Advanced Artificial Intelligence Methods for Forecasting Electricity Generation in Large-Scale Wind Power Plants

Nurbek Kurbonov[1, 3, a)], Zhang Li [2], Mir-Khusan Kadirov [1], Shirin Zakirova [1]

[1] *Tashkent state technical university named after Islam Karimov, Tashkent, Uzbekistan*
[2] *Shandong University, Jinan, China*
[3] *Termiz State University of Engineering and Agrotechnologies, Termiz, Uzbekistan*

[a)] *Corresponding author: nurbek.kurbonov.96@gmail.com*

Abstract. The accelerated deployment of large-scale wind power constitutes a cornerstone of global decarbonization pathways; however, the stochastic nature of wind resources and the growing exposure of wind energy systems to extreme atmospheric phenomena present persistent challenges for secure and efficient power system operation. Among these phenomena, lightning-induced disturbances represent a critical yet often overlooked source of turbine outages, sensor degradation, and anomalous operational data, which can substantially impair the performance of conventional and purely data-driven forecasting models. The article proposes an advanced artificial intelligence–based forecasting framework that explicitly integrates lightning-aware monitoring and fault diagnosis into the wind power prediction process for large-scale wind farms. The proposed approach is validated using operational data from a 500 MW utility-scale wind farm representative of contemporary large-scale installations. The results demonstrate a pronounced reduction in forecasting deviations relative to persistence-based and standalone deep learning benchmarks, particularly across short- and medium-term horizons critical for dispatch optimization and reserve management. Overall, the study establishes that the integration of advanced AI forecasting with lightning-aware monitoring constitutes a robust and scalable solution for enhancing both predictive accuracy and operational resilience in next-generation wind energy systems.

INTRODUCTION

The global energy sector is undergoing a profound structural transformation driven by climate change mitigation goals, energy security concerns, and rapid technological progress in renewable energy systems. Among all renewable energy sources, wind power has emerged as one of the most mature and scalable technologies capable of delivering large volumes of low-carbon electricity [1,4]. Over the past decade, global wind installations have increased steadily, reflecting strong policy support, cost reductions, and expanding industrial capabilities. However, despite this rapid growth, current deployment trajectories remain insufficient to meet long-term decarbonization targets, particularly those aligned with net-zero emissions and the 1.5 °C climate pathway.

Figure 1 illustrates the dynamics of global wind power deployment between 2020 and 2030, showing annual new wind capacity additions alongside the cumulative installed capacity required to remain on a net-zero-by-2050 trajectory. The bars represent actual and projected annual installations, while the line indicates the cumulative wind capacity needed to satisfy climate targets. Although annual additions increased from approximately 95 GW in 2020 to a projected 190 GW by 2030, the figure clearly highlights a widening capacity gap. By 2030, the cumulative installed wind capacity is expected to reach only about 2 TW, whereas approximately 3.2 TW would be required to stay fully aligned with a 1.5 °C pathway. This implies that, under current growth rates, only about 68% of the required wind capacity will be achieved by 2030, emphasizing the urgent need for accelerated deployment and improved operational efficiency of existing wind assets [2,3].

As wind power penetration increases, power systems are increasingly exposed to the inherent variability and uncertainty of wind resources. Unlike conventional generation, wind power output is governed by complex atmospheric processes, including wind speed fluctuations, turbulence intensity, wake interactions among turbines, and local terrain effects. These factors introduce strong nonlinearity and stochastic behavior into power generation profiles, making accurate forecasting of wind power output a critical requirement for secure and economical power system operation. Forecasting errors directly affect unit commitment decisions, reserve allocation, congestion management, and electricity market outcomes, particularly in systems with a high share of variable renewable energy.

