Dear Dr. Gattuso,

We are happy to submit a revised version of our manuscript “Acidification of the Nordic Seas”. The manuscript has undergone major revisions, and it has improved much from the last version.

Major changes include:

- We have included plots of H+ concentration changes in the supplementary material.
- We have assessed the sensitivity of our results to the choice between pH and H+ as the main variable in subsection 4.3.
- We now discuss the implications of using emission driven runs.
- We have clarified that, and why, we work with a 200 m upper layer when we analyze the observational data.
- We have worked on the introduction and the conclusions to enlighten what new our study brings to the scientific community, and how it relates to other studies.
- Please note that we have slightly changed the title after the suggestion of reviewer 3.
- We have revised the language/structure of the text to make it flow better.
- Please note that we found a minor bug in the code for creating Figure 14. When correcting this the residual became smaller compared to the last version of the manuscript.
- We slightly changed the region for which we calculate the aragonite saturation horizon depths in figure 12. In the last version it was for the entire Nordic Seas. Because cold-water corals are constrained to Atlantic Water, we excluded polar waters in the updated version of the manuscript by excluding areas west of 0 °E and 64 °N. The effect on our results are minor.
- We have included the freshwater decomposition of pH drivers in the main-manuscript.

Please find our responses (in blue) to all the reviewers comments (in black) at the following pages.

We want to thank you and the three reviewers for all the time you put into this manuscript.

Best regards,
Filippa Fransner and co-authors

REFEREE #1:

First, a huge ‘Thank you’ to the authors for this review. The invested work is immense and must have taken a lot of time. The improvement is clearly visible. Great job. It was a pleasure to re-read it. I have a few important general comments and then just a list of minor comment, which are rather suggestions. If you do not agree with the minor suggestions, you can keep it the way you wrote it. My only ‘major’ point is the choice of the surface layer (first 200 m). If that point is addressed, I am looking forward to seeing this great work published.

We want to thank referee 1 for going into such depth and detail also for this second revision, and putting so much time into reviewing the manuscript. The comments have been very helpful and they lead to a great improvement of the manuscript.

We have addressed the comments as described below.

General comments

1) Using the word “present”: The use of the word present for the period from 1981 to 2019 is still confusing to me. For example, the abstract states that pH has dropped by 0.06 from pre-industrial to present. It sounds as if present is now (2021), In the next sentence you use the period between 1981 to 2019 and say that pH has decreased by 0.10. How can the decrease be larger in only a small part of the period? Logically this means that pH has increased by 0.04 between 1850 and 1981. When you refer to until “present” you refer to the period form 1850-1981. This is highly misleading in the abstract and has to be explained here. You wrote this very clear in the Summary, maybe make it similar in the abstract. However, I would suggest not using the word present for the time from 1981 to 2019 but rather call it the period of regular ocean observations. Something like: “From pre-industrial to the beginning of regular ocean observations (1850-1981), pH is estimated to have dropped by 0.06 on average… From 1981 to 2019, when observations of the ocean are available, pH has dropped by another 0.10…”. In Fig. 4 present times refers to the period from 1996-2005, again different.

We agree with the reviewer that the use of the word “present” was unclear in the abstract, which has been thoroughly revised in the latest version of the manuscript. We have also gone through the manuscript to clarify this where needed. Additionally, In many places we have replaced present with present-day, which by definition refers to a longer period of time. This is to avoid that present being interpreted as “today” or “this year”.

2) It is often written “uptake of anthropogenic CO2”. However, Anderson and Olsen (2002) suggest that the anthropogenic CO2 is taken up by the ocean South of Nordic Seas and mostly advected laterally into the Nordic Seas, where some of it is even lost to the atmosphere, similar to the Arctic Ocean (Terhaar et al., 2019) and the Barents Sea in particular (Terhaar et al. 2020b). Maybe just rephrase it carefully to “mainly driven by an increase in anthropogenic carbon”.

Thank you for this remark. We agree that it is better to write ”increase in anthropogenic carbon”, instead of “uptake”. We have corrected this.

3) The difference between total, natural, and anthropogenic CO2 is not always clear to me. It becomes particularly confusing when opposing the abstract and the second paragraph of the Introduction, where you write that the Nordic Seas are a sink of CO2. However, this uptake does not enhance acidification if it is part of the natural equilibrium, because the Nordic Seas have always taken up CO2 because they are colder and transported that water away. It hence does not affect acidification (which by definition is a change of the natural state). Could you try to explain this precisely?

We have revised this part of the introduction (lines 45-51), and also added some text about the advective supply of anthropogenic CO2. We prefer not to introduce “natural” and “total” CO2, as we do not use these later on in the manuscript. We think that could introduce some confusion.

4) Why do you use the Mauna Loa data and not the global data (tab next to Mauna Loa)? Mauna Loa is around 1 ppm higher than global numbers and thus increases the difference between the ocean and atmosphere of total CO2 in the Nordic Seas.

In our manuscript we are using the Mauna Loa record to determine the annual growth rate in atmospheric CO2, This is not different across the global set of stations (https://gml.noaa.gov/ccgg/about/global_means.html). The absolute value, which indeed is different, does not play an important role in our calculations. We clarified this on lines 148-150.

Further, while the Mauna Loa record is based on direct measurements, the global estimate is a result of several steps of data processing. In addition, new sites have been added to the global record over the years. Using the Manual Loa record is therefore more straightforward.

5) I would still suggest writing the Methods mostly in passive tense as the book mentioned by James Orr suggests.

We have rewritten it to mostly passive tense. In some cases, especially where we explain certain choices/assumptions that we have made, we keep an active tense, as appropriate.

6) The surface layer of 200m seems to be far too deep to me, especially in fresh Arctic waters. The argument that primary production influences the first 200m does hold either in my opinion as different processes act at the surface (production) and below the surface (remineralization). Especially as the authors mentioned several times in the answers that long-term changes are the main point of the manuscript, the argument of the seasonal cycle is somehow empty. Averaging over 200 m can make it difficult to develop a process understanding. In case the 200 m should remain, I would at least like to see if things change drastically by choosing 50 m and 100 m to understand how sensitive are your results to the choice of the depth? I am especially thinking about section 3.2. You say that past changes are strongest in the Atlantic water and less strong in the Arctic waters. Maybe that is just because Arctic waters are much more stratified and by using 200 m depth, the depths from 100-200 m are much less affected. However, that may not be the case at the surface.

For the analysis of the discrete measurements, we made the choice to have a 200 m thick upper layer to minimize the number of seasonally affected layers, i.e., layers where seasonal undersampling could have an effect on our results. If we would have used 0-100 m and then 100-500 then both would be seasonally affected (since the winter mixed layer goes down to approximately 200 m) with more uncertain trends. The point is that we prefer to have only one layer where the trends need to be interpreted with more caution due to seasonal undersampling.

We have added this explanation on lines 228-231.

