Spatio-Temporal Prediction of Solar Irradiance and Wind Speed Using Machine Learning: A Case Study in Tunisia

Amel Chebbi[1]*, Wissem Zghal[2,3], and Mounir Ben Aissa[3,4]

[1] Laboratoire Optimisation, Logistique et Informatique Décisionnelle (OLID), ISGIS, University of Sfax, Sfax 3021, Tunisia

[2] Laboratory of Electromechanical Systems (LASEM), National Engineering School of Sfax, University of Sfax, Sfax 3038, Tunisia

[3] Higher Institute of Industrial Management of Sfax, University of Sfax, Sfax 3021, Tunisia

[4] OASIS Laboratory, ENIT, Tunisia

DOI: https://doi.org/10.51244/IJRSI.2026.1310000170

Received: 20 October 2026; Accepted: 25 October 2026; Published: 12 November 2026

ABSTRACT

Accurate prediction of solar irradiance and wind speed is essential for renewable energy resource assessment and the planning of photovoltaic (PV), wind, and hybrid PV-wind systems.

This study investigates a spatio-temporal machine learning framework for predicting meteorological variables using observations reported as collected by the National Institute of Meteorology (INM) of Tunisia from 1 September 2023 to 1 September 2024. The dataset covers five Tunisian locations and includes global solar irradiation, wind speed, air temperature, and geographic coordinates. Three regression algorithms , Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) are considered to model the nonlinear relationships between meteorological conditions and renewable energy resources. The comparative results reported in the source manuscript show that XGBoost achieves the highest coefficient of determination (($R^2 = 0.812$)), followed by Random Forest (($R^2 = 0.803$)) and Decision Tree (($R^2 = 0.732$)). XGBoost also yields the lowest reported mean absolute error (MAE = 168.146) and root mean square error (RMSE = 429.362).

However, the available results do not clearly specify the target variable, measurement units, or validation protocol, preventing these performance metrics from being attributed independently to solar irradiance and wind speed. Accordingly, the study establishes a methodological framework for evaluating both prediction tasks and highlights the need for target-specific, temporally consistent validation. Once validated, the forecasts can support renewable resource assessment and preliminary PV-wind system design, providing a basis for subsequent energy yield estimation and techno-economic optimization.

Keywords: Solar irradiance, wind speed forecasting, machine learning, XGBoost, Random Forest, Decision Tree , spatio-temporal modelling, renewable energy.

INTRODUCTION

The increasing integration of renewable energy technologies into modern energy systems has strengthened the need for reliable prediction of weather dependent resources. Solar photovoltaic (PV) and wind energy are particularly sensitive to meteorological variability, making accurate estimation of global solar irradiance and wind speed essential for renewable-resource assessment, preliminary system sizing, and energy planning decisions. Solar forecasting research has progressed from conventional statistical approaches towards machine learning (ML) methods capable of representing nonlinear relationships between meteorological variables and solar-resource availability [1], [2]. Recent reviews further emphasize the importance of data

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

quality, feature selection, forecasting horizon, and model evaluation when comparing solar-prediction approaches [3], [4]. Similar challenges arise in wind forecasting, where atmospheric variability, incomplete input information, and nonlinear dependencies complicate the prediction of wind speed and power [5],[7].

Despite these advances, predictive performance across multiple locations remains a significant methodological challenge. Solar irradiance exhibits pronounced daily and seasonal cycles, while wind speed can fluctuate rapidly under changing atmospheric conditions. Moreover, geographical differences may affect the statistical relationships between meteorological variables and renewable-resource availability. Consequently, models trained using observations from several stations must distinguish temporal patterns from site specific characteristics. A model that performs well on randomly selected observations from familiar locations may not necessarily generalize to future periods or previously unseen sites. Forecast evaluation must therefore be designed according to the intended application, with chronological validation for future prediction and, where feasible, location-based testing for spatial transferability [8], [9]. These considerations are particularly important when meteorological records cover a limited period or a relatively small number of stations.

Machine learning offers a practical framework for investigating such nonlinear and heterogeneous relationships. Decision Tree (DT) regression provides an interpretable baseline by partitioning the predictor space into rule based regions. Random Forest (RF) combines predictions from multiple randomized decision trees, helping reduce the variance associated with individual trees [10]. Gradient boosting methods build successive learners to improve prediction, while XGBoost incorporates regularization and computationally efficient tree boosting [11], [12]. Recent comparative studies and reviews demonstrate the continued relevance of these approaches and more advanced learning architectures in solar and wind forecasting [3],[7], [13]. Nevertheless, algorithm performance is context dependent , no single method can be considered universally superior independently of the dataset, prediction target, and validation procedure.

In this context, the present study investigates a comparative spatio-temporal ML framework for predicting global solar irradiance and wind speed at five locations in Tunisia. The empirical basis consists of meteorological observations reported in the supplied manuscript as collected by the National Institute of Meteorology (INM) of Tunisia between 1 September 2023 and 1 September 2024. The available variables include global solar irradiation, wind speed, average air temperature, and geographical coordinates. The study compares DT, RF, and XGBoost and aims to establish a transparent basis for assessing their predictive behaviour across the available locations.

