

Wind power prediction interval estimation method using wavelet-transform neuro-fuzzy network

Feng Ji^{a,*}, Xingguo Cai^a and Jihong Zhang^b

^a*School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, China*

^b*Shengli Oil Field Electric Power Management Corporation, Dongying, China*

Abstract. Wind power point forecasting is the primary method to deal with its uncertainty. However, in many applications, the probabilistic interval of wind power is more useful than traditional point forecasting. Methods to determine the probabilistic interval of wind power point forecasting value is very essential to power system operations. Based on the bootstrap method, this paper proposed a wavelet transform combined with a neuro-fuzzy network model to estimate the prediction interval of wind power. In the model, to account for the ramp event of wind power series, a wavelet-based ramp event was used and the moving block bootstrap method, which considers the dependence of wind power series, was used to construct sampling datasets. Then, the bootstrapped datasets were estimated by a neuro-fuzzy network inference system. A case study provided a 90% confidence level of prediction intervals, which was constructed to examine the effectiveness of the model.

Keywords: Prediction interval, wind power, bootstrap, ramp event, neuro-fuzzy network

1. Introduction

With the rapid increase of wind turbine installed capacity in power systems, wind energy has become the most important and efficient renewable energy from a marginal energy source. However the variability and uncertainty of wind power brings some new challenges to load balance, economic dispatch and unit commitment in power system operations. Reliable and accurate wind power forecasting systems are the primary method to deal with the uncertainty of wind power, which can smooth operation of power systems.

Based on the data used in the forecasting system, there are two primary groups of forecasting methods

for wind power: physical methods and data mining methods [1, 8]. Physical methods use meteorological data and physical laws to obtain wind speed and wind direction information, which is then transformed into corresponding wind power based on the power curves of wind turbines. Data mining methods use historical datasets of wind power or wind speed to extract typical features and develop prediction models [2]. A variety of intelligent algorithms such as neural networks [16], fuzzy neural networks [25] and support vector machines [18] have been proposed for nonlinear forecasting model in the literatures.

Traditionally, most research has focused on developing accurate point forecasting methods for wind power [3]. However, due to the chaotic nature of meteorological systems and the non-stationarity of wind power series, the output of traditional point forecasting methods can hardly correlate to the real measured values,

*Corresponding author. Feng Ji, School of Electrical Engineering and Automation, Harbin Institute of Technology, 92 West Dazhi Street, Nan Gang District, Harbin, China. Tel./Fax: +86 0451 8641 3641; E-mail: jifeng_sy@163.com.

which may lead to uncertainties and risks for power system operations [9]. And Ferreira and Gama pointed out that dealing with the sudden and large changes of ramp events in wind power has become the major issue in wind power generation [10]. Previous Bossavy and Girard further pointed out that the deterministic point forecast models based on minimum least square criterion have poor performance on forecasting of ramp events [4]. Probabilistic interval forecasting provides a more effective approach to estimate and quantify the potential uncertainty of wind power than point forecasting method [11]. Therefore, the prediction of intervals of wind power generation considering ramp events is a significant research topic.

Custom interval estimation methods were based on the predefined probability distribution, so that many applications of wind power forecasting require the knowledge of the probability distribution of forecasting error before generating the interval. However, in practice, the wind power forecasting error is generally acquired from the real prediction value and measured values of one wind farm. The forecasting method and the topographic feature of the wind farm have remarkable influence on forecasting error. When the wind farm participates in the electric market with some bidding strategy as an independent participant, the statistical method of forecasting error distribution may no longer be valid. The bootstrap method as a statistical inference approach can be used to achieve interval estimation without a predefined probability distribution. However, the custom bootstrap method does not consider the dependence of data. Wind power series are a typical dependence time series, and have some special dependence features; especially the ramp event should be considered when using the bootstrap method to estimate the interval of wind power output.

Ramp forecasting is a relatively new study field compare to general wind power forecasting. Ramp is simply defined as a sudden change in wind power which has large amplitude in a relatively short period of time, but there is no consensus on the accepted formal definition of a ramp event [23]. So far, much literature has defined a ramp event as an indicator function by a specified threshold, but Gallego and Costa reported that wavelet transform can describe wind power ramp characterizations more efficiently than many ramp binary definitions using indicator functions [12].

