Response to the comments from Reviewer #2

We are grateful to the reviewer for the constructive and careful review. The constructive suggestions have helped improved our manuscript. The reviewer's comments are italicized and our responses immediately follow.

Reservoirs represent an important but difficult issue for hydrological modelling. This paper presents a method for ensemble streamflow forecasting at the hourly timescale considering the effects of cascade reservoirs. The method makes use of TIGGE-ECMWF meteorological forecasts, CSSPv2 land surface model and LSTM deep learning model. Through the case study of a reservoir in China, the method is shown to reduce probabilistic and deterministic forecast errors. In general, the paper is well-written with results clearly presented.

There are five comments for further improvements of the paper.

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

1) First of all, more details on the contribution of this paper can be added. As is illustrated in the introduction, the proposed method is built upon the CSSPv2 land surface model (Yuan et al., 2018). What are the limitations of the previous model? Can the limitations be illustrated through some diagnostic plots? Such analysis would make the contribution of this paper more convincing.

Response: Thanks for your comments. In this study, we combined the newly developed CSSPv2 land surface hydrological model with ECMWF meteorological forecasts and the LSTM machine learning model to develop a Meteo-Hydro-LSTM forecasting framework for flooding forecasts over a cascading reservoir catchment. We have clarified in the revised manuscript as follows:

"In this study, we combine the machine learning with hydrometeorological approach for hourly streamflow forecast over a cascade reservoir catchment located in southwestern China. 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. We aim to improving flood forecast over the cascade reservoir catchment by integrating meteorological forecast, hydrological modeling and machine learning."

2) Second, the method is demonstrated for one reservoir. In the meantime, the "study area" section illustrates that there are ten cascade reservoirs in the Hongshuihe hydropower base. Is it possible to select another 2-3 reservoirs to show the robustness of the proposed method? It is noted that the additional case study reservoirs can be elsewhere and are not necessarily located in the Hongshuihe region.

Response: Thanks for your suggestions. Our ultimate goal is to develop a forecast system that can consider both upstream and downstream reservoirs. However, as the first step, we focus on assessing the added value of integrating meteorological forecast, hydrological modeling and machine learning in the flood forecasting. 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."

3) Third, Figure 8 presents an interesting illustration of the time lag between Longtan outflow and Yantan inflow. This lag is largely due to the flowing distance between the two reservoirs. Meanwhile, the section of methods does not tell how the river flow is *considered in the method. Is it performed by routing or hydro-dynamic simulation? How are the parameters determined?*

Response: Thanks for your comments. The river flow was calculated by a routing model employed the kinetic wave equation as covariance function, which was solved via a Newton algorithm. The parameters include slope, river density, roughness, width, and depth. These parameters were calibrated to match the hourly observed streamflow at Yantan hydrological gauge. We have clarified the routing model as follows:

"We used a high-resolution elevation database (hereafter referred to as DEM30) for sub-grid parameterization and figured out the initial values of these river channel parameters. We first extracted the slope angle and the natural river flow path from DEM30, and then identified the accurate river network using a drainage area threshold of 0.18 km^2. River density and bed slope values for each 5km grid were calculated as:

$$rivden = \sum l \,/A \tag{1}$$

$$bedslp = mean(\tan(\beta)) \tag{2}$$

where *rivden* is the river density (m/km^2), *bedslp* is the river channel bed slope (unitless), A is the area of a 5km grid (km^2), $\Sigma\, l$ is the total river channel length (m) within the grid, β is the slope angle (radian) for each river segment located in the grid.

Other river channel parameters were estimated by empirical formulas (Getirana et al., 2012; Luo et al., 2017) as follows:

$$W = 1.956 \times A_{acc}^{0.413} \tag{3}$$

$$H = 0.245 \times A_{acc}^{0.342} \tag{4}$$

$$n = 0.03 + (0.05 - 0.03)\frac{H_{max} - H}{H_{max} - H_{min}} \tag{5}$$

where W, H and n are river width (m), depth (m) and roughness (unitless) for each 5km grid; A_{acc} means the upstream drainage area (km^2); H_{max} and H_{min} refer to the maximum and minimum values of river depth calculated by Eq. (4).

Through a trial-and-error procedure, we calibrated these river channel parameters to match the simulated streamflow with observed hourly records at Yantan hydrological gauge."

4) *Fourth, lead time plays an important part in forecast verification as forecast skill tends to decrease with the increase lead time. Meanwhile, the simulations shown in Figures 4 and 5 seem to have nothing to do with lead time. Please present some plots of ensemble forecasts at different lead times*

Response: Sorry for the confusion we made in the manuscript. The simulations shown in Figures 4 and 5 are driven by the observed meteorological forcings in order to evaluate the performance of the CSSPv2 land surface hydrological model. They are not "real" forecasts. The performance of the ensemble streamflow forecasts are shown via CRPS plots in Figure 7. We have clarified in the revised manuscript as follows:

"Figures 4 and 5 show the results after the calibration of the routing model, where CSSPv2 is driven by observed meteorological forcings to provide streamflow simulations and compare against observed streamflow at Yantan hydrological gauge."