The principal contribution is a multi-location assessment that treats solar-irradiance prediction and wind-speed prediction as distinct but complementary tasks. In contrast to comparisons based solely on aggregate scores, the proposed evaluation framework emphasizes target-specific metrics, consistent preprocessing, clearly defined input and output variables, and validation that respects temporal ordering.

The coefficient of determination ((R^2)), mean absolute error (MAE), and root mean square error (RMSE) are suitable complementary indicators when calculated and interpreted separately for each target [8], [9]. The analysis also identifies the limits imposed by the available observation period and the need for further evidence before claiming generalization to unseen locations. The aggregate scores reported in the source manuscript are therefore regarded as preliminary until the corresponding target-specific results and validation protocol are verified.

By applying this framework to a Tunisian case study, the research seeks to clarify how established tree-based algorithms can be used to investigate spatial and temporal variations in renewable-energy resources. Reliable forecasts could subsequently inform preliminary PV and wind resource assessment, while electrical-energy estimation and techno-economic system optimization would require additional modelling of technology characteristics, conversion efficiency, and energy demand. The study thus provides a reproducible methodological starting point for connecting meteorological prediction with renewable-energy planning without overstating the evidence currently available.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

MATERIALS AND METHODS

presents a machine learning-based model designed to predict key climatic parameters by leveraging geographic inputs such as longitude and altitude. This approach enables estimation of localized weather conditions in regions lacking direct measurement stations. The model integrates algorithms including Random Forests and linear regression to accurately forecast variables such as wind speed, solar irradiation, and temperature. These parameters play a critical role in the optimal design and sizing of renewable energy systems, facilitating reliable assessments without dependence on physical sensor networks.

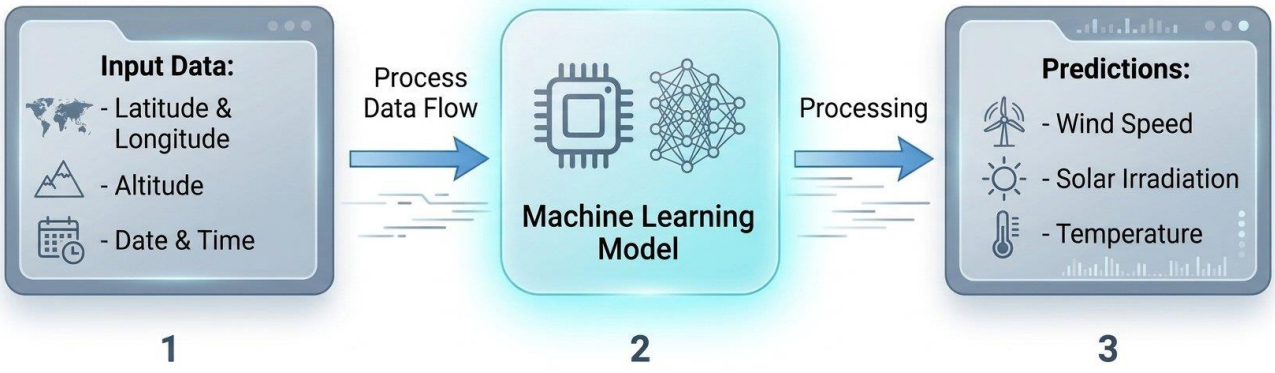


Figure 1: Process of an Artificial Intelligence (AI) model

Study area and data source

The source manuscript reports daily meteorological data obtained from Tunisia's National Institute of Meteorology for the period from 1 September 2023 to 1 September 2024. Five geographic locations are represented by the coordinates listed in Table 1. The reported variables include global solar irradiation, wind speed, and average air temperature. Latitude and longitude are used to represent location; calendar features such as day, month, and year may be derived from the observation date if those fields are available in the original data file.

Table 1. Geographic coordinates

	Latitude	**Longitude**
Zone 1	34.7394361	10.7604024
Zone 2	33.8878082	10.10044
Zone 3	34.881181	9.5263598
Zone 4	35.1687646	8.8365654
Zone 5	35.7707582	10.8280511

Machine Learning Models and Forecasting Methodology

Overview of the Forecasting Framework

The proposed forecasting framework is based on supervised machine learning techniques designed to predict key meteorological variables influencing renewable energy production. The methodology combines

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

meteorological observations and geographical information to establish nonlinear relationships between climatic conditions and renewable energy resources.

The workflow of the proposed approach consists of five main stages:

- Data acquisition from multiple Tunisian climatic regions.
- Data preprocessing and normalisation.
- Feature selection and spatial information integration.
- Machine learning model training and validation.

Three machine learning algorithms were selected for comparative evaluation: Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). These algorithms were chosen because they represent different levels of model complexity and have demonstrated promising performance in renewable energy forecasting applications.

