Clim. Past Discuss.,
https://doi.org/10.5194/cp-2018-120-RC1, 2018

Climate of the Past Discussions

Interactive comment on "Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling" *by* Robert Tardif et al.

Anonymous Referee #1

Received and published: 19 October 2018

General comments

This manuscript describes a new annually resolved 2000-year reanalysis, for which paleodata are assimilated into a climate simulation ensemble, using a Kalman filtering method. The paper is well written and especially interesting from the methodological points of view. First, the way seasonal information is assimilated into annual averaged and second, because the performance of multiple proxy data sets is compared. Probably, data-assimilation techniques are the future in the field of climate reconstruction and hence the work is well suited for Climate of the Past. The use of the data assimilation method with an appended state to account for seasonality is innovative. Thus, it should not be hidden in the appendix but appear more prominent in the main text, written in a way that is understandable to the average climate of the past reader and not

Printer-friendly version

Discussion paper

just data assimilation specialists. Otherwise, it is not clear to the reader how seasonal information can be used if the state vector just contains annual averages.

In contrast to the methodological part and sound statistical validation, the paleoclimatic discussion needs to go more into depth and the authors need to be more critical about their own results. The authors compare their new data set to the previous version (Hakim et al. 2016) and conclude this new version would be an overall improvement. This conclusion is based on validation statistics in the 20th century. However, the new data set lost all multi-decadal to centennial variability in the global mean temperature and does not show a warmer medieval period nor a cooler “little ice age” anymore. The authors do not discuss this issue at all. The paper suggests that this new reanalysis version would present the more likely global mean temperature evolution of the past 2000 years although it is in contrast to what most other reconstructions and paleodata records suggest. In the introduction of this study, the authors write: “Hakim et al. 2016 [is] . . . in good agreement with previous reconstructions of northern hemisphere mean temperature” and this first version had similar low frequency variability as previous reconstructions. The loss of low-frequency variability is most likely a consequence of the proxy data sets used, because this is the major change in the new version. Many of the tree-ring chronologies in Breitenmoser et al. 2014 are not climate sensitive at all or moisture sensitive. As precipitation does not show any low-frequency variability in contrast to temperature, it is a logical consequence that using covariance information from moisture sensitive trees to correct temperature data leads to a loss of low frequency variability. A second reason may be the use of proxy data with dating uncertainties, such as ice cores, in an annual reconstruction. These proxies probably do not have age errors in the 20th century validation period but become just noise if they have an age offset of one or a few years further back in the past. The authors just conclude that using moisture sensitive data leads to improved reconstruction skill, although this is only true in the 20th century validation period but not in the pre-instrumental period, most user of this data set will be interested in.

In the current version, the global mean temperature evolution is the reappearance of the famous “hockey stick” in climate science. After all the discussion, the hockey stick was rising 20 years ago, I would not publish this as a state-of-the-art temperature reconstruction, especially not without a discussion and not if it is an artefact of unscreened input data.

I see two options, the first would involve minor revisions and the second major revisions:

1. It must be stated prominently (already in the abstract) that this reanalysis should not be used or considered to have the correct multi-decadal and centennial variability and that the global mean time series over the last 2000 year potentially has serious issues. The discussion needs to include all problems of data set, too and ideas how to overcome them in the future. In general, the paper should be put more into a context of methodological improvements to achieve better products in the future instead of claiming this would be nearly the prefect reanalysis for the past 2000 years.

2. A proper screening of the data needs to be introduced that prohibits the assimilation of non-climatic information, which has just spurious correction with observations in the short window of overlapping instrumental and proxy data (a minimum of 25 data points has been used in this study, page 5, line 2). These records will have little weight in the assimilation procedure due to large residual variance, but hundreds of little errors probably produce significant noise. Probably, precipitation limited proxies and proxies with age errors have to be removed or treated specifically, too. These are just some ideas and it will need many improvements and new experiments to find and solve the problem.

