https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Weather and air pollution influences on solar energy performance in West Africa: A Bayesian nonlinear mixed-effects approach

Konin Pierre-Claver Kakou[1,2,4], Dungall Laouali[2], Boko Aka[3], and Georg Frey[4]

[1]Doctoral Research Program in Climate Change and Energy, West African Science Service Centre on Climate Change and Adapted Land Use (WASCAL), BP 10662 Niamey, Niger
[2]Département de Physique, Université Abdou Moumouni, BP 10 896 Niamey, Niger
[3]Institut de Recherche sur les Energies Nouvelles, 02 BP 801 Abidjan 02, Côte d'Ivoire
[4]Chair of Automation and Energy Systems, Saarland University, Saarbrücken Campus, 66123 Saarbrücken, Germany

Correspondence: Konin Pierre-Claver Kakou (pierre-claver.kakou@uni-saarland.de)

Abstract. In the context of the global shift toward an energy transition in which solar power plays an increasingly significant role, solar resource forecasting has drawn considerable attention from researchers and stakeholders. Accurate predictions enable better planning, controlled integration of solar energy and improved project profitability. In this regard, understanding how individual predictors influence solar radiation is crucial for selecting relevant inputs, reducing computational costs and enhancing model performance. However, this relationship is still frequently modelled as linear, an assumption that neglects the complex, nonlinear and hierarchical interactions that characterize atmospheric processes. In this study, a Bayesian mixed-effects model was developed to assess how meteorological variables and air pollution affect solar energy generation, using ground-based observations and satellite-derived atmospheric data. The Bayesian framework incorporates prior knowledge, quantifies uncertainty and captures spatiotemporal variability. The results show that the proposed framework captures the nonlinear effects of predictors on solar radiation and outperforms generalized linear and additive models. Significant station-level random effects highlight the importance of local environmental characteristics in multi-site modeling, suggesting architectures like graph neural networks may be advantageous. Temperature, humidity and cloud cover are the primary drivers of global horizontal irradiance, with $PM_{2.5}$ showing a notable impact under cloud-free conditions. Addressing this gap in nonlinearity is significant because capturing multi-scale dependencies is essential not only for advancing predictive modelling frameworks, but also for improving the physical interpretability of solar radiation dynamics and enabling more robust integration of solar resources into energy systems. The findings extend to similar tropical climates and the model can be adapted to diverse regions and data sources to support solar energy optimization.

1 Introduction

In the global context of the energy transition, solar energy stands out as one of the most promising solutions to address climate challenges and growing electricity demand. The number of photovoltaic installations continues to increase, particularly in West African countries which, despite their strong solar potential, still lag significantly in adopting this resource.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Among the key activities that determine the successful installation and operation of a photovoltaic farm, optimal site selection and accurate solar resource prediction play a central role. Forecasting solar radiation is an important step for optimal planning and efficient management of solar energy systems. Reliable estimation of solar radiation not only enhances energy production but also reduces costs and facilitates the integration of this renewable energy into power grids.

Various methods have been developed to predict solar radiation, ranging from physics-based methods to data-driven techniques (Paletta et al., 2023; Kumar et al., 2020; Stüber et al., 2021). Among the latter, machine learning methods and artificial neural networks have gained significant popularity due to their ability to process large datasets and capture complex relationships between variables. Artificial neural networks (ANN), support vector machines (SVMs) and random forest (RF) models have been widely employed due to their high flexibility in learning from diverse input features (Gupta et al., 2021; Voyant et al., 2017). ANN models, in particular, have demonstrated substantial effectiveness in modeling solar radiation. Moreover, SVMs and their Wavelet-coupled variants have been shown to handle small datasets with high precision, especially when forecasting daily solar radiation (Deo et al., 2016). Hybrid models that combine machine learning techniques with physical models have also been developed to integrate meteorological data with solar radiation prediction, yielding better performance than standalone machine learning models (Hedar et al., 2021). Several studies have found that incorporating environmental factors such as air pollutants and aerosol optical depth into machine learning models significantly enhances predictive performance. By considering pollutant data, such as fine particulate matter ($PM_{2.5}$) or nitrogen dioxide (NO_2) concentrations, machine learning models can more accurately capture the impacts of atmospheric pollution on solar radiation levels (Ding et al., 2024). These factors, especially in urban or industrial regions, have a marked impact on solar radiation reduction due to their effect on solar beam attenuation.

The effectiveness of these models hinges on the careful selection of input features. Identifying the most relevant meteorological and pollutant variables could substantially enhance the accuracy and reliability of predictions (Zhang et al., 2022; Zamora et al., 2005; Jeon et al., 2022; El Alani et al., 2021). The selection of appropriate explanatory variables requires a clear understanding of their impacts on solar radiation. Key factors such as temperature, cloud cover, humidity and $PM_{2.5}$ are frequently studied in this context. Recent studies have explored the relationships between predictors and solar radiation. For example, (Birdal, 2024) stated that combining air pollution data with weather conditions data had enabled the prediction of solar irradiance with higher accuracy. (Huang et al., 2021) reviewed various machine learning algorithms employed in predicting solar radiation, emphasizing the significance of selecting appropriate predictors such as sunshine duration, land surface temperature and visibility. Furthermore, the selection of input parameters must be location-specific, as an input parameter that enhances model performance in one region may be unsuitable for another (Gürel et al., 2023). Regarding air pollutants as input variables, other components could be considered, such as black carbon, which, to the best of our knowledge, has not been used as a predictor in any study. However, this pollutant could significantly influence solar radiation due to its light-absorbing properties.

Concerning the estimation of the impact of predictors on solar radiation, few studies have been conducted. One such study applied correlation analysis and a structural equation model, identifying solar zenith angle, cloud cover, aerosols and airmass as significant factors affecting solar radiation (Zhu et al., 2020). In another study, Alam et al. (2019) modeled the clearness

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

index using meteorological predictors (sunshine hours, temperature difference, cloudiness) via multiple regression (regression with several predictors). These studies are often constrained by their focus on linear relationships or their inability to explicitly account for hierarchical effects arising from spatiotemporal variability, such as site-specific differences. This underscores the need for a more advanced approach that integrates mixed-effects modeling with nonlinear relationships to effectively capture the complexity of interactions between predictors and solar radiation. Moreover, the proposed models sometimes produce deterministic predictions, presented as single-point values emerging from a black box, which the user is expected to trust without any indication of their reliability. This approach can give a false sense of precision and confidence, even in situations where the data are noisy, scarce, or fall outside the model's training domain. The lack of uncertainty quantification also limits the ability to assess the risk associated with a prediction, which is crucial not only in fields such as medicine and finance, but also in meteorology and solar energy management (Natras et al., 2023; Sacco et al., 2022; Arora and Ceferino, 2023).

The urgency of understanding the effects of meteorological and pollution parameters on solar irradiance, particularly for photovoltaic (PV) efficiency, is further justified by the lack of reliable data in certain regions, such as West Africa. This data scarcity necessitates reliance on satellite-based and reanalysis products for solar energy prediction, assessment and management. However, the challenge of accurately parameterizing these weather and aerosol variables in such products, due to limited data availability and an incomplete understanding of their regional dynamics, reduces their performance in estimations. In this regard, studies have found that satellite-based products perform worse in tropical regions than in high-latitude areas, attributing this discrepancy to the high solar irradiance and frequent cloud formation in the tropics (Peng et al., 2019; Bilal et al., 2019). Similarly, Danso et al. (2019) highlighted the frequent formation of low-level clouds in southern West Africa, which satellites struggle to capture accurately. Finally, the authors of the present study found that satellite products (CAMS and SARAH-3) and reanalysis datasets (ERA5 and MERRA-2) showed poor agreement with ground-based measurements, not only under cloudy-sky conditions but also during the Harmattan season in Côte d'Ivoire, when dust and other air pollutants reach their peak concentrations (Kakou et al., 2025).

To address the gaps mentioned above, the present study proposes the use of a Bayesian nonlinear mixed-effects model to assess the impacts of temperature, humidity, wind speed, cloud cover and $PM_{2.5}$ concentrations on solar radiation across the various climatic zones of Côte d'Ivoire. This approach fills the existing gaps in the field by enabling the estimation of predictor effects while accounting for variations in measurements across stations and time periods, as well as by quantifying the associated uncertainties. The study is based on the hypothesis that a nonlinear relationship exists between global horizontal irradiance (GHI) and the selected predictors, with inter-station variability influencing the magnitude of these effects.

2 Materials and Methods

2.1 Presentation of the Area of Study

This study investigates twelve meteorological measurement stations distributed across several climatic zones of Côte d'Ivoire (see Figure 1). The geographical coordinates of these stations are listed in Table 1. Côte d'Ivoire is located in West Africa, between latitudes 4° and 11° North and longitudes 2° and 8° West. It has a tropical climate (classified as Af, Am and As/Aw

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

in the Köppen-Geiger system (Beck et al., 2018)) characterized by two main seasons: a wet season from April to October and a dry season from November to March. The average annual temperature ranges from 24°C to 32°C, with coastal areas being cooler than the northern regions. The average annual rainfall varies significantly across the country, from about 1,200 mm in the north to over 2,000 mm in the southern coastal areas (Akobé et al., 2025).

In terms of energy, even though the country has high solar energy potential, with an average daily global horizontal irradiance (GHI) of around 5.0 kWh/m^2, Côte d'Ivoire has long relied primarily on natural gas and hydropower. As of the end of 2023, the installed capacity of the electricity generation fleet is 2,907 MW (69% thermal, 30% hydroelectric and 1% solar) (Bank, 2025). However, the country has recently begun investing in renewable energy, particularly solar power, for which it has set ambitious goals, in order to diversify its energy mix and meet growing demand (Sustainable Energy for All, 2012; Assoumou and McIsaac, 2022). Currently, government measures aim to increase the share of renewable energy in the national energy mix, aiming to reach 42% by 2030 (International Finance Corporation, 2018; Oxford Business Group, 2019).

