TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Các mô hình mạng nơ ron hồi quy dùng cho dự báo điện gió

Nguyễn Văn Ninh, Nguyễn Xuân Thậm, Thân Văn Thông, Đoàn Thanh Bảo, Lương Ngọc Toàn, Lê Tuấn Hộ*

Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 14/04/2025; Ngày sửa bài: 06/08/2025;
Ngày nhận đăng: 12/08/2025; Ngày xuất bản: 28/10/2025

TÓM TẮT

Các giá trị lịch sử của công suất phát điện gió thường được sử dụng trong hầu hết các mô hình và phương pháp dự báo điện gió trong các tài liệu. Tuy nhiên, các thông số vận hành có thể ảnh hưởng đến kết quả dự báo chẳng hạn như tốc độ gió, góc pitch, nhiệt độ môi trường, vị trí nacelle, và hướng gió chưa được xem xét trong các phương pháp hiện tại. Do đó, mục tiêu cơ bản của bài báo này là đề xuất các mô hình mạng nơ ron hồi quy dùng cho dự báo công suất phát điện gió có xem xét các tham số vận hành này. Nghiên cứu này xem xét dữ liệu công suất phát điện gió và các tham số vận hành tương ứng từ máy phát tuabin gió số 5 của 1 trang trại gió. Dữ liệu từ ngày 01 tháng 7 năm 2024 đến 31 tháng 7 năm 2024 được thu thập từ hệ thống SCADA. Đầu tiên, mô hình mạng nơ ron tự hồi quy phi tuyến có các đầu vào được áp dụng để dự báo công suất phát điện gió. Thứ hai, mô hình mạng hồi quy lớp được sử dụng để dự báo công suất phát điện gió. Thứ ba, mô hình mạng nơ ron trễ phân tán được dùng để dự báo công suất phát điện gió. Thứ tư, mô hình mạng nơ ron trễ thời gian được huấn luyện để ước lượng điện gió. Cuối cùng, các mô hình mạng nơ ron hồi quy này được so sánh để xác định mô hình dự báo công suất phát điện gió tốt hơn khi xét theo các tiêu chí sai số tuyệt đối trung bình, sai số phần trăm tuyệt đối trung bình và sai số bình phương trung bình.

Từ khóa: *Dự báo điện gió, mạng nơ ron tự hồi quy phi tuyến có các đầu vào, mạng nơ ron hồi quy lớp, mạng nơ ron trễ phân tán, mạng nơ ron trễ thời gian.*

*Tác giả liên hệ chính.
Email: tuanhole@qnu.edu.vn*

QUY NHON UNIVERSITY
JOURNAL OF SCIENCE

Recurrent neural network models for wind power forecasting

Van Ninh Nguyen, Xuan Tham Nguyen, Van Thong Than, Thanh Bao Doan, Ngoc Toan Luong, Tuan Ho Le*

Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: 14/04/2025; Revised: 06/08/2025;
Accepted: 12/08/2025; Published: 28/10/2025

ABSTRACT

The historical values of wind power generation are generally utilized in most forecasting models and methods in the literature. Unfortunately, the operational parameters such as wind speed, pitch angle, ambient temperature, nacelle position, and wind direction may affect the forecasting results. Therefore, the primary objective of this paper is to propose recurrent neural network models for wind power generation forecasting considering these operational parameters. In this study, the wind power generation data and the associated operational parameters from the wind turbine generator 05 of a wind farm are investigated. The data from July 1st, 2024 to July 31st, 2024 is collected from the SCADA system. Firstly, the nonlinear autoregressive neural network with external input is applied to make the wind power generation prediction. Secondly, the layer recurrent neural network model is employed to forecast wind power generation. Thirdly, the distributed delay neural network model is implemented to predict wind power generation. Fourthly, the time delay neural network model is trained to estimate the wind power. Finally, these recurrent neural network models are compared to determine the better wind power generation forecasting model in terms of mean absolute error, mean absolute percent error, and root mean square error.

Keywords: *Wind power forecasting, nonlinear autoregressive neural network with external input, layer recurrent neural network, distributed delay neural network, time delay neural network.*

1. INTRODUCTION

The rapid increase in energy demand has driven the search for alternative energy sources, in addition to traditional ones that are depleting and causing pollution issues. Wind power is a clean and renewable source. According to the Global Wind Report 2024 (by Global Wind Energy Council), it is shown that 2023 recorded the highest number of new installations in history for onshore wind (over 100 GW) and the second highest for offshore wind (11 GW). Wind energy installations will increase from a level of 117 GW in 2023 to at least 320 GW of annual installations by 2030. Actually, wind power generation plays a significant role in electricity supply. However, wind energy integration into power systems presents inherent unpredictability because of the intermittent nature of wind energy.[1] As wind energy makes significant penetration into the electricity grid, the need for accurate predictions of wind power generation becomes critical and urgent.[2,3] To address these challenges, wind power forecasting (WPF) has emerged as a valuable solution. Consequently, numerous WPF models and methods have been proposed and implemented in the literature.

