IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 June 2026, accepted 9 July 2026, date of publication 13 July 2026, date of current version 20 July 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3712689

RESEARCH ARTICLE

Physics-Guided Residual Learning for Short-Term Solar Irradiance Forecasting

TAREQ ANWAR SHIKDAR[1] **AND HANNU LAAKSONEN**[2], **(Member, IEEE)**
[1]School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland
[2]School of Technology and Innovations, Flexible Energy Resources/Electrical Engineering, University of Vaasa, 65200 Vaasa, Finland

Corresponding author: Hannu Laaksonen (hannu.laaksonen@uwasa.fi)

ABSTRACT Accurate short-term solar irradiance forecasting is essential for reliable operation of renewable-dominated energy systems, where scheduling, flexibility management, and grid operation depend on prediction stability. Solar irradiance prediction remains challenging because it is influenced by both deterministic solar radiation behavior and uncertain atmospheric variations. Existing physical and data-driven approaches often struggle to represent both regular irradiance patterns and rapid weather-related fluctuations simultaneously. This paper proposes a physics-guided residual learning framework that separates these components through a sequential forecasting strategy. A deep temporal learning model first captures regular radiative behavior using meteorological observations and solar geometry features, while a secondary learning stage models the remaining structured forecasting errors. The final prediction combines the initial forecast with the learned residual correction. The proposed method is evaluated on a long-term, high-latitude solar dataset with chronologically separated data. The results demonstrate improved forecasting reliability across seasonal and weather-dependent conditions, providing a practical approach for renewable energy operation and digital energy applications.

INDEX TERMS Artificial intelligence, deep learning, energy management, forecasting, machine learning, neural networks, photovoltaic systems, renewable energy sources, smart grids, solar energy.

NOMENCLATURE

Symbol	Description.
I_t	Measured solar irradiance at time step t.
$\hat{I}_{base,t}$	Initial irradiance prediction generated by the primary forecasting model.
$\hat{I}_{final,t}$	Final corrected irradiance forecast.
e_t	Actual residual error between measurement and base prediction.
$\hat{e}_t$	Estimated residual correction generated by the residual learner.
X_t	Multivariate input feature vector at time step t.
t	Time index.
N	Number of samples.
L	Model loss function.
θ	Trainable model parameters.

The associate editor coordinating the review of this manuscript and approving it for publication was Gerard-Andre Capolino.

I. INTRODUCTION

The increasing integration of photovoltaic generation is transforming modern power systems toward renewable-dominated operation. As renewable penetration increases, accurate short-term solar irradiance forecasting becomes increasingly important for reliable grid operation, energy scheduling, reserve management, and flexibility coordination. Forecast uncertainty directly affects operational decisions such as storage dispatch, electricity market participation, and balancing actions, especially in power systems where conventional flexibility resources are gradually reduced [1], [2], [3], [4], [5], [6].

Solar irradiance forecasting remains challenging because the observed irradiance signal results from multiple interacting physical processes. A significant part of irradiance variation follows deterministic patterns associated with solar geometry, seasonal cycles, and predictable radiative behavior [7], [8], [9]. However, real operating conditions introduce additional uncertainty caused by atmospheric variations, cloud movement, and local meteorological changes, resulting in rapid fluctuations and irradiance ramps [10], [11], [12]. Traditional physical models can effectively represent deterministic solar behavior but often exhibit reduced accuracy under complex atmospheric conditions. Conversely, statistical and machine-learning-based approaches improve nonlinear predictive capability but may struggle to maintain robustness under changing weather conditions [13], [14], [15], [16].

Deep learning methods have recently shown strong potential for renewable energy forecasting because of their ability to model complex nonlinear temporal relationships. Recurrent neural networks, convolutional architectures, and hybrid temporal models have demonstrated improved forecasting accuracy compared with conventional statistical approaches [17], [18], [19], [20], [21]. More recently, attention- and transformer-based architectures have been introduced to capture long-range dependencies in energy time-series forecasting [22], [23], [24]. Although these methods improve overall prediction accuracy, a single forecasting model must generally learn both regular solar patterns and irregular atmospheric variations simultaneously. This can lead to smoothing effects and reduced ability to represent sudden irradiance changes under highly variable weather conditions [11], [20], [25].

To address these limitations, hybrid physical and data-driven forecasting approaches have been widely investigated. Existing studies have incorporated solar geometry information, clear-sky indicators, numerical weather information, and meteorological variables into learning models to improve generalization capability [23], [26], [27], [28]. Furthermore, residual correction, ensemble learning, boosting strategies, and stacked forecasting architectures have been explored to reduce systematic prediction errors and improve forecasting performance [29], [30], [31]. These studies demonstrate that secondary learning stages can effectively compensate for the limitations of individual forecasting models.

However, most existing hybrid forecasting approaches mainly focus on feature-level integration, model combination, or statistical error reduction. Residual components are commonly treated as numerical correction terms to improve accuracy, while less attention has been paid to organizing the forecasting process according to different sources of irradiance variability. Therefore, further investigation is required into structured residual modeling strategies that combine physical understanding of solar behavior with data-driven correction while maintaining computational efficiency.

This work does not aim to introduce a fundamentally new machine learning algorithm. Instead, the contribution lies in developing a physics-guided residual learning framework that reorganizes existing learning components through task decomposition. The proposed framework separates short-term irradiance forecasting into two complementary stages. The first stage uses a temporal deep learning model to represent the regular evolution of irradiance associated with solar geometry and meteorological patterns. The second stage applies ensemble-based residual learning to model remaining structured forecasting errors that are difficult to capture using the primary model alone.

Compared with conventional hybrid or residual-correction approaches, the proposed framework emphasizes functional separation between the primary predictor and the residual learning stage. The residual component is considered to potentially contain useful information rather than only random error. Its relationship with weather-dependent variability is investigated through statistical evaluation, clearness-index-based analysis, seasonal assessment, and feature importance interpretation.

The proposed framework is validated using long-term high-latitude solar observations with chronological data separation to ensure realistic evaluation conditions. The main contributions of this study are summarized as follows:

- A physics-guided residual learning framework is developed to separate regular irradiance forecasting and remaining error correction through a sequential learning structure.
- A hybrid forecasting strategy combining temporal deep learning and ensemble-based residual modeling is introduced to improve short-term prediction reliability without increasing primary model complexity.
- Residual behavior is analyzed using statistical evaluation, weather-regime assessment, and feature importance analysis to investigate remaining forecasting patterns.
- Comprehensive evaluation across seasonal conditions and clearness-index regimes is performed to assess forecasting robustness for renewable energy applications.

Overall, this study presents a structured residual learning approach that combines a physical understanding of solar irradiance variability with data-driven forecasting methods. The proposed framework provides an interpretable and computationally practical solution for improving short-term solar forecasting reliability in renewable energy systems. It should be emphasized that this work does not aim to introduce a fundamentally new machine learning algorithm. Instead, the contribution lies in organizing existing learning components through a physics-guided task decomposition strategy that separates regular irradiance prediction and residual correction.

II. LITERATURE REVIEW

This section reviews prior research relevant to sector-coupled volatility in campus microgrids, multi-horizon forecasting techniques, hybrid battery energy storage systems, and multi-objective optimization approaches. The discussion highlights methodological limitations in existing studies and establishes the motivation for the integrated framework proposed in this paper.

A. PHYSICAL SOLAR IRRADIANCE MODELS

Initial approaches to solar irradiance forecasting relied primarily on deterministic physical formulations derived from solar geometry and atmospheric radiative transfer theory. Clear-sky and transposition models estimate irradiance using astronomical relationships and optical air-mass assumptions, while numerical weather prediction (NWP) models extend forecasting capabilities through meteorological simulation [32], [33], [34]. These methods are effective for representing long-term trends and seasonal variability because solar position and extraterrestrial radiation follow predictable periodic behavior.

Nevertheless, physical models experience significant degradation under real atmospheric conditions. Local cloud motion and rapidly evolving atmospheric states introduce high-frequency variability that cannot be resolved by deterministic formulations [35]. Consequently, forecast bias and ramp prediction errors remain substantial, particularly in high-latitude climates characterized by frequent cloud transitions and strong seasonal contrasts. For operational short-term forecasting, deterministic approaches alone are therefore insufficient.

B. CONVENTIONAL MACHINE LEARNING METHODS

Machine learning models such as support vector regression, random forests, and gradient boosting have been widely applied to solar irradiance forecasting. These methods improve predictive accuracy compared to purely physical approaches by capturing nonlinear relationships between meteorological variables and irradiance observations.

However, conventional machine learning models typically rely on limited temporal context and often struggle to generalize across diverse atmospheric conditions. In particular, their ability to capture rapid irradiance fluctuations caused by cloud dynamics is limited, reducing forecasting reliability in highly variable environments.

C. DEEP LEARNING FORECASTING

Deep learning has recently become the dominant paradigm for renewable energy forecasting due to its ability to learn nonlinear temporal patterns. Recurrent neural networks, particularly long short-term memory (LSTM) and gated recurrent unit (GRU) architectures, capture temporal dependencies in irradiance time series [40], [41], [42]. Convolutional neural networks (CNN) extract short-term temporal features, and hybrid CNN–LSTM structures combine feature extraction with sequential memory to improve predictive accuracy [43]. More recent developments employ attention mechanisms and transformer-based architectures to capture long-range dependencies and improve generalization [44]. Although these approaches significantly reduce average prediction error, they frequently produce smooth forecasts and underestimate rapid irradiance caused by moving clouds [45]. Consequently, deep learning models improve statistical metrics but still exhibit reliability limitations under stochastic atmospheric variability. Recent advancements in solar forecasting have increasingly leveraged transformer-based architectures due to their superior ability to model long-range temporal dependencies in complex energy time series. For instance, multi-scale spatio-temporal graph fusion frameworks have demonstrated strong capability in capturing spatial–temporal correlations in photovoltaic generation [25], while transformer-based hybrid models utilizing shifted-window attention mechanisms have further improved forecasting accuracy under dynamic atmospheric conditions [28]. Informer architecture, in particular, has been widely adopted for efficient long-sequence time-series forecasting, addressing the computational limitations of conventional transformers [23]. Comprehensive reviews also confirm the growing dominance of transformer-based approaches in renewable energy forecasting, highlighting their robustness and scalability across diverse applications [46].

In parallel, hybrid forecasting frameworks integrating deep learning with ensemble and boosting techniques have shown significant improvements in predictive performance and reliability. These approaches combine the nonlinear feature-extraction capabilities of deep neural networks with the error-correction strength of ensemble learners, enabling better handling of stochastic atmospheric variability [10], [14]. Such hybrid methodologies are particularly effective under highly variable weather conditions, where single-model approaches often fail to capture abrupt fluctuations in irradiance.

D. HYBRID AND PHYSICS-INFORMED APPROACHES

To overcome the limitations of purely data-driven models, hybrid physical–AI approaches have been investigated. Physics-guided residual learning neural networks incorporate solar geometry and atmospheric indicators to enhance generalization, while combined NWP–machine learning frameworks aim to reduce systematic bias [47], [48], [49]. These methods demonstrate improved robustness compared with standalone deep learning models. However, existing hybrid methods typically integrate physical knowledge only as additional input features. The remaining prediction error is usually treated as random noise. In practice, forecasting residuals contain structured patterns associated with unresolved atmospheric processes such as cloud intermittency and local turbulence [50]. Current literature rarely explicitly learns this residual component, leaving potentially useful information unexploited.

E. RESIDUAL LEARNING, ENSEMBLE, AND ADVANCED HYBRID FORECASTING APPROACHES

Recent developments in renewable energy forecasting have increasingly focused on advanced hybrid architectures, ensemble learning strategies, and attention-based models to improve prediction reliability under variable atmospheric conditions. Transformer-based architectures have attracted significant interest for their ability to capture long-range temporal dependencies in solar forecasting applications. For example, Mercier et al. proposed a Solar Irradiance Anticipative Transformer framework, demonstrating the ability of attention mechanisms to extract complex temporal irradiance patterns [51]. Similarly, comparative studies involving transformer networks, long short-term memory models, and Light Gradient Boosting Machine approaches have highlighted the importance of evaluating different learning architectures under uncertain and noisy conditions for solar forecasting [52].

