Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2019-177-AC1, 2019

Biogeosciences Discussions

Interactive comment on "Simulation of factors affecting *E.huxleyi* blooms in arctic and subarctic seas by CMIP5 climate models: model validation and selection" *by* Natalia Gnatiuk et al.

Natalia Gnatiuk et al.

natalia.gnatiuk@niersc.spb.ru

Received and published: 5 August 2019

Dear Prof. Heinze,

Thank you very much for your assistance with our manuscript: "Simulation of factors affecting E.huxleyi blooms in arctic and subarctic seas by CMIP5 climate models: model validation and selection" by N. Gnatiuk, I. Radchenko, R. Davy, E. Morozov, and L. Bobylev. We are very thankful to the anonymous reviewers and greatly appreciate their valuable comments and suggestions that considerably improve our manuscript. Regarding the comments from both reviewers, we decided to refuse from focusing on the E. huxleyi problematics per se and concentrate on the main goal of the manuscript,

which is a selection of those climate models that simulate most efficiently the state of abiotic parameters relevant to living conditions of the phytoplankton communities inherent in a number of seas at subpolar and polar latitudes. Accordingly, the Introduction is thoroughly recast. As both reviewers suggested us to improve the "Results and Discussion" section, we will move Figures 4, 6, and 7 to section "Materials and Methods" in order to better describe our methodology of climate models selection. Also, Figures 3, 5, and 8 are deleted as they are either a mere modification of presentations of some other akin figures or their presence in the manuscript is not so important. We also decided to add a new figure (#Figure 7) to section "Results and Discussion" as both reviewers suggested to cover in discussions all studied parameters and seas. So, Figure 7 displays a spatial distribution of biases in five parameters between models and reanalyses in six target seas. The biases are averaged over the vegetation season and 1979/1993-2005 period. Here we give an example of that figure illustrating sea surface salinity. We further will improve the section "Results and Discussion" following the comments from both reviewers. Please kindly find attached the responses to the reviewers and effected revisions, as well as a detailed specification of the changes we introduced. We are looking forward to hearing from you considering these changes and await further instructions. On behalf of the paper's co-authors

Best regards, Natalia Gnatiuk (and co-authors)

Response to Reviewers and Proposed Revisions:

We are very grateful to both reviewers for their constructive and valuable comments and very useful suggestions, which will greatly help in improving our manuscript. Concerning the comments of both reviewers on the choice of factors controlling phytoplankton blooms in general, and coccolithophore in particular, and other reviewers' comments directly related to the coccolithophore blooms, we fully agree with the arguments provided by reviewer #2 in section "general comments": "By not having a primary focus on E.huxleyi blooms in the Introduction, the reader will be able to recognize the wider implications of this extensive intercomparison of climate models – it will also alleviate

some of the major issues of neglecting “what else” underpins coccolithophore blooms and their occurrences.”

Actually, the main objective of the study was to analyze how CMIP5 climate models reproduce different oceanographic and meteorological parameters in the arctic and subarctic seas as well as to form a methodology for validation and selection of the optimal model sub-set. To have practical use of the results we have chosen for case study oceanographic and meteorological parameters that influence coccolithophores blooms in studied arctic and subarctic seas. Due to the fact that we did not consider in the article all the factors (including biotic ones) that influence coccolithophores bloom, we mistakenly paid too much attention to coccolithophores and the factors affecting their blooms. This resulted in shifting the paper’s focus away from the main goal of the study, i.e. to develop a methodology of validation and selection of climate models that simulate most accurately the abiotic conditions within the target marine areas. To mend the situation, we decided to refuse from focusing specifically upon the issue of coccolithophore blooms and put at the forefront the methodology of validation and selection of climate models. In the absence of a close connection to coccolithophores, the article indeed gains greater clarity and becomes focused on the substance of the research done on the comparative effectiveness of global climate models for specific marine objects. We corrected the manuscript according to the recommendations of the reviewers and tried to make the goal of our research clear and precise. Below we have presented all the answers to the comments and all text changes. We earnestly thank the reviewers for their critical comments.

Reviewer #1

General comments G.C.1: In my opinion, the choice of the variables to be validated in the models is incomplete and the chosen subset is not obvious when thinking about the factors controlling phytoplankton blooms. First, all possible drivers should be carefully introduced in the introduction (it is nowhere clearly stated in the manuscript what factors impact phytoplankton dynamics in general). Important variables such as (macro/micro)

nutrients and carbonate chemistry are not considered at all and it is not even thoroughly discussed why. Motivating the choice of drivers to be analyzed in this paper simply by referring to an individual study (Kondrik et al., 2019) which is currently in review for publication in Biogeosciences is not sufficient in my opinion. Additionally, reviewers of the manuscript by Kondrik et al. (2019) have raised similar concerns with regard to the chosen drivers considered in the analysis. In my view, the choice made for the manuscript at hand is a missed opportunity as the evaluated models can provide more comprehensive information on factors impacting phytoplankton/coccolithophore growth than the variables the authors chose here.

G.C.1 answer: As mentioned above, we believe that focusing on factors affecting coccolithophore blooms will take the readers away from the true purpose of the publication. Also, as we provide an incomplete list of factors controlling phytoplankton blooms, we have decided not to use in the manuscript formulations like "factors controlling/affecting coccolithophore blooms". Instead, in the article, we will emphasize that we analyze the models for a number of meteorological and oceanographic parameters. We believe in this case the results may be of interest to a wider readership.

We would like to add that at the beginning of the research we also planned to consider different biogeochemical variables. But CMIP5 models have only monthly outputs for ocean biogeochemical variables, whereas, in this study to develop a methodology for selecting climate models, we employed daily data. Besides, we did not find sufficiently reliable observational data or reanalysis for 20-years period and for all 6 study seas for validation climate models.

G.C.2: In the current version of the manuscript, a discussion of the results is completely missing. While section 3 is called “Results & Discussion”, it currently only represents a description of the Figures, without putting the results into the context of previous literature or how the results impact the overall motivation for the study (assessing the potential future development of E. huxleyi blooms). It is not clear to me e.g. what modellers should take away from their analysis. In a revised version of the manuscript, I

suggest to include a thorough discussion on e.g. the sensitivity of the resulting model combinations on chosen thresholds in the ranking, the impact of the identified model biases on coccolithophore blooms, the impact of neglecting important forcing factors (nutrients, carbonate chemistry) or biotic interactions (which cannot be assessed with this approach as opposed to when coccolithophores are included as an explicit functional type to the model).

G.C.2 answer: We agree that the discussion of the results is not properly presented and needs to be improved.

G.C.3: Regarding choices of the presentation of the results, I would personally like to see more than just temperature in the Barents Sea to be included in the main text (the choice of the figures could be reconsidered, especially given the title of the manuscript). The current choice makes it very hard for the reader to assess how the representation of present-day coccolithophore blooms in these models is potentially affected by biases of all variables impacting phytoplankton dynamics (and not just temperature). Including more detail in the study at hand will also make the assessment in a follow-up paper on future changes easier.

