Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-262-AC1, 2020

Biogeosciences Discussions
Open Access
EGU

Interactive comment on "Linking intrinsic and apparent relationships between phytoplankton and environmental forcings using machine learning – What are the challenges?" *by* Christopher Holder and Anand Gnanadesikan

Christopher Holder and Anand Gnanadesikan

cholder2@jh.edu

Received and published: 24 September 2020

Author responses to Anonymous Referee 1

In the following responses RC stands for Referee Comment and AR stands for Author Response. For sections where draft paragraphs for the revised manuscript are included, the beginning and end of the draft paragraphs are denoted with *BD* (Begin Draft) and *ED* (End Draft).

For detailed descriptions of the tables and figures included with this Author Response,

Printer-friendly version

Discussion paper

please see the Supplemental PDF included with this Author Response.

RC0: The authors try to find a ML method that can establish and help to explain the link between intrinsic and apparent relationships of phytoplankton and environmental forcing. Three different methods were tested: Multiple Linear Regression (MLR), Random Forests (RFs) and Neural Network Ensembles (NNEs). The tests were provided on three different Scenarios. The authors found that the NNEs reproduce well the observed biomass (biomass estimated from intrinsic relationships) based on the knowledge of only apparent relationships when both relationships operate on the same spatial and temporal timescales. All methods fail when the intrinsic and apparent relationships operated on different timescales. However, the main authors' conclusion is that ML methods still can give an information on shapes of intrinsic relationships. Using the Earth System Models (ESM) in their third Scenario they show that the ML methods can extract useful information from this model that can used for an examination of interaction between input variables.

The article arises an interesting subject. However, it misses a clear explanation how the trained data and data for validation were constructed; the explanation of the role of input data, especially the physical meaning of the choice to fix them at 25th, 50th and 75th percentile for a sensitive analysis. I also think that the authors did not use all possible ML capacities for example, test more hidden layers in NNs, or add few more input variables, environmental ones, maybe a temperature.

Also, the main results that the NNEs can reproduce the general shape of intrinsic relationships does not have a systematic character: it is clearly seen that in some cases the NNEs shapes reproduce a different behavior compare to observation data. If the authors want to keep this conclusion, they will have to provide an additional analysis of conditions under which their conclusion is persistent. The article in its current state needs a serious major correction.

AR0: We want to thank Anonymous Referee #1. We have found the comments and

Printer-friendly version

Discussion paper

suggestions they provided to be very helpful in restructuring this manuscript.

RC1: It is hard to understand reading the introduction what is the main aim of the article. I have found three:

AR1: We understand how some of these statements may have caused confusion as to the main purpose of the paper. We address these in the following author responses AR1.1, AR1.2, AR1.3, and AR1.4.

RC1.1: "A significant challenge that remains is determining how intrinsic relationships found in the laboratory scale up to the apparent relationships observed at the ecosystem scale (i.e., scaling the small to the large)."

AR1.1: It was our intent that this served as more of a big-picture gap that has yet to be filled/solved. With this serving as the big-picture statement, it serves as a leadup to rest of the introduction. In this paper, we aren't necessarily trying to answer the entirety of this big-picture statement; rather, we are attempting to answer part of it by examining some smaller aspects of it.

RC1.2: "What is less clear is: 1. Can robust relationships be found? 2. If so, what methods are most skillful in finding them? 3. How do you interpret the apparent relationships that emerge when they diverge from the intrinsic relationships we expect?"

AR1.2: This statement was intended to serve as a link between the two paragraphs. However, we can see the confusion, especially given that the questions are numbered which would typically signify importance. We have removed this statement in its original paragraph and have incorporated the questions into the third statement.

RC1.3: "To investigate when and why the link between intrinsic and apparent relationships break. . ."

AR1.3: We have incorporated the questions from the second statement into the third statement to highlight the main purposes of the paper. The modified structure of the introduction that will be included in the revised manuscript as indicated in the draft

below in AR1.4.

RC1.4: If the two first citations can be linked to the title “What are the challenges?”, the last one, what the authors actually did, is not according to the title “What are the challenges?”. I will advise to authors modify the structure of the introduction and emphasize what is the main idea of this article.

