

Dear Editor,

We would like to thank the Reviewers for their constructive feedback in this second round of revisions. We have addressed each Reviewer's comments in our responses below.

In this second round of revisions, we have made the following changes:

- Included error plots in Fig. 1
- Changed the gray standard deviation regions around the curves in the sensitivity analyses to be the same color as the curves they represent
- Included mention of Appendix C in the main text
- Made minor corrections (RC1-3) suggested by Referee 2 (Luke Gregor)

We greatly appreciate your time, feedback, and consideration of our manuscript.

Thank you,

Christopher Holder

Author Responses to Comments from Anonymous Referee 1 (Round 2)

As with our previous Author Responses in Round 1, we continue the use abbreviations and colors for ease of reading. In the following responses RC stands for Referee Comment and AR stands for Author Response. Referee Comments are in black-colored font and Author Responses are in red-colored font. Unless specified otherwise, references to figure/table numbers and line numbers refer to those of the revised manuscript (Round 1). When we reference the updated revised manuscript that will be submitted along with these author responses, we refer to it as revised manuscript (Round 2).

RC0: Authors made significant improvements of their manuscript. I am appreciated that the authors took into account my previous comments. The manuscript is more completed now and shows meaningful results.

The title of the manuscript is more appropriate. Also, the introduction and conclusion are well structured, and it is much easier to understand the main goal of this work and its results. Authors conclude that the NNEs can provide an important information on the link between intrinsic and apparent relationships that provide more qualitative information than quantitative. It was not obvious in previous version of the manuscript.

The additional table 1 is very useful, as well as the Appendix B.

The results on figure 7 are very important and interesting. I want to congratulate the authors on this idea. Also new interaction plots help to enhance the main results and give the article completeness.

Thank you for the explanation of why you choose the 25th, 50th and 75th percentiles in your work to test the effect of limitations.

AR0: We want to thank Referee 1 for providing additional comments and feedback in this second round of review.

RC1: In the Eq.9 what do authors mean by min_s and max_s? Is it -1 and 1? Please clarify.

AR1: In the revised manuscript (Round 2), we have clarified that the value of min_s is -1 and the value of max_s is 1.

RC2: It can be better to present the maps of differences on Fig. 1. For example, keep Figure 1a like it is and show the differences on Fig.1 b, c and d between true values and ML methods.

AR2: We do agree with your comment that maps of differences can be informative and have incorporated them as part of Fig. 1. Rather than replace Fig. 1 b, c, and d with contour plots showing the errors, we included the error plots below Fig. 1 b, c, d for each ML method. It occurred to us that although it was clear from the error plots that the NNE and RF had the lowest errors, it still seemed like the original prediction contour plots also helped to demonstrate that these two methods did well at reproducing the original patterns. So rather than remove the original prediction contour plots, we added Log₁₀ Absolute Error plots beneath the prediction contour plots.

We chose to use the Log_{10} Absolute Error metric so all the error plots could use the same color scale. The error values of the RF and NNE compared to the errors of the MLR differed by several orders of magnitude.

Additionally, we revised the text to include mention of the newly included plots.

RC3: On Fig. 2, 4 and 5 it would be useful to add the corresponding colours on the grey areas around curves.

RC3: We updated the grey areas around the curves in Fig. 2, 4, and 5 to be colored according to the color of the line they represent. For example, the grey standard deviation region around the green NNE line in Fig. 2 is now light green instead of grey. We also did this for the grey standard deviation regions around the curves in Fig. 11.

RC4: Fig. 8, 9 and 12 miss a colorbar.

AR4: For the interaction plots in Fig. 8 and 12, the color of the plots corresponds to the biomass concentration. Since the z-axis is already the biomass concentration, we chose not to include a colorbar because we thought it would be redundant. The main purpose of allowing the surface plots to be colored from blue to yellow was to introduce contrast into the figures so we could see and interpret them more effectively.

We did not include colorbars for the contour plots in Fig. 9 and 12 because each contour plot is normalized according to its own probability density function. The plots would have been much less informative if all the contour plots were on the same colorbar scale, since the scale for one contour plot had values of 10^4 and another contour plot had values up to 10^{15} . Additionally, the main purpose of the contour plots was to show the regions with the highest density of observations, so a colorbar did not seem completely necessary and we feared it might distract from the overall figure. However, in lieu of a colorbar we did make sure to state the meaning of the colors in the description of the appropriate figures (blue for few observations up to red for many observations).

RC5: I did not notice any mention of the Appendix C in the text.

AR5: We have included mentions in the main text of the figures in Appendix C in the revised manuscript (Round 2).

Author Responses to Comments from Referee 2 – Luke Gregor (Round 2)

As we did above with our Author Responses to Referee 1, we will continue the use of abbreviations and colors for ease of reading. For details about the abbreviations and colors, along with how we reference the manuscript in our responses, please see the red text directly underneath the heading “Author Responses to Comments from Anonymous Referee 1 (Round 2).”

RC0: The authors have taken great care to implement the many suggestions made by the reviewers in the first round of revision. Thank you for implementing these. As a result, I find that the manuscript is much easier to follow and the presentation of the results (figures and tables) is clear – a fantastic improvement. I would also like to commend the authors for including results that may not appear to be a “success”. I have added my comments below. They are only technical corrections. Suggested insertions are shown as italics.

AR0: We want to thank Luke Gregor (Referee 2) for their additional feedback and comments.

RC1: L241: hours-days → *hours to days*

AR1: We replaced the text as suggested.

RC2: L401: “essentially captured all of it” could be strengthened with a quantitative addition of ($R^2 > 0.99$).

AR2: We revised the text as suggested.

RC3: L487: capitalized “No” should be de-capitalized

AR3: We revised the text as suggested.

RC4: Fig 3, 7, 8: Would it be possible to show the same y-scale for the subplots? A suggestion. If the authors feel that this dilutes the message they are trying to convey (the shape of the curve), then do not change.

AR4: Ideally, we would have all the subplots on the same y-scale. However, as you already correctly pointed out, this makes it difficult to discern the shape of the curve in many of the subplots. Additionally, since the color scale corresponds to the biomass concentration (which is also on the z-axis), the colors of the surface plots become affected as well. These two aspects combined make it difficult to interpret the plots when they are on the same y-scale. For example, when we tried to have the subplot of Fig. 3 n on the same y-scale as the subplot of Fig. 3 m, the surface of Fig. 3 n appeared as a mostly flat blue plane since its biomass values were much lower than the biomass values in Fig. 3 m.

RC5: Fig7: A comment that might be useful for future experiments. The underestimation of biomass is the largest for irradiance. Could using the daytime-equivalent irradiance improve estimates? (i.e., do not include nighttime for averaging)

AR5: That is an excellent suggestion and one that we will consider for future experiments. It is reasonable to assume that using the daytime-equivalent would increase the average since all the zero values from nighttime would not be considered in the averaging.

Author Responses to Comments from Anonymous Referee 3 (Round 2)

Referee 3 selected the option “accepted as is” for the revised manuscript (Round 1), so no additional comments or feedback were provided by Referee 3. We want to thank Referee 3 for taking the time to review our manuscript and for their helpful feedback and comments from the first round of revisions.