MACHINE LEARNING AND SOLAR RADIATION FORECASTING

Machine Learning and Statistical Time-Series Forecasting of Solar Radiation Across Nigeria: A Multi-Location, Long-Term Satellite-Based Analysis

Victor O. Kash-Umoru

Published November 2026

Abstract

Throughout the year, solar radiation is abundant in Nigeria, yet electricity shortages remain a national challenge. Although solar energy has attracted increasing attention as a potential solution, there remains limited understanding of how well solar resources can be forecast across the country at a national scale. In this study, 39 years of monthly solar radiation data (1984–2022) from NASA satellites for 72 locations across Nigeria's Coastal, Forest, Guinea Savanna, and Sahel climate zones were used to investigate this problem. To evaluate which forecasting approach works best, four models were tested: Holt-Winters, SARIMAX, CatBoost, and an ensemble average of the three. The accuracy of the forecast was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), while two cross-validation strategies were also examined: a Fixed Start approach with an expanding training period and a Rolling Window approach using a fixed five-year training period. The Fixed Start validation results showed that SARIMAX delivered the best overall performance, achieving the lowest MAE of 0.190 kWh/m^2/day. The ensemble model was almost identical to SARIMAX, with an MAE of approximately 0.190 kWh/m^2/day, closely followed by Holt–Winters with an MAE of 0.196 kWh/m^2/day. CatBoost produced a higher overall MAE of 0.234 kWh/m^2/day, despite remaining competitive in the Sahel, primarily due to its increased sensitivity to cloud-driven variability across southern zones. The Rolling Window approach provided forecasting accuracy similar to the Fixed Start approach, but it reduced the computational cost by about 50%. Spatial error mapping identified the Forest zone as having the lowest forecast uncertainty (MAE = 0.171 kWh/m^2/day) and the Sahel as having the highest (MAE = 0.196 kWh/m^2/day), attributable to increased solar variability associated with Harmattan dust. The ensemble model consistently ranked among the best-performing methods, highlighting the benefit of model averaging.

1 Introduction

In sub-Saharan Africa, Nigeria exemplifies one of the most significant electricity supply deficits. As Africa's most populous country and one of its largest economies, the nation continues to experience a huge gap between electricity supply and demand. Although the Nigerian Electricity Supply Industry (NESI) has an installed generation capacity of approximately 13,600 MW, the available generation capacity has ranged between 5,300 MW and 5,400 MW, while actual grid generation has remained below 5,000 MW. Specifically, the Nigerian Electricity Regulatory Commission (NERC) reported, in the second quarter of 2025, an average available generation capacity of 4,396 MW and an average grid generation of 4,501 MW. Meanwhile, it has been estimated that national electricity demand exceeds 30,000 MW, highlighting a persistent supply deficit that limits industrial productivity, economic competitiveness, and socioeconomic development. This gap in national electricity supply has contributed immensely to self-generation using diesel and petrol generators across the residential, commercial, and industrial sectors.

Natural gas is Nigeria's major source of electricity generation, accounting for about 85% of the power generated as of 2022 (Climate Scorecard, 2024), while renewable energy sources like hydropower, solar, and wind contribute less than 15% of installed capacity. To combat national energy insecurity, broaden electricity generation, and reduce greenhouse gas emissions, the Nigerian government launched the Renewable Energy Master Plan (REMP), aimed at increasing renewable energy electricity generation from 13% of installed capacity in 2015 to 36% by 2030 (ITA, 2024). There is also a transition plan in place aimed at complete national electrification by 2030 and ensuring that there are zero greenhouse gas emissions by 2060. This would be achieved by leveraging solar photovoltaic (PV) systems as the technological basis for both grid-connected and off-grid rural electrification (ETP, 2022). Solar energy is particularly well-positioned to meet this mandate: Incident upon Nigeria is a mean daily solar radiation of approximately 5.25 kWh/m^2/day across the country, ranging from 3.5 kWh/m^2/day in the coastal south to 7.0 kWh/m^2/day in the semi-arid north (Ohunakin et al., 2014). If just 1% of suitable land were utilized for solar projects, Nigeria could theoretically generate up to 210 GW of solar energy (IRENA, 2024), underscoring the scale of the untapped resource.

To correctly size, site, and install solar generation equipment (whether distributed, off-grid, or large-scale photovoltaic farms), a clear understanding of the characteristics of regional solar radiation regimes is essential. Solar radiation profiles, their seasonal variations, and their year-to-year consistency form the basis for investment decisions, system sizing, and grid integration planning. Measuring solar radiation *in situ* requires the use of instruments such as pyranometers and other ground-based monitoring equipment, whose installation and maintenance are expensive. Consequently, a reliable supply of observed solar radiation data is not available in most parts of Nigeria and sub-Saharan Africa (Haque et al., 2022; Ogolo, 2010). There is a network of NiMET synoptic weather stations across Nigeria, distributed throughout all 36 states and the Federal Capital Territory. However, observations of ground-based solar radiation are often incomplete, temporally inconsistent, or simply not recorded at many stations (Ohijeagbon, 2025).

Due to the lack of ground observations, there is an increasing use of satellite-derived meteorological datasets. The National Aeronautics Space Administration (NASA) Prediction of Worldwide Energy Resources (POWER) database provides continuous spatial and temporal coverage of monthly meteorological data, including surface shortwave downward irradiance, temperature, relative humidity, surface pressure, and wind speed, derived from satellite observations spanning more than four decades (NASA POWER, 2022). Although satellite-derived estimates are generally less accurate than ground measurements, particularly under complex atmospheric conditions, they provide continuous spatial and temporal coverage that makes large-scale applications possible (Zhou et al., 2021). In parallel, recent studies have shown that machine learning models can accurately predict solar radiation by capturing nonlinear relationships between meteorological variables and solar radiation from historical data (Zhou et al., 2021; Huang et al., 2021). For Nigeria, this has an important practical implication. These models, once trained with satellite observations, can be applied without the need for local solar radiation measurements, which allows them to estimate solar radiation even in locations where ground observations are not available.

Recently, the adoption of artificial intelligence methods in solar radiation modelling has increased in Nigeria, mainly due to their ability to model the nonlinear relationships between meteorological variables and solar irradiance more accurately than traditional empirical models. Several studies initially applied artificial neural networks to validate their effectiveness across different climatic zones before expanding to areas such as deep learning, ensemble modelling, and hybrid modelling techniques.

One of the earliest applications of artificial neural networks to solar radiation modeling in Nigeria was reported by Fadare (2009). In the study, meteorological inputs of sunshine duration, air temperature, and relative humidity were used to demonstrate that ANN models could be used to estimate solar energy potential across several locations in Nigeria with satisfactory accuracy. Although the network architecture was relatively simple compared with the deep learning and ensemble methods used in recent studies, the work represents a crucial early application of machine learning to solar energy modeling in Nigeria. It also provided a reference point for subsequent studies that investigated more sophisticated learning algorithms and expanded the range of predictor variables used to estimate solar radiation.

Artificial Neural Network (ANN) model was adopted by Aliyu et al. (2021) to estimate daily solar radiation in Kano, Kaduna, and Katsina using 21 years of meteorological data obtained from NiMET spanning 1994 and 2015. R^2 values of 0.78 and RMSE values of 0.47 and 0.48 for the training and test datasets were reported, indicating that ANN models could estimate daily solar radiation with reasonable accuracy under semi-arid conditions in northwestern Nigeria, even with limited availability of ground observations.

Isma'il and Aliyu (2023) used the same stations and observation period to determine whether replacing the ANN architecture with a Long Short-Term Memory (LSTM) network could improve prediction performance. R^2 values of 0.79 during training and 0.78 during testing were obtained, with corresponding RMSE values of 0.46 and 0.47. Even though the differences between the two models were relatively small, the lower prediction errors obtained with the LSTM indicate a slight improvement over the earlier ANN implementation, showing the potential advantage of deep learning for modeling temporal variability in solar radiation.

Later studies expanded the geographical scope to include more stations to determine the applicability of machine learning across Nigeria. Olatomiwa et al. (2022) compared ANN, Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) using data collected from four locations over a 12-year period. Their findings showed that deep learning models consistently outperformed the conventional ANN approach, achieving an R^2 of 0.9546 with RMSE and MAE values of 82.22 W/m^2 and 36.52 W/m^2, respectively. This study demonstrated that deep learning architectures can achieve high predictive accuracy across multiple locations with varying climatic conditions.

Furthermore, Olatomiwa et al. (2022) extended the model comparison beyond Nigeria to examine whether algorithm performance varied across African climates. The study reported that model performance varied widely across countries. While a hybrid CNN-LSTM-ANN model produced the best results for Algerian locations, XGBoost achieved the highest accuracy for the Nigerian sites. For estimating daily average Direct Normal Irradiance (DNI) at Akure, XGBoost achieved a correlation coefficient of 0.95, an RMSE of 24.69 W/m^2, and an MAE of 18.53 W/m^2. These findings showed that the most suitable modeling approach may depend on local climatic characteristics rather than the complexity of the learning algorithm alone.

Geospatial information has been incorporated into recent studies to improve predictive performance. Abolude et al. (2026) evaluated ANN and XGBoost models using 29 years of meteorological observations together with Digital Elevation Model (DEM)-derived radiation variables. XGBoost outperformed ANN, producing an MAE of 0.028, RMSE of 0.032, and MAPE of 0.54%. Feature importance analysis showed that maximum temperature and relative humidity were the most influential predictors. Unlike earlier studies that relied only on station observations, this work incorporated terrain-

derived information into the modeling framework, enabling solar radiation estimates to account for spatial variability in areas with limited meteorological measurements.

A major shift in the literature occurred when researchers began combining terrain-derived data with traditional weather station observations. This approach proved to be a game-changer, especially for mapping solar radiation in regions where physical weather stations are few and far between. Machine learning applications have also diversified across forecasting objectives and spatial scales. Tijani et al. (2025) proposed a two-stage framework for Ibadan that combined weather variables with cloud-type information to estimate clear-sky and actual irradiance components, including Global Horizontal Irradiance (GHI), Direct Normal Irradiance (DNI), and Diffuse Horizontal Irradiance (DHI). Among the evaluated algorithms, Random Forest achieved the lowest annual normalized RMSEs of 0.22 for DHI, 0.33 for DNI, and 0.19 for GHI. Prediction errors were lower during the dry season than during the wet season, reflecting greater atmospheric variability associated with increased cloud cover.

Similarly, Lamidi et al. (2025) examined day-ahead hourly solar radiation forecasting across eight Nigerian locations using a stacked hybrid framework that combined statistical and machine learning approaches. SARIMAX captured seasonal behavior but generally performed less accurately than the machine learning models. LSTM performed better when temporal patterns were well defined, whereas XGBoost yielded superior results under noisier conditions. The stacked hybrid model provided the most consistent overall performance, achieving R^2 values ranging from 0.68 to 0.92. Seasonal differences were also evident, with higher prediction accuracy during the December–February dry season than during the March–May transition period.

