Journal of Climate Change
ISSN 2395-7611 Volume 12 (2026) No. 1. pp. 148–167

https://doi.org/10.70917/jcc-2026-009
Article

AI-Augmented Clear-Sky Algorithms for High-Granularity Solar Irradiance Modelling

Kudzanayi Chiteka[1] and C.C. Enweremadu [1,*]

[1] Department of Mechanical, Bioresources and Biomedical Engineering, University of South Africa, Science Campus, Florida 1710, South Africa
* Correspondence author: enwercc@unisa.ac.za

Abstract: Accurate short-term solar irradiance forecasting is important for reliable photovoltaic operation in regions affected by rapid cloud movement and atmospheric variability. However, clear-sky models cannot adequately respond to cloud-driven irradiance fluctuations, while purely data-driven models may suffer from weak physical consistency and reduced generalisability. The present study developed a physics-informed hybrid framework in which clear-sky irradiance was used as the physical baseline, while machine learning was employed to learn the residual atmospheric deviations caused by clouds, aerosols, moisture and near-surface meteorology. The proposed model was evaluated against persistence, clear-sky, AI-only and standalone Gated Recurrent Unit benchmark models using 10-min-ahead global horizontal irradiance prediction. During the Roodepoort summer test period, the hybrid model achieved an RMSE of 79 W/m^2, MAE of 59 W/m^2, MBE of −8 W/m^2 and R^2 of 0.91, corresponding to RMSE reductions of 32%, 44%, 27% and 18% when compared with the persistence, clear-sky, AI-only and Gated Recurrent Unit models, respectively. Furthermore, the model maintained seasonal RMSE values between 79 and 96 W/m^2 without retraining, retained useful transferability at external Southern African Universities Radiometric Network stations with a mean external RMSE of 95.5 W/m^2 and R^2 of 0.86, and produced calibrated 95% prediction intervals with 93% empirical coverage. These findings showed that clear-sky constrained residual learning can improve deterministic accuracy, uncertainty reliability and seasonal robustness, thus supporting more reliable photovoltaic forecasting, grid integration and operational energy management under variable sky conditions.

Keywords: Hybrid solar forecasting; Physics-informed machine learning; Clear-sky modelling; Solar irradiance prediction; Atmospheric features; Performance forecasting.

1. Introduction

The increasing global energy demand, coupled with the need to reduce greenhouse gas emissions and dependence on fossil fuels, has intensified the transition towards renewable and sustainable energy systems. Studies have indicated that this transition is being advanced through different low carbon pathways, including green hydrogen for energy sector decarbonisation (Bakhsh and Cantino, 2025), sand based concentrated solar power systems for reducing fuel consumption and carbon dioxide emissions (Al-Ansary et al., 2025), solar thermal technologies and systems (Mandal and Reddy, 2025), and emerging photovoltaic technologies such as perovskite-silicon tandem solar cells (Riffat and Reza Samaei, 2025). Furthermore, artificial intelligence has become increasingly important in renewable energy systems, especially for forecasting, optimisation, energy storage management and smart grid control, with specific applications in solar energy, wind energy and integrated energy systems (Razak et al., 2025). Therefore, the transition towards renewable energy does not only depend on generation technologies, but also on intelligent modelling approaches that can improve prediction, decision making and grid reliability.

The rapid adoption of solar photovoltaic (PV) systems has increased the need for accurate short-term solar irradiance prediction to ensure grid stability and reliable energy supply (Odejobi et al., 2024). Solar irradiance variability in small time intervals is governed by meteorological states, which introduce rapid power fluctuations in solar generation affecting PV integration and grid operation (Sasaki et al., 2024). In regions characterised by rapidly changing weather, prediction errors can easily lead to operational challenges and grid instability (Chandel and Roy, 2023).

Solar irradiance prediction methods are commonly classified as physical, statistical, machine learning and hybrid approaches (Peiró-Signes et al., 2025). Physical models, including clear-sky radiative transfer models and numerical weather prediction methods, explicitly represent atmospheric processes governing solar radiation attenuation (Sultana and Tsutsumida, 2024). Clear-sky models such as the Reference Evaluation of Solar Transmittance, 2-band model (REST2) and Ineichen Perez provide useful irradiance estimates under cloud-free conditions, however, their inability to respond to cloud-induced variability and short-term atmospheric disturbances limits their application in high-frequency photovoltaic forecasting (Gueymard and Ruiz-Arias, 2015; Yang, 2020). Although approaches based on numerical weather prediction explicitly resolve cloud microphysics and radiative transfer processes, their high computational cost and typically coarse temporal resolution often limit their suitability for high-frequency, operational photovoltaic forecasting (Kreuwel et al., 2023; Mathiesen et al., 2013).

Statistical and time-series models are known to be computationally efficient and easy to implement. However, their underlying assumptions of linearity and weak stationarity restrict their ability to represent the nonlinear and nonstationary solar irradiance dynamics introduced by rapidly varying cloud fields and aerosol loading (Langella et al., 2016; Thaker and Höller, 2022). As a result, machine-learning techniques are now increasingly acceptable in solar irradiance prediction, especially when they are deliberately coupled with physical inputs to accurately capture nonlinear atmospheric phenomena associated with clouds and aerosols (Chandel and Roy, 2023; Sultana and Tsutsumida, 2024). These ML methods can effectively learn complex nonlinear relationships from large multivariate datasets, thereby outperforming traditional statistical approaches when meteorological and atmospheric predictors, including cloud cover, aerosols and humidity, are incorporated into the predictive model (Pereira et al., 2024; Thaker and Höller, 2022). However, their application in operational solar forecasting may be limited when the models are trained without explicit physical constraints, since this can reduce physical consistency under seasonal and atmospheric regimes that are poorly represented in the training data (Bai et al., 2025; Papachristopoulou et al., 2024).

Hybrid and physics-informed modelling approaches are important because they combine the interpretability of physical formulations with the adaptive capability of data-driven algorithms. Recent reviews have shown that short-term solar irradiance forecasting has increasingly shifted towards neural network and hybrid deep learning models, including long short-term memory (LSTM), gated recurrent unit (GRU), convolutional neural network (CNN), generative adversarial network (GAN), attention-based models and other combined architectures, although their performance still depends on the forecasting horizon, weather classification and evaluation metrics (Assaf et al., 2023; Paletta et al., 2023). Previous studies have also shown that irradiance prediction improves when radiative transfer knowledge, clear-sky constraints, cloud-based numerical weather prediction outputs and machine learning models are combined, rather than relying on standalone physical or purely data-driven approaches (Chauvin et al., 2018; Fernández-Peruchena et al., 2017; Mathiesen et al., 2013; Yang, 2020). Similarly, hybrid deep learning and multi-site studies have improved short-term solar irradiance and photovoltaic power forecasting through feature extraction networks, recurrent neural networks, attention mechanisms, transformer-based structures, multi-location learning and all sky imager based intra hour forecasting (Brahma and Wadhvani, 2020; Richardson et al., 2019; Salman et al., 2024; Siddiqa et al., 2025; Yan et al., 2020). These studies show that hybrid solar forecasting is already established, although the use of clear-sky physics as the central residual baseline for high-granularity operational forecasting still requires further development, consistent with residual learning logic in sustainable energy forecasting (Razak et al., 2026).

Despite these advances, several limitations remain important for high-granularity photovoltaic forecasting. Many existing approaches still focus on clear-sky estimation, irradiance downscaling, or short-term forecasting at temporal scales that may not fully capture rapid cloud induced variability required for operational photovoltaic control (Chauvin et al., 2018; Fernández-Peruchena et al., 2017; Kreuwel et al., 2023). Furthermore, broader transferability, uncertainty calibration and operational deployment remain active challenges, especially where dedicated imaging systems or dense ground-based networks are not available (Brahma and Wadhvani, 2020; Paletta et al., 2023; Richardson et al., 2019). Physical information is also often incorporated as an auxiliary input, post processing reference, or model feature, rather than as a constraining residual learning baseline around which the forecasting problem is formulated (Sultana and Tsutsumida, 2024; Yang, 2020). Therefore, a gap remains for a high-granularity forecasting framework that integrates clear-sky residual learning, multi-source atmospheric predictors, regime and season stratified validation, ensemble-based uncertainty calibration and interpretable model evaluation within one operational workflow.

The present study therefore developed a physics-informed hybrid solar irradiance forecasting framework for 10-min ahead prediction at Roodepoort, South Africa. The framework used a Linke turbidity clear-sky model as the physical reference within a residual learning formulation. The originality of the study is not based on presenting hybrid solar forecasting as an entirely new concept, but on integrating clear-sky constrained residual learning, freely available satellite and reanalysis atmospheric predictors, uncertainty calibration, interpretability and external station validation within one physically constrained workflow for high-granularity irradiance forecasting. This paper is organised as follows: Section 2 presents the methodological approach and data used in the study, Section 3 presents and discusses the results, while Section 4 concludes the paper and provides the key findings and future research directions.

2. Methodology

2.1. Data and Clear-Sky Physical Baseline

Data used in this study were obtained from a combination of ground-based irradiance measurements and available satellite data. Ground-measured global horizontal irradiance (GHI) was measured at Roodepoort, South Africa (26.1473° S, 27.9010° E; 1720 m a.s.l.) using a Davis Vantage Pro2 logged at 1 minute resolution and was subsequently aggregated to a 10-minute modelling interval adopted as the common time step for training and evaluation. The analysis period spanned from 1 January 2021 to 31 December 2023. After timestamp alignment, aggregation and quality control screening, the effective missing fraction was 5.3%. Missing GHI target values were not reconstructed. During aggregation to the 10-min modelling interval, only intervals with sufficient valid 1 min observations were retained, while duplicate timestamps, incomplete aggregation windows and physically unrealistic measurements were excluded. The physical screening rule removed observations satisfying GHI < 0 Wm^{-2} or GHI > $1.2G_{on}cos\theta_z$, where G_{on} and θ_z are the extraterrestrial irradiance normal to the sun's rays and the solar zenith angle, respectively. Sensor uncertainty was linked to calibration status, instrumental accuracy, installation condition, levelling, cleaning and maintenance history of the Davis Vantage Pro2. Since no post deployment recalibration record or reference pyranometer comparison was available for Roodepoort, possible sensor drift, dome soiling, levelling error and maintenance uncertainty were treated as observational limitations. Therefore, the reported forecasting errors should be interpreted within the uncertainty limits of the ground sensor.

