ARTICLE

Inferential Statistics and Machine Learning Models for Short-Term Wind Power Forecasting

Ming Zhang, Hongbo Li and Xing Deng[*]

School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang, 212003, China
*Corresponding Author: Xing Deng. Email: xdeng@just.edu.cn

ABSTRACT

The inherent randomness, intermittence and volatility of wind power generation compromise the quality of the wind power system, resulting in uncertainty in the system's optimal scheduling. As a result, it's critical to improve power quality and assure real-time power grid scheduling and grid-connected wind farm operation. Inferred statistics are utilized in this research to infer general features based on the selected information, confirming that there are differences between two forecasting categories: Forecast Category 1 (0–11 h ahead) and Forecast Category 2 (12–23 h ahead). In z-tests, the null hypothesis provides the corresponding quantitative findings. To verify the final performance of the prediction findings, five benchmark methodologies are used: Persistence model, LMNN (Multilayer Perceptron with LM learning methods), NARX (Nonlinear autoregressive exogenous neural network model), LMRNN (RNNs with LM training methods) and LSTM (Long short-term memory neural network). Experiments using a real dataset show that the LSTM network has the highest forecasting accuracy when compared to other benchmark approaches including persistence model, LMNN, NARX network, and LMRNN, and the 23-steps forecasting accuracy has improved by 19.61%.

KEYWORDS

Wind power forecasting; correlation analysis; inferential statistics; neural network-related approaches

1 Introduction

Coal, petroleum and gas, among other non-renewable resources, will significantly contaminate the human living environment. Wind energy has gotten a lot of attention as a renewable, inexhaustible, and unlimited free energy source. Wind power is valuable not only because it is a renewable energy source, but also because of the megawatt scale of available wind turbines, easy operation, low maintenance costs and even government incentives [1–5]. According to estimates from wind power generating experts, around 2% of the sun's radiant energy is converted into wind energy each year, with installed capacity of up to 10 TW and predicted to increase even faster in the future [6–9]. In comparison to typical thermal power generation, wind energy as a type of green energy in renewable energy can lower power system running costs. As a result, a number of countries are promoting large-scale wind power development. However, the system's wind power quality is seriously impacted by the inherent randomness, intermittence, and volatility

of wind power generation, and uncertainty is introduced to the system's optimal dispatch. As a result, accurate wind speed prediction can benefit not only the quality of energy, but also real-time power grid scheduling and wind farm grid-connected operation [10–14]. One of the most extensively used approaches for predicting wind speed is the neural network. Although a single hidden layer feed-forward network can be used as an efficient predictive model to fit any complex function, constructing a reasonable network model and parameters have used prior knowledge of operators is extremely difficult for researchers and engineers to achieve accurate and satisfactory results [15–18]. Wind speed is simply a time signal with several frequency components, and its spectrum may be divided into two parts: amplitude and phase. Wavelet transformation is a characterisation method that is commonly used to match input signals by scaling the parent wavelet's oscillating pattern type. Similarly, the data decomposition method may properly reflect the signal's time-frequency characteristics and assess the signal's characteristics in the time-frequency domain, allowing for signal analysis at different resolutions. For example, wavelet transformation can properly reflect signal characteristics, assess signal characteristics in the time domain and frequency domain based on a variety of resolutions. Wind power time series can be treated of a layered overlays of several frequency components with varying levels of volatility and periodicity. If multi-layer decomposition is used, a resolution with similar frequency characteristics to each decomposition component may be identified, and the resolution at different scales as well as a suitable analytical procedure can be raised. As a result, the high-precision model is developed based on the properties of each frequency component. Precision wind energy forecasting can combine many volatile power sources at all levels of the transmission and distribution networks [19–24]. Short-term wind power forecasting with high accuracy might be considered an effective method for reducing grid integration and energy trading challenges [25–31]. The short-term wind is random, whereas the long-term wind follows a continuous probability density function, often known as the "Weibull distribution." Physical models and statistical methods are the mainly two kinds of short-term forecasting methodologies. The former method necessitates a great deal of physical data about the wind turbine, but the later usually treated as a soft computing method, is more adaptable and simple to apply in practice.

The accuracy of wind speed and wind power predictions is usually influenced by the surface wind, precipitation probability, maximum temperature, and even the conditional probability of frozen precipitation. In short-term wind power forecasting, wind speed is the most important meteorological component. Stetco et al. [32] provided a bibliographical assessment of general trends in the realms of wind speed and wind power forecasting. For the wind speed forecasts, numerical wind speed predictions based on Kalman filtering [33] and atmospheric models [34] with varied horizontal resolution capabilities were applied. Neural networks in combination with nearest neighbor search methods [35] were used to predict the output power of a specific wind farm using evolutionary optimization algorithms. The current state of hybrid solar-wind power generation system simulation, control, and optimization is described in [36,37]. The sensitivity of conventional generation and transmission were investigated as an example of alternative methodology using a static linear programming model. Deng et al. [36] developed a model for analyzing system optimum configurations based on the probability of power supply loss in hybrid solar-wind systems (LPSP). Deng et al. [38] investigates the uncertainty of wind power forecasting (WPF) using a proposed stochastic model. Tayab et al. [39] proposed some strategies for forecasting grid loads, with two main topics being discussed: short-term load forecasting (STLF) and the influence of anthropologic and structural variables on forecasting accuracy. The following issues are what we want to address in this paper based on the above discussion:

    1) The methods to select the variables among the many available meteorological variables which has a substantial impact on the output power prediction accuracy, should be considered.

    2) The short-term wind power forecasting category and error distribution in benchmark models.

This paper is organized as follows: data description and preprocessing, correlation analysis and neural network-related approaches are provided in Section 2. Section 3 introduces short-term wind power forecasting results obtained by the benchmark approaches, such as persistence, LMNN, NARX network, RNN and LSTM network, and all the performance of employed approaches via an illustrative example are also demonstrated. In Section 4, proposed results and prospective research issues are summarized and discussed.

2 Proposed Approaches for Wind Power Forecasting

The basic data description and distribution, data preprocessing methods, forecasting categories, and forecasting form are presented first in Fig. 1 of the paper's processing diagram. Second, correlation analysis between variables is used to select the input variables for describing output power with the fewest input variables. Furthermore, a heatmap of the correlation matrix between all of the available variables in Table 1 is shown, and a summary of the data distribution is provided by wind rise related to wind speed. Finally, five neural network-related benchmark approaches, such as the persistence model, LMNN, NARX network, LMRNN, and LSTM network, are shown to illustrate the short-term wind power forecasting accuracy.