For the model data, on the other hand , we want to show surface data so that it can easily be compared to other (future) modelling studies where it is standard to show the surface fields.

One may then ask the question: the aim of the manuscript is to put the observed changes into perspective to climate change. Is such a comparison valid if using the 200 upper meters for the observations, and only the surface for the model?

- The main comparison between the observational and model data is provided in figure 4, where we also use the 200 upper meters of the model. We refer to this when we put the different time periods in context to each other.
- The surface maps are then used to get an understanding of regional differences (which we do not touch upon when analyzing the observational data due to the different data coverages in the various regions)
- Furthermore, the past and future drops in pH are very close when comparing the mean changes over the 200 upper meters, and the changes in the surface, indicating that the upper 200 m are broadly representative of surface. Very shallow surface layers, as in the Arctic proper, hardly occur in the Nordic Seas, here, winter mixing is usually deeper than 200 m, except for in the East Greenland Current

We have added some text on this on lines 278-281, 380, and 420, to make this clear.

Additionally, we realized that we have introduced some confusion by using “surface waters” both when discussing changes related to the discrete measurements and changes related to model projections. In sections 5.3 and 5.5, the results show the actual surface (i.e. 0m), while for the discrete measurements (Section 5.4) it is the upper 200m (Please note that in figure 4 the model data have been averaged over the upper 200 m). To avoid this confusion, we now use “upper layer” when we talk about the 0-200m layer, and “surface” when we refer to 0m.

7) Please check the labeling of tables and figures. Sometimes numbers are missing or the S for supplement is missing.

Thank you for noticing this. We have corrected the incorrect labels.

8) Overall, a lot of discussion is already happening in the results section. It would be great to make a clear cut between discussion and results although I understand that it is difficult.

We have merged the sections “Results” and “Discussion” to one “Results and Discussion”.

9) Did you subtract a piControl run to detrend your historical and future simulations? Especially in the deep ocean, the opposing trends in CT and AT maybe just a drift?

We checked if there is any drift in pH the pre-industrial control run to verify that the model is sufficiently spinned up. The change in pH in the pre-industrial control run is very small (more than one order of magnitude less) compared to the changes in the historical run and in the future projections, showing that model drift has a negligible impact on our results. We added some text on this on lines 173-176.

Minor comments

1) Line 2: This first sentence reads strange. The Nordic Seas are vulnerable to ocean acidification (1) at the surface because they already have low saturation states because of low temperatures and (2) below the surface because they have deep winter mixing and hence transport CT to the deeper ocean quickly in comparison to other parts of the ocean. The presence of cold-water coral reefs makes them not more vulnerable but gives one example of the impact of ocean acidification. Krill in the Southern Ocean or sea butterflies are also in danger because of ocean acidification. I suggest rewriting the first sentence.

We have rewritten the sentence after the suggestion of the reviewer.

2) Line 9: Suggest changing “relatively deep” to “below the surface”.

We now write “In some regions, the pH decrease ca be detected down to 2000m”

3) Line 10: Is the significant decrease everywhere or on average or in subregions?

We do not anymore mention “significant” in the abstract, and only mention the range of observed rates over the different subregions.

4) Line 15: Suggest adding “all” before cold-water corals.

done

5) Line 22: Please add globally somewhere here. As your paper is about the Nordic Seas, the reader may think that you are already presenting local results.

done

6) Line 22: The decrease of 0.1 is until which year? 2020? 1981?

We now write from pre-industrial to present days. Because we write “approximately 0.1”, we think that we do not need to be more precise.

7) Line 23: Suggest deleting “this” as there is only one ocean acidification.

We rephrased this sentence following the reviewer’s suggestion (see below).

8) Lines 23-26: Suggest rephrasing this sentence: It is not the pH drop that causes the reduction in $CaCO_3$. Maybe write: “Furthermore, the increasing CT also causes a reduction

in CaCO3 and hence poses a serious threat to marine organisms that have shells and skeletons consisting of CaCO3, such as pteropods and corals". Thus, you have old information (CT increase) before new (reduced CaCO3) and before the consequences (organisms).

Thank you for this suggestion, we used it in the text.

9) Line 37: All surface waters or only Arctic or Atlantic ones?

We rephrased to: "The surface water pCO 2 is generally lower than that of the atmosphere, making the Nordic Seas important sinks for atmospheric CO 2 ."

10) Line 40: Are the references for all three processes or can you add one reference for the primary production, one for cooling and one for the Arctic Ocean waters?

In some of the papers several processes are discussed, so we prefer to put all references at the end of the sentence. We noted that we had made one mistake and put Olsen et al., 2008 instead of Anderson and Olsen, 2002. We have corrected this.

11) Line 41: Do you have a reference for the deep-water formation in these regions?

We added references.

12) Line 47-49: Suggest moving this up in the paragraph as it is a general characteristic. You could write it as follows: The Nordic seas are particularly vulnerable due to low saturation states because of their low temperatures. Then, you continue with the Cant increase that worsens everything and you explain the increase in Cant with the circulation. I think that flows better.

Thank you for the suggestion, we have revised this paragraph.

13) Line 52: Acidification rate is not precise here as it includes saturation states and pH. Please say what this refers to.

Done

14) Line 66: It does not increase it, it creates it in the first place, right?

We have reformulated this part to avoid any confusion related to the choice of words.

15) Line 93: More than only two types of CaCO3 exist in seawater. These two are the most abundant ones.

We have re-written the sentence to mention that these two types are the most abundant ones.

16) Line 95: Not all CaCO3 dissolves. Maybe write: "When Omega of a CaCO3 mineral is less than one, the water is corrosive, and that mineral starts to dissolve"

Done

17) Lines 123/124: Is there a reason why you give the uncertainties sometimes in absolute numbers and sometimes in relative numbers?

Yes, this is what is given in the Olsen et al., 2019 paper as the consistency of the GLODAPv2 data. The values are based on the pre-defined minimum adjustment limits used in the secondary quality control routines and are therefore relative for those variables for which a potential bias is expected to vary with the concentration of the property measured and absolute for those variables where a potential bias is expected to remain constant over the range of concentrations encountered in the ocean.

18) Figure 1: Consider changing the colormap as it now suggest a hard cut at 2000 m.

We now use a blue-only colormap.

19) Line 172: Why is this behind the model data and not part of the observational data or at least behind it?

Thank you for this observation, it makes more sense to have it after the subsection on observational data. We moved it there.

20) Line 175: Suggest deleting: "It is important to mention that"

We removed this phase as there are certainly several regions with poor data coverage, and not only this one.

21) Line 184: Do you know the pre-industrial CO2 level of the model?

We are now providing the modelled CO_2 concentrations in Table S1 in the supplementary material.

22) Line 207: How much is relatively small?

After the request of James Orr we now put figures showing change in H+ in the supplementary material. This sentence has been removed. We also included a new subsection (4.3) with figure 2 where we discuss the choice between H+ and pH.

