Reviewers' comments in blue. Our responses in black. Yellow highlighting emphasises revision undertaken.

Reviewer 1:

This manuscript presents an in-depth study on projecting future hydrological changes in China using the Joint UK Land Environment Simulator (JULES). The study focuses on high-resolution simulations (0.25°) to assess runoff variability and flood and drought risks under medium (SSP245) and high (SSP585) emission scenarios. The results indicate significant regional variations in runoff, with wetter conditions projected in eastern and southern basins and contrasting seasonal patterns between northern and southern China. Overall, the manuscript is well written and organized, and it makes a significant contribution to understanding future hydrological changes in China. However, it should address the limitations below to enhance the reliability and practical relevance of the findings.

Thank you for taking your time to review our manuscript.

Major comments:

1.The study relies on a limited set of observational data for model calibration and validation, which may affect the accuracy and reliability of the simulations. Accurate calibration and validation are critical for ensuring the reliability of hydrological models. The limited data might lead to significant errors in the projections, undermining the study's overall credibility.

Thank you for your valuable feedback. We acknowledge that the quantity and quality of observational data are critical for accurate calibration and validation of hydrological models. However, it remains difficult to access and use observed discharge data in China (Lin et al., 2023). We used the best available observational data to calibrate and validate our model. While the limited dataset may introduce some level of uncertainty, we conducted validation at multiple sites to assess the model performance. Although the number of sites is limited, the results indicate that the model performs acceptably within the available dataset.

We agree that more extensive observational data would improve the accuracy and reliability of the model. We have already included the discussion about discharge data in Lines 365-367.

We plan to add more discussion in Section 4.4 as following: Due to the difficulty in obtaining gauge discharge data in China (Lin et al., 2023), we used limited observational data to calibrate and validate the JULES model. Although the number of sites is limited, the results indicate that the model performs acceptably within the available dataset. If more site data distributed across various regions of China can be obtained and applied to calibration and validation, the model performance could be further improved.

Lin, J., Bryan, B. A., Zhou, X., Lin, P., Do, H. X., Gao, L., Gu, X., Liu, Z., Wan, L., Tong, S., Huang, J., Wang, Q., Zhang, Y., Gao, H., Yin, J., Chen, Z., Duan, W., Xie, Z., Cui, T., … Yang, Z. (2023). Making China's water data accessible, usable and shareable. Nature Water, 1(4), Article 4. https://doi.org/10.1038/s44221-023-00039-y

2. The influence of hydraulic engineering structures, such as dams and reservoirs, on runoff patterns is not considered, potentially overlooking significant anthropogenic impacts on hydrological processes. Human activities significantly alter hydrological processes. Ignoring these factors can lead to unrealistic projections and misinform policy decisions regarding water management and infrastructure planning.

Thank you for your comments. Our study primarily focuses on an understanding of the impacts of climate change on hydrological processes. Investigating how hydraulic structures impact hydrological processes is beyond our scope. Consequently, we did not incorporate the effects of hydraulic engineering structures into our model.

We agree that hydraulic engineering structures have effects on runoff patterns. We have already included it in Lines 368-372.

We plan to enhance Section 4.4 as following: Additionally, this study did not consider the influence of hydraulic engineering on runoff, which could potentially alter the rainfall-runoff response. Our study primarily focuses on understanding the impacts of climate change on hydrological processes. Investigating how hydraulic structures affect such processes is beyond our scope. Consequently, we did not incorporate the effects of hydraulic engineering structures into our model. Future research could involve integrating data on dams, reservoirs, and other hydraulic structures into hydrological models to assess their effects on runoff dynamics. This approach could investigate how human activities impact hydrological processes and contribute to flood vulnerability.

3. The study uses three selected GCMs deemed suitable for China, but a broader range of models might provide a more comprehensive assessment of future hydrological changes and reduce model selection bias. While the selected GCMs are suitable, including a wider range of models would enhance the robustness of the projections and provide a more comprehensive picture of potential hydrological changes.

