Final Authors Comments to Reviewer 3

Comment on egusphere-2022-449
Anonymous Referee #3

Referee comment on "Seasonal forecasting skill for the High Mountain Asia region in the Goddard Earth Observing System" by Elias Charbel Massoud et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-449-RC3, 2022

This paper examines the seasonal prediction skill of the NASA's Goddard Earth Observing System (GEOS) S2S prediction system over High Mountain Asia for a series of hydrological variables. A set of observational data and the MERRA-2 and ERA-5 re-analyses are used as benchmarks.

Monthly means of reforecasts over the period 1981 to 2016 are analysed.

The paper provides a clear and well written description of the system and presents a clear analysis the forecasts of these hydrological variables. It provides a solid detailed reference to further assess more specific aspects of the performance of the GEOS S2S prediction system over HMA. The Figures are clear but, while many detailed descriptions are provided, I find the paper a little thin on new science and on understanding potential sources of predictability. At the very least, some more discussion points should be included.

Author Comments: We very much thank the reviewer for the time spent on our manuscript. We will aim to include more discussion points that address this concern in the next version of the paper, including the addition of evaluation for specific sub-regions as well as the addition of new text in various parts of the manuscript.

MAIN COMMENTS

1) The skills over this large, heterogenous region are small, especially when measured against independent data. I wonder if such small skills are relevant at all. The prediction system may have higher skill in some variables, in more limited domains or at certain times of year.

Low skill is not unexpected given the large area, the varied topography (from lowlands to high mountains) and land cover, and the different regional climates of this large region. For example, only a small Southeastern part of the Tibetan Plateau (and hence, only parts of the HMA) is influenced by the Indian summer monsoon (ISM).

It is well-known that the skills depend on the verification data, and they would be higher, in most cases, when verified MERRA-2 given the parent model. It would be of interest to verify some variables against several datasets, or better, against merged datasets that

take into account uncertainties in the various observations. I realise that such merged dataset might not exist over this region, but the point could be mentioned in the Discussion.

Author Comments: We thank the reviewer for this insightful comment. We do understand how the skills reported here can seem minor and not relevant, however we would like to point out that the skills shown are based on anomaly skill (e.g., anomaly correlation and unbiased RMSE). Therefore, any skill that is above 0 anomaly correlation is relevant, since it essentially means that it is more skillful than just guessing the mean climate. For S2S predictions at a range of 1-3 months, we believe that this is rather useful and promising.

As for gaining results in the more limited domains, our results in Figures 5-12 show maps of the skill and the ensemble spreads for the whole domain, which can be used to analyze spatial differences in the results and to pinpoint specific regions with higher or lower skills in general. Figures such as these can be used to interpret the skill level for different regions within the domain, such as low- or high- elevation regions. For example, Figure 7 shows that there is a higher ensemble spread as well as a higher error for temperature in the India subregion compared to other regions in the domain. To help further satisfy this concern, we will provide new plots showing the evaluation metrics for smaller sub-regions within the domain and will provide more discussion about these results.

Regarding the use of multiple data sets or perhaps a merged data product, this is the reason we use multiple sources of observed and verification data in our study. Relying on solely one source of information may be misleading, hence we used 2 different products for each climate variable to get a sense of the uncertainty in the forecast skill for each variable. We are not aware of a merged data product for this region, perhaps the MERRA-2 product is a close example of what the reviewer is describing here. In the discussion, we will add statements about the potential of having a data set like this and how it can be helpful for the evaluation.

2)Concerning actual societal needs, the reliability of such forecasts is a question of utmost importance that needs to be addressed in a probabilistic context. Is it possible to quantify the reliability of the forecasts with the current system using standard metrics? At least, the outlook could be mentioned in the Discussion.

