The Journal of Engineering

The 7th International Conference on Renewable Power Generation (RPG 2018)

Spatio-temporal Markov chain model for very-short-term wind power forecasting

eISSN 2051-3305
Received on 29th October 2018
Accepted on 10th January 2019
E-First on 18th June 2019
doi: 10.1049/joe.2018.9294
www.ietdl.org

Yongning Zhao[1] ✉, Lin Ye[1], Zheng Wang[2], Linlin Wu[3], Bingxu Zhai[4], Haibo Lan[4], Shihui Yang[5]

[1]*College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, People's Republic of China*
[2]*China Electric Power Research Institute, Haidian District, Beijing 100192, People's Republic of China*
[3]*Electric Power Research Institute, State Grid Jibei Electric Company Limited, Xicheng District, Beijing 100045, People's Republic of China*
[4]*State Grid Jibei Electirc Company Limited, Xicheng District, Beijing 100053, People's Republic of China*
[5]*Global Energy Interconnection Research Institute Co, LTD, No.18 Binhe Road, Future Science City, Changping, Beijing 102209, People's Republic of China*
✉ *E-mail: zyn@cau.edu.cn*

Abstract: Wind power forecasting (WPF) is crucial in helping schedule and trade wind power generation at various spatial and temporal scales. With increasing number of wind farms over a region, research focus of WPF methods has been recently moved onto exploring spatial correlation among wind farms to benefit forecasting. In this study, a spatio-temporal Markov chain model is proposed for very-short-term WPF by extending the traditional discrete-time Markov chain and incorporating off-site reference information to improve forecasting accuracy of regional wind farms. Not only are the transitions between the power output states of the target wind farm itself considered in the forecasting model, but also the transitions from the output states of reference wind farms to that of the target wind farm are introduced. The forecasting results derived from multiple spatio-temporal Markov chains regarding different reference wind farms over the same region are optimally weighted using sparse optimisation to generate forecasts of the target wind farm. The proposed method is validated by comparing with both local and spatio-temporal WPF methods, using a real-world dataset.

1 Introduction

Wind energy has seen its rapid growing around the world over the past few decades. According to the Global Wind Energy Council, the global cumulative installed wind power capacity increased from 2.4 GW in 2001 to 487 GW in 2016 [1]. However, wind power has very limited dispatchability and controllability, due to the naturally stochastic and intermittent features of wind [2]. Thus, large-scale wind power integration has brought big challenges to the safety and stability of power system operations, since the power supply and demand has to be balanced in real time [3]. Wind power forecasting (WPF) is generally recognised as one of the most effective tools to deal with the problems brought by wind power, via helping both system operators and market traders make optimal decisions under uncertainties [4].

The WPF has been widely researched and a variety of methods as well as commercial products have been developed for power system and electricity markets applications. The traditional WPF methods are mainly classified into statistical, physical and hybrid groups. Review of details about different classes of WPF methods are given among others in [5–7]

Most of the existing WPF methods only consider local onsite information (e.g. historical measurements, weather forecasts) of the target wind farm to be forecasted. This kind of methods are essentially based on temporal correlation of time series and generally straightforward and simple to implement, but obviously, the forecasting accuracy will be limited due to insufficient input information. Nevertheless, the ever-increasing number of wind farms over a region is bringing great opportunities for further improving the WPF accuracy. In fact, spatial correlation between regional wind farms always exists as a result of spatially coherent evolution of a weather system [8]. Thus, it is intuitively beneficial by using off-site information from neighbouring or correlated wind farms as inputs to forecast a target wind farm [9, 10]. To this end, many researchers has been focusing on exploring spatial-temporal interdependence structures between wind farms and applying them to improve the WPF performance [11].

Initially, different artificial intelligence models [12, 13] are developed for both wind speed and WPF using data measured at several neighbouring sites. Moreover, regime-switching space-time methods [14], multichannel adaptive filters [15] and graph learning analysis [16] are also among approaches proposed for spatio-temporal WPF. These methods are feasible only when very few target wind farms are to be forecasted since tuning of their model structures for different wind farms is an important but very tedious process. Therefore, they are almost impracticable in the case of a modern power system with tens or even hundreds of wind farms.