AR1.4: We have changed the title of the manuscript to be more reflective of the main points. The current draft title replacement is “Can machine learning extract the mechanisms controlling phytoplankton growth from large-scale observations? – A proof of concept study”

Additionally, a modified introduction will be included in the revised manuscript. The portion of the introduction that highlights the main ideas of the article is in draft form below for reference:

BD To investigate when and why the link between intrinsic and apparent relationships break, we try to answer two main questions in this paper: 1. Can ML techniques find the correct underlying intrinsic relationships and, if so, what methods are most skillful in finding them? 2. How do you interpret the apparent relationships that emerge when they diverge from the intrinsic relationships we expect?

In addressing the first question, we first needed to demonstrate that we had an ML method that would correctly extract intrinsic relationships from apparent relationships. We constructed a simple model in which the intrinsic and apparent relationships operated on the same time and spatial scale and were only separated by a scaling factor, but in which the environmental drivers had realistic inter-relationships. Having a better handle on the results from the first question, we were able to move onto the second question where we look at where the link between intrinsic and apparent relationships break. We modified the first scenario to allow the intrinsic and apparent relationships to operate on different timescales – allowing us to evaluate the impact of time-averaging on the retrieval of intrinsic relationships. Finally, we conduct a proof-of-concept study

Printer-friendly version

Discussion paper

with real output from an ESM. *ED*

RC2: The conclusion also has to be rewritten in clear way: what are the challenges authors have found to link intrinsic and apparent relationships? The authors said: “Our main objective in this manuscript was to use ML to determine under what conditions intrinsic and apparent relationships between phytoplankton are no longer equal. . .” This objective was not clear in the introduction, and again, does not correspond well to the title of the article.

AR2: The revised manuscript will include a modified Conclusion section to better highlight the main objectives and points of the paper. A draft version of the new title and objectives are listed in AR1.4.

RC3: I advise you also to avoid the non-explained abbreviations in the abstract, like line 21: “ESM”, line 28: “MLR”.

AR3: We have removed the non-explained abbreviations from the abstract. They have been replaced with their unabbreviated definitions. This will be corrected in the revised manuscript.

RC4: It would be better not mentioning the results in the introduction (lines 100-103), and instead prepare your readers for the structure of the article.

AR4: We agree with this comment. The lines mentioning the results in the introduction have been removed. This will be corrected in the revised manuscript.

RC5: In the section “2. Methods” I suggest you provide a scheme or formula what exactly you were using as input/predictors and output/target and how it links to your equations 1 and 2 etc. This will simplify the understanding. It is especially important for the description of the Scenario 2. It is unclear, are the values of biomass with which the authors compare their results calculated based on hourly values? Is the biomass for target in learning algorithm smoothed or calculated based on smoothed predictors? It will strongly affect the results.

AR5: A new table will be included in the revised manuscript. The draft version is included (Table 1) which provides the following details for each scenario: the predictor variables, the target variable, the equations used in each scenario, a description of the source file, and a description of the scenario, such as how the target biomass was calculated.

RC6: The word “target” was not used in your article. It is common word in the ML domain and can also help to better understand the method especially for readers who only start to use ML techniques. Also, I would suggest to use the word “validation” instead of “testing”.

AR6: We agree with this comment and have replaced the term “response variable(s)” with the term “target variable(s)”. Additionally, we have replaced the term “testing subset(s)” with the term “validation subset(s)”. This correction will be reflected in the revised manuscript.

RC7: To perform a sensitive analysis the authors fixed two of three predictors at different percentiles. It misses the explanation why 25th, 50th and 75th percentiles were chosen, what is the physical meaning of this choice and how it influences the results? For example, does 75th percentiles represent the extremes and what does it mean? I think that it is important to explain it to better understand the results. Please, clarify that the sensitivity analysis was done already on trained ML model.

AR7: The main reason for choosing these particular percentile values was so that we could examine conditions in a domain space that may be found as an actual observation. Additionally, we wanted to avoid extreme percentiles (1st and 99th) because the standard deviation in the predictions of the trained ML models is very large at these extremes.

We have also clarified that the sensitivity analysis was performed on the trained ML models. The revised version of the manuscript will include these changes, along with more details about how the sensitivity analysis was performed and explain the reasons

Printer-friendly version

Discussion paper

for why we chose particular percentile values.

