Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2020-7-RC2, 2020
© Author(s) 2020. This work is distributed under
the Creative Commons Attribution 4.0 License.

Nonlinear Processes in Geophysics
Discussions

Interactive comment on "Applications of matrix factorization methods to climate data" *by* Dylan Harries and Terence J. O'Kane

Anonymous Referee #2

Received and published: 20 July 2020

This study compares several dimension reduction methods in scenarios with or without a dominant large-scale mode separating from the smaller scales. The introduction section provides a nice review of different dimension reductions methods for finding dominant modes of climate variability. The paper is mostly well written and results are interesting and provides useful suggestions for choosing a proper method for different climate variability analysis scenarios. I have only a few minor comments listed as follows.

1. The conclusion section does mention the caveat of neglecting the role of time dimension. The SST case has a dominant ENSO signal that is mostly a time oscillation of a fixed spatial pattern, while in the geopotential height case there are traveling wave signals with spatially changing patterns. For the later case, can the time dimension be

included in the analysis so that a more physically interpretable mode can be found? Can including time dimension directly in the d-dimensional data produce different results from applying temporal regularization?

2. From the description in section 2, it is a little hard to conceptualize the key differences between the AA and CC methods, and their advantages over the k-mean clustering method. Could you list the cost function and constraints for each method in a succinct manner and highlight their differences?

3. Comparing Figures 9 and 13, the behavior of CC and AA finding basis with more extreme departures from mean than k-mean clustering is the same for both the SST and Z cases. Does one case prefer bases with larger departures from mean than the other? Or, is the magnitude of bases functions less important than their alignment with the actual physical modes?

4. For the Z case, both AA and CC methods are not aligned with the PCA bases functions. Do you have any insights of which method is superior in this case? Or, are they all not finding the physical modes because of missing the time dimension?

5. For the RMSE for reconstructed data (Fig 10), is there a similar plot for the geopotential height case?

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2020-7, 2020.