To this end, multivariate linear regression approach is introduced for efficient high-dimensional WPF. To avoid over-fitting problems and make models more interpretable, sparse modelling techniques are usually used to force regression coefficients of some less important variables to be zeros. Typically, sparse vector autoregression (VAR) models are frequently studied in spatio-temporal WPF. Dowell and Pinson [17] applied a two-stage sparse VAR for very-short-term probabilistic WPF by using the partial spectral coherence and some basic statistics to determine zero coefficients. Cavalcante *et al.* [18] described a forecasting methodology that explores a set of different sparse structures for VAR models based on the Least Absolute Shrinkage and Selection Operator (LASSO) framework. Zhao *et al.* [19] presented a correlation-constrained and sparsity-controlled VAR model by transforming the VAR optimisation into a mixed-integer non-linear programming, which allows both freely controlling sparsity and incorporating expert knowledge on spatial correlation into the forecasting model.

However, regardless of that various VAR-based model were developed for spatio-temporal WPF, the Markov chain has not attracted much attention in this field. A Markov chain is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. The discrete-time Markov chain has been already used in relevant literature for the generation of synthetic wind speed and wind power time series [20, 21] as well as very-short-term WPF [22, 23] and achieved good performance. However, as has been mentioned, the traditional Markov chains

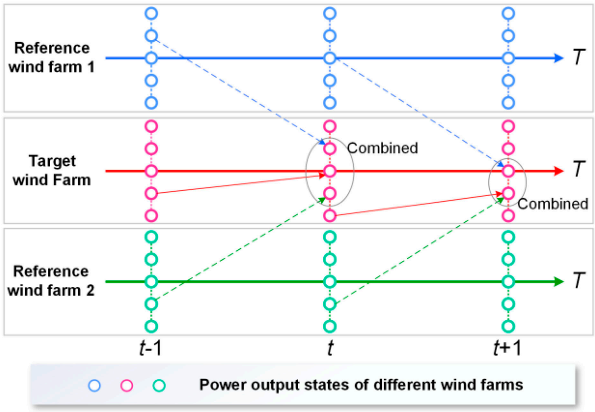

Fig. 1 *Principle illustration of the STMC*

were only used in local WPF by considering only on-site historical data.

In geostatistical modelling, the spatial Markov chain, or the Markov chain geostatistics was developed based on the multi-dimensional Markov chain random field theory [24]. The spatial Markov chain moves or jumps in a space and decides its state at any unobserved location through interactions with its nearest known neighbours in different directions. Inspired by this idea, a first-order discrete spatio-temporal Markov chain (STMC) model is developed for very-short-term WPF. In this model, the spatio-temporal transitions between the wind power output states of reference wind farms and a target wind farm are additionally considered as an extension of traditional Markov chain that only considers temporal states transitions. The forecasts derived from spatio-temporal Markov regarding multiple reference wind farms are optimally combined using sparse modelling to obtain the final forecasts for target wind farm. A case study is carried out to demonstrate the effectiveness of the proposed method.

2 Spatio-temporal Markov chain model

The discrete-time Markov chain (TMC) is a stochastic process with the property that if, for all $t \geq 1$, the probability distribution of state S_{t+1} is determined by the state $S_t$of the process at time t, and does not depend on the past values of S_x for $x \leq t-1$.

The STMC is essentially an extension of the TMC. It considers the transition from the previous output state of a reference wind farm to the next state of a target wind farm, but not only consider the temporal state transitions of the target wind farm itself. The principle of the STMC is illustrated in Fig. 1.

In the following, all the Markov chains are discussed in terms of the first-order and discrete models.

2.1 Partition of wind power output states

Suppose N wind farms are spatially dispersed over a region. The wind power output of wind farm i at time t is $y_{i,t}$, $i = 1,2,\ldots,N$. Before building a Markov chain model, all the wind power values are normalised into the range of [0,1] by their corresponding nominal wind farm capacity:

$$p_{i,t} = \frac{y_{i,t}}{C_i} \tag{1}$$

where $p_{i,t}$ is the normalised wind power output of $y_{i,t}$, and C_i is the nominal installed capacity of wind farm i.

A state is referred to as an interval and the wind power values that lie in a same interval are regarded as the same state. In this paper, all the intervals have same width. With a predefined state width, the state of wind farm i can be expressed as

$$S_i^k = \begin{cases} [(k-1)d, kd), & 1 \leq k \leq k-1 \\ [(k-1)d, kd], & k = k \end{cases} \tag{2}$$

where K is the total number of states of a wind farm, and d is the width of a state interval. Note that the state width or the number of states for different wind farms could be set as different values, but here they are unified as the same value for all wind farms. The resulted states partition for wind farm i is $\{S_i^1, S_i^2, \ldots, S_i^K\}$ and their corresponding state intervals are $\{[0,\ 1/K),\ [1/K,\ 2/K),\ \ldots,\ [(K-1)/K,\ 1]\}$.

