Modeling Wind Energy Production Forecasting using Machine Learning: An In-depth Analysis of Wind Farms in Morocco

Mohamed Bousla
Innovating Technologies Team, National School of Applied Sciences, Tetouan, Abdelmalek Essaadi University, Morocco
mohamed.bousla@etu.uae.ac.ma (corresponding author)

Mohamed Belfkir
United Arab Emirates University, Al Ain, Abu Dhabi, United Arab Emirates
m_belfkir@uaeu.ac.ae (corresponding author)

Omar Elharrouss
United Arab Emirates University, Al Ain, Abu Dhabi, United Arab Emirates
o_elharrouss@uaeu.ac.ae

Ahmed Sadki
ENS, Abdelmalek Essaadi University, Tetouan, 93020, Morocco
ahmed.sadki@etu.uae.ac.ma

Ali Haddi
Innovating Technologies Team, National School of Applied Sciences, Tetouan, Abdelmalek Essaadi University, Morocco
ahaddi1@uae.ac.ma

Youness El Mourabit
Innovating Technologies Team, National School of Applied Sciences, Tetouan, Abdelmalek Essaadi University, Morocco
youness.elmourabit@uae.ac.ma

Badre Bossoufi
LIMAS Laboratory, Faculty of Sciences Dhar El Mahraz, Sidi Mohammed Ben Abdellah University, Fez 30050, Morocco
badre.bossoufi@usmba.ac.ma

Abderrahman Mouradi
ENS, Abdelmalek Essaadi University, Tetouan, 93020, Morocco
abderrahman.mouradi@uae.ac.ma

Abderrahman Elkharrim
ENS, Abdelmalek Essaadi University, Tetouan, 93020, Morocco
abderrahman.elkharrim@uae.ac.ma

Received: 20 January 2025 | Revised: 20 March 2025 | Accepted: 2 April 2025

ABSTRACT

Accurate forecasting of wind energy production is essential for the stable integration of renewable energy sources into power grids, especially given the inherent variability of wind conditions. This study evaluates the effectiveness of Transformer-based models for improving wind energy forecasting accuracy, compared to traditional methods such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRUs). Unlike the conventional sequential models, the Transformer models leverage an advanced attention mechanism, which processes all time steps simultaneously rather than sequentially, thereby efficiently capturing complex, long-term dependencies within the data. To conduct this analysis, we utilized a dataset collected from an operational wind farm located in Tetouan, northern Morocco, covering the period from 2019 to 2020. The experimental results show that the Transformer model consistently outperformed the traditional methods, achieving Mean Squared Error (MSE) of 0.275, 0.234, and 0.221, and Mean Absolute Error (MAE) of 0.305, 0.296, and 0.284 for daily, weekly, and monthly forecasting horizons, respectively. Specifically, the Transformer model achieved approximately a 10% reduction in Mean Absolute Percentage Error (MAPE) compared to the LSTM model. These findings demonstrate the substantial advantage of Transformer-based approaches in wind energy forecasting and underline their potential to significantly enhance the reliability of renewable energy integration into modern power grids.

Keywords-ML; transformer-based models; RNN; LSTM; GRU; wind power forecast; wind energy

I. INTRODUCTION

Climate change is driving a major transformation in the global energy sector, with an urgent need to accelerate the adoption of renewable energy. In this context, wind energy plays a central role due to its widespread availability, relatively low maintenance costs, and absence of pollutant emissions [1, 2]. This transition aims not only to protect the environment but also to meet the growing demand for clean and sustainable energy. Among the renewable energy sources experiencing the strongest growth in the 21st century, wind energy stands out particularly for its integration potential into modern energy systems [3, 4]. However, this integration faces a significant challenge: the natural variability of wind, which considerably complicates electricity production forecasting and optimal management of electrical grids [5]. Indeed, the intermittency of wind resources necessitates reliable forecasting to ensure grid stability [6]. For instance, irregular wind production can create imbalances between electricity supply and demand, leading to disruptions in the network and even power outages. Forecasting complexity is further exacerbated by the influence of multiple meteorological parameters such as atmospheric pressure, temperature, and humidity [7, 8]. Accurate forecasts are therefore essential to effectively balance production and consumption and strengthen the stability of energy systems [9].

Currently, wind energy forecasting approaches can be primarily classified into three categories: physical models, statistical models, and Artificial Intelligence (AI)-based models [10-12]. Physical models, based on atmospheric laws, are suitable for long-term forecasts but involve high computational costs and require continuous data collection, limiting their real-time application [13, 14]. Statistical models, leveraging historical correlations between wind speed and power generation, are effective in short-term forecasting but remain vulnerable to unpredictable meteorological fluctuations [15-17].

Recent advances in AI have significantly transformed forecasting methods, providing robust alternatives to traditional approaches [18]. For example, models such as XGBoost [19], simultaneously optimize prediction speed and accuracy. The Extreme Learning Machine (ELM), valued for its simplicity, is particularly effective in short-term forecasting [20, 21]. Artificial Neural Networks (ANNs), including Deep Neural Networks (DNNs), also offer considerable flexibility in modeling nonlinear relationships in wind data [22, 23]. Similarly, Recurrent Neural Networks (RNNs), particularly modern architectures like Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM), effectively capture complex temporal dependencies for short-term forecasts [24-29].

Despite these advancements, precisely managing temporal variability in forecasting remains a major challenge [30, 31], which is the research gap targeted by this study: exploring the application of advanced Transformer-based models, which remain relatively unexplored in this specific context, to significantly improve wind energy forecasting accuracy. The originality of this study specifically lies in employing Transformer models, renowned for their exceptional capability in managing complex and lengthy sequences, yet still underutilized in wind energy forecasting. These models are applied to an actual operational wind farm, ensuring practical relevance and transferability of the results obtained.

