Interactive comment on "A 439-year daily discharge dataset (1861–2299) for the upper Yangtze River, China" by Chao Gao et al.

General Comments

Overall quality of the paper

Firstly, thanks the authors for this manuscript that includes a good hydrological modeling effort by driven Global Climate Models (GCMs) under climate change scenarios. The paper presents multi-hydrological modeling by inputting GCMs data at daily time scale and impacts of climate change on Yangtze River, China.

The methods of the paper don't include scientific newness. The methods and the materials aren't described in sufficient detail, there are some deficiencies which are remarked at related points. Despite these, it is generally clear, well-structured and well-written but the language could be improved some in the sense of fluency and clarity. And, the manuscript is suitable in terms of themes of ESSD journal. I think it could be accepted after revision based on the review.

<u>Advantages of the manuscript</u>

- Multi-hydrological model approach
- GCMs under climate change scenarios
- Internal consistency evaluation of hydrological models by using GLEAM evapotranspiration data

<u>Disadvantages of the manuscript</u>

- The lack of multi-site (which have different geo-climatic characteristics) comparison)
- Use of 1990 year land use data for particularly future scenarios
- The deficiency of comparison of GCMs based on observed data
- The poor comparison of hydrological models based on observed data

Specific Comments

The title could be edited more suitably for the main purpose of the paper. In the title, "hydrological modeling and GCMs" expressions could be used.

In the abstract; the objective , the original value and the widespread impact of the study could be expressed better.

The introduction could be handled for meaning unity and logical continuity.

In the section of study area, Fig. 1 must be pointed out and river basin characteristics (elevation, slope, aspect, land cover, soil type, etc.) must be given shortly to understand the hydrological cycle of the basin.

In the methods, a flowchart included all processes could be given for the sake of clarity. Moreover, GIS and remote sensing techniques used could be mentioned shortly in this section.

In section 3.1, some information about downscaling, bias correction and data assimilation processes (details, techniques, etc.) of GCMs should be given. And it should be explained that how GCMs grid data was inputted into the hydrological models (as lumped data?). In addition, more details about GCMs (model structures, etc.) and RCP scenarios should be presented.

In section 3.2, please provide details (precipitation, temperature, etc.) of observed daily meteorological data and explain how the data of 189 ground-based stations was inputted into the hydrological models (mean or spatial interpolation/extrapolation of the data?). Furthermore, please give reference/references of flood events (1870 and 1931) because of measurements beginning in 1939.

In section 3.3, please give more details about GLEAM and the assimilation technique. And please present the comparison results of GLEAM and 4 hydrological models for internal consistency of the models in the results section.

In section 3.4, please present input and output data of the models, model structures (loss and transform methods, soil moisture accounting, snow melting-accumulation, evapotranspiration, groundwater, canopy interception etc.) and parameter sets.

In section 3.5, please indicate how many parameters were calibrated for each model and what is the range for each parameter?

In section 3.6, please clarify the problem of "is the use of 1990 year land use convenient for particularly future scenarios?". In conclusions and discussion section, it is stated that 1990 year land use is one of the uncertainties and human interferences have escalated since the 1990s. In my opinion, using 1990 land use data is a big handicap of this study that is not convenient for future modeling. 1990 land use data can be considered suitable for only piControl and historical periods. Modeling of future flows by using the land cover of 1990, when the industrial development has not started, will have misleading results, especially in human impact scenarios. The use of 2019 year land cover data would be more appropriate when modeling the future and this will make the results more consistent. I don't accept the pretext of not available by reason of the fact that current land use data can be easily obtained from various sources.

In section 4.1 and 4.3, while determining the trends, was statistical trend analysis applied? Or how were the trends determined? By only visual technique?