*Corresponding author.
Email: tuanhole@qnu.edu.vn

https://doi.org/10.52111/qnjs.2025.19505

Based on the forecasting time horizon, WPF can be categorized into ultra short-term, short-term, medium-term, and long-term. Various types of forecasting models and methods are developed for wind power generation time series. The traditional statistical models and methods are usually applied by using the previous historical data to perform a forecast. In WPF, the statistical models are commonly applied as exponential smoothing approach,[4,5] autoregressive,[6] autoregressive moving average (ARMA),[7,8] autoregressive integrated moving average (ARIMA).[9,10] Statistical models are more user-friendly and cost-effective to develop than other types of models. Statistical methods primarily rely on historical wind data to forecast the upcoming few hours, making them suitable for short-term predictions. While statistical forecasting models such as ARIMA and exponential smoothing are simple and interpretable, they suffer from limitations such as linearity assumptions, difficulty incorporating multiple external variables, and poor scalability to nonlinear or high-dimensional problems. Besides, artificial intelligence (AI)-based models are another research direction in WPF. Fundamentally, AI-based models, when adequately trained, have the potential to outperform traditional statistical models in forecasting accuracy. Depending on the structure of neural networks (NNs), various models are proposed in wind power generation forecasting. Feed-forward NN models are used to make the wind power prediction such as multi-layer perceptron[11] and back-propagation (BP) NN.[12] Another kind of NN model with feedback known as the recurrent NN (RNN) model is also widely applied in this research direction. The RNN model-based WPF methods are Elman NN,[13] layer RNN,[14] nonlinear autoregressive NN,[15] long short-term memory (LSTM),[16] bidirectional LSTM,[17] gated recurrent unit,[18] and echo state network.[19] Deep learning models such as LSTM, bidirectional LSTM, gated recurrent unit, and echo state network have demonstrated improved forecasting performance compared to earlier NN architectures. Moreover, support vector machine,[20] gradient boosting regression tree algorithms,[21] and ensemble model[22] belonging to machine learning are also implemented in WPF. These machine learning models have also been shown to outperform NN models in specific research cases. Other attempts have been made to combine different models or methods in various ways to improve the forecasting results. These include combinations such as autoregressive fractionally integrated moving average and least square support vector machine,[22] boosting algorithm and ARMA model,[23] hybrid CEEMDAN-EWT deep learning method,[24] and neuro wavelet and LSTM models.[25] These combined models have also demonstrated improved forecasting performance compared to each individual model. In these existing studies, most WPF models address the wind power generation time series data issues without taking the operational parameters into account. These operating parameters have a direct impact on the power output of wind turbines in real-world conditions; therefore, they influence the accuracy of forecasting results. Some significant operational parameters can be considered wind speed, pitch angle, ambient temperature, nacelle position, and wind direction. Therefore, in this paper, one of the six turbines of a wind farm, located on the south-central coast, Vietnam with a 114-metre height and a 132-metre rotor is considered. These operational parameters and wind power generation time series data are collected at 10-minute intervals from the SCADA system. The dataset spans from July 01st, 2024 to July 31st, 2024. This data is divided into three different case studies. Subsequently, numerous RNN models are proposed for wind power forecasting, including nonlinear autoregressive NN with external input (Narxnet), layer RNN (Layrecnet), distributed delay NN (Distdelaynet), and time delay neural network (Timedelaynet). These models incorporate both operational parameters and wind power generation data. To evaluate the performance of WPF models, evaluation criteria such as mean absolute error (MAE), mean absolute percent error (MAPE), and root mean square error (RMSE) are employed. Based on the final results, the most effective forecasting model can be determined. An overview of wind power generation forecasting using RNN models can be represented in Figure 1.

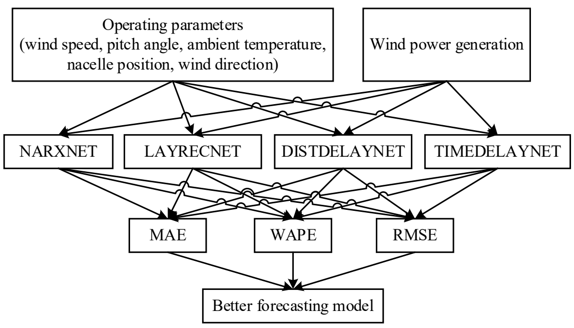

Figure 1. Overview of wind power generation forecasting using RNN models.

2. METHODOLOGY

2.1. Nonlinear autoregressive neural network with external input

Narxnet is a nonlinear autoregressive model with exogenous inputs commonly used in time series modeling. Narxnet is first formally proposed and popularized in the mid-1990s by Lin *et al.*[26] In this model, the current value of a time series depends on both past values of the same series and current and past values of exogenous series. The Narxnet model, applied to time series forecasting, can be mathematically represented as follows:

$$y_t = f(y_{t-1}, y_{t-2}, ..., x_t, x_{t-1}, x_{t-2}, ...) \quad (1)$$

where f denotes an unknown nonlinear function (i.e., transfer function or activation function), y_t denotes the predicted value of the time series data of y at a discrete time t, and x_t denotes the externally determined variable. The Narxnet model can be illustrated in Figure 2.

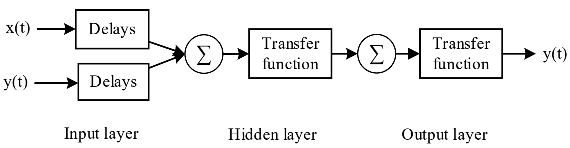

Figure 2. Narxnet model.

2.2. Layer recurrent neural network

Layrecnet is another type of RNN models. The concept of Layrecnet model is proposed by Schmidhuber.[27] In this architecture, the output of the hidden layer is fed back to the input layer with delays. As a result, the network is capable of exhibiting an infinite dynamic response to time series input data. The Layrecnet model can be illustrated in Figure 3.

