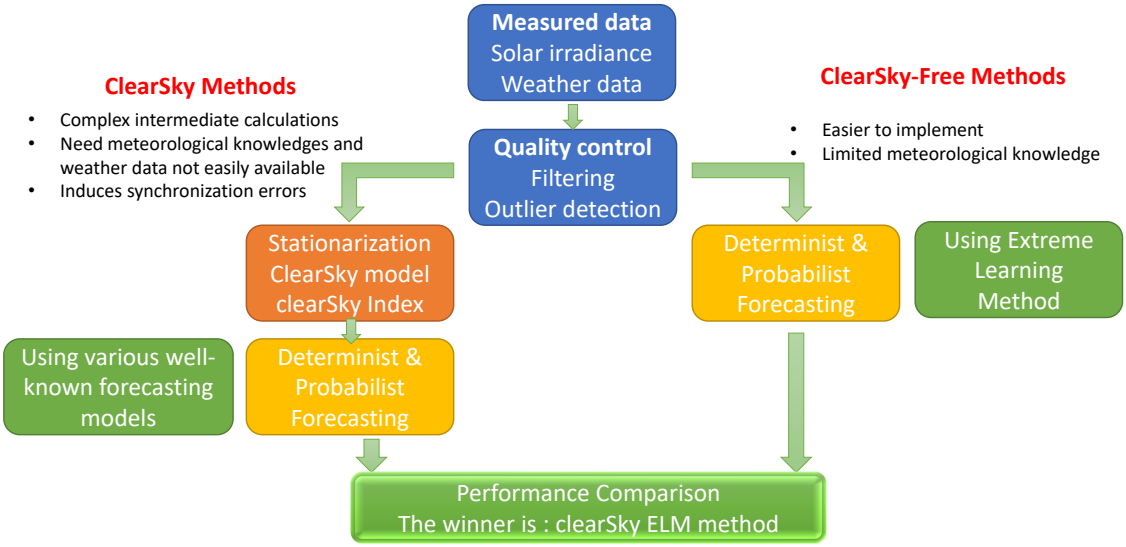
On the Importance of Clearsky Model in Short-Term Solar Radiation Forecasting

Cyril Voyant, Milan Despotovic, Gilles Notton, Yves-Marie Saint-Drenan, Mohammed Asloune, Luis Antonio García Gutiérrez

Graphical Abstract

On the Importance of Clearsky Model in Short-Term Solar Radiation Forecasting

Cyril Voyant, Milan Despotovic, Gilles Notton, Yves-Marie Saint-Drenan, Mohammed Asloune, Luis Garcia-Gutierrez



Highlights

On the Importance of Clearsky Model in Short-Term Solar Radiation Forecasting

Cyril Voyant, Milan Despotovic, Gilles Notton, Yves-Marie Saint-Drenan, Mohammed Asloune, Luis Garcia-Gutierrez

- A novel method using raw `GHI` and machine learning, outperforming Clearsky-Based models.
- Clearsky-Free and Clearsky-Based models in comparison concerning deterministic and probabilistic concerns.
- Avoiding the Clearsky Index reduces dependency on complex intermediate calculations.
- Enables smarter grids, energy trading, and renewable integration.

On the Importance of Clearsky Model in Short-Term Solar Radiation Forecasting

Cyril Voyant[a], Milan Despotovic[b,c], Gilles Notton[c], Yves-Marie Saint-Drenan[a], Mohammed Asloune[c], Luis Garcia-Gutierrez[c]

[a] *OIE Laboratory, Mines-PSL, Sophia-Antipolis, F-06904, Antibes, France*
[b] *Faculty of Engineering, University of Kragujevac, Kragujevac, 34000, Serbia*
[c] *SPE Laboratory, UMR CNRS 6134, University of Corsica Pasquale Paoli, Ajaccio, 20000 , Corsica, France*

Abstract

Clearsky models are widely used in solar energy for many applications such as quality control, resource assessment, satellite-base irradiance estimation and forecasting. However, their use in forecasting and nowcasting is associated with a number of challenges. Synchronization errors, reliance on the Clearsky index (ratio of the global horizontal irradiance to its cloud-free counterpart) and high sensitivity of the clearsky model to errors in aerosol optical depth at low solar elevation limit their added value in real-time applications. This paper explores the feasibility of short-term forecasting without relying on a clearsky model. We propose a Clearsky-Free forecasting approach using Extreme Learning Machine (`ELM`) models. `ELM` learns daily periodicity and local variability directly from raw Global Horizontal Irradiance (`GHI`) data. It eliminates the need for Clearsky normalization, simplifying the forecasting process and improving scalability. Our approach is a non-linear adaptative statistical method that implicitely learns the irradiance in cloud-free conditions removing the need for an clear-sky model and the related operational issues. Deterministic and probabilistic results are compared to traditional benchmarks, including ARMA with `McClear`-generated Clearsky data and quantile regression for probabilistic forecasts. `ELM` matches or outperforms these methods, providing accurate predictions and robust uncertainty quantification. This approach offers a simple, efficient solution for real-time solar forecasting. By overcoming the stationarization process limitations based on usual multiplicative scheme Clearsky models, it provides a flexible and reliable framework for modern energy systems.

Keywords: Solar Irradiance Forecasting, Clearsky Models, Clearsky-Free Approach, Extreme Learning Machine (`ELM`)

1. Introduction

Solar irradiance forecasting is essential for integrating solar energy into grids, ensuring stability, facilitating the energy balance between production and consumption and addressing variability (Azarpour et al., 2022). Precise forecasts optimize production, storage, and distribution, supporting the transition to sustainable energy systems (Yang et al., 2022; Lauret et al., 2012). Despite significant advancements in the field, forecasting accuracy remains constrained by the complex and dynamic nature of solar irradiance. The interaction between atmospheric conditions and temporal variability introduces uncertainties that complicate real-time applications and long-term planning. This study explores these uncertainties and investigates innovative solutions to enhance scalability, minimize reliance on traditional normalization methods, and improve overall forecasting reliability.

1.1. Importance and Challenges of Solar Irradiance Forecasting

Solar energy is expected to reach over 1,500 GW_p of PV capacity by 2030 tripling in just 10 years, which highlights the critical need for accurate forecasting to balance energy supply, optimize grid performance, and decrease dependence on fossil fuels (International Energy Agency (IEA), 2021). Real-time grid management and energy markets require accurate short-term predictions (range from about 30 min to 5-6 hours) since errors in these forecasts can result in inefficiencies and increased costs (Diagne et al., 2013; Voyant et al., 2017). However, forecasting global solar irradiance presents significant challenges due to the inherent variability and non-stationarity of solar irradiance data. In order to solve these challenges, irradiance data are often normalized. One way to normalize solar irradiance data involves utilizing extraterrestrial irradiance that is defined as solar irradiance at the topmost layer of the Earth's atmosphere. This extraterrestrial irradiance can be calculated using a formula that accounts for geometric relationship between the Sun and the Earth.

Normalizing with extraterrestrial irradiance results in the development of forecasting models utilizing the clearness index (Sanfilippo et al., 2016; Akarslan et al., 2018). Most of the traditional approaches to solar irradiance forecasting rely on Clearsky model, computing an intermediate stationary form (called Clearsky Index or CSI), as they typically produce better forecasts compared to models using clearness index (Lauret et al., 2022). Methods employing Clearsky models (Hoyos-Gómez et al., 2022; David et al., 2016; Frimane et al., 2022; Shan et al., 2022; Gneiting et al., 2023) seek to use solar irradiance in cloud-free (clear sky) conditions and subsequently enhance forecasts by incorporating observed meteorological data characterizing the sky state and content in gas and particles with a seasonal and daily variation. Traditional forecasting methodologies face critical limitations, including inaccurate Clearsky models, elevated computational expenses associated with Numerical Weather Prediction (NWP) systems, and a limited ability to adjust to changing local meteorological conditions (Gueymard, 2008; Sobri et al., 2018). In order to overcome these limitations, this study proposes a framework that is based on a Clearsky-Free, data-driven, methodology that employs Extreme Learning Machines (ELM). This methodology simplifies the forecasting process, enhances adaptability, and provides accurate, real-time predictions that are tailored to modern energy systems by explicitly modeling Global Horizontal Irradiance (GHI) with enriched periodic and probabilistic features.

1.2. Challenges with Clearsky-Based Models and the CSI

Clearsky models, such as McClear (Lefèvre et al., 2013), REST2 (Gueymard, 2008) and SPARTA (Ruiz-Arias, 2023), are widely used for calculating solar irradiance under cloud-free conditions. These models provide essential baselines using celestial mechanics and interpreting atmospheric effects related to aerosols, water vapor and ozone, with broad applications in climate science, agriculture, solar energy, and air quality monitoring (Allen et al., 1998; Voyant et al., 2017; Sobri et al., 2018). Despite their versatility, Clearsky-based models (*i.e.,* using the Clearsky model as a stationarization method) exhibit significant limitations, particularly in real-time operational contexts.:

- *Synchronization Issues*: Mismatched measurements and model outputs, especially at low solar elevations when shading and solar mask can intervene, lead to inaccuracies and reduced reliability (Diagne et al., 2013);
- *Dependence on Atmospheric Inputs*: Parameters such as aerosol optical depth and water vapor content are widely available thanks to atmospheric composition model or reanalyses such as CAMS or MERRA2. However, the accuracy of these parameters can be poor in some regions leading to unreliable clear sky irradiance estimation citepgueymard2021solar;
- *Gaps in Coverage*: CSI derived from these models (as the ratio of GHI to Clearsky irradiance), is undefined during nighttime and prone to significant errors during transitional periods such as dawn and dusk (Kumar et al., 2020; Voyant et al., 2017);
- *Limited Adaptability*: Clearsky models struggle to represent diffuse irradiance at low solar elevations and to adapt to localized, dynamic weather conditions, particularly in complex terrains (Santamouris et al., 1990);
- *Surface Orientation Biases*: For forecasting PV power or solar radiation on tilted surface, the irradiance provided by Clearsky model needs to be projected on a inclined plane using so-called transposition model. The limited knowledge of the anisotropy of the diffuse radiance can introduce a bias that propagate through the forecasting pipeline (Ma et al., 2020).

For photovoltaic systems, these limitations are further compounded by dynamic factors such as inverter efficiency, maintenance schedules, and panel degradation, which Clearsky models cannot capture (Chodakowska et al., 2024). As a result, these models, like they are actually used in prediction, could add complexity without significantly improving the accuracy of forecasts. The limitation of Clearsky-Based models underscore the need for alternative approaches where possible. New Clearsky-Free methodologies should demonstrate that it is possible to bypass clearsky model and its limitations by directly leveraging raw irradiance data (e.g., GHI, GTI, BNI, DHI) without relying on theoretical Clearsky baselines (Ruiz-Arias and Gueymard, 2018). An explanation of the stationary process in solar time series prediction is provided in Appendix A, emphasizing that this approach captures the stochastic nature of solar irradiance, enhances robustness and adaptability, and eliminates pre-processing and post-processing steps such as CSI normalization and rescaling, thereby simplifying the forecasting process. These Clearsky-Free methods, powered by advanced

Table 1: Comparison of Clearsky-Based Models and Clearsky-Free Models

Aspect	Clearsky-Based Models	Clearsky-Free Models
Atmospheric inputs Dependence	High aerosols, water vapor and ozone data	Low uses lagged GHI
Applicability in real-time	Limited synchronization and pre-processing	High directly processes raw data
Computational complexity	High advanced numerical models	Low simplified data-driven models
Adaptability to local dynamics	Limited struggles with localized weather patterns	High handles localized and nonlinear variations
Accuracy of forecasts	Moderate sensitive to errors in atmospheric inputs	High leverages endogenous, real-time data
Infrastructure requirements	High powerful computing, robust data pipelines	Low works with simpler configurations
Robustness to data gaps	Low errors propagate in case of missing inputs	High less affected by incomplete datasets
Optimal use case	For theoretical studies or well-instrumented sites	For dynamic and data-scarce environments

machine learning techniques like ELM, represent a significant shift in solar irradiance forecasting. They overcome the key limitations of traditional models for time-series based short-term forecast, providing scalable, accurate, and efficient solutions tailored to modern energy systems (Chodakowska et al., 2024; Voyant et al., 2017).