23) Line 212: Can you just say why you used exactly these values?

We used these values because they are reasonable representative values of the Nordic Seas silicate and phosphate concentration (averaged over the whole dataset). We now describe it in the manuscript.

24) Line 239: Please precise what you mean by 'it' and please add a number to the equation.

Done

25) Line 276: Maybe consider subscripts that are understandable without having to look it up instead of 1 and 2.

We changed subscript 1 to ‘point’ (for discrete data) and 2 to ‘field’ (for mapped fields).

26) Line 290: Can you just briefly motivate the choice of these two periods please.

These periods were selected to include the productive season. We chose the longer period to include the spring bloom and the summer production, and additionally the summer period only because this is the time of lowest nutrients/DIC values. We added some text on this.

27) Line 300: Can you state why you think that they are of minor importance?

Here we put a reference to section 4.1, where this is discussed/shown.

28) Figure 2: In the legend you write present (1981-2019), but the tick in the figure suggest that you are showing 1980-2019. And it there a reason why you are showing the model data and not the change to 2002?

Thank you for spotting this error. We remade the figure so that the subplots exactly fits the time periods.
We show the model data because we want to show that the model performs reasonably well in the surface layer

29) Line 303: Suggest replacing “Before going into regional details of pH changes,” by “First,”

We now write “Here we give an overview of the upper layer, taken to be the upper 200m for both model and observations, pH changes”

30) Line 303: Global or Nordic Seas?

We added “the Nordic Seas” to the sentence.

31) Line 306: These numbers are now calculated by adding the change to 2002 or it is still the absolute model output?

It is the modelled pH output, we added a sentence to explain this.

32) Line 308: I think it is rather for a few locations than a few years. It is likely one cruise and if it is at the end of the range of pH it does not agree with the average.

We agree that this is probably the case. We removed this part. We think that the text on lines 392-395 describe this well.

33) Lines 310-312: Your samples at the beginning of the time period seem to be biased towards high pH regions. That could explain the difference in the trend. If you show the trend for the last 20 years, you will probably get a much better agreement. Maybe worth to think about mentioning that instead of writing that the data can be questioned. It is great data, no reason to make it look bad.

It was not our intention, indeed, to question the quality of the data, but rather their representativeness for the whole Nordic Seas. We concur with your example, and added it to the text.

34) Line 313: Suggest deleting “as expected”

Done

35) Line 315: Suggest addition “CMIP5” in front of model ensemble range.

This part has been revised. We now talk about inter-model spread instead of model ensemble range. We think that adding CMIP5 in front of “inter-model spread” would disturb the flow of the text. It is enough that the reader gets this information from section 3.3.

36) Lines 323-331: This paragraph is about the difference between Nordic Seas and global ocean pH. Starting it with 1850 gives the impression to the reader that we are now looking into the past but that is not the main point. I suggest starting the paragraph with what it is mainly about. That improves the flow. Just moving ‘in 1850’ to the end of the sentence may do the job.

This is a good point, we did as suggested.

37) Lines 333-334: How did you chose the salinity? Wekerle et al. (2016) have a regional definition? You show Arctic waters at the Norwegian coast. That makes not much sense… Did you include these waters as Arctic waters into the scatter plot?

The definition of polar waters /Atlantic Water varies slightly in the literature. We added two references using this specific definition (Malmberg and Desert 1999, and Nondal et al., 2009). It is right that when using a purely salinity-based definition the Norwegian coastal water are classified as polar waters. However, we prefer to keep this grouping of the water masses because the effect of the spatial varying salinity is the largest in the low salinity waters, including the Norwegian coastal water (compare figure 3 d and e, and 3g and h, in the revised manuscript). In the revised manuscript we refer to the watermasses with a salinity <34.5 as low-saline waters. We clarify that the low saline waters include the polar waters that are found in the northwestern part of the domain, and the Norwegian coastal waters that are confined to the Norwegian coast.

38) Line 339: This seems to be contra-intuitive. First you state cold water holds more CT and then you say polar waters hold less? It is because of low pCO_2 after a long time after sea ice or because of the low alkalinity?

Both the low pCO2 and low alkalinity (salinity) results in a low CT. We have added a sentence to clarify this.

39) Line 342: Suggest replacing nonphysical by another word. If it happens in the real world, there must be a mechanism.

This paragraph has been removed because we think that it did not add anything to the manuscript. We think that figure 3 and Table 3 in the current manuscript is enough to illustrate the role of the different drivers in the spatial variability of pH.

40) Line 355: Do pH and pCO2 not always correlate, not only in the Nordic Seas?

This is true, we removed this part, we think that it does not add anything to the paper.

41) Line 365: Bring the salinity earlier when you describe the results and not after the discussion

Done

42) Lines 346-366: This paragraph is very verbose. Is it possible to condense the information?

This is very true, we removed quite a large part of the paragraph.

43) Line 371: Can you just explain why they reinforce each other?

We removed this part, we think that it did not give any added value.

44) Line 376: How can it be more important somewhere if everywhere all other contributions already explain 100%?

This is a result of the rounding, the salinity adds less than 1% to the explanation of the variability. We added a note on this.

45) Line 387: Why do you not just use the original data for the calculation? Thus, you do not need to make this statement.

In this section we prefer to use the depths obtained from interpolation in the figure, to be consistent with the figure. The original data (GLODAPv2) has a much coarser resolution.

46) Line 391: This should be figure 6, I think.

Yes, we had made a mistake with the labeling. This is now corrected.

47) Figure 6: Why are you still not using the whole range on the y-axis? Now, I cannot see the data well enough.

We want to keep the same ranges on the y-axes in all subplots to make it easier for the reader to compare the panels and to put the data in the different subplots in relation to each other.

48) Line 392: Maybe the Barents Sea trend isnot there because the water is too stratified? A surface trend may exist while waters below 50 or 100 m remain unaffected? The same happens probably in the Norwegian basin (lines 394-395)

Winter mixing is deeper that 200 m in both of these regions, so we do not expect that the upper 50 m trend should be different from the 50-200 m trend (we did verify this by calculating the trends also for the upper 50 m).
Regarding the comparison with Skjelvan et al., (2014), they also used the upper 200 m for the surface layer. We clarified this in the manuscript.

49) Line 393: between ±0.2 and ±0.8

Correct, we changed this.

50) Line 429: Maybe add that the largest decline in Atlantic waters is mostly due to the effect that polar waters are already close to zero and the decline in units of saturation state are becoming smaller and smaller.

We introduced a line where we say that the reasons behind this is discussed in Section 5.7.2. The reason to the larger decline of Aragonite saturation state in the Atlantic waters in esmRCP8.5 compared to esmRCP2.6 is that in esmRCP8.5 the warming is more uniform over the Nordic Seas compared to esmRCP2.6, where it mainly occurs in the Atlantic waters. The warming has a positive effect on the aragonite saturation, meaning that in esmRCP2.6 the spatially different warming reinforces the effect of the spatially varying CT increase.

