Predictive Solar Tracking Using Machine-Learned Effective Sun Position: Multi-Horizon Forecasting and Energy Yield Assessment in the Niger Delta

Wisdom Akpan Bassey[1]

Department of Electrical/Electronic Engineering,
University of Uyo, Akwa Ibom, Nigeria
wizdomie@gmail.com

Nseobong I. Okpura[2]

Department of Electrical/Electronic Engineering,
University of Uyo, Akwa Ibom, Nigeria
nseobongokpura@uniuyo.edu.ng

Unwana I. Ibanga[3]

Department of Electrical & Electronics Engineering
University of Uyo, Akwa Ibom State, Nigeria
unwanaiwok@uniuyo,edu,ng

Umoren Mfonobong A.[4]

Department of Electrical /Electronic Engineering,
University of Uyo, Akwa Ibom State Nigeria
mfon4gigis@yahoo.com

Abstract

Predictive solar tracking is most useful when the controller can respond to atmospheric changes without abandoning the known astronomical solar path. This study assesses multi-horizon forecasts of an irradiance-optimal effective sun position for Uyo, Nigeria, using 15 years of hourly NASA POWER data from 2010 to 2024. XGBoost, LSTM, Bi-LSTM, Transformer and a Hybrid Physics-ML residual model were evaluated at 1, 3, 6 and 24 h horizons, with persistence used as the baseline. The Hybrid model produced the lowest combined effective-position RMSE at every horizon: 1.089° at 1 h, 1.647° at 3 h, 2.341° at 6 h and 4.087° at 24 h. A pvlib-based 10 kWp simulation gave mean annual AC energy values of 14,287 kWh/year for fixed tilt, 17,621 kWh/year for SPA-based single-axis tracking and 18,241 kWh/year for 1 h predictive tracking. The 1 h predictive case added 620 kWh/year above SPA tracking in the simulation, but the additional benefit declined sharply as forecast lead time increased. The findings support short-horizon atmospheric correction as a supervisory layer around deterministic solar tracking, while field measurements remain necessary before the simulated increment can be treated as a site-level operational gain.

Keywords: predictive solar tracking; multi-horizon forecasting; effective sun position; energy yield; pvlib; machine learning; Niger Delta; photovoltaic systems

1. Introduction

Predictive solar tracking combines short-term forecasting with mechanical orientation control. Solar-forecasting literature consistently identifies lead time, baseline choice, atmospheric inputs and application-specific evaluation as important design considerations [1], [2]. Recurrent, ensemble and attention-based models have all been used for photovoltaic forecasting, although forecast accuracy matters most when it improves the operating objective of the energy system [3], [4].

The pvlib ecosystem provides established open-source implementations for solar-position calculation, tracker geometry, irradiance transposition and photovoltaic performance modelling [5], together with data-access tools for solar-resource analysis [6]. NASA POWER provides long-duration meteorological and solar-resource records for locations where dense ground measurements are unavailable. A validation study in Ghana found NASA POWER data useful for regional solar-resource assessment in tropical West Africa, while also showing why local measurements remain important for site-specific work [7].

The forecasting problem is treated as physics-guided residual learning. Astronomical solar motion is calculated first, and the learned stage estimates the atmosphere-dependent departure from that reference. This arrangement follows the broader idea of combining physical structure with data-driven learning [8], while the present Hybrid model remains a residual model rather than a physics-constrained neural network. Transformer methods have also been applied to short-term photovoltaic forecasting [9], [10], [13]. Solar-tracking studies show that tracking gain varies with latitude, tracker geometry, controller design and operating conditions [11], [12]. The XGBoost, LSTM and Transformer components follow the established model families described in [14], [15] and [16].

The study evaluates 1, 3, 6 and 24 h forecasts of an irradiance-optimal effective sun position and carries the forecasts into a 10 kWp photovoltaic simulation. Forecast outputs are mapped through the horizontal single-axis tracker geometry so that angular accuracy can be assessed in terms of energy yield. The analysis separates the larger gain produced by astronomical tracking from the smaller incremental gain associated with atmosphere-aware predictive correction.

2. Materials and Methods

2.1 Overall Methodological Framework

Astronomical solar position and atmosphere-dependent variables provide the information used to estimate the effective-position target. Forecasts are produced 1, 3, 6 and 24 h ahead, converted through the tracker geometry and evaluated in a pvlib-based photovoltaic simulation. Persistence provides the baseline. The sequence from data preparation to energy-yield assessment is summarized in Figure 1.

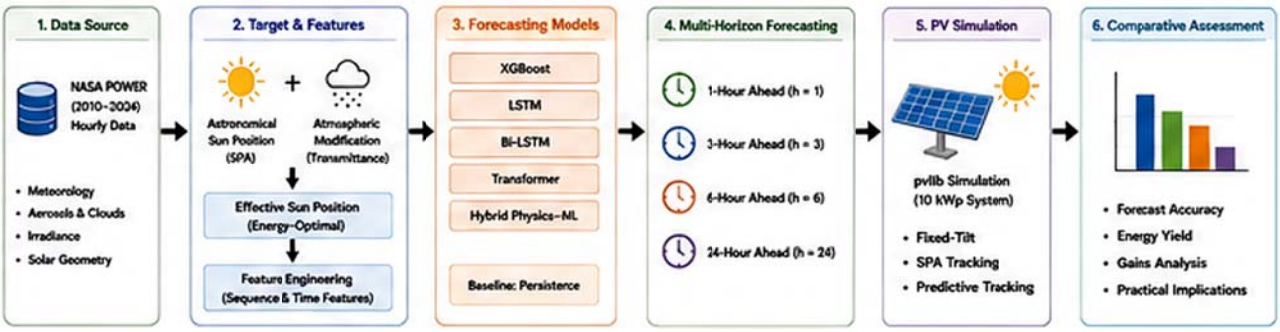

Figure 1. Overall workflow linking multi-horizon forecasting to PV energy-yield assessment.

2.2 System Architecture

The implementation is organized into data ingestion, preprocessing, forecasting, tracker control, photovoltaic simulation and results analysis. Effective-position forecasts are converted to the orientation required by the horizontal single-axis tracker before plane-of-array irradiance, DC power and AC energy are evaluated. The SPA trajectory provides the physical reference for both conventional tracking and the Hybrid residual model. The functional arrangement is presented in Figure 2.