2.1 Data Description and Preprocessing

The dataset is provided by the Software Engineer Divyam Khandelwal [40] and download from the Github, which is a time series with Power (MW), Wind direction 100 m (deg), Wind speed 100 m (m/s), Air temperature 2 m (K), Surface air pressure (Pa) and Density hub height (kg/m∧3), for the time period from 2012.01.01 to 2012.12.31. The detailed information of the considered dataset is given in Table 1.

The values of 'Hour' ranges from 0 to 23, which indicates the number of hours-ahead needed to be forecasted in short-term. For convenience, all the dataset objects are converted in to the standardized ISO 8601 format by following the processing procedure proposed by data scientist Jon Lo. Correspondingly, the forecasting category is split into two following categories:

    1) Category 1: 1 h to 12 h ahead data
    2) Category 2: 13 h to 24 h ahead data

Assume variables listed in Table 1, such as wind power $y_t^{(pow)}$, wind speed $x_t^{(win)}$ and wind direction $x_t^{(dir)}$ etc. at t-th time, then the forecasting is organized by

$$y_{t+k}^{(pow)} \doteq \left[y_t^{(pow)}, \ldots, y_{t-m_y}^{(pow)}, \ldots, x_t^{(win)}, \ldots, x_{t-m_x}^{(win)}, x_t^{(dir)}, \ldots, x_{t-m_u}^{(dir)}, \ldots \right] \tag{1}$$

where p_y and p_x are positive integrate, indicating the model lags associated to wind power $y_t^{(pow)}$ and wind speed $x_t^{(win)}$ at t-th time, respectively. $y_{t+k}^{(pow)}$ and $\left(y_t^{(pow)}, \ldots, y_{t-p_y}^{(pow)}, \ldots, x_t^{(win)}, \ldots, x_{t-p_x}^{(win)} \right)$ are the outputs and inputs, respectively. $k \in N^+$ is a factor to measure the k-ahead forecasting interval used to represent the short-term, median-term or long-term forecasting.
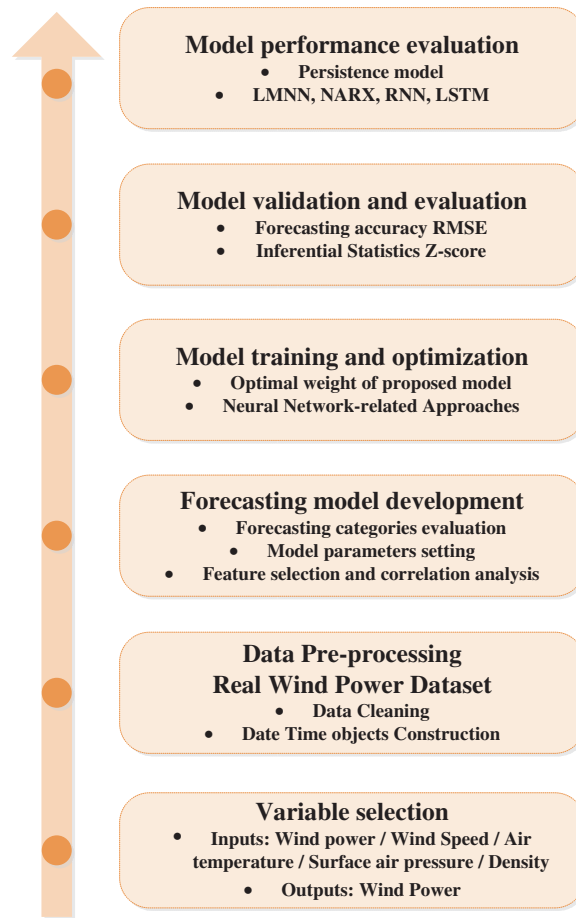
Figure 1: The processing diagram of this paper

Table 1: Data description and distribution

Variables	Power (MW)	Wind direction 100 m (deg)	Wind speed 100 m (m/s)	Air temperature 2 m (K)	Surface air pressure (Pa)	Density hub height (kg/m∧3)
Count	105120	105120	105120	105120	105120	105120
Mean	7.3838	163.4756	8.2059	288.4757	93868.7282	1.1182
Std	5.9499	94.6037	3.6458	10.9733	622.2715	0.0436
Min	0.0000	0.0050	0.0060	259.5210	91703.5520	1.0250
25%	1.5170	90.2425	5.276000	279.5260	93464.1680	1.0830
50%	6.2090	170.7265	8.112500	288.9200	93853.2800	1.1140
75%	13.9120	211.5322	10.998000	297.0010	94254.2960	1.1510
Max	16.0000	359.9940	28.720000	314.7810	96132.1040	1.2580

2.2 The Correlation Analysis between Variables

An ideal variable input is one that is extremely informative, especially when it is independent of each other, has a good number, and can be utilized to generate a set of variable

interpretations. As a result, the ideal input variables will have the fewest input variables to represent the characteristics of the output variables, which promotes neural network structural design and promotional capacities. For linear argument selection approaches, there are forward-back and step-by-step regression methods. In reality, selecting procedures for nonlinear arguments remains a major challenge. Researchers are gradually learning various climate characteristics of wind power, as well as the feedback impact of wind power, radiation, and precipitation, mainly to the use of ground observation data. Station observation, on the other hand, has its own intractable flaws, such as wind power overlap error, weather dependence, and observation area constraints.

$$\text{cov}(X_i, X_j) = \frac{\sum_{k=1,\dots,m} (x_{ki} - \bar{x}_{.i})(x_{kj} - \bar{x}_{.j})}{m-1} \tag{2}$$

where $X = [X_1, X_2, \cdots, X_n]^T$, $\{x_{k.} = [x_{k1}, x_{k2}, \cdots, x_{kn}]^T | 1 \le k \le m\}$, X_i and X_j represent two arbitrary variables selected from the Table 1, $\bar{x}_{.i}$ and $\bar{x}_{.j}$ are the expectation value of the time series $x_{k.}$, and m is the length of the given wind power data. Correlation coefficient $\text{cov}(X_i, X_j)$ is used to reflect the degree of correlation between different variables. The product difference approach is used to determine the correlation coefficient, which represents the degree of correlation between different wind power variables and their respective mean values. Because the mathematical expectation of the variable cannot be determined, the $m-1$ reflects the sample mean of the random variable.

