

Article

Benchmarking Tree-Based Artificial Intelligence Models for Multi-Resolution Solar Irradiance Forecasting Across Various Sky Conditions in Arid Climates

Hasanain A. H. Al-Hilfi ¹, Farhad Shahnia ^{2,*} , Seyit Alperen Celtek ³ , Amirmehdi Yazdani ²  and Hai Wang ² 

¹ IT and Communication Center, University of Basrah, Basra 61004, Iraq; hasanain.alhilfi@uobasrah.edu.iq

² School of Engineering and Energy, Murdoch University, Perth 6150, Australia; amirmehdi.yazdani@murdoch.edu.au (A.Y.); hai.wang@murdoch.edu.au (H.W.)

³ Department of Energy Systems Engineering, Karamanoglu Mehmetbey University, Karaman 70100, Türkiye; salperenceltek@kmu.edu.tr

* Correspondence: f.shahnia@murdoch.edu.au

Abstract

Integrating solar power into electricity grids requires accurate short-term forecasting of the global horizontal irradiance to accurately predict the expected solar power generation. This paper compares five tree-based machine learning models against a Persistence baseline for multi-resolution forecasting in arid climates. A 13-year dataset from Basra, Iraq, has been employed in this study for verification purposes, and the models are tested across various very-short- to short-term forecasting horizons of 5, 10, 15, 30, and 60 min. Unlike most existing studies that focus on single forecasting horizons or mixed climatic conditions, this work systematically benchmarks multi-resolution irradiance forecasting under distinct sky conditions in a hot arid environment using a strict anti-data-leakage framework. To avoid data leakage in these models, feature engineering has used only lagged inputs. The dataset has been split into three groups for training, validation, and testing (respectively 70, 15, and 15% of the entire available dataset). The models were then tested separately under clear, partly cloudy, and cloudy skies. Numerical studies prove that picking the best model depends heavily on the forecast horizon. For very-short-term predictions, the Persistence model was competitive ($RMSE = 21.32 \text{ W/m}^2$), while the Gradient Boosting model proved slightly more accurate ($RMSE = 17.65 \text{ W/m}^2$). For the 60 min horizon, the boosting models took a clear lead. The HistGradientBoosting model resulted in a 67% reduction in the RMSE compared to the Persistence baseline. Also, the top-performing model changed depending on the weather and the time scale. Gradient Boosting was the clear winner for short-term clear sky forecasts, while XGBoost handled the longer horizons. Partly cloudy skies showed a rotating mix of different boosting algorithms taking the lead. However, studies show that when skies were fully overcast, complex machine learning models fail to capture chaotic patterns, making the simple Persistence baseline a necessary reliability safeguard. The results reveal that no single model consistently dominates all forecasting horizons and weather conditions, highlighting the necessity of adaptive model selection for operational solar forecasting. These findings highlight the importance of horizon- and weather-adaptive model selection for operational solar forecasting. Rather than relying on a single universal algorithm, grid operators in arid regions can improve forecasting reliability by dynamically selecting models based on prevailing sky conditions and forecast horizons.

Keywords: global horizontal irradiance (GHI); multi-resolution solar forecasting; arid climate; GradientBoosting; HistGradientBoosting; XGboost; LightGBM



Academic Editor: Tapas Mallick

Received: 18 May 2026

Revised: 14 June 2026

Accepted: 22 June 2026

Published: 29 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

To reduce the global warming impacts caused by CO₂ and other greenhouse gases, the energy sector has started a major shift in its energy production mix [1]. In this context, solar energy systems have been introduced as an essential solution and a fundamental component of the global renewable market [2]. However, the integration of large-scale solar systems into electrical power networks suffers from inherent and unexpected fluctuations due to the passing of clouds in the sky. These energy fluctuations are considered a significant challenge for electrical network stability and power and market operations. This fluctuation is rooted in solar irradiance. While diurnal and seasonal cycles cause predictable changes, the main source of unpredictable fluctuation comes from weather factors, particularly cloud dynamics and aerosols [3,4]. Therefore, the accurate forecasting of global horizontal irradiance (GHI)—the solar radiation that falls horizontally on the surface of the Earth—has now become a necessity for efficient grid management, improved power plant operation, and effective contribution in the energy sector. Accurate GHI data are particularly critical for photovoltaic system modeling and parameter identification, where precise inputs drive the reliability of power output estimations [5]. In this context, short-term solar forecasting, ranging from minutes to a few hours, is particularly crucial for the real-time operations of electrical power networks [6].

To address this challenge, several techniques have emerged. Numerical weather prediction models have been introduced as physics-based techniques, while autoregressive, moving average, and hybrid models such as ARIMA have been adopted as statistical time-series analysis techniques. Recently, machine learning (ML) methods have been suggested for solving prediction issues related to random-pattern data such as time-series data, with solar irradiance being a prime example [2,7–9].

1.1. Related Works

The evolution of solar forecasting techniques has transitioned from physical and statistical methods to advanced ML approaches. This transition is driven by the increasing need for higher accuracy in grid management and the growing complexity of energy systems. For long-term forecasting, numerical weather prediction methods are considered valuable tools; however, their spatial and temporal resolution is insufficient for the short-term, minute-scale forecasting required for grid stability [10]. Traditional statistical techniques, such as autoregressive models, usually struggle to capture complex and non-linear data that characterize solar irradiance [11]. These models typically assume linearity and stationarity, assumptions that frequently break down in the volatile atmospheric conditions of arid regions.

While physical and statistical techniques face inherent limitations, ML models have become popular for modeling complex, non-linear trends directly from historical data without needing any physics equations [12,13]. As an example, early ML solar forecasting models, such as artificial neural networks and support vector machines, clearly outperformed statistical methods by effectively learning hidden patterns in the data [2]. However, these early models often suffered from the ‘black-box’ problem; i.e., they functioned as systems with hidden internal logic, making it difficult for users to interpret exactly how input variables lead to specific predictions. They also require extensive data normalization and hyperparameter tuning to avoid overfitting [14].

To address these challenges, the focus of the research community has shifted toward ensemble tree-based methods. Ensemble tree-based ML techniques such as Random Forest, various Gradient Boosting methods, XGBoost and LightGBM have also demonstrated high accuracy in predicting data with random characteristics, and they have been reported to have the ability to address high dimensional feature spaces [15,16]. Random Forest, for

example, utilizes bagging techniques to reduce variance, while boosting methods iteratively correct errors from previous iterations [17]. Similarly, the recent literature points to the strong performance of ensemble methods in arid climates. In Saudi Arabia, for instance, XGBoost clearly outperformed other techniques in rooftop PV forecasting, achieving an R^2 of 0.975 despite desert weather challenges [18].

In addition to tree-based methods, deep learning models like long short-term memory (LSTM) networks and convolutional neural networks have become popular because they are great at tracking patterns over time [19]. However, there is a trade-off: while these models are powerful, they are often criticized for needing significant computing power and being difficult to interpret compared to tree-based ensembles. Other studies suggest that while LSTMs excel with long data sequences, Gradient Boosting techniques often deliver better accuracy for standard datasets and train much faster [20]. This makes tree-based models a practical choice for real-world forecasting, where speed and the ability to explain the results are critical.

More recent studies of time-series solar forecasting evaluate these tools specifically for grid stability and reducing energy fluctuations. For smart grids, fast models like XGBoost, LightGBM, and CatBoost have been reported to provide the accurate forecasts needed for good energy management [12]. Instead of relying heavily on weather factors, ref. [13] has evaluated ML models using a dataset with six input variables, including humidity, ambient temperature, wind speed, visibility, and cloud ceiling, to predict an output variable like actual solar energy generation. In the same way, ref. [21] has analyzed five ML models like XGBoost and LightGBM to forecast output power using different internal features and found that CatBoost was the optimal model. Using a 1 min data resolution to forecast solar radiation collected in Santo Domingo, ref. [22], found that a Histogram-Based Gradient Boosting model was the best model compared with other ML models, obtaining a root mean squared error (RMSE) = 56.73 W and $R^2 = 0.964$.

Wider comparisons were made between XGBoost, LightGBM, and LSTM, using historical weather data to forecast the output power of PV systems, by [16] and the study shows that the XGBoost model was the most accurate model using different errors metrics. These results were confirmed in a parallel study conducted in Spain, employing hourly solar data to predict output energy over a three-year period, where the study found an RMSE of 11.042 W and R^2 of 0.999 [23]. Recent comparative studies in similar climates have evaluated a wide range of algorithms. For instance, a study conducted in Morocco compared six ML algorithms, including SVR, ANN, decision trees, Random Forest, and XGBoost, for predicting solar energy production [24]. The study concluded that the artificial neural network (ANN) was the most effective predictive model, outperforming tree-based methods like Random Forest and XGBoost, with the lowest RMSE and highest R^2 for daily forecasting. These findings highlight the strength of neural networks, while also confirming that tree-based ensembles remain highly competitive and widely adopted benchmarks in the field.

Furthermore, hybrid techniques combining signal decomposition methods (like Wavelet Transform or Empirical Mode Decomposition) with ML predictors have shown promising results in stabilizing forecasts during fluctuating weather [25]. By decomposing the original irradiance series into stable sub-series, these hybrid models can better handle noise, though they add complexity to the modeling pipeline.

Testing models with real-time weather data often causes data leakage and artificially high accuracy scores. To avoid this trap, ref. [26] has evaluated models using a strictly lagged 15 min dataset and reported that the CatBoost model achieved highly precise results (RMSE of 4.06 MW and R^2 of 0.9106). This demonstrates that ML models are highly effective when data leakage is prevented. However, they still face limitations under certain

weather conditions. For very-short-term forecasts, the Persistence baseline remains highly competitive. This highlights the need for researchers to rigorously compare complex ML algorithms to prove they provide a true advantage over simple baselines.

High aerosol loads and extreme heat are some important climatic features in arid regions [27]. However, most current studies ignore these aspects because they focus heavily on temperate and tropical zones [28,29]. This oversight is critical, as the dust storms and hazy conditions typical of arid environments can cause sudden drops in sunlight that standard models often fail to predict [30]. Some studies have incorporated complex variables such as aerosol optical depth, water vapor content, and ozone thickness into their models too [31,32]. New evidence, however, shows that standard weather data alone can provide highly accurate results, and the best set of input features simply changes depending on the algorithm used [33–35].

1.2. Identified Research Gaps

While solar forecasting has seen major improvements, important gaps still exist. First, most studies limit their assessment to individual forecasting resolutions [36–38]. Few studies compare models across multiple time scales. Researchers rarely test advanced ML models against a simple Persistence baseline to prove that they are better. Assuming a single model dominates all time scales needs more investigation. This is because the forecasting task completely changes; fast, noise-affected shifts are very different from slower baseline patterns.

Second, while numerous studies have investigated the impact of weather conditions on forecast accuracy, detailed investigations comparing the effects of atmospheric factors such as clear, partly cloudy, and overcast conditions across different time resolutions remain limited. While clouds are known to reduce accuracy [3,4,39], it is rare to find studies that check if clear sky models keep their specific advantage during partly cloudy or overcast conditions. As such, the exact impact of weather conditions on solar forecasting accuracy is still unclear.

Furthermore, some important climatic features in arid regions—such as high aerosol loads and extreme heat—are still not widely represented in studies, even though they have a significant impact on forecasting. As such, although dust storms are a common challenge in arid regions such as the Middle East and central parts of Australia, their specific impact on the accuracy of solar forecasting by ML models is rarely tested in the literature. This gap is evident in arid and dusty regions, such as the Middle East or the central parts of Australia, where unique weather conditions may lead to forecasting challenges that are insufficiently addressed by models developed for other climates [1,40].