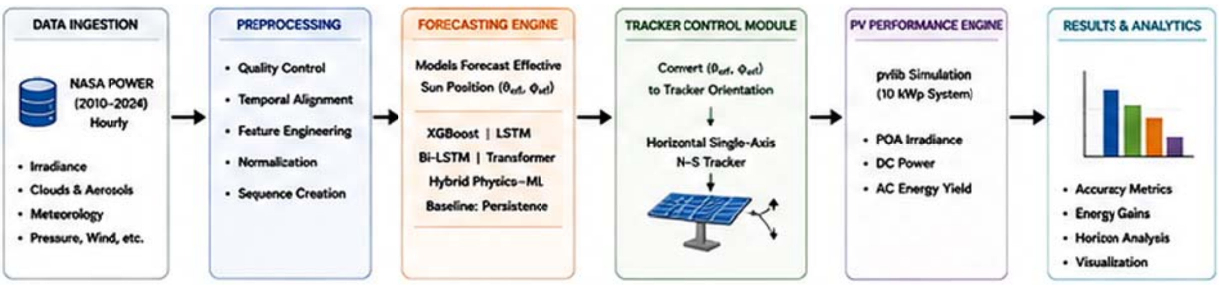

Figure 2. System architecture for multi-horizon predictive solar tracking and energy-yield assessment.

2.3 Architectural Diagrams of Forecasting Models

The forecasting comparison includes recurrent, bidirectional recurrent, attention-based and tree-based methods. LSTM and Bi-LSTM process historical sequences, the Transformer uses positional information with stacked self-attention, and XGBoost uses boosted regression trees [14], [15], [16]. The Hybrid Physics-ML model starts from SPA-derived solar geometry and uses XGBoost to estimate the residual correction associated with the irradiance-optimal target. The model families and their data paths are summarized in Figure 3.

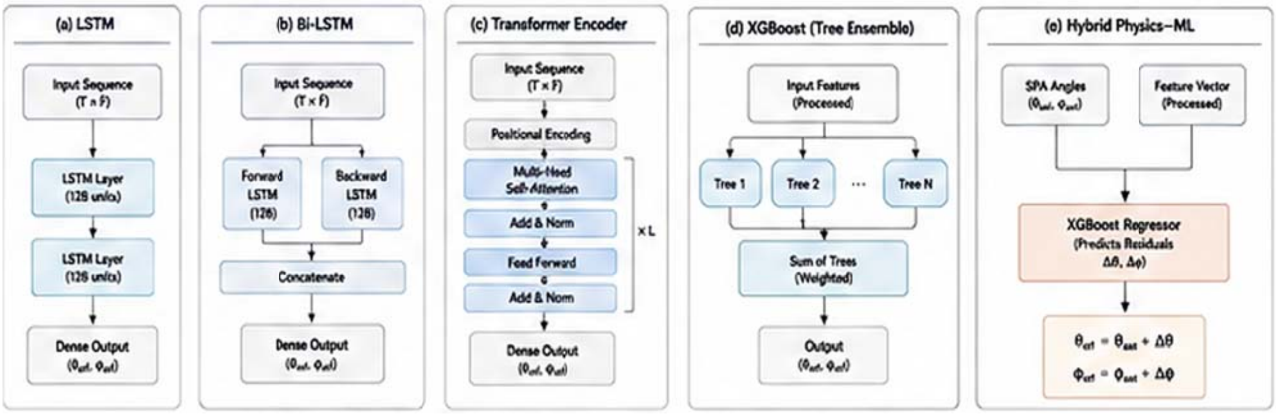

Figure 3. Architectural diagrams of the forecasting models evaluated in this study.

2.4 Study site and data source

The study site is the Faculty of Engineering, University of Uyo, Akwa Ibom State, Nigeria, at 5.041280°N and 7.974°E. Hourly NASA POWER records from 1 January 2010 to 31 December 2024 form the source dataset, giving 131,496 timestamps and 14 atmospheric or meteorological variables. NASA POWER combines satellite-derived solar information with model-assimilated meteorological products for renewable-energy applications [17]. Regional evidence from Ghana supports its use for tropical West African solar-resource studies, while local measurements remain preferable when site-specific validation is required [7].

Table 1. Atmospheric and meteorological variables used in the study

Parameter	Description	Unit	Mean	Min	Max
AOD_55	Aerosol optical depth at 550 nm	unitless	0.591	0.11	3.49
ALLSKY_SFC_SW_DWN	All-sky surface shortwave downward irradiance	Wh/m²	188.61	0.00	981.42
ALLSKY_SFC_SW_DNI	All-sky direct normal irradiance	Wh/m²	78.87	0.00	1134.53
ALLSKY_SFC_SW_DIFF	All-sky diffuse horizontal irradiance	Wh/m²	105.85	0.00	702.55
CLRSKY_SFC_SW_DWN	Clear-sky surface shortwave downward irradiance	Wh/m²	260.31	0.00	1002.45
CLRSKY_SFC_SW_DNI	Clear-sky direct normal irradiance	Wh/m²	204.80	0.00	880.85
T2M	2 m temperature	°C	26.08	14.72	34.57
RH2M	2 m relative humidity	%	87.02	29.02	100.00
PS	Surface pressure	kPa	100.65	99.97	101.35
PRECTOTCORR	Corrected precipitation	mm/h	0.316	0.00	48.20
WS2M	2 m wind speed	m/s	1.506	0.00	5.01
WD2M	2 m wind direction	°	227.94	0.00	359.80
CLOUD_AMT	Cloud amount	%	74.64	0.00	100.00
QV2M	2 m specific humidity	g/kg	18.26	7.30	22.55

The 15-year record is summarized in Table 1. Mean cloud amount is 74.64%, mean relative humidity is 87.02%, and mean aerosol optical depth at 550 nm is 0.591. These values characterize a humid, cloud-prone environment with marked changes in direct and diffuse irradiance between wet and dry periods. The stated coordinate, study period and NASA POWER variable codes identify the source series used in this work, while the same time convention, screening rules and preprocessing procedure are required for a like-for-like reconstruction.

2.5 Research procedure

The analysis followed five connected stages. Data acquisition and quality checks established a chronological hourly record. Cyclic time variables and astronomical quantities were derived only from information available at each timestamp. Solar position supplied the deterministic geometric reference, while the irradiance-based objective defined the effective-position target. Training and model selection used the training and validation intervals, and 2023 to 2024 was reserved for final evaluation. Angular forecasts were then converted through the tracker geometry and used in the photovoltaic simulation.