2.3 Neural Network-Related Approaches for Forecasting

The robustness of the artificial neural network can be determined by the network parameters and the specific morphology of the error surface around the sample (ANN). The network parameters can be coupled to the sample extreme points to make the network more resilient, and the resulting error surface distribution is generally flat. It is critical to evaluate the network output's resilience, which may help address practical difficulties and improve the network's promotion ability and application prospects. The input is effectively a set of feature vectors composed of the available variables from Table 1, and a typical neural network input-output mapping is given by

$$f(x) = \sum_{j=1}^{m} w_j \varphi(||x - x_j||) \tag{3}$$

where m, w_j and $\varphi(||x - x_j||)$ are the number of hidden nodes in hidden-layers, the adjustment weight and a set of m arbitrary functions, respectively. In order to speed up convergence, the weights of the output layer are frequently modified quickly using a linear optimization technique, while the activation functions of the hidden layer are processed slowly using nonlinear optimization strategies. An isotropic Gaussian mapping defined by

$$\delta = \frac{d_{\max}}{\sqrt{2P}} \tag{4}$$

with standard deviation is typically used according to the spread of the centers, where $d_{\max}$ is the maximum distance between the chosen centers. The least mean square approach used for the weight adjustment defined by

$$\Delta W_k = \eta(d_k - W_k^T \Phi)\Phi \tag{5}$$

are applied to the weights of output layers, where W_k and d_k are the corresponding linear weight and the desired output, respectively. If a high-accuracy forecasting model is used, the time-delay must be taken into consideration while estimating wind power in the short future. NARX (Nonlinear autoregressive with external input) helps improve historical data memory and is an important component of a dynamic neural network. The NARX network's training function is a Bayesian regularization based on Levenberg-Marquardt optimization that updates the network's weight and bias values and minimizes a linear combination of squared errors and weights to ensure that it is properly generalized. Typical performance functions are usually measured in practice by

$$MSE = \frac{1}{N} \sum_{i=1,\ldots,N} e_i^2, \quad e_i = t_i - y_i, \quad i = 1, \ldots, N \tag{6}$$

$$MSW = \frac{1}{N} \sum_{i=1,\ldots,N} \omega_j^2 \tag{7}$$

$$MSEreg = \gamma MSE + (1 - \gamma) MSW \tag{8}$$

where γ is the performance ratio. MSE and MSW are the mean sum of squares of network error and biases, respectively.

2.4 Performance Evaluation Metrics

The mean square error (RMSE) is used to predict the degree of discreteness or deviation between the desired output and forecasted ones, to measure the accuracy of the prediction, defined by

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1,\ldots,n} (y_{tf} - \hat{y}_{tr})^2} \tag{9}$$

The new performance function RMSE leads the network to have smaller weights and biases, resulting in a smoother network response and less overfitting of the equation. In its most primitive sense, an RBF neural network has three layers with only one hidden layer that executes a nonlinear transition from input space to hidden space. It has a higher learning efficiency and function approximation than the BP network.

3 Experiments

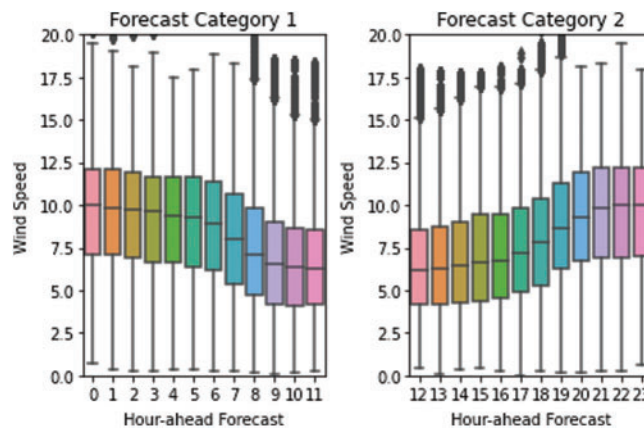### 3.1 Analysis of the Forecasting Categories and Data Distribution

The pd.DataFrame.merge are applied to mearge the training sample with wind speed and wind power, and the mean wind speed, wind power and wind direction is 8.1951, 163.3769 and 0.461702, respectively. The curve of wind speed data for the whole year from January 2012 to December 2012 is shown in Fig. 2, and two forecasting categories with respect to wind speed are visualized in Fig. 3.

The distribution pattern of wind speed under different forecasting categories is generally different, regardless of whether longer or shorter forecasting is used, and the inferential statistics results in Section 3.3 can still confirm the highlighted issue. The wind rise in relation to wind speed is shown in Fig. 4, which suggests a concise view that wind speed and direction are commonly dispersed at 0~15 m/s (about 60%–70%). The wind speed has been concentrated

in three directions: 135–180 degrees, 225–260 degrees, and 270–350 degrees. The geological or meteorological aspects may be to blame for these discrepancies in data distributions.
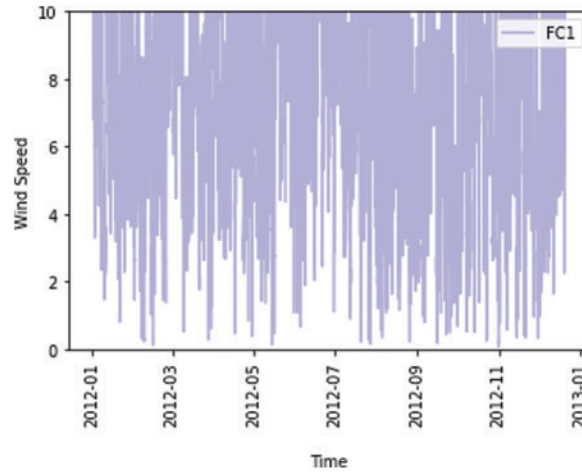
Figure 2: The wind speed trajectory of the whole year in 2012