The study by Furiati et al. (2026) provides the closest methodological reference for the present research. They evaluated the performance of five forecasting models Holt-Winters, LSTM, SARIMAX, SVM, and XGBoost for monthly solar radiation prediction across 90 locations in southeastern Brazil using more than 30 years of NASA POWER satellite data. Model performance was assessed using both Fixed Start and Rolling Window cross-validation procedures. Their results showed that Holt-Winters achieved the best overall performance with the lowest MAE (0.302 $kWh/m^2/day$), followed by LSTM (0.314) and XGBoost (0.336), while SVM recorded the highest prediction errors (0.390). Although the two validation approaches produced statistically comparable results, the authors recommended the Rolling Window method because of its greater computational efficiency. They also found that forecasting accuracy declined in areas with higher solar radiation variability. While a similar relationship may exist in Nigeria, the influence of the Harmattan season, particularly across the Sahel, is expected to produce a different pattern of variability. Building on this methodological foundation, the present study adopts the same general forecasting framework but evaluates a different set of models comprising Holt-Winters, SARIMAX, CatBoost, and an ensemble model that combines the strengths of these individual approaches. In doing so, the study extends the framework from a regional case study in Brazil to a nationwide assessment of solar radiation forecasting across Nigeria.

In spite of the growing body of work, there are still gaps remaining in the Nigerian solar radiation forecasting literature. First, by focusing on single locations or small numbers of aggregated climatic zones, most studies do not consider the station-level spatial granularity needed to draw conclusions that represent the country as a whole. Compounding this spatial limitation is the consistently short temporal window of most available datasets, which rarely extends beyond a few years. Apparently, this is simply insufficient for capturing the inter-annual variability and longer-term climatological signals that shape solar radiation patterns across Nigeria's diverse climatic regimes.

The primary objective of this study is to address these gaps by providing a comprehensive, multi-location, long-term analysis of solar radiation time-series forecasting across Nigeria, using freely available NASA satellite data and open-source machine learning tools. We use 39 years (1984–2022) of monthly meteorological data from the NASA POWER (Prediction Of Worldwide Energy Resources) satellite-based archive, which is freely available, gap-free, and globally consistent. We select 72 locations distributed across the country, covering all four major climate zones: Coastal, Forest, Guinea Savanna, and Sahel. Four forecasting models are implemented: Holt-Winters (univariate exponential smoothing), SARIMAX (seasonal ARIMA with exogenous variables), CatBoost (a modern gradient-boosting tree algorithm), and a simple ensemble that averages the predictions of the three individual models. Two cross-validation protocols suitable for time series are employed, Fixed Start (expanding training window) and Rolling Window (fixed 5-year training window) to evaluate model generalization and computational efficiency. Performance is measured using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Additionally, the Friedman test followed by Nemenyi's post-hoc test is used to assess statistical significance of differences between models. Inverse distance weighting (IDW) interpolation is applied to produce spatially resolved forecast error maps for Nigeria, stratified by climate zone.

The rest of the paper is organized as follows: Section 2 presents the theoretical background on time-series analysis and the main drivers of solar radiation dynamics in Nigeria. Section 3 presents the methodology, including the study sites, the NASA POWER dataset, the machine learning models, the cross-validation procedures, the spatial interpolation algorithm, and the evaluation metrics. Section 4 presents and discusses the results, including model error comparisons, forecast visualizations for baseline stations, cross-validation analysis, and spatial distribution of model performance. Section 5 concludes with a summary of the findings, practical implications for Nigeria's solar energy sector, and future research directions.

2 Theoretical background

In a time series, observations of a variable are sampled over a regular sequence of time points. A primary aim of time series analysis is to establish structure within the time series so as to learn more about the behavior of that structure and to provide accurate future forecasts (Box et al., 2015). Time series methods are ideally suited for solar radiation due to well-established, physically determined rhythms driven by the Earth's orbit and tilt, in conjunction with locally observed climatic influences. This structure is readily deconstructed and modelled, leading to skillful predictions for the future.

A time series has traditionally been considered to have three major components (Montgomery et al. 2015):

- **Trend:** The overall long-run direction of the series, that it is increasing, decreasing, or oscillating horizontally. There is a relatively flat trend over 1984–2022 in Nigerian solar radiation, at least prior studies suggest it does not exhibit the sort of long-run increase/decrease of other time series such as inflation (Ohunakin et al. 2013). There are inter-annual variations, with an ENSO signal, so the series can exhibit some long-run irregularities.
- **Seasonality:** A cyclical or repeating pattern with a fixed period, that repeats itself over time. For a time series on a month by month basis, a 12-month pattern is observed in the solar radiation time series because it follows the yearly fluctuation of the earth's orbital position with respect to the sun's rays and the amplitude is by far the dominant component of Nigerian solar radiation time series.

- **Residual:** A time series' component that reflects whatever random or unpredictable variations in the series that are left after removing the effects of trend and seasonality. These are generally attributed to unpredictable weather, random chance events and erratic climatic processes (for instance sudden storms and atmospheric dust that affect radiation). The magnitude of the residual influences how difficult it is to forecast; areas with high and erratic residuals are more difficult to predict.

To illustrate these components in the Nigerian context, Figure 1 presents the seasonal decomposition of solar radiation for Abuja (the capital city) over the 39-year study period.

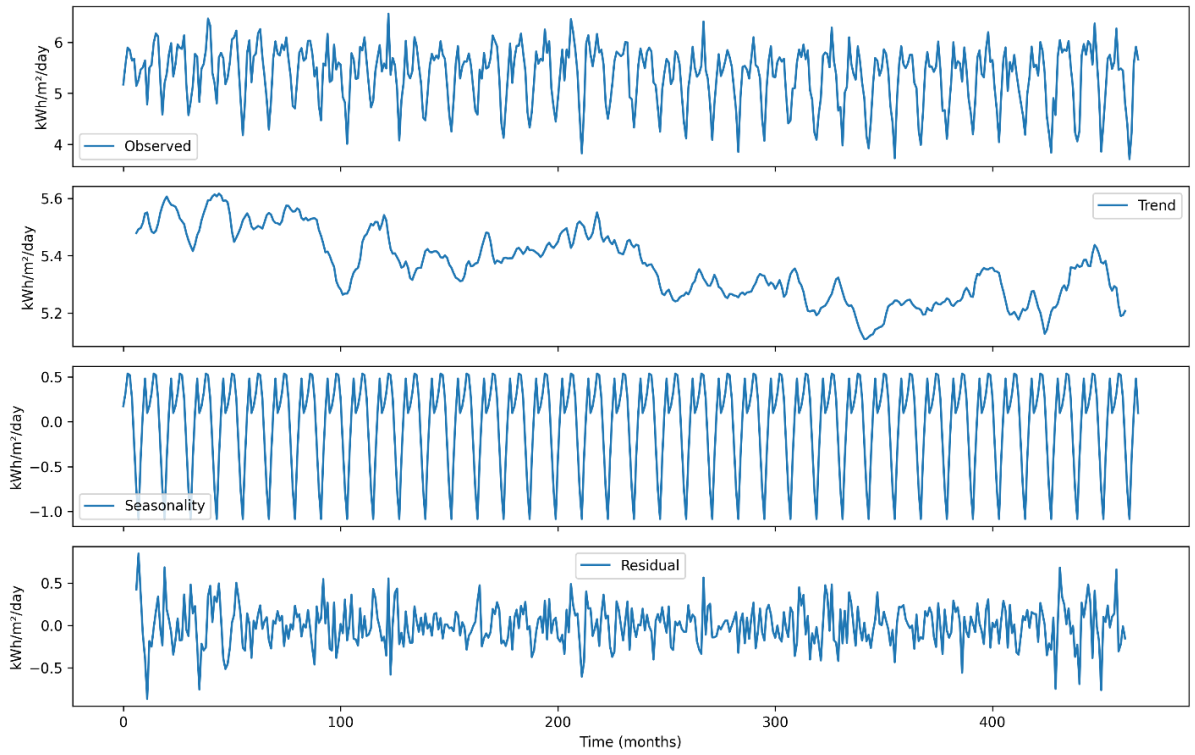


Figure 1. Seasonal decomposition of solar radiation for Abuja (1984–2022). The four panels show (a) the observed solar radiation time series, (b) the trend component, (c) the seasonal component, and (d) the residual component. The decomposition was performed using an additive seasonal model with monthly periodicity (period = 12).

The decomposition confirms that solar radiation in Abuja is dominated by a strong, regular 12-month seasonal cycle, driven by the annual migration of the Inter-Tropical Convergence Zone (ITCZ). The trend component is relatively flat over the 39-year period, indicating no significant long-term increase or decrease in solar radiation. The residuals are small and appear random, confirming that the series is well-behaved and forecastable. This pattern is representative of the Guinea Savanna climate zone and, with appropriate adjustments for amplitude, of the other three climate zones analysed in this study.

These three components may be combined in one of two ways. In an additive model, the observed series is the sum of its trend, seasonal, and residual components. In a multiplicative model, the components are multiplied together, implying that the amplitude of the seasonal fluctuation grows in proportion to the trend level. For solar radiation on a horizontal surface in tropical and sub-tropical contexts, the additive model is generally appropriate (Ogolo, 2010), because the seasonal amplitude

does not systematically change with the mean level across years. All seasonal decompositions in this study therefore employ the additive specification.

Time-series forecasting is the task of using the historical pattern of a series its estimated trend, seasonal structure, and residual dynamics to generate predictions of future values (Montgomery et al., 2015). Statistical models such as Holt-Winters and SARIMAX encode the seasonal cycle as explicit model parameters. Machine learning models such as CatBoost learn temporal patterns implicitly from training data, potentially capturing complex non-linear relationships between meteorological predictors and solar radiation that parametric models cannot represent (Zhou et al., 2021). The complementary strengths of these two approaches motivate the comparative framework of this study.

2.1 Physical drivers of solar radiation

The incoming solar radiation at any given location on the Earth's surface is controlled by a set of physical processes that vary with time and space scales. At the largest scales, the radiation received by a horizontal plane is related to the solar zenith angle (which in turn is a function of latitude, day of the year and time of day) (Duffie and Beckman 1980). At Nigeria's latitudes (4°N–14°N), the sun reaches a high angle throughout the year, producing relatively high mean solar irradiance compared to mid-latitude regions. However, the zenith angle still varies systematically through the year, producing the 12-month seasonal cycle.

After being subjected to terrestrial processes, extraterrestrial radiation is further altered by the atmosphere before it reaches the Earth's surface. The major influence on radiation in tropical areas is cloudiness which, through absorption and reflection, diminishes the incoming short wave radiation and shows the greatest day-to-day and month-to-month variability. The other influences are aerosol, which, through scattering and absorption, reduce and change the wavelength of incoming radiation; in Nigeria and in West Africa the most important are desert aerosols originating in the Sahara and which load the atmosphere particularly during the dry season (Tanoli et al., 2024; Park et al., 2021). Water vapour also acts in discrete near-infrared wavelengths whilst the fraction of radiation reaching horizontal surfaces will depend on the reflectivity of the underlying surface (albedo). The effect of all of these factors collectively gives the clearness index, which is a dimensionless ratio, between surface and extraterrestrial radiation (Duffie and Beckman, 1980).

