

We thank the reviewers for their positive comments. In the following, we report our proposed replies to the specific points from each referee. If accepted, these changes will be properly included in the revised version.

Response to Reviewer #1:

The paper presents and evaluates an updated version of the marine biogeochemical model ECOSMO II, that now also includes a parameterisation for a variable C:Chl ratio of phytoplankton. The biogeochemical is coupled to the HYCOM ocean model, configured for the Arctic Ocean. Model evaluation of a default setup is carried out against observed nutrients and Chl a from in situ measurements and remote sensing, and complemented by comparison to literature values of primary production. Model sensitivity to biogeochemical parameters is evaluated in two further experiments. In general, all model configurations perform quite well with regard to the observations, and improvement in surface chlorophyll when parameters are adapted to the hydrodynamics of the Arctic.

In general, the paper is well written and I very much appreciate the thorough comparison to observed tracers. One drawback is the - to my eyes - somewhat incomplete description of phytoplankton (see (1) below), and the slightly confusing description of how observations were used for comparison (2). I further encourage the authors to improve on the structure of the paper (3) and to extend a bit on the comparison of the three different experiments (see my comment no 4 below).

(1) For better understanding of how phytoplankton responds to nutrient and light limitation I suggest to present the equation(s) for phi (in Eqns 1 and 2). Without knowledge about this functional form it is difficult for the reader to understand its dynamics, in particular as DS2013 might not be accessible to everyone. Further, does the C:Chl ratio (or the amount of chlorophyll in phytoplankton) affect the light attenuation and P-I curve, i.e. the of phytoplankton growth rate?

We agree that the term ‘phi’ (growth limitation) is necessary to be present in a model description paper and understand the referee’s concern that the DS2013 paper may not be accessible for everyone. For this reason, we will include $\phi_{P_j} = \min(\alpha(I), \beta_N, \beta_P, \beta_{Si})$ as the general equation for phi and expand this equation to include all of the necessary formulations for each individual term in the equation totalling to additional 6 equations. We believe these additions will address the request of the referee. The additional equations are as follows and each equation define a limitation respective to light or nutrient:

$$\begin{aligned} \alpha(I) &= \tanh(\varphi I(x, y, z, t)) \\ \beta_N &= \beta_{NH_4} + \beta_{NO_3} \\ \beta_{NH_4} &= NH_4 / (NH_4 + r_{NH_4}) \\ \beta_{NO_3} &= (NO_3 / (NO_3 + r_{NO_3})) \exp(-\gamma NH_4) \\ \beta_{PO_4} &= PO_4 / (PO_4 + r_{PO_4}) \\ \beta_{Si} &= Si / (Si + r_{Si}) \end{aligned}$$

This following equation will replace the PAR equation in the manuscript following ‘*In relation to the addition of a prognostic chlorophyll a state variable, photosynthetically active radiation $I(x,y,z,t)$ at depth undergoing attenuation was modified to have chlorophyll a in the exponential term:*’

$$I(x, y, z, t) = 0.42 * I_s(x, y) \exp \left(-k_w z - k_{chl} \int_z^0 \sum_{j=1}^2 Chl_{P_j} dz \right)$$

For the reader to follow better, PAR will be removed from the equations and will be given as $I(x, y, z, t)$ suggesting a point in 3D and time. PAR is dependent on surface radiation, I_s , water and chlorophyll specific attenuation constants and CHL concentration. C:Chl ratio indirectly affects the light attenuation as Chl is included there. We believe the equation in this form better relates the reader with CHL vs light attenuation.

The model does not relate CHL to P-I curve at the moment, but it will be included in future iterations of ECOSMO.

(2) The description of type of observations, and how these were used for model evaluation (section 4.1) is somewhat confusing (see also below, specific comments).

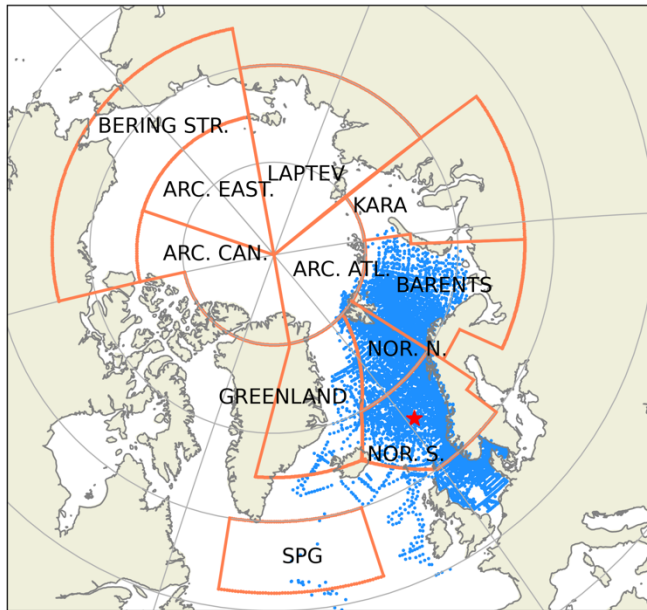
Also

Line 220-221: "Nitrate, silicate, phosphate and chlorophyll a in situ data from Institute of Marine Research (2018) were used for the statistical evaluation of the model results." Where were these data used and what is the difference to the comparison against WOA2013 data?