Figure 3: Two forecasting categories

3.2 Correlation Analysis between Variables

Figs. 5–6 show the correlation analysis and the accompanying heatmap of correlation matrix of the available variables in Table 1. To illustrate correlation estimation between distinct variables, we introduced Pearson correlation coefficients in Fig. 5. It is simple to see that there's a significant correlation between wind speed and output power, with a coefficient of 0.96, fitting the high correlation's range of 0.9–1.0, notably the (a) and (b) of Fig. 6, which show that the coefficient of wind power and wind speed are the highest. Because their correlation coefficients (about 0.0051) are the lowest, there is a weak link between wind power and density hub height (kg/m^3), which is followed by wind direction and air temperature 2 m (K) (correlation coefficients is about 0.0092 and 0.0051, respectively). This is particularly evidence that the randomness, intermittence, and seasonality of natural wind speed, as well as the wind power of wind turbines proportional to

wind turbines, and the output voltage of wind turbines, are all closely related to wind speed fluctuations. Wind speed, to be more specific, has a considerable impact on wind power forecasting accuracy.
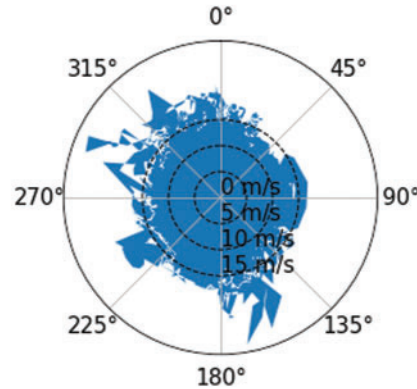

Figure 4: Wind rose of wind speed

3.3 Inferential Statistics and Performance Evaluation

Inferred statistics are statistical methods for inferring population characteristics from selected samples. It have been used to compare forecasting categories to see if there are differences between them: Forecast Category 1 (0–11 h ahead) and Forecast Category 2 (12–23 h ahead). Assume the μ_1 and μ_2 are respectively, the mean of two outlined forecasting categories, the null hypothesis is

$$H_0 : \mu_1 = \mu_2; \quad H_1 : \mu_1 \neq \mu_2$$

and the z-tests for the null hypotheses is setting as Z = 1.96 and significance level $\alpha = 0.05$, based on the evluation results, Z-test statistic formular is defined as $Z = \frac{m-\mu}{s} \cdot \sqrt{n}$.

Z scores is 36.6413 which is greater than z-tests scores Z = 1.96. This indicates that the two forecasting models are significantly different. In addition, as the number of forecasting steps increases, the accuracy of the forecasting will decrease dramatically. In short-term wind power forecasting, a tiny error can nonetheless result in large forecasting inaccuracies. This also shows that there is a distinction between two types of short-term wind power predictions. Table 2 shows the short-term wind power forecasting results based on benchmark methodologies. In Table 2, LMNN: Multilayer Perceptron with LM learning methods; NARX: Nonlinear autoregressive exogenous neural network model; LMRNN: RNNs with LM training methods; LSTM: Long short-term memory neural network.

