

***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**

Luke Gregor (Referee)

luke.gregor@usys.ethz.ch

Received and published: 31 August 2020

1 Overview

The study by Holder and Gnanadesikan tries to assess if machine learning is able to extract the intrinsic relationship between phytoplankton growth and limiting nutrients and light from observed concentrations of nutrients and light intensity. This topic was investigated with three experiments of increasing complexity asking the following ques-

[Printer-friendly version](#)

[Discussion paper](#)



tions (with my brief understanding of the outcomes):

1. Are ML methods able to extract the relationship from observations at all at instantaneous time scales?
Yes, but NNE is better at extracting the relationship than RF despite both achieving fair results
2. If time scales are averaged, can the relationships still be extracted?
Not very well. In most cases the estimated half-saturation is lower than it should be. I.e. even the better of the two ML methods, NNE, is not very accurate.
3. Can the approach work in a more complex model setup where biomass losses are also accounted for?

While I appreciate the question the study is asking and think that this work is important, I found that the manuscript was not very easy to follow (my summaries of the results above might illustrate this). Part of the difficulty may be that the topic is not within my immediate field of expertise, but then I feel there are stylistic changes to be made that will improve the manuscript. I have overall comments in the document below and I linked a **PDF document with comments at the very end of this document** (I used Adobe Reader). I hope these comments help improve the flow of the manuscript.

2 Title

The title can be improved. To someone who is not familiar with the “intrinsic” and “apparent” terminology, the title is not informative. Something along the lines of : *Can machine learning extract the mechanisms of phytoplankton growth from large-scale observations?*

BGD

Interactive
comment

Printer-friendly version

Discussion paper



3 Abstract

The use of “intrinsic” and “apparent” relationships this early in the manuscript made it difficult to understand the study as I am not familiar with the terminology.

4 Introduction

I don't have major concerns with the introduction and it builds a good case for why this study is relevant.

The questions posed (L72-75) and ideas presented (L100-102) are useful in framing the study but are not carried clearly through the manuscript. It would be very useful for the reader to have these questions and ideas as a guide for why each experiment was performed. For example, L72-75 from the basis of experiment 1, but these questions are not explicitly answered in the discussion. And lines 100-102 form the basis of the design for experiment 2.

5 Methods

Regarding the structure of the methods section, I have the following suggestions:

- There is no overview of the methods. I think this would be useful in addition to an accompanying diagram outlining all the experiments and the use of the machine learning approaches used. It would help the reader understand the flow of the study.
- BLING is used throughout the study, albeit with different outputs from the model,

BGD

Interactive
comment

Printer-friendly version

Discussion paper



but it may make sense to introduce the model before the experiment configurations are described.

- It would make sense to formalise the following structure for each experiment:
 - A brief introduction to the experiment
 - HEADING for data
 - HEADING for Machine learning parameterisation / application

From a methodological point of view

- In experiment 2, the authors create hourly data by simulating variability of light conditions. The data are then averaged again to create daily, weekly and monthly data. If I understand correctly, the hourly data is analogous to the data used in experiment 1 - i.e. there is no temporal averaging in the “apparent data”. It would be much more methodologically consistent to use the hourly data in experiment 1 and easier for the reader to follow. Either, the authors should implement this, or should make this explicit and state the reason that a separate experiment is needed.
- Another question is regarding the model: what is the variability of the nutrients at a daily resolution (native model resolution), and the averaged resolutions (weekly, monthly). Show some violin/box plots for the normalised data.
- I still don't fully understand what the predictors and target variables are for each experiment and what is the role of the intrinsic? From what I understand, predictors are always the “apparent” data and biomass is the target. The intrinsic is what describes the relationship between the biomass and the “apparent data”. Please make this more clear. Addressing the points above in the structure section will help with this.

[Printer-friendly version](#)

[Discussion paper](#)



6 Results

- The authors should only show NNE results for experiment 2 (figure 4). Is there a reason for this? My presumption is that the intrinsic relationship estimated by RF for micronutrients is poor, thus only NNE is shown. This should be cleared up (unless I missed this).
- From what I understand, the half-saturation constants are the metric for whether the method is able to capture the intrinsic from the apparent. Make this much more clear - also in the abstract

7 Discussion

- The subheadings could be the questions posed in the introduction (see my previous comments on this section). This would help guide the reader
- I think the authors should make the point that given the simplicity of the definition of biomass, one would expect the ML methods to perfectly represent the Michaelis-Menten curves. The authors do correctly state that RF is less likely to estimate accurately as the method is not able to extrapolate. This then increases the importance of showing the distributions of the training and test data set distributions. A further comment: what is the envelope around the estimated curves and why is there a large variability for the NNE at larger values?
- The discussion around scenario/experiment 3 is not clear and I don't feel that there is a take-home message after reading this section.

BGD

Interactive
comment

Printer-friendly version

Discussion paper



8 Tables and Figures

The captions are not standalone for both figures and tables.

The reader needs to know what the target variable in each table is and there are no units.

What is the envelope around the dashed lines.

Please also note the supplement to this comment:

<https://bg.copernicus.org/preprints/bg-2020-262/bg-2020-262-RC2-supplement.pdf>

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

BGD

Interactive
comment

Printer-friendly version

Discussion paper