Meteorological, cloud, aerosol and atmospheric composition variables were obtained from ERA5 (https://cds.climate.copernicus.eu/datasets), MODIS (https://modis.gsfc.nasa.gov/data/) and CAMS (https://atmosphere.copernicus.eu/data) products and harmonised with the 10-min GHI modelling interval. These predictors were used as slowly varying atmospheric descriptors, while high-frequency irradiance variability was represented through clear-sky solar-geometry variables and lagged ground-measured GHI features. Temporal consistency was verified by checking timestamp continuity, duplicate records, uniform 10-min spacing after aggregation, and chronological ordering before the training, validation and test datasets were separated. Short-term temporal dependencies were represented using lagged irradiance inputs selected through autocorrelation analysis, while continuous predictors were standardised using z score normalisation based on training period statistics only. The same mean and standard deviation were then applied to the validation, test and external site datasets. Lagged GHI was incorporated at the 10-min modelling interval using $k = 6$ input terms, defined as $[GHI_t, GHI_{t-10}, GHI_{t-20}, GHI_{t-30}, GHI_{t-40}, GHI_{t-50}]$, representing a 60 min memory window for predicting GHI_{t+10}. Only information available up to forecast issue time t was used, while ERA5, CAMS and MODIS predictors were causally aligned using product values with valid times less than or equal to t. Rows with incomplete lag histories, unavailable external predictors or lag windows crossing dataset split boundaries were removed to prevent information leakage.

Since GHI usually exhibits strong short-horizon autocorrelation under many sky conditions, baseline benchmarking was performed using a horizon-matched persistence reference, defined as $GHI_{pers(t+\Delta t)} = GHI_t$. The persistence model was evaluated on the same aligned timestamps as the proposed models to avoid performance inflation due to temporal persistence alone. Deterministic clear-sky radiative quantities and solar-geometry terms were evaluated consistently at the prediction time step. To assess transferability beyond Roodepoort, external validation was performed using publicly available ground-measured irradiance data from two independent South African Universities Radiometric Network (SAURAN) stations. The selected stations were Port Elizabeth, a coastal city site at latitude −34.0085907°, longitude 25.66526031° and elevation 35 m, and Upington, an inland desert-region site at latitude −28.489639°, longitude 21.519908° and elevation 884 m. These stations were not used during model training, validation or model selection. For each site, local coordinates were used to recalculate solar-geometry variables and Linke-turbidity clear-sky irradiance. The SAURAN GHI records were quality screened, aligned to the same 10-min modelling interval used for Roodepoort, and merged with the corresponding ERA5, CAMS and MODIS predictors. The same preprocessing steps, predictor construction and evaluation metrics were then applied to both external stations to ensure comparable validation under contrasting coastal and inland atmospheric conditions.

A Linke-turbidity based clear-sky baseline was used rather than a full standalone implementation of either the European Solar Radiation Atlas (ESRA) or Ineichen and Perez model. The formulation retained the main physical elements of these approaches, including solar geometry, pressure adjustment, optical air mass correction and Linke turbidity based atmospheric attenuation (Ineichen and Perez, 2002; Polo et al., 2009). The Linke turbidity factor, T_L, was obtained from a monthly climatological dataset for each study site and assigned to all 10-min records within the corresponding month. It was therefore treated as a monthly atmospheric descriptor, not as a 10-min observation. The same procedure was applied to Roodepoort, Port Elizabeth and Upington so that the clear-sky baseline was generated using site-specific solar geometry and monthly turbidity conditions. Solar geometry parameters, including the solar zenith angle θ_z and extraterrestrial normal irradiance G_{on}, were determined using standard solar geometry formulations. Under clear-sky conditions, DNI_{cs} was computed using the modified Beer-Lambert expression in Equation (1), where T_L represents the atmospheric turbidity relative to a clean dry atmosphere, P is the surface atmospheric pressure, and $m$is the relative optical air mass. The extraterrestrial horizontal irradiance was defined separately as $G_o = G_{on} \cos\theta_z$, while GHI_{cs} was obtained from the projected direct component and the modelled diffuse component as shown in Equation (3).

$$\mathrm{DNI}_{cs} = G_{on} \exp\left[-0.8662\, T_L\, m \left(\frac{P}{1013.25}\right)\right], \tag{1}$$

$$m = \frac{1}{cos\theta_z + 0.50572(96.07995 - \theta_z)^{-1.6364}} \tag{2}$$

$$GHI_{cs} = DNI_{cs} cos\theta_z + DHI_{cs}. \tag{3}$$

The relative optical air mass was determined from the Kasten and Young, (1989) formulation in Equation (2), where θ_z is the solar zenith angle expressed in degrees in the empirical correction term. The exponent term $(96.07995 - \theta_z)^{-1.6364}$ accounts for the nonlinear increase in atmospheric path length at high solar zenith angles. The clear-sky global horizontal irradiance was then computed using GHI decomposition, where the projected direct-beam component on the horizontal plane, $DNI_{cs}\cos\theta_z$, was added to the diffuse horizontal component, DHI_{cs}, as shown in Equation (3).

2.2. Physics-Informed Hybrid Learning Framework

Feature engineering and model development were designed as a unified physics-informed learning framework coupling atmospheric predictors with clear-sky radiative constraints. Predictor variables were grouped according to their physical role in modulating surface irradiance and are shown in Table 1. These comprised clear-sky radiative quantities and solar geometry, cloud properties, aerosol and atmospheric composition variables, near-surface meteorological conditions, and temporal memory features derived from lagged irradiance. Spatial collocation of gridded predictors to the measurement site was implemented by extracting ERA5 and CAMS fields at the surrounding grid points and applying bilinear interpolation to the site coordinates. MODIS cloud retrievals were collocated using the nearest valid footprint, with local averaging applied where multiple valid pixels were available in the vicinity of the site. All predictors were temporally aligned to the 10-min modelling interval using a strictly causal procedure, where only predictor values available at or before the forecast issue time were used. For each forecast issued at time t, ERA5, CAMS and MODIS predictors were selected only from product records with valid times less than or equal to t. ERA5 and CAMS hourly variables were therefore resampled to the 10-min interval using the most recent available valid record, rather than by centred interpolation between past and future hourly values. MODIS overpass derived variables were assigned only after the corresponding overpass time had occurred. No MODIS value from a future overpass was used to estimate predictors before that overpass time. This ensured that the predictor vector used to forecast GHI_{t+10} contained only information available at the forecast issue time.

Table 1. Predictors used in the study: Solar geometry, ERA5 meteorology and cloud variables, MODIS cloud properties, CAMS aerosols, and lagged GHI.

Predictor group	Predictor variables reported as model inputs	Source or native resolution
Target	Forecast target: GHI_{t+10}, representing the 10-min ahead global horizontal irradiance	Ground station, Davis Vantage Pro2, originally recorded at 1 min and processed to the 10-min modelling interval
Solar geometry and extraterrestrial terms	Solar zenith angle (θ_z), solar azimuth (γ_s), extraterrestrial normal irradiance (G_{on}), extraterrestrial horizontal irradiance (G_o), and relative air mass (m)	Derived and computed at the modelling time step
Clear-sky radiative baseline	Clear-sky global horizontal irradiance (GHI_{cs}), clear-sky direct normal irradiance (DNI_{cs}), and clear-sky diffuse horizontal irradiance (DHI_{cs})	Derived using the Linke-turbidity clear-sky formulation and computed at the modelling time step
ERA5 meteorology and moisture	Air temperature, dew-point temperature, surface pressure, wind components, and total column water vapour	ERA5, hourly, approximately 0.25° grid
ERA5 cloud state	Total cloud cover, low cloud cover, mid cloud cover, and high cloud cover	ERA5, hourly, approximately 0.25° grid
MODIS cloud properties	Cloud fraction, cloud optical thickness, cloud-top pressure, and cloud-top temperature	MODIS, discrete overpasses, pixel or footprint scale
CAMS aerosols and composition	Aerosol optical depth at 550 nm and aerosol speciation contributions	CAMS, hourly, approximately 0.4° to 0.75° grid, depending on product
Temporal memory features	Six historical GHI terms at 10-min spacing, defined as $(GHI_t), (GHI_{t-10}), (GHI_{t-20}), (GHI_{t-30}), (GHI_{t-40}), and (GHI_{t-50})$. These were selected from autocorrelation analysis and used as fixed-length short-term memory inputs for predicting (GHI_{t+10}).	Derived from ground-station GHI and processed to the 10-min modelling interval

The use of multi-source atmospheric predictors introduced uncertainty associated with temporal alignment, spatial representativeness, satellite retrieval accuracy and possible harmonisation artefacts. ERA5, CAMS and MODIS variables were available at coarser spatial or temporal resolutions than the ground measured GHI data and were therefore treated as contextual atmospheric indicators rather than exact point observations above the radiometer. ERA5 and CAMS predictors were aligned using the most recent valid records available at the forecast issue time, while MODIS overpass variables were assigned only after the corresponding overpass had occurred. Although this causal alignment prevented future information leakage, it could not fully resolve rapid local cloud transitions, cloud-edge effects or small-scale aerosol gradients, especially under broken cloud conditions. This limitation was partly reduced by using ground measured GHI as the target variable, including lagged GHI terms to capture local irradiance memory, and evaluating model performance under clear, low cloud, broken cloud and overcast regimes.