Forecasting results 1–24-steps ahead wind power forecasting is obtained by LSTM network and tabulated in Table 3. The forecasting results denote that the forecasting performance deteriorates and the forecasting accuracy decreases with the increase in forecasting-steps. The slight inaccuracy in wind speed forecasting usually translates to large errors in wind power predictions. This means that, in addition to the proposed approach in this research, the forecasting model should be capable of error correction, dynamical feedback, and adaptive adjustment.
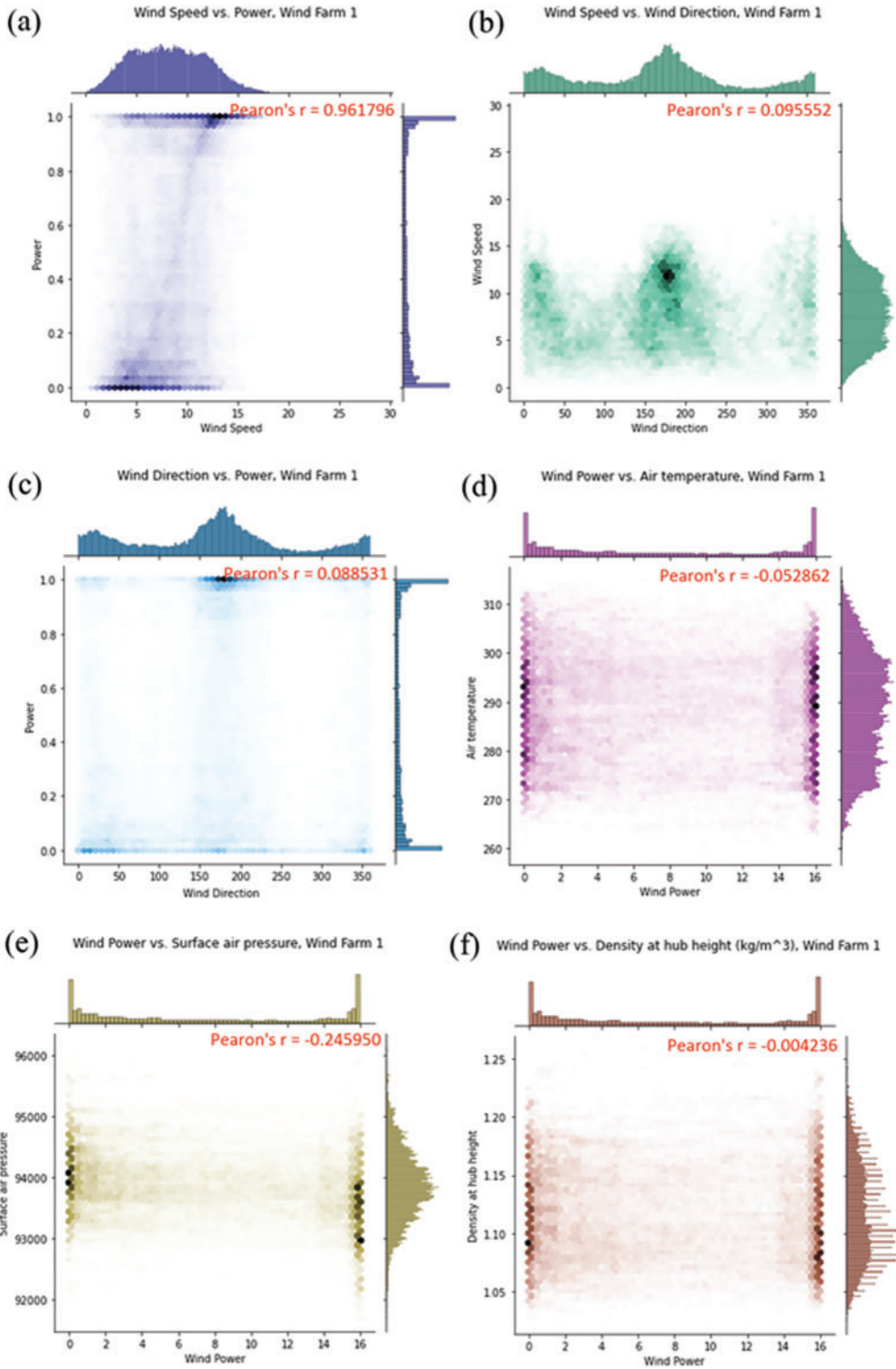
Figure 5: Visualization of correlation analysis (a) Wind speed *vs.* wind power; (b) Wind speed *vs.* wind direction; (c) Wind direction *vs.* wind power; (d) Wind power *vs.* air temperature; (e) Wind power *vs.* surface air pressure; (f) Wind power *vs.* hub height
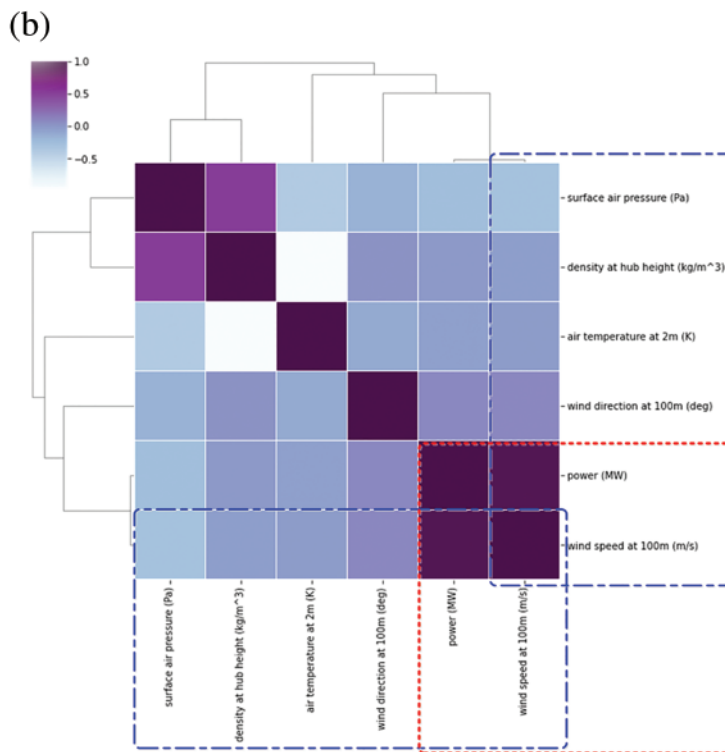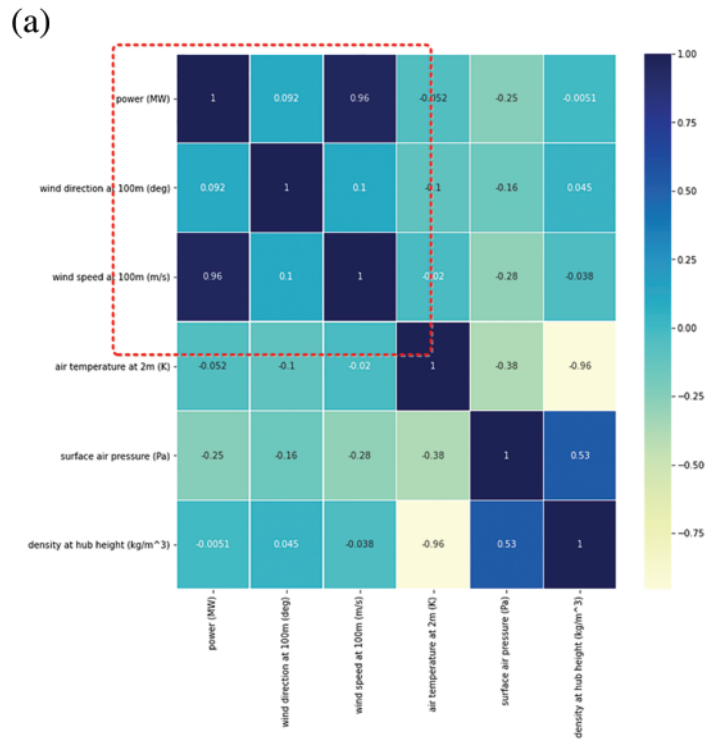
Figure 6: Heatmap of correlation matrix (a) Heatmap of correlation matrix; (b) Heat map derived from correlation matrix

Table 2: Forecasting accuracy comparison

Methods	1-step (5 min)	6-step (30 min)	12-step (1 h)	23-step (2 h)
Persistence	0.3028	0.3417	0.5436	0.6867
LMNN	0.1998	0.2708	0.2795	0.2821
NARX	0.1842	0.2135	0.2457	0.2691
LMRNN	0.1863	0.2027	0.2513	0.2746
LSTM	0.1817	0.1875	0.2247	0.2692

**Table 3:** (1–24-steps ahead) forecasting accuracy by LSTM network

FS	Training RMSE	Testing RMSE	Trainable param #
1-step	0.3562	0.1817	199
2-step	0.3577	0.1830	216
3-step	0.3590	0.1836	242
4-step	0.3610	0.1860	271
5-step	0.3618	0.1872	295
6-step	0.3623	0.1875	319
7-step	0.3624	0.1885	343
8-step	0.3637	0.1902	367
9-step	0.3640	0.1907	391
10-step	0.3642	0.1919	415
11-step	0.4253	0.2242	439
12-step	0.4253	0.2247	463
13-step	0.4256	0.2256	487
14-step	0.4258	0.2275	511
15-step	0.4261	0.2291	535
16-step	0.4271	0.2295	559
17-step	0.4290	0.2307	583
18-step	0.4293	0.2312	607
19-step	0.4298	0.2313	631
20-step	0.4391	0.2352	655
21-step	0.4551	0.2548	679
22-step	0.4588	0.2564	703
23-step	0.4611	0.2692	727
24-step	0.4975	0.3302	751