Wavelet transform is a relatively new mathematical approach for signal analysis and wavelet analysis, has been used in many research fields in the last few decades. The time-frequency decomposition of

wavelet analysis has been successfully applied in image processing [19], signal de-noising [24] and condition monitoring [20]. Wavelet combined with some expert models such as neural networks was also used efficiently in stochastic time-series forecasting [21]. In the electrical engineering field, the wavelet transform has been successfully used to forecast electrical load and electricity price. Zhang reported a neural-wavelet model to determine short-term electrical load forecasting [6]. And a wavelet transform combined with ARIMA and GARCH models was developed for day-ahead electricity price forecasting [28]. Because the wavelet transform can efficiently describe wind power ramp characterizations [12] and wavelet transform can effectively deal with the dependence of wind power series when use with the bootstrap method, this paper combines a wavelet transform with a fuzzy neural network model to predict interval of wind power considering ramp events.

The rest of the paper is organized as follows. Section 2 introduces the wavelet-based ramp event characterization. Section 3 describes the phase space reconstruction prediction interval formulation of wind power. A wavelet transform fuzzy neural network model is developed in Section 4. Comprehensive numerical studies are implemented and analyzed in Section 5. Finally, the conclusion is drawn.

2. Wavelet-based wind power ramp event characterization

2.1. Wavelet transforms

Wavelet is a set of mathematical functions which can scale and shift in a time period. The shifted and scaled mathematical function of wavelet is denoted as the mother wavelet, which can be expressed as in the following equation [17]:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

where b represents the translation parameter and a denotes the dilation or scaling parameter. Figure 1 depicts some basic wavelets.

2.2. Wavelet-based ramp event

The ramp event makes reference to a large variation in wind power output that is observed on a wind farm over a short period of time. To characterize a ramp event,

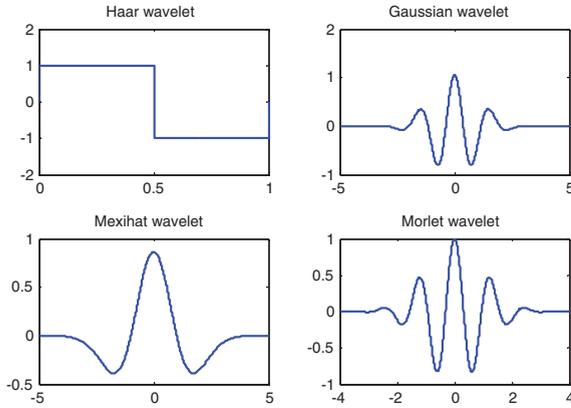


Fig. 1. Basic wavelets.

the following three parameters are usually employed [10, 23]: direction, duration and magnitude.

Many studies have relied on the definition of specific thresholds to decide whether a certain event can be considered a ramp in a binary manner. However, the thresholds which should be set by the end-user arise from different ramp definitions. Gallego and Costa developed a wavelet-based ramp event description which provides an alternative to avoiding the sensitivity inherent to the binary classification [12]. Figure 2 depicts the coefficients of wind power time series after haar wavelet transform.

As shown in Fig. 2, the ramp event of wind power time series is visually defined by the coefficients of wavelet transform, i.e. the coefficients of wavelet transform contain information regarding a ramp event. Therefore, the reconstruction signal of the coefficients was used to predict intervals in the following section.

3. Prediction interval formulation and neuro-fuzzy networks

3.1. Prediction intervals

Given a set of distinct pairs

$$T = \{(x_i, t_i) \in \mathbf{R}^{n+1}, i = 1, 2, \dots, N\} \quad (2)$$

the measured data can be modeled by a regression prediction model with noise for each input x_i ,

$$t_i = f(x_i) + \varepsilon(x_i) \quad (3)$$

where t_i is the i th measured target, function f denotes the underlying true regression function, x_i denotes rele-