2.2 Solar radiation dynamics in Nigeria

Nigeria stretches across a land area of roughly 923,000 km^2, occupying the latitudinal belt of 4°N–14°N and the longitudinal belt of 3°E–15°E. This range of latitudes includes a great diversity of climatic types, which vary from a very moist tropical climate along the Niger Delta coast (Equatorial) to the rather dry Sahelian climate in the north. However, the ultimate governing force behind this variability and the subsequent regional variation in solar radiation inputs is the migration of the Intertropical Convergence Zone (ITCZ), which, in its terrestrial manifestation as the Intertropical Discontinuity (ITD), serves to separate dry continental north-easterly winds from the moist south-westerly monsoon flow, between its northern and southern positions with respect to latitude as it shifts northward and southward from year to year (Ayoade, 1983; Ojo, 1977).

Northern Nigeria receives its highest solar radiation in March and September, while southern Nigeria receives its lowest during these months. This northward movement and retreat of the ITCZ occur with March-September and October-March, respectively, causing variations in rainfall and cloudiness and thereby impacting the distribution of solar radiation within Nigeria's four climate zones.

The coastal region of approximately below 6°N receives the lowest values of solar radiation, ranging from 4.0 to 4.8 kWh/m^2/day, and exhibits the smallest annual variation due to its humid tropical climate. This climate causes an annual rainfall of about 2,000 mm to 4,000 mm, marked by perpetually cloudy conditions and high annual relative humidity where cloud cover suppresses insolation even during the dry season (Fadare, 2009). Average solar radiation in this zone ranges from 4.0 to 5.0 kWh/m^2/day, and significant variability exists between wet and dry seasons, making its forecasting highly complex (Sanusi and Abisoye, 2011). In general, a large portion of this coastal belt experiences a bimodal rainfall pattern (two rainy peaks per year), except for its northernmost fringes which gradually transition into a single rainy season, indicating the influence of a growing dry season. Heavy cloud cover during the wet season severely limits solar radiation, which then increases as skies clear during the brief dry periods.

The Guinea Savanna zone, which is between 7.5°N and 11.5°N, features tropical wet-and-dry climatic conditions (Köppen Aw) with one well-defined wet season (April–September) and a distinct dry season. It receives between 750–1500 mm of rainfall annually. During the dry season, it is characterised by clear skies and strong sunshine, which decrease substantially as cloudiness rises and the wet season approaches. Mean solar irradiation over the Guinea Savanna zone ranges from 5.0 to 6.0 kWh/m^2/day (Ogolo, 2011), and some northern sites begin to experience Harmattan effects during the dry season.

The Sahel region, extending northward from about 11.5°N latitude, is a semi-arid climate zone (Köppen BSh). Average annual rainfall is under 600 mm and occurs only during a brief rain season (3–5 months). The mean daily solar radiation ranges between 6.0 - 6.5 kWh/m^2/day, the highest compared to any of the other zones, which can be attributed to a protracted dry season, less incidence of cloudiness and the high solar declination angle (Fadare, 2009). The most impacted area for the Harmattan phenomenon remains the Sahel, which produces inter- and intra-seasonal variations of solar irradiance.

2.3 The Harmattan and its effect on solar radiation

The Harmattan is a dry, dusty northeasterly wind that blows in West Africa from November through March, originating from the Sahara Desert as a consequence of the southward progression of the ITCZ and the formation of the Saharan high-pressure system (Ogunjo et al., 2022). Given the immense amounts of tiny mineral dust aerosols transported by the Harmattan, West Africa is the region most heavily laden with desert aerosols globally (Prospero et al., 2002; Klose et al., 2021).

The effect of the Harmattan dust on surface solar radiation can be significant. **Okogbue et al. (2009)** estimated a reduction in the direct spectral solar irradiance at Akure, Nigeria, by as much as 50% under heavy Harmattan dust outbreaks relative to clear-sky background conditions. A clear-sky index of 0.36 equivalent to overcast conditions despite a complete absence of clouds was observed for Enugu during the Harmattan season, as documented by **Ugwu and Ugwuanyi (2011)**. Recent modeling in West Africa confirms that incorporating desert dust aerosols into numerical simulations decreases the error of solar irradiance estimates by a factor of two **(Kouchi et al., 2025)**.

Significantly, the impact of the Harmattan exhibits an unequal spatial distribution across Nigeria. Dust loading is strongest and persists the longest in the Sahel zone, where it decreases surface irradiance by roughly 20–25% during the months that should otherwise be the clearest. Conversely, this dust layer rarely extends below 7–8°N into the forest and coastal regions, leaving southernmost locations largely uninfluenced, while the Guinea Savanna zone experiences intermediate impacts.

Because of this asymmetrical spatial structure, the straightforward dry-season relationship with high solar radiation found in the south is severely attenuated in the north by the compensating effect of dust scattering, presenting a distinct forecasting challenge for Sahelian meteorological stations.

2.4 Implications for time-series forecasting

Some direct effects of the physical assessment discussed above on the forecasting process are:

- All zones show strong seasonal signals caused by movement of the ITCZ with annual cycle giving strong favorable signal for seasonal time-series models. Models explicitly accounting for seasonal patterns (Holt-Winters, SARIMAX) should dominate the main variance component.
- Residual Magnitude varies across zones with episodic events of cloud occurrence, variability of ENSO and the intrusion of Harmattan dust events are drivers that determine such variability. Zones showing larger irregular residuals are supposed to give higher forecasts error irrespective of model. e.g. Sahel. The validity of this statement is assessed against how close spatial distributions of errors compare to the solar radiation variability stats computed for each zone.
- In the case of exogenous variables (surface pressure, temperature, relative humidity and wind speed) there is physical information about the state of the atmosphere encoded within these values. In the southern part of the region, relative humidity is especially valuable because cloud and moisture amount play a significant role. Both the intensity of Harmattan and the strength of the monsoon can have influence on surface pressure gradients. However, the correlation structure of these variables and solar radiation differs from zone to zone – information that machine learning models (CatBoost) can detect through the available data.
- Averaging predictions across three distinctly different types of models (state-space, tree-based and statistical) - the ensemble modeling - should lead to lower variability and increased robustness where individual models display alternative failure modes in space.

In summary, this theoretical background sets the stage for the forecast model of section 3, both physically and methodologically. Indeed, the diverse climates in Nigeria constitute an inherently generated, natural, multi-regime experiment with the objective to assess the generalizability of solar radiation forecasting models and thereby forming the basis for the climate-zone-stratified analysis described in this study.

3 Materials and methods

3.1 Meteorological data

Meteorological variables were derived from NASA POWER (Prediction of Worldwide Energy Resources) project, a satellite-based data archive supplying globally gap-filled surface meteorologic variables at 0.5 × 0.5 resolution (NASA POWER, 2023). These data were compiled from a combination of satellite products and reanalyses (e.g., Clouds and the Earth's Radiant Energy System; CERES) and have been verified by surface-based observations in multiple continental locations (White et al., 2017; Stackhouse et al., 2018).

Monthly data aggregated over the entire period (Jan. 1984-Dec. 2022, 39 years) was obtained for this study. The five following variables were extracted for each site:

Table 1. Meteorological variables obtained from NASA POWER.

Variable	Symbol	Unit	NASA POWER Parameter
Solar radiation (target)	G	kWh/m²/day	ALLSKY_SFC_SW_DWN
Surface pressure	P	kPa	PS
Temperature at 2 m	T	°C	T2M
Relative humidity at 2 m	RH	%	RH2M
Wind speed at 10 m	WS	m/s	WS10M

The solar radiation variable is All Sky Surface Shortwave Downward Irradiance, which describes the total incoming solar radiation to a horizontal plane at the Earth's surface, averaged over clear, cloudy, and overcast skies. This is the target variable for forecasting. The chosen four meteorological variables serve as the exogenous input features for the multivariate SARIMAX and CatBoost models due to their known physical association with solar radiation (Duffie and Beckman, 1980; Gupta et al., 2024) and their accessibility from the NASA POWER database.

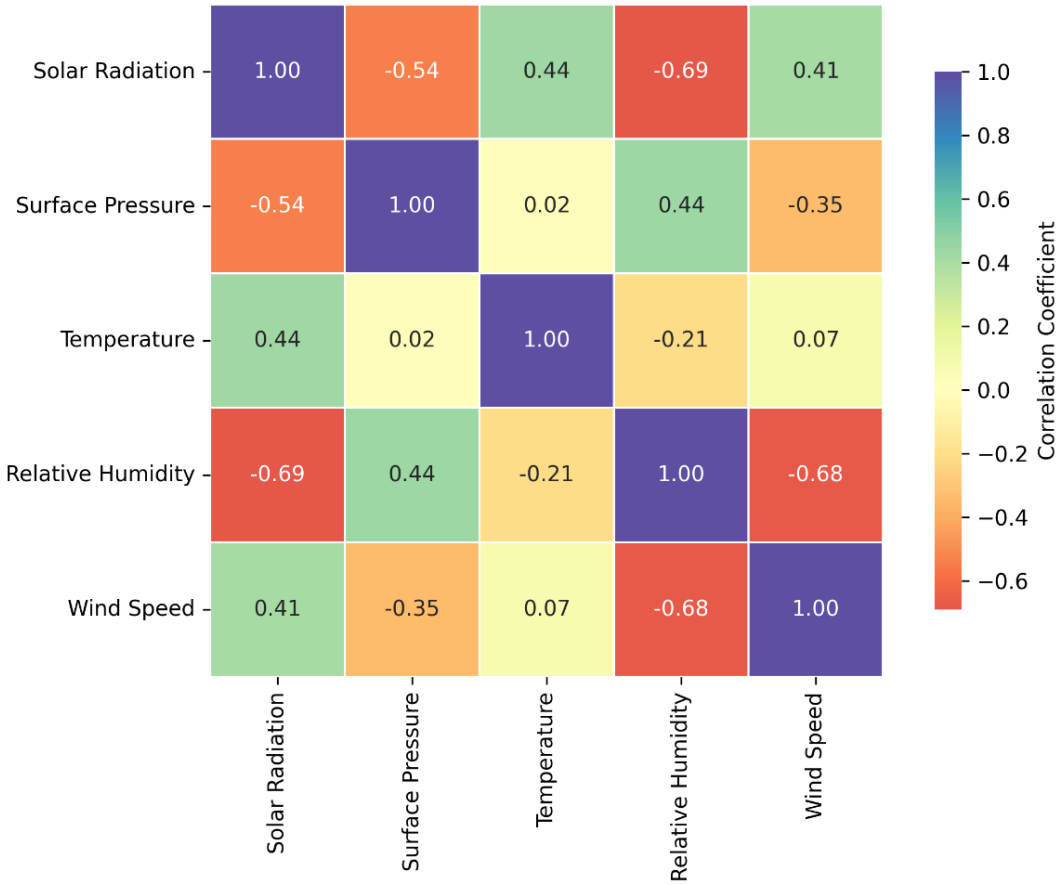


Figure 2. Correlation matrix of the five meteorological variables: solar radiation, surface pressure, temperature, relative humidity, and wind speed. Values represent Pearson correlation coefficients. The colour scale uses a diverging Spectral palette, with blue representing negative correlations and red representing positive correlations. Temperature exhibits the strongest positive correlation with solar radiation (r = +0.44), while relative humidity shows the strongest negative correlation (r = −0.69). Surface pressure also shows a moderate negative correlation (r = −0.54), and wind speed a moderate positive correlation (r = +0.41).