Decision Tree (DT)

Decision Tree is a non-parametric supervised learning algorithm that partitions the input feature space into a set of decision regions through recursive binary splitting. The model predicts the target variable by following a sequence of decision rules derived from the training data.

At each node, the algorithm selects the feature that maximizes information gain while minimizing prediction error. For regression problems, the splitting criterion is generally based on minimizing the Mean Squared Error (MSE):

$$MSE = (1/N) \Sigma (yi - \hat{y}i)^2 \qquad (1)$$

where yi represents the observed value and ŷi denotes the predicted value.

Decision Trees offer several advantages, including interpretability, low computational cost, and the ability to model nonlinear relationships. However, individual trees are often sensitive to noise and may suffer from overfitting when the tree depth becomes excessively large.

Random Forest Model

Random Forest is an ensemble learning algorithm that combines multiple Decision Trees to improve predictive accuracy and robustness. During training, each tree is generated using a bootstrap sample of the original dataset, while a random subset of features is considered at each split.

The final prediction is obtained by averaging the predictions of all individual trees:

$$\hat{y} = \frac{1}{T}\sum_{i=1}^{T} y_i \qquad (2)$$

where:

- T is the total number of trees,
- y_i is the prediction generated by tree i.

The ensemble strategy reduces variance and improves generalization performance compared with a single Decision Tree.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

XG Boost Model

XGBoost (Extreme Gradient Boosting) is an advanced ensemble learning algorithm based on gradient boosting principles. The model sequentially constructs decision trees, where each new tree attempts to minimize the residual errors produced by the previous ensemble.

The objective function can be expressed as:

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)}) + \sum_{k=1}^{K} \Omega(f_k) \quad (3)$$

where:

- l de notes the loss function,
- f_k represents the individual trees,
- Ω is a regularization term controlling model complexity.

The inclusion of regularization mechanisms constitutes one of the major advantages of XGBoost because it helps reduce overfitting while maintaining high predictive accuracy. Compared with conventional boosting algorithms, XGBoost provides faster training, improved scalability, and enhanced predictive performance, making it particularly suitable for renewable energy forecasting applications.

Hyperparameter Configuration

Table 2: Hyperparameter configuration of the investigated models.

Parameter	**DT**	**RF**	**XGBoost**
Max Depth	10	20	8
Number of Trees	1	160	160
Learning Rate	–	–	0.05
Min Samples Split	2	2	2
Subsample	–	–	0.9
Colsample Bytree	–	–	0.9

Hyperparameter values were selected after several preliminary experiments aimed at achieving a balance between prediction accuracy and model complexity.

Attention was given to preventing overfitting while maintaining satisfactory generalization performance.

Training and Validation Procedure

The dataset was divided into training and testing subsets using an 80/20 ratio. The training subset was employed for model development, whereas the testing subset was exclusively reserved for performance evaluation.

Model performance was assessed using four statistical indicators:

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{x}_i - x_i| \quad (4)$$

- Mean Squared Error (MSE)

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - x_i)^2 \quad (5)$$

- Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - x_i)^2} \quad (6)$$

- Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad (7)$$

These metrics provide complementary information regarding prediction accuracy and model reliability.

RESULTS AND DISCUSSION

The prediction framework developed in this study relies on meteorological data representing the temporal variability of solar irradiance, wind speed, and ambient temperature under Tunisian climatic conditions. These variables were selected because they characterize the main environmental factors influencing photovoltaic and wind energy production. Solar irradiance determines the solar energy available for photovoltaic conversion, while wind speed governs the energy captured by wind turbines. Ambient temperature is also considered because it affects photovoltaic module efficiency and provides additional information about local weather conditions. The meteorological dataset is subsequently used to develop and evaluate machine-learning models, particularly XGBoost, for predicting solar irradiance and wind speed. Ambient temperature is included as an additional meteorological variable, subject to its availability and its actual role in the model configuration.

Comparing Different Algorithms

The study evaluates the predictive performance of three machine learning model Decision Tree, Random Forest, and XGBoost in forecasting solar radiation and wind speed. The assessment employs widely used evaluation metrics, including the coefficient of determination (R^2), mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE), as summarized in Table 3.

Table 3: Performance Metrics

Algorithm	**R^2**	**MAE**	**MSE**	**RMSE**
Decision tree	0.732	259.358	356049.809	596.699
Random Forest	0.803	170.531	196205.434	442.951
XGBoost	0.812	168.146	184351.688	429.362

XGBoost demonstrates superior performance, evident from its highest R^2 value of 0.812, which indicates better model fit and its ability to capture complex nonlinear relationships within the data. Moreover, XGBoost attains the lowest errors across MAE, MSE, and RMSE metrics, underscoring its enhanced

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

accuracy and reliability over Random Forest and Decision Tree models in renewable energy forecasting tasks.

