IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 27 April 2026, accepted 15 May 2026, date of publication 19 May 2026, date of current version 1 June 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3694780

RESEARCH ARTICLE

Multi-Scale Convolutional Chebyshev Kolmogorov-Arnold Networks Using Conformal Prediction for Photovoltaic Power Forecasting

LUCAS DE AZEVEDO TAKARA[1], RODRIGO CLEMENTE THOM DE SOUZA[2], VIVIANA COCCO MARIANI[3,4], (Member, IEEE), LEANDRO DOS SANTOS COELHO[1,4], (Member, IEEE), AND STEFANO FRIZZO STEFENON[4]

[1]Graduate Program in Electrical Engineering, Federal University of Paraná, Curitiba 81530-000, Brazil
[2]Graduate Program in Computer Science, State University of Maringá, Maringá 4200-465, Brazil
[3]Graduate Program in Mechanical Engineering, Federal University of Paraná, Curitiba 81530-000, Brazil
[4]Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, 1959-007 Lisbon, Portugal

Corresponding author: Viviana Cocco Mariani (viviana.mariani@ufpr.br)

This work was supported in part by the Coordination for the Improvement of Higher Education Personnel (CAPES) through Transformative Agreement; and in part by the National Council for Scientific and Technological Development (CNPq), Brazil, under Grant 314389/2023-7-PQ, Grant 313169/2023-3-PQ, Grant 407453/2023-7-Universal, Grant 402015/2025-8-Universal, and Grant 305910/2024-8-PQ).

ABSTRACT Accurate photovoltaic (PV) power forecasting remains challenging due to nonlinear dynamics, rapid weather variations, and multi-scale temporal dependencies. This paper proposes ConvChebyKAN, a hybrid deep learning architecture that integrates multi-scale causal convolutions, Chebyshev polynomial parameterized Kolmogorov Arnold networks (KANs), and conformal prediction to improve both deterministic and probabilistic forecasting performance. The multi-scale convolutional module captures short-term fluctuations and long-term temporal dependencies, while the Chebyshev-based KAN improves nonlinear representation capability and numerical stability. A multi-objective hyperparameter optimization strategy based on Pareto efficiency and TOPSIS ranking is used to jointly optimize forecasting accuracy, prediction interval sharpness, and coverage reliability. The model is evaluated using real photovoltaic generation and meteorological data from four seasonal datasets and forecasting horizons from 1 to 48 hours ahead. ConvChebyKAN reaches an MAE of 149.13 kWh, RMSE of 227.34 kWh, R^2 of 98.2%, and sMAPE of 6.21%, outperforming KAN and other machine learning models. These results show that the proposed architecture effectively captures multi-scale temporal patterns while providing reliable uncertainty estimates, supporting PV forecasting applications in power system operation and energy management.

INDEX TERMS Conformal prediction, deep learning, Kolmogorov-Arnold networks, multi-objective optimization, photovoltaic power forecasting.

I. INTRODUCTION

The increasing penetration of photovoltaic (PV) generation has amplified the operational challenges associated with its inherent variability, motivating the development of accurate and reliable short-term forecasting methods [1]. Although forecasting performance has improved substantially in recent years, PV generation remains difficult to predict due to the combination of smooth diurnal cycles and abrupt, short-lived fluctuations caused by rapidly changing weather

The associate editor coordinating the review of this manuscript and approving it for publication was Jolanta Mizera-Pietraszko.

conditions. These dynamics require models capable of capturing multi-scale temporal behavior while remaining robust to noise and inconsistent atmospheric patterns [2].

Standard machine learning (ML) and deep learning (DL) models, including neural networks [3], convolutional [4], transformer [5], and other models, provide strong baselines, yet often face persistent limitations. First, many architectures struggle to simultaneously learn long-term diurnal structure and sharp local transients, leading to error accumulation across forecasting horizons [6]. Second, models that rely solely on fixed activation functions or spline-based parameterizations may become numerically unstable or overly smooth under noisy PV measurements. Also, uncertainty quantification is frequently treated as a separate task, resulting in unreliable or poorly calibrated prediction intervals that limit practical use in grid operation and energy management [7]. Ahmad et al. [8] highlight that the use of artificial intelligence (AI) models can significantly improve the accuracy of solar irradiance forecasts. This enables more efficient operation of PV systems, reducing waste and increasing grid reliability. Recently, Wang et al. [9] proposed a physics-informed neural network with temperature correction to improve PV power prediction under varying meteorological conditions.

These gaps motivate the need for a unified forecasting framework that combines expressive functional approximation, efficient multi-scale feature extraction, and principled interval calibration. To address this challenge, this study proposes a hybrid model named Convolutional Chebyshev Kolmogorov-Arnold network (ConvChebyKAN) that integrates causal multi-scale convolutions with a flexible Kolmogorov-Arnold network (KAN) and a Chebyshev-parameterized essence. The convolution module processes lagged PV power and meteorological variables (clouds, irradiance, temperature, and wind) over multiple window lengths to capture both rapid fluctuations and longer trends using historical data, while the functional-approximation layers represent broader nonlinear patterns with improved numerical stability [10].

KAN enhances standard multilayer Perceptrons (MLPs) by replacing fixed activation functions with learnable, univariate spline operators along the edges [11]. This architectural shift leverages the Kolmogorov-Arnold theorem to offer a superior function approximation, leading to models that are often more compact, achieve comparable accuracy, and possess inherent interpretability due to the direct visualization of each univariate function [12]. Previous work has examined KAN for nonlinear regression, surrogate modeling [13], and time-series forecasting in general [14], while only a few studies have applied it to PV power forecasting [15], [16], Chebyshev-parameterized variants, in particular, remain unexplored.

While KAN offers strong nonlinear approximation capabilities, its direct application to time-series forecasting remains limited due to the lack of mechanisms for hierarchical temporal feature extraction. The proposed ConvChebyKAN architecture addresses this limitation by integrating causal multi-scale convolutions with Chebyshev-parameterized functional mappings, enabling simultaneous modeling of short-term fluctuations and long-term diurnal dynamics.

ConvChebyKAN combines multi-objective hyperparameter optimization with uncertainty quantification through conformal prediction, building 90% reliable prediction intervals on a training set. The optimization minimizes the mean absolute error (MAE) of validation and the prediction interval metrics' width and absolute coverage error, filtering candidate configurations through a Pareto set and ranking them with the TOPSIS to balance accuracy, interval narrowing, and coverage.

Data from a Brazilian PV plant are used, covering four seasons and six forecast time horizons from 1 to 48 hours ahead. ConvChebyKAN is compared with ML and DL models, and forecast performance is evaluated using MAE, mean squared error (RMSE), coefficient of determination (R^2), and symmetric mean absolute percent error (sMAPE), while uncertainty quantification is assessed using forecast interval sharpness and marginal coverage.

The main contributions of this study are summarized as follows:

- Proposal of a hybrid forecasting model that integrates multiscale causal convolutions with a Chebyshev-parameterized KAN, called ConvChebyKAN, enhancing the model's ability to jointly capture smooth diurnal behavior and rapid transient fluctuations in PV energy forecasting.
- Development of a functional structure possible by Chebyshev parameterization, which mitigates the problems of numerical instability and excessive smoothing commonly observed in traditional deep neural networks and spline-based activation schemes.
- The use of a multi-objective learning strategy and model selection based on Pareto efficiency and TOPSIS classification allows balanced optimization of deterministic accuracy, interval sharpness, and coverage across multiple forecast horizons.
- A principled distribution-free interval estimation approach using split conformal prediction to generate well-calibrated prediction intervals, ensuring reliable coverage under diverse irradiance conditions.
- A seasonal and multihorizon evaluation, demonstrating that the ConvChebyKAN model delivers improvements in performance metrics relative to established ML and DL baselines.
- Analysis of horizon-dependent forecasting behavior across four distinct seasonal regimes, revealing performance degradation patterns and identifying horizon ranges where hybrid architectures outperform other forecasting methods.

Thus, this study addresses three persistent challenges in PV forecasting: (i) capturing multi-scale temporal dynamics,

(ii) ensuring numerical stability in functional neural architectures, and (iii) producing calibrated uncertainty estimates.

The remainder of this paper is organized as follows. Section II reviews related work concerning PV forecasting using ML, hybrid, and ensemble models. Next, Section III details the proposed methodology, which covers the data, the models, and the multi-objective conformal framework. Section IV then presents the numerical results and offers a comparative analysis. To conclude, Section V provides a synthesis of our results and implications.

II. RELATED WORKS

This section reviews recent advances in PV power forecasting, emphasizing comparative evidence across modeling paradigms. The literature is organized into four main categories: (i) ML, (ii) DL, (iii) hybrid and ensemble models, and (iv) probabilistic and multi-objective forecasting. This structure highlights recurring methodological limitations and clarifies how the proposed model builds upon and extends existing work.

ML models, such as support vector machines (SVMs) [17], artificial neural networks (ANNs), and tree-based ensembles [18], established the foundation for nonlinear PV forecasting. SVM and ANNs components have been integrated with sky-imaging techniques to mitigate weather-driven variability [19], while decision-tree ensembles have been applied to ultra-short-term irradiance forecasting with fast retraining and feature selection [20]. Physics-informed strategies combining ML models with simplified physical representations have also demonstrated consistent benefits for multi-day horizons [21]. Although effective under stable weather conditions, these approaches often struggle with rapid transients and exhibit limited capacity to model long-range dependencies-issues that motivated the transition toward DL-based architectures.

DL models have advanced PV forecasting by capturing richer temporal patterns and integrating exogenous meteorological signals [22]. Convolutional neural network (CNN) models have been used to extract sky-image features for short-term forecasting [23], while recurrent neural network (RNN) and long short-term memory (LSTM) architectures have proven effective for medium- and long-term horizons, including monthly forecasting across multiple sites [24]. Gated recurrent unit and extreme gradient boosting (GRU-XGBoost) hybridization has shown improvements in multi-day hourly forecasting [25].

Specialized CNN variants, such as SolarNet [26], incorporate parallel pooling to enhance robustness across seasons. Studies comparing a range of statistical and DL models [27] show that DL methods generally outperform classical baselines, but performance remains sensitive to the choice of horizon, input composition, and irradiance regime. Despite these gains, deep networks may oversmooth fine-scale fluctuations or require careful architectural tuning to avoid overfitting and numerical instability gaps that motivate hybrid designs integrating multiscale feature extraction and more flexible nonlinear approximators.

Hybrid models combine complementary forecasting components to better handle the multi-scale nature of PV data. LSTM combined with a discrete grey model has been used for day-ahead forecasting under non-ideal conditions [28], while CNN-LSTM ensembles have been applied to multi-dataset irradiance prediction [29]. Autoregressive integrated moving average [30] (ARIMA) with ANN hybrids have shown gains in daily radiation forecasting [31]. More recent architectures combine LSTM with temporal convolutional networks (TCN) to capture both long- and short-term dynamics [32]. Bi-LSTM ensembles incorporating weather, panel characteristics, and spatiotemporal interactions have been proposed for short- and long-term prediction [33]. Hybrid models that decompose signals before forecasting, such as variational mode decomposition (VMD) [34] or adaptive mode decomposition with ML optimization [35] show strong performance by separating smooth and oscillatory components.

Recent studies emphasize the role of hybrid and ensemble architectures in advancing PV power forecasting. Hybrid recurrent-convolutional models such as the ensemble-optimized BiGRU of Dai et al. [36] combine feature selection, attention mechanisms, and bagging to improve short-term accuracy across sites with differing environmental profiles. Multimodal ensemble strategies further reinforce this trend: Ma et al. [37] demonstrate that combining meteorological variables, all-sky imagery, and power measurements within a stacked ensemble enhances ultra-short-term prediction performance. Multi-scale frameworks, including physics-constrained decomposition models [38] and wavelet-informed Transformer architectures [39], leverage structured temporal representations to enhance stability under variable irradiance. However, these hybrid and multi-scale approaches often require complex coordination between multiple components and still lack a unified framework capable of jointly capturing multiscale temporal structure, nonlinear expressiveness, and uncertainty characterization.

Ensuring reliable performance under rapidly changing atmospheric conditions remains an important challenge, even as the previous studies primarily generate deterministic forecasts. Ensemble aggregation [36], [37] provides partial robustness by reducing sensitivity to noise, missing information, and localized disturbances, while decomposition-based architectures [38] and wavelet-Transformer hybrids [39] improve stability by isolating scale-specific components and clarifying the influence of meteorological drivers. Yet these methods address uncertainty only implicitly and do not yield calibrated uncertainty sets. This limitation motivates the integration of post-hoc, distribution-free techniques, such as conformal prediction, which can complement deterministic forecasting architectures with adaptive and valid uncertainty intervals without modifying their internal structure.

Multi-objective optimization has emerged as a complementary direction: PV forecasting systems optimized for Pareto tradeoffs have shown improved accuracy and robustness across horizons [34], while energy management studies incorporate ML-based forecasting into cost-optimization frameworks [40]. Despite these advances, most probabilistic or multi-objective methods treat uncertainty estimation as a separate post-processing step rather than integrating it into the forecasting architecture.

Based on previous literature, it is noted that forecast accuracy improves with models that explicitly capture nonlinear temporal dynamics, integrate multiscale information, or combine complementary learning modules. However, existing approaches generally face limitations such as (i) difficulty in modeling smooth diurnal cycles and abrupt, short-duration transients in a single architecture, (ii) limited numerical flexibility of deep networks, which often rely on fixed activation functions or spline-based parameterizations with stability constraints, and (iii) uncertainty quantification treated as an external component, leading to inconsistent or poorly calibrated forecast intervals.

Motivated by these gaps, the present study proposes a unified forecasting model that integrates multiscale causal convolutions, a Chebyshev-parameterized KAN, and a multi-objective conformal forecasting framework. The hybrid model ConvChebyKAN aims to provide stable, interpretable, and accurate deterministic forecasts, offering uncertainty estimates across different horizons and seasonal regimes.

III. MATERIALS AND METHODOLOGY

This section presents the methodological framework adopted for PV energy forecasting. It covers the complete workflow, including data acquisition and preprocessing, with emphasis on outlier treatment and data normalization, followed by the feature selection procedure. The theoretical basis and experimental design of the proposed model are then detailed, and the benchmark models used for comparative evaluation are introduced.

A. DATA ACQUISITION AND EXPLORATORY DATA ANALYSIS

The subsections that follow detail the data preparation procedures adopted in this study. The characteristics of the data sources are first described to contextualize the dataset, followed by the data-cleaning steps used to ensure integrity and consistency. Subsequently, all variables are standardized through z-score scaling to place them on a comparable scale, and a feature selection strategy based on the ϕ_K correlation is applied to identify the most informative predictors for model development.

This study uses an original time-series dataset from the PLIN Energia solar plant in Loanda, Paraná, Brazil (latitude −22.90835°, longitude −53.112785°; elevation 495 m), whose geographic position is shown in Figure 1, with the red marker indicating the plant's exact location. The plant consists of five 1 MW lots, with a total inverter capacity of 5 MW and an installed capacity of 6.65 MWp, comprising 16 400 PV modules rated at 400 W each.

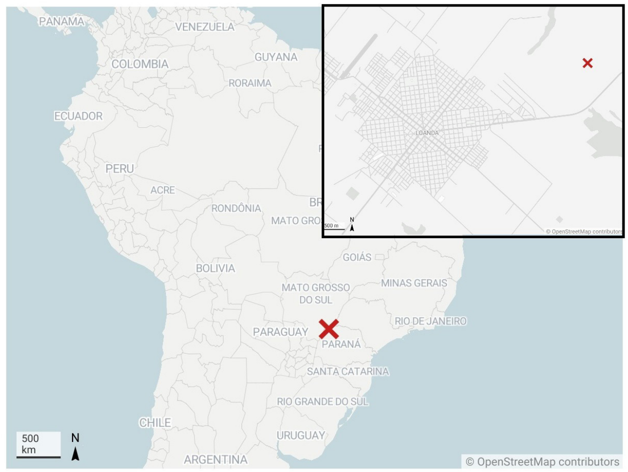

FIGURE 1. Geographical location of the PLIN solar plant.

Measurements were collected at 1-minute frequency from December 21, 2021, to December 20, 2022. The record is segmented into summer, autumn, winter, and spring, with 90, 91, 94, and 89 days, respectively. Daily analysis is restricted to 06:00-18:00 to ensure consistent diurnal coverage. Seasonal sample counts are 1,080 (summer), 1,104 (autumn), 1,128 (winter), and 1,068 (spring). Series were resampled to 1-hour intervals to facilitate analysis while preserving diurnal structure.

Plant-side variables comprise PV power (kWh) and panel temperature (°C). Meteorological exogenous variables were obtained from the Open-Meteo Historical Weather API (Application Programming Interface) and time-aligned to the plant measurements at matched timestamps (Loanda, PR). The weather set includes relative humidity (2 m), cloud cover, wind gusts (10 m), wind speed (10 m), air temperature (2 m), and shortwave radiation (global horizontal irradiance). All variables were concatenated into a single time-aligned dataset. The overview of the plant is shown in Figure 2.

FIGURE 2. Aerial view of the PV arrays of the PLIN solar plant.

Table 1 presents the descriptive statistics of PV power generation and associated meteorological variables across four seasons. Each dataset includes global horizontal irradiance

TABLE 1. Descriptive statistics of seasonal environmental and PV power variables.

