



POWER PREDICTION IN A WIND FARM USING MACHINE LEARNING MODELS

Jessica N. Castillo¹, Gabriel Pesántez², Víctor Hugo Medina Matute³ and Jose E. Naranjo⁴

¹Facultad de Ciencias de la Ingeniería y Aplicadas, Universidad Técnica de Cotopaxi, UTC, Latacunga, Ecuador.

²Facultad de Informática y Electrónica, Escuela Superior Politécnica de Chimborazo, Riobamba, Ecuador.

³Facultad de Ciencias de la Ingeniería y Aplicadas, Universidad Técnica de Cotopaxi, UTC, Latacunga, Ecuador.

⁴Facultad de Ciencias de la Ingeniería y Aplicadas, Universidad Técnica de Cotopaxi, UTC, Latacunga, Ecuador.

¹<https://orcid.org/0000-0002-3120-7229>^{ID}, ²<https://orcid.org/0000-0003-4037-3859>^{ID}

³<https://orcid.org/0000-0002-6149-453X>^{ID}, ⁴<https://orcid.org/0000-0002-2884-1667>^{ID}

Email: *jessica.castillo@utc.edu.ec, gabriel.pesantez@esepoch.edu.ec, victor.medina@utc.edu.ec, jose.naranjo0463@utc.edu.ec

ARTICLE INFO

Article History

Received: February 12, 2026

Reviewed: April 19, 2026

Accepted: May 19, 2026

Published: June 30, 2026

Keywords:

Machine Learning;

Electrical Power;

Wind Farm;

Wind energy;

Renewable energy.

ABSTRACT

This paper presents the development and evaluation of an advanced power prediction system for a wind farm using machine learning and deep learning models, based on historical meteorological data. The analyzed dataset comprises 25,596 hourly records corresponding to the period 2022–2024, which were subjected to a rigorous process of cleaning, physical validation, and generation of derived variables, such as air density and power according to the Betz limit. Based on a Pearson correlation analysis, wind speed, solar radiation, and air temperature were identified as the most influential variables in power generation. Multiple predictive models were trained and compared, including linear regression, Random Forest, XGBoost, MLP neural networks, LSTM, and Transformers, evaluated using RMSE and coefficient of determination (R^2). Additionally, lagged variables were incorporated to improve prediction at 1-, 3-, and 6-hour horizons. The results show that a hybrid LSTM + XGBoost architecture, based on residual error correction, consistently outperforms individual models, achieving an RMSE of 61.98 W and an R^2 of 0.779 in one-hour prediction. Finally, the proposed model was implemented in an operational environment, demonstrating its viability for real-world energy management and planning applications.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution with International License (CC BY 4.0).

I. INTRODUCTION

Wind energy serves as a critical renewable resource for fostering sustainable development and mitigating global carbon emissions [1]. However, the inherent variability and stochastic nature of wind patterns present significant hurdles for accurate wind power forecasting, which is vital for maintaining grid stability and ensuring efficient energy management [2]. This complexity drives the continuous need for sophisticated prediction systems that can maximize resource utilization and optimize power grid operation [3]. The evolution of WPP has seen a transition toward advanced Deep Learning and Machine Learning models. For instance, regression trees provide a versatile alternative for forecasting output across various types of renewable energy plants [4]. In the realm of layout optimization, Artificial Neural Networks combined with genetic algorithms are employed to resolve design challenges and evaluate control benefits [5]. Furthermore, researchers emphasize the necessity of spatially transferable models to avoid the high computational expense associated with recalibrating parameters at new geographical locations [6]. Data-driven frameworks using machine learning wake models and Computational Fluid Dynamics (CFD) simulation data have also improved layout efficiency in large-scale wind farms [7]. Performance improvements in the state of the art are often achieved through specific algorithm selection and hybridization. Extreme Gradient Boosting (XGBoost) demonstrates higher accuracy and faster training times compared to traditional ANN for predicting power under yaw-controlled conditions [8]. Recent studies also highlight the efficacy of Gradient Boosting regressors when enriched with reanalysis meteorological data and weather forecasts [9]. To address the lack of historical data in newly constructed wind farms, few-shot learning approaches incorporating Time-series Generative Adversarial Networks are currently being developed [10].

Simultaneously, the identification and cleaning of abnormal wind power data resulting from sensor failures remain a primary concern for maintaining the integrity of prediction models [11]. Novel methods such as Generative Critic Networks coupled with Multi-Task Learning allow for high-quality interval forecasts across neighboring wind farms [12]. Therefore, the present work aims to develop a system for predicting the power generated by a wind power plant based on available meteorological variables. To this end, a structured approach is followed, encompassing data cleaning, correlation analysis, the training of regression models, and the validation of results.

II. DATA CLEANING AND PREPARATION

Preprocessing was performed to rename columns, convert data types, remove null values, and calculate new derived variables such as air density and power according to the Betz limit. The final columns of the dataset are shown in Table 1.

Table 1: Variables available after data cleaning.

Statistics	Fecha_Hora	Temperatura_C	Humedad_Relativa	Velocidad_Viento	Presion_AireSeco_Pa	Densidad_Aire	Potencia_W	Potencia_Betz_W
count	25596	25596.000000	25596.000000	25596.000000	25596.000000	25596.000000	25596.000000	25596.000000
mean	2023-06-30 13:51:27.482419	6.852035	85.008327	2.458722	71030.935737	0.890324	115.157889	68.288628
min	2022-01-01 00:00:00	-2.430000	31.290000	0.150000	70476.167791	0.860954	0.016047	0.009516
25%	2022-10-02 16:45:00	5.170000	78.630000	1.570000	70930.174030	0.884504	18.765575	11.127986
50%	2023-06-28 23:30:00	6.480000	88.275000	2.459506	71031.183668	0.891389	71.225671	42.236823
75%	2024-03-28 12:15:00	8.500000	93.600000	3.300000	71129.346041	0.895958	172.706470	102.414937
max	2024-12-31 23:00:00	16.140000	99.900000	4.920000	71737.976814	0.921243	579.088300	343.399362

Source: Authors, (2026).

A statistical summary was then performed, confirming that the dataset is extensive (more than 25,000 records) and covers a wide range of environmental conditions.

III. CORRELATION ANALYSIS

In order to select the variables that most influence the power generated, Pearson's correlation matrix was calculated between all available variables (Figure 1). This analysis identified significant relationships between the target variable (Potencia_W) and others such as wind speed, radiation, and temperature.

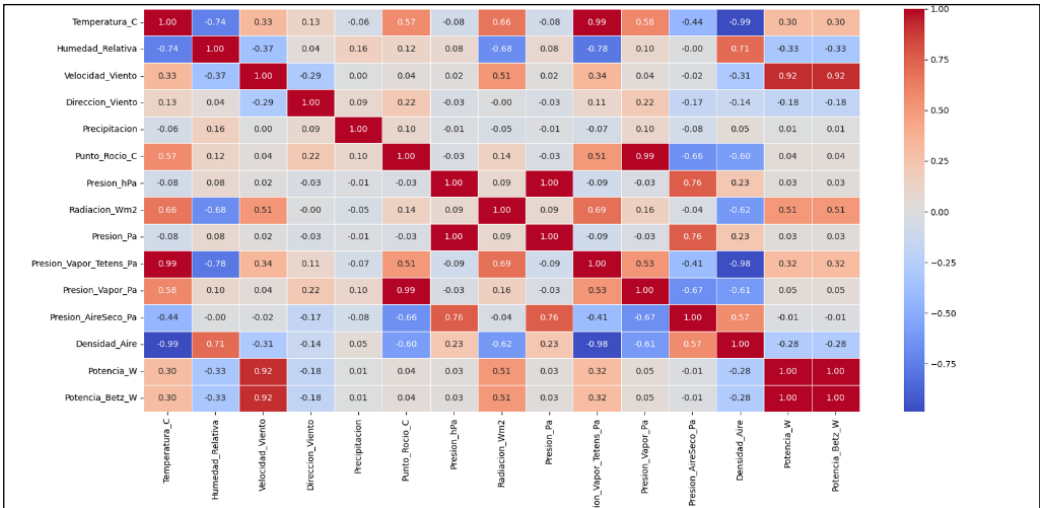


Figure 1: Correlation matrix between variables.
Source: Authors, (2026).