II. WIND DATASET COLLECTION AND PROCESSING

The wind data used in this research were obtained from the Tetouan wind farm, located in northern Morocco near the city of Tetouan. Situated on the scenic hill of Tetouan, this wind farm includes 40 turbines strategically positioned to optimize the region's abundant wind resources. The precise geographical coordinates of the farm are 35°35'54.4"N and 5°34'38.2"W. The Tetouan wind farm is distinguished by its technical attributes as well as its beneficial environmental and economic impacts due to the favorable wind conditions in this coastal area. The strategic arrangement of the turbines along the ridge optimizes energy output while reducing environmental impact, demonstrating the farm's commitment to sustainable energy practices.

Between 2019 and 2020, comprehensive measurements of wind speed (m/s), wind direction (degrees), and power output

(W) were recorded at 10-min intervals. This data collection was made possible through cup anemometers and mechanical wind vanes installed at a height of 80 m, allowing for the capture of wind variations at an optimal altitude for energy production. These instruments were specifically chosen for their robustness and precision, ensuring reliable measurements despite the region's variable weather conditions. The dataset includes 105,120 data points for each wind parameter, providing an extensive view of wind conditions over a prolonged period. The data collected from the 40 turbines were meticulously assembled for model training and evaluation. This comprehensive information is essential for evaluating wind energy generation in the area during the designated period, contributing to a better understanding of wind variations and their impact on energy production.

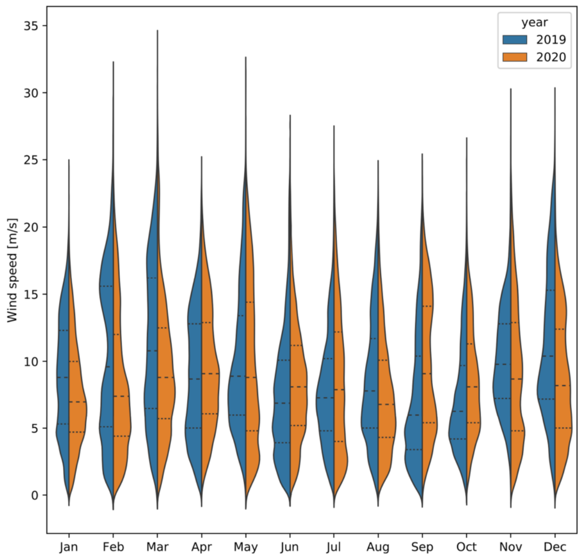

Fig. 1. Monthly wind speed of the 40 Tetouan wind turbines.

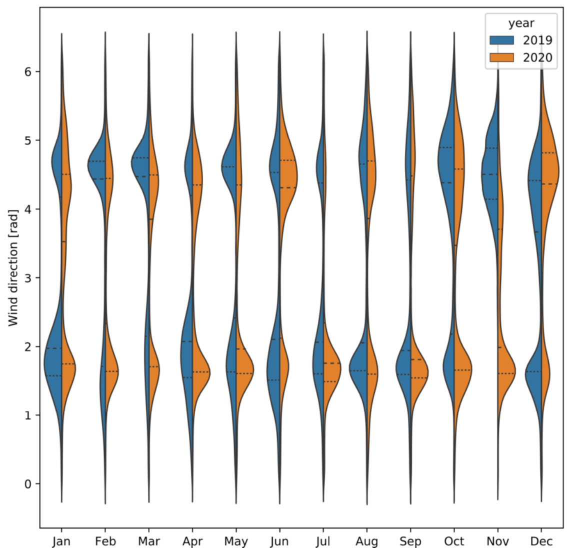

Fig. 2. Monthly wind direction of the 40 Tetouan wind turbines.

Figures 1 and 2 illustrate the monthly variation in wind speed and direction. These parameters exhibit significant fluctuations from month to month, reflecting changing climate conditions. The variations in wind speed and direction are influenced by various atmospheric factors, including temperature differentials, pressure systems, and local geographical features. To ensure the accuracy of the results, the data underwent a cleaning process to remove outliers.

A. Data Normalization and Cleaning

To adapt to the dynamic nature of wind power production, several data preprocessing steps were implemented to optimize model training, accelerate convergence speed, and enhance prediction accuracy. These optimizations aim to enable the models to respond effectively to variations in production patterns while minimizing the impact of abnormal data on forecast precision.

1) Data Normalization

Given the high variability of parameters such as energy production, wind speed, and direction, data normalization was essential to stabilize the models' learning process and ensure more consistent forecasts. A normal quantile transformation was applied, adjusting each feature to follow a normal distribution. This process helped mitigate the impact of extreme values and disperse frequently observed values, thereby facilitating the identification of meaningful patterns within the data. Normalization was implemented using the widely recognized Scikit-learn library, known for its robustness and extensive application in Machine Learning (ML) [32]. By harmonizing the distribution of input data, this transformation not only accelerated model convergence but also improved overall forecast accuracy by reducing potential biases caused by scale differences among features.

2) Outlier Detection and Handling

Managing outliers was a crucial step in ensuring model reliability and eliminating noise that could hinder performance. For wind direction, values exceeding 360° were removed to maintain physical consistency, as this range is the natural limit for direction measurements. For energy production, only positive values were retained, as negative values typically indicate measurement errors or turbine malfunctions. A filter was applied to exclude these outliers, ensuring that only reliable and representative data points were used in model training and evaluation. Through this cleaning process, an overall data efficiency of 78% was achieved, guaranteeing the quality of the information utilized by the models.

3) Impact of Preprocessing Steps on Model Performance

The preprocessing steps significantly contributed to improving model performance. Normalization reduced the risk of learning biases by balancing the distribution of values, enabling the models to better focus on underlying relationships in the data without being influenced by extreme variations. Additionally, managing outliers enhanced forecast robustness by preventing measurement errors and anomalies from disrupting learning. This data cleaning ensured that models were more precise and reliable in forecasting wind energy production, increasing their ability to handle natural variations in wind conditions. In summary, the data normalization and cleaning process optimized the quality of input data and ensured more accurate results. These preprocessing steps play a fundamental role in enhancing the performance of ML models by enabling better data utilization while minimizing the influence of anomalies.