2.6 Astronomical reference and effective-position target

Solar zenith and azimuth are calculated for every timestamp with the pvlib solar-position implementation [5], using the Solar Position Algorithm as the astronomical reference [18]. Hourly temperature and pressure are supplied to the refraction calculation. Samples with solar zenith equal to or greater than 90° are excluded from target evaluation, although night-time observations may remain in historical sequence inputs. This screening leaves 57,148 daytime records for target-based analysis.

The effective position is represented as a virtual zenith-azimuth pair whose corresponding admissible single-axis orientation maximizes simulated plane-of-array irradiance under the available all-sky irradiance components. It does not introduce a second actuator degree of freedom. The predicted virtual direction is passed through the same horizontal north-south single-axis geometry used for the SPA reference, which produces the tracker rotation command. Astronomical position supplies the geometric reference, while direct and diffuse irradiance determine the energy-optimal correction within the admissible tracker range.

Chronological separation is maintained between predictors and targets. Irradiance at the target timestamp is used only when constructing the effective-position label; it is not supplied as future information to the forecasting models. For an h-hour forecast, the feature history ends at time t and the target corresponds to time t+h.

$$\Delta\theta(t) = \theta\mathrm{eff}(t) - \theta\mathrm{SPA}(t) \quad (1)$$

$$\Delta\varphi(t) = \mathrm{atan2}\{\sin[\varphi\mathrm{eff}(t) - \varphi\mathrm{SPA}(t)], \cos[\varphi\mathrm{eff}(t) - \varphi\mathrm{SPA}(t)]\} \quad (2)$$

Equations (1) and (2) define the residuals used by the Hybrid model. The zenith residual is the difference between the irradiance-optimal effective zenith and the SPA zenith. Azimuth is treated as a circular variable through the atan2 formulation in Equation (2), preventing a crossing between 0° and 360° from being interpreted as a large error. Trigonometric evaluation uses angular values in a consistent internal unit, and the resulting residual is reported in degrees.

2.7 Chronological partitioning, leakage control and metrics

The dataset is partitioned chronologically into 2010 to 2020 for training, 2021 to 2022 for validation and model selection, and 2023 to 2024 for final testing. Random shuffling is not used. Scaling, feature transformations and model-selection decisions are fitted on the training interval and assessed on the validation interval before the test period is examined. Sequence models receive only observations available at the forecast origin. Bidirectional processing is confined to the historical input window and does not use samples after the prediction origin. This evaluation structure follows the causal requirements emphasized in short-term solar forecasting [1], [4].

$$\mathrm{MAE} = (1/n)\ \Sigma_{i=1}^{n} |y_i - \hat{y}_i| \quad (3)$$

$$\mathrm{RMSE} = \sqrt{[(1/n)\ \Sigma_{i=1}^{n} (y_i - \hat{y}_i)^2]} \quad (4)$$

$$R^2 = 1 - [\Sigma_{i=1}^{n} (y_i - \hat{y}_i)^2 / \Sigma_{i=1}^{n} (y_i - \bar{y})^2] \quad (5)$$

MAE and RMSE are reported in degrees, and R^2 is used as a descriptive measure of fit. Percentage angular error is not emphasized because the denominator can approach values that make percentage measures unstable or physically uninformative. Engineering interpretation is therefore based on angular error, forecast skill relative to persistence and the simulated energy yield associated with each tracking strategy.

2.8 Forecasting models and horizons

Table 2. Forecasting models and evaluation horizons

Model	Forecast formulation	Horizons
Persistence	Current effective position	1, 3, 6, 24 h
XGBoost	Boosted-tree direct forecast	1, 3, 6, 24 h
LSTM	24 h historical sequence	1, 3, 6, 24 h

Model	Forecast formulation	Horizons
Bi-LSTM	24 h historical sequence	1, 3, 6, 24 h
Transformer	48 h historical sequence	1, 3, 6, 24 h
Hybrid Physics-ML	SPA reference plus learned residual	1, 3, 6, 24 h

All models in Table 2 are evaluated at the same four lead times and on the same chronologically isolated test period. LSTM and Bi-LSTM use 24 h historical windows, while the Transformer uses a 48 h historical window. The architecture recorded in Figure 3 uses two 128-unit LSTM layers for the LSTM model and 128-unit forward and backward recurrent branches for the Bi-LSTM model before dense output. Bidirectional processing remains inside the historical input window, preserving prospective forecasting causality [4], [9], [10]. Model-selection decisions are made on the validation interval rather than on the final test period.

2.9 Predictive tracking and PV energy-yield simulation

The application layer represents a 10 kWp monocrystalline photovoltaic system composed of 25 modules rated at 400 Wp. A constant inverter efficiency of 97.5%, a 2% soiling loss and a 1% wiring loss are applied to every configuration. Fixed tilt is 5° facing south. Conventional tracking follows horizontal single-axis geometry driven by SPA solar position, while predictive tracking passes the forecast effective position through the same geometry. These common assumptions keep the comparison focused on orientation strategy. Solar geometry and photovoltaic calculations are implemented with pvlib [5], [6], [18]. Results from 2023 and 2024 are reported as mean annual energy, obtained by averaging the annual outputs of the two test years. The resulting yields are comparative simulation outputs rather than field-certified production estimates.

$$\text{RMSE\_comb} = \sqrt{(\text{RMSE\_}\theta^2 + \text{RMSE\_}\varphi^2)} \tag{6}$$

$$\text{SS} = 1 - (\text{RMSE\_model} / \text{RMSE\_base}) \tag{7}$$

$$\text{CF} = \text{E\_AC} / (\text{P\_rated} \times 8760) \tag{8}$$

$$\text{G\_yield} = 100 \times (\text{E\_track} - \text{E\_fixed}) / \text{E\_fixed} \tag{9}$$

Equation (6) combines zenith and azimuth RMSE into a single angular index. Equation (7) defines forecast skill relative to persistence, Equation (8) defines annual capacity factor, and Equation (9) gives energy-yield gain relative to fixed tilt. Angular accuracy and simulated energy production are kept as separate measures because a small change in angular error does not necessarily translate into a proportionate change in annual energy.