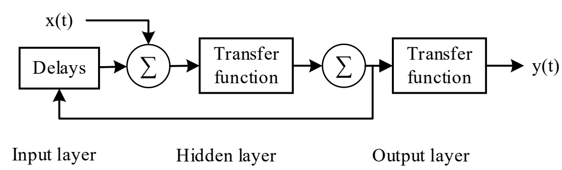

Figure 3. Layrecnet model.

2.3. Distributed delay neural network

Distdelaynet is another type of RNN model. The concept of Distdelaynet model is proposed by Waibel *et al.*[28] The input and hidden layers have a tap delay line associated with them. Therefore, the network exhibits a finite dynamic response to time series input data. The Distdelaynet model can be illustrated in Figure 4.

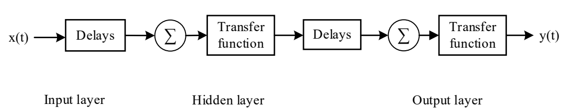

Figure 4. Distdelaynet model.

2.4. Time delay neural network

Timedelaynet is another type of RNN model. The concept of Distdelaynet model is also proposed by Waibel *et al.*[28] In this NN model, the input layer has a tap delay line associated with it. Therefore, the network exhibits a finite dynamic response to time series input data. The Timedelaynet model can be illustrated in Figure 5.

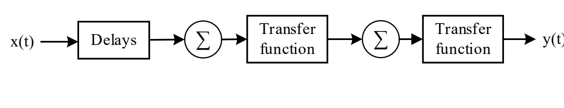

Figure 5. Timedelaynet model.

2.5. Transfer function

In NN models, transfer functions (also known

https://doi.org/10.52111/qnjs.2025.19505

as activation functions) play a crucial role in introducing nonlinearity and determining the output of each neuron. Commonly, three types of transfer functions are employed, each serving a specific purpose within the network architecture:

Linear (i.e., purelin):

$$f = x\,. \tag{1}$$

Hyperbolic tangent sigmoid (i.e., tansig):

$$f = \frac{2}{1+e^{-2x}} - 1\,. \tag{2}$$

Log-sigmoid (i.e., logsig):

$$f = \frac{1}{1+e^{-x}}\,. \tag{3}$$

2.6. Training algorithm

The training process of a NN begins with the random initialization of connection weights, which serve as the adjustable parameters controlling how input signals are transformed through the network layers. These weights are then iteratively updated to minimize a predefined cost (or loss) function, which quantifies the discrepancy between the network's predicted output and the actual target value.[29] This minimization is typically performed using gradient-based optimization algorithms that adjust the weights in the direction that reduces the error.

Several well-established training algorithms are employed for this purpose, each with distinct convergence characteristics and computational requirements:

- Resilient BP (i.e, trainrp): Focuses on the sign of the gradient rather than its magnitude to achieve stable convergence, especially in noisy or flat error surfaces.
- Bayesian regularization BP (i.e, trainbr): Incorporates a regularization term into the cost function to prevent overfitting by balancing model complexity and data fitting.
- BFGS quasi-Newton BP (i.e., trainbfg): Utilizes second-order derivative approximations to speed up convergence, particularly effective in medium-sized networks.
- Levenberg–Marquardt BP (i.e., trainlm): Combines the advantages of the Gauss–Newton algorithm and gradient descent, offering fast convergence and high accuracy for smaller datasets or networks.

The choice of training algorithm can significantly affect the learning efficiency, generalization ability, and computational cost of the neural network model.

2.7. Evaluation criteria

In this study, three evaluation criteria are employed to assess the performance of the forecasting models. These criteria are widely adopted in time series forecasting tasks, as they provide comprehensive insights into prediction accuracy and error distribution. Specifically, the evaluation metrics used are as follows:

MAE:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|A_t - F_t\right| \tag{4}$$

where A_t is the actual value, F_t is the forecasted value, and n is the number of observations.

WAPE:

$$WAPE = \frac{\sum_{t=1}^{n}\left|A_t - F_t\right|}{\sum_{t=1}^{n}\left|A_t\right|}\,. \tag{5}$$

RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(A_t - F_t\right)^2}\,. \tag{6}$$

3. CASE STUDIES

The time-series data from the wind turbine no. 05 (3.5 MW, 114-meter hub height, 132-meter rotor diameter) in a wind farm were collected from July 01st, 2024 to July 31st, 2024 (i.e., case study 3) with 10-minute intervals. Two investigated sub-periods during this timeframe are from July 01st, 2024 to July 07th, 2024 (i.e., case study 1) and from July 25th, 2024 to July 31st, 2024 (i.e., case study 2). The total number of observations in the three case studies is provided in Table 1.

Table 1. Number of observations.

Data	Case study 1	Case study 2	Case study 3
Number of observations	1008	1008	4464

The wind speed, pitch angle, ambient temperature, nacelle position, and wind direction from July 01st, 2024 to July 31st, 2024 are illustrated in Figures 6, 7, 8, 9, and 10, respectively.

