Notes on colours:
- Referee's comments are given in black.
- Authors' comments are given in red.
- Texts from the revised manuscript are given in blue.

Referee 1

Interactive comment on "CSIB v1: a sea-ice biogeochemical model for the NEMO community ocean modelling framework" by Hakase Hayashida et al.

A. Randelhoff (Referee)
achim.randelhoff@takuvik.ulaval.ca
Received and published: 19 September 2018

Hayashida et al. report here on their implementation of a sea ice biogeochemical model into the NEMO framework. The strength of NEMO is the ability to compare different submodules, and thus to isolate the overall impact of very specific parameterizations and submodels. This is an important step in refining existing models and towards singling out future research directions in this ambitious field of mechanistic modelling of (Arctic) ocean biogeochemistry.
I found the paper generally well-written and laid out clearly; find some suggested improvements below.

We thank the referee for his positive comments and thorough feedback on our manuscript. Below, we provided our responses to the referee's comments and revised our manuscript based on these comments as much as possible.

Major issues:
Section 3.1: I am not entirely convinced by your analysis. The "break points" you see in the time series appear to be quite arbitrary; they probably make sense to you based on your familiarity with models and previous studies, but a reader might want to see a statistical analysis that supports your claims. That being said, I am not entirely sure to what extent you need these spinup times for your later analysis; you are probably able to carry through the rest of your paper without most of the claims put forward in this section. If they are important in their own right, I would recommend considering a more rigorous presentation.

We agree that the break points discussed in the time series during spin up are quite arbitrary and that it would be more desirable to make these claims using statistical analysis. Our analysis here is qualitative, and so chose the break points based on the findings from previous model studies. As the referee also pointed out, this spin-up section does not affect the conclusions of the rest of the paper. However, presenting the results during model spin-up can be helpful for future studies. Therefore, we decided to move this section into an appendix section (Appendix B), and revised the text accordingly.

The paper should probably also explain more generally why it focuses on the period

1969-79 for spin-up and 1979 for all experiments, especially since e.g. snowfall climatology (and likely all other data) are much scarcer for that period than later ones. The Arctic also looked very different in 1979 from what it is today, so you should give a good reason if you expect that your comparisons from the 1970s are relevant today as well.

We agree that the model can be better evaluated for more recent period than the period considered in the present study. We extended our simulation up to 2015 and evaluated with more recent observational data (Hayashida 2018, PhD thesis), but these results are planned to be published as separate papers considering the contents of the present study, which is intended as a model description paper. We revised the manuscript to clarify these points (P2 L31):

"We note that this study is intended as a model description paper, and the analysis focuses on results for the year 1979, corresponding to the end of a decadal model spin up. The analysis of the simulation beyond 1979, in which more observational data are available for evaluation (Hayashida 2018b), is planned to be published as a journal article separately."

Hayashida, H. (2018): Modelling sea-ice and oceanic dimethylsulfide production and emissions in the Arctic, PhD thesis, University of Victoria, Victoria, BC, Canada, http://dspace.library.uvic.ca/handle/1828/10486

Minor comments and suggestions:
General:
Some of your figures have colorbars that are of very little use (examples: Fig. 8, Fig. 14b+c, Fig. 16c, ...) because they are scaled linearly and some extreme regions mask the variability over most of the map. Assuming you produce your figures using matplotlib, you could look into https://matplotlib.org/users/colormapnorms.html.

Re: Fig.8, 14b: Using a linear scale for the colorbar allows us to easily distinguish between high and low productive regions, where the contribution of the latter region to the total production (area integrated) is negligible. Hence, we would prefer not to change these figures.

Re: Fig.14c: We revised the figure by reducing the range of the colorbar scale so that the contrast is stronger.

Re: Fig.16c: We revised the figure by adopting the log scale and changing the colormap to "YlBuGn".

You mix present and past tense occasionally (e.g. Section 3.3.1; p20 line 9ff; ...).

Revised to use present tense consistently in Results and Discussion section.

Specific:
p1
15 spell out LIM2/3, PISCES

Revised as suggested.

p2
23 "horizontal transport of ...": Technically correct (here as elsewhere), but it might be more intuitive to explain that you mean transport with sea ice drift.

Revised as suggested by adding "associated with sea ice drift" at the end of the sentence.

27 You may want to quickly mention in this section why you develop another sea ice BGC model after several others already exist. I assume it is because the NEMO framework allows model intercomparison, your overarching goal, and it did not have any yet?

Yes. NEMO does not have any sea-ice biogeochemical model component (except the offline version of Tedesco et al. 2017, as noted on P2 L19 in the original manuscript) and also having one for NEMO allows intercomparison. We revised the manuscript to mention about intercomparison in this paragraph (P2 L27):

"These implementations allow more realistic simulation of sea-ice biogeochemistry and intercomparison of process-based ice algae models."

p3
Tab. 1: Specify that i0 is for the uppermost layer, be it snow or ice or both (which I think it is based on p25, l.14, as Maykut & Untersteiner had apparently measured it for snow-free surfaces, but correct me if I am wrong). Mention briefly (here or in the text) what you mean by "ice algal skeletal layer".

The referee is correct that i0 defined in the paper refers to either snow or ice surface. However, in this Table, we compare i0 for snow surface among studies. We revised the table caption and first row by replacing "i0" with "i0 (snow surface)" to clarify that we are comparing i0 for snow surface. And throughout the manuscript, we replaced "i0" with "i0 for snow surface" for clarification where applicable.

Replaced "the thickness of ice algal skeletal layer" with "the vertical extent of the biologically-active layer at ice base".

9ff. " a thin layer...": Is this synonymous with the "scattering layer" of sea ice optics? What happens when there is less than 10 cm of snow+ice? Does this fraction penetrate only below 10 cm, after which different attenuation coefficients are applied, or is transmissivity below this "thin layer" set to 1?

Yes, the surface thin layer defined in the present study is synonymous with the scattering layer of sea ice optics. When the snow+ice thickness is less than 10 cm, the penetrating fraction (i0) enters the underlying seawater. The Beer-Lambert law is applied for the penetrating fraction. We revised the schematic to clarify these points (Figure 1). We also revised the manuscript to use the term surface scattering layer (SSL) for consistency with the sea ice optics literature.

11 Two times "this" is repetitive

Removed as suggested.

p4
Fig. 1: Again, is this "surface thin layer" the same as the scattering layer? In this schematic, you could also indicate if there is additional attenuation below this "surface thin layer".

Yes, the surface thin layer is identical to the surface scattering layer. In the revised figure caption for Fig.1, we indicate that the radiation penetrating below the SSL attenuates following the Beer-Lambert law.