51) Line 434: This shows that 200 m may be too deep.

Please see our discussion under your major point above.

52) Line 442: Different in which way?

We clarified this (lines 520-524).

53) Lines 500-507: Could the sampling also cause the trend in alkalinity as it did for pH above?

This is probably not the case as we also see the strong trend in the 200-500m layer. This is discussed at line 557-563 .

54) Lines 511-515: In addition to Shu et al. (2018), you should also cite Woosley et al (2020) for the effect of freshening on CT uptake (line 513) and Terhaar et al. (2021) for the effect of freshening on ocean acidification in polar waters (line 515, line 523, line 526).

Thank you for these references, we added them.

55) Figure S5: Could you show the pH change with respect to GLODAPv2 in 2002 for each model? That would allow to clearly see how large the trends are in comparison?

We have added a table (S6) showing the mean pH change in three depth layers obtained from the GLODAP-climatology in combination with the modelled DIC and AT change in the different ESMs.

56) Lines 600/601: Can you compare two different periods? The atm. CO2 increase is exponential, so one would expect the trend since 1991 to be even stronger. On the other hand, the Nordic Sea trend may be biased high because of sampling biases in high pH regions at the beginning of the observational period.

We removed the comparison with the global data.

References

Anderson, L. G., and Olsen, A., Air–sea flux of anthropogenic carbon dioxide in the North Atlantic, Geophys. Res. Lett., 29(17), 1835, doi:10.1029/2002GL014820, 2002.

Terhaar, J., Orr, J. C., Gehlen, M., Ethé, C., and Bopp, L.: Model constraints on the anthropogenic carbon budget of the Arctic Ocean, Biogeosciences, 16, 2343–2367, https://doi.org/10.5194/bg-16-2343-2019, 2019.

Terhaar, J., Torres, O., Bourgeois, T., and Kwiatkowski, L.: Arctic Ocean acidification over the 21st century co-driven by anthropogenic carbon increases and freshening in the CMIP6 model ensemble, Biogeosciences, 18, 2221–2240, https://doi.org/10.5194/bg-18-2221-2021, 2021.

Wekerle, C., Wang, Q., Danilov, S., Schourup-Kristensen, V., von Appen, W.-J., and Jung, T. (2017), Atlantic Water in the Nordic Seas: Locally eddy-permitting ocean simulation in a global setup, J. Geophys. Res. Oceans, 122, 914– 940, doi:10.1002/2016JC012121.

Woosley, R.J. and Millero, F.J. (2020), Freshening of the western Arctic negates anthropogenic carbon uptake potential. Limnol Oceanogr, 65: 1834-1846. https://doi.org/10.1002/lno.11421

REFEREE #2

Overall the reviewers have addressed my comments on the previous version well, here I have just a few extra minor issues that should be taken care of before publication.

Line numbers refer to the manuscript without tracked changes.

We are grateful to reviewer 2 for re-reading the manuscript and identifying these issues. We have addressed all of them as described below.

Abstract

Line 2: consider 'Due to' instead of 'With...', because the latter doesn't imply causality.

Done

Line 2: Nordic Seas 'are', not 'is'.

Done

Line 3: and 'their' impact, not 'its' impact.

We decided to remove this part of the sentence to obtain a better flow.

Line 6: 'meter' => 'm' and add a space between '2000 m'.

Done

Line 6: consider 'well below' => 'still deeper than', for clarity.

This part of the abstract has been removed to obtain a better flow.

Line 7: add apostrophe after 'the Nordic Seas'

Also this part has been removed.

Line 8: by 'significant', do you mean 'statistically significant' or 'considerable/important'?

We mean statistically. We no longer mention significant in the abstract.

Line 9: 'until the year of 2100' => 'by the year 2100'.

This part has been removed.

Line 10: 'which are close' => 'which is close'.

Done

Line 12: 'undersaturated in aragonite' => 'undersaturated with respect to aragonite'.

Done

Line 13: remove apostrophe after Nordic Seas -or- add 'the' before Nordic Seas
Done

Introduction

Line 20: 'of which about 20% HAS been taken up'

Done

There are too many of these sort of language errors for me to continue to point them out, but the whole manuscript needs careful copy-editing.

We have carefully gone through the manuscript to identify language errors.

Eq. (4): this is the equation for Free scale pH, but you actually work with Total scale pH in the rest of the study. Same comment applies on line 239 for the dpH term.

We revised equation 4 to show pH on total scale. For the drivers we consider the contribution of sulphate to be negligible and do not include it in the equations. We added a sentence on this.

Line 83: AT is the difference between the sum of proton acceptors (weak bases) and the sum of proton donors (weak acids), not just the former.

The reviewer is correct. For simplicity, we now write that AT is mostly determined by bicarbonate and carbonate.

Line 93: CaCO3 might be virtually all aragonite and calcite, but there are other forms in seawater too.

We revised this line to write that they are the two most abundant forms.

Methods

Line 193: does CT/AT literally mean CT divided by AT or is it shorthand for something else?

We do not refer to the ratio, only that we include both CT and AT. We have clarified this by writing CT+AT.

Line 274: "including a correlation term would decrease the uncertainty". This statement is not true, or at least not certain. It depends on the sign of the uncertainty correlation term: if positive, then the uncertainty in e.g. pCO2 would be reduced, but if negative it would be increased. For other variables, with different patterns of variability in CT-AT space, the opposite could apply.

The statement has been revised as following: “The correlation between uncertainties in AT, CT were set to 0. This is a reasonable assumption given that CT and AT are measured on different instruments using different analytical methodologies. In addition, including a positive correlation term would decrease the overall uncertainty and we prefer a potential overestimation”

Discussion

Lines 469, 471: missing figure number (11?)

There was a bug in our latex document that caused this, this has now been fixed.

Results

Line 323: still a rogue "units" after a pH value here.

we removed the remaining “units”

Figure 3: subplots (c) onwards should be equal aspect and might benefit from explicitly drawing the 1:1 line.

Thank you for this comment. We remade the scatter plots so that they have an equal aspect ratio. It makes the figure much better!

Conclusion

Line 628: final sentence does not make sense.

The conclusions has undergone a major revision, and this sentence has been removed.

REFEREE #3:

In my review of the original manuscript, I raised a number of concerns, which the authors have tried to address in their revised manuscript along with their response to the reviewers. They have dedicated much effort to this task, which is commendable. However, I still have some non-trivial concerns with the revised manuscript.

We are grateful to James Orr for this thorough review, and all the time he put into the manuscript. The comments are very helpful and we think that working through them has led to a much improved manuscript.

We have addressed the comments as described below.

