Reply to Reviewer #1 (Report #2)

Title: Leveraging multi-model season-ahead streamflow forecasts to trigger advanced flood preparedness in Peru

Author(s): Colin Keating, Donghoon Lee, Juan Bazo, Paul Block

MS No.: nhess-2021-25

The authors would like to again thank Reviewer #1 for the feedback on our revised manuscript. Our specific replies are denoted in blue color and revised manuscript text is denoted by italics.

1. With all the additions from the previous revision round, the paper is a little lengthy now and the chapters could be separated better. Especially the chapters 3.1 and 3.2 repeat content that is visible in the tables, and describe things that have not been used (e.g. l. 264 "Selecting SST regions based on the preseason state of the Niño 1+2 anomaly index instead of MEI did not materially change results at Piura", l. 305–307 "A quantile mapping approach (…) did not substantially differ (…)". Slightly cutting these unnecessary parts should be enough.

We agree with the Reviewer that the manuscript has increased in length, and have accordingly cut the sentences mentioned above as well as additional text to remove information repeated in tables and elsewhere in the manuscript.

2. Please either merge "Results" and "Discussion" to "Results & Discussion" or separate more strictly. The Discussion chapter should give an honest evaluation of the overall results, and put them into context by citing relevant literature. In my opinion, chapter 4.2 is Methods rather than Results, while 4.1 contains Discussion parts (e.g. comparison to Bazo et al. in l. 354). The Discussion then presents 3 more figures. Some repetitions could be avoided from merging the sections.

We have merged chapters 4 and 5 into "Results and Discussion" as suggested by the Reviewer and have moved section 4.2 to the methods section.

3. I doubt that the term "principal component regression PCR" is adequate to describe your method. In my understanding, the term PCR suggests that a regression is applied in principal component space, by which I mean that all variables have been included in the PCA, and the resulting regression is then transformed back into the original feature space. You are using only 1 PC of selected variables and most other predictors are regular variables. I would rather write of linear regression and a PC predictor component. Also I find it a bit pointless to always stress the "multiple" linear regression, as most people doing linear regression use multiple predictors. It's ok if you just use your abbreviation MLR. The formulation in l. 271 "coupled principal component analysis and multiple linear regression" is wordy.

We have revised several instances of "multiple linear regression" to "MLR." Additionally, we have clarified our statistical method by revising lines 267-281 to

The statistical forecast is composed of sub-models built only on data from years in a particular climate state, as represented by the preseason (3-month average) value of MEI. This produces two sub-models for the Marañón River at San Regis and three for the Piura River at Puente Sánchez Cerro. Each sub-model leverages a principal component regression (PCR) framework to predict seasonal (3-month) average streamflow derived from daily observations obtained from SENAMHI as described in Sect. 2.5. In this framework, a principal component analysis is conducted on eligible predictors (Table 2) which are first scaled to have a unit variance. A subset of PCs is retained according to North's Rule-of-Thumb (North et al., 1982) for input into a MLR model, however in all cases just one PC is retained, yielding a linear model of the form:

$$y_t = \beta_0 + \beta_1 PC_1 + e \,, \tag{1}$$

where y_t is observed seasonal streamflow in year t, β_0 is the intercept, β_1 is a fitted regression coefficient, and e is the residual or error. Predictors may be eligible for inclusion in some sub-models and not others, subject to their correlation with streamflow in that phase (Table 3). To be included, the predictor in question must be both significantly correlated with streamflow across all years and significantly correlated with streamflow in the subset of phase-specific years. A hindcast assessment is conducted by evaluating each year in the historical record using the appropriate sub-model to predict seasonal streamflow. For example, in 1998, the preseason (NDJ) average MEI value is 2.43, thus the positive phase sub-model is selected to predict Piura River FMA streamflow.

We prefer to keep the naming convention of PCR because we believe it aligns with prior literature's use of the term (e.g., Delorit and Block, 2017; Lins, 1985; Mortensen et al., 2018)

4. Please introduce abbreviations at the first occurrence of the term, and then always use the abbreviation afterwards. "Multiple linear regression (MLR)", "Threat Score (TS)", and others.

We thank the Reviewer for catching this and have abbreviated to MLR (l. 268, 270 and 279); EAP (l. 414); TS (l. 476, 480); POD (l. 480); FAR (l. 480).

5. Table 5 and Table 6 could be merged when arranging in rows rather than columns, similar to Table 3 (which looks good now). When doing so, it is much easier to visually compare the different models by all metrics. Consider to highlight the best score per metric in bold font.

We have merged Tables 5 and 6 (reproduced below) and have highlighted the best score per metric and site in bold. We agree this is a preferred illustrative approach.

Table 5: Mean RPSS, Pearson correlation coefficient, POD, FAR and TS for each location and forecast approach. Bold text indicates best score metric per site (ties between two models are both bolded).

	Statistical		**GloFAS**		**Multi-model**	
	Piura	Marañón	Piura	Marañón	Piura	Marañón
RPSS	**0.43**	**0.67**	0.18	0.25	**0.43**	**0.67**
Correlation	0.91	0.95	0.91	0.84	**0.94**	**0.96**
POD	**0.63**	**1**	0.38	0.5	0.5	**1**
FAR	0.29	**0.2**	0.25	0.5	**0**	**0.2**
TS	0.5	**0.8**	0.33	0.33	0.5	**0.8**

6. l. 46-54 The thematic jump from exposure/vulnerability to high temperatures in London requires rephrasing. As the rest of the article is about floods, the sentence should start with something like "In the context of heatwaves in London, (…)"

We agree and have revised lines 49-52 to

> *In the context of heatwaves in London, actions to reduce vulnerability for high-risk groups, such as ensuring indoor temperatures are below 26°C, are triggered when a forecast indicates temperatures of at least 32°C during the day and at least 18°C at night (Public Health London, 2018).*

7. l. 65 needs a comma. In addition, consider to split the long sentence after "protocols" (l.66)

Lines 64-68 have been revised to

> *In addition to short term weather forecasts, which are commonly viewed as skillful, medium to long range climate forecasts have also been demonstrated to improve preparedness protocols, resulting in reduced mortality, morbidity, and resource demands (Braman et al., 2013). However, their applications have been limited predominantly as a result of moderate forecast performance and significant uncertainty.*

8. l. 94 the assumption here is actually that the errors in individual models are uncorrelated. Correlated errors would not cancel out.

We agree and have revised (see comment 9).

9. l. 95 full stop after "individual model". Please rephrase the subsequent sentence.

We have revised lines 95-97 to

> *Multi-model techniques have been developed based on the assumption that individual model errors are uncorrelated, in which case a multi-model average*

could provide greater skill than any individual model. Options for combining models include equal weighting, linear regression, or Bayesian methods (e.g., Gneiting and Raftery, 2005).

References

Delorit, J. and P. Block (2017). Evaluation of model-based seasonal streamflow and water allocation forecasts for the Elqui Valley, Chile. Hydrology and Earth System Sciences, 21, 4711-4725. doi.org/10.5194/hess-21-4711-2017.

Lins, H. F. (1985). Interannual streamflow variability in the United States based on principal components. Water Resour. Res., 21, 691–701.

Mortensen, E., Wu, S., Notaro, M., Vavrus, S., Montgomery, R., De Piérola, J., … Block, P. (2018). Regression-based season-ahead drought prediction for southern Peru conditioned on large-scale climate variables. Hydrology and Earth System Sciences, 22(1), 287–303. https://doi.org/10.5194/hess-22-287-2018.