Furthermore, current studies rarely look at how combining shifting weather patterns affects model performance across different time scales. This gap matters most for arid regions such as the Middle East and central parts of Australia. These areas face dust storms and unique cloud patterns that can easily disrupt model performance.

Finally, beyond feature selection, a major problem with data leakage exists in the current research. Many studies use current weather data as input, which artificially boosts their accuracy scores. A fair test requires using only past, lagged data to match real-world conditions. Ultimately, this proves that choosing the right algorithm matters much more than just feeding the model a massive number of input variables.

1.3. The Key Contribution of This Work

To address the above gaps, this paper evaluates and compares the performance of five ML models (i.e., Random Forest, Gradient Boosting, HistGradientBoosting, XGBoost, and LightGBM) against a Persistence baseline using a large-scale dataset collected over

13 years. The comparison has been conducted for multi-resolution GHI forecasting (across five distinct short-term resolutions: 5, 10, 15, 30 and 60 min) in Basra, Iraq (a hot and arid climate, the focus of this research). As such, unlike the existing research, this study has focused on arid climates only and has focused on very short to short horizons. These are the critical focus areas of this study compared to the previous techniques and studies. The paper aims to determine the accuracy and reliability of these ML models in solar forecasting for changes across different time resolutions to enable electricity grid operators to properly select suitable ML models at various horizons in such areas.

In summary, the primary objectives of this research are to identify the optimal tree-based model for each specific forecast horizon in an arid climate and to develop a decision framework for grid operators to adaptively select models based on current weather conditions. As such, the key contributions of this paper to the research field can be summarized as follows:

1. We assess forecasting models using a large-scale dataset collected over 13 years, which is processed and enhanced through feature engineering, strictly utilizing lagged inputs to prevent data leakage, across five distinct short-term resolutions: 5, 10, 15, 30 and 60 min.
2. We determine model accuracy changes across different time resolutions and guide model selections.
3. We evaluate model reliability by testing models across specific weather conditions (clear, partly cloudy, and overcast skies) to categorize the best-performing ML model for each weather pattern at different time resolutions.
4. We compare all models against a Persistence baseline to set a clear standard. This gives grid operators a practical guide for short-term solar forecasting in arid areas, helping to increase solar energy use in these regions.

It should be acknowledged that this study intentionally focuses on a single arid location to isolate the interaction between forecast horizon, sky condition, and model structure without introducing confounding climatic variability. As such, while the numerical values may be site-specific, the observed relative performance trends among tree-based models are expected to remain relevant across arid regions with similar atmospheric characteristics, such as high clearness indices, dust loading, and extreme temperatures. It must be emphasized that similar arid environments may differ in aerosol composition, cloud dynamics, and synoptic patterns. Therefore, while relative trends are expected to generalize, quantitative validation across multiple regions remains necessary.

The remainder of the paper is arranged as follows: Section 2 provides a relevant literature review, Section 3 discusses the materials and methodologies, Sections 4 and 5 introduce and discuss the experimental results, and Section 6 presents the conclusions and observations.

2. Dataset, Data Collection and Processing

This section introduces the study area focused on in this research and discusses the employed dataset and applied data processing approach.

2.1. Study Area and Data Collection

This study employs satellite-derived meteorological data from Basra, Iraq (approximately 30.5° N, 47.8° E). The city has a hot, arid climate that makes forecasting difficult due to high aerosol loads and dust storms, even though it has excellent solar energy potential. The employed dataset contains high-temporal-resolution solar irradiance estimates from the Solcast database [41], covering a 13-year period from 2007 to 2010 and from 2017 to 2025. This long timeframe enables the studied ML models with enough data to learn historical climate shifts, yearly variations, and recent weather patterns. The time-series data consists

of five separate datasets, each corresponding to a 5, 10, 15, 30, or 60 min resolution. It is important to note that a temporal gap exists in the dataset between 2011 and 2016 due to the unavailability of consistent satellite-derived records for the study site during that period. This gap was handled by treating the data as two continuous blocks during preprocessing, ensuring that no artificial temporal dependencies were created across the missing years. This approach is valid because the physical processes governing solar irradiance in arid climates exhibit stationarity, and the extensive 13-year record provides sufficient statistical depth to capture long-term climate variability. Furthermore, since the utilized ML models rely on short-term lagged inputs (e.g., $t - 1$ and $t - 2$) rather than long-term sequential dependencies, the absence of data between 2011 and 2016 does not disrupt the feature–target relationship required for training. The data volume within each continuous block is sufficient to capture the diurnal and seasonal dynamics necessary for robust model learning.

It should be noted that the satellite-derived irradiance data were selected in this study due to their long temporal consistency and availability over extended periods, which are essential for the multi-year, high-resolution benchmarking in this research. Satellite-derived datasets are particularly valuable in arid regions, where dense and long-term ground-based irradiance measurement networks are often limited or unavailable. In addition, such datasets are widely adopted in operational solar forecasting and practical grid management applications. It is important to acknowledge that satellite-derived irradiance estimates inherently carry uncertainties compared to ground-based measurements. For the Solcast dataset utilized, validation studies typically report a mean absolute error (MAE) in the range of 5–10% for GHI under general conditions, with higher errors possible during periods of high aerosol loading or cloud opacity. Specifically, satellite retrieval algorithms can struggle to distinguish between thick cloud covers and heavy dust storms, which are frequent in the study region. This may lead to biases in the input features during such extreme events. These uncertainty bounds provide context for the forecasting results, although the primary focus of this study is the relative performance comparison between models using a consistent data source.

2.2. Data Quality and Statistical Analysis

The employed 13-year dataset was rigorously analyzed to ensure reliability. Basra serves as a representative case study for global arid regions, a fact substantiated by the descriptive statistics presented in Table 1. The analysis of a comprehensive 13-year dataset confirms that the region's climatic features are stable, long-term characteristics rather than short-term anomalies. This extensive record validates the Persistence of thermal extremes, with maximum air temperatures reaching 52 °C, and distinct aridity, evidenced by a low mean relative humidity of 28.46%. Furthermore, the long-term data highlights the region's specific solar dynamics, particularly the dominance of clear sky conditions indicated by a high mean clear sky index of 0.92. To ensure the statistical consistency of the dataset across the temporal gap, the key meteorological parameters (mean GHI and clear sky index) of the pre-gap period (2007–2010) were compared with the post-gap period (2017–2025). The analysis confirmed a high degree of similarity in the distributions of both periods, validating that the gap does not introduce significant distributional bias. This temporal depth ensures that the observed atmospheric patterns are reliable indicators of arid environments, making the findings highly applicable to similar regions. The absence of bias is also verified through a chronological validation strategy in which the training dataset will contain samples from both sides of the gap. If the temporal gap had introduced a significant bias or inconsistency in the training data, the models would have struggled to generalize to the unseen test data. However, later, the results will show that the models

achieved high prediction accuracy on the test set. This strong performance on unseen data serves as empirical verification that the training data—including the segments surrounding the gap—are consistent and suitable for robust model development.

Table 1. Descriptive statistics of key meteorological variables.

Weather Parameter	Unit	Mean	Standard Deviation	Minimum	First Quartile of Data (25%)	Median Quartile of Data (50%)	Third Quartile of Data (75%)	Maximum
GHI	W/m ²	440.07	295.07	1.00	163.00	436.00	690.00	1055.00
DNI	W/m ²	357.74	270.84	0.00	49.00	391.00	591.00	949.00
Diffuse horizontal irradiance (dhi)	W/m ²	192.14	108.17	1.00	110.00	189.00	271.00	595.00
Global tilted irradiance (gti)	W/m ²	504.74	321.99	0.00	184.00	552.00	799.00	1088.00
Air temperature	°C	31.51	10.78	−1.00	22.00	32.00	41.00	52.00
Relative humidity	%	28.46	20.88	4.40	12.70	20.70	37.70	100.00
Wind speed at 10 m	m/s	4.80	2.54	0.00	2.80	4.50	6.50	14.60
Wind direction at 10 m	°	254.90	89.67	0.00	185.00	298.00	315.00	360.00
Cloud opacity	%	8.21	18.63	0.00	0.00	0.00	3.20	97.00
Albedo	-	0.27	0.03	0.21	0.25	0.28	0.29	0.30
Clear sky index	-	0.92	0.19	0.02	0.97	1.00	1.00	1.00

Data quality is especially critical in arid regions, where dust storms and aerosols can easily introduce noise into the measurements [42]. To accurately reflect operational conditions, the descriptive statistics presented in Table 1 are calculated exclusively for daytime hours ($GHI > 0$). Predicting zero irradiance during the night is a trivial task that does not require complex machine learning models. As such, including data corresponding to $GHI = 0$ would artificially lower the error metrics (e.g., RMSE) without reflecting the model’s actual capability to predict solar power generation. Also, excluding nighttime zeros prevents model bias, allowing the algorithms to focus exclusively on complex day-time dynamics. In addition, Figure 1 illustrates the strong inverse relationship between temperature and humidity, a classic signature of an arid climate that distinguishes this study site from temperate zones [43].

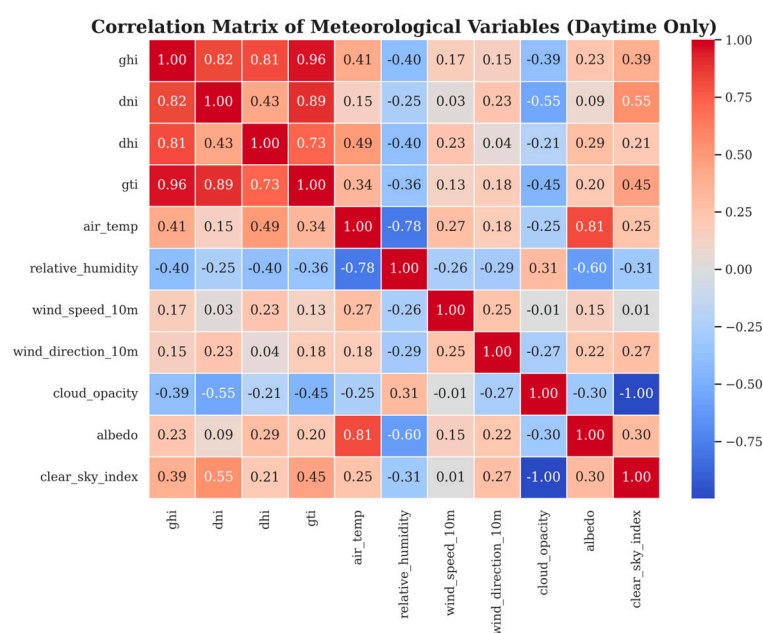


Figure 1. Correlation matrix of meteorological variables.

Table 1 provides a clear picture of the daytime environment for the study site. The GHI has a mean of 440.07 W/m², a standard deviation of 295.07 W/m² and a median

of 436 W/m^2 . The median value aligns closely with the mean, indicating a symmetrical distribution of solar intensity during the day. This is a stark contrast to 24 h statistics, where nighttime zeros drag the median down; here, the statistics reflect actual working conditions for a solar plant. The high standard deviation reflects the significant natural volatility of solar energy in this region. This large deviation is driven by the wide range of irradiance values from morning to noon and the stark contrast between clear skies and intermittent cloud cover, fluctuating between the 25th percentile (163 W/m^2) and the 75th percentile (690 W/m^2) as clouds pass over.

The dataset also captures the extreme climate nature of the study site. The mean air temperature is 31.51°C , but maximum temperatures reach a scorching 52°C , a defining characteristic of the region [42]. In contrast, the relative humidity remains low, averaging just 28.46%.