Alongside advances in deep learning, ensemble-based machine learning methods have demonstrated strong performance in solar energy forecasting. Hanif et al. investigated solar forecasting using modified artificial neural networks and Light Gradient Boosting Machine models, showing the effectiveness of gradient boosting methods in capturing nonlinear meteorological relationships [53]. Furthermore, stacked ensemble strategies that combine Gradient Boosting, Extreme Gradient Boosting, and meta-learning approaches have been proposed to improve the accuracy of short-term solar power forecasting in smart grid applications [54].

Hybrid approaches that combine physical knowledge with data-driven models have also received increasing attention. Physics-guided residual learning ensemble forecasting frameworks integrate domain knowledge with machine learning methods to improve the robustness and interpretability of predictions [55]. Moreover, recent studies that combine deep learning with ensemble models, including recurrent networks and tree-based approaches, have demonstrated that complementary learning strategies can improve solar irradiance prediction compared with individual models [56].

Although these studies confirm the effectiveness of transformer models, ensemble learning, and hybrid forecasting frameworks, most existing approaches primarily focus on improving prediction accuracy by increasing model complexity, integrating features, or combining statistical models. Residual correction is typically treated as a post-processing step for reducing numerical forecasting errors. In contrast, the present work focuses on a physics-guided residual learning structure in which the forecasting task is separated into regular irradiance prediction and remaining-error correction stages. The objective is not to replace existing forecasting algorithms but to improve reliability by structuring complementary learning components.

F. RESEARCH GAP AND LIMITATIONS OF EXISTING METHODS

From the reviewed literature, solar irradiance forecasting approaches can generally be categorized into four groups: deterministic physical models, statistical learning methods, deep learning architectures, and hybrid physical–data-driven frameworks. Deterministic models based on solar geometry and radiative transfer theory effectively capture periodic radiative behavior and seasonal solar motion. However, they perform poorly in the presence of stochastic atmospheric variability because rapid cloud dynamics and local meteorological disturbances cannot be fully captured by deterministic formulations.

Data-driven machine learning models improve predictive accuracy by learning nonlinear relationships between meteorological variables and irradiance observations. Nevertheless, classical machine learning approaches often struggle to generalize across diverse weather regimes and rely on short-term persistence patterns, which limits their robustness during rapid atmospheric transitions.

Recently deep learning architectures have demonstrated superior ability to model nonlinear temporal dependencies in irradiance time series. However, these models frequently produce overly smooth forecasts and tend to underestimate abrupt irradiance ramps caused by cloud movement. This smoothing effect arises because a single deep network attempts to simultaneously represent both deterministic solar radiative behavior and stochastic atmospheric variability within a single mapping.

Hybrid forecasting frameworks partially address these limitations by integrating physical descriptors such as solar geometry or clear-sky indices into data-driven predictors. Although these approaches improve generalization, they typically treat the remaining prediction error as random noise to be minimized rather than as an informative signal. In practice, forecasting residuals often contain structured patterns associated with unresolved atmospheric dynamics, including cloud intermittency and rapid irradiance ramps. However, most existing studies do not explicitly model this residual component or exploit its underlying physical meaning.

Consequently, several important challenges remain unresolved in current solar irradiance forecasting research:

1) Lack of explicit separation between deterministic solar radiative behavior and stochastic atmospheric variability.
2) Loss of high-frequency ramp information due to smoothing behavior in deep learning predictors.
3) Limited robustness across diverse atmospheric regimes and seasonal conditions.
4) Underutilization of structured information contained in forecast residuals, which are typically treated as random noise.

These limitations reduce the reliability of forecasts in operational energy systems, where accurate ramp prediction is critical. Applications such as battery dispatch, reserve activation, and grid balancing require forecasts that not only minimize average error but also capture rapid irradiance fluctuations caused by atmospheric variability. Addressing these challenges requires forecasting frameworks that explicitly distinguish between deterministic solar dynamics and stochastic atmospheric disturbances while effectively utilizing structured information present in prediction residuals.

G. NOVELTY OF THIS WORK

The primary novelty of this study does not lie in introducing a new standalone learning algorithm, but in redefining the role of residual learning within solar irradiance forecasting. Conventional hybrid and boosting-based approaches treat residuals purely as statistical artifacts to be minimized iteratively, without considering their physical origin. As a result, residual modeling in existing frameworks remains an optimization-driven process rather than a physically interpretable learning task.

In contrast, this work introduces a physics-aligned task decomposition paradigm, in which the forecasting problem is explicitly separated into two complementary components based on the underlying generation mechanisms of solar irradiance. The deep temporal network learns deterministic radiative behavior governed by solar geometry and smooth atmospheric evolution, while the residual learning stage is reformulated into a dedicated atmospheric variability inference module that captures structured stochastic disturbances. This conceptual shift transforms residual learning from a generic error-correction mechanism into a structured secondary learning problem, in which the residual is no longer treated as noise but as a physically meaningful signal associated with unresolved atmospheric processes, such as cloud intermittency and irradiance ramps. Unlike conventional boosting or stacking frameworks, where multiple learners attempt to approximate the same target function, the proposed approach enforces functional specialization between model components, ensuring that each learning stage focuses on a distinct physical aspect of the irradiance generation process. This separation reduces representational conflict within a single model and enables improved capture of high-frequency variability without increasing primary model complexity. Therefore, the novelty of this work lies in:

- Introducing a physics-guided decomposition of forecasting tasks rather than feature-level hybridization
- Reformulating residual learning as a structured atmospheric process modeling problem
- Establishing a sequential specialization framework that aligns machine learning architecture with physical irradiance dynamics

This paradigm provides a new perspective on hybrid forecasting design, in which model architecture is guided not only by predictive performance but also by the physical structure of the underlying system.

III. METHODOLOGY

Building upon the limitations identified in Section II, the proposed framework introduces a residual-corrected hybrid architecture that separates deterministic and stochastic irradiance components. Figure 1 illustrates the architecture of the proposed hybrid residual-learning forecasting framework. Unlike conventional boosting or stacking frameworks that sequentially minimize prediction error using homogeneous learners, the proposed approach is motivated by the physical structure of solar irradiance variability. Multivariate time-series input features, including meteorological variables, solar geometry descriptors, and historical irradiance observations, are first processed by a deep temporal forecasting model based on a CNN–BiLSTM network [40], [42]. This stage primarily captures deterministic radiative dynamics governed by solar geometry and periodic atmospheric behavior and produces an initial irradiance estimate $\hat{I}_t^{DL}$.

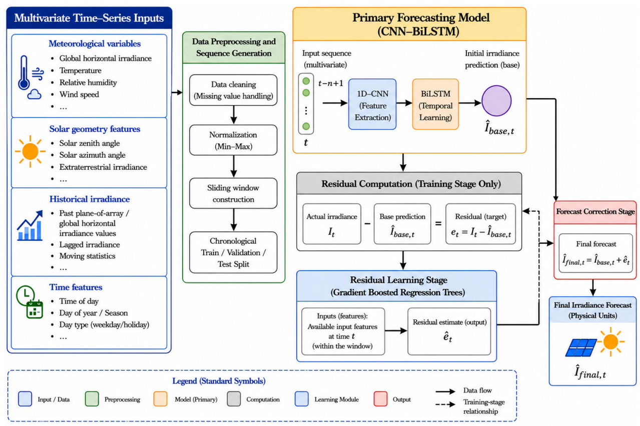


FIGURE 1. **Proposed physics-guided residual learning framework for short-term solar irradiance forecasting with sequential deep learning and residual correction.**

The prediction error is then computed as the difference between the measured irradiance and the deep model output, $e_t = I_t - \hat{I}_t^{DL}$. Instead of treating this residual as random noise, it is provided to a secondary learning stage consisting of gradient-boosted regression trees [57]. This residual learning model captures remaining structured forecasting errors that are associated with weather-dependent irradiance variations and rapid irradiance changes and produces a residual estimate $\hat{e}_t$.

Finally, a hybrid combination layer integrates the deterministic prediction and the learned residual component to generate the final forecast $\hat{I}_t = \hat{I}_t^{DL} + \hat{e}_t$. By separating predictable solar behavior from stochastic atmospheric effects, the framework improves forecasting reliability while maintaining model interpretability.

A. PROBLEM DEFINITION

Let I_t denote the measured plane-of-array (POA) irradiance at time step t. Given a multivariate historical observation window of length L,

$$X_{t-L:t} = \{x_{t-L}, x_{t-L+1}, \ldots, x_t\} \tag{1}$$

The objective is to predict irradiance over a future horizon H:

$$\hat{I}_{t+1:t+H} = F\left(X_{t-L:t}\right) \tag{2}$$

Solar irradiance can be interpreted as the superposition of a deterministic radiative component and a stochastic atmospheric component [32], [33], [34], [35]:

$$I_t = I_t^{det} + I_t^{sto} \tag{3}$$

The deterministic term is governed by solar geometry and periodic meteorological behavior, whereas the stochastic term arises from cloud dynamics and unresolved atmospheric disturbances.

Instead of approximating I_t Using a single mapping, the proposed method decomposes the prediction task into two sequential stages:

$$\hat{I}_t = \hat{I}_t^{DL} + \hat{\epsilon}_t \tag{4}$$

where $\hat{I}_t^{DL}$ is the deep temporal prediction and $\hat{\epsilon}_t$ is a learned residual correction representing stochastic variability

B. DEEP TEMPORAL PREDICTION MODEL

The first stage estimates the smooth deterministic component of irradiance using a hybrid convolutional–recurrent architecture. The convolutional layer extracts short-term temporal patterns from the input sequence [42]:

$$Z_t = \text{Conv1D}\left(X_{t-L:t}\right) \tag{5}$$

This representation is then processed by a bidirectional long short-term memory network (BiLSTM) [40], which captures long-range temporal dependencies:

$$H_t = \text{BiLSTM}\left(Z_t\right) \tag{6}$$

The deep model output is obtained through a fully connected mapping:

$$\hat{I}_t^{DL} = WH_t + b \tag{7}$$

This stage learns deterministic periodic solar behavior, including diurnal cycles, seasonal trends, and smooth meteorological variations.

C. RESIDUAL LEARNING

After the deep temporal prediction stage, the remaining forecasting error is defined as

$$\epsilon_t = I_t - \hat{I}_t^{DL} \tag{8}$$

where I_t represents the measured irradiance and $\hat{I}_t^{DL}$ denotes the prediction obtained from the CNN–BiLSTM deep learning model.

In conventional forecasting approaches, this error is often treated as random noise [23]. However, due to atmospheric variability, particularly during rapid cloud transitions, the residual signal often contains structured temporal patterns. To capture this remaining information, a residual prediction module is introduced. In this study, the residual component is modeled using a gradient boosting regression ensemble based on decision trees, which learns the systematic prediction errors of the deep temporal model. The residual estimator is expressed as

$$\hat{\epsilon}_t = G(Z_t^{res}) \tag{9}$$

where $G(\cdot)$ denotes the gradient boosting regression model and Z_t^{res} represents the feature set used by the residual learner. The input features include:

- physical irradiance estimate (POA_{PHYS})
- estimated PV cell temperature ($TCELL_{PHYS}$)
- clearness index (KT)
- ambient temperature ($T2M$)
- wind speed ($WS10M$)
- lagged irradiance features (24-hour persistence)
- temporal encodings representing hour-of-day and day-of-year.

These predictors provide information about short-term atmospheric evolution responsible for rapid irradiance fluctuations.

The final hybrid irradiance forecast is obtained by combining the deep learning prediction with the residual correction term

$$\hat{I}_t = \hat{I}_t^{DL} + \alpha\hat{\epsilon}_t \tag{10}$$

where α is a correction coefficient controlling the influence of the residual learner on the final prediction. Rather than fixing this parameter arbitrarily, α is determined through validation-based tuning. A grid search over $\alpha \in [0, 1]$ with increments of 0.05 is performed on the validation dataset, and the value that minimizes validation RMSE is selected. The optimal value obtained in this study is $\alpha = 0.15$.

The hybrid forecasting framework is trained using a two-stage sequential procedure. First, the CNN–BiLSTM model is trained on the training dataset to produce baseline irradiance predictions. Second, the residual errors between the predicted and measured values are computed and used as targets to train the gradient-boosted residual model. The final hybrid prediction is then obtained by applying the residual correction term as given in (10).