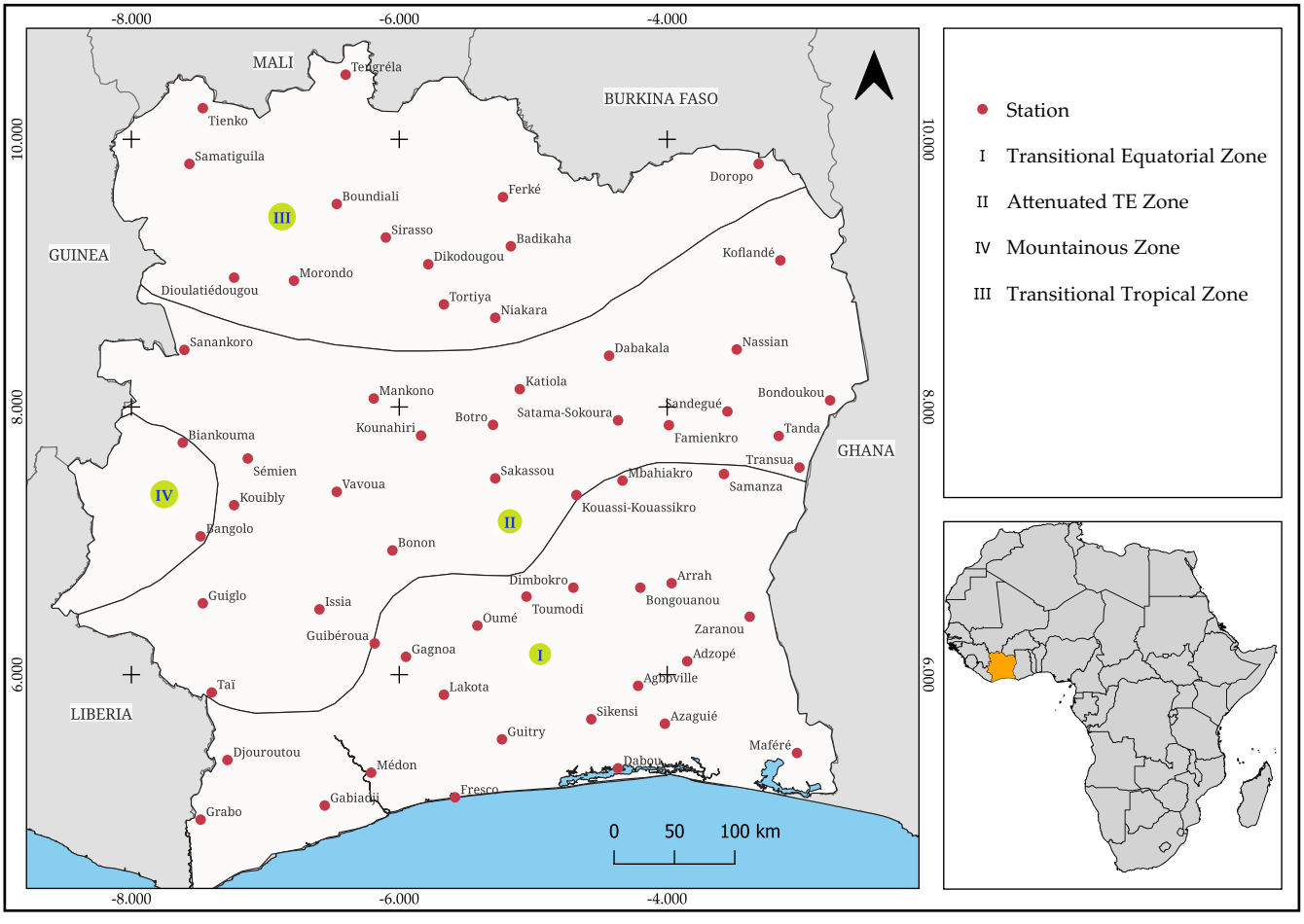


Figure 1. Map of the study area.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Like most countries in the West African subregion, Côte d'Ivoire faces relatively high levels of air pollution (Kouadio et al., 2019; Gnamien et al., 2023), particularly in urban and industrial areas, where the main sources include vehicle emissions, industrial activities, biomass burning and natural dust. The latter two are also the primary sources of air pollution in rural areas. Although several studies have addressed this issue and major initiatives such as the DACCIWA project (Gnamien et al., 2021) and the RAF7016 project of the International Atomic Energy Agency (Popouen et al., 2021) have been conducted to better understand pollution dynamics and aerosol levels in the region, much work remains to be done to achieve a comprehensive and reliable mapping of air pollution across the country.

2.2 Data Description

The objective in this study is to assess the individual and collective effects of various meteorological variables and air pollutants on global horizontal irradiance (GHI) in Côte d'Ivoire. To achieve this, GHI data, along with meteorological data including temperature, humidity and wind speed, were collected from twelve monitoring stations distributed across urban and rural areas of Côte d'Ivoire (see Table 1). These stations are managed by SODEXAM, the meteorological agency of Côte d'Ivoire. The data collection process is described in detail in (Sawadogo et al., 2023; Kakou et al., 2025).

Table 1. Spatial Distribution of Stations across Côte d'Ivoire.

Station	Latitude	Longitude	Station	Latitude	Longitude
Adzope	6.10	-3.85	Samatiguila	9.82	-7.57
Boundiali	9.52	-6.47	Tanda	7.78	-3.17
Danane	7.25	-8.15	Tengrela	10.48	-6.40
Fresco	5.08	-5.58	Toulepleu	6.57	-8.40
Kani	8.45	-6.60	Toumodi	6.58	-5.05
Katiola	8.13	-5.10	Vavoua	7.37	-6.47

The dataset covers the period from 23 July 2018 to 31 December 2020. In addition, cloud cover data and air pollution data including Dust surface mass concentration at the particle size of 2.5 micrometers or smaller ([DU]), Organic carbon surface mass concentration ([OC]), Black carbon surface mass concentration ([BC]), Sea salt surface mass concentration at the particle size of 2.5 micrometers or smaller ([SS]) and Sulfate surface mass concentration ([SO4]) were retrieved from satellite products. All these concentrations were converted from kilogram per cubic meter (kg/m^3) into microgram per cubic meter (μg/m^3) for easy handling. Then, the concentration of $PM_{2.5}$ were obtained through Equation 1 (Collow et al., 2022).

$$[PM_{2.5}] = [DU] + f_{rh,oc} \cdot [OC] + f_{rh,bc} \cdot [BC] + f_{rh,ss} \cdot [SS] + 1.3756 \cdot f_{rh,so4} \cdot [SO4], \quad (1)$$

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

where $f_{rh,species}$ is the hygroscopic growth factor of the species, which quantifies how much the aerosol particle radius grows when exposed to a certain relative humidity (RH) due to water uptake.

$$f_{rh,species} = 1 + (((\frac{R_{Wet\,RH}}{R_{Dry\,RH}})^3 - 1) \cdot \frac{\rho_{Water}}{\rho_{Dry\,species}}), \quad (2)$$

$R_{Wet\,RH}$ and $R_{Dry\,RH}$ are the radii at wet and dry relative humidity, respectively. ρ_{Water} is the water density. $\rho_{Dry\,species}$ is the dry density for the species. The multiplier 1.3756 accounts for the conversion of sulfate ion mass (SO_4^{2-}) to the mass of ammonium sulfate $(NH_4)_2SO_4$, which is the form of sulfate considered in $PM_{2.5}$ mass calculations. The mass of organic matter included in this formula is represented by [OC], which is multiplied by the OM/OC ratio ($f_{OM:OC}$). This ratio depends on location, season and time (Font et al., 2024; Rajput, 2018) and typically takes values of 1.6 ± 0.2, 2.1 ± 0.2 and up to 2.6 for urban aerosols, non-urban aerosols and biomass burning aerosols, respectively (Turpin and and, 2001; Navinya et al., 2020). The distribution of particulate pollution in Côte d'Ivoire is location-specific, but is largely dominated by biomass burning, contributing approximately 40% to the organic carbon concentration in Abidjan (southern region) and 70.7% in Korhogo (northern region) (Gnamien et al., 2023). Accordingly, a value of $f_{OM:OC} = 2.3$ was used in this study. For the calculation of hygroscopic growth factors, the aerosol species densities used in this thesis are those provided by Collow et al. (2023). These are $1800\,kg/m^3$ for carbonaceous aerosols, $2200\,kg/m^3$ for sea salt and $1700\,kg/m^3$ for sulfate. The effective radii were obtained from the GMAO website (NASA Global Modeling and Assimilation Office, 2024), which is part of NASA's Integrated Empirical Science Architecture (IESA) data-sharing portal and provides access to datasets related to aerosol optical properties. The files `optics_BC.v2_3.nc`, `optics_OC.v2_3.nc`, `optics_SS.v5_3.nc` and `optics_SU.v4.nc` contain effective radius (rEff) values for different relative humidity (RH) levels, ranging from 0.0 to 0.9 in steps of 0.1 or from 0.00 to 0.99 in steps of 0.01. When a specific RH value was not available, the closest available value within the range was used. The effective radius at dry relative humidity corresponds to the value at RH = 0.

Fractional cloud cover (CFC) data were sourced from COMET2 product provided by EUMETSAT's Satellite Application Facility on Climate Monitoring (CM SAF). The dataset comprises hourly data from 1 January 2015 to 31 December 2020, derived from MVIRI-SEVIRI instruments aboard METEOSAT (Version 002), with a latitude-longitude grid resolution of $0.05° \times 0.05°$ and covering the METEOSAT full disk (Europe, Africa and the Atlantic Ocean) (Stöckli et al., 2024; Karlsson et al., 2023). Pollution data were obtained from the M2T1NXAER dataset, a two-dimensional, hourly time-averaged data collection within the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2). MERRA-2 is NASA's latest global atmospheric reanalysis for the satellite era, produced by the Global Modeling and Assimilation Office (GMAO) using the Goddard Earth Observing System Model, Version 5.12.4 (GEOS 5.12.4). The dataset spans from 1980 to the present and is updated approximately three weeks after the end of each month. Each data field is time-stamped with the central time of the corresponding hour, beginning at 00:30 UTC (Global Modeling and Assimilation Office (GMAO), 2015).

These satellite-based data were processed to achieve spatio-temporal alignment with ground measurements. To prevent the natural attenuation of solar radiation caused by the diurnal cycle from biasing results, the model was built based on data collected between 10:00 AM (UTC) and 3:00 PM (UTC), which is the period when GHI typically reaches its highest values in Côte d'Ivoire. Furthermore, since temperatures in Côte d'Ivoire are generally considered reasonable within the range of $10°C$ to

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

45°C (Akobé et al., 2025; Yapo et al., 2019), all values outside this range were excluded from the dataset to ensure consistency. After applying these initial filters, we addressed potential anomalies by removing near outliers using the interquartile range (IQR) method. Specifically, the lower and upper bounds for outliers were defined as shown in Equations 3.

$$L_{Bound} = Q_1 - 1.5 \times IQR, \quad U_{Bound} = Q_3 + 1.5 \times IQR, \tag{3}$$

where Q_1 is the 25^{th} percentile, Q_3 is the 75^{th} percentile and IQR $= Q_3 - Q_1$ represents the interquartile range. Data points falling outside these boundaries were identified as outliers and removed from the dataset.

2.3 Model Specification

In this study, a Bayesian nonlinear mixed-effects (BNLME) model was used to estimate the influence of meteorological and air pollution variables on GLobal Horizontal Irradiance. The choice of this particular model was motivated by the limitations observed in other candidates. First, a linear regression model might have sufficed; however, such models are inadequate for capturing the nonlinear relationships between predictors (temperature, cloud cover, air pollutants) and reponse (GHI), although they may be useful for simple exploratory analyses or approximately linear relationships. Nonlinear regression models can address nonlinearity but fail to account for random effects and group-level dependencies. Another option could be machine learning techniques, including random forests, gradient boosting and neural networks. However, these methods, dedicated to pure predictive tasks, often lack interpretability, struggle with uncertainty quantification and cannot incorporate prior knowledge. These limitations of alternative models are summarized in Table 2.

Table 2. Limitations of alternative models for estimating nonlinear mixed effects.

Model	Limitation
LR	Fails to capture nonlinear relationships.
NLR	Lacks the capability to handle random effects.
ML	Lacks interpretability for purely predictive tasks.
Dynamic	Lacks flexibility for nonlinear effect variability.
ST	Does not focus on interactions between covariates.
FNLME	Cannot directly include prior information.

Nonlinear mixed effects model, generally defined by Equation 4, provide an effective framework for analyzing scenarios in which the temporal evolution of individual responses is hypothesized to conform to specific nonlinear functions grounded in scientific theory. This approach is employed in applications where individual responses display systematic nonlinear behavior governed by a finite set of underlying parameters (Lindstrom and Bates, 1990; Davidian and Giltinan, 2003).

$$\begin{aligned} y_{ij} &= f(\phi_i, x_{ij}) + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \\ \phi_i &= A_i\beta + B_i b_i, \quad b_i \sim \mathcal{N}(0, \tau^2) \end{aligned} \tag{4}$$

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

where y_{ij} is the response variable for observation j in group i, $f(\cdot, x)$ is a nonlinear function of predictor verctor (x_{ij}). ε_{ij} represents an intra-group noise term. A_i and B_i are design matrices associated to fixed effects and random effects, respectively. β are the fixed effect coefficients, b_i denotes the random effect coefficients for group i.