3. Results

3.1 Forecast accuracy across horizons

Combined RMSE rises steadily with forecast lead time. Hybrid Physics-ML gives the lowest error in Table 3 at every horizon. XGBoost records 1.921° at 1 h, slightly above the 1.842° persistence error, but improves on persistence from 3 h onward. Sequence-based and physics-guided models therefore give the lowest short-horizon errors in this dataset.

Forecast error rises appreciably as the prediction horizon lengthens (Figure 4). Hybrid Physics-ML and the Transformer remain the two most accurate models across all four horizons, and both retain lower 24 h error than persistence. The energy analysis nevertheless indicates that the residual angular advantage at 24 h adds very little annual energy.

Hybrid Physics-ML retains positive skill over persistence at every horizon, reaching 40.9%, 49.2%, 54.3% and 51.4% at 1, 3, 6 and 24 h, respectively (Table 4). XGBoost is slightly worse than persistence at 1 h but becomes progressively better at the longer horizons.

The skill profiles remain positive for Hybrid Physics-ML and the Transformer across all four lead times (Figure 5). Hybrid Physics-ML records the highest skill throughout. XGBoost is the only learned model with negative skill at 1 h, which confirms the value of retaining a simple persistence benchmark in short-horizon forecasting.

Table 3. Combined effective-position RMSE across forecast horizons

Model	1 h (°)	3 h (°)	6 h (°)	24 h (°)
Persistence	1.842	3.241	5.124	8.412

Model	1 h (°)	3 h (°)	6 h (°)	24 h (°)
XGBoost	1.921	2.812	3.914	6.124
LSTM	1.521	2.241	3.187	5.412
Bi-LSTM	1.347	2.014	2.887	5.014
Transformer	1.187	1.784	2.541	4.412
Hybrid Physics-ML	1.089	1.647	2.341	4.087

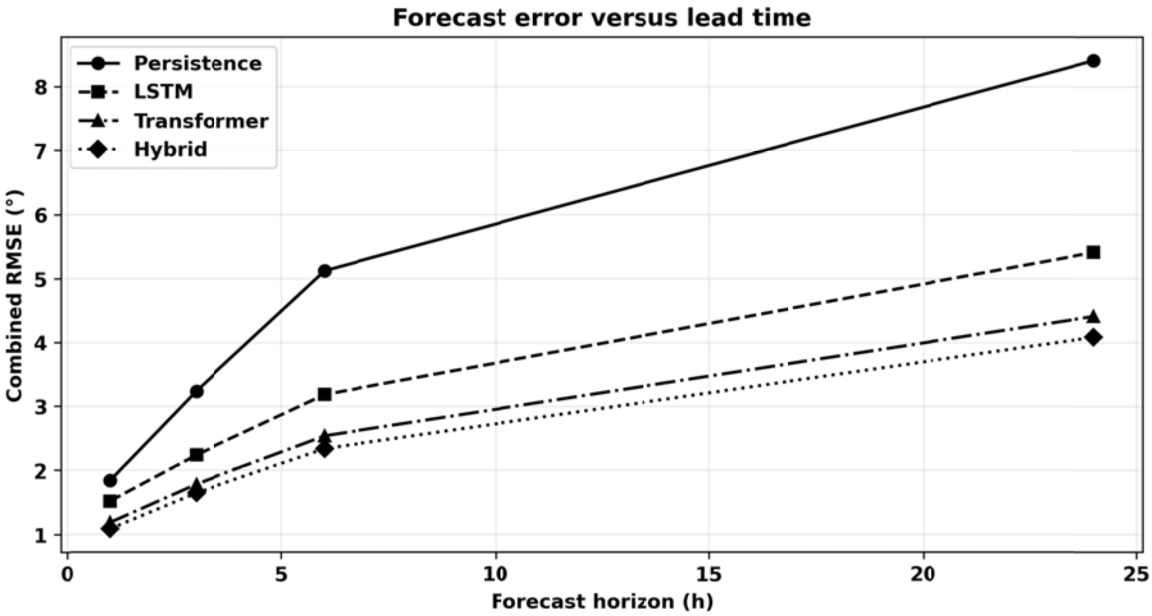

Figure 4. Combined effective-position RMSE as forecast lead time increases.

.

Table 4. Forecast skill relative to persistence across forecast horizons

Model	1 h (%)	3 h (%)	6 h (%)	24 h (%)
XGBoost	-4.3	13.2	23.6	27.2
LSTM	17.4	30.9	37.8	35.7
Bi-LSTM	26.9	37.9	43.7	40.4
Transformer	35.6	45.0	50.4	47.6
Hybrid Physics-ML	40.9	49.2	54.3	51.4