9 I think I understand what you are saying, but in my mind the point of mechanistic modelling is exactly to include more and more parameterizations. Is PISCES' performance not good enough to justify the extra computational costs it would imply as compared to CanOE?

As the reviewer points out, extra computational costs can be justified if the model performs better. CanOE has been developed as an advanced/more sophisticated biogeochemical model for the Canadian Earth System Model version 5 (CanESM5), compared to the simpler NPZD version CMOC (Canadian Model of Ocean Carbon) used in CanESM2 with the idea to limit the complexities as much as possible, while allowing to address insufficiencies in CMOC (e.g. single NPZD model are either tuned towards low nutrient or high nutrient oceans). PISCES includes complexities that do not necessarily warrant the additional computational resources, as well as some which have limited foundations. Preliminary results of a model intercomparison study (Steiner et al., in prep.) for primary production in the Arctic do not necessarily suggest superior performance for PISCES. We would also argue that the point of mechanistic modelling is to improve the parameterizations in a way so they better represent the real world and improve the correspondence between model and observations. That does not necessarily mean including more and more parameterizations. Some parameterisations might be more accurate but do very little to the overall model improvement.

15-16 I can live with the sentence but found it a tad vague, probably because I did not understand what you mean by "state of the ecosystem". In terms of nutrient budget modelling, part of the ecosystem *is* the sulfur cycle so the inverse statement is a tautology.

Replaced "state of the ecosystem" by "conditions of primary and secondary producers".

p5
11-12 43% in units of W/m^2? I also assume you mean "43% of the downwelling shortwave radiation reaching the sea surface"

Revised as suggested.

p6
Fig. 2: "stoichiometry": stoichiometries

Revised as suggested.

p7
4 Plase explain what "[e]ddy diffusion tendencies" are. How are they computed (explicitly)?

Eddy diffusion tendencies are the rate of change in a sea-ice biogeochemical state variable due to horizontal transport by unresolved motions (the second term in Equation 2). They are computed by evaluating the second order diffusive operator using the Crank-Nicholson method. We revised the manuscript to add this information (P7 L23):

"Diffusion is computed within the ice pack by evaluating the second-order diffusive operator using the Crank-Nicholson scheme (Crank and Nicolson, 1996), while it is set to zero at the ice edge."

10 By "concentrations", I assume you mean sea ice bulk, not brine concentrations. Since salinity gets a special treatment, are there problems of mass conservation during ice freezing because of changes in ratios of nutrients to salinity?

Yes, the term "concentrations" refers to sea ice bulk concentrations. As noted in the original manuscript (P7 L16), Equation 3 does violate mass conservation. However, it has negligible impacts on ocean biogeochemistry given the relatively-thin sea ice biologically-active layer (3 cm). Salinity and nutrients in sea ice are modelled separately. While salinity is non-zero throughout the ice column, nutrients (and other sea-ice BGC variables) are represented as having zero concentration above the biologically-active bottom-ice layer. We revised the manuscript to clarify that it is the bottom 3 cm of newly-formed ice that has the same concentration as the underlying seawater (P7 L28) and that the concentration above is zero (P8 L9):

"The bottom 3 cm of newly-formed ice is assumed to contain the same concentrations of biogeochemical state variables as those in the underlying water column." (P7 L28)

"Above the bottom sea-ice biogeochemical layer, the concentrations are set to zero for all biogeochemical tracers." (P8 L9)

11 Add "the concentration (X) of any ..."

Revised as suggested.

18-19 "minimum biomass threshold": I assume this threshold is also arbitrary, not based on field measurements?

The threshold is taken from our earlier model study (Mortenson et al. 2017) which is based on observed range (Garrison et al. 1983). Thus, it is not entirely arbitrary. We revised the manuscript to clarify this point (P8 L7):

"This threshold is derived based on the observed range of ice algal biomass in young sea ice (Garrison et al., 1983) and by assuming a fixed carbon-to-chlorophyll ice algal cell quota (Mortenson et al., 2017)."

Garrison et al. (1983): A physical mechanism for establishing algal populations in frazil ice. Nature, *306*(5941), 363.

29 You say you use the molecular diffusive exchange of nutrients, but my impression is your model would not resolve the molecular sublayer (a few mm from the ice-ocean interface). Without having checked I assume you use combined turbulent-molecular diffusion coefficients, but you may want to include the right references here, especially since such coefficients have never been measured (as far as I know) for tracers other than momentum, heat, and salinity.

As the referee pointed out, it is the combined turbulent-molecular diffusion; the effects of turbulence are accounted as the molecular sublayer is parameterized as a function of friction velocity (Equation 27

of Mortenson et al. 2017), and the molecular diffusion coefficient is derived from measurements for dissolved silica in seawater at 2 degree-C (Rebreanu et al. 2008). We added this discussion in the revised manuscript (P8 L23):

"For 1), the effects of turbulence are approximated by parameterizing the molecular sublayer as a function of friction velocity, and molecular diffusion is calculated using the observed diffusion coefficient of dissolved silica measured in seawater at 2 ∘C (Rebreanu et al., 2008)."

Rebreanu et al. (2008), The diffusion coefficient of dissolved silica revisited. *Marine chemistry*, *112*(3-4), 230-233.

30 Flooding due to negative sea ice freeboard does not count towards flushing?

Flooding due to negative sea ice freeboard is accounted and considered as part of surface ablation in the text. We added this information in the revised manuscript (P8 L22):

"flushing of these variables by flow of water through the ice from rainfall and surface melting (including flooding due to negative freeboard)."

p8
2-3 "designed to be the most realistic ...": "realistic" is as such a bit vague and you might want to rephrase as something like "thought to be most realistic among all choices considered in this paper".

Revised as suggested.

p10
1 "initialized to arbitrarily low values": I do not understand what "arbitrarily low" means.

Replaced "arbitrarily low values" with "very low values (e.g. 0.01 mmol C m-3 for the carbon contents of phytoplankton, zooplankton, and detritus)".

5 "while the other two boundaries (along North America and Eurasia) were assumed to be closed": So how was riverine (freshwater) input distributed into the ocean?

We distribute riverine freshwater flux into specified grid cells based on Dai and Trenberth (2002) as described in the original manuscript (P10 L13). It is independent of lateral boundaries. We realize that this can be better explained visually, and thus, included a figure displaying the locations of "river mouths" where riverine input is deposited in the revised manuscript (Figure 5 and P11 L23 of the revised manuscript):

"Figure 5 shows the seasonal and interannual variability (a and b) and spatial distribution (c) of the total discharge over the pan-Arctic."

13-14 "... was neglected, and therefore ... not addressed": This is a tautology. Is there a reason why? Lack of data? Too small? Too hard?