Reviewer-3 has also raised this issue. We will explain the use of validation data better in Section 4.1 opening the section with the following:

‘the model simulations were evaluated using three different datasets, each of which comprises observed inorganic nutrients and chlorophyll a : (1) World Ocean Atlas 2013 (WOA13; Garcia et al., 2013), (2) Institute of Marine Research (2018) data (IMR18), (3) ESA Ocean Colour CCI v5.0 (OC CCI; Ocean Colour Climate Change Initiative; Sathyendranath et al., 2019).’

This will inform the reader from the start what data is used. In the following paragraphs, we will go through each of them in detail and explain how and for what purpose they were used. In summary, WOA13 is used for evaluating monthly and regional averages, whereas the IMR dataset is used to evaluate the model point-by-point collocating model against in situ data points. The extensive data from IMR18 allowed us to perform the statistical analysis whereas WOA13 was used for a visual comparison. OC CCI data was used for an additional statistical analysis supporting IMR18 data. To further assist the structure of the observed data, we will add the following figure in the Appendix which includes the profile locations for the IMR data.



We would like to answer item 3 and 4 and the following comments together.

(3) Some sentences of the model setup (section 3) already anticipate the results (e.g., lines 183-186: "Using lower ... column stabilizes." and 189-191: During the continuous ... will replace the parameterization set in EXP2."). I would suggest to reorganise the paper a bit (see also comment (4)), and draw conclusions only from the model evaluation.

Sentences anticipating the results will be removed. Please also refer to our extensive comments on the subject below.

(4) I enjoyed reading the analysis in section 4.3. I think this analysis could be complemented by contrasting it with the results obtained with EXP2 and EXP3. For example, Table 2 shows that EXP1 always performs best with regard to the correlation coefficient, in 7 (8) cases out of 12 with regard to the normalised StdDev (Bias%), and in half of the cases with regard to RMSE. Likewise, it performs always best (regardless of metric) for phosphate and nitrate, and in half of the cases with regard to chlorophyll. Even if some differences to EXP2 and EXP3 are only small, I think this could be discussed a bit more (so far, there seems only be a sentence in line 451), and contrasted with the outcome presented in Tables 4 and 5, which indicate that EXP3 performs best with regard to surface Chl a. What is the reason for these differences? Is it because Table 4 and 5 only refer to surface values?

Line 278: "0.6-0.72" - Table 2 gives 0.79-0.83 for EXP1: confusion with EXP3?

Line 279ff: For the biases Table 2 gives the relative bias, but in the text you refer to absolute biases, which is somewhat confusing.

Line 280: "2.47-3.34" - this seems to be the bias of EXP3.

Line 290: "0.74-0.78" - again, for EXP3?

Line 300: "0.81-0.89" - again EXP3?

Line 311: "0.97-1.2" - EXP3?

Line 451: "EXPeriment statistics for inorganic nutrients are very similar in all experiments (Table 2)." - Overall, to me it seems as if EXP1 performs better than EXP2 and EXP3 (see above my comment no. (4))

Referee #2 has also raised similar concerns. We agree that our review of the model results and discussions on the statistics of different experiments were misleading. Therefore, the text on Section 4.3 will be extensively reviewed. Going through the comments and the manuscript, we understand how the text reads to be in favor of EXP3 which was not our intent. We value all the experiments as the parameter sets in these experiments are being actively used, and our intention in this manuscript was to document the model performance for each experiment. For this reason, along with many other revisions in the manuscript, we will be clearly stating the use cases for each experiment in the introduction and the model setup, we will reference what each experiment represents (i.e. EXP1 for the original ECOSMO, EXP2 for the operational model prior to July 2021 and EXP3 for the current operational model after July 2021).

To achieve an objective comparison among the experiments, the text regarding the statistics to which the referee refers now will include results from all the experiments, not only EXP1 which will also correct the confusion stemming from EXP1 vs EXP3 as stated by the referee. We have also noted the EXP1 better performs well in many aspects and each of these will be documented in the text for both the statistical analysis and the monthly average comparison to WOA (as we will also note in the text that EXP1 averages perform better for nitrate and phosphate) and discuss possible reasons. EXP2 will also be included in these new additions.

The contrast between performances among EXP1 and EXP3 with reference to statistical analysis against IMR and satellite data will be discussed. Here are some examples we plan to add:

"Experiments were generally comparable when the model results were regionally and monthly averaged. Notable differences were found for the mid-summer nitrate and phosphate concentrations for the Barents, Norwegian and Greenland Seas, as the drawdown of these nutrients was better resolved by EXP1 compared to climatology as EXP1 summer nutrients were lower than in EXP2 and 3. A possible reason why EXP1 has larger drawdown of nutrients during mid-summer is the higher photosynthesis efficiency applied in EXP1 resulting in higher uptake of nutrients and higher zooplankton grazing rate applied to EXP2 and 3 resulting in higher top-down pressure to phytoplankton preventing phytoplankton biomass to uptake more nutrients."