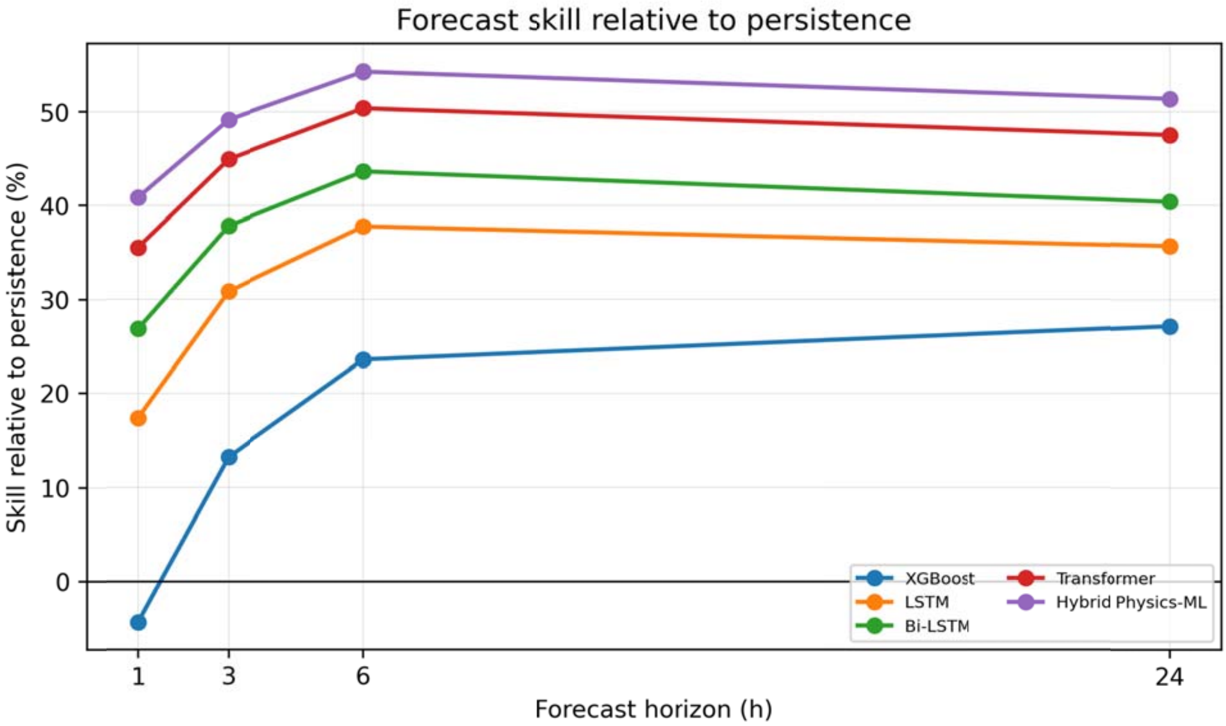

Figure 5. Forecast skill of the evaluated models relative to persistence.

3.2 Simulated energy yield

Most of the simulated energy improvement comes from geometric tracking rather than predictive correction. SPA tracking raises mean annual AC energy by 3,334 kWh/year above fixed tilt, while the 1 h predictive case adds another 620 kWh/year above SPA tracking. Relative to fixed tilt, the corresponding gains are 23.3% and 27.7% (Table 5).

The additional predictive benefit falls rapidly as lead time increases (Figure 6). At 24 h, predictive tracking produces 17,641 kWh/year, only 20 kWh/year above SPA tracking. Under the stated simulation assumptions, short-horizon correction therefore offers a clearer engineering benefit than day-ahead predictive orientation.

The predictive contribution decreases from 620 kWh/year above SPA at 1 h to 433 kWh/year at 3 h, 200 kWh/year at 6 h and 20 kWh/year at 24 h. Only 3.2% of the 1 h incremental benefit remains at 24 h (Table 6).

Beyond the 3 h horizon, increasing angular error removes most of the orientation advantage over SPA. The additional simulated AC energy consequently falls from 620 kWh/year at 1 h to 20 kWh/year at 24 h (Figure 7), even though the learned models still retain some forecast skill over persistence.

Table 5. Simulated annual energy yield for fixed and tracking configurations

Configuration	POA (kWh/m²/yr)	AC energy (kWh/yr)	Gain vs fixed (%)	CF (%)
Fixed tilt, 5° south-facing	1,652	14,287	0.0	16.3
SPA-based HSAT	2,041	17,621	23.3	20.1
Predictive ML, 1 h	2,118	18,241	27.7	20.8
Predictive ML, 3 h	2,097	18,054	26.4	20.6
Predictive ML, 6 h	2,071	17,821	24.7	20.3
Predictive ML, 24 h	2,044	17,641	23.4	20.1
Ideal dual-axis reference	2,287	19,841	38.9	22.6

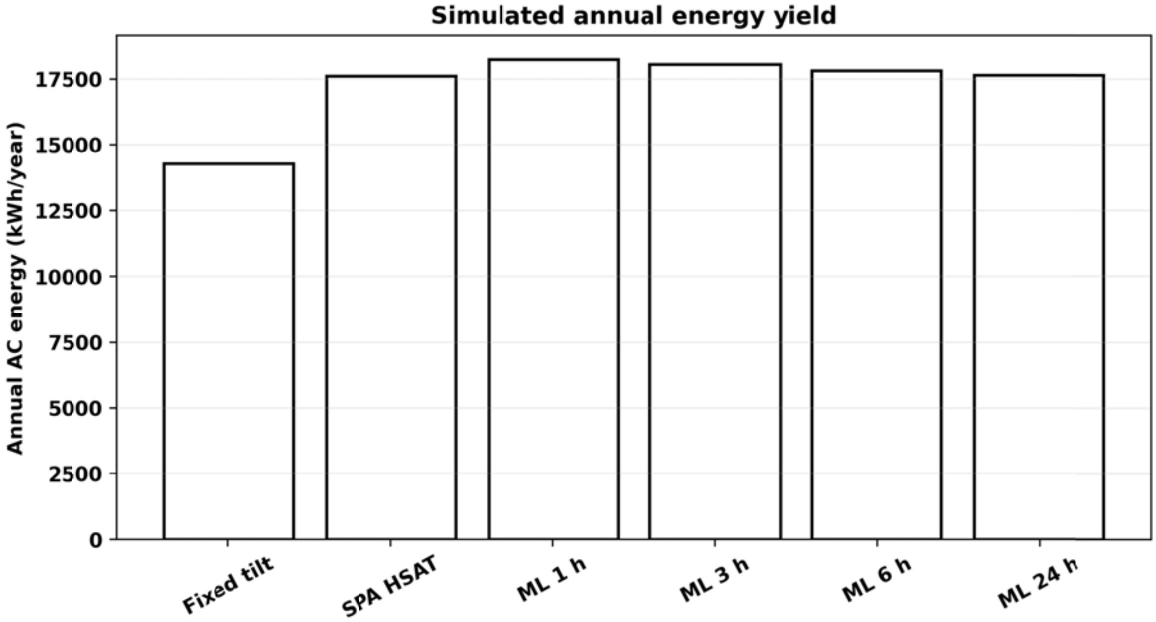

Figure 6. Simulated annual AC energy for fixed, astronomical and predictive tracking.

Table 6. Incremental predictive energy benefit relative to SPA-based tracking

Forecast horizon	Increment over SPA (kWh/yr)	Increase over SPA (%)	1 h benefit retained (%)
1 h	620	3.52	100.0
3 h	433	2.46	69.8
6 h	200	1.14	32.3
24 h	20	0.11	3.2

