Response to the comments from Reviewer #4

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.

This is an interesting manuscript that is well structured and presented. The topic is worthy of publication in HESS. However, before publication please address the following comments.

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

1) Line 201: What is the full name of GR4J? Please check throughout (e.g., TIGGE, ECMWF, CMA, SCE-UA).

Response: Thanks for your comments. We have now provided the full names as follows:

"For example, Humphrey et al. (2016) showed that their combined Bayesian artificial neural network with the modèle du Génie Rural à 4 paramètres Journalier (GR4J) approach outperforms the GR4J model in monthly streamflow forecasting."

"We use the meteorological hindcast data from European Centre for Medium-Range Weather Forecasts (ECMWF) model that participated in the THORPEX Interactive Grand Global Ensemble (TIGGE) project to drive a newly developed high-resolution land surface model, named as the Conjunctive Surface-Subsurface Process model version 2 (CSSPv2, Yuan et al., 2018), to provide runoff and streamflow forecasts, and correct the forecasts via LSTM model."

"… China Meteorological Administration (CMA)."

"In this paper, we calibrated CSSPv2 model against monthly estimated runoff to simulate the natural hydrological processes by using the Shuffled Complex Evolution (SCE-UA) approach (Duan et al., 1994)."

2) Lines 179-182: Please list the values or ranges of the calibrated parameters in a table.

Response: Thanks for your suggestion. We have made a new table as follows:

Table 2. Descriptions of calibrated parameters

Parameters	Range
Maximum velocity of baseflow (mm/day)	0.00000116 ~ 0.000579
Fraction of maximum velocity of baseflow where non-linear baseflow begins	0.001 ~ 0.99
Fraction of maximum soil moisture where non-linear baseflow occurs	0.2 ~ 0.99
Variable infiltration curve parameter	0.001 ~ 1
River width (m)	0 ~ 101.16
River depth (m)	0 ~ 6.46
River density (km/grid)	0.049 ~ 1.03
River roughness	0.033 ~ 0.05
River slope	0.015 ~ 0.47

3) Line 189: How to perform the bias correction of the TIGGE-ECMWF forecast forcing

Response: Thanks for your comments. We have clarified as follows:

"The resolution of TIGGE-ECMWF grid data is 0.25°, so the data was interpolated to 5km grid to drive the CSSPv2 model. We calculated both observations' and TIGGE-ECMWF's yearly average precipitation and temperature, and performed a bias correction by adding back the difference (for temperature) or multiplying back the ratio (for temperature) to match the observations' averages."

4) Line 190: Please give a detailed description of the workflow in Fig. 2.

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

“Figure 2 shows the procedure of hindcasts: the calibrated CSSPv2 model was first driven with observation dataset to generate initial hydrological conditions (soil moisture, surface water, etc.) for each forecast issue date, then CSSPv2 model was driven with forecast data (TIGGE-ECMWF or ESP) at every forecast issue date with the generated initial conditions to perform a 7-day hindcast.”

“The network was trained on sequences of April to September in 2013-2017, with six historical streamflow observations and one forecast interval streamflow to predict the total streamflow at each forecast time step (Figure 2).”

“In addition, we also tried a LSTM streamflow forecast approach which only uses 6-hour historical streamflow data as inputs, and the experiment was termed as LSTM (Table 2). The process of LSTM is similar to Meteo-Hydro-LSTM but without the forecast interval streamflow, which is also shown in Figure 2.”

5) Line 204: Please check the equation number

Response: The equation number has been checked and revised.

6) Line 205: Please clarify the input and output variables here.

Response: We have clarified them as follows:

“where $\mathbf{q_0}$, $\mathbf{q}$, $\mathbf{q_{max}}$ and $\mathbf{q_{min}}$ are the normalized variable, input variable, the maximum and minimum of the sequence of the variable.”

7) Lines 215: The meaning of the F(y) in Eq. 1 is not clear. Also for ref in Eq. 2. Please clarify the meaning of the variables in each equation.

Response: We have revised as follows:

“The Continuous Ranked Probability Score (CRPS) was calculated as follows:

$$\boldsymbol{CRPS} = \int_{-\infty}^{\infty} [\boldsymbol{F}(\boldsymbol{y}) - \boldsymbol{F_o}(\boldsymbol{y})]^{\mathbf{2}}, \tag{7}$$

where

$$F_o(y) = \begin{cases} \mathbf{0}, & y < observed\ value \\ \mathbf{1}, & y \geq observed\ value \end{cases} \quad (8)$$

is a cumulative-probability step function that jumps from 0 to 1 at the point where the forecast variable y equals the observation, and $\boldsymbol{F(y)}$ is a cumulative-probability distribution curve formed by the forecast ensembles."