1.3. Opportunities for Clearsky-Free Models

The limitations of CSI-based forecasting methods highlight the need for alternatives that bypass Clearsky models. These Clearsky-independent approaches predict solar irradiance variables directly, without relying on theoretical baselines offers significant advantages with regard to real-time forecasting. Clearsky-Free models avoid the need for atmospheric inputs that are often subject to errors, such as aerosol optical depth and related synchronization issues. Instead, these models rely on endogenous information, such as lagged values of GHI, in order to simplify the forecasting task and avoid cascading errors (Li et al., 2016). They provide continuity for every hour of the day, even at the critical low-light sunrise and sunset hours, therefore improving grid management-energy storage (Diagne et al., 2013). The hypothesis explored in this paper is that some adaptative statistical model can learn implicitely from past observations the effects of aerosol, ozone and water vapour on the solar irradiance, making the use of a clearsky model unnecessary. Correctly learning this component of atmospheric extinction is not a trivial task, since it is strongly modulated by the effect of clouds, which is several orders of magnitude greater.Not all statistical forecast approaches are capable of learning reliably the effect of aerosol, ozone and water vapour in varying weather conditions. We expect that advanced machine learning techniques, including Extreme Learning Machines, are able to capture complex temporal dependencies scalably and efficiently. Probabilistic extensions (Lauret et al., 2017) enable quantification of uncertainty, allowing for superior operational decision-making (Huang et al., 2006). Table 1 provides an overview of the key differences between Clearsky-Based and Clearsky-Free approaches. This comparison underscores how Clearsky-Free methods tackle important challenges, making them especially effective for applications that demand adaptability, scalability, and efficiency. Moreover, such a Clearsky-Free approach have additional benefits: it would be less computationally complex, making them more robust and adaptive.

1.4. Objective of the Study

Despite their utility, Clearsky-Based models face critical limitations in real-time forecasting. These include inaccuracies in atmospheric inputs, reliance on complex pre-processing steps, and limited adaptability in dynamic or data-sparse environments. Addressing these challenges requires an approach that is simpler, more robust, and less dependent on extensive datasets. The objectives of this study are:

- Develop a Clearsky-Free forecasting framework that eliminates the need for Clearsky normalization and minimizes reliance on atmospheric data.
- Leverage historical and endogenous features, such as lagged Global Horizontal Irradiance (GHI) values, to simplify and streamline the forecasting process.

- Design a solution that is adaptable and capable of performing effectively in environments with sparse or incomplete data.

The proposed framework will be evaluated on three criteria: accuracy, to ensure precise solar irradiance forecasts; robustness, to deliver reliable performance across varying climatic conditions and geographic regions; and parsimony, to achieve high forecasting quality with minimal data and low computational complexity. By addressing the shortcomings of traditional Clearsky-Based models, this study positions Clearsky-Free approaches as practical, scalable, and efficient solutions for modern energy systems. The emphasis on simplicity, adaptability, and data efficiency makes this framework particularly suitable for both data-rich and resource-constrained environments. The structure of this paper is as follows: Section 2 describes the methodology, including data collection, quality control, and the modeling framework. Section 3 presents and discusses the results, covering both deterministic and probabilistic forecasting approaches, as well as comparisons to benchmark models. Finally, Section 4 provides the conclusions and outlines future research directions.

2. Methodology

To effectively address the limitations of Clearsky-Based methods and leverage the potential of Clearsky-Free models, a robust methodology is required. This section details the approach adopted, beginning with the collection and preprocessing of high-resolution data to ensure the accuracy and reliability of the forecasting methodology.

2.1. Data Collection and Quality Control (QC)

This study employs a univariate and endogenous forecasting approach, where Global Horizontal Irradiance (`GHI`) predicts itself. The dataset is sourced from the `SIAR` network of agroclimatic weather stations in Spain (Despotovic et al., 2024). It includes high-resolution `GHI` records from 76 stations over a four-year period, with a 30-min temporal resolution. This dataset offers the fine granularity required for short-term forecasting while encompassing a diverse range of climatic zones, from arid regions to highly variable coastal and mountainous areas, ensuring broad applicability (Vicente-Serrano et al., 2022). To ensure data reliability, a strict quality control process proposed by Garcia-Gutierrez et al. (2022) was applied. This included detecting and rectifying anomalies such as sensor malfunctions, extreme weather impacts, or inconsistent measurements. The `McClear` model, recognized for its accurate estimation of clear-sky irradiance, was used as a reference for validating the `GHI` data (Lefèvre et al., 2013). Since `McClear` operates in hindcast mode (where required atmospheric inputs are only available two days after the observation date), it provides highly reliable theoretical clear-sky estimates, making it a robust basis for identifying and correcting errors in the historical dataset. However, this operational constraint limits its use in real-time forecasting. Special attention was paid to sunrise and sunset periods, where low solar elevation angles are known to introduce higher uncertainty (Herrería-Alonso et al., 2020). To account for temporal dependencies in `GHI`, lagged input variables were created, enabling dynamic adaptation to the inherent variability of solar irradiance (Yang et al., 2022). The dataset was divided into three years for training and validation and one year for testing, ensuring robust evaluation of the model's generalization capabilities under both stable and dynamic weather scenarios. In this context, the `ELM` model was trained using a carefully selected set of input features. The primary predictor is historical `GHI`, using past values to forecast future irradiance. Additionally, the clear-sky index k_t, defined as the ratio between the measured `GHI` and its theoretical counterpart from `McClear`, was included to provide an implicit reference for atmospheric attenuation effects. Temporal features, such as the hour of the day, day of the year, and seasonal indicators, were incorporated to capture diurnal and seasonal variations. Although no direct aerosol measurements are included, their impact on solar radiation is implicitly accounted for through the clear-sky `GHI` from `McClear`, which integrates aerosol optical properties from `CAMS` reanalysis data. By structuring models around historical `GHI` and clear-sky references, they captures complex dependencies between atmospheric conditions and solar irradiance without requiring explicit `AOD` measurements. This approach simplifies the forecasting pipeline while maintaining high performance. Furthermore, it allows for a reliable evaluation of the necessity of clear-sky in forecasting accuracy.

2.2. Model Description

This study evaluates a diverse set of models for solar irradiance forecasting, combining deterministic and probabilistic approaches to address the challenges of traditional Clearsky normalization. These models are

designed to capture temporal dependencies and the inherent periodic variability of solar irradiance data. Naive predictors serve as benchmarks and include:

- Persistence (`P`): Assumes future `GHI` is equal to the most recent observed value (Yang et al., 2022);
- Clearsky Model (`CS`): Uses theoretical Clearsky irradiance while ignoring cloud attenuatation Lefèvre et al. (2013);
- Smart Persistence (`SP`): Combines persistence with deviations from theoretical Clearsky irradiance to account for diurnal patterns (Diagne et al., 2013);
- Climatology-Persistence `CLIPER`: Blends climatological averages with persistence to exploit autocorrelation (Yang, 2019);
- Exponential Smoothing (`ES`): Applies a weighted average to recent deviations from Clearsky irradiance (Makridakis et al., 1998);
- Autoregressive AR(2)-like (`ARTU`): Incorporates temporal dependencies and accounts for uncertainty (Voyant et al., 2022b);
- Combination Model (`COMB`): Aggregates `SP`, `CLIPER`, `ES`, and `ARTU` using ensemble methods (Voyant et al., 2022b).

The reference model in the deterministic case is an autoregressive `AR(p)` model computed with classical least square optimisation. In this model, the `GHI` at time y_t is expressed as a linear combination of its past values: $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$, where ϕ_i are the coefficients, p is the model order, and ϵ_t represents the error term or residual. The coefficients ϕ_i are estimated by minimizing the sum of squared residuals $\min_\phi \sum_{t=1}^{T} \left(y_t - \sum_{i=1}^{p} \phi_i y_{t-i}\right)^2$. This `AR` model serves as a benchmark for assessing the performance of machine learning-based approaches in forecasting. It will be used in two versions: `AR` (without clear sky) and `rAR` (with clear sky). See Sections 3.1.1 and 3.1.2 for more details. The Extreme Learning Machine (`ELM` or `EL`) model is the central predictive framework in this study, providing a scalable and computationally efficient solution. It relies on a single-hidden-layer neural network architecture with sigmoid and Gaussian activation functions to capture both linear and non-linear relationships in the data. The model configuration includes:

- Input and hidden layer optimization: For each forecast horizon, the number of input neurons and hidden neurons is adjusted from in-sample data (the best `nRMSE` defines the best configuration). For one particular site with high variability for example, the 1-hour horizon uses in average 78 inputs and 472 hidden neurons, while the 5-hour horizon uses 138 inputs and 450 hidden neurons. The hidden layer size is scaled approximately four times the input size for robust performance (Huang et al., 2006).
- Training process: Multiple random initializations are used to ensure robustness. Output weights are optimized using Ridge regression, $\beta = (H^\top H + \lambda I)^{-1} H^\top Y$, where H is the hidden layer output matrix, λ is the regularization parameter, and Y is the target vector (Zheng et al., 2023).
- Model selection: The configuration with the lowest normalized root mean square error (`nRMSE`) on the validation set is selected for testing.

A detailed `ELM` definition is given in Appendix B. Note that `ELM` provides a significant advantage over deep learning models in terms of both training and inference speed. Unlike iterative backpropagation-based methods, `ELM` training is performed through a single matrix inversion step, resulting in high computational efficiency. Furthermore, due to its simple architecture consisting of a single hidden layer, real-time inference is enabled, with predictions being generated in a few seconds. In this study, It was verified that incoming data can be processed for a dataset of approximately 70 sites, with an execution time of less than 30 seconds, confirming the model's suitability for operational solar forecasting applications where rapid response times are required. Unlike physics-based models that integrate explicit aerosol-related parameters, the ELM operates in a purely data-driven manner. It statistically extracts the most relevant relationships between past observations and future irradiance without requiring a precise estimation of atmospheric attenuation mechanisms. This means that the model indirectly captures the impact of aerosols, clouds, and other atmospheric phenomena based on historical patterns but does not explicitly differentiate their contributions.

Quantile Regression (QR) is employed as a reference probabilistic forecasting method due to its robustness and established performance in uncertainty quantification. QR predicts conditional quantiles of the response variable, allowing for the generation of reliable prediction intervals without assumptions about the underlying data distribution. The method optimizes an asymmetric loss function defined as:

$$L_\tau(u) = \begin{cases} \tau u & \text{if } u \geq 0 \\ (\tau - 1)u & \text{if } u < 0 \end{cases} \tag{1}$$

where $u = y - \hat{y}$ is the residual and τ represents the quantile level. By solving this minimization problem using linear programming, QR constructs prediction intervals that align with target coverage levels (Koenker, 2005). In addition to QR, a non-parametric methodology is used to derive prediction intervals directly from residuals of deterministic forecasts (Lauret et al., 2017). Lookup tables are generated during the training phase, capturing empirical relationships between forecast errors and observed values (Voyant et al., 2022a). This approach avoids predefined assumptions about error distributions, making it particularly robust in dynamic atmospheric conditions. More information are given in Appendix C. The methodology used during simulations offers several key advantages. It provides comprehensive benchmarking by incorporating both naive predictors and advanced models, enabling rigorous evaluation. The framework is capable of handling high-dimensional data and varying forecast horizons with minimal computational overhead. Probabilistic forecasting is robustly implemented through QR alongside non-parametric approaches for enhanced flexibility.