A draft rewrite of the portion of the paper describing the sensitivity analyses is listed below:

BD Following this, a sensitivity analysis was performed on the trained ML models. We allowed one predictor to vary across its min-max range while holding the other two input variables at specific percentile values. This was repeated for each predictor. This allowed us to isolate the impact of each predictor on the biomass – creating "cross-sections" of the dataset where only one variable changes. For comparison, these values were also run through Eq. 1 and 2 to calculate the true intrinsic response of how the simple phytoplankton model would behave. This allowed us to view which of the models most closely reproduced the underlying intrinsic relationships of the simple phytoplankton model.

For the sensitivity analysis, we chose to hold the predictors that weren't being varied at their respective 25th, 50th, and 75th percentile values. We chose to use these percentile values for several reasons:

1. It allows us to avoid the extreme percentiles (1st and 99th). As we approach these extremes, the uncertainty in the predictions grows quite rapidly because of the lack of training samples within that domain space of the dataset. For example, there are no observations which satisfy the conditions of being in the 99th percentile of two variables simultaneously. This extreme distance outside of the training domain leads to standard deviations in predictions that are too large to provide a substantial level of certainty about the ML model's predictions.

2. Similar to the idea that we can avoid the extremes, we also chose these values as they are quite typical values for the edges of box plots. Generally, values within the range of the 25th to 75th percentiles are not considered outliers. Along those lines, we wanted to examine the conditions in a domain space that might be found in an actual observation, with the reasoning that if there was high uncertainty in the ML predictions

at these more moderate levels, there would be even higher uncertainty towards the extremes. *ED*

RC8: Did the authors try to increase the number of hidden layers in their NNs? It is known that the introduction of more hidden layers can improve the results. It would be interesting to see if there is any effect from the number of hidden layers in this particular problem.

AR8: In responding to this particular comment, we are assuming the Referee is referring to the outcomes of Scenario 2 since the performance in Scenario 1 was already robust.

For our manuscript we chose to use single hidden layer neural networks with 25 nodes in the hidden layer for several reasons:

1. We ran a diagnostic test to determine the ideal number of nodes for each neural network in Scenario 2 (Table 2). From the R2 values in Table 2, it can be observed that the performance of each NNE begins to plateau around 10-15 nodes. However, the inclusion of additional nodes up to about 25 nodes did not cause the training times to increase drastically. We chose to include these extra nodes due to their moderate training times and robust performance.

2. All three Scenarios showed ideal performances in the 10-20 node range, but similar to the first point, we chose to include extra nodes due to their moderate training times and good performance. Additionally, this allowed us to be consistent in the architecture of the neural networks across all the Scenarios and helped to minimize the differences between them.

3. We considered adding more hidden layers to the neural networks (Table 3), but the slight increase in performance did not seem worth while due to significantly increased training times. As can be observed in Table 3, the performance did not significantly increase from a single hidden layer to when a second layer with additional nodes was

Printer-friendly version

Discussion paper

included (0.9700 vs 0.9722-0.9726).

4. Although some applications may require additional hidden layers, it is generally accepted that one hidden layer can approximate most functions provided there are sufficient nodes in the hidden layer and the nodes utilize squashing functions, such as sigmoid functions (Hornik et al., 1989).

Hornik, K., Stinchcombe, M. and White, H.: Multilayer feedforward networks are universal approximators, Neural Networks, 2(5), 359–366, doi:10.1016/0893-6080(89)90020-8, 1989.

RC9: The authors did not provide how they scaled their variables (lines 337-341). This procedure is known as normalization of variables. Normalization ensures that all predictors fall within a comparable range and avoids giving more weight to predictors with large variability ranges.

AR9: We have included the equations that we used to scale and unscale the values of each variable (Eq. 1 and 2; attached to this Author Response). The revised version of the manuscript will include these equations, along with the details pointed out by Referee 1 in RC9. A draft version for the revised manuscript is below:

BD Each variable was scaled between -1 and 1 based on its respective maximum and minimum (Eq. 1).

(See Eq. 1 below)

Where V is the value of the variable being scaled, S stands for the scaled value, and U represents the unscaled value. This step ensures that no values are too close to the limits of the hyperbolic tangent sigmoid activation function, which would significantly increase the training time of each NN. Additionally, this normalization ensures that each predictor falls within a similar range so more weight is not provided to variables with larger ranges. These scalings were also applied to the RF and MLR methods for consistency between methods which did not affect the results of either method (results

not shown). The results presented in this paper were then transformed back to their original scales to avoid confusion from scaling (Eq. 2).