Lack of adequate data to prescribe riverine concentrations of biogeochemical variables in the model domain. We revised the sentence accordingly (P11 L25):

"The river discharge of biogeochemical state variables was neglected due to the lack of adequate data."

p11
Tab. 3: Typo in units of rn_ahtrc_0, should probably be m2s-1.

Revised as suggested.

6 "modified": How? Were only the data quality flags adjusted?

We contacted Clark Pennelly who provided the data and confirmed that there was actually no modification to the atmospheric data values; the only changes were indexing and ordering of latitudinal coordinates and renaming variables. We revised the text slightly to indicate that there was no modification to the data values including the missing data flags (P11 L34):

"As a substitute, we used a version provided by Clark Pennelly at the University of Alberta (personal communication) which addressed the missing data flag errors without any modifications to the atmospheric data (the only changes were indexing and ordering of latitudinal coordinates and remaining variables)."

10 So just to be sure; You used the 1979 snowfall for all years 69-78, but for all the other variables you use the 69-78 time series data?

The 1979 snowfall and total precipitation are used for 1969-1978, while for all the other variables we use the 1969-1978 time series data. We revised the manuscript to clarify this point (P12 L5):

"However, in EXP0, we prescribed the total precipitation and snowfall for 1979 repeatedly for the simulation over the period 1969-1978, while keeping the remaining atmospheric variables the same as the original DFS dataset."

15-16 You may want to consider additionally archiving the current version (e.g. using a doi) upon publication.

As suggested, we archived the current version and produced a DOI. We added this reference to the revised manuscript (P12 L12):

"For a complete list of the parameters, readers are referred to the source-code archive (Hayashida, 2018)."

18-19 "were adjusted to improve": How adjusted, and by what measure did you check that they "improved" sea ice volume etc.? Maybe insert reference to later if this is part of the discussion of the model runs.

We adjusted these parameters by running a number of simulations, every time with a different combination of the parameters. By "improve", we meant their comparison with PIOMAS. We revised this sentence accordingly (P13 L1):

"The other two parameters (hiccrit and pstar) were adjusted to improve the fit with the PIOMAS data product (Section 2.6) in terms of sea-ice volume and extent for 1979 (Section 3.1.1)."

19-20 "were adjusted to simulate reasonable": Same as in the previous sentence.

By reasonable, we mean compared to previous studies. As suggested, we inserted a reference to the later section where we show comparison in the revised manuscript (P13 L3):

"Lastly, two parameters of CanOE (Tref and chldeg) were adjusted to simulate reasonable annual primary production in the Arctic Ocean (Section 3.2)."

p12
18 Repetition: "the the Pan-Arctic"

Revised as suggested.

p13
7 "diagnose potential drifts": Unclear to me. In addition, how do you separate this (effects such as potential mass non-conservation in some tracers) from the inherent "spin-up dynamics" (meaning the adjustment from some relatively arbitrary initial condition to a state that is permitted by model dynamics)?

As discussed in response to the major comment 1 above, our simulation is too short to separate the drifts from the inherent spin-up dynamics quantitatively. Thus, we removed this sentence and revised this section (Appendix B).

p15
4ff. What is the reason for not simply masking the respective regions from the PIOMAS dataset in order to compare your model outputs across the same regions?

As suggested, we interpolated the PIOMAS product onto our model grid to perform grid-to-grid comparison over the same domain (Figure 7), and accordingly updated the time series comparing the sea ice volume and extent between our model and PIOMAS (Figure 6). Because the spatial coverage of SIIV3 differs from our model domain, we excluded its ice extent time series from the figure in the revised manuscript.

28 Is the occurence of such thick ice off Siberia in PIOMAS discussed in any of the literature about PIOMAS? I am mostly asking out of curiosity, I agree with your conclusion that it is likely an artifact.

This thick ice off Siberia is also present (although smaller magnitude) in PIOMAS for August 1993 (Figure 10b of Zhang and Rothrock (2003)). However, it was not discussed in that literature. We also did not find any other literature discussing this feature.

Zhang and Rothrock (2003), Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Monthly Weather Review, 131*(5), 845-861.

p18
13 "confined to shelf regions" (excluding the Barents Sea)

Revised as suggested.

29 space between 'Figure' and '8d'

Revised as suggested.

29 "in both qualitative and quantitative ...": I feel "in quantitative ..." is enough already
as it entails the other.

Removed "qualitative".

p20
Section 3.4: "subsurface chlorophyll a maximum": Maxima at around 5–10 m are hardly
comparable to the severals tens of meters usually found in the Arctic; splitting by Atlantic/Pacific
sectors might be worthwhile here due to their very different hydrographies,
just as extending the profiles in the plots deeper (to e.g. 40 m).

We agree that averaging over smaller regions would make the results closer to what we find in
observations. We looked at the vertical time series at various locations and confirmed that the
subsurface chlorophyll maxima (SCM) can be found at deeper depths, comparable to observations (see
the figure below, which shows the vertical time series at the grid cell corresponding to 170 degree W
and 70 degree N in the Chukchi Sea. The SCM depth (>20 m) is comparable to observations, e.g.,
Brown et al. 2015). A detailed discussion on SCM simulated by CanOE is included in Steiner et al. (in
prep.).

Brown et al. (2015), Characterizing the subsurface chlorophyll a maximum in the Chukchi Sea and
Canada Basin. *Deep Sea Research Part II: Topical Studies in Oceanography*, *118*, 88-104.

However, the focus of the analysis in this section is to quantify the impacts on air-sea gas (DMS) flux
at a larger scale (pan-Arctic), and so the pan-Arctic averaging was used. We revised the manuscript to
discuss this point (P19 L22):

"Note that the meltwater lens and the subsurface maxima are respectively thicker and shallower than
those observed by field measurements (e.g., Brown et al., 2015) because of averaging over the pan-
Arctic domain. The purpose of this spatial averaging is to quantify the impacts at a larger scale rather
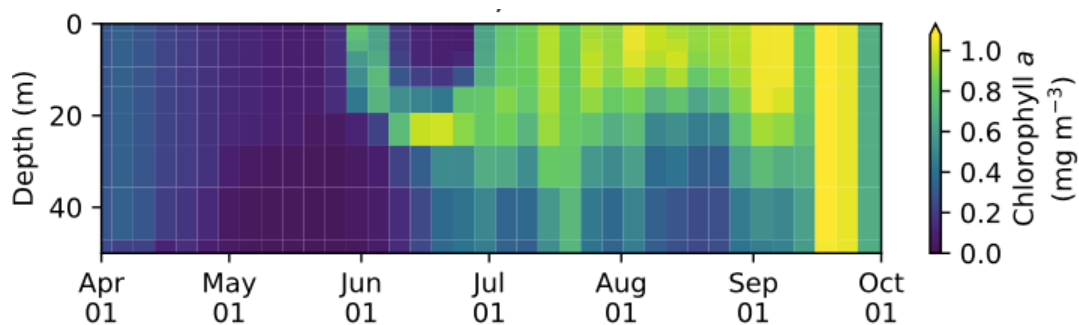than assessing localized effects."