Season	Samples	Features	Acronym	Unit	Mean	Std	Minimum	Maximum	Skewness	Kurtosis
Spring	1,068	PV power	PV power	kWh	2741.5401	1509.7301	42.5901	5035.0301	-0.2701	-1.1001
		G. horiz. irradiance	GHI	W/m^2	563.5001	286.5301	33.0000	1010.4001	-0.2301	-1.1801
		Wind gusts	WG	km/h	25.3401	4.5801	13.7001	33.9601	-0.1601	-0.5001
		Wind speed	WS	km/h	10.0101	2.5201	4.6201	14.7401	-0.0701	-0.7901
		Relative humidity	RH	%	61.6901	12.7201	40.0000	92.0000	0.3401	-0.4901
		Panel temperature	T_{panel}	°C	40.0801	10.4901	21.1901	57.7801	-0.1301	-1.0901
		Air temperature	T_{air}	°C	25.2501	2.9601	18.4201	29.8401	-0.4401	-0.6801
Summer	1,080	PV power	PV power	kWh	2639.9801	1408.1501	53.6601	4887.3101	-0.3801	-1.1401
		G. horiz. irradiance	GHI	W/m^2	584.4701	298.8101	10.8001	1000.2001	-0.3801	-1.1001
		Wind gusts	WG	km/h	23.6201	4.5801	13.3001	31.7001	-0.2001	-0.5701
		Wind speed	WS	km/h	8.8701	2.1901	3.7001	12.8201	-0.2101	-0.5401
		Relative humidity	RH	%	56.6401	14.2101	36.0000	92.2001	0.6701	-0.2801
		Panel temperature	T_{panel}	°C	45.4801	11.4401	25.2701	63.1701	-0.2701	-1.1901
		Air temperature	T_{air}	°C	28.6001	3.2701	21.3001	33.4201	-0.5501	-0.6601
Autumn	1,104	PV power	PV power	kWh	2376.0301	1645.7901	0.0000	5000.1101	-0.0601	-1.3801
		G. horiz.irradiance	GHI	W/m^2	393.6701	259.2101	0.0000	808.4001	-0.1601	-1.3801
		Wind gusts	WG	km/h	25.0801	5.1301	11.5001	34.2001	-0.3101	-0.2801
		Wind speed	WS	km/h	10.6801	2.4401	5.1801	15.2801	-0.2801	-0.6201
		Relative humidity	RH	%	63.5001	12.0501	45.0000	92.0000	0.4301	-0.5801
		Panel temperature	T_{panel}	°C	32.9801	10.5801	16.4501	53.4201	0.1801	-1.0901
		Air temperature	T_{air}	°C	22.9701	3.6001	15.1001	29.4001	-0.1101	-0.6401
Winter	1,128	PV power	PV power	kWh	2466.2401	1587.9601	0.0101	4741.3901	-0.2801	-1.3901
		G. horiz. irradiance	GHI	W/m^2	392.3601	253.2401	0.0000	780.4001	-0.2101	-1.4001
		Wind gusts	WG	km/h	27.1401	5.2601	12.6001	36.1601	-0.3401	-0.1301
		Wind speed	WS	km/h	11.9401	2.5801	5.8001	16.3801	-0.4201	-0.2801
		Relative humidity	RH	%	55.2301	12.8301	34.0000	86.8001	0.4801	-0.4201
		Panel temperature	T_{panel}	°C	34.2901	10.5801	15.8201	50.1801	-0.3301	-1.2401
		Air temperature	T_{air}	°C	22.6601	3.5801	15.4001	28.6001	-0.2701	-0.9401

(GHI), panel temperature (T_{panel}), air temperature (T_{air}), wind speed (WS), wind gusts (WG), and relative humidity (RH). The seasonal samples range from 1,068 to 1,128 observations. Results show that spring exhibits the highest mean PV power (2,741.54 kWh), while summer presents the highest irradiance (584.47 W/m^2), corresponding to elevated panel and air temperatures. In contrast, autumn presents the lowest irradiance and PV output, while winter demonstrates slightly higher wind speeds and gusts, indicating greater atmospheric turbulence. Relative humidity remains within moderate ranges, approximately 55.23-63.50%, throughout the year, with minimal skewness and kurtosis, suggesting near-normal distributions.

GHI exhibits a distinct seasonal pattern, reaching its maximum during summer, mean of 584.47 W/m^2, followed by spring, 563.50 W/m^2, while autumn and winter present substantially lower and comparable averages, 393.67 and 392.36 W/m^2, respectively. The mean PV power generation mirrors this trend, attaining its highest values in spring and summer, 2,741.54 and 2,639.98 kWh, respectively, and declining during winter and autumn to 2,466.24 and 2,376.03 kWh. The slightly reduced PV output in summer compared to spring is consistent with the higher average module temperature observed in this period, 45.48 °C in summer versus 40.08 °C in spring, which negatively affects conversion efficiency. The variability of these parameters also reflects the seasonal regimes, GHI has its largest standard deviation in summer, 298.81 W/m^2, whereas PV power output shows greater variability in autumn and winter, 1,645.79 and 1,587.96 kWh, respectively, with summer displaying comparatively more stable generation levels.

Wind conditions have marked seasonal variability, intensifying during the colder months and weakening in summer. Relative humidity reaches its maximum in autumn, 63.50%, and minimum in winter, 55.23%. The statistical distributions of PV power, global horizontal irradiance, and temperature are generally mildly left-skewed with light tails and negative kurtosis, while RH exhibits positive skewness across all seasons. Near-zero minimum values of GHI and PV power during autumn and winter indicate the occurrence of daytime periods with very low solar irradiance under overcast conditions. Overall, the data highlight clear seasonal dependencies between irradiance, temperature, and PV power generation, confirming that thermal and radiative effects significantly influence PV performance and system efficiency.

The ridgeline plots in Figure 3 visualize the seasonal probability distributions of all variables, complementing the descriptive statistics presented previously. Rather than summarizing central tendencies, these density curves reveal the spread, multimodality, and distributional shifts that characterize each season, providing additional insight into the variability of irradiance, temperature, wind conditions, and PV output.

For PV power and GHI (Figures 3a and 3b, respectively), spring displays a distinct concentration of higher values, reflecting the combination of clear-sky conditions and moderate temperatures that support efficient energy conversion. Summer shows a wider distribution of values because high module temperatures reduce the efficiency of PV energy,

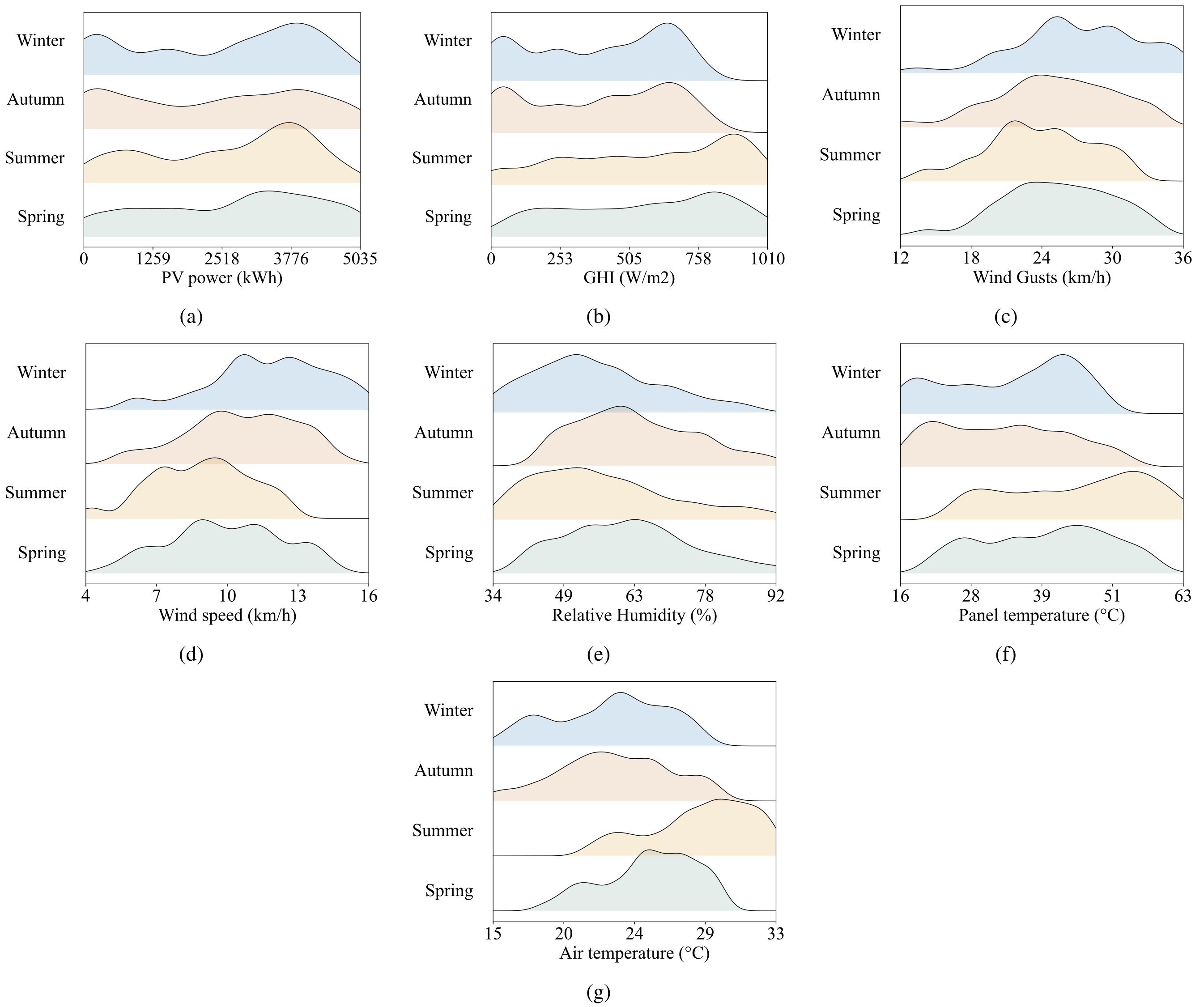

FIGURE 3. Ridgeline density plots of the seasonal distributions for: (a) PV power, (b) GHI, (c) wind gust, (d) wind speed, (e) relative humidity, (f) panel temperature, and (g) air temperature.

and the daily irradiance fluctuates more strongly, leading to greater dispersion in the PV power. The air and panel temperature distributions (Figures 3g and 3f, respectively) reflect this behavior, with summer showing an extended range towards higher temperatures and reduced cooling. Winter remains limited to lower temperatures because weaker sunlight and colder ambient air constrain thermal conditions.

Wind speed and gusts (Figures 3d and 3c, respectively) shift to higher and more variable values in winter, consistent with the more active weather patterns of the cold season, when stronger pressure gradients and frontal passages increase wind intensity. In summer, both variables remain concentrated at low values due to calmer and more stable atmospheric conditions. Relative humidity (Figure 3e) shows a wider distribution in autumn, consistent with frequent transitions between warm and cold air masses, while winter presents a narrower range at lower humidity levels because colder air retains less water vapor.

1) DATA CLEANING

Outliers were detected using the interquartile range (IQR) method applied independently to each hour of the day. For a given hour h, the first and third quartiles computed over the training window are denoted as $Q_1(h)$ and $Q_3(h)$, respectively, with

$$\mathrm{IQR}(h) = Q_3(h) - Q_1(h). \tag{1}$$

Observations at hour, h, were flagged as outliers when $x_t < Q_1(h) - 1.5,\mathrm{IQR}(h)$ or $x_t > Q_3(h) + 1.5\mathrm{IQR}(h)$. Detected outliers were removed and replaced with the hour-of-day mean computed on the training split, preserving the typical diurnal profile while repairing isolated anomalies.

2) NORMALIZATION

To normalize the data, all variables were scaled using the z-score defined by

$$z_t = \frac{x_t - \mu}{\sigma}, \tag{2}$$

where μ and σ are the mean and standard deviation, respectively, estimated on the training split for each variable individually.

3) FEATURE SELECTION

The correlation (ϕ_K) is employed as the ranking criteria for feature selection [41], motivated by the need to go beyond the limitations of the standard Pearson correlation, which measures the strength of linear relationships, often insufficient for capturing the complex, non-linear dependencies inherent in PV power generation [42]. For each candidate predictor X_j (with j indexing all predictors) and the target variable Y, the relationship is represented through an $r \times c$ contingency table constructed using adaptive binning for continuous variables. Given a sample size n, the observed Pearson chi-square statistic χ^2_{obs}, and the corresponding independence (bias-corrected) term χ^2_{ind}, the ϕ_K value is derived accordingly.

$$\phi_K(X_j, Y) = \sqrt{\frac{\max(0,\ \chi^2_{\text{obs}} - \chi^2_{\text{ind}})}{n \min(r-1, c-1)}}, \tag{3}$$

where r and c are the numbers of bins/categories for X_j and Y, respectively, and $\min(r-1, c-1)$ is the effective degrees of freedom.

Figure 4 presents the ϕ_{K} correlation matrices by season, illustrating the degree of association between meteorological parameters and PV power generation. The selected predictors were GHI, module, and air temperatures (T_{panel}, T_{air}), wind speed and gusts (WS, WG), and relative humidity (RH). The following predictors: global radiation, cloud-cover fraction, surface pressure, and precipitation were also evaluated but removed due to redundant associations.

In Spring, PV relates most to T_{panel} at 0.846 with GHI close at 0.820, and WG reaches 0.687, indicating meaningful wind effects in this season. In Summer, GHI leads at 0.879 while T_{panel} is 0.842 and RH and T_{air} rise to 0.666 and 0.677, pointing to stronger atmospheric modulation when irradiance is high. In Autumn, T_{panel} is 0.814, and GHI is 0.804, a narrow gap that suggests thermal and irradiance drivers are nearly balanced as conditions cool. In Winter, T_{panel} is 0.864 and GHI is 0.852, and WS climbs to 0.703, the seasonal maximum, highlighting that wind contributes more variance alongside the core irradiance-temperature pair in cold conditions.

B. PROPOSED MODEL ConvChebyKAN

The forecasting stage is structured around the proposed ConvChebyKAN architecture, which constitutes the central contribution of this study and is presented in detail in the following subsection. To provide a comprehensive performance assessment, a diverse set of comparative forecasting models is additionally employed, spanning linear regularized regressors, ensemble-based learners, and state-of-the-art deep-learning approaches. ConvChebyKAN is formulated as an extension of the KAN. For completeness, the original KAN formulation is presented first to establish the mathematical structure upon which the proposed architecture is developed. The subsequent subsection details the modifications introduced in ConvChebyKAN, including the incorporation of causal multi-scale convolutions and the replacement of spline-based edge functions with Chebyshev polynomial parameterizations.

KAN was proposed by Liu et al. [43], which has its foundation in the Kolmogorov-Arnold representation theorem [44], [45] stating that any multivariate continuous function can be decomposed into a sum of univariate continuous functions. KANs represent an appropriate alternative to traditional MLPs and their variants, offering wide-ranging application potential in tasks like time-series forecasting [46], [47], [48], [49], [50]. Mathematically, the Kolmogorov-Arnold representation theorem can be expressed as an approximation [51]:

$$f(x_1, x_2, x_3, \ldots, x_n) = \sum_{q=1}^{2n+1} \phi_q \left(\sum_{p=1}^{n} \varphi_{q,p}(x_p) \right), \tag{4}$$

where each $\phi_q : \mathbb{R} \to \mathbb{R}$ and $\varphi_{q,p} : [0,1] \to \mathbb{R}$ is a continuous univariate function that maps input variables into a result through a sum of simpler continuous functions. The expression $\sum_{p=1}^{n} \varphi_{q,p}(x_p)$ shows how the univariate functions are combined. The quantity of these univariate functions is inherently limited by the number of input features. Specifically, for a model with n input features, only $2n+1$ univariate functions are required to accurately represent a continuous function.