The response variable is modeled as a nonlinear function of fixed effects, which represent population-level parameters and random effects, which capture individual-specific deviations. The most often used approaches estimate these parameters by maximizing the likelihood function, often employing numerical optimization techniques due to the complexity introduced by nonlinearity and random effects. However, these approaches rely on point estimates and lack robustness in handling parameter uncertainty.

Unlike these alternatives, Bayesian approaches offer a better uncertainty quantification through posteriors distributions, using prior knowledge. In the context of estimating the influence of meteorological and pollution predictors (temperature (T), humidity (H), wind speed (W), cloud cover (C) and $PM_{2.5}$ concentrations (P)) on GHI, Equation 4 translates, for a given observation y_{it} in station i at time t, into the form stated in Equation 5.

$$\begin{aligned} y_{it} &\sim \mathcal{N}(\mu_{it}, \sigma_{it}^2) \\ \mu_{it} &= f(T_{it}, H_{it}, W_{it}, C_{it}, P_{it}, t, \Theta, \alpha_i), \end{aligned} \tag{5}$$

Θ is the vector of global parameters, α_i represents the random effect for station i and $f(\cdot)$ is a combination of nonlinear functions (sigmoids, quadratics and seasonal harmonics), each representing the effect of a predictor on GHI.

Some structural issues with GHI make direct regression less reliable. For example, GHI(t) depends heavily on the solar zenith angle, which introduces non-stationarity and makes it difficult to compare values across different times or locations—even under perfectly clear skies (Stein et al., 2012). As a result, many studies choose to model the clear-sky index k_{it} instead. This index (defined in Equation 6) represents the fraction of potential clear-sky irradiance that actually reaches the ground (Amaro e Silva et al., 2024; Lauret et al., 2022). Its values generally fall between 0 and 1, which makes it more suitable for Beta regression.

$$k_{it} = \frac{y_{it}}{I_{it}^{\text{cs}}} \in (0, 1) \tag{6}$$

where I_{it}^{cs} is the theoretical clear-sky GHI at station i and time t, calculated using models such as Ineichen and Perez (Perez et al., 2002) or McClear clear-sky model (Lefèvre et al., 2013). Although the values of k are theoretically bounded between 0 and 1, values greater than 1 can be observed in practice. To enable the use of a Beta distribution, the k values were constrained between 0.01 and 0.99. Built upon these considerations, Equation 5 is reformulated as shown in Equation 7.

$$\begin{aligned} k_{it} \mid \mu_{it}, \phi &\sim \text{Beta}[\mu_{it}\phi, (1-\mu_{it})\phi] \\ \text{logit}(\mu_{it}) &= f(T_{it}, H_{it}, W_{it}, C_{it}, P_{it}, t, \Theta, \alpha_i), \end{aligned} \tag{7}$$

where ϕ is the precision parameter of the Beta distribution. The logit link function is used to map the mean response μ_{it} from the (0, 1) interval to the entire real line, facilitating the regression analysis.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

To determine the shape of each component of $f(\cdot)$, Physics knowledge, existing literature and descriptive statistics were used. For each covariate, scatter plots were drawn between the clear-sky index k and the covariate in order to identify potential nonlinear relationships. In addition to the scatter plots, several regression curves were fitted to the data to determine the most appropriate functional form. These included linear, quadratic, exponential, logarithmic, sigmoid and locally weighted scatterplot smoothing (LOWESS) regression curves (Cleveland, 1979; Cleveland and Devlin, 1988). Although the latter has the advantage of not requiring the specification of a function to fit a model across the entire dataset, it does not yield a regression function that can be easily expressed using a mathematical formula (NIST/SEMATECH, 2022). Therefore, it was used solely as a reference for evaluating the other regression curves. For each of them, the Pearson correlation coefficient R was calculated to quantify the quality of fit. The curve with the highest R value was retained to represent the effect of the covariate on k and by extension on GHI. The graphs in Figure 2 illustrate this methodology. Additionally, seasonal patterns were incorporated using harmonic functions to account for periodic variations in solar radiation throughout the year. Based on these analyses, the final model structure was defined as shown in Equation 8 (Ospina and Ferrari, 2010; Congdon, 2014).

$$
\begin{aligned}
k_{it} \mid \mu_{it}, \phi &\sim \text{Beta}(\mu_{it}\phi, (1-\mu_{it})\phi) \\
\text{logit}(\mu_{it}) &= \alpha_0 + b_i + \beta_T\, \ell_T(T_{it};\gamma_T,\delta_T) + \beta_H\, \ell_H(H_{it};\gamma_H,\delta_H) \\
&\quad + \beta_W\, \ell_W(W_{it};\gamma_W,\delta_W) + \beta_C\, \ell_C(C_{it};\gamma_C,\delta_C) \\
&\quad + \theta_{P1} P_{it} + \theta_{P2} P_{it}^2 + s_{\text{year}}(d_t)
\end{aligned} \tag{8}
$$

The nonlinear functions ℓ_X for $X \in \{T, H, W, C\}$ are:

$$\ell_X(x;\gamma_X,\delta_X) = (1+\exp\{-\gamma_X(x-\delta_X)\})^{-1}$$

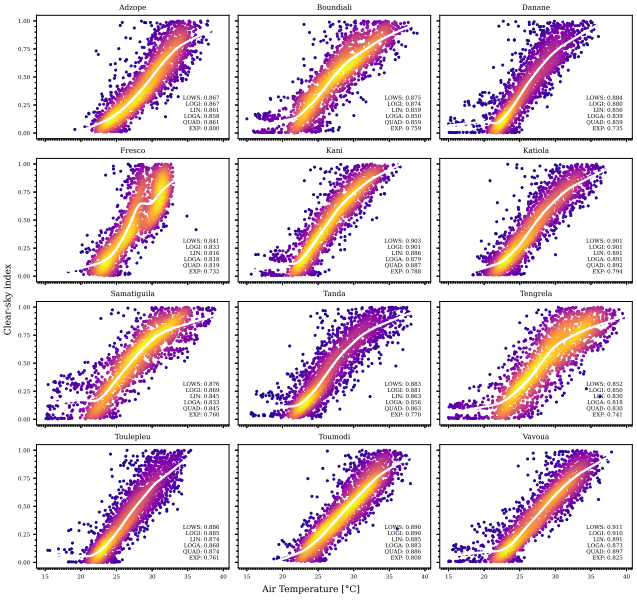
The seasonal Fourier harmonics (Hridoy, 2025; Bazrafshan et al., 2025):

$$s_{\text{year}}(d_t) = \sum_{m=1}^{H} \left[a_m \sin\left(\frac{2\pi m\, d_t}{365.24}\right) + b_m \cos\left(\frac{2\pi m\, d_t}{365.24}\right) \right], \quad H = 1,2,3$$

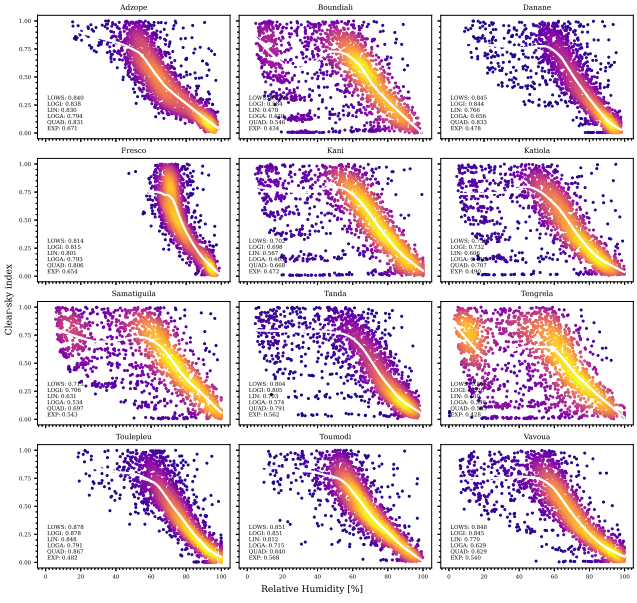
Random effect:

$$b_i \sim \mathcal{N}(0, \sigma_b^2)$$

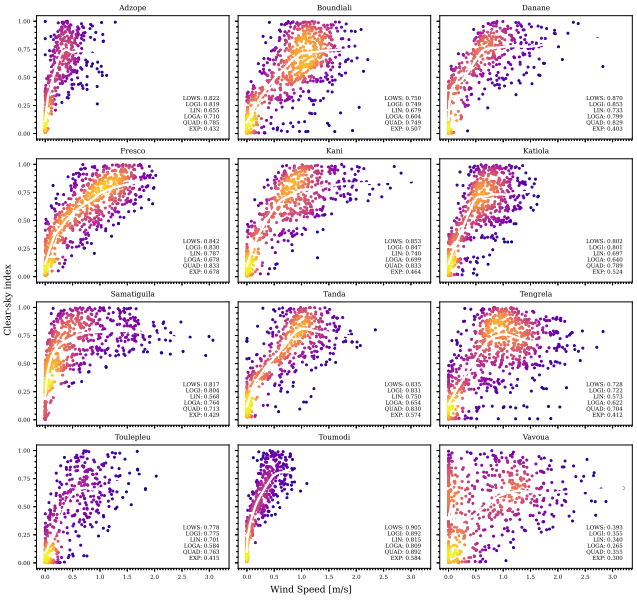
https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026



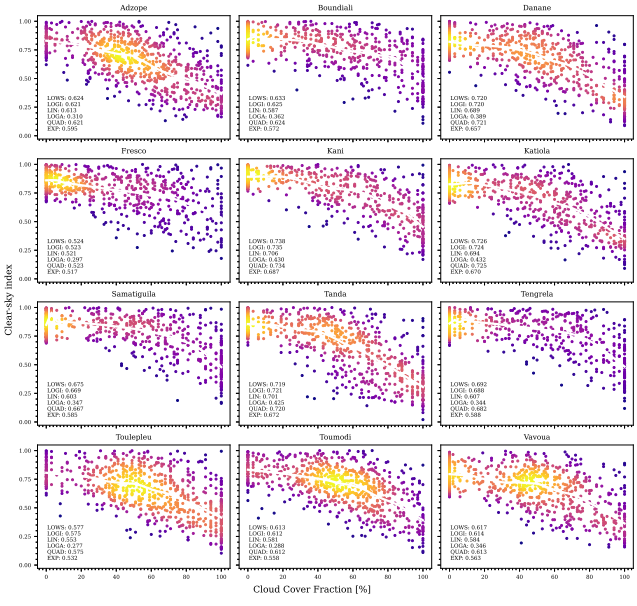
(a) k vs Temperature



(b) k vs Humidity

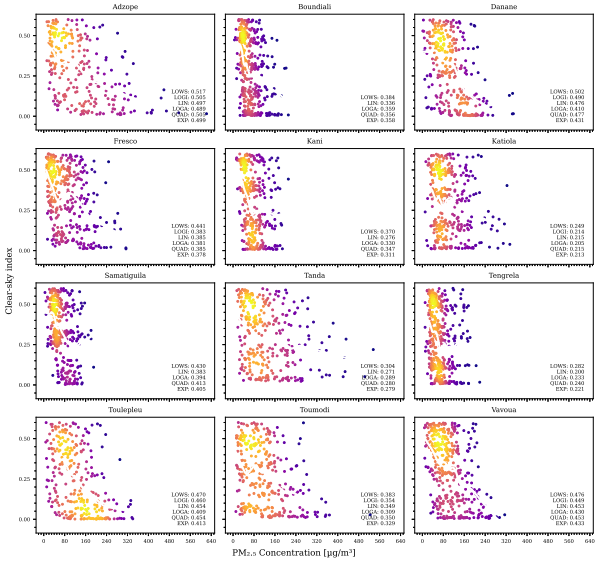


(c) k vs Wind Speed



(d) k vs Cloud Cover

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026



(e) k vs $PM_{2.5}$ Concentrations

Figure 2. Scatter plots and regression curves between the clear-sky index and the covariates. Only the LOWESS curves are shown for clarity. For the other regression curves, Pearson correlation coefficients (R) are displayed. For temperature (a) and relative humidity (b), the data considered are from 7 AM to 5 PM. For cloud cover (d), only the data between 12 PM and 2 PM revealed a significant relationship. For wind speed (c), clear-sky data based solely on cloud cover ($CF \leq 30\%$) were considered. Finally, for $PM_{2.5}$ concentrations (e), data from the Cloudy-Clear scenario (KtC-FCl) were used (see Section 2.4). The following abbreviations represent the Pearson correlation coefficients: LIN (linear), QUAD (quadratic), EXP (exponential), LOGA (logarithmic), LOGI (sigmoid) and LOWS (locally weighted scatterplot smoothing).