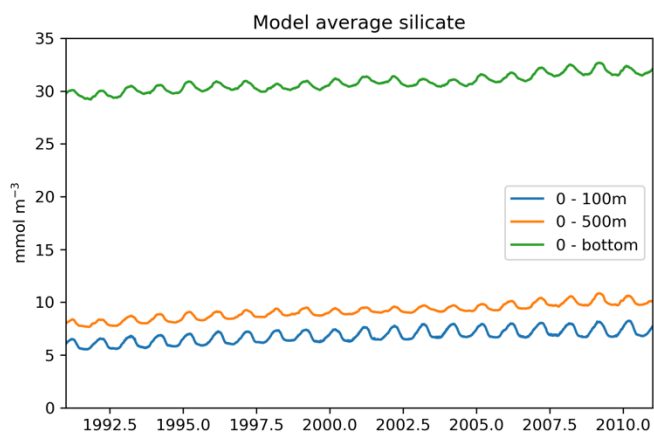
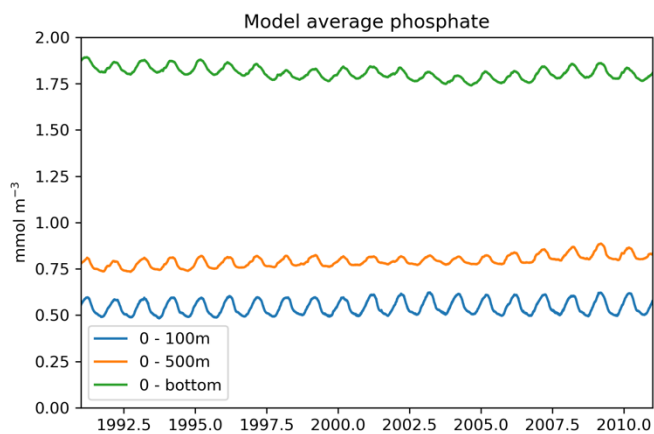
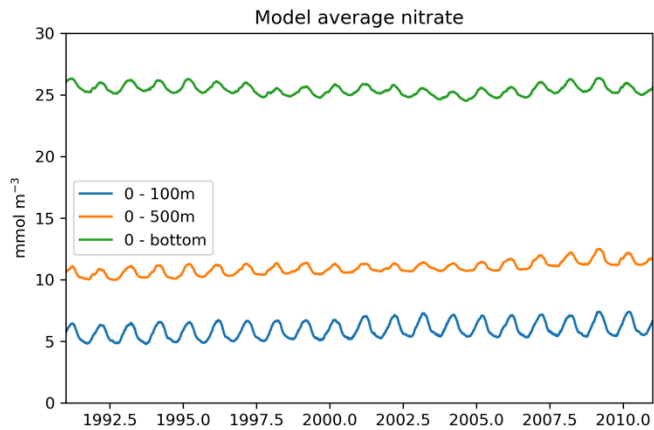
"Among the experiments, all perform very similar in terms of nutrient correlations while for nitrate and phosphate, EXP1 performs slightly better in terms of stds, EXP3 performs slightly better for silicate for the Barents Sea and EXP2 for the Norwegian Sea, though the differences among the experiments were almost negligible. Similarly, EXP1 perform better in terms of %

bias and RMSE for nitrate and phosphate, and EXP3 perform better for silicate. The slightly better performance of EXP1 for nitrate and phosphate is also evident in summer averages when compared to climatology as mentioned before (Figs. 3a and 3c). The model performance for silicate when using monthly averages shows even less differences among the experiments, however, the EXP3 is slightly closer to WOA timeseries compared to EXP1 (Fig. 3b)."

Line 36: Here I would prefer a more specific link to a model application at copernicus (I had to search around a bit). The DOI will be added.

Line 172: Evaluation after just two years of spinup seems to be quite soon - do you have any indication if the model drift (with regard to the BGC components) has decreased to some specific value?

Below we provide model average (spatial and vertical) evolution of nutrients between 1991 and end of 2010 during which the statistical analyses were performed. Inspecting the first years, the model does not have any abnormal drifts that could be interpreted as a the spinup prior to the analyses were sufficient. We note that, unless the referee has objections, we will not include the following figures in the manuscript.



Line 201-202: "The dynamics shown in Appendix A1 is expected to be valid for each model point." - How is this expectation justified?

The sentence will be changed with the following: "The dynamics shown in Appendix A1 is representative for regions with similar plankton dynamics (e.g. Norwegian Sea, Barents Sea), thus can be used as a showcase for the new chlorophyll a specific addition."

Line 229 "Processed model chlorophyll" - What does "processed" mean in this context? (we were aiming for the meaning 'post-processed', but the sentence is modified to: "We used this

dataset for the years 1998-2010. Chlorophyll a and kd490 were remapped to the model grid and the model chlorophyll a was averaged to 1/kd490 (m) depth.”)

Figure 3 caption and elsewhere: "WOA18" - above you refer to WOA2013 (The text is made consistent throughout and states WOA13).

Line 383-383: "silicate concentrations ... in the climatological data." - are discharge rates into the Arctic very different for different types of nutrients? (i.e., outside the assumed stoichiometry)? Perhaps a note on this would be interesting ...