Fig. 9 gives the impression that the surface mixed layer is not deep enough (compared to what I think observations would show), hence surface mixing might be too weak in the model. Could this also be the reason for strong DMS gradients in the upper 10-15 meters?

As in the previous comment, Fig.9 shows pan-Arctic averages which tend to average/smear out mixing depths which could be much higher at specific locations and points in time.

Do you think including under-water PAR irradiance in Fig. 9 could help with the interpretation of the results?

We do not think that the addition of PAR to Fig.9 would aid in the interpretation of the results. It is already clear from the existing figure that the SCM is associated with the depletion of surface nitrate (chlorophyll maxima following the nitracline).

24ff. I am not sure what these gradients in DMS tell me. You state "DMS flux would be underestimated", but can you spell out what exactly is being underestimated? The real flux, the modelled flux? Is the DMS flux calculation currently based on the 1-m concentration or the 12-m concentration, and why does the formula not use the one that would be appropriate for your model? How many percent would the underestimation be? 15% under- (over-?) estimation in the surface ocean DMS concentration as such does not sound so bad, but the effect on the flux probably depends on the air concentration.

Here, we meant that the modelled flux would be underestimated if the model had coarse vertical resolution (i.e. 12 m), as the averaging over 12 m dilutes the DMS concentration. The difference in DMS concentration linearly translates into the modelled flux by definition of the flux paramterization. The parameterization would be independent of model vertical resolution. We agree that the flux depends on the atmospheric concentration as well, which can be important if atmospheric boundary layers are low and the atmospheric concentration is high (Steiner & Denman, 2008). However, in general, the atmospheric concentration of DMS is orders of magnitudes smaller than the surface-water concentration, therefore should have negligible effect on the flux difference between the cases. Neglecting these atmospheric conditions is a common approach for ocean models that are not coupled to atmospheric models. We revised this paragraph accordingly (P20 L4):

"Here, the averaging over a thicker layer results in dilution of the DMS concentration in the uppermost layer represented in the model. Considering that this difference is present primarily during the ice melt period, and therefore that the sea-surface DMS is released into the atmosphere, the modelled sea-to-air DMS flux would be underestimated by a similar amount in the absence of fine vertical resolution in the upper water column."

Steiner and Denman (2008): Parameter sensitivies in a 1-D model for DMS and sulphur cycling in the upper ocean. Deep Sea Research Part I: Oceanographic Research Papers.

p22
8 "mm d^(-1)": Does this mean amount of meltwater equivalent, snow...?

Yes. We added this information in the revised manuscript (P20 L18):

p23
5 "extremely-low": extremely low

Revised as suggested.

14ff. So if these model-internal parameters can have such a big and non-intuitive effect on accumulated snow depth, why do you tune/adjust/modify the input snowfall dataset instead of the model parameters? (I also kept wondering, what happens to the thickness of the snow cover during ice dynamics (i.e. convergence)? Is total snow mass being conserved?)

The model sensitivity to the parameter nn_fsbc is somewhat unexpected because snow accumulation should not be sensitive to this parameter as long as the frequency of computation of surface boundary conditions (defined by nn_fsbc) is higher than that of the input snowfall dataset. While tuning this parameter did improve the simulated snow depth (as demonstrated by EXP1 and EXP2), this tuning is quite arbitrary without known constraints and therefore is not preferable. Furthermore, the tuning might have other implications which we did not assess in the present study. On the other hand, the usage of high-frequency atmospheric forcing is desirable simply because it is more realistic. We revised the manuscript to add this discussion (P22 L13):

"This high sensitivity to the choice of nn_fsbc is somewhat unexpected given that the tested range (1-10 time steps, equivalent to 20-200 minutes) is far less than the temporal resolution of the CORE-II dataset. A more detailed analysis of the model sensitivity to nn_fsbc is outside the scope of this study. Nevertheless, this analysis suggests that the issue with the usage of monthly or climatological-daily snowfall dataset can be resolved by tuning this parameter (as demonstrated in EXP1 and EXP2). However, the tuning of this parameter without known constraints is quite arbitrary and might have other implications for modelled dynamics. The usage of high-frequency atmospheric forcing dataset is recommended whenever possible to prevent the issue discussed here."

Total snow/ice mass is conserved in LIM2 and the convergence is simulated for sea-ice physical variables.

Section 4.2: Briefly remind the reader what i0 is.

Revised as suggested.

p25
14 "This value" referring to Castellani et al.'s 0.3?

Yes. We replaced "This value" with "This value (0.3)" in the revised manuscript.

p26
Section 4.3: I am a bit confused as to what you mean by "neglecting" advection/eddy diffusion. What happens instead when ice concentration changes from or to zero in a grid cell? Are sea ice BGC parameters somehow reset, do they pick up from where they were last time, whenever ice re-appears? I am especially thinking of nutrient and

biomass budgets. Are these still being conserved?
Again, note my earlier reservation about calling this "advection", you may want to specify that you are talking about moving sea ice, and hence moving BGC variables around with the ice (I think).

Neglecting advection/eddy diffusion means tendencies of BGC variables associated with horizontal motions of are artificially suppressed. Comparing this sensitivity run with the standard run allows us to quantify the contribution of horizontal transport to the overall budget. When ice concentration changes from zero to a non-zero value, sea-ice BGC concentrations are set to those in the underlying water column as described in Sec.2.3.2 of the original manuscript. As described in that section, this formulation does violate the mass conservation but the effect is small given the thin-layer (3 cm) of sea-ice BGC. When ice concentration changes to zero, sea-ice BGC concentrations are set to zero as they are all lost in the grid cell.

As suggested, we revised the manuscript to specify that we are talking about moving sea ice (P26 L5):

"These results indicate that the overall effect of horizontal transport associated with moving sea ice over the pan-Arctic is an increase in these quantities."

Section 4.4: Which parameter(s) is/are being modified concretely to "neglect" "the shading"?

We neglected the shading effect by setting the light extinction coefficient for ice algae to zero. We added this information in the revised manuscript (P26 L20):

"In EXP5, the shading effect of ice algae on light transfer through the ice is artificially suppressed in order to assess its impact on under-ice NPP. Effectively, this is done by setting the light extinction coefficient for ice algae to zero (Equation 15 of Mortenson et al. (2017))."

p28
31ff I think you can be more explicit here: What I understand is that ice algae shading can affect pelagic bloom *timing*, but will not affect annual pan-Arctic NPP. Regarding "patchiness of ice algal distribution": If you mean patchiness at the subgrid scale then I do not think your model accounts for this anyway; so perhaps it should not be part of the argument.