Then the representative value of a state can be obtained by averaging the wind power values within its corresponding interval

$$s_i^k = \frac{\sum p_{i,t}}{D_i^k}, p_{i,t} \in S_i^k \tag{3}$$

where the D_i^k is the total number of values within the interval of S_i^k. The representative states vector of wind farm i is $\boldsymbol{s}_i = (s_i^1, s_i^2, \ldots, s_i^K)^{\mathrm{T}}$.

2.2 Spatio-temporal transition probability matrices

To obtain the forecasts for target wind farm i using a reference wind farm j, the spatio-temporal transition probability matrix between the two wind farms needs to be generated by training data.

Based on the training data and the results of states partition, count the number T_{ji}^{km} of occurrences of the spatio-temporal state transitions that the output state of reference wind farm j at time t is S_j^k and the output state of target wind farm i at time $t+1$ is S_i^m, $m = 1,2,\ldots, K$. In addition, also count the number T_{ji}^{km} of occurrences of the state transitions that the output state of reference wind farm j at time t is S_j^k and the output state of target wind farm i at time $t+1$ is any value. Then the probability of state transition from S_j^k of wind farm j to S_i^m of wind farm j can be calculated by

$$q_{ji}^{km} = \frac{T_{ji}^{km}}{T_{ji}^{k}} \tag{4}$$

It should be noted that in the above procedure it allows $i = j$. In that case, the procedure will degenerate into generating pure temporal state transition probabilities regarding the target wind farm itself, which is exactly the traditional discrete-time Markov chain.

The above procedure is repeated until all possible transitions between the states given by the state partition are retrieved and counted. The spatio-temporal state transition probability matrix from wind farm j to wind farm i can be achieved as

$$\boldsymbol{Q}_{ji} = \begin{bmatrix} q_{ji}^{11} & q_{ji}^{12} & \cdots & q_{ji}^{1K} \\ q_{ji}^{21} & q_{ji}^{22} & \cdots & q_{ji}^{2K} \\ \vdots & \vdots & \ddots & \vdots \\ q_{ji}^{K1} & q_{ji}^{K2} & \cdots & q_{ji}^{KK} \end{bmatrix} \tag{5}$$

where $\boldsymbol{Q}_{ji} \in \mathbf{R}^{K \times K}$. When $i = j$, $\boldsymbol{Q}_{ji}$ will become temporal transition probability matrix. As the state transition from wind farm j to wind farm i is not equivalent to the state transition from wind farm i to wind farm j, the derived matrix $\boldsymbol{Q}_{ji}$ is not necessarily equal to $\boldsymbol{Q}_{ij}$.

The same approach can be applied to generate the spatio-temporal transition probability matrices $\boldsymbol{Q}_{ji}(j = 1, 2, \ldots, N)$ between target wind farm i and all reference wind farms (including the target wind farm itself).

2.3 WPF model using spatio-temporal Markov chain

Based on the generated transition matrix $\boldsymbol{Q}_{ji}$, the representative states vector $\boldsymbol{s}_i$ of target wind farm i, and the output state S_{jt} of wind farm j at time t, the one-step-ahead forecast for wind farm i based on wind farm j is formulated as

$$\hat{p}_{i,t+1|j,t} = \boldsymbol{Q}_{ji}^{r_{j,t}} \cdot \boldsymbol{s}_j \tag{6}$$

where $r_{j,i}$ is the index of the state S_{jt}, and $\boldsymbol{Q}_{ji}^{r_{j,t}}$ is the $r_{j,t}$th row of transition matrix $\boldsymbol{Q}_{ji}$.