To illustrate the practical behavior of the proposed formulation, the temporal evolution of irradiance prediction is shown in Fig. 2 for a representative mixed-cloud day. The black curve represents the measured irradiance, the blue dashed curve corresponds to the CNN–BiLSTM prediction, and the red solid curve denotes the final hybrid forecast. While the deep learning model captures the overall diurnal pattern, it tends to smooth rapid irradiance variations and underestimate peak values around noon. After applying the residual correction, the hybrid prediction closely follows the measured signal and accurately captures abrupt irradiance ramps due to cloud transitions. This observation confirms that the residual learning stage models structured atmospheric disturbances rather than random noise, supporting the decomposition framework expressed in (4)–(10).

To improve reproducibility and provide a clear overview of the experimental configuration, Table 1 summarizes the main implementation details of the proposed residual-corrected forecasting framework, including input preparation, preprocessing, primary forecasting model, residual learning stage, and evaluation strategy.

D. LEAKAGE PREVENTION AND RESIDUAL TARGET GENERATION

Because the proposed framework introduces a secondary residual learning stage, careful attention was given to preventing information leakage between training and evaluation

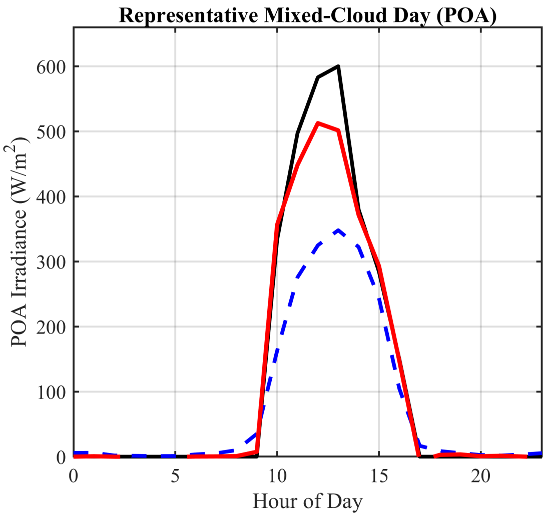


FIGURE 2. Time-series comparison of measured irradiance, baseline CNN–BiLSTM prediction, and residual-corrected hybrid forecast for a representative mixed-cloud day.

TABLE 1. Summary of model training and configuration parameters.

Component	Configuration
Input Data	Multivariate meteorological, solar geometry, and historical irradiance features
Data Split	Chronological train–validation–test separation
Preprocessing	Missing value handling, normalization, and sliding-window sequence generation
Primary Model	CNN–BiLSTM temporal forecasting model
Residual Model	Gradient-boosted regression tree learner
Residual Target	Difference between measured irradiance and base prediction
Optimization	Validation-based hyperparameter tuning
Implementation Environment	MATLAB
Evaluation	Seasonal, clearness-index, residual, and statistical analysis

periods. The complete dataset was separated chronologically into training, validation, and testing subsets without random shuffling to preserve the temporal nature of solar irradiance forecasting.

The primary forecasting model was first trained independently using the training dataset. After obtaining the baseline predictions, the residual target was calculated as:

$$e_t = y_t - \hat{y}_{base,t} \tag{11}$$

where y_t represents the measured irradiance value and $\hat{y}_{base,t}$ represents the corresponding prediction generated by the primary forecasting model at the time step t. The calculated residual sequence was then used as the target output for training the residual learner.

Note that the residual correction stage does not use future irradiance measurements during prediction. During validation and testing, the primary model first generates the baseline irradiance forecast, and the residual learner estimates the correction using only available input information. The measured irradiance values from the validation and testing periods are used exclusively to calculate evaluation metrics after the prediction process is complete.

This chronological separation ensures that the reported improvement originates from the residual learning process rather than temporal leakage between training and testing data. Future work will further extend this evaluation through rolling-origin validation and external multi-year datasets to assess broader generalization capability.

E. HYBRID LEARNING INTERPRETATION

The proposed framework decomposes the forecasting problem into two physically meaningful learning tasks:

- The deep temporal model captures deterministic radiative behavior governed by periodic solar motion and smooth meteorological evolution.
- The residual model learns remaining structured prediction errors that are statistically associated with abrupt irradiance variations.

Unlike conventional hybrid approaches that simply combine heterogeneous features within a single predictor, the proposed method adopts a sequential specialization strategy. The deep temporal model first learns deterministic solar radiative behavior governed by solar geometry and smooth meteorological evolution. The second learning stage does not attempt to relearn the entire irradiance signal; instead, it focuses exclusively on structured prediction error, enabling correction of systematic deviations without increasing the complexity of the deep model. Furthermore, unlike conventional residual-learning pipelines that treat prediction errors as purely statistical artifacts, the proposed framework investigates whether the residual component contains structured information related to unresolved weather-dependent variability rather than treating all remaining errors as random noise. This physically motivated decomposition aligns the machine learning architecture with the underlying generation mechanisms of solar irradiance, resulting in improved forecasting reliability across varying weather regimes.

IV. DATASET AND FEATURE ENGINEERING

A. DATA SOURCE AND STUDY PERIOD

The proposed forecasting framework was evaluated using a long-term solar irradiance dataset spanning January 2016 to December 2023. The dataset consists of hourly observations combining satellite-derived atmospheric measurements with solar geometry calculations[11]. The target variable is the plane-of-array (POA) irradiance, which represents the effective solar radiation incident on the surface of a photovoltaic module and is directly relevant to photovoltaic energy production modeling [32].

To ensure realistic forecasting conditions, the dataset was divided chronologically rather than randomly. The earliest observations were used for model training, followed by a validation period used for model tuning, while the most recent portion was reserved exclusively for performance evaluation. This chronological splitting prevents information leakage and reflects practical deployment scenarios in which future data are unavailable during model development. A summary of the dataset characteristics, model/training setting, and partitioning strategy is presented in Table 2.

TABLE 2. Dataset characteristics, Input features, and Model configuration summary.

Category	Item	Description / Variables
Site information	Location	Lappeenranta, Finland
	Coordinates	61.0550° N, 28.1896° E
	Data source	NASA POWER
Temporal properties	Resolution	Hourly
	Study period	Jan 2016 – Dec 2023
	Splitting method	Chronological (Train–Validation–Test)
Target variable	Output	Plane-of-array irradiance (W/m^2), Cell temperature (°C)
Temporal features	Cyclical encoding	Hour (sin, cos), Day-of-year (sin, cos)
Meteorological features	Atmospheric state	T2M, RH2M, WS10M, PS
Radiative features	Solar condition	Clearness index (KT), solar geometry
Historical features	Persistence	Lagged irradiance (24 h)
Derived features	Dynamics	Moving averages (MA3, MA6), slope indicators
Deep learning model	Architecture	Hybrid 1D CNN–BiLSTM
	CNN layers	2 layers (kernel sizes 3, 5; 64 filters)
	Recurrent layer	BiLSTM (96 units)
	Activation	ReLU
	Regularization	Dropout (0.1–0.2), Layer normalization
Training settings	Optimizer	Adam
	Learning rate	3×10^{-3}
	Batch size	256
	Epochs	40
Residual learning	Algorithm	Gradient Boosting (LSBoost)
	Trees	300–400
	Learning rate	0.05
	Input features	Engineered feature table
Hybrid formulation	Final model	$\hat{Y} = \hat{Y}\_DL + \alpha\cdot\hat{\varepsilon}$
	Coefficient α	Validation-tuned (α = 0.15)
	Training strategy	Sequential (deep → residual)

B. INPUT VARIABLES

The input feature set was designed to represent both deterministic solar radiation behavior and stochastic atmospheric variability. Temporal periodicity was encoded using sinusoidal transformations of the hour of day and day of year, enabling the model to learn daily and seasonal cycles without artificial discontinuities at boundary transitions [40]. Meteorological variables describing the atmospheric state were included to capture scattering and absorption effects influencing solar radiation, thereby providing indirect information about cloud formation and atmospheric transparency.

In addition to meteorological observations, radiative indicators derived from solar measurements were incorporated to characterize short-term irradiance dynamics. Historical irradiance observations provide persistent information, while atmospheric descriptors such as the clearness index reflect the deviation between measured and theoretical radiation conditions [33]. Together, these variables allow the model to distinguish predictable solar motion from weather-driven variability. The selected predictors and their categories are summarized in Table 1.

C. FEATURE ENGINEERING

To improve temporal representation, additional predictors were derived from the original measurements. Lagged irradiance observations were introduced to capture the persistence behavior inherent in solar radiation sequences. Moving statistical descriptors were computed to represent short-term atmospheric stability, while first-order temporal differences were included to characterize rapid irradiance ramps associated with cloud transitions. Periodic encodings further ensured continuous representation of diurnal and seasonal cycles [35].

These engineered features guide the deep learning component in modeling smooth, deterministic radiative behavior, while the residual learning stage focuses on structured stochastic deviations arising from atmospheric variability. This separation supports the proposed decomposition of solar dynamics into predictable and weather-driven disturbances.

D. DATA PROCESSING

All input variables were normalized using min–max scaling parameters computed exclusively from the training subset, and the same transformation was applied to the validation and test data to avoid information leakage and maintain consistency. Normalization ensures stable gradient behavior and prevents variables with larger numerical ranges from dominating the learning process. Night-time observations were retained to preserve natural temporal continuity and enable the model to learn zero-irradiance conditions without artificial filtering. After preprocessing, the dataset contained no missing values, providing uninterrupted sequential input for the forecasting model.

E. DATA LEAKAGE VERIFICATION

Given the substantial performance improvement achieved by the proposed hybrid framework, particular attention was paid to preventing data leakage during model development. The dataset was strictly divided using chronological splitting, ensuring that future observations were never used during training or validation.

All input features were constructed using only past and current information available at the prediction time. Lagged irradiance variables, moving averages, and slope-based features were computed exclusively from historical data, without incorporating any future values.

Furthermore, normalization parameters were derived solely from the training dataset and subsequently applied to validation and test sets, preventing statistical leakage across partitions.

The residual learning stage was trained sequentially using prediction errors generated only from the training data, ensuring that no information from the test set influenced model calibration.

These precautions guarantee that the reported performance improvement reflects genuine predictive capability rather than unintended information leakage.

F. SEQUENCE FORMULATION

The forecasting task was formulated as a multivariate sequence-to-sequence prediction problem [42]. A sliding temporal window of past observations was used as input to estimate irradiance over a subsequent forecasting horizon. This formulation allows the model to capture temporal dependencies and evolving atmospheric conditions while preserving the chronological consistency required for real-world deployment. The chosen input window provides sufficient historical context for learning short-term irradiance dynamics and supports operational forecasting applications such as energy scheduling and flexibility management.

V. RESULTS AND PERFORMANCE EVALUATION

A. OVERALL FORECASTING ACCURACY

The predictive performance of the proposed hybrid framework was evaluated against the standalone deep learning model using standard statistical metrics. The numerical results are summarized in Table 2. For plane-of-array (POA) irradiance, the hybrid model reduces RMSE from 0.0809 to 0.0249 and MAE from 0.0404 to 0.0119, while increasing the coefficient of determination from $R^2 = 0.8468$ to 0.9855. The prediction bias is also effectively eliminated (from 0.0050 to −0.0000). These improvements indicate that the residual learning stage corrects systematic deviations produced by the deep temporal model rather than merely smoothing random noise.

It should be noted that the statistical metrics reported in Table 3 correspond to normalized irradiance values used during model training. For visualization and interpretability, the figures in the following subsections present prediction errors in physical irradiance units (W/m^2) after reversing the normalization. The consistency of this improvement is visually confirmed in Fig. 3, which presents the measured versus predicted POA irradiance. The baseline model exhibits considerable dispersion around the one-to-one reference line, particularly at medium and high irradiance levels where atmospheric variability is strongest. After residual correction, the hybrid predictions form a narrow band closely aligned with the ideal line, demonstrating substantially reduced variance and improved agreement across the full irradiance range.

TABLE 3. Forecasting performance comparison between baseline and proposed models.

Variable	Model	RMSE	MAE	Bias	R²
POA Irradiance	Deep Learning (Base)	0.0809	0.0404	0.0050	0.8468
	Hybrid (Proposed)	**0.0249**	**0.0119**	**−0.0000**	**0.9855**
Cell Temperature	Deep Learning (Base)	0.0452	0.0331	0.0099	0.9160
	Hybrid (Proposed)	0.0452	0.0331	0.0099	0.9160

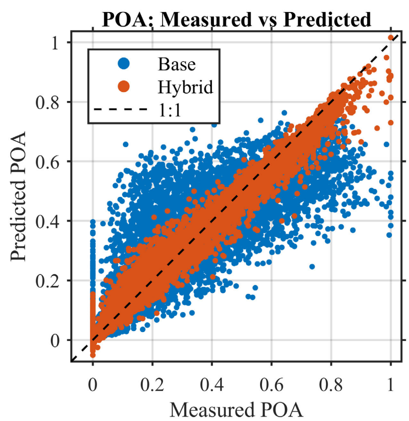


FIGURE 3. Scatter plot comparing measured and predicted POA irradiance for the baseline CNN–BiLSTM and proposed hybrid residual-learning models.

