STM JOURNALS

International Journal of
Energy and Thermal Applications

ISSN: 3108-1479
Volume 4, Issue 2, 2026
July–December
DOI (Journal): 10.37591/IJETA

https://journals.stmjournals.com/ijeta

Review IJETA

Assessing the ARIMA Model for Estimating Monthly Solar Irradiance from Approximated Daily Data in Uyo, Nigeria

Samuel Edet Bassey*

Abstract

This study assesses autoregressive integrated moving average (ARIMA) models for capturing, predicting, and visualizing solar irradiance patterns in Uyo, Nigeria. Previous studies specific to Uyo solar irradiance forecasting explored regression models, which are not suitable for advanced analysis. This gap presented a challenge for accurate energy forecasting, grid integration, and planning. This study utilizes a dataset of approximated solar irradiance values, trained and tested using ARIMA for suitability. The methodology comprises the study location and ARIMA model development. Eighteen ARIMA models were evaluated using the grid search method. The study found ARIMA (2, 1, 2) with a low Akaike Information Criterion (AIC) of 229.556 as the best model for capturing a 24-row, 288-cell solar irradiance observation. The statistical performance analysis reveals that the dataset has a mean ranging from 123 to 192 W/m², a standard deviation between 210 and 300 W/m², and a maximum of 585 to 911 W/m² across the months of January through December, with a total row count of 24 and 288 cells. The Augmented Dickey-Fuller (ADF) test produced an ADF statistic of -4.457189 with a p-value of 0.000235, while the differenced series yielded an ADF statistic of -2.580365 and a p-value of 0.097124, confirming stationarity of the irradiance series. Forecast values from step 25 to step 29 indicate that irradiance might peak slightly above 2 W/m². Evaluation of eighteen ARIMA models using the Akaike Information Criterion showed that ARIMA (2, 1, 2) achieved the lowest AIC value of 229.556, outperforming competing models such as ARIMA (2, 1, 1) with 230.906 and ARIMA (2, 1, 0) with 231.019. The best model forecast values from step 25 to step 29 show a pattern that will be maintained over future predictions, although the curve will be smoother, while ARIMA (1, 1, 1) indicates a constant forecast with no change. Correlation, distribution, and residual plots confirmed that the data were properly captured. This is suitable for advanced analysis, as the study provides a background for further research using long-term datasets to improve prediction accuracy and support solar energy planning in Uyo, Nigeria.

Keywords: Solar radiation, autoregressive integrated moving average (ARIMA), solar irradiance forecasting, time series analysis, renewable energy, Uyo, Nigeria

***Author for Correspondence**
Samuel Edet Bassey
E-mail: sammiebassey5@gmail.com

Research Scholar, Department of Mechanical and Aerospace Engineering University of Uyo, Akwa Ibom State, Nigeria

Received Date: July 07, 2026
Accepted Date: September 23, 2026
Published Date: October 10, 2026

Citation: Samuel Edet Bassey. Assessing the ARIMA Model for Estimating Monthly Solar Irradiance from Approximated Daily Data in Uyo, Nigeria. International Journal of Energy and Thermal Applications. 2026; 4(2): 65–72p.

INTRODUCTION

Solar energy, as a renewable energy source, is cost-effective, flexible, affordable, low-maintenance, and has low carbon emissions. It is an expanding energy option that can be easily transmitted and distributed over a wide area. Photovoltaic energy technology is an area of solar energy technology concerned with the interaction of electrons with radiant energy in semiconductor materials. This radiant energy possesses photons that can be reflected, absorbed, and transmitted

© STM Journals 2026. All Rights Reserved

through the cell. Photovoltaic energy, founded as an effect in 1839 by Edmond Becquerel and developed by Heinrich Hertz in the 1870s, is responsible for the silent conversion of light to electricity through semiconductors spanning diverse material groups such as voluminous materials (mono- or multi-crystalline silicon), thin-layer films (amorphous silicon/germanium, etc.), nanomaterials, polymers, and new materials. Its standalone system can provide electricity storage during peak demand when incorporated with battery storage.

