Clim. Past Discuss., 8, C110–C113, 2012
www.clim-past-discuss.net/8/C110/2012/

Climate of the Past Discussions

CPD

8, C110–C113, 2012

Interactive Comment

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.

Y. H. Yamazaki (Referee)

hiro.yamazaki@ncl.ac.uk

Received and published: 8 March 2012

This article proposes a new methodology for a numerical "ranking" of climate model simulations using a statistically sound distance measure against a set of paleoclimate proxy-inferred temperature data, presenting a thorough description and discussions. It also provides concrete examples of its use in the context of ensemble climate modelling for the last millennium.

The main part of the proposed methodology is based on the conventional quadratic distance measure of time series. Nevertheless, in order to accommodate a variety of spatiotemporal climate proxy data that are fragmented in both time and space, detailed

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

and careful discussions are provided to establish a fair or unbiased ranking, with clear warnings for its application. It also includes the recommendation of an additional use of correlation-based pre-screening in order to avoid simplistic application of the proposed procedure that could produce meaningless results out of noises. Furthermore, it provides a discussion of different calibration procedures of proxy data with clear conclusions for the suitable choice to obtain unbiased model-data comparison. The detailed discussions and proposals in this article are likely to form an important basis for the further investigations of paleoclimate modelling studies. It is both topical and highly relevant for the current issues around the climate change of the last millennium and beyond, and is well worth publication after some clarifications.

The most negative aspect is its sheer volume of text to explain the statistical procedure, together with another segment of lengthy description to explain the modelling experiment that was employed to demonstrate the use of the statistical procedure. It appears that the latter part was written both as an application guide of the proposed procedure as well as a new scientific investigation of unresolved issues for the past climate forcing. This makes Section 9 particularly difficult to follow, and the summary in Section 10 is heavily biased towards the scientific finding from Section 9.

I would think the length of the first part is reasonable given the complexity and vast advancement of the new methodology, besides the discussions therein are generally clear and easy to follow. On the other hand, the lengthy description of published COSMOS millennium model runs and the scientific issues surrounding the forcing uncertainty should perhaps be deemphasised and a more focus should be placed herein to present the second part as a concrete and concise application guide (see an alternative suggestion below).

Overall, an minor revision would be needed before final publication.

Major suggestions for the revision.

a) Due to the large volume of the description of the statistical procedure, it would be

CPD
8, C110–C113, 2012

beneficial to summarise the advantages of this methodology, together with the likely pitfalls and limitations for its practical application.

b) If the authors intend to emphasise the new scientific findings in Section 9 regarding the causes of the past climatic changes, I would recommend this part to be published as a separate article. This will allow more thorough discussions of the limitations in the current modelling study and available paleoclimate data including the climatic forcing.

c) The final conclusion should cover all major findings from all parts of the article.

d) p268-269: Assumed (near) linearity between forcing and response imposes a serious practical limitation. In other words, climate forcings used in recent modelling studies and their responses cannot be expected to be linear, even approximately. This assumption will be particularly misleading to interpret the results of model run ensemble that employs a scaled forcing scenarios to account for uncertainty, which is expected to produce a nonlinear response. In recent GCM experiments such as COSMOS, the volcanic forcing is implemented as the combination of the optical depth in a narrow band of spectrum and a representative particle size of volcanic aerosols, whereas the CO2 forcing can be specified as the actual concentration in the atmosphere or the emission rate; the orbital forcing is provided as spatiotemporal variation of insolation. Therefore, the complex issue against this assumption should probably be treated outside the statistical framework by defining a "forcing" that is expected to cause approximately linear response on the target climatic parameter, leaving a clear statement in this article as a caveat or a practical requirement.

Minor comments

1) p266 L12: The requirement for this assumption is not clear. For example, is the correlation test discussed in Section 8 sufficient? Or, should models be tested using a distance measure before this methodology is applied?

2) p267 L16: Use of "dash" and "equal sign" are confusing here and thereafter. Either

CPD

8, C110–C113, 2012

use a different symbol for itemisation or avoid the equal sign.

3) p268 L5: Although “a proxy for the true temperature” is an appropriate term, it is potentially confusing as the term proxy is often used as a common term to describe the row proxy data of a climatic parameter before calibration. I would think something along the line of “reconstructed temperature from climate proxy data” or “proxy-inferred temperature” would make it clearer.

4) p268 L12: Define “actual” temperature (probably in terms of measured or proxy-inferred temperature)

5) p268 L4: The assumed type of correlation measure should be stated.

6) p273, L21: The title of Section 4 is vague. Perhaps the term “calibrate” or its variant should be included.

7) p280, Section 6: The use of unforced model runs to represent stochastic components are not new but a rather common practice in detection and attribution studies. Referencing past literature would be beneficial here. This also applies to the proposal in p278, L6, although the discussion therein is in a specific context of how to compute the weights.

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