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 reviewers' comments and suggestions that considerably improve our manuscript.

Please kindly find attached the responses to the reviewer 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)

General comments (lines 8-34 in the reviewer's report)

Summary

The principal concerns of the third reviewer predominately reside in justification of both the major incentive of our study and the selection of factors controlling phytoplankton blooms. Also, we explain the reasons for confining our consideration solely to some specific sea areas within the seas addressed.

We fully acknowledge the righteousness of these objections. The matter is that after the first round of reviewing we acquiesced to the recommendation of one of the reviewers to shift the focus of the paper from *E.huxleyi* to the phytoplankton in general and for model selection methodology. Having done this, we actually weakened the major incentive of the study and justification of locations of the specific marine areas that were addressed.

Therefore, we decided to return to the initial version of the paper. However, we further substantiated the study incentive, our reasoning of confining to selected parts of the seas, and the expediency of employment of the concrete set of forcing factors (FFs) that condition/govern the growth of *E.huxleyi*. We extended the number of

FFs up to 10 and included nutrients and some parameters related to the carbon chemistry system in the surface ocean.

Accordingly, we renewed our selection of models from the CMIP5 archive. The respective results are presented and discussed.

Along with the above changes, in the revised paper we also made multiple corrections to the text following the detailed comments of the third reviewer. Below, the respective answers/explications are given point by point.

Line number	Comment	Reply
	Abstract	
10	I suggest to rephrase "sea areas" and say "differ *from*" instead of "differ *in comparison with*".	This sentence has been rephrased.
11	You don't actually test different methodologies in your manuscript, so to say that you want to find the "methodology of selection" in the abstract is misleading in my opinion	We agree. This sentence has been rephrased.
17	Please be more precise and state i) sea surface temperature and ii) salinity averaged over the top 30m – right now it is not clear from the abstract whether you use sea surface salinity or top 30 m averaged salinity.	This requirement is fulfilled.
20	What is "high-skillful"? This is rather subjective, isn't it?	We mean skillful according to our selection procedure. We have corrected the text to reflect this.
21	I don't understand "mutually optimal combinations» - can you rephrase?	We mean that there is no one combination of models which is selected as being the most skillful for different variables. We have corrected the text to clarify this.
23	"a unique best model subset" – best in terms of what?	Best in terms of fit to the reanalysis according to our model

		selection procedure. We have corrected the text accordingly.
24	It has to become clear from the abstract whether this "7-11 models" is a subjective choice of yours or whether it is based on some objective criteria. It currently reads like the former.	It is an objective result of the thresholds set in the model selection procedure. We have clarified this in the text.
27	I am not convinced that the fact that the model subsets outperform the full-model average proves that *your* selection method should be used in the future. The logic in this sentence doesn't hold for me as you have not shown in your manuscript how a different selection method would not achieve the same thing.	We are not comparing model selection procedures here so we do not say that our procedure should necessarily be favoured over other methods, but rather that model selection using this method is preferable to using the full model ensemble. We have corrected the text to clarify this.
	I miss a statement in the abstract regarding why you focus on the Arctic/subarctic. The climate models give global projections and differ not only in this region. In my opinion, you should better motivate the focus area by e.g. stating that this region is projected to be affected especially severely by climate change. Do the models differ most in that area?	Justification of selection of Arctic and sub-Arctic seas is presently given in Abstract and Introduction.
	Introduction	
Page 2		
2	I suggest to avoid saying "prediction of the future status" and instead use "projections" here. Additionally, please state here already what the components of the climate system are.	Agreed. This has been changed.
5	"more realistically" than what?	Deleted.
6	Please say "model output" here instead of "model data"	The recommended change is done throughout the entire text.
9	"usually"? I don't understand the use here. Are you	Agreed, the sentence is rephrased.

	referring to the historical period here or the projections? This needs to be made clearer here.	
10-11	Same as above: historical period? Additionally, it did not become clear from the first paragraph that the models' performance varies across regions and target variables. Therefore, that a different set of models should be used for "each specific task and region" does not logically follow from above. What does "reliable" mean in this context?	Now we mention "historical period" and the explanation why the models should be sea area-specific is presently given above these lines, the word "reliable "is removed.
	Furthermore, after making the distinction between "historical period" and "future projections" clearer in this paragraph, it is important to point out here that you assume that the models that perform "best" (whatever that means) over the historical period also produce the "best" projections. I have seen the sentence along these lines in L. 29-30, but I think it is more appropriate here already.	The remark is accepted and the sentence from Lines 29-30 is presently placed earlier in the text.
12-13	What is a "reliable approach"? How does a model "most efficiently" simulate the state of a variable? I suggest to rephrase.	Done: the sentence is rephrased.
13-14	What are the "abiotic parameters" you're referring to here? They are not introduced. Furthermore, why do we want to understand how phytoplankton dynamics might be affected by future climate change? Why subpolar and polar latitudes?	Yes, this was unclear. We have rewritten this section to clarify what parameters we chose and why, and to clarify the purpose of these choices.
15-16	This sentence does not make sense to me.	As Introduction has been essentially changed (the major incentive of study is distinctly described), this part hopefully became comprehensible.
19	What do you mean by "in case it is not possible to calibrate a model for a selected region"? Additionally, why "furthermore"? Don't you have this exact information in the sentence before? I don't understand.	Agreed. This sentence was superfluous and was removed.

23	“best” regarding what? Please be more precise here. Additionally, I think “Taylor et al.,2001” is not a suitable reference here as this paper does not discuss the choice of a model subset.	"Best" is taken to mean most skillful in simulating historical climate with respect to observations. We have added this comment. Agreed. Here we make a mistake, it should be cited article of Taylor et al., 2012
25-26	Please add references to i) and ii).	The references are provided in the previous sentence, and the two successive sentences are united.
28	Please add a reference to this statement. What is a “poor reproduction of the interannual variability”? The sentence starting with “so that…” is not a sentence, please check the grammar.	Agreed. This sentence was superfluous and was removed.
30	The “therefore” does not make sense here. Please double-check the logical flow between these sentences.	Done: the sentence is rephrased.
31	What does “properly” mean here? Be precise.	This sentence has been rephrased.
32	The choice of the parameters is still not clear. If these are still motivated by the fact that you ultimately want to assess the impact of climate change on phytoplankton dynamics, choosing salinity & wind speed rather than nutrients does not make sense to me. Irrespectively, these variables are not properly introduced or motivated here.	The number of parameters and their choice was reconsidered, respective changes were entered.
Page 3		
3-6	What is the disadvantage of using this approach? Or differently asked: Why did you not choose to do it that way?	These sentences are removed as they refer to studies that address solely atmospheric variables.
7	Why would one select “a unique combination of models for each study variable”? For what? I have trouble following the logic here.	This part has been rephrased. The selection of unique combinations was motivated by the necessity to reveal how well each of the models simulates each

		factor.
10	What are "appropriate models"?	The sentence is reworded and in the following part of the text we avoid using such adjectives.
3-24	It is not clear to me at all what the advantages and disadvantages of each of these approaches are. Instead of simply stating what author X did, rather state *why* they chose that approach and what they could not assess with it or what the limitation of the respective approach was. Thereby, you can better motivate why you chose *not* to follow any of the listed approaches – which is currently not clear from what you write.	We moved this part to Methods section and rephrased.
25	After what you have written in the previous paragraph, this sounds too much like an arbitrary decision to me. Why did you choose to do what you did? In what way does that improve previously applied methodology?	The entire paragraph is revised explaining the incentive of our decision. This part moved to Methods section.
30	Until here, I have assumed that you focus on the global scale – mentioning the subareas in L. 14 of p.2 came out of nowhere at that location (see comment above) and is not linked to the rest of the introduction until here. Please rewrite the introduction, so that your focus area naturally results from the introduction. Why do we care especially about the performance of climate models in polar regions?	Introduction is essentially rewritten, and necessary explications of the study focus on subpolar and polar regions is justified.
30	In what way do the subareas "differ in physical and geographical conditions"? This statement is too vague. Is it any problem that the subareas you choose are linked through circulation? You don't discuss anywhere in the manuscript how e.g. biases in circulation in the North Atlantic could simultaneously affect the North Sea, the area north of Iceland ("Greenland"), the Norwegian Sea, and the Barents Sea. As an example, if a model does not	It is true that biases in circulation patterns likely affect more than one sea but how these biases affect the climate in a given sea is strongly controlled by the local conditions e.g. by the amount of sea ice cover, the presence of mountains etc. We have added comments

	reproduce the observed circulation in the Norwegian Sea, causing biases in e.g. temperature and salinity there, these biases will also impact the Arctic (e.g. the Barents Sea), don’t they? Do you really think that it is appropriate to treat these areas as independent when they are clearly linked to each other?	addressing this.
31	What are “stable localizations”? How is “intense growth” defined? It is still not clear to me why you do that.	We mean the regions in which blooms of plankton regularly occur. We have corrected the text accordingly.
Page 4		
2	No. Until here, it is not at all clear that the variables you chose are the variables controlling phytoplankton blooms, as you still don’t mention important variables in the introduction (e.g. nutrients). Referring to one of your answers in the first round of reviews, namely that the temporal resolution of some of the drivers of phytoplankton dynamics not considered here is not high enough in the model output for the analysis you perform, I was going to recommend to add a statement along these lines somewhere in the text (probably in the method section). However, then I realized you claim to only use monthly data in the method section, so I don’t understand why you cannot assess the nutrient fields. Every reader familiar with phytoplankton dynamics will wonder why you did not consider obvious variables like nutrients, if your goal is to assess the possible implications of climate change on phytoplankton blooms.	We agree. Introduction has been essentially changed. Also, we explained the number of variables and included nutrients.
2-4	This statement does currently not make sense to me. Please clarify *how exactly* your results can improve	Explication is provided.

	ecological models as I don't see how that could be.	
Methods		
Page 4		
7	Be consistent: "34" or "thirty-four" – compare to introduction. Additionally, check grammar throughout the text regarding "GCMs outputs".	This is done throughout the text.
10	I suggest to not use "SSS" when not referring to sea surface salinity. Check throughout the text that you're consistently stating "salinity averaged over top 30m" and not "sea surface salinity" (e.g. in Figure captions).	This is done throughout the text.
11	"known to affect the phytoplankton life cycle" is too vague. E.g., how do wind speed and ocean current speed affect the phytoplankton life cycle? Additionally, the references you cite are all coccolithophore papers, so if you want to make a general statement about phytoplankton here, I suggest to add references of other phytoplankton types.	Again, the respective explanation is now provided in Introduction
12	What do you mean by "availability" here? Please be more precise. This statement is repetitive with the statement in L. 8 – please revise.	We mean availability of model output: not all models have scenario data available or have output of all the variables we assess here. This has been clarified in the text.
24	You state in your answer to the first round of reviews that you wanted to move from only focusing on coccolithophores to phytoplankton as a whole – then why do you still choose your subareas based on *E. huxleyi* blooms? You could easily detect total phytoplankton blooms based on satellite imagery.	As the initial focus on *E. huxleyi* is returned, this comment is no more relevant
27	This information needs to come much earlier, namely in	Presently this information is

	the introduction.	actually given in Introduction.
29	Please rephrase “target water areas”.	Rephrased.
29	What is a “steady localization”?	This has been rephrased: it is the regions where phytoplankton blooms have often occurred in the historical record.
Page 5		
2-5	It is still unclear to me why you do that rather than assessing a larger subarea. Furthermore, you still don’t define what a bloom is for you here.	We added motivation to the text about why we chose specific areas in which blooms of E. huxleyi regularly occur. We clarify through the text that we consider blooms of *E. huxleyi.*
5-9	This statement does not make sense to me. Please rephrase to make clearer how the comparison of models you do here can e.g. be used in “marine ecology related issues” and in “forecasting of the region-specific climate interactions during the vegetation season”. What do you mean exactly?	The sentence is rephrased and any ambiguity is avoided.
10	What are the “regional features of distributions”?	Done: the sentence is rephrased.
17	What does the CPI tell you? Why did you compute it? Additionally, please introduce the short form “CPI” when first discussing the index (introduction). Same applies to other variables throughout the manuscript, such as “sea surface temperature”, where you jump back and forth between the long name and the short form. Please be consistent.	This index widely used in climatology studies for model evaluation and weighted projections. We rephrased this part and mentioned more publications in the text. Done with regard to both CPI and other short form throughout the paper.
14-26	Contrary to what you state in your answer to the reviews	In the revised paper we

	(see also my comment to L. 2 of p. 4), here you state that you mostly use monthly averaged model output. Based on this information, I understand less and less why you disregard e.g. nutrients in your analysis.	significantly extended the number of forcing factors: presently they encompass nutrients as well.
28	Please delete “proposed and”, simply state “we employed”.	“proposed” is deleted
28-29	Coming back to an earlier comment of mine: the percentile score-based model ranking method still seems like a random approach to me – be explicit in your manuscript regarding what the individual pieces of this approach (each of which had been applied before in other studies) could not assess (or where their limitations were) and how combining the individual pieces solved these limitations.	The required justification is presently given in Introduction
Page 6		
1	amplitude or magnitude (throughout the text)?	To avoid misunderstanding, we used “spread” throughout the entire text
3	What is the “total score”? You currently don’t define this.	The missing definition is presently provided
6	Please use “SST”.	done
8	Replace “less than“ by “top”. And please avoid subjective language here (“models considered *very good*”). Strictly speaking, just because a model ranks amongst the top 25% does not mean the respective model does a particularly good job at reproducing the observations – the respective model is only the “*best*” relative to the other models relative to the other models you assessed.	done
14	Please replace “analyzed”, as this description is too vague.	The word “analysed” is changed
17	What is the “amplitude bias”? I don’t understand.	The sentence is rephrased

20	What is "much higher"? Please be precise.	The respective number is given
21	Models	changed
22	What does "placed beyond it" mean? Please rephrase.	The sentence is changed
13-27	Shouldn't all of this be part of the results section?	It was suggested by previous reviewer – to place this part in Methods section. This is an example of applying our approach, and we do not show such results for all variable. That is why we decided to leave it here.
Results & Discussion		
Page 6		
30	"SST" etc. have been introduced before. Please avoid redundancies.	Done throughout the text
Page 7		
8	not an optimal	We rephrased it.
9	What does "properly simulating" mean?	The sentence is reworded to avoid the word "properly"
11	What bias do you mean here? Median? Spatial variability?	The word "bias" is replaced and the sentence reads differently now.
29(p.6) - 12(p.7)	Can you make more general statements that make it easier for the reader to extract the most important information? Or is the main information really only "different models are *best* for different parameters in different subareas"? I suggest to reverse the order then and start the section with the main information to be taken away (L. 8-12) and then explain in more detail how this can be seen.	This part of the text is rephrased.
13	what does "equally reproduce" mean here?	Changed: reproduced equally well
14-16	Please add references here.	References are provided

16	please rephrase “marine tracts”.	“marine tracts” is removed
16	I don’t understand the logic why the error should increase towards the poles because of the convergence of meridians.	It was deleted.
16-18	This logic also doesn’t hold. Why do you think that? I don’t see why you would expect a different set of models to perform “best” in one area and then others in a different area, simply because the two areas differ in their physical environment. Please explain. Referring to an earlier comment, you nowhere in the manuscript explain how the subareas are different in e.g. their physical circulation.	We explained it in 2.2.1. Study regions.
19	represent *the* five	The description of Figure 5 (maps) was removed because we decided to replace it with boxplots. Thus, we replaced Figure 5 and the Supplements with one figure with boxplots.
21	Do you mean a “lower” bias?	
23	What exactly do you mean by “insufficiently accurate parametrization of river runoff”? Don’t the models account for river runoff? You need to elaborate to make your argument clearer. Consider adding references here.	
25	But can’t you assess this in more detail? I mean, you have analyzed the current speeds, have you also assessed the circulation as a whole? You can figure out what the cause is with the available model output.	
29	“[…] and the biases *are* much higher”	
25-29	I currently miss a discussion on the causes of the biases in SDSR.	
30	Avoid subjective language, such as “simulate well”. Furthermore, the description and discussion of the biases in WS is rather short compared to that of the other variables. Try to make it more balanced across the variables.	

32	Avoid statements like “quite good results”. Simply state the result and let the reader decide whether that is good or bad.	
32	observed or obtained?	
19(p.7) - 13(p.8)	The order of variables here is not intuitive as you jump between oceanographic and meteorological ones. I suggest to reorganize.	
Page 8		
4-5	Again, did you look into that?	The description of Figure 5 (maps) was removed because we decided to replace it with boxplots. Thus, we replaced Figure 5 and the Supplements with one figure with boxplots.
10	Use your short forms. Furthermore, “rather large biases” is subjective, please avoid statements like this.	
14	What do you mean by “to examine our percentile score-based”? Be more precise what you examine.	We deleted Figure 6 because this comparison can be seen in Figure 6 – new version (boxplots).
17-24	I am not sure I get what to take away from this.	
19	Do you mean Fig. 6?	
25	what does “somewhat better” mean? What bias (i.e. median, variability…)?	
Conclusions		
7	Check grammar “no any optimal”.	word “any” is removed
10	What are “quite good results”? This is subjective.	Changed
11	The start of the sentence does not make sense to me. What do you mean?	The sentence is deleted.
13	Delete “well”	deleted
16-18	Please avoid 1-sentence paragraphs.	done
19	What “range of different factors” are you referring to here?	Changed the entire paragraph

20	I don't see how the proposed method is "enhancing the model selection procedure".	It was changed.
	Why is there no discussion on the phytoplankton application in the conclusion section? I think this should re-appear here if this motivated the study.	A respective paragraph is presently added.
Figures/Tables		
Fig. 3	*marine* vegetation season; Add unit to SST bias in the caption. I suggest to increase the line thickness of the median bias. Don't capitalize "standard deviation".	It was changed to "E. huxleyi bloom season". The second suggestion is accepted.
Fig. 5a-5e	I find it very confusing that you use "Fig. 5a", "Fig. 5b" etc. to refer to different figures rather than panels within the figures. I suggest to number the figures and then use the letters to refer to panels within each figure. As you nowhere show the pattern of the respective variable that we would expect from the reanalysis, I suggest to at least add the average of the reanalysis under each panel (as you have done for the model ensemble). Additionally, I suggest to refer back to the method section for the exact periods considered or to restate them in the caption.	Figures 5,6 and Supplements are replaced by a single figure with a box-plot, which more informative and more laconic.
Fig. 5a	You don't show sea surface salinity, do you? Define the bias (model-reanalysis or the reverse?) in the caption of *each* figure.	
Fig. 5e	Isn't it also the current direction that matters (see comments above). Have you assessed the circulation patterns as a whole?	
Fig. 6b	What are the "errors"	

Simulation of ~~oceanographic~~factors affecting *E. huxleyi* blooms in Arctic and ~~meteorological parameters in arctic and subarctic~~sub-Arctic seas by CMIP5 climate models: model validation and selection

Natalia Gnatiuk[1], Iuliia Radchenko[1], Richard Davy[2], Evgeny Morozov[1], Leonid Bobylev[1]

[1]Nansen International Environmental and Remote Sensing Centre, St. Petersburg, 199034, Russia
[2]Nansen Environmental and Remote Sensing Center, Bergen, N-5006, Norway

Correspondence to: Natalia Gnatiuk (natalia.gnatiuk@niersc.spb.ru)

Abstract.