As a solar technology, it is concerned with radiation, irradiance, insolation, radiant existence, air mass, and the solar constant, among others. Although vast applications show considerable prospects, its integration into energy infrastructure faces numerous drawbacks, mainly due to its intermittent nature, hence the need for climate/weather modelling. Approaches developed to capture this through time series forecasting, regression, and correlation models are included in this study. Pelkowski (2009) [1] introduced an approach for deriving and interpreting polynomial relationships between ground-level solar irradiance and relative sunshine duration. Paulescu et al. (2021) [2] used polynomial regression techniques to determine clear-sky and cloud transmittance through the Angstrom–Prescott equation. In a related study, Paulescu et al. (2016) [3] created thirty-three Angstrom–Prescott regression equations and tested them against data from fifty-nine European stations. In Nigeria, Augustine and Nnabuchi (2009) [4] formulated regression equations to predict monthly mean global solar radiation and sunshine duration in Calabar, Enugu, and Port Harcourt, drawing on seventeen years of meteorological records. Njau (1996) [5] derived Angstrom–Prescott equations from relevant physical parameters, whereas Liu et al. (2012) [6] estimated model parameters through detailed climatic investigations at eighty locations across China. Mabasa et al. (2020) [7] used mathematical models to estimate solar irradiation as an alternative to data from the South African Weather Service, noting that the existing meteorological stations were insufficient for comprehensive climatic studies. Almorox et al. (2005) [8] developed monthly predictive equations using daily global radiation data from Toledo, Spain. Iradukunda and Chiteka (2015) [9] established annual Angstrom–Prescott coefficients based on data from twenty-nine stations in Zimbabwe. In the Nigerian context, Ekanem, Akpam, and Utah (2016) [10] used fourteen years of monthly mean daily global solar radiation data from Uyo to develop linear, quadratic, and cubic Angstrom–Prescott models. Ibeh et al. (2012) [11] predicted monthly global solar radiation using multilayer perceptron neural networks alongside the Angstrom–Prescott model. Augustine and Nnabuchi (2010) [12] also applied meteorological variables such as global solar radiation, sunshine duration, maximum temperature, cloudiness index, and relative humidity to generate several linear and multilinear regression models. Taken together, these studies show that regression-based and artificial intelligence techniques can effectively predict solar energy conditions across a range of climatic regions [13, 14].

Although previous studies have provided valuable methodologies for solar energy forecasting, they rely on correlation and regression models that are not as advanced as autoregressive integrated moving average (ARIMA). This study provides a methodology to assess different ARIMA models, proposing the best-fitted model to capture a one-year solar irradiance pattern required for weather modelling and planning in Uyo, Nigeria, with accuracy.

METHODOLOGY

Study Location

Uyo is a landlocked city in the southern part of Nigeria; it is the capital city of Akwa Ibom State, Nigeria. This city is located in the tropics between 5.05°N and longitude 7.93°E, with an elevation of 45 m above sea level. According to the Köppen-Geiger climate classification system, Uyo features a tropical monsoon climate (AM). This climate type is characterized by significant rainfall throughout most of the year, with a short dry season and consistently warm temperatures driven by its proximity to the equator. The climate in Uyo is warm and humid year-round. Average monthly temperatures range from approximately 24.1℃ (75.4°F) in the coldest months (July and August) to 27.5℃ (81.6°F) in the warmest month (February). The city experiences high relative humidity, often exceeding 80% for most

© STM Journals 2026. All Rights Reserved

of the year, with humidity levels peaking around 88%–91% from June to October. Uyo receives substantial annual precipitation, with averages reported between 2499 mm and 3015 mm. The warmest months are typically from June to September, while the drier period occurs between December and February.