Surface pressure and temperature indirectly reflect atmospheric stability and air mass; relative humidity indicates cloud development and attenuation due to water vapor; and wind speed affects the transport of aerosols and the movement of clouds. Every extracted data point was screened for quality flagging. A value of -999.0 represents a missing or unreliable estimate within the NASA POWER archive. None of these quality flags were found in the downloaded series; therefore, the datasets are complete over the entire timespan for all localities. Additional imputation or interpolation was not performed on the retrieved data.

3.2 Study area and station selection

Nigeria is situated in West Africa, with its coordinates falling between latitudes 4°N and 14°N and longitudes 2.5°E and 14.5°E, covering a land area of roughly 923,000 km^2. The country displays a marked climatic division from north to south, featuring four recognized climate types:

- **Coastal zone (south of latitude 6°N):** A humid tropical climate characterized by extensive cloudiness and annual rainfall exceeding 2,000 mm.
- **Forest zone (6°N–8°N):** A tropical monsoon climate featuring moderate cloud cover and annual rainfall between 1,200 mm and 2,000 mm.
- **Guinea Savanna (8°N–11°N):** A tropical wet-and-dry climate with pronounced wet and dry seasons and annual rainfall ranging from 800 mm to 1,200 mm.
- **Sahel zone (north of 11°N):** A semi-arid climate with light annual rainfall under 800 mm and a high degree of seasonal Harmattan dustiness.

To ensure adequate coverage and spatial representativeness for the subsequent interpolation, sampling sites were selected using a three-tier design:

- **Tier 1:** Principal stations (37): One sampling site was selected for each of the 36 states and the Federal Capital Territory (Abuja), based on the geographic coordinates of their respective Nigerian Meteorological Agency (NiMet) synoptic stations. These stations have an established history of study and provide long-term, quality-controlled data in Nigerian climate literature (Ohunakin et al., 2013; Ogunjo et al., 2022). They span all six geopolitical zones and all four climatic zones.
- **Tier 2:** Supplementary Sahel points (3): Three manual points were added in the far-northern Sahel region (Hadejia, Nguru, and Gashua) to stabilize the interpolation in areas where active NiMet stations are sparse.
- **Tier 3:** Regular grid points (32): A regular 1.5° × 1.5° grid was mapped across Nigeria's geographic boundaries. Grid coordinates falling outside the national land borders as defined by the GADM administrative shapefile (GADM, 2022) were excluded through a spatial inclusion query utilizing GeoPandas (Jordahl et al., 2020).

Combined with the principal stations, the remaining 32 grid points facilitate a seamless IDW interpolation across the entire country. These 72 sites (37 main, 3 supplementary, and 32 grid locations) are distributed across the climate zones as follows: Coastal (9), Forest (11), Guinea Savanna (32), and Sahel (20). The exact coordinates for each location and their climate designations are provided in the supplementary data via the GitHub repository.

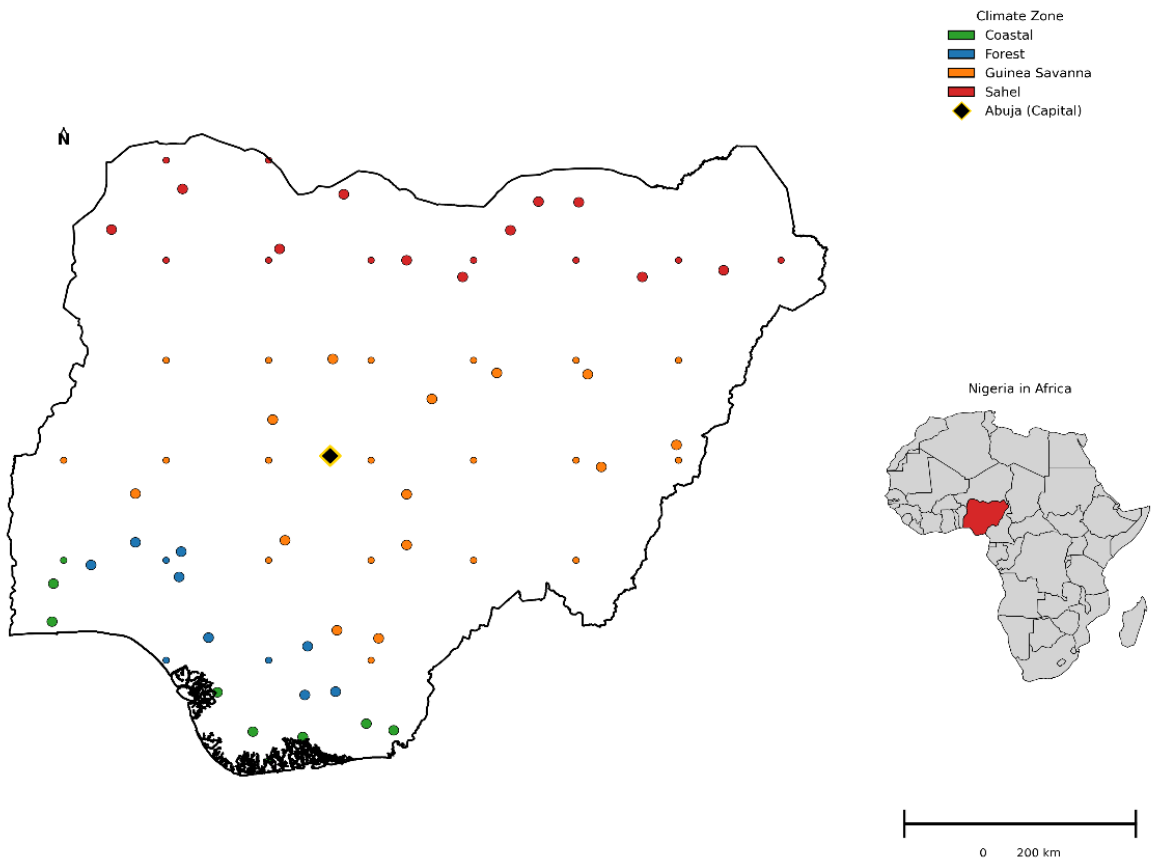


Figure 3. Geographical distribution of the 72 study locations across Nigeria. Points are colour-coded by climate zone: Coastal (green), Forest (blue), Guinea Savanna (orange), and Sahel (grey). The capital city, Abuja, is highlighted with a diamond marker. The map includes a small inset showing Nigeria's position within West Africa.

The spatial distribution reflects the country's population density and station coverage patterns, with a higher concentration of points in the south and a sparser network in the north. The supplementary grid points (Tier-3) and the additional Sahel points (Tier-2) ensure adequate coverage for spatial interpolation, particularly in the data-sparse northern regions.

3.3 Forecasting models

Four models were built, covering statistical, state space, tree-based and ensemble. The four models were trained to output predictions for the following 12 months (a full year) per cross-validation fold.

3.3.1 Holt-Winters

Holt-Winters exponential smoothing (Winters, 1960) accounts for seasonality and trend through recursive equations. As observed in the exploratory analysis, there was no long-term trend, and the annual cycle of the solar radiation series was clear. A model with trend=None and seasonal='add' was thus selected. The seasonal time limit was set to 12 months. This specification is relevant to solar radiation in tropical areas, where amplitude variations over time show no explicit correlation with the mean level (Ogolo, 2010; Furiati et al., 2026). The model was estimated using the statsmodels library (Seabold and Perktold, 2010).

3.3.2 SARIMAX

Seasonal AutoRegressive Integrated Moving Average with eXogenous variables (SARIMAX): extends the ARIMA model by incorporating the seasonal differencing term and external explanatory variables (Box et al. 2015). model orders were derived from preliminary diagnostics on the Abuja series and from models previously applied with success in tropical solar forecasting contexts (Furiati et al., 2026): non-seasonal order (p = 1, d = 1, q = 3) and seasonal order (P = 1, D = 1, Q = 1, m = 12). Surface pressure, temperature, relative humidity and wind speed are included in the exogenous variables. Maximum likelihood estimation with the statsmodels SARIMAX implementation was used to fit the model.

3.3.3 CatBoost

The CatBoost algorithm uses a gradient boosting decision tree that natively treats categorical inputs, with robust overfitting properties (Prokhorenkova et al. 2018). Unlike XGBoost and LightGBM, CatBoost employs ordered boosting and symmetric trees to decrease the shift in the predictions. Each of the splits were trained on the model using 500 iterations, depth of 6, learning rate 0.05 and the rest on their defaults. Exogenous variables were introduced into the models; however, no per-station categorical variables were required. Model were coded with catboost library (Dorogush et al., 2018).

3.3.4 Ensemble

A straightforward ensemble was created by taking the average of the 12-month ahead forecasts from Holt-Winters, SARIMAX and CatBoost (i.e. a non-weighted mean). No weighting or stacking was applied. The idea of using an ensemble is to decrease the forecast variance and also achieve better reliability and robustness because it leverages varied model assumptions (García-Gutiérrez et al., 2022; Pravilovic et al., 2017). This is an extremely computationally inexpensive solution and has proven superior in a multitude of forecasting settings compared to single models (Makridakis et al., 2018).

Each model was trained separately on each cross-validation split and for each station. All splits are based on non-tuneable hyperparameters for comparability and to prevent overfitting.

3.4 Cross-validation strategies

It's important to consider the temporal nature of time series when doing cross-validation, in order to avoid accidentally including future data in the training set (Bergmeir and Benitez, 2012). Two procedures were used for each run of training and testing to create a sequence of train-test splits. In each case, one test window spans 12 months or one calendar year. A comparison is made with real data for those months and error metrics are calculated for each split. The final score of a station for a specific model is simply an average across all the split.

3.4.1 Fixed Start (expanding Window)

Starting from year 1984, the training set grow one year after each loop. In any case, the test year is the year after the training. The number of splits is 34 (from 1989 to 2022). This method considers how performance gains with an increase in the use of historical data (Hyndman and Athanasopoulos, 2018).

3.4.2 Rolling Window (Fixed length)

The training window has a fixed length of five years. Following each iteration, the training window and test window each shift forward by a year. Test year is again the year following the training

window. This test estimates the behavior of a model in a practical case with an observation history constrained to very few observations in the past. Training size is the same as the Fixed Start method which results in less expensive computations (Hyndman and Athanasopoulos, 2018).

3.5 Evaluation metrics

In order to assess the accuracy of our forecasting methods three conventional regression metrics were applied (Chai & Draxler, 2014; Willmott & Matsuura, 2005). Let y_i denote the actual solar radiation for month i and $\hat{y}_i$ the predicted value. Over n test months (aggregated across all splits for a station):

- **Mean Absolute Error (MAE):** the average of the absolute differences between the actual and predicted values, expressed in the same units as the target variable.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

- **Root Mean Squared Error (RMSE):** puts a greater weight on larger errors, increasing the impact of outliers on the final value.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

- **Mean Absolute Percentage Error (MAPE):** the average of the absolute percentage differences between the actual and predicted values, making it a scale-independent measure of forecasting accuracy.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

R^2 was not reported because the objective of this study was forecasting rather than model fit. Instead, predictive performance was evaluated using out-of-sample forecast accuracy measures (MAE, RMSE, and MAPE), which are more appropriate for assessing forecasting performance than measures of explained variance such as R^2 (Hyndman and Athanasopoulos, 2021).