The BNLME model was implemented with version 5.25.1 of Python's `PyMC` package (Abril-Pla et al., 2023). First, five variants were constructed, as specified in Table 3. The first model (BRI) is a baseline model for comparison, which includes only random intercepts and no covariates. The second model (BFE) adds a basic Beta regression and includes raw covariates. Building on the previous models, the third model (FES) incorporates seasonality to assess its contribution. The fourth model (SNL) introduces non-linear covariate effects using sigmoid functions. Finally, the fifth model (NLH3) explores richer seasonality by incorporating additional harmonics ($H = 3$).

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Table 3. BNLME model scenarios of and the different levels added at each stage.

Model	Rnd Eff.	Lin. Cov.	NL Cov.	Season.
BRI	Yes	No	No	No
BFE	Yes	Yes	No	No
FES	Yes	Yes	No	Yes
SNL	Yes	Yes	Yes	Yes
NLH3	Yes	Yes	Yes	Yes

To validate the model, the best variant was selected by comparing leave-one-out cross-validation (LOO-CV) scores using the expected log pointwise predictive density (ELPD) (Vehtari et al., 2017; Spiegelhalter et al., 2002; The ArviZ Developers, 2025). The selected scenario was subsequently compared with other modeling approaches, including the Generalized Linear Model (GLM) and the Generalized Additive Model (GAM). For this comparative evaluation, performance metrics including the coefficient of determination (R^2), the mean absolute difference (MAD), the root mean squared difference (RMSD) and the combined performance index (CPI) were used. They are defined in Equations 9, 10, 11 and 12, respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{9}$$

$$\text{MAD} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{10}$$

$$\text{RMSD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{11}$$

$$\text{CPI}[\%] = \frac{1}{4} \times (\text{KSI} + \text{OVER} + 2 \times \text{rRMSD}) \tag{12}$$

where y_i is the observed value, $\hat{y}_i$ is the predicted value, $\bar{y}$ is the mean of observed values and n is the number of observations. KSI (Kolmogorov-Smirnov Index), OVER metric and rRMSD (relative root mean squared difference) are defined in Kakou et al. (2025); Espinar et al. (2009); Gueymard (2014).

The overall methodological workflow is summarized in Figure 3.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

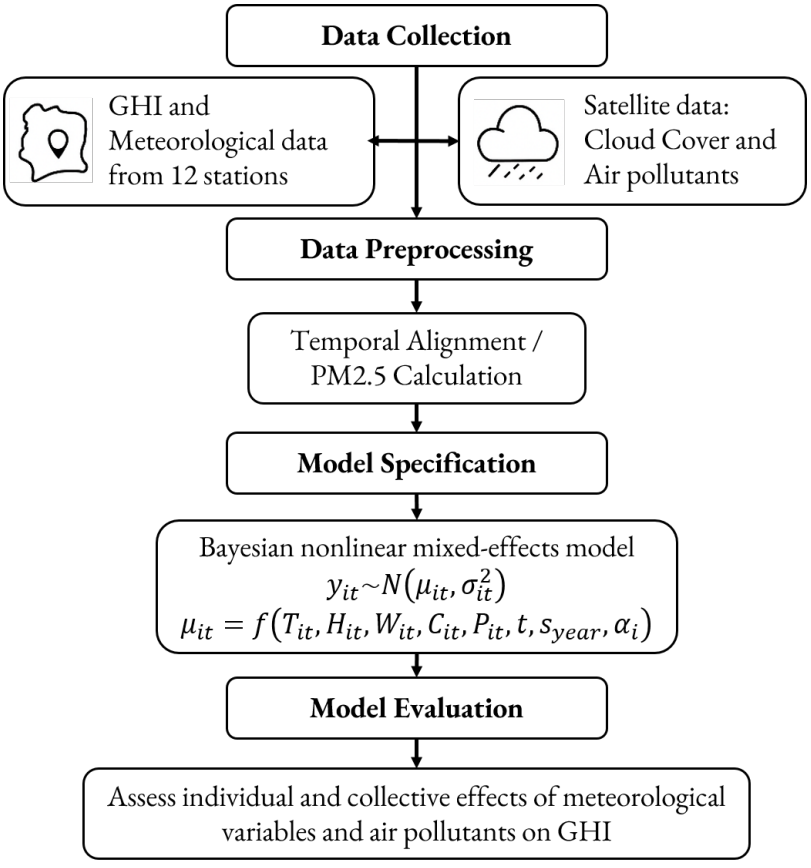


Figure 3. Workflow of the methodology.

2.4 Sky conditions

Cloud cover and the clearness index (K_t) defined as the ratio of the measured ground-level GHI to the extraterrestrial GHI, are two parameters commonly used in the literature to characterize sky conditions. Various definitions rely on the cloud cover fraction, establishing thresholds to categorize Clear, Partly Cloudy/Partly Sunny, Mostly Cloudy and Cloudy skies Correa et al. (2022); National Weather Service (n.d.). For instance, a study on the clear-sky bias of satellite land surface temperature estimates considered a grid point as clear if the cloud fraction was below 30% and cloudy if above 70% Ermida et al. (2019). Similarly, for definitions based on the clearness index, thresholds also vary from one study to another. In Tham et al. (2009), for example, the authors indicated that a clearness index exceeding 0.75 is indicative of clear-sky conditions. In this study, we adopted a cloud cover fraction below 30% and a clearness index above 0.6 Kakou et al. (2025) as both indicating clear-sky conditions. We defined four sky conditions: Cloudy-Cloudy (KtC-FC), where both parameters predict a cloudy sky, Cloudy-Clear (KtC-FCl), where K_t predicts a cloudy sky, but the cloud cover fraction predicts a clear sky, Clear-Cloudy (KtCl-FC), where K_t predicts a clear sky, but the CFC predicts a cloudy sky and Clear-Clear (KtCl-FCl), where both parameters predict a clear sky.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

3 Results and Discussion

3.1 Descriptive Statistics and Exploratory Analysis

Descriptive statistics of the variables considered in this study are summarized in Table 4. For clarity, only the mean and standard deviation (std) values are included in the table. Mean temperature ranges from $25.81 \pm 3.54\,^{\circ}$C in Danané to $28.62 \pm 5.31\,^{\circ}$C in Tengrela. The minimum and maximum temperatures ($10.00\,^{\circ}$C and $41.04\,^{\circ}$C) were recorded in Samatiguila and Tengrela, respectively. Highest variability this variale is observed at Boundiali, Tengrela and Samatiguila.

Mean wind speed varies from 0.26 ± 0.23 m/s to 0.90 ± 0.61 m/s. Minimum values are consistently 0 m/s across all stations, while maximum values range from 0.97 m/s (Toumodi) to 2.83 m/s (Danané). Relatively low values were observed throughout the year, particularly in Vavoua and Tanda. Similar values were found in Abidjan, Korhogo, Bouaké, Yamoussoukro and Man by Kone et al. (2024); Kouadio et al. (2024).

Relative humidity is generally high in Côte d'Ivoire, with mean values ranging from $55.40 \pm 26.74\%$ to $80.63 \pm 9.15\%$. Northern locations such as Tengrela, Boundiali, Samatiguila, Kani and Katiola are the driest, with minimum relative humidity values of 0.34%, 4.44%, 5.35%, 0.36% and 6.74%, respectively. Maximum relative humidity values reach 100% at most stations, indicating saturation and the presence of high specific humidity levels.

Cloud cover fractions exhibit high variability, with minimum and maximum values of 0% and 100% observed across all stations. Mean cloudiness is lower in northern stations such as Tengrela ($44.62 \pm 36.30\%$) and Boundiali ($47.72 \pm 35.75\%$), indicating generally clearer skies in these regions.

Fine particulate matter ($PM_{2.5}$) concentrations are higher than the World Health Organization (WHO) guidelines for annual mean values—$10\,\mu g/m^3$ (2005) and $5\,\mu g/m^3$ (2021) (World Health Organization, 2021). Mean $PM_{2.5}$ values range from $16.64 \pm 10.90\,\mu g/m^3$ to $38.32 \pm 31.28\,\mu g/m^3$, indicating substantial variability. Comparable, though slightly lower, values were found by Bahino et al. (2024), likely due to the overestimation tendency of satellite or reanalysis data compared to in-situ observations.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Table 4. Diurnal Mean and Standard deviation of Temperature (T [°C]), Wind Speed (W [m/s]), Humidity (H [%]), GHI (y [kW/m^2]), Cloud Cover Fraction (C [%]) and $PM_{2.5}$ Concentration P[μg/m^3] across stations.

Station	T		W		H		y		C		P	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
Adzope	28.00	3.68	0.47	0.36	74.03	15.23	0.26	0.21	62.37	27.35	22.14	16.71
Boundiali	27.90	4.88	0.89	0.54	59.17	24.58	0.38	0.27	47.72	35.75	36.70	21.97
Danane	25.81	3.54	0.86	0.66	77.23	14.95	0.30	0.24	63.12	30.22	21.39	18.59
Fresco	26.75	2.75	0.90	0.61	80.63	9.15	0.35	0.25	54.46	32.33	16.64	10.90
Kani	27.42	4.48	0.71	0.47	67.13	22.86	0.39	0.28	52.45	35.32	31.31	22.93
Katiola	26.96	4.01	0.53	0.33	69.32	20.31	0.31	0.23	57.12	32.87	30.53	22.62
Samatiguila	27.33	5.55	0.35	0.35	66.19	25.31	0.39	0.28	48.62	36.84	31.32	19.69
Tanda	26.95	3.98	0.47	0.41	74.16	17.53	0.30	0.23	58.69	31.93	30.78	21.43
Tengrela	28.62	5.31	0.71	0.60	55.40	26.74	0.39	0.28	44.62	36.30	37.96	22.03
Toulepleu	26.96	3.93	0.43	0.34	76.66	17.31	0.31	0.24	56.14	30.67	38.32	31.28
Toumodi	28.22	3.94	0.26	0.23	72.67	16.15	0.31	0.23	59.90	28.69	25.20	20.15
Vavoua	27.26	4.22	0.60	0.53	73.00	17.62	0.30	0.22	60.29	30.19	22.66	18.56

Regarding the response variable, maximum values peak at $1.08\,\mathrm{kW/m^2}$ in Boundiali and the minimum remains low across all sites $0.01\,\mathrm{kW/m^2}$. Mean GHI varies between 0.26 ± 0.21 and $0.39 \pm 0.28\,\mathrm{kW/m^2}$.