Liu et al. [43] expanded the Kolmogorov-Arnold representation to networks of arbitrary depth. This representation is extended by composing the outer functions ϕ_q with the inner functions $\varphi_{q,p}$ across layers. The network architecture is defined by an integer array $[n_0, n_1, \ldots, n_L]$, where n_L represents the number of neurons in the L-th layer. Each layer transforms an input vector $\boldsymbol{x}_l \in \mathbb{R}^{n_l}$ into an output vector $\boldsymbol{x}_{l+1} \in \mathbb{R}^{n_{l+1}}$:

$$\boldsymbol{x}_{l+1} = \begin{pmatrix} \varphi_{1,1,l} & \cdots & \varphi_{1,n_l,l} \\ \vdots & \ddots & \vdots \\ \varphi_{n_{l+1},1,l} & \cdots & \varphi_{n_{l+1},n_l,l} \end{pmatrix} \boldsymbol{x}_l. \tag{5}$$

The function matrix for each layer l is designated as Φ_l, with

$$\Phi_l = \begin{pmatrix} \varphi_{l,1,1}(\cdot) & \cdots & \varphi_{l,1,n_l}(\cdot) \\ \vdots & \ddots & \vdots \\ \varphi_{l,n_{l+1},1}(\cdot) & \cdots & \varphi_{l,n_{l+1},n_l}(\cdot) \end{pmatrix}. \tag{6}$$

In each l-th layer, the activation value of a neuron in the following layer is computed as the sum of the post-activation values from the previous layer. This hierarchical composition

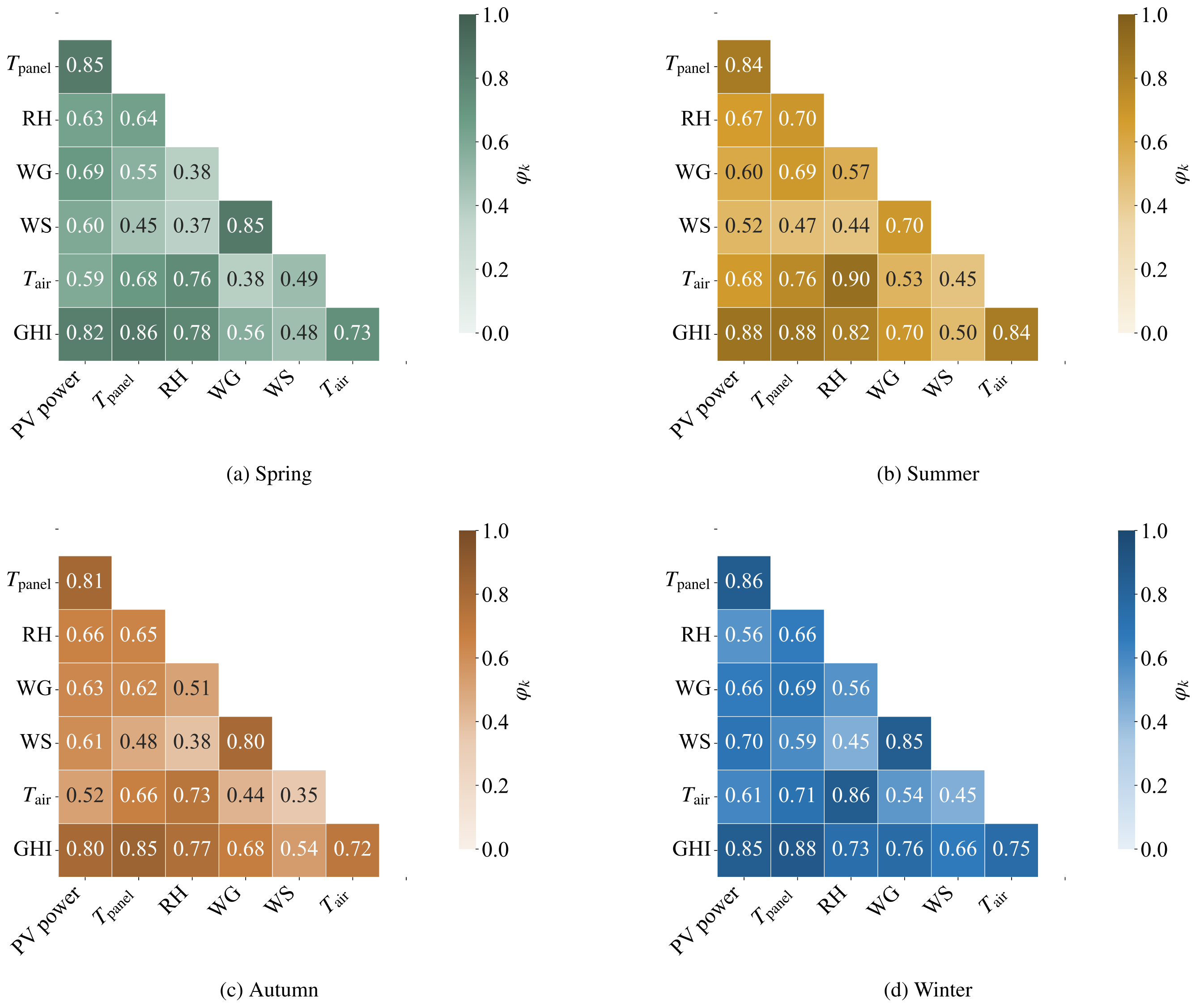

FIGURE 4. $\phi_{\mathbf{k}}$ **correlation matrices by season.**

allows KAN to approximate complex multivariate functions via a sequence of recursive operations. The overall mapping can be written compactly as

$$\mathrm{KAN}(x) = \left(\Phi_{L-1} \circ \Phi_{L-2} \circ \cdots \circ \Phi_0\right)\boldsymbol{x}. \tag{7}$$

The mathematical description of the layerwise aggregation follows:

$$f(\boldsymbol{x}) = \sum_{i_L=1}^{n_L} \phi_{i_L}\left(x_{i_L}^{(L)}\right), \tag{8}$$

where every intermediate scalar activation $x_{i_L}^{(L)}$ is generated recursively by summing nonlinear transformations applied to outputs from the prior layer:

$$x_{i_L}^{(L)} = \sum_{i_{L-1}=1}^{n_{L-1}} \varphi_{L-1,i_L,i_{L-1}}\left(x_{i_{L-1}}^{(L-1)}\right). \tag{9}$$

In the original implementation of KAN [43], each layer combines spline functions (piecewise polynomials) with sigmoid linear unit (SiLU) activations, enabling adaptable function approximation. A general L-layer KAN can be expressed elementwise as

$$x_{l+1,j} = \sum_{i=1}^{n_l} \phi_{l,j,i}\left(x_{l,i}\right), \tag{10}$$

where $\phi_{l,j,i}$ are spline-based univariate activation functions defined along the edges. These layer-by-layer transformations generate the final output.

Figure 5 illustrates a schematic of the KAN architecture, which incorporates the layered transformation process outlined earlier. Each level in the KAN framework applies learnable, spline-based functions to inputs, replacing fixed weights with adaptable nonlinear mappings. This enables

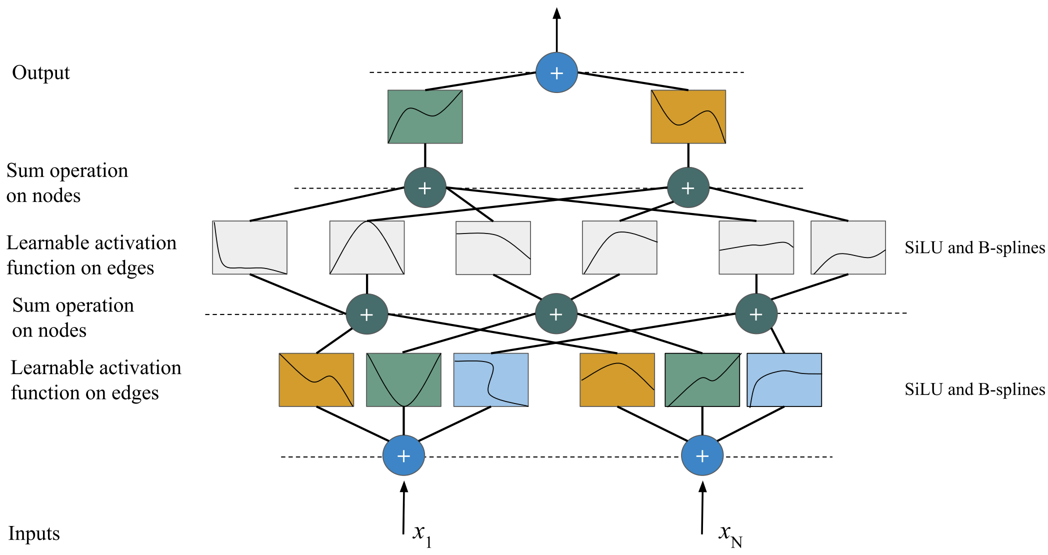

FIGURE 5. Structural overview of the KAN architecture.

KANs to approximate complex, nonlinear functions by capturing intricate dependencies across multiple levels.

The hybrid ConvChebyKAN model proposed in this study extends the spline-based KAN architecture through two targeted enhancements for time-series forecasting. First, a causal multi-scale one-dimensional convolutional (Conv1D) front-end is introduced to extract hierarchical temporal representations from the input sequence while strictly enforcing causality, ensuring that predictions depend exclusively on past observations. Second, the spline-based functional mappings within the KAN backbone are reformulated using a Chebyshev polynomial parameterization. This substitution leverages the orthogonality and favorable conditioning of the Chebyshev basis to mitigate numerical instabilities and overfitting commonly associated with spline interpolation. The additive compositional structure inherent to the KAN framework is preserved, allowing ConvChebyKAN to maintain its theoretical expressiveness while improving robustness, stability, and generalization performance.

Let $\boldsymbol{x}_t \in \mathbb{R}^F$ be the vector of F exogenous variables at time t, and define the length-L lag window $\boldsymbol{X}_{t-L+1:t} = [\boldsymbol{x}_{t-L+1}, \ldots, \boldsymbol{x}_t] \in \mathbb{R}^{L\times F}$. The Conv1D front end stacks R strictly causal temporal convolutions along the time axis of $\boldsymbol{X}_{t-L+1:t}$. Each block uses a kernel of length k and a dilation d, with causality enforced by prefixing $d(k-1)$ zeros so that indices beyond time t are never accessed. A generic update at time t can be written as

$$\boldsymbol{h}_t = \sum_{i=0}^{k-1} \boldsymbol{W}_i \boldsymbol{x}_{t-id} + \boldsymbol{b}, \text{ with } d(k-1) \text{ zeros padded at the start,} \tag{11}$$

where $\boldsymbol{W}_i$ and $\boldsymbol{b}$ are learnable parameters. Multi-scale receptive fields arise from using multiple (k_r, d_r) configurations, enabling the model to capture both smooth diurnal dynamics and sharp high-frequency fluctuations.

In the backbone, each KAN edge function is reparameterized using Chebyshev polynomials of the first kind. Let $T_s(u)$ denote the polynomials defined on $[-1, 1]$ with

$$T_0(u) = 1, \quad T_1(u) = u, \quad T_{s+1}(u) = 2uT_s(u) - T_{s-1}(u). \tag{12}$$

Each univariate mapping takes the form

$$\phi_{\ell,j,i}(x_{\ell,i}) = \sum_{s=0}^{m} a_{\ell,j,i,s} T_s\big(\alpha_{\ell,i} x_{\ell,i} + \beta_{\ell,i}\big), \tag{13}$$

where the coefficients $a_{\ell,j,i,s}$ and scaling parameters $(\alpha_{\ell,i}, \beta_{\ell,i})$ are learned. The Chebyshev expansion improves conditioning, avoids knot selection, and reduces sensitivity to boundary behavior.

Figure 6 presents the full ConvChebyKAN architecture, showing the lagged inputs, causal convolutional pathways, and the Chebyshev-parameterized KAN backbone. In summary, ConvChebyKAN preserves the functional design of KAN while introducing multi-scale causal convolutions and a polynomial reparameterization that improves numerical stability and forecasting performance.

C. BASELINE MODELS

For comparative evaluation, a set of established forecasting models is employed to contextualize the performance of the proposed ConvChebyKAN architecture. The benchmark group includes the KAN [43], [44], [45], [46], [47], [48], [49], [50], [51], used here as the principal nonlinear reference

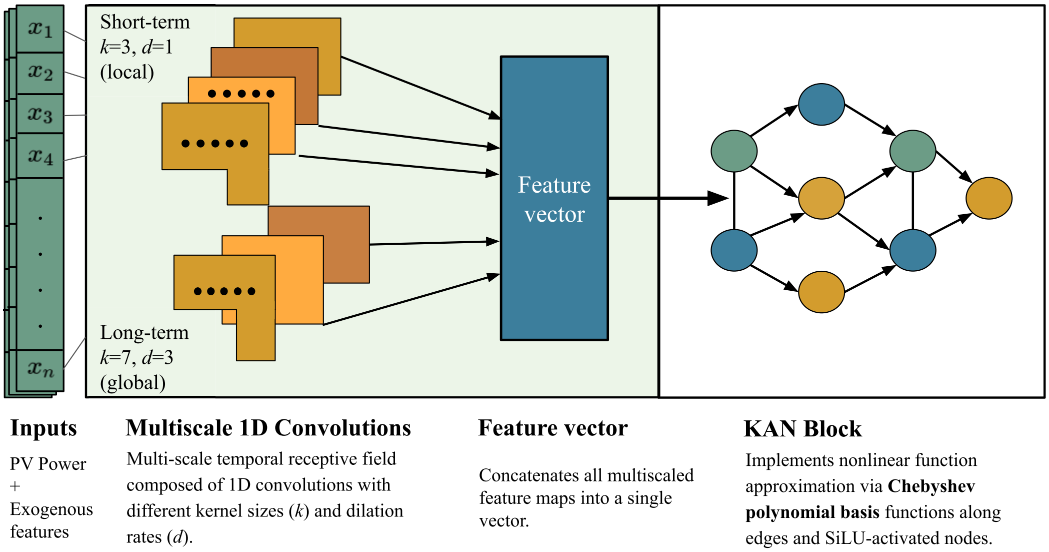

FIGURE 6. Structural overview of the ConvChebyKAN architecture.

due to its functional-compositional formulation. In addition, classical linear models such as Ridge regression [52] and LASSO [53] are considered to assess the contribution of regularization and sparsity in high-dimensional predictor spaces. Ensemble-based approaches [54], including AdaBoost are incorporated to capture nonlinear relationships through boosted decision structures. An MLP is used as a baseline neural estimator, providing a generic nonlinear counterpart devoid of explicit temporal modeling.

To represent state-of-the-art deep-learning architectures for time-series forecasting, the evaluation further includes neural hierarchical interpolation for time series (NHITS) [55], [56], [57], which combines hierarchical multi-rate processing, and patch time series transformer (PatchTST) [3], [58], a transformer-based model leveraging patch tokenization to capture long-range temporal dependencies.

D. PERFORMANCE METRICS

The selection of the optimal forecasting model was guided by MAE, and after three other statistical performance indicators, including RMSE, R^2, and sMAPE were evaluated [59].

The MAE is defined as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n} |\hat{y}_t - y_t|. \tag{14}$$

where n denotes the total number of test predictions, y and $\hat{y}$ are the observed and forecast PV power, respectively, and $\bar{y}$ denotes the sample mean of the test set. This metric quantifies the average absolute forecast deviation, expressed in the original units of PV power, providing a directly interpretable measure of error magnitude [60]. In contrast to squared-error metrics, MAE is less sensitive to infrequent, extreme forecast errors.

The RMSE is given by:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} (\hat{y}_t - y_t)^2}. \tag{15}$$

By squaring the errors prior to averaging, RMSE assigns a disproportionately higher weight to larger deviations. This property is advantageous in applications where high-magnitude errors are associated with significant operational costs or safety risks [61]. RMSE is reported in the original units, facilitating direct comparison with physical system tolerances [42].

The R^2 is calculated as:

$$R^2 = 1 - \frac{\sum_{t=1}^{n} (\hat{y}_t - y_t)^2}{\sum_{t=1}^{n} (y_t - \bar{y})^2}. \tag{16}$$

This statistic represents the proportion of variance in the observed data explained by the forecasts. It permits scale-free comparison across different datasets or forecasting contexts. Note that R^2 values can be negative, indicating that the model's predictive performance is inferior to that of a simple mean predictor. This can occur during prolonged low-variance periods or when forecasts are systematically biased.

The sMAPE is defined as:

$$\mathrm{sMAPE} = \frac{100\%}{n}\sum_{t=1}^{n} \frac{|\hat{y}_t - y_t|}{(|y_t| + |\hat{y}_t|)/2}. \tag{17}$$

sMAPE expresses the forecast error as a percentage, normalized by the average of the observed and predicted values. This scale-invariance allows for meaningful comparisons across time series of differing magnitudes. In practice, a small constant $\varepsilon > 0$ is often added to the denominator to ensure numerical stability when both $|y_t|$ and $|\hat{y}_t|$ are near zero.

E. PREDICTIVE UNCERTAINTY METRICS

Split conformal prediction [62] is applied to the disjoint calibration set. The nonconformity scores are defined as the absolute residuals:

$$s_t = |y_t - \hat{y}_t|, \tag{18}$$

where y_t denotes the observed value and $\hat{y}_t$ the corresponding point forecast at time t. The conformal radius $\hat{q}_{0.10}$ is then estimated as the empirical 0.90 quantile of the nonconformity scores $\{s_t\}$ computed over the calibration set. Prediction intervals (PI) for a new forecast $\hat{y}_\tau$ are constructed as:

$$\mathrm{PI}_\tau = [\hat{y}_\tau - \hat{q}_{0.10},\ \hat{y}_\tau + \hat{q}_{0.10}], \tag{19}$$

This procedure preserves the original point forecasts and, under the assumption of exchangeability between future observations and the calibration sample, guarantees marginal coverage at the nominal level in finite samples.

To evaluate the interval forecasts, the statistical properties of the prediction intervals are assessed. Let $C_t = [L_t, U_t]$ denote the prediction interval generated at time t, where L_t and U_t represent the lower and upper bounds, respectively. The *marginal coverage* (MC) quantifies the empirical probability that the true observation falls within its corresponding prediction interval:

$$\mathrm{MC} = \frac{1}{n}\sum_{t=1}^{n} \mathbb{I}\{y_t \in C_t\}, \tag{20}$$

where $\mathbb{I}\{\cdot\}$ denotes the indicator function, y_t is the true observation at time t, and n is the total number of observations. This metric evaluates the statistical calibration of the uncertainty quantification. The model is considered well-calibrated when its marginal coverage approximates the nominal confidence level α (e.g., $\alpha = 0.9$ for 90% coverage) [63].

The *average sharpness* (AS) measures the mean width of the prediction intervals:

$$\mathrm{AS} = \frac{1}{n}\sum_{t=1}^{n} (U_t - L_t), \tag{21}$$

where U_t and L_t are defined as above. Sharpness characterizes the precision of the predictive uncertainty, with narrower intervals indicating more precise predictions [64]. For a fixed coverage level, intervals with smaller widths are preferred as they convey more actionable information. Consequently, coverage and sharpness metrics should be evaluated jointly to characterize the fundamental trade-off between reliability and precision in conformal prediction frameworks.

