Review of Mortensen et al. (2017) in HESSD

A.) General Comments

Mortensen et al. (2017) presents a new statistically-based seasonal forecast model for prediction of the three month wet season (January-March) in Southern Peru, with a focus on drought. The statistical model is based on principal component regression (PCR) using 11 large-scale climate predictors derived from SST or SLP fields, such as El Niño 3.4. There is good justification for using PCR as the authors highlight it removes the well-known issue of multicollinearity suffered by traditional multiple-linear regression models. The model building and skill evaluation is robust. A long 51-year hindcast period was used, cross-validation for both model building and skill evaluation exercised, and multiple evaluation metrics considered such as correlation, RPSS, and tercile contingency tables as well as for a more extreme precipitation drought category. The hindcast results showed low to moderate skill improvements over climatology as well as a simpler ENSO based model, highlighting a more complex climate-precipitation link than often assumed thus the added benefit of considering a wider set of predictor variables. The model developed has direct practical application in Southern Peru, a region where the availability of seasonal forecasts is limited.

I found this paper interesting from a seasonal forecast and drought prediction perspective given the limited available alternatives in Southern Peru and indeed in many other regions of the world at this lead time and think it deserves publication in the HESS special issue on 'Sub-seasonal to seasonal hydrological forecasting'. There are however several aspects of the paper that need revised before publication. These mainly concern the lack of international literature cited throughout on seasonal/statistical forecasting, the non-standard HESS layout, and a few other minor comments/suggestions that I've outlined below together with some suggestions for improvement.

B.) Specific Comments

1. Throughout the text, including the title, the term 'drought' is used with many statements making the connection to a "hydrologic extreme" (Pg1; L17) and for example "the city's water supplies were reduced" (Pg2; L13-14) etc. It took me until the last line of the introduction (excluding the abstract) to realise it was prediction of meteorological drought (precipitation) rather than hydrological drought that was being pursued. Clearly improving prediction of meteorological drought is still valid, but precipitation deficit does not necessarily propagate to a soil moisture, streamflow, and/or groundwater drought, which are more societally relevant. Often there are more complex processes at play, including temperature/evapotranspiration feedbacks. See Van Loon (2015) for more detail. This should be acknowledged and please be more explicit about the focus on meteorological drought throughout. I think if you are more explicit within the abstract and main text I would not be pedantic about asking to change the title, as it is a good title.

2. I am not convinced the layout of the paper fits reasonably with traditional HESS style. Nor does it help convey the story of the paper as good as it could. I acknowledge the story here is a little more complex and awkward to fit into exactly the traditional style, for example I think it is appropriate to have Sect. 3 as 'Southern Peru Rainy Season and Large-scale Climate Influences', it works well. However:

    a) The introduction does not do the paper justice as it fails to clearly establish core research aims/objectives/questions. The fundamental finding of the paper is that the newly created PCR model was found to be more skilful that a climatological forecast AND a simpler Niño 3.4

index-based model forecast, especially for dryer conditions – Is this not the foundation of your research question(s)? If so, needs to be in the Introduction.

b) Even though drought prediction remains largely unexplored in Peru, there is no background nor reference to the international literature on general seasonal forecasting methods in the introduction section, nor previous work done internationally on statistical forecasting. For example, what is the justification for selecting a statistical forecasting approach over others (i.e. lack of climate/hydrological modelling, limited hydrological data,...)?

c) Methods and Results are scattered over several sections. Reforming methods into a 'Section 4 Methods' and 'Section 5 Results' together with the use of sub-headings would help. I especially found it frustrating to have Sect. 7 after the results section. Surely this should not be hard to have the results divided by sub-heading for 'Sect. 5.1 Season-ahead' and 'Sect. 5.2 Extended Lead Time and Spatial Disaggregation of Regional predictions', or something similar.

d) The 'Summary and Discussion' section does an excellent job of outlining the practical implications of the work. However, it does not discus results in light of the international forecasting literature. Is the degree of increased skill on par with other areas/forecasting approaches?

3. Pg3; L5: The elevation range in Peru is substantial. It would therefore be beneficial for an international audience to provide the mean or median and the range of elevation for the 29 precipitation stations.

4. Pg 4; L3: what is the average correlation and is the method used Pearson?

5. Pg 4; 21: Agree a focus on JFM is justifiable. To help convince the reader that forecasting the wettest season is relevant to water resources/drought perhaps worth mentioning that it is during the wet season that reservoir/aquifer stores are replenished for use during dryer summer months. Being able to skilfully forecast anomalously low precipitation for the wet season is indeed valuable.

6. Pg4; Fig. 2 caption: Add "…using data from 29 precipitation stations in Sect. 2".

7. Pg5; L2: Referring to both EOF and PCA throughout. Stick with one to avoid confusion.

8. Pg5; L7: The Eklundh and Pilesjö (1990) reference is for Ethiopia. Should reference go after "homogeneity" instead?

9. Pg5 & 6; Fig. 3 & 4: It was not clear to me why PCA was used for the observed JFM precipitation totals? What purpose does it serve if the main target for the PCR model is for areal averaged precipitation totals anyway? Also, it is not stated what the physical interpretation of PC2 and PC3 are in this context (i.e. Pg9; L26 & Pg10; L9). As it stands Fig. 3 does not really add anything. It is too difficult to see any difference the size of the red dots. Perhaps adding a scale and/or some gradual colour scale would help? Could you include what elevation threshold the topographic shading represents?

10. Pg6; L12-13: Am I correct in thinking none of the precipitation stations used here are within the rain shadow?

11. Pg6; 20: Are strong El Niño's associated with only low precipitation or are they associated with actual societally impactful droughts such as problems with agriculture, water supply etc.?