A similar behavior is observed for cell temperature in Fig. 4. While the baseline model already achieves good performance (RMSE = 0.0452, $R^2 = 0.9160$), the hybrid framework maintains this accuracy without degradation, confirming that residual learning improves irradiance estimation without negatively affecting temperature prediction stability.

Overall, the combined quantitative comparison in Table 2 and graphical evidence in Figs. 3-4 indicate that residual correction improves forecasting reliability within the evaluated dataset.

B. COMPARATIVE ANALYSIS WITH BASELINE FORECASTING MODELS

To provide a comprehensive and fair benchmarking context, the proposed framework is compared with representative models from different forecasting paradigms. The persistence model serves as a naïve baseline widely used in solar

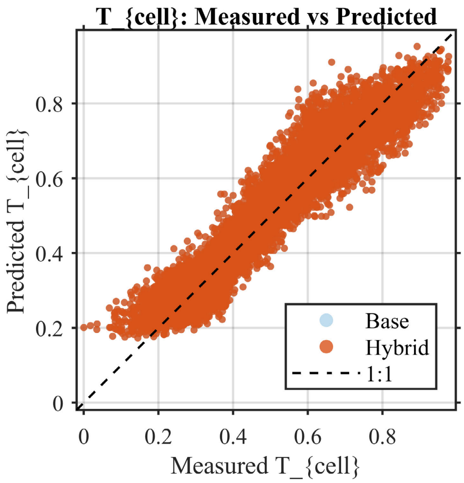


FIGURE 4. Scatter plot comparing measured and predicted photovoltaic cell temperature for the baseline CNN–BiLSTM and proposed hybrid residual-learning models.

forecasting to assess improvement over simple temporal continuity assumptions. Gradient boosting regression is a strong, widely adopted machine learning approach for tabular meteorological data, known for its robustness and competitive performance in energy forecasting tasks. The CNN–BiLSTM model is selected as a standard deep learning baseline due to its ability to capture both local temporal patterns and long-term dependencies in irradiance time series. Together, these baselines provide a balanced comparison across classical, machine learning, and deep learning approaches commonly reported in recent literature. Fig. 5 summarizes the overall forecasting accuracy on the test dataset using RMSE, MAE, and R^2 metrics. The persistence model exhibits the largest prediction error because it fails to capture atmospheric variability. Machine learning and deep learning approaches significantly improve performance, while the proposed hybrid framework achieves the lowest RMSE and highest coefficient of determination.

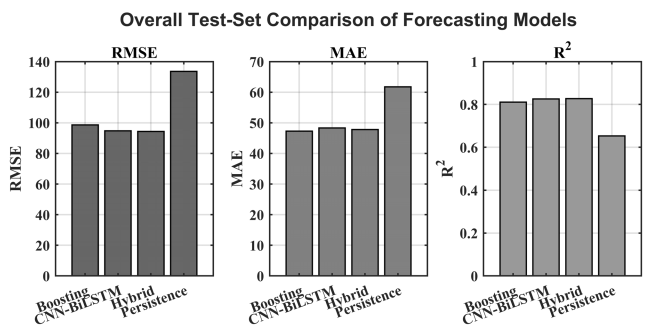


FIGURE 5. Comparison of forecasting model performance on the test dataset using RMSE, MAE, and R^2 evaluation metrics.

Fig. 6 presents predicted versus measured POA irradiance for all forecasting models. The persistence baseline shows substantial dispersion from the ideal one-to-one reference line, particularly at medium and high irradiance levels. The boosting model improves prediction accuracy but still exhibits noticeable variance. The CNN–BiLSTM model produces tighter clustering by capturing temporal dependencies. The hybrid model further reduces dispersion and forms the narrowest prediction band around the reference line, demonstrating improved predictive stability across the entire irradiance range.

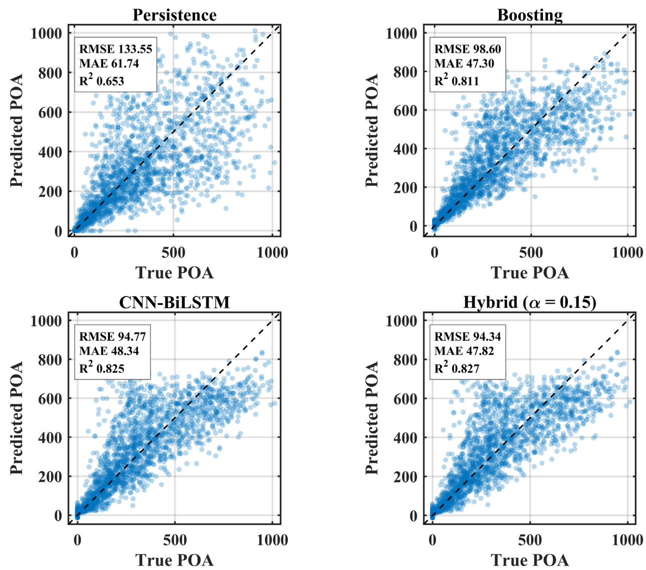


FIGURE 6. Parity plots comparing measured and predicted POA irradiance for the Persistence, Boosting, CNN–BiLSTM, and proposed Hybrid models.

The statistical distribution of forecast errors is shown in Fig. 7. The persistence model yields the widest error distribution, indicating greater forecasting uncertainty. Both boosting and CNN–BiLSTM models significantly reduce error variance. The hybrid framework produces the most concentrated distribution around zero, confirming that the residual learning stage effectively eliminates systematic forecasting deviations.

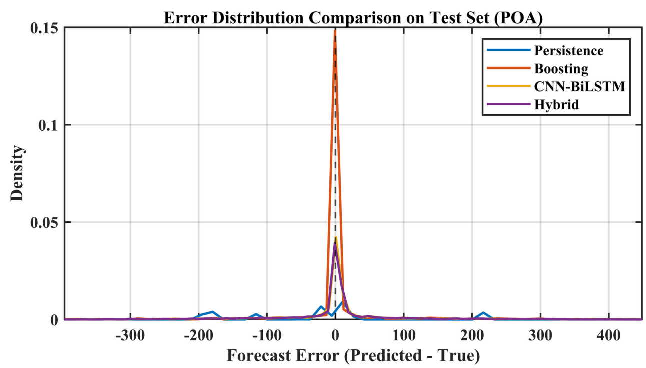


FIGURE 7. Comparison of forecast error distributions for the Persistence, Boosting, CNN–BiLSTM, and proposed Hybrid forecasting models.

To evaluate model robustness under different atmospheric conditions, forecasting accuracy was analyzed across clearness-index regimes. Fig. 8 presents RMSE values under clear and partly cloudy conditions. The persistence model performs poorly in all regimes due to its inability to capture cloud-driven variability. Machine learning and deep learning models significantly reduce forecasting error. Among the evaluated baseline models, the hybrid framework achieves the lowest RMSE and highest coefficient of determination.

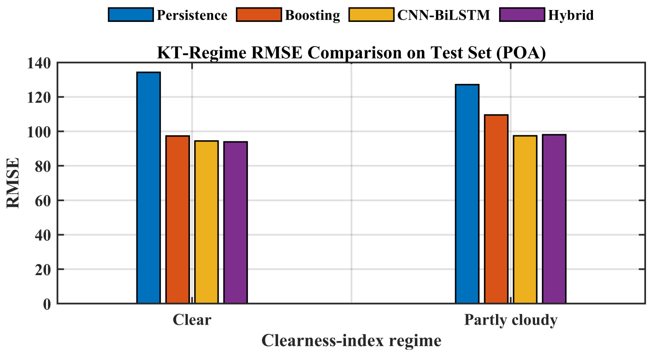


FIGURE 8. Omparison of RMSE under clear and partly cloudy clearness-index regimes for the Persistence, Boosting, CNN–BiLSTM, and proposed Hybrid models.

Residual autocorrelation analysis was performed to investigate the temporal structure of forecast errors. Fig. 9 compares the residual autocorrelation functions of the CNN–BiLSTM and hybrid models. The baseline deep learning model exhibits noticeable correlation at several temporal lags, indicating unresolved temporal dependencies. After applying residual correction, the hybrid model significantly suppresses these correlations, suggesting that the residual learner successfully captures structured atmospheric variability.

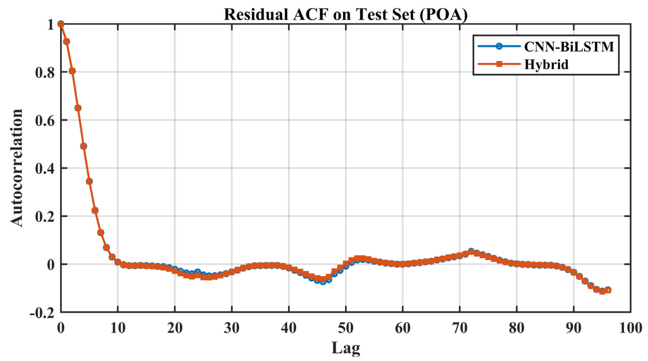


FIGURE 9. Residual autocorrelation (ACF) comparison of irradiance forecasting errors for the baseline CNN–BiLSTM and proposed Hybrid models.

The residual correction strength parameter α was tuned using validation data. Fig. 10 shows the relationship between α and validation RMSE. The optimal value is obtained at $\alpha = 0.15$, which minimizes forecasting error. Larger values lead to over-correction, while smaller values underutilize residual information.

It should be emphasized that the baseline CNN–BiLSTM model was implemented using the same input features, training–validation–test split, and preprocessing pipeline as the proposed hybrid framework. Therefore, the observed

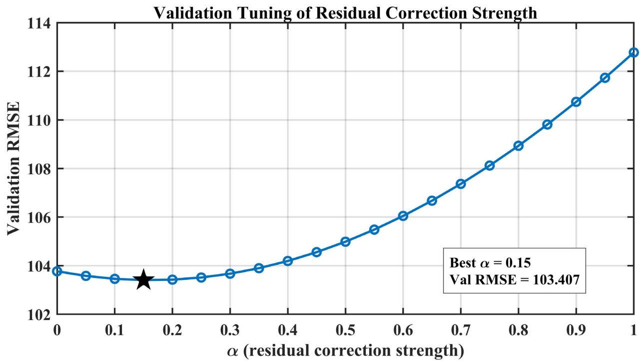


FIGURE 10. Validation-based tuning of the residual correction coefficient (α) for the proposed hybrid forecasting model.

performance improvement is not due to differences in data handling or feature availability but arises solely from the proposed residual learning formulation

C. ERROR DISTRIBUTION ANALYSIS

Average error metrics alone cannot fully describe forecasting reliability, particularly for highly variable solar radiation. Therefore, the probability density distribution of prediction errors was analyzed, as illustrated in Fig. 11.

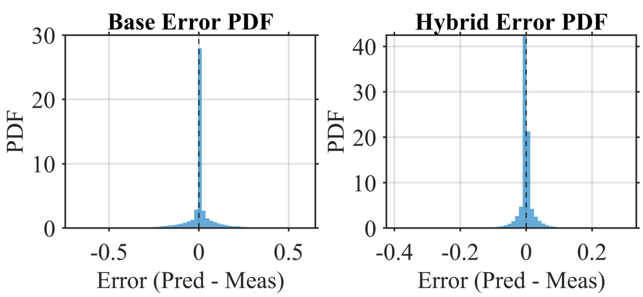


FIGURE 11. Comparison of forecast error probability density functions (PDFs) for the baseline CNN–BiLSTM and proposed Hybrid models.

The baseline deep learning model produces a relatively wide distribution of errors, with visible dispersion around zero, indicating inconsistent prediction behavior across different atmospheric conditions. In contrast, the proposed hybrid framework generates a significantly narrower and more symmetric distribution centered close to zero error. This behavior is consistent with the reduction in RMSE from 0.0809 to 0.0249 and in bias from 0.0050 to approximately 0, as reported in Table 2.