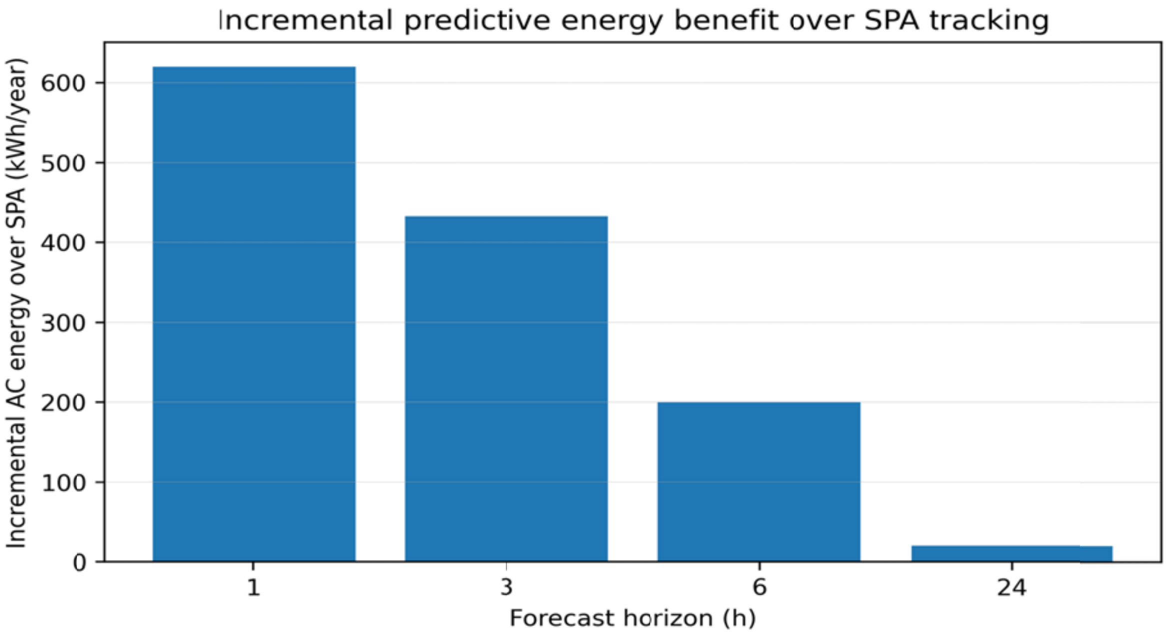

Figure 7. Incremental annual AC energy provided by predictive tracking above the SPA-based tracker.

3.3 Seasonal performance and comparative model behaviour

Seasonal diagnostic values for the 2023 to 2024 test period are summarized in Table 7. Errors increase during the wetter parts of the year, while the Hybrid model retains the lowest reported zenith MAE in every seasonal group. Peak wet-season error is highest for all three selected models, consistent with stronger irradiance variability during heavy cloud cover at Uyo. These values are descriptive test-period diagnostics and are not used for the horizon-specific model ranking reported in Table 3.

Relative reductions preserve the same seasonal ordering (Table 8). Hybrid zenith MAE is 11.0% to 20.9% lower than the Transformer result and 32.6% to 42.3% lower than LSTM across the reported seasonal groups. The advantage remains during the peak wet season, when mean cloud amount reaches 88.3%. Zenith MAE rises from the dry season into wetter periods for all three models, while the Hybrid curve remains below the Transformer and LSTM curves throughout the seasonal partition (Figure 8).

Table 7. Seasonal descriptive diagnostic error of selected models over the 2023 to 2024 test period

Period	**Cloud (%)**	**Hybrid zenith MAE (°)**	**Hybrid azimuth MAE (°)**	**Transformer zenith MAE (°)**	**LSTM zenith MAE (°)**
Dry season, Nov to Feb	48.2	0.412	0.498	0.521	0.714
Early wet season, Mar to May	81.6	0.687	0.812	0.841	1.124
Peak wet season, Jun to Sep	88.3	0.814	0.974	0.987	1.312
Transition, Oct	84.4	0.712	0.851	0.874	1.184
Full test period	73.7	0.634	0.751	0.712	0.941

Table 8. Relative reduction in seasonal descriptive Hybrid zenith MAE over the 2023 to 2024 test period

Period	**Cloud (%)**	**Hybrid MAE (°)**	**Reduction vs Transformer (%)**	**Reduction vs LSTM (%)**
Dry season, Nov to Feb	48.2	0.412	20.9	42.3
Early wet season, Mar to May	81.6	0.687	18.3	38.9
Peak wet season, Jun to Sep	88.3	0.814	17.5	38.0
Transition, Oct	84.4	0.712	18.5	39.9
Full test period	73.7	0.634	11.0	32.6

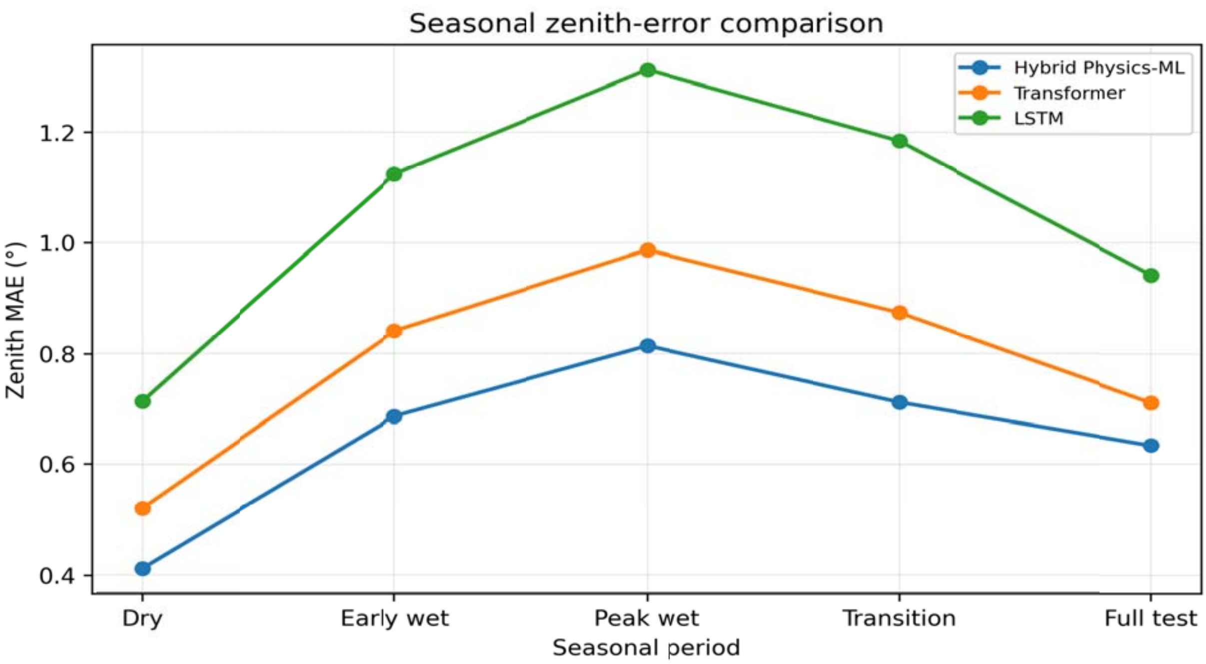

Figure 8. Seasonal descriptive zenith-MAE comparison for the Hybrid, Transformer and LSTM models over the final test period.