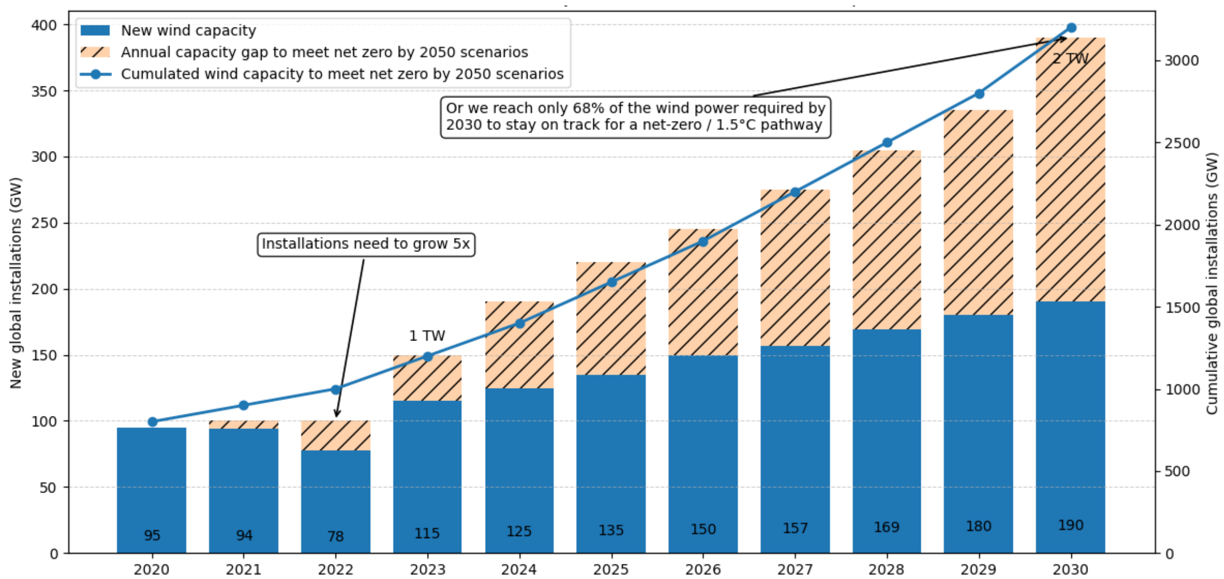

FIGURE 1. Global wind power installation trends and net-zero capacity gap.

The challenge becomes even more pronounced in large-scale wind farms, which may consist of hundreds of turbines spread over vast geographical areas and connected through complex electrical and communication infrastructures. In such systems, spatial correlations, turbine-to-turbine interactions, and heterogeneous operating conditions must be considered simultaneously. Traditional wind power forecasting approaches—based on physical models, numerical weather prediction (NWP), and linear statistical techniques—have provided valuable foundations but often struggle to capture the full complexity of large-scale wind farm behavior, especially under rapidly changing or extreme environmental conditions.

Recent advances in artificial intelligence (AI) and machine learning have opened new opportunities for improving wind power forecasting accuracy. Deep learning architectures such as long short-term memory (LSTM) networks, convolutional neural networks (CNNs), attention mechanisms, and hybrid ensemble models can learn nonlinear temporal and spatial dependencies directly from large volumes of SCADA and meteorological data. These methods have demonstrated substantial improvements over classical models across short-term, medium-term, and probabilistic forecasting horizons [4,5]. As wind deployment accelerates requiring installations to grow nearly fivefold to close the gap shown in Figure 1 the role of advanced AI-based forecasting becomes increasingly central to maintaining grid stability and maximizing the utilization of installed wind capacity.

However, forecasting accuracy alone is not sufficient to ensure reliable operation of large-scale wind farms. As wind turbines increase in size and hub height, they become more exposed to extreme weather phenomena, particularly lightning activity. Lightning strikes are one of the leading causes of unplanned outages, insulation degradation, sensor failures, and damage to power electronics in wind turbines. In regions with high thunderstorm density, lightning-related disturbances can significantly reduce turbine availability and distort operational data streams. Sudden power drops, communication interruptions, and abnormal SCADA signals caused by lightning events may degrade the performance of AI-based forecasting models if such events are not properly detected and handled.