2.3. Evaluation Metrics

The performance of the forecasting models is assessed using both deterministic and probabilistic metrics to capture accuracy, bias, and reliability. This comprehensive evaluation framework is critical for validating the efficacy of Clearsky-Free methodologies in solar irradiance forecasting. For deterministic forecasting, the following metrics are employed:

- Normalized Root Mean Square Error (nRMSE) quantifies the average deviation of predictions from observed values, with a focus on penalizing larger errors, $\texttt{nRMSE} = \mathrm{E}[y]^{-1} \cdot \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$ where y_i represents the observed values, $\hat{y}_i$ the predicted values, N the total number of observations and $\mathrm{E}[y]$ represents the expected value or mean of the random variable y (Makridakis et al., 1998);
- Normalized Mean Absolute Error (nMAE) provides a robust measure of accuracy, less sensitive to outliers compared to nRMSE, $\texttt{nMAE} = \mathrm{E}[y]^{-1} \cdot \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$;
- R-Squared ($\mathtt{R}^2$) measures the proportion of variance in the observed data explained by the model, providing an indicator of goodness-of-fit, $\mathtt{R}^2 = 1 - (\sum_{i=1}^{N}(y_i - \hat{y}_i)^2)/(\sum_{i=1}^{N}(y_i - \mathrm{E}[y])^2)$;
- Normalized Mean Bias Error (nMBE) captures systematic bias in the forecasts (Herrería-Alonso et al., 2020), $\texttt{nMBE} = \mathrm{E}[y]^{-1} \cdot \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)$.

For probabilistic forecasting, the evaluation framework includes:

- Continuous Ranked Probability Score (CRPS) compares the cumulative distribution functions (CDFs) of predicted and observed values, $\texttt{CRPS}(F, y) = \int_{-\infty}^{\infty}\left(F(z) - \mathbf{1}_{z \geq y}\right)^2 dz$, evaluating the entire predictive distribution, where $F(z)$ is the forecast CDF, and y is the observed value. This study computes CRPS using centiles, offering finer granularity than quartiles for solar irradiance forecasts (Gneiting and Raftery, 2007). The quantile-based CRPS method used during this study is detailed in Appendix D;
- Prediction Interval Coverage Probability (PICP) evaluates the reliability of prediction intervals by measuring the percentage of observed values falling within these intervals, $\texttt{PICP} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}_{\underline{y_i} \leq y_i \leq \overline{y_i}}$, where $\underline{y_i}$ and $\overline{y_i}$ are the lower and upper bounds of the prediction interval (Lauret et al., 2017);
- Mean Interval Length (MIL) assesses the sharpness of prediction intervals, providing an indicator of interval width, $\mathrm{MIL} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{y_i} - \underline{y_i}\right)$;
- Interval Score (IS) evaluates the quality of prediction intervals by balancing two factors: sharpness (ensuring the intervals are as narrow as possible, measured by the MIL) and coverage reliability (Penalizing predictions when observed values fall outside the prediction intervals). The Interval Score

is defined as: $\text{IS} = \text{MIL} + \frac{2}{\alpha}\sum_{i=1}^{N}\left[\mathbf{1}_{y_i<\underline{y_i}}\left(\underline{y_i}-y_i\right)+\mathbf{1}_{y_i>\overline{y_i}}\left(y_i-\overline{y_i}\right)\right]$, where α is the nominal coverage level of the prediction interval, and $\mathbf{1}$ denotes the indicator function (equals 1 if the condition is true and 0 otherwise) (Gneiting and Raftery, 2007).

The evaluation framework uses deterministic metrics like `nRMSE` and `nMAE`, along with probabilistic metrics like `CRPS` and `PICP`, to assess model reliability, accuracy, and practical relevance. The Mann-Whitney U test (Mann and Whitney, 1947), a non-parametric method, is used to compare distributions without assuming normality, making it ideal for scenarios where error distributions deviate from Gaussian norms. The test measures shifts in forecast error distributions, with a p-value below the standard threshold of 0.05 indicating a significant difference. For example, Model B outperforms Model A with a p-value of 0.021, indicating superior performance with consistently lower forecast errors.

3. Results and Discussion

In this section, the results for both deterministic and probabilistic forecasting are presented and analyzed. The `ELM` configuration (size of input and hidden nodes), was optimized for each prediction horizon using the Nelder-Mead simplex algorithm (Nelder and Mead, 1965), with an average of 85 input nodes and 430 hidden nodes. The ridge parameter is set to $\lambda = 0.2$, and we arbitrarily select a 60%-40% hybrid transfer function mix of sigmoidal and Gaussian functions, respectively, to ensure both flexibility and efficiency. Training involved 96 runs (with in-sample data) to identify optimal weights of `ELM`. While simulations used a high-performance computing system (1840 CPU cores and 6.4 TB of RAM), `ELM` required less than a minute for one year of training and forecasting. The longer total time (under 10 min per site with the possibility to parallelize certain tasks) was mainly due to the computational demands of `QR`. Violin plots will be used to compare distributions across groups. They combine boxplots and kernel density plots, showing data spread, density, and central tendency. This allows for intuitive visualization of differences in shape, range, and median between groups.

3.1. Deterministic Case

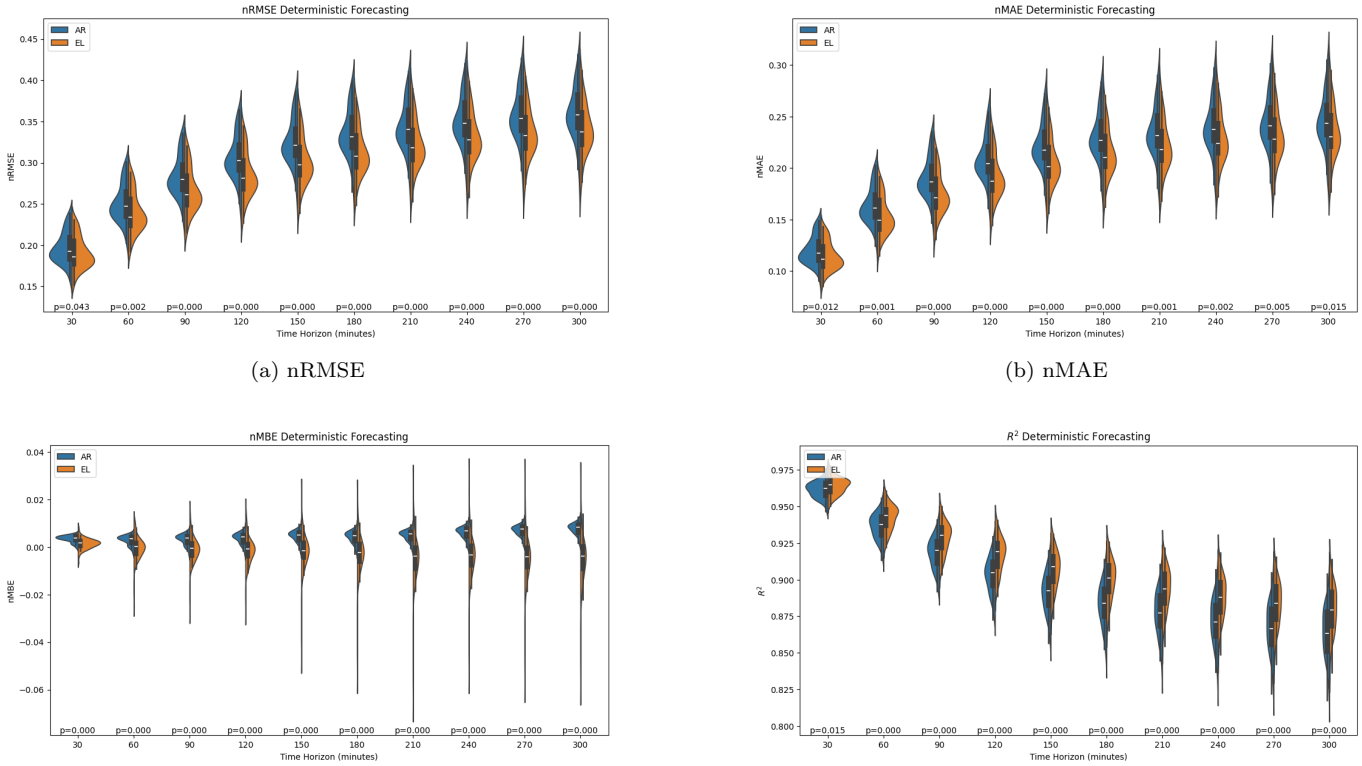
Point prediction, or deterministic prediction, involves generating a single estimated value for a future variable without accounting for uncertainty or variability. This approach is crucial as it enables precise and timely decision-making in applications such as solar energy management, where accurate forecasts can optimize resource utilization and enhance system efficiency.

3.1.1. Justifying the Use of `ELM`

Initially, it is essential to validate the use of `ELM` against a purely linear autoregressive `AR`(p) model, optimizing p for each site and forecasting horizon using a brute force method comparing in-sample results for $p = 1$ to $p = 96$ (p median $= 58$). The comparison of performances of both models is operated on raw data (without employing Clearsky models). Note that p can't be defined from Partial Autocorrelation Function (`PACF`) while this parameter is periodic with raw data. Performance of the deterministic forecasting, across various metrics and time horizons are presented in Figure 1. The resulting p-values from Mann-Whitney U test are displayed below the respective plots for each time horizon in the figures. The results show that `ELM` consistently outperforms `AR` across all error metrics (`nRMSE`, `nMAE`, `nMBE`, $\mathtt{R}^2$) and time horizons, with very low p-values confirming the significance of these differences. Therefore, `ELM` is clearly the better choice for more accurate, reliable, and unbiased predictions.

3.1.2. Validation Against Reference and Naive Predictors

To validate the `ELM` Clearsky-Free approach, predictions are compared to classical solar irradiance tools, which are defined as reference (`rAR`) and naive models (`SP, P,` and `CS`) derived from the `McClear` Clearsky model. The order p of the `rAR`(p) model is optimized using a `PACF`-based approach: the first minimum of the `PACF` function determines the maximum lag to consider, which corresponds to p (p median $= 6$). As shown in Figure 2, `ELM` consistently outperforms all naive models (`SP, P,` and `CS`) across all time horizons (30, 180, and 360 min) with highly significant differences ($p < 0.001$). While `ELM` also surpasses `rAR` at 180 and 360 min ($p < 0.001$), the difference is not significant for the 30-min horizon ($p = 0.239$). Among the models, `EL` ranks as the best, followed by `rAR, SP, CS,` and finally `P`.



(a) nRMSE

(b) nMAE

(c) nMBE

(d) R^2

Figure 1: Evaluation metrics for deterministic forecasting over the 76 `SIAR` Stations. Each plot illustrates a specific metric (`nRMSE`, `nMAE`, `nMBE`, `R²`) across 10 time horizons (30 to 300 min). The comparison includes two models: `AR` (blue) and `ELM` (orange).