3.2 Model Validation

First, several scenarios of the BNLME model were compared using LOO-CV scores to select the best, as shown in Table 5 and Figure 4. The `elpd.loo` column indicates the estimated log predictive density obtained through LOO-CV. A higher value means better out-of-sample predictive performance. Among the evaluated scenarios, the NLH3 model exhibited the best out-of-sample predictive performance, with the highest `elpd.loo` (9552.44) and a model weight of 1.0. Although slightly more complex (`p.loo` = 34.76), it substantially outperformed all alternatives. For instance, the next-best model (SNL) had an `elpd.loo` 381 points lower, with a negligible model weight, suggesting that NLH3 is strongly favored. NLH3 incorpores richer seasonal structure through the inclusion of three harmonics ($H = 3$), which likely enabled it to capture more complex temporal patterns in the data. The addition of seasonality in the FES scenario improves performance compared to the basic Beta regression (BFE). Introducing non-linear effects in the SNL scenario brings further improvement, but it is the inclusion of richer seasonal components in the NLH3 scenario that leads to the best overall performance. In contrast, the baseline model BRI, which includes only random effects and no covariates, shows the poorest performance, which evidence the influence of the predictors considered in this study, as well as the seasonal structure, on the variability of GHI.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Table 5. Comparison of scenarios. The best-performing scenario is italicized.

Scenario	elpd.loo	p.loo	elpd.diff	weight	se	dse
NLH3	*9552.44*	*34.76*	*0*	*1.0*	*97.92*	*0*
SNL	9171.27	30.02	381.16	0.0	94.31	28.40
FES	8673.12	25.33	879.32	0.0	90.63	39.74
BFE	8362.55	23.68	1189.89	0.0	88.63	48.59
BRI	2389.60	13.82	7162.84	0.0	55.19	105.24

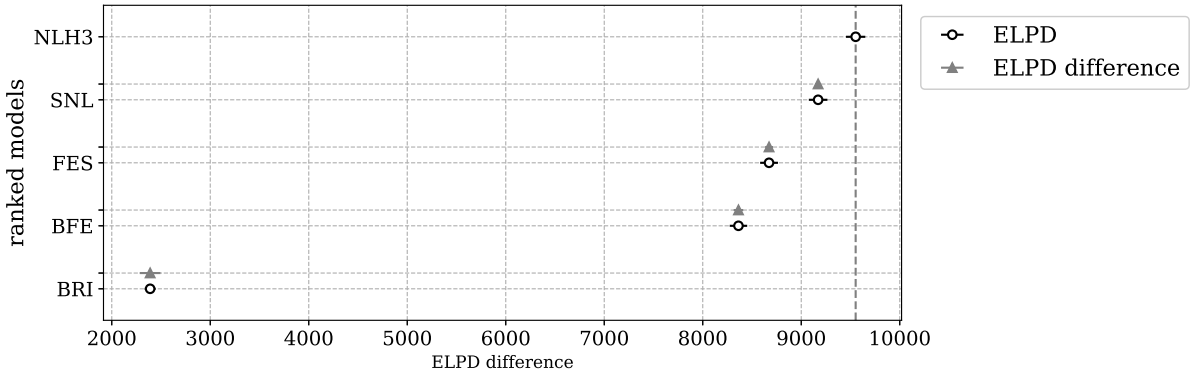


Figure 4. Scenario Ranking based on ELPD Difference (high is better).

To ensure that the model aligns well with the data, generalizes appropriately and produces reliable inferences, a validation based on the posterior distribution of the parameters was carried out for the best-performing scenario (Figure 5). The MCMC diagnostics indicate that the model achieved good convergence. Trace plots show well-mixed, stationary chains with no visible drift and strong overlap across parameters, suggesting effective sampling and stability in the posterior draws. The posterior marginal densities are unimodal and concentrated (left column of Figure 5). The model intercept is centered around -1.0 on the logit scale, implying a moderate baseline mean on the response scale where the model estimates an average k_{it} of approximately 0.27 when all predictors are set to their reference values and random effects are zero. This result is consistent with several studies that have shown clear-sky conditions are rare in Côte d'Ivoire (Danso et al., 2019; van der Linden et al., 2015; Moron et al., 2023; Kakou et al., 2025).

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

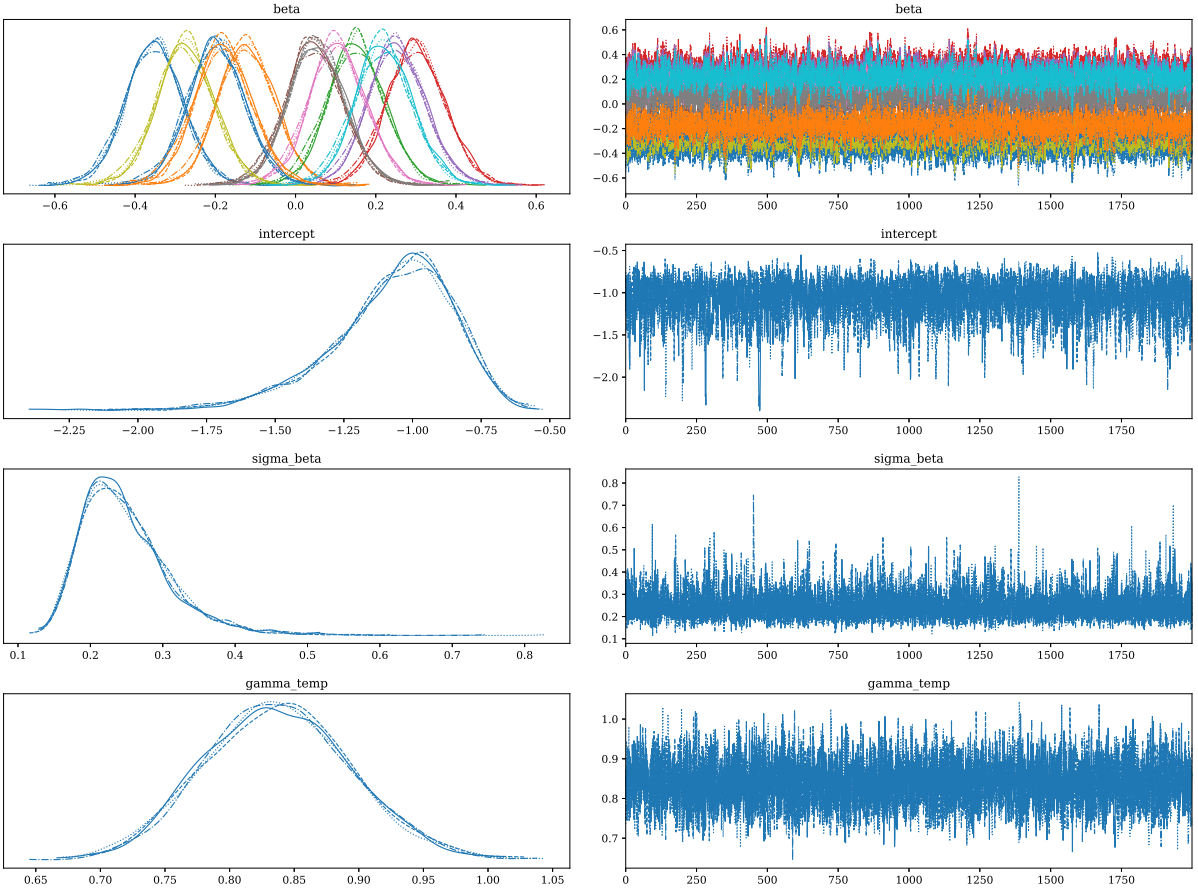


Figure 5. Posterior density and trace plots for selected parameters of the NLH3 model. Each row corresponds to one parameter. The left panels show the posterior marginal densities and the right panels show the traceplots for all MCMC chains. The well-mixed, stationary traces and overlapping densities indicate good convergence and stable posterior sampling for all parameters.

The station-level random effects b_i range between -0.4 and $+0.3$ on the logit scale, indicating a spatial heterogeneity among stations. On the mean scale of μ_{it}, this range corresponds to a difference of approximately 0.13 in the average clear-sky index between the cloudiest and clearest stations. Negative values ($b_i < 0$) characterize stations with lower mean clear-sky conditions, possibly due to higher cloud frequency, dust load and aerosol, whereas positive values ($b_i > 0$) correspond to brighter and more stable atmospheric conditions. The relatively small amplitude of these random effects (with a standard deviation σ_b of approximately 0.25) suggests that the stations in Côte d'Ivoire share a common underlying atmospheric pattern, with only moderate local variations in irradiance.

These results, along with the $\hat{R}$ values between 1.00 and 1.01 for all parameters (see Table 7), indicate that the model outputs are reliable and suitable for interpretation, subject to posterior predictive validation.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Next, NLH3 (the best-performing scenario) was compared to two other models: a generalized linear model (GLM) using a binomial family with a logit link function and a linear additive model (GAM) using a smoothing spline for each predictor. The GAM was implemented using the `LinearGAM` class from the `pyGAM` Python module (Servén and Brummitt, 2018), while the GLM was fitted using the `glm` function from `statsmodels.formula.api` package (Seabold and Perktold, 2010). Both models included the same covariates and random effects as NLH3.

Comparing Bayesian and non-Bayesian models presents a methodological challenge, as they rely on fundamentally different statistical paradigms. For example, in terms of parameters, Bayesian models treat them as random variables with prior distributions, whereas frequentist models treat them as fixed but unknown quantities. Regarding uncertainty, the Bayesian paradigm expresses it through the posterior distribution of parameters, while the frequentist approach represents 'uncertainty' using confidence intervals or standard errors. Nevertheless, several rigorous approaches exist to compare models across paradigms, particularly in terms of predictive performance or model complexity (Hespanhol et al., 2019; Flor et al., 2020; Seide et al., 2020). In this study, the coefficient of determination (R^2), the Mean Absolute Difference (MAD), Root Mean Squared Difference (RMSD) and Combined Performance Index (CPI) were also used to compare the predictive performance of the three models. The comparison results are summarized in Table 6.

Table 6. Model comparison. NLH3 outperformed GAM and GLM in terms of R^2, MAD, RMSD and CPI, indicating superior predictive performance and model fit.

Model	R^2	MAD	RMSD	CPI
GAM	0.674	0.095	0.014	0.075
GLM	0.663	0.097	0.015	0.075
NLH3	0.746	0.082	0.011	0.065

NLH3 outperformed both GAM and GLM models across all evaluation metrics. This superior performance can largely be attributed to the fact that NLH3 employs the correct bounded Beta likelihood. GLMs that assume Gaussian distributions or rely on naive transformations may suffer from misspecification. On the other hand, GAMs based on splines do not enforce the characteristic saturating shape imposed by the logistic form. Moreover, NLH3 incorporates physically meaningful, parametric nonlinear functions (sigmoids) that naturally capture threshold effects and saturation. Additionally, NLH3 benefits from hierarchical pooling through station-specific random effects and seasonal variation is captured using harmonic terms.

3.3 Influence of Weather and Pollution Variables

We now examine the influence of the predictors on the clear-sky index and, by extension, on the global horizontal irradiance (GHI). To do so, we analyze the results of the NLH3 model implementation, as reported in the posterior summaries in Table 7. The parameter estimates are provided on the logit scale, along with their standard deviations (SD), 94% highest density intervals (HDI), Monte Carlo standard errors (MCSE), effective sample sizes (ESS) for both the bulk and tail of the posterior and the potential scale reduction factor ($\hat{R}$).

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Table 7. Posterior summaries of NLH3 model for clear-sky index.

