Biogeosciences Discuss., 10, C4265–C4268, 2013
www.biogeosciences-discuss.net/10/C4265/2013/

Biogeosciences
Discussions
Open Access

Interactive comment on "Distribution of phytoplankton functional types in high-nitrate low-chlorophyll waters in a new diagnostic ecological indicator model" *by* A. P. Palacz et al.

A. P. Palacz et al.

arpa@aqua.dtu.dk

Received and published: 15 August 2013

Here we provide a detailed response (in black) to Referee's comments (in blue italics) on our original manuscript. We are grateful to the reviewer for the constructive criticism which led to an improved version of this manuscript.

"The driving forces/factors which can alter the phytoplankton functional type are complicated. It depends on a lot of factors including the surrounding ecological systems. The authors provides SST, PAR, Wind speed, MLD and Chl as input variables. In my mind, even if the SST, PAR, wind speed and MLD are the same, the Chl maybe different, especially in different areas, so as the phytoplankton functional types."

We agree with the reviewer and would like to point out that this is exactly what the PhytoANN was designed to capture. Chl as an indicator has been assigned a relatively strong weight and it will be used to distinguish between PFT biomass levels based on the difference in Chl alone even when all other indicators are equal.

"The authors gave a reference which also used the ecological indicators to estimate four phytoplankton functional types but that algorithm is limited to one study area and the latitude, longitude and the time are considered as inputs. Have the authors tried to add the latitude, longitude or time as inputs to see if the model performs better?"

We agree with the reviewer that latitude, longitude and time would likely further increase the accuracy of the model as was previously shown, for instance in Raitsos et al. (2008). However, we opted to purposefully omit them in this study because of two main reasons. Firstly, these geographic indicators do not meet the criteria of Link et al. (2010) (as mentioned on page 6, line 25 in the original manuscript) and go against the philosophy of our ecological indicator model, even if it is trained and applied to the same region. Including geographical information and time, which had a strong weight in the model used by Raitsos et al. (2008), could prevent the model from capturing any geographical shifts, as instead of the PFTs reacting mainly to changes in the environmental conditions at a specific location, the model would just predict the PFTs based on the typical (climatological) cycles at that specific location, estimated from latitude, longitude and time. Secondly, using these inputs would prevent us from applying the model outside of the latitude and longitude range included in the training domain because this type of ANN does not perform well when extrapolating beyond the range of values seen during the training phase. Furthermore, time may particularly limit the use of the model in other areas because phytoplankton phenology and seasonal cycles in the physical variables are for example very different in the equatorial Pacific to that in the North Atlantic. Including time information trained to one area is very likely to cause spurious results when applied to another. A condensed version of this discussion is included in Section 2.1 in the revised manuscript.

"The authors divided the confirmatory data into three groups, the training, the validation and the testing. However, I did not find the exact precision for each group of the data, nor the precision for the exploratory data. In the reference [1: Identifying four phytoplankton functional types from space: An ecological approach], they indicated that model could discriminate four major phytoplankton functional types (diatoms, dinoflagellates, coccolithophores, and silicoflagellates) with an accuracy of more than 70%. The authors described the algorithms performances a lot. However I could not find the quantitatively description for the algorithm's performance."

We thank the reviewer for pointing this out. In the original manuscript, we have included a quantitative description of the algorithm's overall precision (all PFTs from all training regions) in Table 3. In the revised manuscript, we add another table in which we provide basic statistics on the algorithm performance during training, evaluation and testing for individual PFTs grouped by testing regions. We do not provide any statistics for exploratory regions because of the hypothesized mismatch between PhytoANN and NOBM in these regions. We evaluate PhytoANN results in exploratory regions by comparing them with in situ and remote sensing observations rather than NOBM results which misrepresent PFT community in these regions to some extent.

"The algorithm developed here based on the NOBM data of the four areas in the Atlantic Ocean. That's why the algorithm performances are similar in these areas as indicated in Fig 6. The authors pointed out that the algorithm developed here shows some different characteristics with NOBM on the areas only for exploratory analysis. The results are possibly interesting, but not clearly presented. The authors used a box-average of long term annual mean contribution to compare these two model performances. Scattering figure/figures with monthly data may be better for the performance comparison."

Apart from the long term annual mean contributions in Figure 6, there are plots of PFT distributions in all exploratory regions on a monthly basis presented in Figure 7. We believe that the four panels in Figure 7 are sufficient to illustrate the differences in PFT

distribution between NOBM and the PhytoANN. Additional scatter plots suggested by the reviewer would include more points from within the boxes but at the expense of adding noise and covering (but not changing) the general patterns on which we base our main discussion points.

Interactive comment on Biogeosciences Discuss., 10, 8103, 2013.