The forecasts for the target wind farm i based on all the reference wind farms with their corresponding spatio-temporal state transition matrices are acquired by using (6). The final forecast for target wind farm is obtained by weighting all the forecasts regarding N reference wind farms, i.e.

$$\begin{aligned}\widehat{p}_{i,t+1} &= \beta_1^i \widehat{p}_{i,t+1|1,t} + \cdots + \beta_j^i \widehat{p}_{i,t+1|j,t} + \cdots + \beta_N^i \widehat{p}_{i,t+1|N,t} \\ &= \boldsymbol{\beta}_i^{\mathrm{T}} \widehat{\boldsymbol{p}}_{i,t+1|N}\end{aligned} \tag{7}$$

where β_j^i is the weighting coefficient for quantifying the devotion of reference wind farm j to wind farm i, $\boldsymbol{\beta} \in \boldsymbol{R}^N$ is weighting coefficient vector, and $\widehat{p}_{i,t+1|N} = (\widehat{p}_{i,t+1|1,t}, \widehat{p}_{i,t+1|2,t}, \ldots, \widehat{p}_{i,t+1|N,t}) \in \boldsymbol{R}^N$ is the forecasts vector whose entries are the forecasts made by all N reference wind farms.

The forecasting error at time $t+1$ is

$$\begin{aligned}e_{i,t+1} &= p_{i,t+1} - \widehat{p}_{i,t+1} \\ &= p_{i,t+1} - \boldsymbol{\beta}_i^{\mathrm{T}} \widehat{\boldsymbol{p}}_{i,t+1|N}\end{aligned} \tag{8}$$

To optimise the coefficient vector of the forecasting model, the sum of squared errors should be optimised over the period of training dataset, that is

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_i = \min_{\boldsymbol{\beta}_i} E_i &= \min_{\boldsymbol{\beta}_i} \sum_{t=1}^{L-1} (e_{i,t+1})^2 \\ &= \min_{\boldsymbol{\beta}_i} \sum_{t=1}^{L-1} (p_{i,t+1} - \boldsymbol{\beta}_i^T \widehat{\boldsymbol{p}}_{i,t+1|N})^2\end{aligned} \tag{9}$$

where L is the length of training time series. As the forecasting model involves a large number of wind farms, a l_1-norm regularisation is further added to help avoid over-fitting and make the model more computationally efficient and interpretable. Equation (9) is modified as

$$\widehat{\boldsymbol{\beta}}_i = \min_{\boldsymbol{\beta}_i} \sum_{t=1}^{L-1} (p_{i,t+1} - \boldsymbol{\beta}_i^{\mathrm{T}} \widehat{\boldsymbol{p}}_{i,t+1|N})^2 + \lambda \parallel \boldsymbol{\beta}_i \parallel_1 \tag{10}$$

where $\parallel \cdot \parallel$ is the l_1-norm of a vector, and λ is shrinkage parameter that balances between the estimation error and the degree of sparsity of the solution.

There have been many algorithms for solving optimisation problems like (10), such as coordinate descent algorithm and alternating direction method of multipliers.

3 Results and discussions

3.1 Data preparation and benchmarking models

A whole-year wind power dataset of 100 wind farms over a region is used to testify the proposed methods. The time resolution is 15 min. Each wind farm time series contains 35,040 data points and is divided into three consecutive parts, including 10,000 data points for training, 10,000 data points for validation (parameter optimisation) and the remaining 15,040 data points for out-of-sample testing.

Different forecasting methods are also implemented as benchmarks, which include local WPF models and spatio-temporal models. The local models include persistence model and autoregression (AR) model, which use only local onsite information while the spatio-temporal models including VAR model and LASSO-based VAR (LASSO-VAR) model use both onsite and offsite spatial information to improve forecasting accuracy. The readers are referred to [18] for mode detail about the LASSO-VAR in WPF. Both the STMC and the LASSO-VAR models are solved by the Matlab package 'Glmnet'. The VAR