$$GHI_{t+10} = GHI_{cs,t+10} + \Delta GHI_{t+10}, \tag{4}$$

where GHI_{t+10} is the measured all-sky global horizontal irradiance at the 10-min ahead forecast time, $GHI_{cs,t+10}$ is the corresponding clear-sky global horizontal irradiance, and ΔGHI_{t+10} is the residual between the measured all-sky irradiance and the clear-sky baseline. Therefore, the residual is defined as the atmospheric deviation from clear-sky conditions, rather than the measured irradiance itself. This residual may be negative during cloud or aerosol attenuation and may become positive during short cloud-edge enhancement events. The residual component was modelled using a nonlinear function represented by Equation (5).

$$\Delta \widehat{GHI}_{t+10} = f(\boldsymbol{x}(t); \boldsymbol{\theta}), \tag{5}$$

where $\boldsymbol{x}(t)$ is the physics-informed feature vector available at forecast issue time t, $\boldsymbol{\theta}$ represents the model parameters, and $\Delta\widehat{GHI}_{t+10}$ is the predicted residual irradiance. The final hybrid forecast was then obtained by adding the predicted residual correction to the clear-sky baseline. This formulation was model agnostic and allowed the nonlinear residual function, (·), to be represented using complementary learning models. In the proposed hybrid framework, the residual ensemble consisted of LSTM, Random

Forest and Gradient Boosting models, which were trained to predict ΔGHI relative to the clear-sky baseline. The LSTM model captured short-term irradiance dependence caused by cloud passage, aerosol variation and atmospheric persistence, while Random Forest and Gradient Boosting captured nonlinear interactions among cloud, aerosol, meteorological and clear-sky predictors.

The framework was implemented using offline training and online inference. The models were trained using the historical training set and were then applied repeatedly for 10-min ahead forecasting without retraining at each forecast step. The input sequence length was fixed at six-time steps, corresponding to the six lagged GHI terms defined earlier and a 60 min memory window at the 10-min modelling interval. The LSTM architecture consisted of an input sequence layer, one LSTM hidden layer with 64 memory units, a dropout layer with a dropout rate of 0.20, one fully connected dense layer with 32 neurons, and a linear output layer for predicting the residual irradiance component. The model was trained using the Adam optimiser with a learning rate of 0.001, while mean squared error was used as the loss function. Training was performed using a batch size of 32 for a maximum of 100 epochs. Early stopping was applied by monitoring the validation loss, with a patience of 10 epochs, and the model weights corresponding to the lowest validation loss were restored for final testing. This stopping rule was used to reduce overfitting while preserving temporal generalisation.

The Random Forest and Gradient Boosting models were used as complementary nonlinear ensemble learners to capture non sequential interactions among cloud, aerosol, meteorological and clear-sky predictor variables. Hyperparameter tuning was performed using a validation-based search procedure, while the independent test dataset was reserved strictly for final evaluation. For the LSTM model, the candidate settings included hidden units of 32, 64 and 128, dropout rates of 0.10, 0.20 and 0.30, dense layer sizes of 16, 32 and 64 neurons, learning rates of 0.0001 and 0.001, and batch sizes of 16, 32 and 64. For the Random Forest model, the tuning process considered the number of trees, maximum tree depth, minimum samples per leaf and number of features evaluated at each split. For the Gradient Boosting model, the tuned parameters included the number of boosting stages, learning rate, maximum tree depth and minimum samples split. The final configuration for each model was selected based on the lowest validation RMSE. The final hybrid residual estimate was then obtained by averaging the residual predictions from the LSTM, Random Forest and Gradient Boosting models before adding the residual correction to the clear-sky GHI baseline. This allowed the hybrid forecast to combine temporal sequence learning with nonlinear predictor interaction modelling.

2.3. Model Validation and Performance Evaluation

Model validation followed a temporally blocked hold-out strategy designed to assess predictive performance, robustness and generalisability without information leakage. The model was trained using the summer period from 1 December 2021 to 28 February 2022, while the validation period covered 1 December 2022 to 28 February 2023. The validation set was used for hyperparameter selection, early stopping and model selection. After model selection, the trained model was evaluated without retraining over the remaining held-out seasonal periods in 2023. For seasonal analysis, the year was divided into summer, autumn, winter and spring according to the Southern Hemisphere seasonal calendar, namely summer: December to February, autumn: March to May, winter: June to August, and spring: September to November. The seasonal results therefore represent fixed-model evaluation over season-specific held-out subsets, rather than model retraining for each season. This chronological separation preserved the time-series structure of the irradiance data and allowed the model to be evaluated under atmospheric conditions not used during training. The validation set was used to monitor LSTM convergence, apply early stopping, restore the best-performing weights and guide final model selection before independent testing. Rolling-window retraining was not implemented because the aim was to evaluate the robustness of a fixed physics-informed residual-learning framework, rather than to improve performance through repeated seasonal updating. This design therefore provided a stricter test of whether the proposed framework could generalise beyond the atmospheric regime used for model development. The final hybrid model was evaluated against four benchmark models, namely a horizon matched persistence model, a standalone clear-sky irradiance model, a direct AI-only machine learning model and a standalone GRU model. These models were used only for comparison and were not included in the proposed hybrid residual ensemble. The proposed hybrid ensemble was formed only from the LSTM, Random Forest and Gradient Boosting residual learners. The AI-only benchmark was implemented as a direct data-driven ensemble using Random Forest and Gradient Boosting models trained to predict GHI 10-minutes ahead directly from the atmospheric, meteorological, cloud, aerosol, solar-geometry and lagged GHI predictors listed in Table 1. Unlike the proposed hybrid model, the AI-only benchmark did not learn the residual term relative to the clear-sky baseline and did not add a learned residual correction to clear-sky GHI. Therefore, the AI-only model served as a conventional machine-learning benchmark for testing whether direct data-driven prediction alone could match the physically constrained residual-learning framework.

A standalone gated recurrent unit (GRU) model was also implemented as a contemporary deep learning benchmark. The same predictor set, six-step input sequence, and training, validation and test partitions used for the proposed framework were adopted. The GRU architecture consisted of one hidden layer with 64 units, dropout of 0.20, one dense layer with 32 neurons and a linear output layer for predicting the 10-min ahead GHI. The model was trained using the Adam optimiser, mean squared error loss, a learning rate of 0.001, a batch size of 32 and a maximum of 100 epochs. Early stopping was applied using validation loss with a patience of 10 epochs. The GRU model was not included in the hybrid residual ensemble but was used to assess whether the proposed physics-informed residual structure provided additional value over a recurrent data-driven benchmark.

The predictive performance was evaluated usingstandard irradiance forecasting metrics. The following metrics were used to evaluate and validate the performance of the model; Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), and coefficient of determination (R^2) and these were computed as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(GHI_i - \widehat{GHI}_i\right)^2}, \tag{6}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|GHI_i - \widehat{GHI}_i|, \tag{7}$$

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(GHI_i - \widehat{GHI}_i), \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(GHI_i - \widehat{GHI}_i)^2}{\sum_{i=1}^{N}(GHI_i - \overline{GHI})^2}. \tag{9}$$

where N is the number of samples, i is the sample index, GHI_i and $\widehat{\mathrm{GHI}}_i$ are the observed and predicted GHI values, respectively, and $\overline{GHI}$ is the mean observed GHI over the evaluation set. The hybrid model performance was benchmarked against the persistence baseline using a performance score (S_p);

$$S_p = 1 - \frac{RMSE_m}{RMSE_p} \tag{10}$$

where $RMSE_m$ and $RMSE_p$ are the RMSE values for the model and the persistence baseline, respectively. The persistence baseline was defined using the same forecasting horizon Δt as, $GHI_{p(t+\Delta t)} = GHI_t$ with $\Delta t = 10\text{min}$.

2.4. Uncertainty, Sensitivity, and Interpretability Analysis

Uncertainty quantification was performed using the hybrid residual ensemble with Monte Carlo dropout applied only to the LSTM residual learner during inference. This approach was selected because it provides a practical and computationally efficient way of estimating predictive uncertainty without requiring a separate probabilistic model or repeated full retraining. For each Monte Carlo pass, dropout was activated in the trained LSTM branch to generate a stochastic residual estimate, while the Random Forest and Gradient Boosting residual estimates remained fixed as deterministic ensemble components. The stochastic LSTM residual estimate was then averaged with the deterministic tree-based residual estimates and added to the clear-sky GHI baseline. This produced a distribution of hybrid GHI forecasts from which the mean prediction, standard deviation and prediction intervals were calculated. Therefore, the uncertainty estimate represents the propagation of LSTM dropout uncertainty through the hybrid residual ensemble, rather than dropout being applied to all ensemble members. A normal-approximation prediction interval (PI), shown in Equation (11), was used to summarise the resulting predictive uncertainty. In this expression, $\widehat{GHI}_t$ is the predicted global horizontal irradiance at time t, $\sigma_{\widehat{\mathrm{GHI}}_t}$ is the ensemble prediction standard deviation, α is the significance level, and $z_{1-\alpha/2}$ is the standard normal critical value for the two-sided $(1-\alpha)$ prediction interval.

$$PI_{(1-\alpha),t} = \left[\widehat{GHI}_t - z_{(1-\alpha/2)}\,\sigma\widehat{GHI}_t\,, \widehat{GHI}_t + z_{(1-\alpha/2)}\sigma\widehat{GHI}_t,\right] \tag{11}$$

$$Imp(X_j) = RMSE_{perm(j)} - RMSE_{base}, \tag{12}$$

$$\widehat{GHI}_t = \phi_0 + \sum_{j=1}^{M} \phi_j \left(x_{j,t}\right) \tag{13}$$

$$S_{X,t} = \frac{X_t}{\mathrm{GHI}_{cs,t}} \frac{\partial GHI_{cs,t}}{\partial X_t}. \tag{14}$$

Probabilistic calibration was further assessed using interval-based diagnostics. For each nominal prediction interval, the prediction interval coverage probability (PICP) was calculated as the proportion of observed GHI values falling between the lower and upper prediction bounds. The average coverage error (ACE) was then computed as the difference between empirical and nominal coverage. The Winkler score was also used to jointly assess interval sharpness and reliability, since it penalises prediction intervals that are either too wide or fail to contain the observed value. Calibration curves were generated by comparing nominal and empirical coverage across multiple confidence levels. Mean prediction interval width (MPIW) was used to measure the average width of the prediction intervals in W/m^2, while prediction interval normalised average width (PINAW) was used to express the interval width relative to the irradiance range. Together, PICP, ACE, MPIW, PINAW and Winkler score were used to assess the reliability, sharpness and practical usefulness of the probabilistic forecasts.