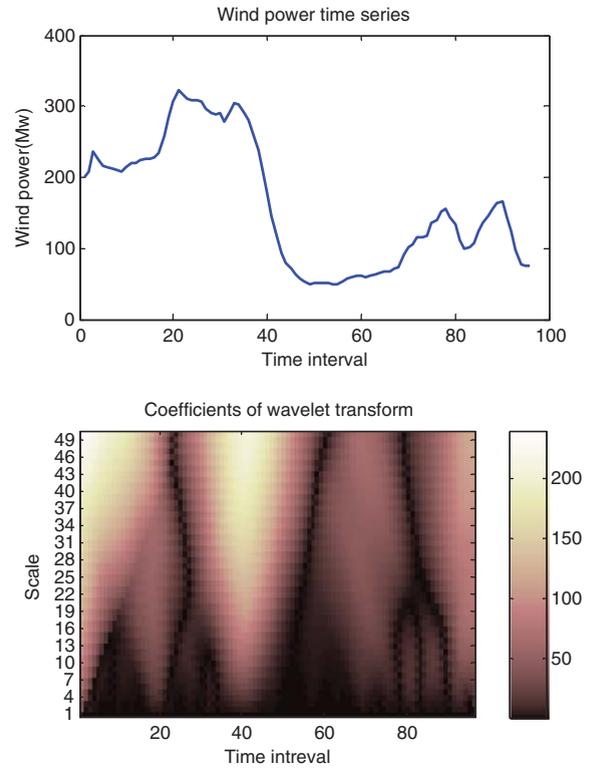


Fig. 2. Wind power time series and coefficients after haar wavelet transform.

vant input variables, and ε denotes the random variation. Theoretically, the trained regression function $\hat{y}(x_i)$ of a prediction model tries to fit the true regression function f , but there are some unavoidable errors between the output of true regression function f and the trained regression function $\hat{y}(x_i)$. Therefore, the prediction or forecasting error can be written as follows:

$$t_i - \hat{y}(x_i) = [f(x_i) - \hat{y}(x_i)] + \varepsilon(x_i) \quad (4)$$

where $t_i - \hat{y}(x_i)$ denotes the error between the measured data and the output of the trained regression function $\hat{y}(x_i)$, and $f(x_i) - \hat{y}(x_i)$ denotes the error between the trained regression function and the true regression function.

In real applications, the true regression function is difficult to obtain, so it is important to quantify the accuracy of our trained regression prediction model. For regression prediction problems, we can distinguish two different aspects: confidence intervals and prediction intervals. Confidence intervals are developed to measure the variance between the true regression function $f(x_i)$ and the trained regression function $\hat{y}(x_i)$, i.e. $f(x_i) - \hat{y}(x_i)$ in Equation (4). The prediction intervals

are developed to measure the variance between the measured data t_i and the output of the trained regression function $\hat{y}(x_i)$. Based on Equation (4), the variance of the total prediction errors σ_t^2 can be expressed as:

$$\sigma_t^2(x_i) = \sigma_{\hat{y}}^2(x_i) + \sigma_{\varepsilon}^2(x_i) \quad (5)$$

where the term $\sigma_{\hat{y}}^2(x_i)$ originates from the variance of prediction model uncertainty, and $\sigma_{\varepsilon}^2(x_i)$ is the noise variance between the measured data and the true regression function. As described by Equation (5), the prediction intervals contain the uncertainty of prediction model and the data noise.

A prediction interval with $(1 - \alpha)\%$ confidence level of the measured target t_i can be expressed as $I_t^\alpha(x_i)$:

$$I_t^\alpha(x_i) = [L(x_i), U(x_i)] \quad (6)$$

such that the coverage rate is equal to:

$$P(L(x_i) < t_i < U(x_i)) = 1 - \alpha \quad (7)$$

where $L(x_i)$ denotes the lower bound of the prediction interval and $U(x_i)$ denotes the upper bound of the prediction interval which, can be obtained as follows:

$$L(x_i) = \hat{y}(x_i) - z_{1-\alpha/2}(\sqrt{\sigma_t^2}) \quad (8)$$

$$U(x_i) = \hat{y}(x_i) + z_{1-\alpha/2}(\sqrt{\sigma_t^2}) \quad (9)$$

where $z_{1-\alpha/2}$ is the critical value of the standard Gaussian distribution.

3.2. Neuro-fuzzy network

Due to the universal approximation and learning capabilities, neural network models are often developed and used for point forecasting and prediction. A neuro-fuzzy network is “a fuzzy inference system that uses learning algorithm derived from or inspired by neural network theory to determine its parameters by processing data samples” [14]. The fuzzy inference system is defined by a set of given fuzzy rules which have formal structures of the form “IF... Then...” that associates a condition to a consequence by means of fuzzy sets.

The neuro-fuzzy network used as a prediction model for prediction intervals is the adaptive network-based fuzzy inference system [18] which is depicted in Fig. 3.