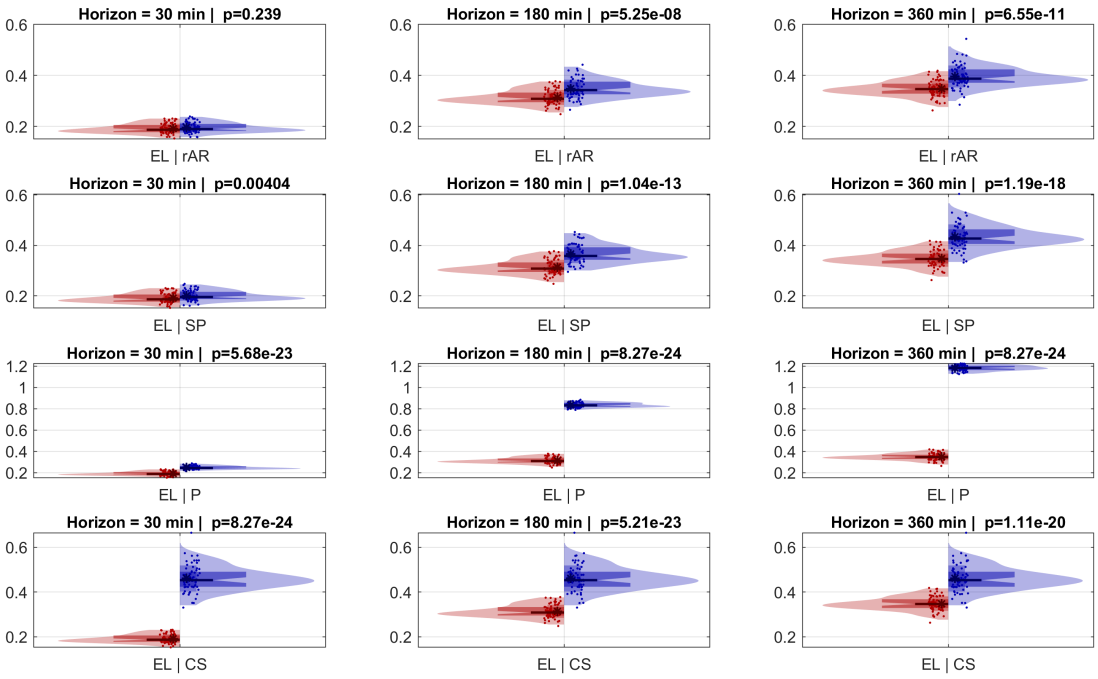


Figure 2: Comparison of the `nRMSE` for deterministic forecasting for time horizons 30 min, 180 min, and 360 min for the `EL` model (left) and four predictors (right): `rAR, SP, P,` and `CS`.

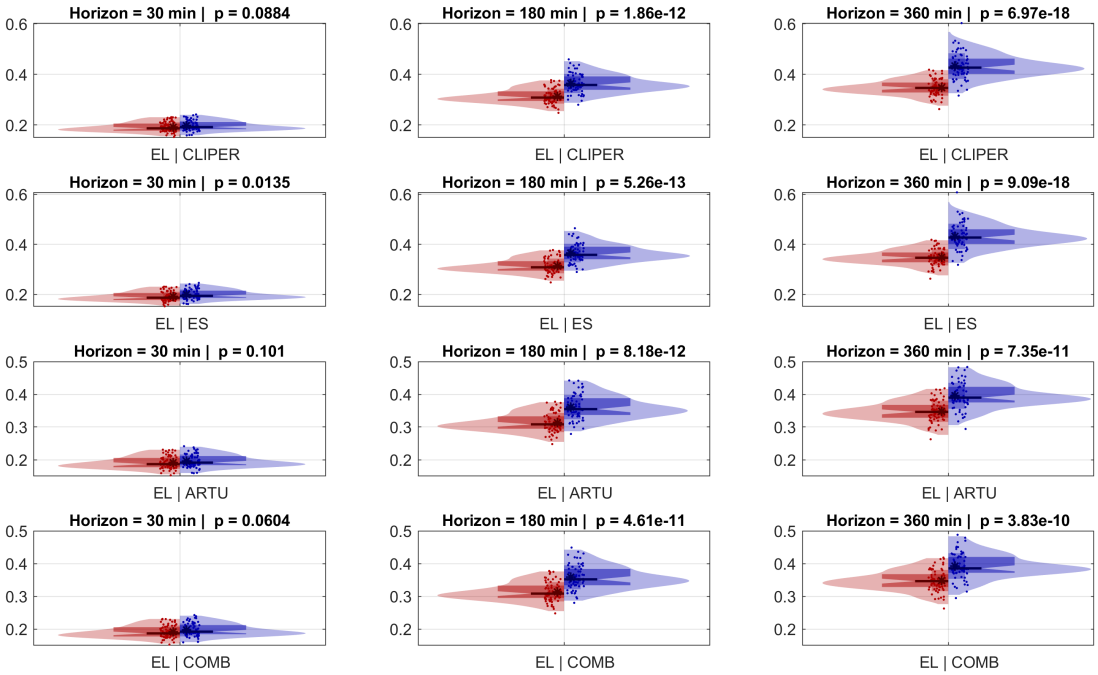


Figure 3: Comparison of the `nRMSE` for deterministic forecasting for time horizons 30 min, 180 min, and 360 min for the `EL` model (left) and four benchmark predictors (right): `CLIPER`, `ES`, `ARTU`, and `COMB`.

3.1.3. Benchmark Expansion and ELM's Dominance

Expanding the benchmark with additional models strengthens the conclusions regarding `ELM`'s superiority (Figure 3). It consistently outperforms all benchmark models (`CLIPER`, `ES`, `ARTU`, and `COMB`) across medium (180 min) and long (360 min) horizons, with highly significant improvements ($p < 10^{-11}$). At the short 30-min horizon, the differences are less pronounced, with some comparisons showing non-significant p-values (e.g., `CLIPER` and `ARTU`), though `EL` still performs competitively. Overall, `ELM` is the most reliable and effective model, particularly excelling in medium and long-term forecasts.

3.1.4. Ranking in Deterministic Case

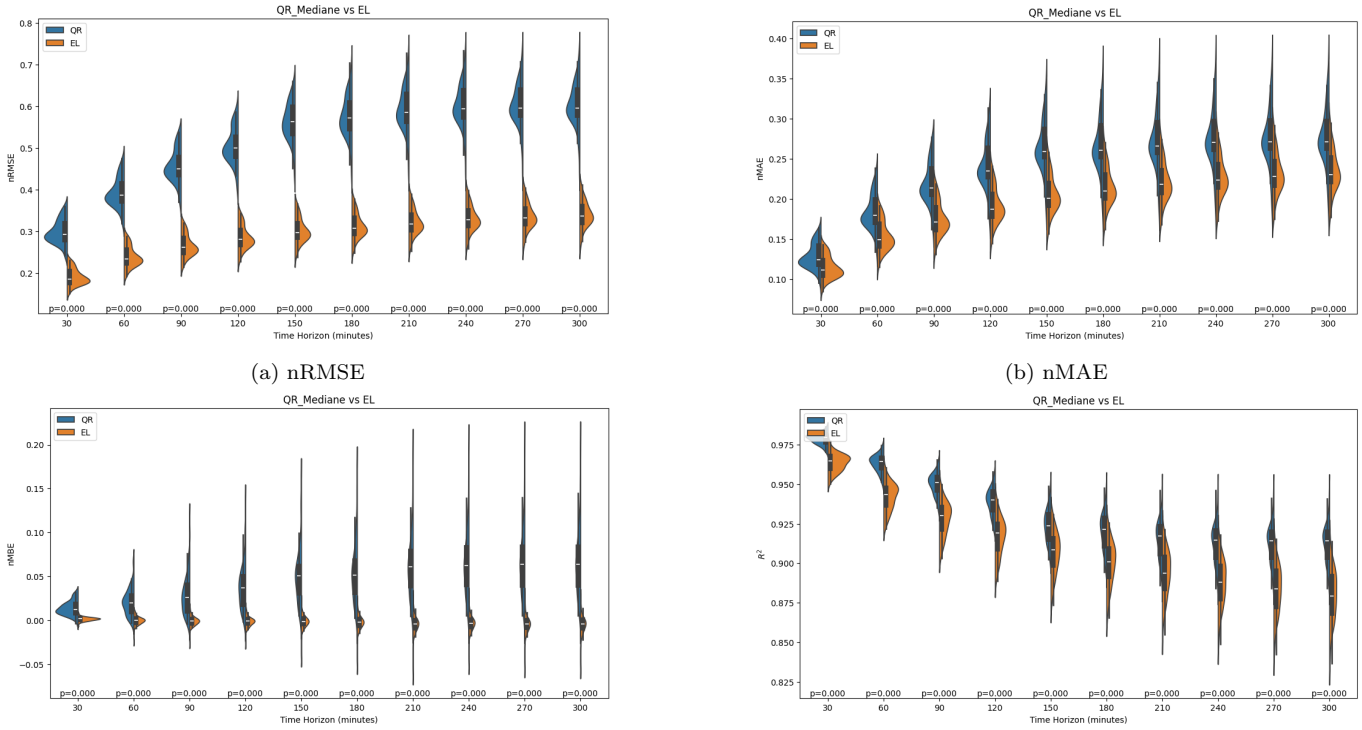
In conclusion, the `ELM` model demonstrates superior performance across all time horizons, particularly excelling in medium and long-term forecasts. While differences are less pronounced at the 30-min horizon, `ELM` remains competitive, solidifying its position as the most reliable and effective forecasting model. Appendix E provide additional information allowing to finally rank all the tested models in the case of point prediction. From best to worst is as follows: `EL`, `AR`, `rAR`, `CLIPER`, `SP`, `ARTU`, `ES`, `COMB`, `CS` and `P`. This validates our assumption that a well-chosen adaptative statistical model can efficiently learn the attenuation due to aerosol, water vapour and ozone from previous measurements. The comparison with simpler models further demonstrates that this can only be achieved using advanced statistical models. A comparison with approaches based on clearsky model finally shows that, when learned from past observations, cloud-free atmospheric modelling is more accurate that when this information is drawn from a Clearsky model.

3.2. Probabilistic Case

Including probabilistic forecasts complements deterministic predictions by quantifying uncertainty, which is critical for optimizing energy systems. This approach allows better risk management, improved resource allocation, and enhanced decision-making under variability, leveraging advanced statistical methods to capture forecast distributions and tail risks.

3.2.1. Median Estimation from Probabilistic Case

One advantage of working directly with probabilistic forecasts is that estimating the median, or 0.5 quantile, inherently provides a deterministic prediction. In the case of Quantile Regression (`QR`), this facilitates seamless integration of probabilistic and deterministic approaches. Similarly, for `ELM`, the methodology based on lookup tables ensures that both probabilistic and deterministic forecasts are generated concurrently, enhancing efficiency and flexibility. In this section, and particularly in Figure 4, the results comparing the medians obtained from `QR` and `ELM` are presented. Overall, `ELM` outperforms `QR` Median across most metrics and time horizons, particularly in medium (180 min) and long-term (360 min) forecasts. For `nRMSE` and