Complementary approaches were employed to model sensitivity and interpretability. The permutation importance of the tree-based models was quantified using the increase in RMSE after predictor permutation, as shown in Equation (12). In this equation, $Imp(X_j)$ is the raw permutation importance of the j^{th}predictor, $RMSE_{perm(j)}$is the RMSE obtained after randomly permuting predictor X_j, and $RMSE_{base}$ is the original RMSE before permutation. To express the results as relative contribution percentages, the raw permutation importance values were normalised against the total positive importance across all predictors in the analysed set. Therefore, the reported feature-importance values are unitless relative percentages and sum to 100% within each feature group or variable-level analysis. SHapley Additive exPlanations (SHAP) values were used only to support the physical interpretation of the ranking, not to calculate the percentage contribution values. Equation (13) was applied for neural network models and SHAP values were employed to attribute predictions to individual inputs where ϕ_0 is the baseline prediction, $x_{j,t}$ is the j^{th} input feature at time t, $\phi_j(x_{j,t})$ is the SHAP value representing the contribution of $x_{j,t}$ to the predicted irradiance, and $M$is the total number of input features. Perturbation of clear-sky parameters and computing of partial derivatives was done to assess the sensitivity to physical parameters as indicated in Equation (14), where $S_{X,t}$ is the normalised sensitivity of $\mathrm{GHI}_{cs,t}$ to the clear-sky parameter X, and $X \in T_L, \theta_z, P, m, G_0$. This normalised form allows the relative influence of different clear-sky parameters to be compared despite differences in their physical units. This sensitivity analysis was used to compare the relative influence of the clear-sky physical parameters on GHI_{cs}, while model robustness was evaluated separately through the reported sky-condition and seasonal performance comparisons.

3. Results and discussion

3.1. Predictive Performance of the Hybrid Model

The predictive performance of the proposed hybrid model was evaluated against four reference benchmarks during the Roodepoort summer test period. These included a horizon matched persistence model, a standalone clear-sky irradiance model, a purely data-driven AI model and a standalone GRU model. The persistence model was included as a short horizon operational benchmark because GHI usually exhibits strong autocorrelation at the 10-min forecasting scale. The clear-sky model represented the deterministic physical baseline, while the AI-only and standalone GRU models represented data-driven benchmarks without explicit clear-sky residual constraints. The GRU model was therefore used only to test whether recurrent temporal learning alone could match the proposed physics-informed residual ensemble. This benchmarking structure therefore allowed the proposed framework to be assessed against temporal persistence, standalone physical modelling, conventional machine learning and contemporary recurrent deep learning architecture.

Table 2. Performance comparison for Roodepoort (summer period).

Model	RMSE (W/m²)	MAE (W/m²)	MBE (W/m²)	R²
Persistence	116	87	−5	0.81
Clear-sky	142	110	−45	0.72
AI-only	108	83	−22	0.83
GRU	96	72	−14	0.87
Hybrid	79	59	−8	0.91

As shown in Table 2, the hybrid model achieved the strongest overall predictive performance during the Roodepoort summer test period, with the lowest RMSE and highest R^2 among all benchmark models. Based on the RMSE values in Table 2, the hybrid framework reduced error by 32%, 44%, 27% and 18% relative to persistence, clear-sky, AI-only and GRU models, respectively. This indicates that the clear-sky constrained residual formulation provided additional predictive value beyond temporal persistence, standalone physical modelling, direct data-driven learning and recurrent temporal modelling. The hybrid model also showed the smallest bias magnitude, thus indicating improved overprediction control relative to the other data-driven benchmarks. Therefore, the results show that the main advantage of the proposed framework came from combining temporal memory, atmospheric predictors and a physically meaningful clear-sky baseline.

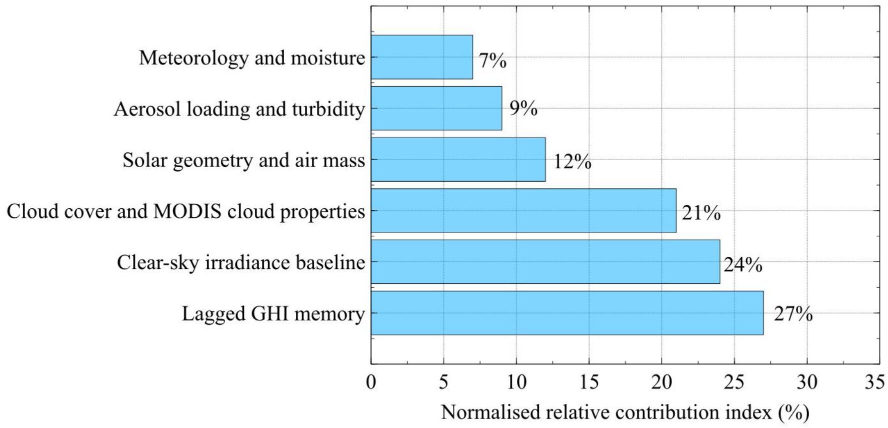

Figure 1. Explainable AI-based relative contribution of the main predictor groups used in the hybrid residual learning framework.

These results are consistent with recent solar radiation studies, such as Nallakaruppan et al., (2026), who reported a best Random Forest performance of R^2 = 0.9028 for solar radiation estimation. The comparable R^2 obtained in the present study shows that the proposed hybrid framework is competitive with recent machine learning based solar radiation studies. Compared with recent transformer based forecasting models, the proposed hybrid model also showed competitive performance for high-granularity solar irradiance prediction. Mercier et al., (2023) also reported transformer based RMSE values of 76.94 W/m^2 for one step ahead forecasting and 112.00 W/m^2 for a 15 min ahead case, while Zhuang et al., (2024) further reported a lower seasonal RMSE range of 44.24 to 63.93 W/m^2 using a specialised transformer architecture. Therefore, the present model performs within the range of recent transformer-based approaches, while retaining the added advantages of physical consistency, interpretability and seasonal robustness.

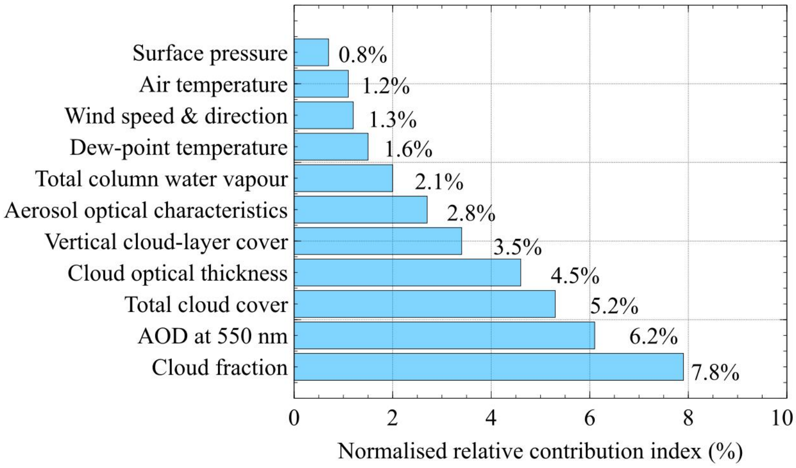

Figure 2. Variable-level atmospheric predictor sensitivity of the hybrid forecasting model. The values represent normalised permutation-importance percentages.

The interpretation of the hybrid model was further supported using an explainable AI-based contribution analysis, which examined the relative influence of the main predictor groups (Figure 1). Lagged GHI memory remained important for 10-min ahead forecasting, while the clear-sky irradiance baseline and cloud-related predictors showed that the model also used physical and atmospheric information beyond past irradiance values. This is important because rapid changes in irradiance are mainly controlled by cloud movement, cloud-edge effects and deviations from clear-sky conditions. The lower relative influence of aerosol loading, turbidity, meteorology and moisture suggests that these variables played a secondary role in the analysed dataset, although they remain relevant for explaining atmospheric attenuation under hazy or humid conditions. The explainable AI analysis therefore confirms that the improved forecasting performance was not produced by statistical fitting alone, but by the combined influence of irradiance memory, clear-sky physical constraints and meteorologically meaningful atmospheric predictors.

The group-level explainable AI analysis was expanded into a variable-level sensitivity analysis to provide a more detailed assessment of individual atmospheric predictor contributions. Each atmospheric predictor was independently perturbed using the permutation importance procedure, while SHAP attribution was used to support the physical interpretation of the ranking. The results are presented in Figure 2.