The rule set expressed by the model follows the schema:

IF x is A_i , THEN

$$y_1 = a_{i1}, \dots, y_m = a_{im} \quad (10)$$

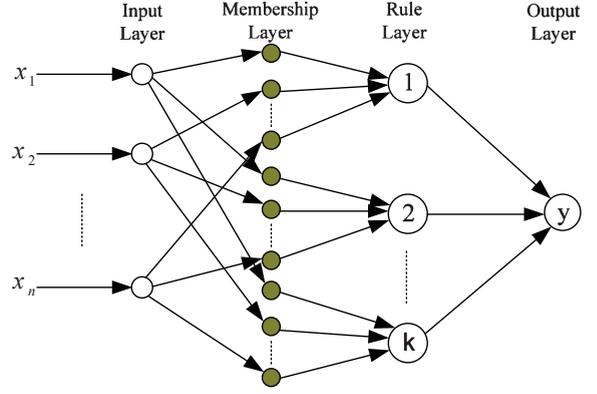


Fig. 3. Adaptive network-based fuzzy inference system.

where $x \in \mathbf{R}^n$ is an input variable, $A_i \in \mathbf{R}^n$ are n -dimensional fuzzy sets characterized by membership functions $\mu_{A_i} \in [0, 1]$, y_i are output variables and $a_{ij} \in \mathbf{R}$ are constants. When we assume the system has only one output, i.e. $m = 1$, multi-dimensional fuzzy sets are often represented as conjunctions of one-dimensional fuzzy sets for readability pursuits [13]. In such case, the input variable is decomposed into one-dimensional variables $x = (x_1, x_2, \dots, x_n)$ and the rule condition is rewritten as: $x_1 = A_{i1}, x_2 = A_{i2}, \dots, x_n = A_{in}$, where A_{ij} are one-dimensional fuzzy sets with membership functions $\mu_{A_{ij}} \in [0, 1]$ that are calculated in the membership layer. As a consequence, the multidimensional membership function is computed in the rule layer by aggregating one-dimensional membership functions with t-norm operator as follows:

$$\mu_{A_i} = \mu_{A_{i1}} \cdot \mu_{A_{i2}} \cdot \dots \cdot \mu_{A_{in}} \quad (11)$$

Once the membership values μ_{A_i} are available, the model output is inferred by the output layer with the following weighted average formula:

$$y = \frac{\sum_{i=1}^n y_i \times \mu_{A_i}}{\sum_{i=1}^n \mu_{A_i}} \quad (12)$$

In this paper, the more popular Takagi-Sugeno fuzzy rule [26] is used to replace the formal IF-THEN rule. The Takagi-Sugeno fuzzy rule can be expressed as:

IF x_1 is A_{i1}, x_2 is A_{i2}, \dots, x_n is A_{in} , THEN

$$y_i = a_{0i} + a_{1i}x_1 + \dots + a_{ni}x_n \quad (13)$$

4. Prediction intervals construction

4.1. Moving block bootstrap method

Compared to other approaches, the bootstrap approach is able to flexibly approximate the non-constant variance and heterogeneous noise, so bootstrap as an approach has been popularly used to obtain prediction intervals for neural network based methods [5]. However, the standard bootstrap assumes that the observations in the data are independent and identically distributed [7]. The moving block bootstrap method extends the standard bootstrap method in include correlated observations [15]. To maintain the temporal dependence in data, sequential observations are randomly sampled in blocks, rather than single observations. Mathematically, the overlapping blocks with l observations can be expressed as:

$$X_i = (x_i, x_{i+1}, \dots, x_{i+l-1}), i = 1, 2, \dots, n - l + 1$$

There will be $n-l+1$ blocks for a time series with n observations. Then, the sampled blocks are combined to form a time series of length n . One of the blocks must be truncated to form a sample of length n when n/l is not an integer. The entire procedure is repeated B times to form the datasets $\{D\}_{b=1}^B$ required for training models. A time series with a length of nine and a of block length three was used to illustrate the schematic diagram of moving block bootstrap in Fig. 4,

4.2. Prediction intervals construction

The bootstrap-based approach assumes that an ensemble of neural network models will reach approximation of true regression of the measured targets. According to this assumption, the true regression is estimated by averaging the point forecasts of B neural networks, expressed as:

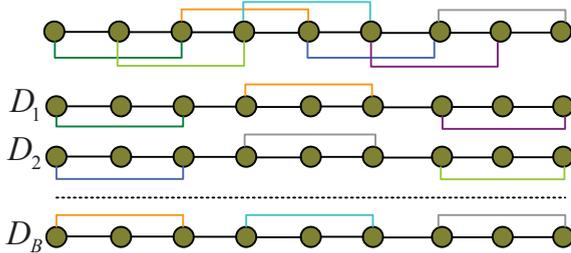


Fig. 4. The schematic diagram of moving block bootstrap.