As can be seen, wind speed has a direct correlation of 0.92 with the power generated, confirming its primary role. A significant correlation was also observed with radiation (0.51) and, to a lesser extent, with temperature (0.30). Therefore, these three variables were selected as inputs for the model. Although vapor pressure exhibited a slightly higher Pearson correlation with generated power than air temperature (0.32 vs. 0.30), it was not included as an input variable because it is not an independent atmospheric driver, but rather a derived quantity strongly dependent on temperature and relative humidity. Including vapor pressure would therefore introduce redundancy and potential multicollinearity without providing additional physical or predictive information beyond that already captured by temperature.

IV. MODELING AND PREDICTION

Four different regression models were trained and compared: linear regression, Random Forest, XGBoost, and an MLP neural network. To do this, the data was divided into training and test sets (80/20). The results were evaluated using root mean square error (RMSE) and coefficient of determination R^2 metrics. The values are shown in Table 2.

Table 2: Model evaluation results.

Model	RMSE	R^2
Regresión Lineal	47.17	0.8520
Random Forest	0.42	1.0000
XGBoost	1.36	0.9999
Red Neuronal (MLP)	6.48	0.9972

Source: Authors, (2026).

The XGBoost model showed an excellent balance between accuracy and generalization, with very low error and an almost perfect fit. Linear regression, although functional, fails to capture the nonlinearity of the phenomenon. Figure 2 shows how nonlinear models (Random Forest, XGBoost, and MLP) fit the actual values more accurately. Linear regression underestimates high values due to its functional form. Although the Random Forest model achieved the lowest RMSE and a perfect coefficient of determination ($R^2 = 1.0000$), it was not selected for further analysis. Such a perfect fit strongly suggests overfitting to the training data, implying that the model may be capturing noise or data-specific patterns rather than generalizable relationships. Consequently, its excellent performance on the test split does not necessarily guarantee robust predictive capability when exposed to new, unseen real-world data, which limits its suitability for operational forecasting.

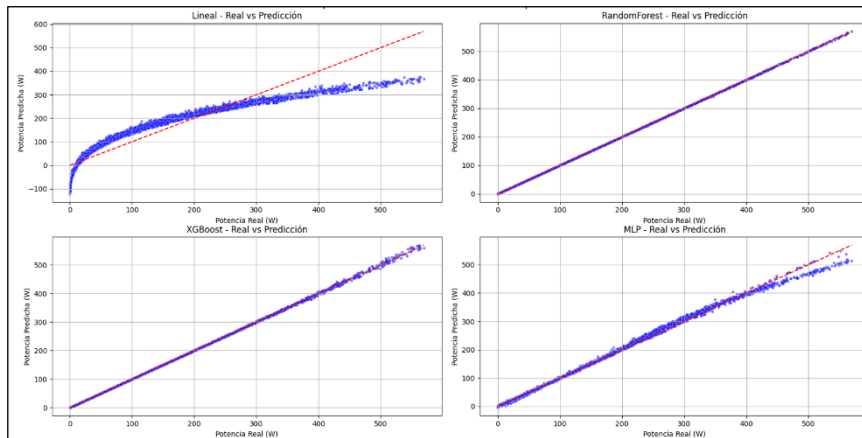


Figure 2: Graphical comparison between actual values and predictions.

Source: Authors, (2026).

V. USER INTERFACE AND IMPLEMENTATION

To facilitate the use of the model, a web interface was developed using Streamlit (Figure 3). This allows the user to enter environmental conditions (wind speed, temperature, and solar radiation) and obtain the predicted power immediately.



Figure 3: Prediction interface developed.

Source: Authors, (2026).

In addition, an automatic recording system was implemented that saves each prediction made along with the conditions entered and the time of calculation (Table 3). This allows prediction histories to be generated for further analysis.

Table 3: Example of recorded predictions.

Velocidad_Viento	Radiacion_Wm2	Temperatura_C	Potencia_Predicha_W	Fecha_Hora
2.5	100	8.0	74.79	2025-06-29 09:12:00
4.8	800	14.0	358.69	2025-06-29 09:15:00

Source: Authors, (2026).

VI. TIME-SERIES POWER FORECASTING

In order to anticipate wind power generation behavior, a time-series forecasting model was implemented using available historical data. The strategy consisted of reorganizing the information so that current environmental conditions served as input, and the power generated one hour later as expected output. The model selected was an XGBoost regressor, trained with the variables: wind speed, solar radiation, and air temperature. The model obtained an RMSE of 80.11 W and a coefficient of determination R^2 of 0.63, indicating moderate short-term prediction capability. Figure 4 shows a comparison between the actual power generated and the power predicted by the model for a subset of data. Although there are occasional differences, the model correctly follows the general trends of growth and decline, showing an acceptable correlation for operational estimation applications.

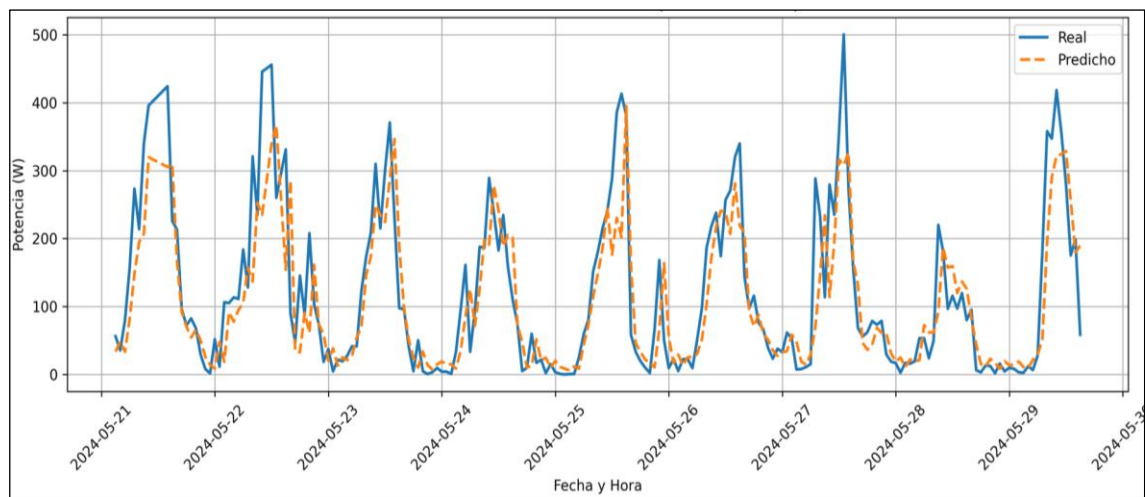


Figure 4: Prediction of future power one hour ahead using the XGBoost model.

Source: Authors, (2026).