ARIMA Workflow

The ARIMA model was implemented using the Python libraries Pandas, NumPy, Matplotlib, and Statsmodels. The series was visualized and differenced while monitoring its ADF value and p-value. A model (1, 1, 1) was tested to check its errors, predict five additional steps, and assess its effectiveness before using grid search over p (0–3), d (0–2), and q (0–3), checking the optimal combination that would yield a lower Akaike Information Criterion (AIC) index. The general ARIMA (p, d, q) model combines three components:

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \Delta^d X_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

AutoRegressive (AR) Component: $X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \epsilon_t$

Integrated (I) Component: $Y_t = \Delta^d X_t$;

Moving Average (MA) Component: $X_t = \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$

Where, p: Number of lag observations (AR order); d: Degree of differencing (I order); q: Size of moving average window (MA order); X_t: Current value of the series; $\emptyset_i$: AR coefficients; ϵ_t : White noise (random error).

STATISTICAL PERFORMANCE ANALYSIS

Dataset Description

Table 1 presents the descriptive statistics of the solar irradiance dataset used for modelling, covering the twelve months of the year from January to December. The dataset comprises a total row count of 24 observations, yielding 288 cells across all months. The mean solar irradiance values range from a minimum of 123.03 W/m² in August to a maximum of 191.53 W/m² in February, indicating moderate seasonal variation across the year. The standard deviation values are notably high, ranging between 210.40 W/m² in July and 299.48 W/m² in February, which suggests considerable dispersion and variability in the irradiance data throughout the year. The minimum values across all months are extremely low, approximately 1.0E-5 W/m², reflecting periods of near-total darkness or negligible solar radiation, typically occurring during nighttime hours. The median values range from 16.105 W/m² in August to 25.072 W/m² in February, while the 25th percentile values remain consistently low across all months, ranging from 0.0234 to 0.0365 W/m². Conversely, the 75th percentile values are substantially higher, ranging from 189.97 W/m² in August to 295.75 W/m² in February. The maximum values recorded range from 585.15 W/m² in August to 910.966 W/m² in February, demonstrating the peak irradiance levels attainable in Uyo. This wide range between minimum and maximum values, combined with the high standard deviation, confirms the highly variable and intermittent nature of solar irradiance in the study location, which justifies the need for robust time series modelling approaches such as ARIMA.

Table 1. Description of data.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
count	24	24	24	24	24	24	24	24	24	24	24	24
mean	165.20	191.53	167.30	174.64	170.56	156.70	134.56	123.03	166.25	172.89	189.90	176.62
std	258.31	299.48	261.59	273.07	266.69	245.01	210.40	192.37	259.95	270.33	296.93	276.16
min	1.2E-5	1.4E-5	1.2E-5	1.3E-5	1.2E-5	1.1E-5	9.8E-6	8.9E-6	1.2E-5	1.3E-5	1.4E-5	1.3E-5
25%	0.0315	0.0365	0.0319	0.0333	0.0325	0.0299	0.0256	0.0234	0.0317	0.0329	0.0362	0.0337
50%	21.626	25.072	21.900	22.861	22.327	20.512	17.615	16.105	21.763	22.632	24.859	23.120
75%	255.09	295.75	258.33	269.66	263.37	241.96	207.78	189.97	256.71	266.96	293.23	272.72
max	785.736	910.966	795.71	830.62	811.22	745.29	640.00	585.15	790.72	822.31	903.21	840.04

© STM Journals 2026. All Rights Reserved

Stationarity Assessment

Figure 1 displays the irradiance series plot of the original data alongside its differenced irradiance series. The stationarity of the time series was assessed using the Augmented Dickey-Fuller (ADF) test. The ADF statistic for the original series was -4.457189, with a corresponding p-value of 0.000235. Since this p-value is considerably less than the conventional significance level of 0.05, the null hypothesis of non-stationarity was rejected, indicating that the original irradiance series was stationary. Following differencing, the differenced series produced an ADF statistic of -2.580365 with a p-value of 0.097124. Although this p-value is slightly above the 0.05 threshold, the ADF statistic remains negative, and the differencing operation was applied to further stabilize the mean and variance of the series. Both the p-value and ADF statistic collectively indicate that the data exhibit stationarity, providing a valid foundation for fitting ARIMA models. The visual is included. It is next to Figure 1. The differenced irradiance series supports these statistical findings, as the differenced series appears to fluctuate around a constant mean with relatively stable variance over time.