We will add: *“As the model is relaxed towards the climatology at the Bering Strait through a sponge layer in the model domain, the high overall nutrient concentrations near the Bering Sea and especially the high silicate concentrations at the Siberian coast due to higher Si/N ratio of Pacific origin water masses compared to the Atlantic water masses, and the addition of high Si/N ratio river discharge is reflected in the model annual averages.”*

Lines 443-446: It is not really clear to me what you want to say with this sentence - could it be rephrased? (will be rephrased to: EXP1 show a very fast primary production response to light availability during the spring bloom period with notably higher chlorophyll a concentrations compared to the observations (results not shown) evident in the high %biases, whereas chlorophyll a concentrations in EXP2 and EXP3 are closer to observed values during spring bloom.)

Lines 478-479: "using realistic ... spin-up period" - what are the realistic constant values? and: constant with respect to what - over time? over depth? (we agree that the sentence is misleading. We will correct it with: “We performed a 27-year run starting in 1990 using WOA2013 profiles from January climatology for the biogeochemical variables ... “)

Other minor comments will be corrected following the referee’s suggestions.

Response to Reviewer #2:

The paper presents a description and evaluation of a new configuration of the biogeochemical model ECOSMO II (ECOSMO II(CHL)) in which they have included chl in the three functional phytoplankton types as prognostic variables. Furthermore, the paper presents a comparison of model experiments using three different parameter sets. In the appendix, the authors have also included a comparison between ECOSMO II and ECOSMO II(CHL) in a configuration where the models are coupled to the 1d physical model GOTM.

The paper is well written and the methodology ambitious. However, I would like to have seen a larger focus on the comparison between ECOSMO II and ECOSMO II(CHL). Furthermore, apart from two sentences (451-453) I find no discussion on why EXP1 seems to perform better

for nitrate and phosphate (Fig 3 and Table 2) than EXP 2 and 3. I find that the paper could be published in GMD after some minor revisions.

Lines 139-144. What limited the concentrations from becoming too small before you added this? Also, you state that: “The minimum concentration at which the loss terms are calculated are...”. Should it say: “The minimum concentration at which the loss terms are switched off”?

There was no limit before the addition of the on/off switch, which quickly became a problem for a well-mixed open ocean. Hence we added this modification. It was a necessity for preventing a very late spring bloom, as the biomasses reached almost 0 after every winter. We agree with your suggestion and will rephrase the sentence accordingly.

Fig.3: It is not obvious to me that EXP2 and 3 perform better than EXP1. It looks the opposite from this figure.

Referee #1 has also raised similar concerns. We would like to share the same answer with both Reviewer #1 and #2. We agree that our review of the model results and discussions on the statistics of different experiments were misleading. Therefore the text on Section 4.3 will be extensively reviewed. Going through the comments and the manuscript, we understand how the text reads to be in favor of EXP3 which was not our intent. We value all the experiments as the parameter sets in these experiments are being actively used, and our intention in this manuscript was to document the model performance for each experiment. For this reason, along with many other revisions in the manuscript, we will be clearly stating the use cases for each experiment in the introduction and the model setup, we will reference what each experiment represents (i.e. EXP1 for the original ECOSMO, EXP2 for the operational model prior to July 2021 and EXP3 for the current operational model after July 2021).

To achieve an objective comparison among the experiments, the text regarding the statistics that the referee refers to now will include results from all of the experiments, not only EXP1 which will also correct the confusion stemming from EXP1 vs EXP3 as stated by the referee. We have also noted that EXP1 performs better in many aspects and each of these will be documented in the text for both the statistical analysis and the monthly average comparison to WOA (as we will also note in the text that EXP1 averages perform better for nitrate and silicate) and discuss possible reasons. EXP2 will also be included in these new additions.

The contrast between the performances of EXP1 and EXP3 with reference to statistical analysis against IMR and satellite data will be discussed. Here are some examples we plan to add:

“Experiments were generally comparable when the model results were regionally and monthly averaged. Notable differences were found for the mid-summer nitrate and phosphate concentrations for the Barents, Norwegian and Greenland Seas, as the drawdown of these nutrients was better resolved by EXP1 compared to climatology as EXP1 summer nutrients were lower than in EXP2 and 3. A possible reason why EXP1 has larger drawdown of nutrients during mid-summer is the higher photosynthesis efficiency applied in EXP1 resulting in higher uptake of nutrients and higher zooplankton grazing rate applied to EXP2 and 3 resulting in higher top-down pressure to phytoplankton preventing phytoplankton biomass to uptake more nutrients.”

“Among the experiments, all perform very similar in terms of nutrient correlations while for nitrate and phosphate, EXP1 performs slightly better in terms of stds, EXP3 performs slightly better for silicate for the Barents Sea and EXP2 for the Norwegian Sea, though the differences among the experiments were almost negligible. Similarly, EXP1 perform better in terms of % bias and RMSE for nitrate and phosphate, and EXP3 perform better for silicate. The slightly better performance of EXP1 for nitrate and phosphate is also evident in summer averages when compared to climatology as mentioned before (Figs. 3a and 3c). The model performance for silicate when using monthly averages shows even less differences among the experiments, however, the EXP3 is slightly closer to WOA timeseries compared to EXP1 (Fig. 3b).”