The variable-level results showed that cloud- and aerosol-related predictors had the strongest atmospheric influence on the hybrid forecast. Cloud fraction recorded the highest contribution of 7.8%, followed by aerosol optical depth at 550 nm, total cloud cover and cloud optical thickness, with contributions of 6.2%, 5.2% and 4.5%, respectively. This indicated that short-term GHI prediction was mainly controlled by cloud shading, aerosol attenuation and cloud optical state. Moisture and meteorological variables made lower but meaningful contributions, with total column water vapour, dew point temperature, wind components, air temperature and surface pressure contributing between 0.8% and 2.1%. These lower values suggest that their effects on GHI were mainly indirect through humidity, boundary layer state, cloud advection and atmospheric density. Therefore, the sensitivity results confirmed that the hybrid model was mainly influenced by physically relevant cloud and aerosol predictors, while secondary meteorological variables supported the representation of local atmospheric variability. Figure 3 further illustrates the temporal behaviour of the models during a representative partly cloudy period. The clear-sky model followed a smooth diurnal irradiance pattern but remained insensitive to short-term cloud passages. The AI only model responded more closely to the measured trend but still showed overshoot and temporal lag during rapid irradiance ramps. The GRU model improved the temporal response but did not consistently capture the full timing and magnitude of sharp irradiance reductions and recoveries. However, the hybrid model followed both the timing and magnitude of the measured irradiance reductions and recoveries more closely. This behaviour is important for high-granularity photovoltaic forecasting because short-term cloud-edge transitions are among the main causes of rapid PV power fluctuations. Therefore, the value of the hybrid framework is not only reflected in the lower aggregate error, but also in its improved ability to reproduce physically meaningful irradiance dynamics under variable sky conditions.

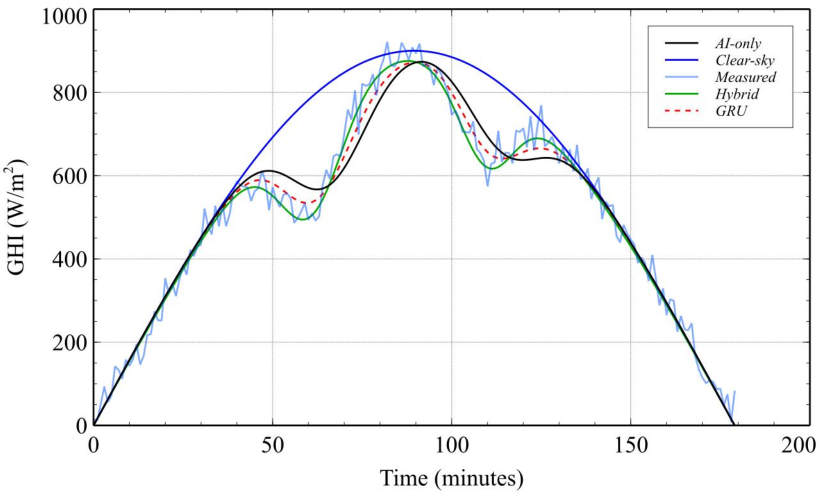

Figure 3. Time series comparison of measured GHI and predicted irradiance from the clear-sky, AI-only, GRU, and hybrid models during a representative partly cloudy period, showing the ability of the hybrid model to more closely track short-term irradiance reductions and recoveries.

3.2. Performance Under Different Sky Conditions

The model was further assessed under different atmospheric regimes to determine whether the observed improvement was consistent under changing sky conditions. Prediction errors were stratified using ERA5 cloud cover fractions and classified into clear, low cloud, broken cloud, and overcast regimes. This classification was important because cloud state directly controls the degree of irradiance variability, hence providing a more physically meaningful assessment of model robustness than aggregate error metrics alone.

As shown in Figure 4, the hybrid model achieved the lowest RMSE across all sky regimes, followed by the GRU model. Under clear conditions, the hybrid model recorded an RMSE of 32 W/m^2, compared to 39 W/m^2 for GRU, 48 W/m^2 for the AI-only model and 61 W/m^2 for the clear-sky model. This showed that even under weak cloud forcing, the hybrid framework corrected deviations caused by turbidity, aerosols, moisture and small atmospheric variations that were not fully represented by the idealised clear-sky formulation. The GRU model provided a stronger temporal benchmark than the generic AI-only model because it captured sequential irradiance behaviour more effectively. However, it remained less accurate than the hybrid model because it did not include an explicit clear-sky residual baseline. Therefore, the performance gap between GRU and the hybrid model indicated the added value of combining temporal learning with clear-sky residual correction.

The performance difference became more pronounced under low cloud and broken cloud conditions, where irradiance variability was stronger and more difficult to predict. In these regimes, the clear-sky model produced higher RMSE values because it could not respond to intermittent shading, cloud-edge enhancement and rapid transitions between direct and diffuse irradiance. The AI-only model reduced the error by capturing nonlinear relationships between irradiance and atmospheric predictors, while GRU further improved the response through short-term temporal learning. However, both data-driven models remained limited during rapid irradiance ramps, thus indicating weaker physical control under short cloud transition events. The broken cloud regime produced the largest errors for all models, with RMSE values of 230 W/m^2 for the clear-sky model, 170 W/m^2 for the AI-only model, 148 W/m^2 for GRU and 130 W/m^2 for the hybrid model. This confirmed that broken cloud conditions were the most difficult forecasting regime because alternating shading, cloud-edge enhancement and irradiance recovery require both the timing and magnitude of short-term GHI changes to be captured accurately.

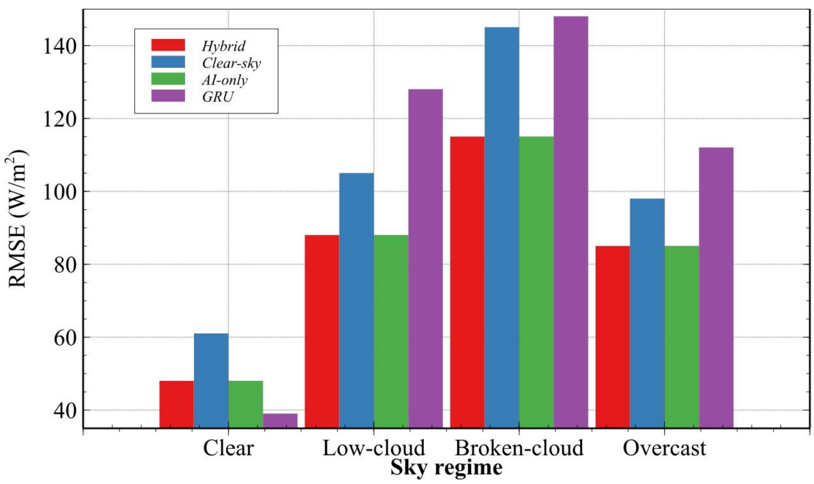

Figure 4. RMSE performance of the clear-sky, AI-only, GRU and hybrid models under clear, low cloud, broken cloud and overcast conditions, showing the stronger advantage of the hybrid framework under variable cloud regimes.

The hybrid model showed the strongest advantage under broken cloud conditions because the clear-sky residual learning structure preserved the physical irradiance envelope while allowing the learning component to correct rapid atmospheric departures. This improvement is operationally important for photovoltaic systems, where rapid irradiance ramps can affect short-term power stability and grid management decisions. Under overcast conditions, the differences between the models became smaller because diffuse irradiance dominated and sharp irradiance ramps were less frequent. However, the hybrid model still maintained the lowest RMSE, indicating that the combined use of clear-sky physics and atmospheric predictors improved the representation of both highly variable and more stable sky conditions. The present results are also comparable to those of Yan et al., (2020), who reported 10-min RMSE values of 39.72, 38.82, 49.49 and 11.44 W/m^2 for spring, summer, autumn and winter, respectively, using a GRU attention-based irradiance forecasting model.

3.3. Seasonal Robustness Beyond the Summer Training Period

Model performance was assessed across season-specific held-out subsets without retraining, using the Southern Hemisphere seasonal calendar: summer from December to February, autumn from March to May, winter from June to August and spring from September to November. This was done to evaluate whether the proposed hybrid framework could preserve prediction accuracy beyond the summer conditions used during model training and validation.

As shown in Figure 5 and Table 3, the hybrid model demonstrated the most stable seasonal performance among all evaluated models. Its RMSE increased from 79 W/m^2 in summer to 96 W/m^2 in spring, representing an increase of 17 W/m^2, or about 21.5%. The R^2 value also decreased slightly from 0.91 to 0.87. This indicated that the summer trained hybrid model retained useful predictive capability under other seasonal conditions, although the effects of seasonal changes in solar elevation, atmospheric path length, cloud morphology, aerosol loading and moisture conditions were not completely removed.

The AI-only model showed a stronger seasonal decline, with RMSE increasing from 108 W/m^2 in summer to 140 W/m^2 in winter. This represented an increase of 32 W/m^2, or about 29.6%, thus indicating stronger dependence on season specific statistical patterns learned during training. The GRU model improved seasonal stability compared with the AI-only model, with RMSE increasing from 96 W/m^2 in summer to 120 W/m^2 in winter. However, it still remained less stable than the hybrid model. This showed that temporal learning alone was not sufficient to fully account for seasonal changes in solar geometry, atmospheric path length, cloud structure and aerosol loading. The remaining performance gap between the GRU and hybrid models therefore highlighted the added value of the clear-sky residual formulation, where the physical irradiance baseline helped stabilise the forecast across changing seasonal conditions.