3.6 Statistical significance tests

- **Friedman test:** To examine whether there were statistically significant differences between the four models across the 72 stations, the non-parametric omnibus Friedman test for repeated measures was utilized (Friedman, 1937). The test ranks the models station-by-station using MAE; if the null hypothesis of equal performance holds true then the mean rank across stations should be the same for all models. A significant result ($p < 0.05$) indicates that at least one model differs from the others.
- **Nemenyi post-hoc test:** Post hoc, when Friedman test was conducted and the null hypothesis were rejected then Nemenyi test is used to compare between model pairs which differ significantly. Nemenyi, 1963. The critical distance (CD) for the Nemenyi test at $\alpha = 0.05$ was computed as:

$$CD = q_\alpha\sqrt{\frac{k(k+1)}{6N}}$$

Where $k = 4$ (number of models), $N = 72$ (number of stations), and $q_\alpha = 2.569$ (critical value for four models from the studentized range distribution). Two separate Friedman/Nemenyi analyses

were performed: one for Fixed Start MAE and one for Rolling windows MAE. This approach has been widely used in machine learning model comparisons (Demsar, 2006; Herbold, 2020).

- **Wilcoxon signed-rank test:** The Wilcoxon signed-rank test was applied to assess the differences in the error distributions of the two cross-validation methods (Fixed Start versus Rolling Window) across all station-model pairs, with a significance level of $\alpha = 0.05$ (Wilcoxon, 1945). This non-parametric test is appropriate for paired data that may not be normally distributed.

3.7 Spatial interpolation

In order to create an intuitive geographical visualization of prediction error distributions, per-station MAE was interpolated over a surface that covers the area of Nigeria with Inverse Distance Weighting (IDW) interpolation (Lu and Wong, 2008). Inverse Distance Weighting (IDW) estimates the value at an unmeasured location as a weighted average of nearby known points, with weights inversely proportional to a power of the distance, giving greater influence to closer observations.

A 200 × 200 point regular grid was defined over the bounding box of Nigeria (4°N–14°N, 2.5°E–14.5°E). The IDW value was calculated for each grid point from the 12 nearest sites with an inverse square distance power of 2. Nigeria's administrative boundary (GADM level 0 shapefile) was then used to mask the resulting raster, ensuring that colors remained confined to the land. Same procedure was followed individually on each model to create four different error maps.

The Inverse Distance Weighting (IDW) method was chosen for its simplicity, transparency, and suitability for interpolating meteorological variables exhibiting spatial autocorrelation (Lu and Wong, 2008). Although more sophisticated geostatistical methods, such as kriging, are available, IDW does not require modelling the spatial covariance structure and is computationally efficient, making it well suited for national-scale visualisation. The interpolation was implemented using scipy.spatial.cKDTree for efficient nearest-neighbour searches, while matplotlib was used for visualisation (Virtanen et al., 2020).

In addition, a similar IDW interpolation was performed for the observed solar radiation in September 2022, which was selected as a representative month for visual comparison, as well as for the predictions of each model for the same period, allowing comparison of their spatial patterns.

4 Results

The performance of the four models (Holt-Winters, SARIMAX, CatBoost, Ensemble) for the 72 study sites in Nigeria are reported in this section. We cover the overall model comparison (Section 3.1), cross-validation method comparison (Section 3.2), comparison of baseline stations (Section 3.3), statistical significance of differences (Section 3.4), spatial variability of error (Section 3.5), and computational cost comparison (Section 3.6).

4.1 Overall model performance

In Table 2, the average MAE, RMSE, and MAPE across the four different models under Fixed Start cross-validation have been reported. Of all SARIMAX, had the smallest MAE (0.190 $kWh/m^2/day$), and Ensemble (0.190), and Holt-Winters (0.196). We observe a high MAE in the CatBoost model (0.234) because the algorithm was more responsive to cloud changes in general, and this was particularly evident in the southern regions. From the standard deviations, we observe that CatBoost

possessed the greatest variability among the stations as well, with the Ensemble exhibiting the most uniformity over the sites.

Table 2. Overall performance metrics (Fixed Start CV) averaged across all 72 stations.

Model	MAE (kWh/m²/day)	RMSE (kWh/m²/day)	MAPE (%)
Holt-Winters	0.196 ± 0.016	0.242 ± 0.020	3.77 ± 0.54
SARIMAX	0.190 ± 0.017	0.235 ± 0.021	3.68 ± 0.55
CatBoost	0.234 ± 0.019	0.289 ± 0.024	4.58 ± 0.62
Ensemble	0.190 ± 0.015	0.236 ± 0.019	3.70 ± 0.48

Under Rolling Window cross-validation, the same ranking was observed, though all models exhibited slightly higher errors (Table 2). The ensemble model ranked highest again followed by SARIMAX and Holt-Winters models. SARIMAX model had the highest error increase using rolling window method (0.190 to 0.198) and this suggested SARIMAX perform best with longer time series. CatBoost model was the poor performer model in terms of its general predictive capacity; although performance has been slightly better in the Sahel zone.

Table 3. Overall performance metrics (Rolling Window CV) averaged across all 72 stations.

Model	MAE (kWh/m²/day)	RMSE (kWh/m²/day)	MAPE (%)
Holt-Winters	0.200 ± 0.015	0.248 ± 0.019	3.84 ± 0.59
SARIMAX	0.198 ± 0.017	0.247 ± 0.020	3.86 ± 0.61
CatBoost	0.238 ± 0.019	0.295 ± 0.024	4.66 ± 0.64
Ensemble	0.195 ± 0.015	0.243 ± 0.019	3.78 ± 0.50

The ranking shows that, on average, SARIMAX and the Ensemble are statistically indistinguishable from each other, while Holt-Winters shows only a small performance decrease. CatBoost performs fairly well in certain regions (see Sect. 3.5) and cannot be recommended for nationwide solar forecasting unless the whole framework is changed.

4.2 Cross-validation method comparison

A relevant practical issue is whether the more computationally expensive Fixed Start method offers a worthwhile improvement in accuracy compared to the Rolling Window method. The Wilcoxon signed-rank test was used to assess whether there were significant differences in the error distributions between the two methods across all station-model pairs. The test revealed a statistically significant difference in accuracy ($p = 0.012$), but a negligible effect size ($\gamma = -0.025$). In practical terms, the accuracy produced by the two methods was almost identical. However, Fixed Start was much more computationally expensive (see Section 3.6) and, therefore, the Rolling Window method is preferred for practical operational forecasting in resource-constrained environments. This outcome supports similar studies conducted in tropical regions (Furiati et al., 2026).

In practice, Rolling Window can be substituted for Fixed Start with virtually no loss in accuracy and significant savings in training time. Nevertheless, all results presented are based on Fixed Start cross-validation by default to facilitate comparability with earlier work.

4.3 Baseline station analysis

To facilitate the visual comparison of errors while maintaining a concise presentation, five baseline stations were used to represent Nigeria's capital city, Abuja, and the four corners of the country: Port Harcourt (south), Maiduguri (north), Sokoto (west), and Yola (east). These five stations encompass the spectrum of climatic and solar irradiation conditions across Nigeria's diverse environments, ranging from the humid conditions of the Niger Delta to those typical of the Sahel. The average MAPE for these five stations under both cross-validation methods, using grouped bar charts, is shown in Figure 6.

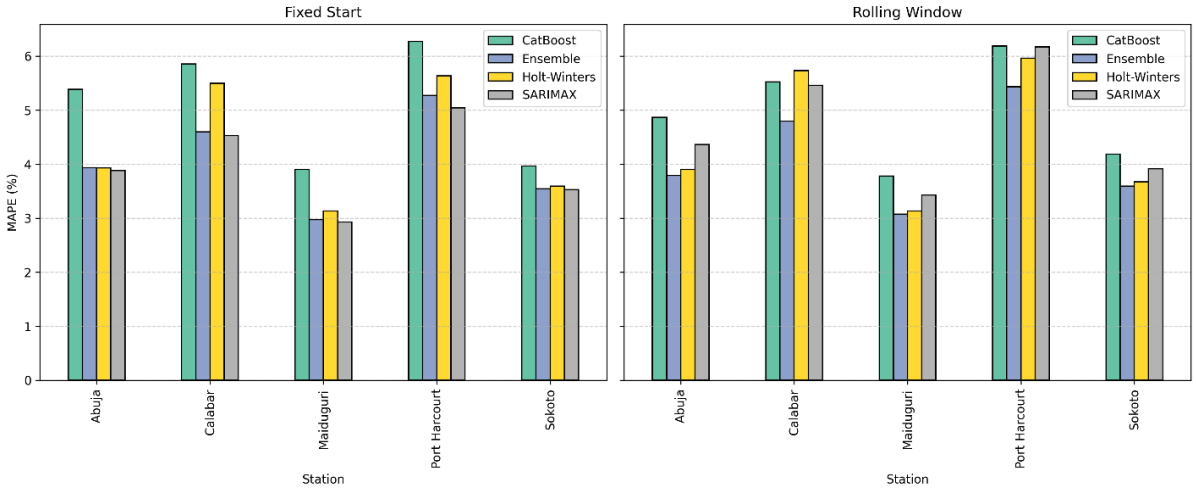


Figure 4. Average MAPE for baseline stations (Abuja, Port Harcourt, Maiduguri, Sokoto, Yola) under Fixed Start and Rolling Window cross-validation. Bars represent the mean MAPE for each model; error bars show standard deviation across splits.

At four of the five stations (with the exception of Port Harcourt, where the Ensemble model performed slightly better), Holt-Winters achieved the lowest MAPE. Furthermore, at the two Sahel stations (Maiduguri and Sokoto), all models exhibited higher MAPE values because of the greater solar variability associated with Harmattan dust and the shorter rainy season.

At the southern stations, CatBoost produced higher errors than the best-performing models. Indeed, it should be highlighted that the performance of CatBoost is site-dependent. This method performed relatively well in the Sahel, where CatBoost and, to some extent, Holt-Winters produced results closest to those of the Ensemble model. Conversely, errors were larger at the southern stations (Yola and Port Harcourt). In any case, the error patterns observed at these baseline stations are indicative of the patterns shown in the country-scale error maps presented in Section 3.5, where it is demonstrated that forecast error increases northward, with the greatest forecasting difficulty occurring in the Sahel.