Training, validation and testing samples of the wind power and wind speed are shown in Fig. 7, and the training and validation curve of the cost funtion (1-step to 24 steps ahead forecasting) is provided in Fig. 8. There is a considerable difference between the two forecasting groups, as discussed previously. Because the persistence model assumes that the present data and the predictor do not change, and infers the predicted value based on inferential analysis, taking into account the 23-steps ahead forecasting outcomes, its forecasting accuracy is considerably lowered with RMSE 0.6867. When compared to the findings obtained by the other five forecasting models, this model has the lowest forecasting accuracy. The LSTM and NARX models share the best predicting results overall with the other approaches because they incorporate the delay and feedback mechanisms of wind power time series and boost the recall ability of historical data. One of the most generally used models of circulatory neural networks is the long-term short-term memory (Long Short-Term Memory, LSTM) network. This addresses two major flaws in simple circulatory neural networks: exploding gradients (i.e., it is trivial to generate infinite and non-values, resulting in data overflow owing to the bigger gradient value) and vanishing gradients (i.e., the learning ability of the model decays and the quality of learning is reduced when gradient values are small or even tend to zero). As a result, the five benchmark techniques have the best predicting accuracy, which is 62.43%, 8.54%, 4.12% and 6.05% lower than the other four models.
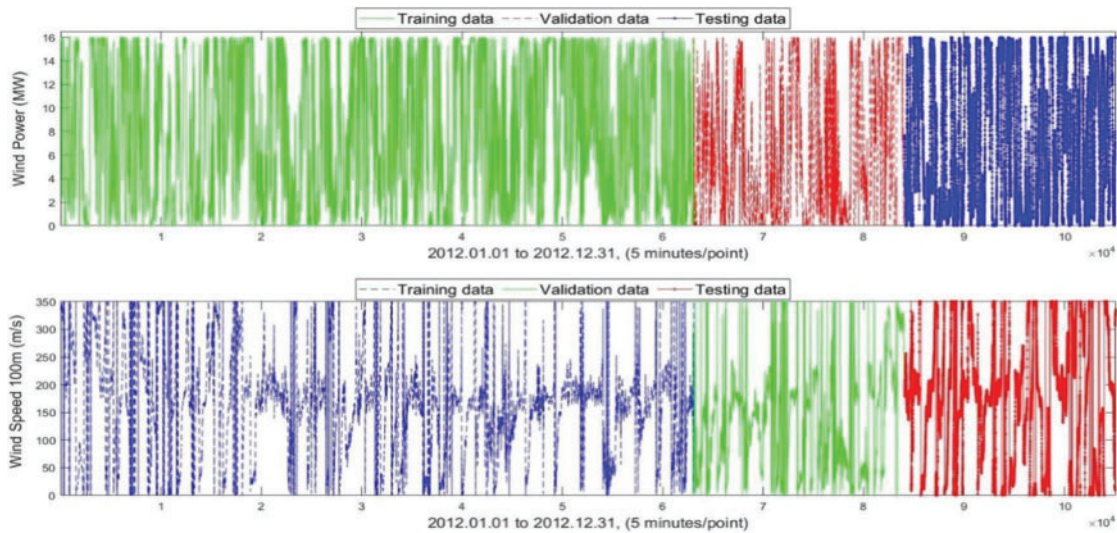
Figure 7: Training, validation and testing samples of the wind power and wind speed
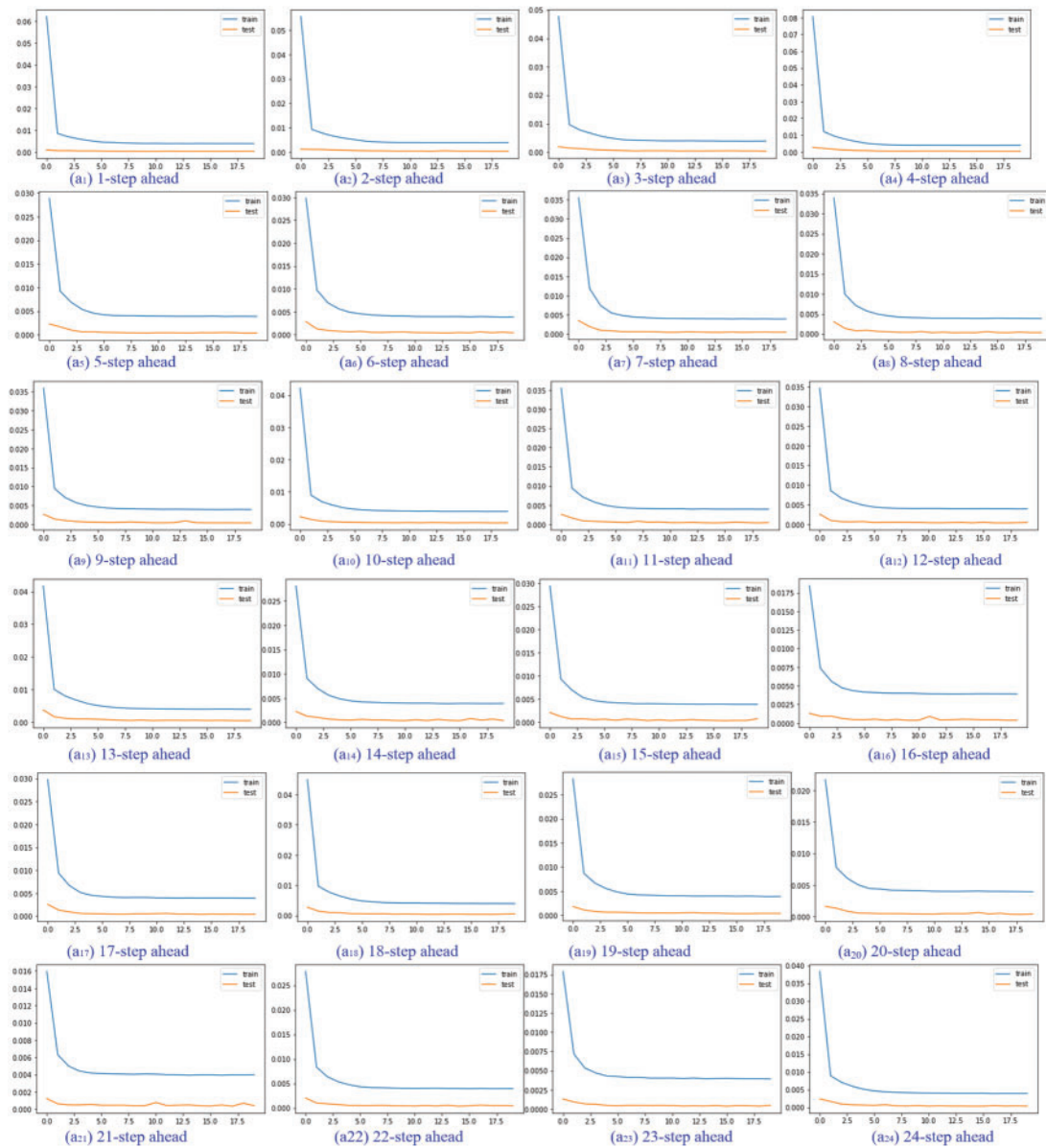
Figure 8: The training and validation curve of the cost function (1-step to 24 steps ahead forecasting)