VII. COMPARISON OF APPROACHES FOR FUTURE PREDICTION

To improve the estimation of power generated one hour in the future, three approaches with different input configurations were evaluated:

1. **Model A (base):** Uses current environmental variables as input: wind speed, temperature, and solar radiation.
2. **Model B (simplified):** Uses only wind speed, as it is the variable with the highest direct correlation with power.
3. **Model C (with delays):** Adds the weather conditions from one hour ago (lag delays = 1) as inputs, in addition to the current conditions.

Table 4 shows the results obtained for each model, evaluated using the root mean square error (RMSE) and the coefficient of determination (R^2).

Table 4: Comparison of one-hour future prediction models.

Modelo	Variables used	RMSE (W)	R^2
A	Speed, Radiation, Temperature	80.11	0.6303
B	Wind Speed only	82.87	0.6045
C	Speed, Radiation, Temperature + Delay (t-1)	78.62	0.6440

Source: Authors, (2026).

The results indicate that the base model (A) already provides a reasonable estimate, but that including weather conditions from the immediate past (model C) slightly improves accuracy. In contrast, reducing the inputs to only one variable (model B) limits performance, despite its high correlation.

VIII. POWER FORECASTING IN A WIND FARM WITH THE INCLUSION OF LAGGED VARIABLES

With the aim of improving the predictive capacity of the future power model, a strategy based on the inclusion of variables with time lags was implemented. This methodology allows the model to take into account the recent behavior of weather conditions and electricity generation, providing temporal context to the prediction process. Lags of 1, 3, and 6 hours ($t-1$, $t-3$, $t-6$) were considered for the most relevant variables identified in previous phases: wind speed, solar radiation, and power generated. These new variables were added to the dataset as additional inputs, maintaining the power generated at 1h, 3h, and 6h in the future as target outputs. The model used was again a multi-output XGBoost, trained with the expanded data. The evaluation was performed using the RMSE and R^2 metrics, as in the previous approaches. The results indicate a substantial improvement over longer horizons, as shown in Table 5.

Table 5: Comparison of model performance with and without lagged variables.

Horizonte	RMSE without delays	R^2 without delays	RMSE with delays	R^2 with delays
1 hour	79.01 W	0.6404	77.13 W	0.6573
3 hours	106.43 W	0.3475	99.05 W	0.4349
6 hours	117.07 W	0.2107	108.68 W	0.3198

Source: Authors, (2026).

At 1 hour, the model with lags maintains a performance very similar to the previous one, with a slight improvement. However, at 3 and 6 hours, there is a significant jump in the ability to track trends. The model without lags tended to produce smoothed curves that did not capture the peaks and troughs of generation well. By incorporating recent history, the model with lags manages to more accurately track abrupt and sustained variations over time, as can be seen in Figure 5.

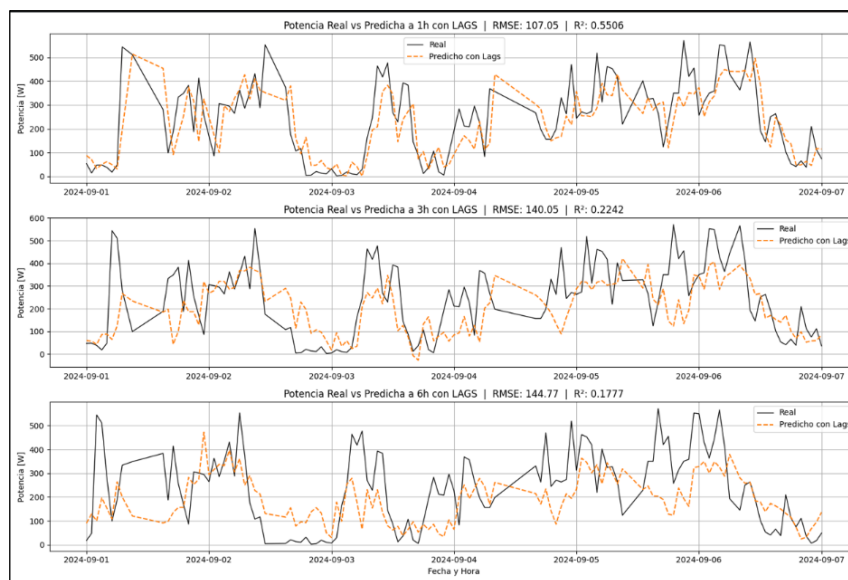


Figure 5: Graphical comparison between actual and predicted power at 1h, 3h, and 6h using a model with lagged variables.

Source: Authors, (2026).

This behavior is especially important in operational applications, where predicting sudden drops or increases in power several hours in advance can impact dispatch decisions and grid stability. Visual analysis confirms that predictions with lags are better able to track actual dynamics, even under highly variable conditions. In summary, the introduction of lagged variables into the future prediction model has proven to be an effective improvement, particularly for horizons longer than one hour. This technique adds memory to the system and allows for better modeling of the temporal patterns that govern wind generation. During the visual analysis of this figure, several abrupt low points were observed in the actual power curve. This prompted an additional review of the quality of the original data, identifying records where the power was close to zero despite sufficient wind or solar radiation for expected generation. These records were marked as potentially erroneous and removed in a second phase of advanced cleansing. After applying this additional cleansing, the original model was retrained without lags to assess the exclusive impact of this cleansing. The results are shown in Table 6.

Table 6: Comparison of model performance without lags before and after advanced cleansing.

Horizon	Original RMSE	Original R^2	RMSE with cleaning	R^2 with cleaning
1 hour	79.01 W	0.6404	79.26 W	0.6387
3 hours	106.43 W	0.3475	105.75 W	0.3570
6 hours	117.07 W	0.2107	118.12 W	0.1977

Source: Authors, (2026).

As can be seen, advanced cleaning does not significantly alter the model's short-term results, but it slightly improves stability at 3 hours, where a small gain in R^2 and a reduction in RMSE are obtained. Although the numerical improvement is modest, this step ensures the elimination of outliers that could distort the interpretation of the results or introduce noise into the training. This strengthens the quality of the dataset as a basis for future implementations.

IX. LSTM MODEL AND ADVANCED VERIFICATION

After implementing tree-based models and classic regressors, we proceeded to test architectures more suitable for time series, in this case an LSTM (Long Short-Term Memory) neural network. Two versions were explored: a basic one and an improved one, evaluating their performance for power prediction at 1h, 3h, and 6h.

IX.1 BASIC LSTM ARCHITECTURE

A simple LSTM network was designed with a hidden layer and multivariable output. The data was normalized with MinMaxScaler, and sequences with a 6-hour delay were used as input. The model was trained for 20 epochs with no apparent overfitting. Figure 6 shows a visual comparison of this model's predictions for the three time horizons.