This improved performance arises primarily from XGBoost's gradient boosting framework, which iteratively minimizes prediction errors and effectively handles complex interactions compared to the averaging approach of Random Forest and the singular structure of Decision Trees. Such findings align with prevailing literature advocating for XGBoost's utilization in high-precision energy modeling and prediction scenarios.

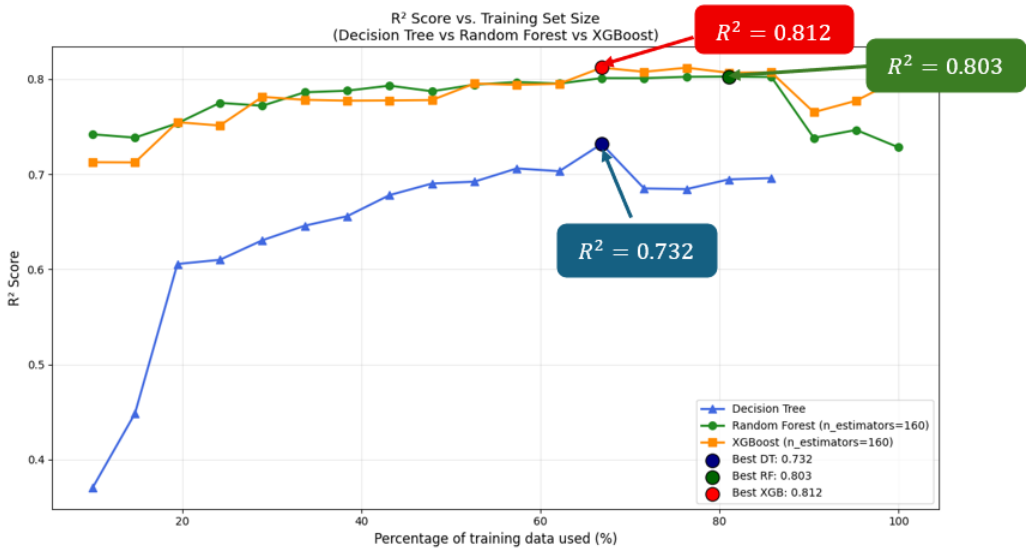


Figure 2: R^2 for different models

The results clearly indicate that the XGBoost model outperforms both the Random Forest and Decision Tree models in terms of predictive accuracy. This is evidenced by XGBoost's highest coefficient of determination (R^2) value, which reflects its enhanced ability to capture complex nonlinear relationships between the input variables and the target outcomes. The gradient boosting framework employed by XGBoost iteratively corrects errors made by previous trees, thereby refining the model's predictions and increasing reliability. Consequently, XGBoost delivers more precise and robust forecasts compared to the ensemble averaging approach of Random Forest and the simpler structure of Decision Trees. These findings align with current literature that highlights XGBoost's superior performance, particularly in scenarios involving intricate data patterns and the necessity for high prediction precision

Comparative Analysis of Observed and Predicted Meteorological Variables

Figure 3 compares observed and predicted air temperature, wind speed, and global solar irradiation over the period from September 1, 2023, to September 1, 2024, The model was trained using data from four sites, while the fifth site was selected for validation to demonstrate the temporal and spatial consistency of the predictions and the model's ability to generalize to previously unseen locations. in Zone 1(33.8878082° N, 10.10044° E). The objective is to evaluate the model's ability to reproduce meteorological variations under the climatic conditions of the study area.

The results suggest good agreement between observed and predicted values, particularly for temperature and global solar irradiation, indicating that the model captures the main temporal and seasonal patterns. However, some discrepancies remain in wind speed predictions, potentially due to atmospheric turbulence, local environmental conditions, and measurement uncertainties.

These findings highlight the potential of the proposed model for renewable energy forecasting, particularly in photovoltaic and wind energy applications. Nevertheless, quantitative assessment using MAE, RMSE,

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

and (R^2), alongside validation on independent datasets, is required to confirm its predictive accuracy, reliability, and generalization capability across different climatic conditions.