Line 320: Is it not the net primary production your extracting? You don't explicitly model respiration so what you model is gross primary production (photosynthesis) minus respiration i.e. the net?

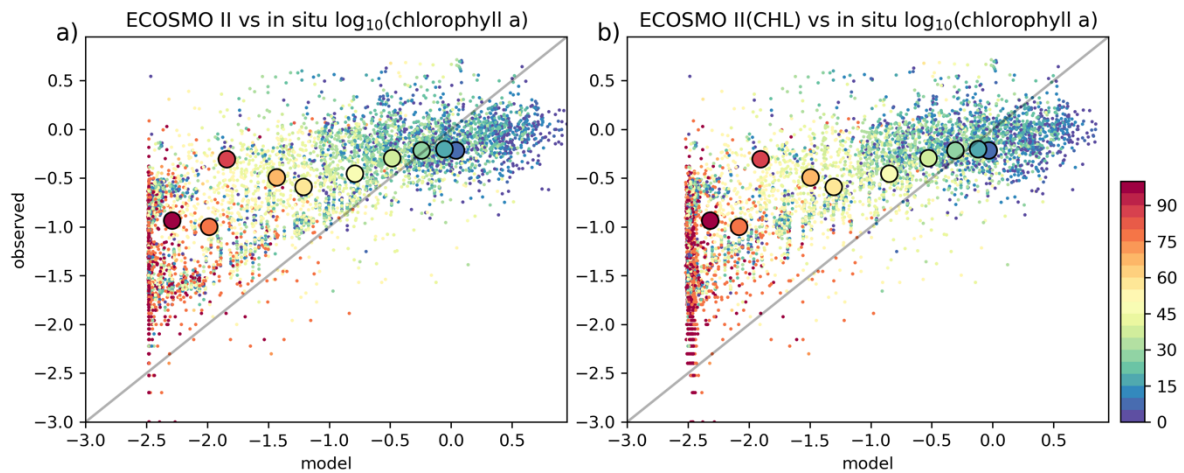
We model gross primary production and do not include respiration as a loss term in the equations. Since respiration is not included in the equations, modifying the gross primary production with an arbitrary multiplier to include respiration (e.g. $0.9 \cdot \text{GPP}$) when comparing to observations would add inaccuracy to our results. As the PP observations have a large range, removing 10% PP from the model GPP will still be within the observed PP ranges, thus we settled on providing GPP in the results and noted this in the section with the sentence “Because the model does not have an explicit term for respiration, we can only extract gross primary production from the model, which is then compared to observations. “

Lines 388-389: What is the reason for the difference in N/P ratios in the Barents sea between WOA and model, I guess mostly as a result of the difference in nitrate in this area?

The model uses Redfield stoichiometry for production and remineralization in the water column. This strict control on N:P ratios in the organic compartments of the model will restrict dynamical changes in the pelagic N:P ratios. The model's tendency to form a Redfield ratio will be counteracted by the relaxation to open boundary inorganic nutrient conditions of ocean climatology, the river nutrient climatology and the benthic nutrient dynamics.

Fig. A1: It would be interesting to see a plot comparing ECOSMOII and ECOSMOII(CHL) to observations so as to get a clearer view of the benefits of variable C:Chl ratios.

We understand that a visual element is necessary to support the statistics table that was included in the manuscript. For this reason, we will add the following figure in the revised manuscript which depicts the data points that were used to calculate the statistics. To distinguish depth layers easily, data points were averaged for each 10 meter depth interval and were depicted with larger markers with color coding for that depth interval.



We will add the following sentences to support this figure: *“While both simulations are statistically similar in general, especially in the deeper euphotic zone (40 – 80 m), ECOSMO II(CHL) statistically performs better at 0 - 20 m and 20-40 m range (Table A1) for almost all statistical quantities. The data that was used to calculate the statistics in Table A1 is visualised in Figure A2 using a scatter plot of the modelled and observed chlorophyll a, which confirms the values in Table A1 showing a slightly better performance of ECOSMO II(CHL) near the surface (0 - 20 m range). 30 m average point is visually slightly better for ECOSMO II, which probably reflects the better % bias performed for 20 – 40 m range (Table A1). While overall the model performance improves, further modifications to either model parameters or formulation should be done for the future iterations of ECOSMO as below 40 meters, the model has not gained a significant improvement suggesting that the chlorophyll a dynamics should be improved for low-light conditions.”*

Other minor comments will be corrected following the referee’s suggestions.

Response to Reviewer #3:

The manuscript describes and evaluates the coupled HYCOM-biogeochemistry ECOSMO II model, configured for the Arctic Ocean. Its evaluation includes some suggestions for fine-tuning of the parameterization according to the geographical characteristics of the expanded domain. The technical description of the ECOSMO II model is well written and constitutes a good baseline for the modeling community. The evaluation of the model results, against observational values, is also of value for a quantitative validation of the model and it is well complemented by the commentary and the figures therein.