The contraction of the error distribution demonstrates that the residual learning stage effectively removes systematic deviations rather than merely averaging errors. In particular, the reduction of heavy tails indicates improved stability during rapidly changing irradiance conditions caused by cloud intermittency. This confirms that the hybrid model not only improves accuracy but also enhances prediction reliability, which is essential for operational photovoltaic scheduling and energy flexibility management

D. REPRESENTATIVE DAY EVALUATION

Beyond statistical accuracy, forecasting reliability must be evaluated from temporal behavior. The time-series comparison shown in Fig. 2 illustrates model performance under clear, partially cloudy, and overcast conditions.

The deep learning model captures the overall diurnal irradiance trend but smooths rapid variations, particularly during cloud transitions. This leads to an underestimation of sharp irradiance ramps and delayed recovery after transient shading events.

The proposed hybrid framework significantly improves short-term dynamics by correcting systematic deviations in the base prediction. The model closely follows sudden increases and decreases in irradiance, indicating that the residual learning stage captures stochastic atmospheric variability rather than only long-term patterns.

Such improvement is operationally important, as accurate ramp prediction directly influences battery dispatch decisions, reserve activation, and grid balancing performance in renewable-dominated power systems.

E. SEASONAL PERFORMANCE

Forecasting reliability under varying solar elevation and atmospheric regimes was evaluated through seasonal performance analysis. The quantitative results are summarized in Table 4, while the graphical comparison is illustrated in Fig. 12.

TABLE 4. Seasonal irradiance forecasting performance evaluation.

Season	Model	RMSE	R^2
Winter	Deep Learning (Base)	0.0452	0.7903
	Hybrid (Proposed)	**0.0159**	**0.9740**
Spring	Deep Learning (Base)	0.0813	0.8678
	Hybrid (Proposed)	**0.0195**	**0.9924**
Summer	Deep Learning (Base)	0.0914	0.8639
	Hybrid (Proposed)	**0.0279**	**0.9873**
Autumn	Deep Learning (Base)	0.0952	0.7406
	Hybrid (Proposed)	**0.0325**	**0.9697**

The proposed hybrid framework consistently improves irradiance prediction accuracy in all seasons. The baseline model exhibits large seasonal variability, with higher errors during autumn and summer due to unstable atmospheric conditions and cloud intermittency. After residual correction, the prediction error decreases substantially, and the coefficient of determination increases across all seasons.

The improvement is particularly significant in challenging low-irradiance conditions. During winter, the RMSE decreases from approximately 0.045 to 0.016, while the coefficient of determination increases from about 0.79 to 0.97. Similarly, autumn performance improves from RMSE ≈ 0.095 to 0.033 and R^2 ≈ 0.74 to 0.97, demonstrating the residual learning stage's ability to compensate for stochastic atmospheric attenuation. Spring and summer also show strong gains, confirming that the model adapts effectively across different solar elevation angles.

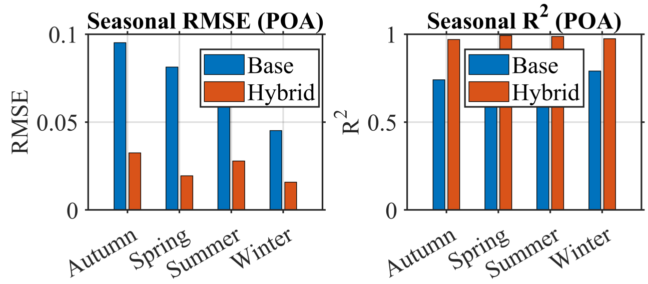


FIGURE 12. Seasonal comparison of RMSE and coefficient of determination (R^2) for POA irradiance prediction using the baseline CNN–BiLSTM and proposed Hybrid models.

The bar charts in Fig. 12 visually confirm this behavior: hybrid predictions (orange bars) consistently reduce RMSE and increase R^2 compared with the baseline model (blue bars). The uniform improvement across all seasons indicates that the proposed decomposition successfully separates deterministic solar motion from weather-driven variability, resulting in robust year-round forecasting performance

F. WEATHER REGIME ROBUSTNESS

To further evaluate reliability under different atmospheric conditions, the dataset was divided into clearness index (KT) regimes: overcast [0.0, 0.3), partly cloudy [0.3, 0.6), and clear sky [0.6, 1.0]. The first two intervals are left-closed and right-open to avoid overlapping between adjacent regimes, while the final interval is right-closed to include the physical upper bound KT = 1.0. The quantitative results are summarized in Table 5, and the graphical comparison is shown in Fig. 13.

TABLE 5. Forecasting performance under clearness index regimes (POA Irradiance).

KT Range	Model	RMSE	Bias
[0.0, 0.3) (Overcast)	Base	0.0947	0.0101
	Hybrid	**0.0317**	**0.0001**
[0.3, 0.6) (Partly Cloudy)	Base	0.0632	0.0043
	Hybrid	**0.0191**	**−0.0002**
[0.6, 1.0] (Clear Sky)	Base	0.0747	0.0011
	Hybrid	**0.0201**	**0.0000**

The baseline deep learning model exhibits significant performance degradation under cloudy conditions, where irradiance variability is dominated by stochastic cloud attenuation. In the overcast regime, the RMSE decreases from approximately 0.0947 to 0.0317 after applying the hybrid correction, and the prediction bias is reduced from 0.0101 to nearly zero. Similar improvements are observed in partially cloudy conditions, where RMSE decreases from 0.0632 to 0.0191, and bias is effectively eliminated.

Even under clear-sky conditions, where deterministic solar behavior dominates, the hybrid model maintains higher accuracy, reducing RMSE from 0.0747 to 0.0201. This indicates that the residual learning stage does not interfere with predictable radiative patterns but only compensates for unresolved atmospheric effects.

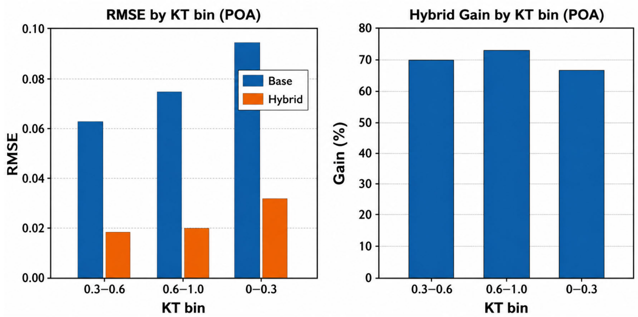


FIGURE 13. **RMSE comparison and hybrid performance gain across different clearness-index (KT) regimes for POA irradiance forecasting.**

The bar charts in Fig. 13 visually confirm that hybrid predictions consistently reduce error magnitude and remove systematic bias across all regimes. These results indicate that the residual model learns remaining structured forecasting patterns rather than duplicating the information already represented by the primary temporal model.

G. ABLATION STUDY

To quantify the contribution of each component in the proposed framework, an ablation analysis was conducted by comparing three configurations: the standalone deep temporal model (Deep/Base), the deep model followed by residual correction without fusion weighting (Deep + Residual ML), and the complete hybrid framework. The results are presented in Figure 14.

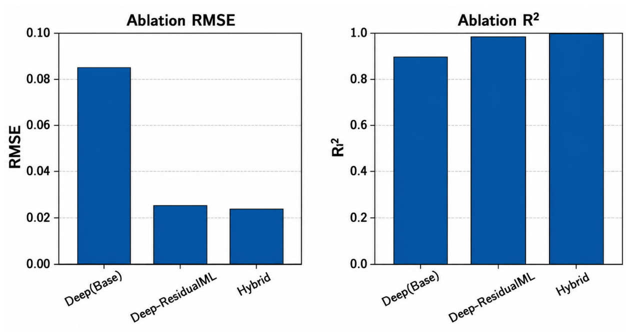


FIGURE 14. **Ablation comparison of RMSE and coefficient of determination (R^2) for the Deep, Deep + Residual, and proposed Hybrid models.**

As shown in Fig. 14 (left), the standalone deep model yields the highest prediction error, with an RMSE of approximately 0.081, indicating that a single nonlinear mapping cannot simultaneously model both periodic solar behavior and stochastic atmospheric fluctuations. Introducing the residual learning stage dramatically reduces the error to about 0.025, demonstrating that a substantial portion of the forecast error is structured and can be learned separately. The full hybrid configuration maintains the same low RMSE while stabilizing prediction behavior.

The improvement is also evident in the coefficient of determination shown in Fig. 14 (right). The deep model achieves an R^2 of roughly 0.85, whereas incorporating residual correction increases R^2 to nearly 0.99, confirming a significantly better representation of irradiance variability. The hybrid configuration preserves this high explanatory capability while improving robustness.

These results indicate that the performance gain is not due to increased model complexity but due to functional task decomposition. The deep network captures smooth deterministic irradiance evolution, while the residual learner models stochastic atmospheric disturbances, validating the proposed sequential specialization strategy.

H. STATISTICAL SIGNIFICANCE ANALYSIS

To verify that the observed improvement is not due to random variation, a paired bootstrap statistical test was conducted on the absolute forecast errors of the baseline deep model and the proposed hybrid framework. For each resampling iteration, the mean difference between the absolute errors $| e_{Base} | - | e_{Hybrid} |$ was computed, producing the empirical distribution illustrated in Figure 15.

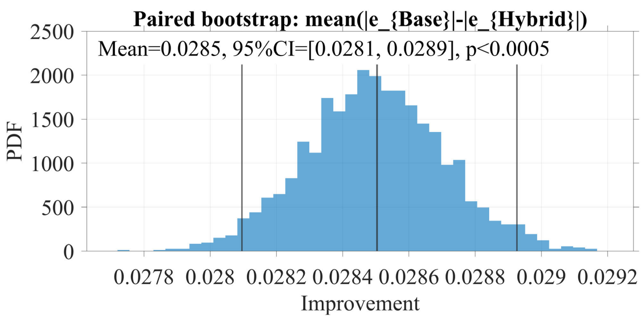


FIGURE 15. **Paired bootstrap analysis demonstrating the statistical significance of the prediction error improvement achieved by the proposed Hybrid model.**

The bootstrap distribution is centered on a positive value, indicating that the hybrid model reduces systematic error. The mean improvement is 0.0285, with a 95% confidence interval of [0.0281, 0.0289] and a significance level of $p < 0.0005$. Since the confidence interval does not include zero and the p-value is far below conventional thresholds, the null hypothesis of equal predictive accuracy can be rejected. This result confirms that the performance gain is statistically significant rather than a consequence of sample variability. Therefore, the hybrid framework consistently improves irradiance forecasting accuracy, reinforcing the conclusion that residual learning captures structured atmospheric effects beyond the capabilities of the standalone deep learning model.

I. COMPUTATIONAL EFFICIENCY

The computational cost of the proposed framework was evaluated by measuring the average inference time per sample, as shown in Fig. 16. The hybrid residual stage requires approximately 0.27 ms per sample, indicating negligible additional overhead compared to the deep temporal predictor. This confirms that the residual correction mechanism improves forecasting accuracy without compromising real-time applicability, making the approach suitable for operational energy management and embedded deployment scenarios. The proposed forecasting framework was implemented and evaluated in MATLAB. Model training was performed offline, while the forecasting stage only requires the sequential execution of the trained temporal forecasting model, followed by the residual correction stage. Therefore, the additional computational requirement introduced by the residual learner is limited compared with increasing the complexity of the primary forecasting architecture. The reported inference time represents the average prediction time per sample during the testing phase.

For metric representation, input and output variables were normalized during model development to improve numerical stability. However, final forecasting evaluation was performed after converting the predicted values back to the original physical scale. Therefore, reported physical-unit results represent actual irradiance values, while normalized metrics are provided only for relative comparison where specified.

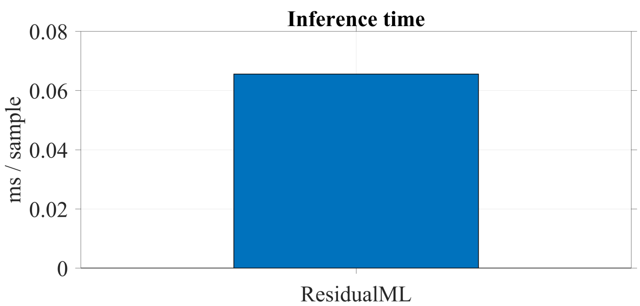


FIGURE 16. Computational efficiency evaluation of the proposed residual-corrected forecasting framework showing the average inference time per testing sample (ms/sample).