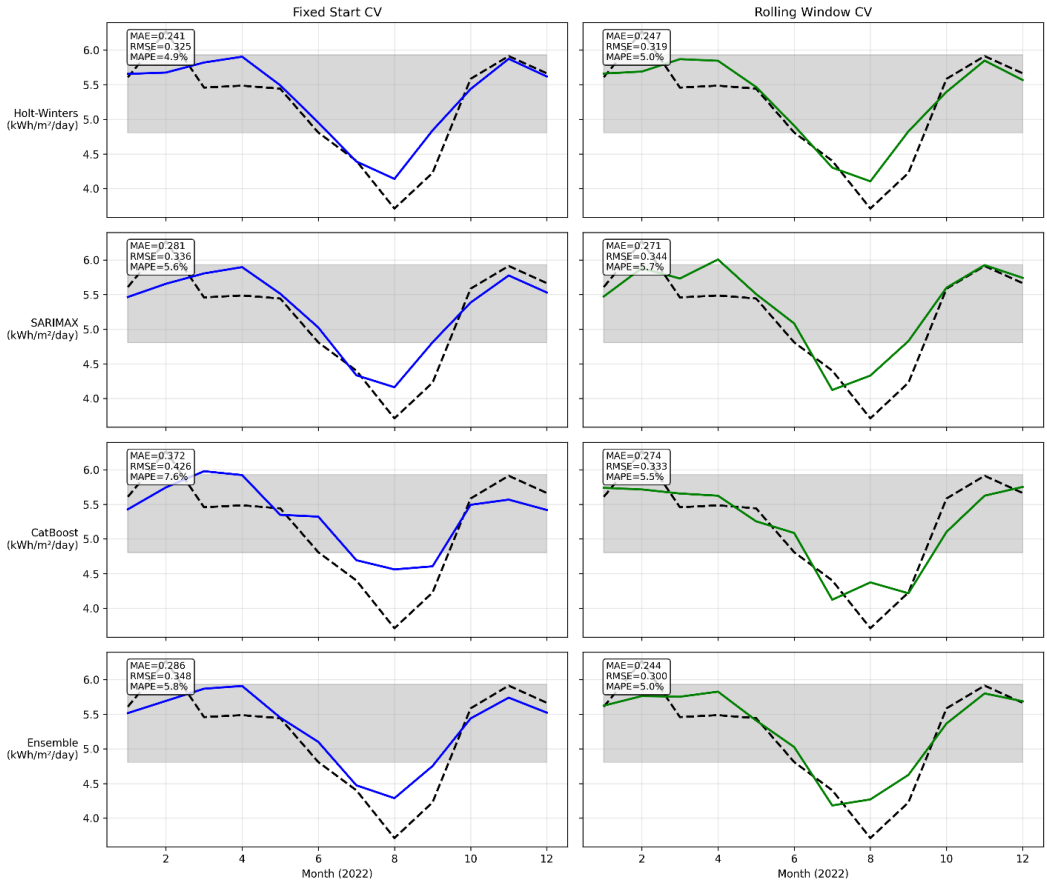


Figure 5. Solar radiation predictions for Abuja (2022) under Fixed Start and Rolling Window cross-validation. Each row corresponds to a model: (a) Holt-Winters, (b) SARIMAX, (c) CatBoost, (d) Ensemble. Left column: Fixed Start CV; right column: Rolling Window CV. The grey band represents ±1 standard deviation of the historical solar radiation (1984–2021). Dashed black line: actual values; coloured line: predictions. MAE and RMSE values are annotated in each subplot.

All models capture the seasonal cycle accurately, with the Ensemble showing the closest agreement with the observed values. Under Rolling Window CV, the predictions are slightly more volatile, consistent with the shorter training window. The grey band (historical ±1 std dev) provides a benchmark; predictions falling within this band are considered satisfactory. CatBoost exhibits the largest deviations during the dry season (November–February), likely reflecting its sensitivity to Harmattan dust variability, which is not explicitly modelled.

The spatial pattern of errors across the baseline stations prefigures the national-scale error maps presented in Section 4.5: errors increase northward, with the Sahel zone being the most challenging to forecast.

4.4 Statistical significance tests

The Friedman test was applied to determine whether significant differences exist among the four models across the 72 locations. Under Fixed Start CV, the test rejected the null hypothesis of equal performance ($\chi^2 = 154.0$, $p < 0.001$). The Nemenyi post-hoc test, with critical distance CD = 0.55 at $\alpha = 0.05$ (computed for $k = 4$ models and $N = 72$ locations), revealed that:

- **SARIMAX, Ensemble, and Holt-Winters form a single equivalence group** – their performance differences are not statistically significant.

- **CatBoost is significantly worse** than all three top models ($p < 0.05$).

Under Rolling Window CV, the Friedman test again rejected the null hypothesis ($\chi^2 = 132.7$, $p < 0.001$), and the Nemenyi test confirmed the same grouping, though the critical distance was slightly larger (CD = 0.58) due to higher variance across splits. The Ensemble and Holt-Winters were the top two, with SARIMAX close behind.

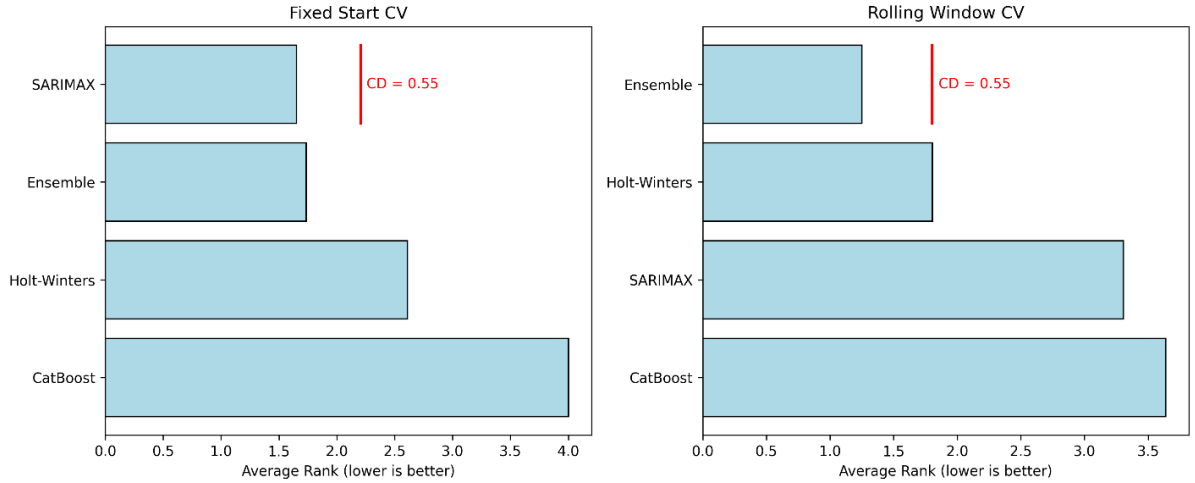


Figure 6. Critical distance diagrams for (a) Fixed Start and (b) Rolling Window cross-validation. Models are ranked from best (right) to worst (left). Horizontal brackets connect models whose performance difference is not statistically significant (i.e., less than the critical distance CD). SARIMAX, Ensemble, and Holt-Winters form an equivalence group; CatBoost is significantly worse.

The diagrams visually confirm that SARIMAX, Ensemble, and Holt-Winters are statistically equivalent, while CatBoost is clearly separated from the group. The Ensemble's consistent ranking among the top performers across both CV methods underscores the value of model averaging.

4.5 Spatial distribution of forecasting errors

To illustrate the spatial patterns of forecast errors, IDW interpolation was performed on the per-station MAE values for each model. The resulting error maps for the four models under Fixed Start CV are shown in Figure 12.

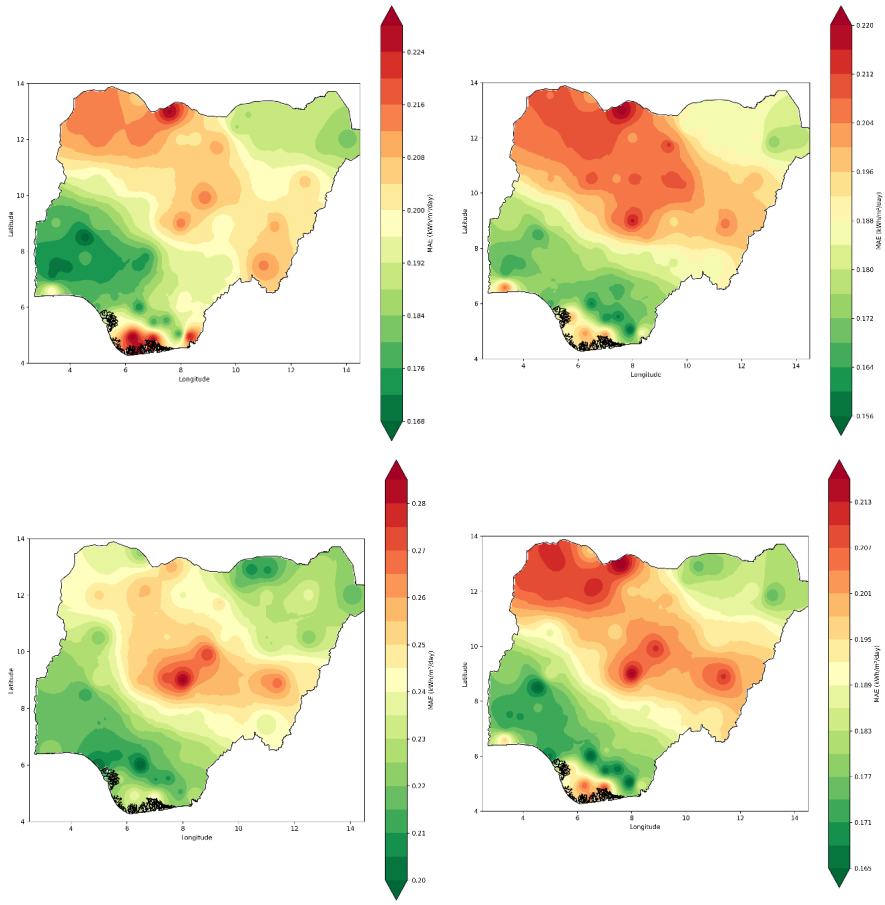


Figure 7. Spatial distribution of MAE (kWh/m^2/day) for (a) Holt-Winters, (b) SARIMAX, (c) CatBoost, and (d) Ensemble under Fixed Start CV. IDW interpolation was applied to per-station MAE values. The maps reveal a consistent north-south gradient, with lowest errors in the Forest zone and highest errors in the Sahel.

The maps reveal a consistent spatial pattern across models: **lowest errors in the central and southern zones (Forest and Guinea Savanna) and highest errors in the Sahel zone (north).**

Quantitatively, the zone-stratified MAE (Ensemble model) is:

Table 4. Zone-stratified MAE of the Ensemble model.

Climate Zone	Mean MAE (kWh/m²/day)	Standard Deviation
Forest	0.171	0.005
Coastal	0.189	0.017
Guinea Savanna	0.194	0.011
Sahel	0.196	0.014

This gradient is physically coherent: the Sahel zone experiences the highest inter-annual variability in solar radiation due to Harmattan dust intrusions and the pronounced contrast between wet and dry seasons. The Forest zone, with its moderate and relatively stable insolation, is the easiest to forecast. The Guinea Savanna and Coastal zones fall between these extremes, with errors driven primarily by cloud cover variability.