A difficulty in the review process is that the input data has not been published, yet. Hence, it is not possible to properly judge the input data base. However, it appears to be basically the Breitenmoser et al. 2014 data set with a few coral and ice core records added. Why do you not simply refer to this first publication and give citations for the

Printer-friendly version

Discussion paper

additional records or wait until the Anderson et al. paper is published?

In general, the decrease of skill further back in time is not discussed sufficiently. It should also be discussed why forcings are not important and what the consequences of unforced simulations ensembles are for the final product, especially further back in the past when the proxy network is sparse. I suggest to evaluate the spatial skill of the reanalysis in the 20th century but with the spatial proxy network at multiple time slices, e.g. 0 AD, 500 AD, 1000 AD, 1500 AD. Additionally, not using forced simulation offers the potential to use them in the validation procedure. It could be checked if temporal and spatial patterns of known past events or periods are well represented in the reanalysis, e.g. spatial moisture distribution after eruptions (Iles and Hegerl, 2015).

Finally, it would be interesting to see a map of the regression residuals to get an idea how many paleodata records have significant influence in the assimilation procedure and which are basically ignored because they have no climate information. Additionally, I would like to know how many records in the PAGES2Kv2 data base have expert information on seasonality? I would be interested to read how well the expert-based seasonality in the PAGES data base agrees with the objective assessment in this study. Probably, the experts did a similar search for highest correlation, maybe just including more possible combinations of growing season months. I was surprised to read that the authors use an extended fall period (JJASON)? Is there any reference for trees which are limited by climatic conditions in these autumn months.

It would be favorable to store the data at a world data center and not at a personal homepage.

Technical corrections

Abstract:

- skill score increase in percent is misleading. It is easy to have a large relative increase if scores were very low in the comparison data set, e.g. in Z500 where CE improves

Printer-friendly version

Discussion paper

from very negative to less negative the increase in percent is large but the skill is still negative!

- be more precise what is meant with “ensemble characteristics”

Introduction:

- Line 17: apart from paleoclimate with annual observations, the forecast model is a third important component (this is even written in the Methods section)

Methods:

- Page 7, line 6ff: I do not understand why the calibration is done with a different gridded data set than the validation. Both data sets a based on largely overlapping instrumental observations and therefore clearly not independent.

- On Page 4 line 14: How many ensemble members?

- On Page 4 line 14 it is written that Hakim et al. 2016 worked “with the same randomly drawn ensemble members used for every year in the reconstruction, whereas on page page 5, line 28 it is written for this study: “each using a different randomly chosen 100–member ensemble”. Are both studies consistent and is each year build on different randomly chosen ensemble members?

- Page 5, line 1: “Only records for which a PSM can be established are shown . . .”. What do you mean by “shown”? There is no reference to a figure. Do you want to say that only records meeting these criteria are assimilated?

- Page 5, line 8: Breitenmoser et al. 2014 is not screened for any climate sensitivity.

- Page 6, Proxy modeling: It should be repeated here over which period the regression coefficients are calculated. As many data points as available in the overlapping period with instrumental data but minimum 25 pairs of x and y?

- Page 7, line 30: Is there a reference that any tree-ring proxy responds to an extended

fall period (JJASON)? The given references point to common growing seasons from May to August in the northern hemisphere. Why are not all combinations of growing season length tested and the optimum is chosen? In the PAGES data base there are also various different length of growing seasons defined.

- Page 8, line 26: “local” should better be “grid box”

The updated reanalysis:

- Page 9, line 5: Can you explain the localization better? Does a cut-off radius of 25000 km mean that each proxy influences basically the entire globe?

- Page 9, line 7: Mention that the reference for the skill score is climatology

Figures:

- some text is too small that it cannot be read in a print version

- figures should have consistent font types

Interactive comment on Clim. Past Discuss., https://doi.org/10.5194/cp-2018-120, 2018.

Printer-friendly version

Discussion paper