(a) nRMSE

(b) nMAE

(c) nMBE

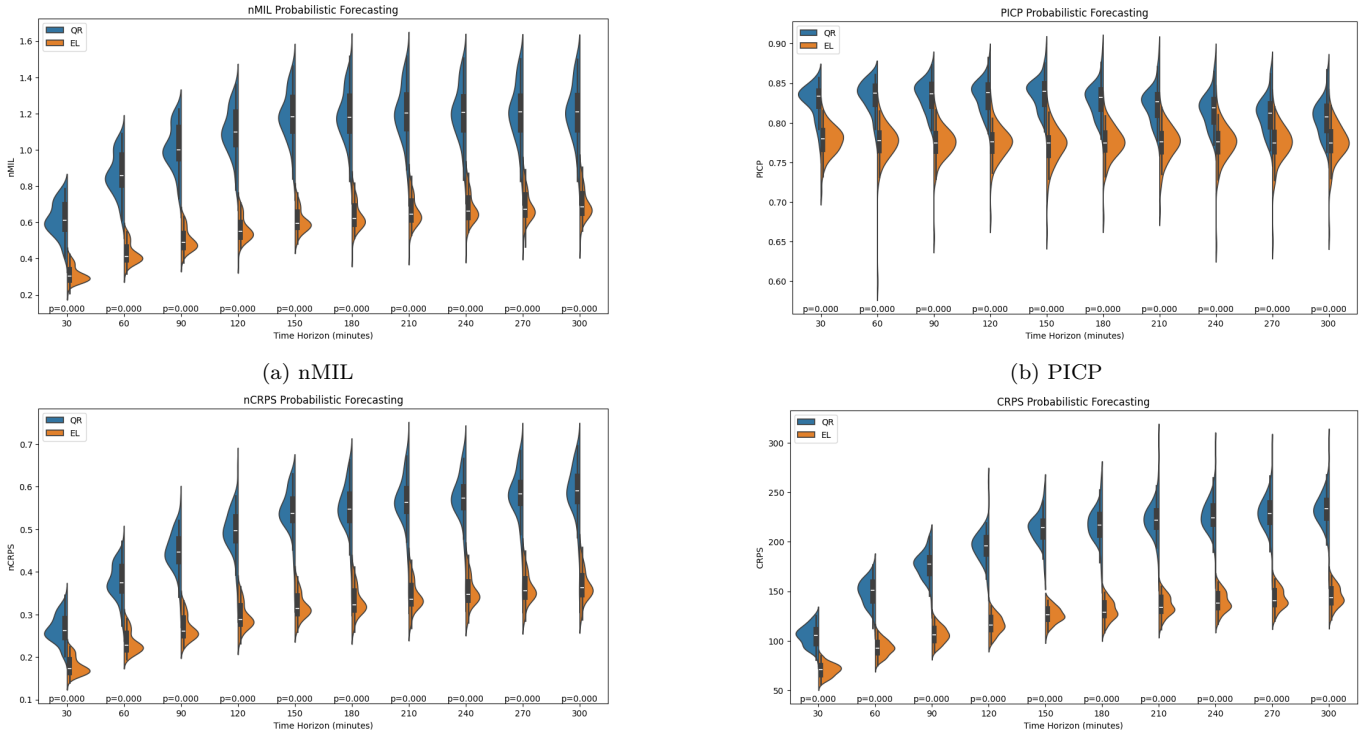
(d) R^2

Figure 4: Evaluation metrics for deterministic forecasting over the 76 SIAR Stations. Each plot illustrates a specific metric (nRMSE, nMAE, nMBE, R^2) across 10 time horizons (30 to 300 min). The comparison includes two models: QR median (blue) and ELM (orange).

nMAE, ELM consistently achieves lower errors, indicating higher predictive accuracy at all horizons. In terms of nMBE, it demonstrates less bias than QR, remaining closer to zero across all horizons. For R^2, QR shows a slight advantage at the 30-min horizon but declines significantly at longer horizons.

3.2.2. QR Versus ELM

Probabilistic error metrics across ten different forecast horizons are presented in Figure 5. The nMIL metric (Figure 5a) evaluates the average width of the prediction intervals, where lower values indicate narrower intervals and thus higher confidence in the forecasts. The results show that both QR and EL maintain consistent interval widths across horizons of 150 min and above, while EL generally provides tighter intervals compared to QR, particularly at longer horizons, indicating higher precision in EL probabilistic forecasts. The PICP metric (Figure 5b) reflects the percentage of actual observations captured within the predicted intervals. For this study, the nominal value is set at $\alpha = 0.2$, corresponding to an 80% prediction interval. A well-calibrated model should have PICP values close to 80%. The results indicate that the QR model consistently achieves PICP values closer to 80%, suggesting better calibration. In contrast, the EL model's slightly lower PICP values highlight a trade-off, as its narrower intervals prioritize precision but might compromise calibration in terms of coverage reliability. The CRPS metric (Figure 5d) evaluates the alignment between forecast distributions and actual observations, where lower values indicate better performance. The EL model consistently outperforms QR across all horizons, indicating that its forecast distributions are closer to actual observations. Similarly, the nCRPS metric (Figure 5c), which normalizes CRPS values, reinforces EL superior performance by highlighting its ability to provide accurate and consistent probabilistic forecasts relative to observed data. For GHI forecasting over a 30-min horizon (mean concerning all sites), the mean Interval Scores (IS) were 115.9 W/m^2 for ELM and 131.2 W/m^2 for QR, as shown in Figure 6. The lower mean IS indicates that ELM is the better model for this horizon. A significance test comparing the IS distributions yielded a p-value of 0.075 for $\alpha < 0.3$, indicating that the difference between ELM and QR is not statistically significant at higher confidence levels ($\geq 70\%$). However, as α increases, the p-value decreases significantly, suggesting that the performance gap becomes statistically significant at lower reliability levels ($< 70\%$). Similar patterns are observed for other forecast horizons, where ELM consistently achieves lower mean IS values than QR. These results, reflected in comparable interval score curves across horizons, confirm that ELM provides superior probabilistic forecasting performance. The statistical significance of the differences,



(a) nMIL

(b) PICP

(c) nCRPS

(d) CRPS

Figure 5: Probabilistic forecasting metrics across 10 time horizons (30 to 300 min, in 30-min increments). Metrics include nMIL, `PICP`, `nCRPS`, and `CRPS` for two models: `QR` (blue) and EL (orange). Statistical p-values for comparison are provided below the violins for each time horizon.

however, varies depending on the reliability level and forecast horizon.

3.2.3. Ranking in Probabilistic Case

The comparison of prediction intervals for `QR` and `ELM` models (Figure 7) highlights the superior performance of `ELM` in probabilistic forecasting. `ELM` provides consistently narrower prediction intervals across all confidence levels (95%, 80%, 50%, and 20%), demonstrating reduced uncertainty and more confident predictions. In contrast, `QR` exhibits significantly wider intervals, particularly at upper than 95% confidence level, indicating higher prediction variance and less precise forecasts. Additionally, the median forecasts (solid lines) from `ELM` align more closely with the observed `GHI` values (dotted red line), showcasing better calibration and accuracy compared to `QR`. While `QR` captures a larger range of possible outcomes due to its wider intervals, this comes at the cost of reduced precision and potentially less actionable predictions. This trade-off between precision and coverage further supports the selection of `ELM` for applications requiring high-confidence, actionable forecasts in energy systems.

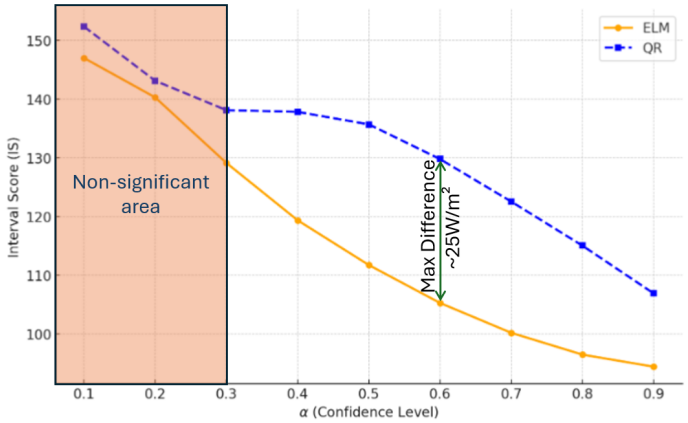


Figure 6: Interval Score (IS) comparison for 30-min horizon and for `ELM` and `QR` models across confidence levels α.

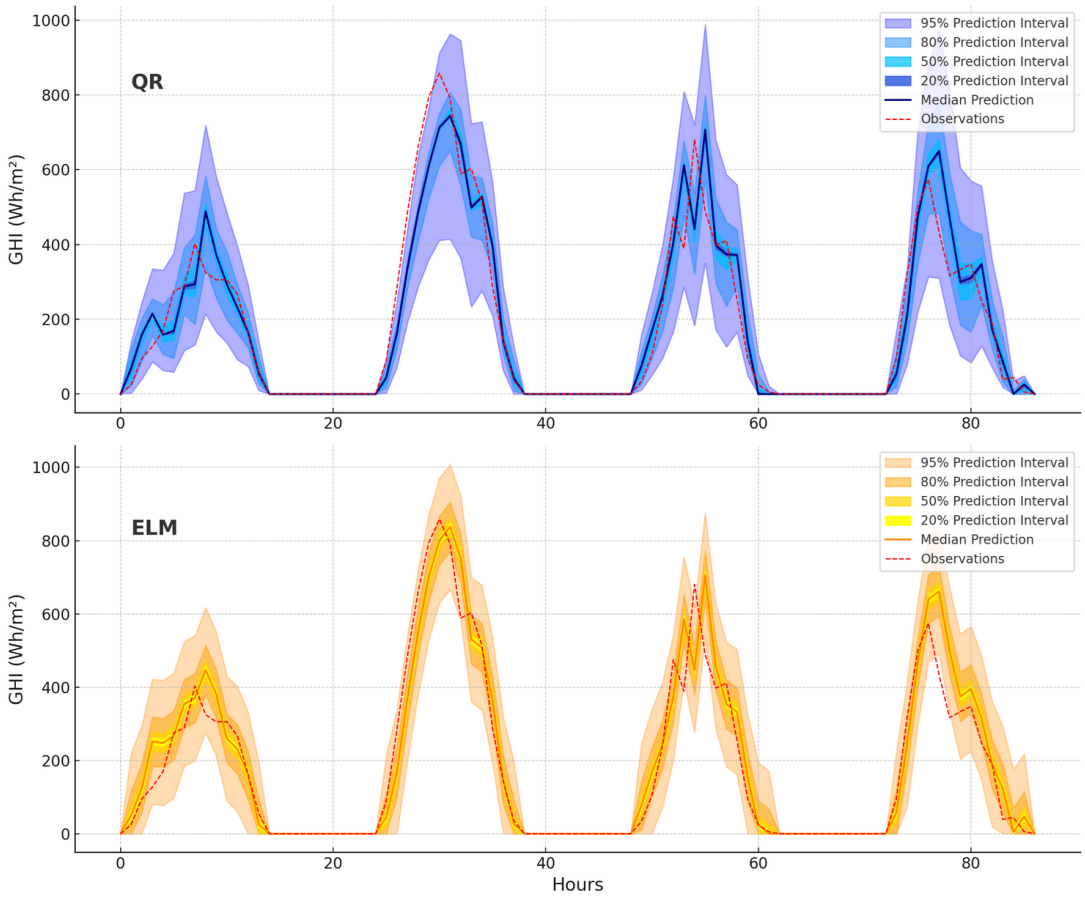


Figure 7: Prediction intervals and median forecasts for Quantile Regression (`QR`, top) and Extreme Learning Machine (`ELM`, bottom) models for `GHI` over an extended period of four days and an horizon 1h.

3.3. Synergies Between Clearsky and Clearsky-Free Approaches

The Clearsky-Free approach demonstrated in this study provides robust results in both deterministic and probabilistic forecasting, but there is substantial potential to further enhance accuracy by integrating the strengths of Clearsky models such as `McClear` and its forecast-oriented counterpart, `CAMS`. These models provide high-quality irradiance baselines, while `ELM` excel at dynamically adapting to local atmospheric variability and temporal dependencies. By coupling these approaches, a hybrid model could outperform traditional methods, leveraging the theoretical precision of Clearsky models and the adaptability of `ELM`. Clearsky models can also enhance the reliability of input data through quality control, improving the overall robustness of the forecasting system. This is particularly beneficial in regions with sparse or unreliable ground-based measurements, where satellite-derived models like `McClear` serve as dependable references. A promising future direction involves creating hybrid models that combine the theoretical baselines of `McClear` with the adaptability of `ELM` for real-time atmospheric variations. Such models would improve solar irradiance forecasts, particularly in regions with complex dynamics. To enhance robustness and generalizability, training datasets should include diverse climatic conditions, and transfer learning techniques should be integrated to enable global applicability without location-specific historical data. These advancements would provide accurate predictions for any irradiance component (e.g., `GHI`, DHI, BNI, GTI) and configurations like tilted or azimuthal orientations. Another key research avenue is refining probabilistic forecasting through advanced uncertainty quantification techniques, such as Gaussian Processes or Bayesian Neural Networks, to improve forecast reliability. Additionally, the application of Clearsky-Free methodologies in decision-support systems for energy grid operators must be explored, ensuring scalability and computational efficiency for real-time operational contexts. By pursuing these directions, Clearsky-Free methodologies can evolve into more reliable and flexible tools for solar irradiance forecasting, addressing diverse operational and environmental challenges.