Author Comments: We thank the reviewer for this comment, and we agree that the reliability or the uncertainty in the reported skill is important especially for societal needs. For this statement, we point to the ensemble spreads shown in Figures 5-12, which show the spread and therefore the 'reliability' of the forecasts from the model. The higher the ensemble spread in these plots, the less certain the various ensemble members are for each climate variable and lead time. Furthermore, the use of multiple data sources in the evaluation also allows us a look at the uncertainty in our results. We will include a statement in the discussion to elaborate on the interpretation of uncertainty and how it impacts societal needs. We will also provide additional text in the discussion section describing the reliability of the forecasts and how that can be estimated using standard reliability metrics. For example, we will provide new figures

showing the comparison between the ensemble spread and the error. Generally, one can compute the spread/error ratio with the goal of that being close to 1; if it is larger than 1 (more spread than error) this is considered "underconfident", and if it is less than 1 this is considered "overconfident" (Fortin et al., 2014). We will provide new plots and relevant references to provide additional and meaningful verification metrics.

3)I wonder about the relationship between surface temperature and the snowpack. Is there is a strong coupling between the two in the forecasts during some months? This could provide a source of skill.

Author Comments: We thank the reviewer for this insightful comment. We do not find any direct relationship between the skill in temperature and how it affects the skill in the snowpack. This is an idea that deserves some discussion, and we will provide some additional details in the paper to discuss this.

4) Improved prediction of the circulation could lead to improved skill. The authors mention the importance of the ISM. I believe that wintertime precipitation over the northern part of HMA is brought by the so-called westerly disturbances. The authors could mention in the Discussion, whether the dynamics and the associated with precipitation is well represented in the forecast.

Author Comments: We again thank the reviewer for this insightful comment. We do report on the importance of the ISM in Section 4.2.2. We will include some details about westerly disturbances and how properly representing that process within the model can impact the forecast skill.

5)There has been a significant effort in recent years to assess the impact of land initialisation (esp. snow, soil moisture) in S2S and seasonal forecasts and some studies are relevant for the HMA region yet there is little mention of that relevant literature.

Koster, R. D., Mahanama, S. P. P., Yamada, T. J., Balsamo, G., Berg, A. A., Boisserie, M., et al. (2011). GLACE2: The second phase of the global land atmosphere coupling experiment: Soil moisture contribution to subseasonal forecast skill. Journal of Hydrometeorology,12(5), 805–822.

Senan, R., Orsolini, Y.J., Weisheimer, A. et al. Impact of springtime Himalayan–Tibetan Plateau snowpack on the onset of the Indian summer monsoon in coupled seasonal forecasts. Clim Dyn 47, 2709–2725 (2016). https://doi.org/10.1007/s00382-016-2993-y

Author Comments: We thank the reviewer for providing these works in the literature. We do report on the impact of land initialization in Section 4.1, and we will also read the works listed here by the reviewer and we will aim to include more comments in the discussion based on these papers.

MINOR COMMENTS

- The words Seasonal forecasts and S2S forecasts seem to be used loosely throughout the paper. The seasonal forecasts are 9-month long but only the first 3 months are analysed. Some operational centers have different set-ups for Seasonal and S2S prediction systems. The authors could double check that S2S is used as it is meant.

Author Comments: We thank the reviewer for this comment. From our understanding, subseasonal forecasting refers to predictions at 10 days to 1 month, and seasonal predictions refer to predictions at 1 month to 9 months lead time. We will do a more thorough check about this definition and make sure our paper is consistent in the terminology throughout.

- It was not clear to me whether total precipitation is liquid precipitation or if it contains also solid precipitation.

Author Comments: The form of precipitation analyzed in this paper is the PRECTOTCORR variable from the GEOS-S2S system, and is derived from the PRECTOT variable, which is total precipitation including rain and snow, i.e., PRECTOT = liquid + solid (total) precipitation. We will include a comment in the paper to better explain this.

- The information on ensemble size should be presented more clearly (Abstract, or Table)

Author Comments: We thank the reviewer for this comment. We will provide a better and more clear description of the ensemble size.

Wording

L58: the foothills of the Himalayas perhaps better than the foot of the Himalayas (?)

Author Comments: We thank the reviewer for this comment. We will fix this in the paper.

L560: precipitation is used twice in same sentence.

Author Comments: We thank the reviewer for this comment. We will fix this in the paper.