

## ***Interactive comment on “Stochastic Ensemble Climate Forecast with an Analogue Mode” by Pascal Yiou and Céline Déandréis***

**Anonymous Referee #2**

Received and published: 19 October 2018

In this manuscript the authors develop an ensemble forecasting system using an analog-based weather generator. They test this ensemble forecasting system for NAO and the temperature at several weather stations. They focus on the forecasts of temporal averages from 5 to 80 days. The forecast is made for each averaging period at the first corresponding lead time. The skill of these forecasts are evaluated through skill scores (the correlation and the continuous rank probability score, CRPS, the latter being well adapted to ensemble forecasts). The authors claim that there is some skill of the temperature and NAO up to seasonal time scales. I am not convinced by the system they propose, nor by the skill they found, for three important reasons:

(i) The system they propose suffers from a very important drawback, which is the progressive convergence toward the climatological mean as illustrated on the right column

C1

of Figure 3. There is only little variability of the forecasts for long time averages, indicating that the ensemble forecast is unreliable. This makes of this system a very poor probabilistic forecasting system, since the forecasts do not span the set of possible values of the observed variable. Reliability is one essential ingredient of ensemble forecasts that can also be easily checked with the decomposition of the CRPS in reliability and resolution. I therefore do not consider this ensemble system appropriate.

(ii) It is not clear at all to me why the authors are looking at the first lead time of the 5, ... 80 days averages. Using this approach, one can certainly expect that if one start from an initial state close to the reality, the forecast of the averages will always be better than the climatological average (provided we have access to an infinite sample). In other words some positive correlation will always be present, even if it is very small. This skill is artificial (due to averaging from the initial state) and I am wondering why the authors did not have looked at the skill of the daily values of NAO or temperatures. My guess is that there is no skill beyond a month or so.

(iii) The analysis of the skill of ensemble forecasts should be done with appropriate tools. The CRPSS is one of them, but it is much more important to look at its decomposition in reliability, resolution and uncertainty. These are standard tools that can be found in classical books or papers (e.g. H. Hersbach, 2000, Weather and Forecasting, 15, 559-570).

Some additional (less important) points

1. The algorithm of page 4 (section 3.2) is far from clear. It would be nice to visualize the algorithm, together with the relations that are used for evaluating the weights.
2. Page 6, line 5. Is  $S=N$ ? This is not clear to me.
3. An additional concern I have is the comparison with the persistence in Figs 4 and 5. It seems to me that the observables based on persistence display a higher variability than the forecasts constructed here (that are converging to the climatology). I there-

C2

fore suspect that the reliability of the persistent forecast is better than the one of the stochastic forecasts (the reliability term in the CRPS decomposition should be smaller for the persistence case), which is not reflected here in the analysis of the CRPSS. I would be very useful to evaluate the different terms of the CRPS to clarify the difference between the two systems. This will allow in particular to clarify why one gets 0.45 for all averages for NAO and why the skill increases for temperature.

Based on these considerations, I do not recommend publication of this manuscript.

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-199>, 2018.