TWO MAJOR CONCERNS

In my original review, my concerns were listed in order of importance. My first concern was that the original manuscript did not consider potential misinterpretation of pH changes, which because of the log scale depend on the initial state of [H+] not just the change in [H+]. The authors response was to (1) mention H+ in the Introduction, (2) add a 4-line paragraph mentioning this concern, and (3) add Figure S17 and Table 7, which show trends in [H+]. In their response they also say "the pH variations in this study are relatively small, and our results do therefore not look significantly different if analyzing H+ instead". They also state that they "remade all our plots showing pH change to verify this", although they only show Figure S17. But in some of the pH maps in the revised manuscript, there are differences in pH across the Nordic Seas of about 0.1 for the preindustrial references state (Figure 4) and about 0.15 for the present-day reference state (Figure 8). These differences appear small, but they imply 30% to 40% differences in terms of the initial value of [H+]. These regional differences in the initial state of surface pH are much larger than the regional differences in the pH change between preindustrial and present (0.02, Fig. 4) and between present and future under RCP2.6 (0.04, Fig. 8); they are comparable to those between present and future under RCP8.5 (0.15, Fig. 10). Therefore differences in the pH reference state (preindustrial or present) shown in the maps of Figs. 4, 8, and 10 will have a substantial effect on the computed pH change. The authors have taken my previous comment too lightly. I asked for major revisions, but they did not comply. I did not ask that they not show pH in the revised manuscript, something they seem to imply in their response. This concern could be properly addressed if the authors would show changes in both H+ and pH on the same set of plots and if they would quantify how much of the regional differences in the pH change are due to differences in the initial state of [H+] and how much is due to the change in [H+].

We understand the concern of James Orr. To support our choice of working with pH only, we have done the following changes in the manuscript:

1) We added figure 2. and section 4.3 where the issue is discussed. In this figure we have plotted the change in pH vs the change in H+ concentration over a range of initial pH values, for six different increases in CT. The figure shows that for the initial pH values that are found in the Nordic Seas (in present climate), the relationship between the pH change and H+ change is approximately linear for the CT increases that we have in this study. This shows that the choice between H+ and pH does not have an important effect on the results in our study. The linear relationship breaks down for larger changes in CT, and/or of the initial pH spans lower pH values. In these cases it is more appropriate to present the results in H+.
2) We added maps for H+ in the Supplementary material (Figs. S17, S19 and S20), showing the H+ concentrations for the various periods of time, together with the change in the H+ concentration, to support the conclusions drawn from figure 2.

My second concern expressed was that the authors were only using 1 model to make projections. In response, the revised manuscript now compares results from multiple CMIP5 Earth system models (ESMs) for some of the analyses. To make this comparison, the authors rely results from those models under emissions-driven, not concentration-driven scenarios. Unfortunately, the emission-driven scenarios mentioned by the authors are not part of the core set of CMIP5 experiments (Taylor et al., 2012). The concentration-driven scenarios were designed for CMIP and IPCC to be the standard way of comparing models because that way the atmospheric CO2 forcing, both radiative and geochemical, remains identical between models. Otherwise, with an emissions-driven scenario the choice of a terrestrial carbon-cycle component of an ESM, may lead to quite unrealistic atmospheric CO2 levels and thus have an undue influence on ocean carbon. That is why the concentration-driven scenarios have been used in the most cited previous studies that compare CMIP5 as well as CMIP6 Earth system models (Bopp et al. 2013, Kwiatkowski et al., 2020). Yes, an emission-driven scenario can be used IN ADDITION to a concentration driven scenario to study carbon-cycle feedbacks, but that is not the focus of the authors in this study. Therefore the authors comparison of the CMIP5 ESMs using emission-driven scenarios alone is poor experimental design.

We thank James Orr for his comment. We agree that when combined, the emission-driven and concentration-driven experiments can be used to quantify the carbon cycle-climate feedback. Nevertheless, the emissions driven runs (at least historical and RCP8.5) are actually listed as CMIP5 core experiments in Taylor et al. (2012, Fig. 2). They are further described as “suitable either for comparison with observations or provide projections.” Below we further elaborate our motivation for using the emission-driven simulation.

The authors further confuse the subject by calling their emission-driven scenarios "RCPs" (Relative Concentration Pathways). Yes, CMIP5 provided emissions for the emission-driven scenarios, such as "esmrcp85", the same set of emissions used to drive the chosen integrated assessment model to come up with the common concentration pathway used by all ESMs under RCP8.5. But the authors cannot refer to their simulations as those from relative concentration pathways (RCPs). The best way to remedy this problem would be for the authors to compare results from the ESMs forced under the actual relative concentration pathways (experiments rcp26, rcp45, rcp85), as has previous work, instead of the emission-driven scenarios (esmrcp26, esmrcp45, esmrcp85).

Before we argue for the use of emission driven runs, we need to give some background on the choices we have done while working on this manuscript:

From the beginning, the primary focus of our study has been to use historical observational data sets to synthesize acidification patterns and their drivers over the past 39 years in the Nordic Seas. The use of the historical run and the future projections from an Earth system model (ESM) is to put these observed changes and its drivers into the climate-change perspective. Initially, we chose to work with one ESM only to facilitate a deeper analysis of regional changes and their drivers in several future scenarios (i.e. to avoid a too long paper). We decided to employ the Norwegian Earth system model due to our in-house expertise with this model, and because we have all required output variables readily accessible. At the start of the analyses of this paper, we had both concentration- and emission- driven runs available. For the emission-driven runs, some model improvements had been made compared to the concentration driven ones, including i) an extended spin-up by a few hundred years and ii) a replacement of the local mass-conservation correction related to surface freshwater fluxes with a global mass correction. The latter leads to improved representations of surface tracer concentration and trends in the deep ocean. Indeed, in an analysis of the output of both runs demonstrated that, the emission-driven NorESM runs revealed a better representation of the carbon system parameters against observations, and therefore is more credible in its projections.

In the first review-round, the reviewers pointed out that we should show how our NorESM simulations perform in comparison to other ESMs, not necessarily in all figures in the paper, but at least where we assess the impact on deep water corals, and maybe also in a supplementary figure. We fully agreed with this, and added simulations from seven emission-driven ESMs to our paper. We chose emission-driven runs for the model ensemble to maintain consistency between our utilised model run and the other ESMs.

We agree that the atmospheric CO2 forcing can vary between emission-driven ESMs. However, we disagree that using emission driven runs is poor experimental design for the following reasons:

Emission-driven simulations are run with a fully coupled, interactive carbon cycle, and therefore include the carbon cycle feedback on the physical climate. Compared to concentration-driven simulations, the emission-driven ones therefore have a more realistic representation of interactions that take place within the climate system, under a given CO_2 emission scenario. We also acknowledge that for this reason, inter-model spread in emission-driven runs is larger than in concentration driven runs (e.g. Booth et al., 2013, Friedlingstein et al., 2014). The choice of multi-model emission-driven ensemble therefore includes a more comprehensive estimate of the effect of model-related uncertainties on climate projections, which is the main point here and which we therefore prefer.