Baseline Forecast Values

Table 2 presents the baseline forecast values generated from steps 25 to 29. The forecasted irradiance values were 1.281019 W/m², 1.902897 W/m², 2.204743 W/m², 2.351253 W/m², and 2.422366 W/m² for steps 25 through 29, respectively. These values exhibit a gradual and consistent upward trend, suggesting that the irradiance might peak slightly above 2 W/m² over the forecast horizon. The incremental increases between successive steps were relatively small, indicating a stable and smooth progression in the forecasted values. These baseline forecasts serve as a reference point against which the performance of the optimized ARIMA models can be compared.

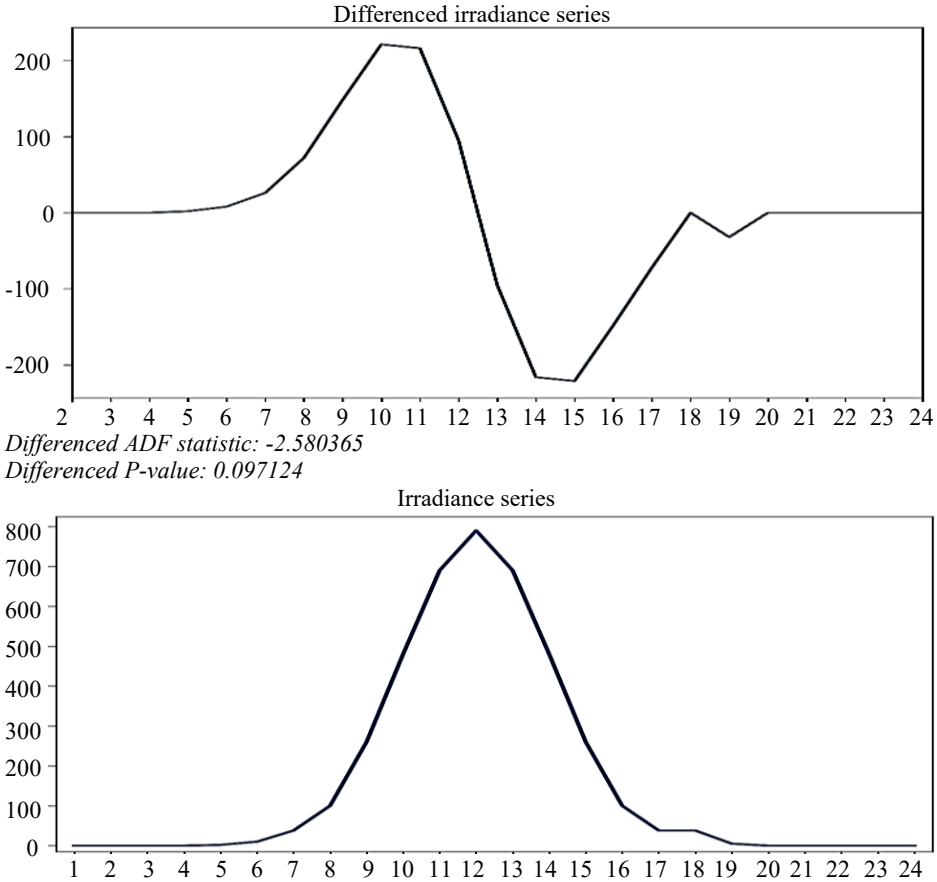


Figure 1. Irradiance series plot.

© STM Journals 2026. All Rights Reserved

Table 2. Forecast values.

Step	Value
25	1.281019
26	1.902897
27	2.204743
28	2.351253
29	2.422366

Table 3. ARIMA models' index.