The reported inference time is the average computational time to generate one irradiance prediction sample during testing, expressed in milliseconds per sample (ms/sample), rather than the total execution time for the entire testing dataset.

VI. MODEL INTERPRETATION AND RELIABILITY ANALYSIS

A. RESIDUAL BEHAVIOR ANALYSIS

To understand the contribution of the residual learning stage, the temporal structure of the prediction error was analyzed using autocorrelation. The residual autocorrelation function (ACF) is shown on Fig. 17.

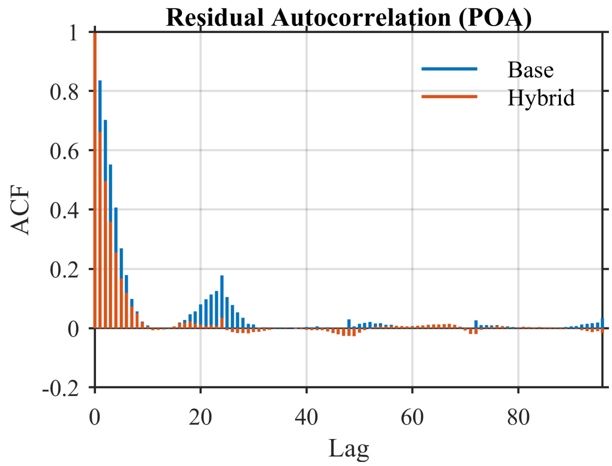


FIGURE 17. Residual autocorrelation function (ACF) of POA prediction error for the base and hybrid models.

The baseline deep learning model exhibits strong positive autocorrelation at short lags and noticeable periodic behavior at longer lags. This indicates that the prediction error contains systematic temporal patterns rather than purely random noise. In other words, the deep temporal model fails to capture certain structured atmospheric dynamics, especially those related to cloud evolution and irradiance ramps.

After applying the residual correction, the autocorrelation magnitude decreases substantially across all lags. The hybrid model residual approaches a near-white-noise distribution, with minimal temporal dependence remaining in the error signal. The disappearance of periodic structures confirms that the second learning stage successfully models the missing atmospheric variability.

These results demonstrate that the residual component represents a learnable stochastic process rather than measurement randomness. The proposed hybrid framework, therefore, separates the forecasting task into two physically meaningful parts: deterministic radiative behavior, handled by the deep network, and stochastic atmospheric disturbances, handled by the residual learner.

To further support the interpretation of the residual component as a structured atmospheric effect, multiple complementary analyses were considered. First, the baseline residual exhibits strong short-lag autocorrelation and periodic behavior, indicating unresolved temporal dependencies rather than random noise. Second, regime-based evaluation across clearness index categories shows that the largest performance improvements occur under overcast and partly cloudy conditions, where irradiance variability is dominated by cloud dynamics. Third, the error dependence on irradiance magnitude reveals increased prediction error during high-irradiance and cloud-edge conditions, which are characteristic of rapid atmospheric transitions. Finally, permutation feature importance analysis demonstrates that the residual model is primarily driven by clearness index, lagged irradiance, moving averages, and irradiance slope, all of which are physically related to short-term atmospheric evolution and irradiance ramp behavior. These results collectively provide empirical evidence that the residual component reflects structured atmospheric variability rather than purely random prediction noise.

B. ERROR DEPENDENCE ON IRRADIANCE MAGNITUDE

The relationship between prediction error and irradiance magnitude is illustrated in Fig. 18 for POA irradiance and Fig. 19 for cell temperature.

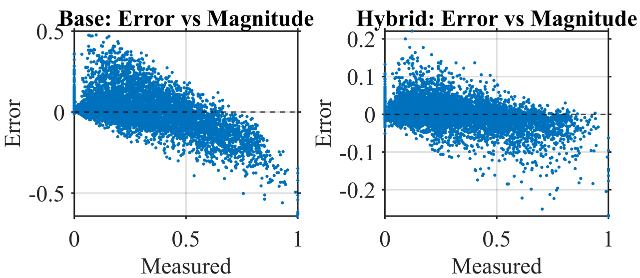


FIGURE 18. POA prediction error vs measured irradiance magnitude.

In Fig. 12, the baseline deep learning model exhibits clear heteroscedasticity: prediction error increases with irradiance level. At high irradiance (normalized POA $\approx$ 0.7–1.0), the error spreads widely and becomes negatively biased, indicating systematic underestimation during cloud-edge transitions with rapid ramps. After applying the residual correction, the hybrid framework produces a much narrower and more symmetric error distribution centered near zero. The spread of error decreases substantially, demonstrating improved stability under rapidly varying atmospheric conditions.

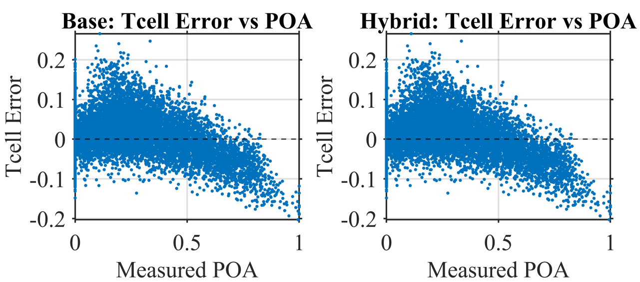


FIGURE 19. Tcell prediction error vs measured irradiance magnitude.

In Fig. 13, a similar dependence is observed for cell temperature error. The base model exhibits irradiance-dependent drift, in which temperature prediction error increases with irradiance. The hybrid model reduces this dependency and stabilizes the error around zero across the entire irradiance range. These results confirm that the residual learner primarily corrects ramp-related deviations rather than steady-state conditions. This behavior is particularly beneficial for grid operation and storage dispatch, since ramp prediction errors contribute disproportionately to balancing requirements.

C. FEATURE IMPORTANCE ANALYSIS

To interpret the behavior of the residual learning stage, permutation feature importance was evaluated for the residual model, as illustrated in Fig. 20.

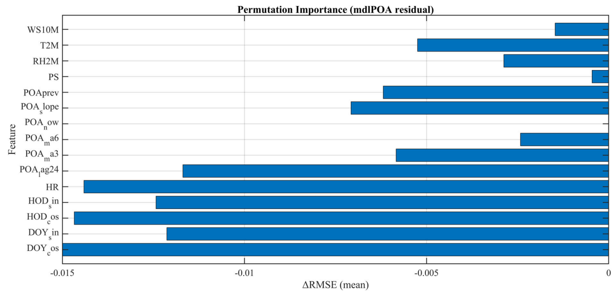


FIGURE 20. Permutational importance of this proposed study.

The results indicate that the most influential predictors are recent irradiance dynamics, particularly the 24-hour lag irradiance (POA_lag24), short-term moving averages (POA_ma3, POA_ma6), and irradiance slope (POA_slope). These variables describe short-term temporal evolution and are strongly associated with cloud-induced ramp events. Periodic encodings (hour-of-day and day-of-year sine/cosine terms) also show high importance, confirming that the residual component still depends on solar geometry when correcting systematic deep-model bias. In contrast, atmospheric variables such as temperature (T2M), relative humidity (RH2M), wind speed (WS10M), and pressure (PS) contribute less but remain non-negligible. This indicates that meteorological conditions influence the magnitude of deviations but do not dominate the correction process. Overall, the importance of distribution demonstrates functional specialization between the two stages: the deep temporal model captures smooth deterministic irradiance evolution. In contrast, the residual model focuses on correcting short-term structured deviations driven by atmospheric variability. Rather than duplicating learning tasks, the hybrid framework assigns complementary predictive roles to each component.

D. FORECAST RELIABILITY AND CALIBRATION

Forecast reliability was evaluated using calibration analysis, shown on Fig. 21. Calibration compares the mean predicted irradiance with the corresponding mean observed irradiance across prediction bins. Ideally, perfectly reliable forecasts lie in the one-to-one reference line.

The baseline deep learning model exhibits a noticeable deviation from the reference line, particularly in the medium- and high-irradiance ranges, indicating systematic miscalibration. In these regions, predicted values do not scale proportionally with observed magnitudes, suggesting overconfident predictions and a biased representation of uncertainty.

E. FAILURE CASES AND REMAINING CHALLENGES

Despite the overall improvement achieved by the proposed hybrid framework, certain forecasting scenarios remain challenging and offer insight into the approach's limitations. The

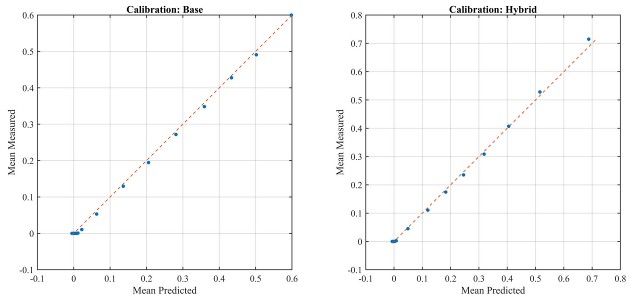


FIGURE 21. Calibration comparison between the baseline deep model and the hybrid model.

most prominent failure cases occur during highly dynamic atmospheric conditions, particularly under rapidly changing cloud cover and cloud-edge transitions. In these situations, irradiance can fluctuate abruptly within short time intervals, leading to sharp ramps that are difficult to capture accurately even with the residual correction stage.

The baseline CNN–BiLSTM model tends to smooth rapid variations due to its temporal averaging, resulting in underestimation of peak irradiance and a delayed response to sudden changes. Although the residual learning component significantly improves ramp tracking, small deviations may still persist when the underlying atmospheric transitions are too abrupt or insufficiently represented in the training data.

Additional challenges are observed under very low irradiance conditions, such as sunrise, sunset, or heavily overcast periods, where the signal-to-noise ratio is low and measurement uncertainty becomes more significant. In these regimes, both the deep model and the residual learner may exhibit increased relative error, although the absolute error magnitude remains small.

Furthermore, the residual model relies on engineered features derived from historical patterns, which may limit its ability to generalize to unseen extreme weather conditions or rare atmospheric events. This highlights the importance of incorporating more diverse training data and potentially integrating external information sources, such as satellite imagery or numerical weather prediction outputs, for improved robustness.

Overall, these failure cases indicate that while the proposed hybrid framework effectively captures structured atmospheric variability, forecasting performance remains constrained under highly non-stationary and extreme conditions. Addressing these limitations represents an important direction for future research.

VII. DISCUSSIONS

A. PHYSICAL INTERPRETATION OF THE HYBRID LEARNING MECHANISM

The magnitude of performance improvement observed in this study may appear unusually high compared to conventional solar forecasting results. This is primarily due to the nature of the baseline deep learning model, which tends to produce smooth predictions and to systematically underestimate high-frequency irradiance variability driven by cloud dynamics. As demonstrated in the residual autocorrelation and error distribution analyses, a significant portion of the prediction error is structured rather than random. The proposed framework explicitly models this structured error as a separate learning task, allowing the residual learner to capture variability that the deep model fails to represent. Consequently, the improvement reflects the recovery of previously unmodeled atmospheric dynamics rather than incremental optimization gains. The results demonstrate that the proposed hybrid framework improves forecasting reliability by separating deterministic radiative behavior from stochastic atmospheric disturbances. The deep temporal model primarily learns periodic solar motion and smooth meteorological evolution, whereas the residual learning stage corrects systematic deviations caused by unresolved cloud dynamics. This explains why the performance improvement is consistent across seasons and weather regimes rather than limited to specific operating conditions.

Conventional deep learning approaches attempt to approximate the entire irradiance signal using a single nonlinear mapping. This forces one model to simultaneously represent predictable solar geometry and unpredictable atmospheric variability. As observed in the representative-day analysis, this leads to over-smoothing and underestimation of ramps. The residual learning stage removes this conflict by assigning stochastic correction to a dedicated model trained on structured forecast errors. Consequently, the forecasts preserve smooth periodic behavior while accurately reproducing rapid irradiance transitions, which are critical for operational energy systems.

The reliability analyses further support this interpretation. Residual autocorrelation results indicate that the baseline error contains non-random temporal structure, while the clearness-index regime analysis shows that the largest improvements occur under cloudy conditions where atmospheric variability dominates. Moreover, the dependence of error on irradiance magnitude and the feature-importance results confirm that the residual learner is driven by variables associated with cloud-induced ramps and short-term atmospheric evolution. These observations collectively indicate that the residual component is consistent with structured atmospheric variability rather than purely random noise. Additionally, the reduced dependence of error on irradiance magnitude demonstrates improved ramp prediction capability. Since ramp events dominate reserve activation and storage dispatch requirements, the primary benefit of the proposed approach is enhanced operational reliability rather than only reduced average error.