The observed warming in the Arctic is more than double the global average and this enhanced Arctic warming is projected to continue throughout the 21st century. This rapid warming has a wide range of impacts on polar and sub-polar marine ecosystems. One of the examples of such an impact on ecosystems is that of coccolithophores, particularly *E. huxleyi*, which have expanded their range poleward during recent decades. The coccolithophore *E. huxleyi* plays an essential role in the global carbon cycle. Therefore, the assessment of future changes in coccolithophore blooms is very important.

Currently, there are a large number of climate models that give projections for various ~~oceanic and~~ oceanographic, meteorological ~~parameters~~, and biochemical variables in the Arctic. However, ~~their estimates often differ in absolute values in specific sea areas in comparison with the~~ individual climate models can have large biases when compared to historical ~~reanalysis data~~observations. The main goal of this research was to ~~find out the methodology of selection of the optimal model~~ select an ensemble of climate models that most accurately reproduces the state of ~~abiotic parameters inherent~~ environmental variables that influence the coccolithophore *E. huxleyi* bloom over the historical period when compared to reanalysis data. We developed a novel approach for model selection to include a diverse set of measures of model skill including the spatial pattern of some variables, which had not previously included in ~~six target arctic and sub-arctic~~a model selection procedure. We applied this method to each of the Arctic and sub-Arctic seas~~, viz. the Barents, Bering, Greenland, Labrador, North and Norwegian seas.~~ in which *E. huxleyi* blooms have been observed. Once we have selected an optimal combination of climate models that most skillfully reproduce the factors which affect *E. huxleyi*, the projections of the future conditions in the Arctic from these models can be used to predict how *E. huxleyi* blooms will change in the future.

Here, we present the validation of 34 CMIP5 atmosphere-ocean General Circulation Models (~~GCM~~GCMs) over the historical period 1979-2005. Furthermore, we propose a procedure of ~~model~~ ranking and ~~selection, which is~~ selecting these models based on the model's skill ~~to represent several~~in reproducing 10 important oceanographic~~ and~~, meteorological ~~parameters in the arctic and subarctic seas: the~~ , and biochemical variables in the Arctic and sub-Arctic seas. These factors include the concentration of nutrients (NO_3, PO_4, and SI), dissolved CO_2 partial pressure, pH, sea surface ~~(i) water~~

temperature~~ and (ii)~~, salinity ~~(~~averaged over the top ~~30 m); (iii)~~30m, 10m wind ~~speed at a height of 10 m above the surface; (iv) ocean~~ and surface current speed~~;~~, and ~~(v) surface~~ downwelling shortwave radiation~~.~~ at sea surface. The validation of the ~~GCMs~~GCMs' outputs against reanalysis data includes analysis of the interannual variability, seasonal cycle, spatial biases and temporal trends of the simulated ~~parameters~~variables. In total, ~~30~~60 combinations of ~~high-skilful~~ models were selected for ~~5~~10 variables over 6 study regions~~.~~ using the selection procedure we present here. The results show that there is no ~~mutually optimal~~ combination of models, nor is there ~~a~~ one ~~top~~ model, that has ~~a~~high skill in reproducing~~ either~~ the regional climatic-relevant features of ~~the whole Arctic region or~~ all combinations of the considered ~~parameters~~variables in target seas. Thereby, an individual subset of models was selected according to our ~~methodology~~model selection procedure for each ~~'~~combination of variable ~~– target sea' combination, a unique best model subset was selected with~~and Arctic/sub-Arctic sea. Following our selection procedure, the number of ~~included~~selected models ~~varying~~in the individual subsets varied from ~~7~~3 to 11.

The paper presents a comparison of the selected ~~best~~ model ~~sub-sets~~subsets and the full-model ensemble of all available CMIP5 models ~~with the respective~~to reanalysis data. The selected ~~best model sub-sets~~subsets of models generally show a better performance ~~vs.~~than the full-model ensemble~~ in more than 70% cases~~. Therefore we conclude that ~~confirms the advisability of using~~within the ~~proposed~~task addressed in this study it is preferable to employ the model ~~ranking method.~~subsets determined through application of our procedure than the full-model ensemble.

1 Introduction

In the last three decades, the Arctic has been warming at more than twice the rate of the global average (Davy et al., 2018; Overland and Wang, 2010). This rapid warming has led to large changes in the physical environment, for example with the loss of sea ice extent and volume (Dai et al., 2019; Kwok, 2018), but it has also had a large impact on the Arctic ecosystem (Hoegh-Guldberg and Bruno, 2010; Johannessen and Miles, 2011). One group of species that have been affected by Arctic warming are coccolithophores such as *Emiliania huxleyi* (hereafter *E. huxleyi*). Reportedly, coccolithophores can affect the carbon and sulphur cycles in the surface ocean, at least within their bloom areas(Balch et al., 2016; Kondrik et al., 2018; Malin et al., 1993; Rivero-Calle et al., 2015; Winter et al., 2013). The effect of these algae on aquatic carbon chemistry results in changes to the carbon fluxes between the atmosphere and ocean (Balch et al., 2016; Morozov et al., 2019; Pozdnyakov et al., 2019; Shutler et al., 2013). Additionally, they contribute to the generation of sulfate aerosols, which scatter solar radiation in the atmosphere and act as cloud condensation nuclei, enabling cloud formation (Malin and Steinke, 2004). Therefore, the coccolithophores are responsible for both warming and cooling effects on the global climate (Charlson et al., 1987; Wang et al., 2018a, 2018b).

Of all the coccolithophores, *E. huxleyi* is the most abundant and productive calcifying organism in the world ocean (McIntyre and Bé, 1967). It is a planktonic species growing at practically all latitudes (Brown and Yoder, 1994; Iglesias-

Rodríguez et al., 2002; Moore et al., 2012) and in the eutrophic to oligotrophic marine waters (Paasche, 2001). The property of this photosynthesizing aquatic organism to produce not only organic carbon, but also calcite, i.e. particulate inorganic carbon (PIC), imparts to *E. huxleyi* a special importance for the global ocean carbon cycle, and, through intricate interactions, for CO_2 exchange fluxes between the ocean and atmosphere (Kondrik et al., 2019; Morozov et al., 2019; Shutler et al., 2013). Moreover, *E. huxleyi* blooms are known to *i*) affect not only the carbon but also sulphur cycles in the surface ocean, at least within bloom zones, and arguably *ii*) contribute to the generation of sulfate aerosols, which eventually enable cloud formation (Malin and Steinke, 2004). This gives *E. huxleyi* blooms a definite climatic dimension in the overall environmental impact of this phenomenon. The scale of the impact should indeed be very significant: such blooms not only release into the water huge amounts of PIC, in some cases reaching nearly one million tons (Balch et al., 2016; Kondrik et al., 2018; Rivero-Calle et al., 2015), but they are very extensive typically covering marine areas in excess of many hundred thousand, sometimes up to one million, square kilometres. Besides they occur annually across the world ocean (Brown and Yoder, 1994; Iglesias-Rodríguez et al., 2002; Moore et al., 2012). Since changes of the regional climate have influenced the ecosystems of the Arctic seas, coccolithophores, particularly *E.huxleyi*, have increasingly expanded their range into Polar waters (Henson et al., 2018; Rivero-Calle et al., 2015; Winter et al., 2013), which is thought to be due to climate warming (Fernandes, 2012; Flores et al., 2010; Kondrik et al., 2017; Okada and McIntyre, 1979; Winter, 1994).

Although *E. huxleyi* cells can adapt to diverse environmental conditions, the blooms of this alga exhibit remarkable inter-annual variations in extent, intensity and localization (Balch et al., 2012; Iida et al., 2002; Kondrik et al., 2017; Morozov et al., 2013; Smyth et al., 2004). Importantly, the aforementioned spatio-temporal variations inherent in *E. huxleyi* blooms prove to be specific to individual marine environments, which indicates that *E. huxleyi* growth is generally conditioned by multiple forcing factors (FFs) acting through feedback mechanisms. Reportedly, the observed spatio-temporal variations are primarily driven by changes in water surface temperature (SST), salinity, levels of photosynthetically active radiation (PAR) and nutrients/micronutrients availability, such as nitrate (NO_3), silicate (SI), ammonium (NH_4), phosphate (PO_4) and iron (Fe) (Iglesias-Rodríguez et al., 2002; Krumhardt et al., 2017; Lavender et al., 2008; Zondervan, 2007). However, it has been found that, in addition to the above FFs, the water column stratification and wind speed at 10m above the surface (WS) also condition the growth of *E. huxleyi:* a decrease in wind stress leads to formation of a shallow mixed layer and retaining of algal cells within the zone of high levels of PAR (Raitsos et al., 2006). The intensity of water movements in general, and specifically water advection driven by ocean surface currents (OCS), was also highly consequential in this regard (Balch et al., 2016; Pozdnyakov et al., 2019). Among the other factors affecting *E. huxleyi* blooms are carbonate chemistry variables such as CO_2 partial pressure in the water, pCO_2, and pH, which are considered to be very important (Tyrrell and Merico, 2004). There has been speculation that the ongoing increase in atmospheric CO_2 should damp/inhibit the growth of coccolithophores (Rivero-Calle et al., 2015), however, this is not supported by multiple observations (Kondrik et al., 2017; Morozov et al., 2013).

As the above FFs are susceptible to climate change, these factors are expected to exert their combined influence on the intensity, spatial extent, and possibly the seasonal duration of *E. huxleyi* blooms in the future. Given that the environmental influence of this phenomenon has both climatological and biogeochemical dimensions at least on a synoptic scale, it appears important to envisage how it will evolve in the mid-term future. This can be done using either biological, e.g., (Gregg et al., 2005) or statistical, e.g., (Pozdnyakov et al., 2019) *E. huxleyi* bloom models, for which the prospective tendencies in FFs are employed. In turn, the tendencies in the FFs can be obtained from climate model output.

Today atmosphere-ocean coupled climate models are state-of-the-art tools for the ~~prediction~~projection of the future ~~status of the~~ climate~~ system components~~ on decadal and centennial time scales ~~(Otero et al., 2018; Taylor et al., 2012).~~(Otero et al., 2018; Taylor et al., 2012). In particular, the modern coupled atmosphere-ocean General Circulation Models (GCMs) include processes that govern the ~~main climate system components such as~~interactions between the ocean, atmosphere, ~~ocean,~~ land ~~and~~, sea~~-~~ ice, and ~~therefore, represent more realistically~~ the ~~processes of their interactions. Thus, the~~carbon cycle. The fifth phase of the Coupled Model Intercomparison Project (CMIP5) gives the opportunity to use ~~data of~~the model output from more than 30 GCMs ~~(Taylor et al., 2012).~~(Taylor et al., 2012). The GCMs provide a large number of ~~the~~ meteorological~~ and~~, oceanographic ~~parameters allowing to perform a~~and biochemical variables and so facilitate the comprehensive assessment of possible climate change impacts on marine ecosystems in the future. However, ~~most of~~ the studies ~~addressing~~which have evaluated the CMIP ~~models intercomparison show~~model's historical simulations have shown that the ~~GCMs~~model outputs ~~usually vary significantly~~have a large spread compared to natural variability (Almazroui et al., 2017; Fu et al., 2013; Gleckler et al., 2008). ~~Therefore, it is important to find a reliable approach for both model quality intercomparison and selection of optimal models for each specific scientific task and region~~The full CMIP5 model ensemble has been found to be skillful at simulating continent-wide surface air temperature, and therefore useful for making robust assessments at these scales (IPCC, 2013). However, model skill at smaller spatial scales, such as for the Arctic, or even for specific Arctic seas, varies considerably from region to region and for different model variables (Overland et al., 2011). Therefore, it is important to find an approach for both model evaluation (comparison with historical climate) and selection of optimal models for each specific scientific task and region that gives a skill score to each model which encompasses all the relevant model variables and properties that are important for the scientific question to be addressed.

The main goal of the paper is to ~~find~~quantify how well CMIP5 models reproduce the main forcing factors (FFs) that influence coccolithophore blooms in the Arctic and sub-Arctic seas. We propose a ~~reliable~~new approach for ~~CMIP5 model selection, in particular, those climate models that simulate most efficiently the state of abiotic parameters relevant to living~~ranking and selecting CMIP5 models for their skill in capturing the historical environmental conditions ~~of phytoplankton communities inherent in a number of seas at subpolar and polar latitudes~~in the Arctic and sub-Arctic seas (viz. the Barents, Bering, Greenland, Labrador, North and Norwegian seas). ~~Such~~We have chosen such a specific task ~~is selected~~ as a case study in order to select model output to ~~have the results that would be applied for projections of abiotic factors affecting the dynamics of phytoplankton communities.~~

~~It is well established that the method of ensemble averaging can be used~~drive a model of coccolithophore blooms to ~~reduce systematic model biases~~predict how these will change in the ~~individual climate models (Flato et al., 2013; Gleckler et al., 2008; Knutti et al., 2010a, 2010b; Pierce et al., 2009; Reichler and Kim, 2008). Furthermore, in case it is not possible to calibrate a model for a selected region, one of the main recommendation from climate model developers is to take into consideration more than one climate model (Flato et al., 2013; Gleckler et al., 2008; Knutti et al., 2010b; Pierce et al., 2009). There are two main approaches to employing climate model ensembles: (i) use of the full ensemble average data (Flato et al., 2013; Gleckler et al., 2008; Knutti et al., 2010b; Reichler and Kim, 2008); and (ii) selection of an ensemble of the best models from the entire set of available climate models based on a comparison with observational data for a historical period (Herger et al., 2018; Knutti et al., 2010b; Taylor, 2001). These two approaches are equally used depending on a specific scientific task: (i) full ensemble averaging for future trends analysis, and (ii) selection of the best models ensembles for regional climate features analysis. However, when there are many climate models available (e.g., in our study the number of models available varied from 25 to 30 depending on the climate variable), then the averaging method will result in very strong smoothing of data, and poor reproduction of the interannual variability. So that only the long-term trend of a given variable will be well captured.~~future. We assume that a climate model that successfully represents the present-day conditions will also ~~succeed in the future projections. Therefore, we chose the second approach, e.g., a selection of climate models that properly simulate the current regional features, including the spatial distribution, of the meteorological and oceanographic 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 the appropriate methodology for selection of the best model ensembles.~~be skillful in future projections. Therefore, we select models based upon the validation of the models within the historical period.

~~There are many approaches for the selection of an optimal set of climate models. One approach suggests choosing the models with focus only on some key climatological parameters, such as air temperature, precipitation and sea level pressure (Almazroui et al., 2017; Duan and Phillips, 2010; Pierce et al., 2009; Sarr and Sarr, 2017). This approach assumes that if the models skillfully reproduce these key parameters, they also must be good at reproducing the regional climate in general. Another approach, which is used in this study, is to select a unique combination of models for each study variable (Agosta et al., 2015; Anav et al., 2013; Fu et al., 2013; Gleckler et al., 2008). In order to select such a unique combination of models, it is necessary, firstly, to perform a validation of climate models through comparing GCMs outputs with the respective observations over a historical period, and then to identify the appropriate climate models based on statistical measures, i.e. to sort or rank the tested models. However, there are no generally accepted solutions for this task. For example, Agosta et al. (2015) ranked the CMIP5 models using only one statistical metric, viz, a climate prediction index, which is the ratio of the root mean square error to the standard deviation of observation data. Gleckler et al. (2008) evaluated the CMIP5 models and~~

~~ranked them through analyzing the climatology of the annual cycle, inter annual variability, and relative errors. They found that the performance of the analysed models varies for different parameters. Das et al. (2018) assessed 34 CMIP5 models using the following three criteria: the mean seasonal cycle, temporal trends, and spatial correlation. On this basis the models were selected using a cumulative ranking approach. Fu et al. (2013) and Ruan et al. (2019) applied a score-based method using multiple criteria for the assessment of CMIP5 model performance: mean value, standard deviation, normalized root mean square error, linear correlation coefficient, Mann-Kendall test statistic Z, Sen's slope, and significance score. Further, Ruan et al. (2019) selected the top 25% ranked CMIP5 models for composing a multi-model ensemble for air temperature projections over the Lower Mekong Basin. Fu et al. (2013) and Ruan et al. (2019) ranked the employed models using a weight criterion from 0.5 to 1.0. Ruan et al. (2019) reported that the introduction of multiple criteria results in less uncertainties in the models' performance in comparison with the respective observation data. However, Fu et al. (2013) and Ruan et al. (2019) did not consider the feature of spatial distribution of variables.~~

~~We decided to compile and improve the previously applied approaches that is to employ the multiple criteria ranking method following Fu et al. (2013) and Ruan et al. (2019) studies but (i) taking into consideration the Agosta et al. (2015) climate prediction index, (ii) analysing the features of spatial distribution of target variables (spatial biases and trends), (iii) ranking the models with the percentile method (25^{th},50^{th}, 75^{th}) that is widely used in statistical analysis, and, finally, (iv) selecting the top 25% ranked CMIP5 models following Ruan et al. (2019).~~

~~As the target arctic and subarctic seas differ in physical and geographical conditions, we performed the validation and selection model procedure for each sea individually. Moreover, we analyzed the specific marine areas with the stable localizations of intense growth of phytoplankton species both in spring (e.g. diatoms) and in summer-autumn (e.g. coccolithophores Kondrik et al., 2017; Smyth et al., 2004). Thus, the target regions permitted to identify the CMIP5 models that represented most closely the cumulative state of the physical environmental factors (abiotic parameters) characterizing the conditions, under which the aforementioned blooms occurred. Such a specific task eventuated in the results that can be useful for further improvements of marine ecological models encompassing the phytoplankton community as well as for modelling the dynamics of physical parameters relevant to surface water environment at high-latitude seas under conditions of changing climate.~~

2 Materials and method

2.1 Data

~~Thirty-four~~34 CMIP5 ~~GCMs~~GCMs' outputs for the historical period 1979-2005 were used in this study. The data are freely available on the ESGF portal (https://esgf-node.llnl.gov). The list of climate models used is presented in Table 1. We analyzed five oceanographic and meteorological variables, namely the sea surface temperature (SST~~) and~~), salinity averaged over 0-30 m (~~SSS~~SS_{30m}), surface wind speed at a height of 10 m (WS), ocean surface current speed (OCS), and shortwave downwelling solar radiation (SDSR~~).~~); and 5 biochemical variables, namely concentration of nutrients (NO3, PO4, and SI), dissolved CO_2 partial pressure (pCO_2), and pH. These ~~abiotic parameters~~forcing factors (FFs) are known to affect the phytoplankton life cycle in sub-polar and polar latitudes ~~(Iglesias-Rodríguez et al., 2002; Raitsos et al., 2006; Winter et al., 2013). The availability of the CMIP5 GCMs analysed in this study are listed in Table1: in total, we used 25 models for OCS, 28 for SSS, SST, SDSR, and 30 for WS. For validation of the climate models outputs~~(Iglesias-Rodríguez et al., 2002; Raitsos et al., 2006; Winter et al., 2013). The analyzed CMIP5 GCMs are listed in Table1: in total, we used outputs of 25 models for OCS, 28 for SS_{30m}, SST, and RDSR, 30 for WS; 11 for PO_4, 13 for SI and pH, 15 for pCO_2, and 16 for NO_3. The number of models employed is different and was dictated by their availability on the ESGS portal. For validation of the climate models outputs, we used atmospheric and oceanic reanalyses: (*i*) Era-Interim from the European Centre for Medium-Range Weather Forecasts (https://apps.ecmwf.int) (Dee et al., 2011) for ~~the surface wind speed at 10 m, sea surface temperature, and shortwave downwelling solar radiation~~SST, WS, and RSDS for the period from 1979 to 2005; and (*ii*) GLORYS2V4 from the European Copernicus Marine Environment Monitoring Service (http://marine.copernicus.eu) for the ~~sea surface salinity and ocean surface current speed for the period 1993-2005.~~SS_{30m}, OCS, and FREEBIORYS2V4 reanalyses for biochemical variables (Perruche, 2018) for the period 1993-2005. The period for verification of the employed climate models was chosen based on the length of the reanalysis data and the limitations inherent in the “historical” runs of the GCMs, which usually terminate in 2005. The selected reanalyses are widely used in the literature and have been shown to be consistent with independent observational data (Agosta et al., 2015; Dee et al., 2011; Garric et al., 2017; Geil et al., 2013).