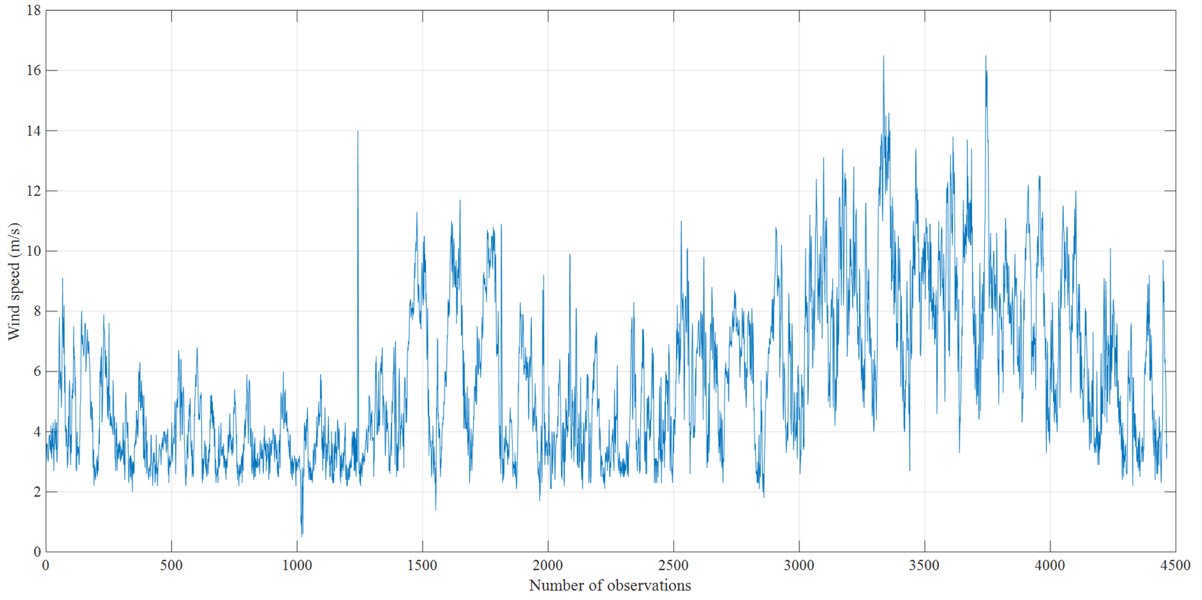

Figure 6. Wind speed.

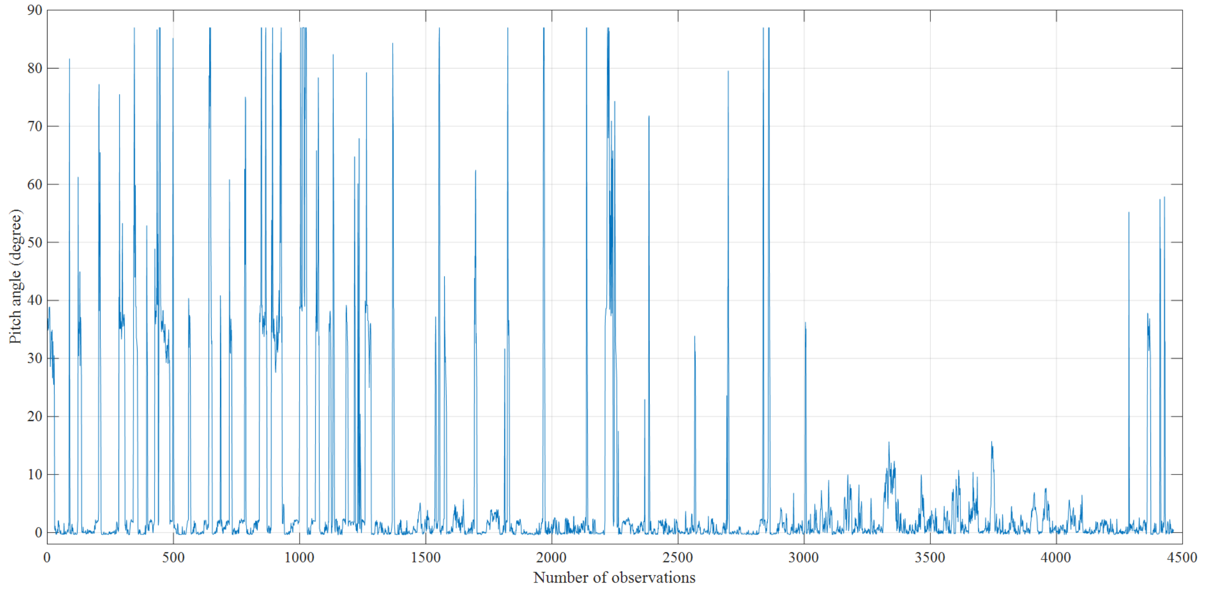

Figure 7. Pitch angle.