(See Eq. 2 below)

Where the letters represent the same values as those in Eq. 1. *ED*

RC10: It is hard to agree that the NNEs and RFs represent well the behavior as for example on Figure 3 in left column the NNEs show a strong increase at the end of the macronutrient range that does not in agreement with the observation values; and there is a false decrease in the middle column at 25th percentile.

AR10: These behaviors likely have several causes:

1. In the areas of the dataspace where there are few or no observations, the NNEs are less certain about the predictions. This can be noted as the large gray areas in the middle column for the 25th percentile in Fig. 3. Additionally, the uncertainties in the previously mentioned subplot may seem large due to the magnitude of the predictions for those conditions. For example, the y-axis of the middle column for the 25th percentile of Fig. 3, shows the biomass in the range of 10-7. The other percentile plots in the middle column of Fig. 3 show biomass in the range of 10-6, a full order of magnitude larger than the biomass in the 25th percentile of the same column.

2. Since the lines in the sensitivity analyses for the NNEs are the average response of ten individual neural networks, it's possible that the NNE line in the sensitivity analyses could be pulled higher or lower in areas of higher uncertainty due to extreme outlying predictions. This might explain the behavior of the increase in biomass at the end of the macronutrient range in the 75th percentile subplot of the left column in Fig. 3. Additionally, since the random forest is an average of 500 trees it may be less susceptible to extreme outlying predictions, which would help to explain why that behavior is not seen in the random forest predictions for the previously mentioned macronutrient subplot.

We have only listed some of the potential causes, so others may exist to explain the

deviations in behavior. We will address these in more detail in the revised manuscript.

RC11: It is interesting to know if the authors have an idea about new parameter or variable that can bring back the information on hourly variability lost due to the time-averaging to improve the results?

AR11: We have considered a couple of ideas to attempt to bring back the variability lost due to the time-averaging:

1. We implemented a method that allowed us to visualize how the apparent relationships change from the hourly timescale through to the monthly averaged timescale.

To capture the apparent relationships at various timescales between the hourly and monthly timescales, we averaged the hourly dataset over a range of hourly timespans. Specifically, we averaged over timescales including 1 hour (original hourly set), 2, 3, 4, 6, 8, 12, 24, 48, 72, 168 (weekly), and 720 (monthly) hours. It was necessary to average over timespans that were multiples of, or divisible by, 24. Sets of hours that did not meet this criteria meant that hours from one day were ultimately being averaged with hours from another day. For example, using a 7 hour timespan for averaging would mean that the last three hours of Day 1 were then being averaged with the first four hours of Day 2.

This new set of averaged timescales was then used to train NNEs with one NNE corresponding to each averaged timescale. To be consistent, the NNEs were trained according to the same specifications listed in the original manuscript (10 individual NNs in each NNE, 25 nodes in hidden layer, one hidden layer, three predictors, stopping criteria, etc.).

We then performed sensitivity analyses on each of the trained NNEs to see the apparent relationships for each averaged timescale and set the percentile values for the other variables at their 50th percentile (median). To visualize all the timescales at once, we plotted them on contour plots (Fig. 1). The left three plots (Fig. 1 a, c, e) show the

full span on the timescales we tested. The greatest changes in the apparent relationships were in the first 24 hours, which is difficult to see because of the difference in magnitude of the hourly and monthly averaged timescales. Since there is little change in the apparent relationships past the 24-hour averaged dataset, we decided to examine the relationships of the timescales at and below 24 hours (Fig. 1 b, d, f). Again, what we are seeing is that the more the dataset is averaged, the more variability is lost, and the greater the underestimation of the relationships. However, what we also see is that many of the apparent relationships below 12 hours are fairly close to the 1-hour apparent relationships. It may be possible that one may be able to use this type of diagnostics test to know learn the necessary frequencies to recover the true relationships.

2. We included additional variables and tested different percentile values.

We included the length of day (the number of hours in a day that a particular location received sunlight above 0 W m-2) as a predictor for each of the time-averaged datasets of Scenario 2 and trained a new set of NNEs. With the exception of the number of predictors, the NNEs followed the same specifications listed in the original manuscript (10 individual NNs in each NNE, 25 nodes in hidden layer, one hidden layer, stopping criteria, etc.).

