

Comparative assessment of statistical, Machine Learning, and Hybrid models for global solar irradiation forecasting in a tropical climate: A case study of Man (Côte d'Ivoire, West Africa)

Souleymane Tuo^{1,*}, Josée N'gbesso Yao¹, Mahama Bamba¹, Simon Abé Yapi², Boko Aka³

¹Faculty of Sciences and Technologies, Université de Man, Côte d'Ivoire, P.O. Box 20 Man

²Laboratory of technologies, Université Félix Houphouët Boigny de Cocody, Abidjan, Côte d'Ivoire, P.O. Box 34 Abidjan 01, Côte d'Ivoire

³Novel Energies Research Institute, Université Nangui Abrogoua, Abidjan, Côte d'Ivoire, P.O. 801 Abidjan, Côte d'Ivoire

Received: 9 February 2026 / Received in revised form: 20 April 2026 / Accepted: 15 June 2026

Abstract:

Accurate solar irradiation forecasting is essential for the reliable integration of photovoltaic energy into power systems, particularly in tropical regions where atmospheric conditions exhibit strong temporal variability. However, the nonlinear and multiscale nature of solar radiation dynamics makes forecasting performance highly dependent on the prediction horizon. This study investigates global solar irradiation forecasting in Man, Côte d'Ivoire, using a comparative framework involving statistical (ARMA), machine learning (K-NN, Random Forest, and XGBoost), and deep learning (LSTM and CNN-LSTM) models. Forecasts were developed for one-hour-ahead (H+1), one-day-ahead (D+1), and one-month-ahead (M+1) horizons using meteorological data from NASA's POWER database. Model performance was evaluated using RMSE, nRMSE, MAE, MAPE, and MBE. Feature importance analysis based on Random Forest and the Diebold–Mariano test were used to assess the contribution of input variables and compare model performance. The results reveal a strong scale-dependent forecasting behavior. For the H+1 and D+1 horizons, the CNN-LSTM model achieved the highest predictive accuracy, with nRMSE values of 9.84% and 13.15%, respectively, significantly outperforming the persistence benchmark ($p < 0.001$). At the M+1 horizon, the K-NN model provided the best performance, with an nRMSE of 4.49%. These findings demonstrate that no single model is universally optimal across temporal scales. Instead, forecasting performance is governed by the interaction among temporal resolution, atmospheric variability, data availability, and model complexity. The proposed scale-aware framework supports reliable solar forecasting and photovoltaic energy integration in tropical and sub-Saharan regions.

Keywords : Solar irradiation forecasting; Machine learning; Hybrid deep learning; Tropical climate; Photovoltaic energy; Man (Côte d'Ivoire).

* Corresponding author:

Email address: tuosouleyman@yahoo.fr (S. Tuo)

<https://doi.org/10.70974/mat10126119>

List of Abbreviations

Abbreviation	Definition
ARMA	Auto-Regressive Moving Average
AI	Artificial Intelligence
GHI	Global Horizontal Irradiance
k -NN	K Nearest Neighbors
LSTM	Long Short-Term Memory
CNN-LSTM	Convolutional Neural Network combined with LSTM
MAE/MAPE	Mean Absolute (Percentage) Error
MBE	Mean Bias Error
MSE/RMSE	Mean Squared Error Root Mean Squared Error
nRMSE	Normalized Root Mean Squared Error
MLP	Multilayer Perceptron
ReLU	Rectified Linear Unit
Tanh	Hyperbolic Tangent
NASA	National Aeronautics and Space Administration

1. Introduction

Solar energy is a renewable energy source of vital importance in combating climate change and reducing the dependence on fossil fuels [1]. Being abundant and accessible, solar energy has become an attractive alternative for power generation in many parts of the world. According to the International Energy Agency (IEA), the potential of solar energy is immense, with the capacity to supply up to 16% of the world's electricity production by 2050 [2]. In Africa, and particularly in Côte d'Ivoire, this challenge takes on critical dimensions. Historically dependent on hydropower and gas [3,4], the country is accelerating its energy transition with recent initiatives such as the inauguration of its first photovoltaic solar power plant in Boundiali [5] and the development of additional solar energy projects that have been announced [6]. However, the intermittency of solar irradiation poses significant challenges for its efficient integration into power grids [7,8]. The ability to accurately forecast solar radiation is essential for maximizing the efficiency of photovoltaic systems, minimizing operating costs, and improving power grid management. Accurate forecasts enable grid operators to better integrate variable solar power into the energy mix and ensure a stable power supply [9]. In addition, accurate prediction of solar irradiation is important for agricultural, weather, and climate applications, influencing crop planning, water management, and the study of the effects of climate change [10].

In Africa, solar irradiation prediction is of particular importance due to the continent's high solar potential and the need to improve access to electricity. Despite its enormous solar energy potential, Africa remains largely under-exploited in this field. Energy poverty and lack of access to electricity in many African countries are evi-

dence of this under-utilization of solar-powered systems. It is estimated that over 600 million people in Africa lack access to electricity [11]. Given the underdevelopment of solar energy in Africa, models that accurately predict solar irradiation are needed to encourage the use of solar energy on the continent.

Several studies have used different techniques to predict solar irradiation. For example, Huang *et al.* [12] combined techniques based on sky imagery and satellite data to predict solar irradiation, obtaining accurate results due to the use of high spatial and temporal resolution data.

Numerical weather prediction (NWP) models have also been widely used for solar irradiation prediction. Lorenz *et al.* [13] demonstrated the effectiveness of NWP models for solar irradiation forecasting in Europe using the WRF (Weather Research and Forecasting) model, producing hourly forecasts with high accuracy. However, these models require considerable computational resources and are sensitive to initial conditions.

Methods based on the time series formalism, such as intelligent persistence, Markov chains, and autoregressive moving average (ARMA) models, are also commonly used. Ahmed *et al.* [14] applied intelligent persistence to forecast short-term solar irradiance in Australia, showing significant improvements over simple persistence. Liao *et al.* [15] used Markov chains to forecast solar radiation in Germany, with promising results. Voyant *et al.* [16] used an ARMA model to forecast solar radiation in France, showing that this method could outperform persistence models under a variety of climatic conditions.

Machine learning models have gained popularity for solar irradiation prediction due to their ability to handle large datasets and capture complex non-linear relationships. For example, Breiman [17] introduced random forests, a technique that aggregates predictions from multiple decision trees to improve the robustness and accuracy of predictions. Chen and Guestrin [18] developed XGBoost, a boosting technique that has shown exceptional performance in predicting solar irradiance.

Artificial neural networks (ANNs), a subset of machine learning, are also widely used to predict solar irradiation. Kalogirou [19] used forward-propagating neural networks to estimate solar irradiance in Cyprus, showing a significant improvement over traditional methods. Recurrent neural networks (RNNs) and LSTMs (Long Short-Term Memory), introduced by Hochreiter and Schmidhuber [20], are designed to handle temporal sequences and capture long-term dependencies. These models have been adapted to predict solar radiation with increasing success, as illustrated by the study by Jayasankar *et al.* [21], who used a deep LSTM model to predict solar radiation in Chennai, India. Their approach, based on recurrent neural networks, showed a significant improvement in forecast accuracy by capturing complex relationships in weather data, outperforming traditional methods. Furthermore, Yadav and Chandel [22], in their 2014 review, examined artificial neural network (ANN)-based techniques to identify suitable methods available in the literature

for solar irradiation prediction. Their study shows that artificial neural network techniques offer superior accuracy compared with conventional methods.

However, it is important to note that local geographical and climatic characteristics significantly influence the performance of prediction models. For example, a model that performs well in a region with a stable climate may not perform well in a region with a variable climate [23]. Furthermore, the quality and quantity of historical data available for a specific region can also affect model accuracy [16]. Consequently, it is crucial to adapt prediction models to local conditions in order to obtain reliable and accurate results.

In this context, the aim of this study is not only to compare the performance of several solar irradiation forecasting models, but also to investigate how the forecasting horizon influences the relative effectiveness of statistical, machine learning, and deep learning approaches in a tropical environment. Seven forecasting models were selected for this purpose: ARMA, K-Nearest Neighbors (K-NN), Random Forest, XGBoost, Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM), and the hybrid Convolutional Neural Network–Long Short-Term Memory (CNN-LSTM) architecture. These models were chosen because they represent different levels of modeling complexity and learning paradigms, ranging from traditional statistical methods to advanced deep learning techniques.

The models were trained and evaluated using satellite-derived meteorological data obtained from NASA's POWER database. In addition to historical solar irradiation measurements, several meteorological variables were incorporated to assess the contribution of exogenous atmospheric information to forecasting performance. The analysis was conducted for three forecasting horizons: hourly (H+1), daily (D+1), and monthly (M+1). This multi-horizon framework makes it possible to examine how the predictive behavior of different models evolves as the temporal scale changes from highly dynamic short-term forecasting to seasonally driven long-term forecasting.

A key motivation of this work is that the physical processes governing solar irradiation variability differ substantially across temporal scales. Short-term forecasting is strongly influenced by rapidly changing atmospheric conditions such as cloud cover, humidity, and aerosol concentrations, whereas longer forecasting horizons are increasingly dominated by seasonal and climatological patterns. Consequently, the most appropriate forecasting model may vary according to the prediction horizon, making the identification of scale-adapted modeling strategies an important scientific and operational challenge.

To ensure a comprehensive assessment of model performance, forecasting accuracy was evaluated using multiple statistical metrics, confidence intervals derived from time-series cross-validation, and the

Diebold–Mariano test to determine whether differences in predictive accuracy were statistically significant. Furthermore, feature importance analysis was performed to identify the meteorological variables that contribute most to forecasting skill and to improve the interpretability of the proposed models.

This study contributes to the existing literature in three main ways. First, it provides a comprehensive comparison of statistical, machine learning, and deep learning approaches for solar irradiation forecasting in a tropical region of West Africa. Second, it investigates the scale-dependent nature of forecasting performance across hourly, daily, and monthly horizons. Third, it evaluates the impact of exogenous meteorological variables and statistically validates model superiority using rigorous hypothesis testing procedures.

By addressing these objectives, the study contributes to a better understanding of solar irradiation predictability in tropical climates and provides valuable insights for the development of reliable forecasting systems to support the integration of photovoltaic energy into power systems in Côte d'Ivoire and other sub-Saharan regions facing similar climatic conditions.

2. Materials and methods

2.1. Presentation of the study area

The town of Man, in Côte d'Ivoire, is located in the western part of the country, at latitude 7.4125° N and longitude -7.5542° W. This region is characterized by a varied landscape comprising mountains, tropical forests, agricultural plantations, and mountainous terrain with peaks reaching up to 1000 m in altitude. Figure 1 shows the location of the Man region in Côte d'Ivoire.

2.2. Data collection

This study relies exclusively on satellite-derived solar irradiation and meteorological data from NASA's POWER (Prediction of Worldwide Energy Resources) portal, due to the unavailability of ground-based measurements for the target region. The dataset includes key variables such as temperature, surface pressure, relative humidity, wind speed, and global solar radiation under both clear-sky and all-sky conditions.

Although satellite-derived datasets may introduce uncertainties compared to *in situ* measurements, their use is widely accepted in solar energy research, particularly in data-scarce regions such as sub-Saharan Africa. NASA POWER data are generated through a combination of satellite observations, reanalysis products, and validated radiative transfer models, providing physically consistent and globally homogeneous estimates of meteorological variables. This ensures temporal continuity, spatial coverage, and internal consistency across variables, which are essential properties for machine learning-based forecasting studies.

