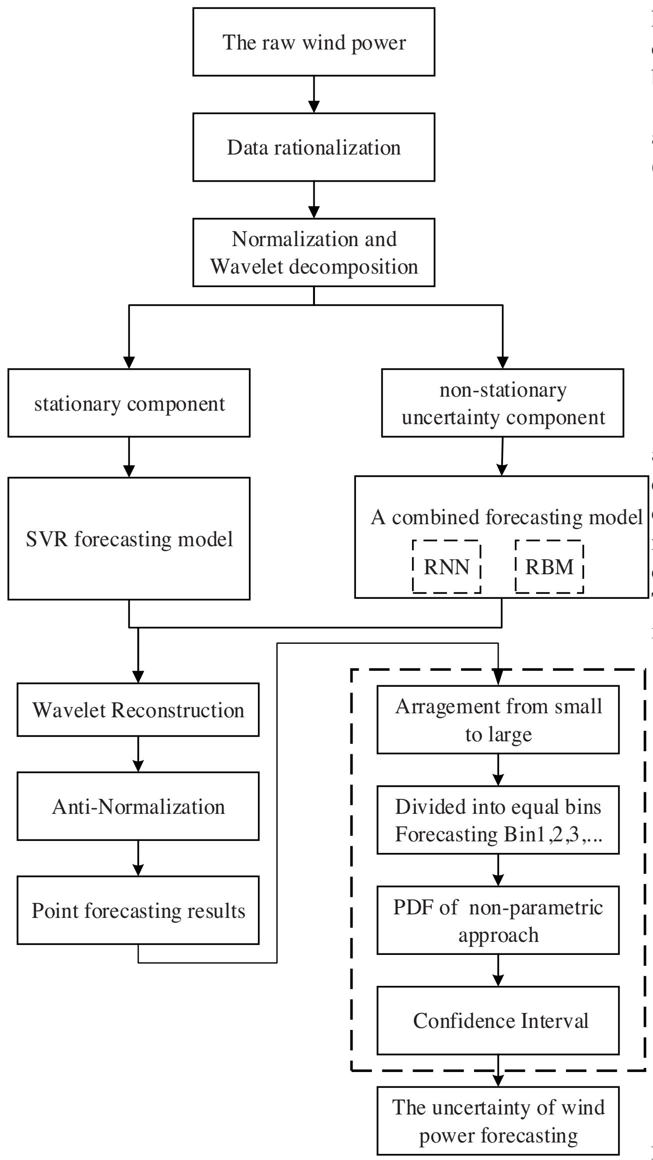

Fig. 4. Schematic diagram of the proposed forecasting approach

B. Experiment and evaluation criterion

Wind farm is situated on Zhangjiakou, Hebei Province, where wind power data is collected for analysis with a 5-second resolution. This wind farm has a rated capacity of 110.5MW with 130 Vestas 850 wind turbines. The requirement of time scale for power system regular scheduling is usually 15 minutes resolution, so raw wind power data is averaged in the 15-minute resolution.

Since there is no definite method for the selection of the hidden units and the number of nodes in deep learning approach, here the number of DBN hidden units is three and the number of nodes in every hidden unit is 100. The number of forecasting bins is 9.

The criteria for comparing the performance are mean absolute error (MAE) and mean absolute percentage error (MAPE) as follows:

$$MAE = \frac{1}{m}\sum_{t=1}^{m}\left|y_{real} - y_{forecast}\right| \tag{10}$$

$$MAPE = \frac{1}{m}\sum_{t=1}^{m}\frac{\left|y_{real} - y_{forecast}\right|}{y_{real}} \tag{11}$$

C. Experiment analysis

The power interval segmentation needs to be divided according to the actual situation. If the number of sample points divided into one bin is small, the estimation effect of the kernel density estimation is not ideal at this time, and the bin of a small number of sample points needs to be merged with the previous one, and the combined bin is used for kernel density estimation. The different power bins of kernel density estimation are shown in Figure 5.

