Response to Referee Prof. Flato

Format key:

Normal text is considered general discussion

Italic text quotes the Referee

Bold text indicates additions to the manuscript

We wish to thank Referee Prof. Flato for his quick response and constructive review of our submission.

Prof. Flato suggests two separate point of consideration:

*"I have a slight quibble with the use of the word 'uncertainty' in this paper, and in particular the extent to which reducing 'spread' is equivalent to reducing 'uncertainty'. This may be largely a semantic issue, but it is not obvious to me that reducing spread *necessarily* reduces uncertainty (in the sense of the confidence one has, or should have, in a prediction or projection). Spread is of course directly related to uncertainty, and the partitioning suggested by Hawkins and Sutton yields considerable insight into the sources of uncertainty and how these change over time. But I think one has to be a bit careful in equating reduced spread with reduced uncertainty (and by extension, enhanced confidence) as is done here. One can readily construct schemes that reduce spread (e.g. discarding all models but one), but don't really reduce uncertainty. Perhaps a few sentences on this topic could be added?"*

> Firstly, our use of the word uncertainty in this context is perhaps a little enthusiastic as we do indeed equate a reduction in model 'spread' with a reduction in 'uncertainty'. Prof. Flato states a simple example where this would not be a valid statement which we consider a helpful point from which to adjust and clarify our terminology. But also note the additional comments in our reply to Dr Massonnet about testing our uncertainty estimates which, within the limitation of the models examined, appear largely reliable.
>
> We will add **potential** to "reducing uncertainty" and "increased confidence" to highlight our slight hesitation with such claims. We will clarify our use of the word 'uncertainty' with the following sentences located in Sect. 4.4 of the manuscript:
>
> **An additional source of uncertainty that we neglect here is the PIOMAS calibration uncertainty emerging from the choice of atmospheric reanalysis and ice model tuning. This could be assessed by sampling the different versions of the PIOMAS reanalysis described in Lindsay et al. (2014).**
>
> **In the following sections, we equate reducing model spread with reduced uncertainty. While some of the outlier simulations of SIT are now more similar to the multi-model mean, this doesn't necessarily equate to reduction in uncertainty. The initial selection of GCMs may not have been representative, or**

all of the GCMs from CMIP5 may have some inherent systematic biases, reducing the spread of which wouldn't help sample future observations.

..

Secondly, Prof. Flato rightly points out two points of confusion in Fig. 3:

"I did note two things related to Figure 3 however: – the caption states that 'ice-free' is defined as the "first occurrence ... below 0.15m", but the legend gives a range of years of 'ice-free year'. I didn't understand this.

– the legend indicates no change in 'ice-free' year for the high-mean (blue) example when the multiplicative correction is applied (compare Fig 3a and 3c) even though the curve is obviously shifted downward. I suspect a typo in the legend. The same applies to Fig 3b and 3d where again the ice-free year for the blue curve is unchanged."

The reason for the confusion in the first point is primarily due to inadequate explanation of what the dates below 'ice-free' represent in this figure. This is rectified by adding the following sentence to the caption:

"Ice-free" is here defined as the first occurrence of an ensemble member below 0.15 m. **Shown is the "ice-free" ensemble range, i.e. the year of the first ensemble member to be "ice-free" to the last ensemble member to be "ice-free".**

Secondly Prof. Flato also notes that the "ice-free" statistics are identical on comparing Fig 3a with 3c, and Fig 3b with 3d. This is in fact not a typo and a true representation of our "ice-free" criterion, this is partly coincidence and partly due to how the four correction methods shown manipulate the time series. While Prof. Flato rightly points out that the curves have obviously shifted, the "ice-free" date remains the same. This is shown by examining when the thin coloured lines cross 0.15 m. This is an important point that Prof. Flato observes, the following paragraph is added to Sect. 3.4 to highlight this behaviour:

Comparing the ensemble range in projected ice-free date between the correction methods it is apparent that although the shapes of time-series have qualitatively changed this does not always result in a different range in projected ice-free date. For example on comparing the high mean – high variance GCM (blue) between (a) to (c) and (b) to (d); this is partly coincidence and partly due to how the four correction methods shown manipulate the time series. The MAVRIC method (e) results in a unique set of ice-free dates. This is an important attribute that the MAVRIC method displays, as the ice-free date is of vital importance to stakeholders in the Arctic and more basic methods of bias correction fail to appropriately impact on this parameter.

We again thank Referee Prof. Flato for his quick response and constructive review of our submission.

Kind Regards,

N. Melia, K. Haines, and E. Hawkins