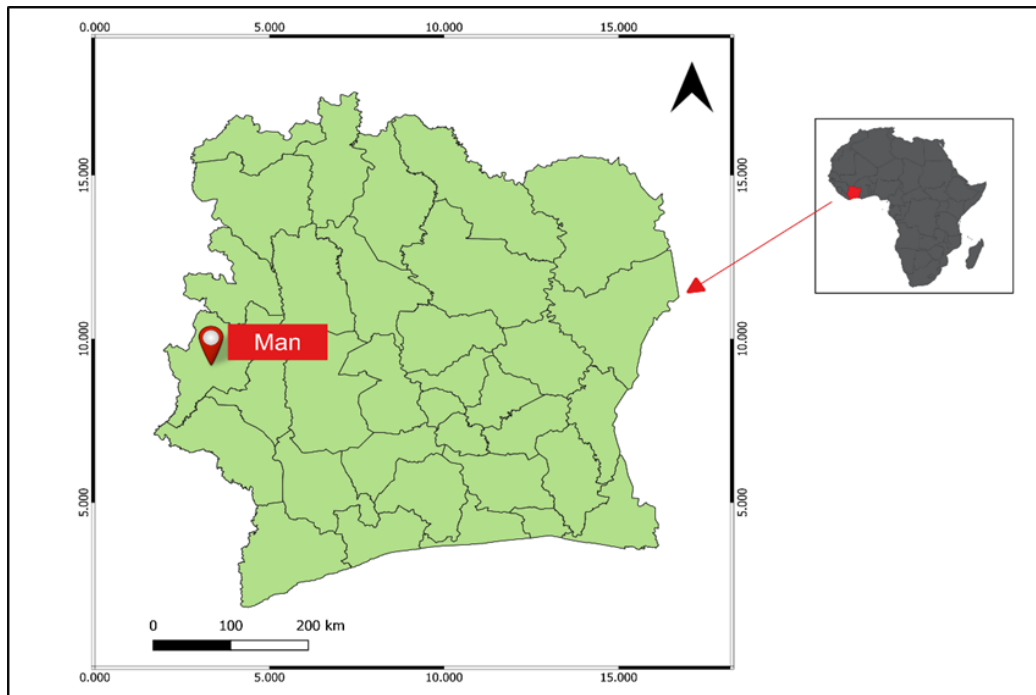


Fig. 1. Location of the Man region in Côte d'Ivoire (West Africa).

Furthermore, the objective of this work is comparative rather than one of absolute calibration. All forecasting models are trained and evaluated on the same dataset under identical conditions, ensuring a fair and unbiased benchmarking framework. In this context, potential systematic biases in satellite data do not affect the relative performance comparison between models, as these biases are equally propagated across all approaches. Therefore, the absence of ground-based measurements does not compromise the validity of the study, but rather reflects a realistic constraint commonly encountered in developing regions, while still allowing robust methodological conclusions.

To evaluate model robustness at different time scales, the data are structured into three forecasting horizons:

- **Hourly horizon (H+1):** Hourly data covering the period 2017–2020 are used for short-term forecasting. After restricting observations to daytime hours (06:00–18:00), the dataset contains approximately 17 520 samples.
- **Daily horizon (D+1):** Daily observations from 2000 to 2020 are used for medium-term forecasting, resulting in 7 665 samples.
- **Monthly horizon (M+1):** Monthly aggregated data over the period 2000–2020 are used for long-term forecasting, yielding 252 samples, of which 200 are used for training neural network-based models.

The temporal resolution of the data (hourly, daily, and monthly) makes it possible to compare model performance on different time scales, and thus to study the impact of data granularity on the accuracy of solar radiation forecasts.

2.3. Data pre-processing

Before applying the data to machine learning models for training, it is essential to carry out pre-processing to facilitate learning and improve model robustness. Pre-processing is tailored to the specific characteristics of the data in the context of time series modeling. In this study, the following steps were followed:

- **Data cleansing:** Raw data often contain missing values or outliers that can bias the model results. Records containing multiple missing values were identified and deleted. Outliers were detected using the boxplot method [24]. Over a period of twenty-one years, *i.e.*, a total of approximately 7 665 data points (365 days per year for 21 years), only 50 values were identified and considered outliers, representing less than 1% of the total dataset. Given their small number and the uncertainty regarding their nature, they were not removed. These values could be the result of extreme weather conditions and could contain important information for the prediction model. Consequently, they were retained in the dataset without further processing.
- **Data transformation:** To ensure the quality and relevance of the data used for model training, a standardization process was applied. This involves transforming the data so that it follows a normal distribution, with a mean of zero and a standard deviation of one. Standardization is an important step in data preparation, as it harmonizes the different scales of variables and facilitates the convergence of learning algorithms. The standardization formula is:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

where z is the standardized value, x is the original value, μ is the mean of the original values, and σ is the standard deviation of the original values.

- **Selection of input features:** The selection of input variables is a key step in the construction of efficient prediction models, as it reduces model complexity while retaining the most influential variables. Several methods have been proposed in the literature for performing this selection. For example, LASSO (Least Absolute Shrinkage and Selection Operator) regression is a popular method that penalizes regression coefficients to favor the selection of the most significant variables, thereby improving accuracy while reducing the risk of overfitting [25]. Decision trees are also widely used for feature selection, as they identify the most influential variables based on their ability to partition the data optimally [17]. In addition, the WEKA tool, which applies machine learning techniques, is commonly used for variable selection, particularly for models based on neural networks and decision trees. Yadav *et al.* [26] demonstrated the effectiveness of WEKA for input variable selection in solar radiation prediction models, using filtering and optimization algorithms to assess variable importance.

In this study, the Random Forest method was used to select input features. This approach involves training the Random Forest algorithm on all available features. Each input variable was evaluated in terms of its contribution to reducing prediction error. The variables with the highest importance scores were selected as the most significant input variables for the prediction models. According to Genuer *et al.* [27], this method is particularly effective for identifying the most important variables, as it enables the processing of large datasets while reducing the risk of overfitting.

By applying this method, the relevant features for the prediction of global solar irradiation were identified, thus improving model accuracy.

For this study, lag variables were created for the previous ten days for a set of meteorological variables as well as for the target variable, global solar irradiation (ALLSKY_SFC_SW_DWN). The variables used include: ALLSKY_SFC_SW_DWN (Global Solar Irradiation, $\text{kW h m}^{-2} \text{ d}^{-1}$), ALLSKY_SFC_KT (Clarity Index, dimensionless), ALLSKY_SFC_UV_INDEX (Surface UV Index, dimensionless), T2M (Temperature at 2 m, $^{\circ}\text{C}$), TS (Terrestrial Surface Temperature, $^{\circ}\text{C}$), PRECTOTCORR (Corrected Precipitation, mm d^{-1}), QV2M (Specific Humidity at 2 m, g kg^{-1}), RH2M (Relative Humidity at 2 m, %), PS (Surface Pressure, kPa), WS10M (Wind Speed at 10 m, m s^{-1}), WD10M (Wind Direction at 10 m, degrees), and WS2M (Wind Speed at 2 m, m s^{-1}).

A Random Forest selection method was then ap-

plied to identify the variables whose past measurements are most relevant to the prediction of global solar radiation (ALLSKY_SFC_SW_DWN).

- **Data stationarization:** To deal with seasonal variations and long-term trends in global solar irradiation, the data were stationarized using the clear-sky index methodology. The clear-sky index K_C is defined as the ratio between solar irradiance measured on the ground and the radiation modeled under clear-sky conditions [28]. For global radiation (R_g):

$$K_C = \frac{R_g^{\text{measured}}}{R_g^{\text{clear-sky model}}} \quad (2)$$

This index enables the generation of a new time series suitable for most forecasting models based on the study of time series and the modeling of stochastic processes [28]. In this stationarization methodology, the K_C index is predicted and then post-processed to recover the desired quantity, *i.e.*, global irradiation.

- **Lag variables and temporal features:** To capture the temporal dynamics of time series, lag variables were created. These lag variables enable models to take into account past values of input features, which is particularly relevant for LSTM and ARMA models that exploit temporal dependencies.

2.4. Prediction models

2.4.1. Long Short-Term Memory (LSTM)

LSTM is a type of recurrent neural network (RNN) capable of learning long-term dependencies in time series. It is particularly effective for modeling sequential data and handling vanishing gradient problems, through its unique design with memory cells [20].

2.4.2. Multi-Layer Perceptron (MLP)

The Multi-Layer Perceptron (MLP) is a type of artificial neural network that uses hidden layers between inputs and outputs to capture complex relationships between variables. It is widely used for various prediction tasks due to its flexibility and ability to learn non-linear representations [29].

2.4.3. K-Nearest Neighbors (K-NN)

K-NN is a non-parametric algorithm that predicts samples according to the k nearest neighbors in feature space. It is simple to understand and implement, and proves effective for classification and regression tasks with moderately sized datasets [30].

2.4.4. XGBoost

XGBoost is an optimized implementation of the gradient boosting algorithm, which improves model performance by combining several weak decision trees to form a robust model. It is known for its speed and efficiency in handling large datasets and for achieving top performance in data science competitions [18].

2.4.5. Random Forest

Random Forest is an ensemble of decision trees, where each tree is built from a bootstrap sample of the original dataset. It uses the average of individual tree predictions to improve accuracy and reduce overfitting [17].

2.4.6. AutoRegressive Moving Average (ARMA)

The ARMA model combines autoregressive (AR) components and moving average (MA) components to model stationary time series. It is widely used in economics and meteorology to predict future trends based on past values and residual errors [31].

2.4.7. CNN-LSTM

The CNN-LSTM model combines the deep learning capabilities of Convolutional Neural Networks (CNNs) to extract local features from data with the Long Short-Term Memory (LSTM) network's ability to capture temporal dependencies. This model is particularly effective for complex spatiotemporal time series [32].

2.4.8. Persistence Model

The persistence model is one of the simplest baseline approaches used in time series forecasting. It assumes that the future value of a variable is equal to its most recently observed value. In mathematical terms, the prediction at time $t + 1$ is set equal to the value at time t . Despite its simplicity, this model is often difficult to outperform in very short-term forecasting problems and therefore serves as a strong reference benchmark for evaluating more advanced predictive models.

2.4.9. Mean Model

The mean model is another basic baseline approach in time series forecasting. It predicts future values as the arithmetic mean of all historical observations available in the training dataset. This approach assumes that the series is stationary around a constant average value and does not account for temporal dynamics or trends. Although simplistic, the mean model provides a useful reference to assess whether more complex models are able to capture meaningful patterns beyond a constant-level prediction.

2.5. Model development

The prediction models were developed following specific architectures adapted to the characteristics of the data. For each model, with the exception of the ARMA model, the parameters were tuned using Sklearn's Grid-SearchCV method. This approach optimizes the hyperparameters by testing different combinations to find the one offering the best predictive performance [33].

2.5.1. Model architecture

a. Long Short-Term Memory (LSTM)

To model the temporal dynamics of solar irradiation, an LSTM network was implemented. This architecture is well-suited for time series forecasting due to its ability to capture long-term dependencies through gated memory mechanisms [20, 34].