Table 9. Implementation structure of the principal forecasting models

Model	**Forecast formulation**	**Historical input**	**Predicted quantity**

Model	Forecast formulation	Historical input	Predicted quantity
XGBoost	Boosted-tree direct forecast	Forecast-origin feature vector	Effective-position target
LSTM	Recurrent sequence forecast	24 h sequence	Effective-position target
Bi-LSTM	Bidirectional recurrent sequence forecast	24 h historical sequence	Effective-position target
Transformer	Attention-based sequence forecast	48 h historical sequence	Effective-position target
Hybrid Physics-ML	SPA reference plus learned residual	Forecast-origin features and SPA reference	Zenith and circular-azimuth residuals

The principal modelling choices are consolidated in Table 9 so that forecast formulation, historical input length and predicted quantity can be compared directly across the five learned models. Hardware-dependent execution time is deliberately excluded because it is not a portable property of a model.

At 1 h, Hybrid Physics-ML records the lowest RMSE, 1.089°, and the highest skill over persistence, 40.9%. The Transformer follows at 1.187° RMSE and 35.6% skill, while XGBoost is slightly worse than persistence (Table 10). The 1 h RMSE comparison is presented in Figure 9. Accuracy is emphasized because execution time depends on hardware, software and measurement procedure and cannot be interpreted as a portable model property.

Table 10. One-hour accuracy of the principal forecasting models

Model	1 h RMSE (°)	Skill vs persistence (%)	1 h ranking	Baseline comparison
XGBoost	1.921	-4.3	5	Below persistence
LSTM	1.521	17.4	4	Above persistence
Bi-LSTM	1.347	26.9	3	Above persistence
Transformer	1.187	35.6	2	Above persistence
Hybrid Physics-ML	1.089	40.9	1	Above persistence

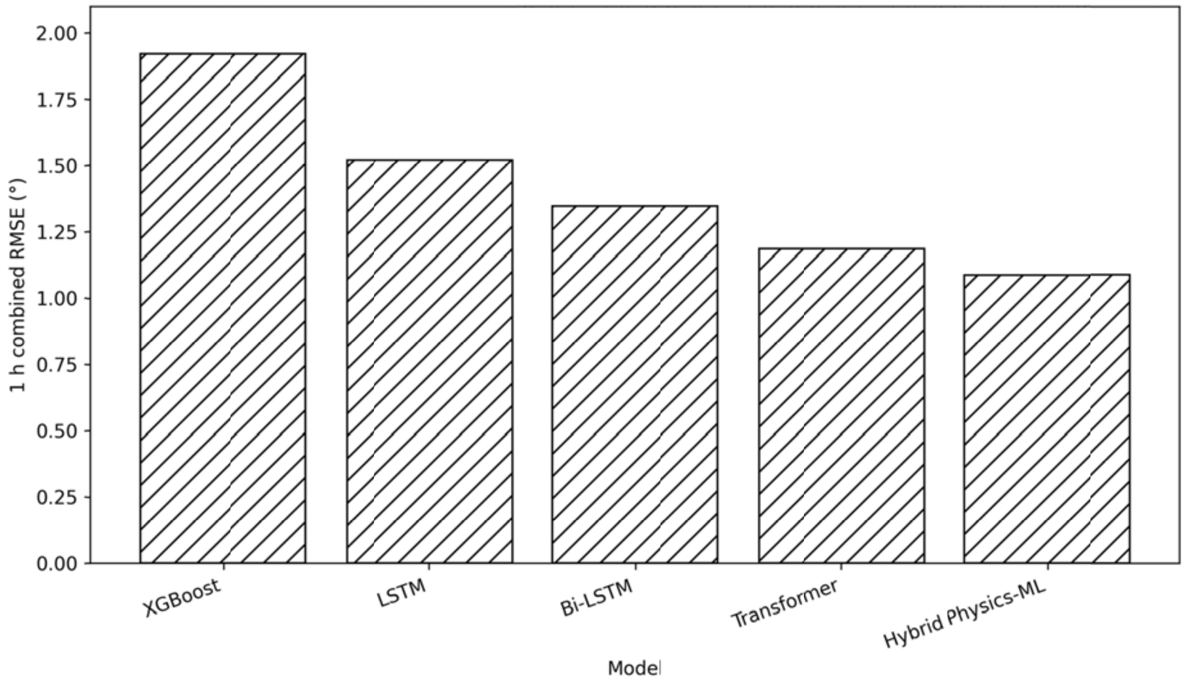

Figure 9. One-hour RMSE comparison for the principal forecasting models.

4. Discussion

Two effects are evident in the energy results. Geometric tracking provides the larger gain, increasing mean annual simulated output from 14,287 to 17,621 kWh/year. Atmosphere-aware predictive correction adds a smaller increment, with the 1 h case reaching 18,241 kWh/year. Published solar-tracking studies report the same broad pattern that tracking can provide substantial gains while the exact magnitude remains dependent on site conditions and system design [11], [12].

Forecast error increases with lead time, consistent with short-term solar-forecasting studies [1], [4]. The Hybrid formulation begins with deterministic solar geometry and learns only the residual associated with the irradiance-optimal target, reducing the trajectory component that must be represented by the data-driven stage. The Transformer remains competitive on this dataset [9], [10], [13], but its lower error relative to the other purely data-driven models does not yield more simulated energy than the Hybrid model. A forecast has engineering value only when its angular improvement changes the tracker command enough to increase incident irradiance.

The energy values are simulation outputs based on the NASA POWER record and the stated pvlib assumptions. Auxiliary tracker energy, mechanical backlash, stow events, row-to-row shading, terrain effects and detailed inverter clipping are outside the model. The 620 kWh/year increment is therefore a simulated control benefit rather than a measured field gain. Hourly NASA POWER data also cannot capture every sub-hourly cloud transition encountered by a physical tracker. Field validation should record local irradiance, realized tracker angle, actuator energy and AC output before the predictive increment is generalized to an operating installation.