We knew from previously testing the higher percentile values (>90th percentile) for the sensitivity analyses that the higher percentiles led to greater uncertainty in the predictions (Fig. 2). Although the true relationships were captured in the 100th percentile plots, insofar as they were within the gray standard deviations of predictions, the uncertainty in the predictions had a range equal to or greater than the response (Fig. 2 g, h, i).

When we included length of day as a predictor and used the high percentile values of the other variables, we found that the 100th percentile plots captured the true relationship and the gray standard deviation ranges decreased (Fig. 3 i, j, k, l). Furthermore,

Printer-friendly version

Discussion paper

all of three of the time-averaged datasets were able to capture the correct relationships.

We will include this new result, with some discussion, in the revised manuscript.

RC12: The article misses the total results for Scenarios 2 and 3 like on Figure 1.

AR12: The contour plots of Figure 1 were mostly intended to provide a visual representation of the difference in performance of the ML methods compared to multiple linear regression. Table 1 serves the same purpose Figure 1, but with performance metrics instead visual representation. Additionally, Table 1 provides quantifiable measures of each method's performance, whereas Figure 1 only allows for a visual qualitative measure. Since Table 1 serves as a quantifiable metric of performance we chose to use only tables for Scenarios 2 and 3, instead of contour plots like those in Figure 1.

RC13: In the Discussion of results for Scenario 3 the authors reasoned about BLING model behavior and did not mention their results. It would be interesting to know the authors' thoughts about NNEs behavior on Figure 7 middle column at 75th percentile.

AR13: From our best estimation, we assume that Referee 1 is indicating that we do not mention the specific results of any one particular ML method. For example, we didn't differentiate between the responses found by RFs and NNEs.

In the Discussion section of Scenario 1, we mention that RF does not possess the same extrapolation capabilities as NNEs. Because of RF's inability to extrapolate, we focused our discussion of Scenario 3 on the relationships found by NNEs. However, it is now clear to us that we should either specify that point in the Discussion of Scenario 3 or discuss the relationships found by each method. The revised version of the manuscript will include the previously mentioned specification or will discuss the relationships found by each individual method.

RC14: I have also found that the authors wrote too much on things that they did not do: lines 171-177.

AR14: We originally included this section because partial dependence plots (PDP)

are a common visualization technique used in ML, especially for random forest (RF) analysis. In previous versions of this manuscript, initial comments (before submission to any journal) typically cited the use of PDPs vs sensitivity analyses as being an important distinction. So we felt it necessary to provide a brief explanation of how the use of sensitivity analyses differs from PDPs.

RC15: Lines 260 – 273: I suggest to rewrite these two paragraphs, there is a mention of results that have not yet been presented; and I do not see the necessity to mention that the authors "previously had little experience with ML."

AR15: We understand the confusion with these paragraphs. In past submissions, we were asked several questions:

1. Why had we chosen to use Random Forests and Neural Network Ensembles, instead of other machine learning algorithms? 2. Why did we choose to compare the performance of the machine learning algorithms to Multiple Linear Regression?

We tried to address these questions with these two paragraphs (lines 260-273), but we acknowledge that we may have included additional details that may not be needed. To address these two questions and still include details we find necessary, we have shortened the first of the two paragraphs mentioned in the Referee Comment. This condensed version will be included in the revised manuscript. A draft of the condensed paragraph is included below:

BD We chose to use Random Forests (RFs) and Neural Network Ensembles (NNEs) in this manuscript. Although other ML methods exist, the list of possible choices is rather long. It was decided that the number of ML algorithms being compared would be limited to RFs and NNEs given their popularity in studying ecological systems. Additionally, we chose to compare the performance of the ML techniques to the performance of Multiple Linear Regression which allows for a comparison of linear and non-linear methods. *ED*

Printer-friendly version

Discussion paper

Regarding the second paragraph, comments by other Referees have suggested the second paragraph (lines 270 to 273) remain unaltered. It provides an explicit statement by us that we are not trying to invalidate results or discourage the use of MLR in marine ecological systems.

RC16: Please use the figures captions like “a”, “b” etc.

AR16: All figures that have subplots now have letter captions for each of the subplots. Additionally, the figure descriptions will refer to the letter captions for subplots instead of their location within each figure. For example when referring to a subplot, it is cited as “Box a” instead of the “top-left subplot in Fig. XX”. These corrections will be reflected in the revised manuscript.

RC17: Please, expand the figure captions, for example on Figure 2 it would be good to add that the black line is estimated from Eq. 1.