The proposed model consists of two stacked LSTM layers, each with 50 units, followed by a dense out-

put layer. The network was trained using the Mean Squared Error (MSE) loss function and optimized with the Adam algorithm [35]. A combination of L1 and L2 regularization was applied to improve generalization [36]. Training was performed over 20 epochs.

LSTMs are composed of cells capable of retaining information over long periods through special gating units. The mathematical expressions associated with LSTM units are:

- **Forget gate (F_t):** This gate decides which information from the previous state should be forgotten.

$$F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f) \quad (3)$$

- **Input gate (I_t):** This gate determines which new information will be stored in the cell state.

$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \quad (4)$$

- **Memory modulation gate ($\tilde{C}_t$):** This gate creates new candidate values that can be added to the cell state.

$$\tilde{C}_t = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \quad (5)$$

- **Cell state (C_t):** The cell state is updated by combining the old cell state, modulated by the forget gate, with the new candidate cell state, modulated by the input gate.

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (6)$$

- **Output gate (O_t):** This gate determines the next hidden state.

$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \quad (7)$$

- **Hidden state (H_t):** The hidden state is updated based on the cell state and the output gate.

$$H_t = O_t \odot \tanh(C_t) \quad (8)$$

where X_t is the input at time t ; H_{t-1} is the hidden state at time $t - 1$; W_{xi} , W_{xf} , W_{xo} , and W_{xc} are the weights associated with the input; W_{hi} , W_{hf} , W_{ho} , and W_{hc} are the weights associated with the hidden state; and b_i , b_f , b_o , and b_c are the biases.

These equations show how LSTMs manage information at each time step, allowing for accurate modeling of time series such as global solar irradiation.

To implement these equations, the Keras library, specialized in deep learning, was used. Keras provides tools to create and train advanced neural networks by leveraging backend frameworks such as TensorFlow, CNTK, or Theano [18].

Hyperparameter Tuning: Model performance was improved through a grid search over key hyperparameters:

- Activation function: {ReLU, Tanh, Sigmoid}

- LSTM units: {32, 50, 64, 100}
- Learning rate: {0.0001, 0.001, 0.01}
- Batch size: {16, 32}

The optimal configuration obtained is: activation: Tanh; LSTM units: (50, 50); learning rate: 0.01; batch size: 16.

b. Multi-Layer Perceptron (MLP)

An MLP was employed as a baseline model to capture non-linear relationships between input variables and solar radiation. The MLP is a feedforward neural network composed of fully connected layers, widely used for regression tasks.

The initial architecture considered in this study consisted of a single hidden layer with 50 neurons. To improve model performance, a grid search procedure was subsequently conducted to explore a range of architectural and training hyperparameters.

Hyperparameter Tuning: The following hyperparameter space was investigated:

- Hidden layer sizes: {(50), (100), (50, 50), (100, 50), (100, 100)}
- Activation functions: {ReLU, Tanh, Sigmoid}
- Solver: {Adam, SGD}
- Regularization parameter (α): {0.00001, 0.0001, 0.001, 0.01}
- Initial learning rate: {0.0001, 0.001, 0.01}

The optimal configuration obtained is: hidden layer size: (50, 50); activation function: Sigmoid; solver: Adam; regularization parameter (α): 0.01; learning rate: 0.01.

c. K-Nearest Neighbors (K-NN)

The K-NN algorithm is characterized by a major parameter, k . The algorithm selects the k nearest neighbors of each observation to predict solar radiation. Euclidean distance is used as a measure of similarity [30].

Initially, a fixed number of neighbors was considered. However, to improve predictive performance, a grid search combined with a 5-fold time series cross-validation procedure was conducted to identify the optimal hyperparameters.

Hyperparameter Tuning: The following parameter space was explored:

- Number of neighbors (k): {3, 5, 7, 9, 11, 13, 15, 20}
- Weight function: {uniform, distance}
- Distance metric parameter (p): {1, 2}

The parameter $p = 2$ corresponds to the Euclidean distance, while $p = 1$ corresponds to the Manhattan distance. The best-performing configuration obtained is: $k = 10$; weights: uniform; distance metric: Euclidean ($p = 2$).

d. XGBoost

The Extreme Gradient Boosting (XGBoost) algorithm was used as a powerful ensemble learning method for regression. XGBoost is based on gradient boosting of decision trees and is particularly effective for capturing complex non-linear relationships while controlling overfitting through regularization.

Hyperparameter Tuning: The following hyperparameter space was explored:

- Number of trees (**n_estimators**): {50, 100, 200, 300}
- Learning rate: {0.01, 0.05, 0.1, 0.2}
- Maximum tree depth: {3, 5, 7, 9}
- Minimum child weight: {1, 3, 5}
- Subsample ratio: {0.6, 0.8, 1.0}
- Column sampling (**colsample_bytree**): {0.6, 0.8, 1.0}
- L1 regularization (**reg_alpha**): {0, 0.1, 1, 10}
- L2 regularization (**reg_lambda**): {0, 0.1, 1, 10}

The optimal hyperparameters obtained are: **n_estimators**: 100; learning rate: 0.05; max depth: 3; min child weight: 5; subsample: 0.8; **colsample_bytree**: 1.0; **reg_alpha**: 0; **reg_lambda**: 0.1.

e. Random Forest

The Random Forest model used in this study comprises 50 trees. Each tree is built using a subset of the features, and hyperparameters such as the number of trees and maximum depth are optimized with GridSearchCV. The Gini impurity criterion is used to measure partitioning quality. This approach has proved effective in numerous environmental prediction studies [37].

Hyperparameter Tuning: The following parameter space was explored:

- Number of trees: {50, 100, 200}
- Maximum depth: {None, 10, 20, 30}
- Minimum samples split: {2, 5, 10}
- Minimum samples per leaf: {1, 2, 4}
- Maximum features: {sqrt, log2, None}

The best-performing hyperparameters are: number of trees: 50; maximum depth: None; minimum samples split: 10; minimum samples per leaf: 4; maximum features: sqrt.

f. Convolutional Neural Network combined with LSTM (CNN-LSTM)

A hybrid CNN-LSTM architecture was implemented to jointly capture local patterns and temporal dependencies in solar radiation time series. This type of model has demonstrated strong performance in sequential data analysis by combining convolutional feature extraction with recurrent learning capabilities.

The final architecture consists of the following components: a Conv1D layer with 32 filters, a kernel size of

2, and ReLU activation, used to extract local temporal features from the input sequences; an LSTM layer with 50 units, responsible for modeling long-term dependencies; and a fully connected output layer with a single neuron for regression.

Hyperparameter Tuning: The following parameter space was explored:

- Number of filters: {32, 64}
- Kernel size: {2, 3}
- LSTM units: {50, 100, 150}
- Learning rate: {0.0001, 0.001, 0.01}

The model is compiled with the MSE loss function and uses the Adam optimizer with a learning rate of 0.001. The model is trained over 20 epochs.

g. AutoRegressive Moving Average (ARMA)

In this study, an ARMA model was employed. Following the verification of stationarity of the solar radiation time series using the Augmented Dickey–Fuller (ADF) and KPSS tests, an ARIMA(p, d, q) modeling framework was adopted. Given that the series is stationary in level, the differencing order was set to $d = 0$, reducing the model to an ARMA(p, q) specification.

The model identification procedure was conducted in two stages. First, a set of autoregressive models AR(p) was estimated by varying p within the range [5, 20], while fixing $q = 0$. For each specification, the Akaike Information Criterion (AIC), introduced by Akaike [38], was computed, and the optimal order p was selected as the value minimizing the AIC.

In the second stage, the selected value of p was fixed, and the moving average order q was varied within the interval [0, 10] to estimate a set of ARMA(p, q) models. The AIC was again used to assess whether the inclusion of moving average terms improves model performance.

Model selection was based on the minimization of the AIC, while adhering to the principle of parsimony. The final model was subsequently subjected to diagnostic checking and used for forecasting purposes. The corresponding results are presented in Section 3..

2.5.2. Training process

The dataset was split chronologically to preserve temporal dependencies inherent in solar radiation time series. The final partition consists of 70% training data, 20% validation data for hyperparameter tuning, and 10% independent test data. This temporal split ensures that future information is not used during model training, thereby preventing data leakage.

To ensure robust model evaluation and reduce overfitting risk, cross-validation was performed using a 5-fold time series cross-validation strategy [39]. The training data were split into five sequential folds, where each fold was in turn used as a validation set while the preceding folds were used for training. This process was repeated until each fold had been used once as the validation set.

Importantly, the test set (10%) was strictly isolated before any pre-processing, feature selection, or model

training steps. This guarantees an unbiased evaluation of the final model performance on unseen data.

2.6. Model evaluation and statistical validation

To evaluate the performance of different models for global solar radiation prediction, several commonly used metrics were applied. These metrics quantify the accuracy of each model by measuring the difference between predicted and observed values. The metrics used are:

- **Root Mean Squared Error (RMSE):** RMSE is the square root of the MSE. It quantifies the amplitude of the error while taking into account large deviations. This metric is particularly useful in the context of solar radiation prediction, where large errors can have a significant impact on results [40].

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (9)$$

where Y_i is the observed value, $\hat{Y}_i$ is the predicted value, and n is the total number of observations.

- **Normalized Root Mean Squared Error (nRMSE):** The nRMSE is the RMSE normalized by the mean of the observed values, allowing easier comparison between different datasets or models.

$$\text{nRMSE} = \frac{\text{RMSE}}{\bar{Y}} \times 100 \quad (10)$$

- **Mean Absolute Error (MAE):** This metric calculates the average of the absolute errors between predictions and actual values. It provides a more robust measure with respect to outliers than the RMSE and gives a clear indication of the average error, which is relevant to assessing the overall accuracy of the model in practical application contexts [41].

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (11)$$

- **Mean Absolute Percentage Error (MAPE):** The MAPE measures the error as a percentage, thus providing a normalized view of the error. This metric is useful to understand the magnitude of the error in relation to the actual values.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100 \quad (12)$$

- **Mean Bias Error (MBE):** The MBE measures the average bias of predictions, indicating whether predictions are generally underestimated or overestimated. A positive MBE value means that, on average, the cumulative predicted data overestimates the cumulative actual measurements,

while a negative value indicates the opposite.

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i) \quad (13)$$

- **Performance gain (relative improvement in nRMSE):** To better assess the improvement of advanced models over reference approaches, the relative gain in nRMSE is computed as follows:

$$\text{Gain}(\%) = \frac{\text{nRMSE}_{\text{ref}} - \text{nRMSE}_{\text{model}}}{\text{nRMSE}_{\text{ref}}} \times 100 \quad (14)$$

where $\text{nRMSE}_{\text{ref}}$ corresponds to the benchmark model (mean or persistence), and $\text{nRMSE}_{\text{model}}$ corresponds to the evaluated model.

- **Confidence intervals:** To assess the reliability and robustness of the performance metrics obtained through the 5-fold time series cross-validation procedure, 95% confidence intervals were computed using the Student's t -distribution. Let $x_1, x_2, \dots, x_n$ denote the metric values obtained across the validation folds. The confidence interval around the sample mean is calculated as:

$$\text{CI}_{95\%} = \bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} \quad (15)$$

where $\bar{x}$ is the sample mean, s is the sample standard deviation, n is the number of folds, and $t_{\alpha/2, n-1}$ is the critical value of the Student's t -distribution with $n - 1$ degrees of freedom.

