Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2019-45-RC2, 2020

Nonlinear Processes in Geophysics
Discussions

Interactive comment on "Post-processing of seasonal predictions – Case studies using the EUROSIP hindcast data base" *by* Emmanuel Roulin and Stéphane Vannitsem

Anonymous Referee #2

Received and published: 16 February 2020

Review of the manuscript "Post-processing of seasonal predictions - Case studies using the EUROSIP hindcast data base" by Emmanuel Roulin and Stéphane Vannitsem

The authors evaluate the performance of the seasonal hindcasts from EUROSIP multi-model system for a number of variables and regions in Europe, and consider various post-processing techniques and various multi-model combinations.

I was overwhelmed with the sheer amount of presented information but I don't feel I learned anything valuable from the study. I don't exactly know what the goal of the study is. If it is to demonstrate that some regions in Europe have some skill on seasonal time scales, then I don't think there is really need to consider all these multitude of

the results (2-3e4 in the author's estimation). This particular point could have been demonstrated with a very small subset of the model results. On the other hand, if the point of the paper to demonstrate advantages or disadvantages of a particular post-processing method, then after reading the manuscript I still have any idea which method is recommended. It is also not clear what multi-model combination is to be used. It is all look rather random and disorganized.

A sheer amount of the considered statistics presents a selection problem. Even if there were no skill in none of the model predictions, some of the skill estimates will be positive just by chance due to sampling variability. Given the large amount of statistics considered ($\sim$2-3e4), it is not hard to imagine that many hindcasts will appear skilful, even in the absence of true skill. Many if not all of the skill estimates summarized in Tables are likely to be inflated and are not accurate representations of the true skill.

I don't believe it is sufficient for a scientific paper just to give a description of the results, without providing any useful insight. No new post-processing methodology is suggested. No new scientific insights are gained. Language needs some polishing as well.

I can't recommend this paper for publication.

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2019-45, 2019.