G.C.3 answer: Based on one example - the temperature in the Barents Sea, we aimed to describe the methodology of climate model validation and selection in detail. Of course, from the point of view of assessing how climate models represent present-day coccolithophore blooms, such choices of presentation of the results is very uninformative. But we have tried to illustrate with this example, each step of the model's validation and the selection and to show the spread of the model's values for each selection criterion. Using this approach, we intended to prove the need for a comprehensive analysis that is not confined solely either to the seasonal cycle or inter-annual variability or trends or spatial errors.

G.C.4: Overall, I think that the literature review in the introduction on factors controlling coccolithophore blooms in the arctic/subarctic (or North Atlantic) and possible drivers

for observed changes in their distributions is not comprehensive enough in its current form. In my detailed comments below, I suggested a few papers that could be considered in my view – a result of a very brief literature search I have done (as I am not 100% familiar with the literature of the arctic/subarctic), but this list is by no means exhaustive. The authors should revise the manuscript accordingly, as this might also help to motivate why certain variables are (or are not) considered in their study.

G.C.4 answer: We will work on the Introduction part.

G.C.5: Throughout the manuscript, the writing needs to be more concise and to the point. Often, it is not clear to the reader why certain information is given, i.e. what the relevance is for the study at hand or what the take-away message is (see detailed comments below of e.g. the introduction). The authors should especially revise the result section, which is currently a list of brief descriptions of the Figures without making it clear enough why they were chosen to be included and what the key message for each Figure is, which makes this section quite hard to read in its current form. Ultimately, all figure captions are currently incomplete as they do not describe what is actually shown in the respective figures. G.C.5 answer: We will improve the manuscript and the Results and Discussion section accordingly.

Detailed comments of Reviewer #1:

Abstract:

D.C.1 p. 1, L. 9: Please add ocean acidification here. It is not only the effect of global warming that can be expected to impact future coccolithophore blooms.

D.C.2 90 p. 1, L. 10: I find this statement on the aim of the paper misleading, as none of these models (or very few) includes an explicit parametrization of coccolithophores to the best of my knowledge. I think it you need to state more clearly what exactly you do here. You don't actually assess the blooms, but only how well the models reproduce the present-day environmental conditions that favour coccolithophore blooms. Also,

please be precise here what you mean by “optimum combination”.

D.C.3 p. 1, L. 11: This last statement of the paragraph is misleading because you don’t actually address this in the study. Please point out that this is future work that can/will be done following this study. Additionally, please add a “potential” in the last part of the sentence 100 “[. . .] potential future changes [. . .] can be assessed.”

D.C.1-3 Answer:

We thank the reviewer for the comments and re-wrote the paragraph: “Currently, there are a large number of climate models that give projections for various oceanic and meteorological parameters in the Arctic. However, their estimates often differ in absolute values in individual sea areas in comparison with the historical reanalysis data. The main goal of this study was to identify the optimal ensemble of models that most accurately reproduce the selected abiotic parameters inherent in six selected seas, viz. – Barents, Bering, Greenland, Labrador, North and Norwegian Seas.”

Specific comments of Reviewer #1:

p. 1, L. 14: Please delete the “complex” or describe what methodology you’re using. I suggest to rephrase to something along these lines “Here, we present the validation of 34 CMIP models over the historical period. Furthermore, we present the procedure for model selection, which is based on their skill to represent important forcing factors for coccolithophore blooms.”

Answer:

We agree and rephrased the sentence in the following form: “Here, we present the validation of 34 CMIP models over the historical period. Furthermore, we present the procedure for model selection, which is based on the models’ skill to represent some important abiotic factors controlling phytoplankton blooms in the selected oceanographic and meteorological in arctic and subarctic seas: sea surface (i) temperature and (ii) salinity (averaged over the top 30 m); (iii) wind speed at a height of 10 m above the

surface; (iv) ocean surface current speed; and (v) surface downwelling shortwave radiation.”

p. 1, L. 15: Are these five factors really known to be the dominant factors impacting coccolithophore dynamics in the arctic and subarctic?

Answer:

We consider these five factors as one of the important abiotic controlling factors in the studied seas without a direct reference to coccolithophore ecology.

p. 1, L. 16: The chosen set of environmental factors to be validated in the models is not obvious to me. There is no rationale from my point of view as to why one would completely neglect nutrient fields and carbonate chemistry in the validation of the models (see general comments and detailed comments further down). Furthermore, I suggest to include a brief description what the environmental conditions wind speed, current speed, and salinity are proxies for as phytoplankton growth in these models is not a direct function of these variables.

Answer:

We agree that nutrients and carbonate chemistry are important for phytoplankton. However, we aim to select the appropriate CMIP5 models for one of the important abiotic factors. We added the following sentence: “Such parameters as wind speed, current speed and salinity do not explicitly affect the phytoplankton bloom development, however, they reflect the turbulence in the ocean upper layer and therefore they are liable to influence the nutrient availability and increase of solar radiation penetration during steady upper water stratification.”

p. 1, L. 16: Please check throughout the text: are you using sea surface salinity (as stated e.g. in the abstract and introduction) or the salinity averaged over the top 30m (as e.g. stated in section 2.1 or the caption of Figure 9)?

Answer:

We thank the reviewer for this comment. We added this information in the Abstract: “Furthermore, we present the procedure for model selection, which is based on their skill to represent important abiotic controlling factors for phytoplankton blooms in the selected oceanographic and meteorological in arctic and subarctic seas: sea surface (i) temperature and (ii) salinity (averaged over the top 30 m); (iii) wind speed at a height of 10 m above the surface; (iv) ocean surface current speed; and (v) surface downwelling shortwave radiation.”

p. 1, L. 20: “best models” in what respect? Please be precise here.

Answer:

We decided to delete the following sentence: “The selection of the best models was performed separately for each study area in the Barents, Bering, Greenland, Labrador, North and Norwegian Seas and for each of the five forcing factors affecting the coccolithophore blooms”.

p. 1, L. 22: I don’t understand this the statement about “30 combinations of most-skillful models were selected”. Selected for what? Additionally, I suggest to state how many models are considered within each combination.

p. 1, L. 23: “common” is used in what sense here? I don’t understand this. How do you define “high skill”? This is rather subjective. I suggest to rephrase.