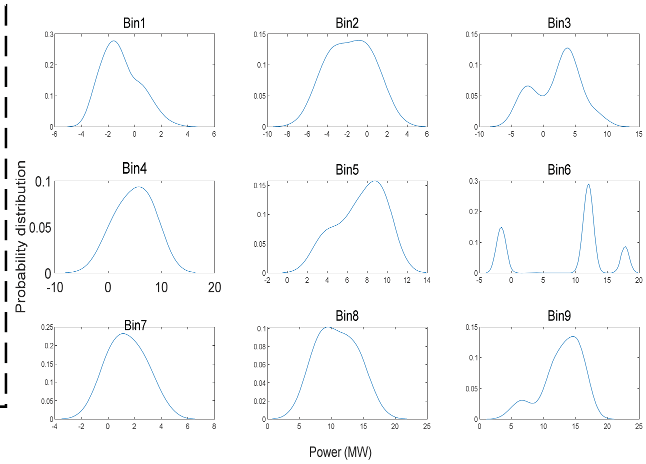

Fig. 5. The different power bins of kernel density estimation

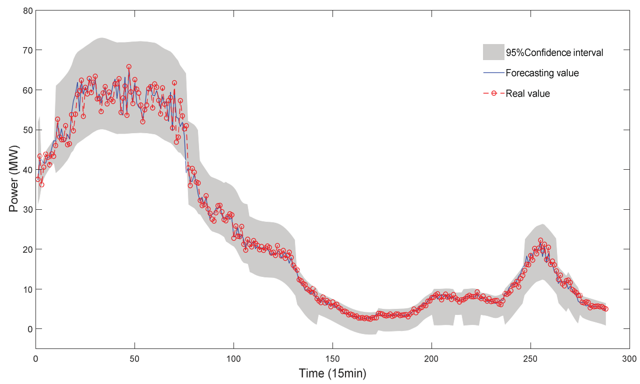

Fig. 6. Comparison of forecasting results with real values

It can be seen from figure 6 that the forecasting value and the real value curve are basically the same, and the forecasting accuracy is satisfactory. The gray interval in this figure is the confidence interval when the confidence level is 95 percent. It can be concluded that the probabilistic performances obtained from the proposed approach are satisfactory.

As shown in figure 7 and figure 8, a combined forecasting approach, a PSO-LSSVM model and a single DBN are used to forecast the wind power separately for comparison purposes.

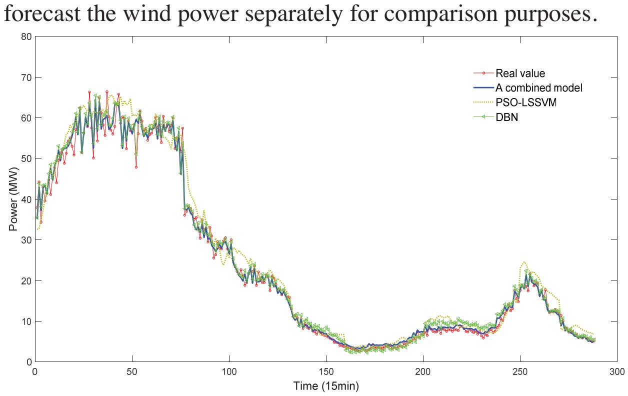

Fig. 7. Comparison of wind power results with three forecasting approaches

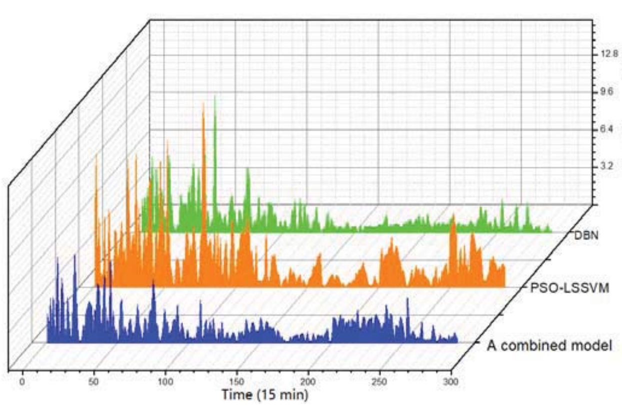

Fig. 8. Comparison of forecasting absolute errors with three approaches

This paper proposed approach is effective in nonlinear mapping learning, and the overall accuracy is relatively stable, and error is at a low level. The PSO-LSSVM model is a multi-input single-output model. The training process requires less data than the neural network. When the forecasting step is short, the accuracy is better, but with increase of the step length, the error is also increased accordingly. The nonlinear mapping effect of the single DBN deep learning model is satisfactory, but the data decomposition may change some intrinsic contact information, and the model training is not particularly sufficient, and the forecasting result is slightly worse than the proposed approach. The single DBN takes a long time to execute.

