

Review comments for **Skill of seasonal flow forecasts at catchment-scale: an assessment across South Korea** by Lee et al.

In this paper, the authors have conducted an evaluation of seasonal hydrological forecasts using Tank model driven by the ECMWF SEAS5 seasonal weather forecasts in various South Korean catchments. The topic is timely and brings insights to the field of hydrological forecasting. However, I have identified several areas that require attention to ensure the manuscript meets the standards of the journal.

1. The manuscript currently needs a clearer explanation of the methodologies employed, which also relates to the second point of terminology use. Some parts of method description need to be expanded. For example, the calculation of contributions from each variable needs more details. When examining the contribution of forecasted precipitation by substituting the other two variables with observations, please clarify if this was done for all ensemble members, and is it skill that you are compared on or score?
2. Consistent Use of Terminology: The paper misuses the word “*skill*” in different ways which should be sometimes specified by *performance or quality*, and this inconsistent use lead to confusion unfortunately. There seems to be also some ambiguity in the use of “*skill*” and “*score*”. For instance, “theoretical skill” and “actual skill”, are they referred to CRPS or CRPSS? If they are skills, could you specify which benchmark is being used to calculate them? If they are scores, then I suggest not using skill in the names. I would suggest the authors to clearly distinguish these terms and maintain consistent usage throughout the manuscript to improve clarity.
3. Abbreviation Usage: The manuscript needs a thorough review to ensure that all abbreviations are properly introduced upon their first occurrence. Additionally, to avoid redundancy, each abbreviation should only be defined once. Moreover, certain abbreviations have been assigned multiple meanings within the manuscript (for instance, CRPS at Line 183 and Line 258). This presents a significant source of confusion.
4. Actual Skill and Theoretical Skill: The authors raise an interesting point about the significance of using actual skill over theoretical skill to provide more insights for water resource management on whether to use SFFs and when. However, in the analysis, this is conducted by calculating CRPSS, with ESP as benchmark, thus the use of either actual or theoretical references does not play such a big role, as long as the benchmark is using the same reference as the forecasts. On the other hand, the information gained from theoretical skill in this paper is to validate the performance of the hydrological model by showing its proximity to the actual skill (or perhaps more appropriately, the “actual score”). In this case, it didn’t really reflect the argument of providing significant information for the users.

Line by Line comments:

Line 23, “actual skill” here sounds ambiguous since there is no other information explaining this term, which might lead to misunderstanding.

Line 25, please add brief information on the methods that you use to get the conclusion that precipitation is the most important variable.

Line 57, this is the first time that ESP is mentioned (excluding abstract), therefore full explanation is needed here.

Line 77, to my knowledge, the reference Pechlivanidis et al., 2020 is not using ESP in the analysis, therefore cannot support the argument here.

Line 93, I'm a little bit suspicious on this sentence here that "only a few studies" have used SEAS5 for seasonal hydrological skill assessment. For example, the reference you mentioned before Pechlivanidis 2020 is actually using SEAS5 at higher spatial resolution.

Line 139, what is the criteria of dividing the four seasons, are they based on precipitation or flow?

Line 142, the information of annual variability is not shown in the figure but only in the text, right?

Line 143, typhoon and monsoon might not need to start with an uppercase character here.

Line 149, the abbreviation of KMA should be noted in the previous sentence when it is firstly mentioned.

Line 169, regarding SEAS5 data, here the period 1993-2020 is mentioned, but in the method part and in Figure 2, based on my understanding, the forecast period is 2011 to 2020. Please clearly specify this.

Line 181, SFFs has been mentioned many times already.

Line 183, here CRPS is referred to as skill but later it is referred to as score (Line 258).

Line 188, the plot needs to be improved. To calculate CRPS needs the forecast (either ESP or SEAS5) and the reference (either real or pseudo observation), therefore the arrows should lead from corresponding systems to the box of CRPS. However, this is not systematically shown in the plot.

Line 190, to my knowledge there is SEAS5 forecasts with higher spatial resolution that is available.

Line 205, a potential problem for linear scaling on precipitation is, it might generate very large values. Have you had any solutions to avoid this?

Line 247, as defined in Eq.4?

Line 265, what does SPF stand for? Or maybe you mean SFFs? Otherwise please add the full name for the abbreviation.

Line 270 and Line 258, redundant information.

Line 275, Major does not need an uppercase here.

Line 275, here the CRPS of ESP is calculated using real observation as reference, it is correct?

Line 285, here comes the explanation of SPF, but it is already mentioned many times before this.

Line 310, here I would strongly recommend to distinguish skill from score, since you have CRPSS later which are actually skills, but here these are scores.

Line 327, this part should be described in method session, and more details are needed for fully understanding.

Line 498, are these conclusions from Figure 8? Considering there are only two dry years and two wet years, the conclusion needs to be drawn carefully, otherwise it's not very scientifically valid.

Figure S1, please explain which benchmark is used here to calculate from CRPS to skill (skill-3P, skill-T).