Parameter	Mean	SD	HDI 3%	HDI 97%	MCSE (Mean)	MCSE (SD)	ESS (Bulk)	ESS (Tail)	$\hat{R}$
alpha0	-1.075	0.229	-1.529	-0.712	0.004	0.005	3162	3368	1
beta_T	4.279	0.360	3.634	4.959	0.005	0.004	5572	4720	1
beta_H	-1.646	0.094	-1.818	-1.467	0.001	0.001	6334	4981	1
beta_W	0.908	0.218	0.564	1.308	0.004	0.005	3832	3590	1
beta_C1	-0.150	0.007	-0.164	-0.136	0	0	10431	6267	1
beta_C2	-0.052	0.007	-0.066	-0.039	0	0	10418	6265	1
beta_P1	0.010	0.011	-0.012	0.031	0	0	7267	5855	1
beta_P2	0.009	0.003	0.003	0.016	0	0	7998	6198	1
delta_T	1.337	0.160	1.038	1.626	0.002	0.002	5853	4602	1
delta_H	1.256	0.035	1.191	1.320	0	0	6743	4818	1
delta_W	-1.285	0.323	-1.899	-0.755	0.006	0.005	3824	3482	1
a1	-0.095	0.032	-0.152	-0.034	0	0	6512	5513	1
b1	-0.207	0.012	-0.230	-0.184	0	0	8634	6283	1
a2	-0.070	0.011	-0.091	-0.051	0	0	9519	6478	1
b2	0.183	0.024	0.139	0.228	0	0	6392	5322	1
a3	-0.139	0.014	-0.166	-0.113	0	0	7102	5781	1
b3	0.160	0.010	0.142	0.179	0	0	9397	6670	1
sigma_b	0.248	0.062	0.147	0.358	0.001	0.001	7045	4676	1
gamma_T	0.837	0.057	0.735	0.946	0.001	0.001	5786	5352	1
gamma_H	3.684	0.222	3.280	4.113	0.003	0.002	6871	5331	1
gamma_W	1.311	0.215	0.920	1.709	0.003	0.002	4461	4719	1
phi	14.010	0.174	13.694	14.343	0.002	0.002	11744	5682	1

All $\hat{R}$ values lie between 1.00 and 1.01, which shows that the MCMC chains have converged well. It is also worth noting that all ESS values exceed 3,000 (ranging from 3,162 to 11,744), demonstrating high sampling efficiency and a large number of effectively independent posterior draws. Moreover, MCSE values are near zero (between 0 and 0.006), so Monte Carlo uncertainty in posterior means and SDs is negligible. These MCMC diagnostics collectively confirm a well-behaved posterior and support the reliability of the resulting inferences.

The relationships between the clear-sky index and the predictors (temperature, humidity, wind speed and cloud cover) were modeled using sigmoid functions denoted by $\ell_X(x;\gamma_X,\delta_X)$ (see Section 2.3). Since this function is monotonic in x when $\gamma_X > 0$ and the logit link is also monotonic, the sign of β_X determines the direction of the effect on the expected clearness index. Specifically, $\beta_X > 0$ implies an increase in μ_{it} as X increases, while $\beta_X < 0$ implies the opposite. Thus, an increase

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

in temperature ($\beta_T = 4.279$, 94% HDI: [3.634, 4.959]) and a decrease in relative humidity ($\beta_H = -1.646$, 94% HDI: [-1.818, -1.467]) are associated with a higher clear-sky index, which aligns with atmospheric physics. Higher temperatures promote cloud dissipation and drier air is generally linked to clearer conditions. Similarly, an increase in wind speed ($\beta_W = 0.908$, 94% HDI: [0.564, 1.308]) is associated with a higher clear-sky index, likely due to the dispersion of particles and moisture in the air. However, the literature presents mixed findings regarding the effect of wind speed on solar radiation. For instance, Bett and Thornton (2016) found a negative, albeit modest, correlation ranging from -0.4 to -0.2 in Britain, which they attributed to the tendency for windy days to be cloudier, particularly in Atlantic-influenced regions. These conditions are associated with higher wind speeds (Kazil et al., 2016), whereas the average in Côte d'Ivoire is very low (Kone et al., 2024; Kouadio et al., 2024). Other studies explain this anticorrelation by noting that higher wind speeds can transport aerosols, increasing aerosol optical depth (AOD) and thereby reducing surface solar radiation under certain conditions (Smirnov et al., 2012).

An increase in cloud cover fraction ($\beta_C = -1.154$, 94% HDI: [-1.710, -0.668]) leads to a reduction in the clear-sky index and consequently in GHI, consistent with the well-established understanding that clouds obstruct solar radiation. However, the attenuating effect of cloud cover is weaker than that of relative humidity. This can be explained by the fact that cloud cover directly impacts the amount of solar radiation reaching the surface. High humidity, on the other hand, often contributes to cloud formation, which in turn reduces surface solar radiation. Moreover, humidity increases the absorption and scattering of radiation in the atmosphere, even in cloud-free conditions, through hygroscopic effects. Additionally, in tropical regions like Côte d'Ivoire, humidity remains high throughout most of the year (typically 80–90% or more), whereas cloud cover is more variable. In modelling point of view, the correlation between humidity and cloud cover could lead the model to distribute the attenuation effect between the two variables, thereby reducing the estimated coefficient for one when the other is included. Lastly, ground-based humidity measurements may be more reliable than satellite-derived cloud cover estimates.

In the literature, fine particulate matter ($PM_{2.5}$) is widely reported to attenuate solar radiation by scattering and absorbing sunlight (Hu et al., 2025; Shim et al., 2025; Song et al., 2022). In this study, we approximate aerosol effects using $PM_{2.5}$ concentrations. The relationship between $PM_{2.5}$ and the clear-sky index is modeled using a quadratic function. The estimated coefficients, $\beta_{P1} = 0.010$ (94% HDI: [-0.012, 0.031]) and $\beta_{P2} = 0.009$ (94% HDI: [0.003, 0.016]), suggest a slight positive correlation between $PM_{2.5}$ and the clear-sky index. This implies that the attenuating effect of $PM_{2.5}$ is not strongly supported by the data. Most stations in this study are located in rural, low-industrialized areas where $PM_{2.5}$ levels tend to be relatively low, which may explain the weak or absent correlation. Indeed, Song et al. (2022) demonstrated that a certain concentration threshold is necessary for $PM_{2.5}$-induced attenuation to be noticeable. In their study in Hong Kong, they found that when $PM_{2.5}$ concentrations exceeded approximately 33.5 $\mu g/m^3$, global horizontal irradiance (GHI) decreased by over 5% in most months. Furthermore, the $PM_{2.5}$ data used in this study are derived from reanalysis and satellite products, which may not capture fine-scale local variations in aerosol concentrations accurately.

However, under low cloud cover conditions (cloudless but overcast skies, or KtC-FCl), the relatively high Pearson correlation coefficients between $PM_{2.5}$ and solar radiation in scatter plots, along with their regression curves, suggest that $PM_{2.5}$ may still exert a non-negligible influence on solar radiation under clear-sky conditions. This effect could be masked by the presence of clouds in other situations.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

3.4 Limitations and Future Work

Some limitations should be acknowledged in this study. First, the $PM_{2.5}$ and cloud cover data were obtained from satellite products and comined with in-situ measurements. Using $PM_{2.5}$ concentrations as a proxy for aerosol content, instead of aerosol optical depth (AOD), may miss certain components of atmospheric pollution such as dust and other particulates that can affect solar radiation. Second, computational constraints limited the Bayesian inference process to a subsample of 2,000 observations, rather than using the full dataset.

Future studies could also use AOD data from ground-based sun photometers (AERONET) or satellite products such as MODIS and MISR, which may provide a more comprehensive measure of aerosol impact. Also, including additional meteorological variables, such as precipitation and atmospheric pressure, may help develop a more complete understanding of the drivers of solar radiation variability in tropical regions.

4 Conclusions

This study developed a Bayesian hierarchical model to analyze the influence of weather variables and air pollution on global horizontal irradiance via the clear-sky index. The model incorporated temperature, relative humidity, wind speed, cloud cover fraction and $PM_{2.5}$ concentrations as predictors, along with seasonal harmonics and station-specific random effects to account for temporal and spatial variability. The best-performing scenario including non-linear effects and three seasonal harmonics, effectively captures the nonlinear influence of predictors on solar radiation and outperformed simpler generalized linear and additive models. Moreover, station-level random effects, driven by local environmental characteristics, are non-negligible and should be accounted for when training models across multiple sites; architectures such as graph neural networks may be more suitable for this purpose. While temperature, humidity and cloud cover emerge as the primary drivers of global horizontal irradiance, a noticeable influence of $PM_{2.5}$ is observed under cloud-free conditions. Future work could expand the analysis to more industrialized areas, incorporate additional atmospheric variables and explore alternative computational methods to enhance model efficiency and complexity.

Declaration of generative AI utilisation in the manuscript preparation process

During the preparation of this work, the author(s) used ChatGPT to assist with English translation, Python code debugging, and identifying relevant references. All content generated with this tool was reviewed and edited by the author(s), who take full responsibility for the accuracy and integrity of the article.

Author contributions. Konin Pierre-Claver Kakou: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Visualization, Writing - original draft. Dungall Laouali, Boko Aka and Georg Frey: Review, Supervision.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Competing interests. The authors declare no competing interests.

Code/Data availability

The current version of *Bayesian Global Horizontal Irradiance Modeling: Code & Data* is available from the project website at https://zenodo.org/records/17818456 under the Creative Commons Attribution 4.0 International License. The exact version of the model used to produce the results in this paper is archived on Zenodo under DOI 10.5281/zenodo.17818456, as are the input data and the scripts needed to run the model and reproduce all figures and simulations presented in this work (Kakou, 2025).

Acknowledgements. The authors sincerely thank the West African Science Service Centre on Climate Change and Adapted Land Use and the German Federal Ministry of Education and Research for their financial support, which made this research possible. We also gratefully acknowledge the National Meteorological and Aeronautical Services of Côte d'Ivoire for providing the essential data used in this study.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

References

Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fonnesbeck, C. J., Kochurov, M., Kumar, R., Lao, J., Luhmann, C. C., Martin, O. A., Osthege, M., Vieira, R., Wiecki, T., and Zinkov, R.: PyMC: a modern, and comprehensive probabilistic programming framework in Python, PeerJ Computer Science, 9, e1516, https://doi.org/10.7717/peerj-cs.1516, 2023.

Akobé, E. Y., Diawara, A., Yoroba, F., Kouassi, B. K., Yapo, A. L. M., Diba, I., Kouadio, K., Tiémoko, D. T., Koné, D. I., and Diedhiou, A.: Trends in the Occurrence of Compound Extremes of Temperature and Precipitation in Côte d'Ivoire, Atmosphere, 16, https://doi.org/10.3390/atmos16010003, 2025.

Alam, F., Rehman, S. U., Rehman, S., Jahangir, M., Shoaib, M., Siddiqui, I., and Ulfat, I.: Empirical model development for the estimation of clearness index using meteorological parameters, Turkish Journal of Electrical Engineering and Computer Sciences, 27, 4429–4441, https://doi.org/10.3906/elk-1903-27, 2019.

Amaro e Silva, R., Benavides Cesar, L., Manso Callejo, M. A., and Cira, C.-I.: Impact of Stationarizing Solar Inputs on Very-Short-Term Spatio-Temporal Global Horizontal Irradiance (GHI) Forecasting, Energies, 17, https://doi.org/10.3390/en17143527, 2024.

Arora, P. and Ceferino, L.: Probabilistic and machine learning methods for uncertainty quantification in power outage prediction due to extreme events, Natural Hazards and Earth System Sciences, 23, 1665–1683, https://doi.org/10.5194/nhess-23-1665-2023, 2023.

Assoumou, E. and McIsaac, F.: Côte d'Ivoire's electricity challenge in 2050: Reconciling economic development and climate commitments, Energy Policy, 160, 112 681, https://doi.org/10.1016/j.enpol.2021.112681, 2022.

Bahino, J., Giordano, M., Beekmann, M., Yoboué, V., Ochou, A., Galy-Lacaux, C., Liousse, C., Hughes, A., Nimo, J., Lemmouchi, F., Cuesta, J., Amegah, A. K., and Subramanian, R.: Temporal variability and regional influences of PM2.5 in the West African cities of Abidjan (Côte d'Ivoire) and Accra (Ghana)††Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4ea00012a, Environmental Science: Atmospheres, 4, 468–487, https://doi.org/10.1039/d4ea00012a, 2024.