B. Training and Testing Dataset

To avoid any potential bias during the training process, the pre-processed data were divided into two distinct subsets: 70% of the data was utilized for model training and the remaining 30% was used for model evaluation and final prediction generation. This systematic approach ensures a thorough analysis of the models and their respective capabilities, allowing for an objective assessment of their effectiveness.

III. MACHINE LEARNING-BASED ENERGY PREDICTION MODELS

In wind energy forecasting, it is crucial to examine data pertaining to wind velocity and orientation. We evaluated several ML models, including RNN, LSTM, GRU, and Transformer-based models. We used a novel approach using Transformers to enhance predictive accuracy. All models were trained and evaluated on consistent datasets, guaranteeing identical data partitions and training parameters.

A. Deep Neural Network Architecture

The DNN used in our research is structured with an input layer that integrates essential data such as wind speed and direction. Subsequently, there are three hidden layers with 128, 32, and 32 neurons, respectively. The output layer has a solitary neuron tasked with forecasting active wind energy (y). To maintain consistent mean and variance of inputs across layers and to address normalization issues, all hidden layers use the Scaled Exponential Linear Unit (Selu) activation function, whilst the output layer utilizes a linear activation function. The network's weights were randomly initialized from a truncated normal distribution centered at zero, with a width of $1/N$, where N is the number of input units. This DNN architecture is adept at identifying intricate patterns in the data, making it particularly useful for forecasting active wind power and other regression-related tasks.

B. Recurrent Neural Networks

As mentioned above, the data for this study were collected every 10 minutes. Variations in wind speed, direction, and operating circumstances at each moment influence the measurements for the subsequent time period. This establishes temporal relationships and patterns in the data that conventional ML models, such as SVM, find challenging to capture effectively. To address these issues, we used RNNs, which are especially designed to process sequential input via the integration of feedback connections. The fundamental equations that regulate RNNs are:

- Hidden State :

$$h_t = f\left(w_h h_{t-1} + w_x x_t + b\right) \tag{1}$$

where h_t represents the hidden state at time t, f denotes an activation function (like the hyperbolic tangent or ReLU), w_h is the weight matrix for the previous hidden state, w_x is the weight matrix for the input, and b is the bias term.

- Output :

$$y_t = w_y h_t + b_y \tag{2}$$

where y_t is the output at time t and w_y and b_y are the weight matrix and bias for the output, respectively.

C. Long Short-Term Memory

LSTMs represent an enhanced iteration of RNNs, designed to mitigate the vanishing gradient issue often seen in conventional RNNs. The system comprises memory cells and many gating mechanisms (input gate, forget gate, and output gate) that facilitate the model's ability to ascertain which information to retain or eliminate.

- Input Gate:

$$i_t = \sigma\left(w_i x_t + U_i h_{t-1} + b_i\right) \tag{3}$$

- Forget Gate:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{4}$$

- Output Gate:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{5}$$

- Cell State :

$$Ct = f_t . C_{t-1} + i_t Tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{6}$$

- Hidden State :

$$h_t = o_t . Tanh(C_t) \tag{7}$$

where σ is the sigmoid function, and W and U are the respective weight matrices.

The use of gates allows LSTMs to retain information over long periods, which is essential for time-dependent data where past events can impact future predictions. Thus, LSTMs are particularly suited for predicting wind power.

D. Gated Recurrent Units

GRUs are an optimized variant of LSTMs that consolidate certain elements of the input and forget gates into a single update gate. This reduction lowers the model's complexity while yet proficiently capturing long-term interdependence.

- Update Gate:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{8}$$

- Reset Gate:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{9}$$

- Cell State:

$$\widetilde{h_t} = Tanh(W_h x_t + U_h(r_t h_{t-1}) + b_h) \quad (10)$$

- Hidden State:

$$h_t = (1 - z_t).h_{t-1} + z_t.\widetilde{h_t} \quad (11)$$

where z_t and r_t are the update and reset gates, respectively.

GRUs are often faster to train than LSTMs due to their simpler architecture, while still achieving comparable performance in time series forecasting tasks, making them useful for wind power forecasting.

E. Transformer-Based Models

With the introduction of attention modules, Transformer-based models have demonstrated remarkable performance in Natural Language Processing (NLP) [33]. Researchers have begun to integrate self-attention into various time series architectures to further enhance forecasting performance. This has led to the emergence of Transformers, which have proven effective in time series forecasting, including energy forecasting [34]. In our study, we adapted the Informer method to assess the impact of Transformers on wind data forecasting. The fundamental equation of the attention mechanism in Transformers is given by:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (12)$$

where Q, K, and V represent the query, key, and value matrices, respectively, and d_k is the dimension of the keys. This approach allows for capturing long-term relationships in time series data more efficiently than by the previous methods. By combining these different approaches, our goal is to optimize the accuracy of predictions concerning the energy produced by wind, while accounting for the dynamic variations in wind and environmental conditions.

In summary, each model adopts a specific approach to processing temporal data, with distinct strengths and limitations. RNNs and their variants (LSTM and GRU) excel at capturing short-term dependencies due to their sequential structure but often face challenges with long-term dependencies, especially due to the vanishing gradient problem. LSTM and GRU models partially address this limitation with gating mechanisms that allow information to be retained over longer sequences. In contrast, Transformers stand out by incorporating an attention mechanism capable of capturing long-term dependencies simultaneously and in parallel, making their architecture particularly suitable for predictions in complex time series.