The miscoverage for the nominal 90% coverage level is defined as:

$$\mathrm{Miscoverage}_{0.90} = \left|\mathrm{MC}^{\mathrm{cal}} - 0.90\right|. \tag{22}$$

The $\mathrm{MC}^{\mathrm{cal}}$ represents the empirical marginal coverage computed over all calibration points under the one-step-ahead forecasting scheme. To obtain a more granular and lower-variance estimate of the residual distribution on the calibration set, forecasts are generated using a one-step-ahead strategy with fully overlapping origins. This ensures that every timestamp in the calibration set contributes to the collection of nonconformity scores $\{s_t\}$.

In the multi-objective hyperparameter tuning process, the prediction-interval metrics, specifically AS and $\mathrm{Miscoverage}_{0.90}$, are evaluated over the entire calibration set simultaneously. This design stabilizes the estimation of the conformal quantile $\hat{q}_{0.10}$ and yields more precise estimates of both marginal coverage and interval width. While this one-step-ahead calibration prioritizes the stability of the empirical quantiles, it should be noted that error accumulation at longer horizons may influence the reported marginal coverage values, making horizon-specific conformal calibration a relevant direction for future research.

F. EXPERIMENTAL SETUP

The input signals include lagged PV power and, when applicable, time-aligned lagged meteorological exogenous variables previously mentioned. The baseline machine-learning models: AdaBoost, LASSO, Ridge, and MLP, together with PatchTST and NHITS, are trained in a univariate configuration using only lagged PV power. In contrast, KAN and ConvChebyKAN use multivariate inputs that combine lagged PV power with the exogenous variables listed in Table 1.

This configuration aligns with the modeling paradigms commonly adopted in the literature for these respective architectures. The baseline models are employed to establish performance benchmarks by capturing autoregressive temporal dependencies from historical lags. Conversely, the KAN-based architectures are leveraged specifically for their capacity to interpret sequential complex, high-dimensional interactions between PV power generation and exogenous meteorological features.

The input window length L defines the sequence of values provided to each model by specifying how many past observations are used as predictors. For instance, when $L = 3$, the model receives the three most recent observations at times $t-3$, $t-2$, and $t-1$ to forecast values beyond t. In the multivariate setting, the same window is applied to every input channel so that, for instance, $L = 3$ provides $\{\mathrm{PV}_{t-3:t-1}\}$ together with aligned triplets for the previously mentioned exogenous variables. The value of L is selected separately for each model through hyperparameter tuning. All models are evaluated at forecasting horizons of 1, 3, 6, 12, 24, and 48 steps ahead.

1) FORECASTING STRATEGIES

Two distinct multi-step forecasting strategies are established: recursive forecasting and the multiple-input multiple-output (MIMO) approach. The primary difference lies in the mechanism by which multi-step horizons are generated. The selection of a strategy is dictated by the output structure of the specific model and its inherent capability to predict multiple steps concurrently [65].

- **Recursive forecasting**: A one-step-ahead predictive model is first estimated, then the forecast obtained for time $t + 1$ is subsequently fed back into the model as an input to generate the prediction for $t + 2$. This iterative process continues until the forecasting horizon H is reached, thereby constructing the complete forecast vector in a stepwise manner. Since each forecast depends on previously predicted values, errors may propagate and accumulate over successive steps. This recursive forecasting strategy is applied to the baseline ML models AdaBoost, LASSO, Ridge, and MLP, which are inherently single-output. Training a one-step model and iteratively extending it across horizons provides a consistent approach to multi-step forecasting without the need for multi-output architectures [66].
- **MIMO**: A single model maps the input window $\mathbf{x}_{t-L+1:t}$ directly to the entire future horizon $\hat{\mathbf{y}}_{t+1:t+H}$ in one forward pass. All forecast steps are generated at once from the same set of past observations, avoiding recursive feedback and limiting stepwise error accumulation while allowing dependencies across horizons to be learned jointly. Within this framework, KAN and ConvChebyKAN are trained with multivariate inputs composed of lagged PV power and the previously mentioned exogenous variables. PatchTST and NHITS follow the same strategy while using a univariate input that contains only PV power lags [67].

Figure 7 illustrates the multiple input multiple output configuration adopted for PV power forecasting, where the input consists of a window of length L, constructed from hourly daytime observations recorded between 7 and 18 h, resulting in twelve samples per calendar day. Each step represents one hour in the filtered series; hence, a single daytime period corresponds to twelve samples.

The model produces an H-step forecast vector in a single forward pass, predicting all future values simultaneously from the same input block without any recursive feedback. Three forecasting horizons are considered to examine the relationship between input size and prediction range. For $H = 48$, equivalent to four daytime days ahead, the input length equals the forecast horizon ($L = 48$). Similarly, for $H = 24$, the same relationship is maintained over two daytime days. Conversely, for $H = 12$, the model predicts one daytime day ahead while using an input twice as long ($L = 24$). This extended input window enables the model to capture longer-term contextual dependencies and enhances temporal consistency across consecutive days.

2) HYPERPARAMETER TUNING

Hyperparameter optimization is performed with the Bayesian optimization algorithm [68], considering the tree-structured Parzen estimator (TPE) [69]. TPE is a sequential model-based optimization approach that probabilistically models "good" and "bad" regions of the search space, dynamically proposing hyperparameter configurations with high expected improvement based on previous trials. The three-criteria minimization over the hyperparameter space is performed using the same calibration metrics defined in Sec. III-E:

$$\min_{\theta\in\Theta}\big(\mathrm{MAE}_{\mathrm{val}}(\theta),\ \mathrm{AS}_{\mathrm{cal}}(\theta),\ \mathrm{Miscoverage}_{0.90}(\theta)\big), \tag{23}$$

where θ denotes a hyperparameter configuration, $\mathrm{MAE}_{\mathrm{val}}$ is the validation MAE, $\mathrm{AS}_{\mathrm{cal}}$ is the *Average Sharpness* computed on the calibration set, and $\mathrm{Miscoverage}_{0.90}(\theta)$ is defined in Eq. (22) as $\left|\mathrm{MC}^{\mathrm{cal}}(\theta)-0.90\right|$, with $\mathrm{MC}^{\mathrm{cal}}$ denoting the Marginal Coverage on the calibration set.

To enhance computational efficiency, trials are executed concurrently through multiprocessing. In this distributed TPE framework, multiple workers interface with a shared study database, each executing an identical processing pipeline. The pipeline comprises the following steps: feature standardization using statistical parameters derived from the training set, model training on the training set, computation of validation MAE by aggregating all validation forecasts, calibration of prediction intervals on the dedicated calibration set, and finally, logging the three objective values.

The TPE sampler operates by constructing kernel density estimates from the historical records of completed trials. Subsequently, it samples new hyperparameter configurations that maximize the expected improvement (EI) criterion, effectively balancing exploration and exploitation in the search space. Following the evaluation, non-dominated solutions are identified to form the Pareto-optimal set. The TOPSIS [70] is then applied to rank these Pareto-optimal configurations, enabling the selection of a single optimal hyperparameter set for each distinct seasonal period and forecasting horizon.

3) TOPSIS RANKING

TOPSIS ranks alternatives by favoring those that are simultaneously closest to an ideal point and farthest from an anti-ideal point in objective space. After Pareto filtering, the remaining trials are scored on the three minimized objectives. Each objective column is normalized and equally weighted to form a three-dimensional vector v_i for configuration i. The ideal point a^* collects the columnwise minima and the anti-ideal a^- the columnwise maxima. The ranking is driven by the closeness coefficient

$$C_i = \frac{\|v_i - a^-\|_2}{\|v_i - a^-\|_2 + \|v_i - a^*\|_2}, \tag{24}$$

which increases as a configuration moves nearer the ideal and farther from the anti-ideal. Candidates are ordered by C_i in descending order; the top-ranked configuration is retained per season and horizon.

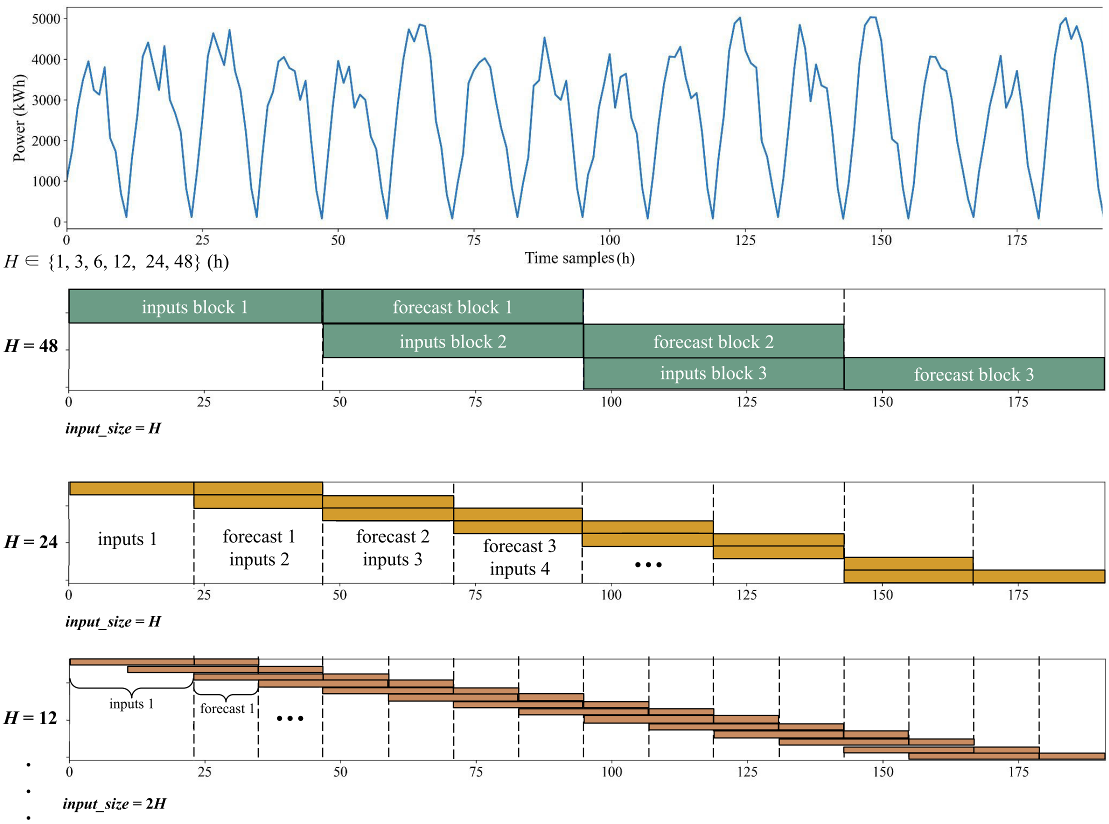

FIGURE 7. Schematic representation of the MIMO framework for multi-horizon forecasting.

4) CROSS VALIDATION STRATEGY

The forecasting procedure follows strict temporal causality and blocks any information leakage across all splits. The time series is divided into four contiguous segments: training, validation, calibration, and test. At each forecast origin, the model consumes an input window of length L built from hourly daytime observations between 7 and 18 h, which yields twelve samples per daytime day, and outputs a H-step forecast vector. All preprocessing statistics, such as the parameters used for standardization, are computed on the training segment only and then held fixed for the remaining stages.

During validation and test, the forecast origins advance with stride H. This produces non-overlapping H-step forecast blocks. The choice of stride H prevents reuse of future observations inside a block and avoids double counting of the same outcomes. It replicates deployment, where each decision yields a full H-step forecast and the next decision is taken only after those H steps have elapsed. Metrics are computed by pooling all forecast-observation pairs from the validation and test segments under this non-overlapping protocol.

During calibration, the forecast origins advance with unit stride. This produces overlapping input windows and, therefore, a residual at every timestamp. The choice of unit stride increases the number of residuals, which yields more stable empirical quantiles for conformal prediction. The denser sampling reduces variance in the estimated quantiles and improves the reliability of the target coverage. Final test metrics are reported by pooling all forecast-observation pairs over the test segment under the same causal protocol. Figure 8 presents the complete workflow of the proposed framework, encompassing data preprocessing, feature engineering, model training, multi-objective hyperparameter optimization via Pareto screening and TOPSIS ranking, conformal prediction interval calibration, and final performance evaluation.

IV. RESULTS AND DISCUSSION

This section presents the results, covering the computational setup, seasonal data partitioning, and hyperparameter configurations. Multi-objective optimization trade-offs are discussed, followed by forecast performance comparisons with baselines across seasons and horizons. Representative test cases and horizon-specific performance are also analyzed.

A. COMPUTATIONAL SETUP

Experiments were run in JupyterLab on Ubuntu 22.04.3 with Python 3.10, using a workstation with an Intel Core i9-10900X CPU, 32 GB RAM, and an NVIDIA RTX 2080

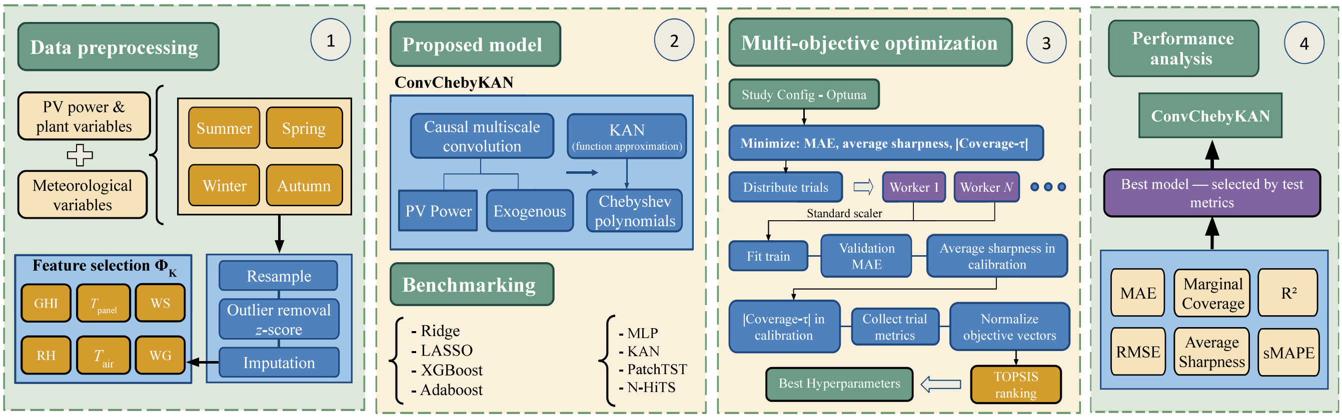

FIGURE 8. Comprehensive experimental and modeling workflow for PV power forecasting.

TABLE 2. Chronological partitioning of seasonal datasets into training (54%), validation (18%), calibration (10%), and testing subsets (18%), 12 samples/day.

Season	Samples	Training	Validation	Calibration	Testing
Spring	1,068	576	192	108	192
Summer	1,080	588	192	108	192
Autumn	1,104	612	192	108	192
Winter	1,128	600	228	108	192

SUPER GPU. Classical baselines (linear and tree models) used MLForecast with scikit-learn, while sequence models (KAN, ConvChebyKAN, NHITS, PatchTST) used NeuralForecast (PyTorch). ConvChebyKAN, not yet in NeuralForecast, was implemented from scratch by extending BaseModel with Chebyshev KAN layers and a multi-scale causal Conv1d front-end. Hyperparameters were optimized via TPE, and rolling-origin cross-validation was applied. Training followed default library practices.

B. STRUCTURED DATA SPLITTING

Each seasonal time series was chronologically partitioned into four non-overlapping and contiguous segments: training, validation, calibration, and testing. The models were fitted using the training subset, while the validation subset was employed to assess point-forecast accuracy during the hyperparameter optimization process. The calibration subset was used exclusively to adjust the prediction intervals, and the testing subset was reserved for the final model evaluation. This partitioning strategy ensures the prevention of information leakage, preserves temporal causality, and is applied consistently across all seasons and forecast horizons.

Table 2 describes the distribution of samples used for model training, validation, calibration, and testing across the four seasonal datasets, considering a sampling rate of 12 observations per day. The number of total samples per season ranges, with proportional splits ensuring temporal consistency and preventing data overlap between subsets.

The training subset represents approximately half of each seasonal dataset, providing sufficient data for model fitting. The validation subset contains around 18% of the data, used for hyperparameter optimization and point-forecast performance assessment. The calibration subset, about 10%, is dedicated exclusively to tuning prediction intervals, while the remaining samples form the testing subset for unbiased evaluation. This consistent partitioning framework across all seasons preserves chronological order and ensures comparability of forecasting results while maintaining temporal causality and preventing information leakage.

C. ANALYSIS OF MODEL HYPERPARAMETERS

The hyperparameter search space was systematically explored using the TPE over 100 full iterative trials, without employing any early stopping mechanism. To improve computational efficiency, the procedure was parallelized across 4-6 worker nodes, each assigned batches of five trial configurations, enabling up to 20-30 simultaneous model fittings and evaluations. Selection of configurations was guided by a composite loss function designed to approximate the Pareto frontier across three critical metrics: point forecast accuracy (measured by MAE), sharpness of the 90% prediction intervals (mean width), and calibration (empirical coverage relative to the nominal 0.90 level). This multi-objective strategy ensured the identification of a model exhibiting a balanced trade-off among these properties, rather than one over-optimized for a single criterion.