Revised as suggested to be more explicit by adding the word "timing".

By "patchiness of ice algal distribution", we do not mean sub-grid scale, but the pan-Arctic distribution is patchy (i.e. confined to shelf regions) as described earlier in this section. We clarified this point in the revised manuscript (P28 L24):

"However, given the patchiness of ice algal distribution (mostly confined to shelf regions) and the control of the light through the open-water fraction, the impact of the shading on the pan-Arctic under-ice annual NPP is negligible."

p29
Fig.15: I think including ice+snow transmissivity, or snow depth, or lead fraction, or something else of the sort should be included here to separate the rise in under-ice

PAR into the two factors "increasing sun angle" and "more transparent ice cover" as
the season progresses.

While separating the increase in under-ice PAR into the two factors can be useful, we do not think that
including any of those suggested variables would allow us to distinguish the two because those
suggested variables are dependent on both of the two factors. For example, increasing sun angle would
lead to more melting which leads to more transparent ice cover. Thus, we kept the figure as is.

p30
4 "were necessary to properly simulate": I am unsure what "properly" means here.

Replaced "properly" with "adequately".

Referee 2

Interactive comment on "CSIB v1: a sea-ice biogeochemical model for the NEMO community ocean
modelling framework" by Hakase Hayashida et al. Anonymous Referee #2
Received and published: 2 October 2018

This paper describes one more Pan-Arctic coupled model. I think it is a well-written paper and it seems
to fit the scope of the journal. I have a few general comments/questions (below) and several minor
comments/corrections made directly on the paper pdf (attached). I think this paper may be accepted
after minor to moderate modifications. I suggest that authors address my general comments below to
help the reader understanding better some of the modeling options taken here. This can be done with
some small addition of text to the original manuscript. I also suggest that authors have a look at my
minor comments/questions and choose the best way to address them. In general these should be quite
easy to handle.

We thank the referee for his/her positive comments and constructive feedback on our manuscript. We
revised the manuscript based on the comments/corrections made directly on the submitted manuscript
as much as possible. Please see the attached submitted manuscript where we added our responses to the
referee's comments/corrections.

Below we provide responses to the general comments/questions.

General comments/questions
1) I think that the effort made here to test the model and compare it with observations is quite
important. This is frequently lacking in modeling studies that emphasize obtained results without a
proper assessment of model performance. The modes implemented here is compared with observations
temporarily and spatially (both horizontally and vertically). I think this is a good example. I guess
authors could improve a bit the comparison by including some statistical measures of model
performance such, as for example, the Nash Sutcliffe model efficiency and the Percentage model bias
synthesized in Allen et al. (2017). In this case they could perhaps make comparisons across time and
space simultaneously and come up with some objective qualitative assessment of model performance.

We appreciate the suggestion to perform more objective analysis of our model results that can be helpful for readers. As suggested, we quantified the model performance using the percentage model difference as a measure. This measure was used to compare: 1) the annual-mean ice volume and extent (Section 3.1.1; P14 L17); and 2) the March- and September-mean ice thickness distributions (Section 3.1.2; P16 L2 and L12) between our model and PIOMAS in the revised manuscript. In short, the comparison 1 was better than the comparison 2, indicating that our model agrees well with PIOMAS in terms of simulating the overall structure, but not so (although still reasonable; 30-40% difference) in spatial patterns.

2) Why a "new" Pan-Arctic model? I think it would help if authors justified the reasons for selecting a specific sea-ice biogeochemistry model, especially considering that the selected model simulates only bottom-ice biogeochemistry while, since the 90s, several authors adopted vertically resolved sea-ice biogeochemistry models, suggesting the importance of the stocks of algae, nutrients, etc., in upper ice layers through their contribution to vertically integrated production (e.g. Arrigo et al., 1993; Vancoppenolle et al., 2010; Pogson et al., 2011; Duarte et al., 2015). I have the impression that the emphasis on bottom sea-ice biogeochemistry comes from the larger availability of studies on land-fast ice, with a typical large accumulation of ice algae at the bottom few centimeters. However, studies in the pack ice over the open ocean show quite a different picture, where maximum may occur at various depths (e.g. Melnikov et al., 2002; Olsen et al., 2017).

The purpose of our model development work here is its in-line coupling implementation specifically into NEMO, which has not been done previously.

While vertically resolved sea-ice biogeochemistry models are more desirable, we argue that its implementation into a 3-D modelling framework is impractical due to computational costs. Note that all of the previous model studies that adopted the vertically resolved sea-ice models are based on 1-D column frameworks. To the best of our knowledge, the CICE model is the only model system that has the capability to vertically resolve sea-ice BGC (Jeffery et al. 2016). However, other 3-D sea-ice physical models, such as the LIM model used in the present study, do not resolve multi vertical layers (only 2 sea-ice layers), hence vertically resolving sea-ice BGC seems impractical. We revised Section 2.3 to explain this point and acknowledge the fact that biomass above the bottom ice layer can be substantial (P7 L5):

"Sea-ice biogeochemical processes are assumed to take place in a layer of fixed thickness at the ice base. Hence, this bottom- ice biogeochemical layer is not explicitly modelled and does not correspond to one of the two ice layers in LIM. Although algal biomass in ice core samples above this layer can be substantial (e.g., Melnikov et al., 2002; Olsen et al., 2017), resolving vertical distributions of sea-ice biogeochemistry in 3-D models is computationally impractical at present."

Jeffery et al. (2016): Biogeochemistry of Cice: The Los Alamos Sea Ice Model Documentation and Software User's Manual Zbgc_colpkg Modifications to Version 5

3) Why testing the model for a period when available data is much less than in recent years and, therefore, it becomes much more difficult to properly evaluate model performance? In fact and with regard to the biogeochemical data, author's comparisons with other data sources may be biased by the differences in the temporal frames of various studies.

We agree that the model can be better evaluated for more recent period than the period considered in

the present study. We extended our simulation up to 2015 and evaluated with more recent observational data (Hayashida 2018, PhD thesis), but these results are planned to be published as separate papers considering the contents of the present study, which is intended as a model description paper. We revised the manuscript to clarify these points (P2 L31):

"We note that this study is intended as a model description paper, and the analysis focuses on results for the year 1979, corresponding to the end of a decadal model spin up. The analysis of the simulation beyond 1979, in which more observational data are available for evaluation (Hayashida 2018b), is planned to be published as a journal article separately."