ARIMA models	AIC value	ARIMA models	AIC value	ARIMA models	AIC value
0, 0, 0	284.415	0, 1, 2	233.698	1, 1, 1	236.874
0, 0, 1	263.345	1, 0, 0	262.137	1, 1, 2	234.510
0, 0, 2	251.320	1, 0, 1	250.752	2, 0, 0	239.519
0, 1, 0	249.109	1, 0, 2	244.791	2, 0, 1	240.692
0, 1, 1	238.581	1, 1, 0	238.991	2, 0, 2	239.061
2, 1, 0	231.019	2, 1, 1	230.906	2, 1, 2	229.556

Best ARIMA (2, 1, 2) AIC = 229.556

Table 4. Best model forecast values.

Step	Value
25	7.202710
26	28.300456
27	41.694997
28	39.249299
29	28.267817

ARIMA Model Selection and Evaluation

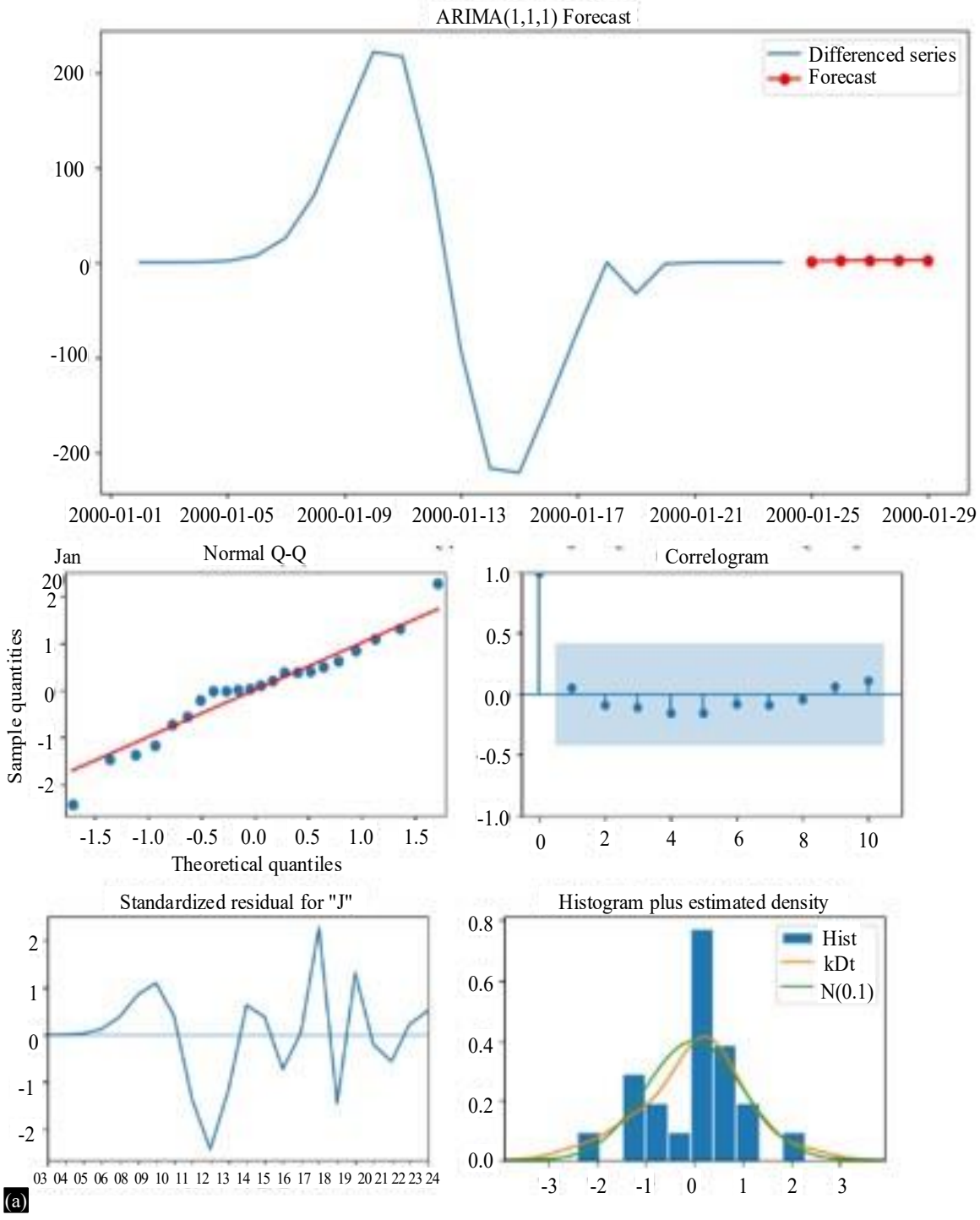
Table 3 presents the AIC values for 18 different ARIMA models evaluated using the grid search method. The models tested span a range of autoregressive (p), integrated (d), and moving average (q) orders, specifically combinations where p, d, and q each range from 0 to 2. The AIC values range from a minimum of 229.556 to a maximum of 284.415, providing a clear basis for model comparison and selection. The ARIMA (0, 0, 0) model, which represents a purely random process with no autoregressive or moving average components, yielded the highest AIC value of 284.415, confirming its unsuitability for capturing the underlying structure of the irradiance data. As model complexity increased, the AIC values generally decreased, indicating improved model fit. The ARIMA (0, 1, 2) model achieved an AIC of 233.698, while ARIMA (1, 1, 1) recorded 236.874 and ARIMA (1, 1, 2) recorded 234.510. The ARIMA (2, 1, 0) model produced an AIC of 231.019, ARIMA (2, 1, 1) yielded 230.906, and ARIMA (2, 1, 2) achieved the lowest AIC value of 229.556. Based on the principle that lower AIC values indicate better model performance, ARIMA (2, 1, 2) was selected as the best model for capturing the solar irradiance patterns in Uyo. This model incorporates two autoregressive terms, one degree of differencing, and two moving average terms, striking an optimal balance between model complexity and goodness of fit.