Bank, W.: M300 AES Compact Côte d'Ivoire, Report, World Bank / Mission 300 (M300), Côte d'Ivoire, https://africa-energy-portal.org/sites/default/files/2025-02/m300_aes_compact_civ-27012025_0.pdf, 2025.

Bazrafshan, O., Zamani, H., Farokhzadeh, B., and Caloiero, T.: A Spectral Analysis-Driven SARIMAX Framework with Fourier Terms for Monthly Dust Concentration Forecasting, Earth, 6, https://doi.org/10.3390/earth6040123, 2025.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Scientific Data, 5, 180 214, https://doi.org/10.1038/sdata.2018.214, 2018.

Bett, P. E. and Thornton, H. E.: The climatological relationships between wind and solar energy supply in Britain, Renewable Energy, 87, 96–110, https://doi.org/10.1016/j.renene.2015.10.006, 2016.

Bilal, B., Rune, G., and Tobias, B.: Solar radiation estimation at high latitudes: Assessment of the CMSAF databases, ASR and ERA5, Solar Energy, 182, 397–411, https://doi.org/10.1016/j.solener.2019.02.058, 2019.

Birdal, R. G.: The Influence of Air Pollution Concentrations on Solar Irradiance Forecasting Using CNN-LSTM-mRMR Feature Extraction, Computers, Materials and Continua, 78, 4015–4028, https://doi.org/10.32604/cmc.2024.048324, 2024.

Cleveland, W. S.: Robust Locally Weighted Regression and Smoothing Scatterplots, Journal of the American Statistical Association, 74, 829–836, https://doi.org/10.1080/01621459.1979.10481038, 1979.

Cleveland, W. S. and Devlin, S. J.: Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting, Journal of the American Statistical Association, 83, 596–610, https://doi.org/10.1080/01621459.1988.10478639, 1988.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Collow, A., Buchard, V. J., Chin, M., Colarco, P. R., Darmenov, A. S., and Silva, A. M. D.: Supplemental Documentation for GEOS Aerosol Products, *GMAO Office Note No. 22 (Version 1.0)*, Goddard Space Flight Center, https://gmao.gsfc.nasa.gov/pubs/office_note, document ID: 20220017880. Acquired: November 28, 2022, 2022.

Collow, A., Buchard, V. J., Chin, M., Colarco, P. R., Darmenov, A. S., and Silva, A. M. D.: Supplemental Documentation for GEOS Aerosol Products, Tech. Rep. 22 (Version 1.0), GMAO (Global Modeling and Assimilation Office), Greenbelt, Maryland, United States, available online: https://gmao.gsfc.nasa.gov/pubs/office_notes (accessed on 11 March 2025), 2023.

Congdon, P.: Modelling changes in Arctic Sea Ice Cover: an application of generalized and inflated beta and gamma densities, Journal of Statistical Distributions and Applications, 1, 3, https://doi.org/10.1186/2195-5832-1-3, 2014.

Correa, L. F., Folini, D., Chtirkova, B., and Wild, M.: A Method for Clear-Sky Identification and Long-Term Trends Assessment Using Daily Surface Solar Radiation Records, Earth and Space Science, 9, e2021EA002 197, https://doi.org/https://doi.org/10.1029/2021EA002197, e2021EA002197 2021EA002197, 2022.

Danso, D. K., Anquetin, S., Diedhiou, A., Lavaysse, C., Kobea, A., and Touré, N. E.: Spatio-temporal variability of cloud cover types in West Africa with satellite-based and reanalysis data, Quarterly Journal of the Royal Meteorological Society, 145, 3715–3731, https://doi.org/10.1002/qj.3651, 2019.

Davidian, M. and Giltinan, D. M.: Nonlinear models for repeated measurement data: An overview and update, Journal of Agricultural, Biological, and Environmental Statistics, 8, 387, https://doi.org/10.1198/1085711032697, 2003.

Deo, R. C., Wen, X., and Qi, F.: A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset, Applied Energy, 168, 568–593, https://doi.org/10.1016/j.apenergy.2016.01.130, 2016.

Ding, Y., Wang, Y., Li, Z., Zhao, L., Shi, Y., Xing, X., and Chen, S.: Improving Solar Radiation Prediction in China: A Stacking Model Approach with Categorical Boosting Feature Selection, Atmosphere, 15, https://doi.org/10.3390/atmos15121436, 2024.

El Alani, O., Abraim, M., Ghennioui, H., Ghennioui, A., Ikenbi, I., and Dahr, F.-E.: Short term solar irradiance forecasting using sky images based on a hybrid CNN–MLP model, Energy Reports, 7, 888–900, https://doi.org/10.1016/j.egyr.2021.07.053, technologies and Materials for Renewable Energy, Environment and Sustainability, 2021.

Ermida, S. L., Trigo, I. F., DaCamara, C. C., Jiménez, C., and Prigent, C.: Quantifying the Clear-Sky Bias of Satellite Land Surface Temperature Using Microwave-Based Estimates, Journal of Geophysical Research: Atmospheres, 124, 844–857, https://doi.org/10.1029/2018JD029354, 2019.

Espinar, B., Ramírez, L., Drews, A., Beyer, H. G., Zarzalejo, L. F., Polo, J., and Martín, L.: Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations, Solar Energy, 83, 118–125, https://doi.org/10.1016/j.solener.2008.07.009, 2009.

Flor, M., Weiß, M., Selhorst, T., Müller-Graf, C., and Greiner, M.: Comparison of Bayesian and frequentist methods for prevalence estimation under misclassification, BMC Public Health, 20, 1135, https://doi.org/10.1186/s12889-020-09177-4, 2020.

Font, A., F. de Brito, J., Riffault, V., Conil, S., Jaffrezo, J.-L., and Bourin, A.: Calculations of the conversion factor from organic carbon to organic matter for aerosol mass balance, Atmospheric Pollution Research, 15, 102 301, https://doi.org/10.1016/j.apr.2024.102301, 2024.

Global Modeling and Assimilation Office (GMAO): M2T1NXAER. Version 5.12.4. MERRA-2 tavg1_2d_aer_Nx: 2d, 1-Hourly, Time-averaged, Single-Level, Assimilation, Aerosol Diagnostics V5.12.4, Archived by National Aeronautics and Space Administration, Goddard Earth Sciences Data and Information Services Center (GES DISC), https://doi.org/10.5067/KLICLTZ8EM9D, geospatial dataset covering the time range 1980-01-01 to 2022-01-17. Includes aerosol diagnostics for atmospheric and air quality studies., 2015.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Gnamien, S., Yoboué, V., Liousse, C., Ossohou, M., Keita, S., Bahino, J., Siélé, S., and Diaby, L.: Particulate Pollution in Korhogo and Abidjan (Cote d'Ivoire) during the Dry Season, Aerosol and Air Quality Research, 21, 200 201, https://doi.org/10.4209/aaqr.2020.05.0201, 2021.

Gnamien, S., Liousse, C., Keita, S., Silué, S., Bahino, J., Gardrat, E., Kassamba-Diaby, M., Ochou, A., and Yoboué, V.: Chemical characterization of urban aerosols in Abidjan and Korhogo (Côte d'Ivoire) from 2018 to 2020 and the identification of their potential emission sources, Environ. Sci.: Atmos., 3, 1741–1757, https://doi.org/10.1039/D3EA00131H, 2023.

Gueymard, C. A.: A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects, Renewable and Sustainable Energy Reviews, 39, 1024–1034, https://doi.org/10.1016/j.rser.2014.07.117, 2014.

Gupta, A., Gupta, K., and Saroha, S.: Solar Irradiation Forecasting Technologies: A Review, Strategic Planning for Energy and the Environment, 39, 319–354, https://doi.org/10.13052/spee1048-4236.391413, 2021.

Gürel, A. E., Ağbulut, Ü., Bakır, H., Ergün, A., and Yıldız, G.: A state of art review on estimation of solar radiation with various models, Heliyon, 9, e13 167, https://doi.org/10.1016/j.heliyon.2023.e13167, 2023.

Hedar, A.-R., Almaraashi, M., Abdel-Hakim, A. E., and Abdulrahim, M.: Hybrid Machine Learning for Solar Radiation Prediction in Reduced Feature Spaces, Energies, 14, https://doi.org/10.3390/en14237970, 2021.

Hespanhol, L., Vallio, C. S., Costa, L. M., and Saragiotto, B. T.: Understanding and interpreting confidence and credible intervals around effect estimates, Brazilian Journal of Physical Therapy, 23, 290–301, https://doi.org/10.1016/j.bjpt.2018.12.006, 2019.

Hridoy, M. B.: An exploration of modeling approaches for capturing seasonal transmission in stochastic epidemic models, Mathematical Biosciences and Engineering, 22, 324–354, https://doi.org/10.3934/mbe.2025013, 2025.

Hu, A., Duan, Z., Zhang, Y., Huang, Z., Ji, T., and Yin, X.: Impact of PM2.5 Pollution on Solar Photovoltaic Power Generation in Hebei Province, China, Energies, 18, https://doi.org/10.3390/en18154195, 2025.

Huang, L., Kang, J., Wan, M., Fang, L., Zhang, C., and Zeng, Z.: Solar Radiation Prediction Using Different Machine Learning Algorithms and Implications for Extreme Climate Events, Frontiers in Earth Science, Volume 9 - 2021, 596 860, https://doi.org/10.3389/feart.2021.596860, 2021.

International Finance Corporation: Côte d'Ivoire: Unlocking Private Investment to Achieve 42% Renewable Energy by 2030, https://www.ifc.org/en/insights-reports/2018/cote-d-ivoire-unlocking-private-investment, [Online; accessed 19-September-2025], 2018.

Jeon, H.-J., Choi, M.-W., and Lee, O.-J.: Day-Ahead Hourly Solar Irradiance Forecasting Based on Multi-Attributed Spatio-Temporal Graph Convolutional Network, Sensors, 22, https://doi.org/10.3390/s22197179, 2022.

Kakou, K. P.-C.: Bayesian Global Horizontal Irradiance Modeling: Code & Data (v1.0), https://doi.org/10.5281/zenodo.17818456, 2025.

Kakou, P.-C. K., Laouali, D., Aka, B., Osei, J. A., Ette, N. F. K., and Frey, G.: Multi-Timescale Validation of Satellite-Derived Global Horizontal Irradiance in Côte d'Ivoire, Remote Sensing, 17, https://doi.org/10.3390/rs17060998, 2025.

Karlsson, K.-G., Stengel, M., Meirink, J. F., Riihelä, A., Trentmann, J., Akkermans, T., Stein, D., Devasthale, A., Eliasson, S., Johansson, E., Håkansson, N., Solodovnik, I., Benas, N., Clerbaux, N., Selbach, N., Schröder, M., and Hollmann, R.: CLARA-A3: The third edition of the AVHRR-based CM SAF climate data record on clouds, radiation and surface albedo covering the period 1979 to 2023, Earth System Science Data, 15, 4901–4926, https://doi.org/10.5194/essd-15-4901-2023, 2023.

Kazil, J., Feingold, G., and Yamaguchi, T.: Wind speed response of marine non-precipitating stratocumulus clouds over a diurnal cycle in cloud-system resolving simulations, Atmospheric Chemistry and Physics, 16, 5811–5839, https://doi.org/10.5194/acp-16-5811-2016, 2016.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Kone, D., Tuo, S., Coulibaly, K. A., Olivier Gore, B. T., Dje, K. B., Traore, M., and Aka, B.: Assessment of wind potential energy of four major cities in Cote d'ivoire using satellite data from 2015 to 2022, Journal of King Saud University - Science, 36, 103 579, https://doi.org/10.1016/j.jksus.2024.103579, 2024.

