Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-102-AC3, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "An experimental seasonal hydrological forecasting system over the Yellow River basin – Part II: The added valuefrom climate forecast models" *by* Xing Yuan

X. Yuan

yuanxing@tea.ac.cn

Received and published: 1 May 2016

I am very grateful to the reviewer for the positive and careful review. The thoughtful comments have helped improve the manuscript. The reviewer's comments are italicized and my responses immediately follow.

1. Page 2, line 15: A reference is neededhere to support this statement.

Response: I will add the references "Pappenberger et al., 2008; Swinbank et al., 2016".

2. Page 2 line 16: change "flooding forecast" to "floodforecasting".

Response:I will revise it as suggested.

3. Page 2, line 25: Is NMME qualified to be called “open source”? Itsforecasts are made available to the research community, but the system itself is notopen source, is it?

Response: I did not say that the NMME models are “open source”. I called it “an open source of multimodel seasonal climate hindcast datasets” in the paper. Those hindcast datasets are made available to the public by the IRI personnel through the NMME project.

4. Page 3, line 28: If Yellow river basin is HEAVILY managed, I wonder if such activities can be simply represented by a linear regression in the postprocessing procedure. The probability distribution will be highly distorted as the goal of water resource management over the river is to do flood control and irrigation withdraw. Thus observed flow is much more steady (less variant) with less extremes duringboth dry and wet conditions. Linear regression is typically used between variables thatare normally distributed. Can you commend on this? This the linear regression is notsuitable here, it needs to be corrected.

Response: Thanks for the important comment. Actually the naturalized and observed streamflow datasets do show the characteristics in the upper reaches of the basin as the reviewer’s comment: the observed flow is much more steady (less variant) with less extremes during both dry and wet seasons (e.g., Lanzhou station in Figure 7). But this is not the case in the lower reaches, where the observed streamflow is consistently lower than the naturalized streamflow due to heavy human water consumption. To account for the seasonality in the water management, the linear regression is applied for each calendar month, where the water allocations during different years are similar and stable. Therefore, the linear regression method can be used to correct the systematic biases. However, I agree with the reviewer that it has drawbacks for correcting the nonlinear errors, where I mentioned it in the manuscript. With the water allocation and consumption data collected in the future, more sophisticated method should be implemented in the forecasting system. I will revise the manuscript as follows: “For

Printer-friendly version

Discussion paper

the upper reaches, the observed flow is much more steady (less variant) with less extremes during both dry and wet seasons; but for the lower reaches, the observed streamflow is consistently lower than the naturalized streamflow due to heavy human water consumption."

And the disadvantage of the linear regression method will be discussed at the end of the paper as follows: "(1) a linear time series post-processing model, although considering the seasonality in the water subtraction by calibrating the parameters against observed streamflow month by month, is not sufficient to simulate and forecast a hydrological system with intensive human interventions because of the nonlinearity and nonstationarity. Either connecting with a seasonally dependent water subtraction submodel based on the subtraction statistics or explicitly representing the human intervention processes in the forecasting system is not only necessary to further reduce the uncertainty in the hydrological models, but also to facilitate the understanding of the hydrological predictability with human dimension".

5. Page 4, line 31: why don't you use lapse rate correction here when binlinear interpolate the temperature forecast from models?

Response: The systematic bias (including the topography induced temperature bias) can be easily corrected by using the quantile-mapping method that is used in this study, i.e., mapping the forecast into the observed climatology with lapse rate correction.

6. Page 4 line 32: When you say "all ensemble member", are you referring to all members from one individual model or from the entire NMME ensemble?

Response: As I mentioned in the manuscript, "for each calendar month and each NMME model…" So it refers to all members from one individual model. The rationale is that different models have different climatology that affects the robustness, so I chose to correct them individually.

7. Page 5 linear 10: Again, I don't think the linear regression model is suitable or good

enough to represent the human component of the hydrological system.

Response: I did not clarify that the linear regression saves the day. The statement actually clarifies that the hydrological post-processing is necessary to bridge the gap between the observed streamflow and a hydrological model calibrated against the naturalized streamflow. As I respond above, the disadvantage of the linear regression model will be discussed. But the linear regression does make the hydrological simulations closer to the observation over the river basins with human interventions.

