Response to the comments from Reviewer #3

We are grateful to the reviewer for the constructive and careful review. We have incorporated the comments to the extent possible. The reviewer's comments are italicized and our responses immediately follow.

The manuscript "Ensemble streamflow forecasting over a cascade reservoir catchment with integrated hydrometeorological modeling and machine learning" integrate the meteorological forecast, land surface hydrological model simulation and machine learning to forecast hourly streamflow over the Yantan catchment, and the results show that the flood forecast has been significantly improved. This work is very meaningful and the paper has been well-written. I therefore recommend this paper resubmitted after minor revisions. My comments are listed as follows:

Response: We would like to thank the reviewer for the positive comments. Please see our responses below.

1) The number of ESP members is 40, while the number of members used from TIGGE-ECMWF is 51, which is not the same as ESP, Will the evaluation results over predicted due to the number of members used from TIGGE is larger?

Response: Thanks for your comments. First, the meteorological forcings from the same date as the hindcast start date to the next 9 days of each year (excluding the target year) were selected as the ESP forcing in this study. Although more members could be added to ESP, it will increase the overlap between different hindcasts due to limited historical samples. Second, larger sample of ESP means that more days far from the start date would be included, which may degrade the performance of ESP. Third, 40 members should be comparable to 51 members given the short hindcast periods.

2) The upstream basin streamflow are used as CSSPv2 model inputs to provide the upstream inflow information. In the hindcast experiments, the upstream outflow inputs used are forecasts or observations? When lacking upstream outflow prediction, how does the system operate?

Response: Thanks for your comments. In the hindcast experiments, the upstream outflow inputs used as inputs are observations to exclude the forecast uncertainty from upstream streamflow for the evaluation of the meteo-hydro-LSTM forecasting system. In the real-time forecasting system, we may use the upstream outflow from the previous day or the LSTM forecasts. We have clarified this caveat in the discussion as follows:

"This study mainly focused on exploring the added values of meteorology-hydrology coupled forecast and LSTM forecast in a non-closed catchment, so the forecast uncertainty from upstream outflow was ignored by using the observed outflow. In the future, the upstream outflow forecast is planned to include, but this requires the development of upstream hydrometeorological forecast capability, as well as the reservoir regulation forecast that is very challenging. The artificial intelligence (AI) techniques are expected to complement the physical model for reservoir regulation forecast." (L396-L403 in the tracked version of the revised manuscript)

3) Is it possible to forecast the rainfall-streamflow using meteorological forcing from a closed watershed controlled by the Yantan station and correct it by LSTM instead?

Response: Thanks for your comments. Yes, it is possible. But besides meteorological forecasts, the reservoir regulations should be incorporated or predicted, which is very challenging due to limited data and human interventions. Moreover, the motivation of this study is mentioned in the abstract as follows:

"For a cascade reservoir catchment, the discharge of the upstream reservoir contributes to an important part of the streamflow over the downstream areas, which makes it tremendously hard to explore the added value of meteorological forecasts. Here, we integrate the meteorological forecast, land surface hydrological model simulation and machine learning to forecast hourly streamflow over the Yantan catchment, where the streamflow is influenced both by the upstream reservoir water release and the rainfall-runoff processes within the catchment.”

As a first step, here we evaluate the performance of this meteo-hydro-LSTM coupled system. Future studies would incorporate upstream forecasts through a physical-statistical hybrid approach.

4) The upstream inputs are also essential to this forecast system, hope to see some evaluation about this element.

Response: The upstream inputs used in this system are observations, please see our response to your comment #2.

5) The calibration results evaluated by NSE shows a worse result in the upstream grids than downstream ones. Is it possible to improve the calibration results by increasing the Iteration times set in the SCE-UA methods?

Response: Thanks for your comments. We have clarified in the revised manuscript as follows:

“Higher NSE in the upstream part of Jiazhuan station (Figure 1) is due to more humid climate (not shown), where hydrological models usually have better performance over wetter areas. For the downstream areas with less precipitation, the higher NSE is related to the higher percentage of sand in the soil (not shown). Under the same meteorological conditions, there is higher hydraulic conductivity with higher sand content (Wang et al., 2016), and it yields less runoff under infiltration excess, which is more suitable for the saturation excess-based runoff generation for the CSSPv2 model (Yuan et al., 2018).” (L281-)

Figures R1 and R2 show the distributions of precipitation and sand percentage. According to Figure R3, most grids reach its best performance within 2000 iteration times.