Thank you for your comments. We chose six CMIP6 GCMs based on their performance in representing regional climate variability and their availability in high temporal resolution suitable for our hydrological model. While using more GCMs can indeed provide a broader range of uncertainties, we need to make a trade-off between the number of GCMs and the computational resources required. The ideal situation is to select a few models that can represent the majority of the GCMs and have a small bias from the observations.

Our first step in GCM selection was to identify six GCMs based on previous studies that demonstrated their good performance for precipitation and temperature in China (Yang et al., 2021; Lu et al., 2022; Jia et al., 2023). Next, we downscaled precipitation from these six GCMs and compared the bias with ERA5 datasets. Therefore, the GCMs we selected have biases that are as small as possible when compared to the observed data. These steps have been already included in the manuscript (Section2.2, Lines 128-137).

We will revise our manuscript to include evidence that the GCMs we selected can represent the majority of the GCMs. We downloaded 19 models (including ACCESS-CM2, ACCESS-ESM1-5, CanESM5, CMCC-ESM2, CNRM-ESM2-1, EC-Earth3, EC-Earth3-Veg, FGOALS-g3, GFDL-ESM4, INM-CM4-8, INM-CM5-0, MIROC-ES2L, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NorESM2-LM, NorESM2-MM, TaiESM1, and UKESM1-0-LL) according to the list in NEX-GDDP-CMIP6 (Thrasher et al., 2022) and calculated the daily average precipitation and temperature from 1959 to 2014 of all GCM ensemble means and selected GCM ensemble means. The selected GCMs mean temperature is 6.07 °C while the mean for all GCMs is 6.03 °C. The selected GCMs mean precipitation is 2.08 mm/day while the mean for all GCMs is 2.45 mm/day. Therefore, the GCMs we selected are representative of the performance of most GCMs.

To present this evidence, we plan to include the following text in Section 2.2: To ensure the selected GCMs represent the performance of most GCMs, daily average precipitation and temperature from 1959 to 2014 for the selected GCM ensemble means were compared with the ensemble means of 19 GCMs. These GCMs include ACCESS-CM2, ACCESS-ESM1-5, CanESM5, CMCC-ESM2, CNRM-ESM2-1, EC-Earth3, EC-Earth3-Veg, FGOALS-g3, GFDL-ESM4, INM-CM4-8, INM-CM5-0, MIROC-ES2L, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NorESM2-LM, NorESM2-MM, TaiESM1, and UKESM1-0-LL.

And we plan to include the following text in Section 3.2: The selected GCMs mean temperature is 6.07 °C while the mean for all 19 GCMs is 6.03 °C. The selected GCMs mean precipitation is 2.08 mm/day while the mean for all 19 GCMs is 2.45 mm/day. Therefore, the GCMs we selected are representative of the performance of most GCMs.

Jia, Q., Jia, H., Li, Y., & Yin, D. (2023). Applicability of CMIP5 and CMIP6 Models in China: Reproducibility of Historical Simulation and Uncertainty of Future Projection. Journal of Climate, 36(17), 5809–5824. https://doi.org/10.1175/JCLI-D-22-0375.1

Lu, K., Arshad, M., Ma, X., Ullah, I., Wang, J., & Shao, W. (2022). Evaluating observed and future spatiotemporal changes in precipitation and temperature across China based on CMIP6-GCMs. International Journal of Climatology, 42(15), 7703–7729. https://doi.org/10.1002/joc.7673

Yang, X., Zhou, B., Xu, Y., & Han, Z. (2021). CMIP6 Evaluation and Projection of Temperature and Precipitation over China. Advances in Atmospheric Sciences, 38(5), 817–830. https://doi.org/10.1007/s00376-021-0351-4

4. The ERA5 dataset, known to overestimate precipitation in some regions of China, is used for downscaling, potentially leading to inaccuracies in runoff projections. The use of ERA5 data may lead to overestimated runoff, especially in northern and western regions. This can affect the accuracy of flood and drought risk assessments.

Thank you for your insightful comments. ERA5 generally overestimates precipitation in the northern and western regions of China, but it can capture seasonal variations and the broad spatial distributions in both magnitudes and trends (Sun et al., 2021; Zhou et al., 2023). We have already included this information in the discussion section in the manuscript (Section 4.1, Lines 324-326).