The lower and upper bounds of the confidence interval are given by:

$$\text{CI}_{\text{lower}} = \bar{x} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} \quad (16)$$

$$\text{CI}_{\text{upper}} = \bar{x} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} \quad (17)$$

A confidence level of 95% was adopted throughout this study.

3. Results and discussion

3.1. Stationarity analysis of the time series

The stationarity of the solar radiation time series was assessed using two complementary statistical tests: the Augmented Dickey–Fuller (ADF) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. The results are summarized in [table 1](#).

The ADF test yields a test statistic of -7.97 , which is significantly lower than the critical values at the 1%, 5%, and 10% significance levels (-3.43 , -2.86 , and -2.57 , respectively). The associated p -value is extremely small ($\approx 2.77 \times 10^{-12}$), well below the 5% significance threshold. Therefore, the null hypothesis of non-stationarity is strongly rejected, indicating that the series is stationary according to the ADF test.

In contrast, the KPSS test produces a test statistic of 0.066 , which is lower than all critical values at the 10%, 5%, 2.5%, and 1% levels (0.347 , 0.463 , 0.574 ,

and 0.739 , respectively), with a p -value of 0.1 . Consequently, the null hypothesis of stationarity cannot be rejected, further confirming the stationarity of the series.

The consistency between both tests, despite their opposite null hypotheses, provides strong evidence of stationarity. While the ADF test rejects the presence of a unit root, the KPSS test supports the assumption of stationarity.

Overall, the results reported in [table 1](#) indicate that the solar radiation series is stationary in level, implying constant statistical properties over time, such as mean and variance. Therefore, no transformation such as differencing is required prior to time series modeling.

3.2. ARMA Model identification and selection

The identification of the optimal model was carried out in two stages based on the analysis of the Akaike Information Criterion (AIC). First, as shown in [figure 2](#), the evolution of the AIC was examined as a function of the autoregressive order p for $\text{AR}(p)$ models with $q = 0$. A decreasing trend in AIC was observed as p increased, indicating an improvement in model fit.

However, the rate of decrease became negligible beyond $p = 10$, suggesting that additional parameters do not significantly enhance model performance. Based on the minimum AIC and the principle of parsimony, the optimal autoregressive order was selected as $p = 10$.

In the second stage, as illustrated in [figure 3](#), the value of $p = 10$ was fixed, and the AIC was evaluated as a function of the moving average order q . A substantial decrease in AIC was observed between $q = 0$ and $q = 1$, indicating that the inclusion of a first-order moving average component significantly improves the model. For higher values of q , the AIC shows only minor variations, although the minimum value is reached at $q = 7$.

However, the improvement in AIC between $q = 1$ and $q = 7$ remains limited relative to the increase in model complexity. In accordance with the principle of parsimony, a lower-order model is preferred when it provides comparable performance. Therefore, the model with $q = 1$ was retained as a suitable compromise between goodness-of-fit and simplicity.

Based on these results, the selected model is $\text{ARMA}(10, 1)$.

3.3. Selection of variables

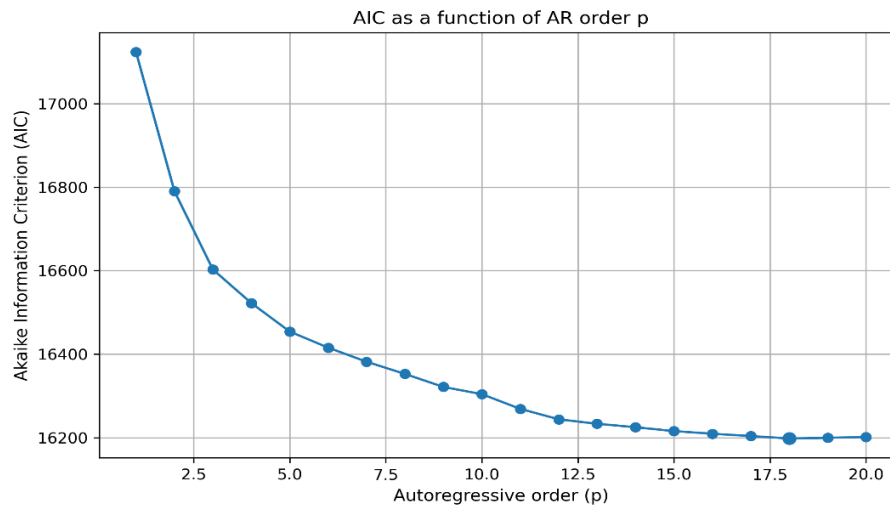
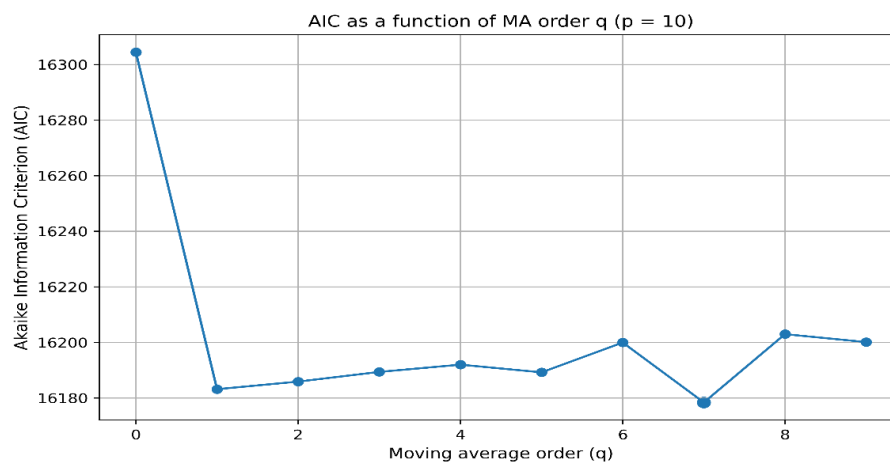
The Random Forest feature selection procedure identified six variables: WD10M (Wind Direction at 10 m), ALLSKY_SFC_KT (Clarity Index), PRECTOTCORR (Corrected Precipitation), ALLSKY_SFC_UV_INDEX (Surface UV Index), ALLSKY_SFC_SW_DWN (Global Solar Radiation), and PS (Surface Pressure). Among these, global solar radiation (ALLSKY_SFC_SW_DWN) and wind direction at 10 m (WD10M) were the most important.

[Table 2](#) presents the performance of the Random Forest model using different combinations of variables: only the previous values of global solar radiation, all available variables, and the variables selected by Random Forest.

Table 1

Summary of stationarity test results.

Test	Statistic	p -value	Null hypothesis	Decision	Conclusion
ADF (Augmented Dickey–Fuller)	−7.97	2.77×10^{-12}	Non-stationarity	Reject H_0	Series is stationary
KPSS	0.066	0.10	Stationarity	Fail to reject H_0	Series is stationary

**Fig. 2.** Evolution of the Akaike Information Criterion (AIC) as a function of the autoregressive order p for $AR(p)$ models (with $q = 0$).**Fig. 3.** Evolution of the Akaike Information Criterion (AIC) as a function of the moving average order q for $ARMA(p, q)$ models (with $p = 10$).**Table 2**

Comparison of Random Forest model performance at the H+1 horizon according to the variables used.

Configuration	nRMSE	MAPE
Univariate (ALLSKY SFC_SW_DWN)	11.5	8.71
All available variables	11.16	8.42
Selected variables	11.08	8.34
Values are expressed in %		

The results show that the use of all variables improves performance compared to a univariate approach. However, the selection of variables using the Random Forest method provides slightly better performance (nRMSE = 11.08%, MAPE = 8.34%) while reducing model complexity. Figure 4 illustrates the relative importance of the selected variables, while table 3 presents their corresponding importance values in percentage.

Table 3

Relative importance of input variables in the predictive model.

Variable	Importance (%)
ALLSKY_SFC_SW_DWN	32.0
WD10M	9.7
PS	8.5
ALLSKY_KT	7.5
ALLSKY_SFC_UV_INDEX	7.5
PRECTOTCORR	7.0
WS10M	6.0
WS2M	5.5
T2M	4.0
QV2M	3.5
RH2M	3.4
TS	3.3

3.4. Performance of the models

The performance of different models for global solar radiation prediction was evaluated and compared over three time horizons: H+1 (one hour ahead), D+1 (one day ahead), and M+1 (one month ahead). The metrics used include RMSE, nRMSE, MAE, MAPE, and MBE. The results are reported in [table 4](#). The top three values of nRMSE and MAPE for each horizon are highlighted: green for the first, yellow for the second, and red for the third.

3.4.1. H+1 Horizon

For the H+1 horizon, the CNN-LSTM hybrid model showed the best performance with an RMSE of 55.18 W h m^{-2} , an nRMSE of 9.84%, and an MAE of 40.96 W h m^{-2} . The LSTM model followed closely with an RMSE of 56.08 W h m^{-2} and a MAE of 40.22 W h m^{-2} . The MLP and Random Forest models also showed good performance, but slightly lower, with RMSE of 58.46 W h m^{-2} and 59.37 W h m^{-2} , respectively. The ARMA and K-NN models achieved the lowest performance, with RMSE of 66.92 W h m^{-2} and 63.69 W h m^{-2} . The LSTM and CNN-LSTM model forecasts for the H+1 horizon are shown in [figure 5](#).

3.4.2. D+1 Horizon

For the D+1 horizon, the CNN-LSTM model once again outperforms the other models, with an RMSE

of 658.1 W h m^{-2} and an nRMSE of 13.15%. The performances of the LSTM and XGBoost models were relatively close, with nRMSE of 13.77% and 13.92%, respectively. The K-NN and Random Forest models had the highest errors, with nRMSE of 14.31% and 14.24%, respectively.

The LSTM and CNN-LSTM model forecasts for the D+1 horizon are illustrated in [figure 6](#).

3.4.3. M+1 Horizon

For the M+1 horizon, the K-NN model stands out from the LSTM and CNN-LSTM hybrid models by achieving the lowest values of RMSE, nRMSE, MAE, and MAPE. This may be due to the size of the training sample, limited to 200 data points, which could have hampered the ability of LSTM and CNN-LSTM models to generalize effectively. Indeed, neural networks can be sensitive to the amount of data, especially when it is insufficient to allow optimal learning.

The K-NN achieved an RMSE of 225.9 W h m^{-2} , an nRMSE of 4.49%, and an MAE of 188.6 W h m^{-2} . Random Forest follows with an RMSE of 235.2 W h m^{-2} and an nRMSE of 4.6%. XGBoost and LSTM models also showed competitive performance with nRMSE of 5.05% and 5.02%, respectively, while ARMA and MLP models showed the highest errors with RMSE of 275.6 W h m^{-2} and 263.4 W h m^{-2} , respectively.

The LSTM, CNN-LSTM, and K-NN forecasts for the M+1 horizon are shown in [figure 7](#).

3.4.4. Performance analysis compared to Benchmark models

After comparing the different forecasting models, it is relevant to assess their added value against commonly used benchmark approaches in solar radiation forecasting. For each forecasting horizon (H+1, D+1, and M+1), the statistical ARMA model as well as the best-performing machine learning and deep learning models identified earlier are compared with reference models based on historical mean and persistence. This comparison makes it possible to quantify the performance gains achieved by advanced methods and to evaluate their ability to improve forecasting accuracy compared to simpler approaches ([Table 5](#)).