Table 3 details the hyperparameter search ranges and season-specific averages, with H denoting the forecast horizon in hours and kH representing an input window of k times the horizon. Optimal windows converged to $2H$ in spring and summer, $5H$ in autumn, and $4H$ in winter, reflecting increased historical memory needs in autumn and winter. The ConvChebyKAN model employs a shallow, causal multi-scale Conv1D front-end with dilations of 1-2 to cover the input window without skipping significant history. Autumn configurations include a third kernel-scale branch to capture longer patterns, while summer and winter

TABLE 3. Search space range and optimized hyperparameters to adjust the prediction models.

Model	Hyperparameter	Search Space	Spring	Summer	Autumn	Winter
ConvChebyKAN	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	5H	4H
	hidden_size	$[64, 256]$	128	64	128	64
	degree	$\{4, 6, 8\}$	6	8	8	4
	n_hidden_layers	$\{1, 2, 3\}$	1	1	1	1
	learning_rate	$[1e{-}4, 1e{-}2]$	3.4567e-4	2.5210e-3	2.4816e-4	9.5923e-3
	batch_size	$\{32, 64, 128\}$	128	64	64	64
	max_steps	$[250, 2000]$	750	500	250	1000
	ms_depth	$\{1, 2, 3\}$	2	2	3	2
	ms_base_channels	$\{16, 24, 32, 48, 64\}$	16	48	24	64
	ms_channel_decay	$[0.5, 1.0]$	0.75	1.0	1.0	0.5
	ms_dilation_base	$\{1, 2, 3\}$	1	2	1	2
	ms_kernel_scales	Conditional	[0.6, 0.2]	[0.7, 0.3]	[0.7, 0.3, 0.1]	[0.7, 0.3]
KAN	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	5H	4H
	hidden_size	$[64, 256]$	128	128	128	128
	n_hidden_layers	$\{1, 2, 3\}$	1	1	1	1
	learning_rate	$[1e{-}4, 1e{-}2]$	7e-4	1.2e-3	9e-4	1.5e-3
	batch_size	$\{32, 64, 128\}$	64	64	64	64
	max_steps	$[200, 2000]$	500	500	500	500
	use_exog	{True, False}	True	True	True	True
MLP	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	5H	4H
	layers	$\{1, 2, 3, 4\}$	2	3	2	3
	hidden_size	$[64, 512]$	256	256	256	256
	learning_rate	$[1e{-}4, 1e{-}2]$	8e-4	1.1e-3	6e-4	1.3e-3
	batch_size	$\{32, 64, 128\}$	64	64	64	64
	dropout	$[0.0, 0.5]$	0.1	0.1	0.1	0.1
Ridge	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	2H	2H
	alpha	$[1e{-}6, 10]$	1.0	1.0	1.0	1.0
	fit_intercept	{True, False}	True	True	True	True
	normalize	{True, False}	False	False	False	False
LASSO	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	5H	4H
	alpha	$[1e{-}6, 1]$	0.001	0.001	0.001	0.001
	fit_intercept	{True, False}	True	True	True	True
	max_iter	$[1000, 10000]$	5000	5000	5000	5000
AdaBoost	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	5H	4H
	n_estimators	$[50, 500]$	200	200	200	200
	learning_rate	$[0.01, 1.0]$	0.08	0.12	0.10	0.06
	\textit{loss} (regressor)	{linear, square, exponential}	linear	linear	linear	linear
	base_estimator	DecisionTreeRegressor	DTR	DTR	DTR	DTR
	max_depth (tree)	$[1, 10]$	3	3	3	3
NHITS	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	5H	4H
	n_blocks	$[1, 4]$	2	2	2	2
	n_layers	$[1, 4]$	2	2	2	2
	hidden_size	$[64, 512]$	256	256	256	256
	stack_types	{identity, trend, seasonal}	identity	identity	identity	identity
	learning_rate	$[1e{-}4, 1e{-}2]$	8e-4	1.0e-3	1.2e-3	9e-4
	max_steps	$[200, 2000]$	500	500	500	500
PatchTST	*input_size*	[H,2H,3H,4H,5H,12H]	2H	2H	5H	4H
	patch_len	$\{8, 12, 16, 24\}$	16	16	16	16
	stride	$\{1, 2, 4, 8\}$	2	2	2	2
	d_model	$[64, 512]$	256	320	256	384
	n_heads	$[2, 8]$	4	6	8	4
	n_layers	$[2, 8]$	3	4	5	4
	learning_rate	$[1e{-}4, 1e{-}3]$	5e-4	6e-4	4e-4	7e-4
	batch_size	$\{32, 64, 128\}$	64	64	64	64

increase base channel counts to enhance feature extraction for complex dynamics. Depth remains limited, *ms_depth* = 2, except 3 in autumn, with capacity expanded via channels and scale branches. The Chebyshev KAN backbone maintains a single hidden layer (hidden size: 64-128) with seasonally adjusted polynomial degrees: 6-8 in spring/summer for sharp nonlinearities, and 4 in winter for smoother mappings. Capacity is thus selectively increased based on seasonal data demands, using stable Chebyshev polynomial expansions instead of deeper architectures.

D. PARETO-OPTIMAL TRADE-OFFS IN PV POWER FORECASTING

Figure 9 illustrates the seasonal multi-objective Pareto fronts obtained for the day-ahead PV power forecasting task. Each panel corresponds to a different season (a) spring, (b) summer, (c) autumn, and (d) winter, and displays the trade-offs among three objectives, the Mean Absolute Error of validation (MAEval), representing point-forecast accuracy; average sharpness (AS), quantifying the mean width of the 90% prediction intervals, and

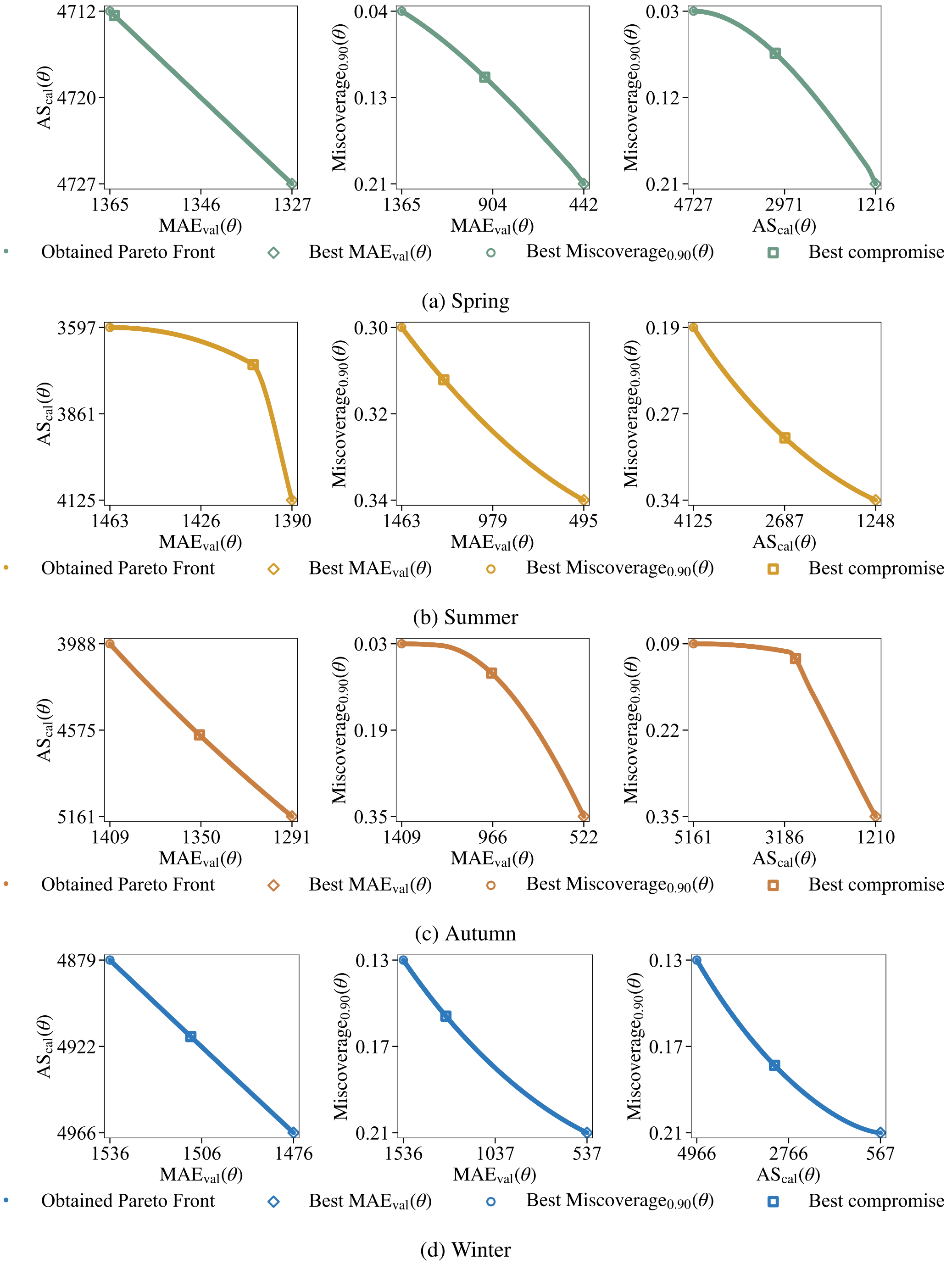

FIGURE 9. **Seasonal multi-objective Pareto fronts for day-ahead forecasting.**

miscoverage (|Coverage-0.90|), measuring deviation from the nominal confidence level. The figure highlights reference points in each Pareto set: the best solutions for MAEval, $\text{Miscoverage}_{0.90}$, AS, and the TOPSIS-selected compromise configuration. The ConvChebyKAN model consistently achieves balanced performance, with the TOPSIS-selected solutions residing in regions of the Pareto fronts that ensure robustness without sacrificing interval calibration.

In spring, the Pareto front reveals a gradual trade-off between interval sharpness and coverage, with MAEval remaining relatively stable. As AS decreases, the miscoverage initially changes slowly, then rises sharply beyond a turning point, reflecting the well-known calibration-sharpness trade-off described in forecast evaluation literature. This non-linear evolution follows the calibration-sharpness relation described by Gneiting et al. [71], where after a certain point, gains in sharpness require concessions in

TABLE 4. Performance evaluation of photovoltaic power multi-step ahead forecasting models in the spring.

Horizon	Model	MAE (kWh)	RMSE (kWh)	R^2	sMAPE (%)	AS (kWh)	MC	Rank
1	**ConvChebyKAN**	**295.4678**	**395.3195**	**0.9220**	**13.5220**	**1034.9924**	**0.8273**	**1**
	KAN	315.3585	418.3429	0.9126	18.2178	1241.2069	0.8750	2
	AdaBoost	372.8494	469.5077	0.8900	25.1325	1612.4219	0.9219	4
	MLP	382.5030	517.0284	0.8666	22.2360	1735.9633	0.9062	5
	LASSO	395.1310	484.4780	0.8828	25.2508	1560.6283	0.8802	6
	Ridge	395.1313	484.4787	0.8828	25.2508	1560.6434	0.8802	7
	NHiTS	440.2443	589.8768	0.8263	22.2516	1428.8570	0.8229	8
	PatchTST	627.2130	790.5633	0.6880	40.5779	2464.4189	0.8906	9
3	**ConvChebyKAN**	**333.8132**	**442.3221**	**0.9024**	**16.1551**	**1281.4295**	**0.8333**	**1**
	KAN	356.4062	463.6894	0.8927	19.5477	1815.5239	0.9375	2
	MLP	380.1286	498.7848	0.8758	24.5773	1799.9961	0.9427	3
	AdaBoost	402.4005	512.7377	0.8688	26.0039	1705.2498	0.9062	4
	NHiTS	420.2042	559.5597	0.8437	23.7292	1637.9427	0.8750	6
	LASSO	461.2337	569.8049	0.8379	28.3151	1780.2712	0.8750	7
	Ridge	461.2343	569.8054	0.8379	28.3151	1780.2788	0.8750	8
	PatchTST	507.6535	651.5478	0.7881	35.4150	1817.7613	0.8646	9
6	**ConvChebyKAN**	**354.7490**	**465.4762**	**0.8918**	**16.6573**	**1919.3470**	**0.9531**	**1**
	KAN	405.0616	520.4060	0.8648	20.3505	1997.7083	0.9323	2
	MLP	409.9367	518.0087	0.8661	24.4934	1861.8853	0.9271	3
	AdaBoost	412.0846	516.8327	0.8667	26.4130	1729.7852	0.9167	4
	PatchTST	416.7383	532.5983	0.8584	24.0133	1956.2045	0.9427	5
	NHiTS	476.8011	651.1020	0.7884	22.7261	1862.1311	0.8594	7
	Ridge	482.3509	591.7518	0.8252	28.6774	1843.7033	0.8698	8
	LASSO	482.3510	591.7519	0.8252	28.6774	1843.6964	0.8698	9
12	**ConvChebyKAN**	**393.5884**	**492.3516**	**0.8790**	**19.0802**	**1975.3478**	**0.9531**	**1**
	MLP	418.0161	536.3240	0.8564	23.6877	1800.0140	0.8958	2
	AdaBoost	422.3480	533.4900	0.8579	26.6893	1796.6199	0.9062	3
	KAN	430.4755	534.3074	0.8575	20.5704	1916.7109	0.9479	5
	PatchTST	432.3014	557.8083	0.8447	21.0493	1881.9165	0.8906	6
	Ridge	485.3560	601.4972	0.8194	28.2188	1862.9875	0.8698	7
	LASSO	485.3561	601.4971	0.8194	28.2188	1862.9823	0.8698	8
	NHiTS	531.8877	690.5563	0.7620	22.5632	1952.5295	0.8542	9
24	**ConvChebyKAN**	**368.6006**	**475.4663**	**0.8871**	**21.0046**	**2062.0320**	**0.9479**	**1**
	KAN	416.0615	522.7086	0.8636	22.2020	2001.6840	0.9167	2
	AdaBoost	446.2757	558.7434	0.8442	28.0611	1827.7034	0.8802	4
	NHiTS	468.4812	616.6018	0.8102	25.0397	1920.1696	0.8906	5
	MLP	470.4949	574.3334	0.8353	24.7866	2321.4958	0.9635	6
	Ridge	477.1019	578.5592	0.8329	29.3832	1943.8271	0.8906	7
	LASSO	477.1024	578.5596	0.8329	29.3832	1943.8244	0.8906	8
	PatchTST	507.4750	664.3783	0.7797	22.4346	1859.3819	0.8177	9
48	**ConvChebyKAN**	**373.6880**	**491.9773**	**0.8792**	**22.0732**	**1773.2514**	**0.9010**	**1**
	KAN	456.7255	591.1185	0.8256	25.8263	2027.8717	0.9010	2
	PatchTST	416.8599	583.3415	0.8301	23.0375	1686.9418	0.8594	3
	NHiTS	439.0792	582.5543	0.8306	20.2913	1774.4460	0.8490	4
	AdaBoost	443.6414	554.4423	0.8466	28.2032	1706.0731	0.8750	5
	MLP	489.5793	606.3197	0.8165	27.8453	2116.6771	0.9010	6
	Ridge	496.9172	601.4054	0.8195	31.3643	1960.7781	0.8958	7
	LASSO	496.9186	601.4073	0.8195	31.3643	1960.7830	0.8958	8

calibration, even if accuracy is essentially stable. The shape of the front indicates that accuracy gains can be maintained up to a threshold before calibration begins to deteriorate.

For summer, the front exhibits a nearly linear relationship between AS and miscoverage, suggesting a proportional compromise between interval width and reliability. The limited dispersion of MAEval values indicates that point-forecast accuracy is less sensitive to hyperparameter variations under high-irradiance, low-variability conditions typical of this season. The selected compromise lies on an intermediate segment of the front, balancing precision and uncertainty calibration.

In autumn, a broader Pareto region emerges, reflecting increased variability in both AS and miscoverage. The slopes of the front vary across subregions, implying that certain configurations prioritize sharper intervals at the expense of calibration, while others achieve more balanced uncertainty quantification. The weak correlation between MAEval and the other objectives highlights the independence of deterministic accuracy from interval calibration during transition seasons characterized by unstable irradiance.

In winter, all three objectives evolve in a coordinated manner, with minimal variation in MAEval and near-parallel movement of AS and miscoverage. The compactness of the front and the proximity of reference points suggest that the optimization yielded a set of solutions of comparable quality, where further tuning offers diminishing returns. This coherence likely reflects the lower irradiance dynamics and more predictable generation profiles in wintertime.

TABLE 5. Performance evaluation of photovoltaic power multi-step ahead forecasting models in the summer.