Indeed, concentration-driven runs are useful to identify the sources of the inter-model spread in an individual model component as the carbon cycle climate feedback is excluded. This is what is done in many model-intercomparison studies prepared for CMIP and IPCC, where the scope partly is to understand the sources of the model-differences. But, the aim of our study is not to understand and attribute the source of inter-model spread in the projected acidification, i.e. it was never designed to be a model-intercomparison study. Here, we use model simulations primarily to (i) project the rates of future acidification and assess potential impacts on cold-water corals under different scenarios (e.g., high- vs low-CO_2) and (ii) elucidate the mechanism governing these rates. Nevertheless, we agreed in the first review round to add results from other models as suggested by the reviewer "… to show at least where the NorESM1-ME model is situated relative to other Earth system models." We agree that it is useful to give the reader an idea of model-related uncertainties in the climate projections.

We would like to add that, despite many model-intercomparison studies prepared for CMIP and IPCC use concentration-driven runs, there are several biogeochemistry model-intercomparisons that do not focus on carbon-cycle feedbacks but still employ emission-driven runs that have been published (e.g., Zhao et al., 2014, Kessler and Tjiputra, 2016; Wang et al., 2016; Oschlies et al., 2017).

We are confident that our preference of emission-driven runs over concentration-driven runs is valid, and have decided to stay with this in the manuscript. However, since the reviewer has raised this key issue, we have decided to clarify and discuss the projection uncertainties associated with emission-driven runs, and we have therefore done the following changes in the revised manuscript:

1. We have replaced all mentions of RCP with esmRCP when talking about the emission-driven runs
2. We have estimated the impact of the modelled deviation in atmospheric CO2 in our emission driven NorESM-runs, from the concentration driven ones, by adding Table 1.
3. We added pH-timeseries from concentration-driven ESM's in figure S5.
4. We discuss the implications of emission vs concentration driven runs at lines 169-174, and 172-182 in the section about models, and on lines 386 and 405 in section 5.2.

References:

- Kessler, A. and Tjiputra, J.: The Southern Ocean as a constraint to reduce uncertainty in future ocean carbon sinks, Earth Syst. Dynam., 7, 295–312, https://doi.org/10.5194/esd-7-295-2016, 2016.
- Zhao, F. and Zeng, N.: Continued increase in atmospheric CO_2 seasonal amplitude in the 21st century projected by the CMIP5 Earth system models, Earth Syst. Dynam., 5, 423–439, https://doi.org/10.5194/esd-5-423-2014, 2014.

- Wang, L., Huang, J., Luo, Y. *et al.* Narrowing the spread in CMIP5 model projections of air-sea CO_2 fluxes. *Sci Rep* **6,** 37548 (2016). https://doi.org/10.1038/srep37548
- Oschlies Andreas, Duteil Olaf, Getzlaff Julia, Koeve Wolfgang, Landolfi Angela and Schmidtko Sunke, 2017: Patterns of deoxygenation: sensitivity to natural and anthropogenic drivers, Phil. Trans. R. Soc. A., 375.
- Booth, B. B. B., Bernie, D., McNeall, D., Hawkins, E., Caesar, J., Boulton, C., Friedlingstein, P., and Sexton, D. M. H.: Scenario and modelling uncertainty in global mean temperature change derived from emission-driven global climate models, Earth Syst. Dynam., 4, 95–108, https://doi.org/10.5194/esd-4-95-2013, 2013.
- Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks, *Journal of Climate*, *27*(2), 511-526. https://journals.ametsoc.org/view/journals/clim/27/2/jcli-d-12-00579.1.xml

Overall, it appears to me that this manuscript still would still require substantial revision, and thus I cannot recommend it for publication at this time.

OTHER CONCERNS (of moderate or minor importance, language suggestions)

Title: The title "Nordic Seas Acidification" does not quite work because "Nordic Seas acidifcation". For the same reason, we cannot say "Oceans Acidfication". Thus I suggest to change the title to "Acidification of the Nordic Seas".

Thank you for this remark. We have done as suggested.

ABSTRACT

L2: change "Nordic Seas is" to "Nordic Seas are"

Done

L3-5: The second sentence is hard to understand because of too many clauses and commas. It should be rewritten. A solution that could work would be simply to delete ", and its impact on cold-water corals,". Moreover, this paper does not investigate impacts on cold-water corals, only the aragonite saturation state of the waters that corals may be exposed to. In addition, cold-water corals are mentioned later in the abstract so there is no need to try to squeeze that in here.

We followed the suggestion to remove “and its impact on cold-water corals”.

L6-7: The authors make contradictory statements about impacts of the shoaling ASH on cold-water corals (compare L6-7 to L14-15).

The abstract has been revised, these contradictory statements are no longer there.

L8: "significant" should not be used here (or elsewhere) unless it is replaced by "statistically significant" and the authors tell us with respect to what it is statistically significant.

We have rewritten the abstract so that we no longer use the word significant.

L10: same problem as L8

Solved as described above.

L14: change "to be lifted to" to "to shoal by"

The abstract has undergone a major revision, and we no longer mention the actual depths of the saturation horizon obtained with the model projections.

L16-17: delete the qualitative ending ", which to some extent is opposed by increasing alkalinity".

This qualitative statement is no longer in the abstract.

INTRODUCTION

The objectives at end of Introduction (just before subsection 1.1) should be refined. Is the justification for this study just that the authors want to look at rates of acidification in both models and observations? Would they expect those rates to be so different from previous estimates. If so, why would they differ substantially. Unfortunately, the authors fail to return to this so-called gap in the Discussion, i.e., by comparing their results to the previous estimates of Skogen et al. (2014, 2018).

Thank you for this remark. The justification of this study is to get an overview of the pH changes, and their drivers, from pre-industrial to 2100 for various scenarios, and to get a more in-depth understanding of how this varies within the Nordic Seas and over depth, by compiling different kinds of data-sources.

We revised the last part of the introduction by:

- describing more in detail the work of Skogen et al., 2014 and 2018
- refining the objectives to make it more clear what new this study brings.

We have also revised the conclusions and the discussion by putting our work in context to previous work.

L34: change "its" to "their"

Done

L35: It would be clearer if the authors changed "characterized" by "dominated"

Done

L39: change "comes as a result of" to "results from"

Done

L42: change "and consequently help" to ", helping"

Done

L44-45: Confusing. I think the authors should replace "would ultimately lead to early and relatively large detection" with "leads to higher".

We agree that this was confusing. We decided to remove the link to the acidification in this phrasee, and put it later.

L46: change the verbose "have negative impacts on" to "degrade"

We put the link to the cold-water corals in a separate phrase.

L52. "Acidification rates" is ambiguous. All readers will not get that you mean pH change. Please refer to pH change explicitly.

Done

SUBSECTION 1.1
- You cannot have a subsection 1.1 if you do not have a subsection 1.2

We made it to a separate section.