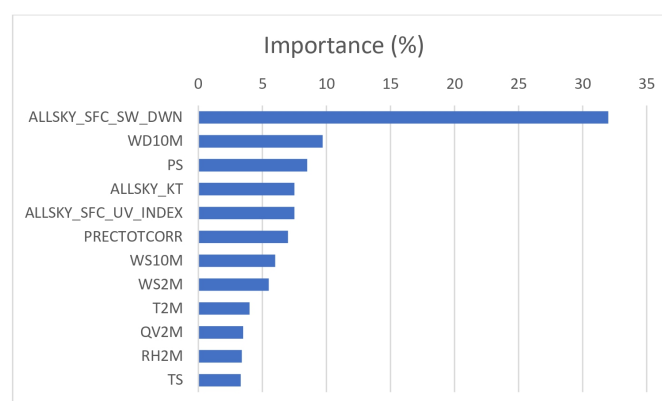


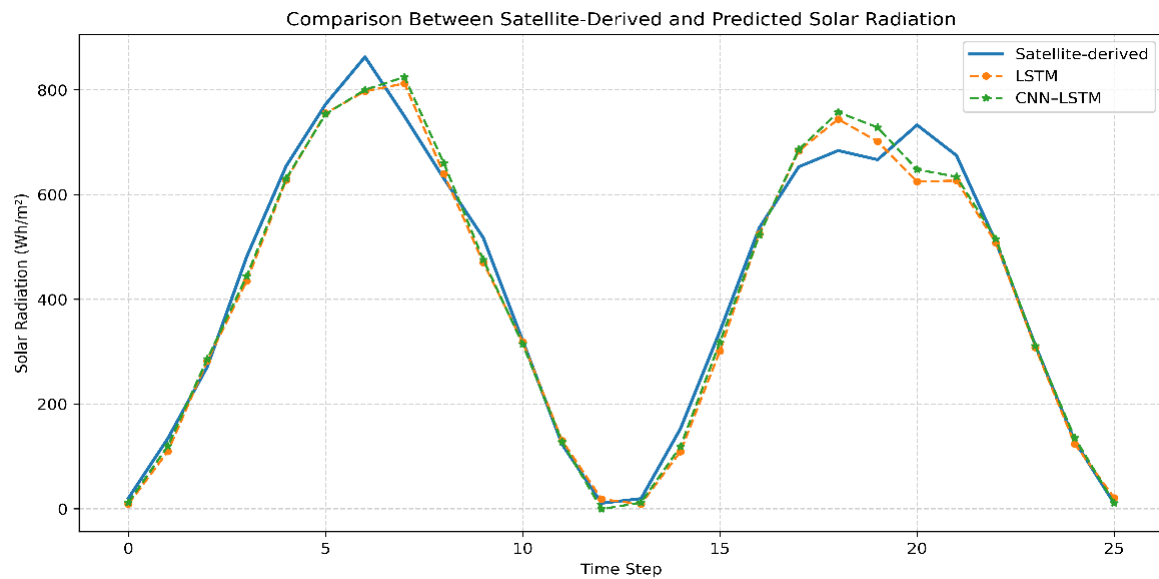
Fig. 4. Relative importance of each variable selected by the Random Forest algorithm.

Table 4

Forecasting errors by horizon for the different models.

Model	RMSE $\pm$ 95% CI (5-fold CV)	nRMSE (%)	MAE	R^2	MBE
H+1 Horizon					
K-NN	63.69 $\pm$ 0.48	11.89	46.29	0.902 $\pm$ 0.001	-3.1
ARMA	66.92 $\pm$ 0.59	12.49	48.46	0.877 $\pm$ 0.004	-1.789
XGBoost	60.16 $\pm$ 0.49	11.23	43.82	0.912 $\pm$ 0.003	-1.196
R_FOREST	59.37 $\pm$ 0.45	11.08	40.82	0.943 $\pm$ 0.004	4.32
MLP	58.46 $\pm$ 0.44	10.91	40.72	0.930 $\pm$ 0.003	0.946
LSTM	56.08 $\pm$ 0.44	10.47	40.22	0.937 $\pm$ 0.003	-12.785
CNN-LSTM	55.18 $\pm$ 0.74	9.84	40.96	0.951 $\pm$ 0.004	0.264
D+1 Horizon					
K-NN	715.5 $\pm$ 4.35	14.31	551.3	0.287 $\pm$ 0.006	38.7
ARMA	700.7 $\pm$ 6.22	14.02	536.5	0.345 $\pm$ 0.009	-0.4
XGBoost	696.4 $\pm$ 7.87	13.92	533.38	0.328 $\pm$ 0.007	0.0001
R_FOREST	711.8 $\pm$ 4.69	14.24	543.9	0.331 $\pm$ 0.009	0.0252
MLP	711.1 $\pm$ 3.44	14.23	546.7	0.295 $\pm$ 0.010	82.7
LSTM	688.2 $\pm$ 3.93	13.77	529.9	0.340 $\pm$ 0.008	-11.5
CNN-LSTM	658.1 $\pm$ 3.96	13.15	502.5	0.350 $\pm$ 0.008	39.6
M+1 Horizon					
K-NN	225.9 $\pm$ 3.1	4.49	188.6	0.678 $\pm$ 0.018	0.2
ARMA	275.6 $\pm$ 4.3	5.45	232.8	0.459 $\pm$ 0.016	13.1
XGBoost	254.5 $\pm$ 2.3	5.05	207.7	0.591 $\pm$ 0.036	-17.4
R_FOREST	235.2 $\pm$ 2.2	4.60	197.5	0.660 $\pm$ 0.013	-3.3
MLP	263.4 $\pm$ 2.0	5.23	216.3	0.630 $\pm$ 0.029	9.1
LSTM	253.7 $\pm$ 2.1	5.02	296.6	0.636 $\pm$ 0.018	-14.7
CNN-LSTM	251.0 $\pm$ 2.4	4.98	197.5	0.630 $\pm$ 0.035	-22.7

RMSE, MAE and MBE are expressed in W h m^{-2} . CI: confidence interval; RMSE: root mean square error; MAE: mean absolute error; MBE: mean bias error.

**Fig. 5.** Comparison between satellite-derived and predicted solar irradiation for the hourly horizon (H+1).

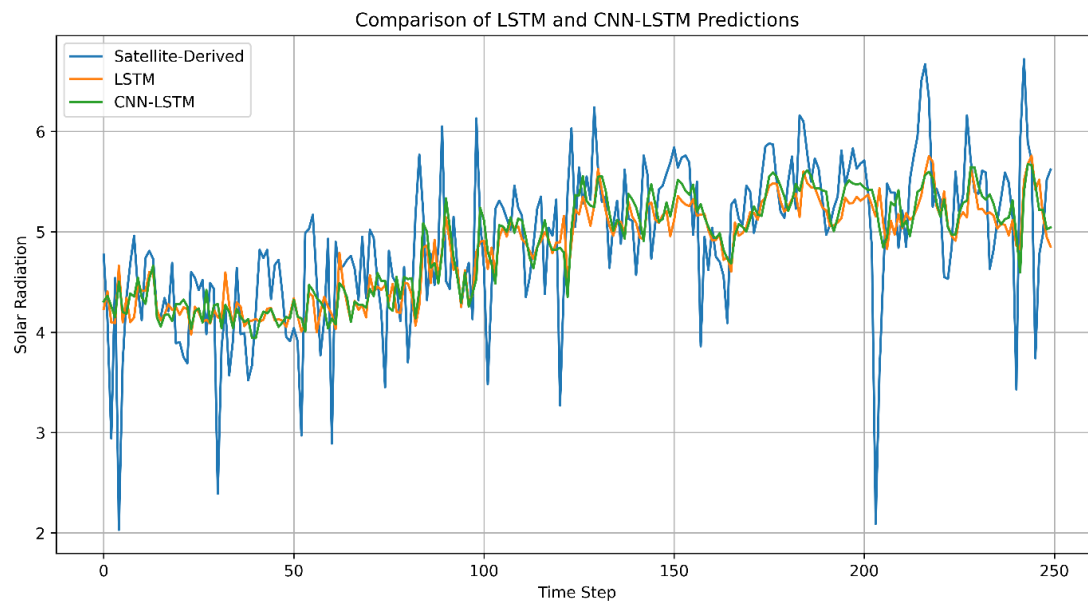


Fig. 6. Comparison between satellite-derived and predicted solar irradiation for LSTM and CNN-LSTM models for the daily horizon (D+1).

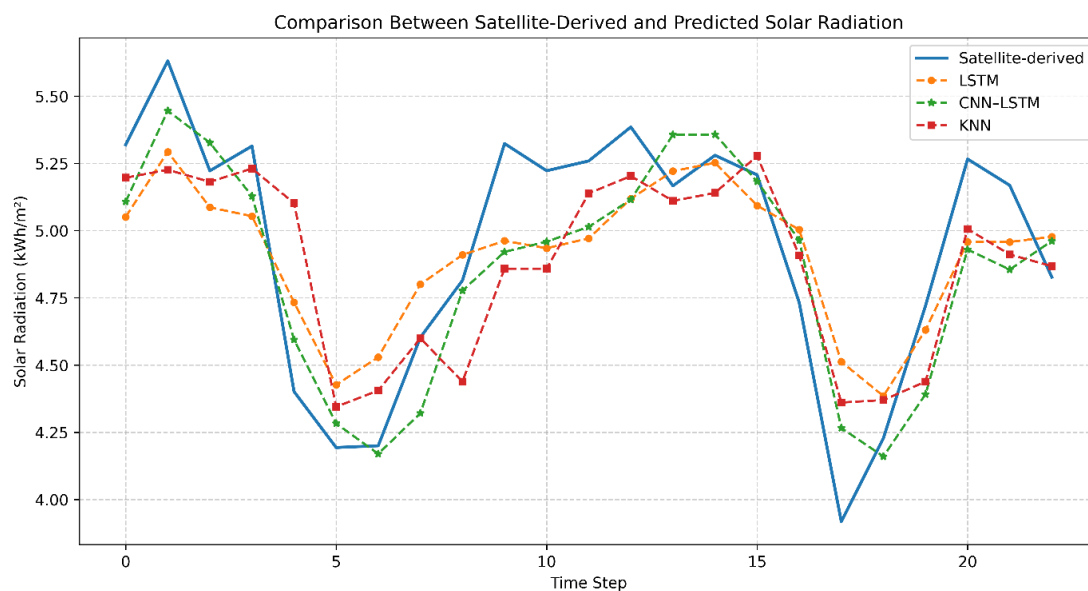


Fig. 7. Comparison between satellite-derived and predicted solar irradiation for LSTM, CNN-LSTM, and K-NN models for the monthly horizon (M+1).

The results in [table 5](#) clearly demonstrate the advantage of the advanced models used in this study compared to benchmark models (historical mean and persistence), regardless of the forecasting horizon.

At the H+1 horizon, the improvements are the most significant. The CNN-LSTM model achieves a strong reduction in errors, with a 78.1% improvement over the mean model and 69.6% over the persistence model. ARMA and Random Forest also confirm substantial reductions exceeding 60%, highlighting their ability to effectively capture short-term temporal dependencies.