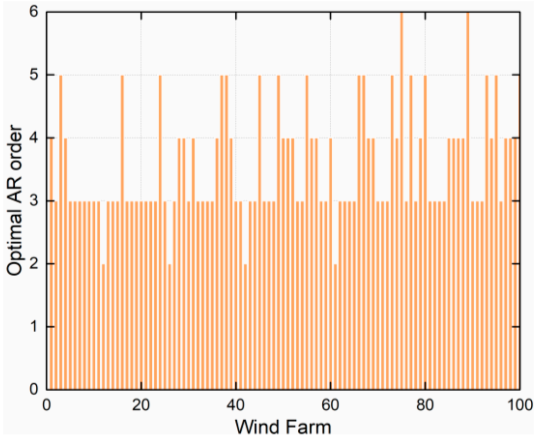

Fig. 2 *Optimal order of AR model for each wind farm*

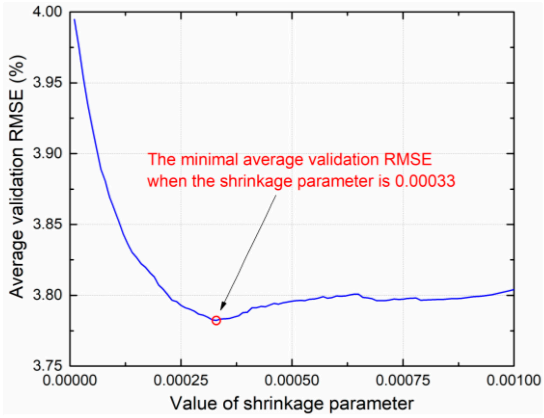

Fig. 3 *Average validation RMSE of the LASSO-VAR with varying value of its shrinkage parameter*

results are achieved by setting the shrinkage parameter as zero. The root mean squared error (RMSE) and mean absolute error (MAE) are used as metrics to evaluate these methods [25].

It is recommended that the wind power time series should be transformed to Gaussian ones by using Logit transformation [26]. However, the ranges of wind power values may be different for different wind farms, which can make it inconvenient to demonstrate the proposed method. Thus, the Logit transformation is not applied in this paper for any WPF model for the sake of fair comparisons.

3.2 Parameter setting

Partial auto-correlogram is an effective tool to determine the orders of AR time series models. The partial auto-correlogram calculated from training data is used for each of the 100 wind farms to determine their optimal orders, which are shown in Fig. 2.

It is shown that the AR models for different wind farms have different optimal orders. Each AR model is trained using its corresponding optimal order. However, a spatio-temporal model requires a unified order for all wind farms, thus the optimal orders of VAR and LASSO-VAR are set as a moderate value of 3, according to Fig. 2.

Another important parameter of LASSO-VAR is its shrinkage parameter. This parameter is optimised using validation data. The trend of average validation RMSE of 100 wind farms with the varying values of the shrinkage parameter is given in Fig. 3. The RMSE firstly decreases and then increases, with the optimal value achieved when the parameter is 0.00033. So, the optimal shrinkage parameter of the LASSO-VAR is set as 0.00033.

The trend of average validation RMSE of STMC model with varying value of λ is also given in Fig. 4.

It can be seen that the trend of average validation RMSE of the STMC is very similar to that of the LASSO-VAR. The RMSE reaches its minimum at $\lambda = 0.00025$. Thus, the λ is finally set as

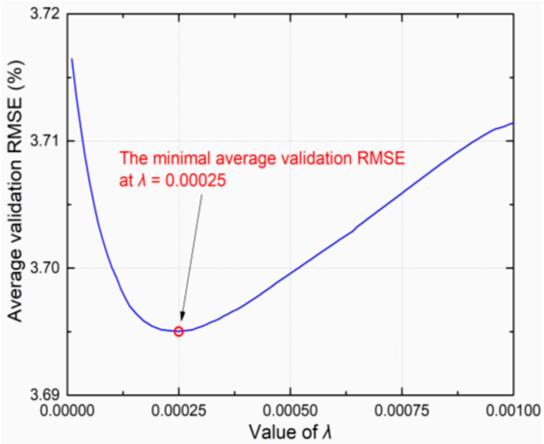

Fig. 4 *Average validation RMSE of the STMC with varying value of λ*

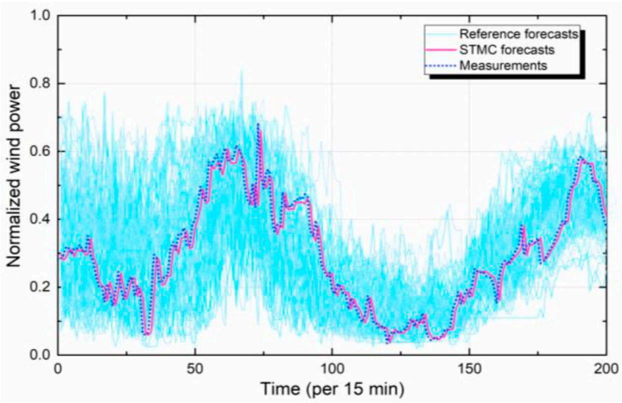

Fig. 5 *Selected segment of STMC forecasting results*

Table 1 Average testing RMSE and MAE of different methods

Methods	Average RMSE, %	Average MAE, %
persistence	5.1446	3.0013
AR	5.1160	2.9868
VAR	7.2854	4.9009
LASSO-VAR	5.6989	3.6178
STMC	4.8758	2.9652

0.00025. The other parameter of the STMC, i.e. number of partitioned states is set as $K = 100$.

3.3 Forecasting results and analysis

A selected segment of the final STMC forecasting results for a specific wind farm is demonstrated in Fig. 5, along with the measurements and 100 reference forecast curves obtained by the spatio-temporal transition probability matrix regarding each reference wind farm.