2.2 Methods for model selection

It is well established that the method of ensemble averaging can be used to reduce systematic model biases in the individual climate models (Flato et al., 2013; Gleckler et al., 2008; Knutti et al., 2010; Pierce et al., 2009; Reichler and Kim, 2008). There are two main approaches to employing climate model ensembles: (i) use of the full-ensemble average data for future trends analysis (Flato et al., 2013; Gleckler et al., 2008; Knutti et al., 2010; Reichler and Kim, 2008); and (*ii*) selection of an ensemble of the models from the entire set of available climate models yielding the best fit to the observational data for a historical period (Herger et al., 2018; Knutti et al., 2010; Taylor et al., 2012). We chose the second approach for analysing the ability of GCMs to reproduce main forcing factors (FFs) that influence *E. huxleyi* bloom: nutrient concentrations (nitrates, phosphates, silicates), salinity averaged over the top 30 m (SS_{30m}), sea surface temperature (SST), wind speed (WS), downwelling shortwave radiation at the surface (RSDS), pH, *p*CO2, and ocean current speed (OCS).

There are many different approaches to ranking and selection climate models following validation with historical observations. For example, Agosta et al. (2015) ranked the CMIP5 models using only one statistical metric, viz, a climate prediction index (CPI), “which is widely used in climatology studies for model evaluation and weighted projections” (Connolley and Bracegirdle, 2007; Franco et al., 2011; Murphy et al., 2004). Gleckler et al. (2008) evaluated the CMIP5 models and ranked them by analyzing the climatology of the annual cycle, inter-annual variability, and relative errors. They found that the performance of the analyzed models varied for different variables. Das et al. (2018) assessed 34 CMIP5 models using the following three criteria: the mean seasonal cycle, temporal trends, and spatial correlation. On this basis, the models were selected using a cumulative ranking approach. Fu et al. (2013) and Ruan et al. (2019) applied a score-based method using multiple criteria for the assessment of CMIP5 model performance: mean value, standard deviation, normalized root mean square error, linear correlation coefficient, Mann-Kendall test statistic Z, Sen's slope, and significance score. Further, Ruan et al. (2019) selected the top 25% ranked CMIP5 models by applying a weight criterion from 0.5 to 1.0 to the different measures. Ruan et al. (2019) reported that the introduction of multiple criteria results in fewer uncertainties in the models' performance in comparison with the respective observation data.

Having tested the approaches cited above, we developed our own methodology which combines elements from some of these. We employ the multiple criteria ranking method following Fu et al. (2013) and Ruan et al. (2019) studies but with the following modifications: (*i*) we took into consideration the Agosta et al. (2015) climate prediction index, (*ii*) analyzed the features of spatial distribution of target variables (spatial biases and trends), (*iii*) ranked the models with the percentile method (25^{th},50^{th}, 75^{th}) that is widely used in statistical analysis, and, finally, (*iv*) we selected the top 25% ranked CMIP5 models following Ruan et al. (2019).

2.2.1 Study regions

~~The target regions are six arctic and subarctic seas: the Barents, Bering, Greenland, Labrador, North and Norwegian seas. Only specific areas were selected in each target sea relying on~~The target regions are six Arctic and sub-Arctic seas: the Barents, Bering, Greenland, Labrador, North and Norwegian seas, where *E. huxleyi* blooms regularly occur (Kondrik et al., 2017). As mentioned above, the reason we chose the listed seas was that, in the context of global climate change, the Arctic and sub-Arctic seas have experienced the most pronounced changes in environmental variables due to the Arctic amplification. In addition, the target seas differ in physical and geographical conditions, which strongly affect their climate. While they are linked by common circulation patterns, e.g., with the warm air advection coming into the Arctic from the Atlantic Ocean, how this circulation affects the climate in a given sea is strongly affected by the local conditions. Therefore, we performed the validation and selection model procedure for each sea individually. Only specific areas within which intense growth/blooms of *E. huxleyi* frequently occur were selected in each sea according to the results obtained by Kazakov et al. (2018)~~ for the coccolithophore *E. huxleyi* blooms~~ based on the Ocean Colour Climate Change Initiative dataset version 3.0 (https://esa-oceancolour-cci.org/) for the period from 1998 to 2016. ~~The selection~~A comparison of the ~~listed seas~~area-averaged values for the entire sea and ~~the specific areas within them was prompted by several reasons: firstly, in the context~~only for the region of ~~global climate change,~~ the ~~subarctic and arctic seas are characterized by one~~regular occurrence of ~~the most pronounced changes in environmental parameters due to the so called Arctic amplification, and, secondly, in the target water areas, summer-autumn phytoplankton blooms (e.g.~~ *E.* *huxleyi*~~) have~~ blooms showed a ~~steady localization, while~~

~~in other parts of the investigated seas the localization of phytoplankton blooms is variable from year to year.~~significant difference. For ~~identifying the specific~~example, it is about 2^{o}C degrees among all models for SST in the Barents Sea where the *E. huxleyi* blooms cover the largest area of the sea compared to other seas. To identify the relevant study areas~~, on the~~ from a raster image ~~with~~that contained all blooming events ~~during~~ over the period 1998-2016, we ~~masked~~selected those polygons ~~that confine the territories seas~~ where blooms occurred for more than one 8-day period (Fig. 1). ~~Besides, the periods for~~ For model validation ~~were selected based~~we focused on ~~a~~ sea-specific blooming periods~~, which include all summer months and, in some cases, beyond them~~: June-September for the Barents and Labrador seas, June-August for the Greenland Sea, May-July for the North Sea, May-August for the Norwegian Sea, and January-December for the Bering Sea (Kazakov et al., 2018). Thus, ~~it is noteworthy that~~ the results of the ~~performed comparison of models~~model validation can be used not only in terms of marine ecology-related issues (i.e. carbon cycle chemistry, water acidity, nutrients availability, etc.) but also for the purposes of forecasting ~~of the~~ region-specific climate~~ interactions during the vegetation season, taking into account that the selection of the climate models was carried out individually for each sea/sea zone~~-driven feedbacks between the environmental factors governing *E. huxleyi* growth.

2.~~3~~2.2. Model evaluation ~~metrics~~measures

The CMIP5 climate models were validated against ~~the~~ reanalysis data in order to assess how well they reproduce the regional features of ~~distribution~~ the selected ~~parameters/~~variables. The validation methodology for the ~~GCMs~~GCMs' outputs included the analysis of the climatological-mean seasonal cycle, interannual variability and trends, and analysis of the spatial ~~distributions~~distribution of climatological-mean biases and trends for selected ~~parameters~~variables averaged over the blooming period in each sea.

a) *The seasonal cycle* was analyzed using the multi-year averaged monthly variables for all months of the year (i.e., a sample size of 12). Basic statistical measures were calculated, such as the root-mean-square deviation (RMSD), the correlation coefficient (r), and the standard deviation (SD) (Fu et al., 2013; Gleckler et al., 2008; Kumar et al., 2015; Ruan et al., 2019). In addition, following Agosta et al. (2015) we calculated the climate prediction index (CPI) for the seasonal cycle, which is a ratio of the model root mean square error to the standard deviation of observation data. This model evaluation ~~statistics~~statistic weighs the simulated data against the observations and is often used to validate model ~~data~~output (Agosta et al., 2015; Golmohammadi et al., 2014; Moriasi et al., 2007; Murphy et al., 2004; Stocker, 2004).

b) *The interannual variability* of the ~~parameters~~variables was analyzed based on monthly variables solely for blooming periods (the sample size varied according to sub-region and ~~parameter~~variables combination, e.g., a sample size for SST in the Barents Sea was 108~~—~~ monthly variables from June to September during 1979-2005). The same statistical measures for analysis of the seasonal cycle were used, viz. RMSD, r, SD, and CPI.

c) *The spatial distribution of biases and trends* between the model outputs and the reanalysis data were calculated for temporal-averaged data in each grid point of the ~~target~~ marine zone considered in this study.

2.~~4.~~2.3 Percentile ~~score-based model~~ ranking ~~method~~approach

For ranking models and selection of the ~~best~~ model sub-set, we ~~proposed and~~ employed the percentile ~~score-based model~~ ranking ~~method~~approach, which is a compilation of the previously applied model ranking and the selection approaches with some modifications (see also ~~Introduction).~~2.2 Methods for model selection). Following Fu et al. (2013) and Ruan et al. (2019), we used ~~the~~ multiple criteria for model selection (RMSD, r, SD). Following Agosta et al. (2015) we ~~analysed~~analyzed the climate prediction index (CPI~~),~~) and considered the differences in spatial distributions of biases and trends between the model outputs and the respective reanalysis data. Further, we ranked the models based on the percentile method (25th, 50th, 75th) for each ~~obtained~~ statistical ~~metrics~~measure based on the amplitude of its values. Finally, we selected the top 25% ranked CMIP5 models following Ruan et al. (2019) for each considered oceanographic and meteorological ~~parameter~~variables, and the target region. Thus, for example, for a sample of 28 models, the top 25% is a sub-set of 7 models that showed the best total~~-~~ score~~.~~ (the sum of every score of statistical measures, see Tab. 2). However, if two or more models show the same score, they are all ~~are~~ included in the selected ~~best~~ model sub-set. Thus, the number of ~~included models in~~ selected ~~best model subsets varying~~models varies from ~~7~~3 to 11.

Figure 2 illustrates an example of the percentile ~~score-based~~ ranking approach applied to the RMSD of ~~the sea surface temperature~~SST in the Barents Sea. We divided the ~~obtained~~ statistical measures into 4 groups based on the amplitude of the values and assigned a score to each model according to its group: (*i*) ~~models considered as~~ very good ~~(less than~~models (top 25th percentile of the distribution of the statistical ~~metrics distribution~~measures) were given a score of 3; (*ii*) good models (between 50th and 25th percentile) were given a score of 2; (*iii*) satisfactory models (between 75th and 50th percentile) were given a score of 1; and (*iv*) unsatisfactory models (more than 75th percentile) were given a score of 0. In the case of the correlation coefficient, it is vice versa, very good models with correlations scores above 0.75 were ranked with a score of 3, and so forth.

For ranking models based on the ~~obtained~~ differences in the spatial ~~distributions~~distribution of biases and trends between model outputs and reanalysis, we ~~analysed~~used the absolute values of the ~~median~~mean and the ~~amplitude~~spread of the spatial variation in model biases. For example, Figure 3 displays the box plots of spatial variability in SST biases relevant to the ~~target~~studied area in the Barents Sea for the ~~vegetation~~blooming season (June-September) and the study period 1979-2005. The ~~median~~mean bias varies from -6.6 (model #20) to 1.5 ~~K~~^{o}C (model #24) among the models, whereas the ~~amplitude bias~~spread yielded by the model and that from observations has a wide spread of values from 7.3 (model #21) to 16.5 ~~K~~^{o}C (model #3). Thus it can be concluded from Fig. 3 that the analysis of spatial distribution of biases is very important, e.g., if we compare model #2 (ACCESS1-3) with model #3 (CanESM2), we can see that the ~~medians~~means of these two models have a small difference (0.28 ~~K~~^{o}C), while, the ~~amplitude~~spread of spatial values for model #3 is much higher (by ~6 ^{o}C) than that for model #2. Application of the percentile ~~score-based method~~ranking approach to ~~modes~~model #2 (ACCESS1-3) and

#3 (CanESM2) resulted in inclusion of only the former in the ~~best~~ model sub-set~~, whereas the latter was placed beyond it~~ (Fig. 4).

Table 2 presents all calculated statistics that were used to rank GCMs for SST in the Barents Sea as well as the final total score for each model. The spread of the total assigned scores is from 9 to 35. Based on this range we selected the top 25% of GCMs. Thus, the best model ensemble for SST for the Barents Sea is the 8-model set: ACCESS1-0; ACCESS1-3; GFDL-CM3; HadGEM2-ES; MIROC-ESM; MIROC-ESM-CHEM; MPI-ESM-LR; MPI-ESM-MR. ~~Additionally, we identified the top model for SST in this region – MIROC-ESM.~~ The same procedure was performed for other target seas~~/zones~~ and variables.

3 Results and discussion

The results of model validation and ranking, as well as the selected ~~best~~ CMIP5 model ~~sub-sets for five oceanographic and meteorological variables, viz. the 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)~~subsets in the Barents, Bering, Greenland, Labrador, North and Norwegian seas are presented in Fig. ~~4.~~4 (for 5 oceanographic, and meteorological variables), and Fig. 5 (for 5 biochemical variables). Each number ~~of~~in the heat ~~map~~maps shows the final skill score for ~~one~~each combination of model~~-~~, variable~~ intersection.~~, and sea. For each individual column, ~~its own~~a colour gradation was applied based on our percentile ranking approach~~;~~: therefore, the same numbers can have different colours on the heat ~~map~~maps. For example, for OCS in the Barents Sea, the spread of the final model scores is from 7 to 26, whereas for ~~SSS~~SS_{30m} it is from 8 to 34. Therefore, even model #3 CanESM2 has the total score 26 for ~~SSS~~SS_{30m} (which is higher than that (25) for OCS), this model was not included in the ~~SSS best~~ SS_{30m} selected model sub-set and ~~has~~is coloured red~~ color~~, whereas for OSC it is included in the ~~best~~selected model sub-set and ~~has~~highlighted in green ~~color~~colour. The final skill scores of ~~the~~those models, which were ~~selected as~~included in the ~~best~~ model sub-sets are ~~highlighted~~marked in bold blue, and their total number is indicated at the bottom of each column.

Analysing the heat ~~map~~maps, one can conclude~~,~~ that there is no ~~an optimal~~ model ensemble, or ~~a one top~~ single model, which could ~~properly~~equally well simulate all ~~parameters~~variables over the different target seas~~/regions~~. However, some climate models show good results for many cases, e.g., ACCESS1-3; ACCESS1-0; GFDL-CM3; GISS-E2-R; GISS-E2-R-CC; HadGEM2-AO; HadGEM2-CC; HadGEM2-ES; INMCM4~~.~~; MPI-ESM-LR; MPI-ESM-MR. The models that have ~~higher biases~~the lowest total scores across the majority of the target regions are CMCC-CM; FGOALS-g2; IPSL-CM5A-LR; IPSL-CM5A-MR; IPSL-CM5B-LR; MIROC5; MRI-ESM1.

~~Such heterogeneity of climate models ability to equally reproduce the regional climate features residing in different seas can be explained by various reasons. Climate models are often tuned to adequately reproduce global processes and globally averaged values. An insufficient number of long time series of observations is available for model calibration, especially for~~

~~marine tracts. GCMs errors increase to the poles because of the convergence of meridians at the poles. In addition, the target arctic and sub-arctic seas are essentially different in terms of their physical and geographical conditions, which could also cause the ability of the GCMs to reproduce well the conditions in some seas and fail in others.~~

~~In order to analyse how well the selected best-model sub-sets represent five studied parameters, we analysed the spatial distribution of biases between the selected model ensemble and the respective reanalysis data in six target seas, viz, the Barents, Bering, Labrador, Greenland, Norwegian and North seas (Figure 5a-e). Thus, fewer biases in SSS are determined in the case of the Labrador, Greenland and Norwegian seas (±0.5 psu); high positive biases observed in the Bering Sea next to the coastline: up to 1.5-4 psu, this overestimation is possibly due to insufficiently accurate parameterization of the river runoff in the sub-arctic region (Figure 5a). 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 some overestimation of river runoff or underestimation of salty atlantic water. The selected CMIP5 models simulate SDSR (Figure 5b) well almost in all target seas: the biases in SDSR in the Barents Sea vary from 5 to 14 W m^{-2} (≈4-10 %), in the Bering Sea – from 2 to10 W m^{-2} (≈2-9 %), in the Greenland Sea – from 0 to 12 W m^{-2} (≈0-7 %), in the North Sea – from 1 to17 W m^{-2} (≈0-7 %), in the Norwegian Sea – from 4 to 9 W m^{-2} (≈2-5 %), only in the Labrador Sea the CMIP5 models overestimate SDSR and the biases much higher – from 20 to 29 W m^{-2} (≈11-15 %). The selected GCMs simulate WS well in all studied seas: the biases in WS are not more than 1 m s^{-1}, only in some places of the Bering and North Seas' coastal regions, the biases in WS simulations are up to about 1.5 m s^{-1} (Figure 5c). Concerning SST, we also obtained quite good results for the selected models. Low biases were observed mainly over the entire territory of the North and Norwegian seas constituting ±0.5° C (Figure 5d). Near the English Channel models overestimate the temperature by ≈2° C in the North Sea probably due to the influence of warm water from the English Channel, and models slightly underestimate the temperature by ≈1° C near the coastline in the Norwegian Sea. In the Labrador Sea, the CMIP5 models simulate SST with lower biases in the northern and north-western parts of the sea – ±0.5° C (Figure 5d). However, in the southern and south-western parts of the sea, the models underestimate SST by ≈1-2° C, which is possibly due to the influence of the cold Labrador Current. In the Greenland Sea, the models underestimate SST by ≈1-1.5° C in the west probably also due to the influence of the cold Greenland Current and overestimate SST by ≈2° C in the south apparently due to overestimation of contribution of the warm Atlantic water (the North Atlantic Current). In the Barents Sea, the models overestimate north-western part of the sea probably due to the influence of the warm atlantic water, and in the southern part of the study area, the models underestimate SST by ≈1-2° C probably due to some underestimation of the influence of coming warm atlantic waters. Finally, the CMIP5 models simulate the surface ocean current speed with rather large biases, especially in the Bering Sea and closer to the Bering Strait (-0.19…0.14 m s^{-1}), where the models mainly overestimate OCS (Figure 5e). Smaller biases in the modeling of the OCS by CMIP5 models found for the Barents and Greenland seas – from -0.06 to 0.03 m s^{-1}. The biases in the other studied seas vary from -0.17 to 0.06 m s^{-1}.~~

~~To examine our percentile score-based model ranking method we analysed the spatial distribution of biases and trends for the full-model ensemble, selected best-model sub-set and top-model vs. reanalysis data for each target sea and parameter~~

~~combination. Figure 6 illustrates the case for SST in the Barents Sea, and in the Supplements we present maps for all variables and target regions. As seen in Fig. 6a, the full 28 model ensemble underestimates the SST in the target region while the top model, MIROC-ESM, overestimates it. The selected 8-model ensemble shows smaller biases (± 1 K) in SST for the most part of the sea. Illustrated in Fig. 5b, the spatial distribution of SST trends (the difference between model data and reanalysis data) indicates that the full 28-model ensemble overestimates the trends for the whole sea (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 it reveals reanalysis small trends (± 0.01 K yr^{-1}) that are similar to Era-Interim. As for the selected 8-model ensemble, the spatial variability of errors in trends in SST varies from -0.01 to 0.06 K yr^{-1}, although for the major part of the study region the errors are in the range -0.01 to 0.02 K yr^{-1}. Analysis of comparison of all selected model sub-sets (see Supplements) shows that, in general, the selected best-model ensemble assures somewhat better performance (with regard to the biases between model and reanalysis data) than either the full-model ensemble or the single top-model do. Comparing the full-model ensemble, selected sub-set models or/and top-model performance in terms of biases and trends, the selected best-model ensembles are more skilful in parameter simulations, respectively in 74% (biases) and 83% (trends) cases. The performance of the selected models proved to be equal to the full-model ensemble and top-model efficiency, respectively in 13% (biases) and 10% (trends) cases, and they are less skilful in the simulations in 13% (biases) and 7% (trends) cases.~~

Such heterogeneity in the ability of climate models to reproduce the climate features in different seas can be partly explained. Climate models are often tuned to adequately reproduce global processes and globally averaged values (Mauritsen et al., 2012; Schmidt et al., 2017). An insufficient number of long-time series of observations is available for model calibration, especially for marine waters. There are also very limited observations of climate processes in the Arctic which limit model development for the Arctic environment (Vihma et al., 2014).