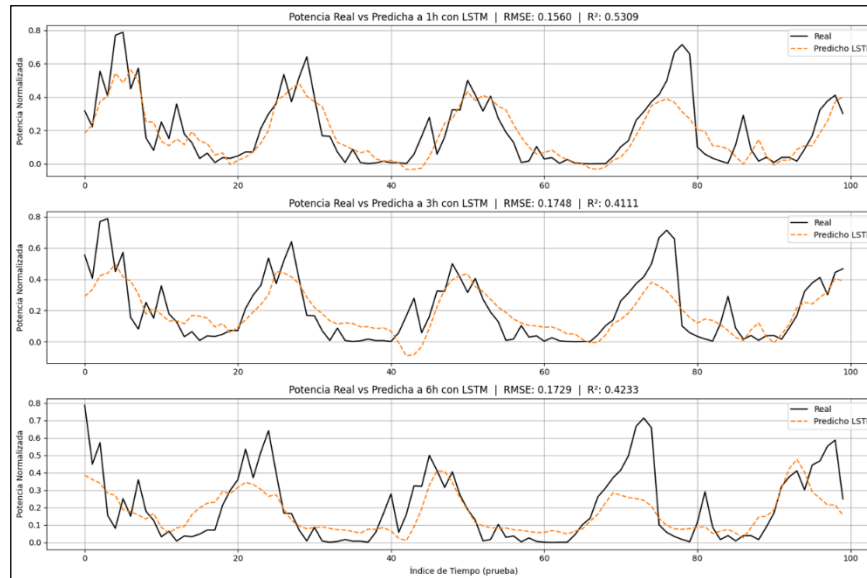


Figure 6: Graphical comparison between actual and predicted power at 1h, 3h, and 6h with basic LSTM.
Source: Authors, (2026).

The quantitative performance of the basic LSTM model for each prediction horizon is summarized in Table 7, where the normalized RMSE and the coefficient of determination (R^2) are reported for the test dataset.

Table 7: Performance metrics of the basic LSTM model at different prediction horizons.

Horizon	RMSE (norm.)	R^2
1h	0.1560	0.5309
3h	0.1748	0.4111
6h	0.1729	0.4233

Source: Authors, (2026).

IX.2 IMPROVED LSTM

To improve accuracy, a deeper architecture with two LSTM layers (64 and 32 neurons), Dropout regularization (0.2), and EarlyStopping with a patience of 5 epochs was tested. The 1-hour prediction is shown in Figure 7.

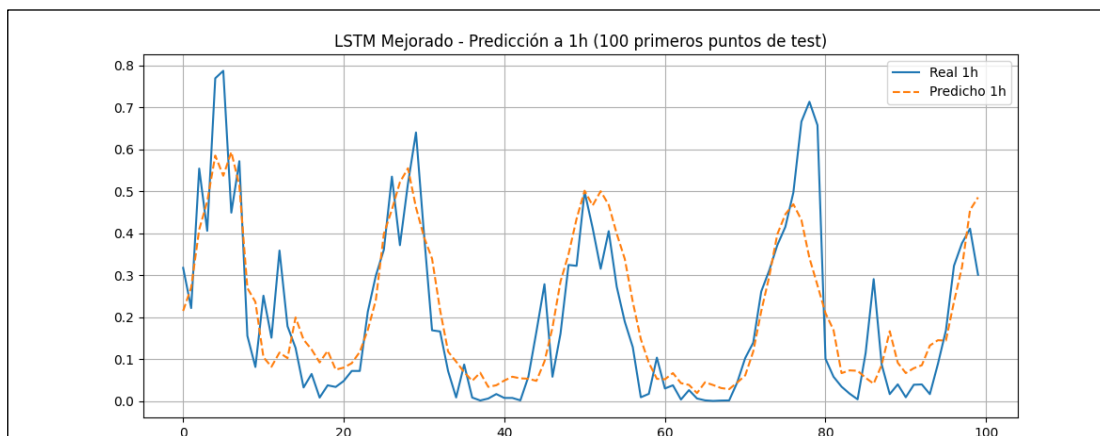


Figure 7: 1-hour prediction curve using Enhanced LSTM.
Source: Authors, (2026).

Compared to previous models, the enhanced LSTM outperforms XGBoost in 3-hour and 6-hour predictions and approaches its performance in the 1-hour horizon. The metrics achieved are shown in Table 8.

Table 8: Performance metrics of the improved LSTM model at different prediction horizons.

Horizon	RMSE (norm.)	R ²
1h	0.1544	0.5403
3h	0.1705	0.4395
6h	0.1716	0.4315

Source: Authors, (2026).

IX.3 DATASET QUALITY VERIFICATION

Given that the predictions had certain errors at low power points, a visual and statistical verification of the integrity of the final dataset (datos_limpios_avanzado.csv) was performed. Figure 8 shows the histogram of the power generated.

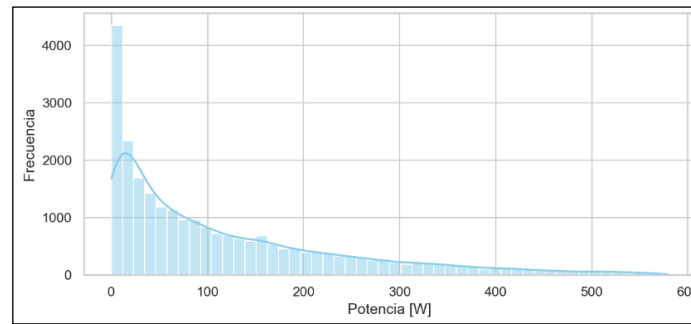


Figure 8: Distribution of power generated.

Source: Authors, (2026).

The most relevant physical relationships were also analyzed. This can be seen in Figure 9.

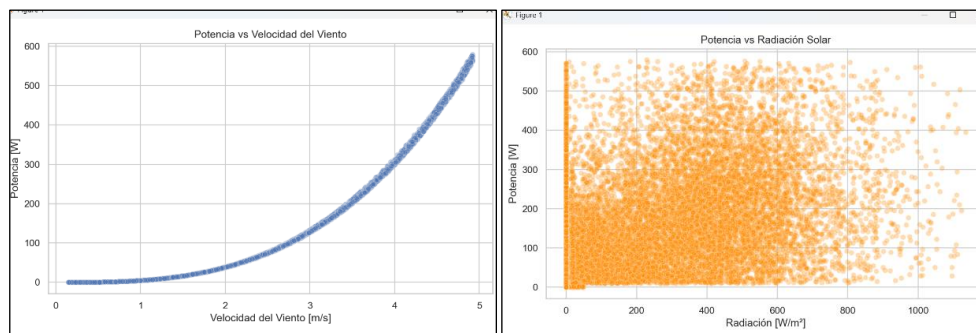


Figure 9: Power vs. Wind speed (left) and Power vs. Solar radiation (right).

Source: Authors, (2026).

The curve with respect to wind showed ideal parabolic behavior, while greater dispersion was observed with solar radiation. The latter could be due to instrumental noise or conditions not captured by the available variables. Finally, the average daily curve was evaluated (Figure 10), confirming the expected shape with maximum production during daylight hours.

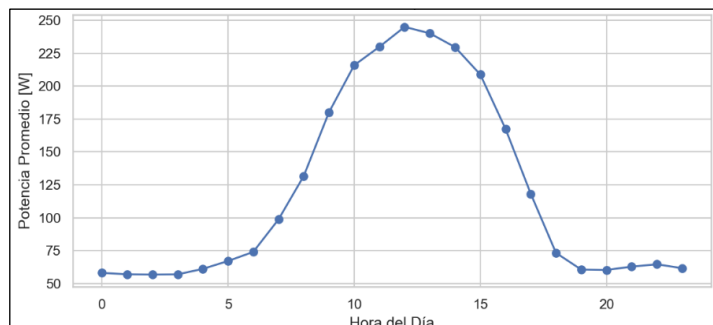


Figure 10: Average power per hour of the day.

Source: Authors, (2026).

Validation of the dataset showed that the data are physically consistent. It was confirmed that it is not necessary to eliminate all values close to zero, but only those that are not justified by physical conditions. This criterion was successfully applied in advanced cleaning and allowed the LSTM model to achieve competitive predictions, especially in the medium term.