In section 4.2, please explain what the model performance threshold is. Is it for good performance? Please give reference/references for it (Table 4). In addition, please comment why SWIM and VIC models had lower performance in the validation period. Moreover please provide Q90 results in Fig. 5 which shows the scatter plots. So, in this figure, the coefficient of determination (R^2) of 4 hydrological models should be given for significance of the scatter plots. Likewise, WLS values of extreme conditions should be presented for the calibration period. For this section, finally, please show the mean annual evapotranspiration outputs of the other three models based on sub-basins in Fig. 7 and compare all of the models with GLEAM based on statistical and/or mathematical criteria. It is better to see the all models in comparison. Even further, VIC and GLEAM could be compared based on grid by means of a confusion matrix.

In competing interests section, please give information of author contribution for each author in favor of justice.

In Fig. 1, please present DEM to get to know the basin better. And please specify how many grid cells are in the basin.

In titles of Fig. 2 and Fig. 3, “seasonal” expressions should be replaced with “monthly”.

Although the study is daily flow modeling, the graphical representation in Fig. 6 is given on monthly time scale since the application period is very long. If so, please express this.

In Fig. 8, are the values mean outputs of hydrological models and GCMs (a-b)? And there is a noteworthy decreasing in discharge (d), why? Please comment. And why is “2100-2169” period not available (c-d)?

In Table 4, it is interesting that the performance criteria are so close between the cal-val periods. Please comment what the reason.

Technical Corrections

P3 L6 and L7: “0.2 ^{o}C/10a” and “-11 mm/10a” → Please present as open.

P4 L27: “(Moriasi et al., 2007)” must be placed in the end of the sentence.

P5 L21-L22 and P6 L2-L3: “seasonal” → “monthly”

P6 L27: “models could reproduce the monthly river flow …” → Please explain which one of “monthly modeling” or “daily modeling and transforming into monthly” was performed.

P7 L13 and L14: “with the exception of RCP 6.0” → Please comment.

P7 L14 and L15: “an increasing trend in the 21st century, turning into a decreasing trend in the 22nd century.” → Why? Please explain considering the precipitation.

P7 L18: “11,517 m^3s^{-1}” → “11,517 m^3s^{-1} (363.2 billion m^3)”. For comparison of observation value stated in the section of study area.

P7 L23: “(except for RCP 4.5)” → Why? Please interpret.

P8 L30: “For the calibration” → “For the calibration and the validation periods”

P9 L1: “a cross-validation method” → Please give info. Leave-one-out or k-fold technique? If leave-one-out technique, it is problem of independence of validation dataset.

P9 L6 and L7: “the simulated extreme peak values in the 1930s, 1950s and 1990s were also in good agreement with the historical documented records …” → Please add “except 1998 flood” for objectiveness and being non-manipulative. It is specified that the modeled peak flow is 36,000 m^3/s (in **P6 L20**) and the observed value is 68,500 m^3/s (in **P4 L9**) for the 1998 flood event.

The references in **P12 L5 and P12 L8** must be replaced due to the alphabetical order.

Recommendations

- The model performance values for the discharge and the evapotranspiration could be combined to consider the internal consistency of the model. As well in the calibration period, the model could be calibrated according to both the discharge and the evapotranspiration (multi-criteria calibration).
- The hydrological models can be performed in different areas in point of multi-site approach.

Data Set Control-Evaluation

In accordance with the aim-scope of the journal, the dataset control process of the review is rather important case. For future reuse and reinterpretation, I checked the quality of the datasets which are available at the relevant web link.

I think there is potential of the data being useful in the future but there are only simulated discharge file (xls) as time series. The observed data (discharge, precipitation, temperature, etc.), the data of GCMs (precipitation, temperature, etc. for pi and RCP scenarios), the evapotranspiration data of 4 hydrological models and GLEAM (as time series), the grid evapotranspiration data (GLEAM and VIC) and the other grid data used (DEM, soil and land use data) must be also presented for both reproducibility of scientific and usefulness of data. In addition, to perform tests for data quality, the above mentioned data sets are necessary. The availability of these datasets are important for usefulness and completeness, too. And at the aim and scope web page of the journal, the expression of “each article should publish as much data as possible” supports the completion of deficiencies.

At the datasets web link page, in Table 3 of README.pdf file, “Future” words for 4 hydrologic models were miswritten.