IV. EVALUATION METRICS

To evaluate the accuracy and robustness of the considered forecasting models in the context of wind energy prediction, three performance indicators were used: Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics provide a comprehensive overview of each model's performance, highlighting their specific strengths and limitations. To ensure consistent and comparable results, the calculations were conducted in a standardized hardware environment, using a DELL XPS 15 9520 equipped with a 12th-generation Intel Core i7-12700H processor, featuring 14 cores with a maximum frequency of 4.7 GHz, 16 GB of DDR5 RAM running at 4800 MHz, an NVIDIA GeForce RTX 3050 Ti graphics card with 4 GB of GDDR6 memory, and a 1 TB M.2 PCIe NVMe SSD. The reported training time excludes data loading time, as data were directly read from the SSD to optimize processing efficiency. This standardized hardware configuration ensures a rigorous evaluation of the models in terms of accuracy and performance.

Each of these metrics plays an essential role in evaluating the performance of forecasting models, and their combined use provides a comprehensive and nuanced view of the accuracy of forecasts in the context of wind data analysis.

A. Mean Absolute Error (MAE)

The MAE is a particularly useful Key Performance Indicator (KPI) for assessing the accuracy of forecasts in the field of data modeling. This metric is highly valued for its simplicity and ease of interpretation, as it directly expresses the deviation between forecasts and actual observations without considering the sign of the errors. The MAE is calculated by:

$$MAE = \frac{1}{n}\sum^{n}\left|y_i - f\left(x_i\right)\right| \quad (13)$$

where y_i represents the actual energy production value, while $f(x_i)$ denotes the predicted value by the model for the input x_i. A low MAE indicates that the forecasts are very close to the actual values, which is desirable for any forecasting application. Furthermore, the MAE is particularly advantageous in contexts where errors are evenly distributed, as it does not favor larger errors over smaller ones, unlike other metrics such as MSE.

B. Mean Squared Error (MSE)

The MSE is another crucial metric used to evaluate forecast accuracy. Unlike the MAE, which measures the average of absolute errors, the MSE focuses on the average of the squared errors between the actual and predicted values. This approach amplifies the impact of large errors, meaning that the MSE is sensitive to significant fluctuations in the data. In fact, larger errors are assigned greater weight, which can be beneficial in scenarios where minimizing major errors is essential [35]. The MSE is defined by:

$$MSE = \frac{1}{n}\sum^{n}\left(y_i - f\left(x_i\right)\right)^2 \quad (14)$$

A low MSE suggests that the model has small errors, which is a positive indicator of its performance. Due to its sensitivity to large errors, the MSE is often used in contexts where precision is paramount and where significant deviations can have serious consequences, such as in forecasting wind energy production.

C. Mean Absolute Percentage Error (MAPE)

MAPE is a widely recognized and used metric for assessing forecast accuracy. The MAPE calculates the average of absolute errors expressed as a percentage, thus normalizing the errors relative to the actual values. This makes this metric particularly useful when comparing forecasts across different datasets or time periods, as it provides a relative measure of accuracy. The MAPE formula is:

$$MAPE = \frac{1}{n}\sum^{n}\left|\frac{y_i - f\left(x_i\right)}{y_i}\right| \times 100 \quad (15)$$

where y_i represents the measured actual power energy, while $f(x_i)$ corresponds to the power energy predicted by the model. In terms of interpretation, a model displaying a MAPE of less than 10% is generally considered highly accurate. A MAPE between 10% and 20% indicates a good forecasting model, while a MAPE of 20% to 30% is regarded as reasonable. A MAPE greater than 50% suggests that the model is inaccurate and requires improvements.

V. RESULTS AND DISCUSSION

As shown in Table I, the proposed model consistently outperformed the sequential approaches of RNN, LSTM, and GRU, with particularly significant improvements observed in weekly and monthly forecasts. Specifically, the Transformer model achieved a MAPE approximately 10% lower than that of the LSTM, demonstrating its enhanced ability to adapt to wind variability and to extract relevant temporal patterns over extended horizons. Visual comparisons of model performance are presented in Figures 3-5.

TABLE I. EVALUATION OF FORECASTING ACCURACY AND COMPUTATIONAL EFFICIENCY AT DAILY, WEEKLY, AND MONTHLY SCALES

Method	MAE	MSE	MAPE	Time/epoch (s)
Daily				
RNN	0.462	0.418	3.659	325 ± 10
LSTM	0.879	0.997	1.021	590 ± 10
GRU	0.354	0.308	2.952	650 ± 20
Informer	0.305	0.275	2.877	1510 ± 5
Weekly				
RNN	0.421	0.381	4.167	321.5 ± 20
LSTM	0.7762	0.950	1.105	550 ± 10
GRU	0.303	0.247	2.866	580 ± 10
Informer	0.296	0.234	2.418	910 ± 30
Monthly				
RNN	0.404	0.345	4.0045	300 ± 10
LSTM	0.722	0.957	1.010	530 ± 7
GRU	0.296	0.233	2.854	500 ± 100
Informer	0.284	0.221	2.17	900 ± 20