It can be seen that the measurements and the STMC forecasts can be well encompassed by the 100 reference forecast curves. The STMC forecasting model can extract useful patterns from abundant reference information. Therefore, it is expected that the STMC forecasts for the target wind farm can be improved by optimising the weights of these individual forecasts.

The average testing RMSE and MAE of 100 wind farms by using different methods are provided in Table 1. The proposed STMC is the best one among all these forecasting methods. The bad performance of the VAR model can be explained by the over-fitting problem in high-dimensional modelling. However, the LASSO-VAR is always reported to perform quite well in literature, while in this case it does not behave as expected.

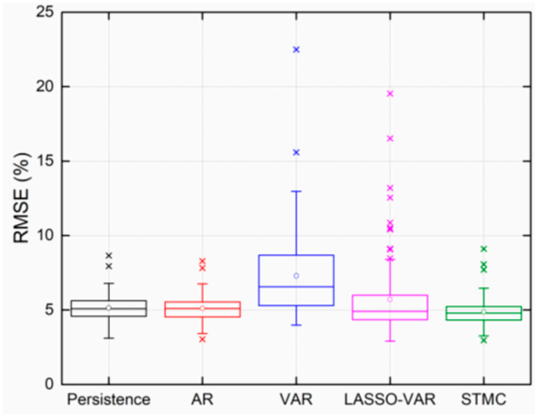

Fig. 6 *RMSE boxplots of 100 wind farms for different methods*

In order to provide a more comprehensive predictive performance analysis, the box plots of RMSE of 100 wind farms for different methods are depicted in Fig. 6.

The stars in the figure are outliers and the circles are the average RMSE, which have been given in Table 1. It is shown that the VAR and the LASSO-VAR present very large forecasting errors for several wind farms, in comparison with other methods. This is the main reason that leads to higher average RMSE values of the two methods. However, it should be noted that the accuracies of the LASSO-VAR for at least half number of wind farms are competitive with that of the STMC.

4 Conclusions

In this paper, a spatio-temporal Markov chain model is proposed for very-short-term WPF of a large number of wind farms over a region. This model is obtained by extending the traditional discrete-time Markov chain into a spatio-temporal framework, which aims to apply spatial correlation between wind farms to benefit forecasting.

The proposed model is verified on a real-world data from 100 wind farms. The model is more accurate and more stable in terms of forecasting errors when compared with other methods. Some further work could be done to improve or extend the proposed method:

(i) The impact of the number of partitioned states on the predictive performance needs to be investigated to find the optimal number of states.
(ii) The number of partitioned states can be set as different values for different wind farms.
(iii) The forecasting in longer time horizons needs to be evaluated.

5 Acknowledgments

This work was supported by the National Key R&D Program of China (Technology and application of wind power / photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200) and eponymous Complement S&T Program of State Grid Corporation of China (SGLNDKOOKJJS1800266).

6 References

[1] Sawyer, S., Dyrholm, M.: '*Global wind report-annual market update 2016*' (Global Wind Energy Council (GWEC), Brussels, Belgium, 2016)

[2] Li, Z., Ye, L., Zhao, Y., *et al.*: 'Short-term wind power prediction based on extreme learning machine with error correction', *Prot. Control Mod. Power Syst.*, 2016, **1**, (1), pp. 1–8

[3] Zhao, Y., Ye, L., Wang, W., *et al.*: 'Data-driven correction approach to refine power curve of wind farm under wind curtailment', *IEEE Trans. Sustain. Energy*, 2018, **9**, (1), pp. 95–105

[4] Pinson, P.: 'Wind energy: forecasting challenges for its operational management', *Stat. Sci.*, 2013, **28**, (4), pp. 564–585

[5] Jung, J., Broadwater, R.P.: 'Current status and future advances for wind speed and power forecasting', *Renew. Sustain. Energy Rev.*, 2014, **31**, pp. 762–777