Kouadio, J.-M. S., Néné, F. D., Grafouté, M., N'guessan, A., Yeboua, S. P. C., and Yao, N.: Harnessing the wind energy potential in Yamoussoukro, the Economic Capital of Côte d'Ivoire, Heliyon, 10, e30 170, https://doi.org/10.1016/j.heliyon.2024.e30170, epub 2024 May 1, 2024.

Kouadio, K., Adje, J., Maesano, C., Veronique, Y., and Liousse, C.: PM2.5 concentration and Health Impact Assessment in Abidjan Hospital, Côte d'Ivoire, Environmental Epidemiology, 3, https://doi.org/10.1097/01.EE9.0000608208.12827.c0, 2019.

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

Lauret, P., Alonso-Suárez, R., Le Gal La Salle, J., and David, M.: Solar Forecasts Based on the Clear Sky Index or the Clearness Index: Which Is Better?, Solar, 2, 432–444, https://doi.org/10.3390/solar2040026, 2022.

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., and Morcrette, J.-J.: McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions, Atmospheric Measurement Techniques, 6, 2403–2418, https://doi.org/10.5194/amt-6-2403-2013, 2013.

Lindstrom, M. J. and Bates, D. M.: Nonlinear Mixed Effects Models for Repeated Measures Data, Biometrics, 46, 673–687, https://doi.org/10.2307/2532087, 1990.

Moron, V., Camberlin, P., Aellig, R., Champagne, O., Fink, A. H., Knippertz, P., and Philippon, N.: Diurnal to interannual variability of low-level cloud cover over western equatorial Africa in May–October, International Journal of Climatology, 43, 6038–6064, https://doi.org/10.1002/joc.8188, 2023.

NASA Global Modeling and Assimilation Office: Aerosol Optics from the GOCART Model, available online: https://portal.nccs.nasa.gov/datashare/iesa/aerosol/AerosolOptics/ (Accessed on 11 March 2025), 2024.

National Weather Service: Did You Know? - Sky Condition Definitions, https://www.weather.gov/psr/didyouknow, accessed: 2025-01-19, n.d.

Natras, R., Soja, B., and Schmidt, M.: Uncertainty Quantification for Machine Learning-Based Ionosphere and Space Weather Forecasting: Ensemble, Bayesian Neural Network, and Quantile Gradient Boosting, Space Weather, 21, e2023SW003 483, https://doi.org/10.1029/2023SW003483, 2023.

Navinya, C. D., Vinoj, V., and Pandey, S. K.: Evaluation of PM2.5 Surface Concentrations Simulated by NASA's MERRA Version 2 Aerosol Reanalysis over India and its Relation to the Air Quality Index, Aerosol and Air Quality Research, 20, 1329–1339, https://doi.org/10.4209/aaqr.2019.12.0615, 2020.

NIST/SEMATECH: 4.1.4.4. LOESS (aka LOWESS), https://www.itl.nist.gov/div898/handbook/pmd/section1/pmd144.htm, https://doi.org/10.18434/M32189, part of the NIST/SEMATECH e-Handbook of Statistical Methods; last updated April 27, 2022; accessed September 4, 2025, 2022.

Ospina, R. and Ferrari, S. L. P.: Inflated beta distributions, Statistical Papers, 51, 111–126, https://doi.org/10.1007/s00362-008-0125-4, 2010.

Oxford Business Group: Diversifying the mix: Solar energy and biomass plants are targeted for development alongside new dams in an effort to increase energy capacity, https://oxfordbusinessgroup.com/reports/cote-divoire/2019-report/economy/diversifying-the-mix-solar-energy-and-biomass-plants-are-targeted-for-development-alongside-new-dams-in-an-effort-to-increase-energy-capacity, [Online; accessed 19-September-2025], 2019.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Paletta, Q., Terrén-Serrano, G., Nie, Y., Li, B., Bieker, J., Zhang, W., Dubus, L., Dev, S., and Feng, C.: Advances in solar forecasting: Computer vision with deep learning, Advances in Applied Energy, 11, 100 150, 2023.

Peng, X., She, J., Zhang, S., Tan, J., and Li, Y.: Evaluation of Multi-Reanalysis Solar Radiation Products Using Global Surface Observations, Atmosphere, 10, https://doi.org/10.3390/atmos10020042, 2019.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F.: A new operational model for satellite-derived irradiances: description and validation, Solar Energy, 73, 307–317, https://doi.org/https://doi.org/10.1016/S0038-092X(02)00122-6, 2002.

Popouen, A. J., Djagouri, K., AGBO, D. D. A., Koua, A. A., and Monnehan, A. G.: Concentration levels of PM_{2.5}, PM₁₀ and Black Carbon in the Industrial Area of Yopougon, Abidjan, Côte d'Ivoire, International Journal of Physics, 9, 90–95, https://doi.org/10.12691/ijp-9-2-4, 2021.

Rajput, P.: OM/OC Ratio of Polar and Non-Polar Organic Matter during Wintertime from Indo-Gangetic Plain: Implications to Regional-Scale Radiative Forcing, Aerosol Science and Engineering, 2, 153–164, https://doi.org/10.1007/s41810-018-0032-6, 2018.

Sacco, M. A., Ruiz, J. J., Pulido, M., and Tandeo, P.: Evaluation of machine learning techniques for forecast uncertainty quantification, Quarterly Journal of the Royal Meteorological Society, 148, 3470–3490, https://doi.org/10.1002/qj.4362, 2022.

Sawadogo, W., Bliefernicht, J., Fersch, B., Salack, S., Guug, S., Diallo, B., Ogunjobi, K., Nakoulma, G., Tanu, M., Meilinger, S., and Kunstmann, H.: Hourly global horizontal irradiance over West Africa: A case study of one-year satellite- and reanalysis-derived estimates vs. in situ measurements, Renewable Energy, 216, 119 066, https://doi.org/10.1016/j.renene.2023.119066, 2023.

Seabold, S. and Perktold, J.: Statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, 2010.

Seide, S. E., Jensen, K., and Kieser, M.: A comparison of Bayesian and frequentist methods in random-effects network meta-analysis of binary data, Research Synthesis Methods, 11, 363–378, https://doi.org/https://doi.org/10.1002/jrsm.1397, 2020.

Servén, D. and Brummitt, C.: pyGAM: Generalized Additive Models in Python, https://doi.org/10.5281/zenodo.1208723, 2018.

Shim, J., Park, S., and Song, D.: Impact of particulate matter (PM10, PM2.5) on global horizontal irradiance and direct normal irradiance in urban areas, Building and Environment, 271, https://doi.org/10.1016/j.buildenv.2025.112610, publisher Copyright: © 2025, 2025.

Smirnov, A., Sayer, A. M., Holben, B. N., Hsu, N. C., Sakerin, S. M., Macke, A., Nelson, N. B., Courcoux, Y., Smyth, T. J., Croot, P., Quinn, P. K., Sciare, J., Gulev, S. K., Piketh, S., Losno, R., Kinne, S., and Radionov, V. F.: Effect of wind speed on aerosol optical depth over remote oceans, based on data from the Maritime Aerosol Network, Atmospheric Measurement Techniques, 5, 377–388, https://doi.org/10.5194/amt-5-377-2012, 2012.

Song, Z., Wang, M., and Yang, H.: Quantification of the Impact of Fine Particulate Matter on Solar Energy Resources and Energy Performance of Different Photovoltaic Technologies, ACS Environmental Au, 2, 275–286, https://doi.org/10.1021/acsenvironau.1c00048, 2022.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A.: Bayesian measures of model complexity and fit, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64, 583–639, https://doi.org/10.1111/1467-9868.00353, 2002.

Stein, J. S., Hansen, C. W., and Reno, M. J.: Global horizontal irradiance clear sky models: implementation and analysis, Tech. rep., Sandia National Laboratories, https://digital.library.unt.edu/ark:/67531/metadc831646/, report, March 1, 2012. Accessed September 3, 2025. Available from University of North Texas Libraries, UNT Digital Library., 2012.

Stöckli, R., Bourgeois, Q., Tetzlaff, A., Schöder, M., and Hollmann, R.: CM SAF Cloud Fractional Cover dataset from Meteosat First and Second Generation - Edition 2 (COMET Ed. 2), https://doi.org/10.5676/EUM_SAF_CM/CFC_METEOSAT/V002, 2024.

https://doi.org/10.5194/egusphere-2025-6054
Preprint. Discussion started: 26 February 2026

Stüber, M., Scherhag, F., Deru, M., Ndiaye, A., Sakha, M. M., Brandherm, B., Baus, J., and Frey, G.: Forecast Quality of Physics-Based and Data-Driven PV Performance Models for a Small-Scale PV System, Frontiers in Energy Research, Volume 9 - 2021, 639 346, https://doi.org/10.3389/fenrg.2021.639346, 2021.

Sustainable Energy for All: Rapid Assessment Gap Analysis Côte d'Ivoire, https://www.seforall.org/sites/default/files/l/2015/05/Cote_dIvoire_RAGA.pdf, [Online; accessed 19-September-2025], 2012.

Tham, Y., Muneer, T., and Davison, B.: A generalized procedure to generate clear-sky radiation data for any location, International Journal of Low-Carbon Technologies, 4, 205–212, https://doi.org/10.1093/ijlct/ctp022, 2009.

The ArviZ Developers: arviz.compare — ArviZ documentation, https://python.arviz.org/en/stable/api/generated/arviz.compare.html, version 0.22.0, 2025.

Turpin, B. J. and and, H.-J. L.: Species Contributions to PM2.5 Mass Concentrations: Revisiting Common Assumptions for Estimating Organic Mass, Aerosol Science and Technology, 35, 602–610, https://doi.org/10.1080/02786820119445, 2001.

van der Linden, R., Fink, A. H., and Redl, R.: Satellite-based climatology of low-level continental clouds in southern West Africa during the summer monsoon season, Journal of Geophysical Research: Atmospheres, 120, 1186–1201, https://doi.org/10.1002/2014JD022614, 2015.

Vehtari, A., Gelman, A., and Gabry, J.: Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC, Statistics and Computing, 27, 1413–1432, https://doi.org/10.1007/s11222-016-9696-4, 2017.

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

World Health Organization: WHO Global Air Quality Guidelines: Particulate Matter ($PM_{2.5}$ and PM_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide, World Health Organization, Geneva, ISBN 9789240034228, https://www.who.int/publications/i/item/9789240034228, 2021.

Yapo, A. L. M., Diawara, A., Yoroba, F., Kouassi, B. K., Sylla, M. B., Kouadio, K., Odoulami, R. C., and Tiémoko, D. T.: Twenty-First Century Projected Changes in Extreme Temperature over Côte d'Ivoire (West Africa), International Journal of Geophysics, 2019, 5610 328, https://doi.org/10.1155/2019/5610328, 2019.

Zamora, R. J., Dutton, E. G., Trainer, M., McKeen, S. A., Wilczak, J. M., and Hou, Y.-T.: The Accuracy of Solar Irradiance Calculations Used in Mesoscale Numerical Weather Prediction, Monthly Weather Review, 133, 783 – 792, https://doi.org/10.1175/MWR2886.1, 2005.

Zhang, X., Zhang, M., Cui, Y., and He, Y.: Estimation of Daily Ground-Received Global Solar Radiation Using Air Pollutant Data, Frontiers in Public Health, Volume 10 - 2022, 860 107, https://doi.org/10.3389/fpubh.2022.860107, 2022.

Zhu, T., Guo, Y., Wang, C., and Ni, C.: Inter-Hour Forecast of Solar Radiation Based on the Structural Equation Model and Ensemble Model, Energies, 13, https://doi.org/10.3390/en13174534, 2020.