IX.4 TECHNICAL NOTE ON POWER VALUES EQUAL TO OR CLOSE TO ZERO

During data exploration and verification, it was observed that a considerable proportion of records have power values close to zero. This situation is to be expected in wind or hybrid generation systems and does not necessarily represent an error. There are entirely valid physical conditions in which the power generated can be zero, for instance:

- Nighttime hours, in the absence of solar radiation.
- Wind speeds below the minimum cut-off speed (typically < 2.5 m/s).
- Periods of scheduled downtime or system standby conditions.

However, records were also detected in which the power was ≈ 0 even though the weather conditions were sufficient to produce energy. Specifically, combinations such as the following were identified:

- Power ≈ 0 , but with wind > 2.5 m/s
- Power ≈ 0 , but with radiation > 200 W/m²

This subset was considered anomalous, as it contradicts the expected physical behavior of the system. As a corrective measure, advanced cleaning was implemented in the script `01_limpieza_avanzada.py`, where only these inconsistent records were removed and those with physical justification were retained. Therefore, the use of zero or very low power values within the model is fully justified, insofar as they represent real scenarios of the system: conditions of low available energy. This decision also allows the model to learn to correctly predict moments without generation, which is essential for real energy forecasting applications.

X. HYBRID ARCHITECTURE

After observing room for improvement in the LSTM model's predictions, a hybrid architecture was proposed that combined the temporal learning capacity of the LSTM with a decision tree model to correct the systematic errors of the former. This scheme consisted of two stages. In the first stage, an LSTM network was retrained on the normalized power series, with a 1-hour prediction horizon. In the second stage, the residual error between the LSTM prediction and the actual power was calculated, and an XGBoost model was trained to predict this error, using current meteorological variables, the LSTM prediction, and the time of day as input. The final prediction was obtained by adding the LSTM prediction to the correction estimated by the XGBoost model. Figure 11 shows the comparison between the actual values, the original LSTM prediction, and the adjusted prediction of the hybrid model. A visible improvement in peak tracking can be seen, as well as an overall reduction in dispersion. The metrics confirmed this improvement: the hybrid model achieved an RMSE of 62.12 W and a coefficient of determination R^2 of 0.7781, clearly surpassing the values obtained by the LSTM model without adjustment. This indicates that the hybrid approach significantly improves predictive power, especially in sections where the LSTM made systematic errors.

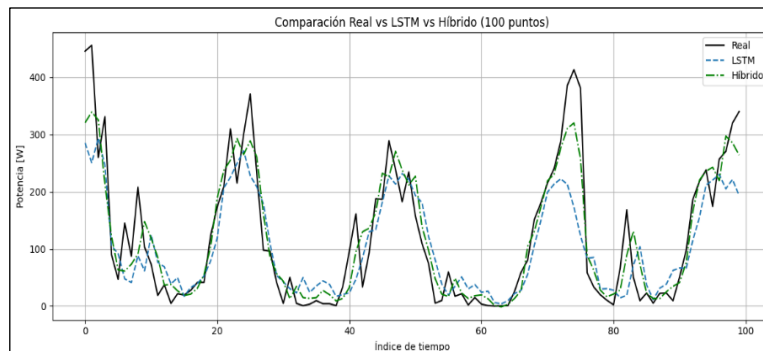


Figure 11: Comparison between actual power, LSTM prediction, and hybrid prediction (100 samples from the test set).

Source: Authors, (2026).

Given the good performance observed in the LSTM + XGBoost hybrid model at 1 hour, the same methodology was extended to prediction horizons of 3 and 6 hours. New LSTM models were trained for each time horizon, and in each case an XGBoost model was used again to predict residual errors and perform the final adjustment. The performance of the hybrid LSTM + XGBoost model across different prediction horizons is summarized in Table 9. The results show that the hybrid approach consistently improves predictive accuracy as the forecast horizon increases, maintaining a strong capacity to follow the underlying power generation trends and reducing the systematic deviations observed in the standalone LSTM predictions.

Table 9: Performance metrics of the hybrid LSTM + XGBoost model at different prediction horizons.

Horizon	RMSE (W)	R^2
1 hora	61.98	0.7791
3 horas	77.56	0.6541
6 horas	82.83	0.6052

Source: Authors, (2026).

These results show that the hybrid approach maintains good performance even over longer horizons, where traditional models tend to lose accuracy. In all cases, the hybrid model outperforms its unadjusted versions.

XI. TRANSFORMERS

With the aim of exploring more advanced architectures for multi-horizon power forecasting, it was decided to evaluate the performance of a model based on Transformers. These models, originally proposed for natural language processing, have recently shown great potential for time series forecasting tasks due to their ability to model long-term relationships, their efficient parallelization, and their flexibility in working with multivariate data. In this implementation, a simple Transformer was constructed using Keras' MultiHeadAttention layer. The model was trained on the same input sequences as the LSTM model, with normalized meteorological variables and a 1-hour prediction horizon. Unlike the LSTM, the Transformer does not rely on implicit sequential order, so internal position encoding was considered as part of the design. Figure 12 shows the comparison between the actual values and the prediction obtained by the Transformer model. In terms of performance, however, the results obtained were inferior to those achieved by the previously explored architectures:

- **RMSE:** 98.59 W
- **R²:** 0.4412

These values indicate a loss of accuracy compared to the basic LSTM model, and even more so compared to the hybrid model with XGBoost. Possible reasons include the lack of explicit temporal encoding, the need for a larger amount of data to generalize correctly, and the sensitivity of this type of architecture to hyperparameter optimization. Although it did not outperform the previous models, this test was useful in confirming that the LSTM + XGBoost hybrid architecture is better suited to the available dataset. It thus established itself as the most robust strategy within the context of this project.

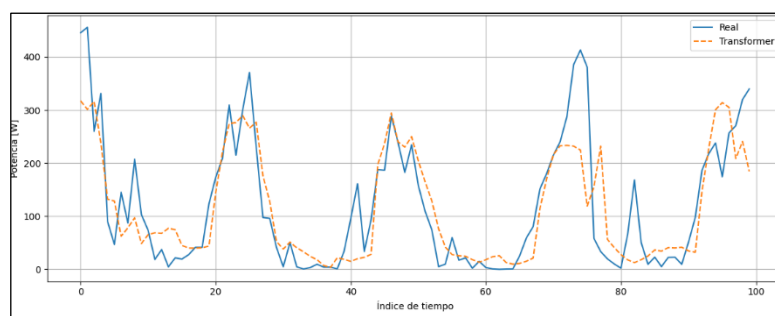


Figure 12: Power prediction using Transformer (100 samples from the test set).

Source: Authors, (2026).

In order to evaluate the direct applicability of the trained hybrid model (LSTM + XGBoost), a final script was developed for the actual prediction of wind power at 1-hour intervals. This stage represents the transition from the experimental approach to an operational scenario, where the model can receive new data and deliver automated predictions. A recent sample of the dataset (new_data.csv) was used as simulated input. The script 24_real_hybrid_prediction.py performs the following steps:

1. Loads and sorts the new data chronologically.
2. Applies the same normalization used in training, using the previously saved MinMaxScaler.
3. Generates multivariate sequences to feed the LSTM network.
4. Performs the initial prediction with the saved LSTM model.
5. Calculates the residual error correction using the XGBoost model.
6. Adds both outputs to obtain the final hybrid prediction.
7. Denormalizes the results to express them in watts.
8. Exports a file named hybrid_prediction_1h.csv with the predicted power for each instant.

This generated file is a useful tool for integration with external visualization or energy management platforms, enabling short-term monitoring and forecasting based on recent data. This implementation validates that the developed model not only achieves good results in controlled tests, but is also fully ready for use with new data in real-world environments.

