This review concentrates on the statistical part of the manuscript.

The manuscript raises important point on the problem of interpreting nonlinear process with linear models. Which is alarmingly common.

The manuscript acts as method paper by introducing new statistical modelling framework on the solar energy related data. It applies data from Côte d'Ivoire as their test data, but the method would be applicable also in any other land area. However, the manuscript needs major revision before I could recommend it for publication. See my specific comments below.

We sincerely thank the reviewer for the careful evaluation of our manuscript and for recognizing both the importance of investigating nonlinear atmospheric processes and the potential value of the proposed statistical framework. We appreciate the reviewer's positive assessment that the work addresses an important methodological challenge and that the proposed approach has applicability beyond the case study presented here.

We also appreciate the reviewer's detailed comments, which have been invaluable in improving the manuscript. In response, the manuscript has undergone a substantial revision. The methodology has been fundamentally redesigned and now introduces a Bayesian Nonlinear Hierarchical Structural Model (BNHSM) that replaces the previous Bayesian Nonlinear Mixed-Effects (BNLME) formulation. The revised manuscript also adopts the clearness index as the response variable, explicitly accounts for nonlinear relationships using Bayesian penalized splines, incorporates interaction effects among key atmospheric variables and considerably strengthens the physical interpretation of the statistical results while carefully distinguishing conditional associations from causal effects. In addition, the study has been extended from a single-country application to a comparative analysis across two contrasting climatic regions, Côte d'Ivoire and Germany, thereby considerably broadening the scope and generality of the work.

We have carefully addressed each of the reviewer's specific comments individually below.

Specific comments

lines 59-65: You cannot refer regression models and linear mixed models as black box as they are deterministic models with clear equational for that can be solved (often numerically, but still). Though, I agree with the need for proper uncertainty estimates.

We sincerely thank the reviewer for this important observation. We agree that referring to regression and linear mixed-effects models as *black-box* models was inappropriate, as these models are transparent, mathematically well defined, and highly interpretable. As part of the major revision, the

Introduction has been substantially rewritten and this statement has been removed. The revised manuscript now motivates the proposed Bayesian framework by emphasizing its ability to quantify parameter and predictive uncertainty, flexibly model nonlinear relationships and account for hierarchical data structures. This concern has been fully addressed in the revised manuscript.

lines 153-162: the data are limited quite much. Why so strict limitations with time, should you not examine the whole daylight period to see the overall solar energy production.

We sincerely thank the reviewer for this valuable comment. We agree that restricting the analysis to a limited period of the day could unnecessarily reduce the representativeness of the dataset when the objective is to characterize the full variability of surface solar irradiance. As part of the major revision, the data preprocessing strategy has been completely redesigned. The previous temporal filtering, the associated temperature filtering and IQR-based outlier removal procedure, has been removed. The revised study now analyzes the complete set of valid daytime observations, and accounts for the deterministic effects of solar geometry through the clearness index and explicit temporal components within the Bayesian hierarchical model.

Lines 163-171: nice overview of reasons why commonly applied models do not work here, I would add GLMM (sometimes useful but limited uncertainty quantification) and GAM (which you apply later in the manuscript) to the list.

Table 2: if you are not using the abbreviations later use the full names of the models here for better readability. (as a frequentist, for FNLME I would not consider missing prior specification as limitation, actually quite opposite as I am not comfortable with the idea of strong priors defining the outcome of the models).

We sincerely thank the reviewer for these valuable suggestions. We agree that the discussion of alternative statistical approaches in the original manuscript could have been more comprehensive, particularly regarding generalized additive models (GAMs) and generalized linear mixed models (GLMMs). We also appreciate the comment concerning the presentation of model names and the discussion of prior information in relation to frequentist nonlinear mixed-effects models.

As part of the major revision, however, this section has been completely rewritten. The revised Introduction no longer presents a comparative catalogue of alternative modelling approaches or the corresponding table of model limitations. Instead, it focuses on motivating the proposed Bayesian Nonlinear Hierarchical Structural Model (BNHSM) from the perspective of the scientific problem,

namely, the need to jointly represent nonlinear atmospheric relationships, hierarchical dependence across stations and countries, interaction effects and principled uncertainty quantification.

Equations 7 and 8: Justify the use of logit link.

We sincerely thank the reviewer for this valuable comment. We agree that the original manuscript did not sufficiently justify the choice of the logit link function. As part of the major revision, the model formulation has been substantially redesigned and the methodological section has been completely rewritten. The revised manuscript now explicitly motivates the choice of the logit link within the Bayesian Beta regression framework by noting that it maps the conditional mean of the clearness index, which is naturally bounded on the open interval (0,1), onto the real line while ensuring that predicted values remain within the admissible range. This choice is standard for Beta regression and facilitates the interpretation and estimation of nonlinear additive effects within the hierarchical Bayesian model. Consequently, the concern has been addressed in the revised methodology.

Lines 216-217: With which variables were he correlations calculated? Overall using only R as measure of goodness-of-fit is not enough, it looks only one aspect of the fit.

Lines 218-219: justify the use of harmonic functions for seasonality.