Horizon	Model	MAE (kWh)	RMSE (kWh)	R^2	sMAPE (%)	AS (kWh)	MC	Rank
1	**ConvChebyKAN**	**186.1185**	**245.7234**	**0.9645**	**12.2944**	**846.7591**	**0.9010**	**1**
	NHITS	215.7567	282.9230	0.9530	17.5143	958.3295	0.8958	3
	KAN	220.2146	283.0379	0.9529	19.5277	877.3270	0.9010	4
	LASSO	220.4224	278.5237	0.9544	19.9177	878.7296	0.8646	5
	Ridge	220.4225	278.5245	0.9544	19.9176	878.7353	0.8646	6
	AdaBoost	248.8720	310.5732	0.9433	18.6483	801.8960	0.9323	7
	MLP	254.9085	318.8183	0.9403	18.1080	1038.0981	0.9375	8
	PatchTST	463.8373	593.3195	0.7931	42.0462	1826.0701	0.8698	9
3	**ConvChebyKAN**	**203.4190**	**267.0563**	**0.9581**	**14.0460**	**858.8028**	**0.8854**	**1**
	MLP	245.0352	309.6623	0.9437	18.6232	1224.1292	0.9635	3
	LASSO	247.9249	312.9411	0.9425	21.7800	1079.6639	0.8906	4
	Ridge	247.9250	312.9415	0.9425	21.7800	1079.6646	0.8906	5
	KAN	249.2785	315.6649	0.9414	23.1471	998.9204	0.9115	6
	AdaBoost	258.3142	323.7524	0.9383	18.8865	803.4378	0.9167	7
	PatchTST	272.0619	347.1911	0.9292	24.2732	1068.7775	0.8750	8
	NHITS	275.0734	371.3124	0.9190	17.7158	925.1512	0.8021	9
6	**ConvChebyKAN**	**204.6557**	**266.3218**	**0.9583**	**14.2466**	**775.8420**	**0.8542**	**1**
	KAN	210.2043	276.3198	0.9551	14.1751	899.2634	0.8906	2
	MLP	238.7168	304.1082	0.9457	16.2256	1004.1581	0.9167	3
	Ridge	243.2984	308.2796	0.9442	21.0336	1131.4609	0.9271	4
	LASSO	243.2985	308.2794	0.9442	21.0337	1131.4588	0.9271	5
	PatchTST	247.4918	318.5724	0.9404	21.2281	913.6484	0.8438	6
	AdaBoost	258.4227	323.9861	0.9383	18.8865	804.6329	0.9167	8
	NHITS	262.2763	333.4334	0.9347	18.9310	975.7318	0.8490	9
12	**ConvChebyKAN**	**212.8830**	**277.5857**	**0.9547**	**14.2071**	**873.8399**	**0.8802**	**1**
	KAN	219.4327	283.7901	0.9527	16.5511	783.2921	0.8542	2
	MLP	229.0617	293.8774	0.9493	14.8866	820.1150	0.8229	3
	PatchTST	234.3841	298.6805	0.9476	15.9205	879.9149	0.8490	4
	Ridge	243.7159	308.0829	0.9442	21.0943	1132.2365	0.9427	5
	LASSO	243.7160	308.0828	0.9442	21.0944	1132.2349	0.9427	6
	NHITS	255.5845	325.8355	0.9376	17.3814	857.3572	0.8281	8
	AdaBoost	262.9459	329.4697	0.9362	19.0187	807.4361	0.9167	9
24	**ConvChebyKAN**	**217.5289**	**281.5726**	**0.9534**	**16.1773**	**728.4365**	**0.8177**	**1**
	KAN	236.5077	304.7176	0.9454	17.8319	902.1016	0.8594	2
	PatchTST	238.3529	315.8949	0.9414	16.1926	832.4216	0.8229	3
	Ridge	260.1642	325.8537	0.9376	22.8027	1105.3671	0.9062	5
	LASSO	260.1645	325.8541	0.9376	22.8028	1105.3682	0.9062	6
	AdaBoost	265.0587	333.1943	0.9348	19.0760	801.1194	0.8906	7
	MLP	315.7727	396.4519	0.9076	26.2746	1518.4238	0.9479	8
	NHITS	319.5083	427.4770	0.8926	25.9921	1221.8963	0.8698	9
48	**ConvChebyKAN**	**202.7850**	**257.1514**	**0.9611**	**16.5851**	**1891.8387**	**0.9792**	**1**
	PatchTST	251.1806	321.7287	0.9392	21.4342	768.0173	0.7292	2
	AdaBoost	269.4882	340.1745	0.9320	19.2000	752.7301	0.8646	4
	NHITS	273.4360	354.8701	0.9260	17.6329	790.0528	0.7500	5
	KAN	297.6856	379.2951	0.9155	22.4674	855.8012	0.7240	6
	Ridge	305.9506	375.5466	0.9171	25.5121	1165.1808	0.8906	7
	LASSO	305.9516	375.5479	0.9171	25.5122	1165.1855	0.8906	8
	MLP	306.3788	395.7880	0.9080	26.5406	910.1863	0.7552	9

E. ANALYSIS OF PERFORMANCE METRICS

Tables 4-7 summarize the forecasting performance of nine models across six prediction horizons (hours ahead) and four seasonal datasets. For each season, the models are ordered by MAE, which serves as the primary ranking criterion.

The reported metrics include MAE and RMSE as measures of point-forecast accuracy, R^2 as the explained variance, sMAPE as a scale-free indicator of percentage accuracy, marginal coverage as the empirical proportion of observations within the nominal 90% conformal prediction interval, and average sharpness as the mean interval width. Rank 1 indicates the best-performing model per horizon, and the best values are highlighted in bold. According to the comparative results summarized in the tables, ConvChebyKAN demonstrates the most balanced performance, achieving lower forecast MAE/RMSE, competitive sMAPE, and near-nominal uncertainty calibration with compact conformal intervals.

ConvChebyKAN provides higher predictive stability and reliability compared to baselines. Its hybrid convolutional-Chebyshev architecture enhances local feature extraction while preserving the global functional approximation, enabling effective modeling of the nonlinear and nonstationary dynamics characteristic of solar irradiance. This design supports a favorable trade-off between accuracy (MAE, RMSE, sMAPE) and interval-based uncertainty assessment (marginal coverage, average sharpness), aligning with previous evidence that hybrid or physics-informed architectures tend to outperform purely data-driven models under variable PV conditions [29]. Recent developments in

TABLE 6. Performance evaluation of PV power multi-step ahead forecasting models in the autumn.

Horizon	Model	MAE (kWh)	RMSE (kWh)	R^2	sMAPE (%)	Width (kWh)	MC	Rank
1	**ConvChebyKAN**	**237.1533**	**369.9253**	**0.9508**	**21.5821**	**1083.4997**	**0.8698**	**1**
	KAN	262.8621	389.0483	0.9456	25.4494	1055.2537	0.8646	2
	MLP	278.6858	423.1400	0.9356	29.8781	1125.6583	0.8333	3
	Ridge	321.9247	441.0508	0.9301	42.2365	1938.2864	0.9471	7
	LASSO	321.9232	441.0491	0.9301	42.2360	1942.2864	0.9471	6
	AdaBoost	389.2692	459.8477	0.9240	44.7569	1555.6843	0.9167	8
	NHITS	280.3901	434.4652	0.9321	25.1737	924.8640	0.7760	4
	PatchTST	515.5863	695.1685	0.8262	53.4764	1576.0422	0.8490	9
3	**ConvChebyKAN**	**302.0142**	**435.2930**	**0.9372**	**24.5085**	**1188.3822**	**0.8333**	**1**
	KAN	320.9494	466.8447	0.9216	28.5225	1605.3364	0.9010	2
	MLP	358.1507	513.6505	0.9051	30.9658	1946.8567	0.9427	4
	Ridge	433.1383	577.4948	0.8801	47.1581	1872.7239	0.9010	8
	LASSO	433.1372	577.4930	0.8801	47.1580	1872.7247	0.9010	7
	AdaBoost	458.3120	545.1780	0.8931	47.1988	1664.8711	0.8750	9
	NHITS	352.9948	512.7850	0.9055	33.5982	2623.5741	0.9792	3
	PatchTST	388.2000	543.2292	0.8939	44.9249	2038.5962	0.9479	6
6	**ConvChebyKAN**	**340.9575**	**485.1406**	**0.9154**	**25.5085**	**1842.0498**	**0.9271**	**1**
	KAN	379.2908	534.4039	0.8973	30.9047	1594.0355	0.8802	2
	MLP	416.9150	576.9910	0.8803	31.8565	2474.8500	0.9635	4
	Ridge	476.4920	632.2556	0.8563	48.4389	1906.9978	0.8698	8
	LASSO	476.4919	632.2547	0.8563	48.4389	1907.0001	0.8698	7
	AdaBoost	504.7418	615.8552	0.8636	48.5816	1732.7287	0.8281	9
	NHITS	411.0099	565.9117	0.8849	33.9696	2857.2741	0.9896	3
	PatchTST	425.8583	574.7964	0.8812	35.3946	2486.7130	0.9792	5
12	**ConvChebyKAN**	**367.5636**	**528.1855**	**0.8997**	**36.5629**	**1736.4436**	**0.9010**	**1**
	KAN	418.6467	582.5527	0.8780	38.6731	1452.6746	0.8073	2
	MLP	547.2844	749.8900	0.7978	40.4410	2156.7598	0.8594	7
	Ridge	470.2988	624.6010	0.8597	47.6120	1933.2870	0.8646	3
	LASSO	470.2990	624.6007	0.8597	47.6120	1933.2873	0.8646	4
	AdaBoost	528.8864	643.4396	0.8511	49.2761	1932.8523	0.8490	6
	NHITS	601.9345	792.8710	0.7740	43.1987	2391.4812	0.9219	9
	PatchTST	556.8215	728.0368	0.8094	42.9537	1886.6012	0.8333	8
24	**ConvChebyKAN**	**467.8721**	**642.8139**	**0.8631**	**32.0907**	**1830.8032**	**0.8750**	**1**
	KAN	505.5779	689.0748	0.8293	37.1246	1594.5133	0.7917	2
	MLP	613.6713	855.8172	0.7367	38.1453	2102.4006	0.8177	8
	Ridge	559.5260	723.0348	0.8120	51.3228	2074.4087	0.8490	4
	LASSO	559.5267	723.0348	0.8120	51.3229	2074.4067	0.8490	5
	AdaBoost	587.6524	738.8246	0.8037	51.8596	2490.1419	0.8333	7
	NHITS	659.8724	875.7411	0.7243	45.7581	1875.9242	0.7396	9
	PatchTST	556.8861	782.6454	0.7798	40.3322	1556.6348	0.7135	3
48	**ConvChebyKAN**	**451.2415**	**636.2391**	**0.8658**	**37.5633**	**2246.4227**	**0.9219**	**1**
	KAN	566.8918	781.1090	0.7806	44.9593	2027.6726	0.7917	6
	MLP	671.3198	879.3136	0.7220	50.9706	1903.8900	0.7552	8
	Ridge	556.0931	688.4655	0.8296	52.0621	1885.1589	0.8021	4
	LASSO	556.0942	688.4660	0.8296	52.0622	1885.1586	0.8021	5
	AdaBoost	1139.3728	1571.8884	0.1116	63.2941	3770.5349	0.8385	9
	NHITS	511.6070	695.9587	0.8258	40.1511	2144.5761	0.8646	2
	PatchTST	584.5683	774.6130	0.7843	44.9192	1385.8081	0.6615	7

KAN have also demonstrated that operator-based parameterizations improve time-series modeling by capturing nonlinear temporal dependencies more efficiently [15]. The integration of a Chebyshev polynomial basis with convolutional layers offers additional numerical stability and expressiveness, thus providing a principled framework for multi-scale temporal representation and cross-variable fusion.

In the spring season, ConvChebyKAN achieves the lowest errors and highest R^2 values on all horizons while maintaining favorable sMAPE scores. This period is characterized by rapid irradiance fluctuations and high variability, conditions under which transformer-based and tree-based models typically exhibit greater dispersion and reduced temporal consistency. ConvChebyKAN mitigates such effects through its multi-scale convolutional filters, which stabilize predictions and limit error propagation across horizons. Compared with the baseline KAN, the inclusion of local convolutional filtering enables sharper and well-calibrated conformal intervals, enhancing performance under rapidly changing irradiance regimes.

During summer, which exhibits strong irradiance and regular diurnal patterns, ConvChebyKAN maintains its superior performance across all horizons. The model consistently achieves lower forecast errors and higher R^2 values than both KAN and machine-learning baselines, particularly for extended horizons where long-range temporal dependencies are more relevant. The sMAPE values remain competitive, confirming that improvements are not limited to absolute-scale metrics. The conformal intervals remain narrow and close to the nominal

TABLE 7. Performance evaluation of photovoltaic power multi-step ahead forecasting models in the winter.

Horizon	Model	MAE (kWh)	RMSE (kWh)	R^2	sMAPE (%)	AS (kWh)	MC	Rank
	ConvChebyKAN	**149.1274**	**204.4357**	**0.9826**	**22.7635**	**642.9085**	**0.8303**	**1**
	LASSO	184.7470	240.7615	0.9759	27.1502	847.9640	0.9271	3
	Ridge	184.7473	240.7578	0.9759	27.1492	848.0214	0.9271	4
	KAN	191.0953	261.6907	0.9715	26.4174	874.5645	0.8958	5
1	MLP	193.8399	248.4824	0.9743	27.7469	954.7860	0.9531	6
	AdaBoost	196.1201	243.0927	0.9754	16.6520	768.8721	0.8958	7
	NHITS	369.6438	565.2755	0.8670	42.0472	2064.9859	0.9531	8
	PatchTST	422.7950	540.0568	0.8786	58.3317	1573.0396	0.8906	9
	ConvChebyKAN	**152.5662**	**209.0580**	**0.9818**	**18.6610**	**857.6401**	**0.9427**	**1**
	KAN	164.2363	221.3357	0.9796	22.8958	833.7386	0.9427	2
	NHITS	172.1511	238.2125	0.9764	24.7110	929.1745	0.9531	3
	PatchTST	189.1201	251.9501	0.9736	26.8462	1143.5769	0.9740	5
3	LASSO	196.5121	252.5175	0.9735	27.3307	919.4745	0.9271	6
	Ridge	196.5144	252.5180	0.9735	27.3309	919.5270	0.9271	7
	AdaBoost	209.4796	263.2116	0.9712	17.0879	821.2630	0.8906	8
	MLP	234.6809	371.2831	0.9426	27.5066	1745.6908	0.9844	9
	ConvChebyKAN	**171.8494**	**236.1288**	**0.9768**	**21.6746**	**774.2723**	**0.9115**	**1**
	MLP	182.9765	255.9571	0.9727	22.9295	963.0127	0.9167	2
	KAN	185.8730	250.5141	0.9739	23.9119	843.1483	0.9323	3
	PatchTST	190.1580	250.6741	0.9738	26.0291	991.2161	0.9479	4
6	LASSO	202.1176	270.3388	0.9696	25.4193	953.8164	0.9115	6
	Ridge	202.1182	270.3376	0.9696	25.4178	953.8387	0.9115	7
	AdaBoost	209.8831	263.8722	0.9710	17.0987	851.3516	0.8958	8
	NHITS	237.0408	312.9015	0.9592	25.7000	954.2962	0.8854	9
	ConvChebyKAN	**185.8412**	**249.5109**	**0.9741**	**22.2995**	**807.8808**	**0.8854**	**1**
	KAN	195.7859	279.1484	0.9676	22.1112	856.2154	0.8750	2
	MLP	197.4286	263.5198	0.9711	24.9416	809.2299	0.8906	3
	LASSO	207.5350	283.2273	0.9666	25.1489	995.2234	0.8958	4
12	Ridge	207.5351	283.2265	0.9666	25.1483	995.2375	0.8958	5
	AdaBoost	212.9101	270.4490	0.9696	17.1654	880.1133	0.9219	6
	NHITS	219.9346	288.4546	0.9654	26.5425	846.1206	0.8594	8
	PatchTST	241.2719	316.0952	0.9584	25.5877	896.4682	0.8333	9
	ConvChebyKAN	**217.6717**	**297.6837**	**0.9631**	**23.3702**	**864.6026**	**0.8643**	**1**
	PatchTST	229.4737	312.6304	0.9593	26.7955	853.6581	0.8438	2
	KAN	233.9983	311.2441	0.9597	27.3724	741.9675	0.7656	3
	NHITS	234.6722	295.2472	0.9637	26.1024	972.2614	0.8854	4
24	AdaBoost	246.2995	283.3785	0.9666	29.9327	916.9239	0.8698	6
	MLP	254.3645	319.3251	0.9576	36.0994	940.5771	0.8698	7
	Ridge	363.7562	437.0354	0.9205	35.9479	1868.7328	0.9740	8
	LASSO	363.7586	437.0382	0.9205	35.9481	1868.7447	0.9740	9
	ConvChebyKAN	**230.2497**	**301.9586**	**0.9620**	**24.2916**	**1028.7090**	**0.8954**	**1**
	NHITS	252.9825	320.4053	0.9573	25.4384	888.8972	0.7969	3
	KAN	255.3203	329.0297	0.9549	29.6716	795.5104	0.7865	4
	MLP	272.3915	354.1049	0.9478	38.0815	716.1478	0.6979	5
48	Ridge	272.3958	341.6207	0.9514	29.4745	1090.5761	0.8646	6
	LASSO	272.4033	341.6271	0.9514	29.4710	1090.5912	0.8646	7
	AdaBoost	282.0408	336.2260	0.9529	31.3314	1317.0016	0.8594	8
	PatchTST	288.8158	355.2782	0.9475	30.0907	623.0128	0.6042	9

90% coverage level, indicating accurate uncertainty calibration. This combination of precision and reliability is critical for operational planning during high-generation periods and aligns with previous studies highlighting the role of structured temporal priors in maintaining forecast stability [29].