XII. RESULTS

It is the logical presentation of the results that demonstrate the true contribution of the research and also justify the conclusion. The parts indicated above, Introduction, Material and Methods serve to explain how the results are obtained. Use tables and figures for your explanation. The results must be written in the past, in a brief and simple way. Avoid redundancy. Analyzing and interpreting the results is part of the discussions, their importance, achievements and limitations, highlighting the innovative aspects of the practical applications of the study and the conclusions derived from them, delimiting unresolved issues. If necessary, recommendations can be proposed. The dataset comprised 25,596 hourly records spanning January 2022–December 2024 and was described after data cleaning and feature engineering (air density and Betz-limit power were derived). A Pearson correlation analysis was performed to select inputs for the predictive models: wind speed exhibited a correlation of 0.92 with generated power, solar radiation 0.51 and air temperature 0.30; these three variables were therefore retained as primary predictors. Data were split into training and test sets using an 80/20 partition prior to model training. Four baseline regressors (linear regression, Random Forest, XGBoost and an MLP) were trained and evaluated using RMSE and R² on the test set. XGBoost and tree-based methods achieved the lowest errors in the instantaneous prediction task, while the linear model showed larger deviation at high-power values. For short-term forecasting (one hour ahead) an XGBoost regressor trained with current environmental inputs was evaluated separately and reported an RMSE of 80.11 W and R² = 0.6303.

Three alternative one-hour forecasting configurations were also compared: Model A (wind speed, radiation and temperature), Model B (wind speed only) and Model C (current variables plus 1-hour lag). Model C produced the lowest RMSE and highest R^2 among the three. Temporal lagging was further explored by adding t-1, t-3 and t-6 features for wind speed, radiation and power; models trained with these lagged inputs were evaluated at horizons of 1 h, 3 h and 6 h and showed consistent RMSE reductions and R^2 improvements relative to models without lags, particularly at the 3 h and 6 h horizons. A targeted advanced data cleaning step removed physically inconsistent records (instances of near-zero power despite sufficient wind or radiation); the retrained models after this cleaning exhibited only minor changes in short-term metrics and a small improvement in stability at the 3 h horizon. Sequence models were evaluated next. A basic LSTM (single hidden layer, 6-hour input sequences, MinMax normalization, 20 epochs) was trained and tested for 1 h, 3 h and 6 h horizons using normalized RMSE and R^2 ; an improved LSTM architecture (two LSTM layers 64→32, dropout 0.2, EarlyStopping) was also trained.

The improved LSTM produced marginal gains in normalized RMSE and R^2 over the basic LSTM. A hybrid two-stage architecture was then implemented: an LSTM generated the baseline one-hour prediction and an XGBoost model was trained to predict the LSTM residuals using meteorological inputs, the LSTM prediction and hour-of-day; the final hybrid prediction was obtained by adding the XGBoost residual correction to the LSTM output. The hybrid model yielded substantially lower RMSE and higher R^2 at 1 h and maintained improved performance at 3 h and 6 h compared with the LSTM alone. A Transformer-based model (single MultiHeadAttention block, same input sequences) was also trained for the 1 h task; the Transformer produced inferior performance relative to the LSTM and hybrid approaches on the available dataset. All experiments and metrics, reported in table n, below were computed on the test partition held out from model training. A final prediction script for operational use was developed and executed on a recent data sample; the script applied the same normalization and sequence-generation procedures used in training, produced hybrid one-hour predictions, denormalized the outputs and exported the results to `prediccion_hibrida_1h.csv` for subsequent integration with visualization or energy-management systems.

Table 10: Metrics of all models developed.

Model / Configuration	Prediction horizon	RMSE (W)	R^2
Linear regression (baseline)	instantaneous	47.17	0.8520
Random Forest (baseline)	instantaneous	0.42	1.0000
XGBoost (baseline)	instantaneous	1.36	0.9999
MLP (baseline)	instantaneous	6.48	0.9972
XGBoost (future prediction, Model A)	1 h	80.11	0.6303
Model B (wind only)	1 h	82.87	0.6045
Model C (with lag t-1)	1 h	78.62	0.6440
Without lags	1 h	79.01	0.6404
With lags (t-1,t-3,t-6)	1 h	77.13	0.6573
Without lags	3 h	106.43	0.3475
With lags	3 h	99.05	0.4349
Without lags	6 h	117.07	0.2107
With lags	6 h	108.68	0.3198
After advanced cleaning (1 h)	1 h	79.26	0.6387
After advanced cleaning (3 h)	3 h	105.75	0.3570
After advanced cleaning (6 h)	6 h	118.12	0.1977
LSTM (basic, normalized metrics)	1 h	0.1560 (norm.)	0.5309
LSTM (basic, normalized)	3 h	0.1748 (norm.)	0.4111
LSTM (basic, normalized)	6 h	0.1729 (norm.)	0.4233
LSTM (improved, normalized)	1 h	0.1544 (norm.)	0.5403
LSTM (improved, normalized)	3 h	0.1705 (norm.)	0.4395
LSTM (improved, normalized)	6 h	0.1716 (norm.)	0.4315
Hybrid LSTM + XGBoost (final)	1 h	61.98	0.7791
Hybrid LSTM + XGBoost (final)	3 h	77.56	0.6541
Hybrid LSTM + XGBoost (final)	6 h	82.83	0.6052
Transformer (single block)	1 h	98.59	0.4412

Source: Authors, (2026).

XIII. DISCUSSIONS

By [13] presented a short-term power prediction method for wind-farm clusters that fused multi-source spatio-temporal features through a graph-attention clustering step and a TimesNet predictor, reporting modest but consistent RMSE and MAE reductions across three Chinese regions. [14] combined variational mode decomposition with a bagging extreme learning machine and multi-objective aggregation, obtaining substantial out-of-sample RMSE reductions for very-short and short horizons. [15] proposed an indirect spatio-temporal strategy that classified data by wind direction and applied dedicated multivariate feature extraction, showing statistically superior performance versus several benchmarks. Compared with these studies, which either exploit explicit spatial clustering, advanced signal decomposition, or directional stratification for multi-site/cluster forecasts, our work focused on a single-site prediction using physically derived inputs (air density, Betz limit) and a hybrid LSTM+XGBoost residual correction; consequently, the gains reported in multi-site or decomposition-based approaches suggest promising avenues to extend our hybrid framework by incorporating spatial information or decomposition preprocessing to further reduce errors.

By [16] developed a WPCA feature weighting combined with a PSO-tuned GRU and reported 5–16% MAE and 10–16% RMSE improvements versus conventional models on a real wind farm. [17] evaluated direct and indirect point-forecasting pipelines that integrated several NWP sources and applied automated ML optimization, achieving error reductions in the 8–22% range for a Norwegian farm. [18] designed a physics-informed neural architecture that embedded analytical wake model nodes, reporting over 20% RMSE improvement by explicitly accounting for wake interactions. These contributions reinforce two messages relevant to our study: first, careful feature extraction and hyperparameter optimization (as in WPCA+PSO and BH-XGBoost approaches) can materially improve sequence models such as our LSTM; second, the explicit modeling of physical effects (wake/NWP fusion) could complement our data-driven hybrid residual correction and likely enhance performance in multi-turbine or wake-affected deployments beyond our single-site setting. In turn [19] provided a comprehensive meta-analysis of wind power forecast error metrics and introduced the Error Dispersion Factor (EDF) as a diagnostic that relates nRMSE and nMAE across horizons and farm sizes, offering benchmarking recommendations for forecast reporting. [20] proposed M2STAN, a multi-modal multi-task spatio-temporal attention network that combined graph attention with Bi-GRU for multi-location multi-step prediction and demonstrated superiority over several baselines. [21] coupled physical process models with machine learning using transfer learning strategies to improve wake-affected turbine predictions, showing that hybridization can inherit the strengths of both model classes. While our results emphasized a hybrid LSTM+XGBoost architecture and improvements from lagged inputs and data cleaning, the methodological recommendations and probabilistic/dispersion diagnostics advocated by Piotrowski et al. and the multi-location attention mechanisms developed by Wang and Zhou indicate concrete extensions: adopting normalized error reporting (nRMSE/nMAE/EDF) would facilitate cross-study benchmarking, and embedding spatio-temporal attention or physical coupling could extend our hybrid approach to multi-turbine contexts.