8) Line 291: Is the diurnal cycle in Fig. 8 the climatology averaged result? If so, please give the ranges to indicate the uncertainty. It also looks a bit strange that the rainfall reaches its maximum in the early morning rather than the afternoon in summer. Please clarify.

Response: The time shown in the figure is the universal time. According to Beijing time zone (GMT+8:00), the rainfall reaches its peak at the afternoon, we have modified the figure and revised the descriptions as follows:

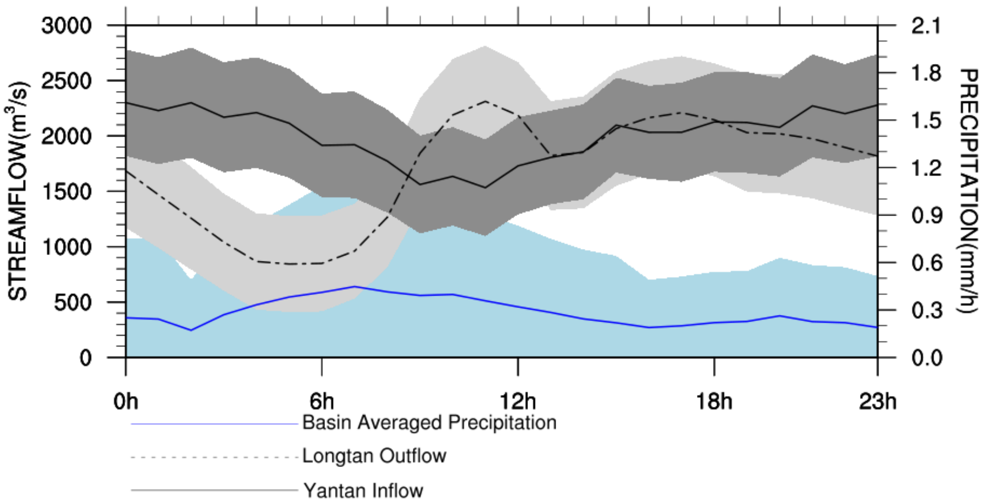

Figure 8. Diurnal cycles of Longtan outflow (m^3/s; dashed black line), Yantan inflow (m^3/s; solid black line) and basin-averaged precipitation (mm/h; blue line), as well as their ranges. The time shown in this figure is universal time.

9) Lines 298-301: Why would the combination of heavy rainfall and the decrease of upstream flow make the hydrological model performance worse? Does the model consider the reservoir?

Response: In this study, we used observed upstream streamflow as inputs while the ensemble meteorological forecasts as forcings. When the upstream flow decreases and rainfall increase, the upstream control on downstream flow decreases, and the influence of interval flow resulted from the meteorology-hydrology forecasts rises. And the uncertainty from meteorological forecasts would propagate to the hydrological forecasts more obviously. This highlights the importance of both the upstream outflow and the accuracy of meteorological forecasts within the catchment.

10) Line 328: Please add some implications of this work and as well as its deficiencies that need to be improved in the last section.

Response: Thanks for your comments. We have added the implications and deficiencies as follows:

"Most cascade reservoirs yet cannot forecast streamflow beyond 6 hours, and the integrated Meteo-Hydro-LSTM approach has potential to improve the forecasts at long leads. 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 facility, 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."

11) Add the units in Figs. 6 and 7.

Response: Revised as suggested.

