Clim. Past Discuss., 2, S694–S696, 2006
www.clim-past-discuss.net/2/S694/2006/

Climate of the Past Discussions

CPD

2, S694–S696, 2006

Interactive Comment

Interactive comment on "Mid-Holocene climate change in Europe: a data-model comparison" *by* S. Brewer et al.

Anonymous Referee #2

Received and published: 19 December 2006

Review of "Mid-Holocene climate change: a data-model comparison"

General comment.

This paper represents an interesting approach towards quantitative comparison of palaeoclimate data and model output - an important area of research that I am happy to encourage. However, I have reservations about the applicability of the method and the interpretation of the results.

Specific comments.

A fundamental problem for the field is that data are generally available on a point-by-point and highly localised basis, whereas model grid points represent some sort of spatial average. A direct comparison will not necessarily be appropriate, as even a model which represents reality perfectly on its resolved scales cannot be expected to match pointwise samples with a higher resolution. This paper illustrates the issue clearly enough with multiple data points frequently disagreeing within the same model grid cell (Fig 2a)! Standard methods for approaching this problem would be to either apply some statistical downscaling of model output (to account for unresolved variability) or else perform some smoothing/interpolation of the data onto the model grid. This manuscript applies a rather different technique which does neither of those things. Unfortunately, I'm not sure that it works, but I hope that the authors can persuade me that it does (or that it can be adapted to do so). Given the "noisy" nature of the data it seems that the model and data represent fundamentally incommensurate parameters - ie, there is a representativity error which has not been considered.

Although the abstract mentions "regions of homogeneous climate change" and the ability of the models to reproduce "the correct location of those regions" there is (unless I have misunderstood something) absolutely no quantitative geographical comparison of models to data, just a brief qualitative discussion. Indeed the geographical segregation of the clusters themselves is rather tenuous and in no way enforced (or even encouraged) by the clustering algorithm, which merely aggregates similar triples of (humidity, degree days, min temp). A "model" which outputs random uncorrelated outputs with the appropriate variance for each of the 3 parameters would probably simulate the clusters under this analysis, whereas a "perfect" model which correctly reproduces the climate on its resolvable scales might well not do so. Furthermore, while the authors make the valid point that a good technique for evaluating models should ideally allow for modest geographical shifts in features of the climate, it is hard to justify the claim that a model which happens to have a single cold wet gridpoint somewhere at the edge of the domain has successfully simulated this feature of climate change (cluster number 3) which is visible in the data across the centre of the area under study. But unless the issue of the different spatial scales of model and data is addressed, it is not even clear that a model ought to have any grid cells which exhibit this behaviour, since there are only a small proportion of data points which do so, and they are frequently adjacent

CPD

2, S694–S696, 2006

Interactive Comment

EGU

to data with different properties.

Does the formalism of fuzzy logic and the use of Hagaman distance bring any real benefits over the use of the more common log-likelihood for Gaussian errors: (m-o)ˆ2/2(s_oˆ2+s_mˆ2)? I realise that your fuzzy triples can be (and are, for the data) asymmetric but other than that I don't really see the point of this aspect of the analysis (even after reading the previous cited papers).

In conclusion, the fundamental question I have is this: is there any way that this method can be used to meaningfully discriminate between and/or improve models? In particular, is there any reason to expect that a model which simulates the data better (by your measures, either before or after the rescaling) is actually providing a better picture of the overall climate fields at its given resolution, than one which does not? I don't see how this can be justified while the issues of geographical location and representativity are ignored.

Technical corrections

Some parts of the description (section 2.4) seem poor. In particular, it is not clear how points can be "randomly assigned to the closest vector" and I can't understand what standard deviation is calculated across the repeated results to justify your selection of clusters (especially as the clusters have no natural order). Table 3 has at least two shades of grey and the scales for a and b in Fig 1 are probably swapped with c and d. Fig 3 needs a more detailed caption (what do the various shapes represent?) and it should be mentioned that Fig 4 represents the model results after rescaling.

Interactive comment on Clim. Past Discuss., 2, 1155, 2006.