Review for "Exploiting the signal to noise ratio in multi-system predictions of summertime precipitation and maximum temperatures in Europe", Acosta-Navarro and Toreti

General comments

This study reveals how the ratio between ensemble mean and ensemble spread in dynamical seasonal forecasts for boreal summer can inform in advance about potentially large temperature departures from climatology. This is a valuable asset, considering the limited predictability over Europe. The paper also indicates that this metric can help identify potential predictability drivers, with the example of a North Atlantic spring SST Atlantic preceding anomalously warm summers over Europe.

Overall the scientific question is simple and clear, and the results are quite convincing. Additionally, the paper is well written and enjoyable to read. This paper deserves to be published in WCD, after considering a few mostly minor revisions.

The introduction and discussion sections are particularly short. It is an asset for the introduction since it is sufficiently precise to properly state the scientific question addressed, and goes straight to the point. On the other hand, the results shown throughout the paper raise a number of questions that would deserve to be at least pinpointed in the discussion, if the authors want to keep the scope as it is. The main point to address concerns the section focusing on predictability source, and is detailed hereafter.
Finally some figures need to be be improved, as suggested below, to facilitate the reading and understanding.

Specific comments

L.42: "Tmax may be seen as a seasonal integrator of heatwave activity".
No doubt that your approach is valid, but wouldn't it have been simpler to use the seasonal average of daily mean temperature, rather than daily maximum temperatures ?
The departure of daily mean temperatures from climatology also permits to identify heatwaves, and it integrates the full diurnal cycle, which may be relevant since night-time temperature is also impactful during heatwaves.

L. 86-88 : This is not obvious from your figures. Considering this comment, we probably expect to see blue shades over NW Europe in fig 2.c but it is barely visible on your maps: the contrast between dark gray and blue is weak. I would suggest either to improve the maps (you may remove country borders, or pick a lighter gray shade for non significant values, or alternatively hatch/stipple where values are statistically significant.) or to provide additional support to your assertion.

If fig. 5d shows the mean 1994 and 2003 JJA, one might wonder if the SST dipole in the Atlantic is present equivalently in both summers, or prevails in one of the two summers. Since spring 2003 does not stand out as a season with particularly high spring SST index (fig. 5c), this question is even more intriguing. You could provide in appendix or supplementary material separate maps for both summers and comment on the specific case of 2003 (weak dipole in spring, stronger in summer ?)

L. 144-146: In my view, the subsampling of members based on the May index for re-forecast initialized in May makes no sense because the ensemble spread is presumably too small. I guess Mid-Atlantic SST anomalies are somewhat persistent, and it takes time for ensemble members to diverge from each other (and from SST initial conditions too). I think a discussion on the spread would be valuable here.

Overall, in section 3.2, the distinction between spring and summer SST patterns is a bit overlooked, and gives an impression of 'cherry picking'. Which one (spring or summer) is better correlated with warm summer anomalies over Europe ? In figure 5 you build your index on spring patterns, then you show summer patterns for 1994 and 2003, but the 2003 spring index does not stand out, and this point is not even commented on. Finally, in the subsampling procedure, you admit that it does not work when the subsampling is based on the May SST index. Besides the question of the relevance of subsampling too early after initialization (see point above), how could your findings help forecasters in real-time ? Would it be more advisable for them to systematically subsample their summer forecast ensemble, even if there is very little match with the SST pattern, or should they (also? instead?) evaluate a kind of distance metric or spatial correlation for each member to the SST dipole pattern, because maybe the SST pattern is not relevant every year.

References:

The reference of the MeteoFrance system is the following
Batté L., L. Dorel, C. Ardilouze, and J.-F. Guérémy, 2017: Documentation of the METEO-FRANCE seasonal forecasting system 8. Météo-France, 36 pp., https://www.umr-cnrm.fr/IMG/pdf/system8-technical.pdf.

L.38 Since you have retrieved the forecast data from the C3S Climate datastore, I think you should cite it and acknowledge properly, as recommended here:

https://confluence.ecmwf.int/display/CKB/
How+to+acknowledge+and+cite+a+Climate+Data+Store+%28CDS
%29+catalogue+entry+and+the+data+published+as+part+of+it

And here: https://cds.climate.copernicus.eu/cdsapp#!/dataset/seasonal-original-single-levels?tab=doc

Technical corrections

L.48 Typo: It should read "from ERSSTv5 (Huang et al., 2017) and ERA5, respectively."

Figure 1 : according to fig. 1c and 1f, a few grid points have SNR values exceeding 1. Thus, this should also show in the colorbar of fig 1b and 1e which looks bounded by 1

Figure 5 : in the caption, the reference to subplot (e) is missing

Figure 6: When zooming in, I can see gray dots everywhere. This figure should be improved. In addition, the color bars are different for temperature and precipitation. The same color bar would help compare the skill improvements for both variables. Finally, the color bars themselves are confusing: they should be improved by aligning the labels with color bounds.

L. 110-111: There is a mistake in the description of the ERA5 gray line in Fig5a/b, which gave me a hard time to understand: it is not the absolute value of the standardized ERA5 Tmax anomalies that you show, but the absolute deviation from mean! Luckily, the caption of figure 5 is okay.

L. 116: Missing specification that this statement refers to fig.5b

L. 117-119 : For these years with high ensemble SNR and high absolute anomalies in observations, I assume that the sign of the model re-forecast anomaly is the same as the sign of the observed anomaly. If so, this should be stated in some way, or overlaid in fig 5a/5b if feasible.

L. 142 : Fig 6a,c and 6b,d, not 5

L.145: typo 're-forecasts'