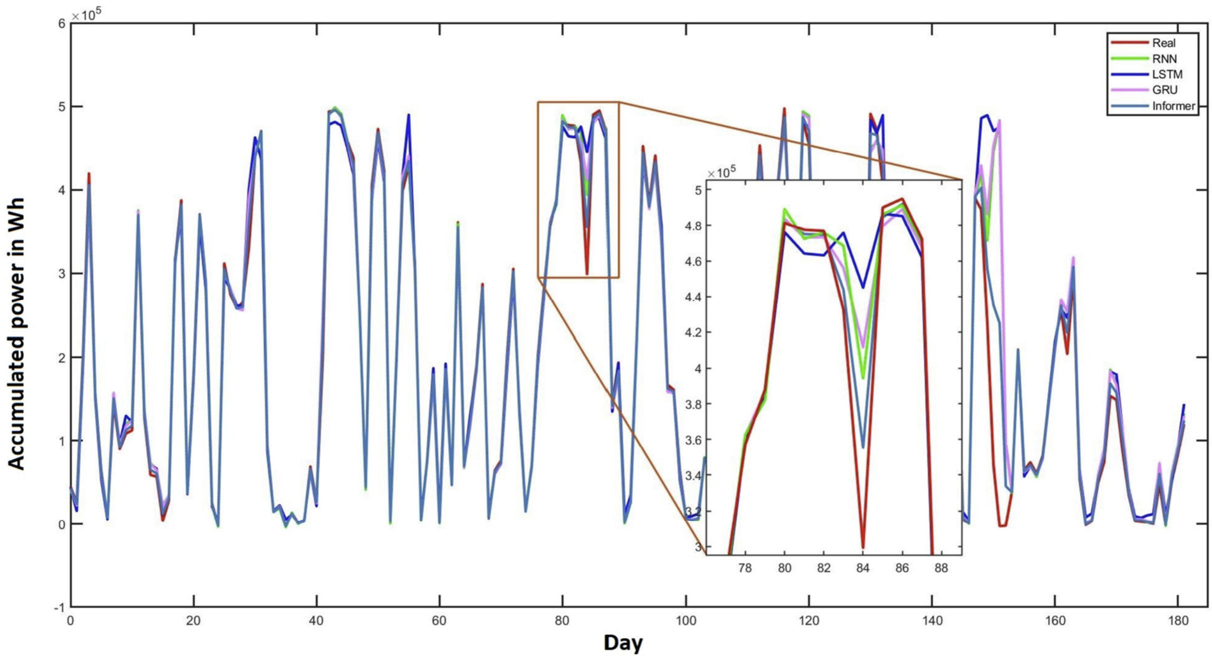

Fig. 3. Comparative analysis of the daily accumulated energy predictions.

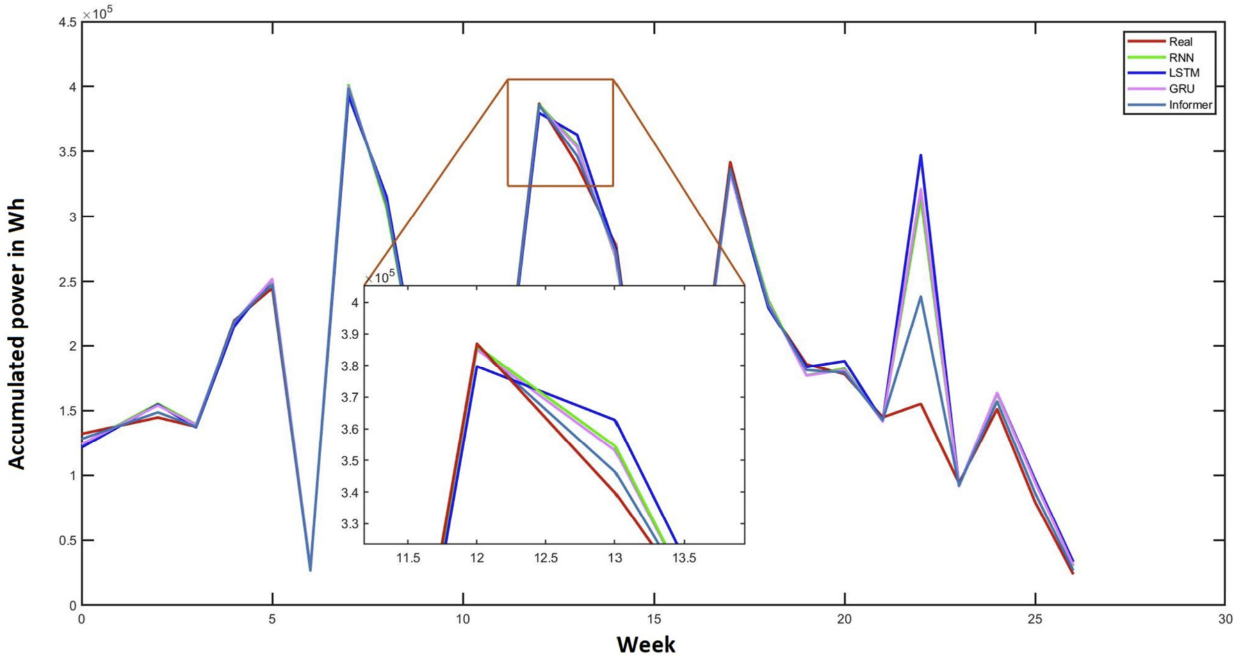

Fig. 4. Comparative analysis of the weekly accumulated energy predictions.

