

Interactive comment on “The Arctic Predictability and Prediction on Seasonal-to-Interannual TimEscales (APPOSITE) data set” by J. J. Day et al.

Anonymous Referee #1

Received and published: 23 November 2015

In this paper a multi-model protocol for analysing potential model predictability is introduced, focusing on the potential predictability of the Arctic sea ice conditions on the seasonal to interannual timescale. The setup of the ensemble simulations is explained as well as the diagnostics used to analyse potential predictability of Arctic sea ice extent and volume. Seven different models have contributed to create a dataset following the basic guidelines of this protocol, with some difference in the more specific details such as ensemble size and number of ensemble start dates. The results for the ensembles of four of these models regarding potential Arctic sea ice predictability have previously been discussed in a paper by Tietsche et al. (2014), while the results for the remaining three models are added to the discussion for this paper.

C3062

In general I appreciate the effort of the authors to make the data available to the broader scientific community and to use this publication as a reference for the setup of the experiment protocol. Analysing potential predictability and the differences therein between GCMs is certainly an important area of research, especially as a tool to inform seasonal prediction systems of the feasibility of future improvements. The paper is generally well written and the structure is straight forward. While I appreciate the authors' choice to keep this publication short and concise, I do have some comments that might increase the length of the paper quite a bit. My main point of critique is that the paper is very close to the previous publication by Tietsche et al. (2014) without presenting a more detailed description of the experimental setup, and without discussing the new results equally detailed as the previous study. Since both aspects are the main points of this paper, they should be extended, still keeping them as separate aspects of the same publication, i.e. first the discussion of the protocol, then the application to the newly contributed models, highlighting the importance of both.

General comments

As a first comment and to repeat my question of the summary, could the authors be more specific regarding the focus of this paper and how it differs from the Tietsche et al. (2014) publication. I assume you want to equally focus on the results for the additional three models as well as on the general setup of the protocol. But at the moment I would claim that both parts are a bit too short and not very detailed.

Some more specific examples regarding the experimental setup:

When you write about the high, low and medium sea ice states used for initialisation, how is that reflected in the actual ensemble start dates? Does this relate to the sea ice volume, the sea ice area or average sea ice thickness? Are they separated in some way in the archiving structure? Are you trying to estimate the impacts of different initial conditions by this approach, even though some models only have 8 different start dates, which would make it difficult to actually assess differences in the predictability

C3063

caused by the initial state?

When you say “well spaced” (page 8815, line 18) how is this defined? Was there a minimum spacing between successive start dates that you have generally defined for all models to insure independence of the initial state?

How was the length of the control run defined? Different models have different spin up times and might take longer to equilibrate. After only 100 years I wouldn't think any model has really equilibrated, as can be seen by the strong drift of most of the models. Could you comment on some of these details, stating advantages and disadvantages of the choices you had to make to generate this dataset. Also, in this context, the time axis for the panels in Figure 1 doesn't make much sense to me. The start date of each model control seems more or less random, even though the text reads they started from (the same?) static state oceanic depth profile.

Were the SST perturbations applied globally, also in areas of sea ice cover?

Regarding the two metrics, were they applied to detrended monthly means? If so, was the detrending based on the control or all ensemble members? It would simplify the explanations for the metrics if you would actually expand the expectation value as was done in Collins (2002), also to show which normalization you chose (what is sigma?).

What kind of significance test was applied to the ACC?

Are there any specific plans to extend this dataset, i.e. to include more models? Or to use this dataset for other predictability studies?

Some more specific examples regarding the results:

The sea ice models in this study differ in many aspects. Could you comment a bit on how this affects the results? For example, do models with similar albedo and melt pond parametrizations produce similar results, or do models with similar sea ice dynamics (number of sea ice classes and so on) produce similar mean states and climate variability? I know this is a difficult questions, since the other model components show

C3064

significant differences as well. However, it would be interesting to know whether some systematic differences can be identified.

Could you please expand the paragraph about the mean state and climate variability. For one, it is not surprising that the mean states of the models are different compared to the mean state of the observations, which have been recorded over a shorter period of time and under transient forcing conditions. Furthermore, could you comment on how model variability and mean state affect the predictability metrics.

What are the consequences of the different drifts in the models? Do you expect a more equilibrated model to provide a more accurate estimate of potential predictability?

Why didn't you apply any of the spatial predictability metrics which were used by Tietche et al. (2014)? What about the other start dates provided, especially January? Since the extended results of this paper are mentioned as one of the two major contributions of this study, it would be nice if the paragraphs about the model results (page 8818) could be expanded, providing more details on the differences and similarities in predictability between the models and possible reasons for that.

Page 8818, lines 12-15: How does this relate to the results of the current study?

Page 8818, line 23: There is always a chance that you remove internal variability by detrending, also for a longer timeseries. It is just less likely.

Page 8818, lines 26-27, and page 8819, lines 1-3: This paragraph is difficult to read. Maybe you could break up the sentences.

Page 8819, lines 6-7: The differences of the mean state and variability between models and observations wasn't discussed in any detail.

Page 8819, line 17: Not really true for E6F (early loss of predictability for sea ice volume; no re-emergence of predictability for NRMSE).

Minor comments:

C3065

Page 8811, line 16: Change to “Unprecedented”, “opportunities”, “businesses”.

Page 8811, line 17: Change to “but has also”.

Page 8811, line 23: “appreciation”.

Page 8812, line 1: What do you mean by “significantly skillful”? Could you also give a reference here?

Page 8812, lines 9-11: Please rephrase this sentence. Be more specific about this “fundamental limit”, which has different timescales for the atmosphere and the sea ice.

Page 8812, lines 20-21: Please expand this. What are the disadvantages of potential predictability studies? How does model uncertainty affect predictability estimates?

Page 8813, line 5: Change to “. . . climate variables as well. In order. . .”.

Page 8813, line 10: Differences in design such as?

Page 8813, line 12: Differences in the results such as?

Page 8813, lines 13-16: Again, could you name some of the differences, either here or before?

Page 8814, line 22: Change to “sea ice”.

Page 8815, line 1: Change to “distribution, as well as”.

Page 8815, lines 11-13: Can you quantify this/be more specific? Does this have consequences for summer sea ice predictability when it comes to different model mean states?

Page 8815, line 20: Change to “depending on”.

Page 8816, line 8: Remove comma at the end.

Page 8816, line 21: Change to “inter-model”.

C3066

Page 8818: Mention Figure 5 again, after first sentence of 3.2 and 3.3.

Page 8819, line 14: Change to “interannual”.

Page 8820, line 7: Change to “constraints.”.

Page 8820, lines 8-11: Could you give a reference here?

Page 8820, line 23: Change to “submodel&frequency”.

Page 8820, line 23 onwards: Check for text size and font here and on the next page.

Page 8820, line 25: Is it “1” (this line) or “r1” (next page, line 1).

Figure 2 and 3: Is the average taken over the entire simulation length or only for the years after the spin-up?

Figure 4: Mention detrending in caption.

Interactive comment on Geosci. Model Dev. Discuss., 8, 8809, 2015.

C3067