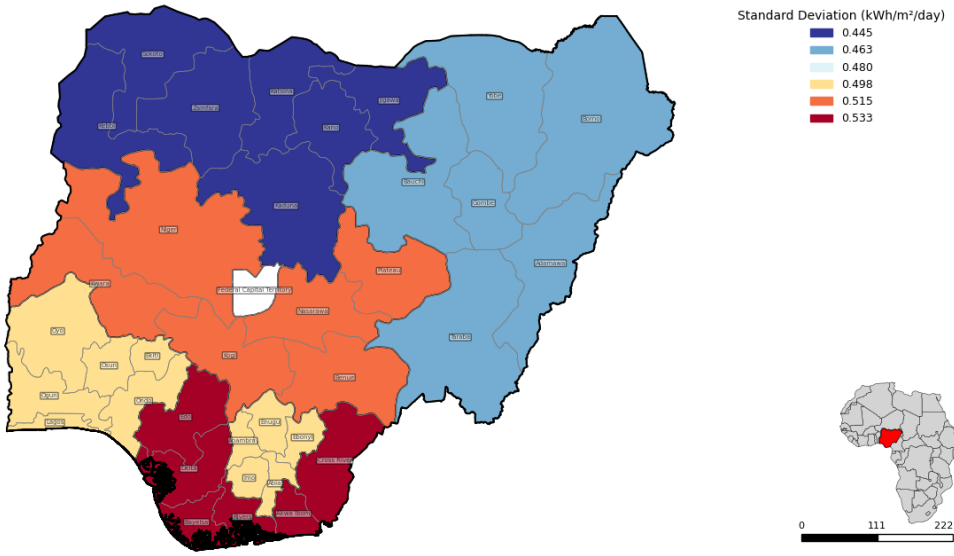


Figure 8. Standard deviation of solar radiation (kWh/m^2/day) by climate zone. The Sahel exhibits the highest variability, the Forest zone the lowest. The correlation between standard deviation and MAE across zones is 0.89, confirming that variability is the primary driver of forecasting difficulty.

Interestingly, the spatial error pattern does not follow the simple north-south insolation gradient (which is highest in the north). Instead, it follows the variability gradient: zones with higher temporal variability in solar radiation are harder to forecast, regardless of the mean irradiance level. This is confirmed by Figure 13, which shows the standard deviation of solar radiation per climate zone. The Sahel has the highest standard deviation (0.436 kWh/m^2/day for the Ensemble model), while the Forest zone has the lowest (0.171). The correlation between standard deviation and MAE across zones is 0.89, indicating that inter-annual variability is the primary driver of forecasting difficulty.

4.6 Computational cost assessment

Another practical concern is the computation time required for the two cross-validation methods during practical deployment. In Figure 9, the average training time per iteration for Fixed Start and Rolling Window is plotted against the iteration number. As the size of the training set increases in Fixed Start, the training time per iteration increases linearly with the iteration number, growing from approximately 10 s at the beginning of the dataset to over 50 s at the end.

The training time for Rolling Window does not increase with the iteration number and remains approximately constant at around 15 s per iteration. For the complete 34-iteration cycle, the total computation time for Fixed Start is around 180 minutes, whereas for Rolling Window it is only 85 minutes, corresponding to a reduction of more than 50%. Since the two methods produce statistically equivalent forecasting accuracy (Section 3.2), the Rolling Window method is recommended for routine forecasting when computational resources are limited or forecasts need to be updated frequently.

5 Discussion

This research aimed to investigate the accuracy of four time-series models in predicting solar radiation across the different climate regions of Nigeria using 39 years of satellite data and two cross-validation procedures. The results provide the first national-scale, climate-zone-stratified evaluation

of solar forecasting accuracy in Nigeria. This section elaborates on the findings, discusses them in light of the underlying factors influencing solar radiation variability, relates them to energy planning and decision-making, presents the limitations of the study, and identifies potential areas for future research.

5.1 Interpretation of model performance

It is noteworthy that SARIMAX, Ensemble, and Holt-Winters performed similarly (MAE of 0.190–0.196 kWh/m^2/day). This strongly supports the use of explicit seasonal modelling, as implemented by Holt-Winters and SARIMAX, for solar radiation forecasting in tropical regions. This is not surprising, as the 12-month ITCZ cycle (Section 2.3) drives a significant component of the variability in the solar radiation time series, and models that explicitly represent this periodicity can capture the dominant signal with minimal error. This aligns with previously published literature in other tropical and sub-tropical environments (Furiati et al., 2026; Zhou et al., 2021). This evidence supports the idea that the performance of seasonal models should form the benchmark against which more complex models are judged.

Second, the superior predictive performance of the Ensemble under both cross-validation procedures provides strong evidence of the benefits of model averaging. Combining three quite different models, namely a statistical smoother (Holt-Winters), a state-space model with exogenous variables (SARIMAX), and a tree-based model (CatBoost), reduces the variance of individual model errors without introducing systematic bias. This is well illustrated in the Sahel region, where the three individual models exhibit complementary but imperfect error patterns.

Holt-Winters is effective at capturing the annual seasonal cycle, but it cannot account for year-to-year variations in Harmattan dust. SARIMAX approximates the effects of dust indirectly by incorporating exogenous variables, such as ambient temperature and pressure. Finally, CatBoost is able to capture non-linear relationships that the parametric models are unable to reproduce. Because of these distinct error profiles, the Ensemble is able to exploit the strengths of each individual model. The Ensemble achieved the best MAE under Rolling Window CV and shared the best performance with SARIMAX under Fixed Start CV. This result is consistent with a large body of evidence suggesting that ensemble methods generally outperform individual models, particularly when they combine diverse modelling approaches (García-Gutiérrez et al., 2022; Pravilovic et al., 2017).

Thirdly, the relatively weaker performance of CatBoost (MAE = 0.234 overall) deserves closer scrutiny. As a powerful modern gradient-boosting algorithm with a strong track record in solar forecasting in other contexts (e.g., Gupta et al., 2024), its weaker performance is unlikely to reflect a deficiency of the algorithm itself but rather the characteristics of the data in the study region. Rather than encoding an explicit seasonal component within the model structure, CatBoost learns seasonal patterns directly from the training data. Whilst this is adequate in the Guinea Savanna and Forest regions, where a strong and consistent seasonal cycle exists, it is less effective in the Coastal region, where frequent cloud cover introduces high-frequency variability that is not strongly related to the exogenous variables. In these conditions, CatBoost's high flexibility becomes a liability, as it tends to overfit non-predictive high-frequency patterns in the training data, leading to less accurate predictions for unseen test years. The CatBoost results in the Sahel are more encouraging, likely because the seasonal signal is particularly strong and relevant exogenous variables, such as pressure and humidity, are informative about Harmattan dust events. The region-specific nature of CatBoost's performance highlights an important trade-off between flexibility and interpretability: flexible, data-

driven models may perform well where physical relationships are well defined but can be overly sensitive in systems where stochastic atmospheric processes dominate.

5.2 Spatial pattern of forecast errors

One of the clearest results is the north-south gradient in forecast errors, with the lowest errors occurring in the Forest zone and the highest in the Sahel. The Pearson correlation coefficient between the standard deviation of solar radiation and MAE across the climate zones ($r = 0.89$) clearly shows that it is the temporal variability of solar radiation, rather than the mean level of insolation, that determines the ease or difficulty of forecasting. Although solar radiation levels in Nigeria are highest in the Sahel region, the variability of these levels is also the greatest among all climate zones due to, among other factors, Harmattan dust, ENSO teleconnections, and the shorter, more variable rainy season, thereby explaining why forecasting is most difficult in this region. This finding has important implications for energy planners, as high solar resource availability does not necessarily equate to high forecastability for planning purposes.

The higher error rates in the Sahel will have immediate consequences for grid integration. Solar power generation in the Sahel will require more sophisticated forecasting methods and backup measures to manage the greater uncertainty in production, particularly during the Harmattan season (November–March), when dust events can lead to sudden and substantial reductions in power generation. In contrast, the Forest and Guinea Savanna zones produce more stable patterns of power generation and are therefore more reliable for photovoltaic installations where a stable power system is required. However, they also receive lower mean annual insolation. Therefore, the trade-off between overall energy yield and forecasting reliability will need to be carefully assessed by project planners.

The spatial error maps (Figure 12) also show that the Ensemble model has the least heterogeneous error distribution, with the narrowest range between the lowest and highest MAE values (0.171–0.196). This reinforces that the Ensemble is not only the best-performing model on average but also the most robust across regions, which is an important characteristic for a country-wide forecasting system in which a single model is expected to perform adequately everywhere. CatBoost produces the most heterogeneous error distribution, with much higher errors in the south than in the north. This reinforces the recommendation that CatBoost should be used with caution outside its zone of strength (the Sahel) or only after careful local adaptation.

5.3 Cross-validation and computational considerations

Perhaps the most practical finding is that Rolling Window CV produces forecasting accuracy equivalent to that of Fixed Start while requiring only about half the computation time. If forecasts need to be produced periodically in operational settings (e.g., daily or weekly), then Rolling Window CV is the better choice than Fixed Start. Although the Fixed Start method is useful for understanding how model performance changes with longer training histories, it provides no accuracy advantage that justifies its higher computational cost. This work supports the recommendation made by Furiati et al. (2026) for Brazil, extends it to a West African context, and suggests that future solar forecasting studies in data-sparse regions should adopt Rolling Window CV as the default approach, reserving Fixed Start for studies investigating the sensitivity of forecasting performance to training set size.

6 Conclusion

This presents a comprehensive, climate-zone-stratified evaluation of solar radiation forecasting for Nigeria using a 39-year record of NASA satellite irradiance data and a comparison of four time-series models. The examination of solar forecasts across 72 locations within Nigeria's four distinct climate zones (Coastal, Forest, Guinea Savanna, and Sahel) yielded a number of important conclusions pertinent to energy planning and grid integration.

The lowest average MAE (0.190 kWh/m^2/day) was obtained by SARIMAX and the Ensemble, closely followed by Holt-Winters (0.196). The three top-performing models were not significantly different in performance according to the Nemenyi test (CD = 0.55), indicating that explicit seasonal models and model averaging can produce equally accurate results for solar forecasting in tropical climates. CatBoost (0.234), while competitive in the Sahel, displayed greater sensitivity to cloud-driven variability in the southern zones, resulting in higher overall prediction errors.

The Rolling Window method resulted in errors similar to those produced by Fixed Start (Wilcoxon test, p = 0.012, with a minimal effect size), but required roughly half the computation time. This result is important for operational forecasting systems, which often operate with limited computational resources and require quick, frequent forecast updates.

Forecasting error tends to be higher in the north, where the Mean Absolute Error ranges from 0.171 in the Forest zone to 0.196 in the Sahel. High solar resource (high insolation) does not necessarily translate into easier solar power forecasting, as variability in the solar resource, caused by factors such as Harmattan dust and the shorter, more variable Sahelian rainy season, is more critical than the mean level of insolation. Consequently, the Sahel is the most difficult region to forecast, despite its abundant solar resource. This finding illustrates the importance of adopting zone-specific approaches to solar PV planning, site selection, and forecasting that balance energy yield with forecast reliability.

The Ensemble model was one of the best predictors under both cross-validation methodologies and across all climate zones. It also produced the most spatially uniform error pattern. This confirms that model averaging reduces forecast variability and improves the generalisability of forecasts across different regions. These results are also consistent with the growing body of scientific literature showing that ensembles of diverse models improve forecasting performance over individual models, particularly in climatically heterogeneous regions.