- This section seems to be a repetition of well-known CO2 system chemistry in the ocean. It is largely textbook material.
Because it is not novel, please just to delete it and cite references where readers can go if they are unfamiliar with seawater CO2 chemistry.

We do think that this section makes an important part of the manuscript because we refer to it several times later on in the manuscript. We therefore decided to keep it.

- Equation (4): I think that if the authors are going to actually give equations with [H+] or pH, they should also mention that in seawater oceanographers have different pH scales, while also pointing out the pH scale that they have adopted for this study (presumably the total scale).

We revised equation 4 to show pH on total scale.

L115: Is it a Norwegian or EU program. Please state.

It is Norwegian, we now state this in the manuscript.

L123-124: What does "considered consistent" mean? The last part of the sentence could be shortened to "4 umol kg-1 for both Ct and At"

Olsen et al.2016 mean consistent among cruises. We added this to the sentence. We revised the last part of the sentence as suggested.

L139: shorten last part of sentence as suggested just above.

Done

Section 2.1.2: see 2nd major comment

Please see our answer under the 2nd major comment.

Section 2.2.2

L204:
- change the verbose "It is important to keep in mind that" to "Because"

We have removed this paragraph. We now discuss the issue in section 4.3

- delete "and that"

See answer above.

L207: Why say "are relatively small". Cannot the authors be more quantitative?

See answer above.

Present

L209-210: Why mention temperature twice in the same sentence?

Thank you for this observation. This has been revised.

L217-218: suggest to change " (they are" with ", being" and to remove the final ")".

Done

L222: comma fault: remove the comma

Done

L227: change "generally is" to "are generally"

This part has been revised, and we do no longer mention the summer mixed layer depth.

L235 (Equation (7)): The authors need to explicitly state how they calculate the partial derivatives.

Done

L239: what is meant by "it"?

We refer to H+ in equation 10 (8 in the last version of the manuscript). This has been clarified.

L248-249: The rates of change of atmospheric xCO2 (ppm) and atmospheric pCO2 (uatm) will not be identical unless the correction factor between them is 1 (e.g., atmospheric pressure of 1 atm and no humidity). In the cold high latitudes the atmospheric pressures are substantially less than 1 on average. The authors could say though that the rates of change of those two variables are proportional.

Thank you for pointing this out, we adopted your suggestion.

Section 2.2.3

L270: in the "MATLAB version of" CO2SYS?

Added

L280-281: While systematic uncertainties would tend to cancel out when calculating differences, random uncertainties would not.

Thank you for this, we adopted the suggestion.

L296: Delete "We note that"

Done

RESULTS

line 315:
- suggest to change "model ensemble range" to "CMIP5 model range". The word ensemble as used by modelers is ambiguous (multiple simulations with the same model or multiple models).

We now write “inter-model spread”.

- The CMIP5 model range in final pH is actually quite large. It would be much tighter if the authors would have compared results from the concentration driven pathways rather than the emission driven pathways.

Please see our answer to your second major concern above.

line 324: Delete "Note that". How much lower?
Done. It is about 0.1 lower, we added this.

line 325: change "the one" to "that"

Done

line 329: The reference to Dai et al. (2019, Nature Comm.) confuses the issue. Consistent with previous studies, Dai et al. define Arctic amplification in terms of surface air temperature, not surface ocean temperature. In the Arctic there has been little warming of surface sea surface temperature because of ice cover; conversely, there has been enhanced warming of surface air temperature, i.e., about twice the global average. But what about the Nordic Seas that are largely ice free even during winter? I would expect the authors to focus on the Nordic Seas, given the title of the paper. Is the historical SST change there higher than the global average?

We do see a faster warming in the Nordic Seas in the esmRCP26 and esmRCP45 scenarios, but not in the esmRCP8.5 nor in the historical (we realize that this was not described in detail enough in the manuscript). It is a good point that the simulated faster warming does not have to be related to Arctic amplification. It could also be related to differences in other physical characteristics (for example stratification) and how it changes with climate.
We realize that this subject is too complex to go into in this paper, and we therefore remove the text related to the faster warming. Instead we now write:

"This is partially a result of the colder waters of the Nordic Seas, which gives them a lower buffer capacity. Additionally, in esmRCP8.5, there is an increase in the pCO_2 undersaturation of the global ocean that increases the global average pH (Fig. S16). Other factors driving this decreasing pH difference between the global ocean and the Nordic Seas can be differential heating. A quantitative assessment of the drivers is beyond the scope of this paper."

Lines 327-330: To make this statement, the authors would need to demonstrate that the SST change in the Nordic Seas is actually higher than the global average. Or it could be documented with appropriate references.

Please see our answer to the previous comment.

line 334: delete "its"

Done

line 337: It would be simpler and clearer if the authors could write "from southeast to northwest".

Done

line 337-338: I do not see evidence in Figure S15 for the authors statement here that CT increases with decreasing T in the Atlantic water. Furthermore, CT is a conservative property with respect to changes in temperature, so the authors would need to be clearer about the processes that make surface CT increase with temperature.

We remade the figure with more contour levels to make the gradients clearer. We agree that the increasing CT with decreasing temperature is not very clear, and we therefore now write “Within the Atlantic water there is a *tendency* of increasing CT with decreasing temperature” We added a sentence about the underlying process.

lines 338-339: Please explain why the CT of polar waters is lower than that of Atlantic waters?

Done

lines 341-343: I have some problems with this statement:
* The statement that "pH decreases with increasing temperature, salinity, CT and AT (... Fig. S16)" is not clear in Fig S16. For Atlantic waters there seems little correlation, while for polar waters, there are vastly different relationships

From the feedback we got from the reviewers on this part, we understood that the text on lines 340-345 in the last version of the manuscript adds little value to the manuscript, and we decided to remove it. We moved table 3 to the supplementary material. We think that the text related to Table 4 and figure 3 (Table 3 and Figure 3 in the revised manuscript) demonstrates what we want to show here.

* The use of "strong" and "significant" in this sentence does not seem to be supported by Fig S16.

This text has been removed as described above.

* The correlations of pH vs salinity, pH vs CT, and pH vs AT look extremely similar, suggesting that the latter two are also driven by changes in salinity.

It does not have to be a direct effect of salinity, but can be a result of the contrasting properties of the polar waters and the Atlantic Water. This text has been removed as described above.

lines 343-345:

* The authors seem to infer that if the drivers were perfectly orthogonal it would be possible to calculate the contributions of each driver from such correlation plots (Fig. S16). But even in this ideal case, I do not think that is possible.

This text has been removed as described above.

* Is not the conclusion at the end of the sentence something we could not already say before this analysis? I think the authors could do better than make this weak statement.

We agree that the analysis on lines 340-345 does not add much to the story, and we therefore removed it. We think that the text related to Table 4 and figure 3 demonstrates what we want to show here.

Line 349:
- change "to the picture" to "contributions"

Done

- the parenthetical statement "(temperature, CT and AT explain all together 89%)" is redundant. It should be deleted.