TABLE I
COMPARISON OF FORECASTING ERRORS WITH THREE APPROACHES

Approach	MAE (MW)	MAPE (%)	Time (s)
Proposed approach	1.08	7.02	89.13
PSO-LSSVM	2.42	11.21	50.41
DBN	1.29	9.03	505.26

As can be seen from Table 1, the proposed approach (i.e. a combined forecasting approach) has a better performance than the other two forecasting model. Compared with the single DBN model, the forecasting accuracy of the combined forecasting approach is high efficiency and high accuracy, which is suitable for the very short-term wind power interval forecast.

V. CONCLUSION

Owing to the economic and environmental benefits, wind power is becoming one of the more promising supplements for power generation, but high penetration of wind power provides a good many of challenges in power system operations and planning because of its uncertain and intermittent nature. Probabilistic forecasting of wind power with high accuracy is a pressing need. A novel combined forecasting approach based on wavelet transform, deep learning and ensemble technique is proposed with the purpose of it.

(1) Raw wind power data is decomposed into the stationary component and the non-stationary uncertainty component via wavelet toolbox.

(2) The stationary component is forecasted by a quick SVR model, and the non-stationary is forecasted by a hybrid deep learning with RNN-RBM.

(3) Extending the result of point forecasting to probabilistic forecasting by the non-parametric approach, which gives an estimate of where the wind power generation will lie within a certain degree of confidence. The confidence interval obtained

can provide more information for power system operations and planning. It is also evident that the proposed approach demonstrates a high potential in power system.

ACKNOWLEDGMENT

The authors gratefully acknowledge Dr. Hao Chen, Mr. Baoxin Hao, Mr. Xuan Chen (State Grid Jiangsu Electric Power Co., Ltd Maintenance Branch Company) for instructive advices and useful suggestions in this paper.

REFERENCES

[1] L. Zhenya, *Global energy internet*. Beijing: China Electric Power Press, 2015.

[2] S. Dan, *Research on the development and policy of new energy industry*. Beijing: China Social Sciences Press, 2015.

[3] Y. Tao, H. Chen and X. Liu, "For wind unstable wind power volatility modeling method of research," *International Journal of Simulation--Systems, Science & Technology,* vol. 16, pp. 9.1-9.6, 2015.

[4] N. E. Administration, "Wind power grid operation in 2017,", 2017.

[5] L. Xie, P. M. S. Carvalho, L. A. F. M. Ferreira, J. Liu, B. H. Krogh, N. Popli, and M. D. Ilic, "Wind integration in power systems: operational challenges and possible solutions," *Proceedings of the IEEE,* vol. 99, pp. 214-232, 2011.

[6] W. Du, J. Bi, T. Wang, and H. Wang, "Impact of grid connection of large-scale wind farms on power system small-signal angular stability," *Csee Journal of Power & Energy Systems,* vol. 1, pp. 83-89, 2015.

[7] D. J. Brayshaw, A. Troccoli, R. Fordham, and J. Methven, "The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: A case study over the UK," *Renewable Energy,* vol. 36, pp. 2087-2096, 2011.

[8] H. Chen, Q. Wan, F. Li, and Y. Wang, "GARCH in mean type models for wind power forecasting," in *Power and Energy Society General Meeting*, 2013, pp. 1-5.

[9] T. Yubo, C. Hongkun, W. Jie, and L. Xin, "Research on the dynamic simulation model of the new-energy power system based on grey-correlation analysis," *Revista Ibérica de Sistemas e Tecnologias de Informação,* vol. 18B, pp. 259-269, 2016.