By [22] enhanced XGBoost via Bayesian hyperparameter optimization and demonstrated robust short-term forecasting performance, particularly during ramp events and low wind speeds. [23] proposed a multivariable hybrid system that combined outlier testing, feature extraction, an incremental kernel ELM predictor, and a bespoke multi-objective optimizer to maximize predictive capabilities over several Turkish datasets. Our findings—where XGBoost formed both a baseline and the residual-correction stage of the hybrid—align with Xiong’s emphasis on careful hyperparameter tuning and with Guo’s focus on outlier detection and multi-objective optimization; therefore, applying Bayesian optimization to our XGBoost residual model and formalizing a multi-objective training criterion (e.g., balancing RMSE and robustness metrics) could plausibly further reduce the hybrid RMSE reported in Table 7. [24] addressed the few-data problem for newly commissioned farms by generating synthetic samples through a WGANP and decomposing them with EEMD before training an evolutionary neural predictor; validation on a Spanish farm showed benefits from data augmentation. [25] proposed a serio-parallel CNNs-LSTM pretraining and partial transfer strategy to adapt models to target farms with scarce histories, demonstrating clear superiority over non-transfer approaches across a swarm of Chinese farms. Although our dataset was relatively large (25,596 hourly records) and did not require synthetic augmentation for model training, these works highlight practical strategies for deploying our hybrid solution in data-poor contexts: GAN-based augmentation or parameter transfer from pre-trained CNN-LSTM extractors could enable rapid adaptation of the LSTM+XGBoost pipeline to new sites where historical measurements are limited. [26] introduced MSTAN, a multi-source temporal attention network that produced probabilistic, multi-horizon density forecasts by selecting NWP driving variables via variable attention and temporal attention modules; their approach outperformed deterministic counterparts across three farms. [27] trained ANN-based wake models using large CFD datasets (ADM-R and RANS outputs) to reproduce velocity and turbulence fields and to improve farm power predictions, validating against LES and onsite measurements.

According to [28] combined genetic algorithms with LSTM (GLSTM) to optimize sequence window size and neuron counts, reporting average improvements of 6–30% across seven European farms and statistical significance versus standard LSTM. Relative to these studies, which emphasize probabilistic prediction, physics-informed wake modeling and structural hyperparameter search, our deterministic hybrid provided strong deterministic accuracy at one-hour and multi-hour horizons but did not quantify predictive uncertainty or explicitly incorporate CFD-driven wake fields or GA-based LSTM tuning; integrating uncertainty estimation and either physics-informed features or structured hyperparameter search represents clear directions to increase the operational robustness and interpretability of our forecasts. [29] developed a Temporal Convolutional Network (TCN) for ultra-short forecasts in a large utility-scale farm and used Taguchi-based orthogonal tuning to obtain hyperparameters that yielded consistent improvements over existing methods across wind speeds. [30] coupled high-resolution NWP with ANN frameworks over complex terrain and tested two hybrid strategies (direct turbine output forecasting and wind-then-power conversions), reporting normalized error improvements and demonstrating viable operational libraries for forecast mode. These results complement our conclusions: TCN architectures and systematic experimental hyperparameter optimization may outperform standard LSTM in sub-hourly contexts (suggesting an architectural alternative to explore), while the demonstrated benefits of NWP coupling in complex terrain indicate that fusion of NWP outputs with our hybrid residual correction could improve medium-range forecasts beyond the one-hour horizon reported in this work. In summary, the reviewed literature spans complementary directions—spatio-temporal attention and multi-site clustering, physics-informed wake modeling, advanced decomposition and augmentation, probabilistic forecasting, and rigorous hyperparameter/optimization strategies—and each study provides methodological prescriptions that map directly onto potential extensions of our LSTM+XGBoost hybrid.

XIV. CONCLUSIONS

From a cleaned and feature-engineered dataset of 25,596 hourly records, models were trained and evaluated showing that a hybrid two-stage architecture (LSTM + XGBoost residual correction) delivered the best deterministic performance at the 1-hour horizon (RMSE = 61.98 W, $R^2 = 0.7791$) and maintained superior accuracy at 3 h (RMSE = 77.56 W, $R^2 = 0.6541$) and 6 h (RMSE = 82.83 W, $R^2 = 0.6052$) compared with single-model alternatives (e.g., XGBoost baseline RMSE = 80.11 W, $R^2 = 0.6303$; Transformer RMSE = 98.59 W, $R^2 = 0.4412$); inclusion of lagged inputs (t-1, t-3, t-6) consistently improved multi-hour forecasts, while targeted advanced cleaning modestly increased stability at intermediate horizons. These results confirm that combining temporal sequence learning with a tree-based residual corrector effectively reduces systematic errors and better captures peaks than standalone sequence or attention models in the present dataset and context.

XV. AUTHOR'S CONTRIBUTION

Conceptualization: Jessica N. Castillo, Jose E. Naranjo, and Gabriel Pesántez.

Methodology: Jessica N. Castillo, Jose E. Naranjo, and Víctor Hugo Medina Matute.

Investigation: Jessica N. Castillo, Gabriel Pesántez, and Víctor Hugo Medina Matute.

Discussion of results: Jessica N. Castillo, Jose E. Naranjo, and Víctor Hugo Medina Matute.

Writing – Original Draft: Jessica N. Castillo and Jose E. Naranjo.

Writing – Review and Editing: Jessica N. Castillo, Jose E. Naranjo, and Gabriel Pesántez.

Resources: Gabriel Pesántez and Víctor Hugo Medina Matute.

Supervision: Jessica N. Castillo.

Approval of the final text: Jessica N. Castillo, Gabriel Pesántez, Víctor Hugo Medina Matute and Jose E. Naranjo.

XVI. ACKNOWLEDGMENTS

This work is part of the research project “*Modelado matemático en ingeniería: enfoques generales y aplicaciones.*” The authors would like to express their sincere gratitude to the Universidad Técnica de Cotopaxi for its institutional support, academic environment, and commitment to fostering scientific research and innovation.

