Clim. Past Discuss.,
doi:10.5194/cp-2016-129-RC1, 2017

Climate of the Past Discussions

Interactive comment on "Reconstructing past climate by using proxy data and a linear climate model" *by* Walter A. Perkins and Gregory J. Hakim

Anonymous Referee #1

Received and published: 12 January 2017

Review on Reconstructing past climate by using proxy data and a linear climate model Walter A. Perkins and Gregory J. Hakim

I find that this article does not belong to the journal of Climate of the Past. Thus I reject the publication. However, I find the idea novel and interesting and I would suggest to submit a revised version to a more theoretical journal.

Here is why it doesn't belong to Climate of the Past: The idea of using LIM as a substitute for otherwise expensive online data assimilation sounds wonderful, and would prove to be it if one didn't need to introduce a parameter a. But as authors showed with a being equal to 1, the results of linear online DA are worse than of offline DA. Then the question I pose to authors is how to choose an optimal a? What would be the criteria for optimal a? Looking at the tables 1 and 2 it seems only 20CR has highest CE, r, and

lowest CRPS at the same a, mainly 0.9. For other models highest CE is reached with one value of a, while highest r with another value of a. Therefore an investigation of how to choose an optimal a is lacking in this paper, and without an optimal a I don't see how to extend this work to practical applications of the past climate state reconstruction. There is yet another manifestation of more theoretical work to be done, mainly in Sec. 4.2, where authors find inconsistency between the best model 20CR in terms of a scalar skill (CE, r, and CRPS) but worse in terms of spatial reconstruction. They propose that it could be due to a short time scale but this could be checked. And again, does this mean that $\alpha = 0.9$ is not optimal for 20CR? Therefore what I suggest is to study the methodology in a theoretical framework by revising the article and submitting it to a theoretical journal.

Other major comments: LIM is calibrated on model 1 (CCSM4) without any data assimilation over a period 850-1850, on model 2 (MPI) without any data assimilation over a period 850-1850, on model 3 with data assimilation (20CR) over a different period 1850-2012, and on a data set over yet a different time period (1950-2010 I would assume, though it is not mentioned in the paper). Thus the models are completely different in terms of the time period, use or not of the data, and only being the data. This makes it hard to compare and draw conclusions. Instead LIM should be calibrated on a model without DA, on the same model with DA, and on observations used in that DA.

As the prior authors used results of the CCSM4 model, the same model they used for LIM calibration. It appears that linear CCSM4 DA provides good results in terms of both scalar skills and spatial reconstruction. Is it because there is less inconsistency? How would it change if the prior was from another model?

In order to provide a fair comparison authors need to include "expensive" online DA (using a nonlinear model instead of LIM).

Minor comments: Page 7, Line 18: Why is there a shift in blending coefficient? This is

again related to my comment on how to choose an optimal a.

Page 7, Line 31: Why is there improvement compared to offline DA even though the trend is largely underestimated for $\alpha = 0.95$?

It would be interesting to introduce another metric – bias – in order to check whether the model either underestimates or overestimates the observed values.

I suggest plotting time series of averaged temperature of different models against observations for best a for CE, for best a for r, and for best a for CRPS.

Interactive comment on Clim. Past Discuss., doi:10.5194/cp-2016-129, 2016.