Clim. Past Discuss., 8, C2867–C2870, 2012
www.clim-past-discuss.net/8/C2867/2012/

Climate of the Past Discussions

Interactive comment on "Skill and reliability of climate model ensembles at the Last Glacial Maximum and mid Holocene" *by* J. C. Hargreaves et al.

Anonymous Referee #2

Received and published: 18 December 2012

Synopsis: This paper presents a comparison between multiple ensembles of paleoclimate simulations and two well-sampled periods with surface temperature proxy-records in the Last Glacial Maximum and the mid-Holocene. The two periods have adequate spatial sampling to compare large-scale response of the climate models with the temperature records. The paper uses three techniques to assess reliability and skill of the ensembles at reproducing the past climate changes. There are two reliability metrics with one metric based on rank histograms (i.e., the location of the observed data within the ensemble spread) and the second based on Taylor diagrams of pattern-based statistics. The skill scores are based on two reference forecast systems. The first reference forecast system predicts a constant temperature change pattern with

equal changes everywhere and the second predicts a spatially resolved pattern, which to first order, captures land-sea contrast and polar amplification. The models are shown to perform well for the LGM with both reliability and skill scores while the mid-Holocene results are generally negative or non-conclusive. The lack of skill for the mid-Holocene is not well-understood with potential problems in the analyses including lack of SST data over much of the ocean, missing forcings, or missing feedbacks in the models. The paper is interesting and significant in its application of techniques typically used in weather forecasting to assess model quality in making predictions.

Overall comments: The paper is well written and well organized and there are very few major comments. There are a few places where additional material would be helpful to the reader as discussed below. A significant issue is that the figure captions are often not descriptive of the figures as stand alone text, which forces the reader to go back and forth between the text and the figures.

Recommendation: Accept with minor revisions.

Minor comments and technical corrections/suggestions:

Abstract:

- Please add the dates for the Last Glacial Maximum and mid-Holocene periods.

- Please define what "regional scales" means in this paper. As a comment on future regional climate change, there is a big difference between 100km and 1000km but both could be considered regional given they are sub-continental scales.

Page 3483, line 11. It would be useful to have a slightly longer discussion of what feedbacks or forcings are considered missing in the models. This could go here in the introduction or later in the conclusions.

Page 3484 ,line 23-4, change to: These diagrams summarize three. . ."

Page 3486, line 29. Please provide an explanation for why a Gaussian error is added

scaled by 1 degree Celsius. Is this purely arbitrary or is the scaling based on some estimated measurement errors?

Page 3487, line 3-4. The figure caption does not match the description here. Please fix.

Page 3488, lines 7-13. It would be useful to have a discussion here on whether the equal weight assumption is consistent with the expected variability in various regions where observations are available. If higher variability in high latitudes is expected like in modern climate, then a brief discussion of this for paleoclimate would be warranted.

Page 3489, lines 6-7. Please be more clear about the random deviates used in this paper. How big are they? and why are they added?

Page 3491, last paragraph. Please point out that the first reference prediction alternative is a very weak one but it is also a reasonable first check. I think this could be clarified.

Page 3491, line 24. Replace tests with test.

Page 3492, line 15. "straightforward to calculate" needs to be backed up with equations. Please add the equations.

Page 3493, line 23. I don't know what "under represented" means or the phrase "resolution at the finest scales there is under-represented." This should be clarified, please. Is this what others call "representativeness error"?

Page 3494, line 19. At the beginning of this paragraph, it should be stated that the results are robust. It weaves and bobs around the issue with lots of example robustness tests until reaching the end but it would be more to the point if this was mentioned at the beginning.

Page 3496, line 5-21. It is not clear why the land+ocean data can produce positive skill while the others individually can both produce negative skill. This is rather confusing at

first glance and can use some clarification.

Page 3497, line 8. Delete "the" before "calculating"

–"–, line 23-25. Here, it should be stated what the full set of eight ensembles were. Something like: "We tested eight ensembles that included. . ."

Page 3498, line 9. Change "analysis" to "analyses"

Line 18-19, Suggest changing "are highly spatially variable" to "have high spatial variability"

Page 3499, line 8. It is not clear again what "representation error" means.

Figures: General comment 1: I suggest that the maps use an equal-area projection so that the tropics (0-30) and extra-tropics (30-90) show that they have roughly equal area. A simple linear in sin(latitude) transformation would be ok.

General comment 2: The filled dots are very difficult to read in the printed version.

Figure 2: Given that "Coldest month" is one of the data sets, it would be useful to show it.

Figure 6: Basic descriptions of the three statistics shown on the Taylor diagram are needed in the caption.

Figure 7. What about the other five ensembles? I suggest showing them despite their negative results.

Interactive comment on Clim. Past Discuss., 8, 3481, 2012.