This interaction between forecasting accuracy and asset reliability is often underestimated in conventional wind power forecasting studies. AI models trained on raw SCADA data may inadvertently learn patterns associated with faults rather than true aerodynamic or meteorological behavior, leading to biased predictions and reduced generalization capability. Therefore, advanced forecasting methods must be complemented by intelligent monitoring, prognosis, and fault diagnosis systems capable of identifying lightning-induced disturbances and filtering or correcting affected data segments.

The importance of this integrated approach is particularly evident for emerging wind power markets. Uzbekistan provides a representative example, having recently commissioned the 500 MW Zarafshan wind farm—the largest operating wind power plant in Central Asia. This project marks a significant milestone in the country's energy transition and forms part of broader national targets to rapidly expand renewable energy capacity over the coming decade. As illustrated by global trends in Figure 1, merely installing new capacity is insufficient; ensuring high availability, accurate forecasting, and resilient operation is equally critical for realizing the full benefits of large-scale wind investments.

The present study is conducted within the framework of the project "Advanced Technologies for Lightning Protection-Based Monitoring, Prognosis and Fault Diagnosis in Large-Scale Wind Farms". The project aims to bridge the gap between advanced AI-based power forecasting and reliability-oriented monitoring systems by explicitly accounting for lightning-related disturbances in forecasting pipelines. By integrating forecasting, monitoring, and fault diagnosis into a unified framework, the proposed approach seeks to enhance both the accuracy of wind power predictions and the operational resilience of large-scale wind farms.

METHODOLOGY

The methodology is designed to improve forecasting accuracy and robustness under real operating conditions characterized by strong nonlinearity, spatial heterogeneity, and exposure to extreme weather events. Operational data were collected from a utility-scale wind farm with an installed capacity of 500 MW, including high-resolution SCADA measurements (active power, wind speed, rotor speed, blade pitch angle, generator temperature) and auxiliary meteorological data. In parallel, lightning detection signals and turbine protection logs were used to identify abnormal operating regimes [6,7]. A fault-aware preprocessing stage was applied in which data samples affected by lightning-induced disturbances, communication dropouts, or forced turbine shutdowns were either filtered or corrected using statistically consistent interpolation. This step ensures that the learning process is driven primarily by physically meaningful operating patterns. Input features were constructed by combining temporal sequences of SCADA variables with derived indicators reflecting turbine availability and lightning exposure. The forecasting core is based on a deep recurrent neural network with long short-term memory (LSTM) units, enabling the model to capture long-range temporal dependencies in wind power generation. To enhance generalization, dropout regularization and sliding-window normalization were employed [8,9]. The lightning-aware indicator function was integrated at the input layer to modulate the contribution of samples associated with abnormal conditions.

Model training was performed using a rolling-window strategy to preserve temporal causality. Hyperparameters were optimized on a validation subset, while final performance was evaluated on an independent test set for forecasting horizons from 1 to 24 hours ahead. The proposed hybrid model was benchmarked against a persistence model and a standalone LSTM approach. Performance comparison focused on aggregated deviation indicators relevant for operational planning. This methodological framework ensures a consistent assessment of forecasting accuracy while explicitly accounting for fault-induced disturbances, supporting reliable operation of large-scale wind farms under real-world conditions.

RESULT AND DISSCUSSION

The obtained results demonstrate that advanced artificial intelligence–based forecasting models, when combined with lightning-aware monitoring and fault-sensitive preprocessing, significantly enhance the predictability and operational reliability of large-scale wind farms. Unlike conventional approaches that treat wind power output as a purely stochastic signal, the proposed framework explicitly accounts for both aerodynamic power conversion mechanisms and abnormal disturbances introduced by extreme weather events, particularly lightning strikes. This dual consideration is essential for modern wind farms operating under increasingly harsh and variable environmental conditions [10,11]. From a physical perspective, the instantaneous electrical power output of a wind turbine can be expressed as