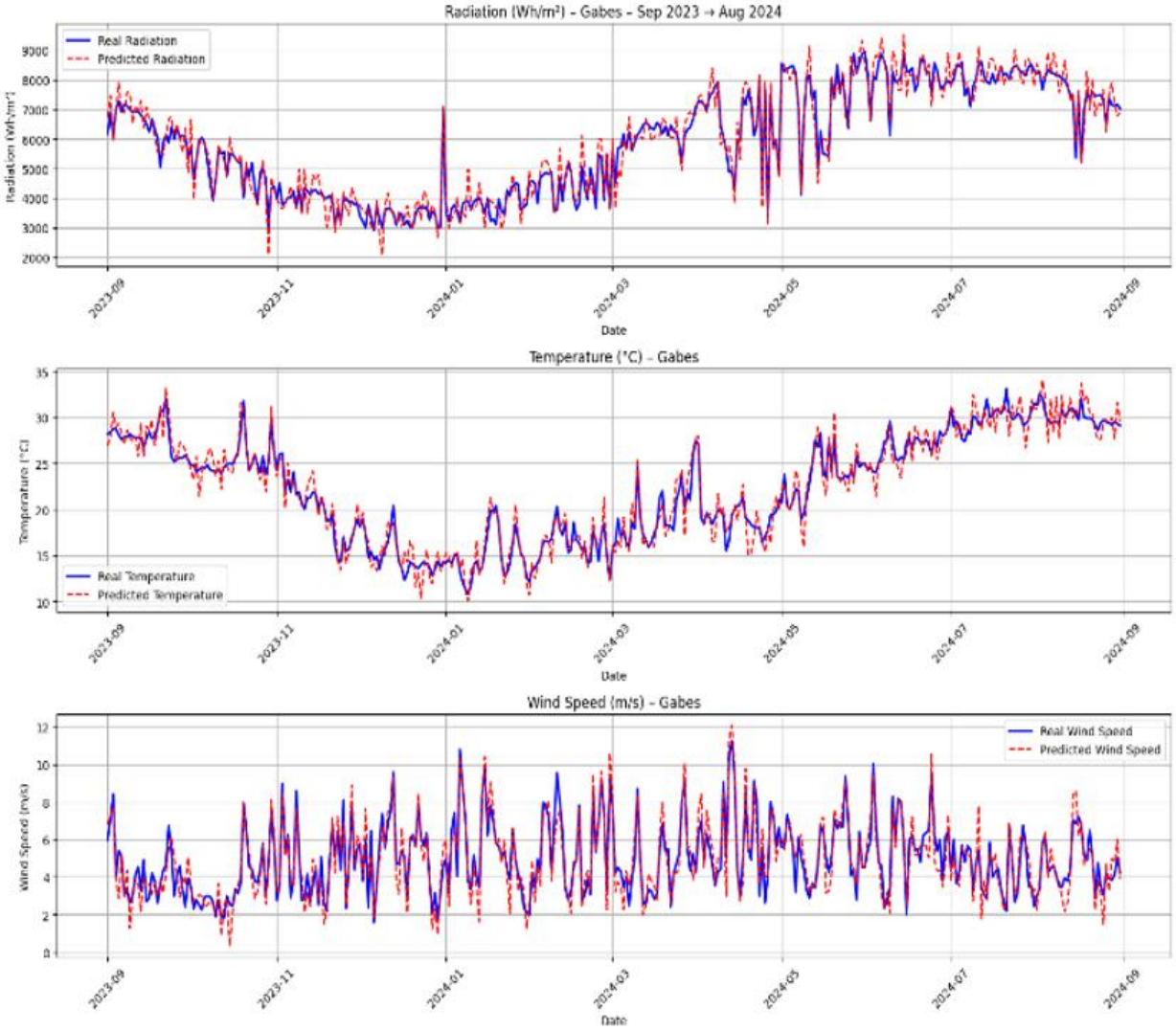


Figure 3: Comparison of Real and Predicted values

Figure 4 provides additional validation of the model's predictive accuracy by comparing actual versus predicted values for temperature, wind speed, and global solar irradiation. This validation uses a randomly selected subsample of more than 30 dates scattered throughout the previously analyzed one-year dataset from Figure 10, thereby assessing the model's generalization capability under non-continuous, sporadic temporal conditions.

The findings reveal consistent alignment between observed and predicted values for temperature and global solar irradiation, reaffirming the model's efficacy in capturing seasonal and meteorological patterns even with intermittent data.

Wind speed predictions exhibit somewhat increased variability, reflecting the inherently stochastic nature of wind dynamics, as commonly reported in meteorological forecasting studies. Despite this, the overall congruence reinforces the robustness and reliability of the model for practical applications where data availability is irregular or fragmented.

This complementary analysis highlights the model's potential for deployment in real-world renewable energy forecasting scenarios characterized by incomplete temporal data, supporting the development of resilient AI-driven predictive systems. Further validation across diverse geographic locations and temporal resolutions would serve to broaden confidence in generalizability and predictive stability.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