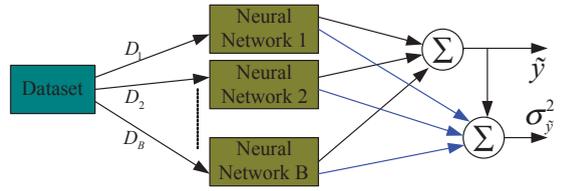


Fig. 5. The framework of B neural networks.

$$\tilde{y}(x_i) = \frac{1}{B} \sum_{b=1}^B \tilde{y}_b(x_i) \tag{14}$$

where $\tilde{y}_b(x_i)$ is the prediction value of the input samples generated by the b th bootstrapped neural network model.

The variance of model misspecification uncertainty can be estimated using the variance of B neural networks outputs:

$$\sigma_{\tilde{y}}^2(x_i) = \frac{1}{B-1} \sum_{b=1}^B (\tilde{y}_b(x_i) - \tilde{y}(x_i))^2 \tag{15}$$

The framework of B neural networks is depicted in Fig. 5.

From Equations (4) and (5), the variance of errors $\sigma_{\tilde{\epsilon}}^2$ can be calculated as follows:

$$\sigma_{\tilde{\epsilon}}^2(x_i) = E(t_i - \tilde{y}(x_i))^2 - \sigma_{\tilde{y}}^2(x_i) \tag{16}$$

A set of variance squared residuals is developed:

$$r_i^2 = \max((t_i - \tilde{y}_i)^2 - \sigma_{\tilde{y}}^2, 0) \tag{17}$$

Associated with the input x_i , we can obtain the transformed datasets:

$$D_{\tilde{\epsilon}} = \{x_i, r_i^2\}_{i=1}^B \tag{18}$$

Then, a new neural network model can be indirectly trained to estimate the unknown values of $\sigma_{\tilde{\epsilon}}^2(x_i)$, so as to maximize the probability of observing samples [27].

4.3. Prediction intervals evaluation

Prediction interval coverage probability (PICP) is introduced to evaluate the performance of prediction interval.

The PICP is measured by counting the number of target values covered by the constructed prediction intervals:

$$PICP = \frac{1}{n} \sum_{i=1}^n c_i \tag{19}$$

where n is the number of test samples, and c_i is the indicator of PICP, defined as follows:

$$c_i = \begin{cases} 1 & t_i \in [L_t^\alpha, U_t^\alpha] \\ 0 & t_i \notin [L_t^\alpha, U_t^\alpha] \end{cases} \quad (20)$$

5. Case study

In the case study, the wind power data of a wind farm with a 114 MW nominal generation capacity was used to test the proposed approach. First, the wind power series was bootstrapped and sampled to obtain the test datasets $\{x_i, t_i\}_{i=1}^B$. In order to contain the dependence of the wind power series, the autocorrelation function which can describe the degree of correlation at different time t_1, t_2 of stochastic signal was used to determine the length of the moving block. The autocorrelation function can be expressed as:

$$\begin{aligned} R(\tau) &= \int_{-\infty}^{\infty} f(t + \tau)f^*(t)dt \\ &= \int_{-\infty}^{\infty} f(t)f^*(t - \tau)dt \end{aligned} \quad (21)$$

The autocorrelation function of the wind power time series is shown in Fig. 6. As shown in Fig. 6, the wind power series has obvious dependence of 16 lengths time interval span, so we chose a moving block with 16 lengths.

The bootstrapped datasets were then processed by wavelet transform to obtain the coefficients to describe the ramp event of wind power. For example, the coefficients of one dataset after wavelet transform by Harr wavelet in shown in Fig. 7.