Answer to above 2 comments:

We thank the reviewer for the above two comments. The sentences are now modified as follows: “In total, 30 combinations of high-skilful models were selected to study of the potential future changes in the abiotic parameters relevant to the phytoplankton blooms. The results show that there is no mutual optimal combination of models for six studied seas, nor is there one top-model, that has a skill in reproducing regional features across the combination of the five considered parameters and all arctic and subarctic seas. Therefore, in the Barents Sea the optimal model ensemble for current

speed - 7, sea salinity - 7, sea temperature - 7, wind speed - 7, solar radiation - 8; in the Bering Sea: current speed - 7, sea salinity - 8, sea temperature - 8, wind speed - 11, solar radiation - 8; in the Greenland Sea: current speed - 7, sea salinity - 11, sea temperature - 8, wind speed - 10, solar radiation - 9; in the Labrador Sea: current speed - 7, sea salinity - 8, sea temperature - 8, wind speed - 10, solar radiation - 8; in the North Sea: current speed - 8, sea salinity - 11, sea temperature - 8, wind speed - 9, solar radiation - 8; in the Northern Sea: current speed - 8, sea salinity - 9, sea temperature - 10, wind speed - 9, solar radiation - 8.”

p. 1, L. 25: What should e.g. modelers conclude from your analysis? I miss a statement on the broader implications of your study in the abstract.

Answer:

We decided to remove this paragraph as Reviewer #2 suggested us to do it.

Introduction:

p. 2, L. 2-5: Please include a brief description on how exactly coccolithophores impact the carbon cycle (as done for the sulfur cycle). Additionally, I don’t think Rivero-Calle et al. (2015) and Winter et al. (2013) are appropriate references for the biogeochemical impact of coccolithophores here, as these only describe changes in the biogeography and occurrence over time. Check e.g. Iglesias-Rodriguez et al. (2002) or Balch (2018) (and references therein) for the biogeochemical imprint of this phytoplankton group.

p. 2, L. 3: Please delete the “additionally”. You describe the impact on the sulphur cycle here.

p. 2, L. 6: It is not only essential to study E. hux. blooms, but coccolithophore blooms in general. E. huxleyi has not yet been introduced in this line of the text. Please change to “coccolithophores” instead of “E. huxleyi”

p. 2, L. 7: Please introduce the abbreviation “E. huxleyi” here.

p. 2, L. 9: Please add a reference to the temperature and salinity tolerance.

p. 2, L. 10: I suggest to add the more recent reference “Krumhardt et al. (2017)” here, as they provide the most recent compilation of the global present-day distribution of coccolithophores (to my knowledge).

p. 2, L. 11: Have coccolithophores really expanded because of ecosystem changes in the Arctic? Don’t you mean “as a result of recent changes in environmental conditions, coccolithophores have expanded poleward”? Please revise the logic in this sentence.

p. 2, L. 12: Henson et al. (2018) is not an appropriate reference here (they don’t talk about 165 the changes in E. huxleyi blooms in the cited paper). Please consider adding e.g. Rivero-Calle et al. (2015) here.

p. 2, L. 12: Winter et al. (2013) suggest that that the poleward expansion is mainly driven by temperature, salinity, or nutrients, but Rivero-Calle et al. (2015) and Krumhardt et al. (2016) suggest that carbonate chemistry matters as well. Please be more comprehensive in the discussion of possible drivers for the expansion.

p. 2, L. 14: Please be more precise here: When you say “E. huxleyi blooms have a high positive correlation with [. . .]”, do you mean the occurrence, their size, their duration. . .?

p. 2, L. 14: The description of controls on E. huxleyi blooms (and causes for its changes) is not comprehensive enough. I only did a very brief 10-minute search in the literature and found a number of papers that could be relevant for the introduction of this paper (only focusing on those of the northern hemisphere, i.e. disregarding the wealth of recent literature on Southern Ocean coccolithophore dynamics, see e.g. Balch et al., 2016, Nissen et al., 2018 and references therein): please have a look at e.g. Daniels et al. (2015), Harada et al. (2012), Oziel et al. (2017), and Smyth et al. (2004) (and references therein). I suggest to first describe the factors that contribute to phytoplankton/coccolithophore blooms in general (these are currently not introduced) and to then discuss what has been suggested for coccolithophores in general and in

the (sub)arctic in particular. Please motivate why you think nutrients and carbonate chemistry are not important as this is not at all obvious.

p. 2, L. 12-20: Please clearly differentiate between discussing drivers of present-day coccolithophore blooms as opposed to possible drivers of observed/future changes in coccolithophore distributions 190 and bloom dynamics.

p. 2, L. 20-32: I find it problematic to focus so much on a single paper here, especially as the discussed paper by Kondrik et al. (2019) has not yet been accepted. One of the main criticisms by the reviewers of that paper was the neglect of important variables as potential drivers of coccolithophore blooms (such as e.g. carbonate chemistry). I think the study at hand can be much more generally motivated, without going into the details of this specific one. To that aim, and similarly to the points raised in the review of Kondrik et al. (2019), the analysis in the manuscript by Gnatiuk et al. should be more comprehensive in the assessment of potential drivers of coccolithophore blooms, especially because the output from models is assessed here, which can provide information on all environmental variables impacting phytoplankton growth. There should not be a a-priori-restriction to the drivers assessed here without giving a good reason to do so. Please revise the introduction and the analysis in that respect.

p. 2, L. 2-32: The whole first part of the introduction does not provide a comprehensive summary of what is known about drivers of coccolithophore bloom dynamics and does not naturally result in the knowledge gap that will be assessed in this study. From my point of view, it should be substantially revised following the comments made above. Additionally, the models are not properly introduced. E.g. no reference to the CMIP is given. Furthermore, it should be clearly stated that none (or maximum a few, to be doublechecked) of the CMIP5 models includes an explicit parametrization of coccolithophores, which is why it is currently only possible to project potential changes of their blooms based on changes in environmental conditions (but note the recent paper by Krumhardt et al., 2019). This comes with the limitation that biotic interactions cannot be assessed, which should be clearly stated in the discussion section (see also

Krumhardt et al., 2017).

Answer to the above comments in the Introduction section:

We thank the reviewer for the valuable and constructive comments for page 2 of the manuscript. In view of the reviewer's comments, we decided to remove the first three paragraphs in the Introduction and we concentrate on the main goal of this manuscript, i.e. a selection and validation of the CMIP5 climate models against the available reanalysis data for one of the important abiotic parameters that influence phytoplankton blooms in six sub-arctic and arctic seas (Barents, Bering, Greenland, Norweigan, North and Labrador Seas). We will add some information similar to - "Today, there are a large number of CMIP5 models that differ in the reliable representation of specific parameters, which mainly characterize the physical conditions (abiotic parameters), forming, among other things, the habitat of biota in the surface layers of the seas at a regional scale. To identify the models from among the CMIP5 most accurately reproducing the values of one or another parameter, we used an approach based on the 25th, 50th and 75th percentile calculations of such statistical metrics as correlation coefficient, standard deviation, root means square deviation, the ratio of root mean square deviation to standard deviation of the observations as well as the spatial bias between the model data and reanalysis and spatial distribution of temporal trend of the parameters (Fu et al., 2013; Gleckler et al., 2008; Murphy et al., 2004; Ruan et al., 2019). This method gives the opportunity to evaluate multiple statistical metrics using the same approach (percentile-based) and further giving to a model a total score that allows us to select the top 25% skilful models for the abiotic parameters and seas."

p. 3: The first paragraph does not link well with the above. Please work on your flow in the introduction. Additionally, this whole page reads like it should be in the method section. Please revise and consider moving at least parts of it to the method section.