AR17: The figure and table captions will be expanded in the revised version of the manuscript to include more specific details. A draft version for Figure 2 is included below as an example.

BD Figure 2: Sensitivity analysis for Scenario 1 showing the true and predicted relationships for how each predictor affects the biomass when the other predictors are set at specific percentiles. The columns correspond to the predictors and the rows correspond with the percentile value at which the other predictors were set. The black line shows the true intrinsic relationship calculated from Eq. 1 and 2. The dashed lines show the predicted apparent relationships for each method (MLR – red; RF – blue; NNE – green). The gray region around the RF and NNE dashed lines shows the standard deviation of the predictions. *ED*

RC18: Figure 5 was mentioned as the last but it is placed before Figure 6 and 7.

AR18: The order of the figures will be corrected in the revised manuscript so each figure is numbered based upon when it is first mentioned in the manuscript.

RC19: Line 149, 156, 168, 356: please avoid to use the sign “ in scientific paper.

AR19: We removed the quotations around the words in the specified lines. Additionally, we have reexamined our usage of other quotation marks throughout the paper and have removed quotations we deemed unnecessary. This will be reflected in the revised manuscript.

RC20: Line 149: sign “ should be before the point .

AR20: We removed the quotations around that particular term. This will be reflected in the revised manuscript.

RC21: Line 206: it feels that “but” should be replaced by “and”.

AR21: We replaced the transition word “but” with the transition word “and”. This change will be reflected in the revised manuscript.

RC22: Line 226, 235, 543: word “just” is unnecessary.

AR22: We removed the word “just” from the suggested lines. This will be reflected in the revised manuscript.

RC23: Line 250: “an ML” should be replaced by “a ML”.

AR23: We replaced all instances of “an ML” with “a ML”. This will be reflected in the revised manuscript.

RC24: Line 356: Miss a figure indication.

AR24: We filled in the missing figure indication. This change will be reflected in the revised manuscript.

Please also note the supplement to this comment:
https://bg.copernicus.org/preprints/bg-2020-262/bg-2020-262-AC1-supplement.pdf

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-262, 2020.

$$V_S = \frac{max_S - min_S}{max_U - min_U}(V_U - min_U) + min_S \quad (1)$$

Fig. 1. Equation 1. Equation used to scale each variable between -1 and 1.

Printer-friendly version

Discussion paper

$$V_U = \frac{max_U - min_U}{max_S - min_S}(V_S - min_S) + min_U \quad (2)$$

Fig. 2. Equation 2. Equation used to scale each variable back to its original values.

Scenario	Predictors	Target	Equations Used	Source File Description	Scenario Description
1	Macronutrient (mol kg^{-1}); Micronutrient (mol kg^{-1}); Irradiance (W m^{-2})	Biomass (mol kg^{-1})	1, 2	Monthly Output from BLING	Nutrient distributions (predictors) from BLING were fed to Eq. 1 and 2 to calculate the biomass (target)
2	Macronutrient (mol kg^{-1}); Micronutrient (mol kg^{-1}); Irradiance (W m^{-2})	Biomass (mol kg^{-1})	1, 2, 5	Daily Output from BLING	1) Hourly values for the predictors were interpolated using the Daily Output of BLING 1a) The macronutrient and micronutrient hourly values were calculated using a standard interpolation between the daily points. 1b) The irradiance hourly values were calculated from Eq. 5 using the value of the BLING daily input, hour of day, time of year, and location. 2) Hourly values of the predictors were fed to Eq. 1 and 2 to calculate hourly values for the biomass (target) 3) Daily-averaged values were calculated by averaging 24 hours for each location through one year 4) Weekly-averaged values were calculated by averaging 168 hour blocks of time for each location through the year 5) Monthly-averaged values were calculated by averaging the number of hours in each month (days per month * 24) for each location through the year 6) The true relationships were calculated by using the range of the hourly values for the predictors and calculating the biomass based on Eq. 1 and 2.
3	Macronutrient (mol kg^{-1}); Micronutrient (mol kg^{-1}); Irradiance (W m^{-2})	Biomass (mol kg^{-1})	6, 7 (Equations within BLING used to determine the biomass)	Monthly Output from BLING	Nutrient distributions from the BLING Output were used as the predictors; Biomass from the BLING Output itself was used as the target

Fig. 3. Table 1: Details for each Scenario.