Furthermore, the clear sky index provides insight into the prevailing weather. With a mean of 0.92, the data confirm that the study site enjoys mostly clear skies during the day. This high clearness is visually supported by the low median cloud opacity (0.00), though the maximum opacity of 97 shows that heavy cloud events do occur. Figure 1 reinforces these relationships, showing that GHI has a strong positive correlation with a direct normal irradiance (DNI) of 0.91 and a clear sky GHI of 0.98. Wind speed and direction are also well distributed, allowing the models to account for wind-driven cooling and dust movement.

Ultimately, this variability—driven by the gap between clear and cloudy states—confirms that simple linear models would struggle to capture the complex dynamics. This necessitates the use of advanced non-linear ensemble methods [44]. The dataset is physically consistent, comprehensive, and fully prepared for training the high-performance ML models.

2.3. Data Preprocessing and Feature Engineering

The GHI data was cleaned by removing anomalies and handling any missing values. To improve the models' predictions, the input variables must be prepared carefully. Most importantly, to prevent data leakage, all weather features were shifted. This means that the models only use past data (time $t - 1$) to predict future GHI (time $t + k$). It is crucial to distinguish between deterministic and stochastic features when addressing data leakage. Solar geometry features (e.g., Zenith, Azimuth) and theoretical clear sky irradiance were calculated for the target prediction time (t). Since these variables are derived from astronomical equations rather than atmospheric measurements, their future values are known with absolute certainty at the time of prediction and thus do not constitute data leakage. Conversely, all stochastic meteorological variables (e.g., measured GHI, DNI, cloud opacity) were strictly lagged ($t - 1$) to ensure that no observed future data entered the training process. Feature engineering was then used to create new inputs with real physical meaning to help the models' learning processes.

To compare the ML models fairly, this study has employed 10 key features, as discussed below, and, to find the most useful input features, their impact on the accuracy error reduction was tested across all five models and time resolutions. The feature importance scores were scaled from 0 to 100. Table 2 lists these 10 engineered features chosen for this study, showing their average normalized scores across all models and time resolutions. These specific features are well established in the existing literature [45,46]. As an example, it can be seen from Table 2 that the Random Forest model measures a feature's importance by calculating how much it lowers the mean squared error (MSE) when splitting data at a node, averaged across all the trees.

Table 2. Top 10 most predictive engineered features across all models and forecast resolutions.

Rank	Feature Name	Frequency *	Avg. Importance #	Description
1	GHI_{lag1}	20	61.28	GHI value from the previous timestep (lagged GHI)
2	GHI_{clear}	20	22.52	Theoretical maximum GHI under clear skies
3	∇GHI	20	5.43	Instantaneous rate of change in GHI
4	k_t	20	1.73	Ratio of measured GHI to clear sky GHI
5	$O_{c,t-1}$	20	1.56	Cloud density from the previous timestep
6	hour	20	1.03	Hour of the day
7	zenith	20	0.9	Solar zenith angle
8	DNI_{clear}	20	0.75	Theoretical Direct Normal Irradiance
9	DNI_{lag1}	20	0.66	Direct Normal Irradiance from previous timestep
10	azimuth	20	0.63	Solar azimuth angle

* Frequency indicates the number of cross-validation folds ($k = 20$) in which the feature was selected. # Importance values represent the mean decrease in model performance (scaled).

It should be highlighted that the aggregated feature importance ranking presented in Table 2 intends to provide a global interpretability overview across models and forecast horizons rather than a model-specific diagnostic. While individual models exhibit minor variations in importance ranking, the dominant predictors (lagged GHI and clear sky components) remain consistent across algorithms. A single lag depth (lag = 1) was selected for the input features. This choice is justified by the high temporal autocorrelation of solar irradiance, where the immediate past value ($t - 1$) captures most of the predictive information for short-term horizons. This behavior is typically confirmed by the sharp cut-off of the partial autocorrelation function after the first lag [47]. While additional lags (such as $t - 2$ or $t - 3$) were considered, preliminary analysis suggested they contributed minimal additional predictive power for these resolutions and risked introducing noise into the models. Instead, the historical trend was effectively captured by the GHI gradient feature, which incorporates information from $t - 2$ to measure the rate of change.

The key features used in this research are defined as follows:

1. GHI_{lag1} : This parameter represents the GHI value from the previous time step, making it a key feature for short-term forecasts [48]. The Persistence model relies entirely on this feature because it assumes that the weather stays the same from one step to the next [49]. This model is expressed by the following [48]:

$$GHI = GHI_{t-1} \quad (1)$$

Therefore, GHI_{t-1} provides vital autoregressive information for statistical and ML models, capturing the inherent temporal inertia of GHI .

2. GHI_{clear} : The precise theoretical calculation of GHI under standard ideal circumstances is executed using the Ineichen–Perez model [50], which provides the essential baseline for calculating atmospheric attenuation, as described by

$$GHI_{clear} = f(solar_{zenith}, atmospheric_{parameters}) \quad (2)$$

3. GHI gradient (∇GHI): This factor measures the instantaneous rate of change in GHI , providing the model with information about recent trends. To strictly prevent data leakage, this gradient is calculated using only historical values available at the time of prediction. It is defined as follows [48]:

$$\nabla GHI(t) = \frac{GHI_{t-1} - GHI_{t-2}}{\nabla t} \quad (3)$$

where ∇t corresponds to the time resolution (e.g., 5, 10, or 60 min). A positive $\nabla GHI(t)$ value indicates a recent increase in irradiance, while a negative value suggests the onset of cloud cover or sunset. Both GHI_{t-1} and GHI_{t-2} are historical values known at the moment of the forecast; thus, this feature is completely safe from data leakage.

4. Clear sky index (K_t): To classify the weather, GHI is normalized to create a clear sky index. The index is calculated from the following [51]:

$$K_t = \frac{GHI}{GHI_{clear}} \quad (4)$$

5. Cloud opacity ($O_{c,t-1}$): This variable measures the cloud density to help determine sky conditions [48]. By using this lagged value, the model checks the sky's previous state to predict sudden solar irradiance fluctuations, making it highly useful for spotting incoming clouds in the volatile weather of the study site.
6. Temporal features (hour, zenith, azimuth): These variables track the sun's physical location. The hour shows the daily cycle of solar intensity. The zenith angle measures the sun's height above the horizon, which controls how much atmosphere the radiation must pass through. The azimuth simply points to the compass direction of the sun. All of these are basic requirements for calculating solar geometry [48]. These parameters were calculated using the specific date and time and geographical coordinates (latitude and longitude) of the study site utilizing standard solar position algorithms.
7. DNI components (DNI_{clear} , DNI_{lag1}): DNI represents the portion of solar radiation coming directly from the solar disk. This parameter provides the forecasting tool with precise details about the solar resources available for direct solar energy, as well as cloud types and thicknesses [52]. DNI_{clear} is the theoretical clear sky DNI and offers a baseline for potential direct irradiance, while the lagged actual DNI (DNI_{t-1}) provides immediate feedback on previous atmospheric transparency. Lagged versions of these components were used to prevent data leakage [48].

To clarify the experimental setup, the forecasting timeline is defined as follows:

1. Current time ($t - 1$): This is the moment when the forecast is made, at which the historical data (e.g., GHI_{t-1} , $O_{c,t-1}$) are available.
2. Target time (t): This is the future time step for which the GHI value is predicted.
3. Forecast horizon: This is the gap between the current time and the target time. Since this study uses datasets with different resolutions, the horizon is equal to the time step of each dataset (i.e., 5, 10, 15, 30, or 60 min).
4. Available input variables: The available input variables include the following:
 - Past data: Stochastic meteorological variables (GHI , DNI , cloud opacity) from time $t - 1$ and earlier.
 - Future data: Deterministic variables (zenith, azimuth, clear sky GHI) calculated for the target time t . These are known in advance because they depend on astronomical equations, not weather.

The final step involves splitting the data into a 70:15:15 percent ratio for training, validation, and testing, respectively, to ensure that the models are evaluated on unseen data, confirming their accuracy and precision. The dataset was split chronologically, rather than randomly, to ensure that all validation and test samples strictly postdate the training data and prevent any form of temporal data leakage.

2.4. Weather Classification Framework

To ensure forecasting accuracy over various weather conditions, the test data were split into specific weather categories. This split was based on the clear sky index (K_t), a ratio

that compares measured *GHI* against the maximum possible clear sky *GHI* to determine sky clarity, given by (4). The dataset was split into three weather categories based on the following K_t thresholds [53,54]:

- Clear sky ($K_t > 0.65$): The sky is mostly clear, allowing maximum sunlight to reach the surface.
- Partly cloudy ($0.3 < K_t \leq 0.65$): Clouds pass frequently, causing sudden and unpredictable changes in solar radiation.
- Cloudy ($K_t \leq 0.3$): Thick clouds block the sun, keeping irradiance levels very low.

Based on these classifications, the trained models were then tested on these specific subsets to see which one performed best in each scenario. The performance metrics (introduced in the next section) directly show the accuracy and reliability of each ML model under these three different weather conditions.

To better understand the dataset's composition, the distribution of weather classes is presented in Table 3. As expected for an arid region, the data is heavily dominated by clear sky conditions, which account for approximately 90% of all daytime samples. In contrast, cloudy periods are relatively rare, representing a small but critical fraction of the data where forecasting is most challenging. The strong dominance of clear sky samples may bias models toward learning stable irradiance patterns, potentially reducing sensitivity to rare cloudy events. Resampling techniques were intentionally avoided in this study to preserve the operational reality of the arid climate, where clear skies are the dominant state. Otherwise, artificially balancing the dataset would have misrepresented the actual distribution of weather events. Instead, the imbalance was addressed through stratified evaluation (as discussed later in Section 4), where performance is assessed separately for each weather class. This imbalance partly can explain the reduced model performance under overcast conditions and highlights the need for targeted modeling or resampling strategies in future work.

Table 3. Distribution of weather classes across resolutions.

Resolution (min)	Clear Sky (%)	Partly Cloudy (%)	Cloudy (%)
5	90.86	5.17	3.97
10	90.85	5.14	4.01
15	81.55	7.43	11.03
30	91.24	4.92	3.84
60	91.19	4.80	4.01

3. ML Models and Performance Metrics

This section discusses the considered ML models in this research and introduces the performance metrics considered in their successful solar irradiance forecasting.

3.1. Evaluated ML Models

This study tests five advanced tree-based algorithm ML models and compares them against a Persistence baseline. These models were chosen because they have been reported to handle large datasets effectively and are highly accurate for solar forecasting [12,13,15]. The selected models are as follows:

1. Random Forest: This ensemble method builds many separate decision trees. Each tree makes a prediction, and the model averages them together to get the final result. This averaging lowers the prediction error and stops the model from overfitting, making it very stable.
2. Gradient Boosting: This model builds decision trees one at a time. Each new tree is trained specifically to fix the mistakes made by the previous ones. This cycle repeats

until the error stops dropping. However, because it keeps trying to fix tiny errors, it can easily overfit if the training data are noisy.

3. Histogram-Based Gradient Boosting (HistGradientBoosting): This model is a faster version of the standard Gradient Boosting model. It works by grouping similar data points into 'bins', which cuts down the processing time. This makes it much quicker to build and test accurate models.
4. XGBoost: This model is a highly efficient boosting algorithm that works well on large datasets [12,16]. It includes built-in regularization that can be adjusted based on the specific problem, reducing the need for early stopping. This helps prevent overfitting and typically leads to better results on new data.
5. LightGBM: This model is a fast Gradient Boosting framework that uses smart techniques like gradient-based one-side sampling and exclusive feature bundling to train much faster and use less memory than XGBoost, especially on large datasets [16]. Instead of growing trees level-by-level, this model grows them leaf-by-leaf. This helps it to find complex patterns quickly, though it requires careful parameter tuning to avoid overfitting.