Answer:

Thank you for the comment. We will improve it.

p. 3, L. 7: How are the "best models" defined here? Please be precise what you mean.

Answer:

We changed to "high-skilled".

p. 3, L. 8: Please revise the sentence "These two approaches usually give a good result". Good in what respect? Please add references.

Answer:

We will modify this sentence and add references.

p. 3, L.11: Choosing this method implies the assumption that whatever model is representing present-day conditions best will also do the best job in projecting these into the future, doesn't it? I think this is important to state here.

Answer:

Thank you for the comment, we modified it as follows: "We assume that a model that successfully represents well the present-day conditions will do it well in the future. Therefore, we chose the second approach – a selection of climate models that properly simulate the regional features (spatial distribution) of the parameters under study (sea surface temperature and salinity, surface wind speed at 10 m, ocean surface current speed, and surface downwelling shortwave radiation). At that, it was important to define an appropriate methodology for the selection of the best model ensembles."

p. 3, L. 16-21: It is not clear to me what the take-away message of this paragraph is. How does the first approach, assessing how well models do in representing air temperature, sea level pressure, and precipitation help in the assessment of environmental factors impacting phytoplankton/coccolithophore growth? Please work on this paragraph and make it more specific to the goal of your work. Consider combining it with the next to avoid having a 2- sentence paragraph with no clear take-away message.

Answer:

We will improve it.

p. 3, L. 21 – p. 4, L.2: Again, the take-away message in context of your specific goals for the paper are not clear. Please re-write.

Answer:

We will do it.

p. 4, L. 3: Why do you conclude that? This is not clear to me from what you have presented so far in the introduction.

Answer:

We will improve it. We rephrased the sentence as follows: “We consider that the second approach is appropriate for our study using a score-based method with multiple criteria.”

p. 4, L. 3-5: You have not presented the differences in environmental conditions of the different focus areas. Please revise the introduction accordingly. I don’t understand what the second half of the sentence means: How can areas have a wide range of parameters?

Answer: We modified the introduction at this stage and we will improve it. Also, we modified the sentence and added additional information: “Since we deal with six arctic and subarctic seas with rather different environmental conditions and different abiotic parameters, it was decided to individually analyze each sea., It should be noted that this approach was applied not only to specific seas, but, moreover, to their specific areas that are the stable localizations of extensive growth of phytoplankton species both in spring (e.g. diatoms) and in summer-autumn (e.g. coccolithophores: Smyth et al., 2004; Kondrik et al., 2017). Therefore, it was of interest to evince the suite of such CMIP5 models that simulate most successfully the state of the physical environmental

factors constituting the abiotic conditions, in which the algal blooms develop. Thus, the obtained results can be useful for further improvement of existing phytoplankton ecological models as well as, in general, for modelling the temporal dynamics of ocean surface physical parameters at high latitudes in light of ongoing climate change.”

Methods

p. 4, L. 17-19: As mentioned before, I don’t understand based on what grounds you neglect the assessment of nutrients and the carbonate chemistry in the models.

Answer:

We agree that nutrients and the carbonate chemistry are important for coccolithophores bloom. We modified the sentence as follows: “As mentioned above, there were 5 parameters selected in arctic and subarctic seas for further comparison of CMIP5 models: sea surface temperature (SST) and salinity averaged over 0-30 m (SSS), surface wind speed at a height of 10 m (WS), ocean surface current speed (OCS), and shortwave downwelling solar radiation (SDSR).”

p. 4, L. 23-25: Did you include regional models, e.g. CORDEX? I can’t find it in Table 1. If you didn’t include those models, don’t make that statement here. I am bit confused. Please distinguish between regional and global models and state which kind you considered.

p. 4, L. 25: Did you only consider global models in the end? This is not clear from your description in this section. Please clarify.

Answer for two above comments:

We did not include regional models. We modified the sentence as follows: “Whereas the regional models have a high resolution of 11-25 km but they simulate only atmosphere or ocean separately and do not cover all six seas within the same model run. Therefore, we applied only global models in the study.”

p. 4, L. 25-26: I suggest to give the range of models available: number available for FFs ranged from X1 for variable Y1 to X2 for variable Y2 (see Table 1). What do you mean by “main characteristics”? Please rephrase.

Answer:

We modified the sentences: “In total, we considered 34 GCMs for the historical experiment, but the number of models available for concrete parameters varies: OCS – 7-8, SSS – 7-11, SST – 8-10, WS – 7-11, and SDSR – 8-9. The list of climate models used, their resolution and their availability for the studied parameters are presented in Table 1.”

p. 5, L. 4: replace “has been shown” by “have been shown”.

Answer:

We changed it.

p. 5, L. 6: Please choose a better description in the title than simply “methods”, maybe something like “model evaluation metrics”?

Answer:

We changed it (considering a suggestion from Reviewer #2) to “The regions under investigation and model evaluation metrics”

p. 5, L. 7: Please rewrite “regions under the study”. Add “Sea” behind “Norwegian”.

Answer:

We changed it: “The regions under study are six arctic and subarctic seas: Barents, Bering, Greenland, Labrador, North and Norwegian Seas.”

p. 5, L. 7-8: How do you define a bloom here? Please state this and add references. Additionally, you don’t state what data you base Fig. 1 on to define the blooms. Please clarify in the text and the figure caption. I suggest to draw the study regions into Fig. 1.

To help the reader localize the different subregions.

Answer:

We changed the sentence as follows: “The areas of study were selected for each sea on the basis of results obtained by Kazakov et al. (2018) for the E.huxleyi blooms based on Ocean Colour Climate Change Initiative dataset version 3.0 (Europea Space Agency) for the period from 1998 to 2016 (Fig. 1).” We changed the figure caption accordingly.

p. 5, L. 11: Do you mean model output here when you say “data? Please clarify.

p. 5, L. 12: I have a hard time believing that the blooming period lasts from January-December in the Bering Sea. What bloom definition is used for this?

