Response to interactive comment by Anonymous Referee #2

The authors have assessed three different methods - Bayesian Joint Probability (BJP), the Snowmelt Runoff Model (SRM) and a hybrid approach (SRM - Ensemble Streamflow Prediction inflow means as additional predictor in BJP approach) for forecasting seasonal streamflow to the two largest dams in the Upper Indus Basin, Pakistan. The authors concluded that BJP approach is simple and it worked well to provide probabilistic seasonal streamflow forecasts. The topic is relevant for publication in HESS. Overall, the paper is well written. I recommend a moderate revision to the manuscript and the following concerns need to be addressed:

Response: Thank you for your encouraging and helpful review.

Major Concerns:

1. Under the BJP approach, was the conditional multivariate normal distributions fit over the entire season or on monthly basis? How many samples were generated through Monte Carlo Simulations under the BJP approach? Provide details.

Response: We will modify the text to include clarifying information: "*The cross-validated BJP forecast performance was assessed for 1975-2015 (41 seasons), with the BJP models calibrated on a seasonal basis (i.e. 40 data points) using 1000 MCMC samples for each of the leave-one-out calibrations*."

2. Page 7, lines 24-27, the skill of using March flow and/or one climate predictor looks very similar to each other. The authors are recommended to use statistical significance test to compare if the skills are significantly different from each other.

Response: We will add uncertainty estimates to the skills scores, from bootstrapping, to aid interpretation of how different the forecast models are from one another.

3. Given that most of the streamflow at Indus River at Tarbela is snowmelt driven, use of a direct or indirect indicator of snow as one of the predictors, along with the projected summer air temperature can improve the forecasting skill. The authors are encouraged to consider global precipitation (for winter) and air temperature forecasts as predictors, which can represent snow as one of the inputs to the model.

Response: We agree that GCM forecasts of precipitation and temperature could potentially be used as predictors, however in this work we are assuming that the water resources practitioners do not readily have access to GCM seasonal climate forecast data (including hindcasts, needed for model establishment). Hence out approach relies on information regarding temperature and precipitation being captured by our selected climate index predictors (statistically). This could be the subject of future research with dynamical models, so we will mention this in our revised discussion.

4. It is not clear why MEI for May and Jun from previous year enhanced the skill score for Indus at Tarbela? Explain.

Response: The MEI (May-June) predictor skill relates to autumn/winter snow accumulation, a lag of 4 months to snow accumulation from October onwards. It is thus not unreasonable that circulation systems bringing moisture into the region during autumn/winter are influenced by the forcing initiated by ENSO processes during the summer.

There are many supporting references presenting details of ENSO/precipitation teleconnections for the region, with several cited in the manuscript. As an example, Mariotti (2007)[2] notes: "*The associated circulation pattern during El Nino (La Nina) involves a southwesterly (northeasterly) moisture flux that brings more (less) moisture into this region* [southwest central Asia]. *This flux flows along the northwestern flank of the large-scale high pressure anomaly over the Indian and western Pacific Oceans, broadly, the western pole of the Southern Oscillation see-saw pattern. Unlike many ENSO teleconnections in which lower pressure and weaker subsidence leads to more precipitation, this mechanism does not require a change in the local pressure, but rather involves a change in the tropical moisture supply to subtropical-midlatitude storms.*"

5. Page 7, line 1, how good or better the skill enhancement is if SSCRSP (or SSRMSE) changes from 21 to 24.3 (within moderate skill range in Table 1)? Does it reduce uncertainty? Clarify.

Response: As described in the verification methods section, improvements in CRPS reflect improvement in accuracy and/or sharpness and improvements in RMSE reflect improvements in accuracy of the median only. So an inference can be made through comparative analysis of the various skill metrics, including the IQR. We acknowledge that it is a small difference and will add uncertainty via bootstrap results to help clarify this.

6. In Table 3, it will be good to know the correlations that are statistically significant (e.g. at 95% confidence interval) based on the sample size.

Response: We will indicate statistical significance in the table.

7. Page 10, lines 2-7, the hypotheses listed are not clear. As mentioned by the authors earlier, it is already known the snowmelt plays an important role for Indus River at Tarbela. So it not a hypothesis. Also, the results indicated that adding NAO, when used as a predictor, did not improve forecasting skill.

Response: We will re-word to avoid the confusion caused by the term "hypothesised". Table 2 shows that the NAO predictor did add some skill, however not as much as the selected ENSO based predictor.

Minor Concerns:

8. Did the models use monthly (or daily) data for the model fitting? If so, it needs to be clearly stated.

Response: The BJP is calibrated to seasonal data (i.e. 41 data points 1975 to 2015). This will be clarified in the text.

RC2:9. Page 6, lines 27 – 30, RMSEP needs to be used instead of RMSE. Also RMSEP needs to be defined in the text.

Response: We will correct and add a citation that includes the derivation of RMSEP (Wang, Q. J., and Robertson, D. E.: Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences, Water Resources Research, 47, 10.1029/2010WR009333, 2011.)

RC2:10. In figures 3a, 4a, 5a and 6a, what are the bounding lines (is it 95% Confidence Interval)?

---

[2] Mariotti, A.: How ENSO impacts precipitation in southwest central Asia, Geophysical Research Letters, 34, 10.1029/2007GL030078, 2007.

Response: As stated in the figure caption, these are Kolmogorov 5% significance bands. We will clarify this in our revised text referring to these figures.