However, it is important to note that the current formulation produces deterministic point forecasts. In practical energy system applications, uncertainty information is often required for risk-aware decision-making, such as reserve allocation and storage scheduling. While the proposed framework improves deterministic accuracy, incorporating probabilistic forecasting techniques, such as prediction intervals or quantile-based models, could further enhance its applicability.

B. COMPARISON WITH EXISTING APPROACHES

Solar forecasting methods generally fall into four categories: deterministic physical models, statistical machine learning methods, deep learning architectures, and hybrid physical–data-driven frameworks [40], [47], [49], [51]. Physical models provide interpretable predictions based on solar geometry but cannot resolve short-term cloud variability [50]. Conventional machine learning models improve local accuracy but lack temporal generalization across atmospheric regimes [51]. Deep learning architectures capture temporal dependencies but tend to underestimate high-frequency irradiance ramps due to smoothing behavior45. Hybrid feature-augmented methods improve robustness but typically integrate all information into a single predictor and treat forecast error as random noise [23], [25], [28]. The proposed framework differs fundamentally in that it explicitly models the residual component as a structured atmospheric process. Rather than enriching inputs for a single model, the forecasting task is decomposed into complementary stages. The deep network focuses on deterministic radiative behavior, while the secondary learner specializes in atmospheric disturbances.

This sequential specialization enables regime-independent improvement across clear, mixed, and overcast conditions, indicating that the residual model learns atmospheric variability rather than duplicating deterministic solar patterns. A summary of the comparison is presented in Table 6.

TABLE 6. Comparison with existing forecasting paradigms.

Method	Per.	Cloud	Ramp	Interp.	Reliab.
H = High, M = Medium, L = Low, P = Poor, I = Improved, S = Strong					
Physical	H	L	P	H	L
ML	M	M	L	L	M
DL	H	M	L	L	M
Hybrid	H	M	I	M	M–H
Proposed	H	H	S	M–H	H

C. LIMITATIONS

Despite the promising performance improvements demonstrated in this study, several limitations should be acknowledged. First, the residual learning component relies on the availability of high-quality historical observations, and its effectiveness may decrease in regions where meteorological measurements or irradiance records are sparse or incomplete.

Second, the proposed framework produces deterministic point forecasts and does not explicitly quantify predictive uncertainty. In energy system applications such as grid operation, reserve scheduling, and storage management, probabilistic forecasting can provide important information about forecast confidence and variability. Methods such as prediction intervals, quantile regression, or probabilistic deep learning models could be incorporated to extend the proposed framework toward uncertainty-aware forecasting.

Third, the present evaluation is conducted using data from a single geographic location in Finland. Because solar irradiance forecasting performance can be influenced by regional climatic characteristics and atmospheric variability, further investigation using multi-site datasets from diverse climatic regions is necessary to assess the proposed approach's generalization capability.

Although the proposed framework demonstrates improved forecasting performance using long-term high-latitude solar observations, the evaluation is limited to a single geographical location. Since irradiance variability and atmospheric behavior differ across climates, further multi-site validation is required before generalizing the findings to other regions. In addition, the present study focuses on deterministic point forecasting. Future work will investigate probabilistic forecasting approaches to quantify prediction uncertainty and support risk-aware renewable energy operation. Although the proposed framework is evaluated against representative forecasting approaches, including persistence, gradient boosting, and temporal deep learning models, the comparison does not include every recently developed forecasting architecture. Recent Transformer-based models, Informer/Autoformer variants, advanced gradient-boosting frameworks such as Light Gradient Boosting Machine and Extreme Gradient Boosting, and other hybrid deep learning approaches have demonstrated strong potential for renewable energy forecasting. Therefore, the objective of the present comparison is not to claim universal superiority over all existing methods, but to demonstrate the effectiveness of the proposed residual-correction strategy under the considered experimental conditions. Future work will extend the benchmarking framework by incorporating additional state-of-the-art architectures and multi-dataset evaluations.

It should be noted that the present study does not directly measure cloud movement or atmospheric physical states. Therefore, the relationship between residual behavior and atmospheric variability is interpreted through indirect statistical evidence, including residual autocorrelation, clearness-index analysis, seasonal evaluation, and feature importance. Direct validation using additional atmospheric observations, such as sky images or satellite-derived cloud information, will be considered in future work.

Finally, although the method significantly improves the reliability of irradiance prediction, its integration into real-time power system operations, such as battery dispatch, reserve scheduling, or grid control, has not been evaluated and remains an important direction for future research.

D. NOVELTY AND PRATICAL IMPLICATIONS

The primary novelty of this study lies in treating forecast residuals as a structured atmospheric process rather than random noise. Instead of combining heterogeneous features into a single predictive model, the forecasting task is decomposed into complementary learning stages aligned with the physical origin of irradiance variability.

This shifts hybrid modeling from feature fusion to functional specialization:

- Deep network models deterministic radiative behavior,
- The residual learner models stochastic atmospheric disturbances.

The resulting framework is more physically interpretable and maintains reliability under rapidly changing weather conditions.

From an operational perspective, improved ramp prediction directly benefits:

- storage scheduling,
- reserve allocation,
- flexibility activation,
- energy trading in renewable-dominated grids.

Therefore, the proposed method functions as a reliability-oriented forecasting layer suitable for digital twin energy systems, where forecast stability is often more critical than marginal reductions in mean error.

VIII. CONCLUSION

This paper introduced a residual-corrected physics-guided hybrid forecasting framework that decomposes solar irradiance into deterministic radiative behavior and stochastic atmospheric variability. A CNN–BiLSTM deep temporal network first models' periodic solar geometry and smooth meteorological evolution, and a gradient-boosted regression trees used as a residual learning stage to model remaining structured forecasting errors. The final forecast is obtained by combining the deep prediction with the learned residual component, allowing each model to specialize in physically meaningful behavior rather than forcing a single network to approximate the entire signal.

The proposed framework demonstrates substantial quantitative improvements over the standalone deep learning model. For plane-of-array irradiance prediction, the hybrid approach reduces RMSE from 0.0809 to 0.0249 ($\approx$ 69% reduction) and MAE from 0.0404 to 0.0119 while increasing the coefficient of determination from $R^2 = 0.8468$ to 0.9855 and eliminating prediction bias (from 0.0050 to $\sim$0.0000). Seasonal evaluation confirms consistent improvement in reliability: winter RMSE decreases from 0.0452 to 0.0159, spring from 0.0813 to 0.0195, summer from 0.0914 to 0.0279, and autumn from 0.0952 to 0.0325, with corresponding R^2 values approaching 0.97–0.99 across all seasons. Weather-regime analysis further shows strong robustness, with overcast-conditioned RMSE dropping from 0.0947 to 0.0317, partly cloudy from 0.0632 to 0.0191, and clear-sky from 0.0747 to 0.0201, demonstrating accurate performance under both deterministic and highly stochastic conditions.

Reliability-oriented analysis supports these numerical gains. The error distribution becomes significantly narrower and symmetric, residual autocorrelation approaches white noise, and calibration curves closely follow the one-to-one reference line. A paired bootstrap test confirms statistical significance, with a mean error improvement of 0.0285, a 95% confidence interval [0.0281, 0.0289], and $p < 0.0005$, indicating that the performance gain is systematic rather than sample-dependent. Importantly, the residual correction adds only negligible computational overhead while maintaining real-time applicability.

Unlike traditional hybrid models that merely fuse features, the proposed method introduces functional specialization: the deep network captures deterministic solar periodicity, while the residual learner models remaining structured forecasting errors associated with weather-dependent irradiance variations. This decomposition preserves smooth diurnal behavior while accurately reproducing abrupt irradiance transitions, which are critical for operational decisions such as battery dispatch, reserve activation, and flexibility scheduling in renewable-dominated grids.

Overall, the study demonstrates that treating prediction residuals as structured learnable information significantly improves both accuracy and reliability without increasing model complexity. The framework, therefore, provides a practical forecasting layer for digital-twin energy systems and photovoltaic operation. Future research will extend the method toward probabilistic forecasting, multi-site spatial generalization, and integration with real-time energy management and storage optimization platforms.

ACKNOWLEDGMENT

The authors used AI-based language tools, including Grammarly, QuillBot, and ChatGPT, solely for improving grammar, spelling, and phrasing during manuscript preparation. The scientific ideas, simulations, analyses, and conclusions presented in this article are entirely them original contributions. All responsibility for the content rests with them.

REFERENCES

[1] E. Kabir, P. Kumar, S. Kumar, A. A. Adelodun, and K.-H. Kim, "Solar energy: Potential and future prospects," *Renew. Sustain. Energy Rev.*, vol. 82, pp. 894–900, Feb. 2018, doi: 10.1016/j.rser.2017.09.094.

[2] Z. Si, M. Yang, Y. Yu, and T. Ding, "Photovoltaic power forecast based on satellite images considering effects of solar position," *Appl. Energy*, vol. 302, Nov. 2021, Art. no. 117514, doi: 10.1016/j.apenergy.2021.117514.

[3] C. Zhang, L. Hua, C. Ji, M. S. Nazir, and T. Peng, "An evolutionary robust solar radiation prediction model based on WT-CEEMDAN and IASO-optimized outlier robust extreme learning machine," *Appl. Energy*, vol. 322, Sep. 2022, Art. no. 119518, doi: 10.1016/j.apenergy.2022.119518.

[4] H. Acikgoz, "A novel approach based on integration of convolutional neural networks and deep feature selection for short-term solar radiation forecasting," *Appl. Energy*, vol. 305, Jan. 2022, Art. no. 117912, doi: 10.1016/j.apenergy.2021.117912.

[5] C. Hu, Z. Cai, Y. Zhang, R. Yan, Y. Cai, and B. Cen, "A soft actor-critic deep reinforcement learning method for multi-timescale coordinated operation of microgrids," *Protection Control Mod. Power Syst.*, vol. 7, no. 1, p. 29, Dec. 2022, doi: 10.1186/s41601-022-00252-z.

[6] S. Zhang and J. J. Q. Yu, "Bayesian deep learning for dynamic power system state prediction considering renewable energy uncertainty," *J. Mod. Power Syst. Clean Energy*, vol. 10, no. 4, pp. 913–922, Jul. 2022, doi: 10.35833/MPCE.2020.000939.

[7] Z. Peng, D. Yu, D. Huang, J. Heiser, S. Yoo, and P. Kalb, "3D cloud detection and tracking system for solar forecast using multiple sky imagers," *Sol. Energy*, vol. 118, pp. 496–519, Aug. 2015, doi: 10.1016/j.solener.2015.05.037.

[8] F. J. Rodríguez-Benítez, M. López-Cuesta, C. Arbizu-Barrena, M. M. Fernández-León, M. Á. Pamos-Ureña, J. Tovar-Pescador, F. J. Santos-Alamillos, and D. Pozo-Vázquez, "Assessment of new solar radiation nowcasting methods based on sky-camera and satellite imagery," *Appl. Energy*, vol. 292, Jun. 2021, Art. no. 116838, doi: 10.1016/j.apenergy.2021.116838.

[9] M. Ajith and M. Martínez-Ramón, "Deep learning based solar radiation micro forecast by fusion of infrared cloud images and radiation data," *Appl. Energy*, vol. 294, Jul. 2021, Art. no. 117014, doi: 10.1016/j.apenergy.2021.117014.

[10] H. Hissou, S. Benkirane, A. Guezzaz, M. Azrour, and A. Beni-Hssane, "A novel machine learning approach for solar radiation estimation," *Sustainability*, vol. 15, no. 13, p. 10609, Jul. 2023, doi: 10.3390/su151310609.

[11] Prediction of Worldwide Energy Resources (POWER). *NASA Langley Research Center (LaRC) POWER Project Funded Through the NASA Earth Science Directorate Applied Science Program*. Accessed: Feb. 10, 2026. [Online]. Available: https://power.larc.nasa.gov

[12] S. Park, Y. Kim, N. J. Ferrier, S. M. Collis, R. Sankaran, and P. H. Beckman, "Prediction of solar irradiance and photovoltaic solar energy product based on cloud coverage estimation using machine learning methods," *Atmosphere*, vol. 12, no. 3, p. 395, Mar. 2021, doi: 10.3390/atmos12030395.