12. Pg 7; Fig. 5: Add units of SST anomalies (i.e. $^{\circ}\mathrm{C}$)? You could improve plot by adding two horizontal lines to represent El Niño/La Niña thresholds (i.e., $\pm 0.5^{\circ}\mathrm{C}$). Also define that you are using

Pearson's correlation coefficient (I presume?) in the first instance (in the text) and define symbol as r. Then use r in every instance throughout the paper for clarity. I note this is done in some places and not in others (e.g. Pg9; L30).

13. Pg8; 5-8: Very short paragraph, better added to the previous one?

14. Pg9; 21: Change "previously identified" with "identified in Sect. 3".

15. Pg10 & 11; Fig. 7 & 8: Worth adding a dot/circle to mark study region on the maps for an international audience?

16. Pg10; Fig. 7: Why using the first PC of regional JFM precipitation instead of just the area-averaged JFM precipitation total? I do note these are very similar and map would look the same.

17. Pg11; L1-5: Delete section as repeated from Pg10.

18. Pg12; L3: "the JFM precipitation series"… but which one? First PC or observed totals?

19. Pg13; Table 1 & L1-4: The use of the asterisks is a little confusing. When I first looked at table 1 I presumed the asterisks was for statistically significant correlations. But are these instead those that are NOT correlated with JFM precipitation? Although you say that all are significant with at least one of the first three PCs. I can see here how perhaps the use of the first three PCs is useful but the reader is left with a bit of a jump to understand this without understanding what PC2 and PC3 represent. Could adding three additional columns to the right hand side of Table 1 for PC1, PC2, and PC3 help with this, then have the asterisks marking any value with statistically significant correlations. This allows the reader to see that perhaps one climate variable is correlated with all four precipitation series, or just e.g. PC3?

20. Pg14; L18-19: Not clear how the ensemble in Fig. 10 was created. More detail needed here. Also, how many ensemble members etc.?

21. Pg15;L6: A few issues with this sentence. Suggest changing to something like: "An RPSS value less than zero signifies no forecast skill over the reference climatology forecast (i.e. it is ….), a value equal to zero for when the forecast is only as skilful as climatology, and values greater than zero represents a skilful forecast. A value of one represents a perfect forecast".

22. Pg15; L9-14: Need more details about the use of 3x3 contingency tables, you might find the Svensson (2016) paper (and references therein) useful for this and as an example of statistical seasonal forecasting more generally. Also, more definition of what is meant by "extremely dry conditions". I know this is mentioned in the results, but it should be here that the methods details are given.

23. Pg15; L16-18: Which combinations of the 11 predictors in Table 1 made it into the final PCR model? I know PCA was used, but can weight be given to original 11 predictor variables? For example, can I tell how important, if at all, Niño 3.4 is to the final model?

24. Pg16; L5-7: The main message I get from Table 2 is that Near normal and Below normal precipitation is good, but it is the above normal that drags the hit rate to 51%.

25. Pg17; L21-27: This is a key conclusion. That the new model is more skilful than both climatology and a simpler ENSO only model is central to the paper. A moderate skill improvement at a seasonal level is still valuable. Also interesting that RPSS is negative for Niño 3.4 only model, why do you think this is? Overall, re-framing the introduction to include this science question at the start would strengthen the paper. This finding is lost as it is, so you need to place this moderate but important improvement in light of the ongoing work in this active area of research internationally, and not only the practical uses of the forecast in S. Peru.

26. Pg17; Sect. 7: I like the idea of extended lead time analysis, but the technical details should be first outlined in the proposed 'Methods' section and results presented under a sub-heading within the 'Results' section.

27. Pg18; Fig. 12: I'm missing how you are going from regional level to station level here. The extended lead time is good, but the spatial disaggregation is the weakest part of the analysis at present.

28. Pg19; L6-9: What statistical test is used to determine if the difference between regional and station correlation values is statistically significant or not?

29. Pg19; Sect. 8: There is no discussion of the key limitations of the forecasting method/model (e.g. poor for above average precipitation). It would be good here to offer some suggested avenues for further research to overcome such methodological limitations.

30. I like the final paragraph on Pg 20 as it highlights well the practical importance of seasonal forecasting using climate information in a region where none is currently available.

C.) Technical Corrections

1. The paper is generally well written, but in places language is a bit colloquial. E.g.
   - **a)** Pg1; L28: Change "vary drastically" to e.g. "vary considerably"
   - **b)** Pg2; L19: Change "wreaked havoc" to e.g. "was particularly severe"
   - **c)** Pg2; L26: Change "The dire" to e.g. "The societally challenging"

2. Pg3; Fig. 1 caption: Do white circles not represent SPCC stations, and blue the SENAMHI?
3. Pg6; 11: Add a comma in (Garreaud, 1999)
4. Pg9; L3: Change "hydrometeorologic" to "hydrometeorological"? Perhaps this term is used in the US?
5. Pg13; Table 1: Add space between 'Time frame'
6. Pg18; L9: Add full stop after "…etc.)".
7. Pg1; 19: Add "regional" and "totals" in front of and after "January-March precipitation", respectively?

References

Svensson, C.: Seasonal river flow forecasts for the United Kingdom using persistence and historical analogues, Hydrol. Sci. J., 61(1), 19–35, doi:10.1080/02626667.2014.992788, 2016.

Van Loon, A. F.: Hydrological drought explained, Wiley Interdiscip. Rev. Water, 2(4), 359–392, doi:10.1002/wat2.1085, 2015.