[6] Zhang, Y., Wang, J., Wang, X.: 'Review on probabilistic forecasting of wind power generation', *Renew. Sustain. Energy Rev.*, 2014, **32**, pp. 255–270

[7] Tascikaraoglu, A., Uzunoglu, M.: 'A review of combined approaches for prediction of short-term wind speed and power', *Renew. Sustain. Energy Rev.*, 2014, **34**, pp. 243–254

[8] Girard, R., Allard, D.: 'Spatio-temporal propagation of wind power prediction errors', *Wind Energy*, 2013, **16**, pp. 999–1012

[9] Tastu, J., Pinson, P., Kotwa, E., *et al.*: 'Spatio-temporal analysis and modeling of short-term wind power forecast errors', *Wind Energy*, 2011, **14**, pp. 43–60

[10] Ye, L., Zhao, Y., Zeng, C., *et al.*: 'Short-term wind power prediction based on spatial model', *Renew. Energy*, 2017, **101**, pp. 1067–1074

[11] Tastu, J., Pinson, P., Trombe, P.J., *et al.*: 'Probabilistic forecasts of wind power generation accounting for geographically dispersed information', *IEEE Trans. Smart Grid*, 2014, **5**, (1), pp. 480–489

[12] Damousis, I.G., Alexiadis, M.C., Theocharis, J.B., *et al.*: 'A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation', *IEEE Trans. Energy Convers.*, 2004, **19**, (2), pp. 352–361

[13] Barbounis, T.G., Theocharis, J.B.: 'Locally recurrent neural networks for wind speed prediction using spatial correlation', *Inf. Sci.*, 2007, **177**, pp. 5775–5797

[14] Gneiting, T., Larson, K., Westrick, K., *et al.*: 'Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method', *J. Am. Stat. Assoc.*, 2006, **101**, pp. 968–979

[15] Dowell, J., Weiss, S., Hill, D., *et al.*: 'Short-term spatio-temporal prediction of wind speed and direction', *Wind Energy*, 2014, **17**, pp. 1945–1955

[16] He, M., Yang, L., Zhang, J., *et al.*: 'A spatio- temporal analysis approach for short-term forecast of wind farm generation', *IEEE Trans. Power Syst.*, 2014, **29**, (4), pp. 1611–1622

[17] Dowell, J., Pinson, P.: 'Very-short-term probabilistic wind power forecasts by sparse vector autoregression', *IEEE Trans. Smart Grid*, 2016, **7**, (2), pp. 763–770

[18] Cavalcante, L., Bessa, R.J., Reis, M., *et al.*: 'LASSO vector autoregression structures for very short- term wind power forecasting', *Wind Energy*, 2017, **20**, (4), pp. 657–675

[19] Zhao, Y., Ye, L., Pinson, P., *et al.*: 'Correlation-Constrained and sparsity-controlled vector autoregressive model for spatio-temporal wind power forecasting', *IEEE Trans. Power Syst.*, 2018, **33**, (5), pp. 5029–5040, doi: 10.1109/TPWRS.2018.2794450

[20] Shamshad, A., Bawadi, M.A., Hussin, W.M.A.W., *et al.*: 'First and second order markov chain models for synthetic generation of wind speed time series', *Energy*, 2005, **30**, (5), pp. 693–708

[21] Tagliaferri, F., Hayes, B.P., Viola, I.M., *et al.*: 'Wind modelling with nested markov chains', *J. Wind Eng. Ind. Aerodyn.*, 2016, **157**, pp. 118–124

[22] Carpinone, A., Giorgio, M., Langella, R., *et al.*: 'Markov chain modeling for very-short-term wind power forecasting', *Electr. Power Syst. Res.*, 2015, **122**, pp. 152–158

[23] Yoder, M., Hering, A.S., Navidi, W.C., *et al.*: 'Short-term forecasting of categorical changes in wind power with markov chain models', *Wind Energy*, 2014, **17**, (9), pp. 1425–1439

[24] Li, W.D.: 'Markov chain random fields for estimation of categorical variables', *Math. Geol.*, 2007, **39**, (3), pp. 321–335

[25] Zhao, Y.N., Ye, L., Li, Z., *et al.*: 'A novel bidirectional mechanism based on time series model for wind power forecasting', *Appl. Energy*, 2016, **177**, pp. 793–803

[26] Pinson, P.: 'Very-short-term probabilistic forecasting of wind power with generalized logit-normal distributions', *J. Royal Stat. Soc. C*, 2012, **61**, (4), pp. 555–576