In order to verify our methodology, we compared selected ensemble with the full model ensemble for the time-averaged spatial distribution of biases, relative to reanalyses data for the historical period (1979/1993-2005), for each study variable in 6 target seas (Fig. 6). The box plots (Fig. 6) show that the selected model ensemble mainly performs better than the full-model ensemble, i.e. mean value (red dot) located closer to the zero line (dashed). The biggest difference between these two approaches obtained for the concentration of Silicium (SI) in favor of the ranking model approach.

Analysing the box plots of the selected model ensemble (Fig. 6), the lower spread of biases is obtained for ocean current speed (OCS), salinity averaged over 30 m (SS_{30m}), and concentration of Silicium (SI). CMIP5 GCMs generally underestimate RSDS, especially over the Labrador Sea. Likewise, GCMs mainly underestimate WS except for the Labrador and Barents seas. For OCS all ensembles have a low spread of biases and its mean value located very close to zero but they have many outliers (black dots). CMIP5 GCMs in different seas show heterogeneous results – they underestimate or overestimate SST, SS_{30m}, and all biochemical variables. Also, Séférian et al. (2013) reported that CMIP5 GCMs differ enormously in biochemical variables but they show fewer biases comparing to the previous model versions (CMIP3) for wind speed. Flato et al. (2013) found that CMIP5 models have higher biases (both positive and negative) for SST in polar regions, and quite large negative biases relative to other latitudes for salinity in the Arctic. Rickard et al. (2016) summarised that oceanographic variables in CMIP5 models reveal better agreement across all models compared to biochemical ones. Lavoie et al. (2013) detected that GFDL and MPI models better represent nitrate concentrations, and GFDL model best represents salinity among other considered models in the Labrador Sea. In our study, these models also selected as best for the Labrador Sea. It is quite difficult to compare obtained results with other already published researches because of using different models or a various number of models in full-ensemble and study regions. Some mentioned authors apply full-model ensemble other select models with better performance, but they didn't compare these two approaches as we have done.

4 Conclusions

~~A~~ In the paper, we presented results of validation of 34 CMIP5 models compared to ERA-Interim, GLORYS2V4 and FREEBIORYS2V4 reanalyses for the historical period (1979/1993-2005). Besides we proposed the percentile ~~score-based model~~ ranking ~~method has been presented~~ approach for selection ~~of optimal model ensembles from a total of 34 CMIP5 models, for five different~~ climate ~~relevant variables (SST, WS, SSS, OCS, SDSR)~~ model sub-sets that most accurately reproduces the state of 10 forcing factors affecting *E.huxleyi* blooms over the historical period in six ~~arctic~~Arctic and ~~subarctic~~sub-Arctic seas, viz. the Barents, Bering, Labrador, Greenland, North, and Norwegian seas. ~~The best model ensembles for each parameter and each target sea were selected (in total 30~~In total 60 combinations of the most-~~skilful~~skillful models were selected (10 variables and 6 target seas) based on different statistical measures: the root mean square error, correlation coefficient, standard deviation, ~~RMSD observations standard deviation ratio,~~climate prediction index (CPI), spatial biases and trends. Our results show that there is no ~~any optimal~~ model ensemble or ~~a one top~~ individual model, which could best simulate all ~~parameters~~variables across all target seas. Despite the fact that the Arctic is often considered as one single region in many studies, our results show that CMIP5 climate models do not have consistent performance across such a large area. However, the selected ~~best~~ model ensembles show ~~quite good~~ results with ~~lesser biases in~~ smaller ~~study regions, i.e., some specific arctic seas~~biases than the full-model ensemble.

~~To assure best implementation of the model selection results, it is essential to select climate models that properly simulate the spatial distribution of the chosen variables. Therefore, we suppose that the spatial distribution of biases and trends in the considered parameters are as well important as other statistical metrics within the framework of the model selection procedure performed. Based on our~~ The results~~, we can also conclude that it is essential not only to analyse spatially averaged values, but also the spatial distribution of their amplitudes.~~

~~The results of examination~~ of the percentile ~~score-based model~~ ranking ~~method~~approach proposed in this paper ~~generally reveal a~~show better performance (mean is closer to zero) of the selected ~~best~~ model ensemble vs. the full-model ensemble ~~or a single best model~~ for different variables and target ~~regions.~~

seas. We can conclude that ~~the range of different factors is important for model selection~~it is important to include a number of different evaluation criteria when selecting the best models from an ensemble, including the spatial pattern of model biases, and that the proposed methodology is a way of ~~enhancing~~improving the model selection ~~procedures sophistication~~procedure that promises a better chance to identify more skillful models for the features we are interested in. ~~Thus, the proposed method can be used for analyses to be done for other seas/regions with the purpose to evaluate the performance climate models in terms of various atmospheric and oceanic parameters at different scales.~~

Given that the environmental impacts of *E. huxleyi* communities are diverse and encompass both climatological and marine ecology dimensions, the established sets of CMIP5 climatological models most closely simulating the environmental conditions under which this taxon grow, open the way for envisaging how this phenomenon will further evolve in light of ongoing climate change. This can be done using *E. huxleyi* bloom model, for which the changes in the forcing factors for *E. huxylei* blooms will be employed. Finally, although the present study has been performed for the coccolithophore *E. huxleyi* which vegetates at Arctic and sub-Arctic latitudes, the reported methodological approach is not algal-specific and can be applied to studies of other algal species composing the phytoplankton communities in the world ocean.

Author contribution

NG, RD, LB: methodology development. NG, IR: development of the paper concept. IR, NG, EM: data processing and figures producing. All authors contributed to the writing and discussion of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

~~*Acknowledgements*~~

Acknowledgments

We express our gratitude for the financial support of this study provided by the Russian Science Foundation (RSF) under the project 17-17-01117.

N. Gnatiuk and I. Radchenko thank Dmitry Pozdnyakov and Dmitry Kondrik for the invitation to participate in the project as well as for very useful discussions of the results obtained.

We acknowledge the members of the Coupled Model Intercomparison Project phase 5, the European Centre for Medium-Range Weather Forecasts, the European Copernicus Marine Environment Monitoring Service, and we extend our gratitude to the modelling groups specified in Table 1).

References

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.

Almazroui, M., Nazrul Islam, M., Saeed, S., Alkhalaf, A. K. and Dambul, R.: Assessment of Uncertainties in Projected Temperature and Precipitation over the Arabian Peninsula Using Three Categories of Cmip5 Multimodel Ensembles, Earth Syst. Environ., 1(2), 23, doi:10.1007/s41748-017-0027-5, 2017.

~~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.~~

Balch, W. M., Drapeau, D. T. and Bowler, B. C.: Step-changes in the physical, chemical and biological characteristics of the Gulf of Maine, as documented by the GNATS time series, Mar. Ecol. Prog. Ser., 450, 11–35, doi:10.3354/meps09555, 2012.

Balch, W. M., Bates, N. R., Lam, P. J., Twining, B. S., Rosengard, S. Z., Bowler, B. C., Drapeau, D. T., Garley, R., Lubelczyk, L. C., Mitchell, C. and Rauschenberg, S.: Factors regulating the Great Calcite Belt in the Southern Ocean and its biogeochemical significance, Global Biogeochem. Cycles, 30(8), 1124–1144, doi:10.1002/2016GB005414, 2016.

Brown, C. W. and Yoder, J. A.: Coccolithophorid blooms in the global ocean, J. Geophys. Res., 99(C4), 7467, doi:10.1029/93JC02156, 1994.

Charlson, R. J., Lovelock, J. E., Andreae, M. O. and Warren, S. G.: Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate, Nature, 326, 655–661, 1987.

Connolley, W. M. and Bracegirdle, T. J.: An Antarctic assessment of IPCC AR4 coupled models, Geophys. Res. Lett., 34(22), L22505, doi:10.1029/2007GL031648, 2007.

Dai, A., Luo, D., Song, M. and Liu, J.: Arctic amplification is caused by sea-ice loss under increasing CO 2, Nat. Commun., 10(1), doi:10.1038/s41467-018-07954-9, 2019.

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.

Davy, R., Chen, L. and Hanna, E.: Arctic amplification metrics, Int. J. Climatol., 38(12), 4384–4394, doi:10.1002/joc.5675, 2018.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137(656), 553–597, doi:10.1002/qj.828, 2011.

~~Duan, Q. and Phillips, T. J.: Bayesian estimation of local signal and noise in multimodel simulations of climate change, J. Geophys. Res., 115(D18), D18123, doi:10.1029/2009JD013654, 2010.~~

Fernandes, M.: The Influence of Stress Conditions on Intracellular Dimethylsulphoniopropionate (DMSP) and Dimethylsulphide (DMS) Release in Emiliania huxleyi., 2012.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C. and Rummukainen, M.: Evaluation of Climate Models, in Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assess- ment Report of the Intergovernmental Panel on Climate Change, pp. 741–866, NY~~. [online] Available from: https://pure.mpg.de/rest/items/item_1977534/component/file_3040450/content (Accessed 15 April 2019),~~., 2013.

Flores, J. A., Colmenero-Hidalgo, E., Mejía-Molina, A. E., Baumann, K. H., H., J., L. and Rodrigues, T.: Distribution of large Emiliania huxleyi in the Central and Northeast Atlantic as a tracer of surface ocean dynamics during the last 25,000 years, Mar. Micropaleontol., 76, 53–66, 2010.

Franco, B., Fettweis, X., Erpicum, M. and Nicolay, S.: Present and future climates of the Greenland ice sheet according to the IPCC AR4 models, Clim. Dyn., 36(9–10), 1897–1918, doi:10.1007/s00382-010-0779-1, 2011.

Fu, G., Liu, Z., Charles, S. P., Xu, Z. and Yao, Z.: A score-based method for assessing the performance of GCMs: A case study of southeastern Australia, J. Geophys. Res. Atmos., 118(10), 4154–4167, doi:10.1002/jgrd.50269, 2013.

Garric, G., Parent, L., Greiner, E., Drévillon, M., Hamon, M., Lellouche, J.-M., Régnier, C., Desportes, C., Le Galloudec, O., Bricaud, C., Drillet, Y., Hernandez, F. and Le Traon, P.-Y.: Performance and quality assessment of the global ocean eddy-permitting physical reanalysis GLORYS2V4., 19th EGU Gen. Assem. EGU2017, Proc. from Conf. held 23-28 April. 2017 Vienna, Austria., p.18776, 19, 18776~~ [online] Available from: http://adsabs.harvard.edu/abs/2017EGUGA..1918776G (Accessed 16 April 2019),~~, 2017.

Geil, K. L., Serra, Y. L., Zeng, X., Geil, K. L., Serra, Y. L. and Zeng, X.: Assessment of CMIP5 Model Simulations of the North American Monsoon System, J. Clim., 26(22), 8787–8801, doi:10.1175/JCLI-D-13-00044.1, 2013.

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.

Gregg, W. W., Casey, N. W. and McClain, C. R.: Recent trends in global ocean chlorophyll, Geophys. Res. Lett., 32(3), L03606, doi:10.1029/2004GL021808, 2005.

Henson, S. A., Cole, H. S., Hopkins, J., Martin, A. P. and Yool, A.: Detection of climate change-driven trends in phytoplankton phenology, Glob. Chang. Biol., 24(1), e101–e111, doi:10.1111/gcb.13886, 2018.

Herger, N., Abramowitz, G., Knutti, R., Angélil, O., Lehmann, K. and Sanderson, B. M.: Selecting a climate model subset to optimise key ensemble properties, Earth Syst. Dyn., 9(1), 135–151, doi:10.3929/ETHZ-B-000246202, 2018.

Hoegh-Guldberg, O. and Bruno, J. F.: The Impact of Climate Change on the World's Marine Ecosystems The role of oxidative stress in differential coral bleaching View project, Science (80-.)., 328, 1523–1528, doi:10.1126/science.1189930, 2010.

Iglesias-Rodríguez, M. D., Brown, C. W., Doney, S. C., Kleypas, J., Kolber, D., Kolber, Z., Hayes, P. K. and Falkowski, P. G.: Representing key phytoplankton functional groups in ocean carbon cycle models: Coccolithophorids, Global Biogeochem. Cycles, 16(4), 47-1-47–20, doi:10.1029/2001GB001454, 2002.

Iida, T., Saitoh, S. and Miyamura, T.: Temporal and spatial variability of coccolithophore blooms in the eastern Bering Sea, 1998-2001, Prog. Oceanogr., 55, 165–175, 2002.

IPCC: Climate change 2013 the physical science basis: Working Group I contribution to the fifth assessment report of the intergovernmental panel on climate change, Cambridge University Press., 2013.

Johannessen, O. M. and Miles, M. W.: Critical vulnerabilities of marine and sea ice-based ecosystems in the high Arctic, Reg. Environ. Chang., 11(SUPPL. 1), 239–248, doi:10.1007/s10113-010-0186-5, 2011.

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.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G. A., Knutti, R., Furrer, R., Tebaldi, C., Cermak, J. and Meehl, G. A.: Challenges in Combining Projections from Multiple Climate Models, J. Clim., 23(10), 2739–2758, doi:10.1175/2009JCLI3361.1, ~~2010a~~2010.

~~Knutti, R., Abramowitz, G., Collins, M., Eyring, V., Gleckler, P. J., Hewitson, B. and Mearns, L.: Good practice guidance paper on assessing and combining multi model climate projections. [online] Available from: http://www.ipcc-wg2.org/meetings/EMs/IPCC_EM_MultiModelEvaluation_MeetingReport.pdf#page=11 (Accessed 10 April 2019b), 2010.~~

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. E., Pozdnyakov, D. V. and Johannessen, O. M.: Satellite evidence for enhancement of the column mixing ratio of atmospheric CO2 over E. Huxleyi blooms, Trans. Karelian Res. Cent. Russ. Acad. Sci., 9, 125–135, 2019.

Kondrik, D. V., Pozdnyakov, D. V. and Johannessen, O. M.: Satellite Evidence that E. huxleyi Phytoplankton Blooms Weaken Marine Carbon Sinks, Geophys. Res. Lett., 45(2), 846–854, doi:10.1002/2017GL076240, 2018.

Krumhardt, K. M., Lovenduski, N. S., Iglesias-Rodriguez, M. D. and Kleypas, J. A.: Coccolithophore growth and calcification in a changing ocean, Prog. Oceanogr., 159, 276–295, doi:10.1016/J.POCEAN.2017.10.007, 2017.

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.

Kwok, R.: Arctic sea ice thickness, volume, and multiyear ice coverage: Losses and coupled variability (1958-2018), Environ. Res. Lett., 13(10), doi:10.1088/1748-9326/aae3ec, 2018.

Lavender, S. J., Raitsos, D. E. and Pradhan, Y.: Variations in the Phytoplankton of the North-Eastern Atlantic Ocean: From the Irish Sea to the Bay of Biscay, in Remote Sensing of the European Seas, pp. 67–78, Springer Netherlands, Dordrecht., 2008.

Lavoie, D., Lambert, N. and Van der Baaren, A.: Projections of future physical and biogeochemical conditions in the Northwest Atlantic from CMIP5 Global Climate Models., 2013.

Malin, G. and Steinke, M.: Coccolithophore-derived production of dimethyl sulphide, in Coccolithophores, pp. 127–164., 2004.

Malin, G., Turner, S., Liss, P., Holligan, P. and Harbour, D.: Dimethylsulphide and dimethylsulphoniopropionate in the Northeast atlantic during the summer coccolithophore bloom, Deep Sea Res. Part I Oceanogr. Res. Pap., 40(7), 1487–1508, doi:10.1016/0967-0637(93)90125-M, 1993.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H. and Tomassini, L.: Tuning the climate of a global model, J. Adv. Model. Earth Syst., 4(3), n/a-n/a, doi:10.1029/2012MS000154, 2012.

McIntyre, A. and Bé, A. W. H.: Modern coccolithophoridae of the atlantic ocean-I. Placoliths and cyrtoliths, Deep. Res. Oceanogr. Abstr., 14(5), doi:10.1016/0011-7471(67)90065-4, 1967.

Moore, T. S., Dowell, M. D. and Franz, B. A.: Detection of coccolithophore blooms in ocean color satellite imagery: A generalized approach for use with multiple sensors, Remote Sens. Environ., 117, 249–263, doi:10.1016/j.rse.2011.10.001, 2012.

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.