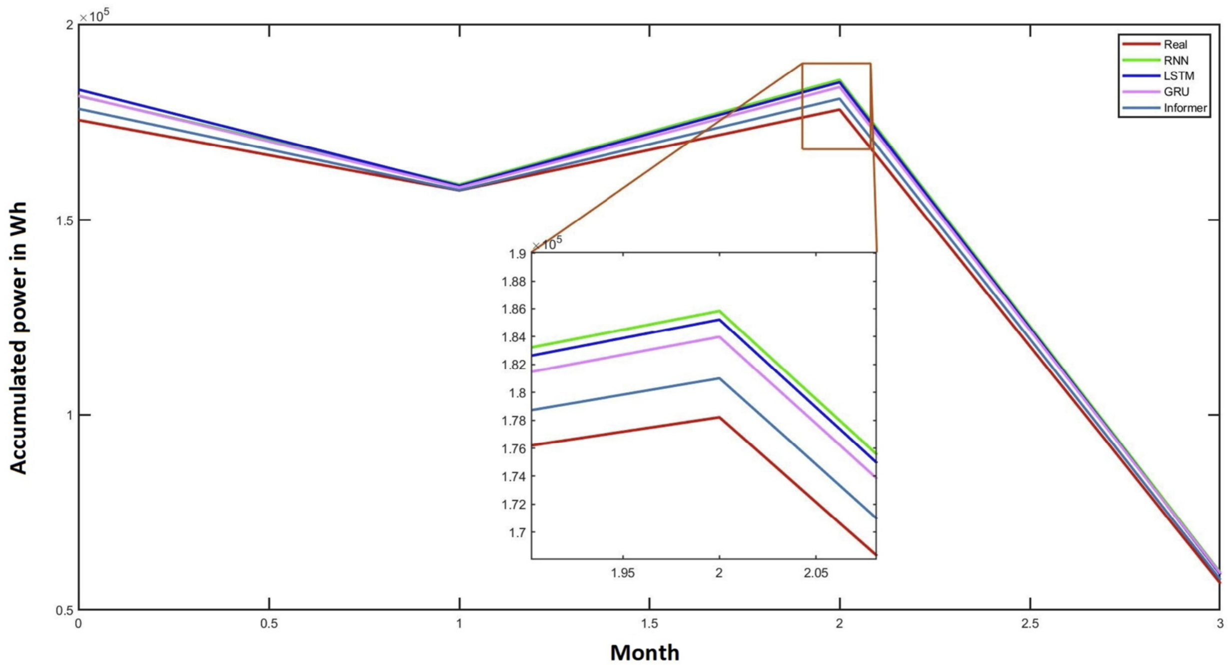

Fig. 5. Comparative analysis of the monthly accumulated energy predictions.

The analysis of daily forecasts (Figure 3) reveals that while all models yield satisfactory results, the proposed model stands out by significantly reducing absolute error. The sequential models tend to over-smooth short-term fluctuations, limiting their responsiveness to sudden changes in wind patterns. In contrast, the transformer model dynamically adjusts its temporal weightings through its attention mechanism, allowing it to respond more effectively to rapid meteorological changes.

The weekly forecasts (Figure 4) confirm this trend and further highlight the transformer model's robustness under operational constraints. A specific incident around week 20, where the grid operator temporarily curtailed wind production, caused notable prediction errors for the sequential models. The Informer, however, successfully mitigated this disturbance thanks to its global learning capacity, enabling it to anticipate and adapt to unexpected fluctuations. This adaptability makes it particularly well-suited for reliable renewable energy integration into the power grid.

Regarding the monthly forecasts (Figure 5), the Informer demonstrates a greater ability to capture seasonal trends, outperforming sequential models that struggle to represent long-term dependencies. The notably lower MSE and MAE values achieved by the Informer confirm its capacity to track meteorological cycles, which is a key advantage for long-term energy planning and effective management of wind power intermittency.

Overall, these findings emphasize the Informer's ability to weigh and prioritize relevant temporal features, allowing it to deliver accurate and stable forecasts across all time scales, even under dynamic and unstable environmental conditions. Its resilience to both operational and weather-related disruptions positions it as a strategic tool for energy system optimization and long-term planning.

VI. CONCLUSION

In this paper, we introduced a novel and comprehensive dataset specifically tailored for wind energy production forecasting, encompassing essential parameters: date, wind speed, wind direction, and energy output. Utilizing this extensive dataset, we conducted a rigorous comparative analysis of advanced deep learning models, namely RNN, LSTM, GRU, and Transformer-based, to assess their predictive capabilities within a realistic operational context.

The acquired results highlight significant accuracy improvements achieved by the Transformer-based model, with performance gains ranging from 5% to 20% over LSTM and GRU models, along with improved computational efficiency [36-38]. The superior performance of the proposed model primarily stems from its attention mechanism, which enables simultaneous analysis of all time steps, effectively capturing intricate long-term temporal patterns that traditional sequential models often struggle with. The findings confirm and build upon the existing literature by demonstrating that GRU models exhibit comparatively higher error rates (MAE and MAPE) for wind speed [39, 40] and LSTM models excel at producing precise prediction intervals suitable for high-precision forecasting tasks [41, 42].

A key contribution of this study lies in demonstrating, through practical application, the tangible advantages and predictive superiority of Transformer-based methods, filling an important research gap. The implications of these findings extend directly to practical wind energy management, facilitating better integration of renewable energy sources into existing power grids, optimized operational decision-making, and potentially significant reductions in energy production costs.

Despite the promising results, our study recognizes the computational intensity of Transformer models as a notable limitation. Future research directions should thus focus on developing optimized Transformer variants or hybrid approaches capable of preserving high predictive accuracy while reducing computational demand. Such advancements will further enhance the practicality and scalability of these forecasting models, driving forward the development of robust, cost-efficient, and environmentally sustainable wind energy solutions essential to global climate change mitigation efforts.

REFERENCES

[1] A. Allouhi, "A hybrid PV/wind/battery energy system to assist a run-of-river micro-hydropower for clean electrification and fuelling hydrogen mobility for young population in a rural Moroccan site," *Journal of Cleaner Production*, vol. 442, Feb. 2024, Art. no. 140852, https://doi.org/10.1016/j.jclepro.2024.140852.

[2] A. El-Maaroufi, M. Daoudi, and R. Ahl Laamara, "Optimal design and techno-economic analysis of a solar-wind hybrid power system for laayoune city electrification with hydrogen and batteries as a storage device," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 136, Dec. 2024, Art. no. 103719, https://doi.org/10.1016/j.pce.2024.103719.

[3] M. Bousla, A. Haddi, Y. El Mourabit, A. Sadki, A. Mouradi, and A. El Kharrim, "Detection and Prevention of Repetitive Major Faults of a WTG by Analysis of Alarms Through SCADA," in *Digital Technologies and Applications*, 2023, pp. 745–752, https://doi.org/10.1007/978-3-031-29857-8_74.

[4] M. Bousla, Y. El Mourabit, and A. Haddi, "Optimizing Wind Farm Maintenance in Northern Morocco through SCADA System Failure Data Analysis," in *2024 International Conference on Computing, Internet of Things and Microwave Systems (ICCIMS)*, Gatineau, QC, Canada, Jul. 2024, https://doi.org/10.1109/ICCIMS61672.2024.10690783.