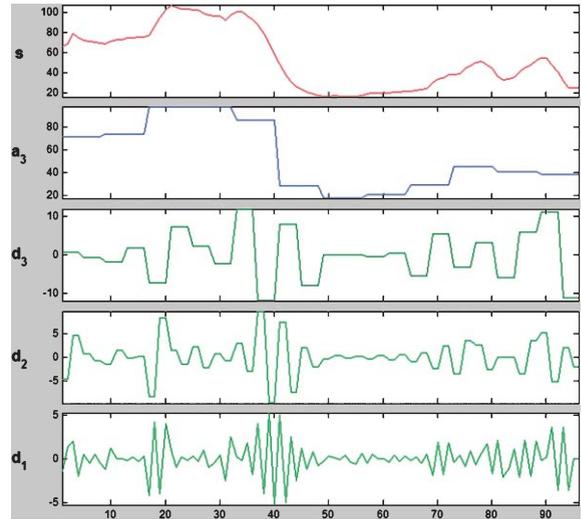


Fig. 7. Coefficients of three-level Harr wavelet transform.

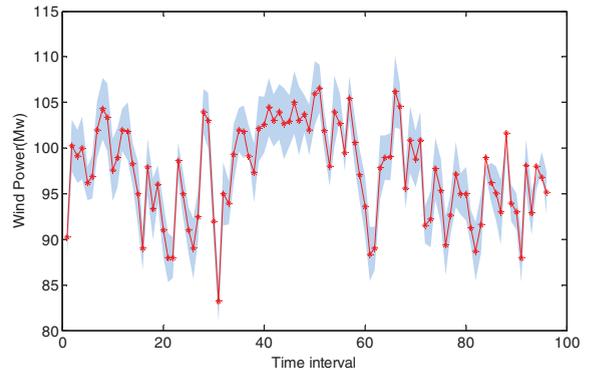


Fig. 8. The prediction intervals of wind power series.

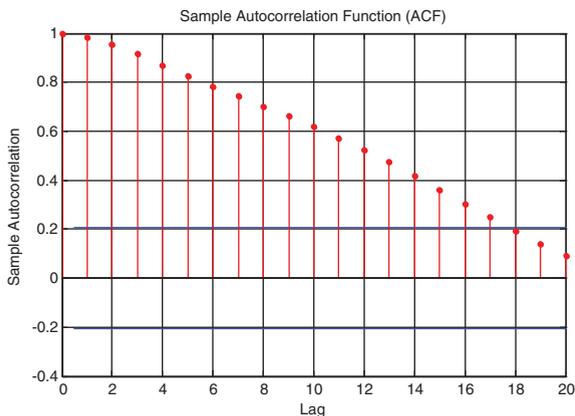


Fig. 6. The autocorrelation function of wind power series.

The coefficients with normal fuzzy membership function as input were trained by the fuzzy inference system. The prediction intervals with a 90% confidence level were predicted by the neural-fuzzy networks, as shown in Fig. 8.

6. Conclusion

Accurate and reliable forecasting of wind power is essential for the optimal management of wind farms and power system operations. However, due to the chaotic nature of meteorological systems and the non-stationarity of wind power series, traditional point forecasting can hardly be accurate. Prediction interval forecasting provides an effective solution to estimate

and quantify the potential impacts and risks facing system operation with wind penetration. In this paper, a neuro-fuzzy network model was proposed to estimate the prediction intervals of wind power. In the model, ramp events which have become a major issue in wind power generation can be flexibly expressed based on the shifted and scaled wavelet. The moving block bootstrap method with a matched length is more reasonable for the dependence of wind power stochastic series. The construction of prediction intervals with a confidence level of 90% based on historical wind power series shows the effectiveness of the neuro-fuzzy network model.