Best Model Forecast Analysis

Table 4 presents the forecast values generated by the best-performing model, ARIMA (2, 1, 2), from steps 25 to 29. The forecasted values were 7.202710 W/m², 28.300456 W/m², 41.694997 W/m², 39.249299 W/m², and 28.267817 W/m² for steps 25 through 29, respectively. These values exhibit a distinct pattern: an initial sharp increase from step 25 to step 27, peaking at 41.694997 W/m², followed by a gradual decline through steps 28 and 29. This pattern suggests that the model anticipates a short-term surge in solar irradiance before returning to lower levels, which is consistent with the intermittent and fluctuating nature of solar radiation in Uyo.

© STM Journals 2026. All Rights Reserved

Figure 2 presents a visual comparison of the ARIMA model forecasts. The forecast from the best ARIMA (2, 1, 2) model from steps 25 to 29 indicates that the same pattern will be maintained over future predictions, although the curve will be smoother than the historical data. In contrast, the ARIMA (1, 1, 1) model shows that the forecast will be constant, indicating no anticipated change over the forecast horizon. This comparison highlights the superior dynamic capturing capability of the ARIMA (2, 1, 2) model relative to simpler alternatives. Additional diagnostic graphs included in Figure 1 show the correlation, distribution, and residuals of the fitted model, all of which confirm that the data were properly captured and that the residuals behaved as white noise, validating the adequacy of the selected model.