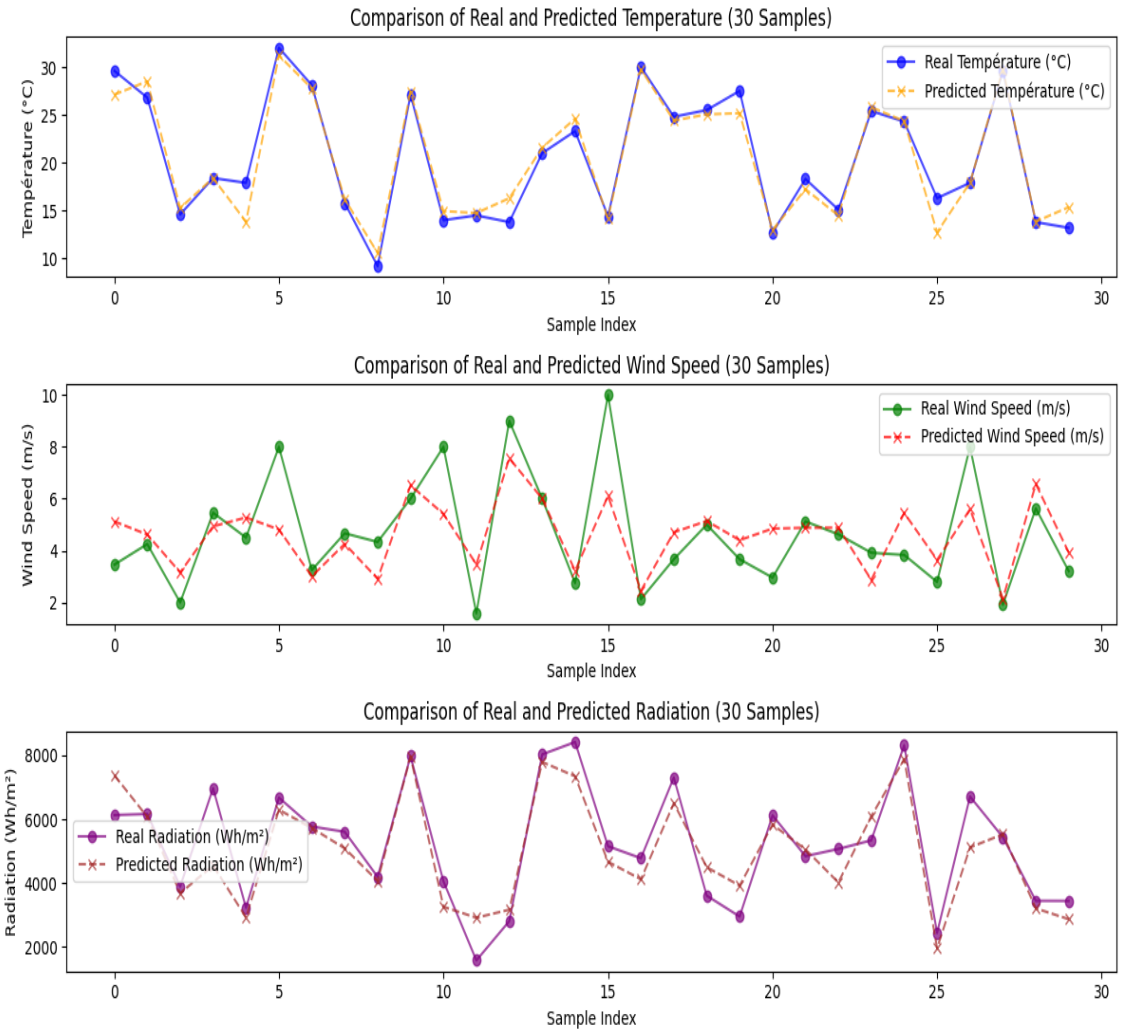


Figure 4: Comparison of Real and Predicted Irradiation (30 dates)

Advanced Prediction of Solar Irradiation, Wind Speed, and Temperature

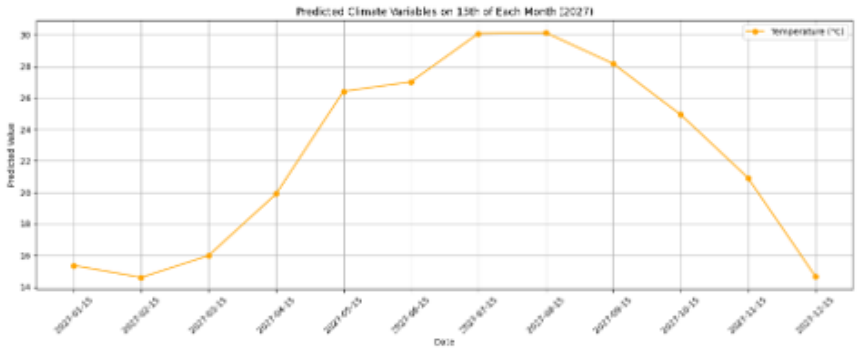
The analysis was further refined by focusing on a geographically fixed reference point located at latitude 34.7449° and longitude 10.7832°. At this specific location, a series of targeted temporal analyses were conducted to evaluate the model's predictions of solar irradiation, temperature, and wind speed across strategically selected dates, as illustrated in Figure 5. This methodological approach was designed to rigorously assess the model's temporal sensitivity its capacity to maintain consistent predictive performance amidst varying seasonal and temporal conditions.

Evaluating temporal sensitivity is vital to verifying the robustness and reliability of forecasting models, especially when applied over extended periods with different meteorological regimes. By investigating the model's behavior across discrete dates that capture seasonal dynamics, fluctuations in weather patterns, and climatic variability, the study ensures that the model can produce accurate and stable forecasts beyond the training data. Such an approach reflects real-world conditions where models must generalize effectively despite inherent temporal heterogeneity.