4. Conclusion

This article presents a groundbreaking method for solar irradiance forecasting without relying on the Clearsky model, offering performance that rivals or surpasses traditional approaches dependent on Clearsky. Specifically, an adaptive statistical model effectively learns the attenuation caused by aerosols, water vapor, and ozone using past measurements. By eliminating intermediate normalization steps, the methodology streamlines the forecasting process and improves reliability, paving the way for more accurate and efficient atmospheric modeling. Leveraging raw `GHI` data and advanced machine learning techniques such as Extreme Learning Machines, it captures solar irradiance's periodicity and local variability, with deterministic and probabilistic forecasting providing robust uncertainty estimates. Conventional Clearsky-Based models are limited by synchronization issues, inaccuracies at night, and uncertainties at low solar elevations, making them inadequate in dynamic environments. Clearsky-Free approaches overcome these challenges. This work critiques in fact the flaws of the `CSI`, particularly its multiplicative framework, which leads to division by zero and timestamp inconsistencies near sunrise and sunset. Note that the filtering data process using a solar angle threshold of 10–20° is insufficiently rigorous. Clearsky-Free methods for sure, or maybe additive models, offer simpler and more robust alternatives. Future improvements could include integrating high-quality Clearsky models (like `McClear` or `CAMS`) directly as input of machine learning frameworks for enhanced accuracy and resilience, paving the way for more effective solar irradiance forecasting systems. The stationarization problem is significantly more complex for the tilted component of global irradiance (GTI), as it involves not only the Clearsky model error but also the additional error introduced by the transposition model. Clearsky-free models can be highly promising for predicting this component, which is directly linked to Photovoltaic production. This will be the focus of future works to explore this possibility further. Collaboration between quality control and advanced Clearsky-Free methodologies significantly improves solar irradiance prediction under challenging atmospheric conditions or sparse data scenarios. While the proposed approach remains statistically viable for operational use, particularly under typical atmospheric conditions, extreme pollution events may challenge its accuracy. In such cases, integrating `AOD`-sensitive models or hybrid approaches could enhance robustness and adaptability. Striking a balance between simplicity, predictive power, and computational efficiency remains a key objective for future developments in Clearsky-Free forecasting frameworks. The implications of these developments are vast. Integrating these methods into decision-support systems for smart grid management, energy trading, and renewable energy integration offers transformative potential. The clears-sky free approach also could support more accurate direct and diffuse solar radiation forecasts, notably for inclined surfaces where applying clear-sky models is complicated. Real-time, accurate, and reliable forecasts can optimize solar energy usage in modern energy systems, contributing to efficient grid operations and accelerating the transition toward a low-carbon energy future.

CRediT Authorship Contribution Statement

CV: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. MD: Writing – review & editing, Methodology. GN: Writing – review & editing, Supervision. YMSD: review & editing, Methodology. MA: Validation, Software, Data curation. LGG: Investigation, Data curation.

Appendix A. About Stationarity in Solar Energy Forecasting Context

In time series analysis, stationarity is a critical concept referring to a process whose statistical properties, such as mean $\mathbb{E}[y_t]$, variance $\mathrm{Var}(y_t)$, and autocovariance $\mathrm{Cov}(y_t, y_{t+h})$, remain invariant over time (Voyant et al., 2017). A stationary series satisfies:

$$\mathbb{E}[y_t] = \mu, \quad \mathrm{Var}(y_t) = \sigma^2, \quad \text{and} \quad \mathrm{Cov}(y_t, y_{t+h}) = \gamma(h), \tag{A.1}$$

where μ and σ^2 are constants, and $\gamma(h)$ depends only on the lag h. Stationarity simplifies statistical modeling by ensuring that the estimated relationships within the data are consistent over time, a prerequisite for many traditional models like `ARIMA`. In the context of solar energy forecasting, time series data are often non-stationary due to periodic components (daily and seasonal cycles), trends, and noise arising from atmospheric conditions. Transformations such as differencing ($y_t - y_{t-1}$) or detrending ($y_t - T(t)$ or $y_t/T(t)$ where $T(t)$ models a deterministic trend) are conventionally employed to enforce stationarity. However, these transformations risk discarding valuable information intrinsic to the series, such as inherent

seasonal patterns or long-term dependencies, particularly in univariate settings. Recent advancements in machine learning (Liu et al., 2024), such as Extreme Learning Machines (ELM), challenge the necessity of strict stationarity, leveragging the raw structure of the data without requiring explicit stationarization (Rahman et al., 2021; Silva et al., 2021; Haixu et al., 2022). The ELM training process, relying on a pseudo-inverse solution for $\boldsymbol{\beta}$ (see Section 2.2), is robust to non-stationary inputs, capturing periodic and stochastic variations directly within the endogenous structure of the series (Pan and Zhao, 2013). In univariate solar irradiance forecasting, the seasonal and diurnal cycles are embedded within the series itself. By feeding p-lagged observations $\{y_{t-1}, y_{t-2}, \ldots, y_{t-p}\}$ as input to ELMs, these models can inherently learn the temporal dependencies and periodicities without explicit detrending or differencing. This approach aligns with the view that stationarization may not be essential in modern forecasting paradigms, particularly when the model architecture itself is designed to extract complex patterns from raw data. Studies show that ELM deliver good accuracy while requiring significantly fewer resources compared to other deep learning models. This makes them particularly effective in operational short-term solar forecasting tasks, achieving good results with greatly reduced computational overhead (Al-Dahidi et al., 2018). The ability to handle non-stationary data directly underscores the potential of ELMs to maybe, redefine conventional practices in renewable energy forecasting.

Appendix B. Overview of the Extreme Learning Machine (ELM) for Solar Forecasting

The ELM architecture consists of three layers: an input layer with N_{input} neurons, a hidden layer with N_{hidden} neurons, and a single-neuron output layer. For the 30-minute forecasting horizon, we use $N_{\text{input}} = 78$ historical features and $N_{\text{hidden}} = 472$ neurons in the hidden layer. The model operates as follows:

- Weight Initialization: Input-to-hidden weights $\mathbf{W} \in \mathbb{R}^{N_{\text{hidden}} \times N_{\text{input}}}$ are randomly drawn from a uniform distribution $W_{ij} \sim \mathcal{U}(-1, 1)$. A bias term $\mathbf{b} \in \mathbb{R}^{N_{\text{hidden}}}$ is also initialized randomly;
- Hidden Layer Activation: Each hidden neuron applies a nonlinear transformation to the input, with a mix of 60% sigmoidal activation and 40% Gaussian radial basis activation $H_i = g(W_i X + b_i)$, where $g(\cdot)$ is Sigmoid Activation (60%; $\frac{1}{1+e^{-x}}$) or Gaussian Activation (40%; $\exp(-\frac{x^2}{2\sigma^2})$ where σ is set to 1 in our implementation. The proportion of sigmoid and Gaussian neurons is controlled by the threshold parameter $T = 0.6$;
- Output Weight Computation: Unlike conventional neural networks, ELM does not require backpropagation. Instead, output weights β are computed analytically using Ridge Regression:

$$\beta = (H^T H + \lambda I)^{-1} H^T Y, \tag{B.1}$$

  where H is the hidden layer matrix and $\lambda = 0.2$ is a regularization parameter;
- Training Procedure: To improve robustness and mitigate variance due to weight initialization, we conduct 200 independent runs ($j \in [1, 200]$) with different initializations of $\mathbf{W}$ and $\mathbf{b}$. The best model is selected using a winner-takes-all strategy based on in-sample validation performance:

$$\beta^* = \underset{\beta_j}{\text{argmin}} \left(\text{nRMSE}(\beta_j)\right); \tag{B.2}$$

- Final Model for Out-of-Sample Testing: The selected model is then evaluated on the out-of-sample test set and compared with benchmark models.

To further clarify the structure of ELM, an example of architecture is shown in Figure B.8: The model consists of $N_{\text{input}} = 78$ input features, $N_{\text{hidden}} = 472$ hidden neurons, and a single output neuron, leading to a total of:

$$\begin{aligned}(N_{\text{input}} + 1) \times N_{\text{hidden}} + (N_{\text{hidden}} + 1) \times N_{\text{output}} &= (78 + 1) \times 472 \\ &+ (472 + 1) \times 1 = 37,401 \text{ parameters.}\end{aligned} \tag{B.3}$$

Among these parameters, 37,368 parameters (99.91%) are randomly initialized and remain fixed throughout training, and 473 parameters (1.26%) are optimized using Ridge Regression. The model includes four key hyperparameters: the number of hidden neurons N_{hidden}, the activation mix ratio T, the ridge regression

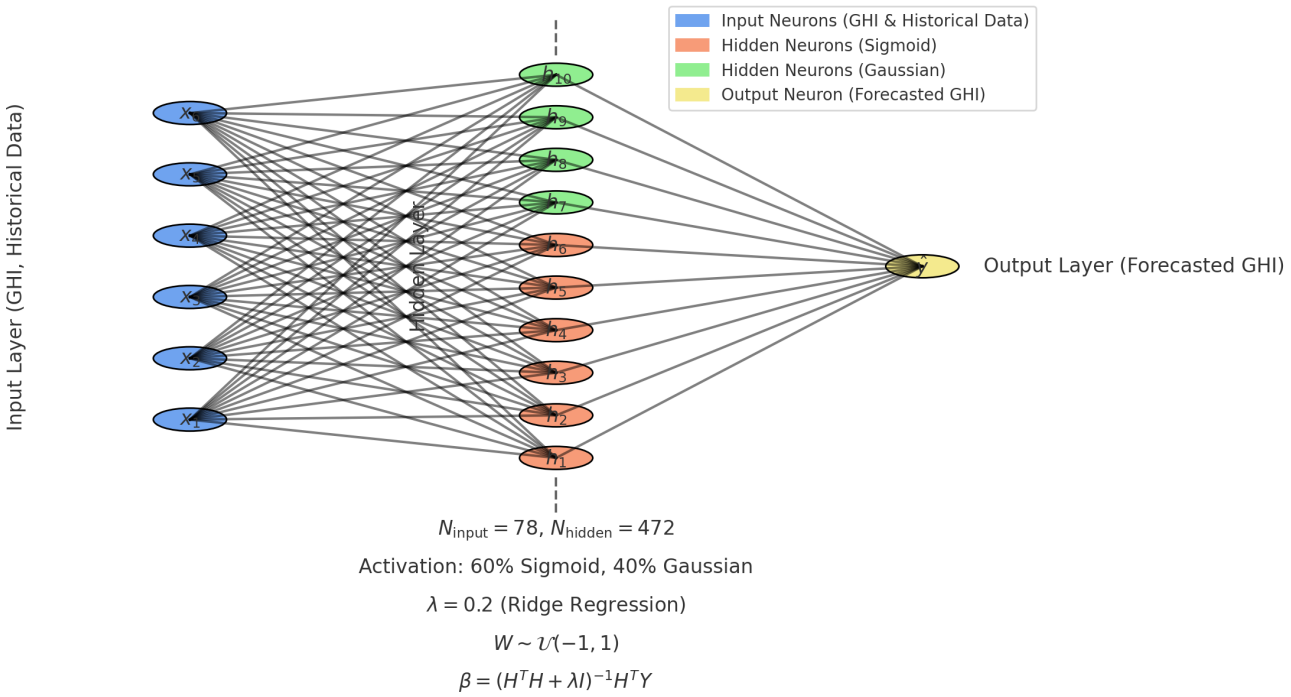


Figure B.8: Schematic representation of the Extreme Learning Machine (`ELM`) architecture for a 30-minute forecast horizon. The model consists of an input layer ($N_{\text{input}} = 78$), a hidden layer ($N_{\text{hidden}} = 472$) with 60% sigmoid and 40% Gaussian neurons, and a single-output neuron predicting future `GHI`.

regularization term λ, and the number of runs N_{runs} for the winner-takes-all selection. Despite its analytical training process, `ELM` can be computationally intensive due to the matrix inversion in the output weight calculation, especially when scaling to larger datasets. Additionally, operational deployment may present challenges for energy operators, particularly in ensuring the robustness of predictions across diverse weather conditions, data availability, and the interpretability of the learned model. These aspects highlight the balance between model complexity and practical usability in experimental forecasting applications. Moreover, while our approach statistically learns the effects of atmospheric variations from historical data, it does not explicitly differentiate aerosol contributions nor adapt to sudden pollution peaks. This is a known limitation, as real-time AOD retrievals are not incorporated. However, given the rarity of extreme pollution events (fewer than five pollution alert days in 2024 in Nice City), this trade-off remains acceptable in the context of operational solar forecasting. Nonetheless, if such events were to become more frequent, future work could explore hybrid approaches integrating real-time aerosol information.