https://doi.org/10.52111/qnjs.2025.19505

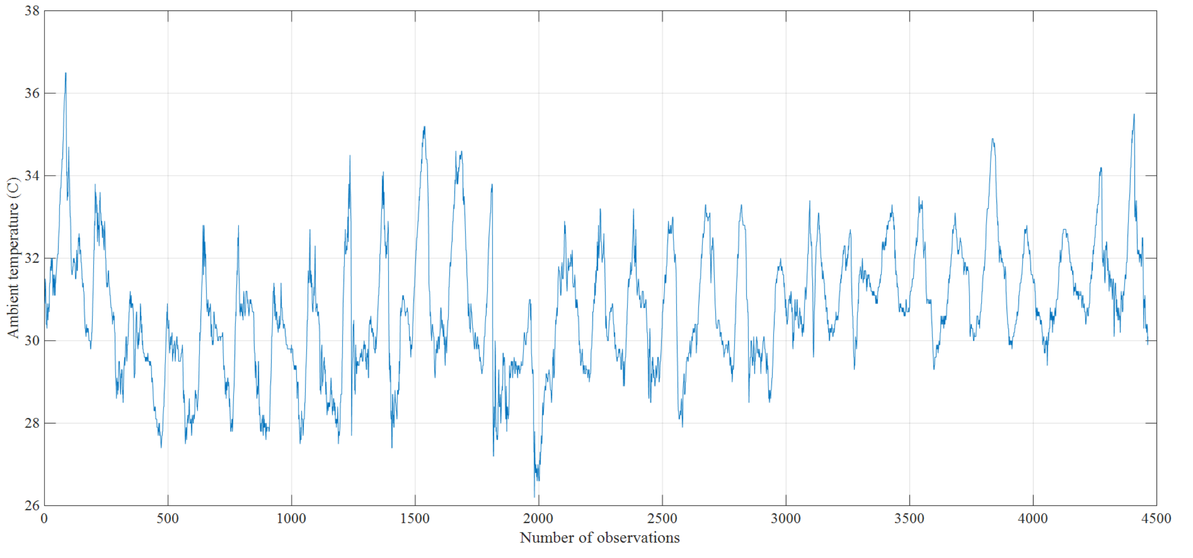

Figure 8. Ambient temperature.

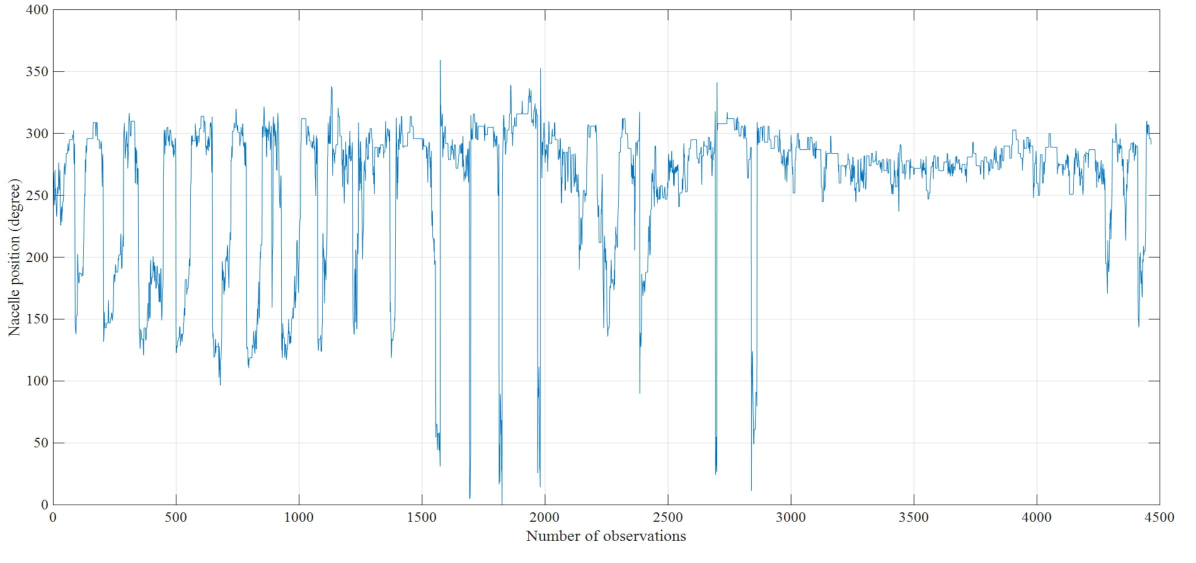

Figure 9. Nacelle position.

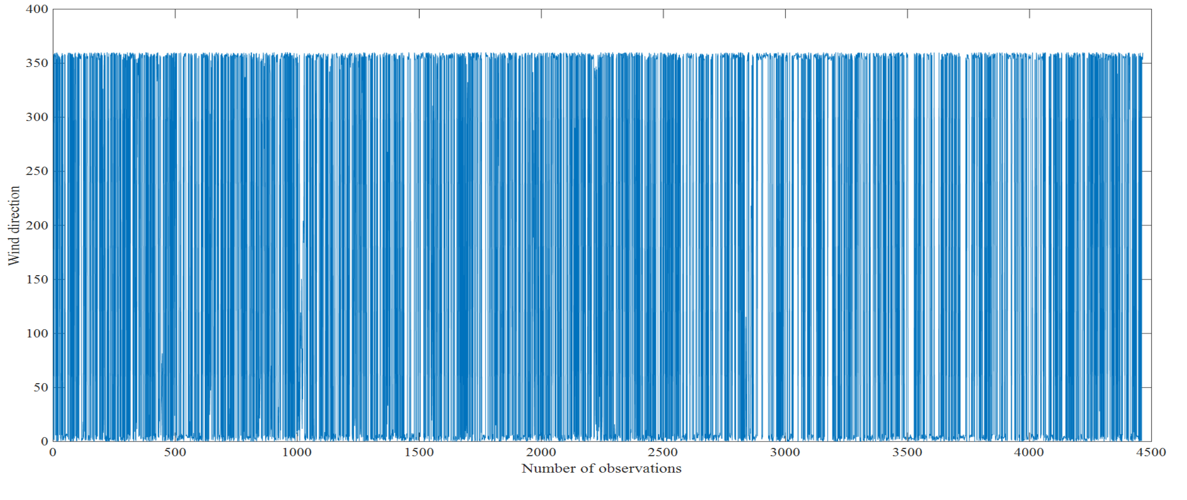

Figure 10. Wind direction.

All the recurrent NN models with one hidden layer have been implemented. Moreover, these models, each configured with one feedback delay in their respective structures, are coded and trained in Matlab software. The transfer function, training algorithms, and number of hidden neurons of all the trained Narxnet, Layrecnet, Distdelaynet, and Timedelaynet models for case studies 1, 2, and 3 are represented in Tables 2, 3, and 4, respectively.