At the D+1 horizon, performance remains superior, although the gains become more moderate. CNN-LSTM improves results by 16.1% compared to the mean and 24.4% compared to persistence, while other models

show improvements generally ranging between 10% and 20%.

At the M+1 horizon, advanced models still maintain a notable advantage. K-NN achieves up to a 45.3% reduction in error compared to the mean and 40.2% compared to persistence, while models such as ARMA and CNN-LSTM still exceed 27% improvement.

Overall, advanced models consistently outperform benchmark approaches, with the highest performance observed at the short-term horizon (H+1). Although the gains decrease as the forecasting horizon increases, they remain significant, confirming the relevance of statistical, machine learning, and deep learning approaches for solar radiation forecasting.

3.4.5. Statistical validation using the Diebold–Mariano test

To verify whether the observed improvements over the persistence model are statistically significant, the Diebold–Mariano (DM) test was applied to the forecasting errors of the selected models (Table 6). For each forecasting horizon, the ARMA statistical model, the best-performing machine learning model, and the best-performing deep learning model were compared against the persistence model. The null hypothesis of the test states that both models under comparison have equal predictive accuracy. A p -value lower than 0.05

leads to the rejection of this hypothesis, indicating that the observed difference in performance is statistically significant at the 95% confidence level.

The results in table 6 show that the performance differences between the advanced models and the persistence model are highly significant at the H+1 and D+1 forecasting horizons. Indeed, all comparisons at these two horizons yield p -values far below the 0.05 threshold, leading to the rejection of the null hypothesis of equal predictive accuracy. These findings confirm that the improvements previously observed in terms of RMSE and nRMSE are not due to chance, but reflect a real enhancement in forecasting performance.

Table 5

Performance gains of advanced models relative to benchmark models.

Model	nRMSE (%)	Gain vs. mean (%)	Gain vs. persistence (%)
H+1 Horizon			
Historical mean	44.99	—	−39.2
Persistence	32.33	28.1	—
ARMA	12.49	72.2	61.4
Random Forest	11.08	75.4	65.7
CNN-LSTM	9.84	78.1	69.6
D+1 Horizon			
Historical mean	15.67	—	9.9
Persistence	17.39	−11.0	—
ARMA	14.02	10.5	19.4
XGBoost	13.92	11.1	19.9
CNN-LSTM	13.15	16.1	24.4
M+1 Horizon			
Historical mean	8.20	—	−9.2
Persistence	7.51	8.4	—
ARMA	5.45	33.6	27.4
K-NN	4.49	45.3	40.2
CNN-LSTM	4.98	39.3	33.7

Gains are expressed in percentage relative to benchmark models.

Table 6

Results of the Diebold–Mariano (DM) test comparing advanced forecasting models with the persistence model for the H+1, D+1, and M+1 prediction horizons.

Horizon	Comparison	DM statistic	p -value	Conclusion
H+1	ARMA vs. Persistence	−16.12	< 0.001	Significant difference
H+1	Random Forest vs. Persistence	−17.13	< 0.001	Significant difference
H+1	CNN-LSTM vs. Persistence	23.60	< 0.001	Significant difference
D+1	ARMA vs. Persistence	−6.15	7.73×10^{-10}	Significant difference
D+1	XGBoost vs. Persistence	−6.15	7.71×10^{-10}	Significant difference
D+1	CNN-LSTM vs. Persistence	−6.24	4.12×10^{-10}	Significant difference
M+1	ARMA vs. Persistence	−0.4734	0.635	No significant difference
M+1	K-NN vs. Persistence	−2.22	0.026	Significant difference
M+1	CNN-LSTM vs. Persistence	−1.70	0.089	No significant difference

DM = Diebold–Mariano test statistic. p -values below 0.05 indicate significant differences.

At the H+1 horizon, the ARMA, Random Forest, and CNN-LSTM models all exhibit highly significant differences compared to the persistence model. This confirms that statistical, machine learning, and deep learning approaches effectively exploit recent information in the data to improve very short-term solar radiation forecasting.

Similar results are obtained for the D+1 horizon. The ARMA, XGBoost, and CNN-LSTM models all show extremely low p -values ($< 10^{-9}$), demonstrating that their performance is significantly superior to that of the persistence model. Although the improvements at this horizon are less pronounced than at H+1, they remain statistically robust.

At the M+1 horizon, the results are more mixed. Only the K-NN model shows a statistically significant improvement over persistence ($DM = -2.22$; $p = 0.026$). In contrast, the differences observed for ARMA ($p = 0.635$) and CNN-LSTM ($p = 0.089$) are not significant at the 5% level. Despite generally better RMSE performance, the gains achieved by these models cannot be considered sufficiently robust to rule out the effect of random variability in the data.

Overall, this analysis confirms the statistical superiority of advanced models over the persistence benchmark for the H+1 and D+1 horizons. For the M+1 horizon, only the K-NN model retains a statistically significant advantage. These results suggest that advanced approaches are particularly effective for short- and medium-term forecasting, while their advantage tends to decrease as the forecasting horizon increases and the uncertainty associated with solar radiation dynamics becomes more pronounced.

3.4.6. Statistical comparison of the CNN-LSTM and LSTM models

To statistically validate the observed differences between the LSTM and CNN-LSTM models, the Diebold–Mariano (DM) test was applied to compare their predictive accuracy. The results ($DM = 3.11$, $p < 0.01$) indicate that the CNN-LSTM model significantly outperforms the LSTM model. This confirms that the performance improvement observed in [table 4](#) is not due to random variation but reflects a genuine difference in predictive capability.

3.5. Discussion

The results demonstrate a clear scale-dependent behavior in solar irradiation forecasting, highlighting the strong interaction between temporal resolution, atmospheric dynamics, and model complexity. This confirms that forecasting performance is not solely model-driven, but fundamentally governed by the physical structure of the solar radiation process across different time scales.

At the H+1 and D+1 horizons, the CNN-LSTM model consistently achieves superior performance, with nRMSE values of 9.84% and 13.15%, respectively. These horizons are dominated by high-frequency atmospheric variability induced by rapidly evolving cloud systems, humidity fluctuations, and aerosol dynamics. Such processes introduce strong non-stationarity and

nonlinear dependencies in the irradiance signal, limiting the effectiveness of linear and purely autoregressive approaches. The superior performance of CNN-LSTM is explained by its ability to jointly model local temporal structures and long-range dependencies: convolutional layers act as adaptive feature extractors for short-term variability, while LSTM units capture sequential dependencies and memory effects in atmospheric evolution. This hybridization provides a more physically consistent representation of irradiance dynamics compared to standalone statistical or machine learning models. These findings are consistent with recent studies demonstrating the effectiveness of recurrent architectures in capturing nonlinear meteorological interactions [42], while hybrid CNN-based frameworks further enhance representational capacity by integrating local pattern extraction [43].

The robustness of these improvements is statistically confirmed by the Diebold–Mariano test, which indicates highly significant differences between CNN-LSTM and the persistence benchmark for both H+1 and D+1 horizons ($p < 0.001$). This provides strong evidence that performance gains are not attributable to stochastic variability but reflect genuine improvements in predictive skill. The observed reductions in error, reaching 69.6% at H+1 and 24.4% at D+1 relative to persistence, further indicate that the model captures physically relevant information embedded in recent atmospheric states. This is particularly relevant in tropical environments such as Man (Côte d’Ivoire), where rapid convective activity and cloud dynamics strongly modulate short-term solar resource availability [44].

Beyond model architecture, the results highlight the critical role of exogenous meteorological predictors. The feature importance analysis reveals that past irradiance and wind direction (WD10M) constitute the dominant drivers of predictive performance. This confirms that solar radiation forecasting is not purely autoregressive but strongly conditioned by atmospheric transport mechanisms and cloud advection processes. Incorporating such variables allows the model to partially account for the physical movement of air masses and cloud structures, leading to improved generalization and reduced input redundancy.

A fundamentally different predictive regime emerges at the M+1 horizon. Here, the K-NN model outperforms all competing approaches, achieving an nRMSE of 4.49%. This shift reflects a transition from short-term chaotic dynamics to a regime dominated by seasonal and climatological structure. Temporal aggregation acts as a low-pass filter, attenuating high-frequency variability and enhancing periodic components of the signal. In this context, forecasting becomes increasingly dependent on pattern similarity rather than dynamic state evolution. The superiority of K-NN indicates that instance-based learning is particularly effective when the system exhibits strong recurrence and low dynamical complexity. Similar behaviors have been reported in long-term solar forecasting studies, where aggregation reduces nonlinear complexity and favors

memory-based approaches over high-capacity neural architectures [45].

The limited performance of CNN-LSTM at the monthly scale is further explained by data constraints. With approximately 200 observations, the monthly dataset is insufficient to fully exploit the parameterization capacity of deep neural networks. In such low-sample regimes, model variance increases and the risk of overfitting becomes dominant, reducing generalization performance. This sensitivity to data scale is well documented in deep learning literature and is particularly critical in environmental applications where long-term observations are often limited [46]. This interpretation is corroborated by the Diebold–Mariano test, which shows no statistically significant difference between CNN-LSTM and persistence at M+1 ($p > 0.05$), whereas K-NN maintains a significant advantage.

Seasonal stratification further reveals systematic regime-dependent errors. All models exhibit higher accuracy during the dry season compared to the rainy season, where convective activity and cloud intermittency introduce additional stochasticity into the irradiance signal. For instance, CNN-LSTM achieves 6.78% nRMSE during the dry season versus 10.05% during the rainy season at H+1. This degradation reflects the increasing unpredictability of cloud radiative effects under high atmospheric instability. The analysis of bias further indicates consistent seasonal asymmetry, with overestimation during dry periods and underestimation during wet periods, suggesting incomplete representation of cloud radiative attenuation processes.

From a methodological perspective, the study highlights a fundamental trade-off between model complexity, data availability, and temporal aggregation. Deep learning models demonstrate clear advantages in high-frequency, nonlinear regimes but become suboptimal under strong aggregation and limited sample sizes. Conversely, simpler non-parametric models such as K-NN exhibit greater robustness in low-variability, seasonally dominated regimes. This confirms that no single model is universally optimal, but rather that forecasting performance emerges from the interaction between data structure and model inductive bias.

Finally, the reliance on NASA POWER reanalysis data introduces inherent uncertainties related to spatial resolution and retrieval algorithms, which may affect absolute error levels. Moreover, the single-site experimental design limits spatial generalization. Future work should therefore focus on multi-site validation, integration of ground-based measurements, and the development of transfer learning frameworks to enhance model portability across climatic zones. Incorporating physically informed deep learning architectures and numerical weather prediction inputs could further improve forecast robustness and support operational photovoltaic integration in West Africa.

Overall, this study demonstrates that solar irradiation forecasting is governed by a strong scale-dependent predictability structure. CNN-LSTM models are most effective under high-frequency atmospheric variability,

while similarity-based approaches such as K-NN become preferable under aggregated, seasonally driven regimes. These findings provide both methodological and physical insights into solar resource predictability in tropical climates, and underline the necessity of scale-aware modeling strategies for reliable renewable energy forecasting.

3.6. Analysis of prediction errors

The results of the model evaluations illustrated through boxplots (Figure 8) allow a detailed analysis of prediction errors. These boxplots show the error distribution for each model, and reveal that advanced machine learning models such as XGBoost, Random Forest, LSTM, and CNN-LSTM far outperform traditional models such as ARMA and K-NN. Indeed, these advanced models show a low variance of errors, indicating better stability and reliability in predictions.