The above five ML models, which represent the current prevailing paradigms in tree-based ensemble learning, have been compared with a Persistence model (baseline). The Persistence model is a fundamental benchmark method used in time-series forecasting and operates in this research based on the assumption that the GHI value does not change over the forecast horizon, predicting that the future value at time $t + k$ is equal to the current value at time t . Despite its simplicity, the Persistence model can serve as a rigorous baseline, particularly for very-short-term horizons in this research, where weather conditions often exhibit temporal inertia.

This study deliberately focuses on tree-based ensemble models to maintain interpretability, computational efficiency, and strict control over data leakage, which are critical requirements for short-term operational forecasting. It should be highlighted that, while deep learning models such as LSTM and convolutional neural networks have also demonstrated strong performance in some forecasting applications, they typically require larger feature sets, additional feature engineering (e.g., sequence window tuning), extensive tuning, and significantly higher computational costs. Since the objective of this study is a controlled benchmark of tree-based models under strict anti-leakage constraints, such comparisons are beyond the scope of this research and are reserved for future work.

3.2. Experimental Setup and Performance Metrics

To ensure fair comparisons, a clear testing setup was used. As mentioned earlier, for each time resolution, the dataset was split into three groups (i.e., 70% for training the models, 15% for tuning parameters, and 15% for the final test evaluation). To guarantee that the results are reproducible, every model was initialized with a fixed random seed (`random_state = 42`). Furthermore, since tree-based algorithms are structurally stable and the dataset is extensive, a single training run was found sufficient to generate consistent and reliable performance estimates. Splitting the data chronologically keeps the timeline intact, which prevents data leakage and reflects real-world forecasting. However, finding the best hyperparameters using Bayesian optimization takes a huge amount of computing power [55]. Bayesian optimization was chosen to manage this computationally heavy task and is designed to optimize black-box functions where evaluating each setting means retraining a model. The employed Bayesian optimization framework consists of two primary parts:

1. In this step, the model uses a Gaussian process to estimate the validation loss for different hyperparameters. It mixes initial assumptions with actual test results to predict both the expected loss and the uncertainty for any setting.
2. An acquisition function: This function uses the surrogate model's guesses to pick the next hyperparameters to test. The expected improvement method is a common choice for this step [55], given by

$$EI(\vec{x}) = E\left[\max(f_{\min} - f(\vec{x}), 0)\right] \quad (5)$$

where $EI(\vec{x})$ is the expected improvement for the hyperparameter configuration, $\vec{x}$ is the vector of hyperparameters being evaluated, E denotes the expectation operator, $f_{\min}$ is the best observed objective function value (minimum validation error) found so far, and $f(\vec{x})$ is the objective function value (validation error) for the current hyperparameters where the best observed value is defined, and the expectation is calculated over the distribution of x . This step-by-step approach efficiently searches through different hyperparameters until it finds the best settings to minimize the validation error. For example, consider an iteration where the best validation RMSE found thus far is $f_{\min} = 30 \text{ W/m}^2$. The surrogate model evaluates a new hyperparameter set $\vec{x}_{\text{new}}$. It predicts that the expected error for this set is $f(\vec{x}_{\text{new}}) = 28 \text{ W/m}^2$. Equation (5) then calculates the expected improvement (EI), representing the probability that this new set will yield an error lower than 30 W/m^2 . If EI is high, the algorithm selects $\vec{x}_{\text{new}}$ for actual training. The Bayesian optimization process was allocated a fixed budget of 100 iterations for each model to ensure fair comparison. This optimization process was performed independently for each forecasting horizon (5, 10, 15, 30, and 60 min) to ensure that the hyperparameters were tailored to the specific dynamics of each time resolution. The optimized hyperparameter settings obtained from this process for each evaluated model are detailed in Table 4. It should be noted that the values in this table are representative examples for the 60 min horizon only, as listing all configurations for every horizon would be too extensive.

After optimization, these five ML models were tested using three standard metrics, i.e., RMSE, symmetric mean absolute percentage error (sMAPE), and the coefficient of determination (R^2), defined by

$$\begin{aligned} RMSE &= \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \\ sMAPE &= \frac{100}{N} \sum_{i=1}^N \frac{2|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i|} \\ R^2 &= 1 - \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \end{aligned} \quad (6)$$

where N is the total number of data points in the test set, y_i is the actual (observed) GHI value at the time step, $\hat{y}_i$ is the predicted GHI value at the time step, and $\bar{y}$ is the mean of the actual observed GHI values.

To match real-world solar panel operations, all metrics were calculated using only daytime data, where $GHI > 0$. This threshold automatically adjusts for seasonal changes in sunrise and sunset, while still including heavily overcast hours where diffuse radiation is present [56,57].

Using these three metrics gives a complete picture of the models' performance, balancing mathematical accuracy with practical worth. Figure 2 illustrates the flowchart of the proposed methodology, summarizing the experimental workflow, from raw data collection to the final performance evaluation.

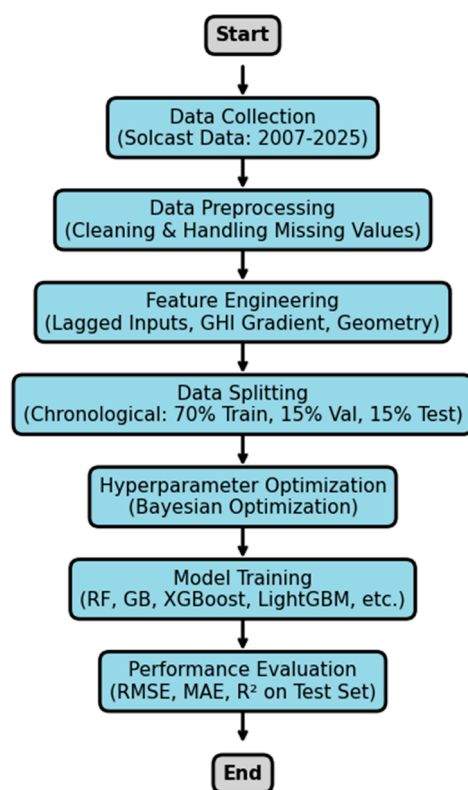


Figure 2. Flowchart of the proposed methodology for multi-resolution solar irradiance forecasting.

Table 4. Representative optimized hyperparameter settings (shown for the 60 min horizon).

Models	Key Parameters	Search Range (Min, Max)	Values
Random Forest	n_estimators	(50, 300)	100
	max_depth	(5, 20)	15
	min_samples_split	(2, 10)	2
	min_samples_leaf	(1, 10)	1
	max_features	(1, 10)	1
Gradient Boosting	learning_rate	(0.01, 0.3)	0.1
	n_estimators	(50, 300)	100
	max_depth	(3, 10)	5
	subsample	(0.7, 1.0)	1
HistGradientBoosting	learning_rate	(0.01, 0.3)	0.1
	max_depth	(5, 20)	15
	max_iter	(50, 300)	100
	min_samples_leaf	(10, 50)	20
XGBoost	learning_rate	(0.01, 0.3)	0.1
	max_depth	(3, 10)	5
	n_estimators	(50, 300)	100
	subsample	(0.7, 1.0)	1
	colsample_bytree	(0.7, 1.0)	1
	reg_alpha	(0, 1.0)	0
	reg_lambda	(0, 5.0)	1
LightGBM	learning_rate	(0.01, 0.3)	0.1
	num_leaves	(20, 50)	31
	max_depth	(3, 10)	5
	feature_fraction	(0.7, 1.0)	1
	bagging_fraction	(0.7, 1.0)	1
	min_data_in_leaf	(10, 50)	20

All experiments in this study were implemented using the Python programming language (version 3.9). The tree-based models were developed using standard libraries: Random Forest, Gradient Boosting, and HistGradientBoosting were implemented using the *scikit-learn* library, while XGBoost and LightGBM utilized their respective native libraries (*xgboost* and *lightgbm*). The Bayesian optimization process for hyperparameter tuning was conducted using the *scikit-optimize* library.

4. Overall Multi-Resolution Performance

The five ML models introduced in Section 3.1 are tested against the Persistence baseline using only daytime data for forecasting horizons of 5 to 60 min. These results represent the overall performance aggregated across all weather conditions and are presented in Table 5. This table lists the corresponding RMSE, MAE, sMAPE and R^2 metrics for each ML model and resolution under the training, validation and test datasets. Figure 3 provides a visual comparison of these results, illustrating the divergence in error trends between the Persistence baseline and the ML models across the different forecasting horizons.

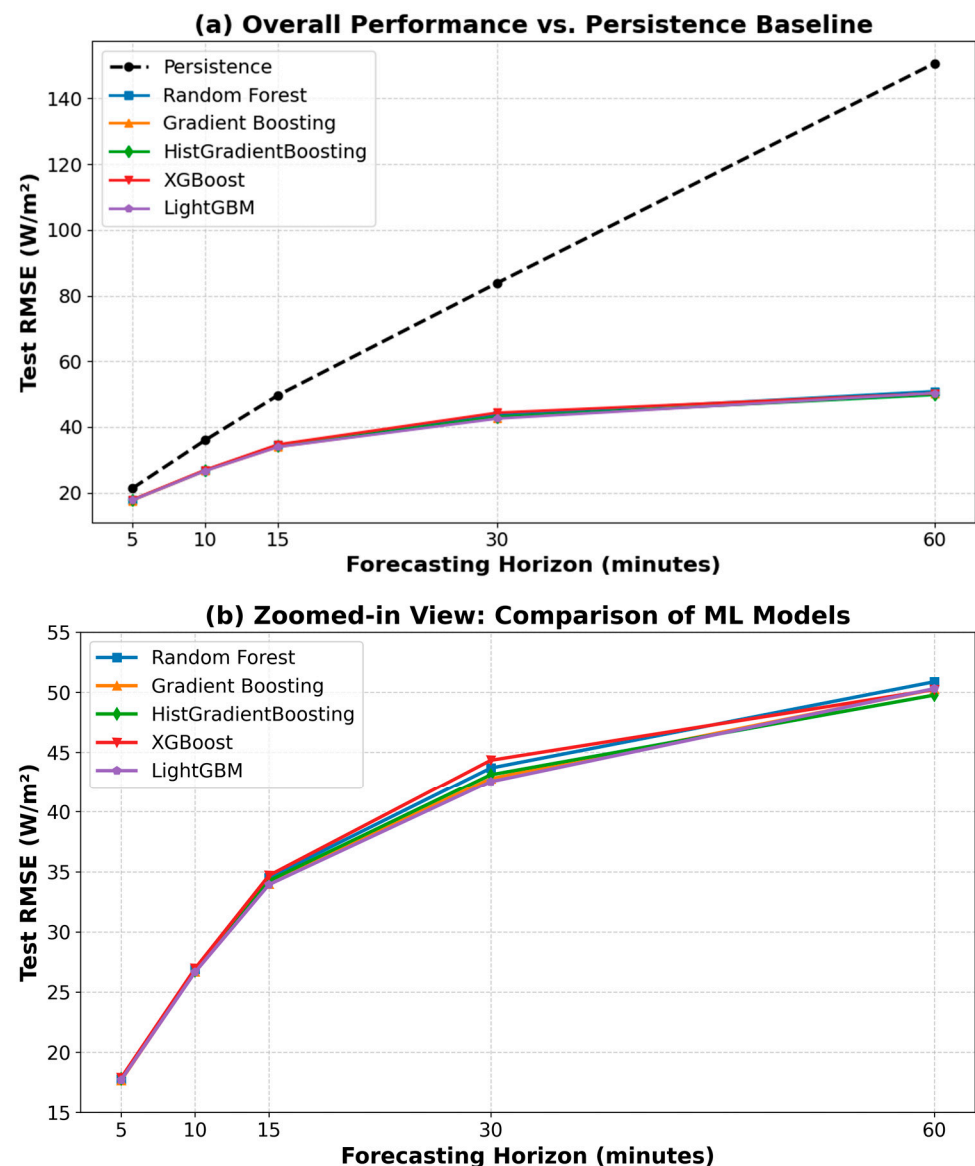


Figure 3. Comparison of model performance across different forecasting horizons: (a) overall RMSE comparison including the Persistence baseline, and (b) zoomed-in view comparing the performance of the ML models.