We sincerely thank the reviewer for these valuable comments. We agree that the original procedure used to select nonlinear functional forms based primarily on Pearson's correlation coefficient was not sufficiently rigorous, and that the rationale for the seasonal harmonic representation required further justification.

This entire model-building strategy has been replaced. The revised Bayesian Nonlinear Hierarchical Structural Model no longer relies on fitting multiple candidate functions (linear, quadratic, exponential, logarithmic, sigmoid, etc.) and selecting one according to Pearson's R. Instead, nonlinear relationships are estimated directly from the data using Bayesian penalized splines. This eliminates the need for a preliminary function-selection procedure. In addition, the treatment of temporal variability has been reformulated and explicitly justified in the revised methodology.

Figure 2 really hard to read. The color scheme is poorly chosen and the numbers within need 300% magnification to be readable with 27” display.

We sincerely thank the reviewer for this observation.

We agree that the original figure was difficult to read and that its primary purpose was to illustrate the preliminary procedure used to select nonlinear functional forms for the predictors. In the major revision, this procedure has been removed because the revised Bayesian Nonlinear Hierarchical Structural Model estimates nonlinear relationships directly using Bayesian penalized splines.

Equation 12: Give the functions for OVBER and rRMSD also here.

We thank the reviewer for this suggestion.

In the revised methodology, the performance metrics and equations concerned, including OVER and rRMSD, are no longer used in the revised manuscript. Equation 12 and the associated discussion have therefore been removed.

Lines 288-295: What are the LOO score? What is elpd.loo and p.loo and what is their difference?

We thank the reviewer for this important comment.

We agree that the original manuscript did not define the leave-one-out cross-validation quantities clearly enough. In particular, elpd.loo represents the estimated expected log predictive density under leave-one-out cross-validation and measures out-of-sample predictive performance, whereas p.loo is an estimate of the model's effective number of parameters and reflects model complexity. These quantities therefore serve different purposes and should not be interpreted as directly comparable measures.

In the major revision, however, the previous comparison among several BNLME scenarios has been removed. The revised manuscript instead presents a dedicated model-evaluation and diagnostics section in which PSIS-LOO, WAIC, their predictive interpretation and the associated Pareto-k diagnostics are explicitly defined before the results are discussed. The terminology and presentation have consequently been clarified in the revised version.

Figure 5 introduces the posterior plots but what kind of prior did you have?

We sincerely thank the reviewer for this important observation.

We agree that the original manuscript introduced posterior distributions without first clearly describing the prior distributions used in the Bayesian model. As part of the major revision, this issue has been fully addressed. The revised methodology now includes a dedicated description of the prior distributions assigned to all model parameters, including the fixed effects, spline coefficients,

interaction terms, random-effect standard deviations and the Beta precision parameter. These priors are introduced before presenting the posterior results.

Line 319 and Table6: What kind of spline functions and other specification did you define for GAM? Gaussian or binary family? That might affect the fit quite drastically and thus also your comparisons.

We sincerely thank the reviewer for this important observation.

We agree that a meaningful comparison with a generalized additive model (GAM) requires a complete specification of the smoothing basis, smoothing parameter estimation, response family, and link function, as these choices can substantially influence model performance. As part of the major revision, however, the comparison with GLM and GAM has been removed. The revised manuscript focuses on the proposed Bayesian Nonlinear Hierarchical Structural Model (BNHSM) and evaluates its performance using Bayesian model validation and diagnostic tools, including posterior predictive checks, Bayesian R^2, PSIS-LOO, WAIC and convergence diagnostics. Consequently, this comment is no longer applicable to the revised manuscript.

Line 377: the model does not suggest correlation, but perhaps association. Correlation is something you calculate usually with Pearson or Spearman method.

We sincerely thank the reviewer for this important statistical clarification.

We agree that the term *correlation* was inappropriate in the context of regression model coefficients, which quantify conditional associations rather than correlation coefficients in the Pearson or Spearman sense. This section has been completely rewritten. The revised manuscript now analyzes aerosol effects using AOD550 instead of PM2.5, and throughout the Results and Discussion we describe the estimated coefficients as conditional associations within the Bayesian hierarchical model rather than correlations or causal effects. This terminology has been adopted across the revised manuscript.

Lines 384-387: I don't understand the meaning of this section. If you have a multivariable model stating something, why to go back to Pearson correlation? Don't you trust your model result?

We sincerely thank the reviewer for this insightful comment.

We agree that this discussion was not fully consistent with the multivariable modelling framework. Once the joint Bayesian model has been fitted, inference should be based on the estimated conditional associations within that model rather than on separate bivariate Pearson correlations, which do not account for the influence of the remaining predictors. This discussion has been removed. The revised manuscript interprets aerosol effects exclusively through the estimates of the Bayesian Nonlinear Hierarchical Structural Model. It therefore provides a statistically consistent interpretation of the results.

The authors would like to express their sincere appreciation to the reviewer for the time and effort they devoted to evaluating this manuscript. Their constructive comments and valuable suggestions were instrumental in improving the quality of this work.