Answer to two above comments:

We modified the sentences and added some information: “The periods of study were selected based on a sea-specific blooming periods which include all summer months and in some cases even early autumnal months: June-September for the Barents and Labrador seas, January-December for the Bering Sea, June-August for the Greenland Sea, May-July for the North Sea, and May-August for the Norwegian Sea (Kazakov et al., 2018). Thus, it is noteworthy that the results of the performed comparison of models can be used not only for the purposes of ecology-related fields of study but the overall forecasting of the region-specific climate interactions during summer season as well, taking into account that the selection of the climate models was carried out individually for each sea.”

p. 5, L. 10-14: I don’t fully understand why you’re restricting the analysis to the times and locations of identified E. huxleyi blooms under present-day environmental conditions for each sea (if the models do not necessarily reproduce the environmental conditions at these exact locations and times). Don’t you want to restrict the analysis to the observed environmental conditions at the times/locations of the blooms (i.e. the

observed environmental niche)? As a consequence, I am wondering why don't you define each subregion as a slightly larger area than currently done.

Answer:

Trying to answer the question that the reviewer has raised here, we entered the following sentences into the text: "The periods of study were selected based on a sea-specific blooming periods which include all summer months and in some cases early autumnal months : June-September for the Barents and Labrador seas, January-December for the Bering Sea, June-August for the Greenland Sea, May-July for the North Sea, and May-August for the Norwegian Sea (Kazakov et al., 2018). Thus, it is noteworthy that the results of the performed comparison of models can be used not only for the purposes of ecology-related fields of study but the overall forecasting of the region-specific climate interactions during Summer season as well, taking into account that the selection of the climate models was carried out individually for each sea."

p. 5, L. 17: The interannual variability of what exactly? The seasonal cycle/amplitude, summer average, average over blooming period, . . .? Please be precise here.

Answer:

We modified the sentence here: "The methodology of the validation of GCMs included the analysis of the climatological-mean seasonal cycle and interannual variability of selected parameters averaged over blooming period in each. "

p. 5, L. 18: The seasonal cycles [. . .]

Answer:

We changed it.

p. 5, L. 19: "but the interannual variability " of what?

p. 5, L. 19: Replace "sea" by "subregion"

Answer for the two above comments:

We changed it to: "The seasonal cycle was analyzed using multi-year averaged monthly variables for all months of year (i.e., a sample size of 12), but the parameters interannual variability of the parameters was analyzed based on monthly values of variables for the blooming periods only (sample size varied according to subregion and parameter combination, e.g., a sample size for SST in the Barents Sea was 108 – monthly variables from June to September during 1979-2005). "

p. 5, L. 23: Can you rephrase "RMSD-observations standard deviation ratio"? I have a hard time understanding what you mean here. Please consider to add the formulas to make it very clear.

Answer:

We changed the sentence: "Additionally, we calculated the ratio of root mean square deviation to standard deviation of the observations (RSR) – one of the model evaluation statistics that weighs the simulated data against the observations (Agosta et al., 2015; Golmohammadi et al., 2014; Moriasi et al., 2007; Murphy et al., 2004; Stocker, 2004)."

p. 5, L. 25-26: Rephrase to something like "For the assessment/evaluation of the interannual variability [. . .]"

p. 5, L. 26: Do you mean the difference in the spatial distribution of temporal trends between the model output and the reanalysis data? This sentence is not clear to me. Please rephrase to clarify.

Answer for the two above comments:

We changed it to the following sentence - "For the assessment of the interannual variability analysis we also calculated the difference in spatial distribution of temporal trends between the model output and the reanalysis data and spatial bias between the model data and reanalysis (Anav et al., 2013; Das et al., 2018; Fu et al., 2013; Gleckler et al., 2008; Kumar et al., 2015; Ruan et al., 2019)."

p. 5, L. 26-27: What exactly is “your percentile score-based model ranking method”? This method is defined nowhere in the method section up to this point. In particular, the description of this ranking method should be very clear (e.g. by including an overview listing the metrics are included in the ranking), as the main result of your study is based on this ranking.

315 p. 5, L. 31: Less than 25% of what? Please be precise. What do you base these thresholds on? It seems rather subjective to me. What is the effect of the choice on the outcome? This needs to be discussed somewhere in the text.

p. 6, L. 4: Again, choosing 25% seems random to me (see previous comment).

Answer for the three above comments:

We will add some definitions of the employed statistical metrics. So far we modified the text as follows: “Further, we applied our percentile score-based model ranking method. This method is based on the representation of the distribution of the statistical metrics that we used for the analysis in this study in accordance with boxplots – the lowest value as minimum, 25th percentile, 50th percentile or median, 75th percentile and the highest value as maximum. Figure 2 shows an example of this approach applied to RMSD of sea surface temperature in the Barents Sea. We divided the statistical measures into 4 groups based on the amplitude of the calculated metrics and assigned a score to each model according to its group: (i) models considered as very good (less than 25th percentile or 25% of the distribution of the statistical metrics) were assigned a score of 3; (ii) good models (between 50th percentile or median or 50% and 25%) were assigned a score of 2; (iii) satisfactory models (between 75th percentile or 75% and 50%) were assigned a score of 1; and (iv) unsatisfactory models (more than 75%) were assigned a score of 0. In the case of the correlation coefficient, it is vice versa, very good models with correlations scores above 0.75 ranked with a score of 3, and so forth. Finally, we summed up the total score for each GCM and selected the optimal ensemble of climate models which we take to be the top 25th percentile or 25% of

GCMs (sub-ensemble) ranked according to their total score (Ruan et al., 2019). This procedure was applied to each parameter and study region.”

Results & Discussion

p. 6, L. 6: Personally, I find it a bit unfortunate that only results for temperature and the Barents Sea are presented in the main text. Isn’t there a better way to synthesize the results and present more than just one tiny subarea and one forcing factor?

Answer:

We agree with the reviewer, and added an additional figure to sections Results and Discussion with the spatial distribution of biases in five parameters between models and reanalysis data in six target seas. The biases are averaged over the vegetation season and the time period 1979/1993-2005. So far we present an example of this kind of figure using sea surface salinity (PSU), averaged over 30 m, for 6 studied seas: “Figure 7. Spatial distribution of biases in five parameters between models and reanalysis in six studied seas averaged over the vegetation season and the time period 1979/1993-2005.” We added an example of the text: “In order to analyse how well the selected ensemble models represent the five studied parameters, we calculated the spatial distribution of biases in the parameters, which were calculated as the difference between selected ensemble models and reanalysis data, in six studied seas - Barents, Bering, Labrador, Greenland, Norwegian and North Seas (Figure 7). Thus, less biases in SSS are determined in the case of the Labrador, Greenland and Norwegian Seas ($\pm$0.5 PSU); low biases prove to be often in the Bering Sea next to coastline - up to 1.5-4 PSU, this overestimation is possibly due to the river runoff from the Kuskokwim mountains, the Alaska and Aleutian ranges. SSS is underestimated in waters next to the coastline in the Barents and North Seas (1.5-2.5 PSU), which is probably due to the overestimation of river runoff or underestimation of salty Atlantic water.”