Table 5. Comparison of RMSE, MAE, sMAPE and R² of the five considered ML models against the baseline model across various horizons.

Resolution (min)	Model	RMSE			MAE			sMAPE			R ²		
		Training	Validation	Test	Training	Validation	Test	Training	Validation	Test	Training	Validation	Test
5	GradientBoosting	10.53	14.82	17.65	4.07	5.28	5.6	2.82	3.71	4.1	0.999	0.998	0.996
	HistGradientBoosting	10.51	14.98	17.78	4.59	5.86	6.19	3.72	4.52	4.89	0.999	0.997	0.996
	LightGBM	10.59	14.84	17.69	4.15	5.35	5.68	3.19	4.03	4.48	0.999	0.998	0.996
	Persistence	15.84	18.88	21.32	12.43	13.45	14.04	9.9	10.53	11.02	0.997	0.996	0.995
	RandomForest	7.69	14.98	17.77	2.62	4.7	4.91	1.38	2.82	3.2	0.999	0.997	0.996
	XGBoost	10.19	15.12	17.86	4.44	5.87	6.2	3.28	4.17	4.52	0.999	0.997	0.996
10	GradientBoosting	19.46	24.07	26.69	7.31	8.74	8.67	4.32	5.16	5.45	0.996	0.993	0.992
	HistGradientBoosting	19.19	24.19	26.77	7.6	9.17	9.12	5.2	6.04	6.3	0.996	0.993	0.992
	LightGBM	19.63	24.02	26.65	7.32	8.73	8.66	4.6	5.31	5.59	0.996	0.994	0.992
	Persistence	30.96	33.72	36.01	24.57	25.62	26.17	17.72	18.38	18.93	0.989	0.987	0.985
	RandomForest	13.28	24.5	26.87	4.67	8.2	7.93	2.38	4.52	4.74	0.998	0.993	0.992
	XGBoost	18.46	24.53	26.95	7.75	9.65	9.56	4.77	5.7	5.98	0.996	0.993	0.992
15	GradientBoosting	27.35	28.54	33.98	10.38	11.13	11.6	5.41	5.95	6.58	0.991	0.991	0.987
	HistGradientBoosting	26.84	28.87	34.25	10.54	11.44	12	6.33	6.91	7.35	0.992	0.991	0.987
	LightGBM	27.72	28.57	33.93	10.43	11.11	11.57	5.61	6.06	6.59	0.991	0.991	0.987
	Persistence	45.9	45.35	49.7	36.63	36.52	38.05	24.25	24.72	25.26	0.976	0.977	0.972
	RandomForest	17.92	29.54	34.51	6.38	10.71	10.91	3.17	5.64	6	0.996	0.99	0.987
	XGBoost	25.2	30.17	34.69	10.55	12.44	12.7	6.01	6.79	7.28	0.993	0.99	0.986
30	GradientBoosting	36.62	42.33	42.8	14.93	16.57	15.5	7.12	7.93	8.03	0.985	0.98	0.98
	HistGradientBoosting	36.65	42.75	43.13	15.13	16.84	15.82	7.94	8.32	8.37	0.985	0.98	0.98
	LightGBM	37.44	42.18	42.55	15.08	16.46	15.4	7.28	7.83	7.91	0.984	0.98	0.98
	Persistence	81.02	82.5	83.82	68.46	69.67	70.26	40.11	40.13	40.5	0.924	0.924	0.923
	RandomForest	20.63	42.91	43.69	8.16	16.05	14.93	3.95	7.41	7.34	0.995	0.98	0.979
	XGBoost	32.7	43.9	44.34	14.58	17.92	16.96	7.3	8.43	8.43	0.988	0.979	0.979
60	GradientBoosting	42.39	50.5	50.26	17.36	20.37	18.7	8.07	9.1	8.92	0.98	0.972	0.972
	HistGradientBoosting	41.07	50.45	49.77	17.3	20.55	18.81	9.23	9.57	9.23	0.981	0.972	0.973
	LightGBM	43.61	50.12	50.31	17.66	20.28	18.81	8.26	9.12	8.95	0.979	0.972	0.972
	Persistence	145.44	148.53	150.75	125.74	128.63	130.49	64.15	62.78	62.3	0.762	0.755	0.751
	RandomForest	22.78	50.93	50.88	9.18	20.05	18.31	4.16	8.8	8.58	0.994	0.971	0.972
	XGBoost	36.71	51.29	50.23	16.34	21.49	19.36	8.24	9.68	9.3	0.985	0.971	0.972

Looking at the absolute errors in Table 5, the ML models clearly outperform the baseline as the forecast horizon gets larger (longer). At 5 min horizons, the Gradient Boosting model demonstrates the lowest RMSE (17.65 W/m^2) compared to the Persistence model (21.32 W/m^2). At 60 min horizons, the Persistence model fails, with its RMSE hitting 150.75 W/m^2 . However, the ML models do not suffer from this issue; the HistGradient-Boosting kept the 60 min RMSE at just 49.77 W/m^2 . This trend is supported by the R^2 values where Persistence drops to 0.751 at 60 min, while the ML models maintain an R^2 above 0.972.

From Table 5, it can also be seen that the Random Forest model is prone to overfitting. For example, at the 60 min horizon, its training R^2 is a perfect 0.994, but drops to 0.972 on the test set. This means that the model has memorized and is biased towards the training data. This behavior is attributed to Random Forest's tendency to fully grow trees and capture noise in highly autocorrelated time-series data. In contrast, boosting models apply sequential regularization, which improves generalization. As such, the Gradient Boosting variants handle this much better. The LightGBM performance, for instance, only drops from 0.979 to 0.972 between its training and testing, proving it generalizes well to the new (untrained) data.

5. Model Performance Under Distinct Weather Conditions

Because solar output changes drastically depending on cloud cover, the models are tested across three specific sky states, i.e., clear, partly cloudy, and overcast. Table 6 breaks down the error metrics for these categories across all time horizons. In this table, 'Overall (Daytime)' combines all weather conditions to show a generalized baseline of accuracy. This is highly useful because a grid operator does not always know exactly when a cloud will appear, and thus they need a model that performs reliably across mixed conditions.

It should be highlighted that the negative R^2 values observed under fully overcast conditions in Table 6 do not indicate numerical instability but rather reflect the absence of a predictable signal in highly stochastic cloud-driven irradiance dynamics. Furthermore, this degradation is likely exacerbated by the uncertainties inherent in satellite-derived data. This is because, when satellite algorithms struggle to distinguish between thick clouds and dust aerosols, the input features become noisy, hampering the ability of tree-based models to learn accurate patterns. Under such conditions, even simple statistical baselines may outperform learned models. In this context, Persistence acts as the 'most reliable option', 'the least error-prone choice' or 'the least poor' option, but it cannot be classified as 'a highly accurate forecasting solution'. Therefore, it can be treated as a necessary reliability safeguard against the large errors that occur when complex models fail to capture chaotic patterns. This highlights the physical unpredictability of irradiance rather than shortcomings of the modeling framework.

The RMSE performance for these different weather scenarios is visually compared in Figures 4–6. As shown in Figure 4, clear skies presented almost no forecasting difficulty. The ML model bars stay very low and flat across all time resolutions. All models were highly accurate, with the Gradient Boosting model achieving the lowest error at 5 min horizons ($\text{RMSE} = 13.30 \text{ W/m}^2$). The Persistence model results also stay very close to the ML models at the 5 min horizon ($\text{RMSE} = 18.20 \text{ W/m}^2$). This visual gap shows that, on stable, sunny days in arid climates, a simple baseline model is still highly competitive.

Table 6. Model performance under different weather conditions.

Resolution (min)	Model	Weather	RMSE	MAE	sMAPE	R ²
5	Persistence	Overall (Daytime)	21.32	14.04	11.02	0.995
		Clear Sky	18.2	13.13	9.25	0.996
		Partly Cloudy	45.15	26.24	29.26	0.881
		Cloudy	32.02	17.36	27.42	0.683
	RandomForest	Overall (Daytime)	17.77	4.91	3.2	0.996
		Clear Sky	13.33	3.12	1.56	0.998
		Partly Cloudy	44.8	25.1	17.01	0.883
		Cloudy	33.23	18.2	23.84	0.659
	GradientBoosting	Overall (Daytime)	17.65	5.6	4.1	0.996
		Clear Sky	13.3	3.89	2.31	0.998
		Partly Cloudy	44.24	24.84	18.24	0.886
		Cloudy	33.04	18.46	28.45	0.662
	HistGradientBoosting	Overall (Daytime)	17.78	6.19	4.89	0.996
		Clear Sky	13.47	4.54	3.09	0.998
		Partly Cloudy	44.25	24.82	20.28	0.886
		Cloudy	33.17	18.41	27.31	0.66
	XGBoost	Overall (Daytime)	17.7	5.72	4.18	0.996
		Clear Sky	13.35	4.03	2.43	0.998
		Partly Cloudy	44.29	24.86	18.15	0.885
		Cloudy	33.1	18.4	27.94	0.661
	LightGBM	Overall (Daytime)	17.69	5.68	4.48	0.996
		Clear Sky	13.34	3.99	2.67	0.998
		Partly Cloudy	44.28	24.82	19.98	0.885
		Cloudy	33.07	18.34	27.16	0.662
10	Persistence	Overall (Daytime)	36.01	26.17	18.93	0.985
		Clear Sky	32.06	25.08	16.48	0.988
		Partly Cloudy	67.4	41.46	44.22	0.735
		Cloudy	54.05	29	41.22	0.099
	RandomForest	Overall (Daytime)	26.87	7.93	4.74	0.992
		Clear Sky	20.08	5.08	2.37	0.995
		Partly Cloudy	64.13	37.76	23.36	0.76
		Cloudy	59	33.43	37.13	−0.074
	GradientBoosting	Overall (Daytime)	26.69	8.67	5.45	0.992
		Clear Sky	19.8	5.89	2.96	0.995
		Partly Cloudy	64.26	37.65	24.25	0.759
		Cloudy	58.85	33.57	40.76	−0.069
	HistGradientBoosting	Overall (Daytime)	26.77	9.12	6.3	0.992
		Clear Sky	19.86	6.37	3.67	0.995
		Partly Cloudy	64.18	37.7	27.48	0.759
		Cloudy	59.47	33.91	41.21	−0.091
	XGBoost	Overall (Daytime)	26.61	8.7	5.41	0.992
		Clear Sky	19.81	5.96	2.91	0.995
		Partly Cloudy	63.9	37.45	24.57	0.761
		Cloudy	58.3	33.15	40.27	−0.049
	LightGBM	Overall (Daytime)	26.65	8.66	5.59	0.992
		Clear Sky	19.81	5.9	3.16	0.995
		Partly Cloudy	64.17	37.66	24.57	0.759
		Cloudy	58.51	33.04	39.18	−0.056

Table 6. Cont.

