Interactive comment on "Skill of a global forecasting system in seasonal ensemble streamflow prediction" *by* Naze Candogan Yossef et al.

Naze Candogan Yossef et al.

ncandogan@hotmail.com

Received and published: 14 February 2017

Anonymous Referee #1:

We thank Referee #1 for her/his kind words and valuable comments.

With regards to the specific comments and corrections, we understand that we inter-mingle the terminology of GMC and NWP. We will clarify this accordingly and provide a consistent description of the meteorological forecasting system used.

The choice of the 20 major river basins is a continuation of earlier analysis (Candogan Yossef et al., 2012, 2013) and this selection covers a wide range of hydro-climatic conditions. We will provide a short explanation. In the original paper, also the location

of the GRDC stations is given. We will summarize this information here and include a brief description of the nature of the datasets used. We will clarify that we compare discharge values at the station location in order to allow us to determine both the theoretical and actual skill from in our analysis.

The referee rightly asks why we have used the ECWF S3 forecast and not the newer, potentially more skilful S4. The only reason is that S4 was not available when the project started. Still, one would expect the S3 temperature and precipitation to be sufficiently accurate. While the forecast skill might improve using S4, we believe that the current study using S3 provides important insights.

We also thank Referee #1 for her/his suggestions to improve the tables and figures in the manuscript. We will clarify in the text and in the table captions which BS is presented. In the main text, we will summarize the tabular information to its essentials and relegate the full tables to an appendix. We will clarify the legend of the global map (Figure 1) as suggested by the referee.

Anonymous Referee #2:

We thank Referee #2 for her/his kind words on the content and organization of our manuscript and the suggestions to improve its formulation. We understand that one of the major comments concerns the terminology used re verification. We value the suggestion to further extricate the BS and interpret it in terms of the forcing and hydrological uncertainty.

In our simulations we used the observed forcing over the period 1979-1984 to spin up the model and then ran the model starting from these initial states with the ECMWF S3 seasonal forecasts for the period 1979-2010. For each of these simulations, we excluded the first two years and limited the subsequent analysis to the period 1981-2010 in order to avoid any further bias. We did not include the 1981 member in the ESP ensembles as suspected by the reviewer and we will clarify this in the revised manuscript. We will clarify this better in the text.

We understand that the number of tables is too large for the main body of the text and we will change it accordingly. The suggestion to relegate the BS values to an appendix may be considered, and the same holds for the different basins. In the design of the figures we experimented with different lay-outs, colour-codings etc. to present the material concisely. We will re-evaluate this and try and consolidate the information to its essentials.

We will comprise a brief description of the 20 major river basins used and underline the continuity with our earlier work. We wished not to dwell too much on points raised in earlier studies and therefore focused on the skill assessment here. However, in light of the distinction between forcing and hydrological uncertainty suggested by Referee #2, aspects of flow regulation can be covered in the discussion. The specific comments raised by Referee #2 will help to improve the clarity of the paper and we will accommodate them within the revised manuscript.

Anonymous Referee #3:

We thank Referee #3 for her/his suggestions and comments that certainly will help us to improve the quality of our manuscript. We will address these comments here and hope to make a stronger case for the conclusions we had drawn originally.

We thank Referee #3 for the number of relevant and interesting articles and we will include these with other references in our introduction.

In the workflow of the ESP we indeed excluded the year under consideration out of the ensembles. It is unfortunate that this is not completely clear from our description and we will improve this in our revision. We acknowledge that the selected 25th and 75th percentiles of the streamflow are not very extreme. However, it should be noted that we consider monthly discharges and, therefore, using the upper and lower 10% over 32 years provide far too small number of events for the analysis.

The suggestion to use the Continuous Ranked Probability Skill Score (CRPSS) is in-

teresting and a viable solution to the BSS. However, we deem that the differences between the thresholds used in the Brier Score (BS) are not too different, and, although the remark justified, the practical implications are not too grave.

As mentioned by the other referees, Referee #3 suggests to move the tables to the Supplementary Information. We will do this accordingly and review, in light of all comments, the possibility to update the figures and to present our findings as clear and concise as possible. We wish to stress that already in writing the current manuscript much time and effort were spent to present the material as briefly as possible without jeopardizing the transparency of our outcomes.

Concerning the verification of S3 dataset, we will include the references to the verification carried out by ECMWF. We can expect the quality of the seasonal forecast to be initially fairly well, conditioned by SSTs and will start to deteriorate later on, as seasonal forecasts start to show signs of systematic model errors after about ten days into the forecast. As mentioned in the manuscript, the ECMWF applied a daily bias-correction based on quantile-quantile transformation but did not introduce any artificial terms in the equations to reduce the drift. In order to account for drift, we applied a bias correction using 12 datasets varying per forecast month, provided by the ECMWF. Therefore, we expect the temperature and precipitation to be reasonably correct, and we believe further verification of S3 forecasts would be beyond the scope of our study.

In terms of technical comments, we will update the description of the BS throughout the manuscript. Similarly, we will remove the occurrence of "actualized" in the text and refer to the Murray-Darling as such throughout. Equally, we will update the abstract.

———————————————