

## Interactive comment on "Pseudo-proxy evaluation of Climate Field Reconstruction methods of North Atlantic climate based on an annually resolved marine proxy network" by Maria Pyrina et al.

## Anonymous Referee #3

Received and published: 22 May 2017

## General comments:

Pyrina et al. test the ability of two climate field reconstruction methods (PCR and CCA) to reconstruct sea surface temperatures in the North Atlantic based on pseudoproxy experiments (PPEs) that replicate the spatial locations and noise characteristics of a small Arctica islandica proxy network. They show that, within the context of these PPEs, both PCR and CCA can produce reasonable skillful reconstructions of sea surface temperatures in parts of this region, but with differences depending on which model/reanalysis dataset is used as the target field. Overall, I think the paper presents some interesting results and is mostly well-written (though see "Technical Corrections" below). I do think the paper could benefit by being a little more explicit about

C1

the research question(s), objective(s), key results, and (especially) the overall significance/contributions of the findings. As it stands, the broader significance of this work is a little unclear. I also have several specific concerns (listed in "Specific Comments") that I think the authors should address prior to publication.

Specific comments:

1) Why choose calibration periods that are earlier than the validation period? Given that every real-world climate reconstruction would be calibrated during the later period during which instrumental data is available, wouldn't it make more sense to calibrate the PPEs using the 1850-1999 period and then validate on the earlier time periods (Medieval, LIA, etc)? To me, that seems more intuitive and would still address the stationarity issue.

2) Is there any reason why these two particular CFR methods (PCR and CCA) were chosen, while other common CFR methods (e.g., RegEM-TTLS and RegEM-ridge) were excluded? I'm not necessarily suggesting that the authors need to redo the analyses with additional CFR methods, but I would at least like to see a little justification for why these methods were chosen while others were excluded.

3) I think it would be useful to include other complementary validation statistics, such as mean bias, coefficient of efficiency (CE), reduction of error (RE), and/or root mean squared error (RMSE). I'm not sure that only correlation and standard deviation ratio are enough for a robust assessment of model performance. A spatial assessment of mean bias and either CE or RMSE could add important information to this study.

4) How is it possible to have a standard deviation ratio greater than 1 for the CCA results? As I understand it, all parametric reconstruction methods will result in at least some variance loss unless the proxy is perfectly correlated with the target climate variable (McCarroll et al. 2015), which would only be the case at the particular grid cells with the noise-free pseudoproxies. Even in other studies that used noise-free pseudoproxies (e.g., Smerdon et al. 2010, 2011), at least some variance loss was observed.

I therefore don't see how the reconstructed SSTs could have grid cells with greater variance than the "observed" (or in this case, model-simulated) SSTs.

5) The authors state that they chose to retain the first 10 PCs in PCR and the first 5 EOFs in CCA. How sensitive are the results to this choice? How were these thresholds chosen? Why were different thresholds chosen for the two CFR methods? As it stands, these seem like arbitrary choices. Did the authors consider more objective criteria for determining these thresholds, such as the "estimated noise continuum" approach used by Mann et al. 2007 or an optimization approach similar to Smerdon et al. 2010?

6) I would like to see some more discussion about the stationarity of the PPEs. Specifically, there is an implicit assumption when creating the pseudoproxies that the proxy response to SST variability is stationary (since the pseudoproxies are just SST+noise). Is this necessarily a realistic assumption for real-world A. islandica, or is it possible that the response of this species to SST variation could be non-stationary (similar to the well-known "divergence problem" in high latitude tree-ring widths)?

Technical corrections:

-Line 32: "loosing" should be changed to "losing"

-Line 39: delete "target to"

-Lines 52-54: I would rephrase this sentence. It is not clear as is. I think "aggravating" is the wrong word choice.

-Line 65: "recent years" should be letter "e" not "c".

-Line 74: delete the comma after "PPEs"

-Lines 97-100: It might be nice to see the proxy locations in a figure instead of just listing the coordinates. I don't see this currently in the figures, and I think it would be helpful for interpretation of the results.

-Line 264: "extend" should be changed to "extent"

-Figure labels: these are labeled "PCA" but throughout most of the paper this CFR method is referred to as PCR. I would suggest changing the figure labels to PCR to match the term used in the main text (though there are some inconsistencies throughout the text regarding use of PCA vs. PCR).

References:

McCarroll, D., Young, G. H., & Loader, N. J. (2015). Measuring the skill of variancescaled climate reconstructions and a test for the capture of extremes. The Holocene, 25(4), 618–626.

Mann, M. E., Rutherford, S., Wahl, E., & Ammann, C. (2007). Robustness of proxybased climate field reconstruction methods. Journal of Geophysical Research, 112, D12109.

Smerdon, J. E., Kaplan, A., Zorita, E., González-Rouco, J. F., & Evans, M. N. (2011). Spatial performance of four climate field reconstruction methods targeting the Common Era. Geophysical Research Letters, 38, L11705.

Smerdon, J. E., Kaplan, A., Chang, D., & Evans, M. N. (2010). A Pseudoproxy Evaluation of the CCA and RegEM Methods for Reconstructing Climate Fields of the Last Millennium. Journal of Climate, 23(18), 4856–4880.

СЗ

Interactive comment on Clim. Past Discuss., doi:10.5194/cp-2017-61, 2017.