Resolution (min)	Model	Weather	RMSE	MAE	sMAPE	R ²
15	Persistence	Overall (Daytime)	49.7	38.05	25.26	0.972
		Clear Sky	45.16	37.38	14.52	0.972
		Partly Cloudy	81.55	56.32	37.77	0.653
		Cloudy	50.13	28.53	101.33	0.192
	RandomForest	Overall (Daytime)	34.51	10.91	6	0.987
		Clear Sky	23.98	6.23	1.86	0.992
		Partly Cloudy	71.34	42.29	19.12	0.734
		Cloudy	54.36	22.86	28.6	0.05
	GradientBoosting	Overall (Daytime)	33.98	11.6	6.58	0.987
		Clear Sky	23.1	7.08	2.18	0.993
		Partly Cloudy	71.58	42.35	19.23	0.733
		Cloudy	53.84	22.76	31.63	0.068
	HistGradientBoosting	Overall (Daytime)	34.25	12	7.35	0.987
		Clear Sky	23.28	7.41	2.33	0.993
		Partly Cloudy	71.66	42.87	19.88	0.732
		Cloudy	54.77	23.56	37.47	0.035
	XGBoost	Overall (Daytime)	33.9	11.66	6.89	0.987
		Clear Sky	23.14	7.17	2.21	0.993
		Partly Cloudy	71.26	42.09	19.13	0.735
		Cloudy	53.51	22.86	34.52	0.079
	LightGBM	Overall (Daytime)	33.93	11.57	6.59	0.987
		Clear Sky	23.2	7.08	2.18	0.993
		Partly Cloudy	71.23	42.21	19.24	0.735
		Cloudy	53.59	22.6	31.72	0.076
30	Persistence	Overall (Daytime)	83.82	70.26	40.5	0.923
		Clear Sky	81.2	70.47	37.57	0.925
		Partly Cloudy	113.24	76.68	75.81	0.28
		Cloudy	96.33	53.3	62.49	−2.071
	RandomForest	Overall (Daytime)	43.69	14.93	7.34	0.979
		Clear Sky	30.7	9.83	3.77	0.989
		Partly Cloudy	100.62	62	32.88	0.432
		Cloudy	123.14	78.38	64.5	−4.018
	GradientBoosting	Overall (Daytime)	42.8	15.5	8.03	0.98
		Clear Sky	29.56	10.54	4.33	0.99
		Partly Cloudy	100	61.11	34.91	0.439
		Cloudy	122.11	77.52	66.77	−3.934
	HistGradientBoosting	Overall (Daytime)	43.13	15.82	8.37	0.98
		Clear Sky	29.52	10.73	4.67	0.99
		Partly Cloudy	100.26	62.1	35.36	0.436
		Cloudy	125.55	80.17	66.79	−4.217
	XGBoost	Overall (Daytime)	42.56	15.46	8.01	0.98
		Clear Sky	29.31	10.5	4.33	0.99
		Partly Cloudy	99.71	61.01	34.73	0.442
		Cloudy	121.7	77.48	66.29	−3.901
	LightGBM	Overall (Daytime)	42.55	15.4	7.91	0.98
		Clear Sky	29.43	10.43	4.26	0.99
		Partly Cloudy	100.39	61.57	33.99	0.434
		Cloudy	119.8	76.92	66.53	−3.75

Table 6. Cont.

Resolution (min)	Model	Weather	RMSE	MAE	sMAPE	R ²
60	Persistence	Overall (Daytime)	150.75	130.49	62.3	0.751
		Clear Sky	152.2	134.62	59.54	0.735
		Partly Cloudy	144.53	102.63	99.6	−0.224
		Cloudy	117.26	63.51	76.73	−3.515
	RandomForest	Overall (Daytime)	50.88	18.31	8.58	0.972
		Clear Sky	37.05	12.48	4.56	0.984
		Partly Cloudy	111.9	70.45	36.06	0.266
		Cloudy	139.71	92.16	73.11	−5.41
	GradientBoosting	Overall (Daytime)	50.26	18.7	8.92	0.972
		Clear Sky	36.17	12.85	4.8	0.985
		Partly Cloudy	111.97	70.34	36.5	0.265
		Cloudy	139.2	93.66	76.19	−5.363
	HistGradientBoosting	Overall (Daytime)	49.77	18.81	9.23	0.973
		Clear Sky	35.95	13.05	5.12	0.985
		Partly Cloudy	111.27	70.3	37.98	0.274
		Cloudy	136.39	91.68	74.4	−5.109
	XGBoost	Overall (Daytime)	50.21	18.72	8.9	0.972
		Clear Sky	36.23	12.91	4.77	0.985
		Partly Cloudy	111.82	70.06	36.76	0.267
		Cloudy	138.4	93.25	75.86	−5.29
	LightGBM	Overall (Daytime)	50.31	18.81	8.95	0.972
		Clear Sky	36.32	13	4.88	0.985
		Partly Cloudy	111.5	69.8	36.29	0.271
		Cloudy	139.33	93.87	75.31	−5.375

Moving to more volatile weather, Figure 5 shows that partly cloudy skies present a much tougher operational challenge due to sudden cloud shadows dropping the GHI. The Persistence model perfectly illustrates this failure, shooting straight up to 144.53 W/m² at the 60 min horizons. However, the Gradient Boosting models handle this volatility much better. HistGradientBoosting cuts the RMSE of the 60 min horizon down to 111.27 W/m². While accuracy naturally drops at extended horizons under these volatile conditions ($R^2 \approx 0.27$), the boosting models track the rapid fluctuations slightly better than the Random Forest model. This makes them highly valuable for real-time grid balancing during sudden clouding events.

Ultimately, under the fully overcast conditions shown in Figure 6, it can be seen that all models struggle to find a mathematical pattern. An interesting trend appears at the 5 min horizons: the Persistence baseline model beats the ML models in RMSE. However, as the forecast extends, it fails, dropping to negative R^2 scores at 30 and 60 min horizons (−2.07 and −3.51). The complex ML models fail in the opposite way. They maintain positive R^2 scores at 5 min horizons, but their RMSE values explode at 60 min horizons (e.g., the Random Forest model predicts 139.71 W/m²). This collapse in R^2 is a known characteristic of chaotic weather data. Because GHI fluctuates randomly under thick clouds, even predicting the historical average becomes highly inaccurate. A negative R^2 here simply proves that no mathematical pattern could have been found in the chaotic cloud movements, making the predictions effectively random. Consequently, RMSE remains the primary metric for assessing reliability under these conditions, as it quantifies the absolute error magnitude (W/m²) without being skewed by the low variance inherent to overcast periods. Therefore, under such chaotic conditions, Persistence is an option rather than a truly accurate tool, serving as a reliability safeguard when complex models fail.

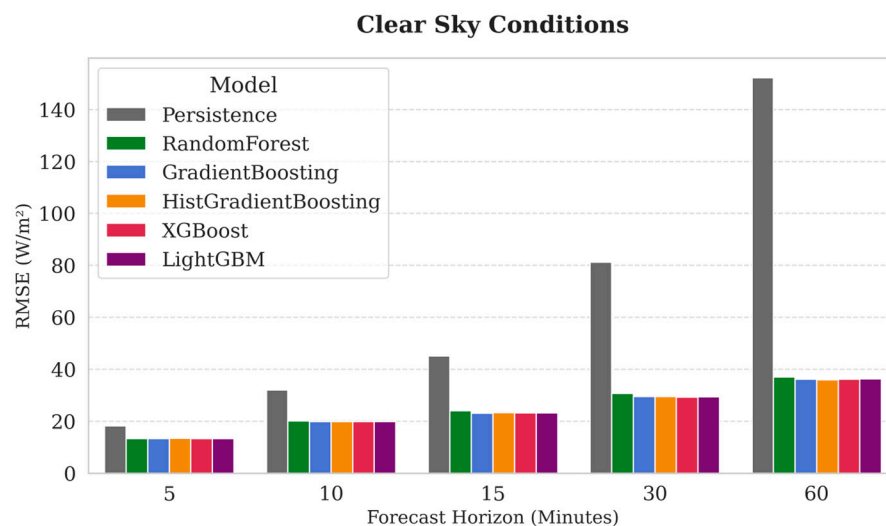


Figure 4. RMSE results for clear sky conditions.

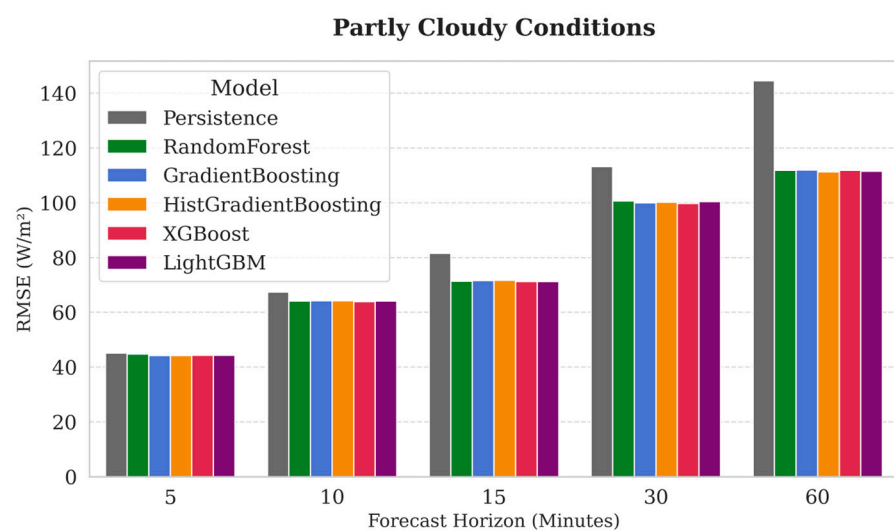


Figure 5. RMSE results for partly cloudy conditions.

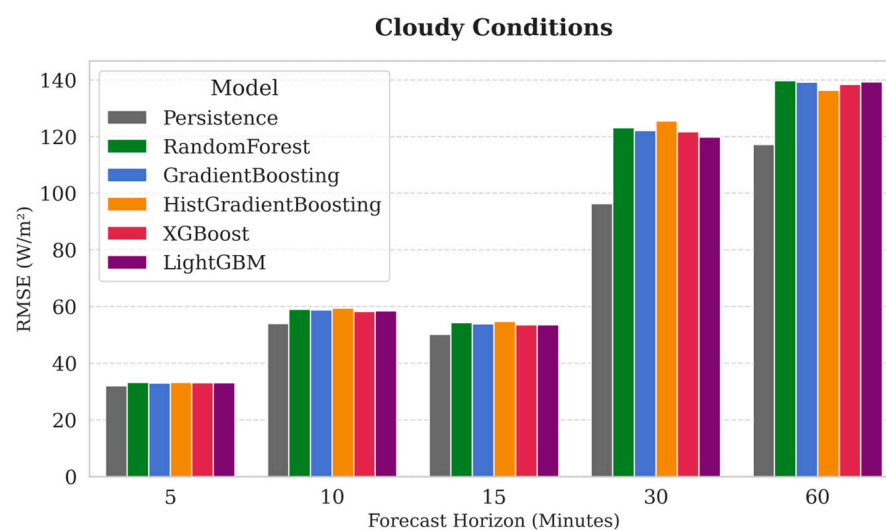


Figure 6. RMSE results for cloudy conditions.