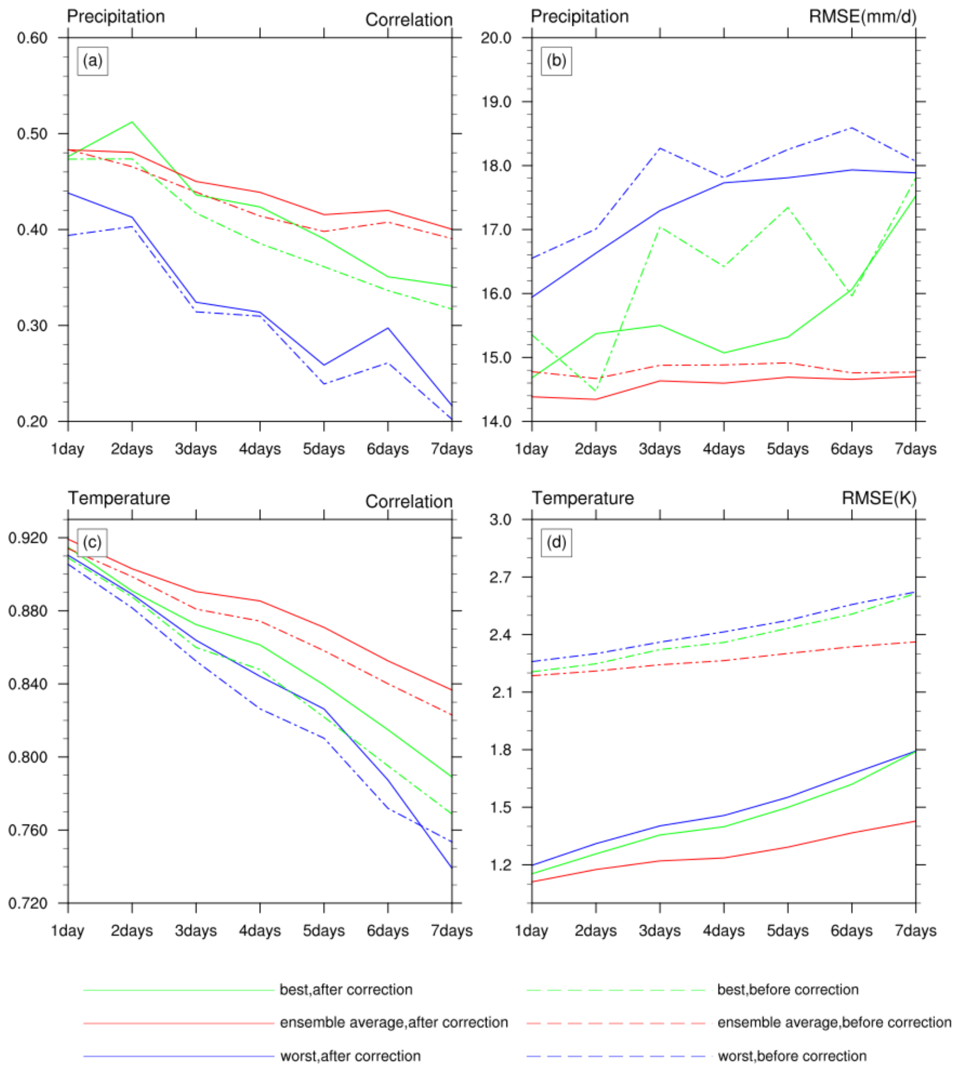

Figure 6. Evaluation of precipitation and temperature hindcasts from TIGGE-ECMWF. The red and blue lines represent the best and worst results among 51 TIGGE-ECMWF ensemble members respectively, and the green lines represent the results for the ensemble means of 51 members. Solid and dashed lines represent the results after and before bias corrections, respectively.

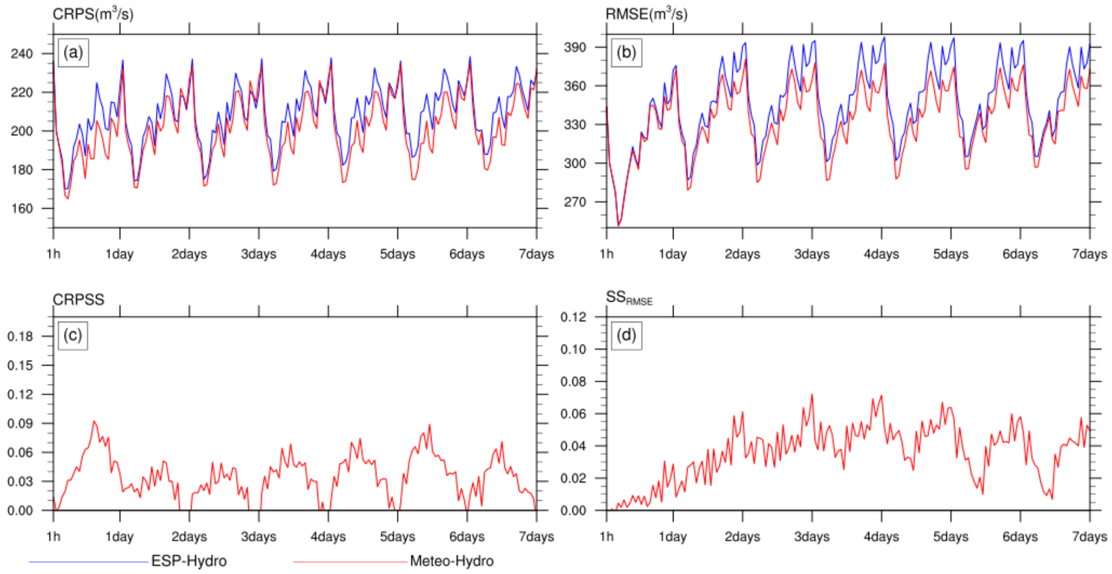

Figure 7. (a) Continuous Ranked Probability Score (CRPS) and (b) Root Mean Squared Error (RMSE) for daily streamflow ensemble forecasts at Yantan gauge. (c) and (d) are the skill score in terms of CRPS and RMSE for Meteo+Hydro, where ESP+Hydro is used as reference forecast.