8. Table 2: Can you actually showhow the two time series of streamflow look like, with a QQ plot or scatter plot? The current illustration is not very convincing.

Response: Thanks for the comment. Actually Fig. 7 shows a few examples of the time series of streamflow from observation and post-processed simulations. And I will add the time series of the simulated streamflow before post-processing and the naturalized streamflow into the same figure for comparison.

9. Page 5 line 32: change “measures” to“metrics

Response: I will revise it as suggested.

10. Equation 1: This equation gives a space-time mixed formula for anomalycorrelation. Later in the paper, AC is also calculated for individual location, so it is necessary to mention that equation 1 can be simplified for such purpose.

Response: As I already mentioned in the manuscript, “If the AC is used for each grid cell within the Yellow River basin (i.e., there is only a summation over time), it is reduced to the Pearson correlation.”

11. Page 6, line10: What is the impact of having different ensemble members in ESP and NMME onthe skill assessment?

Response: To my experience, more ensemble members will result in higher reliability, but not necessarily the sharpness. As we know, the ESP forecast refers to a kind of

Printer-friendly version

Discussion paper

climatological forecast, so it is already the most reliable forecast. Adding more ensembles to the ESP usually does not improve the skill. Currently, the ESP consists of all historical forcings for the target seasons (excluding the target year) during the validation period (1982-2010), if we expand it with more ensemble members (e.g., those forcings before 1982), it has a risk that the results will be even biased if there is a shift in the climate (e.g., decadal variation). Given that the main focus of the paper is the deterministic forecast skill and the setup of the ESP experiment, I think the impact of the ensemble members for the ESP simulation is limited.

12. Figure 1: A pixelited shaded plot probably looks better and easier to read tan the current one. Can you highlight the correlations that are actually statistically significant?

Response: Thanks for the comment. I will revise as suggested. As I mentioned in the manuscript, an anomaly correlation (AC) of 0.05 (as shown in colors) would be statistically significant give the large samples.

13. Page 6, line 16: I don't agree with this assessment. Most models do show the highest skill for month 1, maybe except GFDL. So the lead time is still quite important, maybe just as important as seasonality. If you think there is not dependence on lead time, what could be the cause for that?

Response: I did not state that the lead time is not important. I used "not necessarily" but not the "NOT" in the speculation. It is related to the strong seasonality, i.e., it is usually more skillful during dry season than wet season.

14. Page 6, line 27: It would be interesting to know how much of the improvement in forecast skill is due to increase in the ensemble size.

Response: The reviewer raises an interesting question. Actually in the past, I did some testing to see whether a subset of the NMME ensemble is more skill than the grand NMME ensemble in terms of the deterministic forecast. But it is very difficult to select the optimal subset ensemble members. For the training period, sometimes a subset ensemble is more skillful than the grand ensemble; but it usually does not hold for the verification period. For a real-time forecasting, it is very difficult to select a subset of the NMME models (according to the hindcast) that is consistently more skillful than the grand ensemble mean. Perhaps it is partly because of the similarity of the NMME models that we discussed in a paper (Yuan and Wood, 2012). But again, this is very complicated especially for the precipitation, and it is out of the scope of the paper. So I decided not to include it in the paper. If the reviewer has any suggestions, I am very glad to try it in the future study.

15. Figure 2: The current way of plotting makes the 0.5month value almost invisible. I suggest a pixelited shaded plot.

Response: I will revise it as suggested.

16. Page 7, line 23:I don't think you cannot draw conclusion like this from Figure 4 although this is likelyvery true. Statements like this need to be more careful.

Response: Thanks for the comment. In this section, all hindcasts are verified against VIC offline simulations, i.e., the errors in the hydrological forecast model is neglected. To avoid confusion, I will add a note in the revised manuscript: "(but note that this result may be model dependent since the hydrological hindcasts in this section are verified against VIC offline simulations by neglecting the errors in the hydrological model)"

17. Page 7, line 32: "less"than what?

Response: I will revise it as "As the ICs have less control on the runoff forecasts than the meteorological forcings. . ."