[10] S. A. P. Kani and G. H. Riahy, "A new ANN-based methodology for very short-term wind speed prediction using Markov chain approach," in *Electric Power Conference, 2008. EPEC 2008. IEEE Canada*, 2008, pp. 1-6.

[11] J. Chen, X. Xue, M. Ha, D. Yu, and L. Ma, "Support vector regression method for wind speed prediction incorporating probability prior knowledge," *Mathematical Problems in Engineering,* vol. 2014, pp. 1-10, 2014.

[12] R. Jursa, "Variable selection for wind power prediction using particle swarm optimization," in *GECCO '07 Proceedings of the 9th annual conference on Genetic and evolutionary computation*, 2007, pp. 2059-2065.

[13] H. CHEN, S. GAO, Y. WANG, and J. ZHANG, "Wind power forecasting method based on generalized autoregressive conditional heteroskedasticity with skewness and kurtosis model," *Proceedings of the CSEE,* pp. 3456-3461, 2017.

[14] T. N. Krishnamurti, C. M. Kishtawal, T. E. Larow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, "Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble," *Science,* vol. 285, pp. 1548-1550, 1999.

[15] Y. A. Katsigiannis, A. G. Tsikalakis, P. S. Georgilakis, and N. D. Hatziargyriou, "Improved wind power forecasting using a combined neuro-fuzzy and artificial neural network model," *Lecture Notes in Computer Science,* vol. 3955, pp. 105-115, 2006.

[16] A. Kusiak, H. Zheng and Z. Song, "Models for monitoring wind farm power," *Renewable Energy,* vol. 34, pp. 583-590, 2009.

[17] Z. Guo, W. Zhao, H. Lu, and J. Wang, "Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model," *Renewable Energy,* vol. 37, pp. 241-249, 2012.

[18] L. Wang, L. Dong, Y. Hao, and X. Liao, "Wind power prediction using wavelet transform and chaotic characteristics," in *World Non-Grid-Connected Wind Power and Energy Conference*: IEEE, 2009, pp. 1 - 5.

[19] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science,* vol. 349, pp. 255-260, 2015.

[20] V. I. H. Der, "Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour," *J Meteor,* vol. 14, pp. 160-164, 1957.

[21] E. Magosso, M. Ursino, A. Zaniboni, and E. Gardella, "A wavelet-based energetic approach for the analysis of biomedical signals: Application to the electroencephalogram and electro-oculogram," *Applied Mathematics & Computation,* vol. 207, pp. 42-62, 2009.

[22] J. P. S. Catalao, H. M. I. Pousinho and V. M. F. Mendes, "Hybrid wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal," *IEEE Transactions on Sustainable Energy,* vol. 2, pp. 50-59, 2010.

[23] G. E. Hinton, *A practical guide to rraining restricted boltzmann machines*: Springer Berlin Heidelberg, 2012.

[24] T. Yubo, C. Hongkun, W. Jie, H. Qian, and Y. Ruixi, "Wind power prediction based on deep learning method and its uncertainty," *Journal of the Balkan Tribological Association,* vol. 4, pp. 1148-1156, 2015.

[25] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *Journal of Machine Learning Research,* vol. 2, pp. 999-1006, 2008.

[26] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks,* vol. 17, pp. 113-126, 2004.

Yubo Tao received the Ph.D. degree from School of the engineering, Wuhan University. He is currently an Engineer with the State Grid Jiangsu Electric Power Maintenance Branch Company. His research interests include distributed generation, renewable energy power generation and intelligent operation and maintenance.

Yafei Zhang received the B.S. degree from the School of the Engineering, Wuhan University. He is currently an Engineer with the State Grid Jiangsu Electric Power Maintenance Branch Company.

Bo Chen received the B.S. degree from the Department of Electrical Engineering, Southeast University. He is currently an Engineer with the State Grid Jiangsu Electric Power Maintenance Branch Company.

Shi Jing received the Ph.D. degree from the School of electrical engineering, Huazhong University of Science and Technology. She is now an associate professor with the School of electrical engineering, Wuhan University. Her research interests include renewable energy power generation, high voltage equipment and DC distribution networks.