Morozov, E., Pozdnyakov, D., Smyth, T., Sychev, V. and Grassl, H.: Space-borne study of seasonal, multi-year, and decadal phytoplankton dynamics in the Bay of Biscay, Int. J. Remote Sens., 34(4), 1297–1331, doi:10.1080/01431161.2012.718462, 2013.

Morozov, E., Kondrik, D., SS, C. and Pozdnyakov, D. V.: Atmospheric columnar CO2 enhancement over e. huxleyi blooms:

case studies in the North Atlantic and Arctic waters, Limnol. Oceanogr., 3, 28–33, 2019.

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.

Okada, H. and McIntyre, A.: Seasonal distribution of modern coccolithophores in the western North Atlantic Ocean, Mar. Biol., 54(4), 319–328, doi:10.1007/BF00395438, 1979.

Otero, N., Sillmann, J. and Butler, T.: Assessment of an extended version of the Jenkinson–Collison classification on CMIP5 models over Europe, Clim. Dyn., 50(5–6), 1559–1579, doi:10.1007/s00382-017-3705-y, 2018.

Overland, J. E. and Wang, M.: Large-scale atmospheric circulation changes are associated with the recent loss of Arctic sea ice, Tellus A Dyn. Meteorol. Oceanogr., 62(1), 1–9, doi:10.1111/j.1600-0870.2009.00421.x, 2010.

Overland, J. E., Wang, M., Bond, N. A., Walsh, J. E., Kattsov, V. M. and Chapman, W. L.: Considerations in the Selection of Global Climate Models for Regional Climate Projections: The Arctic as a Case Study, J. Clim., 24(6), 1583–1597, doi:10.1175/2010JCLI3462.1, 2011.

Paasche, E.: A review of the coccolithophorid emiliania huxleyi (prymnesiophyceae), with particular reference to growth, coccolith formation, and calcification-photosynthesis interactions, Phycologia, 40(6), 503–529, doi:10.2216/i0031-8884-40-6-503.1, 2001.

Perruche, C.: PRODUCT USER MANUAL For the Global Ocean Biogeochemistry Hindcast GLOBAL_REANALYSIS_BIO_001_029 Issue: 1.0., 2018.

Pierce, D. W., Barnett, T. P., Santer, B. D. and Gleckler, P. J.: Selecting global climate models for regional climate change studies, Proc. Natl. Acad. Sci., 106(21), 8441–8446~~ [online] Available from: https://www.pnas.org/content/106/21/8441.short (Accessed 10 April 2019),~~, 2009.

Pozdnyakov, D., Kondrik, D., Kazakov, E. and S, C.: Environmental conditions favoring coccolithophore blooms in subarctic and arctic seas: a 20-year satellite and multi-dimensional statistical study, in SPIE: Remote Sensing of the Ocean., 2019.

Raitsos, D. E., Lavender, S. J., Pradhan, Y., Tyrrell, T., Reid, P. C. and Edwards, M.: Coccolithophore bloom size variation in response to the regional environment of the subarctic North Atlantic, Limnol. Oceanogr., 51(5), 2122–2130, doi:10.4319/lo.2006.51.5.2122, 2006.

Reichler, T. and Kim, J.: How Well Do Coupled Models Simulate Today's Climate?, Bull. Am. Meteorol. Soc., 89(3), 303–312, doi:10.1175/BAMS-89-3-303, 2008.

Rickard, G. J., Behrens, E. and Chiswell, S. M.: CMIP5 earth system models with biogeochemistry: An assessment for the southwest Pacific Ocean, J. Geophys. Res. Ocean., doi:10.1002/2016JC011736, 2016.

Rivero-Calle, S., Gnanadesikan, A., Castillo, C. E. Del, Balch, W. M. and Guikema, S. D.: Multidecadal increase in North Atlantic coccolithophores and the potential role of rising CO2, Science (80-.)., 350(6267), 2015.

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.

~~Sarr, A. and Sarr, A.: Multi-Scale Characteristics of Precipitation and Temperature over West Africa Using SMHI-RCA Driven by GCMs under RCP8.5, Am. J. Clim. Chang., 06(03), 455–486, doi:10.4236/ajcc.2017.63024, 2017.~~

Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J. C., Hannay, C., Molod, A., Neale, R. B. and Saha, S.: Practice and philosophy of climate model tuning across six US modeling centers, Geosci. Model Dev., 10(9), 3207–3223, doi:10.5194/gmd-10-3207-2017, 2017.

Séférian, R., Bopp, L., Gehlen, M., Orr, J. C., Ethé, C., Cadule, P., Aumont, O., Salas y Mélia, D., Voldoire, A. and Madec, G.: Skill assessment of three earth system models with common marine biogeochemistry, Clim. Dyn., 40(9–10), 2549–2573, doi:10.1007/s00382-012-1362-8, 2013.

Shutler, J. D., Land, P. E., Brown, C. W., Findlay, H. S., Donlon, C. J., Medland, M., Snooke, R. and Blackford, J. C.: Climate of the Past Geoscientific Instrumentation Methods and Data Systems Coccolithophore surface distributions in the North Atlantic and their modulation of the air-sea flux of CO 2 from 10 years of satellite Earth observation data, Biogeosciences, 10, 2699–2709, doi:10.5194/bg-10-2699-2013, 2013.

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.

Taylor, K. E~~.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res. Atmos., 106(D7), 7183–7192, doi:10.1029/2000JD900719, 2001.~~

~~Taylor, K. E., Stouffer, R. J., Meehl, G. A., Taylor, K. E.~~, Stouffer, R. J. and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bull. Am. Meteorol. Soc., 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Tyrrell, T. and Merico, A.: Emiliania huxleyi: bloom observations and the conditions that induce them, Coccolithophores, 75–97, 2004.

Vihma, T., Pirazzini, R., Fer, I., Renfrew, I. A., Sedlar, J., Tjernström, M., Lüpkes, C., Nygård, T., Notz, D., Weiss, J., Marsan, D., Cheng, B., Birnbaum, G., Gerland, S., Chechin, D. and Gascard, J. C.: Advances in understanding and parameterization of small-scale physical processes in the marine Arctic climate system: a review, Atmos. Chem. Phys, 14, 9403–9450, doi:10.5194/acp-14-9403-2014, 2014.

Wang, S., Maltrud, M. E., Burrows, S. M., Elliott, S. M. and Cameron-Smith, P.: Impacts of Shifts in Phytoplankton Community on Clouds and Climate via the Sulfur Cycle, Global Biogeochem. Cycles, 32(6), 1005–1026, doi:10.1029/2017GB005862, 2018a.

Wang, S., Maltrud, M., Elliott, S., Cameron-Smith, P. and Jonko, A.: Influence of dimethyl sulfide on the carbon cycle and biological production, Biogeochemistry, 138(1), 49–68, doi:10.1007/s10533-018-0430-5, 2018b.

Winter, A.: Biogeography of living coccolithophores in ocean waters, Coccolithophores, 1994.

Winter, A., Henderiks, J~~. and , B~~., Beaufort, L., Rickaby, R. E. M. and Brown, C. W.: Poleward expansion of the coccolithophore Emiliania huxleyi, J. ~~Plankt~~Plankton Res., 36(2), 316–325~~ [online] Available from: https://academic.oup.com~~, doi:10.1093/plankt/~~article-abstract/36/2/316/1500265 (Accessed 3 July 2019),~~fbt110, 2013~~.~~.

Zondervan, I.: The effects of light, macronutrients, trace metals and CO 2 on the production of calcium carbonate and organic carbon in coccolithophores-A review, Deep. Res. II, 54, 521–537, doi:10.1016/j.dsr2.2006.12.004, 2007.

Table 1. CMIP5 models used for simulation of selected ~~parameters~~variables: SST – sea surface temperature in ~~K~~ºC, WS – ~~near-surface~~10 m wind speed in m s^{-1}, SDSR – surface downwelling shortwave solar radiation in W m^{-2}, ~~SSS~~SS_{30m} – sea~~ surface~~ salinity (averaged over top 30 m) in ~~psu~~PSU, OCS – surface ocean current speed in m s^{-1}, concentration of nutrients (NO_3, PO_4, and SI) in mol m^{-3}, dissolved CO_2 partial pressure (pCO_2) in Pa, and pH (models available for respective variable are marked as “+”)

Model	**ID**	**Modelling Center** (acronym, full name, and country)	**Resolution** (ºlon x ºlat)	**SST**	**WS**	**SDSR**	**SS~~S~~$_{30m}$**	**OCS**	NO_3	PO_4	SI	pCO_2	pH
ACCESS1.0	1	CSIRO-BOM, Commonwealth Scientific and Industrial Research Organisation, Australia and Bureau of Meteorology, Australia	1.25 x 1.875	+	+	+	+	+					
ACCESS1.3	2			+	+	+	+	+					
CanESM2	3	CCCma, Canadian Centre for Climate Modelling and Analysis, Canada	2.7906 x 2.8125	+	+		+	+	+			+	+
CMCC-CM	4	CMCC, Centro euro-Mediterraneo sui Cambiamenti Climatici, Italy	0.7484 x 0.75	+	+	+	+	+					
CMCC-CMS	5		3.7111 x 3.75	+	+	+	+	+	+	+		+	+
CNRM-CM5	6	CNRM-CERFACS, Centre National de Recherches Meteorologiques, France and Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique, France	1.4008 x 1.40625	+	+	+	+	+	+	+	+		
CSIRO-Mk3.6.0	7	CSIRO-QCCCE, Commonwealth Scientific and Industrial Research Organization, Australia and Queensland Climate Change Centre of Excellence, Australia	1.8653 x 1.875		+	+	+	+					
EC-EARTH	8	EC-EARTH, EC-EARTH consortium, Europe	1.1215 x 1.125	+									
GFDL-CM3	9	NOAA GFDL, National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, USA	2 x 2.5	+	+	+	+	+					
GFDL-ESM2G	10			+	+	+	+	+	+	+	+	+	+
GFDL-ESM2M	11			+	+	+	+	+	+	+	+	+	+
GISS-E2-H	12	NASA GISS, National Aeronautics and Space Administration, Goddard Institute for Space Studies, USA	2 x 2.5	+	+	+	+	+					
GISS-E2-H-CC	13			+	+	+	+	+	+		+	+	
GISS-E2-R	14			+	+	+	+	+					
GISS-E2-R-CC	15			+	+	+	+	+	+		+	+	
HadCM3	16	MOHC INPE, Met Office Hadley Centre, UK and	2.5 x 3.75		+								

Добавленные ячейки
Добавленные ячейки
Добавленные ячейки
Добавленные ячейки
Добавленные ячейки

HadGEM2-AO	17	Instituto Nacional de Pesquisas Espaciais, Brasil	1.25 x 1.875	+	+	+	+	+					
HadGEM2-CC	18			+	+	+	+	+	+		+	+	+
HadGEM2-ES	19			+	+	+	+	+	+		+	+	+
IPSL-CM5A-LR	20	IPSL, Institut Pierre-Simon Laplace, France	1.8947 x 3.75	+	+	+	+	+	+	+	+		+
IPSL-CM5A-MR	21			+	+	+	+	+	+	+	+		+
IPSL-CM5B-LR	22			+	+	+	+	+	+	+	+		+
MIROC5	23	MIROC, Atmosphere and Ocean Research Institute, the University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	1.4008 x 1.40625	+	+	+	+						
MIROC4h	24		0.5616 x 0.5625		+								
MIROC-ESM	25	MIROC, Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, the University of Tokyo, and National Institute for Environmental Studies, Japan	2.7906 x 2.8125	+	+	+	+						+
MIROC-ESM-CHEM	26			+	+	+	+						+
MPI-ESM-LR	27	MPI-M, Max Planck Institute for Meteorology, Germany	1.8653 x 1.875	+	+	+	+	+	+	+	+	+	+
MPI-ESM-MR	28			+	+	+	+	+	+	+	+	+	+
MRI-CGCM3	29	MRI, Meteorological Research Institute, Japan	1.12148 x 1.125	+	+	+	+	+					
MRI-ESM1	30				+				+	+		+	+
NorESM1-M	31	NCC, Norwegian Climate Centre, Norway	1.8947 x 2.5	+		+	+						
NorESM1-ME	32			+		+	+	+	+	+	+	+	+
INM-CM4	33	INM, Russian Academy of Sciences Marchuk Institute of Numerical Mathematics, Russia	1.5 x 2		+	+							+
FGOALS-g2	34	LASG-CESS, Institute of Atmospheric Physics, Chinese Academy of Sciences; and Tsinghua University, China	2.7906 x 2.8125					+					
Total number of available CMIP5 models				**28**	**30**	**28**	**28**	**25**	**16**	**11**	**13**	**15**	**13**

Table 2. Results of the CMIP5 model performance for SST in the Barents Sea.

~~(~~ Numbers in brackets indicate the models' scores.

~~(~~RMSD is the root-mean-square deviation, ~~K~~°C; r is the correlation coefficient between models and reanalysis; ~~RSR~~CPI is ~~the RMSD-observations standard deviation ratio~~climate prediction index; $|SD_{dif}|$ is the modulus of the standard deviation difference (model minus reanalysis), ~~K~~°C; $|Tr_m|$ is the modulus of spatial trend ~~median~~mean difference (the model minus reanalysis), ~~K~~°C yr^{-1}; $|Tr_a|$ is the modulus of spread of spatial ~~trend amplitude~~trends difference (the model minus reanalysis), ~~K~~°C yr^{-1}; $|Br_m|$ is the modulus of spatial bias ~~median~~mean difference (the model minus reanalysis), ~~K~~°C; $|Br_a|$ is the modulus of spread of spatial biases ~~amplitude~~ difference (the model minus reanalysis), ~~K~~°C).

Model acronym	ID	Seasonal cycle (averaged over the territory)				Interannual variability (averaged over the territory)				Spatial trends (Tr) and biases (Br)				Total score
		RMSD	r	~~RSR~~CPI	$\|SD_{dif}\|$	RMSD	r	~~RSR~~CPI	$\|SD_{dif}\|$	$\|Tr_m\|$	$\|Tr_a\|$	$\|Br_m\|$	$\|Br_a\|$	
ACCESS1-0	1	0,26(3)	0,99(2)	0,13(3)	0,08(3)	1,17(3)	0,68(3)	0,81(3)	0,02(3)	0,06(2)	0,01(3)	0,07(3)	14,7(2)	**33**
ACCESS1-3	2	0,37(3)	0,99(3)	0,19(3)	0,03(3)	1,02(3)	0,75(3)	0,71(3)	0,19(3)	0,01(3)	0,01(3)	0,57(3)	16,1(1)	**34**
CanESM2	3	1,76(2)	0,98(2)	0,88(2)	0,28(0)	2,21(2)	0,64(3)	1,54(2)	1,12(3)	0,10(1)	0,04(3)	0,85(3)	17,2(1)	24
CMCC-CM	4	5,15(0)	0,96(1)	2,58(0)	1,73(1)	7,06(0)	0,28(3)	4,90(0)	0,63(0)	0,06(2)	0,18(0)	6,64(0)	13,1(2)	9
CMCC-CMS	5	4,40(0)	0,97(2)	2,20(0)	1,34(1)	5,94(0)	0,56(3)	4,12(0)	0,59(0)	0,01(3)	0,02(3)	5,58(0)	14,1(2)	14
CNRM-CM5	6	0,64(3)	0,99(2)	0,32(3)	0,55(1)	1,59(3)	0,73(3)	1,10(3)	0,81(2)	0,08(2)	0,00(3)	0,49(3)	16,4(1)	29
EC-EARTH	7	0,41(3)	0,99(2)	0,21(3)	0,13(2)	1,43(3)	0,64(3)	0,99(3)	0,38(3)	0,13(1)	0,12(1)	0,14(3)	18,1(0)	27
GFDL-CM3	8	1,34(3)	0,99(3)	0,67(3)	0,20(3)	1,71(3)	0,80(3)	1,19(3)	0,22(3)	0,00(3)	0,09(1)	1,39(3)	11,1(3)	**34**
GFDL-ESM2G	9	3,23(1)	0,98(2)	1,62(1)	0,27(2)	3,72(1)	0,69(3)	2,58(1)	0,29(3)	0,04(3)	0,04(3)	3,46(1)	13,9(2)	23
GFDL-ESM2M	10	2,60(2)	0,99(2)	1,30(2)	0,61(3)	3,42(2)	0,68(3)	2,37(2)	0,25(2)	0,01(3)	0,08(2)	3,10(2)	15,7(1)	26
GISS-E2-H	11	3,39(1)	0,97(3)	1,70(1)	0,41(3)	4,09(1)	0,83(3)	2,84(1)	0,18(3)	0,05(2)	0,04(3)	3,86(1)	11,4(3)	25
GISS-E2-H-CC	12	3,68(1)	0,96(2)	1,84(1)	0,56(3)	4,62(1)	0,72(3)	3,20(1)	0,12(2)	0,03(3)	0,02(3)	4,36(1)	10,8(3)	24
GISS-E2-R	13	3,34(1)	0,96(2)	1,67(1)	0,04(1)	3,83(1)	0,72(3)	2,66(1)	0,84(3)	0,05(2)	0,07(2)	3,34(1)	15,1(2)	20
GISS-E2-R-CC	14	3,38(1)	0,96(2)	1,69(1)	0,07(1)	3,78(1)	0,75(3)	2,62(1)	0,83(3)	0,03(3)	0,05(2)	3,29(2)	13,6(2)	22
HadGEM2-AO	15	1,28(3)	0,99(2)	0,64(3)	0,01(3)	1,51(3)	0,73(3)	1,05(3)	0,13(3)	0,02(3)	0,05(2)	1,33(3)	19,8(0)	31
HadGEM2-CC	16	1,70(2)	0,99(2)	0,85(2)	0,16(2)	2,34(2)	0,62(3)	1,62(2)	0,35(3)	0,05(2)	0,05(2)	1,66(3)	19,1(0)	25
HadGEM2-ES	17	0,30(3)	0,99(3)	0,15(3)	0,08(3)	0,98(3)	0,77(3)	0,68(3)	0,00(3)	0,05(2)	0,04(3)	0,09(3)	17,5(1)	**33**
IPSL-CM5A-LR	18	3,66(1)	0,98(2)	1,83(1)	0,31(3)	4,59(1)	0,70(3)	3,19(1)	0,18(3)	0,01(3)	0,03(3)	4,32(1)	18,4(0)	22
IPSL-CM5A-MR	19	2,22(2)	0,99(2)	1,11(2)	0,67(1)	2,57(2)	0,73(3)	1,78(2)	0,80(2)	0,06(2)	0,05(2)	1,91(2)	16,0(1)	23
IPSL-CM5B-LR	20	5,03(0)	0,96(1)	2,52(0)	1,71(1)	6,90(0)	0,36(3)	4,79(0)	0,69(0)	0,00(3)	0,03(3)	6,51(0)	17,6(0)	11
MIROC-ESM	21	1,40(3)	0,99(3)	0,70(3)	0,04(3)	1,63(3)	0,82(3)	1,13(3)	0,06(3)	0,01(3)	0,08(2)	1,51(3)	11,8(3)	**35**
MIROC-ESM-CHEM	22	0,97(3)	0,99(3)	0,49(3)	0,05(3)	1,34(3)	0,82(3)	0,93(3)	0,13(3)	0,07(2)	0,05(3)	1,08(3)	15,1(2)	**34**
MIROC5	23	2,42(0)	0,98(2)	1,21(0)	0,51(1)	5,69(2)	0,51(3)	3,95(2)	0,64(2)	0,18(0)	0,08(2)	5,14(0)	19,8(0)	14
MPI-ESM-LR	24	1,27(3)	0,99(3)	0,63(3)	0,04(3)	1,54(3)	0,81(3)	1,07(3)	0,21(3)	0,02(3)	0,04(3)	1,33(3)	16,3(1)	**34**
MPI-ESM-MR	25	0,91(3)	0,99(2)	0,45(3)	0,05(3)	1,47(3)	0,71(3)	1,02(3)	0,11(3)	0,05(2)	0,04(3)	0,96(3)	17,2(1)	**32**
MRI-CGCM3	26	2,88(2)	0,99(3)	1,44(2)	0,08(2)	2,54(1)	0,82(3)	1,77(1)	0,34(3)	0,00(3)	0,07(2)	2,30(2)	11,9(3)	27
NorESM1-M	27	1,53(2)	0,99(2)	0,77(2)	0,76(2)	2,56(2)	0,64(3)	1,78(2)	0,31(2)	0,05(2)	0,07(2)	2,33(2)	13,7(2)	25
NorESM1-ME	28	1,72(2)	0,99(2)	0,86(2)	0,78(2)	2,79(2)	0,57(3)	1,94(2)	0,39(2)	0,02(3)	0,02(3)	2,58(2)	15,0(2)	27