Appendix C. Non-Parametric Prediction Interval Generation

In addition to Quantile Regression, a non-parametric methodology was implemented to derive prediction intervals directly from the residuals of deterministic forecasts (Pinson et al., 2007; Wang et al., 2022). This approach uses a lookup table generated during the training phase (in-sample data), which captures empirical relationships between prediction errors and desired coverage levels (between 0 and 1). During the training phase, the residuals of deterministic forecasts ($y - \hat{y}$) are computed. For each residual, the standard deviation $\hat{\sigma}$ is calculated and depend on the site and the horizon considered. For a given coverage level $(1 - \alpha)$, prediction intervals are defined for out-sample data as $\underline{y}(\alpha) = \hat{y} - k_\alpha \times \hat{\sigma}$, $\overline{y}(\alpha) = \hat{y} + k_\alpha \times \hat{\sigma}$, where $\hat{y}$ is the deterministic forecast, and k_α is a unitless scaling factor derived from the residual distribution to achieve the desired coverage. `PICP` and `nMIL` are then computed according equations in Section 2.3. By interpolating these metrics across coverage levels α, a lookup table is constructed (several values k_α^{lookup}). This table maps each α to a corresponding scaling factor k_α, enabling the generation of prediction intervals without assuming a specific error distribution. During testing, the lookup table is used to compute prediction intervals for a given α. For each forecast $\hat{y}$, the prediction intervals are derived as $\underline{y}(\alpha) = \hat{y} - k_\alpha^{lookup} \times \hat{\sigma}$, $\overline{y}(\alpha) = \hat{y} + k_\alpha^{lookup} \times \hat{\sigma}$. This process yields prediction interval over a nominal coverage $(1 - \alpha)$ according $[\underline{y}(\alpha) < \hat{y} < \overline{y}(\alpha)]_{(1-\alpha)}$. The lookup-table-based methodology provides several notable advantages. By avoiding assumptions about error distributions, it adapts flexibly to diverse atmospheric conditions, ensuring

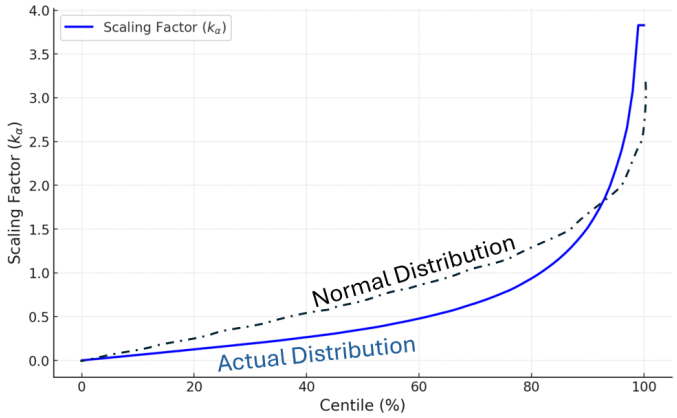


Figure C.9: Lookup table representing scaling factors k_α (mean of all stations concerning 30 min horizon) across centiles, used to compute prediction intervals for probabilistic forecasts.

accurate predictions even in non-stationary environments. Its robustness stems from empirical relationships established between residuals and forecasts during the training phase, allowing for the dynamic generation of reliable prediction intervals. Additionally, this non-parametric approach complements `QR` by offering a computationally efficient alternative for deriving probabilistic forecasts (Perna et al., 2024). Together, the Quantile Regression (`QR`) and the lookup table integrated with the `ELM` model, offer a scalable and practical solution for operational solar irradiance prediction, capable of delivering probabilistic forecasts. The lookup table in Figure C.9 illustrates the scaling factors (k_α) derived during the training phase for each centile (0% to 100%). The curve is smooth, reflecting a well-defined relationship between residual variance and coverage levels. Lower centiles show minimal scaling, while higher centiles exhibit increasing values, consistent with broader prediction intervals required for higher coverage. This structure ensures a dynamic adjustment of prediction intervals, allowing for robust probabilistic forecasts under varying atmospheric conditions. It will be noted that $k_\alpha = 2$ is close to providing a nominal coverage of 95%, as in the parametric Gaussian case.

Appendix D. Quantile-Based Approximation of `CRPS`

In this study, `CRPS` is approximated using quantiles instead of the full `CDF` (Bröcker, 2012). Let Q_α denote the predicted quantile at probability level $(1-\alpha)$, such that Q_α satisfies $P(Y \leq Q_\alpha) = \alpha$, for $\alpha \in [0, 1]$. The quantile-based `CRPS` is computed as the weighted sum of absolute deviations between the predicted quantiles Q_α and the observation y via $\texttt{CRPS}(Q, y) = \sum_{i=1}^{N_\alpha} w_i \cdot |Q_{\alpha_i} - y|$, where N_α is the number of quantiles used (here, $N_\alpha = 101$), and the weights w_i correspond to the quantile resolution ($w_i = \Delta\alpha = 0.01$ for evenly spaced quantiles). This reduces to $\texttt{CRPS}(Q, y) = 0.01 \cdot \sum_{i=1}^{101} |Q_{\alpha_i} - y|$. The classical and quantile-based `CRPS` formulations are related through the fact that the `CDF` can be approximated by quantiles using $F(x) \approx \sum_{\alpha_i \leq x} w_i$. Thus, the quantile-based `CRPS` is a discrete approximation of the continuous integral, where the quantiles Q_α replace the explicit `CDF`. The classical `CRPS` provides a precise evaluation when a continuous `CDF` is available, making it suitable for models that predict full distributions, such as Gaussian processes. In contrast, the quantile-based `CRPS` is particularly efficient for models like Quantile Regression, which inherently provide discrete quantile estimates. This approach avoids the need to reconstruct or approximate the `CDF`, offering a simpler yet robust alternative. The quantile-based `CRPS` also provides practical advantages. By avoiding integration over a continuous distribution, it achieves computational efficiency, making it well-suited for large datasets. Furthermore, it aligns naturally with quantile-based models, such as Quantile Regression, without requiring additional assumptions. This method directly handles probabilistic forecasts expressed as quantiles, simplifying the evaluation pipeline. The accuracy of the quantile-based approximation depends on the resolution of the quantiles (Zamo and Naveau, 2018). For evenly spaced quantiles, as in this study (101 quantiles), the approximation is sufficiently precise for practical purposes. To ensure interpretability, a normalized `CRPS` is also computed as $\texttt{nCRPS} = CRPS \cdot \mathrm{E}[y]^{-1}$. This normalization enables a relative comparison across datasets with different scales.

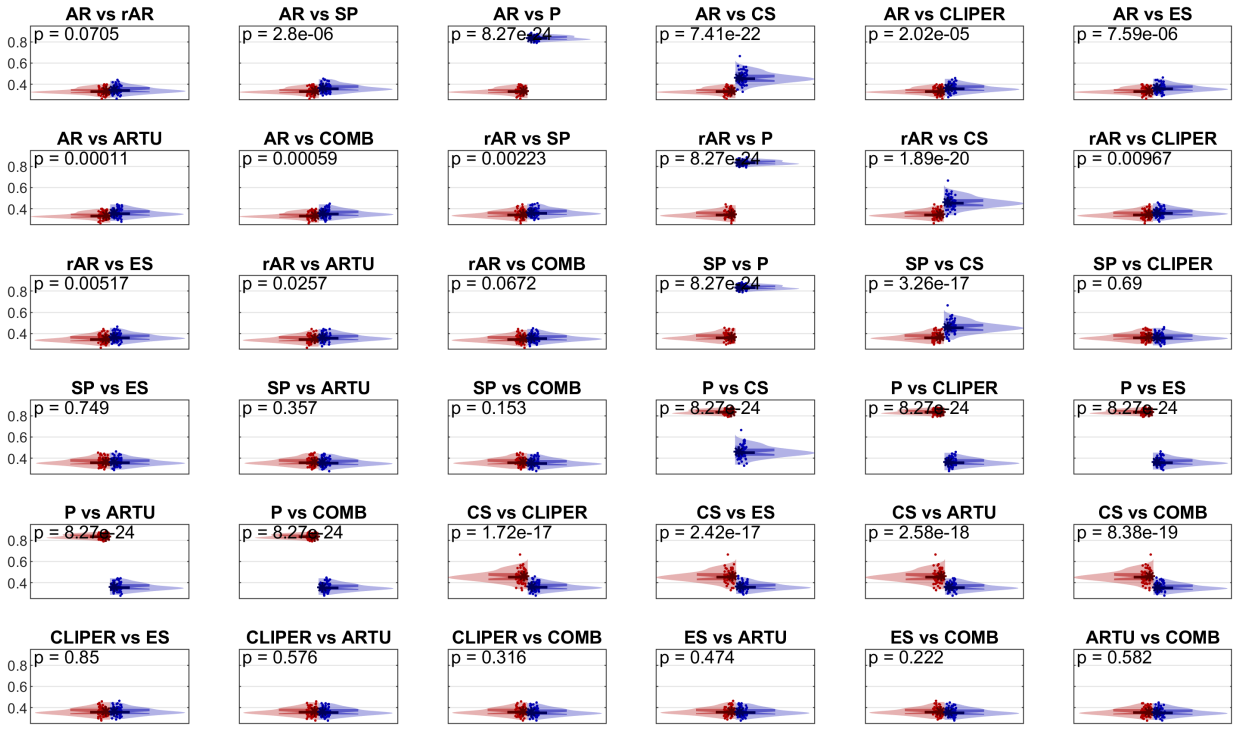


Figure E.10: Cross-comparison of benchmark methods for all sites and a 180-min horizon. Each subplot shows `nRMSE` distributions for a pair of methods, with p-values indicating the significance of the difference.

Appendix E. Analysis of Models Performance

Figure E.10 compares the normalized Root Mean Square Error (`nRMSE`) distributions of the evaluated methods. `AR` consistently outperforms most models, including `COMB` ($p = 0.00059$) and `SP` ($p = 2.8 \times 10^{-6}$), with distributions centered at lower `nRMSE` values. While `AR` and `rAR` show near-equivalent performance ($p = 0.0705$), `AR` holds a slight edge with lower median errors. COMB and `SP` are competitive but fail to match `AR`, especially against weaker methods like `P` and `CS`, which consistently rank as the poorest performers ($p = 8.27 \times 10^{-24}$ and 7.41×10^{-22}, respectively). `CLIPER` and `ES` demonstrate intermediate performance, with no significant differences between them ($p = 0.85$). Overall, `AR` emerges as the most reliable method, closely followed by `rAR`. `CLIPER` and `SP` provide simpler, yet competitive alternatives, while `COMB`, `ES`, and `ARTU` offer strong but more sophisticated solutions. `P` and `CS` remain the least effective models.