XVII. REFERENCES

- [1] A. G. AbdElkader and M. M. Saafan, “Optimizing wind power forecasting with RNN-LSTM models through grid search cross-validation,” *Sustainable Computing: Informatics and Systems*, vol. 45, 2025, doi: 10.1016/j.suscom.2024.101054.
- [2] A. Wan et al., “A novel hybrid BWO-BiLSTM-ATT framework for accurate offshore wind power prediction,” *Ocean Engineering*, vol. 312, 2024, doi: 10.1016/j.oceaneng.2024.119227.
- [3] X. Wang, Y. Hao, and W. Yang, “Novel wind power ensemble forecasting system based on mixed-frequency modeling and interpretable base model selection strategy,” *Energy*, vol. 297, 2024, doi: 10.1016/j.energy.2024.131142.
- [4] J. W. Jayasinghe, P. Ekanayake, O. Panahatipola, C. I. Madhushani, and U. S. Rathnayake, “Forecasting the power generation at renewable power plants in Sri Lanka using regression trees,” *Results in Engineering*, vol. 22, 2024, doi: 10.1016/j.rineng.2024.102111.
- [5] S. Yang, X. Deng, and G. Guangxing, “Machine-learning-based wind farm optimization through layout design and yaw control,” *Renew. Energy*, vol. 224, 2024, doi: 10.1016/j.renene.2024.120161.
- [6] M. E. Arrieta-Prieto and K. R. Schell, “Spatially transferable machine learning wind power prediction models: v-logit random forests,” *Renew. Energy*, vol. 223, 2024, doi: 10.1016/j.renene.2024.120066.
- [7] G. Guangxing, X. Deng, Z. Ti, S. Yang, S. Huang, and Y. Wang, “A data-driven layout optimization framework of large-scale wind farms based on machine learning,” *Renew. Energy*, vol. 218, 2023, doi: 10.1016/j.renene.2023.119240.
- [8] M. E. Nakhchi, S. Win Naung, and M. T. Rahmati, “Wake and power prediction of horizontal-axis wind farm under yaw-controlled conditions with machine learning,” *Energy Convers. Manag.*, vol. 296, 2023, doi: 10.1016/j.enconman.2023.117708.
- [9] R. A. Sobolewski, M. A. Tchakorom, and R. Couturier, “Gradient boosting-based approach for short- and medium-term wind turbine output power prediction,” *Renew. Energy*, vol. 203, pp. 142–160, 2023, doi: 10.1016/j.renene.2022.12.040.
- [10] A. Meng et al., “A novel multi-gradient evolutionary deep learning approach for few-shot wind power prediction using time-series GAN,” *Energy*, vol. 283, 2023, doi: 10.1016/j.energy.2023.129139.
- [11] Y. Wu, J. Zhang, Z. Yuan, and F. Deng, “Review on Identification and Cleaning of Abnormal Wind Power Data for Wind Farms; 风电场风功率异常数据识别与清洗研究综述,” *Dianwang Jishu/Power System Technology*, vol. 47, no. 6, pp. 2367–2379, 2023, doi: 10.13335/j.1000-3673.pst.2022.1681.
- [12] J. Shi, B. Wang, K. Luo, Y. Wu, M. Zhou, and J. Watada, “Ultra-short-term wind power interval prediction based on multi-task learning and generative critic networks,” *Energy*, vol. 272, 2023, doi: 10.1016/j.energy.2023.127116.
- [13] M. Yang, C. Han, W. Zhang, and B. Wang, “A short-term power prediction method for wind farm cluster based on the fusion of multi-source spatiotemporal feature information,” *Energy*, vol. 294, 2024, doi: 10.1016/j.energy.2024.130770.
- [14] M. H. D. M. Ribeiro et al., “Variational mode decomposition and bagging extreme learning machine with multi-objective optimization for wind power forecasting,” *Applied Intelligence*, vol. 54, no. 4, pp. 3119–3134, 2024, doi: 10.1007/s10489-024-05331-2.
- [15] S. Sun, Z. Du, K. Jin, H. Li, and S. Wang, “Spatiotemporal wind power forecasting approach based on multi-factor extraction method and an indirect strategy,” *Appl. Energy*, vol. 350, 2023, doi: 10.1016/j.apenergy.2023.121749.
- [16] Y. Xiao, C. Zou, H. Chi, and R. Fang, “Boosted GRU model for short-term forecasting of wind power with feature-weighted principal component analysis,” *Energy*, vol. 267, 2023, doi: 10.1016/j.energy.2022.126503.
- [17] G. Yakoub, S. Mathew, and J. B. Leal, “Intelligent estimation of wind farm performance with direct and indirect ‘point’ forecasting approaches integrating several NWP models,” *Energy*, vol. 263, 2023, doi: 10.1016/j.energy.2022.125893.
- [18] N. Guo et al., “A physics-inspired neural network model for short-term wind power prediction considering wake effects,” *Energy*, vol. 261, 2022, doi: 10.1016/j.energy.2022.125208.
- [19] P. J. Piotrowski, I. D. S. Rutyna, D. Baczyński, and M. Kopyt, “Evaluation Metrics for Wind Power Forecasts: A Comprehensive Review and Statistical Analysis of Errors,” *Energies (Basel)*, vol. 15, no. 24, 2022, doi: 10.3390/en15249657.

- [20] L. Wang and Y. He, "M2STAN: Multi-modal multi-task spatiotemporal attention network for multi-location ultra-short-term wind power multi-step predictions," *Appl. Energy*, vol. 324, 2022, doi: 10.1016/j.apenergy.2022.119672.
- [21] H. Zhou, Y. Qiu, Y. Feng, and J. Liu, "Power prediction of wind turbine in the wake using hybrid physical process and machine learning models," *Renew. Energy*, vol. 198, pp. 568–586, 2022, doi: 10.1016/j.renene.2022.08.004.
- [22] X. Xiong, X. Guo, P. Zeng, R. Zou, and X. Wang, "A Short-Term Wind Power Forecast Method via XGBoost Hyper-Parameters Optimization," *Front. Energy Res.*, vol. 10, 2022, doi: 10.3389/fenrg.2022.905155.
- [23] H. Guo, J. Wang, Z. Li, and Y. Jin, "A multivariable hybrid prediction system of wind power based on outlier test and innovative multi-objective optimization," *Energy*, vol. 239, 2022, doi: 10.1016/j.energy.2021.122333.
- [24] H. Yin, Z. Ou, Z. Zhu, X. Xu, J. Fan, and A. Meng, "A novel asexual-reproduction evolutionary neural network for wind power prediction based on generative adversarial networks," *Energy Convers. Manag.*, vol. 247, 2021, doi: 10.1016/j.enconman.2021.114714.
- [25] H. Yin, Z. Ou, J. Fu, Y. Cai, S. Chen, and A. Meng, "A novel transfer learning approach for wind power prediction based on a serio-parallel deep learning architecture," *Energy*, vol. 234, 2021, doi: 10.1016/j.energy.2021.121271.
- [26] H. Zhang, J. Yan, Y. Liu, Y. Gao, S. Han, and L. Li, "Multi-source and temporal attention network for probabilistic wind power prediction," *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2205–2218, 2021, doi: 10.1109/TSTE.2021.3086851.
- [27] Z. Ti, X. Deng, and M. Zhang, "Artificial Neural Networks based wake model for power prediction of wind farm," *Renew. Energy*, vol. 172, pp. 618–631, 2021, doi: 10.1016/j.renene.2021.03.030.
- [28] F. Shahid, A. Zameer, and M. Muneeb, "A novel genetic LSTM model for wind power forecast," *Energy*, vol. 223, 2021, doi: 10.1016/j.energy.2021.120069.
- [29] R. Meka, A. Alaeddini, and K. Bhaganagar, "A robust deep learning framework for short-term wind power forecast of a full-scale wind farm using atmospheric variables," *Energy*, vol. 221, 2021, doi: 10.1016/j.energy.2021.119759.
- [30] L. Donadio, J. Fang, and F. Porté-Agel, "Numerical weather prediction and artificial neural network coupling for wind energy forecast," *Energies (Basel)*, vol. 14, no. 2, 2021, doi: 10.3390/en14020338.