© STM Journals 2026. All Rights Reserved

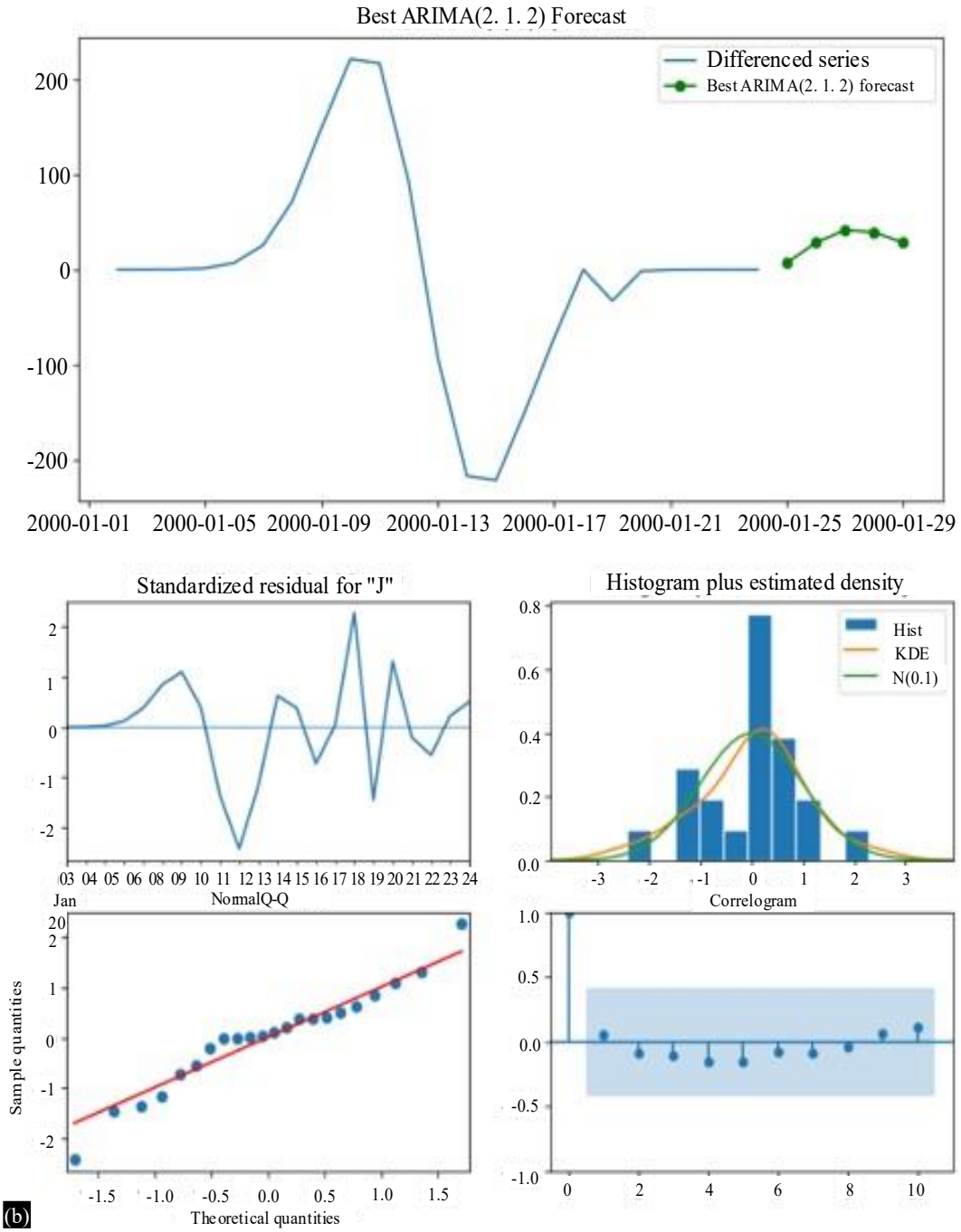


Figure 2. ARIMA model forecast comparison.

CONCLUSION

This study developed and evaluated an ARIMA model for forecasting solar irradiance in Uyo, Nigeria, using monthly solar irradiance data. The statistical analysis successfully characterized the temporal behavior, seasonal variation, and forecasted performance of solar irradiance in the study area. The results revealed that solar irradiance was highest during the dry season (November–February) and lowest during the rainy season (July–August), reflecting the influence of seasonal climatic conditions on the availability of solar energy. The ADF test confirmed the stationarity of the series, and the grid search evaluation of 18 ARIMA models identified ARIMA (2, 1, 2) with an AIC of 229.556 as the optimal model. The best model forecasts revealed a dynamic pattern that maintained the underlying

© STM Journals 2026. All Rights Reserved

structure of the data while producing smoother predictions. These findings demonstrate that ARIMA (2, 1, 2) is suitable for advanced analysis of solar irradiance in Uyo and provides a solid foundation for further research using long-term datasets to improve prediction accuracy and support solar energy planning in the region. The results of this study provide valuable information for renewable energy planning, photovoltaic system design, power generation scheduling, energy policy formulation, and sustainable energy development in Uyo and other locations with similar climatic conditions. Future studies should consider incorporating additional meteorological variables, such as cloud cover, temperature, relative humidity, rainfall, and wind speed, or employing hybrid forecasting techniques that combine ARIMA with artificial intelligence or machine learning models to further improve forecasting accuracy and robustness.