We used three meteorological datasets to drive the model, ERA5 performs best among them. We agree with other scholars that, despite its biases, ERA5 remains one of the best reanalysis datasets available, providing a comprehensive set of variables (Cucchi et al., 2020; Xu et al., 2021). We plan to include this information in Section 4.1: We used three meteorological datasets to drive the model, including ERA5, Global Meteorological Forcing Dataset (GMFD, Sheffield et al., 2006) and China Meteorological Forcing Dataset (CMFD, He et al., 2020). Our evaluation showed that ERA5 performed the best compared to the other datasets. We agree with other scholars that, despite its biases, ERA5 remains one of the best reanalysis datasets available, providing a comprehensive set of variables (Cucchi et al., 2020; Xu et al., 2021).

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., & Buontempo, C. (2020). WFDE5: Bias-adjusted ERA5 reanalysis data for impact studies. Earth System Science Data, 12(3), 2097–2120. https://doi.org/10.5194/essd-12-2097-2020

He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., & Li, X. (2020). The first high-resolution meteorological forcing dataset for land process studies over China. Scientific Data, 7(1), 25. https://doi.org/10.1038/s41597-020-0369-y

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. Journal of Climate. https://doi.org/10.1175/JCLI3790.1

Sun, H., Su, F., Yao, T., He, Z., Tang, G., Huang, J., Zheng, B., Meng, F., Ou, T., & Chen, D. (2021). General overestimation of ERA5 precipitation in flow simulations for High Mountain Asia basins. Environmental Research Communications, 3(12), 121003. https://doi.org/10.1088/2515-7620/ac40f0

Xu, Z., Han, Y., Tam, C.-Y., Yang, Z.-L., & Fu, C. (2021). Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100). Scientific Data, 8(1), 293. https://doi.org/10.1038/s41597-021-01079-3

Zhou, Z., Chen, S., Li, Z., & Luo, Y. (2023c). An Evaluation of CRA40 and ERA5 Precipitation Products over China. Remote Sensing, 15(22), Article 22. https://doi.org/10.3390/rs15225300

5. The manuscript lacks a detailed discussion on the practical implications of the findings for water resource management and climate adaptation strategies in China. While the scientific findings are robust, their practical application in policy and management is crucial. A more detailed discussion could help translate the scientific results into actionable strategies for stakeholders.

Thank you for your suggestions. We plan to add a Section 4.5 to enhance the discussion section of our manuscript:

4.5 Practical implications

Given the projected increase in runoff depth by 7.30 mm per decade under the high emission scenario, water resource managers should prepare for higher water availability, especially in eastern and southern basins. This information is vital for optimising water storage and distribution systems to prevent waste and ensure equitable water distribution.

The expected increase in runoff during spring, summer, and autumn under high emission scenarios necessitates seasonal adaptation measures. For instance, enhanced flood control infrastructure will be essential in the southeast basin to mitigate the heightened flood risk.

The projection of drier winters in southern China and contrasting seasonal patterns between northern and southern regions highlight the need for region-specific drought preparedness strategies. Investments in drought-resistant crops and efficient irrigation systems will be crucial in northern China.

Minor comments:

1. Line 61-62， please specify how coarse it is.

Thanks for your suggestion. We plan to add the resolution information in the sentence in Lines 59-62 as following: However, their analysis mainly based on results from CMIP5, CMIP6, Inter-Sectoral Impact Model Inter-Comparison Project (ISMIP2a) and Global Land Data Assimilation System (GLDAS), the resolutions of their runoff projections were coarse (⩾0.5°).

2. Line 83, it should be Section 2.1.

Thanks for your careful review. We will revise it.

3. In figure 1, it is recommended to add 10 dashed lines at present.

Thanks for your suggestion. We will update the 9 dashed lines used currently to 10 dashed lines.

4. I would recommend to add the control area for the selected gauges somewhere.

Thanks for your suggestion. We have carefully reviewed the GRDC dataset, and unfortunately, it does not include explicit control area information for the gauges. However, we have provided the locations of the gauges in Fig. 1 as a reference.