



EGUsphere, referee comment RC2  
<https://doi.org/10.5194/egusphere-2022-419-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on egusphere-2022-419**

Anonymous Referee #2

---

Referee comment on "Long-term evaluation of the Sub-seasonal to Seasonal (S2S) dataset and derived hydrological forecasts at the catchment scale" by Marianne Brum and Dirk Schwanenberg, EGU sphere, <https://doi.org/10.5194/egusphere-2022-419-RC2>, 2022

---

### **Summary:**

The authors present an evaluation of forecast skill of an sub-seasonal 2 seasonal prediction for two regions in Central Germany. The analysis encompasses nine numerical weather prediction (NWP) models. The authors use a set of well-established forecast skill metrics for this purpose. For both of these regions, the authors analyse meteorological forecast skill for precipitation and temperature. For one of the regions, a hydrologic model (HBV) is applied and forecast skill is evaluated for discharge forecasts. The main findings of the study are that hydrologic forecasts are skillful at longer lead times at which meteorological forecasts are not skillful anymore. The study also shows that forecast have become better over time (i.e., better in 2020 than in 2015) with a larger gain for meteorological variables in comparison to hydrologic variables. The nature of the study is more a report than a scientific exploration. For example, the authors provide no explanation why the forecast skill is higher for discharge than for meteorological variables or why the forecast have gained skill over time. The authors could have performed experiments with HBV to show that it is indeed the hydrologic inertia that results in higher skill at longer leads. Similarly, the authors report that forecast gain over time is present, but do not explain why this is happening. Which changes have been applied to the forecast system that could explain the gain. If no change to the forecast system has been applied, this finding might well be an artifact of the experimental design and truly, no gain is present. These substantial points have to be addressed thoroughly before the manuscript can be published. The manuscript is well-structured and overall well written and easy to follow. I only found the figure captions too short and also lacking an introduction of abbreviations used in figures.

### **Major comments:**

The selection of Rhineland-Pfalz (RLP) for evaluating the meteorological performance seems arbitrary. As seasonal forecast are global in nature, this evaluation could be done at that scale. RLP is also very close to the selected catchment (both are located in

Germany a few hundred kilometers apart) and share very similar climate. Please consider using another setting for the meteorological evaluation.

One of the major points is the gain in forecast performance from the beginning of the forecast period (2015) to the end (2020). 5 years is a very short period that is insufficient to characterize the climatology of the investigated catchment (upper Main in Germany). It is unclear whether the gain in skill is an artefact of catchment observations. In other words, it could well be that the upper Main was easier to forecast in 2020 because of the weather that occurred in 2020 than in 2015. The authors need to elaborate on the reasons why they think it is actually an improvement in the system. It is surprising to me that the meteorological lead-time gain can increase by 20 days over such a short period. This means that a forecast at lead time 40 in 2020 is as skillful as a forecast at lead time 20 in 2015. This would be an enormous improvement.

Table 3: Within this table, only the correlation of forecast feature with the highest correlation is shown and others are not reported. It would be necessary to know the correlation of the other model features too to judge how much better the model feature with the highest correlation really is. Additionally, the authors should conduct a statistical significance test to demonstrate that the outperformance by the model features with the highest correlation is significant.

In general, the Figure captions are too short to understand the Figures and need to be expanded. Additionally, abbreviations used in figures are not explained (kwbc in Figure 4) and coloring of lines is also unclear.

### **Minor edits:**

#### *Abstract*

- should mention Germany as location of catchments
- should provide reference to S2S project

#### *Intro:*

- paragraph starting at line 48 can be merged with paragraph ending at line 33. They cover the same points.

- line 54: which time? the past decades or lead time.

- line 56: what do the authors mean by "definite model predictors"?
- line 58: What do the authors verify here? The sentence seems to be incorrect. Verify is also not an appropriate word in the context of modelling. Maybe better choose validate.

### *Materials and Methods*

- line 74: please modify sentence to make clear which statistic corresponds to which gauge
- line 98: it is not clear how temperature dataset has been interpolated from the stations to the grid scale.
- line 125ff. Some metrics that are introduced here are not used in the results section, for example 'value'. These should be removed.
- line 155: HBV uses a triangular weighting function for the river routing. It is not clear to me how this is used to connect the nodes N0, N1, N2, and N3. It needs to be clarified what these nodes represent.
- Line 157: It is not clear to me how a data assimilation (DA) is run for HBV. HBV is a conceptual hydrologic model where model state variables cannot be mapped to observation-based variables. Please expand. It seems like DA improved the performance (see line 166ff), but the model performance even without DA is very high and certainly within the range deemed as useful for end users. Please clarify why DA is necessary at all for this application.

### *Results:*

- Line 178: it should be stated that Q97 is not shown
- Line 180: The sentence starting at 'Persistence, on the...' is not clear to me. Could you please rephrase.
- Line 198: please state that this statement is only valid for the meteorological forecasts.
- Figure 3: it is confusing that the name for both gauges contain Schwuerbitz in the name.
- Figure 3: kwbc is not explained
- Line 239ff: It is unclear to me how the composite-model lead-time gain is calculated.