Table 2. Information of the recurrent NN models in case study 1.

Model	Transfer function	Training algorithm	Hidden neurons
Narxnet	Tansig	Trainbr	47
Layrecnet	Tansig	Trainbr	30
Distdelaynet	Tansig	Trainbr	143
Timedelaynet	Logsig	Trainbr	253

Table 3. Information of the recurrent NN models in case study 2.

Model	Transfer function	Training algorithm	Hidden neurons
Narxnet	Logsig	Trainbr	21
Layrecnet	Tansig	Trainbr	79
Distdelaynet	Tansig	Trainlm	232
Timedelaynet	Tansig	Trainbr	301

Table 4. Information of the recurrent NN models in case study 3.

Model	Transfer function	Training algorithm	Hidden neurons
Narxnet	Tansig	Trainlm	32
Layrecnet	Tansig	Trainbr	48
Distdelaynet	Logsig	Trainlm	372
Timedelaynet	Tansig	Trainbr	180

The performance evaluation results of all RNN forecasting models across the three case studies are presented in Tables 5, 6, and 7, respectively. These tables provide a comprehensive comparison based on multiple error metrics, allowing for a thorough assessment of each model's forecasting accuracy and consistency under varying data conditions. By analyzing these evaluation values, the relative effectiveness of each RNN architecture can be quantitatively compared across different time periods.

Table 5. Evaluation criteria of the proposed forecasting models in case study 1.

Model	MAE	WAPE(%)	RMSE
Narxnet	56.0914	22.3584	87.4324
Layrecnet	18.6198	7.4220	24.9867
Distdelaynet	65.9784	26.2994	130.9846
Timedelaynet	49.9778	19.9215	98.3547

Table 6. Evaluation criteria of the proposed forecasting models in case study 2.

Model	MAE	WAPE(%)	RMSE
Narxnet	269.7987	17.0885	382.9613
Layrecnet	80.9957	5.1301	104.7831
Distdelaynet	341.8460	21.6519	490.9236
Timedelaynet	207.4276	13.1381	289.5628

Table 7. Evaluation criteria of the proposed forecasting models in case study 3.

Model	MAE	WAPE(%)	RMSE
Narxnet	168.6443	18.3242	263.3885
Layrecnet	41.3159	4.4892	58.1676
Distdelaynet	248.0907	26.9565	388.8692
Timedelaynet	164.6041	17.8852	256.3703

https://doi.org/10.52111/qnjs.2025.19505

Based on the results presented in Table 5, the Narxnet model yields improved forecasting performance over the Distdelaynet model, with lower error values across all three metrics (MAE = 56.0914, WAPE = 22.3584%, RMSE = 87.4324 vs. MAE = 65.9784, WAPE = 26.2994%, RMSE = 130.9846). Furthermore, the Timedelaynet model demonstrates superior predictive accuracy compared to Narxnet, achieving a lower MAE of 49.9778, WAPE of 19.9215%, and RMSE of 98.3547. Among all models, the Layrecnet model achieves the best performance in this case study, with significantly lower errors (MAE = 18.6198, WAPE = 7.4220%, RMSE = 24.9867), indicating its strong capability in capturing the temporal dynamics of wind power data.

Similarly, Table 6 shows that the Layrecnet model (MAE = 80.9957, WAPE = 5.1301%, RMSE = 104.7831) continues to outperform the Timedelaynet (MAE = 207.4276, WAPE = 13.1381%, RMSE = 289.5628), Narxnet (MAE = 269.7987, WAPE = 17.0885%, RMSE = 382.9613), and Distdelaynet models (MAE = 341.8460, WAPE = 21.6519%, RMSE = 490.9236), further confirming its robustness across different time periods.

In Table 7, the evaluation results also confirm the superiority of the Layrecnet model in the third case study. It achieves the lowest error values (MAE = 41.3159, WAPE = 4.4892%, RMSE = 58.1676), outperforming Timedelaynet (MAE = 164.6041, WAPE = 17.8852%, RMSE = 256.3703), Narxnet (MAE = 168.6443, WAPE = 18.3242%, RMSE = 263.3885), and Distdelaynet (MAE = 248.0907, WAPE = 26.9565%, RMSE = 388.8692).

Figures 11, 12, and 13 illustrate the comparison between the actual wind power generation data and the forecasting results from the four models across case studies 1, 2, and 3, respectively. The actual wind power values are shown as solid blue lines, while the predicted values generated by the Narxnet, Layrecnet, Distdelaynet, and Timedelaynet models are represented by dashed lines in red, blue, purple, and light blue, respectively.

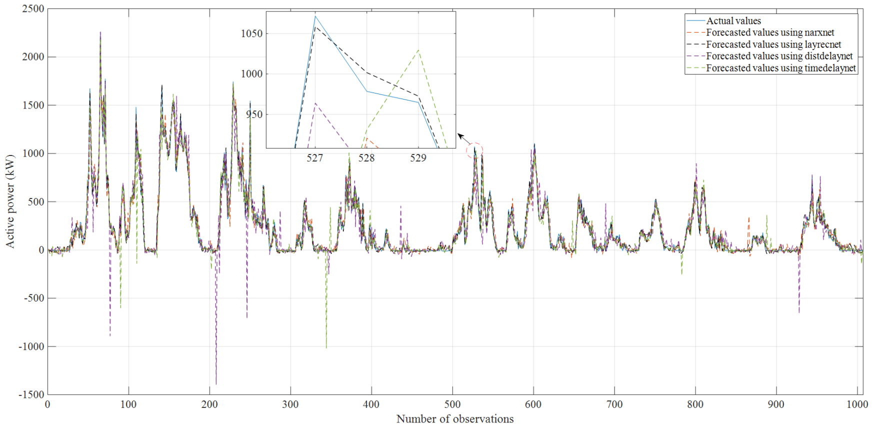

Figure 11. The forecasted results of wind power generation in case study 1.