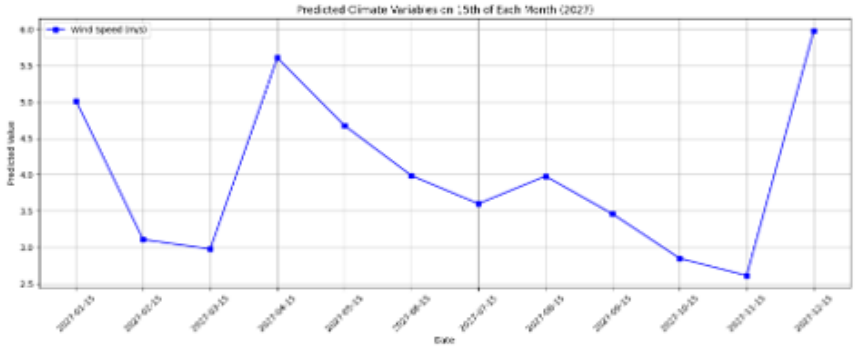
Findings from this temporal analysis underscore the model's ability to capture critical meteorological variations consistently, highlighting strengths in predicting temperature and solar irradiation with stable accuracy.

The temporal sensitivity assessment thereby supports the practical applicability of the model in renewable energy forecasting, encouraging future research to expand validations across diverse temporal scales and geographic locations for enhanced reliability and extrapolation.

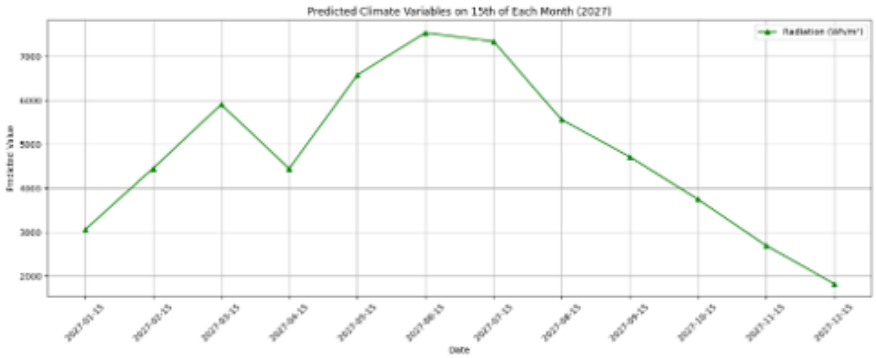
ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026



a. Prediction temperature



b. Prediction wind speed



c. Prediction global solar irradiance

Figure 5: Forecasting Meteorological Variables for the Location

(34.7449° N, 10.7832° E)

CONCLUSION

This study proposes a spatio-temporal machine learning framework for predicting global solar irradiation and wind speed across five geographically distributed locations in Tunisia, addressing a fundamental challenge in renewable energy assessment: the reliable characterization of meteorological resources under spatially and temporally variable climatic conditions. By comparing Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), the study establishes a comparative framework for investigating the suitability of tree-based learning algorithms for renewable resource prediction using meteorological observations collected over the period from 1 September 2023 to 1 September 2024.

The principal contribution of this research lies in the integration of spatial information, temporal meteorological variability, and comparative machine learning evaluation within a unified framework specifically applied to Tunisian renewable energy resources. Unlike an evaluation based exclusively on a single location or a single prediction algorithm, the proposed approach considers five geographically distinct

sites and emphasizes the importance of distinguishing solar irradiance prediction from wind speed prediction. This distinction is essential because the two variables are governed by different physical processes and require separate performance assessments before their predictions can be reliably used in renewable energy applications. Furthermore, the study establishes a methodological basis for evaluating the generalization capabilities of predictive models through chronological validation and location-specific performance analysis.

The comparative results reported in the available manuscript indicate that XGBoost achieved the strongest aggregate performance, with (R^2=0.812), (MAE=168.146), and (RMSE=429.362), outperforming Random Forest and Decision Tree according to the reported metrics. These results suggest that gradient-boosted tree ensembles may represent nonlinear meteorological relationships more effectively than a single decision tree in the investigated dataset. However, the scientific robustness of this conclusion depends on confirming the target variables, measurement units, data partitions, and evaluation protocol associated with these values. Accordingly, the reported performance establishes a preliminary comparative indication rather than definitive evidence of independently validated accuracy for both solar irradiance and wind speed.

A second contribution is the establishment of a reproducibility-oriented evaluation strategy**.** The proposed framework highlights the need for separate prediction tasks, chronological training and testing, consistent evaluation metrics, transparent hyperparameter selection, and comparisons with simple forecasting baselines. Incorporating these elements into the final experimental implementation would make it possible to distinguish genuine predictive improvements from potential effects of temporal dependence, data leakage, or differences in target scaling. The resulting methodology would also facilitate comparisons between geographical locations and provide a more reliable basis for evaluating the transferability of machine learning models to meteorological conditions not represented in the training data.