REFERENCES

1. Pelkowski J. A physical rationale for generalized Ångström–Prescott regression. Sol Energy. 2009;83(7):955-63. doi:10.1016/j.solener.2008.12.011.
2. Paulescu M, Stefu N, Calinoiu D, Paulescu E, Pop N, Boata R, et al. Ångström–Prescott equation: physical basis, empirical models and sensitivity analysis. Renew Sustain Energy Rev. 2016;62:495-506. doi:10.1016/j.rser.2016.04.012.
3. Paulescu M, Badescu V, Budea S, Dumitrescu A. What Ångström–Prescott equation tells us about the cloud and clear-sky climatologies? Theor Appl Climatol. 2022;147:239-50. doi:10.1007/s00704-021-03805-8.
4. Augustine C, Nnabuchi MN. Relationship between global solar radiation and sunshine hours for Calabar, Port Harcourt and Enugu, Nigeria. Int J Phys Sci. 2009;4(4):182-8. doi:10.5897/IJPS.9000481.
5. Njau EC. Generalised derivation of the Angstrom and Angstrom-Prescott equations. Renew Energy. 1996;7:105-8. doi:10.1016/0960-1481(95)00109-3.
6. Liu X, Xu Y, Zhong X, Zhang W, Porter JR, Liu W. Assessing models for parameters of the Ångström–Prescott formula in China. Appl Energy. 2012;96:327-38. doi:10.1016/j.apenergy.2011.12.083.
7. Mabasa B, Lysko MD, Tazvinga H, Mulaudzi ST, Zwane N, Moloi SJ. The Ångström–Prescott regression coefficients for six climatic zones in South Africa. Energies. 2020;13:5418. doi:10.3390/en13205418.
8. Srivastava RC, Pandey H. Estimating Angstrom-Prescott coefficients for India and developing a correlation between sunshine hours and global solar radiation for India. ISRN Renew Energy. 2013;2013:403742. doi:10.1155/2013/403742.
9. Iradukunda C, Chiteka K. Angstrom-Prescott type models for predicting solar irradiation for different locations in Zimbabwe. Stroj Vestn J Mech Eng. 2023;69(1-2):32-48. doi:10.5545/sv-jme.2022.331.
10. Ekanem IE, Akpan UE, Utah EU. Analysis of global solar radiation in Uyo, Akwa-Ibom State, Nigeria using temperature models. World J Appl Sci Technol. 2016;8(2):146-53.
11. Ibeh GF, Agbo GA. Comparison of empirical and artificial neural network models for the correlation of monthly average global solar radiation with sunshine hours in Minna, Niger State, Nigeria. Int J Phys Sci. 2012;7(8). doi:10.5897/IJPS11.1750.
12. Augustine C, Nnabuchi MN. Analysis of some meteorological data for some selected cities in the Eastern and Southern zone of Nigeria. Afr J Environ Sci Technol. 2010;4(2):92-9. doi:10.5897/AJEST09.092.
13. Liu Y, Tan Q, Pan T. Determining the parameters of the Ångström–Prescott model for estimating solar radiation in different regions of China: calibration and modeling. Earth Space Sci. 2019;6:1976-86. doi:10.1029/2019EA000635.
14. Wan KKW, Tang HL, Yang L, Lam JC. An analysis of thermal and solar zone radiation models using an Ångström–Prescott equation and artificial neural networks. Energy. 2008;33:1115-27. doi:10.1016/j.energy.2008.01.015.

© STM Journals 2026. All Rights Reserved