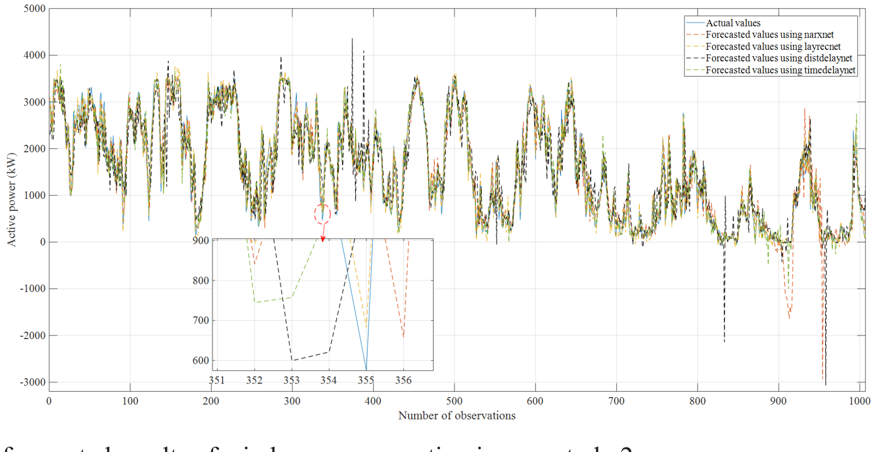

Figure 12. The forecasted results of wind power generation in case study 2.

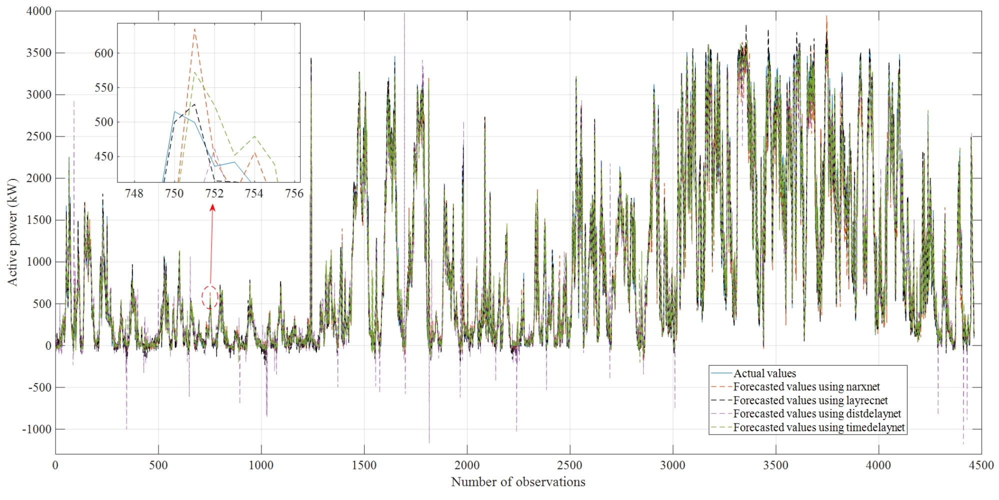

Figure 13. The forecasted results of wind power generation in case study 3.

4. CONCLUSION

In this paper, operational parameters such as wind speed, pitch angle, ambient temperature, nacelle position, and wind direction are taken into account for wind power generation forecasting. A wind farm comprising six turbines, each with a capacity of 3.5 MW, a hub height of 114 meters, and a rotor diameter of 132 meters, is considered. The time series data of turbine no. 05 from July 01st, 2024 to July 31st, 2024 is collected. Several RNN architectures consisting of Narxnet, Layrecnet, Distdelaynet, and Timedelaynet are proposed as alternative approaches for wind power generation forecasting. Among these models, the Layrecnet model demonstrates superior performance WPF results compared to these other models in terms of MAE, WAPE, and RMSE. For further study, the optimal structures of these models can be identified to provide better solutions. Moreover, to improve model reliability, anomalous data points and outliers are identified and filtered out prior to training the forecasting models.

Acknowledgments

This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2024.847.18.

REFERENCES

1. F. Ullah, X. Zhang, M. Khan, M. S. Mastoi, H. M. Munir, A. Flah, Y. Said. A comprehensive review of wind power integration and energy storage technologies for modern grid frequency regulation, *Heliyon*, **2024**, *10*(9), e30466.

2. G. Li, J. Shi. On comparing three artificial neural networks for wind speed forecasting, *Applied Energy*, **2010**, *87*(7), 2313-2320.

3. J. Lerner, M. Grundmeyer, M. Garvert. The importance of wind forecasting, *Renewable Energy Focus*, **2009**, *10*(2), 64-66.

4. A. Kusiak, Z. Zhang. Short-horizon prediction of wind power: a data-driven approach, *IEEE Transactions on Energy Conversion*, **2010**, *25*(4), 1112-1122.