From an applied perspective, validated forecasts of solar irradiance and wind speed can provide the meteorological inputs required for subsequent photovoltaic and wind energy modelling. Their integration with photovoltaic conversion models, wind turbine power curves, technical constraints, and system losses would enable more realistic estimates of renewable electricity production. In a further development, these forecasts could be incorporated into the techno-economic optimization of standalone PV, wind, and hybrid PV-wind systems, considering energy demand, storage capacity, supply reliability, investment costs, operation and maintenance expenditures, and environmental indicators. Nevertheless, these downstream benefits remain prospective and require explicit validation through energy-system simulation and optimization.

Several limitations define the scope of the present findings. The dataset covers approximately one year and five locations, limiting the assessment of interannual climatic variability and the representativeness of the results at the national scale. In addition, the aggregate performance metrics available in the manuscript do not establish separate predictive accuracy for the two target variables. Future research should therefore extend the observation period, increase the number of meteorological stations, validate each target independently, and evaluate performance across different seasons and locations. Further investigations should also compare optimized XGBoost models with temporal learning approaches, including Long Short-Term Memory networks, when sufficient sequential data are available, and quantify the uncertainty associated with the resulting forecasts.

Ultimately, this research provides a methodological foundation for the data-driven characterization of solar and wind resources in Tunisia. Its main scientific value is not limited to identifying a potentially effective prediction algorithm; it also emphasizes the importance of combining geographical information, appropriate temporal validation, reproducible experimentation, and physically meaningful interpretation of forecast outputs.

Following target-specific validation, this framework could support more reliable renewable resource assessment and serve as a basis for integrating machine learning predictions into multi-objective energy-system design. Such an extension would create a direct methodological link between meteorological forecasting and the techno-economic and environmental optimization of hyrid renewable energy systems.

REFERENCES

1. C. Voyant, G. Notton, S. A. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting: A review," Renewable Energy, vol. 105, pp. 569–582, 2017, doi: 10.1016/j.renene.2016.12.095.
2. M. A. Inman, H. T. C. Pedro, and C. F. M. Coimbra, "Solar forecasting methods for renewable energy integration," Progress in Energy and Combustion Science, vol. 39, no. 6, pp. 535–576, 2013, doi: 10.1016/j.pecs.2013.06.002.
3. K. Barhmi, C. Heynen, S. Golroodbari, and W. van Sark, "A review of solar forecasting techniques and the role of artificial intelligence," Solar, vol. 4, no. 1, pp. 99–135, 2024, doi: 10.3390/solar4010005.
4. "Machine learning algorithms for solar irradiance prediction: A recent comparative study," e-Prime—Advances in Electrical Engineering, Electronics and Energy, vol. 7, Art. no. 100453, 2024, doi: 10.1016/j.prime.2024.100453.
5. D. Alves, F. Mendonça, S. S. Mostafa, and F. Morgado-Dias, "The potential of machine learning for wind speed and direction short-term forecasting: A systematic review," Computers, vol. 12, no. 10, Art. no. 206, 2023, doi: 10.3390/computers12100206.
6. 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, Art. no. 2457, 2023, doi: 10.3390/en16052457.
7. Y. Yang et al., "A survey on wind power forecasting with machine learning approaches," Neural Computing and Applications, vol. 36, pp. 12753–12773, 2024, doi: 10.1007/s00521-024-09923-4.
8. R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021. [Online]. Available:
9. A. M. Tashman, "Out-of-sample tests of forecasting accuracy: An analysis and review," International Journal of Forecasting, vol. 16, no. 4, pp. 437–450, 2000, doi: 10.1016/S0169-2070(00)00065-0.
10. L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
11. J. H. Friedman, "Greedy function approximation: A gradient boosting machine," The Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.
12. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
13. V. K. Saini et al., "Learning based short term wind speed forecasting models for smart grid applications: An extensive review and case study," Electric Power Systems Research, 2023, Art. no. 109502, doi: 10.1016/j.epsr.2023.109502.
14. H. Liu and Z. Zhang, "Development and trending of deep learning methods for wind power predictions," Artificial Intelligence Review, vol. 57, Art. no. 112, 2024, doi: 10.1007/s10462-024-10728-z.
15. O. Khouili et al., "Evaluating the impact of deep learning approaches on solar and photovoltaic power forecasting: A systematic review," Energy Strategy Reviews, vol. 59, Art. no. 101735, 2025, doi: 10.1016/j.esr.2025.101735.
16. "A review on data-driven methods for solar energy forecasting," Applied Energy, vol. 400, Art. no. 126631, 2025, doi: 10.1016/j.apenergy.2025.126631.
17. "Artificial intelligence for wind speed forecasting: A review on multi-scale decomposition and intelligent fusion strategies," Advances in Wind Engineering, vol. 2, no. 2, Art. no. 100055, 2025, doi: 10.1016/j.awe.2025.100055.
18. "Forecast evaluation for data scientists: Common pitfalls and best practices," Data Mining and Knowledge Discovery, 2023, doi: 10.1007/s10618-022-00894-5.