p. 6, L. 9-11: I don’t think a Taylor diagram needs to be explained in the result section. I suggest to rather briefly explain when the agreement between model and reference

data set is good (i.e. how to interpret the plot) instead of simply stating what can be seen (see comment on L. 15-17).

p. 6, L. 11/12: Please add "[. . .] capture the climatological seasonal cycle [. . .]". Furthermore, please explain how it can be seen from the plot that the seasonal cycle is represented better than the interannual variability (see previous comment).

p. 6, L. 14: Are these numbers really unitless? If so, define somewhere that you plot normalized SD and RSMD (method section, consider adding formula there) and state that here by saying e.g. "the SD and RMSD normalized by XX are between . . .". This will help the reader to follow.

p. 6, L. 15-17: This is the information you should start your paragraph with (see previous comments). First explain to the reader how to interpret the plot. However, the statement that "the closer the model data is to the x-axis, the better the correlation coefficient" is not entirely correct, as the correlation coefficient is shown on the radial axis. A point with RMSD/SD/CorrCoeff of 0.1/0.1/0.1 is closer to the x-axis than a point with 1.0/0.8/0.9 (note that this is under the assumption that RMSD is on the x-axis, SD on the y-axis and the correlation coefficient on the radial axis, see comment on the Figure further down)– but the correlation coefficient of the second point is higher. Please be precise in the description. Also, a correlation coefficient is high/low and not good/bad.

p. 6, L. 17: Replace "climate parameter" by "e.g. SST".

p.6, L. 17-18: Please revise the grammar of this sentence.

p. 6, L. 18-20: This statement is redundant with the method section.

p. 6, L. 9-20: For the whole description of the Taylor diagram, please add the names of models here that show the highest/lowest correlation coefficients, RMSD etc. to make it easier for the reader to extract the information from the plot.

Answer to above 8 comments:

Our intention was to describe our methodology using Figures 2-8. We decided to delete the Taylor diagram as it only illustrates the root mean square deviation, standard deviation and correlation together; but we analyzed these statistical metrics separately in the form of a table. Thus, we deleted the second paragraph from page 6 and moved Figures 4, 6, 7 to section Methodology section. In addition, we deleted Figure 5 as well, since the analysis is very similar to Figure 4, and Figure 8.

p. 6, L. 21-23: If you say you show the "spatial distribution", I expect maps. Do you mean the spatial variability of the climatological SST bias across the subregion? Please be more precise throughout the description.

Answer:

We modified the sentence as follows: "Figure 3 illustrates the box plots of the spatial variability of SST biases in the selected area of the Barents Sea for the vegetation season (June-September) during 1979-2005 and the time period 1979-2005."

p. 6, L. 22: I see median biases that are >0 (e.g. for the model 2). Please double-check.

p. 6, L. 24: Do you mean the maximum bias? I don't understand "amplitude bias" (throughout paragraph). Similar to above, please add the names of the models showing the numbers you're stating to make it easier for the reader to find the information you're stating in the plot.

Answer to the above 2 comments:

Thank you for this correction. We changed it as follows: "For model ranking, we analyzed the absolute values of both the median bias and the amplitude (maximum and minimum values difference) of the spatial variation in model biases. The median bias varies from -6.6 (model #20) to 1.5 K (model #24) among the models, whereas the amplitude of values has a wide spread from 7.3 (model #21) to 16.5 K (model #3)."

p. 6, L. 24-25: Please revise this statement, e.g. simply stating that the comparison shows large variability across the models.

Answer:

We will do it.

p. 6, L. 27-29: Where is this seen? This is not included in Fig. 4. If you're referring to a different plot here, please add the reference.

Answer:

We changed it as follows: "After the percentile score-based method was applied to model #2 (ACCESS1-3), it was included in the optimal ensemble, whereas and model #3 (CanESM2) was not included (Figure 8)."

p. 6, L. 29: Similar to above, be more precise in your description. From just the wording "spatial distribution of annual trends") the reader expects maps here, not box plots.

p. 6, L. 31: How are "significant trends" defined here? How can that be seen in the plot? Please be precise. What models show a significant trend? What is an "unrealistic trend" for you here?

p. 7, L. 1: How do you know that? As mentioned above, I think it is important to state in the method section that this is the assumption you make (a model that reproduces the observations best over the historical period (however you define "best"), also gives the "best" projections for the future).

Answer to the above 3 comments:

We deleted the figure plotting the spatial variability of the trends as the procedure of the analysis is similar to the figure displaying the spatial variability of biases. Also, we will move Figure 4 to section Method.

p. 7, L. 9-12: Is the +/- 1K the average over the domain? Currently, the reader at this point has totally forgotten why you're doing this exercise. I suggest to always relate your analysis back to your goal of projecting potential future changes in coccolithophore blooms. I understand that this will be a follow-up paper, but this paper would gain a lot if

you speculated at least. How can these biases be expected to impact these estimates? You could do some basic calculations using a Q10 function (see e.g. Nissen et al., 2018) or a temperature optimum function (see e.g. Krumhardt et al., 2017) describing the impact of temperature on phytoplankton growth.

Answer:

We modified it as follows: "Figure 4 illustrates the spatial distribution of biases for SST between models and reanalysis data for the Barents Sea; biases are averaged over the vegetation season (June-September) and the time period 1979-2005 for the full 28-model ensemble, the selected 8-model ensemble, and the top-model. As seen, the full 28-model set underestimates the SST in the study region while the top-model, MIROC-ESM, overestimates it. The selected 8-model ensemble shows smaller biases ($\pm$ 1 K) in SST for most of the study area in the Barents Sea."

p. 7, L. 13: What error do you mean here? Please be precise and make sure that all the metrics you present are carefully introduced in the method section.

Answer:

We modified the sentence as follows: "The spatial distribution of errors (the difference between model and reanalysis data) in SST trends between models and reanalysis in the study region is presented in Fig. 5."

p. 7, L. 15-16: Please revise the grammar of this sentence.

Answer:

We modified the sentence as follows: "The full 28-model ensemble overestimates trends for the whole study region (model-reanalysis errors are 0.03-0.07 K yr-1), the top-model MIROC-ESM partly underestimates the SST trend, but for the larger area reveals similar to Era-Interim reanalysis small trends ($\pm$ 0.01 K yr-1)."

p. 7, L. 13-17: Similar to above, I don't understand from the current presentation of the

results what these mean.

Answer:

We added a concluding sentence to the paragraph: “The maps show for the Barents Sea the difference between the SST model and reanalysis data averaged over the vegetation season and the time period 1979-2005; this comparison reveals better results for the selected sub-ensemble.”

p. 7, L. 24-28: This is repetitive with the method section and what should be in the figure caption. There is no need to state it this detailed in the main text.