A comparative analysis of the LSTM and CNN-LSTM models highlights the superiority of the CNN-LSTM model. The latter benefits from the integration of convolutional networks and LSTM networks, which allows it to capture both temporal dependencies in the data and important local features. This combination improves the model's ability to provide more accurate predictions, with lower errors and greater consistency in results.

3.7. Seasonal performance analysis

The performance evaluation of the forecast models was also carried out by season (dry season and rainy season) for the D+1 and H+1 horizons. The results are presented in tables 7 and 8.

The CNN-LSTM model showed the best performance for both the D+1 and H+1 horizons, in both dry and rainy seasons. Indeed, for the D+1 horizon, the CNN-LSTM model obtained an nRMSE of 10.63% in the dry season and 13.60% in the rainy season. For the H+1 horizon, the model obtained an nRMSE of 6.78% in the dry season and 10.05% in the rainy season. The LSTM and MLP models also showed good performance, ranking just behind the CNN-LSTM. Variations in performance between seasons highlight the impact of environmental conditions on the accuracy of predictions.

In the rainy season, predictions are less accurate, which is reflected in higher nRMSE values. This decrease in accuracy may be due to the increased variability and unpredictable weather conditions specific to this season.

In addition, models show a tendency to overestimate values in the dry season and underestimate them in the rainy season, as indicated by MBE values.

These variations in performance between seasons highlight the impact of environmental conditions on prediction accuracy. By understanding how model performance varies with seasons, practitioners can better anticipate and manage risks associated with inaccurate forecasts, particularly in critical areas such as renewable energy management or natural disaster prevention.

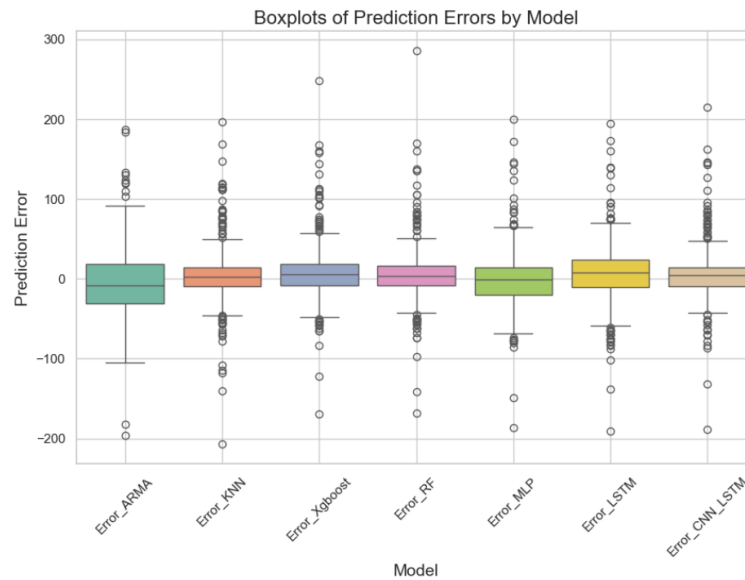


Fig. 8. Distribution of prediction errors for the different models evaluated, illustrated through boxplots.

Table 7

Evaluation of model performance by season for the daily horizon (D+1).

Season	Metric	K-NN	ARMA	XGBoost	R_FOREST	MLP	LSTM	CNN-LSTM
Dry	RMSE	622 ± 3.78	622 ± 7.04	625 ± 5.54	616 ± 4.05	613 ± 2.96	593 ± 3.38	587 ± 3.58
	nRMSE	12.06	12.12	11.95	11.89	11.77	11.50	10.63
	MAE	486	477	470	460	462	462	459
	MAPE	9.42	9.28	9.21	8.98	8.99	8.97	8.27
	R^2	0.289	0.330	0.347	0.333	0.298	0.352	0.353
	MBE	0.12	0.08	0.08	0.08	0.00	0.08	0.04
Rainy	RMSE	722 ± 4.38	702 ± 7.93	704 ± 6.25	726 ± 4.77	725 ± 3.50	702 ± 4.00	595 ± 3.58
	nRMSE	14.53	14.27	14.39	14.65	14.59	14.18	13.60
	MAE	556	537	541	551	555	543	458
	MAPE	11.23	11.20	11.17	11.42	11.41	11.11	10.76
	R^2	0.287	0.327	0.344	0.330	0.295	0.349	0.349
	MBE	0.02	-0.05	-0.03	0.00	-0.08	-0.03	-0.07

RMSE, MAE and MBE are expressed in W h m^{-2} . nRMSE and MAPE are expressed in %.

Table 8

Evaluation of model performance by season for the hourly horizon (H+1).

Season	Metric	K-NN	ARMA	XGBoost	R_FOREST	MLP	LSTM	CNN-LSTM
Dry	RMSE	46.64	48.25	45.92	41.51	39.59	42.20	39.59
	nRMSE	8.02	8.35	7.89	7.14	6.81	7.25	6.78
	MAE	32.36	35.17	33.21	27.69	27.09	26.31	27.31
	MAPE	6.07	6.59	6.22	5.18	5.07	4.93	4.65
	R^2	0.91	0.89	0.92	0.94	0.95	0.95	0.96
	MBE	-2.07	6.77	2.02	0.21	3.83	0.05	-3.35
Rainy	RMSE	69.54	69.54	69.54	69.54	69.54	69.54	58.32
	nRMSE	12.85	13.05	12.11	11.95	11.32	11.98	10.05
	MAE	51.28	51.50	48.35	45.50	44.52	44.39	45.56
	MAPE	10.38	10.42	9.79	9.21	9.01	8.99	8.73
	R^2	0.90	0.87	0.91	0.93	0.93	0.94	0.95
	MBE	-2.71	-1.13	-1.94	0.63	1.84	-1.06	-1.65

RMSE, MAE and MBE are expressed in W h m^{-2} . nRMSE and MAPE are expressed in %.

3.8. Contribution of the study

This study provides a comprehensive multi-model evaluation of solar irradiation forecasting in a tropical environment (Man, Côte d'Ivoire), comparing statistical (ARMA), machine learning (K-NN, Random Forest, XGBoost), and deep learning (LSTM, CNN-LSTM) approaches across multiple forecasting horizons (H+1, D+1, and M+1).

A key contribution is the demonstration of a strong scale-dependent behavior in forecasting performance. Results show that CNN-LSTM is the most effective model for short- and medium-term horizons (H+1 and D+1), while K-NN outperforms more complex models at the monthly scale (M+1). This highlights that forecasting accuracy strongly depends on the temporal resolution of the problem.

The study also quantifies significant improvements over benchmark models (persistence and mean approaches), with gains reaching up to 69.6% in nRMSE, and validates these improvements using the Diebold–Mariano statistical test. In addition, feature importance analysis identifies key exogenous variables, particularly wind direction and past irradiation, as important predictors of solar variability.

Overall, this work provides a scale-aware forecasting framework and practical guidance for selecting appropriate models depending on the forecasting horizon and data characteristics in tropical climates.

3.9. Challenges and limitations of the study

The generalizability of solar irradiation forecasting models remains a key challenge in this field. Local climatic conditions strongly influence model performance, meaning that approaches developed for one region cannot be directly transferred to another without proper adaptation. In addition, the quality and quantity of historical data available for a given location play a crucial role in determining predictive accuracy [16]. This limitation is consistent with the findings of Mukhtar *et al.* [47], who reported that the best-performing model varies significantly across sites, highlighting the need for region-specific calibration.

In this study, it was also observed that advanced deep learning models, particularly CNN-LSTM, achieve strong performance at short- and medium-term horizons. However, their effectiveness is highly dependent on the size of the training dataset. In the case of the monthly forecasting horizon, where the number of observations is relatively limited, the risk of overfitting increases, which can reduce the generalization capability of complex architectures.

Another important limitation relates to scalability and computational complexity. As the volume of meteorological and satellite data increases, forecasting models must be able to efficiently process large datasets while maintaining stability and accuracy. Although the models used in this study provide satisfactory performance, their extension to larger and higher-resolution datasets may require additional optimization in terms of architecture design and computational resources [48].

To address these challenges, hybrid and adaptive

modeling strategies appear promising. In this work, the use of a hybrid CNN-LSTM model demonstrated the benefits of combining spatial feature extraction and temporal dependency learning. More broadly, techniques such as transfer learning and the integration of high-quality local observational data could further improve model robustness and adaptability to different climatic contexts. These approaches not only enhance predictive accuracy but also reduce the need for full model retraining across regions, as highlighted in previous studies [49].

4. Conclusion

This study investigated global solar irradiation forecasting in a tropical environment (Man, Côte d'Ivoire) using a comprehensive comparative framework involving statistical (ARMA), machine learning (K-NN, Random Forest, XGBoost), and deep learning (LSTM and CNN-LSTM) models across three forecasting horizons (H+1, D+1, and M+1). The main objective was to evaluate how model performance varies with temporal scale and to identify the most suitable forecasting approaches for different prediction horizons.

The results demonstrate a clear scale-dependent behavior in forecasting performance. For short- and medium-term horizons (H+1 and D+1), the CNN-LSTM model consistently outperformed all other approaches, achieving the lowest prediction errors due to its ability to capture both nonlinear relationships and temporal dependencies in highly variable atmospheric conditions. At the monthly horizon (M+1), the K-NN model provided the best performance, indicating that similarity-based approaches are more effective when the signal is dominated by seasonal patterns and reduced high-frequency variability.

The study also confirms that the improvements achieved by advanced models over benchmark methods are statistically significant for short- and medium-term forecasts, as demonstrated by the Diebold–Mariano test. In addition, the inclusion of exogenous meteorological variables, particularly wind direction and past irradiation, contributed to improving predictive accuracy and model interpretability.

Overall, the findings highlight that no single model is optimal across all temporal scales. Instead, forecasting performance depends on the interaction between model complexity, data availability, and the temporal structure of solar irradiation. The proposed scale-aware framework provides practical guidance for selecting appropriate forecasting models in tropical climates and supports more efficient integration of photovoltaic energy into power systems in Côte d'Ivoire and similar regions.

Future work could focus on integrating ground-based measurements, improving model transferability across regions, and exploring advanced hybrid or physics-informed machine learning approaches to further enhance forecasting robustness and operational applicability.