		R^2 Values		
		Daily	Weekly	Monthly
Number of Nodes	1	0.5533	0.5472	0.5624
	2	0.7655	0.7705	0.7806
	5	0.9283	0.9248	0.9363
	10	0.9633	0.9628	0.9673
	15	0.9676	0.9678	0.9713
	20	0.9693	0.9694	0.9727
	25	0.9700	0.9702	0.9732
	35	0.9709	0.9709	0.9737
	50	0.9716	0.9715	0.9743

Fig. 4. Table 2: The R2 values for the diagnostic test used to determine the how the number of nodes in the hidden layer of a single layer neural network affected the performance of the Scenario 2 datasets.

		R^2 Values		
		Daily	Weekly	Monthly
Layers and Number of Nodes	25	0.9700	0.9702	0.9732
	25-10	0.9722	0.9724	0.9750
	25-25	0.9726	0.9727	0.9756

Fig. 5. Table 3: The R2 values for the diagnostic test used to determine the how the number of hidden layers and nodes within individual neural networks affected the performance of the Scenario 2 datasets.

Printer-friendly version

Discussion paper

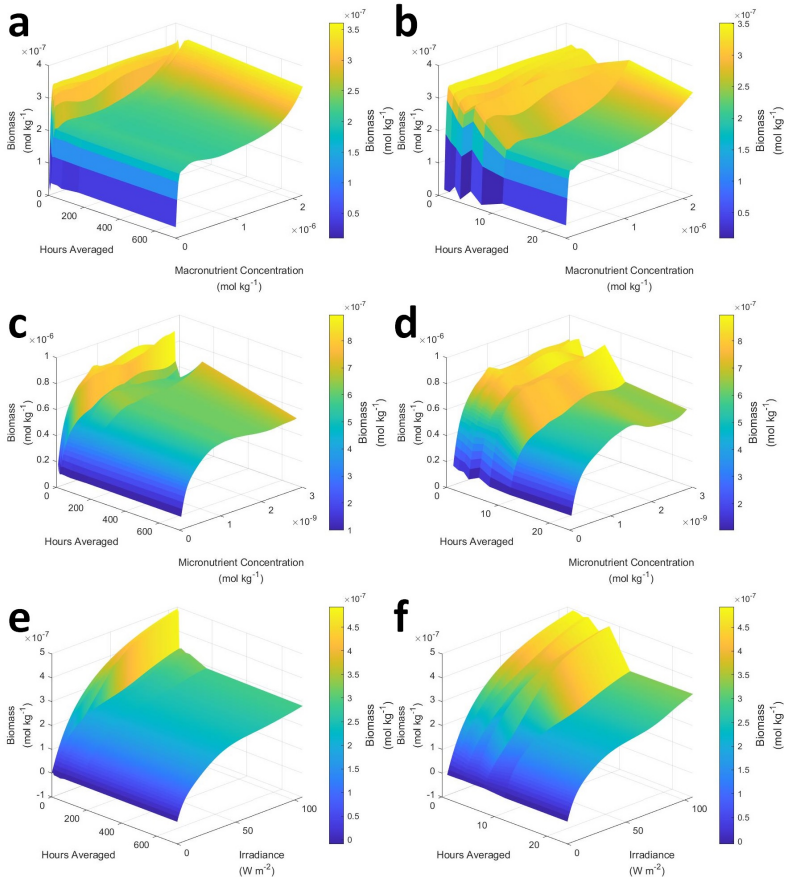

Fig. 6. Figure 1: Contour plots showing the apparent relationships found across different averaged timescales for Scenario 2.

