Review: Assessment of simulations of Arctic sea ice the PlioMIP models.

The article “Assessment of simulations of Arctic sea ice in the PioMIP models” by F.W. Howell et al. analyzes sea ice representation in pre-industrial and mid-Pliocene climates in an ensemble of 8 different global coupled models. The inter-model spread is discussed and possible relations between pre-industrial and Pliocene simulations are investigated.

Since data are very sparse for the Pliocene and basically no robust information about sea ice distribution in the Pliocene is available, we do not gain much information about model performance by simultaneously comparing pre-industrial and mid-Pliocene model simulations. There are several studies existing, which either evaluate a larger number of global models for present day climate (better suited for model evaluation than pre-industrial) or relate behavior of ice properties in present day and future studies in a much more comprehensive way than this study does. Thus, I appreciate that this new version of the article focus more on sea ice conditions under pre-industrial and mid-Pliocene conditions and less on a potential ranking of model’s ability to simulate sea ice.

Unfortunately, large parts of the manuscript are very descriptive and mentioning trivialities; it is not necessary to describe in detail the ice thickness pattern of every single participating model. Instead more analysis should focus on the causes for the model spread and the causes for different pre-industrial – Pliocene differences. Why is the seasonal amplitude of ice extent different? How does the atmospheric circulation look like in the different models both in pre-industrial and pliocene? How does this affect the sea ice thickness distribution in the different models? How does the atmospheric and oceanic heat transport into the Arctic look like and how might it affect sea ice representation?

The argumentation that the representation of NAO or AMO could not be investigated at all due to too short time series is not entirely convincing. I agree 30 years are short but NAO and AMO are strongly affecting sea ice and if we can not say anything about NAO and AMO, then it might also be difficult to make robust conclusions on the sea ice conditions themselves. Here, the question arises why only 30 years from the 500-1000 year simulations are used for the analyses? To make results more robust, a longer period should be used.

I am still not convinced by using CV as measure for the variability even it has been used by Stroeve et al. 2014 – your CV is also not exactly what Stroeve et al. (2014) used and they used it in a somewhat different context. CV assumes that sea ice concentration and thickness variability should linearly increase with ice concentration and thickness, respectively. To my knowledge, no one ever showed this.

In any case, CV can not be a reliable measure for ice variations if ice thickness/ concentration are near 0 as it is in four of your Pliocene-summer simulations. Thus, the conclusions that especially summer ice variability increases in Pliocene compared to the pre-industrial simulations should not be taken based only on CV-values. I would even call this statement wrong.

CV could be used if a) the “normal” standard deviation/ variance of the models is shown as comparison and b) evidence is shown that we can assume a linear growth of ice concentration and ice thickness

variability with increasing ice concentration and thickness, respectively (e.g. cite an article that shows this).

Specific comments:

1. Line 6: I would say that the statement that the model spread is 3 times larger in summer than in the rest of the year is just wrong: Figure 13 shows that ice varies between 0.7 and 2.7m thickness in winter (quite evenly distributed) and between 0 and 2 m in summer (5 out of 8 models between 0 and 0.3 m, the other three about 1, 1.7 and 2.2 m). See discussion above.

2. Line 40 and following, line 306-310: I would suggest making a single section discussing model-setup and the experiments a bit more in detail. What is the difference to the CMIP5-model versions? How have sea ice and SST been prescribed in Pliocene-Experiment 1 (AGCM-simulations) if observations are so uncertain?

3. Line 54: additional to what?

4. Line 54: “reduces”

5. Lines 59-61: Are you sure all models provide ice thickness as grid-box mean and not as mean over the ice-covered part only (which is the case in most CMIP5 models)?

6. Line 83: Figure 1: Extend the plotted area to the south – if this figure does not show the position of the ice edges, it is useless. It does not matter that the other figures show a different area, you could indicate this in the figure caption.

7. Lines 92-98: You should also indicate the observed amplitude for comparison.

8. Line 106: I do not see any ice thickness anomaly at 0E, 80-90N, to be clear it is maybe better to say 180E, 80-90N instead of “Greenwich meridian”.

9. Line 110: mention that ice thickness is likely overestimated at the Siberian coast in MIROC, COSMOS.

10. Lines 111-113: This is not needed. However, if you want to mention that sea ice is thinner in Nordic Seas compared to Siberian coast you should also shortly mention why this is the case.

11. Lines 120-143: This discussion is too detailed and difficult to read. The section should be shortened and the most important points mentioned. The reader can find details in the figure.

12. Line 164: In contrast to this, in the discussion it seems to be stated that CCSM shows a larger ice extent in the mid-Pliocene – please check.

13. Line 187: delete one “amplitude”

14. Line 205: “Many” sounds a bit strange with a total of 8 models whereas 4 do not show any ice in summer.

15. Figure 7: As figure 1: please show a larger area in the plot.

16. Line 218/219: difficult to judge from the Figures 7 and 8, there are very large differences across models in winter as well and Figure 9 shows about the same max-min difference in winter compared to summer. As mentioned earlier I do not think CV should be used and comparing Figures 9 and 12 clearly indicates why. Just from looking at Figure 9: do you really want to suggest that model spread is 3 times larger in summer than in winter?

17. Line 284-296: Please make clearer if this is a summary from Shu et al. 2015. Please specify which time period Shu et al. (2015) analyzed. I would suggest adding these numbers, at least annual mean ice extent to table 2.

18. Lines 299-302: I do not understand this sentence: Please clarify.

19. Lines 303-305: This sentence is not clear.

20. Lines 326-327: Of course they are not the same: maybe better: “…vary strongly: The summer sea ice …”

21. Line 324-325: maybe better: “almost ice free” or “ice free in late summer”

22. Line 326-327: sounds like HadCM3 simulates summer sea ice in the entire Arctic Basin. Ok, not really wrong but actually ice concentration is very low along all ice edges.

23. Line 405: Please specify what you mean with “CMIP5”. Here and elsewhere CMIP5 is compared to pre-industrial simulations: Pre-industrial simulations are also part of CMIP5 as historical and future and many more simulations are. It seems you mean a certain time-period with CMIP5 (historical, satellite period …)?

24. Line 435-440: SST and SAT are not necessarily drivers of sea ice variations but could also be driven by ice variations: One reason for better correlation in Pliocene could be that larger parts of the ocean north of 60N are ice-free for longer periods in the year and could thus warm up much stronger than in the pre-industrial period. The longer, the ice-free period, the more the ocean can warm. In the pre-industrial period instead, summer SST and SAT in the Arctic will almost be very near melting temperature of ice, it does not matter, if sea ice concentration is 100% or 50% in a certain gridbox or smaller region: as long as some ice is left, the ocean can hardly warm up.

25. Lines 453ff: The albedo discussion would fit better into section 4.3.1 “Influence of sea ice models”

26. Lines 521-526: According to the introduction, it is debated if sea ice was seasonal or perennial? Yes, HadCM3 agrees with the findings of perennial Arctic sea ice but the other models agree with findings of Cronin et al., Moran et al and Poyak et al.. Is there any particular reason to believe more in the perennial assumption? Furthermore, HadCM3 shows a very unrealistic sea ice concentration distribution in both

pre-industrial and mid-Pliocene summer, thus even if HadCM3, probably by chance, keeps the points of Knees (2014) at 80N ice-free year around, we can be quite sure that Arctic sea ice distribution will not look like HadCM3. All the years with very low observed ice concentrations (e.g. 2007, 2012) still showed the thickest ice with highest concentration north of Greenland and the Canadian Archipelago.