Hayashida, H. (2018): Modelling sea-ice and oceanic dimethylsulfide production and emissions in the Arctic, PhD thesis, University of Victoria, Victoria, BC, Canada, http://dspace.library.uvic.ca/handle/1828/10486

References Allen, J. I., Holt, J. T., Blackford, J., & Proctor, R. (2007). Error quantification of a high-resolution coupled hydrodynamic-ecosystem coastal-ocean model: Part 2. Chlorophyll-a, nutrients and SPM. Journal of Marine Systems, 68(3-4), 381-404. doi: C2 10.1016/j.jmarsys.2007.01.005 Arrigo, K. R., Kremer, J. N., & Sullivan, C. W. (1993). A Simulated Antarctic Fast Ice Ecosystem. JOURNAL OF GEOPHYSICAL RESEARCH, 98, 17. Duarte, P., Assmy, P., Hop, H., Spreen, G., Gerland, S., & Hudson, S. R. (2015). The importance of vertical resolution in sea ice algae production models. Journal of Marine Systems, 145, 69-90. doi: 10.1016/j.jmarsys.2014.12.004 Melnikov, I. A., Kolosova, E. G., Welch, H. E., & Zhitina, L. S. (2002). Sea ice biological communities and nutrient dynamics in the Canada Basin of the Arctic Ocean. Deep-Sea Research Part I-Oceanographic Research Papers, 49(9), 1623-1649. doi: Pii S0967-0637(02)00042-0 Doi 10.1016/S0967-0637(02)00042-0 Olsen, L. M., Laney, S. R., Duarte, P., Kauko, H. M., Fernandez-Mendez, M., Mundy, C. J., . . . Assmy, P. (2017). The seeding of ice algal blooms in Arctic pack ice: The multiyear ice seed repository hypothesis. Journal of Geophysical Research-Biogeosciences, 122(7), 1529-1548. doi: 10.1002/2016JG003668 Pogson, L., Tremblay, B., Lavoie, D., Michel, C., & Vancoppenolle, M. (2011). Develop- ment and validation of a one-dimensional snow-ice algae model against observations in Resolute Passage, Canadian Arctic Archipelago. Journal of Geophysical Research- Oceans, 116. doi: Artn C04010 10.1029/2010jc006119 Vancoppenolle, M., Goosse, H., de Montety, A., Fichefet, T., Tremblay, B., & Tison, J. L. (2010). Modeling brine and nutrient dynamics in Antarctic sea ice: The case of dissolved silica. Journal of Geophysical Research-Oceans, 115. doi: Artn C02005 10.1029/2009jc005369

Please also note the supplement to this comment: https://www.geosci-model-dev-discuss.net/gmd-2018-191/gmd-2018-191-RC2- supplement.pdf

Please see the attached submitted manuscript where we added our responses to the referee's comments/corrections.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-191, 2018.

Referee 3

Interactive comment on "CSIB v1: a sea-ice biogeochemical model for the NEMO community ocean modelling framework" by Hakase Hayashida et al.
Anonymous Referee #3
Received and published: 4 October 2018
Geoscientific Model Development Discussions (Ms. No. gmd-2018-191) Title: CSIB v1: a sea-ice biogeochemical model for the NEMO community ocean modelling framework Authors: Hakase Hayashida, James R. Christian, Amber M. Holdsworth, Xianmin Hu, Adam H. Monahan, Eric Mortenson, Paul G. Myers, Olivier G. J. Riche, Tessa Sou, and Nadja S. Steiner

#Summary
The authors described a newly developed 3-D sea ice biogeochemical model embedded on the pan-Arctic NEMO and CanOE framework and then evaluated its basic performance with available validation datasets. Several sensitivity experiments were also performed to verify unknown parameter values such as shortwave absorption in the snow surface layer and light shading effect of ice algae. The paper was well written, and most parts of the presented analyses are quite reasonable. On the other hand, the target period of 1970s seems to be too old and short for model evaluation with reliable observational data. Computational cost should not be a primary reason because a simulation for more recent years can also be performed in the same manner. If there are another circumstances, please explain specifically.

We thank the referee for his/her positive and helpful comments on our manuscript. We revised the manuscript to reflect the referee's suggestions as much as possible. Below, we provide our responses to the referee's comments, including the responses to the questions posed in the #Summary section.

#Major Comments
One of my major concerns is that the model target period for 1969-1979 is so old. Such an experiment is regarded as a spin-up one but is not usually chosen for main analyses due to lower accuracy of atmospheric forcing datasets and lack of field measurements. In addition, most results in the sensitivity experiments are seen only in 1979. Hence potential readers cannot judge whether the presented anomalies are typical or unique features. The authors may have insufficient computational resource. Even in that case, they can run the model for more recent years (e.g., from 2000 or 2010) and/or use decadal mean forcing (e.g., 1990s or 2000s). Then discussion with decadal changes would provide more interesting scientific findings. Is it impossible?

We agree that the model can be better evaluated for the more recent period than the period considered in the present study. We did in fact conduct the simulation up to 2015 and performed a detailed evaluation with more recent observational data (Hayashida 2018b PhD thesis). We initially had planned one publication combining the technical information and the full evaluation of model variables, but the amount of information was simply too much to put it into one paper. Hence, we revised it to have one initial paper including the model description as well as several technical aspects (sensitivity experiments) focusing on the spinup period and one paper with detailed evaluation and science content covering the time period 1979-2015. We revised the manuscript to clarify these points (P2 L31):

"We note that this study is intended as a model description paper, and the analysis focuses on results for the year 1979, corresponding to the end of a decadal model spin up. The analysis of the simulation

beyond 1979, in which more observational data are available for evaluation (Hayashida 2018b), is planned to be published as a journal article separately."

Hayashida, H. (2018): Modelling sea-ice and oceanic dimethylsulfide production and emissions in the Arctic, PhD thesis, University of Victoria, Victoria, BC, Canada, http://dspace.library.uvic.ca/handle/1828/10486

Second concern is that the authors show GPP for ice algae and NPP for phytoplankton. It is a little confusing. I know that some reasons are described in the manuscript and that their difference is minor. However, I expect that the authors rerun their model to produce NPP for ice algae or GPP for phytoplankton. It is important to reduce extra cares for potential readers throughout the manuscript.

We appreciate the concern raised by the referee regarding GPP vs NPP. We re-visited the literature and also had discussions with colleagues, and we are now convinced that the first (source) term in the model equation for ice algae is representative of NPP for ice algae. In other words, we now consider what we considered GPP in the original manuscript as NPP. The logic for this decision is as follows:

The specific growth rate of ice algae (d^-1) prescribed in our model is based on Eppley (1972), who derived this parameter based on measurements of particulate matter production. From Sakshaug et al. (1997) and Hashimoto et al. (2005) (quoted below), it is understood that particulate matter production is a measure closer to NPP than GPP. Therefore, we consider that the first term in the model equation for ice algae, the product of growth rate and biomass, as NPP.