[5] X. Zhao, C. Wang, J. Su, and J. Wang, "Research and application based on the swarm intelligence algorithm and artificial intelligence for wind farm decision system," *Renewable Energy*, vol. 134, pp. 681–697, Apr. 2019, https://doi.org/10.1016/j.renene.2018.11.061.

[6] S. Jafarian-Namin, A. Goli, M. Qolipour, A. Mostafaeipour, and A.-M. Golmohammadi, "Forecasting the wind power generation using Box–Jenkins and hybrid artificial intelligence: A case study," *International Journal of Energy Sector Management*, vol. 13, no. 4, pp. 1038–1062, Jun. 2019, https://doi.org/10.1108/IJESM-06-2018-0002.

[7] M. Bousla *et al.*, "Analysis and Comparison of Wind Potential by Estimating the Weibull Distribution Function: Application to Wind Farm in the Northern of Morocco," *Sustainability*, vol. 15, no. 20, Jan. 2023, Art. no. 15087, https://doi.org/10.3390/su152015087.

[8] Y. Amellas, O. E. Bakkali, A. Djebli, and A. Echchelh, "Short-term wind speed prediction based on MLP and NARX network models," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 1, pp. 150–157, Apr. 2020, https://doi.org/10.11591/ijeecs.v18.i1.pp150-157.

[9] Y. Li *et al.*, "An Explainable Deep Learning Model for Daily Sea Ice Concentration Forecast," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–17, 2024, https://doi.org/10.1109/TGRS.2024.3386930.

[10] G. De Moliner, P. Giani, G. Lonati, and P. Crippa, "Sensitivity of multiscale large Eddy simulations for wind power calculations in complex terrain," *Applied Energy*, vol. 364, Jun. 2024, Art. no. 123195, https://doi.org/10.1016/j.apenergy.2024.123195.

[11] S. Rüstemli, O. Güntas, G. Şahin, A. Koç, W. van Sark, and S. Ş. Doğan, "Wind power plant site selection problem solution using GIS and resource assessment and analysis of wind energy potential by estimating Weibull distribution function for sustainable energy production: The case of Bitlis/Turkey," *Energy Strategy Reviews*, vol. 56, Nov. 2024, Art. no. 101552, https://doi.org/10.1016/j.esr.2024.101552.

[12] W. Li, Y. Chong, X. Guo, and J. Liu, "A hybrid wind power prediction model based on seasonal feature decomposition and enhanced feature extraction," *Energy and AI*, vol. 18, Dec. 2024, Art. no. 100442, https://doi.org/10.1016/j.egyai.2024.100442.

[13] S. M. Valdivia-Bautista, J. A. Domínguez-Navarro, M. Pérez-Cisneros, C. J. Vega-Gómez, and B. Castillo-Téllez, "Artificial Intelligence in Wind Speed Forecasting: A Review," *Energies*, vol. 16, no. 5, Jan. 2023, Art. no. 2457, https://doi.org/10.3390/en16052457.

[14] L.-L. Li, X. Zhao, M.-L. Tseng, and R. R. Tan, "Short-term wind power forecasting based on support vector machine with improved dragonfly

algorithm," *Journal of Cleaner Production*, vol. 242, Jan. 2020, Art. no. 118447, https://doi.org/10.1016/j.jclepro.2019.118447.

[15] L.-L. Li, Y.-B. Chang, M.-L. Tseng, J.-Q. Liu, and M. K. Lim, "Wind power prediction using a novel model on wavelet decomposition-support vector machines-improved atomic search algorithm," *Journal of Cleaner Production*, vol. 270, Oct. 2020, Art. no. 121817, https://doi.org/10.1016/j.jclepro.2020.121817.

[16] B. Memon, M. H. Baloch, A. H. Memon, S. H. Qazi, R. Haider, and D. Ishak, "Assessment of Wind Power Potential Based on Raleigh Distribution Model: An Experimental Investigation for Coastal Zone," *Engineering, Technology & Applied Science Research*, vol. 9, no. 1, pp. 3721–3725, Feb. 2019, https://doi.org/10.48084/etasr.2381.

[17] M. Bousla, A. Haddi, Y. E. Mourabit, A. Sadki, A. Mouradi, and A. E. Kharrim, "Optimizing Wind Farm Performance in Northern Morocco with Weibull Distribution," in *Digital Technologies and Applications*, 2024, pp. 454–463, https://doi.org/10.1007/978-3-031-68653-5_43.

[18] S. Hu, Y. Xiang, D. Huo, S. Jawad, and J. Liu, "An improved deep belief network based hybrid forecasting method for wind power," *Energy*, vol. 224, Jun. 2021, Art. no. 120185, https://doi.org/10.1016/j.energy.2021.120185.

[19] Z. Jiang, J. Che, M. He, and F. Yuan, "A CGRU multi-step wind speed forecasting model based on multi-label specific XGBoost feature selection and secondary decomposition," *Renewable Energy*, vol. 203, pp. 802–827, Feb. 2023, https://doi.org/10.1016/j.renene.2022.12.124.

[20] J.-P. Lai, Y.-M. Chang, C.-H. Chen, and P.-F. Pai, "A Survey of Machine Learning Models in Renewable Energy Predictions," *Applied Sciences*, vol. 10, no. 17, Jan. 2020, Art. no. 5975, https://doi.org/10.3390/app10175975.

[21] L. Donadio, J. Fang, and F. Porté-Agel, "Numerical Weather Prediction and Artificial Neural Network Coupling for Wind Energy Forecast," *Energies*, vol. 14, no. 2, Jan. 2021, Art. no. 338, https://doi.org/10.3390/en14020338.

[22] Z. Zhang, H. Dai, D. Jiang, Y. Yu, and R. Tian, "Multi-step ahead forecasting of wind vector for multiple wind turbines based on new deep learning model," *Energy*, vol. 304, Sep. 2024, Art. no. 131964, https://doi.org/10.1016/j.energy.2024.131964.