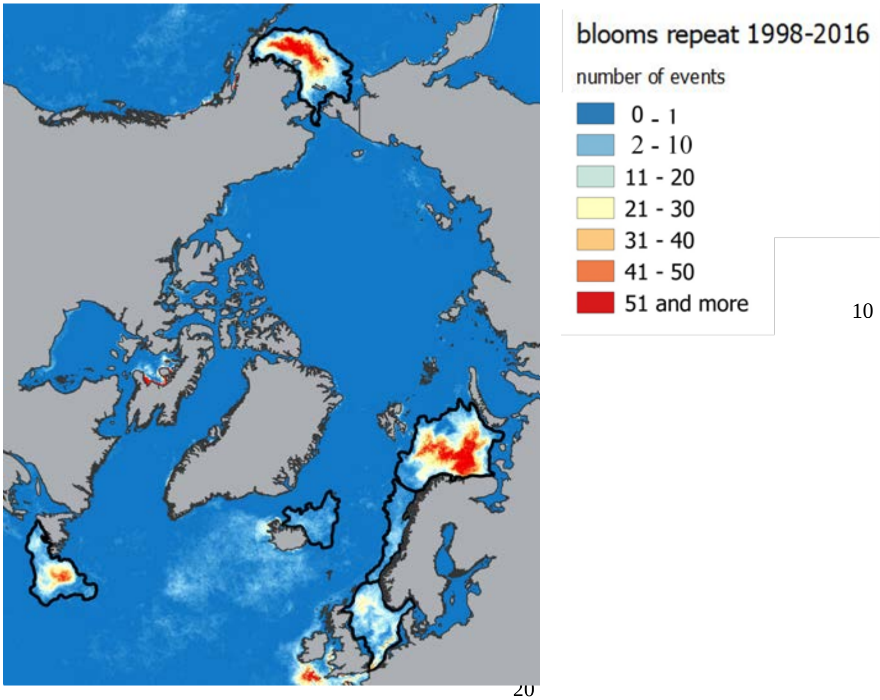


Figure ~~1:~~1. Spatial distribution of *E. huxleyi* blooms occurrence based on the Ocean Colour Climate Change Initiative dataset version 3.0 (Kazakov et al., 2018) **for the Barents, Bering, Labrador, Greenland, North, and Norwegian seas. Black lines confine the territories where blooms occurred more than one 8-day period and show target sea areas.**

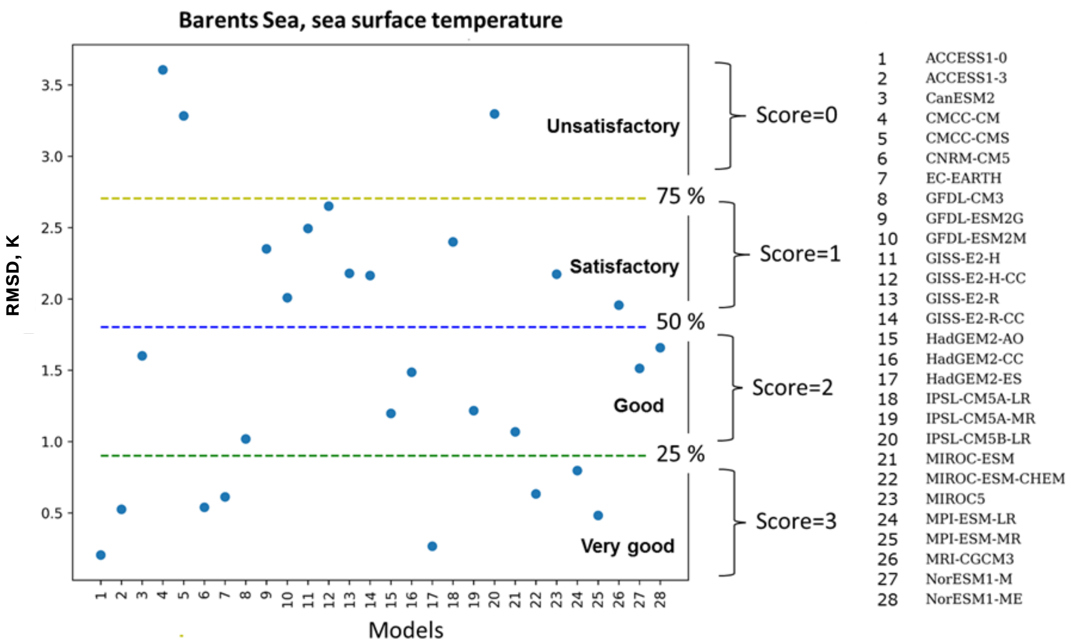
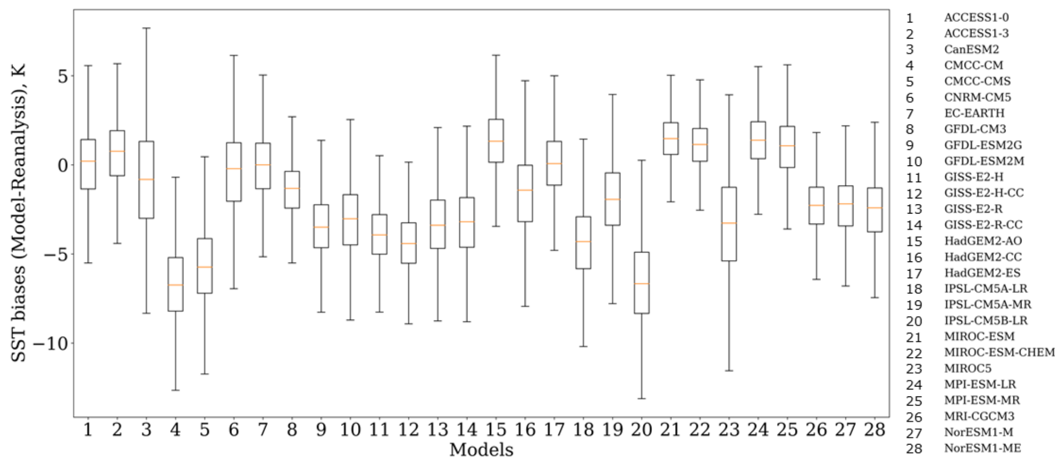


Figure 2~~:~~. A schematic representation of the percentile ~~score-based model~~ ranking ~~method (Division~~approach: division of RMSD values distribution of 28 models into four groups that are limited by 25th, 50th~~-~~, and 75th percentiles and the relative assignment of scores from 3 to 0 to each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



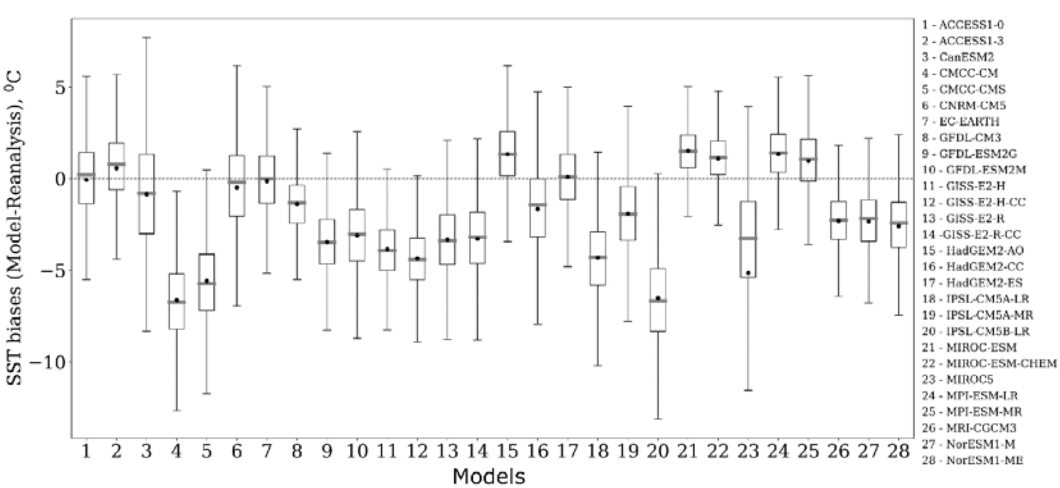


Figure 3~~:~~. Box plots of the spatial variability of SST biases~~,~~ (°C), which are calculated as the difference between the model and reanalysis data in the Barents Sea for *E. huxleyi* bloom season over the ~~vegetation season and the time~~ period from 1979~~-~~ to 2005. Each box spreads from the lower quartile Q1 to the upper quartile Q3 of biases, the ~~orange~~gray lines represent the medians. The

dots show mean values. The lower "whiskers" are represented as Q1-1.5 ~~Standard~~standard deviation and the upper "whiskers" are represented as Q3+1.5 ~~Standard~~standard deviation.

ID	CMIP5 models	Barents Sea					Bering Sea					Greenland Sea					Labrador Sea					North Sea					Norwegian Sea				
		OCS	SSS	SST	WS	SDSR	OCS	SSS	SST	WS	SDSR	OCS	SSS	SST	WS	SDSR	OCS	SSS	SST	WS	SDSR	OCS	SSS	SST	WS	SDSR	OCS	SSS	SST	WS	SDSR
1	ACCESS1-3	23	34	33	28	27	30	23	17	27	24	22	31	26	29	31	27	29	18	30	13	29	30	32	23	27	23	32	36	24	25
2	ACCESS1-0	26	33	34	28	27	27	24	26	26	29	18	31	27	27	33	27	26	22	26	20	31	30	30	23	25	28	31	35	25	24
3	CanESM2	25	26	24	29		27	24	26	14		19	15	30	19		16	29	33	9		26	22	34	18		29	22	35	21	
4	CMCC-CM	7	26	9	23	21	29	22	25	27	14	21	28	16	27	21	23	30	18	20	14	27	23	25	24	8	13	33	22	30	8
5	CMCC-CMS	16	22	14	24	23	29	23	25	28	15	25	33	32	22	16	25	35	15	21	15	24	19	30	25	13	24	31	36	28	14
6	CNRM-CM5	18	31	29	28	13	31	25	26	30	26	21	32	23	26	19	29	30	30	26	29	23	31	30	28	29	25	34	31	27	25
7	CSIRO-Mk3-6-0	20	23		19	21	21	26		31	14	20	35		26	10	21	27		30	17	23	25		24	16	19	33		15	13
8	EC-EARTH			27					27					35					28					30					36		
9	FGOALS-g2	17					4					8					24					11					12				
10	GFDL-CM3	20	32	34	27	23	32	20	32	32	26	19	30	32	21	28	27	25	25	28	28	23	19	31	29	22	26	33	36	27	24
11	GFDL-ESM2G	21	30	23	26	26	29	25	20	30	14	24	27	22	30	24	20	27	29	27	21	22	27	32	27	26	26	33	30	26	25
12	GFDL-ESM2M	15	33	26	27	25	32	20	24	29	20	23	33	23	23	18	27	32	24	27	27	24	18	29	28	28	25	33	33	23	27
13	GISS-E2-H	10	29	25	29	12	26	19	29	30	28	16	32	28	28	25	15	15	14	19	28	20	30	32	28	31	17	33	36	19	34
14	GISS-E2-H-CC	14	24	24	30	12	25	21	32	32	26	13	24	25	28	17	18	23	23	18	19	19	31	32	26	29	20	27	35	26	32
15	GISS-E2-R	19	8	20	26	12	28	25	25	32	29	25	29	28	30	22	22	26	27	26	29	23	28	31	29	30	23	32	33	27	34
16	GISS-E2-R-CC	20	9	22	27	11	29	27	28	32	30	24	28	26	30	25	22	22	30	28	28	22	25	30	30	29	24	35	29	27	29
17	HadCM3				16					28					25					27					27					19	
18	HadGEM2-AO	26	32	31	30	29	30	28	29	32	30	17	23	27	31	33	19	11	30	28	13	28	30	35	20	28	26	31	34	21	31
19	HadGEM2-CC	22	32	25	30	25	29	26	32	30	29	20	19	31	29	33	22	20	30	30	16	29	31	33	28	31	27	32	35	25	32
20	HadGEM2-ES	21	33	33	27	30	25	24	28	30	27	17	25	28	28	33	25	17	26	29	13	28	26	32	29	30	28	30	33	23	32
21	INMCM4				30	32				26	32				16	33				18	30				23	31				24	28
22	IPSL-CM5A-LR	18	12	22	23	29	30	25	34	27	26	18	29	25	19	25	19	31	23	24	26	22	12	21	13	20	17	29	28	17	25
23	IPSL-CM5A-MR	20	18	23	24	29	33	22	32	31	24	17	28	32	27	27	21	27	25	24	23	25	7	26	23	28	25	31	31	18	27
24	IPSL-CM5B-LR	11	9	11	15	27	33	27	22	31	26	15	11	12	18	13	14	21	31	23	19	21	13	18	14	16	12	13	25	14	22
25	MIROC4h				32					18					28					21					27					28	
26	MIROC5		31	14	28	22		14	16	24	31		32	33	28	32		31	19	21	27		25	20	28	25		24	17	25	32
27	MIROC-ESM		31	35	15	26		13	31	33	20		29	22	26	20		30	29	26	9		26	34	16	13		30	34	16	25
28	MIROC-ESM-CHEM		30	34	19	23		15	31	31	21		29	20	25	18		34	28	21	10		28	34	15	18		28	33	16	25
29	MPI-ESM-LR	21	31	34	25	21	32	29	24	31	11	12	33	29	21	19	16	22	21	21	10	26	31	33	27	19	13	31	34	28	23
30	MPI-ESM-MR	17	33	32	24	19	31	28	21	29	15	17	31	31	25	18	12	24	28	20	15	23	31	35	25	18	13	25	35	27	23
31	MRI-CGCM3	26	20	27	13	25	28	28	30	10	26	26	13	25	16	19	21	16	26	14	18	20	29	32	12	28	28	20	33	15	33
32	MRI-ESM1				12					9					11					14					8					16	
33	NorESM1-M		33	25		20		17	24		13		30	26		10		23	23		14		30	34		25		31	33		25
34	NorESM1-ME	23	33	27		23	28	23	23		15	23	31	20		14	27	21	28		10	25	30	31		28	24	35	32		23
	total selected models	7	7	8	7	8	7	8	8	11	8	7	11	8	11	8	8	8	8	10	8	8	11	8	9	8	8	9	10	9	8

30 - selected optimal model ensemble

- score < 25% "very good"

- 25% < score < 75% "good" & "satisfactory"

- score > 75% "unsatisfactory"