5.1. Comparison with Existing Studies

To critically evaluate the performance of the proposed tree-based models, we compared this study’s results with a recent comprehensive study by [24]. That study, conducted in Morocco (a similar semi-arid region), benchmarked ANN against tree-based models like Random Forest and XGBoost, finding that ANN provided the highest accuracy for daily solar forecasting.

The current study complements these findings by focusing on short-term horizons (5 to 60 min) rather than daily averages. While [24] reported ANN as the superior model, this study’s results demonstrate that tree-based models, particularly HistGradientBoosting, achieve exceptional accuracy ($R^2 = 0.973$ for the 60 min horizon) in an arid environment. The slight performance edge of ANNs reported in the literature for daily data often comes with higher computational costs and ‘black-box’ complexity. Conversely, the current findings validate that, for real-time, minute-scale forecasting, tree-based models offer a highly efficient and interpretable alternative, providing near-optimal accuracy with significantly lower latency, which is crucial for operational grid stability.

5.2. Limitations and Future Work

It is important to acknowledge that this study deliberately focused on a single arid location (Basra, Iraq); therefore, the quantitative error metrics reported are site-specific. However, the relative performance trends observed among the tree-based models are expected to remain valid for other arid regions characterized by similar atmospheric dynamics, such as high aerosol loads and distinct clear sky dominance. Future work should focus on validating this benchmarking framework across multiple geographic locations to generalize the findings further. Additionally, investigating hybrid modeling approaches could potentially improve forecasting accuracy under the highly stochastic conditions of fully overcast skies, where current tree-based models struggle.

6. Discussion and Conclusions

The benchmark analysis that was conducted confirmed that no single ML model is best for every scenario. The right choice depends heavily on the analyzed time horizon and the weather conditions. Table 7 summarizes the top-performing models from the five studied ML models for each specific scenario. For example, it can be concluded from this table that the Gradient Boosting model dominates in the overall results at 5 min horizons, while the XGBoost model takes the overall lead at the 15 and 30 min horizons.

Table 7. Determined best model per condition.

Resolution (min)	Clear Sky	Cloudy	Overall (Daytime)	Partly Cloudy
5	GradientBoosting	Persistence	GradientBoosting	GradientBoosting
10	GradientBoosting	Persistence	XGBoost	XGBoost
15	GradientBoosting	Persistence	XGBoost	LightGBM
30	XGBoost	Persistence	LightGBM	XGBoost
60	HistGradientBoosting	Persistence	HistGradientBoosting	HistGradientBoosting

As was highlighted earlier, the negative R^2 values observed under fully overcast conditions reflect the absence of a predictable signal in highly stochastic cloud-driven irradiance dynamics, and, in such conditions, even the simple statistical baseline model outperforms the learned models, acting as a reliability safeguard rather than a truly accurate tool, highlighting the physical unpredictability of irradiance.

Based on these specific results, the following decision framework can be proposed for electrical grid operators when selecting a suitable ML model for the more accurate forecasting of solar generation across their networks at various horizons:

- Very-short-term (5–15 min) horizons: The Gradient Boosting and XGBoost models provide the highest overall accuracy during this timeframe. The Persistence model remains a strong alternative for clear days, but ML models are necessary to handle passing clouds.
- Short-term (30–60 min): The HistGradientBoosting and XGBoost models are the top choices here. At a 60 min resolution, HistGradientBoosting drops the overall RMSE from 150.75 W/m² down to 49.77 W/m². This translates to an error reduction of roughly 67% compared to the Persistence baseline.
- Weather-dependent strategy: Model selection must change based on the sky and weather conditions. For partly cloudy conditions, the best model shifts depending on the exact forecast window (e.g., LightGBM at 15 min resolutions or the HistGradientBoosting model at 60 min resolutions). However, for fully overcast days, the data clearly shows that relying on a simple Persistence model is the safest option to avoid ML models chasing chaotic noise.

In an operational setting, this adaptive selection is feasible by calculating the clear sky index in real time using on-site pyranometer measurements and theoretical clear sky models. This instantaneous classification allows the control system to dynamically switch to the appropriate forecasting algorithm based on current sky conditions. As such, from an operational perspective, it can be concluded that grid operators are able to and should dynamically select forecasting algorithms based on both the forecast horizon and the prevailing sky conditions rather than relying on a single universal model. From a computational standpoint, the inference time for the top-performing models (e.g., HistGradientBoosting) was found to be minimal (typically less than one second on a standard CPU), confirming their suitability for real-time grid applications where forecasts must be generated rapidly.

In summary, this paper aimed to identify the best ML model for predicting GHI parameters in arid environments, and to see if the ideal model shifts across different time frames. After testing multiple models across five distinct horizons, clear patterns emerged to answer both points. The numerical results showed that the best model changes depending on the time resolutions and the weather. For 5 min forecasts, the Gradient Boosting model provided the highest accuracy. As the forecast extended to 10 to 60 min, specific boosting variants—particularly XGBoost and HistGradientBoosting—took the lead. The Random Forest model, on the other hand, struggled with overfitting and consistently produced the highest errors at these longer horizons.

Furthermore, the models reacted differently depending on the sky conditions. The Gradient Boosting model dominated in clear sky scenarios across the board. In contrast, the partly cloudy tests showed that no single boosting model was best for every time resolution. The Gradient Boosting model is better at 5 min resolutions, but the advantage shifted to other models—like XGBoost, LightGBM, and finally HistGradientBoosting—as the forecast extended to 60 min. As such, the results proved that grid operators should avoid a one-size-fits-all approach. By matching the algorithm to the current sky conditions and time resolution, they can make more accurate real-time decisions. However, these results also show a major flaw: when it is fully overcast, complex ML models lose all their advantage. For cloudy skies, they perform worse than a simple Persistence baseline for any forecast longer than 5 min.

It should be noted again that accurate solar forecasting relies on matching the algorithm to the specific weather and time resolution. Since this study looked at the hot, arid climate of Basra, the findings of this research are valid for such environmental conditions.

Author Contributions: Conceptualization, H.A.H.A.-H.; methodology, H.A.H.A.-H.; software, H.A.H.A.-H.; validation, F.S. and S.A.C.; formal analysis, A.Y. and H.W.; resources, H.A.H.A.-H. and F.S.; data curation, H.A.H.A.-H.; writing—original draft preparation, H.A.H.A.-H. and F.S.; writing—review and editing, S.A.C., A.Y. and H.W.; visualization, H.A.H.A.-H.; supervision, A.Y. and H.W.; project administration, F.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw meteorological data presented in this study are publicly available from the Solcast platform (<https://solcast.com/>). The Python code and processed data used to generate the results are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial neural network
DNI	Direct normal irradiance
GHI	Global horizontal irradiance
LSTM	Long short-term memory
MAE	Mean absolute error
ML	Machine learning
MSE	Mean squared error
R ²	Coefficient of determination
RMSE	Root mean squared error
sMAPE	Symmetric mean absolute percentage error

References

- Schaeffer, R.; Schipper, E.L.F.; Ospina, D.; Mirazo, P.; Alencar, A.; Anvari, M.; Artaxo, P.; Biresselioglu, M.E.; Blome, T.; Boeckmann, M.; et al. Ten New Insights in Climate Science 2024. *One Earth* **2025**, *8*, 101285. [CrossRef]
- Chatterjee, S.; Khan, P.W.; Byun, Y.C. Recent Advances and Applications of Machine Learning in the Variable Renewable Energy Sector. *Energy Rep.* **2024**, *12*, 5044–5065. [CrossRef]
- Ye, J.; Chen, J. The Ultimate Meteorological Question from Observational Astronomers: How Good Is the Cloud Cover Forecast? *Mon. Not. R. Astron. Soc.* **2013**, *428*, 3288–3294.
- Shi, C.; Wang, T.; Wang, G.; Letu, H. The Net Warming Effect of Clouds on Global Surface Temperature May Be Weakening or Even Disappearing. *Geosci. Front.* **2025**, *16*, 102107. [CrossRef]
- Adar, M.; Babay, M.A.; Boussif, M.; Khaouch, Z.; Abbassi, Z.; Najih, Y.; Mabrouki, M. Optimization of Photovoltaic System Modelling: A Comparative Study and Experimental Validation Using Bond Graph Methodology and a Genetic Algorithm. In *Applied Mathematics, Modeling and Computer Simulation*; IOS Press: Amsterdam, The Netherlands, 2024; pp. 723–730.
- Kut, P.; Pietrucha-Urbanik, K. Forecasting Short-Term Photovoltaic Energy Production to Optimize Self-Consumption in Home Systems Based on Real-World Meteorological Data and Machine Learning. *Energies* **2025**, *18*, 4403.
- El-Amarty, N.; El Fadili, H.; Bennani, S.D. Accurate Short-Term Solar Irradiance Forecasting with TinyML on Edge Device. In Proceedings of the 2024 International Conference on Circuit, Systems and Communication (ICCS), Fez, Morocco, 28–29 June 2024; pp. 1–6.
- Sward, J.A.; Ault, T.R.; Zhang, K.M. Genetic Algorithm Selection of the Weather Research and Forecasting Model Physics to Support Wind and Solar Energy Integration. *Energy* **2022**, *254*, 124367. [CrossRef]
- Chen, D.; Shi, X.; Jiang, M.; Zhu, S.; Zhang, H.; Zhang, D.; Chen, Y.; Yan, J. Selecting Effective NWP Integration Approaches for PV Power Forecasting with Deep Learning. *Sol. Energy* **2025**, *301*, 113939. [CrossRef]
- Zhang, F.; Hong, X.; Zhao, Z.; Gan, Z.; Ouyang, P.; Xiao, H.; Zhang, R.; Wei, X.; Cai, M.; Lu, F. Short-Term Forecasting of Cloud Physical Properties Based on Fourier Neural Operator Method. *Geophys. Res. Lett.* **2026**, *53*, e2025GL119553.

