Dear Dr. Jean-Pierre Gattuso,

We are happy to submit a revised version of our manuscript “Acidification of the Nordic Seas”. We want to thank you and reviewer 4 for this revision round, which has helped clarifying some unclear points of the manuscript related to the emission-driven runs. We have addressed all the comments of the reviewer as described on the following pages.

Sincerely,
Filippa Fransner and co-authors

I was asked by the editor to comment on the respective merits of emission- versus concentration-driven simulations to address the specific objective to document acidification of the Nordic Seas. Here are my comments, which also address additional points.

We are grateful for the reviewer's work and suggestions that identify weak and unclear points in the manuscript! We have addressed all the points as described below.

First, this study mainly uses model simulations from one single model (i.e., NorESM1-ME). I think this needs to be stated upfront in the abstract. As currently written, the abstract gives the impression that more than one model was used, but that is not the case.

Thank you for this comment! We fully agree with this and have clarified it in the abstract.

Second, the authors have decided to use emission-driven runs and not concentration-driven runs from the NorESM1-ME model. Both emission-driven and concentration-driven simulations have been used in the literature to analyze changes in the Earth system. For ocean acidification studies, however, concentration-driven simulations are usually used especially for studies analyzing past changes (although not exclusively). In my opinion, it is ok using emission driven simulations when analyzing future changes in acidity in the Nordic Seas. However, the choice of using emission driven simulations is a bit problematic when analyzing past changes (i.e., over the historical period). The atmospheric CO_2 over the past 170 years is well known but the simulated atmospheric CO_2 over the historical period in these emission driven runs deviates from the observed one (in this study by 14 ppm in year 2005 according to Table S1), and therefore also the historical changes in acidity. This is especially problematic for Figure 4a and part of Figure 4b, but also for Figure 5 and Figure 6, and should be addressed.

We agree with the reviewer that it is important to analyze the impact of this deviation in the atmospheric CO_2 on simulated past changes in pH. In our previous version of the manuscript, we did this briefly in section 5.2 (first paragraph), where we referred to Table S1. There, we estimated that the deviation in atmospheric CO_2 in NorESM-ME causes a drop in pH that is 0.01 stronger than expected during the period 1850-2005 under actual atmospheric pCO_2 change. This is, however, on the same order of magnitude as the uncertainty in the GLODAPv2 pre-industrial estimate, and it is one order of magnitude less than the actual pH change from 1850 to 2005 (approximately 0.1) Its impact on our results is therefore minor.

To make this analysis more prominent in the manuscript, we now discuss Table S1 earlier, in Section 3.3, where we introduce our choice of emission-driven models. We also mention the impact of this deviation in atmospheric CO_2 on pH in all instances where we illustrate and discuss the historical change: in the figure captions of Figure 4, 5 and 6, and at line 393-395 in section 5.2.

Third, there are only a subset of CMIP5 emission driven simulations available and only for esmRCP8.5. This ‘caveat’ is mentioned in the manuscript. However, the comparison between the NorESM1-ME simulations and the CMIP5 emission driven simulations is currently buried in the supplementary material. I think the paper would become more comprehensive when including the CMIP5 range also in the figures of the main text, for example as vertical bars in Figure 4 (e.g. in year 2100, but also in year 1850 and year 2005). The reader would immediately see where the NorESM1-ME stands in comparison with other comprehensive CMIP5 ESMs. In addition, the paper would benefit from a paragraph, where the comparison with CMIP5 ESMs is described. For example, when looking at Figure S5 and Table S6 it seems to me that the NorESM1-ME model simulates the largest decrease in pH of all CMIP5 models in the upper 2000m. This might be due to stronger simulated atmospheric CO_2 increases than in other models. When using the same forcing (concentration driven runs), the model seems to be more in line with other CMIP5 models. Is this true? In any case, this needs to be discussed in the main text, maybe even highlighted in the abstract.

Thank you for these suggestions! We included the CMIP5 model ranges for 1870, 2005 and 2099 in Figure 4 (We could not do it for 1850 and 2100 because the simulations with some models did not cover these years). We also expanded our text on the inter-model spread in section 5.2 (lines 410-415) where we now mention that NorESM1-ME simulates a stronger decrease in pH than the other models, which most likely is a result of a stronger increase in atmospheric CO_2. As the reviewer suggested, the larger spread in the emission driven runs compared to the concentration-driven ones is most likely a result of the inter-model spread in atmospheric CO_2, which we now mention. A full analysis of the differences between the emission-driven and concentration-driven runs is, however, beyond the scope of this paper.

Forth, why are the emission driven simulations so different than the concentration driven simulations below 3000m (CMIP5 as well as NorESM1-ME; Figure S5g-i)?

Thank you for noticing this! This spread was a result of a mistake in our data handling, which we apologise for. After correcting this, the inter-model spread in the deep is very similar in the emission-and concentration-driven runs. The corresponding figure has been revised accordingly.