4 Conclusion

Inferred statistics are used in this research to confirm that there is a significant difference between the two forecasting groups, i.e., Forecast Category 1 (0–11 h ahead) and Forecast Category 2 (12–23 h ahead) are the two types of forecasts. The wind speed has a significant impact on the forecasting accuracy of wind power when compared to the wind direction, air temperature 2 m (K), surface air pressure (Pa), and density hub height (kg/m^3) based on the correlation analysis. To verify the final performance of the forecasting output, five benchmark methodologies are used:

persistence model, LMNN, NARX network, LMRNN, and LSTM. For accurate and dependable wind power forecasting, we would use dynamical analysis with error correction capability in combination with reinforcement learning in the future.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Hong, Y. Y., Rioflorido, C. L. P. P. (2019). A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. *Applied Energy, 250,* 530–539. DOI 10.1016/j.apenergy.2019.05.044.

2. Watson, S., Moro, A., Reis, V., Baniotopoulos, C., Barth, S. et al. (2019). Future emerging technologies in the wind power sector: A european perspective. *Renewable and Sustainable Energy Reviews, 113,* 109270. DOI 10.1016/j.rser.2019.109270.

3. Javed, M. S., Ma, T., Jurasz, J., Amin, M. Y. (2020). Solar and wind power generation systems with pumped hydro storage: Review and future perspectives. *Renewable Energy, 148,* 176–192. DOI 10.1016/j.renene.2019.11.157.

4. Vargas, S. A., Esteves, G. R. T., Maçaira, P. M., Bastos, B. Q., Oliveira, F. L. C. et al. (2019). Wind power generation: A review and a research agenda. *Journal of Cleaner Production, 218,* 850–870. DOI 10.1016/j.jclepro.2019.02.015.

5. Shair, J., Xie, X., Yan, G. (2019). Mitigating subsynchronous control interaction in wind power systems: Existing techniques and open challenges. *Renewable and Sustainable Energy Reviews, 108,* 330–346. DOI 10.1016/j.rser.2019.04.003.

6. Deng, X., Shao, H., Hu, C., Jiang, D., Jiang, Y. (2020). Wind power forecasting methods based on deep learning: A survey. *Computer Modeling in Engineering & Sciences, 122(1),* 273–302. DOI 10.32604/cmes.2020.08768.

7. Mahela, O. P., Khan, B., Alhelou, H. H., Siano, P. (2020). Power quality assessment and event detection in distribution network with wind energy penetration using stockwell transform and fuzzy clustering. *IEEE Transactions on Industrial Informatics, 16(11),* 6922–6932. DOI 10.1109/TII.9424.

8. Kushwaha, A., Gopal, M., Singh, B. (2020). Q-learning based maximum power extraction for wind energy conversion system with variable wind speed. *IEEE Transactions on Energy Conversion, 35(3),* 1160–1170. DOI 10.1109/TEC.60.

9. Xu, B., Chen, D., Venkateshkumar, M., Xiao, Y., Yue, Y. et al. (2019). Modeling a pumped storage hydropower integrated to a hybrid power system with solar-wind power and its stability analysis. *Applied Energy, 248,* 446–462. DOI 10.1016/j.apenergy.2019.04.125.

10. Navas, R. K. B., Prakash, S., Sasipraba, T. (2020). Artificial neural network based computing model for wind speed prediction: A case study of Coimbatore, Tamil nadu, India. *Physica A: Statistical Mechanics and its Applications, 542,* 123383. DOI 10.1016/j.physa.2019.123383.

11. Zhou, Q., Wang, C., Zhang, G. (2019). Hybrid forecasting system based on an optimal model selection strategy for different wind speed forecasting problems. *Applied Energy, 250,* 1559–1580. DOI 10.1016/j.apenergy.2019.05.016.

12. Sorknæs, P., Djørup, S. R., Lund, H., Thellufsen, J. Z. (2019). Quantifying the influence of wind power and photovoltaic on future electricity market prices. *Energy Conversion and Management, 180,* 312–324. DOI 10.1016/j.enconman.2018.11.007.

13. Zhu, M., Qi, Y., Belis, D., Lu, J., Kerremans, B. (2019). The China wind paradox: The role of state-owned enterprises in wind power investment versus wind curtailment. *Energy Policy, 127,* 200–212. DOI 10.1016/j.enpol.2018.10.059.

14. Ning, C., You, F. (2019). Data-driven adaptive robust unit commitment under wind power uncertainty: A Bayesian nonparametric approach. *IEEE Transactions on Power Systems, 34(3),* 2409–2418. DOI 10.1109/TPWRS.59.

15. Altan, A., Karasu, S., Zio, E. (2021). A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. *Applied Soft Computing, 100,* 106996. DOI 10.1016/j.asoc.2020.106996.

16. Jeon, J., Panagiotelis, A., Petropoulos, F. (2019). Probabilistic forecast reconciliation with applications to wind power and electric load. *European Journal of Operational Research, 279(2),* 364–379. DOI 10.1016/j.ejor.2019.05.020.

17. Meng, H., Wang, M., Olumayegun, O., Luo, X., Liu, X. (2019). Process design, operation and economic evaluation of compressed air energy storage (CAES) for wind power through modelling and simulation. *Renewable Energy, 136,* 923–936. DOI 10.1016/j.renene.2019.01.043.

