Title: A comparison of model simulations of Asian mega-droughts during the past millennium with proxy reconstructions.

By B. Fallah and U. Cubasch

Manuscript reference: cp-2014-73

This work presents an analysis of reconstructions and two ensembles of last millennium simulations focusing on Asian mega-droughts: the PDSI of both model ensembles and of reconstructions is analyzed and compared; a maximum covariance analysis of model PDSI and SSTs is performed; and finally the regime behavior in both model ensembles is also analyzed. The purpose of the manuscript is valid and interesting, however, I think the manuscript it is not at the stage of being published and I would encourage the authors consider the arguments herein and discuss them or implement them if they think they are of use previous to further revision.
In addition to the following comments I would advise the authors to improve the rationale of the paper, to discuss the logics of why each analysis is performed and how the three different analysis connect/relate to each other and jointly help to clarify the dynamics of drought. This discussion is missing. More specific comments are provided in the next pages, however, I would highlight here several issues. The influence of using ensemble averages instead of ensemble members in the two initial analyses instead of the individual members should be argued and discussed. Also the rationale for using ensemble members and ensemble averages in the regime analysis. The benefits and purposes of the approach undertaken should be explained. Also, if the results of the analysis support an agreement in the timing of drought in reconstructions and simulations, the influence of external forcing as the common factor between the reconstructions and the simulations should be discussed. I think that the differences and similarities in the results with the two model ensembles in the three analysis performed should also be discussed.

I hope some of the following comments are of use for the authors in revising the paper.

COMMENTS

_____

1.- Data and methods, Page 2688, line 2-...
The paper uses the ECHAM5/MPIOM and the GISS-E2-R simulations as part of the PMIP3-CMIP5 experiments. While this is true for the GISS-E2-R runs, the forcing specifications of the ECHAM5/MPIOM do not comply with the PMIP3 specifications. This part should be rewritten accordingly.

Line 19-20
'The forcings in the ... ECHAM5/MPIOM simulations' This is nor strictly correct. At least the text should explain/discuss in which sense they are similar. For instance CO_2 in the ECHAM5/MPIOM is calculated interactively. For other forcings it may be worth discussing the differences/similarities.

2.- Section 3, Page 2690, line 1-8.
I like the approach of using model ensembles. However there are some issues related to this approach that I think are worth discussing or at least being considered by the autors. The EOFs in Fig 1 are indeed similar. However the model EOFs are obtained from ensemble average fields in which the internal variability is canceled or diminished by averaging. Reality (i.e. MADA) is comparable in those terms to only one realization of the system, i. e one model simulation. If the MADA is compared to ensemble averages, the behavior of each ensemble member may be an important issue to report. Additionally, the statistics obtained may be not only related to the model behavior but also to the ensemble average itself. For instance, it would be expected that a mode explains more variance if the ensemble average is taken over a larger number of members (like in the case of ECHAM/MPIOM for which the ensemble is larger than GISS-E2-R) or that the pattern correlations are larger. I think it is worth discussing the implications of using model ensembles in this work relative to individual members and how this affects the results.

3.- Mega-droughts, forcing and use of ensemble averages.
Section 3, Page 2690, line 8-10:
'Those periods ... 'active' resp. 'break' phases of the ... Fig. 2)'.
This is interesting. Defining synchronous periods of positive or negative values in the reconstructions and the model simulations relies on the assumption that both the reconstructed and the simulated drought are produced in response to forcing. Would the authors agree on this? If so, I think this should be clearly stated as a rationale from the beginning. Additionally, this would support the use of ensemble averages in comparison to single ensemble members. However it would be important to indicate how closer is the ensemble average to the reconstructed pdsi (if it is at all) relative to the individual ensemble members. How does this influence the comparison in Fig 2? how is the variability in Fig 2 of the ensemble average PCs relative to the individual members.
The coincidence of the timing of mega-droughts during the LIA in the model ensembles and the reconstructions is discussed. This can only happen if both happen in response to external forcing. This should be reinforced by the individual members showing a poorer coincidence and discussing this fact.

Section 3, Page 2690, line 14-21:
It is also important to state how mega-droughts are defined in this paper. I have not seen the definition so far.

There seems to be indeed some agreement in the time evolution of the index in Fig 2, both if we consider the Pcs or if the periods of 'active' and 'break' phases are considered. I would suggest discussing this a bit more in detail. How good is this agreement in terms of correlation, perhaps for high and low (multi-decadal) timescales. How different is the agreement of the ensemble averages relative to the individual members and how important is this for the credibility that the coincidence is not by chance and arguably related to external forcing? Why do the 5 historical mega-droughts not coincide with the minima in both series, or at least in the MADA (red) line?

Discussion of Fig 3: What is the meaning of dots? Fig3 caption indicates that dots stand for grid points that agree with MADA. Can you be more specific? How can it be that some gridpoints in c are in blue (positive) while in MADA are in brown (negative) and still have a dot indicating agreement?
Fig 3c indicates a different pattern to that of MADA (Fig 3a) and GISS (Fig 3b). This is discussed in the text. How can it be that in Fig 2 MADA and ECHAM5 seem to be in phase while GISS is not?
Similarly for the following Fig4-Fig7 panels. The comments on the reconstructed droughts and their importance are welcome, although I would suggest discussing their consistency with Fig 2 in terms of the PC and the 'active' and 'break' phases.