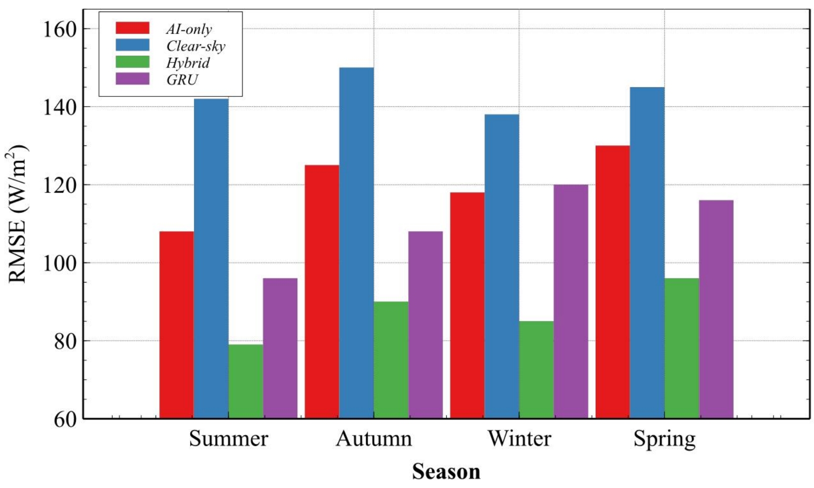

Figure 5. Seasonal RMSE comparison of the clear-sky, AI-only, GRU and hybrid models across summer, autumn, winter and spring, showing seasonal robustness under changing irradiance and atmospheric conditions without additional retraining.

Table 3. Seasonal performance comparison of the clear-sky, AI-only, GRU and hybrid models across summer, autumn, winter and spring at Roodepoort using the 10-min forecasting interval

Season	Model	RMSE (W/m²)	MAE (W/m²)	MBE (W/m²)	R²
Summer	Clear-sky	142	110	−45	0.72
	AI-only	108	83	−22	0.83
	GRU	96	72	−14	0.87
	Hybrid	79	59	−8	0.91
Autumn	Clear-sky	150	116	−42	0.70
	AI-only	125	95	−28	0.80
	GRU	108	80	−18	0.85
	Hybrid	85	64	−10	0.90
Winter	Clear-sky	138	122	−40	0.68
	AI-only	140	105	−30	0.77
	GRU	120	88	−21	0.82
	Hybrid	92	69	−12	0.88
Spring	Clear-sky	152	118	−43	0.69
	AI-only	132	100	−27	0.79
	GRU	116	84	−19	0.84
	Hybrid	96	72	−11	0.87

The clear-sky model remained physically consistent across seasons, but its RMSE values remained high, ranging from 138 W/m^2 in winter to 152 W/m^2 in spring. This showed that deterministic clear-sky physics alone could not adequately represent real sky irradiance variability caused by clouds, aerosols and moisture. Therefore, the seasonal results confirm that the main strength of the hybrid framework was obtained from combining physical consistency with adaptive residual correction. The improved seasonal transferability of the hybrid model can therefore be attributed to its clear-sky irradiance anchoring, which preserved the solar-geometry constraint, while the residual learner adjusted for cloud, aerosol and moisture-driven deviations. This means that the model provided a more transferable forecasting structure without requiring seasonal retraining. The results therefore showed that the hybrid approach improved seasonal robustness by separating the stable solar geometry component from the variable atmospheric component. This was supported by the GRU comparison, where temporal learning improved performance but still showed stronger seasonal degradation than the hybrid model. Hence, physically constrained residual learning provided a more stable framework for high-granularity irradiance forecasting without frequent seasonal retraining.

3.4. Multi-site Transferability Assessment

Although model development was based on the Roodepoort station, external validation was conducted using two independent SAURAN stations to assess whether the trained framework could retain useful predictive skill under different geographic and atmospheric conditions. For each external station, solar geometry variables and Linke turbidity clear-sky irradiance were recalculated using the local coordinates. The atmospheric, aerosol, cloud and meteorological predictors were then aligned to the same 10-min modelling interval used for Roodepoort. This ensured that the trained hybrid residual learning model was tested under comparable temporal conditions, but under different climatic and atmospheric settings.

The external site validation results are presented in Table 4. The proposed framework retained useful predictive skill at both external stations, although performance declined relative to the original Roodepoort test site. The RMSE increased from 79 W/m^2 at Roodepoort to 88 W/m^2 at Upington and 103 W/m^2 at Port Elizabeth, while R^2 values remained 0.88 and 0.84, respectively. The mean external site performance was 95.5 W/m^2 for RMSE, 72 W/m^2 for MAE, −15 W/m^2 for MBE, R^2 of 0.86 and PICP@95 of 0.91. These results indicate promising but limited transferability, rather than full multi-site generalisation.

Table 4. Preliminary external site validation performance of the proposed hybrid residual learning framework across the original Roodepoort test site and two independent SAURAN stations.

Station	**Location type**	**Climatic setting**	**RMSE (W/m²)**	**MAE (W/m²)**	**MBE (W/m²)**	**R²**	**PICP@95**
Roodepoort	Original test site	Inland urban Highveld site	79	59	−8	0.91	0.93
Upington	External SAURAN station	Inland desert-region site	88	66	−12	0.88	0.92
Port Elizabeth	External SAURAN station	Coastal city site	103	78	−18	0.84	0.90
Mean external-site performance	External validation	Two-station external mean	95.5	72	−15	0.86	0.91

The observed performance loss was expected because the external stations were not used during model development and no local retraining was applied. The RMSE increase was approximately 11% at Upington and 30% at Port Elizabeth, while the mean external site RMSE increased by about 21% relative to Roodepoort. This reduction was therefore interpreted as the realistic penalty of applying a fixed model to new atmospheric and geographic conditions. However, the retention of R^2 values above 0.84 and PICP values between 0.90 and 0.92 showed that the model preserved useful deterministic and probabilistic skill outside the original training environment. The stronger performance at Upington suggests that transferability was better under atmospheric conditions that were closer to the Roodepoort setting. In contrast, the larger error at Port Elizabeth indicated that coastal conditions, local cloud dynamics, aerosol loading, moisture behaviour and the spatial representativeness of ERA5, CAMS and MODIS predictors can reduce model accuracy. This is physically reasonable because high-granularity irradiance forecasting is strongly affected by local cloud transitions, which may not be fully captured by gridded atmospheric predictors.

The results therefore provide preliminary evidence that the hybrid framework can be transferred to independent South African stations, but they do not demonstrate full site independence. The clear-sky component improved transferability because it was recalculated for each location, thereby preserving the site-specific solar geometry and irradiance envelope. However, the residual atmospheric correction remained partly dependent on the cloud, aerosol, humidity and meteorological patterns represented during training. A similar site-dependent trend was reported by (Siddiqa et al., 2025), where the solar power forecasting model produced different R^2 values across individual site-specific and combined datasets, ranging from 0.7956 to 0.9692, with a combined dataset value of 0.8967. In the present study, the hybrid model also retained useful external site performance, with R^2 values of 0.88 and 0.84 at the two SAURAN stations, although some loss was observed relative to Roodepoort. This means that transferability was supported by the recalculated physical clear-sky baseline, while the residual atmospheric correction remained dependent on the representativeness of local cloud, aerosol, humidity and meteorological patterns. Thus, limited site-specific calibration or additional local training data may still be required to obtain the highest accuracy at new sites.

The remaining limitation is related to the scale of external validation. Although the use of two independent SAURAN stations improved the evidence for spatial transferability, broader validation across more stations, longer time periods, coastal and inland climates, and different aerosol and cloud regimes would be required before claiming full geographic generalisability. Therefore, the present results should be interpreted as evidence of promising transferability under independent site conditions, rather than proof of universal model applicability.

3.5. Uncertainty and Reliability of Predictions

Besides point-forecasting accuracy, reliable short-term solar irradiance prediction for operational photovoltaic applications also requires well calibrated uncertainty estimates that reflect atmospheric variability and regime transitions. It is therefore not sufficient for a forecasting model to provide only deterministic irradiance values, because operational decisions also depend on how reliable the prediction bounds are during stable and rapidly changing sky conditions. For this reason, uncertainty quantification was integrated into the proposed hybrid framework by applying Monte Carlo dropout to the LSTM residual branch and propagating its stochastic predictions through the hybrid ensemble, while the Random Forest and Gradient Boosting branches remained deterministic.

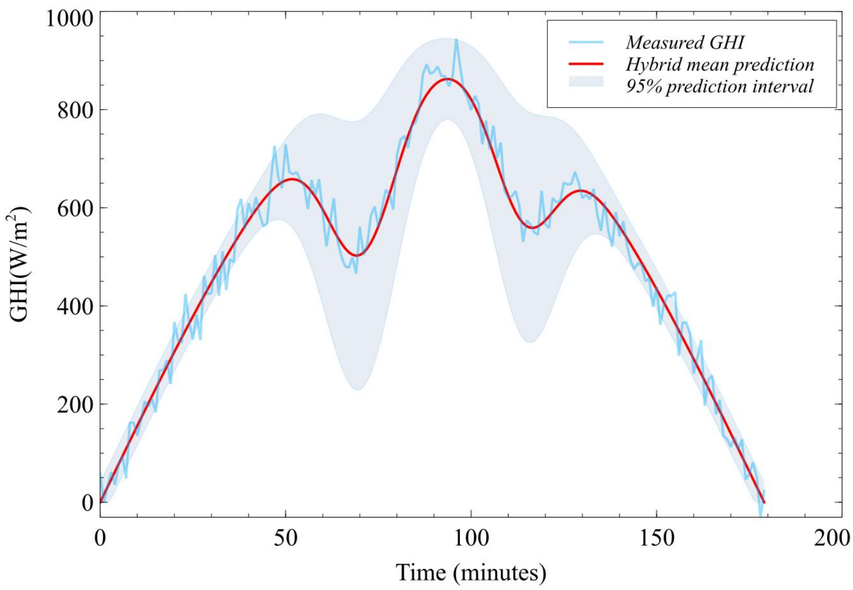

Figure 6. Hybrid model mean prediction and 95% prediction interval during a representative cloud transition event, showing the variation in forecast uncertainty under changing irradiance conditions.