Answer:

We deleted several sentences and modified the paragraph as follows: “The selected optimal CMIP5 model ensembles for the other seas and parameters are presented in Fig. 9. The heat map shows the final model scores, which represent the results of our percentile score-based model ranking approach. From the heat map, we can conclude that there is no optimal model ensemble, or one top-model, which could properly simulate all parameters over study regions. However, some climate models exhibit good results for many cases, e.g., ACCESS1-3; ACCESS1-0; HadGEM2-AO; HadGEM2-CC; HadGEM2-ES; GFDL-CM3; INMCM4; GISS-E2-R; GISS-E2-R-CC. The model that have higher biases across the majority of the study regions are CMCC-CM; FGOALS-g2; IPSL-CM5A-LR; IPSL-CM5A-MR; IPSL-CM5B-LR; MIROC5; MRI-ESM1.”

p. 7, L. 28-30: Does it surprise you that the model combinations vary?

Answer:

In many studies that use climate model data, vast regions are considered, in particular, the entire Arctic. Most studies also use the approach when one set of models is selected for different parameters. Our results confirm that the same model does not properly reproduce the distribution features of all the parameters we examined and is not suitable for the analysis of large regions. It is one of our messages to readers to be

more careful when to choose climate models at the study.

p. 8, L. 3-5: How is “better performance” defined here? Is not clear to me how you conclude this.

Answer:

We modified the sentence as follows: “In general, the selected best-model ensemble shows better performance (with regard to biases between model and reanalysis data) than either the full-model ensemble or the single top-model.”

Conclusions p. 8, L. 13-16: The statement that the Arctic is often considered as a single region in other studies is never made in the introduction, but should be included there as a motivation to look at subregions. Furthermore, you don’t actually assess the whole area, so I suggest to revise this statement, as you don’t actually compare the performance over the whole area to the smaller subareas.

Answer:

We will do it.

p. 8, L. 18: What about the temporal development of the environmental conditions?

Answer:

We suppose that trends are responsible for the temporal development of environmental conditions.

p. 8, L. 18-21: Are more important than what? Please be precise. I cannot follow your logic here. Please revise to clarify, taking also into account the comments I made in the result Section.

Answer:

We modified the sentence as follows: “Therefore, we suppose that the spatial distribution of biases and trends in considered parameters are more important than other

statistical metrics within the framework of the model selection procedure performed."

p. 8, L. 24: And the time series is even shorter for SSS and ocean currents, isn't it? What is "out-of-sample" testing? Please try to avoid introducing concepts in the conclusion section which were not discussed before. Why did you not test by excluding certain time periods from the analysis?

Answer:

We agree with the reviewer and deleted the following sentences: "Due to the short sample period of reanalysis data (1979-2005), we did this evaluation without out-of-sample testing. Definitely, it is better to test any model ranking method on another historical period. It will be possible to consider the period 1950-2014 with the release of new data, e.g., CMIP6, ERA5." We consider the period is very short to be divided into two independent periods for the analysis.

p. 8, L. 27: important for what? Please be precise.

Answer:

We modified the sentence as follows: "We can conclude that the range of different factors is important for model selection, including the spatial pattern of model biases, and the proposed methodology is a way of enhancing the model selection procedures sophistication that promises a better chance to identify more skilful models for the features we are interested in."

p. 8, L. 31: Why only at regional scales?

Answer:

We modified the sentence as follows: "Thus, the proposed method can be used for analyses regarding other regions with the purpose to evaluate climate model performance with respect to various atmospheric and oceanic parameters at different scales."

Figures/Tables

Fig. 1: You don't simply show the "locations of the blooming areas" here, but the spatial distribution of the frequency of blooms. Please be more precise. I suggest to show the subregions in the plot directly (put names and add e.g. a black contour to show the extent). Please add to the caption what data this map is based on and how you define a bloom.

Answer:

We corrected it as follows: "Figure 1: Spatial distribution of the frequency of E. huxleyi blooms based on the Ocean Colour Climate Change Initiative dataset version 3.0 (Kazakov et al., 2018) for the Barents, Bering, Labrador, Greenland, North, and Norwegian Seas."

Fig. 2: Be more precise in caption, a lot of information on what is seen in the plot is missing. What is the unit of the RMSD?

Answer:

We corrected the caption as follows: "Figure 2: A schematic representation of the percentile score-based model ranking method (Division of RMSD values distribution of 28 models into four groups that are limited by 25th, 50th and 75th percentiles and the relative assignment of scores from 3 to 0 for each group accordingly - very good, good, satisfactory and unsatisfactory)."

Fig. 3: The way I know it, a Taylor diagram shows the RMSD (normalized by the standard deviation of the reference data set) on the x-axis, the standard deviation (normalized) on the y-axis and the correlation coefficient on the radial axis. It is not clear to me what exactly you're showing. Please add labels to the plot (y-axis, grey circles) and also say what you're showing in the caption (including units or state if you normalize by something). Also, please add panel labels to the plot and the caption.

Answer:

We decided to delete the Taylor diagram as it only illustrates root mean square devia-

tion, standard deviation and correlation together; however, we analyzed these statistical metrics separately in the form of a table.

Fig. 4 & 5: Possibly replace “distribution” by “variability”? Be precise in what you show. What are the orange line and the whiskers? How is the bias defined (Fig. 4)? What trend is shown Fig. 5; trend in average over blooming period averaged over subregion?)?

Answer:

We corrected the caption of Figure 4 as follows: “Figure 3: Box plots of the spatial variability of SST biases, which are calculated as the difference between the model and reanalysis data, in the Barents Sea over the vegetation season and the time period 1979-2005. Each box spreads from the lower quartile Q1 to the upper quartile Q3 of biases, the orange lines represent the medians. The lower “whiskers” are represented as Q1-1.5. The Standard deviation and the upper “whiskers” are represented as Q3+1.5 Standard deviation.” We decided to delete Fig. 5, since it illustrates a similar analysis procedure as that in Figure 4. After revision, it is Figure 3 that we moved to the section Method.

Fig. 6: Restate blooming period in caption, add unit of SST bias.

Answer:

We changed the caption as follows: “Figure 4: Spatial distribution of biases in SST (K) between model and reanalysis data in the Barents Sea; the biases are averaged over June-September.”

Fig. 7: What error are you showing here? Please add the unit of the SST trend in the caption. The colorbar currently states that you’re showing SST (K) – please double-check. Please restate the blooming period.

Answer:

We corrected the caption as follows: “Figure 5: Spatial distribution of errors, which are calculated as the difference between the model and reanalysis values of annual SST trends (K yr-1) in the Barents Sea (June-September).” We will modify the colorbar caption.

Fig. 8: In my view, it is not really common to plot SST in Kelvin, consider changing it to °C. Please add the units of the variables in the figure caption. Explain what the fit is, exchange “x” and “y” by the actual variables you fitted. Please don’t use black/dashed for all fits, I suggest to change the color of each fit to the color of the respective full time series.