ID	CMIP5 models	Barents Sea					Bering Sea					Greenland Sea					Labrador Sea					North Sea					Norwegian Sea				
		OCS	SS_{30m}	SST	WS	SDSR	OCS	SS_{30m}	SST	WS	SDSR	OCS	SS_{30m}	SST	WS	SDSR	OCS	SS_{30m}	SST	WS	SDSR	OCS	SS_{30m}	SST	WS	SDSR	OCS	SS_{30m}	SST	WS	SDSR
1	ACCESS1-3	**23**	**34**	**33**	28	**27**	30	23	17	27	24	22	**31**	26	**29**	**31**	**27**	29	18	**30**	13	**29**	**30**	32	23	27	23	32	**36**	24	25
2	ACCESS1-0	**26**	**33**	**34**	28	**27**	27	24	26	26	**29**	18	**31**	27	27	**33**	**27**	26	22	26	20	**31**	**30**	30	23	25	**28**	31	**35**	25	24
3	CanESM2	**25**	26	24	**29**		27	24	26	14		19	15	**30**	19		16	29	**33**	9		**26**	22	**34**	18		**29**	22	**35**	21	
4	CMCC-CM	7	26	9	23	21	29	22	25	27	14	21	28	16	27	21	23	**30**	18	20	14	**27**	23	25	24	8	13	**33**	22	**30**	8
5	CMCC-CMS	16	22	14	24	23	29	23	25	28	15	**25**	**33**	**32**	22	16	**25**	**35**	15	21	15	24	19	30	25	13	24	31	**36**	**28**	14
6	CNRM-CM5	18	31	29	28	13	**31**	25	26	30	26	21	**32**	23	26	19	**29**	**30**	**30**	26	**29**	23	**31**	30	**28**	**29**	25	**34**	31	**27**	25
7	CSIRO-Mk3-6-0	20	23		19	21	21	**26**		**31**	14	20	**35**		26	10	21	27		**30**	17	23	25		24	16	19	**33**		15	13
8	EC-EARTH			27					27					**35**					28					30					**36**		
9	FGOALS-g2	17					4					8					24					11					12				
10	GFDL-CM3	20	32	**34**	27	23	**32**	20	**32**	**32**	26	19	30	**32**	21	**28**	**27**	25	25	**28**	**28**	23	19	31	**29**	22	**26**	**33**	**36**	**27**	24
11	GFDL-ESM2G	21	30	23	26	26	29	25	20	30	14	**24**	27	22	**30**	24	20	27	**29**	**27**	21	22	27	32	27	26	**26**	**33**	30	26	25
12	GFDL-ESM2M	15	**33**	26	27	25	**32**	20	24	29	20	**23**	**33**	23	23	18	**27**	**32**	24	**27**	**27**	24	18	29	**28**	28	25	**33**	33	23	27
13	GISS-E2-H	10	29	25	**29**	12	26	19	29	30	**28**	16	**32**	28	**28**	25	15	15	14	19	**28**	20	**30**	32	**28**	**31**	17	**33**	**36**	19	**34**
14	GISS-E2-H-CC	14	24	24	**30**	12	25	21	**32**	**32**	26	13	24	25	**28**	17	18	23	23	18	19	19	**31**	32	26	**29**	20	27	**35**	26	**32**
15	GISS-E2-R	19	8	20	26	12	28	25	25	**32**	**29**	**25**	29	28	**30**	22	22	26	27	26	**29**	23	28	31	**29**	**30**	23	32	33	**27**	**34**
16	GISS-E2-R-CC	20	9	22	27	11	29	**27**	28	**32**	**30**	**24**	28	26	**30**	25	22	22	**30**	**28**	**28**	22	25	30	**30**	**29**	24	**35**	29	**27**	29
17	HadCM3				16					28					25					**27**					27					19	
18	HadGEM2-AO	**26**	32	31	**30**	**29**	30	**28**	29	**32**	**30**	17	23	27	**31**	**33**	19	11	**30**	**28**	13	**28**	**30**	**35**	20	28	**26**	31	34	21	**31**
19	HadGEM2-CC	**22**	32	25	**30**	25	29	**26**	**32**	30	**29**	20	19	**31**	**29**	**33**	22	20	**30**	**30**	16	**29**	**31**	**33**	**28**	**31**	**27**	32	**35**	25	**32**
20	HadGEM2-ES	21	**33**	**33**	27	**30**	25	24	28	30	27	17	25	28	**28**	**33**	25	17	26	**29**	13	**28**	26	32	**29**	**30**	**28**	30	33	23	**32**
21	INMCM4				**30**	**32**				26	**32**				16	**33**				18	**30**				23	**31**				24	28
22	IPSL-CM5A-LR	18	12	22	23	**29**	30	25	**34**	27	26	18	29	25	19	25	19	**31**	23	24	26	22	12	21	13	20	17	29	28	17	25
23	IPSL-CM5A-MR	20	18	23	24	**29**	**33**	22	**32**	**31**	24	17	28	**32**	27	**27**	21	27	25	24	23	25	7	26	23	28	25	31	31	18	27
24	IPSL-CM5B-LR	11	9	11	15	**27**	**33**	**27**	22	**31**	26	15	11	12	18	13	14	21	**31**	23	19	21	13	18	14	16	12	13	25	14	22
25	MIROC4h				**32**					18					**28**					21					27					**28**	
26	MIROC5		31	14	28	22		14	16	24	**31**		**32**	**33**	28	**32**		**31**	19	21	**27**		25	20	**28**	25		24	17	25	**32**
27	MIROC-ESM		31	**35**	15	26		13	**31**	**33**	20		29	22	26	20		**30**	**29**	26	9		26	**34**	16	13		30	34	16	25
28	MIROC-ESM-CHEM		30	**34**	19	23		15	**31**	**31**	21		29	20	25	18		**34**	28	21	10		28	**34**	15	18		28	33	16	25
29	MPI-ESM-LR	21	31	**34**	25	21	**32**	**29**	24	**31**	11	12	**33**	29	21	19	16	22	21	21	10	**26**	**31**	**33**	27	19	13	31	34	**28**	23
30	MPI-ESM-MR	17	**33**	**32**	24	19	**31**	**28**	21	29	15	17	**31**	**31**	25	18	12	24	28	20	15	23	**31**	**35**	25	18	13	25	**35**	**27**	23
31	MRI-CGCM3	**26**	20	27	13	25	28	**28**	**30**	10	26	**26**	13	25	16	19	21	16	26	14	18	20	29	32	12	28	**28**	20	33	15	**33**
32	MRI-ESM1				12					9					11					14					8					16	
33	NorESM1-M		**33**	25		20		17	24		13		30	26		10		23	23		14		**30**	**34**		25		31	33		25
34	NorESM1-ME	**23**	**33**	27		23	28	23	23		15	**23**	**31**	20		14	**27**	21	28		10	25	**30**	31		28	24	**35**	32		23
	total selected models	7	7	8	7	8	7	8	8	11	8	7	11	8	10	9	7	8	8	10	8	8	11	8	9	8	8	9	10	9	8

30 - selected optimal model ensemble

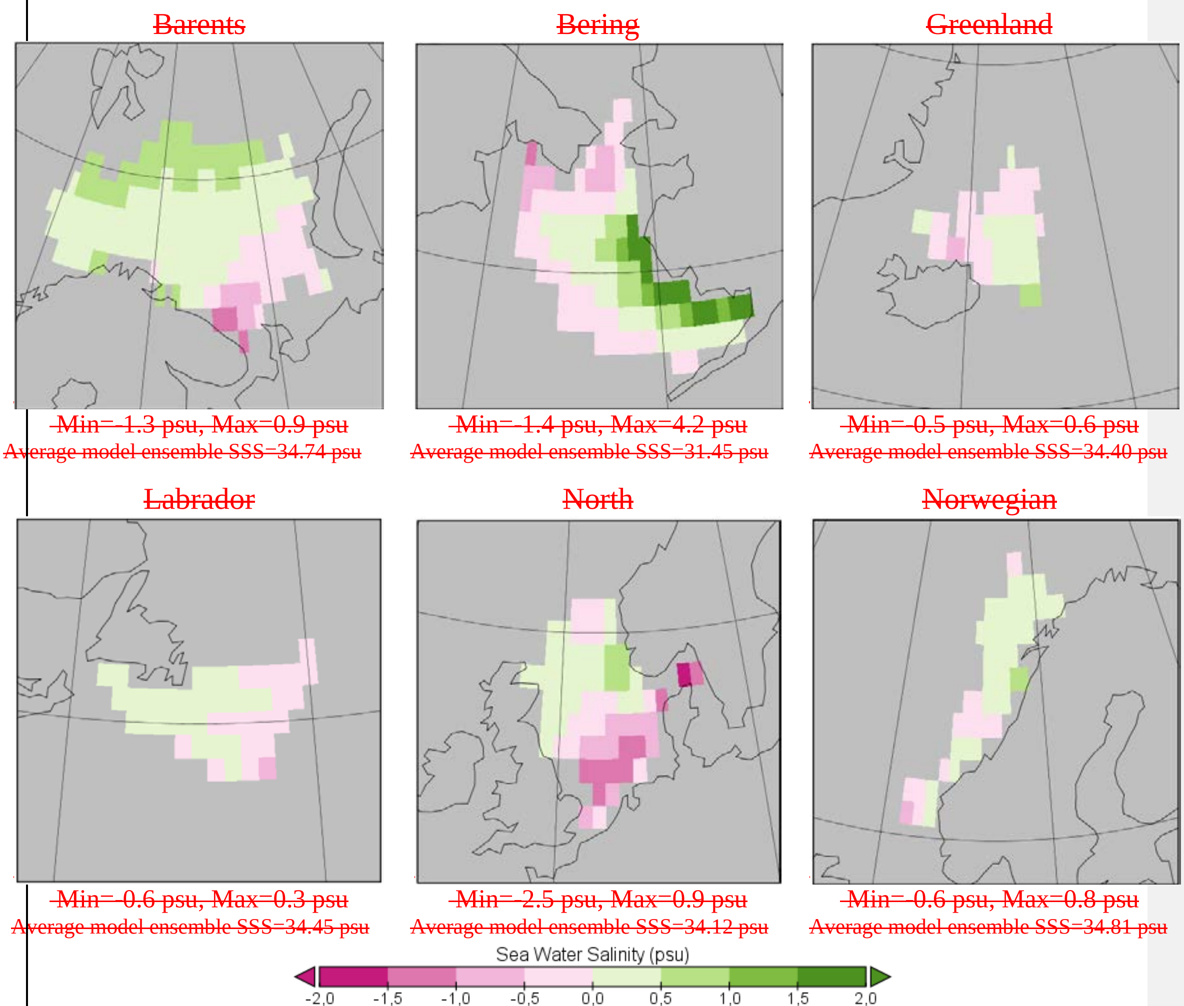
- score < 25% "very good"

- 25% < score < 75% "good" & "satisfactory"

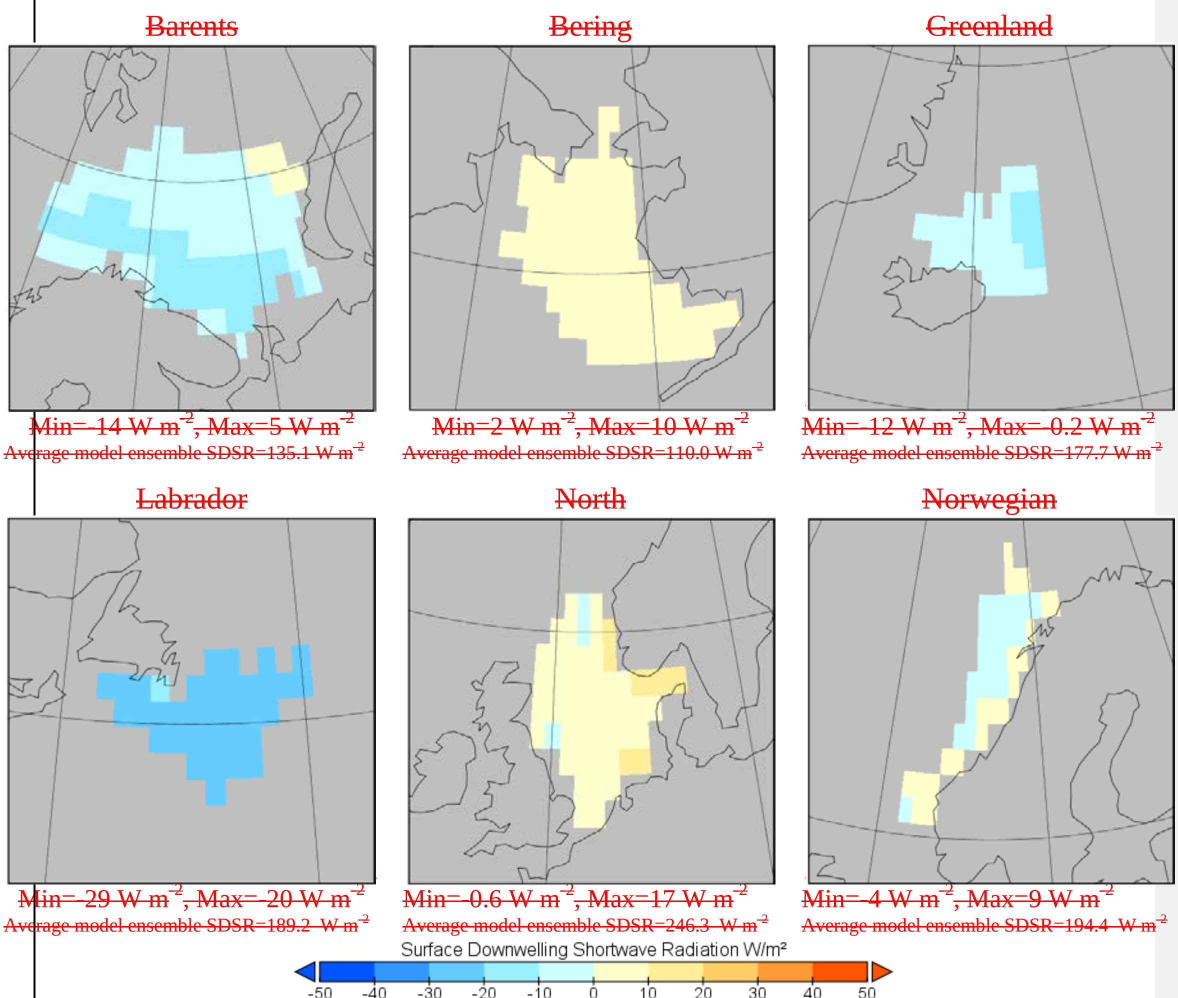
- score > 75% "unsatisfactory"

Figure 4~~:~~. Heat map with the final model scores obtained using the percentile ~~score-based model~~ ranking ~~method~~approach for the ~~five~~5 oceanographic and meteorological variables (sea surface temperature (SST~~, K) and~~), salinity averaged over 0-30 m (~~SSS, psu~~SS_{30m}), surface wind speed at 10 m (WS~~, m s^{-1}~~), ocean surface current speed (OCS~~, m s^{-1}~~), and surface shortwave downwelling solar radiation (SDSR~~, W m^{-2}~~) for the Barents, Bering, Greenland, Labrador, North, and Norwegian seas based on different

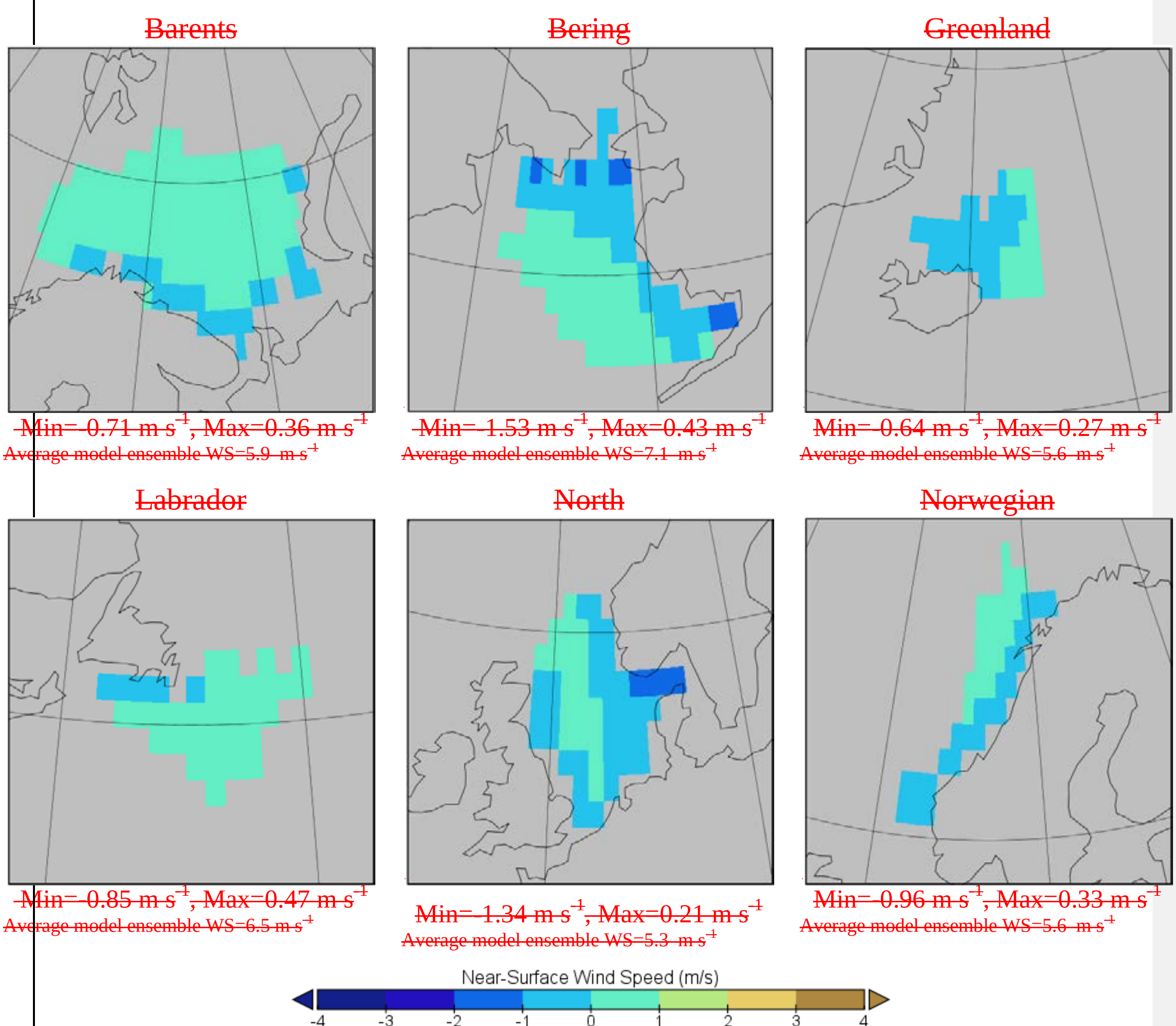
statistical ~~metrics (Figure~~measures (Fig. 2, ~~Table~~Tab. 2). The white areas indicate ~~that the~~a lack of model ~~was not considered due to partial or complete unavailability of hindcasts,~~output for historical and ~~future~~RCP projections (RCP4.5, RCP8.5) in open data~~.~~ sources.