Printer-friendly version

Discussion paper

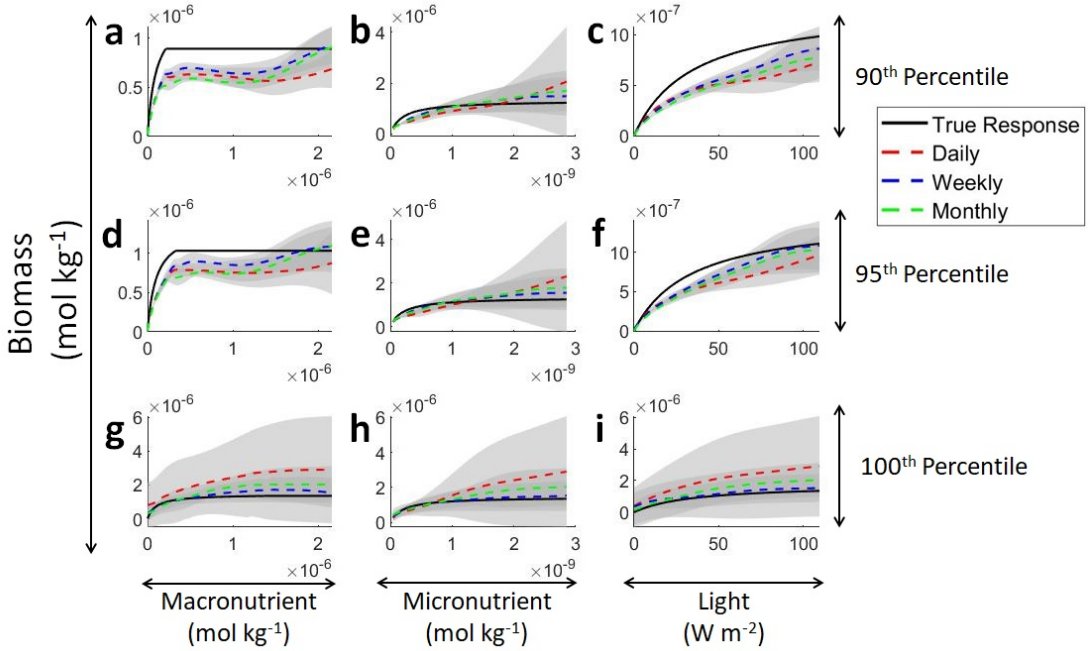

Fig. 7. Figure 2: Sensitivity analysis for Scenario 2 showing the true and predicted relationships for how each predictor affects the biomass when the other predictors are set at higher percentiles.

Printer-friendly version

Discussion paper

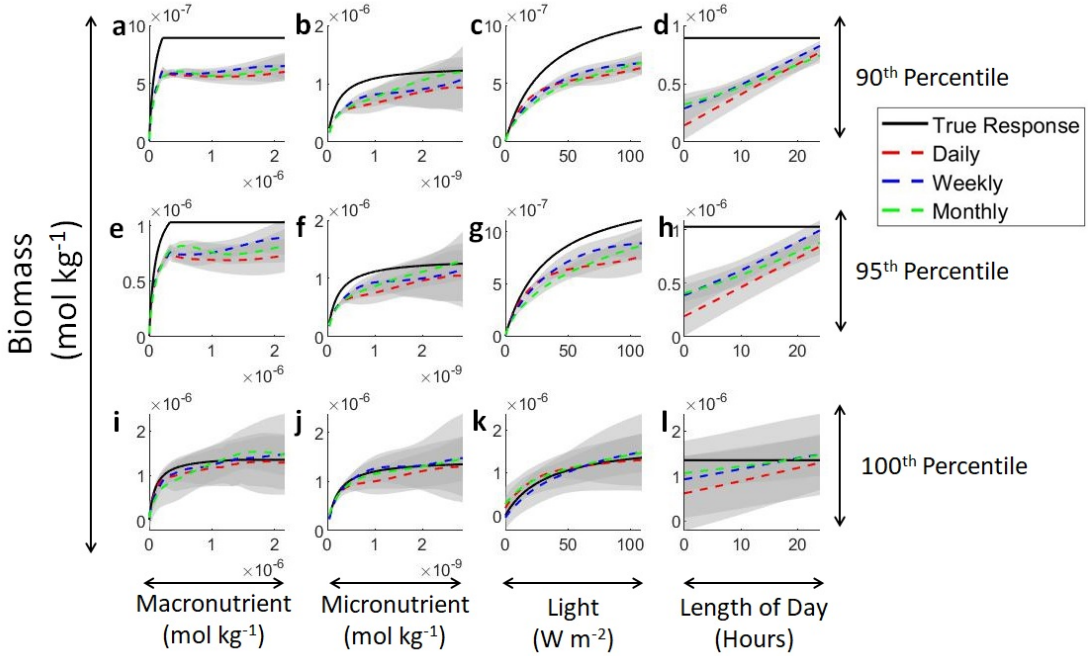

Fig. 8. Figure 3: Sensitivity analysis for Scenario 2 with Length of Day as an added predictor showing the true and predicted relationships for how each predictor affects the biomass using high percentiles.

Printer-friendly version

Discussion paper