References

- [1] A.M. Foley, P.G. Leahy, A. Marvuglia and E.J. Mckeogh, Current methods and advances in forecasting of wind power generation, *Renewable & Sustainable Energy Reviews* **13** (2009), 915–920.
- [2] A. Khosravi, S. Nahavandi and D. Creighton, Prediction intervals for short-term wind farm power generation forecasts, *IEEE Transactions on Sustainable Energy* **3** (2013), 602–610.
- [3] A. Costa, A. Crespo, J. Navarro, et al., A review on the young history of the wind power short-term prediction, *Renewable & Sustainable Energy Reviews* **12** (2008), 1725–1744.
- [4] A. Bossavy, R. Girard and G. Kariniotakis, Forecasting uncertainty related to ramps of wind power production, *Journal of Climate & Applied Meteorology* **2** (2010), 1–9.
- [5] A. Khosravi, S. Nahavandi, D. Creighton and D. Srinivasan, Optimizing the quality of bootstrap-based prediction intervals, *Proceedings of International Joint Conference on Neural Networks* **42**(4) (2011), 3072–3078.
- [6] B. Zhang and Z. Dong, An adaptive neural-wavelet model for short term load forecasting, *Electric Power Systems Research* **59** (2001), 121–129.
- [7] B. Efron and R.J. Tibshirani, *An introduction to the bootstrap*, Chapman and Hall, 1993.
- [8] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang and G. Conzelmann, Wind power forecasting: State-of-the-art 2009, *Argonne National Laboratory, ANL/DIS-10-1, Decision and Information Sciences Division* **32**(2) (2009), 124–130.
- [9] C. Wan, Z. Xu, P. Pinson, Z. Dong and K. Wong, Probabilistic forecasting of wind power generation using extreme learning machine, *IEEE Transactions on Power Systems* **3** (2014), 1033–1044.
- [10] C. Ferreira, J. Gama, L. Matias, A. Botterud and J. Wang, A survey on wind power ramp forecasting, in *Report ANL/DIS 10-13*, Argonne National Laboratory, 2010.
- [11] C. Wan, Z. Xu, P. Pinson, Z. Dong and K. Wong, Optimal prediction intervals of wind power generation, *IEEE Transactions on Power Systems* **29** (2014), 1166–1174.
- [12] C. Gallego, A. Costa, A. Cuerva, L. Landberg, B. Greaves and J. Collins, A wavelet-based approach for large wind power ramp characterization, *Wind Energy* **16** (2013), 257–278.
- [13] C. Mencar, G. Castellano and A. Fanelli, Deriving prediction intervals for neuro-fuzzy networks, *Mathematical and Computer Modelling* **42** (2005), 719–726.
- [14] D. Nauck, F. Klawonn and R. Kruse, *Foundations of neuro-fuzzy systems*, Wiley, 1997.
- [15] E. Carlstein and H.R. Künsch, Matched-block bootstrap for dependent data, *Bernoulli* **4** (1998), 305–328.
- [16] G.N. Kariniotakis, G.S. Stavratskis and E.F. Nogaret, Wind power forecasting using advanced neural networks models, *IEEE Transactions on Energy Conversion* **4** (1996), 762–767.
- [17] I. Daubechies, *Ten lectures on wavelets*, Society for Industrial and Applied Mathematics Press, 1992.
- [18] J. Shing and R. Jang, ANFIS: Adaptive-network-based fuzzy inference system, *IEEE Transaction on Systems Man & Cybernetics* **23** (1993), 665–684.
- [19] K. Amolins, Y. Zhang and P. Dare, Wavelet based image fusion techniques—an introduction, review and comparison, *Journal of Photogrammetry and remote sensing* **62** (2007), 249–263.
- [20] K. Zhu, Y. Wong and G. Hong, Wavelet analysis of sensor signals for tool condition monitoring: A review and some new results, *International Journal of Machines Tools and Manufacture* **49** (2009), 537–553.
- [21] L. Milidiu, R. Machado and R. Renteria, Time-series forecasting through wavelets transformation and a mixture of expert models, *Neurocomputing* **28** (1999), 145–156.
- [22] M.A. Mohandes, T.O. Halawani, S. Rehman and A. Hassian, Support vector machines for wind speed prediction, *Renewable Energy* **29** (2004), 939–947.
- [23] O. Tinghui, Z. Xiaoming and Q. Liang, A survey of wind power ramp forecasting, *Energy and Power Engineering* **5** (2013), 368–372.
- [24] S. Mallat and W. Hwang, Singularity detection and processing with wavelets, *IEEE Transactions on Information Theory* **38** (1992), 617–643.
- [25] T.G. Barbounis and J.B. Theocharis, A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation, *Neurocomputing* **70** (2007), 1525–1542.
- [26] T. Takagi and M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, *IEEE Transactions on Systems Man & Cybernetics* **15** (1985), 116–132.
- [27] T. Heskes, Practical confidence and prediction intervals, *Advances in Neural Information Processing Systems* **9** (1997), 176–182.
- [28] Z. Tan, J. Wang and J. Xu, Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models, *Applied Energy* **87** (2010), 3606–3610.