



EGUsphere, referee comment RC1
<https://doi.org/10.5194/egusphere-2022-183-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-183

Constantin Ardilouze (Referee)

Referee comment on "Exploring the relationship between temperature forecast errors and Earth system variables" by Melissa Ruiz-Vásquez et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-183-RC1>, 2022

Review for "Exploring the relationship between temperature forecast errors and Earth system variables", Ruiz-Vásquez et al.

General assessment :

This manuscript investigates the main sources of temperature subseasonal forecast skill at the global scale. It evaluates the relationships between potential drivers of three kinds (climate, circulation and land surface) derived from multiple observational/reanalysis datasets, and subsequent temperature forecast errors of the ECMWF extended range reforecasts, at different seasons. Overall, climate drivers tend to prevail, but land surface drivers also greatly contribute to forecast errors. Circulation drivers seem less relevant although not everywhere. Finally, based on correlation strength and forecast error amplitude, the authors highlight regions where subseasonal forecast skill could be potentially improved across seasons.

The scope of this study is both original and of great interest for the S2S community. The underlying rationale is relatively simple, but relevant too, so that I consider it an asset here. I would also like to stress that the paper is well written, well articulated, clear and enjoyable to read. My one main concern is about the evaluation of forecast errors. I feel like the metric used is not well suited for such a study where a distinction is made between seasons (see details below). It may have a limited impact in terms of the Spearman correlations found, but probably not on those of section 3.4. I think the authors should consider either correcting this metric or justify its relevance in the light of my comment below.

Main comment:

L. 150-151: I have the feeling that smaller errors found in transition seasons could also be due to the metric used here. This metric is based on departures from annual average temperature (i.e. no annual cycle). Therefore, for a given year, over mid and high latitudes at least, the annual average must be relatively close to fall and spring average temperatures, but substantially higher than winter and lower than summer average temperatures. Consequently, summer and winter forecast departures from annual average temperature must be generally higher (in absolute value) than fall and spring counterparts, both for forecasts and observations. Finally the amplitude range of the resulting forecast errors is probably larger as well.

I am puzzled by this choice of annual average temperature to compute departures here. Why not use weekly (or at least monthly) climatologies based on the 2001-2017 period instead ? This would have avoided the issue of seasonality and that of changing the reference every year.

Two more minor comments on this metric or its interpretation are reported below.

Minor points:

L.66-67: To what extent could the use of 2 model versions affect your results? I understand that initialization is unchanged, but readers not aware of the changes between these 2 model versions might wonder if they concern, say, a land/vegetation scheme or/and a key atmospheric parameterization for example. Such changes are prone to modify the relationships studied in this manuscript. If possible, I would suggest defending this particular point.

L.72: across S2S literature, the definition of leadtime weeks vary: some studies exclude days 1 to 4 after initialization, so that week 1 is defined as the day-5 to day-11 window (e.g. Vitart 2004, de Andrade et al. 2021). Could you specify - and discuss if need be - your method in this respect ?

L.100: I am not sure how T_{for} (annual average) is computed. If I understand well, you have two 6-week forecasts per week, i.e. 104 forecasts x 6 weeks per year. Not to mention the ensemble members. How do you proceed to compute the forecast annual average temperature and ensure it is comparable with observational average (one realization, by essence) ? I would recommend to be more specific, and also to state somewhere in the manuscript that forecasts are actually ensembles, and how these ensembles are handled here. I guess you have been dealing with ensemble means, but this needs to be specified somewhere.

L. 156-158: Agreed, but could it also be related to the greater temperature variability over mid-latitudes? I mean that if you would compute the seasonal average of $(T_{i,for} - T_{for})$

absolute values, for each grid cell, I expect these values to be lower at low latitudes, and therefore, the forecast errors end up lower as well (see e.g. Extended Data Fig. 7 and 8 in Tamarin-Brodsky et al. (2020)). I am not 100% affirmative but since you have not normalized temperature anomalies with their standard deviation, this could explain some (most?) of the meridional gradient depicted in Fig. 2.

Table 1 layout: I would suggest to make it clearer that when the column "Source" is empty, it means "similar source as above". Maybe a double quote could do ? And also, if possible and if allowed by the editor, try to reduce the font to have less line breaks. This would ease the reading.

Figure 6, NA region, DJF season: by eye, significant correlation seems quite unlikely although this may be due to a "Pearson correlation oriented" perception instead of Spearman.

Why not apply a lighter color shade, or a dashed style for instance, to the smoothing lines corresponding to pixels without significant correlation? Alternatively, you could indicate in the subplots the percentage of pixels of each region with significant correlation.

L.159: typo: parenthesis issue

L.292: I am not sure it is correct to describe NAO and MJO as "ocean phenomena"

Figure 6 : last row (SA region): the x-axis tick labels are arguably wrong (no positive values)

Another question that comes to mind when reading your conclusion is the extent to which the same Earth system variables would contribute to explain temperature forecast errors in the same regions for other S2S forecast systems. For example the spatial patterns of subseasonal temperature forecast skill show similarities between models and some predictability drivers are known to impact the same regions for different models (e.g. Ardilouze et al. 2021). I understand this would go way beyond the scope of this study, but I mention it for consideration.

References:

Tamarin-Brodsky, T., Hodges, K., Hoskins, B.J. *et al.* : Changes in Northern Hemisphere temperature variability shaped by regional warming patterns. *Nat. Geosci.* 13, 414–421,

2020

Vitart, F.: Monthly forecasting at ECMWF, *Mon. Weather Rev.*, 132, 2761–2779, 2004

de Andrade, F. M., Young, M. P., MacLeod, D., Hiron, L. C., Woolnough, S. J., and Black, E.: Subseasonal Precipitation Prediction for Africa: Forecast Evaluation and Sources of Predictability, *Weather Forecast.*, 36, 265–284, 2021

Ardilouze, C., Specq, D., Batté, L., and Cassou, C.: Flow dependence of wintertime subseasonal prediction skill over Europe, *Weather Clim. Dynam.*, 2, 1033–1049, 2021