"*Net primary productivity is related to the 'growth rate', which can be defined as the net turnover rate for particulate carbon (not including production of DOC), provided that the cells are in steady-state (balanced) growth (Eppley, 1981)*". --Sakshaug et al. (1997)

"*Generally, the primary production is estimated from the rate of uptake of inorganic carbon into particulate carbon and/or the rate of evolution of oxygen into the water. In incubations of 24 h, the former method is considered to provide the values closest to net primary production (NPP), while the latter comes closest to gross primary production (GPP) (e.g., Falkowski and Raven 1997).*" --Hashimoto et al. (2005)

Eppley (1972): Temperature and phytoplankton growth in the sea.
Sakshaug et al. (1997): Parameters of photosynthesis: definitions, theory and interpretation of results.
Hashimoto et al. (2005): Relationship between net and gross primary production in the Sagami Bay, Japan.

As a result of this decision, we revised the manuscript to replace GPP with NPP for ice algae and also revised Section 2.5.3 (Output) where we describe NPP (P13 L10):

"Ice algal NPP is assumed to equal the growth term in the model equation (Mortenson et al., 2017), as the specific growth rate associated with that term is derived from Eppley (1972). This rate is a measure of particulate production, which is considered to provide values closer to NPP than gross primary productivity (GPP) (e.g., Sakshaug et al., 1997; Hashimoto et al., 2005). Thus, the loss due to respiration is implicitly included in the growth term in the model equation for ice algae."

#Detailed Comments
[Introduction] >Line 6 in Page 2 Terminology of "mechanistic model" is unfamiliar, at least for me.

"Numerical model" is more standard, right?

We replaced 'mechanistic' with 'process-based'. 'process-based' distinguishes the model from other numerical models such as statistical models.

[Section 2] >Section 2.3 What are "unresolved (eddy diffusion) motions of sea ice" described here. I know that eddy diffusion in ocean models represents sub-grid seawater exchange driven by mesoscale eddies. However, lateral exchange of (solid) sea ice packs due to eddy activity hardly occurs except a part of marginal ice zones. Some sea ice models adopt eddy diffusion only to damp numerical instability depending on advection schemes. Do the authors assume other physical processes in the central Arctic for this term? I find a sentence "Readers are referred to Vancoppenolle et al. (2012)". But I appreciate that the authors explain more details, because a sensitivity experiment related with "eddy diffusion" term is presented in this manuscript.

As described by the reviewer, the diffusion term in the LIM model is designed to dampen numerical instabilities but can also be considered as turbulent-like component of sea ice motion. We revised this section accordingly (P7 L21):

"Diffusion, on the other hand, is represents transport by unresolved motions (random component of sea-ice motion analogous to turbulence in fluids; Thorndike, 1986; Rampal et al., 2009, 2016), and is often tuned to improve numerical stability."

>Section 2.3.3 Let me confirm whether ambient temperature controlling algal growth rate is kept at a freezing point of underlying seawater or not.

It is set to the temperature of underlying seawater. We revised the manuscript to clarify this (P8 L14):

"The growth rate of ice algae is dependent on ambient temperature (of underlying seawater),"

[Section 3]
>Line 21 in Page 18 Please correct a unit of Dupont's PP.

Revised as suggested.

>Line 1 in Page 20 What kinds of formulation or parameters are necessary to represent mat and stand contributions? Why did not the authors consider them in the present model?

By mat and strand communities, we are referring to macroscopic ice-algal aggregates (e.g. *Melosira arctica*) that are different from ice-algal blooms not just in terms of size but also in terms of physiology (Assmy et al. 2013). Therefore, we would need to add another state variable in order to represent mat and strand communities. We did not consider them in the present model because their contribution to the pan-Arctic NPP is negligible compared to ice-algal blooms because of their limited area coverage (Assmy et al. 2013). Given this finding, the old estimates by Legendre et al. (1992) quoted in the manuscript are quite speculative also knowing the underlying assumption in their estimates (discussed in detail in Deal et al. 2011). We included this discussion in the revised manuscript for clarification (P19 L8):

"Although the upper end accounts for contribution from mat and strand communities that are not represented in our model, their contribution to the pan-Arctic production should be small as their

spatial distribution is generally localized (e.g., Assmy et al., 2013)."

Assmy et al. (2013): Floating ice-algal aggregates below melting Arctic sea ice. PLOS ONE.

[Section 4]
>Section 4.1 (EXP1 and 2) It took a time for me to understand the purpose of these sensitivity experiments. Is it right that CORE-II provides only monthly mean data for snowfall and that DFS snow- fall has daily mean time series only from 1979? Has daily climatology for 1979-2012 been used in the case of simulation before 1978? If yes, a simulation for more recent years does not cause this problem. Anyway, I recommend that the authors explain backgrounds for this analysis more clearly. In addition, this subsection describes only snow depth comparison. How about impacts of forcing choice on ice algal PP?

Yes, CORE-II provides only monthly mean data, while DFS provides daily mean data for individual years from 1979 and daily-climatological mean data prior to 1979. We revised the manuscript to clarify the purpose of these sensitivity experiments (P20 L12):

"Note that the temporal resolution of the snowfall and total precipitation fields in the CORE-II dataset is monthly. In EXP2, the snowfall and total precipitation fields over the period 1969-1978 are replaced by their respective 1979-2012 daily climatological values as in the original DFS dataset (Dussin et al., 2016). Comparing between EXP0 and EXP1 allows us to assess the impacts of atmospheric forcing (DFS vs CORE-II), while comparing between EXP1 and EXP2 allows us to assess the impacts of snowfall dataset (daily vs daily climatology) on modelled snow depth."

The impacts of forcing choice on ice algal PP were not assessed because we think that models should simulate reasonable snow depth before incorporating sea-ice biogeochemistry, as snow is a driver for light-limited ice algal growth. While we expect that EXP1 and EXP2 would have resulted in higher ice algal PP, quantifying these would not be so valuable because snow depth was not simulated realistically.

>Section 4.2 (EXP3) It may cause misleading that "Using these higher i0 reduces light limitation, and hence enhances ice algal primary production", even though light penetration through snow column is overestimated. Since formulation and parameter values of light limitation term largely differ between models as described in the previous paragraph, light intensity at the skeletal layer is not always directly linked to ice algal PP. I recommend that the authors describe this part more carefully.

As suggested by the referee, we revised the manuscript to describe this part more carefully (P23 L30):

"The overall impact of i0 on ice algal production depends on the choice of formulation and parameter(s) for the light limitation function as discussed previously."