4.- MCA analysis, Figs. 8, 9
Section 4.1, Page 2692
I think this discussion is also interesting, but it should also be improved at various levels. For instance, the correlations of the PDSI series are reported to be significant. The relation with GISS is indeed suggestive. It would be good to indicate significance in table 1 as an alternative to the largest value in each row. Are correlations calculated over annual values or low pass filtered values?. This should be indicated, also if autocorrelation has been taken into account and, if not, I suggest it should be.

Additionally, after reading Section 3 where the synchronicity between Asian drought in reconstructions and simulations is discussed (Fig 2), I think it may be interesting to include in this analysis the reconstructed pdsi in order to link it to changes in simulated SST. This is based on the argument in the previous section that reconstructed and simulated PDSI are related.

I think that another issue that is relevant for this work is to, once again, illustrate or report on the behavior of individual simulations. How does this influence on the reported explained variances (lines 9-10)?. Also, the time series in figures 8 and 9 suggest an influence of external forcing (volcanic) in the mid 15th and early 19th century. I recommend this should be discussed and would likely be more evident in the ensemble average than in the individual simulations.

5.- Regimes, Section 4.2
Page 2693: regarding the different explained variances in both models, this is

noteworthy and deserves some more comments, for instance in relation to the variance explained by the individual ensemble members and the effect that averaging over a different ensemble size may have on the result.

Concerning the analysis performed later with ECHAM5/MPIOM and GISS-E2-R I can think of several issues that I would like the authors to discuss about or consider:

− The discussion about the distribution of the two regimes in Fig 10 and Fig 13 is interesting. In Section 4.2.1 the authors describe the spatial variability of both regimes. I think the figure and discussion would gain from showing the precipitation patters associated to this regime, perhaps also the PDSI from the model. Also the distribution of associated PDSI events to each regime can perhaps be shown if the authors consider it of use. It makes more sense to me to trace the actual behavior of these variables from the available model runs, that in fact show differences between them, than rather argue only from the literature based on different analysis by other authors.

− Regarding the use of pdfs of the regimes for the individual simulations and for the ensemble averages: I have reservations here about the use of these pdfs for the ensemble averages. First I would suggest to argue and discuss the changes from the individual ensemble members to the ensemble average. Second, I would like the authors to discuss the meaning of the regimes obtained from the ensemble average. The model is expected to produce very different states for a given time step in different simulations due to internal variability. What is in this case the meaning of the regime state for a given time step of the ensemble average?. I think discussing this a bit more and why the probability of regime 2 diminishes is pertinent.

− The patterns in Figures 13 and 15 are vaguely described and not discussed. I suggest the authors argue about the differences and similarities in the results obtained from both models. Are these the correct events (5 historical mega-droughts) to consider? (recall the low agreement of the minima in Figs 2,8,9 with model series).

6.- I think conclusions and abstract should be rethought in view of the previous points. I would advise including in the conclusions some cross section discussion about the results obtained in each section and how they complement each other. Also, how the agreement/disagreement between reconstructed and simulated drought is affected by model issues (e.g. ensemble averages vs ensemble members and the benefits of using two ensembles) and how it may be traced to external forcing in different models. The different/similar results obtained by the two models should be discussed.

Minor comments:

- Page 2686, line 16:
  '...that lasting for years...' -> '...that last for years...'
- Page 2686, line 23:
  '... century have been associated with...' -> ' ...century are expected from...'

- Page 2687, line 6-7:
  '... deviation from <u>a/the</u> normal watter ...'
  line 11-12:
  '... at <u>the</u> wrong times...'
  I suggest checking text for grammar. There are more (relatively unimportant) details of this sort that I will exhaustively highlight here, just some of them picked up 'randomly':
  Page 2690, line 1: '... presents <u>a</u> monsoon...'.
  Page 2690, line 6: '... GI<u>SS</u>-E2-R...'.
- Page 2687, lines 15-22:
  The rationale of this part is not clear to me as a reader. I think it is possibly a problem of language/writting and maybe the authors can reconsider improving the writting: the role of ENSO and aridity trends; the inter-annual changes in seasonality are not clear; nor how this leads to the question of forcing vs internal variability.
- Page 2687, lines 25-26:
  Maybe worth mentioning also PMIP3

- Page 2689, lines 21-21:
'The advantage of MCA, compared with coupled ... highly coupled'. I think it would be good referencing this statement to some previous literature so that the reader can have this information for interpretation or potential use. Similarly the comments about MCA or coupled EOFs.

- Page 2689, lines 27-28:
'... remapped to MPIOM's grid'. It would be good stating the specific resolution.

- Page 2698, line 5:
  Get correct reference for IPCC. Also in page 2695, line 20.
- Page 2692, line 12:
  Krishna Kumar et al (2006): I did not find this reference in the list.
- Page 2695, Line17-20. Can you explain/discuss this sentence better? (see point