[23] M. Bousla, M. Belfkir, A. Haddi, Y. El Mourabit, and B. Bossoufi, "Comparison of artificial intelligence approaches for estimating wind energy production: A real-world case study," *Results in Engineering*, vol. 24, Dec. 2024, Art. no. 103626, https://doi.org/10.1016/j.rineng.2024.103626.

[24] Adomako, Abigail Birago, et al. "Deep learning approaches for bias correction in WRF model outputs for enhanced solar and wind energy estimation: A case study in East and West Malaysia." Ecological Informatics (2024): 102898.

[25] U. Singh and M. Rizwan, "SCADA system dataset exploration and machine learning based forecast for wind turbines," *Results in Engineering*, vol. 16, Dec. 2022, Art. no. 100640, https://doi.org/10.1016/j.rineng.2022.100640.

[26] J. He and J. Xu, "Ultra-short-term wind speed forecasting based on support vector machine with combined kernel function and similar data," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, Nov. 2019, Art. no. 248, https://doi.org/10.1186/s13638-019-1559-1.

[27] C.-S. Tu, C.-M. Hong, H.-S. Huang, and C.-H. Chen, "Short Term Wind Power Prediction Based on Data Regression and Enhanced Support Vector Machine," *Energies*, vol. 13, no. 23, Jan. 2020, Art. no. 6319, https://doi.org/10.3390/en13236319.

[28] Y.-D. Syu *et al.*, "Ultra-Short-Term Wind Speed Forecasting for Wind Power Based on Gated Recurrent Unit," in *2020 8th International Electrical Engineering Congress (iEECON)*, Chiang Mai, Thailand, Mar. 2020, https://doi.org/10.1109/iEECON48109.2020.229518.

[29] R. Yu *et al.*, "LSTM-EFG for wind power forecasting based on sequential correlation features," *Future Generation Computer Systems*, vol. 93, pp. 33–42, Apr. 2019, https://doi.org/10.1016/j.future.2018.09.054.

[30] Z. Niu, Z. Yu, W. Tang, Q. Wu, and M. Reformat, "Wind power forecasting using attention-based gated recurrent unit network," *Energy*, vol. 196, Apr. 2020, Art. no. 117081, https://doi.org/10.1016/j.energy.2020.117081.

[31] M. A. Saeed, Z. Ahmed, J. Yang, and W. Zhang, "An optimal approach of wind power assessment using Chebyshev metric for determining the Weibull distribution parameters," *Sustainable Energy Technologies and Assessments*, vol. 37, Feb. 2020, Art. no. 100612, https://doi.org/10.1016/j.seta.2019.100612.

[32] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. null, pp. 2825–2830, Aug. 2011.

[33] H. Zhou *et al.*, "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 12, pp. 11106–11115, May 2021, https://doi.org/10.1609/aaai.v35i12.17325.

[34] S. Mo, H. Wang, B. Li, Z. Xue, S. Fan, and X. Liu, "Powerformer: A temporal-based transformer model for wind power forecasting," *Energy Reports*, vol. 11, pp. 736–744, Jun. 2024, https://doi.org/10.1016/j.egyr.2023.12.030.

[35] Y. Wang, R. Zou, F. Liu, L. Zhang, and Q. Liu, "A review of wind speed and wind power forecasting with deep neural networks," *Applied Energy*, vol. 304, Dec. 2021, Art. no. 117766, https://doi.org/10.1016/j.apenergy.2021.117766.

[36] M. R. Sarkar, S. G. Anavatti, T. Dam, M. Pratama, and B. A. Kindhi, "Enhancing Wind Power Forecast Precision via Multi-head Attention Transformer: An Investigation on Single-step and Multi-step Forecasting," in *2023 International Joint Conference on Neural Networks (IJCNN)*, Gold Coast, Australia, Jun. 2023, https://doi.org/10.1109/IJCNN54540.2023.10191444.

[37] A. Ghasemi and M. Hashemi, "Harnessing the Power of Transformer Learning with Long-Term Memory for Renewable Energy Forecasting," in *2023 IEEE Virtual Conference on Communications (VCC)*, NY, USA, Aug. 2023, pp. 171–176, https://doi.org/10.1109/VCC60689.2023.10475044.

[38] S. Huang, C. Yan, and Y. Qu, "Deep learning model-transformer based wind power forecasting approach," *Frontiers in Energy Research*, vol. 10, Jan. 2023, https://doi.org/10.3389/fenrg.2022.1055683.

[39] M. Robith, R. S. Wibowo, and W. Wibowo, "Short-Term Wind Power Forecasting in East Java Using Gated Recurrent Unit: 3rd International Conference on Electronic and Electrical Engineering and Intelligent System, ICE3IS 2023," *Proceedings - 2023 3rd International Conference on Electronic and Electrical Engineering and Intelligent System*, pp. 58–63, 2023, https://doi.org/10.1109/ICE3IS59323.2023.10335218.

[40] D. Sakubu and Q. Liu, "Microgrid Forecasting Using Multiple Deep learning Models," in *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*, Haikou, Hainan, China, Sep. 2021, pp. 1490–1497, https://doi.org/10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00222.

[41] A. A. Ewees, M. A. A. Al-qaness, L. Abualigah, and M. A. Elaziz, "HBO-LSTM: Optimized long short term memory with heap-based optimizer for wind power forecasting," *Energy Conversion and Management*, vol. 268, Sep. 2022, Art. no. 116022, https://doi.org/10.1016/j.enconman.2022.116022.

[42] A. Banik, C. Behera, Tirunagaru. V. Sarathkumar, and A. K. Goswami, "Uncertain wind power forecasting using LSTM-based prediction interval," *IET Renewable Power Generation*, vol. 14, no. 14, pp. 2657–2667, 2020, https://doi.org/10.1049/iet-rpg.2019.1238.