>Section 4.3 (EXP4) Why is "space opening" necessary for new ice algal growth? It can also be considered that higher algal biomass causes higher PP as long as nutrient is available. Please explain a limitation factor. As mentioned above, the authors should specify more detailed processes of "eddy diffusion" of sea-ice biogeochemical state variables. In addition, can individual impact of this eddy diffusion term be estimated? Sea ice drift patterns also have large interannual variability depending on wind stress fields. Therefore, I am afraid that the presented anomalies in a single year are not representative in the pan-Arctic region.

Space opening helps ice algal growth because the ice algal loss term has quadratic dependency on biomass. Hence, the growth is more efficient at lower biomass. We agree with the referee that the nutrient advection can further promote higher PP in the productive regions, which may partly explain the no change in the nitrate concentration in these region (Fig. 14c in the revised manuscript). We revised the manuscript to clarify these two possible explanations (P26 L13):

"One possible explanation for these spatial differences is that the horizontal transport of sea ice takes ice algae out of regions of high productivity into regions of low productivity. This allows more efficient growth by maintaining the loss due to viral infection and aggregation (represented by the quadratic mortality term in the model) at relatively low values in the productive regions. Another factor is the horizontal transport of nutrients into these regions which are taken up by ice algae and results in the further increase in ice algal production."

As mentioned in the response above, we described sea-ice diffusion more carefully. Here, the focus is on the horizontal transport of sea ice (advection + diffusion) rather than advection vs diffusion, therefore, we did not save and evaluate model output for individual contributions. We agree that the interannual variability can be large due to wind stress fields, and therefore our results are not representative of long-term average. However, the quantities we presented are representative of the pan-Arctic average (for a particular year and using particular forcing and model). We revised the manuscript to note on the interannual variability (P26 L6 as well as P30 L9):

"However, we note that these values could be quite different in other years given the large interannual variability in wind stress fields driving sea ice drift patterns." (P26 L6)

"While we believe that these findings would be qualitatively similar in other years, it would be worthwhile to quantify their interannual variability." (P30 L9)

>Section 4.4 (EXP5) I do not understand that "the reduction in nutrient drawdown under regions of large ice algal biomass enhances nutrient advection into regions of low ice algal biomass". Why is nutrient advection enhanced by biogeochemical processes? Definition of bloom onset using bottom-ice PAR looks unnatural for me, because under-ice NPP is also calculated in the present model. The authors may try to compare with Castellani et al. (2017) more directly. But the target year is quite different between two studies so that this information is not so valuable.

We did more analysis on this and found that nutrient limitation is already so high in these regions of increased NPP that additional supply via advection should not have an impact. We then found that the NPP increase is dominated by small phytoplankton, but the reasons for this response of the modelled ecosystem to a perturbation to light are unclear. In the revised manuscript, we removed the possible explanation about nutrient, added a figure showing the increase in NPP dominated by small phytoplankton, and revised the text accordingly (P26 L29):

"However, in some regions, shading results in a slight increase in under-ice NPP which is dominated by small phytoplankton (Figure 16c). The underlying mechanisms for this response of the modelled ecosystem to a perturbation to light are unclear."

Definition of bloom onset using bottom-ice PAR was used in this analysis in order to directly compare with Castellani et al. (2017). We revised the manuscript to note the difference in the target year between the two studies (P28 L8):

"Furthermore, direct comparison is difficult due to the difference in the target year of simulation; Castellani et al. (2017) simulated 2012, while we consider 1979."

[Section 5]
>Line 5 in Page 30 Again, validation in a single old year is insufficient to appeal the model performance, especially in terms of sea ice volume and extent.

We agree with the referee that a single year comparison is not sufficient to appeal the model performance; this is left for future studies as noted in our response to the referee's major comment 1. We revised the manuscript accordingly (P28 L32):

"Results of the reference simulation (EXP0) were discussed and compared with previous studies, with a focus on the year 1979; more thorough evaluation of the model performance over the recent decades is planned for future studies."

 [Figure]
>Spatial maps I recommend that names of major countries or cities are overlaid.

We revised Figure 5 to add names of major countries/regions.

>Time Series Values of vertical axis should be smart (e.g., every 1 or 10).

We modified the vertical axis of the time series to be spaced uniformly throughout the manuscript.

>Figure 3 Magenta contours are hardly distinguishable from red contours.

We replaced the magenta with cyan and removed the 3000 m isobath, as such detail is not needed for the paper.

>Figure 4 Is total number of vertical layer same between NAA1 and NAA6 versions (i.e., 46 layers)? This figure is confusing. How about replacing vertical axis to water depths?

Yes, the total number of vertical layer is the same (46 layers) between the two configurations. The figure below shows the case where the vertical axis is replaced by water depths. We don't think that this figure is particularly better than the original figure. Comparing each layer between the two configurations in the original figure clearly demonstrates the point that NAA1 has finer resolution than NAA6, so we keep the figure unchanged.
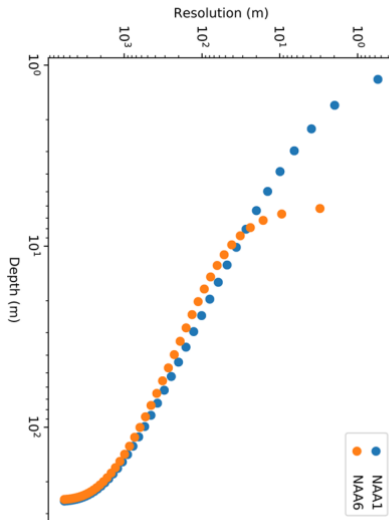
>Figure 6 Although area total amounts are shown in Figure 5, why are only GPP and NPP values in this figure plotted using "per unit area"? I think that the pan-Arctic average value is meaningless because of large spatial variability.

We agree that there is inconsistency in the representation of GPP and NPP between Fig5 and Fig6. However, we argue that either of these (areal integral or areal average) provides the same information about temporal variability by definition (areal integral = areal average * area of Arctic Circle, in this context), and therefore we do not think that they need to be changed especially since there is no observation data to compare the areal integral of GPP or NPP on a seasonal scale (e.g. m-2 d-1). As for the interannual time series (Fig5 in the submitted manuscript), we used annual areal integrals to compare with previous studies (as discussed in the manuscript).

>Figure 10 Please insert a zero line in (b). And I am wondering why the 1-m average minus the 12-m average shows negative for most periods.

We inserted a horizontal line at y = 0 in Fig10b. Regarding the second point, the caption was incorrect; it should have been 12-m average minus 1-m average. We appreciate the referee for catching this error. Both the figure and the caption were revised accordingly.

>Figure 16 Contrast of a color bar in (c) seems to be weak.

We changed the colormap to show more contrast in the revised manuscript.

[Table]
>Table 3 Unit of eddy diffusivity should be m^2/s.

Revised as suggested.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-191, 2018.