18. Page 8, line 25: "representativeness"??

Response: I will revise as suggested.

19. Page 8, line 34: What non-stationary feature are you referring to here? If there is a trend, can you actually tell if it is caused by water withdraw or climate change?

Response: This refers to the human interventions. I will incorporate the time series of naturalized streamflow and the VIC simulations without post-processing in the revised Figure 7. The revised figure will show that the drying trend in the 1980s and 1990s is both caused by climate change and human interventions because the naturalized streamflow also has a drying trend, but is weaker than the observed streamflow.

20. Figure 8: Why not showthe negative part of SS?

Response: There are some small negative values for SS, but are not significantly different from zero. Therefore, they (as well as those small positive values) are not discussed in the paper for investigating the added value from climate-model-based hydrological forecast.

21. Page 10, line 3: IC is important, but not necessary always dominant.

Response: Thanks for the comment. I will replace it with "strong control".

22. Page 10 line 23-25: This conclusion is counter intuitive. Are you saying that if we were to have a perfect land surface model, the climate forecast in hydrological forecasts at long leads would be less useful?

Response: Most previous studies verify the forecast against hydrological model simulations by neglecting the errors in the hydrological models. My statement is that those studies (NOT perfect hydrological model) might underestimate the usefulness of the climate forecasts at long leads. With a perfect hydrological model, the skill for the climate-model-based hydrological forecasting will increase, and so does for the ESP forecasting, so the added value from (usefulness of) climate model is not necessarily increase.

23. Page 10, line 31-32: This addresses a different type of uncertainty. Use of multiple models help to address uncertainties associated with model, not observations.

Response: What I mean is exactly the same as the reviewer. To avoid confusion, I will revise it as follows: "...forecasting with multiple hydrological models might be useful to

Printer-friendly version

Discussion paper

quantify the uncertainty in the hydrological model"

24. Page 11, line 3: This depends on what type of downscaling method to be used, a dynamic downscaling scheme might not suffer the same.

Response: According to my experiences in the dynamical downscaling (Yuan and Liang, 2011; Yuan et al., 2012), neglecting the human component will also affect the performance of dynamical downscaling. This is because most climate models, especially for those used in the seasonal forecasting, do NOT consider the human interventions such as reservoir regulation, irrigation, land use changes and groundwater pumping, and their forecasts may suffer from that.

25. The last paragraph is an interesting discussion, but some of the statements are not directly based on the results of the current research, might need to be revised somehow.

Response: I thank the reviewer for the appreciation. This discussion focus on the representation of human intervention in the hydrological forecasting system, development of the system with multiple hydrological models, prediction of seasonal hydrology within the context of global environmental change, and the interpret of the ensemble hydrological forecast. They are the questions we would like to address in our future study. So I would like to keep them unless the reviewer has specific concerns.

References:

Pappenberger, F., Bartholmes, J., and Thielen, J., et al.: New dimensions in early flood warning across the globe using grand-ensemble weather predictions, Geophys. Res. Lett., 35, L10404, doi:10.1029/2008GL033837, 2008.

Swinbank, R., et al.: The TIGGE project and its achievements, Bull. Am. Meteorol. Soc., 50, 49-67, doi:10.1175/BAMS-D-13-00191.1, 2016.

Yuan, X., and Wood, E. F.: On the clustering of climate models in ensemble seasonal forecasting, Geophysical Research Letters, 39, L18701, doi:10.1029/2012GL052735, 2012

Yuan, X., Liang, X.-Z., and Wood, E. F.: WRF ensemble downscaling seasonal forecasts of China winter precipitation during 1982-2008, Climate Dynamics, 39, 2041-2058, doi:10.1007/s00382-011-1241-8, 2012

Yuan, X., and Liang, X.-Z., 2011: Improving cold season precipitation prediction by the nested CWRF-CFS system, Geophysical Research Letters, 38, L02706, doi:10.1029/2010GL046104.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-102, 2016.