In autumn, when irradiance decreases and atmospheric variability increases, ConvChebyKAN continues to exhibit generalization performance. The extended receptive field provided by its multi-scale convolutional architecture captures slower transitions between cloudy and clear-sky periods, maintaining low MAE, RMSE, and sMAPE values across horizons. Marginal coverage tends to increase slightly, reflecting a calibrated response to greater uncertainty rather than a degradation in predictive accuracy. Compared to KAN and NHITS, ConvChebyKAN maintains a more stable relationship between interval width and empirical coverage, indicating robustness to transitional seasonal regimes.

Lastly, in the winter season, characterized by reduced solar input and shorter daylight periods, ConvChebyKAN once again obtains superior point forecasting and interval-based consistency. The model achieves the lowest MAE and highest R^2 values while preserving competitive sMAPE scores and coverage close to the nominal level. The conformal intervals remain relatively narrow even at longer horizons, suggesting a stable and consistent uncertainty estimation process. Competing models, such as tree-based and transformer architectures, tend to overestimate uncertainty, producing overly wide intervals. ConvChebyKAN effectively mitigates this issue, achieving a balanced trade-off between sharpness and

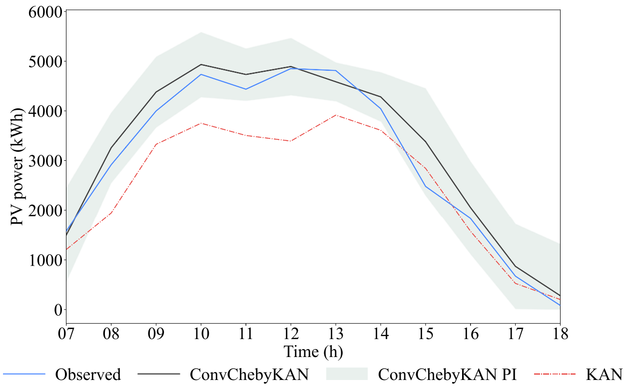

FIGURE 10. PV power forecasting results on the test-set day with the largest MAE at 48-hour-ahead horizon for spring.

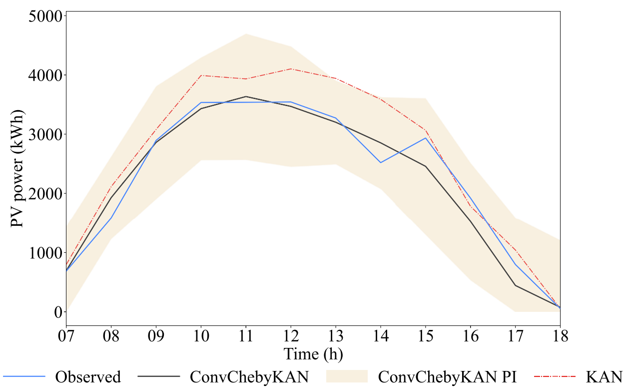

FIGURE 11. PV power forecasting results on the test-set day with the largest MAE at 48-hour-ahead horizon for summer.

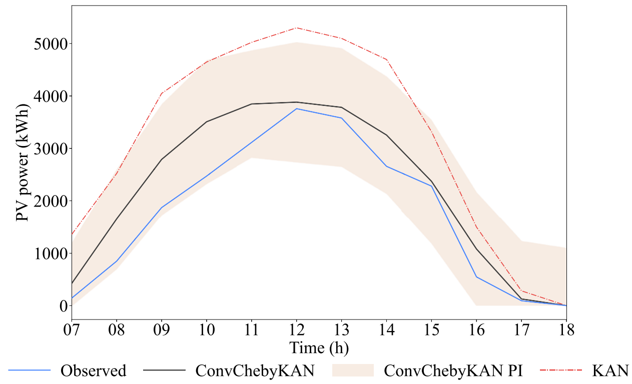

FIGURE 12. PV power forecasting results on the test-set day with the largest MAE at 48-hour-ahead horizon for autumn.

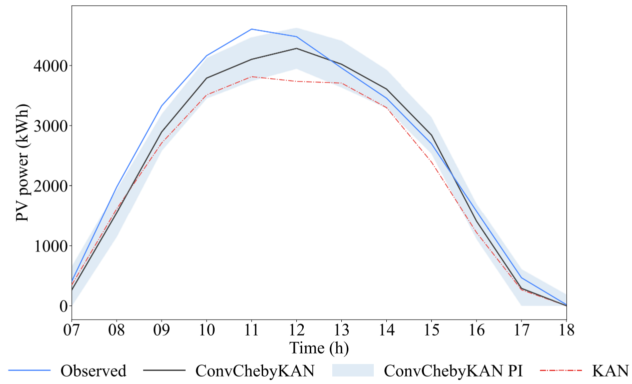

FIGURE 13. PV power forecasting results on the test-set day with the largest MAE at 48-hour-ahead horizon for winter.

calibration. In general, these results confirm the adaptability of the model in diverse seasonal conditions and its capacity to maintain reliable forecasts throughout varying irradiance regimes.

F. ANALYSIS OF MODEL PERFORMANCE

Figures 10 to 13 present the PV power forecasting results for the test-set day exhibiting the largest reduction in MAE achieved by the proposed ConvChebyKAN model relative to the KAN baseline at the 48-hour-ahead prediction horizon. The figure encompasses all four seasons, Figure 10 spring, Figure 11 summer, Figure 12 autumn, and Figure 13 winter, allowing a comparative evaluation of model performance under varying irradiance and atmospheric regimes.

In spring and autumn, characterized by high intra-day variability due to transient cloud cover, ConvChebyKAN effectively suppresses oscillations near peak generation hours and mitigates error accumulation following short-term irradiance drops. This improved temporal coherence results in smoother forecasts that remain closely aligned with observed power output. During summer, when irradiance is strong and diurnal patterns are more regular, ConvChebyKAN maintains alignment even during rapid irradiance fluctuations, avoiding the underestimation bias exhibited by KAN

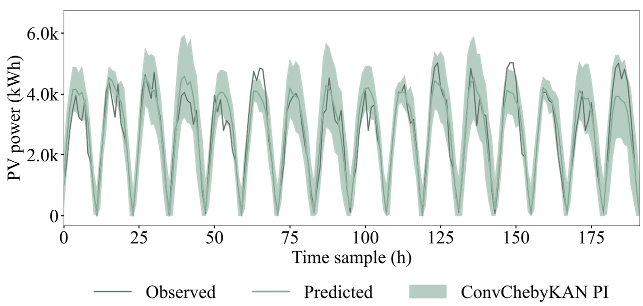

FIGURE 14. Observed and predicted PV power for the test set at 48 hours ahead during the spring season.

during post-dip recovery phases. The model demonstrates superior adaptability to rapid transitions between high and low generation levels, reflecting enhanced responsiveness to meteorological dynamics. In winter, with lower solar input and shorter daylight periods, ConvChebyKAN exhibits accurate delineation of the generation window, with improved synchronization between the predicted and observed start and end times of production. The amplitude bias observed in KAN is notably reduced, indicating a more faithful reconstruction of the diurnal generation profile.

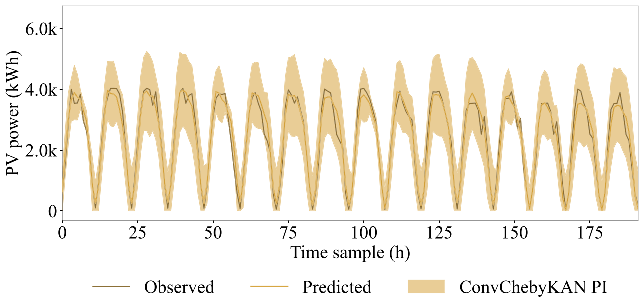

FIGURE 15. Observed and predicted PV power for the test set at 48 hours ahead during the summer season.

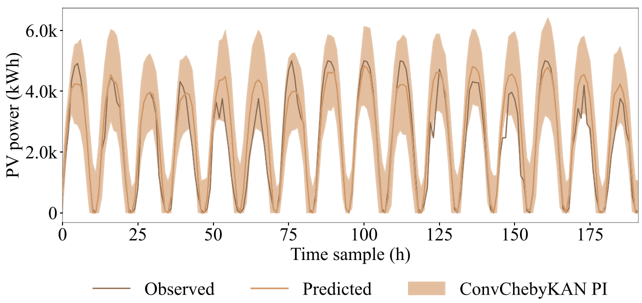

FIGURE 16. Observed and predicted PV power for the test set at 48 hours ahead during the autumn season.

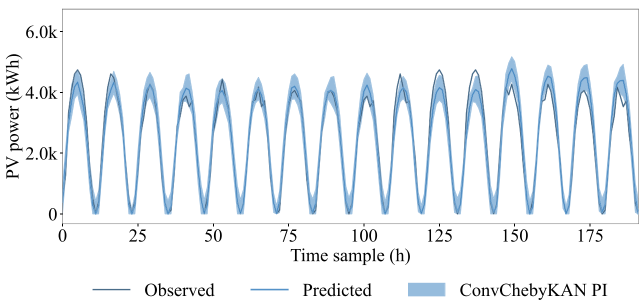

FIGURE 17. Observed and predicted PV power for the test set at 48 hours ahead during the winter season.

Across all seasons, the uncertainty bands generated by ConvChebyKAN remain compact while maintaining coverage close to the nominal 90% confidence level, signifying an effective balance between sharpness and calibration. These qualitative observations corroborate the quantitative results reported in previous section, confirming that the proposed model reduces large local errors, enhances temporal stability, and delivers reliable uncertainty quantification across diverse seasonal and meteorological conditions.

Figures 14 and 15 show the PV power forecasts generated by the ConvChebyKAN model over the last seven days of spring and summer, respectively, while 16 and 17 present analogous results of autumn and winter, respectively, across 48 hours ahead. The figures illustrate how predictive performance evolves with increasing lead time under varying irradiance regimes.

At the longest horizon, 48 hours ahead, forecast degradation becomes more evident, particularly under variable irradiance conditions such as those in spring (see Figure 14) and autumn (see Figure 16), where rapid weather transitions introduce non-linearities difficult to extrapolate. Still, ConvChebyKAN demonstrates consistent performance, with predicted curves maintaining realistic diurnal structure and bounded errors. In summer (see Figure 15), with stable and high irradiance, forecast accuracy remains high even at extended horizons, whereas winter (see Figure 17) exhibits reduced generation levels and slightly larger uncertainty bands due to shorter daylight duration and higher atmospheric variability.

V. CONCLUSION

Accurate PV power forecasting is essential for supporting reliable operation and scheduling in modern power systems with increasing penetration of renewable energy. However, the strong intermittency, nonlinear dynamics, and multi-scale temporal variability of solar generation pose significant challenges to conventional forecasting approaches. To address these issues, this study introduced ConvChebyKAN, a hybrid forecasting framework that integrates a causal multi-scale convolutional architecture with a Chebyshev-parameterized KAN and conformal prediction. The proposed framework was designed to enhance both deterministic forecasting accuracy and probabilistic reliability by combining hierarchical temporal feature extraction with stable functional representation and calibrated uncertainty estimation.

Compared with conventional ML and DL forecasting methods, the proposed framework exhibits several important characteristics:

1) Multi-scale temporal feature extraction. The causal convolutional front-end was designed to capture temporal patterns at different scales, enabling the model to simultaneously represent smooth diurnal generation cycles and rapid fluctuations caused by short-term meteorological variations. This multi-scale structure significantly improved the model's ability to represent the complex dynamics inherent in PV power generation.
2) Stable functional representation using Chebyshev-parameterized KAN. The KAN was parameterized using Chebyshev polynomials instead of traditional spline-based functions. This modification improved numerical stability and provided a more robust nonlinear representation of the relationships between meteorological inputs and PV power output.
3) Balanced optimization of deterministic and probabilistic performance. A multi-objective optimization framework based on Pareto screening and TOPSIS ranking was employed to identify optimal model configurations. This strategy ensured that the selected model simultaneously achieved high deterministic accuracy and reliable probabilistic forecasting performance.

4) Reliable uncertainty quantification through conformal prediction. The integration of split conformal prediction enabled the generation of calibrated prediction intervals without imposing strong distributional assumptions. As a result, the forecasting framework provided reliable probabilistic information that is essential for risk-aware decision-making in power system operations.

Real-world data from a PV plant in Brazil were used to evaluate the proposed approach under multiple seasonal conditions and forecasting horizons. The experimental results demonstrated the following main conclusions:

1) ConvChebyKAN consistently achieved superior deterministic forecasting accuracy. Across all seasons and forecasting horizons, the proposed model produced the lowest MAE and the highest R^2 values compared with benchmark models, including several state-of-the-art ML and DL approaches. Average improvements ranged from 3-12% for day-ahead forecasts and 10-32% for four-day-ahead horizons relative to the baseline KAN model.
2) The multi-scale convolutional architecture effectively captured complex temporal dynamics. The experimental analysis indicated that the convolutional feature extraction stage played a critical role in modeling both rapid fluctuations and smooth periodic patterns in PV generation. This capability significantly enhanced forecasting accuracy across varying climatic conditions.
3) The conformal prediction framework produced reliable and well-calibrated prediction intervals. The generated 90% prediction intervals remained well-calibrated and sufficiently sharp, confirming that the model can provide accurate probabilistic forecasts while maintaining meaningful uncertainty quantification under diverse seasonal conditions.
4) The proposed framework demonstrated strong practical potential for operational PV forecasting. By simultaneously providing accurate point forecasts and reliable prediction intervals, ConvChebyKAN offers valuable support for operational decision-making processes such as generation scheduling, reserve allocation, and renewable integration in modern power systems.

Future work will focus on several extensions of the proposed framework. First, additional datasets from different geographical regions and climatic conditions will be incorporated to further evaluate the generalization capability and robustness of the model. Second, improvements in computational efficiency and model complexity will be explored to enhance real-time forecasting performance. Finally, integrating the proposed forecasting framework into energy management and power system scheduling strategies represents an important direction for future research, enabling a more comprehensive assessment of the operational value of advanced PV forecasting methods in real-world power systems.

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

ACKNOWLEDGMENT

The authors thank CAPES for enabling open access through its transformative agreement with IEEE.

REFERENCES

[1] X. Wang, W. Liu, J. Ni, and M. Zhang, "Sample-wise graph-based multivariate short-term PV power forecasting," *IEEE Trans. Sustain. Energy*, vol. 16, no. 4, pp. 3003–3014, Oct. 2025.
[2] P. Spiru, "Assessment of renewable energy generated by a hybrid system based on wind, hydro, solar, and biomass sources for decarbonizing the energy sector and achieving a sustainable energy transition," *Energy Rep.*, vol. 9, pp. 167–174, Sep. 2023.
[3] L. de Azevedo Takara, A. C. Teixeira, H. Yazdanpanah, V. C. Mariani, and L. dos Santos Coelho, "Optimizing multi-step wind power forecasting: Integrating advanced deep neural networks with stacking-based probabilistic learning," *Appl. Energy*, vol. 369, Sep. 2024, Art. no. 123487.
[4] H. Min and B. Noh, "SolarNexus: A deep learning framework for adaptive photovoltaic power generation forecasting and scalable management," *Appl. Energy*, vol. 391, Aug. 2025, Art. no. 125848.
[5] R. Chen, G. Liu, Y. Cao, G. Xiao, and J. Tang, "CGAformer: Multi-scale feature transformer with MLP architecture for short-term photovoltaic power forecasting," *Energy*, vol. 312, Dec. 2024, Art. no. 133495.
[6] A. Rasouli and M. Rastegar, "An ensemble of deep learning, machine learning, and statistical methods stacked with meta-learning for forecasting net energy consumption in multi-carrier energy systems: Economic impact assessment," *Energy*, vol. 340, Dec. 2025, Art. no. 139172.
[7] Z. Serat, "Optimizing renewable energy systems for 100% clean energy target: A comparative study of solar, hydro, pumped hydro, and battery storage technologies," *J. Energy Storage*, vol. 104, Dec. 2024, Art. no. 114441.
[8] T. Ahmad, D. Zhang, and C. Huang, "Methodological framework for short-and medium-term energy, solar and wind power forecasting with stochastic-based machine learning approach to monetary and energy policy applications," *Energy*, vol. 231, Sep. 2021, Art. no. 120911.
[9] K. Wang, L. Wang, Q. Meng, C. Yang, Y. Lin, J. Zhu, Z. Zhao, C. Zhou, C. Zheng, and X. Gao, "Accurate photovoltaic power prediction via temperature correction with physics-informed neural networks," *Energy*, vol. 328, Aug. 2025, Art. no. 136546.
[10] S. S. Sidharth, A. R. Keerthana, R. Gokul, and K. P. Anas, "Chebyshev polynomial-based Kolmogorov–Arnold networks: An efficient architecture for nonlinear function approximation," 2024, *arXiv:2405.07200*.
[11] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, "KAN: Kolmogorov–Arnold networks," 2025, *arXiv:2404.19756*.
[12] J. P. M. Carvalho, S. F. Stefenon, M. H. D. M. Ribeiro, and K.-C. Yow, "Hybrid Kalman filter-Kolmogorov–Arnold network using local interpretable model-agnostic explanations for photovoltaic power forecasting," *Social Netw. Comput. Sci.*, vol. 7, no. 4, p. 319, Mar. 2026.
[13] V. C. Mariani, C. A. R. do Nascimento, and L. D. S. Coelho, "Multi-objective quantum differential evolution with Kolmogorov–Arnold network surrogate modeling for plate-fin heat exchanger design," *Appl. Thermal Eng.*, vol. 289, Mar. 2026, Art. no. 129553.
[14] J. Liu, F. Yang, and K. Yan, "Multi-scale data fusion and AdaptiveLoss Kolmogorov–Arnold network for multivariate time series forecasting," *Inf. Fusion*, vol. 125, Jan. 2026, Art. no. 103454.
[15] Z. Hou, B. Liu, G. Fan, J. Shi, and Y. Wang, "Short-term photovoltaic power prediction based on KAN model," in *Proc. IEEE 8th Conf. Energy Internet Energy Syst. Integr. (EI2)*, Nov. 2024, pp. 254–260.
[16] Y. Peng, J. Ma, M. Liao, C. Wang, and J. Long, "An improved photovoltaic power forecasting method using variational mode decomposition and Kolmogorov–Arnold networks," in *Proc. 5th Int. Conf. Smart Grid Energy Eng. (SGEE)*, Nov. 2024, pp. 507–512.