Figure 6 shows the mean irradiance prediction of the hybrid model together with the corresponding 95% prediction intervals during a representative cloud transition event. The prediction intervals showed physically meaningful and adaptive behaviour. During stable atmospheric periods, especially near clear-sky intervals dominated by deterministic solar geometry, the intervals narrowed to below 50 W/m², thus indicating high predictability and limited short-term variability. However, during broken cloud conditions, the uncertainty bands expanded substantially, in some cases exceeding 150 W/m², which coincided with cloud-edge effects, intermittent shading and sharp irradiance ramps. This widening was important because it showed that the model did not maintain overconfident predictions when the atmosphere became more variable. In practical terms, narrow prediction intervals indicate periods when photovoltaic output can be scheduled with higher confidence, while wider intervals warn operators that additional reserve margin or flexible battery dispatch may be required.

The prediction intervals shown in Figure 6 achieved an empirical coverage of 93%, thus indicating close agreement between the predicted uncertainty range and the observed irradiance values during the cloud transition event. This result showed that the uncertainty estimates were not excessively narrow or overly conservative. It is therefore important because a useful probabilistic forecast must capture most observed irradiance outcomes while still maintaining intervals that are sufficiently sharp for practical decision making.

Table 5. Uncertainty calibration of the probabilistic forecasts (empirical coverage vs nominal and interval width).

Model	Nominal interval	PICP	ACE	MPIW	PINAW	Winkler score
AI-only	50%	0.42	−0.08	54	0.054	82
AI-only	80%	0.70	−0.10	90	0.090	132
AI-only	90%	0.81	−0.09	110	0.110	163
AI-only	95%	0.88	−0.07	125	0.125	188
Hybrid	50%	0.47	−0.03	64	0.064	77
Hybrid	80%	0.78	−0.02	106	0.106	118
Hybrid	90%	0.88	−0.02	126	0.126	137
Hybrid	95%	0.93	−0.02	140	0.140	152

As shown in Table 5, the AI-only model consistently under-covered the observed irradiance values across the nominal prediction intervals. For example, its PICP was 0.88 at the 95% nominal interval, thus indicating that the uncertainty bounds were too narrow and did not fully capture actual irradiance variability. This under-coverage is important in operational photovoltaic systems because it may create false confidence during cloud transition periods, leading to insufficient reserve planning, poor battery dispatch and weaker preparation for sudden PV output ramps. In contrast, the hybrid model produced better calibrated uncertainty intervals. At the 95% nominal interval, PICP improved from 0.88 for the AI-only model to 0.93 for the hybrid model, while at the 90% interval it improved from 0.81 to 0.88. Although the MPIW increased from 125 W/m^2 to 140 W/m^2 at the 95% level, this was not interpreted as a weakness. Rather, it showed that the hybrid model allowed the uncertainty range to expand when atmospheric variability increased. Therefore, a slightly wider but better calibrated interval was more useful than a narrow interval that frequently missed observed irradiance values.

The lower ACE and Winkler score values further showed that the hybrid model achieved a better balance between reliability and sharpness. The Winkler score decreased from 188 for the AI only model to 152 for the hybrid model at the 95% interval, thus indicating that the hybrid forecast was not simply wider, but less overconfident and less heavily penalised. From an operational photovoltaic perspective, this means that the hybrid uncertainty estimates provided better guidance for generation scheduling, reserve planning and battery operation, especially under broken cloud or ramping conditions. The improved probabilistic behaviour can be attributed to the physics-informed residual learning structure of the hybrid framework. The clear-sky radiative baseline constrained the forecast within a physically realistic irradiance envelope, while the residual learner allowed the intervals to expand under cloud driven attenuation and atmospheric variability. This confirmed that the uncertainty estimates were not governed by statistical dispersion alone but were linked to meaningful physical changes in the irradiance field.

3.6. Computational Cost and Operational Feasibility

The operational feasibility of the proposed framework was evaluated by separating offline model development from online forecasting. Model training was treated as a periodic task carried out when new historical data become available, when site conditions change, or when recalibration is required. The computational assessment was carried out on a Dell Precision Tower 7910 workstation with 64 GB RAM. For the three-year dataset used in this study, preprocessing, timestamp alignment, lag construction and atmospheric predictor merging required approximately 5 min. The LSTM residual model required approximately 15 min for training, while the Random Forest and Gradient Boosting models required approximately 4 min and 7 min, respectively. The standalone GRU benchmark required approximately 13 min under the same data split and training configuration. Once the models were trained, the generation of a single 10-min ahead forecast required approximately 0.8 s, excluding the time needed to download or update external atmospheric products. These timings indicate that the computational burden is concentrated during model development, whereas the trained model can generate 10-min ahead forecasts within the operational forecast cycle.

For operational implementation, each 10-min prediction required recent GHI history, updated solar geometry variables, clear-sky irradiance estimates and aligned ERA5, CAMS and MODIS predictors. As summarised in Table 6, most components had low computational requirements during inference, allowing deterministic and probabilistic forecasts to be generated within the required forecast cycle after model training. However, computational feasibility alone does not guarantee real-time deployability because external atmospheric products may be delayed by product update frequency, data access, spatial interpolation requirements and communication constraints. Delayed cloud, aerosol or moisture information can reduce forecast value during rapidly changing atmospheric conditions. Therefore, practical deployment would require automated data ingestion, predictor synchronisation, quality control and communication reliability. The framework is also scalable to larger photovoltaic monitoring networks because the trained model requires lightweight inference at each forecast step. For multi-site deployment, additional requirements would mainly involve site level preprocessing, including local solar geometry calculation, clear-sky baseline generation, atmospheric predictor alignment and ground measured GHI quality control. These processes can be automated across sites, although deployment may still be affected by missing sensor records, spatial mismatch between gridded predictors and local cloud fields, and the need for periodic recalibration when site conditions change.

Table 6. Computational requirements, operational roles and practical implications of the main components of the proposed hybrid solar irradiance forecasting framework.

Component	Computational requirement	Operational role	Practical implication
Solar geometry and clear-sky baseline	Low	Deterministic baseline calculation at each 10-min step	Suitable for real-time implementation
Lagged GHI features	Low	Captures short-term local irradiance memory	Requires only recent ground-station measurements
ERA5, CAMS and MODIS predictors	Moderate	Provides atmospheric, aerosol and cloud-state information	Main limitation is data latency and spatial/temporal alignment
LSTM residual model	Moderate during offline training, low during inference	Captures temporal irradiance dependence and atmospheric persistence	Does not require retraining at every forecast step
Random Forest and Gradient Boosting	Moderate during offline training, low during inference	Captures nonlinear predictor interactions	Useful for robust ensemble comparison
Ensemble uncertainty estimation	Moderate	Generates prediction intervals for operational risk assessment	Adds reliability information for grid and PV management
Full operational workflow	Low to moderate during inference	Produces 10-min ahead deterministic and probabilistic forecasts	Feasible for near-real-time operational forecasting where all input data streams are available within the forecast cycle

The proposed framework is therefore computationally suitable for short-term operational PV forecasting because model inference is fast after training has been completed. However, this feasibility should not be interpreted as automatic real time deployability. The main practical constraint remains the timely availability of external atmospheric predictors, including ERA5, CAMS and MODIS, together with reliable data ingestion, predictor synchronisation and quality control. Therefore, the framework can support near-real-time PV decision making when the required ground-based and external predictor streams are available within the 10-min forecast cycle.

4. Conclusion

The investigation demonstrated that coupling a clear-sky irradiance baseline with residual machine learning provides an effective approach for 10-min ahead global horizontal irradiance forecasting under variable atmospheric conditions. The proposed framework improved forecasting accuracy, physical interpretability and operational reliability because the deterministic clear-sky component was retained, while the learning models corrected atmospheric deviations caused by clouds, aerosols, moisture and near-surface meteorology. The hybrid framework outperformed the persistence, clear-sky, AI-only and standalone GRU benchmarks, with the strongest improvement observed under broken cloud conditions. The model achieved an RMSE of 79 W/m^2, MAE of 59 W/m^2, MBE of −8 W/m^2 and R^2 of 0.91 during the Roodepoort summer test period. This confirmed that the residual learning structure improved both aggregate error reduction and the representation of rapid irradiance variability, especially when cloud driven attenuation caused sharp departures from clear-sky behaviour.

The seasonal and external site assessments further showed that the model retained useful generalisation ability. Seasonal RMSE remained between 79 and 96 W/m^2 without retraining, while the external SAURAN validation produced a mean RMSE of 95.5 W/m^2 and R^2 of 0.86. These results indicate that the clear-sky constrained residual structure improved robustness under changing solar elevation, cloud morphology, aerosol loading and atmospheric path length, although some performance reduction was still observed under contrasting site conditions. The framework also provided operational value through calibrated uncertainty intervals and lightweight inference. The 95% prediction interval achieved 93% empirical coverage, thus allowing the reliability of each short-term forecast to be assessed during stable and variable sky conditions. Furthermore, most computational cost occurred during offline model development, while online inference remained suitable for near real time photovoltaic forecasting when ground-based and external predictor streams were available within the 10-min forecast cycle.

However, the study still has limitations. Model development and internal validation were based mainly on the Roodepoort station, meaning that the learned residual correction may still reflect local Highveld cloud, aerosol, humidity and meteorological patterns. In addition, the limited number of external SAURAN stations restricts broad climatic generalisation, while sensor calibration, radiometer levelling, cleaning, maintenance, quality control, data latency, temporal alignment and spatial representativeness may influence the reported errors. Future work should therefore extend the framework to more stations and climatic regions, strengthen radiometer calibration and maintenance documentation, include higher resolution cloud and aerosol observations, and evaluate adaptive recalibration strategies for long term operational deployment.

Declaration

AI assistance was limited to grammar editing using ChatGPT. All research analysis and content development were conducted independently by the authors.

Author Contributions

Kudzanayi Chiteka: Conceptualization, Original draft, Methodology, Investigation, Software, Formal analysis. **CC Enweremadu:** Supervision, Validation, Review & editing, Formal analysis, Resources.

Funding

This research received no external funding.

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author.