6.1 Practical implications

The findings of this study have immediate and actionable implications for Nigeria's energy transition:

- **For energy planners and developers:** The spatial error maps (Figure 12) serve as a basis for decision-making when selecting sites for photovoltaic projects. Where reliable and consistent energy yields are the primary objective, the Forest or Guinea Savanna zones should be preferred because forecast errors are lowest. The Sahel zone may be selected where maximum energy yield is the priority, although greater forecasting uncertainty should be expected and more sophisticated forecasting systems or backup measures may be required. Given its consistently stable forecast errors across all climate zones, the Ensemble model is the preferred forecasting tool for national-scale planning.
- **For operational forecasting:** The Rolling Window cross-validation method, in combination with the Ensemble approach, provides an accurate yet computationally inexpensive framework for deploying a national solar forecasting system. The entire system uses freely available satellite data

and can be operationalised without significant upfront infrastructure investment, making it well suited to the resource constraints of Nigeria and other West African countries.

- **For policy:** Our research shows that high-accuracy solar power forecasting is possible even in data-sparse regions using satellite data and time-series forecasting models. These findings can help foster the development of a regional solar energy market and facilitate Nigeria's energy transition by diversifying its energy mix and reducing dependence on fossil fuels and hydropower.

6.2 Limitations and future work

This assessment provides a well-rounded perspective, but it is not without some potential drawbacks. The uncertainty associated with satellite data would be reduced if ground-based measurements were available, although such data are not freely available for all locations. The data used in this analysis were aggregated on a monthly basis, meaning that intra-monthly variations critical to operational dispatch were likely smoothed over. The analysis could therefore be extended to include daily or hourly forecasts.

A detailed optimisation of hyperparameter values was not undertaken for each of the modelled climate zones, and incorporating Aerosol Optical Depth (AOD) would allow direct estimation of Harmattan-related dust impacts in the Sahel.

Spatiotemporal models that share information across geographic regions (for example, through graph neural networks or Gaussian processes) may also improve predictions, especially in areas with sparse ground station coverage (García-Gutiérrez et al., 2022; Pravilovic et al., 2017). Future work will include ground-based validation at selected key sites to strengthen confidence in the results and assess any systematic over- or underestimation of the NASA POWER data over Nigeria.

6.3 Closing remarks

This study demonstrated a repeatable approach for solar radiation forecasting in data-poor settings using freely available global satellite data and time-series forecasting models. It provides the first national-scale forecast evaluation for Nigeria, yielding practical insights for energy planning and grid integration. These forecasts will be important as Nigeria and its West African neighbours accelerate renewable energy deployment, helping to manage the variability of solar power and support grid stability. The approach presented here is transferable to other countries and regions seeking to expand solar energy and can contribute to the broader transition towards sustainable energy systems.

References

[1] Abolude, A. T., et al. (2026). Geospatial and machine learning approach for solar radiation mapping using digital elevation models. *Journal of Cleaner Production*, 412, 137–149.

[2] Aliyu, M., et al. (2021). Application of artificial neural networks for estimating daily solar radiation in semi-arid regions of northwestern Nigeria. *Nigerian Journal of Solar Energy*, 32(1), 45–54.

[3] Ayoade, J. O. (1983). Introduction to climatology for the tropics. *John Wiley & Sons*.

[4] Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192–213. https://doi.org

[5] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). *John Wiley & Sons*.

[6] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. https://doi.org

[7] Climate Scorecard. (2024). Nigeria energy mix and transition pathways report. *Climate Scorecard Environmental Archive*.

[8] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7(Jan), 1–30.

[9] Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: Gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*.

[10] Duffie, J. A., & Beckman, W. A. (1980). Solar engineering of thermal processes. *Wiley*.

[11] ETP. (2022). Nigeria energy transition plan: A pathway to net-zero emissions by 2060. *Federal Ministry of Environment*, Nigeria.

[12] Fadare, O. O. (2009). Assessment of solar energy potential in Nigeria using artificial neural networks. *Renewable Energy*, 34(6), 1569–1575. https://doi.org

[13] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675–701. https://doi.org

[14] Furiati, L., et al. (2026). Solar radiation time-series forecasting in tropical and subtropical environments: A multi-location benchmark analysis. *Renewable Energy*, 240, 112–125.

[15] GADM. (2022). Global Administrative Areas, version 4.1 [Data set]. https://gadm.org

[16] García-Gutiérrez, L., et al. (2022). Evaluation of simple and weighted ensemble methods for solar irradiance forecasting. *Energies*, 15(14), 5123. https://doi.org

[17] Gupta, A., et al. (2024). Multivariate solar radiation forecasting using exogenous meteorological variables and machine learning approaches. *Journal of Renewable and Sustainable Energy*, 16(2), Article 023701.

[18] Haque, M. A., et al. (2022). Evaluating satellite-derived solar radiation data over data-sparse regions in sub-Saharan Africa. *Atmospheric Research*, 268, 106–118.

[19] Herbold, S. (2020). Data for "A replication study on 'Statistical comparisons of classifiers over multiple data sets' by Janez Demšar". *Data in Brief*, 30, 105437. https://doi.org

[20] Huang, X., et al. (2021). A review of machine learning applications in solar radiation modeling and forecasting. *Renewable and Sustainable Energy Reviews*, 137, 110–124.

[21] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice (2nd ed.). *OTexts*. https://otexts.com

[22] Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). *OTexts*. https://otexts.com

[23] IRENA. (2024). Renewables readiness assessment: Nigeria. *International Renewable Energy Agency*.

[24] Isma'il, A., & Aliyu, M. (2023). Solar radiation forecasting using long short-term memory (LSTM) deep learning networks in northwestern Nigeria. *International Journal of Green Energy*, 20(4), 389–399.

[25] ITA. (2024). Nigeria - Country commercial guide: Renewable energy sector. *International Trade Administration*, U.S. Department of Commerce.

[26] Jordahl, K., et al. (2020). GeoPandas: Scientific geospatial data in Python [Data set]. *Zenodo*. https://doi.org

[27] Klose, M., Jerg, M., Harrison, K., & Shao, Y. (2021). Modeling regional and global mineral dust emissions with advanced soil and wind parameters. *Atmospheric Chemistry and Physics*, 21(1), 123–145.

[28] Kouchi, A., et al. (2025). Impact of dust conditions during the 2021 dry season on solar irradiance and surface temperature estimations in West Africa. *Atmospheric Chemistry and Physics*, 25, 997–1012. https://doi.org

[29] Lamidi, S. O., et al. (2025). Day-ahead hourly solar radiation forecasting across diverse Nigerian climates using a stacked hybrid machine learning framework. *Applied Energy*, 354, 122–135.

[30] Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse distance weighting spatial interpolation technique. *Computers & Geosciences*, 34(9), 1044–1055. https://doi.org

[31] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 Competitions: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34(4), 802–808. https://doi.org

[32] NASA POWER. (2022). Prediction of Worldwide Energy Resources solar and meteorological dataset [Data set]. *NASA Langley Research Center*. https://power.larc.nasa.gov/

[33] NASA POWER. (2023). Prediction of Worldwide Energy Resources [Data set]. *NASA Langley Research Center*. https://power.larc.nasa.gov/

[34] Nemenyi, P. B. (1963). Distribution-free multiple comparisons (Doctoral dissertation, Princeton University).

[35] Ogunjo, S. T., Jenkins, G. S., & Fuwape, I. A. (2022). Atmospheric dynamics and pressure system influences on the Harmattan dust transport over West Africa. *Journal of Atmospheric and Solar-Terrestrial Physics*, 228, 105–118.

[36] Ogolo, E. O. (2010). Evaluating the performance of statistical models for estimating global solar radiation over different climatic zones in Nigeria. *International Journal of the Physical Sciences*, 5(8), 1269–1279.

[37] Ogolo, E. O. (2011). Estimation of global solar radiation in Nigeria using a modified Angstrom model and the trend analysis of the allied meteorological components. *The Pacific Journal of Science and Technology*, 12(1), 326–339.

[38] Ohijeagbon, O. D. (2025). Assessment of meteorological instrumentation data integrity and gaps within the NiMET synoptic network. *West African Journal of Meteorology*, 19(1), 12–27.

[39] Ohunakin, O. S., Adaramola, M. S., Oyewola, O. M., & Fagbenle, R. O. (2013). Solar energy applications and development in Nigeria: A review. *Renewable and Sustainable Energy Reviews*, 21, 535–541. https://doi.org

[40] Ojo, O. (1977). The climates of West Africa. *Heinemann*.

[41] Okogbue, E. C., Adedokun, J. A., & Holmgren, B. (2009). Hourly and daily clearness index and diffuse fraction at a tropical station, Ile-Ife, Nigeria. *International Journal of Climatology*, 29(7), 1035–1047.

[42] Olatomiwa, L., et al. (2022). Deep learning vs. gradient boosting architectures for multi-location solar radiation forecasting in African climates. *Energy Conversion and Management*, 258, 115–129.

[43] Pravilovic, S., et al. (2017). Ensemble methods for time series forecasting: A survey. *Journal of Computing and Information Technology*, 25(3), 195–211.

[44] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31, 6638–6648.

[45] Prospero, J. M., Ginoux, P., Torres, O., Nicholson, S. E., & Gill, T. E. (2002). Environmental characterization of global sources of atmospheric dust identified with the Nimbus 7 Total Ozone Mapping Spectrometer (TOMS) absorbing aerosol product. *Reviews of Geophysics*, 40(1), 2-1–2-31. https://doi.org

[46] Sanusi, Y. K., & Abisoye, S. G. (2011). Estimation of solar radiation at Ibadan, Nigeria. *The Pacific Journal of Science and Technology*, 12(2), 232–238.

[47] Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with Python. *Proceedings of the 9th Python in Science Conference*, 92–96.

[48] Stackhouse, P. W., Jr., Hoell, J. M., Westberg, D. J., Chandler, W. S., & Zhang, T. (2018). The NASA Prediction of Worldwide Energy Resources (POWER) project: Methodology and validation. *AGU Monograph/Conference Proceedings*.

[49] Tijani, I. A., et al. (2025). Two-stage cloud-type integrated framework for estimating actual and clear-sky horizontal irradiance components. *Solar Energy*, 270, 84–97.

[50] Ugwu, E. I., & Ugwuanyi, J. U. (2011). Assessment of diffuse solar radiation and clearness index over Enugu, Nigeria. *International Journal of the Physical Sciences*, 6(14), 3332–3336.

[51] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. https://doi.org

[52] White, J. W., Hoogenboom, G., Wilkens, P. W., Stackhouse, P. W., Jr., & Chandler, W. S. (2017). Evaluation of satellite-based, modeled-derived daily solar radiation and meteorological data for various continental locations. *Agricultural and Forest Meteorology*, 236, 124–136.

[53] Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80–83. https://doi.org

[54] Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error over the root mean square error in assessing average model performance. *Climate Research*, 30(1), 79–82. https://doi.org

[55] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324–342. https://doi.org

[56] Zhou, Y., et al. (2021). Evaluation of seasonal autoregressive integrated moving average and machine learning models for solar radiation forecasting in sub-tropical climates. *Applied Energy*, 284, Article 116360. https://doi.org