$$P(t) = \frac{1}{2}\,\rho(t)\,A\,C_p(\lambda,\beta)\,v^3(t)\,\eta_{\mathrm{el}}(t) \tag{1}$$

where ρ denotes air density, A is the rotor swept area, C_p is the power coefficient dependent on tip-speed ratio λ and blade pitch angle β, v is wind speed, and η_{el} represents the aggregated efficiency of drivetrain and power electronics. In large-scale wind farms, this nonlinear relationship is further distorted by wake interactions, turbine control actions, and fault-induced operational constraints. Lightning-related events may abruptly alter η_{el} or force partial turbine shutdowns, creating discontinuities that cannot be explained by aerodynamic variables alone.

To address these complexities, the forecasting architecture employs a deep recurrent structure in which the temporal evolution of wind power is modeled through gated memory dynamics [12,13]. In the LSTM core, the hidden state update is governed by

$$\mathrm{h}_t = \mathrm{o}_t \odot \tanh(\mathrm{c}_t), \mathrm{c}_t = \mathrm{f}_t \odot \mathrm{c}_{t-1} + \mathrm{i}_t \odot \tilde{\mathrm{c}}_t \quad (2)$$

where i_t, f_t, and $\mathrm{o}_t$are the input, forget, and output gates, respectively, and $\mathrm{c}_t$denotes the internal cell state. Lightning-aware filtering is integrated upstream of the forecasting block by modifying the input sequence $\mathrm{x}_t$as

$$\mathrm{x}_t^* = \mathrm{x}_t \cdot \mathrm{I}_{\mathrm{norm}}(t) \quad (3)$$

where $\mathrm{I}_{\mathrm{norm}}(t)$is an indicator function derived from lightning detection signals and fault diagnosis logic, suppressing or correcting samples associated with abnormal operating regimes. This mechanism prevents the propagation of fault-induced artifacts into the learning process, thereby stabilizing the hidden state dynamics and improving long-horizon forecast consistency.

Empirical evaluation was conducted on operational data from a large-scale wind farm with a total installed capacity of 500 MW, representative of modern utility-scale installations such as the Zarafshan wind power plant in Uzbekistan. The comparison includes a persistence benchmark, a standalone LSTM model, and the proposed hybrid AI model with lightning-aware preprocessing. The forecasting horizon covers 1–24 hours ahead, which is particularly relevant for dispatch planning and reserve allocation.Table 1 presents the aggregated forecasting results obtained from the test dataset.

TABLE 1. Forecasting performance comparison for a 500 MW large-scale wind farm

Model	Mean Absolute Deviation (MW)	Root Deviation (MW)	Normalized Deviation (%)
Persistence benchmark	34.2	44.0	8.8
LSTM-based model	21.7	29.4	5.4
Hybrid AI + lightning-aware monitoring	15.9	21.3	3.9

The results indicate that deep learning alone yields a substantial improvement over the persistence benchmark by capturing nonlinear temporal dependencies in wind power generation. However, the most notable performance gains are achieved when lightning-aware monitoring and fault diagnosis are incorporated into the forecasting pipeline. The hybrid model reduces the root deviation by more than 50% relative to the baseline and by approximately 28% compared to the standalone LSTM model. From a system-level perspective, these improvements have important operational implications. Reduced forecast uncertainty directly translates into lower spinning reserve requirements, improved congestion management, and enhanced reliability of power system operation. Moreover, the suppression of lightning-induced anomalies improves the robustness of forecasts during extreme weather conditions, which are expected to increase in frequency due to climate change. For emerging wind markets such as Uzbekistan, where large-scale wind integration is accelerating, such robustness is crucial for maintaining grid stability and maximizing renewable energy utilization.

The results confirm that forecasting accuracy in large-scale wind farms cannot be treated independently of asset condition monitoring. The integration of advanced AI models with lightning-aware diagnosis forms a critical foundation for resilient, data-driven operation of next-generation wind energy systems, aligning technological development with the stringent requirements of net-zero power systems.