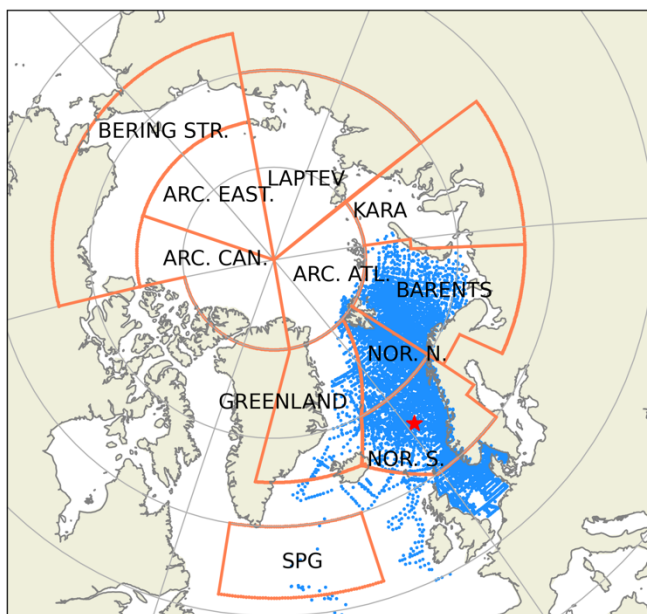
The paper is very well written, but it would benefit, in my opinion, by a better and more organized description of what observations were used in what experiments.

Reviewer #1 has also raised this issue and we share the answer here. We will better explain the use of validation data in Section 4.1 opening the section with the following:

‘the model simulations were evaluated using three different datasets, each of which comprises observed inorganic nutrients and chlorophyll a : (1) World Ocean Atlas 2013 (WOA13; Garcia

et al., 2013), (2) Institute of Marine Research (2018) data (IMR18), (3) ESA Ocean Colour CCI v5.0 (OC CCI; Ocean Colour Climate Change Initiative; Sathyendranath et al., 2019).'

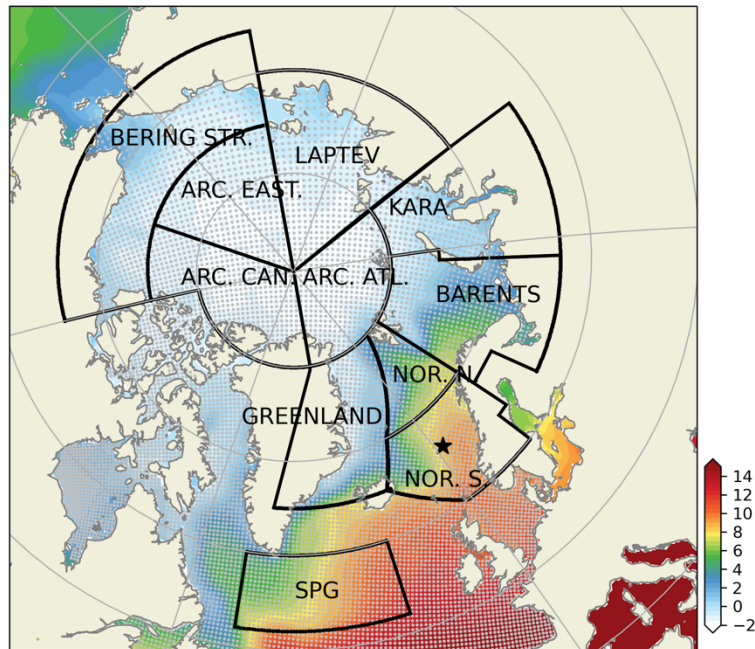
This will inform the reader from the start what data is used. In the following paragraphs, we will go through in detail each of them and how and for what purpose they were used will be given. In summary, WOA13 is used for evaluating monthly and regional averages, whereas IMR dataset is used to evaluate the model point-by-point collocating the model to observed data points. The extensive data from IMR18 allowed us to perform the statistical analysis whereas WOA13 was used for a visual comparison. OC CCI data was used for an additional statistical analysis supporting IMR18 data. To further assist the structure of the observed data, we will add the following figure in the Appendix which includes the profile locations for the IMR data.



The selection of the areas (and experiments) to show and compare seems ad-hoc. What is the justification for these selections?