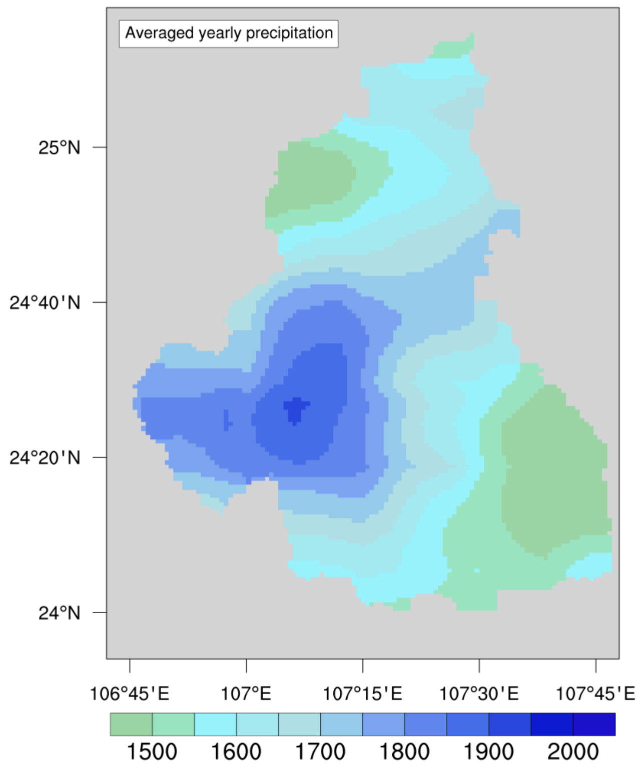

Figure R1. The spatial distribution of average yearly precipitation (mm).

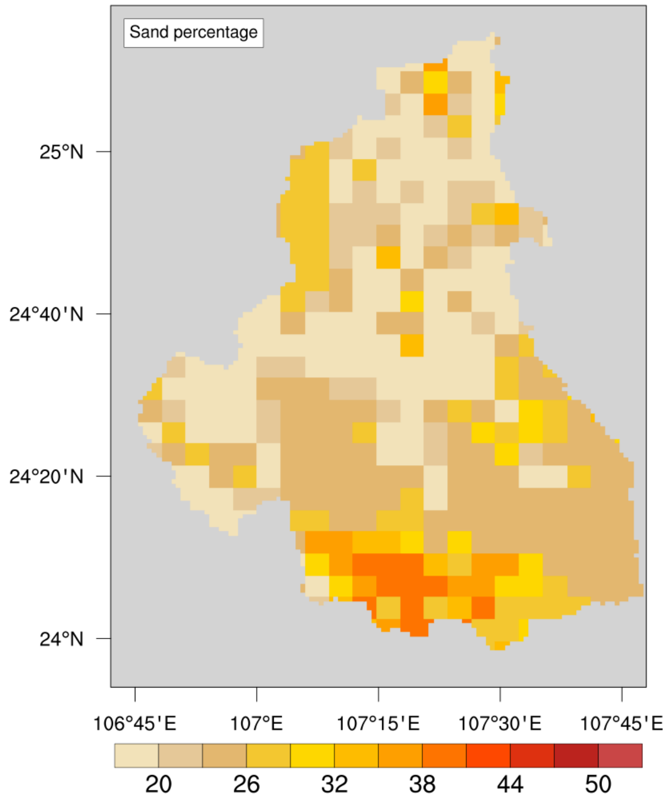

Figure R2. Spatial distribution of sand proportion.

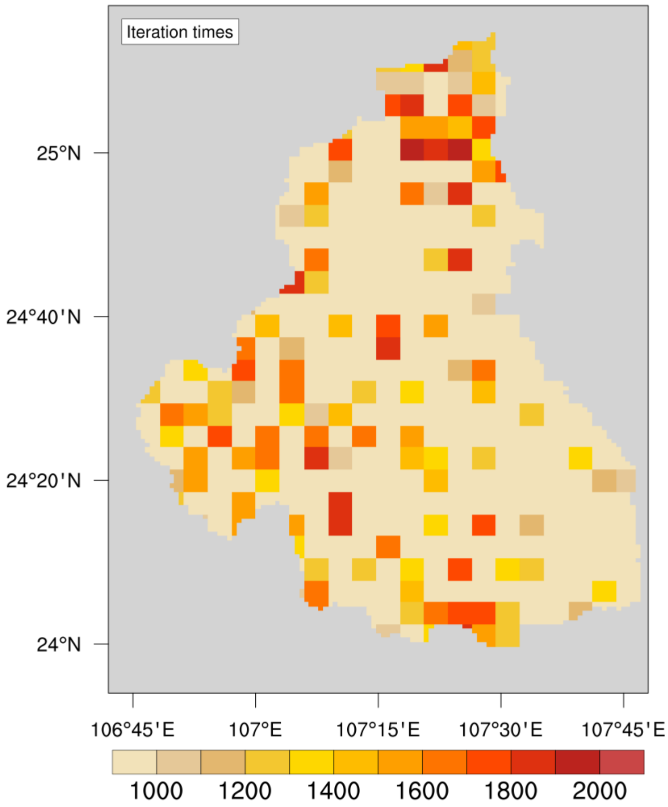

Figure R3. Iteration times when best performance occurred.