5. Practical Applications

A grid-connected campus or commercial array in Uyo could apply the proposed control structure with horizontal single-axis trackers. The controller would compute the SPA orientation locally and apply the 1 h Hybrid forecast as a bounded correction. Mechanical angle limits, a minimum expected-benefit threshold and a data-quality check can be evaluated before a command is issued. When the forecast is unavailable or fails the quality check, the tracker can remain on the SPA trajectory.

The same control principle could be used in solar mini-grids where actuator energy and maintenance affect operating decisions. Short-horizon corrections can be issued only when the predicted irradiance benefit exceeds a deadband, which would reduce unnecessary movement during uniformly clear or uniformly overcast periods. A field test should compare SPA-only and predictive-control operation while recording plane-of-array irradiance, AC energy, actuator energy, commanded angle and realized angle.

6. Conclusion

Multi-horizon effective-position forecasting was evaluated for predictive solar tracking in Uyo using a 15-year NASA POWER record. Hybrid Physics-ML produced the lowest combined RMSE at each forecast horizon, with values of 1.089°, 1.647°, 2.341° and 4.087° at 1, 3, 6 and 24 h. Mean annual simulated output was 14,287 kWh/year for fixed tilt, 17,621 kWh/year for SPA-based tracking and 18,241 kWh/year for 1 h predictive tracking. The additional simulated gain over SPA tracking was 620 kWh/year at 1 h and declined sharply with lead time. Under the stated assumptions, short-horizon atmosphere-aware correction offers the strongest predictive benefit. Field measurements are still required to determine the net gain after mechanical, electrical and site-specific losses are included.

References

[1] D. Yang, W. Wang, C. A. Gueymard, T. Hong, J. Kleissl, J. Huang, M. J. Perez, R. Perez, J. M. Bright, X. Xia, D. van der Meer, and I. M. Peters, "A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality," Renewable and Sustainable Energy Reviews, vol. 161, art. 112348, 2022, doi: 10.1016/j.rser.2022.112348.

[2] D. Yang, W. Wang, and X. Xia, "A Concise Overview on Solar Resource Assessment and Forecasting," Advances in Atmospheric Sciences, vol. 39, no. 8, pp. 1239-1251, 2022, doi: 10.1007/s00376-021-1372-8.

[3] K. J. Iheanetu, "Solar Photovoltaic Power Forecasting: A Review," Sustainability, vol. 14, no. 24, art. 17005, 2022, doi: 10.3390/su142417005.

[4] F. Pandžić and T. Capuder, "Advances in Short-Term Solar Forecasting: A Review and Benchmark of Machine Learning Methods and Relevant Data Sources," Energies, vol. 17, no. 1, art. 97, 2024, doi: 10.3390/en17010097.

[5] K. S. Anderson, C. W. Hansen, W. F. Holmgren, A. R. Jensen, M. A. Mikofski, and A. Driesse, "pvlib python: 2023 project update," Journal of Open Source Software, vol. 8, no. 92, art. 5994, 2023, doi: 10.21105/joss.05994.

[6] A. R. Jensen, K. S. Anderson, W. F. Holmgren, M. A. Mikofski, C. W. Hansen, L. J. Boeman, and R. Loonen, "pvlib iotools: Open-source Python functions for seamless access to solar irradiance data," Solar Energy, vol. 266, art. 112092, 2023, doi: 10.1016/j.solener.2023.112092.

[7] A. D. Quansah, F. Dogbey, P. J. Asilevi, P. Boakye, L. Darkwah, S. Oduro-Kwarteng, Y. A. Sokama-Neuyam, and P. Mensah, "Assessment of solar radiation resource from the NASA-POWER reanalysis products for tropical climates in Ghana towards clean energy application," Scientific Reports, vol. 12, art. 10684, 2022, doi: 10.1038/s41598-022-14126-9.

[8] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nature Reviews Physics, vol. 3, no. 6, pp. 422-440, 2021, doi: 10.1038/s42254-021-00314-5.

[9] J. Wang, W. Hu, L. Xuan, F. He, C. Zhong, and G. Guo, "TransPVP: A Transformer-Based Method for Ultra-Short-Term Photovoltaic Power Forecasting," Energies, vol. 17, no. 17, art. 4426, 2024, doi: 10.3390/en17174426.

[10] G. Piantadosi, S. Dutto, A. Galli, S. De Vito, C. Sansone, and G. Di Francia, "Photovoltaic power forecasting: A Transformer based framework," Energy and AI, vol. 18, art. 100444, 2024, doi: 10.1016/j.egyai.2024.100444.

[11] K. Kumba, P. Upender, P. Buduma, M. Sarkar, S. P. Simon, and V. Gundu, "Solar tracking systems: Advancements, challenges, and future directions: A review," Energy Reports, vol. 12, pp. 3566-3583, 2024, doi: 10.1016/j.egyr.2024.09.038.

[12] M. A. Ponce-Jara, I. Pazmiño, Á. Moreira-Espinoza, A. Gunsha-Morales, and C. Rus-Casas, "Assessment of Single-Axis Solar Tracking System Efficiency in Equatorial Regions: A Case Study of Manta, Ecuador," Energies, vol. 17, no. 16, art. 3946, 2024, doi: 10.3390/en17163946.

[13] M. F. Hanif and J. Mi, "Harnessing AI for solar energy: Emergence of transformer models," Applied Energy, vol. 369, art. 123541, 2024, doi: 10.1016/j.apenergy.2024.123541.

[14] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, pp. 785-794, 2016, doi: 10.1145/2939672.2939785.

[15] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention Is All You Need," in Advances in Neural Information Processing Systems 30, pp. 5998-6008, 2017.

[17] T. Zhang, W. S. Chandler, J. M. Hoell, D. Westberg, C. H. Whitlock, and P. W. Stackhouse Jr., "A Global Perspective on Renewable Energy Resources: NASA's Prediction of Worldwide Energy Resources (POWER) Project," in Proceedings of ISES World Congress 2007, Vols. I-V, D. Y. Goswami and Y. Zhao, Eds. Berlin, Germany: Springer, pp. 2636-2640, 2008, doi: 10.1007/978-3-540-75997-3_532.

[18] I. Reda and A. Andreas, "Solar Position Algorithm for Solar Radiation Applications," Solar Energy, vol. 76, no. 5, pp. 577-589, 2004, doi: 10.1016/j.solener.2003.12.003.