Dear Editor and anonymous Reviewer,

We express our sincere gratitude for the insightful comments and constructive criticisms on our manuscript titled "*Refining marine net primary production estimates: Advanced uncertainty quantification through probability prediction models*" (MS No.: egusphere-2024-3221). In response to your valuable feedback, we have meticulously revised our manuscript to enhance its clarity, coherence, and overall scientific contribution. Specific modifications have been made to address each point raised by the reviewers, and these are detailed in the subsequent pages, where we provide a point-by-point response to your comments. Reviews' comments are in normal text, whereas our responses are in blue.

This revision process has been a collaborative effort among all co-authors, and we believe that the adjustments made significantly improve the manuscript. We are confident that these changes have addressed your concerns and enriched the manuscript.

With kind regards,

Mengyu Xie (on behalf of all co-authors)

Reviewer #1:

Detailed Review for '*Refining Marine Net Primary Production Estimates: Advanced Uncertainty Quantification through Probability Prediction Models*', Jie Niu et al.

1. Line 26: In the abstract, the source of the NPP estimate (i.e., model output or observation) used in the paper should be mentioned.

We have revised the abstract to explicitly mention the source of the NPP estimate and have provided details about the research location and data used in the study (Lines 28-34).

2. Line 30: The author should explain the nature and the sources of uncertainty in NPP estimates. And why it is important.

We have revised the abstract to clarify the sources and nature of uncertainty in NPP estimates and to emphasize their significance. Specifically, we have included

information about the challenges arising from measurement difficulties, errors in satellite-based inversion, and the need for reliable uncertainty quantification to improve ecosystem management and global carbon cycle modeling (Lines 23 - 28).

3. Line 61-61: It is important to mention the recent study in Satyendranath et al. 2020 (Reconciling models of primary production and photoacclimation, Applied Optics)

We thank the reviewer for highlighting this relevant study. In response, we have incorporated a reference to Satyendranath et al. (2020) into the revised manuscript. Specifically, we have added a sentence to emphasize their contribution to improving primary production models by addressing parameter assignment and its impact on reducing uncertainties (Lines 77 - 80).

4. Line 122: Again, it's important to mention why estimating uncertainty is important?

In the revised manuscript (Lines 129 - 135), we have included sentences to emphasize the significance of uncertainty estimation.

5. Line 137-138: Authors should rephrase “discloses the results” to “discusses the results”.

Corrected (Line 153).

6. Line 167-167: Why are these variables (input features) important in terms of estimating NPP?

We have added detailed explanations in the revised manuscript to clarify the relevance and importance of the input variables for estimating NPP (Lines 182 - 181).

7. Line 164: For PAR, SSP, SH and NPP data, authors should mention direct links for the data they used for experiments.

We have provided direct links to the datasets used in this study (Lines 179-182, Lines 195 - 197).

8. Line 783: Table-1: No need to mention the links here, acronyms are sufficient.

We have corrected Table 1 by removing the dataset links and retaining only the acronyms, as suggested.

9. Line 785: Table-2: Authors should be more clear about the “missing quantity” units i.e., days.

Thank you for your reminder. We have updated Table 2 to include the unit "days" for the “missing quantity” column.

10. Line 187: What specific algorithm was applied to make the time series interpolation.

In our research, we used the ‘interpolate’ function from the Python Pandas library, configured with the 'time' method, to perform the time series interpolation. This approach, while classified as linear interpolation, incorporates the time factor, ensuring that the intervals between timestamps are explicitly considered. This feature enhances its suitability for time series data, particularly datasets with periodic variations like those in our study, enabling more accurate estimation of missing values. Although it is computationally simpler than periodic interpolation methods (e.g., Fourier transform or time series models with seasonal decomposition), the ‘time’ method sufficiently captures the periodicity and variations inherent in our dataset, making it both efficient and effective for this application (Line 224-228).

11. Line 198-216: Authors can drop using “NPP” repeatedly, just the algorithm name is sufficient.

Corrected.

12. Line 208-211: It is not clear why CbPM is negatively correlated with AP. Authors should give an explanation.

Thank you for raising this insightful question. In response, we have elaborated on the relationship between AP and CbPM in the revised manuscript, providing an explanation for the observed negative correlation (Lines 260 - 264).

13. Line 223: Typo in equation number.

Corrected. (Line 275)

14. Line 282: It is not clear whether the author had normalised the input features since they are in different scales.

At the beginning of Section 2.3 of the article, it has been clarified that the input data of different scales have been normalized (Line 341).

15. Line 378: Do the authors have any explanation behind finding the lowest CPRS value than the other models?

In the revised manuscript (Section 3.2.2, Lines 460 – 470, and 476 – 485), we have elaborated on potential factors contributing to the lower CRPS value for the CAFE model, in terms of both variance and cumulative distribution function. Also, Figs. S1 to S6 have been added in the SI to better explain the differences among the training and testing datasets of three NPPs.

16. Line 466-467: Applying a low pass filter on the time series is recommended before reaching this conclusion about long-term trend.

We have applied a low-pass filter to the time series data for the three NPPs to isolate the long-term trends. The filtered results have been included in the revised Figure 3 to visually represent the smoothed trends, ensuring the analysis and conclusions are supported by appropriately processed data (Lines 237 – 243).

17. Line 478-481: Any previous studies (reference papers) that can support the statement about Bayesian model performing better in estimating uncertainty?

Thank you for the comment. In the section introducing the Bayesian method (Lines 270 – 275), we have added citations to relevant literature to support the statement about the Bayesian model's superior performance in estimating uncertainty (lines 274– 275).

18. Line 483: What formula did the authors use to estimate the CDFs?

In the revised manuscript, we have added a detailed explanation of the formula used to estimate CDFs (Lines 396 - 407).

19. Line 486-487: As mentioned in the previous comment, the estimation of Train mean NPP and CAFE NPP curves are not clearly mentioned.

Thank you for highlighting this point. In the revised manuscript, we have clarified that since our models generate probabilistic predictions, the curves presented in some figures represent the mean of these predictions. This clarification has been added in Section 2.3 to ensure transparency regarding the methodology and interpretation of the results (Lines 355 - 360).

20. Line 505 "Small" should be replaced by "lower values" for more clarity.

Corrected. (Line 625)

21. Line 509-515: Test mean NPP lying below at lower values and the alteration at higher values is not appearing very significantly. Also, test mean NPP seems to over-estimate at mid-range but this is not the same as seen in the scatter plot (Fig. 6) where it is almost evenly distributed across either side of the 1:1 line.

We appreciate the reviewer's detailed observation. In response, we have revised the text to provide a clearer explanation of the observed patterns in the CDF curves and

their relationship to the scatter plot (Fig. 6). Additionally, we have clarified the interpretation of the differences between the predicted and true value CDFs and provided insights into potential reasons for these discrepancies (Lines 623 - 625).

22. Fig 10: The curves are difficult to distinguish. Different choice of colours recommended.

We appreciate the reviewer's detailed observation. We have revised the colors in Fig 10 to better present the detailed information clearly. However, the contrast is not significant due to the fact that the predicted means of the two models are closer and the folds in the graph overlap more.

23. Fig 10: Capturing the seasonal cycle is fairly easy as most of the input features contain the same signal. To have a better understanding about how good the models are in reproducing the extreme values, authors should plot the anomaly time series by removing seasonal signals overlayed with observation treated in the same way.

Thank you for highlighting this point. We have drawn anomaly time series plot with seasonal signals removed (Figs. S7 and S8), and compared the ability of two probability prediction models to reproduce extreme values.