References

- [1] IEA, Integrating Solar and Wind: Global Experience and Emerging Challenges, International Energy Agency, Paris (2024).
<https://www.iea.org/reports/integrating-solar-and-wind>
- [2] IEA, World Energy Outlook, International Energy Agency, Paris (2024).
<https://www.iea.org/reports/world-energy-outlook-2024>
- [3] W.M. Yeo, *A political economy analysis of the electricity sector in Côte d'Ivoire*, Energy Res. Soc. Sci. 118 (2024) 103833.
<https://doi.org/10.1016/j.erss.2024.103833>
- [4] M. Deshaies, A. Kouadio, Les enjeux du développement des énergies renouvelables pour la production d'électricité en Côte d'Ivoire, Rev. Espace Géogr. Soc. Marocaine 26 (2019).
<https://revues.imist.ma/index.php/EGSM/article/download/15049/8368/37864?>
- [5] CI-Energies, CI-Energies inaugure la centrale solaire de Côte d'Ivoire à Boundiali (2024).
<https://www.cinergies.ci/2024/04/04/ci-energies-inaugure-la-centrale-solaire-de-cote-divoire-a-boundiali/>
- [6] Ministère des Mines, du Pétrole, et de l'Énergie de Côte d'Ivoire, La Côte d'Ivoire renforce sa transition énergétique avec 4 centrales solaires d'une capacité de 210.3 MWc (2025).
<https://www.energie.gouv.ci/actualite/la-cote-divoire-renforce-sa-transition-energetique-avec-4-centrales-solaires-dune-capacite-de-2103-mwc>
- [7] IEA, Efficient Grid-Interactive Buildings: Future of Buildings in ASEAN, International Energy Agency, Paris (2023).
<https://www.iea.org/reports/efficient-grid-interactive-buildings>
- [8] D. Yang, T. Reindl, W. Walsh, *Short-term solar irradiance forecasting based on sky imaging and artificial neural networks*, Renew. Energy 90 (2019) 83–92.
<https://doi.org/10.1016/j.renene.2019.02.056>
- [9] S. Pelland, J. Remund, T. Oozeki, K. De Brabandere, L. Vuilleumier, *Photovoltaic and Solar Forecasting: State of the Art*, IEA PVPS Task 14, Subtask 3.1 (2013).
https://iea-pvps.org/wp-content/uploads/2013/10/Photovoltaic_and_Solar_Forecasting_State_of_the_Art_REPORT_PVPS_T14_01_2013.pdf?
- [10] D. Yang, J. Kleissl, C.A. Gueymard, H.T. Pedro, C.F.M. Coimbra, *History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining*, Sol. Energy 168 (2018) 60–101.
<https://doi.org/10.1016/j.solener.2017.11.023>
- [11] IRENA, Solar PV in Africa: Costs and Markets, International Renewable Energy Agency, Masdar City, United Arab Emirates (2016).
<https://www.irena.org/Publications/2016/Sep/Solar-PV-in-Africa-Costs-and-Markets>
- [12] Y. Huang, Y. Chen, X. Liu, Y. Sun, Y. Wang, Y. Wen, *Solar irradiance prediction using sky images and satellite data*, Energy 195 (2020) 117011.
<https://doi.org/10.1016/j.energy.2020.117011>
- [13] E. Lorenz, D. Heinemann, L. Schneider, *Short-term forecasting of solar irradiance using the WRF model*, Sol. Energy 136 (2016) 251–262.
<https://doi.org/10.1016/j.solener.2016.07.012>
- [14] R. Ahmed, A.M. Siddiqui, M. Raza, *Short-term solar irradiance forecasting using smart persistence model*, Energy Rep. 7 (2021) 237–245.
<https://doi.org/10.1016/j.egy.2021.01.001>
- [15] S. Liao, G. Li, P. Wang, *Solar irradiance forecasting based on Markov chain model*, Appl. Energy 160 (2015) 135–144.
<https://doi.org/10.1016/j.apenergy.2015.09.012>
- [16] C. Voyant, G. Nottton, S. Kalogirou, M.L. Nivet, C. Paoli, F. Motte, A. Fouilloy, *Machine learning methods for solar radiation forecasting: A review*, Renew. Energy 105 (2017) 569–582.
<https://doi.org/10.1016/j.renene.2016.12.095>
- [17] L. Breiman, *Random forests*, Mach. Learn. 45(1) (2001) 5–32.
<https://doi.org/10.1023/A:1010933404324>
- [18] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining, ACM, New York (2016) 785–794.
<https://doi.org/10.1145/2939672.2939785>
- [19] S.A. Kalogirou, *Solar Energy Engineering: Processes and Systems*, Academic Press (2009).
<https://doi.org/10.1016/B978-0-12-374501-9.X0001-5>
- [20] S. Hochreiter, J. Schmidhuber, *Long short-term memory*, Neural Comput. 9(8) (1997) 1735–1780.
<https://doi.org/10.1162/neco.1997.9.8.1735>
- [21] K.C. Jayasankar, G. Anandhakumar, A. Kalaimurugan, *Prediction of solar radiation using deep LSTM-based machine learning algorithm*, J. Environ. Nanotechnol. 13(3) (2024) 01–08.
<https://doi.org/10.13074/jent.2024.09.243806>
- [22] A.K. Yadav, S.S. Chandel, *Solar radiation prediction using artificial neural network techniques: A review*, Renew. Sustain. Energy Rev. 33 (2014) 772–781.
<https://doi.org/10.1016/j.rser.2013.08.055>

- [23] F. Besharat, A.A. Dehghan, A.R. Faghieh, *Empirical models for estimating global solar radiation: A review and case study*, Renew. Sustain. Energy Rev. 21 (2013) 798–821.
<https://doi.org/10.1016/j.rser.2012.12.043>
- [24] J.W. Tukey, *Exploratory Data Analysis*, Addison-Wesley, Reading, Massachusetts (1977).
https://books.google.ci/books?id=UT9dAAAAIAAJ&redir_esc=y
- [25] R. Tibshirani, *Regression shrinkage and selection via the lasso*, J. R. Stat. Soc. Ser. B Methodol. 58(1) (1996) 267–288.
<https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [26] A.K. Yadav, H. Malik, S.S. Chandel, *Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models*, Renew. Sustain. Energy Rev. 31 (2014) 509–519.
<https://doi.org/10.1016/j.rser.2013.12.008>
- [27] R. Genuer, J.M. Poggi, C. Tuleau-Malot, *Variable selection using random forests*, Pattern Recognit. Lett. 31(14) (2010) 2225–2236.
<https://doi.org/10.1016/j.patrec.2010.03.014>
- [28] A. Fouilloy, *Comparaison de méthodes d'apprentissage automatique de prévision de la ressource solaire pour une application à une gestion optimisée des réseaux intelligents*, Ph.D. Thesis, Université de Corse-Pascal Paoli (2019).
<https://theses.hal.science/tel-03479238>
- [29] F. Rosenblatt, *The perceptron: A probabilistic model for information storage and organization in the brain*, Psychol. Rev. 65(6) (1958) 386–408.
<https://doi.org/10.1037/h0042519>
- [30] T. Cover, P. Hart, *Nearest neighbor pattern classification*, IEEE Trans. Inf. Theory 13(1) (1967) 21–27.
<https://doi.org/10.1109/TIT.1967.1053964>
- [31] G.E.P. Box, G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco (1976).
https://books.google.ci/books/about/Time_Series_Analysis.html?id=1WVHAAAAMAAJ&redir_esc=y
- [32] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, W.-C. Woo, *Convolutional LSTM network: A machine learning approach for precipitation nowcasting*, in: Adv. Neural Inf. Process. Syst. (2015) 802–810.
https://papers.nips.cc/paper_files/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf
- [33] Pirjatullah, Kartini, D.T. Nugrahadi, D.T. Muiliadi, A. Farmadi, *Hyperparameter tuning using GridSearchCV on the comparison of the activation function of the ELM method to the classification of pneumonia in toddlers*, in: Proc. 4th Int. Conf. Computer Informatics Engineering (IC2IE), Jakarta, Indonesia (2021) 390–395.
<https://doi.org/10.1109/IC2IE53219.2021.9649353>
- [34] Y. Bengio, P. Simard, P. Frasconi, *Learning long-term dependencies with gradient descent is difficult*, IEEE Trans. Neural Netw. 5(2) (1994) 157–166.
<https://doi.org/10.1109/72.279181>
- [35] D.P. Kingma, J. Ba, *Adam: A method for stochastic optimization*, in: 3rd Int. Conf. Learning Representations (ICLR) (2015).
<https://arxiv.org/abs/1412.6980>
- [36] A.Y. Ng, *Feature selection, L1 vs. L2 regularization, and rotational invariance*, in: Proc. 21st Int. Conf. Machine Learning (ICML) (2004) 78.
<https://doi.org/10.1145/1015330.1015435>
- [37] A. Liaw, M. Wiener, *Classification and regression by randomForest*, R News 2(3) (2002) 18–22.
https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf
- [38] H. Akaike, *A new look at the statistical model identification*, IEEE Trans. Autom. Control 19(6) (1974) 716–734.
<https://doi.org/10.1109/TAC.1974.1100705>
- [39] C. Bergmeir, R.J. Hyndman, B. Koo, *A note on the validity of cross-validation for evaluating autoregressive time series prediction*, Comput. Stat. Data Anal. 120 (2018) 70–83.
<https://doi.org/10.1016/j.csda.2017.11.003>
- [40] C.J. Willmott, *On the validation of models*, Phys. Geogr. 2(2) (1981) 184–194.
<https://doi.org/10.1080/02723646.1981.10642213>
- [41] R.J. Hyndman, A.B. Koehler, *Another look at measures of forecast accuracy*, Int. J. Forecast. 22(4) (2006) 679–687.
<https://doi.org/10.1016/j.ijforecast.2006.03.001>
- [42] E. Abdullah, *Enhancing solar irradiance forecasting using LSTM and meteorological data*, J. Inf. Syst. Eng. Manag. 10(37s) (2025) 161–166.
<https://doi.org/10.52783/jisem.v10i37s.5449>
- [43] B. Ladjal, M. Nadour, M. Bechouat, N. Hadroug, M. Sedraoui, A. Rabehi, M. Guermoui, T.F. Agajie, *Hybrid deep learning CNN-LSTM model for forecasting direct normal irradiance: a study on solar potential in Ghardaia, Algeria*, Sci. Rep. 15 (2025) 15404.
<https://doi.org/10.1038/s41598-025-99626-0>
- [44] A. Demirci, I. Dagal, M. Terkes, U. Cali, *Towards reliable solar power forecasting in Sub-Saharan Africa: An explainable hybrid AI approach for Chad*, Energy Rep. 15 (2026) 108997.
<https://doi.org/10.1016/j.egyr.2026.108997>

- [45] T. Nazih, *Short-term and long-term solar irradiance forecasting with advanced machine learning techniques in Zafarana, Egypt*, Sci. Rep. 15 (2025) 39553.
<https://doi.org/10.1038/s41598-025-24051-w>
- [46] A. Gufron, P.M.P. Garniwa, D.A. Putera, F.A. Suwadana, D. Puspita, H. Lee, I.A. Aditya, S. Supriatna, *A preliminary LSTM-IDW model for spatiotemporal hourly solar radiation estimation in tropical regions*, Renew. Sustain. Energy Transit. 7 (2025) 100105.
<https://doi.org/10.1016/j.rset.2025.100105>
- [47] M. Mukhtar, A. Oluwasanmi, N. Yimen, Z. Qinxiu, C.C. Ukwuoma, B. Ezurike, O. Bamisile, *Development and comparison of two novel hybrid neural network models for hourly solar radiation prediction*, Appl. Sci. 12(3) (2022) 1435.
<https://doi.org/10.3390/app12031435>
- [48] K.C. Jayasankar, G. Anandhakumar, A. Kalaimugan, *Advancements and challenges in solar radiation prediction: A review of machine learning approaches*, J. Environ. Nanotechnol. 13(2) (2024) 60–64.
<https://doi.org/10.13074/jent.2024.06.242736>
- [49] S.J. Pan, Q. Yang, *A survey on transfer learning*, IEEE Trans. Knowl. Data Eng. 22(10) (2010) 1345–1359.
<https://doi.org/10.1109/TKDE.2009.191>