5. X. Zheng, T. Jin. A reliable method of wind power fluctuation smoothing strategy based on multidimensional non-linear exponential smoothing short-term forecasting, *IET Renewable Power Generation*, **2022**, *16*(16), 3573-3586.

6. A. Anwar, A. N. Mahmood. *Enhanced estimation of Autoregressive wind power prediction model using constriction factor particle swarm optimization*, 2014 9th IEEE

Conference on Industrial Electronics and Applications, Hangzhou, China, 2014.

7. J. Wang, Q. Zhou, X. Zhang. *Wind power forecasting based on time series ARMA model*, IOP Conference Series: Earth and Environmental Science, **2018**, *199*(2), 1-6.

8. R. B. Magadum, S. Bilagi, S. Bhandarkar, A. Patil, A. Joshi. *Short-term wind power forecast using time series analysis: Auto-regressive moving-average model (ARMA)*, Recent developments in electrical and electronics engineering: select proceedings of ICRDEEE 2022, Springer Nature Singapore, Singapore, 2023.

9. F. A. Eldali, T. M. Hansen, S. Suryanarayanan, E. K. P. Chong. *Employing ARIMA models to improve wind power forecasts: a case study in ERCOT*, 2016 North American Power Symposium (NAPS), IEEE, Denver, CO, USA, 2016.

10. B. M. Hodge, A. Zeiler, D. Brooks, G. Blau, J. Pekny, G. Reklatis. Improved wind power forecasting with ARIMA models, *Computer Aided Chemical Engineering*, **2011**, *29*, 1789-1793.

11. J. Li, H. Cheng, X. Zhou, M. Wang, Z. Wang. *Optimization of multi-layer perceptron for wind power generation prediction based on improved grey wolf algorithm*, International Symposium on New Energy and Electrical Technology, Springer Nature Singapore, Singapore, 2023.

12. W. Sun, Y. Wang. Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network, *Energy Conversion and Management*, **2018**, *157*, 1-12.

13. J. Huang, R. Qin. Elman neural network considering dynamic time delay estimation for short-term forecasting of offshore wind power, *Applied Energy*, **2024**, *358*, 122671.

14. Z. O. Olaofe, K. A. Folly. *Wind power estimation using recurrent neural network technique*, IEEE Power and Energy Society Conference and Exposition in Africa: Intelligent Grid Integration of Renewable Energy Resources (PowerAfrica), Johannesburg, South Africa, 2012.

15. A. Ahmed, M. Khalid. *A nonlinear autoregressive neural network model for short-term wind forecasting*, 2017 9th IEEE-GCC Conference and Exhibition (GCCCE), IEEE, Manama, Bahrain, 2017.

16. Z. Xiao, F. Tang, M. Wang. Wind power short-term forecasting method based on LSTM and multiple error correction, *Sustainability*, **2023**, *15*(4), 3798.

17. L. Jovanovic, K. Kumpf, N. Bacanin, M. Antonijevic, J. Mani, H. Shaker, M. Zivkovic. *Decomposition aided bidirectional long-short-term memory optimized by hybrid metaheuristic applied for wind power forecasting*, International Conference on Computational Sciences and Sustainable Technologies, Bangalore, India, 2023.

18. Z. Niu, Z. Yu, W. Tang, Q. Wu, M. Reformat. Wind power forecasting using attention-based gated recurrent unit network, *Energy*, **2020**, *196*, 117081.

19. H. Wang, Z. Lei, Y. Liu, J. Peng, J. Liu. Echo state network based ensemble approach for wind power forecasting, *Energy Conversion and Management*, **2019**, *201*, 112188.

20. J. Zeng, W. Qiao. *Support vector machine-based short-term wind power forecasting*, 2011 IEEE/PES Power Systems Conference and Exposition, IEEE, Phoenix, AZ, USA, 2011.

21. S. Park, S. Jung, J. Lee, J. Hur. A short-term forecasting of wind power outputs based on gradient boosting regression tree algorithms, *Energies*, **2023**, *16*(3), 1132.

22. D. Kumar, R. Abhinav, N. Pindoriya. *An ensemble model for short-term wind power*

forecasting using deep learning and gradient boosting algorithms, 2020 21st National Power Systems Conference (NPSC), IEEE, Gandhinagar, India, 2020.

23. X. Yuan, Q. Tan, X. Lei, Y. Yuan, X. Wu. Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine, *Energy*, **2017**, *129*, 122-137.

24. Y. Jiang, X. Chen, K. Yu, Y. Liao. Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm, *Journal of Modern Power Systems and Clean Energy*, **2017**, *5*(1), 126-133.

25. H. H. Aly. *A proposed hybrid deep learning model for wind power forecasting*, 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS), IEEE, Manama, Bahrain, 2024.

26. 26. T. Lin, B. G. Horne, P. Tino, C. L. Giles. Learning long-term dependencies in NARX recurrent neural networks, *IEEE Transactions on Neural Networks*, **1996**, *7*(6), 1329-1338.

27. J. Schmidhuber. Learning complex, extended sequences using the principle of history compression, *Neural Computation*, **1992**, *4*(2), 234-242.

28. A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, K. J. Lang. Phoneme recognition using time-delay neural networks, *IEEE Transactions Acoustics, Speech, Signal Processing*, **1989**, *37*(3), 328–339.

29. T. H. Le, L. Dai, H. Jang, S. Shin. Robust process parameter design methodology: a new estimation approach by using feed-forward neural network structures and machine learning algorithms, *Applied Sciences*, **2022**, *12*(6), 2904.

https://doi.org/10.52111/qnjs.2025.19505