5) *Fifth, CRPS in Figure 7 exhibits some diurnal circle that can relates to the diurnal circle of reservoir inflow/outflow in Figure 8. This result may be due to the setting of the LSTM deep learning model. When preparing streamflow data for LSTM, has the mean been subtracted? Are alternative settings, e.g., subtracting the mean or not, tested for LSTM?*

Response: Thanks for your comments. The results in Figure 7 didn't include the LSTM deep learning model, while they are based on CSSPv2 streamflow forecasts driven by TIGGE-ECMWF meteorological forecasts or climatological forecasts (i.e., ESP). The diurnal cycle of CRPS is related to the upstream water release from Longtan station, and the diurnal cycle of catchment-averaged precipitation. Please see Figure 8 and related text.

When performing LSTM corrections, we didn't subtract the mean streamflow, but normalized the streamflow data with the maximum and minimum streamflow. Please see section 2.2.4 for details.

6) The location map can be improved by illustrating all the reservoirs in the Hongshuihe hydropower base. In addition, the location of the Hongshuihe hydropower base in China can be presented by using an inset plot.

Response: Thanks for the suggestion. We have redrawn Figure 1 as below.

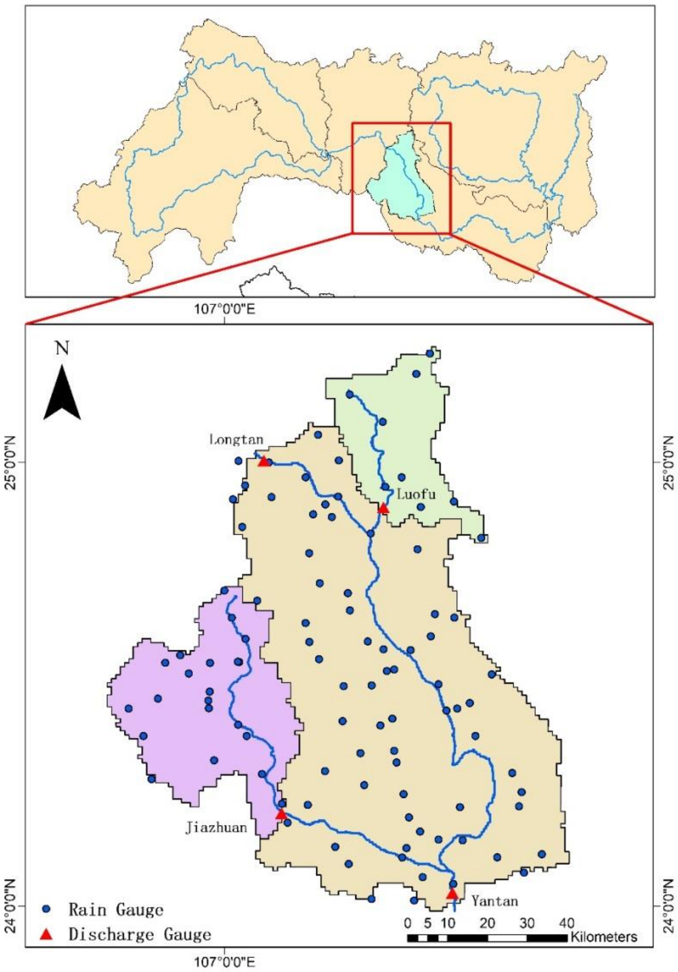

Figure 1. Locations of discharge gauges and rain gauges over the Yantan basin.

7) In Table 1, please illustrate the year/month range and time step for the hydrological dataset.

Response: Thanks for the suggestion. We have revised as below.

Table 1. Information of hydrological datasets

Dataset	Time Range	Time step
Rain Gauge Observation Forcing	2013/1/1 ~ 2017/12/31	Hourly
Longtan & Yantan Discharge Gauge Streamflow data	2013/1/1 ~ 2017/12/31	Hourly
Jiazhuan & Luofu Discharge Gauge Streamflow data	2013/4/1 ~ 2017/9/30	Daily
TIGGE-ECMWF Forecast Forcing	2013/4/1 ~ 2017/9/30	Hourly

References:

Getirana, A. C. V., Boone, A., Yamazaki, D., Decharme, B., Papa, F., and Mognard, N.: The Hydrological Modeling and Analysis Platform (HyMAP): Evaluation in the Amazon Basin, J. Hydrometeorol., 13, 1641–1665, https://doi.org/10.1175/JHM-D-12-021.1, 2012.

Luo, X., Li, H. Y., Ruby, L. L., Tesfa, T. K., Augusto, G., & Fabrice, P., et al.: Modeling surface water dynamics in the amazon basin using mosart-inundation v1.0: impacts of geomorphological parameters and river flow representation. Geosci. Model. Dev., 10(3), 1-42. https://doi.org/10.5194/gmd-10-1233-2017 , 2017.