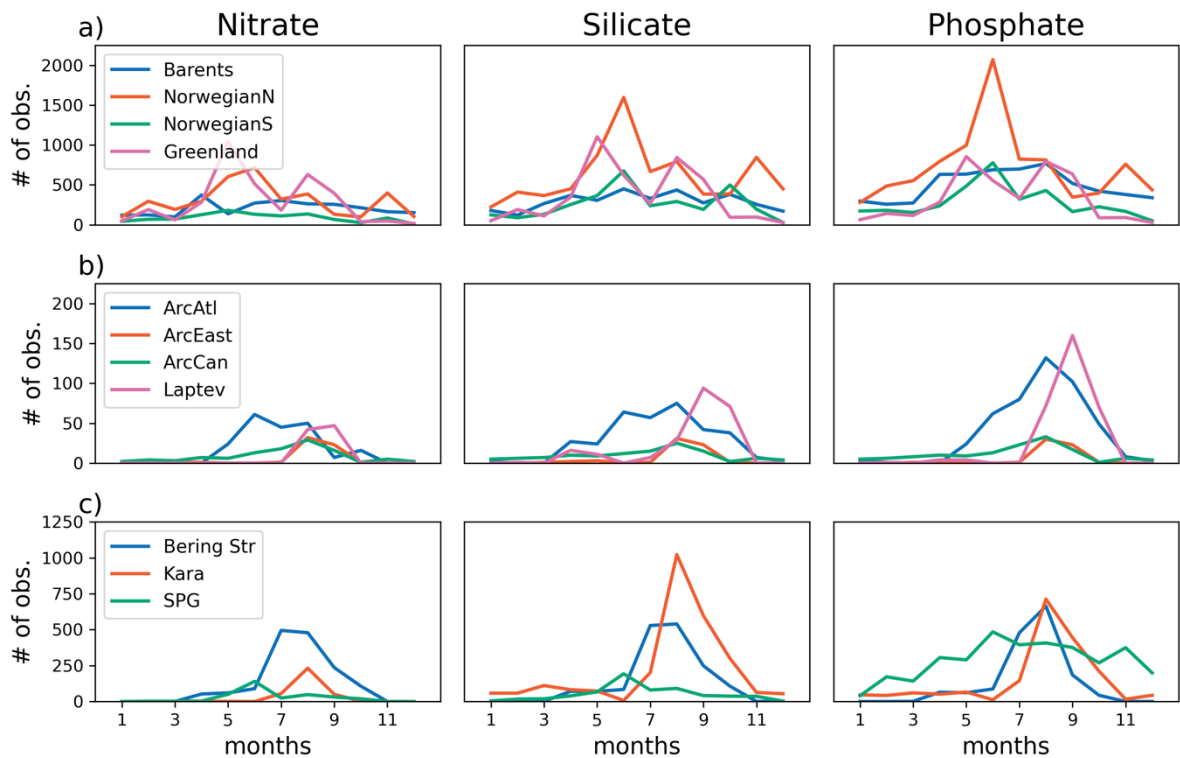
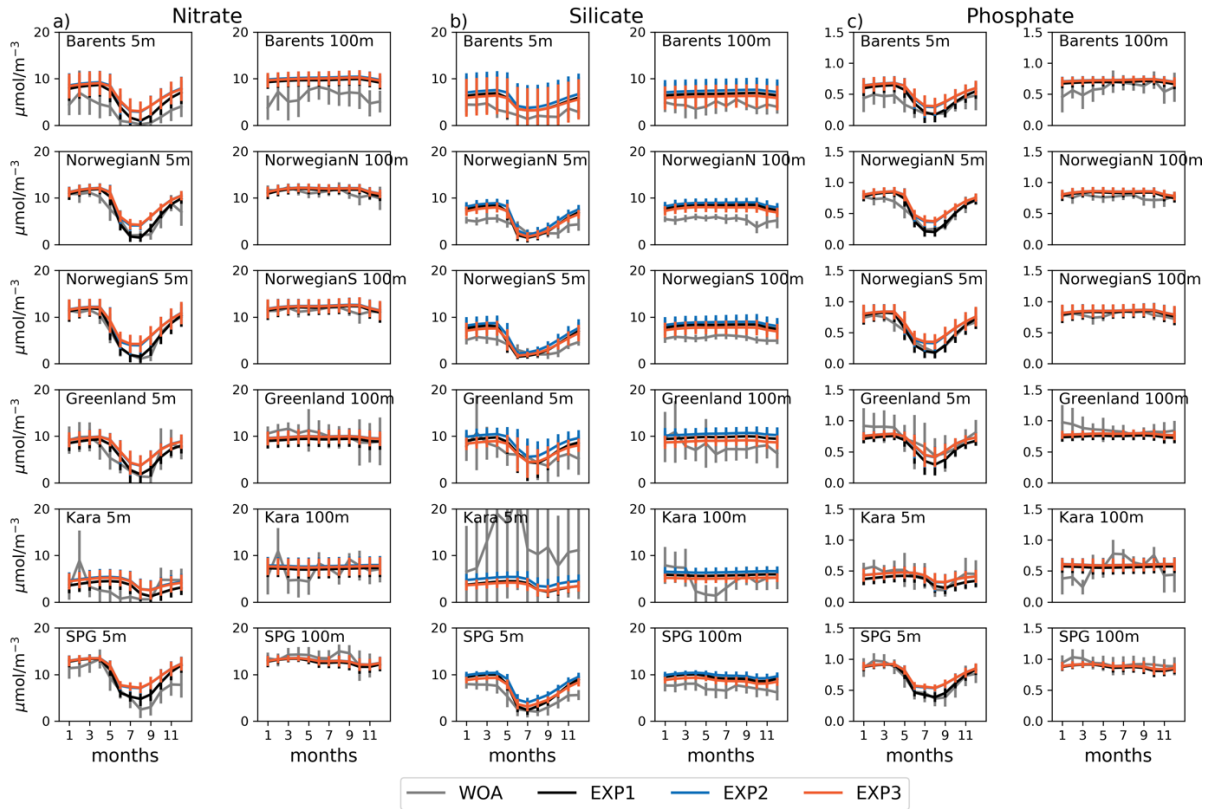
The subregions were defined based on geographical definitions and environmental characteristics such that Barents Sea area defined in this manuscript cover the shelf area south and east of Svalbard, and following the opening to the west (roughly the line we defined, start the deeper ocean that the Norwegian Sea is located. Norwegian Sea is of high interest for our model development and this region in the manuscript extends 20° of latitudes, and such a long latitude coverage can be significant in relation to the timing of the spring bloom, thus in our analyses we divided the region to north and south from halfway. We appreciate you raising concern for a need for a physical setting of the region and added annual temperature climatology to Figure 2 (please see below), and there the reader will see that the Norwegian Sea (N and S) is distinct from the Greenland region where the region borders roughly represent this distinction. Furthermore, the arctic regions (ARC-XXX) were separated from the rest as ARC regions are in general sea-ice covered. You will notice in the figure above that ARC regions are distinctively free of observations, and ARC border separate the observed regions. BERING STR. region is defined to separate it from the rest as it is a