References

Al-Ansary H., Al-Shehri A., Alaqel S., Djajadiwinata E., Saeed R. S., Saleh N. S., El-Leathy A., Danish S., Al-Suhaibani Z. "Economic and environmental assessment of a concentrated solar power system using sand as a heat transfer medium", Green Technology & Innovation, 1, pp. 1–10, 2025. https://dx.doi.org/10.36922/GTI025290009

Assaf A. M., Haron H., Abdull Hamed H. N., Ghaleb F. A., Qasem S. N., Albarrak A. M. "A Review on Neural Network Based Models for Short Term Solar Irradiance Forecasting", Applied Sciences, 13, pp. 8332, 2023. https://doi.org/10.3390/app13148332 .

Bai M., Fang Z., Tao S., Xiang S., Bian J., Xiang Y., Zhao P., Jin W., Weyn J. A., Dong H., Zhang B., Sun H., Thambiratnam K., Zhang Q., Sun H., Zhang X., Wu Q. "SolarSeer: Ultrafast and accurate 24-hour solar irradiance forecasts outperforming numerical weather prediction across the USA", arXiv, 2025. https://doi.org/10.48550/arXiv.2508.03590.

Bakhsh B. J., Cantino G. "European Hydrogen Valleys as catalysts for systemic decarbonization: A sustainability-oriented innovation and multi-level perspective analysis", Green Technology & Innovation, 1, pp. 1–15, 2025. https://doi.org/10.36922/GTI025330012.

Brahma B., Wadhvani R. "Solar Irradiance Forecasting Based on Deep Learning Methodologies and Multi-Site Data", Symmetry, 12, pp. 1830, 2020. https://doi.org/10.3390/sym12111830.

Chandel P., Roy L. "Solar Radiation Prediction Based on Hybrid Machine Learning Technique". In Proceedings of the 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 418–24, 2023. https://doi.org/10.1109/ICTACS59847.2023.10390492.

Chauvin R., Nou J., Eynard J., Thil S., Grieu S. "A new approach to the real-time assessment and intraday forecasting of clear-sky direct normal irradiance", Solar Energy, 167, pp. 35–51, 2018. https://doi.org/10.1016/j.solener.2018.02.027.
Fernández-Peruchena C. M., Gastón M., Schroedter-Homscheidt M., Marco I. M., Casado-Rubio J. L., García-Moya J. A. "Increasing the temporal resolution of direct normal solar irradiance forecasted series", AIP Conference Proceedings, 1850, pp. 140007, 2017. https://doi.org/10. 1063/1.4984515.
Gueymard C. A., Ruiz-Arias J. A. "Validation of direct normal irradiance predictions under arid conditions: A review of radiative models and their turbidity-dependent performance", Renewable and Sustainable Energy Reviews, 45, pp. 379–96, 2015. https://doi.org/10.1016/j.rser.2015. 01.065.
Ineichen P., Perez R. "A new airmass independent formulation for the Linke turbidity coefficient", Solar Energy, 73, pp. 151–57, 2002. https://doi.org/10.1016/S0038-092X(02)00045-2.
Kasten F., Young A. T. "Revised optical air mass tables and approximation formula", Applied Optics, 28, pp. 4735–38, 1989. https://doi.org/10.1364/AO.28.004735.
Kreuwel F. P. M., Knap W., Schmeits M., Vilà-Guerau de Arellano J., van Heerwaarden C. C. "Forecasting day-ahead 1-minute irradiance variability from numerical weather predictions", Solar Energy, 258, pp. 57–71, 2023. https://doi.org/10.1016/j.solener.2023.04.050.
Langella R., Proto D., Testa A. "Solar Radiation Forecasting, Accounting for Daily Variability", Energies, 9, pp. 200, 2016. https://doi.org/10.3390/en9030200.
Mandal P., Reddy K. S. "Techno-economic assessment of solar power tower and solar parabolic dish system in India", GD, 1, pp. 1–17, 2025. https://doi.org/10.17184/eac.9496.
Mathiesen P., Collier C., Kleissl J. "A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting", Solar Energy, 92, pp. 47–61, 2013. https://doi.org/ 10.1016/j.solener.2013.02.018.
Mercier T. M., Rahman T., Sabet A. "Solar Irradiance Anticipative Transformer". In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 2065–74, 2023. https://doi.org/10.1109/CVPRW59228.2023.00200.
Nallakaruppan M. K., Johnson J., Mapari S., Dasgupta A., Udgir S., Narmatha C. "Reliable and efficient solar radiation estimation with the insights of XAI", Scientific Reports, 16, pp. 3549, 2026. https://doi.org/10.1038/s41598-025-33604-4.
Odejobi O. A., Alawode K. O., Lawal M. O. "Efficient Method For Forecasting Solar Irradiance - A Review", Fudma Journal of Sciences, 8, pp. 285–98, 2024. https://doi.org/10.33003/fjs-2024-0806-2786.
Paletta Q., Terrén-Serrano G., Nie Y., Li B., Bieker J., Zhang W., Dubus L., Dev S., Feng C. "Advances in solar forecasting: Computer vision with deep learning", Advances in Applied Energy, 11, pp. 100150, 2023. https://doi.org/10.1016/j.adapen.2023.100150.
Papachristopoulou K., Fountoulakis I., Bais A. F., Psiloglou B. E., Papadimitriou N., Raptis I.-P., Kazantzidis A., Kontoes C., Hatzaki M., Kazadzis S. "Effects of clouds and aerosols on downwelling surface solar irradiance nowcasting and short-term forecasting", Atmospheric Measurement Techniques, 17, pp. 1851–77, 2024. https://doi.org/10.5194/amt-17-1851-2024.
Peiró-Signes A., García-Díaz J. C., Peiró-Signes A. "A Comparative Study of Statistical and Machine Learning Methods for Solar Irradiance Forecasting Using the Folsom PLC Dataset", Preprints, 2025. https://doi.org/10.20944/preprints202506.0827.v1.
Pereira S., Canhoto P., Salgado R. "Development and assessment of artificial neural network models for direct normal solar irradiance forecasting using operational numerical weather prediction data", Energy and AI, 15, pp. 100314, 2024. https://doi.org/10.1016/j.egyai.2023.100314.
Polo J., Zarzalejo L. F., Martín L., Navarro A. A., Marchante R. "Estimation of daily Linke turbidity factor by using global irradiance measurements at solar noon", Solar Energy, 83, pp. 1177–85, 2009. https://doi.org/10.1016/j.solener.2009.01.018.
Razak T. R., Alaoui A. E., Jarimi H., Ul-Saufie A. Z., Ismail M. H., Fuzi M. F. M. "An AI-Based Hybrid Prophet–LSTM Model for Forecasting and Financial Optimisation in Sustainable Energy Grids", AIFSC, 1, pp. 42–60, 2026. https://doi.org/10.65582/aifsc.2026.004.
Razak T. R., Ismail M. H., Darus M. Y., Jarimi H., Su Y. "Artificial Intelligence in Renewable Energy: A Systematic Review of Trends in Solar, Wind, and Smart Grid Applications", Research and Reviews in Sustainability, 1, pp. 1–22, 2025. https://doi.org/10.5334/rss.6.
Richardson W., Cañadillas D., Moncada A., Guerrero-Lemus R., Shephard L., Vega-Avila R., Krishnaswami H. "Validation of All-Sky Imager Technology and Solar Irradiance Forecasting at Three Locations: NREL, San Antonio, Texas, and the Canary Islands, Spain", Applied Sciences, 9, pp. 684, 2019. https://doi.org/10.3390/app9040684.

Riffat J., Reza Samaei S. “Recent Advances in Perovskite–Silicon Tandem Solar Cells: Progress, Challenges, and Pathways to Commercialisation”, EC, 1, pp. 113–26, 2025. https://doi.org/10.61552/EC.2025.009.

Salman D., Direkoglu C., Kusaf M., Fahrioglu M. “Hybrid deep learning models for time series forecasting of solar power”, Neural Computing and Applications, 36, pp. 9095–112, 2024. https://doi.org/10.1007/s00521-024-09558-5.

Sasaki J., Okada M., Utsunomiya K., Yamaguchi K. “A Physics-Informed Machine Learning Approach utilizing Multiband Satellite Data for Solar Irradiance Estimation”, arXiv, 2024. https://doi.org/10.48550/arXiv.2407.04283.

Siddiqa A., Rana N., Khan W. Z., Jeribi F., Tahir A. “Explaining solar forecasts with generative AI: A two-stage framework combining transformers and LLMs”, PLoS One, 20, pp. e0331516, 2025. https://doi.org/10.1371/journal.pone.0331516.

Sultana N., Tsutsumida N. “Combining Machine Learning with Physics-Based Mathematical Models for Near-Real-Time Clear-Sky Irradiance Prediction from Geostationary Satellite Imagery”. In Proceedings of the IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, pp. 7472–77, 2024. https://doi.org/10.1109/IGARSS53475.2024.10642113.

Thaker J., Höller R. “A Comparative Study of Time Series Forecasting of Solar Energy Based on Irradiance Classification”, Energies, 15, pp. 2837, 2022. https://doi.org/10.3390/en15082837.

Yan K., Shen H., Wang L., Zhou H., Xu M., Mo Y. “Short-Term Solar Irradiance Forecasting Based on a Hybrid Deep Learning Methodology”, Information, 11, pp. 32, 2020. https://doi.org/10.3390/info11010032.

Yang D. “Choice of clear-sky model in solar forecasting”, Journal of Renewable and Sustainable Energy, 12, pp. 026101, 2020. https://doi.org/10.1063/5.0003495.

Zhuang Z., Wang H., Yu C. “Short-Term Irradiance Prediction Based on Transformer with Inverted Functional Area Structure”, Mathematics, 12, pp. 3213, 2024. https://doi.org/10.3390/math12203213.