11. Mohanty, P.; Subhadarshini, K.; Nayak, R.; Pati, U.C.; Mahapatra, K. Exploring Data-Driven Multivariate Statistical Models for the Prediction of Solar Energy. In *Computer Vision and Machine Intelligence for Renewable Energy Systems*; Elsevier: Amsterdam, The Netherlands, 2025; pp. 85–101.
12. Aksoy, N.; Genc, I. Predictive Models Development Using Gradient Boosting Based Methods for Solar Power Plants. *J. Comput. Sci.* **2023**, *67*, 101958. [\[CrossRef\]](#)
13. Nguyen, H.N.; Tran, Q.T.; Ngo, C.T.; Nguyen, D.D.; Tran, V.Q. Solar Energy Prediction Through Machine Learning Models: A Comparative Analysis of Regressor Algorithms. *PLoS ONE* **2025**, *20*, e0315955. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Hassija, V.; Chamola, V.; Mahapatra, A.; Singal, A.; Goel, D.; Huang, K.; Scardapane, S.; Spinelli, I.; Mahmud, M.; Hussain, A. Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cogn. Comput.* **2024**, *16*, 45–74. [\[CrossRef\]](#)
15. Azman, M.A.; Jantan, H.; Bahrin, U.F.M.; Kadir, E.A. Solar Power Production Forecasting Model Using Random Forest Algorithm. In *Intelligent Systems Design and Applications*; Springer: Cham, Switzerland, 2023; pp. 135–144.
16. Vargas, J.; Martinez, R.; Loo, L. Enhancing Photovoltaic Energy Forecasting with Machine Learning: A Comparison Study of XGBoost, LightGBM and LSTM Models. In Proceedings of the 2024 IEEE Latin American Conference on Computational Intelligence (LA-CCI), Bogota, CO, USA, 13–15 November 2024; pp. 1–6.
17. Costa, T.; Falcão, B.; Mohamed, M.A.; Annuk, A.; Marinho, M. Employing Machine Learning for Advanced Gap Imputation in Solar Power Generation Databases. *Sci. Rep.* **2024**, *14*, 23801. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Singh, R.; Singh, S.; Gupta, S.; Alotaibi, M.A.; Malik, H. Forecasting Rooftop Photovoltaic Solar Power Using Machine Learning Techniques. *Energy Rep.* **2025**, *13*, 3616–3630. [\[CrossRef\]](#)
19. Mellit, A.; Massi Pavan, A.; Ogliairi, E.; Leva, S.; Lughi, V. Advanced Methods for Photovoltaic Output Power Forecasting: A Review. *Appl. Sci.* **2020**, *10*, 487. [\[CrossRef\]](#)
20. Kina, C.; Tanyildizi, H.; Al Bakri Abdullah, M.M.; Razak, R.A.; Imjai, T. Comparison of Deep LSTM and Machine Learning Models for Predicting Compressive Strength of Fly Ash/Slag-Based Geopolymer Concrete. *Sci. Rep.* **2025**, *15*, 32871. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Levent, I.; Sahin, G.; Isik, G.; van Sark, W.G. Comparative Analysis of Advanced Machine Learning Regression Models with Advanced Artificial Intelligence Techniques to Predict Rooftop PV Solar Power Plant Efficiency Using Indoor Solar Panel Parameters. *Appl. Sci.* **2025**, *15*, 3320. [\[CrossRef\]](#)
22. Ramirez-Rivera, F.A.; Guerrero-Rodriguez, N.F. Ensemble Learning Algorithms for Solar Radiation Prediction in Santo Domingo: Measurements and Evaluation. *Sustainability* **2024**, *16*, 8015. [\[CrossRef\]](#)
23. Saigustia, C.; Pijarski, P. Time Series Analysis and Forecasting of Solar Generation in Spain Using eXtreme Gradient Boosting: A Machine Learning Approach. *Energies* **2023**, *16*, 7618. [\[CrossRef\]](#)
24. Ledmaoui, Y.; El Maghraoui, A.; El Aroussi, M.; Saadane, R.; Chebak, A.; Chehri, A. Forecasting solar energy production: A comparative study of machine learning algorithms. *Energy Rep.* **2023**, *10*, 1004–1012. [\[CrossRef\]](#)
25. Van Poecke, A.; Tabari, H.; Hellinckx, P. Unveiling the Backbone of the Renewable Energy Forecasting Process: Exploring Direct and Indirect Methods and Their Applications. *Energy Rep.* **2024**, *11*, 544–557.
26. Pan, F. Forecasting Solar Energy Generation Using Machine Learning Techniques and Hybrid Models Optimized by War SO. *Informatica* **2025**, *49*.
27. Onaiwu, G.E.; Ayidu, J.N. Impact of Aerosols on Climate Change and Radiative Forcing. *Environ. Policy Rev.* **2025**. [\[CrossRef\]](#)
28. Gawusu, S.; Zhang, X.; Jamatutu, S.A.; Yeboah, E.K.; Tizumah, M.W.; Yakubu, S. Adaptive Solar Energy Modeling for Sustainable Urban Infrastructure: Addressing Non-Linear Conversion Challenges. *Environ. Dev. Sustain.* **2025**, 1–35. [\[CrossRef\]](#)
29. Sharma, H.; Kaur, S. Deep Learning for Sustainable Development Across Climate, Energy, Agriculture and Urban Systems. *Discov. Sustain.* **2025**, *6*, 1408. [\[CrossRef\]](#)
30. Abdessadak, A.; Ghennioui, H.; El Bhiri, B.; Thirion-Moreau, N.; Abraim, M.; Merzouk, S. Assessing the Effects of Dust on Solar Panel Performance: A Comprehensive Review and Future Directions. *Eng. Proc.* **2025**, *112*, 9. [\[CrossRef\]](#)
31. Amara, M.B.; Rdhaounia, E.; Balghouthi, M. Adaptive Solar Irradiance Forecasting in Arid Regions: Enhancing Accuracy with Localized Atmospheric Adjustments. *J. Eng. Res.* **2024**, *13*, 2663–2679. [\[CrossRef\]](#)
32. Mardani, M.; Hoseinzadeh, S.; Garcia, D.A. Developing Particle-Based Models to Predict Solar Energy Attenuation Using Long-Term Daily Remote and Local Measurements. *J. Clean. Prod.* **2024**, *434*, 139690. [\[CrossRef\]](#)
33. Atiea, M.A.; Shaheen, A.M.; Alassaf, A.; Alsaleh, I. Enhanced Solar Power Prediction Models with Integrating Meteorological Data Toward Sustainable Energy Forecasting. *Int. J. Energy Res.* **2024**, *2024*, 8022398. [\[CrossRef\]](#)
34. Sutarna, N.; Tjahyadi, C.; Oktivasari, P.; Dwiyaniti, M.; Tohazen. Feature Optimization for Short-Term Solar Power Forecasting using Bidirectional LSTM Networks. In Proceedings of the 2024 7th International Conference of Computer and Informatics Engineering (IC2IE), Bali, Indonesia, 12–13 September 2024; pp. 1–6.
35. Ali, M.; Souahlia, A.; Rabahi, A.; Guermoui, M.; Teta, A.; Tibermachine, I.E.; Rabahi, A.; Benghanem, M. A Robust Deep Learning Approach for Photovoltaic Power Forecasting Based on Feature Selection and Variational Mode Decomposition. *J. Niger. Soc. Phys. Sci.* **2025**, *7*, 2795. [\[CrossRef\]](#)

36. Zhang, X.; Bose, I. Reliability Estimation for Individual Predictions in Machine Learning Systems: A Model Reliability-Based Approach. *Decis. Support Syst.* **2024**, *186*, 114305. [CrossRef]
37. Hewamalage, H.; Ackermann, K.; Bergmeir, C. Forecast Evaluation for Data Scientists: Common Pitfalls and Best Practices. *Data Min. Knowl. Discov.* **2023**, *37*, 788–832. [PubMed]
38. Jailani, N.L.M.; Dhanasegaran, J.K.; Alkaws, G.; Alkahtani, A.A.; Phing, C.C.; Baashar, Y.; Capretz, L.F.; Al-Shetwi, A.Q.; Tiong, S.K. Investigating the Power of LSTM-Based Models in Solar Energy Forecasting. *Processes* **2023**, *11*, 1382. [CrossRef]
39. Sales, V.G.; Strobl, E.; Elliott, R.J. Cloud Cover and Its Impact on Brazil's Deforestation Satellite Monitoring Program: Evidence from the Cerrado Biome of the Brazilian Legal Amazon. *Appl. Geogr.* **2022**, *140*, 102651. [CrossRef]
40. Emmerson, K.M.; Thatcher, M.; Osbrough, S.; Clarke, J.M. Quantifying Natural Emissions and Their Impacts on Air Quality in a 2050s Australia. *Atmos. Environ.* **2025**, *349*, 121144. [CrossRef]
41. Solcast API Documentation. Available online: <https://solcast.com/> (accessed on 1 December 2025).
42. Al-Timimi, Y.J.; Al-Khudhairy, D.H. Analysis of Air Temperature Trends in Iraq. *J. Environ. Earth Sci.* **2015**, *5*, 14–25.
43. Al-Salihi, A.M.; Mohammed, H.A. Analysis of the Relationship Between Meteorological Parameters and Aerosol Optical Depth over Iraq. *Atmos. Res.* **2020**, *239*, 104923.
44. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine Learning Methods for Solar Radiation Forecasting: A Review. *Renew. Energy* **2017**, *105*, 569–582. [CrossRef]
45. Mohamed, M.; Mahmood, F.E.; Abd, M.A.; Chandra, A.; Singh, B. Dynamic Forecasting of Solar Energy Microgrid Systems Using Feature Engineering. *IEEE Trans. Ind. Appl.* **2022**, *58*, 7857–7869. [CrossRef]
46. Al-Musaylh, M.S.; Al-Daffaie, K.; Downs, N.; Ghimire, S.; Ali, M.; Yaseen, Z.M.; Igoe, D.P.; Deo, R.C.; Parisi, A.V.; Jebar, M.A. Multi-Step Solar Ultraviolet Index Prediction: Integrating Convolutional Neural Networks with Long Short-Term Memory for a Representative Case Study in Queensland, Australia. *Model. Earth Syst. Environ.* **2025**, *11*, 77.
47. Mugware, F.W.; Ravele, T.; Sigauke, C. Short-Term Predictions of Global Horizontal Irradiance Using Recurrent Neural Networks, Support Vector Regression, Gradient Boosting Random Forest and Advanced Stacking Ensemble Approaches. *Computation* **2025**, *13*, 72. [CrossRef]
48. Chou, J.S.; Krang, J.; Limantonio, D.N. Regional Solar Generation Prediction with Metaheuristically Optimized Artificial Intelligence for Sustainable Grid Management. *Renew. Energy* **2025**, *231*, 124641.
49. Mohanasundaram, V.; Rangaswamy, B. Photovoltaic Solar Energy Prediction Using the Seasonal-Trend Decomposition Layer and ASOA Optimized LSTM Neural Network Model. *Sci. Rep.* **2025**, *15*, 4032. [CrossRef] [PubMed]
50. Al-Hilfi, H.A.; Abu-Siada, A.; Shahnia, F. Estimating Generated Power of Photovoltaic Systems During Cloudy Days Using Gene Expression Programming. *IEEE J. Photovolt.* **2020**, *11*, 185–194. [CrossRef]
51. Terrén-Serrano, G.; Martínez-Ramón, M. Deep Learning for Intra-Hour Solar Forecasting with Fusion of Features Extracted from Infrared Sky Images. *Inf. Fusion* **2023**, *95*, 42–61. [CrossRef]
52. Marion, B. A Model for Deriving the Direct Normal and Diffuse Horizontal Irradiance from the Global Tilted Irradiance. *Sol. Energy* **2015**, *122*, 1037–1046. [CrossRef]
53. Kakou, P.C.K.; Laouali, D.; Aka, B.; Osei, J.A.; Ette, N.F.K.; Frey, G. Multi-Timescale Validation of Satellite-Derived Global Horizontal Irradiance in Côte d'Ivoire. *Remote Sens.* **2025**, *17*, 998.
54. Psiloglou, B.E.; Kambezidis, H.D.; Kaskaoutis, D.G.; Karagiannis, D.; Polo, J.M. Comparison between MRM simulations, CAMS and PVGIS databases with measured solar radiation components at the Methoni station, Greece. *Renew. Energy* **2020**, *146*, 1372–1391. [CrossRef]
55. Madhiarasan, M. Bayesian Optimisation Algorithm Based Optimised Deep Bidirectional Long Short Term Memory for Global Horizontal Irradiance Prediction in Long-Term Horizon. *Front. Energy Res.* **2025**, *13*, 1499751. [CrossRef]
56. Mystakidis, A.; Koukaras, P.; Tsalikidis, N.; Ioannidis, D.; Tjortjis, C. Energy Forecasting: A Comprehensive Review of Techniques and Technologies. *Energies* **2024**, *17*, 1662. [CrossRef]
57. Al-Hilfi, H.A.; Abu-Siada, A.; Shahnia, F. Combined ANFIS-Wavelet Technique to Improve the Estimation Accuracy of the Power Output of Neighboring PV Systems During Cloud Events. *Energies* **2020**, *13*, 1613.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.