nutrient relaxation to WAO13 region and LAPTEV and KARA are naturally divided with the surrounding islands. SPG region is a subregion within the subpolar gyre area. Necessary text giving this information will be added to the text.



Why not compare all of the areas in all comparisons and statistics? Seems that it wouldn't add much more content and would be of better use in a comprehensive descriptive assessment of the model for this region at large. Perhaps a way to scope the comparative analysis would be to concentrate on one or two of the subregions of the model domain.

For the statistical analysis, the model comparison to IMR data were restricted by data availability, almost exclusive to the NOR.S, NOR.N and BARENTS regions, which are our main area of interest (Please see the figure that was provided with our first comment to Referee #3 for the location of the observations). This figure will be included in the Appendix). For this reason, the manuscript can only focus on these 3 regions (Table 2, Fig.4). For the regional and monthly averaged time-series plots (Fig.3), we were less restricted as we have climatology to compare with. However, this should be done with caution, as in certain regions the climatology is based on very few observations. To show this, we will add a figure (Fig. A2) that depicts the number of observed points for each region. After our inspection on the number of samples, we decided to include the regions of meaningful coverage in Figure 3, NOR.N, NOR.S, BARENTS, GREENLAND, SPG and KARA. We note that KARA is also limited in data outside summer period. BERING STR is not included in the figure as it is too close to the relaxation zone of the model. The text in the manuscript is extensively modified to account for these changes.



The performance comparison for the model would be well prescribed in a table of EXPT vs. AREA with some objectively-derived performance metric value (i.e. rmse/skill score).

As pointed out above, we can only do the statistical analysis for the regions where we have data and these regions were already covered in the text. Statistics (rmse, bias, correlation and standard deviations for nitrate, silicate, phosphate and chlorophyll a) for EXPT and AREA are given in Table 2. We tried to further extend the chlorophyll error statistics with the satellite data but for this region, satellite chlorophyll is hindered by clouds, ice cover and dark seasons. Therefore, we kept the statistical analysis in Table 4 and 5, Figure 9 consistent with the regions in Table 2 and Figure 3. But following the comments from the other referees, the text will be extensively modified to include results for EXP1, EXP2 and EXP3 and comment on each of their strengths. The modification to Figure 3 now depicts both Norwegian regions, Greenland and Kara Sea which would assist further towards your concerns.

This would allow for a good summary of the many combinations of experiments and areas covered. The many regions, very comprehensive, somewhat dilutes the detailed validation of the model itself, only providing a general descriptive validation with some statistical quantification for some of the areas.

We value the necessity of domain wide analysis as you also suggested here and above, and for this reason we included domain wide model output (Figs. 5 and 8) with region specific model variable quantities for the readers interest and reference for future studies.

While technically, this is a paper of excellent value, scientifically, it could have been further elaborated to elucidate what the model brings in terms of better understanding of the hydrodynamics and biogeochemistry of the Arctic region. It would also benefit from a more detailed description, via equations, diagrams, and text, of the plankton group dynamics and again, the scientific value of those results.

In summary, the paper is very well written and the technical topic is well covered, but the scientific value of the paper could be improved with more details on the model parameterization, inherent errors therein, and the “take home” message of what is learned scientifically.

As a model description paper, the main focus of this study is to provide the technical changes to ECOSMO, present the final state of the model, and evaluate the model performance using different parameterizations. While doing so, we have long-term model time-series spanning for 2 decades for the Nordic Seas and the Arctic. We included the general biogeochemistry of this region within a technical paper context. We plan for further scientific investigations of this data in future studies. However, considering the referee’s comments, we will revise the manuscript towards strengthening its existing scientific messages in the abstract, conclusion and Section 5 will have improved scientific emphasis. Specifically:

- Comment on the value added to the model by the including explicit chlorophyll a
- What do different parameter values mean in terms of model performance and internal dynamics.
- Lessons learned with respect to future evolutions of the model

Reviewer #1 also commented on expanding the equations in the manuscript and we will follow both the suggestions of Reviewer #1 and #3 and detailed descriptions of phytoplankton growth dynamics will be included. Please refer to our answer to Reviewer #1 on the subject.

Other minor comments will be corrected following the referee's suggestions.