References

Akarslan, E., Hocaoglu, F., Edizkan, R., 2018. Novel short term solar irradiance forecasting models. Renew Energy 123, 58–66. doi:10.1016/J.RENENE.2018.02.048.

Al-Dahidi, S., Ayadi, O., Adeeb, J., Alrbai, M., Qawasmeh, B.R., 2018. Extreme learning machines for solar photovoltaic power predictions. Energies 11. URL: https://www.mdpi.com/1996-1073/11/10/2725, doi:10.3390/en11102725.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, 1–300. URL: www.fao.org/4/x0490e/x0490e00.htm.

Azarpour, A., Mohammadzadeh, O., Rezaei, N., Zendehboudi, S., 2022. Current status and future prospects of renewable and sustainable energy in north america: Progress and challenges. Energy Conversion and Management 269, 115945. doi:10.1016/j.enconman.2022.115945.

Bröcker, J., 2012. Evaluating raw ensembles with the continuous ranked probability score. Quarterly Journal of the Royal Meteorological Society 138. doi:10.1002/qj.1891.

Chodakowska, E., Nazarko, J., Nazarko, Ł., Rabayah, H.S., 2024. Solar radiation forecasting: A systematic meta-review of current methods and emerging trends. Energies 17. doi:10.3390/en17133156.

David, M., Ramahatana, F., Trombe, P., Lauret, P., 2016. Probabilistic forecasting of the solar irradiance with recursive arma and garch models. Solar Energy 133, 55–72. doi:10.1016/J.SOLENER.2016.03.064.

Despotovic, M., Voyant, C., Garcia-Gutierrez, L., Almorox, J., Notton, G., 2024. Solar irradiance time series forecasting using auto-regressive and extreme learning methods: Influence of transfer learning and clustering. Applied Energy 365, 123215. doi:10.1016/j.apenergy.2024.123215.

Diagne, M., David, M., Lauret, P., Boland, J., Schmutz, N., 2013. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews 27, 65–76. doi:10.1016/j.rser.2013.06.042.

Frimane, Â., Munkhammar, J., van der Meer, D., 2022. Infinite hidden markov model for short-term solar irradiance forecasting. Solar Energy 244, 331–342. doi:10.1016/J.SOLENER.2022.08.041.

Garcia-Gutierrez, L., Voyant, C., Notton, G., Almorox, J., 2022. Evaluation and comparison of spatial clustering for solar irradiance time series. Applied Sciences 12. doi:10.3390/app12178529.

Gneiting, T., Lerch, S., Schulz, B., 2023. Probabilistic solar forecasting: Benchmarks, post-processing, verification. Solar Energy 252, 72–80. doi:10.1016/J.SOLENER.2022.12.054.

Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102, 359–378. doi:10.1198/016214506000001437.

Gueymard, C.A., 2008. Rest2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – validation with a benchmark dataset. Solar Energy 82, 272–285. doi:10.1016/j.solener.2007.04.008.

Haixu, Y.L., Wu, Jianmin, W., Mingsheng, L., 2022. Non-stationary transformers: Rethinking the stationarity in time series forecasting. Advances in Neural Information Processing Systems 35. doi:10.48550/arXiv.2205.14415.

Herrería-Alonso, S., Suárez-González, A., Rodríguez-Pérez, M., Rodríguez-Rubio, R.F., López-García, C., 2020. A solar altitude angle model for efficient solar energy predictions. Sensors 20. doi:10.3390/s20051391.

Hoyos-Gómez, L., Ruiz-Muñoz, J., Ruiz-Mendoza, B., 2022. Short-term forecasting of global solar irradiance in tropical environments with incomplete data. Appl Energy 307, 118192. doi:10.1016/J.APENERGY.2021.118192.

Huang, G.B., Zhu, Q.Y., Siew, C.K., 2006. Extreme learning machine: Theory and applications. Neurocomputing 70, 489–501. doi:10.1016/j.neucom.2005.12.126.

International Energy Agency (IEA), 2021. Renewables 2021: Analysis and forecasts to 2026. IEA, Paris URL: www.iea.org/reports/renewables-2021. licence: CC BY 4.0.

Koenker, R., 2005. Quantile Regression. Cambridge University Press. doi:10.1017/CBO9780511754098.

Kumar, D.S., Yagli, G.M., Kashyap, M., Srinivasan, D., 2020. Solar irradiance resource and forecasting: a comprehensive review. IET Renewable Power Generation 14, 1641–1656. doi:10.1049/iet-rpg.2019.1227.

Lauret, P., Alonso-Suárez, R., Le Gal, J., Salle, L., David, M., 2022. Solar forecasts based on the clear sky index or the clearness index: Which is better? Solar 2, 432–444. doi:10.3390/SOLAR2040026.

Lauret, P., David, M., Pedro, H.T.C., 2017. Probabilistic solar forecasting using quantile regression models. Energies 10. doi:10.3390/en10101591.

Lauret, P., Rodler, A., Muselli, M., David, M., Diagne, H.M., Voyant, C., 2012. A bayesian model committee approach to forecasting global solar radiation, in: WREF 2012 : World Renewable Energy Forum, Denver, United States. p. 1. URL: https://hal.science/hal-00682217.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter-Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J.W., Morcrette, J.J., 2013. McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. Atmospheric Measurement Techniques 6, 2403–2418. doi:10.5194/amt-6-2403-2013.

Li, D.H., Lou, S., Lam, J.C., Wu, R.H., 2016. Determining solar irradiance on inclined planes from classified CIE (international commission on illumination) standard skies. Energy 101, 462–470. doi:10.1016/j.energy.2016.02.054.

Liu, Z., Cheng, M., Li, Z., Huang, Z., Liu, Q., Xie, Y., Chen, E., 2024. Adaptive normalization for non-stationary time series forecasting: A temporal slice perspective. Advances in Neural Information Processing Systems 36. doi:10.5555/3666122.3666750.

Ma, Y., Zhang, X., Mei, S., Zhen, Z., Gao, R., Zhou, Z., 2020. Ultra-short-term solar power forecasting based on a modified clear sky model, in: 2020 39th Chinese Control Conference (CCC), pp. 5311–5316. doi:10.23919/CCC50068.2020.9189533.

Makridakis, S., Wheelwright, S.C., Hyndman, R.J., 1998. Forecasting methods and applications. 3rd edition ed., Wiley. doi:10.2307/2287014.

Mann, H.B., Whitney, D.R., 1947. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics 18, 50 – 60. doi:10.1214/aoms/1177730491.

Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. The computer journal 7, 308–313. doi:10.1093/comjnl/7.4.308.

Pan, F., Zhao, H., 2013. Online sequential extreme learning machine based multilayer perception with output self feedback for time series prediction. Journal of Shanghai Jiaotong University (Science) 18, 366–375. doi:10.1007/S12204-013-1407-0.

Perna, S., Austnes, P.F., Gerini, F., Chevron, M., Fazio, A.D., Falco, P.D., Paolone, M., 2024. A comparative analysis of empirical copula and quantile regression methods for probabilistic load forecasting, in: 2024 18th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), pp. 1–6. doi:10.1109/PMAPS61648.2024.10667333.

Pinson, P., Nielsen, H., Møller, J., Madsen, H., Kariniotakis, G., 2007. Non-parametric probabilistic forecasts of wind power: required properties and evaluation. Wind Energy 10, 497–516. doi:10.1002/WE.230.

Rahman, M., Shakeri, M., Tiong, S., Khatun, F., Amin, N., Pasupuleti, J., Hasan, M.K., 2021. Prospective methodologies in hybrid renewable energy systems for energy prediction using artificial neural networks. Sustainability 13, 2393. doi:10.3390/SU13042393.

Ruiz-Arias, J.A., 2023. SPARTA: Solar parameterization for the radiative transfer of the cloudless atmosphere. Renewable and Sustainable Energy Reviews 188, 113833. URL: https://linkinghub.elsevier.com/retrieve/pii/S1364032123006901, doi:10.1016/j.rser.2023.113833.

Ruiz-Arias, J.A., Gueymard, C.A., 2018. Worldwide inter-comparison of clear-sky solar radiation models: Consensus-based review of direct and global irradiance components simulated at the earth surface. Solar Energy 168, 10–29. URL: https://www.sciencedirect.com/science/article/pii/S0038092X18301257, doi:https://doi.org/10.1016/j.solener.2018.02.008. advances in Solar Resource Assessment and Forecasting.

Sanfilippo, A., Martin-Pomares, L., Mohandes, N., Perez-Astudillo, D., Bachour, D., 2016. An adaptive multi-modeling approach to solar nowcasting. Solar Energy 125, 77–85. doi:10.1016/J.SOLENER.2015.11.041.

Santamouris, M., Tselepidaki, I., Dris, N., 1990. Evaluation of models to predict solar radiation on tilted surfaces for the mediterranean region. Solar & Wind Technology 7, 585–589. doi:10.1016/0741-983X(90)90067-C.

Shan, S., Li, C., Ding, Z., Wang, Y., Zhang, K., Wei, H., 2022. Ensemble learning based multi-modal intra-hour irradiance forecasting. Energy Convers Manag 270, 116206. doi:10.1016/J.ENCONMAN.2022.116206.

Silva, R.P., Zarpelão, B., Cano, A., Junior, S.B., 2021. Time series segmentation based on stationarity analysis to improve new samples prediction. Sensors (Basel, Switzerland) 21. doi:10.3390/s21217333.

Sobri, S., Koohi-Kamali, S., Rahim, N.A., 2018. Solar photovoltaic generation forecasting methods: A review. Energy Conversion and Management 156, 459–497. doi:10.1016/j.enconman.2017.11.019.

Vicente-Serrano, S., Domínguez-Castro, F., Reig, F., Beguería, S., Tomas-Burguera, M., Latorre, B., Peña-Angulo, D., Noguera, I., Rabanaque, I., Luna, Y., Morata, A., El Kenawy, A., 2022. A near real-time drought monitoring system for spain using automatic weather station network. Atmospheric Research 271, 106095. doi:10.1016/j.atmosres.2022.106095.

Voyant, C., Lauret, P., Notton, G., Duchaud, J.L., Garcia-Gutierrez, L., Faggianelli, G.A., 2022a. Complex-valued time series based solar irradiance forecast. Journal of Renewable and Sustainable Energy 14, 066502. doi:10.1063/5.0128131.

Voyant, C., Notton, G., Duchaud, J.L., García Gutiérrez, L.A., Bright, J.M., Yang, D., 2022b. Benchmarks for solar radiation time series forecasting. Renewable Energy 113, 102–114. doi:10.1016/j.renene.2022.04.065.

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.L., Paoli, C., Motte, F., Fouilloy, A., 2017. Machine learning methods for solar radiation forecasting: A review. Renewable Energy 105, 569–582. doi:10.1016/j.renene.2016.12.095.

Wang, K., Zhang, Y., Lin, F., Wang, J., Zhu, M., 2022. Nonparametric probabilistic forecasting for wind power generation using quadratic spline quantile function and autoregressive recurrent neural network. IEEE Transactions on Sustainable Energy 13, 1930–1943. doi:10.1109/TSTE.2022.3175916.

Yang, D., 2019. Making reference solar forecasts with climatology, persistence, and their optimal convex combination. Solar Energy 193, 981–985. doi:10.1016/j.solener.2019.10.006.

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

Zamo, M., Naveau, P., 2018. Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts. Mathematical Geosciences 50, 209–234. doi:10.1007/s11004-017-9709-7.

Zheng, Y., Ge, Y., Muhsen, S., Wang, S., Elkamchouchi, D.H., Ali, E., Ali, H.E., 2023. New ridge regression, artificial neural networks and support vector machine for wind speed prediction. Advances in Engineering Software 179, 103426. doi:10.1016/j.advengsoft.2023.103426.