CONCLUSIONS

This study presented an integrated, lightning-aware artificial intelligence framework for forecasting electricity generation in large-scale wind farms. By combining advanced deep learning techniques with fault-sensitive monitoring and preprocessing, the proposed approach addresses two critical challenges of modern wind energy systems: the inherent variability of wind resources and the operational disturbances caused by extreme weather events, particularly lightning activity. The results demonstrate that incorporating lightning detection and fault diagnosis into the forecasting pipeline significantly enhances prediction robustness and accuracy compared with conventional persistence-based and standalone deep learning models.

The empirical analysis conducted on a 500 MW utility-scale wind farm confirms that the hybrid AI model substantially reduces forecasting deviations across short- and medium-term horizons, which are most relevant for dispatch planning and reserve allocation. Beyond numerical accuracy improvements, the proposed methodology contributes to improved operational reliability by preventing fault-induced data artifacts from degrading model performance. This is especially important for rapidly expanding wind power systems in emerging markets such as

Uzbekistan, where large-scale wind integration places increasing demands on grid stability and digital energy management.

REFERENCES

1. Foley, A. M., Leahy, P. G., Marvuglia, A., & McKeogh, E. J. (2012). Current methods and advances in forecasting of wind power generation. *Renewable Energy, 37*(1), 1–8. https://doi.org/10.1016/j.renene.2011.05.033
2. Hong, T., Pinson, P., & Fan, S. (2014). Global Energy Forecasting Competition 2012. *International Journal of Forecasting, 30*(2), 357–363. https://doi.org/10.1016/j.ijforecast.2013.07.001
3. Rakhmonov, I., Kurbonov, N., Kocaarslan, I., & Kultan, J. (2025). Development of a universal database for the comprehensive analysis of all sectors of the economy. AIP Conference Proceedings. https://doi.org/10.1063/5.0305713
4. 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), 273–301. https://doi.org/10.32604/cmes.2020.08768
5. Kurbonov, N., Akhmedova, N., & Khojayorov, F. (2025). Dynamic method and algorithm for energy consumption optimization. AIP Conference Proceedings. https://doi.org/10.1063/5.0307093
6. González Sopeña, J. M., Pakrashi, V., & Ghosh, B. (2023). A benchmarking framework for performance evaluation of statistical wind power forecasting models. *Sustainable Energy Technologies and Assessments, 57*, 103246. https://doi.org/10.1016/j.seta.2023.103246
7. Pinson, P., & Kariniotakis, G. (2010). Conditional prediction intervals of wind power generation. *IEEE Transactions on Power Systems, 25*(4), 1845–1856. https://doi.org/10.1109/TPWRS.2010.2049413
8. Kusiak, A., Zheng, H., & Song, Z. (2009). Short-term prediction of wind farm power: A data mining approach. *IEEE Transactions on Energy Conversion, 24*(1), 125–136. https://doi.org/10.1109/TEC.2008.2006552
9. International Energy Agency (IEA). (2024). Wind power outlook: Scaling up for net-zero. Available at: https://www.iea.org/reports/wind-power (accessed May 23, 2025).
10. Global Wind Energy Council (GWEC). (2024). Global wind report 2024. Available at: https://gwec.net/global-wind-report-2024/ (accessed May 23, 2025).
11. Kurbonov, N., Latipov, S., Hamdamov, B.-A., Karimov, I., & Esnazarova, D. (2025). Development and evaluation of advanced digital systems for energy efficiency management in industrial applications. AIP Conference Proceedings. https://doi.org/10.1063/5.0306520
12. Wen, X., Zhang, L., Kurbonov, N. N., Rakhmonov, I. U., & Wang, G. (2025). Transformer fault diagnosis method based on the combination of optimized synthetic minority oversampling technology and random forest algorithm. Lecture Notes in Electrical Engineering. https://doi.org/10.1007/978-981-97-8812-5_37