18. Shahid, F., Zameer, A., Muneeb, M. (2021). A novel genetic LSTM model for wind power forecast. *Energy, 223,* 120069. DOI 10.1016/j.energy.2021.120069.

19. Lima, M. A. F., Carvalho, P. C., Fernández-Ramírez, L. M., Braga, A. P. (2020). Improving solar forecasting using deep learning and portfolio theory integration. *Energy, 195,* 117016. DOI 10.1016/j.energy.2020.117016.

20. Kisvari, A., Lin, Z., Liu, X. (2021). Wind power forecasting–A data-driven method along with gated recurrent neural network. *Renewable Energy, 163,* 1895–1909. DOI 10.1016/j.renene.2020.10.119.

21. Mäkitie, T., Normann, H. E., Thune, T. M., Gonzalez, J. S. (2019). The green flings: Norwegian oil and gas industry's engagement in offshore wind power. *Energy Policy, 127,* 269–279. DOI 10.1016/j.enpol.2018.12.015.

22. Yang, W., Yang, J. (2019). Advantage of variable-speed pumped storage plants for mitigating wind power variations: Integrated modelling and performance assessment. *Applied Energy, 237,* 720–732. DOI 10.1016/j.apenergy.2018.12.090.

23. Naik, J., Dash, P. K., Dhar, S. (2019). A multi-objective wind speed and wind power prediction interval forecasting using variational modes decomposition based multi-kernel robust ridge regression. *Renewable Energy, 136,* 701–731. DOI 10.1016/j.renene.2019.01.006.

24. Shilaja, C., Arunprasath, T. (2019). Optimal power flow using moth swarm algorithm with gravitational search algorithm considering wind power. *Future Generation Computer Systems, 98,* 708–715. DOI 10.1016/j.future.2018.12.046.

25. Weschenfelder, F., Leite, G. D. N. P., da Costa, A. C. A., de Castro Vilela, O., Ribeiro, C. M. et al. (2020). A review on the complementarity between grid-connected solar and wind power systems. *Journal of Cleaner Production, 257,* 120617. DOI 10.1016/j.jclepro.2020.120617.

26. Wang, Y., Hu, Q., Li, L., Foley, A. M., Srinivasan, D. (2019). Approaches to wind power curve modeling: A review and discussion. *Renewable and Sustainable Energy Reviews, 116,* 109422. DOI 10.1016/j.rser.2019.109422.

27. Ghoushchi, S. J., Manjili, S., Mardani, A., Saraji, M. K. (2021). An extended new approach for forecasting short-term wind power using modified fuzzy wavelet neural network: A case study in wind power plant. *Energy, 223,* 120052. DOI 10.1016/j.energy.2021.120052.

28. Han, L., Zhang, R., Wang, X., Bao, A., Jing, H. (2019). Multi-step wind power forecast based on VMD-lSTM. *IET Renewable Power Generation, 13(10),* 1690–1700. DOI 10.1049/iet-rpg.2018.5781.

29. Wang, J., Niu, T., Lu, H., Yang, W., Du, P. (2019). A novel framework of reservoir computing for deterministic and probabilistic wind power forecasting. *IEEE Transactions on Sustainable Energy, 11(1),* 337–349. DOI 10.1109/TSTE.5165391.

30. Abedinia, O., Lotfi, M., Bagheri, M., Sobhani, B., Shafie-Khah, M. et al. (2020). Improved EMD-based complex prediction model for wind power forecasting. *IEEE Transactions on Sustainable Energy, 11(4),* 2790–2802. DOI 10.1109/TSTE.5165391.

31. Mishra, S., Bordin, C., Taharaguchi, K., Palu, I. (2020). Comparison of deep learning models for multivariate prediction of time series wind power generation and temperature. *Energy Reports, 6,* 273–286. DOI 10.1016/j.egyr.2019.11.009.

32. Stetco, A., Dinmohammadi, F., Zhao, X., Robu, V., Flynn, D. et al. (2019). Machine learning methods for wind turbine condition monitoring: A review. *Renewable Energy, 133,* 620–635. DOI 10.1016/j.renene.2018.10.047.

33. Shair, J., Xie, X., Yuan, L., Wang, Y., Luo, Y. (2020). Monitoring of subsynchronous oscillation in a series-compensated wind power system using an adaptive extended Kalman filter. *IET Renewable Power Generation, 14(19),* 4193–4203. DOI 10.1049/iet-rpg.2020.0280.

34. Parks, K., Wan, Y. H., Wiener, G., Liu, Y. (2011). *Wind Energy Forecasting: A Collaboration of the National Center for Atmospheric Research (NCAR) and Xcel Energy (No. NREL/SR-5500-52233).* National Renewable Energy Lab. (NREL), Golden, CO (United States).

35. Bay, C. J., Annoni, J., Taylor, T., Pao, L., Johnson, K. (2018). Active power control for wind farms using distributed model predictive control and nearest neighbor communication. *Annual American Control Conference,* pp. 682–687. Wisconsin Center, United States. DOI 10.23919/ACC.2018.8431764.

36. Deng, X., Shao, H. (2020). Deep learning approach with optimizatized hidden-layers topology for short-term wind power forecasting. *Energy Engineering, 117(5),* 279–287. DOI 10.32604/EE.2020.011619.

37. Li, J., Liu, P., Li, Z. (2020). Optimal design and techno-economic analysis of a solar-wind-biomass off-grid hybrid power system for remote rural electrification: A case study of West China. *Energy, 208,* 118387. DOI 10.1016/j.energy.2020.118387.

38. Deng, X., Shao, H., Wang, X. (2020). Seasonal characteristics analysis and uncertainty measurement for wind speed time series. *Energy Engineering, 117(5),* 289–299. DOI 10.32604/EE.2020.011126.

39. Tayab, U. B., Zia, A., Yang, F., Lu, J., Kashif, M. (2020). Short-term load forecasting for microgrid energy management system using hybrid HHO-FNN model with best-basis stationary wavelet packet transform. *Energy, 203,* 117857. DOI 10.1016/j.energy.2020.117857.

40. Divyam Khandelwal, Wind Power Dataset (2019). https://github.com/divyam-khandelwal/Thesis-Wind-Power-Prediction-Model.