Answer:

We deleted this figure to avoid overloading of the paper with figures as it is not that much important in the manuscript.

Fig. 9: Please briefly summarize what the numbers for each model-variable combination represent and refer back to the method section and Fig. 2. Please also explain in the caption what the white areas are and refer back to Table 1. Please add the units to the variables in the Figure caption.

Answer:

We modified the caption as follows: “Figure 6: Heat map with the final model scores obtained using the percentile score-based model ranking method for the five parameters (sea surface temperature (SST, K) and salinity averaged over 0-30 m (SSS, PSU), surface wind speed at 10 m (WS, m s-1), ocean surface current speed (OCS, m s-1), and shortwave downwelling solar radiation (SDSR, W m-2) for the Barents, Bering, Greenland, Labrador, North, and Norwegian seas based on different statistical metrics (Figure 2, Table 2). The white areas indicate that the model was not considered due to partial or complete unavailability of hindcasts, and future projections (RCP4.5, RCP8.5) data.”

Table 1: Replace “concrete” by “respective. Please define all abbreviations in the Figure caption (e.g. SST, WS. . .) and add units.

Answer:

We corrected accordingly: “Table 1. CMIP5 models used for simulation of selected parameters: SST – sea surface temperature in K, WS – near-surface wind speed in m s-1, SDSR – surface downwelling shortwave solar radiation in W m-2, SSS – sea surface salinity (averaged over 30 m) in PSU, OCS – surface ocean current speed in m s-1 (models available for respective variable are marked as “+”)”

Table 2: Please add units in the Figure caption. What is SDdif? This is never explained in the text (method section). Please be consistent with the use of underscores in caption and Table (e.g. Trm vs Trm). What does “modulus of standard deviation difference” mean? I don’t understand this. Please use the exact same names as introduced in the method section.

Answer:

SDdif is the difference between the model’s standard deviation and the reanalysis’s standard deviation. By “modulus of standard deviation difference” we meant the value regardless of its sign. We improved this part in the section Method. We corrected the caption is corrected as follows: “(Numbers in brackets indicate the models’ scores. RMSD is the root-mean-square deviation, K; r is the correlation coefficient between models and reanalysis; RSR is the RMSD-observations standard deviation ratio; |SDdif| is the modulus of the standard deviation difference (model minus reanalysis), K; |Trm| is the modulus of spatial trend median difference (the model minus reanalysis), K yr-1; |Tra| is the modulus of spatial trend amplitude difference (the model minus reanalysis), K yr-1; |Brm| is the modulus of spatial bias median difference (the model minus reanalysis), K; |Bra| is the modulus of spatial biases amplitude difference (the model minus reanalysis), K).”

References (AR1) Agosta, C., Fettweis, X. and Datta, R.: Evaluation of the CMIP5 models in the aim of regional modelling of the Antarctic surface mass balance, Cryosph., 9, 2311–2321, 2015.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R. and Zhu, Z.: Evaluating the Land and Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models, J. Clim., 26(18), 6801–6843, doi:10.1175/JCLI-D-12-00417.1, 2013.

Das, L., Dutta, M., Mezghani, A. and Benestad, R. E.: Use of observed temperature statistics in ranking CMIP5 model performance over the Western Himalayan Region of India, Int. J. Climatol., 38(2), 554–570, doi:10.1002/joc.5193, 2018.

Gleckler, P. J., Taylor, K. E. and Doutriaux, C.: Performance metrics for climate models, J. Geophys. Res. Atmos., 113(D6), D06104, doi:10.1029/2007JD008972, 2008.

Golmohammadi, G., Prasher, S., Madani, A. and Rudra, R.: Evaluating Three Hydrological Distributed Watershed Models: MIKE-SHE, APEX, SWAT, Hydrology, 1(1), 20–39, doi:10.3390/hydrology1010020, 2014.

Kazakov, E., Kondrik, D. and Pozdnyakov, D.: Spatial data assimilation with a service-based GIS infrastructure for mapping and analysis of E. Huxleyi blooms in arctic seas, in Sixth International Conference on Remote Sensing and Geoinformation of the Environment., 2018.

Kondrik, D., Pozdnyakov, D. and Pettersson, L.: Particulate inorganic carbon production within E. huxleyi blooms in subpolar and polar seas: a satellite time series study (1998–2013), Int. J. Remote Sens., 38(22), 6179–6205, doi:10.1080/01431161.2017.1350304, 2017.

Kondrik, D., Kazakov, E., Chepikova, S. and Pozdnyakov, D.: Prioritization of the vector factors controlling Emiliania huxleyi blooms in subarctic and arctic seas: A multidimensional statistical approach, Biogeosciences Discuss., 1–24, doi:10.5194/bg-2019-104,

2019.

Kumar, D., Mishra, V. and Ganguly, A. R.: Evaluating wind extremes in CMIP5 climate models, Clim. Dyn., 45(1–2), 441–453, doi:10.1007/s00382-014-2306-2, 2015.

Moriasi, D. N., Arnold, J. G., Liew, M. W. Van, Bingner, R. L., Harmel, R. D. and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Am. Soc. Agric. Biol. Eng., 50(3), 885–900, 2007.

Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M. and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430(7001), 768–772, doi:10.1038/nature02771, 2004.

Ruan, Y., Liu, Z., Wang, R. and Yao, Z.: Assessing the Performance of CMIP5 GCMs for Projection of Future Temperature Change over the Lower Mekong Basin, Atmosphere (Basel)., 10(2), 93, doi:10.3390/atmos10020093, 2019.

Smyth, T. J., Tyrrell, T. and Tarrant, B.: Time series of coccolithophore activity in the Barents Sea, from twenty years of satellite imagery, Geophys. Res. Lett., 31(11), n/a-n/a, doi:10.1029/2004GL019735, 2004.

Stocker, T. F.: Models change their tune, Nature, 430(7001), 737–738, doi:10.1038/430737a, 2004.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2019-177, 2019.

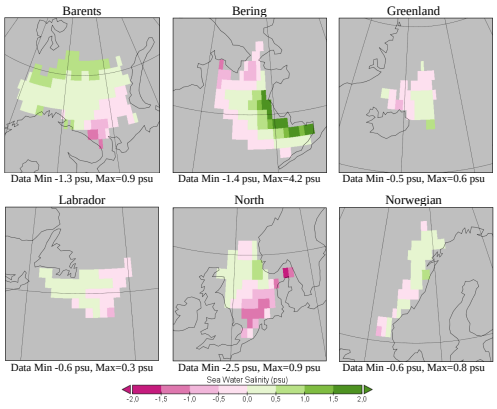

Figure 7. Spatial distribution of biases in five parameters between models and reanalysis in six studied seas averaged over the vegetation season and the time period 1979/1993-2005.

Fig. 1.