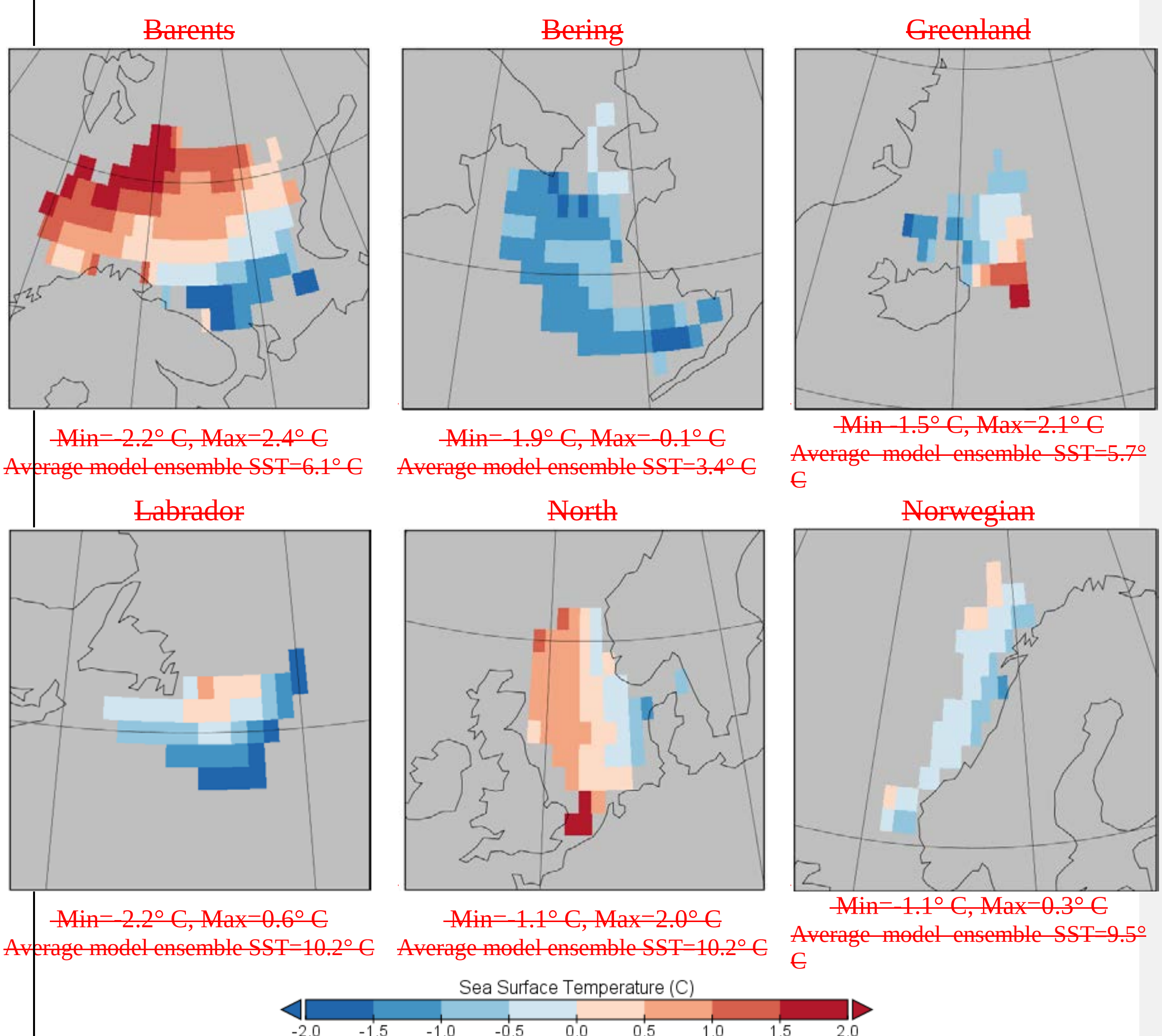
~~**Figure 5a. Spatial distribution of biases in sea surface salinity models and reanalysis in six target seas averaged over the vegetation season and the time period 1993-2005.**~~



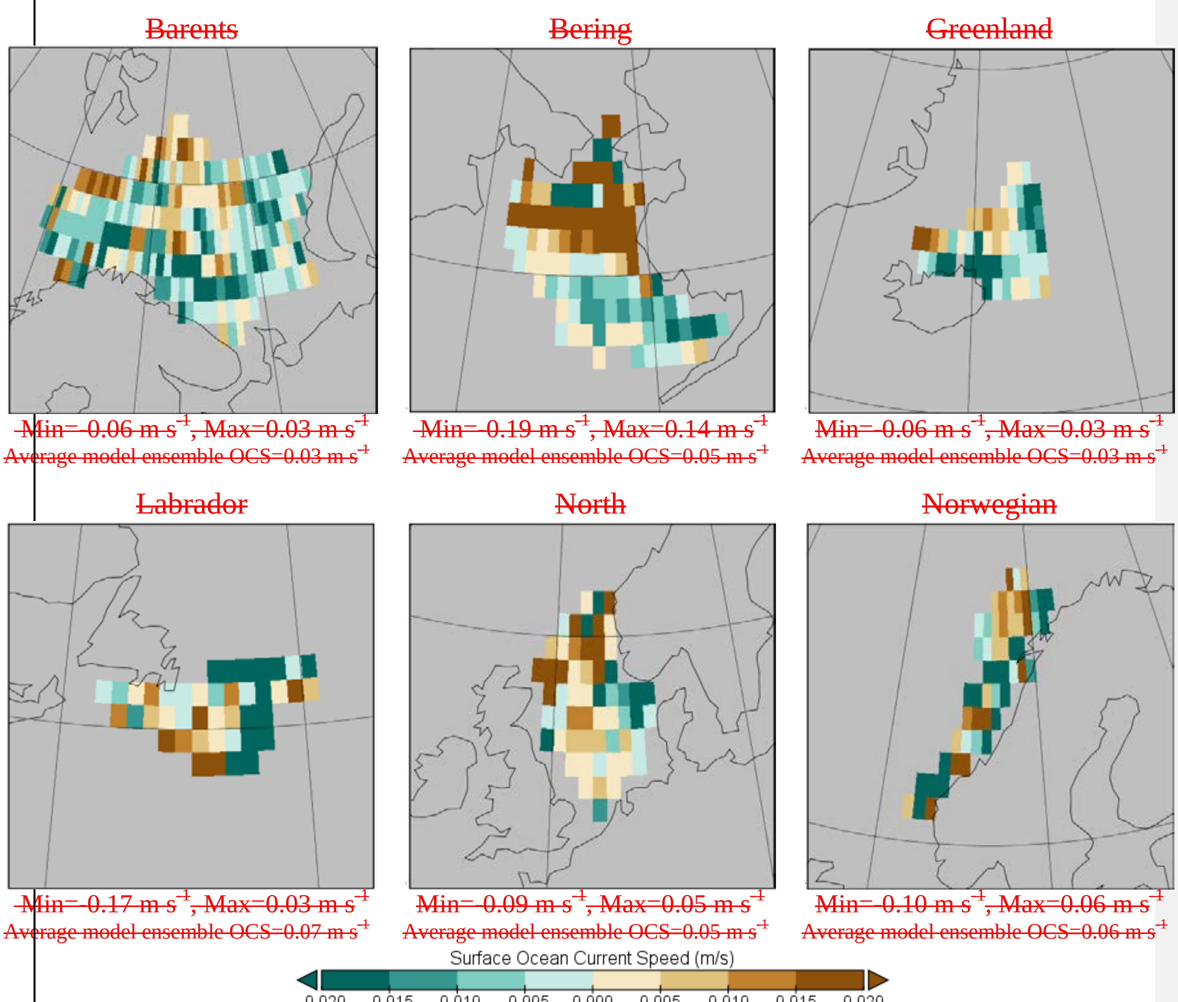
~~**Figure 5b. Spatial distribution of biases in surface downwelling solar radiation between models and reanalysis in six target seas averaged over the vegetation season and the time period 1979-2005.**~~


~~Barents~~
~~Min=-0.71 m s~~$^{-1}$~~, Max=0.36 m s~~$^{-1}$
~~Average model ensemble WS=5.9 m s~~$^{-1}$
~~Bering~~
~~Min=-1.53 m s~~$^{-1}$~~, Max=0.43 m s~~$^{-1}$
~~Average model ensemble WS=7.1 m s~~$^{-1}$
~~Greenland~~
~~Min=-0.64 m s~~$^{-1}$~~, Max=0.27 m s~~$^{-1}$
~~Average model ensemble WS=5.6 m s~~$^{-1}$
~~Labrador~~
~~Min=-0.85 m s~~$^{-1}$~~, Max=0.47 m s~~$^{-1}$
~~Average model ensemble WS=6.5 m s~~$^{-1}$
~~North~~
~~Min=-1.34 m s~~$^{-1}$~~, Max=0.21 m s~~$^{-1}$
~~Average model ensemble WS=5.3 m s~~$^{-1}$
~~Norwegian~~
~~Min=-0.96 m s~~$^{-1}$~~, Max=0.33 m s~~$^{-1}$
~~Average model ensemble WS=5.6 m s~~$^{-1}$
Near-Surface Wind Speed (m/s)
-4
-3
-2
-1
0
1
2
3
4

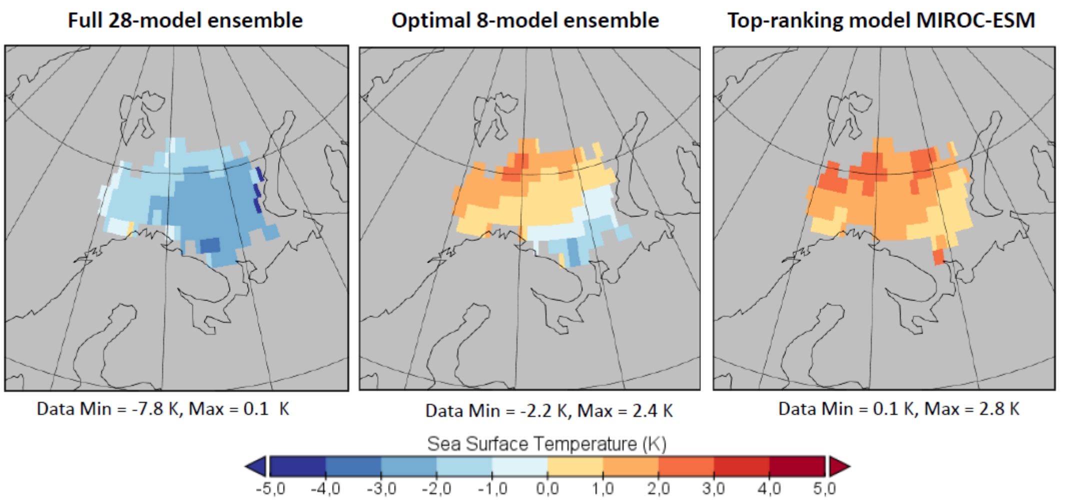
~~Figure 5c. Spatial distribution of biases in near-surface wind speed between selected model ensemble and reanalysis in six target seas averaged over the vegetation season and the time period 1979-2005.~~



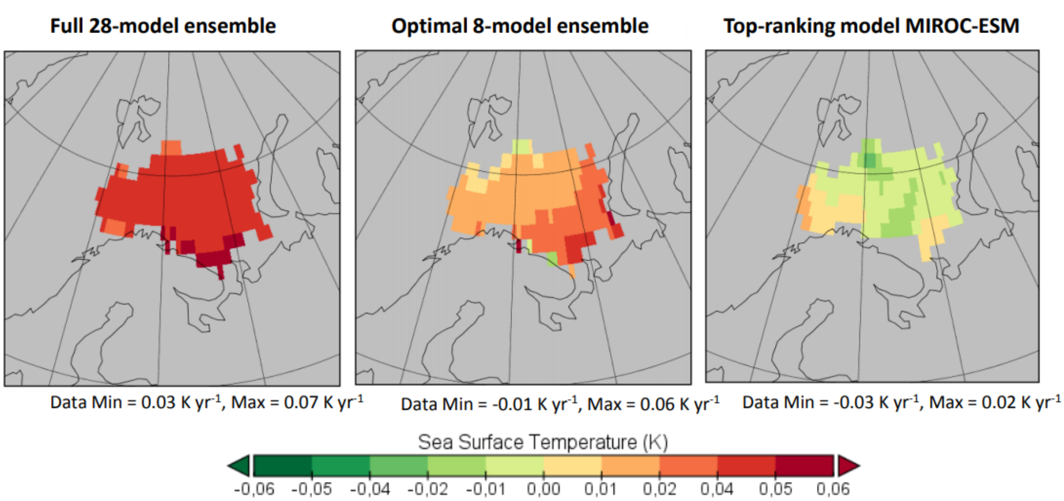
~~Figure 5d. Spatial distribution of biases in sea surface temperature models and reanalysis in six target seas averaged over the vegetation season and the time period 1979-2005.~~



~~**Figure 5e. Spatial distribution of biases in surface ocean current speed models and reanalysis in six target seas averaged over the vegetation season and the time period 1993-2005.**~~



~~**Figure 6a: Spatial distribution of biases in SST (K) between models and reanalysis data in the Barents Sea; the biases are averaged over June-September.**~~



~~Figure 6b: Spatial distribution of errors, which are calculated as the difference between model and reanalysis values of annual SST trends (K yr^{-1}) in the Barents Sea (June-September)~~

ID	CMIP5 models	Barents					Bering					Greenland				
		NO_3	PO_4	SI	pH	pCO_2	NO_3	PO_4	SI	pH	pCO_2	NO_3	PO_4	SI	pH	pCO_2
1	CanESM2	14			22	26	13			11	23	12			27	29
2	CMCC-CESM	9	12		9	25	13	3		10	20	7	13		6	13
3	CNRM-CM5	5	10	30			19	32	20			10	6	29		
4	GFDL-ESM2G	27	22	18	24	29	21	32	24	18	30	17	16	24	9	18
5	GFDL-ESM2M	23	21	26	12	23	18	34	25	13	29	12	11	26	5	16
6	GISS-E2-H-CC	10		26		5	9		21		22	9		27		6
7	GISS-E2-R-CC	11		29		11	11		22		7	10		25		15
8	HadGEM2-CC	12		8	20	26	5		11	8	24	9		8	20	27
9	HadGEM2-ES	12		14	19	27	4		12	7	25	10		8	24	32
10	inmcm4					22					26					27
11	IPSL-CM5A-LR	15	16	27	20		10	33	22	17		11	11	28	11	
12	IPSL-CM5A-MR	18	19	29	19		17	35	21	19		12	9	28	23	
13	IPSL-CM5B-LR	19	15	27	21		10	33	20	13		11	7	29	14	
14	MIROC-ESM					30					28					19
15	MIROC-ESM-CHEM					30					27					21
16	MPI-ESM-LR	23	24	28	27	33	8	32	23	10	24	19	17	23	15	26
17	MPI-ESM-MR	20	25	27	29	32	7	33	26	14	25	17	12	27	16	26
18	MRI-ESM1	18	14		19	26	13	32		6	26	22	22		27	27
19	NorESM1-ME	12	10	22	22	28	7	32	12	18	30	7	8	21	9	19
	total selected models	4	3	3	3	4	4	5	3	3	4	4	3	4	3	5

ID	CMIP5 models	Labrador					North					Norwegian				
		NO_3	PO_4	SI	pH	pCO_2	NO_3	PO_4	SI	pH	pCO_2	NO_3	PO_4	SI	pH	pCO_2
1	CanESM2	14			22	15	20			28	21	16			23	28
2	CMCC-CESM	14	5		6	13	23	6		9	15	9	17		6	14
3	CNRM-CM5	5	4	27			9	14	22			6	8	28		
4	GFDL-ESM2G	17	17	28	25	20	23	22	28	27	24	8	7	21	18	21
5	GFDL-ESM2M	14	10	28	17	15	10	21	21	23	16	16	21	25	10	10
6	GISS-E2-H-CC	16		26		22	24		23		12	12		31		16
7	GISS-E2-R-CC	17		26		30	24		20		20	15		26		25
8	HadGEM2-CC	17		11	18	14	22		12	23	15	12		15	12	16
9	HadGEM2-ES	17		15	23	17	22		12	25	17	13		15	15	20
10	inmcm4					14					21					11
11	IPSL-CM5A-LR	18	3	26	14		24	18	26	21		18	9	30	18	
12	IPSL-CM5A-MR	16	5	29	17		24	20	27	26		20	15	30	17	
13	IPSL-CM5B-LR	11	3	27	15		21	13	23	17		16	10	29	20	
14	MIROC-ESM					18					19					18
15	MIROC-ESM-CHEM					18					14					19
16	MPI-ESM-LR	17	12	34	23	24	27	27	26	30	30	22	24	32	17	22
17	MPI-ESM-MR	18	9	33	23	24	27	23	25	30	33	22	19	30	18	25
18	MRI-ESM1	20	19		16	23	26	10		28	20	21	13		18	18
19	NorESM1-ME	9	14	13	11	15	17	13	26	26	26	15	14	23	7	19
	total selected models	8	3	3	4	4	7	3	5	4	4	4	3	5	6	4

30 - selected optimal model ensemble

- score < 25% "very good"

- 25% < score < 75% "good" & "satisfactory"

- score > 75% "unsatisfactory"

Figure 5. Heat map with the final model scores obtained using the percentile ranking approach for the 5 biochemical variables (concentration of nutrients (NO_3, PO_4, and SI), dissolved CO_2 partial pressure (pCO_2), and pH) for the Barents, Bering, Greenland, Labrador, North, and Norwegian seas based on different statistical measures (Fig. 2, Tab. 2). The white areas indicate a lack of model output for historical and RCP projections (RCP4.5, RCP8.5) in open data sources.

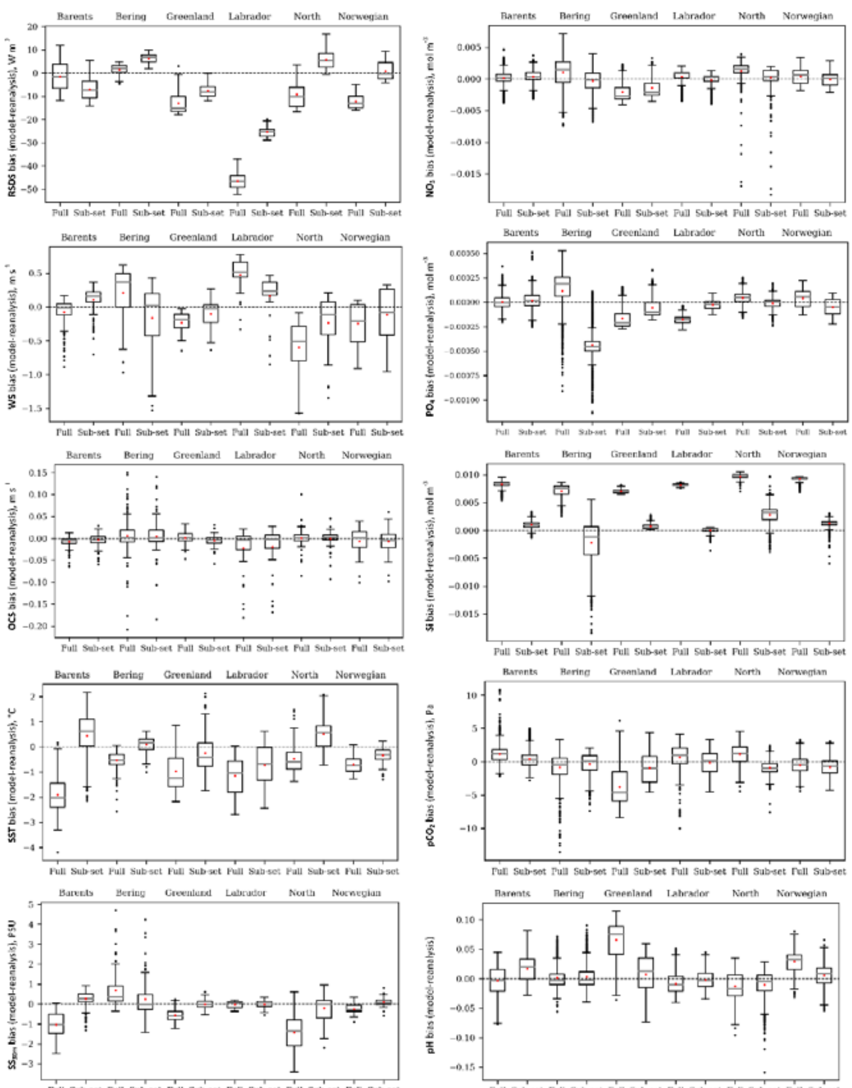


Figure 6. Box plots of the spatial distribution of biases (model ensemble minus reanalyses) of 5 oceanographic and meteorological (left), and 5 biochemical variables (right): sea surface temperature (SST), salinity averaged over 0-30 m (SS_{30m}), surface wind speed at 10 m (WS), ocean surface current speed (OCS), surface shortwave downwelling solar radiation (SDSR), concentration of

nutrients (NO_3, PO_4, and SI), dissolved CO_2 partial pressure (pCO_2), and pH for the Barents, Bering, Greenland, Labrador, North, and Norwegian seas averaged over the study period for comparison of full and selected model ensembles.