[17] X. Cai and R.-J. Wai, "Intelligent DC arc-fault detection of solar PV power generation system via optimized VMD-based signal processing and PSO–SVM classifier," *IEEE J. Photovolt.*, vol. 12, no. 4, pp. 1058–1077, Jul. 2022.

[18] Z. Song, S. Cao, and H. Yang, "An interpretable framework for modeling global solar radiation using tree-based ensemble machine learning and Shapley additive explanations methods," *Appl. Energy*, vol. 364, Jun. 2024, Art. no. 123238.

[19] Y. Chu, M. Li, H. T. C. Pedro, and C. F. M. Coimbra, "Real-time prediction intervals for intra-hour DNI forecasts," *Renew. Energy*, vol. 83, pp. 234–244, Nov. 2015.

[20] J. E. Sanchez-Lopez, J. Solís-García, and J. C. Riquelme, "Semi-real-time decision tree ensemble algorithms for very short-term solar irradiance forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 158, Jul. 2024, Art. no. 109947.

[21] D. V. Pombo, H. W. Bindner, S. V. Spataru, P. E. Sørensen, and P. Bacher, "Increasing the accuracy of hourly multi-output solar power forecast with physics-informed machine learning," *Sensors*, vol. 22, no. 3, p. 749, Jan. 2022.

[22] L. O. Seman, S. F. Stefenon, K.-C. Yow, L. D. S. Coelho, and V. C. Mariani, "Fourier-enhanced sequence-to-sequence latent graph neural networks for multi-node spatiotemporal forecasting in a hydroelectric reservoir," *Eng. Appl. Artif. Intell.*, vol. 167, Mar. 2026, Art. no. 113939.

[23] Y. Sun, V. Venugopal, and A. R. Brandt, "Short-term solar power forecast with deep learning: Exploring optimal input and output configuration," *Sol. Energy*, vol. 188, pp. 730–741, Aug. 2019.

[24] Y. Jung, J. Jung, B. Kim, and S. Han, "Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea," *J. Cleaner Prod.*, vol. 250, Mar. 2020, Art. no. 119476.

[25] Y. Xu, S. Zheng, Q. Zhu, K.-C. Wong, X. Wang, and Q. Lin, "A complementary fused method using GRU and XGBoost models for long-term solar energy hourly forecasting," *Expert Syst. Appl.*, vol. 254, Nov. 2024, Art. no. 124286.

[26] D. Korkmaz, "SolarNet: A hybrid reliable model based on convolutional neural network and variational mode decomposition for hourly photovoltaic power forecasting," *Appl. Energy*, vol. 300, Oct. 2021, Art. no. 117410.

[27] H. Sharadga, S. Hajimirza, and R. S. Balog, "Time series forecasting of solar power generation for large-scale photovoltaic plants," *Renew. Energy*, vol. 150, pp. 797–807, May 2020.

[28] M. Gao, J. Li, F. Hong, and D. Long, "Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM," *Energy*, vol. 187, Nov. 2019, Art. no. 115838.

[29] B. Gao, X. Huang, J. Shi, Y. Tai, and J. Zhang, "Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks," *Renew. Energy*, vol. 162, pp. 1665–1683, Dec. 2020.

[30] L. F. R. Agottani, M. M. Girardi, W. R. de Souza Chaves, L. de Azevedo Takara, L. dos Santos Coelho, and V. C. Mariani, "Multi-step ahead wind power forecasting based on N-BEATS model," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur. (ISGT EUROPE)*, Oct. 2024, pp. 1–5.

[31] B. Belmahdi, M. Louzazni, and A. E. Bouardi, "A hybrid ARIMA–ANN method to forecast daily global solar radiation in three different cities in Morocco," *Eur. Phys. J. Plus*, vol. 135, no. 11, Nov. 2020.

[32] T. Limouni, R. Yaagoubi, K. Bouziane, K. Guissi, and E. H. Baali, "Accurate one step and multistep forecasting of very short-term PV power using LSTM-TCN model," *Renew. Energy*, vol. 205, pp. 1010–1024, Mar. 2023.

[33] R. Mondal, S. K. Roy, and C. Giri, "Solar power forecasting using domain knowledge," *Energy*, vol. 302, Sep. 2024, Art. no. 131709.

[34] J. Wang, Y. Zhou, and Z. Li, "Hour-ahead photovoltaic generation forecasting method based on machine learning and multi objective optimization algorithm," *Appl. Energy*, vol. 312, Apr. 2022, Art. no. 118725.

[35] S. Li, J. Wang, H. Zhang, and Y. Liang, "Solar photovoltaic power forecasting system with online manner based on adaptive mode decomposition and multi-objective optimization," *Comput. Electr. Eng.*, vol. 118, Sep. 2024, Art. no. 109407.

[36] Y. Dai, W. Yu, and M. Leng, "A hybrid ensemble optimized BiGRU method for short-term photovoltaic generation forecasting," *Energy*, vol. 299, Jul. 2024, Art. no. 131458.

[37] Y. Ma, W. Yu, J. Zhu, Z. You, and A. Jia, "Research on ultra-short-term photovoltaic power forecasting using multimodal data and ensemble learning," *Energy*, vol. 330, Sep. 2025, Art. no. 136831.

[38] X. Gao, Y. Zang, Q. Ma, M. Liu, Y. Cui, and D. Dang, "A physics-constrained deep learning framework enhanced with signal decomposition for accurate short-term photovoltaic power generation forecasting," *Energy*, vol. 326, Jul. 2025, Art. no. 136220.

[39] Q. Sun, F. Yan, W. Sun, and Y. Zhou, "DWT-former: Fusing wavelet-based multi-scale features and transformer-based temporal representations for photovoltaic power forecasting," *Energy*, vol. 341, Dec. 2025, Art. no. 139283.

[40] R. Krishna and S. Hemamalini, "Improved TLBO algorithm for optimal energy management in a hybrid microgrid with support vector machine-based forecasting of uncertain parameters," *Results Eng.*, vol. 24, Dec. 2024, Art. no. 102992.

[41] M. Baak, R. Koopman, H. Snoek, and S. Klous, "A new correlation coefficient between categorical, ordinal and interval variables with Pearson characteristics," *Comput. Statist. Data Anal.*, vol. 152, Dec. 2020, Art. no. 107043.

[42] A. C. R. Klaar, S. F. Stefenon, L. O. Seman, V. C. Mariani, and L. D. S. Coelho, "Optimized EWT-Seq2Seq-LSTM with attention mechanism to insulators fault prediction," *Sensors*, vol. 23, no. 6, p. 3202, Mar. 2023.

[43] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, "KAN: Kolmogorov–Arnold Networks," in *Proc. 13th Int. Conf. Learn. Represent.*, 2025.

[44] A. Kolmogorov, "On the representation of continuous functions of several variables by superpositions of continuous functions of lesser variable count," *Doklady Akademii Nauk SSSR*, vol. 114, pp. 369–373, 1957.

[45] V. I. Arnold, "On the representation of continuous functions of three variables by the superposition of continuous functions of two variables," *Math. Notes*, vol. 54, no. 5, pp. 235–236, 1958.

[46] C. Ying, W. Wang, J. Yu, Q. Li, D. Yu, and J. Liu, "Deep learning for renewable energy forecasting: A taxonomy, and systematic literature review," *J. Cleaner Prod.*, vol. 384, Jan. 2023, Art. no. 135414.

[47] M. H. Sulaiman, Z. Mustaffa, M. S. Saealal, M. M. Saari, and A. Z. Ahmad, "Utilizing the Kolmogorov–Arnold networks for chiller energy consumption prediction in commercial building," *J. Building Eng.*, vol. 96, Nov. 2024, Art. no. 110475.

[48] Z. Gao and M. Kong, "MP-KAN: An effective magnetic positioning algorithm based on Kolmogorov–Arnold network," *Measurement*, vol. 243, Feb. 2025, Art. no. 116248.

[49] M. U. Danish and K. Grolinger, "Kolmogorov–Arnold recurrent network for short term load forecasting across diverse consumers," *Energy Rep.*, vol. 13, pp. 713–727, Jun. 2025.

[50] X. Han, X. Zhang, Y. Wu, Z. Zhang, and Z. Wu, "KAN4TSF: Are KAN and KAN-based models effective for time series forecasting?" 2024, *arXiv:2408.11306*.

[51] S. Somvanshi, S. A. Javed, M. M. Islam, D. Pandit, and S. Das, "A survey on Kolmogorov–Arnold network," 2024, *arXiv:2411.06078*.

[52] A. K. M. E. Saleh, M. Arashi, and B. M. G. Kibria, *Theory of Ridge Regression Estimation With Applications*. Hoboken, NJ, USA: Wiley, 2019.

[53] L. S. Coelho, H. V. H. Ayala, and V. C. Mariani, "Co and nox emissions prediction in gas turbine using a novel modeling pipeline based on the combination of deep forest regressor and feature engineering," *Fuel*, vol. 355, 2024, Art. no. 129366.

[54] H. Liu, H. Wu, H. Jin, and Y. He, "Adaptive forecasting of photovoltaic power based on dual-type models' ensemble and online error correction," *Appl. Energy*, vol. 408, Apr. 2026, Art. no. 127397.

[55] C. Challu, K. G. Olivares, B. N. Oreshkin, F. G. Ramirez, M. M. Canseco, and A. Dubrawski, "NHITS: Neural hierarchical interpolation for time series forecasting," in *Proc. AAAI Conf. Artif. Intell.*, vol. 37, 2023, pp. 6989–6997.

[56] L. D. A. Takara, L. F. R. Agottani, C. V. Züge, W. R. de Souza Chaves, L. dos Santos Coelho, and V. C. Mariani, "Deep learning-based probabilistic forecasting of household energy consumption in smart grids," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur. (ISGT EUROPE)*, Oct. 2024, pp. 1–5.

[57] S. F. Stefenon, L. O. Seman, C. K. Yamaguchi, L. D. S. Coelho, V. C. Mariani, J. P. Matos-Carvalho, and V. R. Q. Leithardt, "Neural hierarchical interpolation time series (NHITS) for reservoir level multi-horizon forecasting in hydroelectric power plants," *IEEE Access*, vol. 13, pp. 54853–54865, 2025.

[58] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, "A time series is worth 64 words: Long-term forecasting with transformers," in *Proc. 11th Int. Conf. Learn. Represent.*, 2022.

[59] M. H. D. M. Ribeiro, R. G. da Silva, G. T. Ribeiro, V. C. Mariani, and L. D. S. Coelho, ‘‘Cooperative ensemble learning model improves electric short-term load forecasting,’’ *Chaos, Solitons Fractals*, vol. 166, Jan. 2023, Art. no. 112982.

[60] L. O. Seman, S. F. Stefenon, V. C. Mariani, and L. D. S. Coelho, ‘‘Ensemble learning methods using the hodrick–prescott filter for fault forecasting in insulators of the electrical power grids,’’ *Int. J. Electr. Power Energy Syst.*, vol. 152, Oct. 2023, Art. no. 109269.

[61] S. F. Stefenon, L. O. Seman, L. S. A. da Silva, V. C. Mariani, and L. D. S. Coelho, ‘‘Hypertuned temporal fusion transformer for multi-horizon time series forecasting of dam level in hydroelectric power plants,’’ *Int. J. Electr. Power Energy Syst.*, vol. 157, Jun. 2024, Art. no. 109876.

[62] A. Solari and V. Djordjilović, ‘‘Multi split conformal prediction,’’ *Statist. Probab. Lett.*, vol. 184, May 2022, Art. no. 109395.

[63] L. Sluijterman, E. Cator, and T. Heskes, ‘‘How to evaluate uncertainty estimates in machine learning for regression?’’ *Neural Netw.*, vol. 173, May 2024, Art. no. 106203.

[64] T. Gneiting, F. Balabdaoui, and A. E. Raftery, ‘‘Probabilistic forecasts, calibration and sharpness,’’ *J. Roy. Stat. Soc. Ser. B, Stat. Methodol.*, vol. 69, no. 2, pp. 243–268, Apr. 2007.

[65] S. Ben Taieb, G. Bontempi, A. F. Atiya, and A. Sorjamaa, ‘‘A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition,’’ *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7067–7083, Jun. 2012.

[66] S. Resifi, E. A. Aawar, H. P. Dasari, H. Jebari, and I. Hoteit, ‘‘A novel deep learning approach for regional high-resolution spatio-temporal wind speed forecasting for energy applications,’’ *Energy*, vol. 328, Aug. 2025, Art. no. 136356.

[67] S. Li, L. Guo, J. Zhu, M. Liu, J. Chen, and Z. Meng, ‘‘Short-term multi-step wind speed forecasting with multi-feature inputs using variational mode decomposition, a novel artificial intelligence network, and the polar lights optimizer,’’ *Renew. Energy*, vol. 256, Jan. 2026, Art. no. 123965.

[68] L. O. Seman, A. C. R. Klaar, M. H. D. M. Ribeiro, and S. F. Stefenon, ‘‘Enhanced random vector functional link networks with Bayesian-based hyperparameter optimization for wind speed forecasting,’’ *IEEE Access*, vol. 13, pp. 208105–208122, 2025.

[69] Y. Ozaki, Y. Tanigaki, S. Watanabe, and M. Onishi, ‘‘Multiobjective tree-structured Parzen estimator for computationally expensive optimization problems,’’ in *Proc. Genetic Evol. Comput. Conf.*, Jun. 2020, pp. 533–541.

[70] C. L. Hwang and K. Yoon, *Multiple Attribute Decision Making: Methods and Applications: A State-of-the-Art Survey*, vol. 186. Berlin, Germany: Springer, 1981.

[71] T. Gneiting, A. E. Raftery, A. H. Westveld, and T. Goldman, ‘‘Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation,’’ *Monthly Weather Rev.*, vol. 133, no. 5, pp. 1098–1118, May 2005.

LUCAS DE AZEVEDO TAKARA received the B.E. degree in control and automation engineering from the Instituto Nacional de Telecomunicações (INATEL), Brazil, in 2020, and the M.Sc. degree in electrical engineering from the Federal University of Paraná (UFPR), Brazil, in 2023, where he is currently pursuing the Ph.D. degree in electrical engineering, where his research focuses on machine learning and deep learning methods for time-series forecasting, classification, and decision-making under uncertainty. He has experience in applying artificial intelligence methods to real-world predictive modeling and decision-support problems. His research interests include time-series forecasting, reinforcement learning, deep learning, and generative artificial intelligence.

RODRIGO CLEMENTE THOM DE SOUZA received the B.Sc. degree in computer science from the State University of Maringá (UEM), Brazil, in 2000, the M.Sc. degree in production engineering and systems from the Pontifical Catholic University of Paraná (PUCPR), Brazil, in 2008, and the Ph.D. degree in numerical methods in engineering from the Federal University of Paraná (UFPR), Brazil, in 2013, including a doctoral research period with The University of Manchester, U.K., in 2012. He is currently an Associate Professor with UFPR, a Permanent Faculty Member of the Graduate Program in Computer Science, UEM in the area of computational intelligence, and the Leader of the Natural and Scientific Computing Research Group.

VIVIANA COCCO MARIANI (Member, IEEE) received the B.E. degree in mathematics from the Federal University of Santa Maria, Brazil, and the M.Sc. degree in computer science and the Ph.D. degree in mechanical engineering from the Federal University of Santa Catarina, Brazil. She was a Postdoctoral Researcher with the University of Padua, Italy, in 2018, and INESC TEC, Portugal, in 2025. She is currently a Full Professor with the Department of Electrical Engineering, Federal University of Paraná (UFPR), Brazil. Her research interests include prediction and classification using machine learning models, and optimization techniques applied to engineering problems.

LEANDRO DOS SANTOS COELHO (Member, IEEE) received the B.Sc. degree in informatics and the degree in electrical engineering from the Federal University of Santa Maria, in 1994 and 1999, respectively, and the M.Sc. degree in computer science and the Ph.D. degree in electrical engineering from the Federal University of Santa Catarina, in 1997 and 2000, respectively. He is currently a Full Professor with the Department of Electrical Engineering, Federal University of Paraná (UFPR). He is a member of the Editorial Board of Elsevier's journals.

STEFANO FRIZZO STEFENON received the Ph.D. degree in electrical engineering from the State University of Santa Catarina (UDESC), Brazil, in 2021, and the Ph.D. degree in computer science and artificial intelligence from Università degli Studi di Udine (UNIUD), Italy, in 2025. He was a Postdoctoral Fellow Researcher with the Faculty of Engineering and Applied Science, University of Regina (UofR), Saskatchewan, Canada. He is currently a Professor (Adjunct) with the Lisbon School of Engineering (ISEL), Polytechnic University of Lisbon (IPL), Portugal. His research interests include electrical power systems, DL, computer vision, and time-series forecasting.