# *Interactive comment on* "Assessment of Arctic and Antarctic Sea Ice Predictability in CMIP5 Decadal Hindcasts" *by* C.-Y. Yang et al.

Anonymous Referee #1

Received and published: 18 June 2016

The authors present an evaluation of sea ice covers in decadal hindcasts run with 11 models for CMIP5. They consider September Arctic and Antarctic sea ice concentration and extent, compares with observations over the period 1980-2006 (+10 years), and evaluate the ability of the models (and the multi-model ensemble mean) to capture the raw fields, the trends, regional features and anomalies relative to the trends.

Presenting an evaluation of the CMIP5 decadal experiments is a good thing per se, thus this paper is useful and timely. It actually reads well and presents interesting results. This draft shows that the predictive skill at decadal time scale is poor for both Arctic and Antarctic September sea ice extent. Yet, some skill arises in the Arctic due to the decreasing trend (anthropogenically forced) that models are able to reproduce. In the Antarctic, the lack of predictive skill is attributed to the models not being able to capture the recent increasing trend.

It would have been interesting to compare predictive skill for maximum sea ice covers in both hemispheres, ie September Antarctic sea ice and March Arctic sea ice. Germe et al (2014) wrote a very interesting study on the predictability of winter Arctic sea ice based on a CMIP5 hindcast (not used in the present draft), which gave some clues on possible processes to investigate. September Arctic sea ice extent is in the spotlight, but few skill should be expected from CMIP5 decadal hindcasts besides the trend, due to the date of initialization early in the year (around January 1 for most systems I guess) and the predictability barrier documented in Day et al (2014b).

I noticed a few major points that should be addressed before possible publication : some methodological issues, some unclear figures, and a poor discussion of the results. These points are marked with an asterisk (*), and can be considered as "major comments".

1.Introduction

L51: "ie" -> "eg"

L51-52: the sentence sounds weird. Does it mean that sea ice treatment is important to reduced the remaining "large uncertainties for decadal climate prediction"?

L111: "implemented an experiment" -> "implemented an experimental framework..."

L113: "validation" -> "evaluation"

2.Models and data

*How did the authors select those 11 models ? Why not using e.g. CNRM-CM5 (documented in Germe et al., 2014) or EC-Earth?

*In the first paragraph, the authors explain that there are two approaches for initialization, full-field and anomaly initialization. On L163, they introduce direct vs indirect initializations, which are not defined here.

*There is no information on the start date (ie: time of the year when the hindcast

are initialized). Are all models initialized in early January? Earlier? During January, the Antarctic sea ice cover is low : could it be a reason for the lack of skill (e.g. no persistence possible)?

L139-141: Let's be honest : the main reason why we don't focus on the Antarctic sea ice minimum is because we know the models poorly simulate this minimum. See for instance Turner et al., 2013.

3.Prediction skill of CMIP5 decadal hindcasts

*A discussion on the significance of the predicted trends is missing, as those trends seem to be calculated based on only 6 points (this is actually not well explained)...

L163 : see above : what is "direct and indirect full initialization"?

L165 : what is "improved initialization"? Compared to what?

L172-174 : the formula is wrong. Pbar and Obar should be divided by n to denote an average.

L200-202 : the contradiction in this sentence strikes me.

L210/212 : "lead-time" or "lead time"? The authors should be consistent.

L212-213, "those longer lead times are weighted towards inclusion of more recent years in the observations with accelerated decline of Arctic sea ice". I don't fully understand this sentence.

*L235-239 : it seems to me that a few models are positively biased in the Summer (ie summer Arctic sea ice extent too high), and that is can be due to sea ice still present in the nordic seas. Could it also be due to sea ice presence too further south in the Pacific sector, namely in the Bering sea? Shouldn't the authors add the Bering Sea to their "Pacific" sector?

*L245 : does it mean that the sea ice extent for all lead time equals sea ice extent at

lead time 0? It is not fully clear, as hindcasts are likely initialized near January, and not in September...

L261-263 : so what? This very interesting discussion misses a conclusion.

L297 : I don't see that.

4.Discussion and conclusion

The authors should first remind of the main conclusions of their study, then start the discussion.

L333 : underestimate -> underestimation ; overestimate -> overestimation

*L349-352 : are we sure of the causality? How is the AMOC in uninitialized simulations with CFSv2?

L350-352 : so what? Would anomaly initialization avoid such problem (instead of full initialization)?

L354 : AASW -> AABW

L363-365 : see comment above (349-352).

L425 : "but also variables (...) THAT influence ..."

L434 and sq : the authors should explain better what they mean by "direct and indirect" initialization.

L446 and sq : the authors may wish to refer to the work by Chevallier et al. (2016) on an intercomparison of Arctic sea ice cover in reanalyses, in order to illustrate that there is still some way to go.

Tables and Figures

*Table 1 : the authors should highlight clearly what models used anomaly, direct or indirect full field initialization.

*Figure 2 (7) : I have a methodological concern in figures 2 and 7 : shouldn't the authors show debiased sea ice extent for models that used anomaly initialization?

*Figure 3 : I see dashed areas on the figures that are not explained in the caption. Could the authors add a few word on that, or remove those? This dashed areas sometimes spead much southward than an acceptable model would do for September Arctic sea ice cover. . .

*Figure 4 (9) : the authors should show the 95% confidence intervals.

Figure 12 / 13 : I don't understand if the curves represent all the hindcasts, or an average over the all hindcasts. The caption ("thin gray line represents each ensemble member...") is misleading, and I am not sure there isn't any methodological error here (average over all "member 1", average over all "member 2"...?).

References

Chevallier, M. and coauthors, 2016, Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project, Climate Dynamics, 1-30.