Done

Line 361: remove the comma

This sentence has been removed after the suggestion of reviewer 1 to shorten the paragraph.

line 362: change "relation" to "relationship"

Also this sentence has been removed. Reviewer 1 pointed out that pH and pCO2 always correlate. We therefore thought that it does not bring anything new by going into these details.

line 371: delete "are counteracting, they"

This part has been removed to shorten the paragraph.

line 372: "solubility of OmegaAr"? Do the authors mean "solubility product"?

Yes, this has been corrected.

line 390: What is "Fig. 4.1"? The same problem is found in the title to Table 5.

Thank you for spotting this. It was related to a problem with the labelling in our latex-file. This has been fixed.

line 413:
* "this period of time"? Please be more specific.

We followed the suggestion and have now specified the time.

* "we detect trends in the uncertainties". Some detail is needed.

We added some numbers on the trends.

line 442: The authors refer to "The strong positive trend in pCO2", but it seems they actually mean to refer to Delta pCO2. Unclear.

Thank you for spotting this, we have clarified it.

line 469 and 471: missing Figure number

This was also related to the problem with the labelling in our latex-file. It has been fixed.

line 497: "salinity decomposition" is ambiguous. I would recommend to change that to ""more detailed freshwater decomposition".

We have put the freshwater decomposition figure in the main manuscript. This sentence has been removed.

lines 498-500: I disagree with the authors statement here that seems to say that the uncertainties of freshwater component compromise the assessment of the so-called biogeochemical component. The uncertainties concerning the freshwater component may be large but the freshwater contributions of AT and CT counterbalance one another, making their combined contribution negligible. The uncertainties will typically affect freshwater contributions from AT and CT in equal an opposite ways and thus they don't really matter for the BGC component.

Thank you for pointing this out. We have removed these lines.

line 516: What does "reduces/amplify" mean? Please use the word or words you mean instead of the slash, which is ambiguous.

Here we referred to the effect of increasing alkalinity and salinity. We agree that it was not clear. The paragraph has been revised to make it clearer.

line 535: change "4.5" to "RCP4.5"

It has been changed to esmRCP4.5.

line 549: It would be better to begin the sentence with "Assuming a Redfield ratio of O2:C = 132:106," We adopted the suggestion.

line 553: The errors in the way the authors approximate the partial derivatives (sensitivities) are not mentioned, but are probably as large or larger than the other factors mentioned.

Thank you for this remark, we added as one of the possible reasons behind the residuals.

line 589: ambiguous. What does upper bound mean? shallower or deeper?

We meant shallower, we edited the text to make this clear.

SUMMARY AND CONCLUSIONS

It is not clear in the Discussion nor in the Conclusions how the authors have filled the gap they pointed out in the Introduction, namely the advances made by this study relative to the previous work by Skogan et al. Indeed the findings of Skogan et al. were never really mentioned, so we don't have anything to compare this work to. The authors need to bring this up in the Conclusions. To what extent do the findings of this study confirm or contradict previous work. To what extent do they embark on completely new territory. How has our knowledge been incremented by this study. The Conclusions section mostly just reiterates what was said in previous sections and ends with a caveat, a detail about a minor part of the work. The big picture seems to be missing.

We agree that this is an important part that is lacking in the manuscript. We worked on this by revising the last two paragraphs of the introduction and the conclusions. Specifically, in the conclusions we now put our results into perspective to previous work.

line 611: What is meant by "careful estimates"? Do the authors mean "conservative estimates"? If so, I am surprised. Since NorESM1-ME tends to simulate an ASH that is too shallow, does that not imply a bias in the opposite direction, i.e., corals will be projected to be exposed too soon to undersaturated waters. This it seems not a conservative estimate but rather the opposite.

We agree that the formulation of this sentence was not very good. We reformulated it as follows:

Because NorESM1-ME tends to simulate relatively strong drops in pH and shallow saturation horizons in comparison to our ESM-ensemble for esmRCP8.5, our estimated aragonite saturation horizons for esmRCP2.6 and esmRCP4.5 can be considered to be in the shallow bound of possible future states.

line 614-615: The authors state "The effects of increasing CT is slightly opposed by increasing AT , which partly comes as a result of the increasing salinities". But I find this statement puzzling. Based on their freshwater Taylor expansion (Fig. S19), the change in salinity has an equal and opposite effect on CT and AT. These effects cancel out. The same figure shows that it is the changes in the biogeochemical AT term that partly counterbalances the changes in the biogeochemical CT term. Salinity changes have nothing to do with that.

This is correct, thank you for pointing this out! . It cannot come as a result of increasing salinities, and must be a result of biogeochemical processes. We have revised the sentence.

line 616: I am confused by this phrase: "the impact of temperature in the surface is ambiguous, and even shows a cooling in some places". What is meant by "ambiguous"? Could that phrase be replaced by something simpler such as "at the surface, there is cooling in some places."

We now write that there is no clear temperature change in the upper 200m.

line 629: change the first "to" to "too"

Done

FIGURES

Figure 10: change "rates of" to "the". To the GLODAP estimate for the present, the authors added the change, not the rates of change.

Done

Fig. 11: I think that this figure and the simple Taylor expansion should NOT be included in this paper. It offers no added value and only makes the paper longer. Only the freshwater Taylor expansion should be shown and discussed. Yes the latter is slightly more complex, but it is better because it is able to show that the combined contribution of salinity-driven changes in AT and CT terms is negligible for pH. Focusing on the simple Taylor expansion will confuse readers (see my comment just above concerning lines 614-615). Thus Fig. 11 should be replaced with Fig. S19, and only on the freshwater Taylor expansion should be shown. There is no reason to show the simple Taylor expansion if the freshwater Taylor expansion is presented.

We agree that the freshwater decomposition gives additional information that is worth taking up in the main manuscript. Although the freshwater component is negligible for the pH changes, this decomposition helps us understand what lies behind the changes in CT and AT, and in particular it enlightens the regional differences in CT and AT changes obtained from the model simulations. We have therefore made the following changes in the manuscript:

- we replaced figure 11 with S19 to show the freshwater decomposition, as suggested
- we added one figure showing the decomposition of the modelled surface changes in CT and AT into their respective freshwater and biogeochemical parts. We thought that adding this in figure 12 would result in too many subplots. Because the freshwater forcing mainly is at the surface we did not produce a similar figure for the cross-sections.

- We moved equations 1-4 in the supplementary material to the main manuscript.

Thank you for taking up this point!

Fig. S14: What do the colors mean. There is no information about color codes in the caption nor the legend.

We added a description in the caption.

Fig. S16: in the legend, please change "single grid" to "one grid cell".

Done

TABLES

Table 4: delete the 2nd header line that is in the middle of the table

Done. We also simplified the table by just having one header-row for the “Drivers”.

Tables in the Supplement should include an S before the number, but some do not do so

Done

REFERENCES

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93(4), 485-498.