[13] F. Lin, Y. Zhang, and J. Wang, "Recent advances in intra-hour solar forecasting: A review of ground-based sky image methods," *Int. J. Forecasting*, vol. 39, no. 1, pp. 244–265, Jan. 2023, doi: 10.1016/j.ijforecast.2021.11.002.

[14] T. Niu, J. Li, W. Wei, and H. Yue, "A hybrid deep learning framework integrating feature selection and transfer learning for multi-step global horizontal irradiation forecasting," *Appl. Energy*, vol. 326, Nov. 2022, Art. no. 119964, doi: 10.1016/j.apenergy.2022.119964.

[15] B. K. Puah, L. W. Chong, Y. W. Wong, K. M. Begam, N. Khan, M. A. Juman, and R. K. Rajkumar, "A regression unsupervised incremental learning algorithm for solar irradiance prediction," *Renew. Energy*, vol. 164, pp. 908–925, Feb. 2021.

[16] X. Zhang, F. Fang, and J. Wang, "Probabilistic solar irradiation forecasting based on variational Bayesian inference with secure federated learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7849–7859, Nov. 2021.

[17] H. Zang, L. Cheng, T. Ding, K. W. Cheung, Z. Wei, and G. Sun, "Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning," *Int. J. Electr. Power Energy Syst.*, vol. 118, Jun. 2020, Art. no. 105790, doi: 10.1016/j.ijepes.2019.105790.

[18] W. Kong, Y. Jia, Z. Y. Dong, K. Meng, and S. Chai, "Hybrid approaches based on deep whole-sky-image learning to photovoltaic generation forecasting," *Appl. Energy*, vol. 280, Dec. 2020, Art. no. 115875.

[19] J. O. Kamadinata, T. L. Ken, and T. Suwa, "Sky image-based solar irradiance prediction methodologies using artificial neural networks," *Renew. Energy*, vol. 134, pp. 837–845, Apr. 2019.

[20] Y. Fu, H. Chai, Z. Zhen, F. Wang, X. Xu, K. Li, M. Shafie-Khah, P. Dehghanian, and J. P. S. Catalão, "Sky image prediction model based on convolutional auto-encoder for minutely solar PV power forecasting," *IEEE Trans. Ind. Appl.*, vol. 57, no. 4, pp. 3272–3281, Jul. 2021, doi: 10.1109/TIA.2021.3072025.

[21] W.-C. Tsai, C.-S. Tu, C.-M. Hong, and W.-M. Lin, "A review of state-of-the-art and short-term forecasting models for solar PV power generation," *Energies*, vol. 16, no. 14, p. 5436, Jul. 2023, doi: 10.3390/en16145436.

[22] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017, pp. 4–9.

[23] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 11106–11115, doi: 10.1609/aaai.v35i12.17325.

[24] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, and J. Uszkoreit, "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent.*, 2021, Art. no. 4492.

[25] G. Yang, J. Xiao, C. Zhang, D. Yang, and C. Li, "A real-time photovoltaic power estimation framework based on multi-scale spatio-temporal graph fusion," *Electronics*, vol. 14, no. 22, p. 4492, Nov. 2025, doi: 10.3390/electronics14224492.

[26] A. Nespoli, A. Niccolai, E. Ogliari, G. Perego, E. Collino, and D. Ronzio, "Machine learning techniques for solar irradiation nowcasting: Cloud type classification forecast through satellite data and imagery," *Appl. Energy*, vol. 305, Jan. 2022, Art. no. 117834, doi: 10.1016/j.apenergy.2021.117834.

[27] L. Wang, X. Li, Y. Hao, and Q. Zhang, "Ultra-short-term solar irradiance prediction using an integrated framework with novel textural convolution kernel for feature extraction of clouds," *Sustainability*, vol. 17, no. 6, p. 2606, Mar. 2025, doi: 10.3390/su17062606.

[28] P. Tang, Y. Su, W. Zhao, Q. Wang, L. Zou, and J. Song, "A hybrid framework for photovoltaic power forecasting using shifted windows transformer-based spatiotemporal feature extraction," *Energies*, vol. 18, no. 12, p. 3193, Jun. 2025, doi: 10.3390/en18123193.

[29] F. Wang, Z. Xuan, Z. Zhen, Y. Li, K. Li, L. Zhao, M. Shafie-Khah, and J. P. S. Catalão, "A minutely solar irradiance forecasting method based on real-time sky image-irradiance mapping model," *Energy Convers. Manage.*, vol. 220, Sep. 2020, Art. no. 113075, doi: 10.1016/j.enconman.2020.113075.

[30] J. Simeunovic, B. Schubnel, P.-J. Alet, and R. E. Carrillo, "Spatio-temporal graph neural networks for multi-site PV power forecasting," *IEEE Trans. Sustain. Energy*, vol. 13, no. 2, pp. 1210–1220, Apr. 2022.

[31] L. Cheng, H. Zang, Y. Xu, Z. Wei, and G. Sun, "Probabilistic residential load forecasting based on micrometeorological data and customer consumption pattern," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3762–3775, Jul. 2021, doi: 10.1109/TPWRS.2021.3051684.

[32] E. Lorenz, J. Hurka, D. Heinemann, and H. G. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 2, no. 1, pp. 2–10, Mar. 2009, doi: 10.1109/JSTARS.2009.2020300.

[33] D. Yang, J. Kleissl, C. A. Gueymard, H. T. C. Pedro, and C. F. M. Coimbra, "History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining," *Sol. Energy*, vol. 168, pp. 60–101, Jul. 2018.

[34] R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, "Modeling daylight availability and irradiance components from direct and global irradiance," *Sol. Energy*, vol. 44, no. 5, pp. 271–289, 1990, doi: 10.1016/0038-092x(90)90055-h.

[35] J. Kleissl, M. Lave, M. Jamaly, and J.-L. Bosch, "Aggregate solar variability," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2012, pp. 1–3, doi: 10.1109/PESGM.2012.6344809.

[36] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Prog. Energy Combustion Sci.*, vol. 34, no. 5, pp. 574–632, Oct. 2008, doi: 10.1016/j.pecs.2008.01.001.

[37] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting: A review," *Renew. Energy*, vol. 105, pp. 569–582, May 2017, doi: 10.1016/j.renene.2016.12.095.

[38] U. Kumar and S. Syama, "Support vector machine based short term solar power prediction," in *Proc. 6th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2022, pp. 264–269.

[39] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994, doi: 10.1109/72.279181.

[40] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[41] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Oct. 2014, pp. 1724–1734, doi: 10.3115/v1/d14-1179.

[42] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[43] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 802–810.

[44] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *Int. J. Forecasting*, vol. 37, no. 4, pp. 1748–1764, Oct. 2021, doi: 10.1016/j.ijforecast.2021.03.012.

[45] K. E. Bassey, "Solar energy forecasting with deep learning technique," *Eng. Sci. Technol. J.*, vol. 4, no. 2, pp. 18–32, Jun. 2023, doi: 10.51594/estj.v4i2.1286.

[46] U. Hussan, H. Wang, J. Peng, H. Jiang, and H. Rasheed, "Transformer-based renewable energy forecasting: A comprehensive review," *Renew. Sustain. Energy Rev.*, vol. 226, Jan. 2026, Art. no. 116356, doi: 10.1016/j.rser.2025.116356.

[47] Z. Ren, S. Zhou, D. Liu, and Q. Liu, "Physics-informed neural networks: A review of methodological evolution, theoretical foundations, and interdisciplinary frontiers toward next-generation scientific computing," *Appl. Sci.*, vol. 15, no. 14, p. 8092, Jul. 2025, doi: 10.3390/app15148092.

[48] A. Patil and K. Kulkarni, "A hybrid ML-NWP approach for predicting GHI and power to mitigate penalties under CERC's deviation settlement mechanism," in *Proc. IEEE India Geosci. Remote Sens. Symp. (InGARSS)*, Goa, India, Dec. 2024, pp. 1–4, doi: 10.1109/INGARSS61818.2024.10984316.

[49] C.-L. Huang, Y.-K. Wu, C.-C. Tsai, J.-S. Hong, and Y.-Y. Li, "Revolutionizing solar power forecasts by correcting the outputs of the WRF-SOLAR model," *Energies*, vol. 17, no. 1, p. 88, Dec. 2023, doi: 10.3390/en17010088.

[50] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 13, no. 3, pp. 253–263, Jul. 1995, doi: 10.1080/07350015.1995.10524599.

[51] T. M. Mercier, T. Rahman, and A. Sabet, "Solar irradiance anticipative transformer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Vancouver, BC, Canada, Jun. 2023, pp. 2065–2074, doi: 10.1109/CVPRW59228.2023.00200.

[52] N. T. Nguyen and M. H. Pham, "Robustness of transformer, long short-term memory, and light gradient boosting machine models for solar power forecasting under noisy input conditions," *J. Européen des Systèmes Automatisés*, vol. 59, no. 2, pp. 293–301, Feb. 2026, doi: 10.18280/jesa.590201.

[53] M. F. Hanif, M. S. Naveed, M. Metwaly, J. Si, X. Liu, and J. Mi, "Advancing solar energy forecasting with modified ANN and light GBM learning algorithms," *AIMS Energy*, vol. 12, no. 2, pp. 350–386, 2024, doi: 10.3934/energy.2024017.

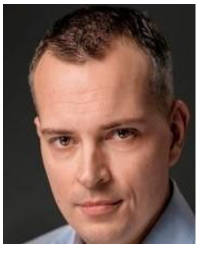
[54] G. Rohini, T. Mariprasth, S. A. Bustanji, and I. Zaitsev, "A stacked gradient boosting–XGBoost ensemble with ridge meta-learner for accurate short-term solar PV power forecasting in smart grids," *Sci. Rep.*, vol. 16, no. 1, p. 16871, Apr. 2026, doi: 10.1038/s41598-026-47042-3.

[55] M. Naghapushanam, B. Jeevarathinam, and C. Sankari, "Physics-informed voting ensemble for solar power generation forecasting: Integrating domain knowledge with machine learning," *Energy Informat.*, vol. 9, no. 1, Jan. 2026, Art. no. 3, doi: 10.1186/s42162-025-00604-7.

[56] D. Díaz-Bedoya, M. González-Rodríguez, X. Serrano-Guerrero, and J.-M. Clairand, "Solar irradiance forecasting with deep learning and ensemble models: LSTM, random forest and extra trees with multivariate meteorological data," *IET Smart Grid*, vol. 8, no. 1, Jan. 2025, Art. no. e70019, doi: 10.1049/stg2.70019.

[57] J. H. Friedman, "Greedy function approximation: A gradient boosting machine.," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi: 10.1214/aos/1013203451.

TAREQ ANWAR SHIKDAR received the B.Sc. degree (Hons.) in electrical and electronic engineering from Leading University, Sylhet, Bangladesh, in 2021, and the Erasmus Mundus Joint master's degree (EMJM) in smart cities and communities (SMACCs) in 2026. During the EMJM program, he specialized in smart grid solutions at the University of Vaasa, Finland, energy efficiency in buildings at the University of the Basque Country (UPV/EHU), Spain, and information and communication technologies at the International Hellenic University (IHU), Greece. His research focuses on AI-driven forecasting, optimization, and digital twin technologies for PV-BESS-integrated smart grids under Nordic operating conditions. He has authored several journal and conference publications and serves as a reviewer for IEEE-indexed conferences. His professional experience includes lecturing and research in energy systems and robotics laboratories. His research interests include smart grids, renewable energy integration, battery energy storage systems, artificial intelligence, digital twins, energy management, and IoT-enabled predictive analytics.



HANNU LAAKSONEN (Member, IEEE) received the M.Sc. (Tech.) degree in electrical power engineering from Tampere University of Technology, Tampere, Finland, in 2004, and the Ph.D. (Tech.) degree in electrical engineering from the University of Vaasa, Vaasa, Finland, in 2011. His employment experience includes a Research Scientist with the VTT Technical Research Centre of Finland and the University of Vaasa. Previously, he was a Principal Engineer with ABB Ltd. in Vaasa. He is currently a Professor of electrical engineering with the University of Vaasa. He is also an Electrical Engineering Department PU-Manager, a Flexible Energy Resources-Research Team Leader, and a Manager of the Smart Energy Transition Master's Program. His research interests include the control and protection of low-inertia power systems and microgrids, active management of distributed and flexible energy resources in future smart energy systems, and future-proof technology and market concepts for smart grids.