Clim. Past Discuss., 8, C93–C97, 2012
www.clim-past-discuss.net/8/C93/2012/

Climate of the Past Discussions

Interactive comment on "Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium" *by* A. Hind et al.

Anonymous Referee #1

Received and published: 4 March 2012

The manuscript presents statistical models for observed and latent variables that play a role when comparing forced or unforced climate simulations with climate proxy data or instrumental records. Based on these a distance measure between simulations and proxies is developed that ranks the simulations in the same order as if the distance to the true, unknown climate was used. The distance measure can be used for multiple proxies, for proxies that represent different seasons, and includes weights that depend on the error of the proxies, which can vary in time. A significance test is then developed for the difference of the distance between simulation and proxies for two different simulations. A significance test is also formulated for the correlation between a simulation and the proxies. The framework is then tested in a pseudo-proxy setup where temperatures from either low or high solar forcing ensemble members of the COSMOS millennium simulations are used to generate pseudo-proxies. In this experiment the method is able to distinguish between the two ensembles if the high solar forcing ensemble is used as the pseudo reality, but not if the low solar forcing ensemble is used.

Although distance measures are a standard concept, this paper constitutes a major contribution to palaeoclimatology, as for the first time the specific form of the distance measure and the calibration of the proxies are developed based on a sound statistical model rather than being chosen ad hoc. The approach formulated here seems fully adequate for the problem and limitations are clearly stated. It is likely to strongly influence simulation-data comparisons in the future. As shown in the pseudo-proxy experiment, the method, when used with actual climate proxy data, has the potential to contribute to answering the open question about the magnitude of the solar forcing. The study also leads to the practically highly relevant statement that not much would be gained from increasing the current number of available temperature proxies, but that developing proxies with a higher signal to noise ratio would be very beneficial for deciding which simulations are closer to reality.

The first part of the paper, in which the statistical framework is explained, is very well structured and in general written very clearly. In this part the reader is guided well through the text and there are only a few places where a bit more explanation for readers without a strong background in statistics would be helpful. The second part, where the pseudo-proxy experiments are discussed is good with respect to content, but unfortunately the writing style is of a lower standard than in the first part. This is reflected for instance by several cases of repeating the same information in different locations, which indicates that the writer partly lost the overview of the text.

In summary, this is an excellent paper with high intellectual clarity and high practical relevance. I thus fully recommend publication after the specific points listed below have been addressed.

Specific comments

1.) page 270, line 9: the text says that ε is uncorrelated with y. Should it not be uncorrelated with τ?

2.) page 271, lines 17-19. The comment on multiple forcings has been made already on page 270, when discussing the forced model. Can it be deleted here?

3.) page 274, line 3: The sign of the first term of the right hand side of the equation seems wrong. This has no consequences for what follows, because after averaging the term is zero.

4.) page 274, line 17: replace 'any bias' with 'any ranking bias'?

5.) page 274, line 19: There are still the random variables x and ξ on the right hand side. Should there not only be expectation values?

If there should be an E(ξ) term, why is it not zero as stated when statistical model 1 is introduced?

6.) p275: Here z0 is regressed on τ (or on y) and the regression relationship is inverted to obtain an estimate z for τ (or y). As the authors point out this estimate contains the full variance of τ (or y) plus noise, so it has a higher variance than τ (or y).

Sometimes in climate science the direct approach is taken, where τ (or y) is regressed on z0. The estimate of τ (or y) has then a lower variance than τ (or y) unless the residual variance is added. This lead to some discussions in the climate literature about 'inflation vs. randomization' (e.g. von Storch, 1999: On the use of inflation in statistical downscaling, J. Climate, 12.).

It would be good to point out that the classical calibration approach used by the authors is different from the inflation approach in the direct regression, although in both cases the estimate contains the full variance of τ (or y).

It would also be good if the authors could comment on how appropriate it is to call the

CPD

8, C93–C97, 2012

Interactive Comment

rescaled proxy z an ‘estimator’ for τ (or y). In the case of low correlations the variance of z will be very high and it seems wrong to call z an estimate in the sense ‘given z0 this is what we expect τ (or y) to be’. In the case of the direct regression the estimator is the expectation value for τ (or y) given z0, and the variance of the PDF is given by the residual variance. What is the interpretation of z in the classical calibration approach? This may all be explained in Brown (1993) or other standard texts, but it would be helpful for readers without a strong statistical background to recall the main points.

7.) page 277: The form of the PDF (eqn. 3) and the way how the weights are obtained should be better explained.

8.) page 280, line 23: There is a left bracket missing in the first term on the RHS.

9.) page 290: It might be good to repeat here that the present study uses only pseudo-proxies to avoid the misunderstanding that the question which solar forcing is more realistic will actually be answered. The point is partly made in the last sentence, but an earlier and clearer statement would be helpful.

10.) Section 9.2 is rather messy and should be rewritten. Problems include:

- MPI-ESM does not include an additional carbon cycle (page 291, line7), JSBACH and HAMOCC are the components of the carbon cycle model.

- Forcings are not added to ‘CTRL reference boundary conditions’; this is a wrong use of ‘boundary conditions’, which are state variables at domain boundaries, not forcings.

- The phrase boundary conditions is also wrongly used on page 292, line 3. The correct term here is initial conditions.

- The discussion of the different solar forcings is not linked to the discussion of the same topic in 9.1. In 9.1 the references that indicate a strong solar forcing are Shapiro et al. (2011) and Friend (2011), so recent studies suggest a stronger than previously estimated forcing. In contrast in 9.2. Bard et al. (2000) is cited, which gives the impression that a strong forcing has only been suggested earlier, but recent studies

CPD
8, C93–C97, 2012

indicate a weak forcing. In 9.1 the forcings are compared in units of W/m2 whereas in 9.2 percentages are used, which makes a direct comparison of the two parts of the text difficult.

- The forcings are discussed twice at the end of page 291 and at the end of page 292.

- The text seems to switch quite randomly back and forth between discussing simulation results and discussing the forcings. A clearer structure would be good.

11.) There are more repetitions in 9.3.Page 293, lines 15-17 and page 294, lines 23-25 are highly redundant.

12.) page 295, lines 3-4. What have these studies found?

13.) page 295, line 27 to page 296, line 4: This has been said earlier and the text should be deleted.

14.) Page 296, lines 4 – 10. Why are these remarks not in 9.3, where the setup is discussed?

15.) Page 296, lines 10-13. The comment on hierarchical distributions is unclear.

16.) page 299, line 13: should say ‘could not have been expected’.

17.) page 300, lines 14-15, This has been said before on page 299 lines 4-6, but without the reference. The two text bits should be combined.

18.) Figures 3 and 5 are rather small and not easy to read. The y-axis range should be changed so there is less white space.

Fig.A1 should have the same y-axis range as Fig. 7

Interactive comment on Clim. Past Discuss., 8, 263, 2012.