Reply to Referee #1

Z. Yin on behalf of all co-authors

1 **"This paper presents a modeling study of the effects of irrigation and dams on streamflow changes in the Yellow River Basin. There are many similar attribution studies in the literature looking at various influencing factors in the study region. Authors argue that streamflow fluctuations are not well examined in previous studies. But I am not convinced that this attempt would lead to a significant advance in this field."**

A: Thank you very much for your comments. It is true that many attribution studies have been performed in the Yellow River Basin (YRB). But different from them, there are three main advantages is this study.

First, novel crop module and China's Plant Functional Types (PFT) map were used in this work. Accurate crop simulation is a precondition of reasonable irrigation estimation. Some previous studies do not have crop simulations and need observed or satellite-based data (e.g., Leaf Areas Index and fraction of photosynthetically active radiation absorbed by green vegetation) to drive their irrigation simulations. Although some Global Hydrological Models and Global Land Surface Models (GHMs and GLSMs) did develop their crop modules, the crop functions, which are always based on C3 grass generics parameterizations, are too coarse to simulation varied crop types and phenology over China. The novel crop module in ORCHIDEE is able to simulate most physical processes throughout the whole crop growth period (Wang, 2016), It has specific parameterizations for wheat, maize and rice, which are the three main staple crops in China, which have been calibrated based on census data (Wang et al., 2017). Moreover, the novel China's PFT map has been developed including the fractions of wheat, maize, and rice based on 1:1 million vegetation map and provincial scale census data from the National Bureau of Statistics. For the first time, the irrigation consumption is estimated based on varied phenology of different crop types in different regions.

Second, we simulate river discharges and dam operations in the YRB and validate them on a recent time period. Some global studies simulated the Yellow River with irrigation and dam operations. But the period of most simulations starts from 1960s or 1970s, when a high proportion of discharges was less affected by dams. In this study, we focus on the period when huge reservoirs (LongYangXia in 1986 and XiaoLangDi in 1999) started regulation. More importantly, we are the first to show the simulated water storage change of reservoirs and to validate it with observations from literature. The correlation coefficient of simulated and observed water storage change of LiuJiaXia and LongYangXia is over 0.9, suggesting that the dam model is able to reproduce dam

operations under climate variations.
Third, detailed diagnosis of anthropogenic factors in the YRB. Many global studies admit the complexity in simulating the streamflows of the YRB (Haddeland et al., 2014; Hanasaki et al., 2018; Wada et al., 2014, 2016). However, rare studies demonstrate where the mismatches from, and whether any key factor or mechanism is missing in the model. Through reviewing literature and reports, we demonstrated several possible important factors (mechanisms) missed in current simulations in the YRB, which are not well represented in GHMs and GLSMs as well. Details are discussed in our reply to Comment 3.

2 **"1. The main drawback of this modeling study lies in the coarse resolution of the simulations. The hydrological modeling community has advanced significantly towards hypo-resolution simulations, especially at the river basin scale. Here, authors conduct the simulations at a spatial resolution of $0.5^\circ \times 0.5^\circ$ in the river basin, using global-scale products for model inputs and validations. I believe authors should utilize local data for configuring their model in this specific river basin, given the availability of various high-resolution meteorological forcing data in China and ET products as well."**
A: To pursue accurate river discharge simulations, many hydrological models used high resolution atmospheric forcing (like 10 km) as driver. However, different from their objective for short-term flood prediction, our aim is to understand the mechanisms and discover missing mechanisms of how human activities affect the discharge fluctuations in the YRB, for which high resolution forcing is not necessary. In fact, our previous study (Xi et al., 2018) utilized 0.1° forcing (Chen et al., 2011) to attribute different factors to the trends of streamflows over China, which showed large overestimation of Yellow River annual discharge. Thus, the crucial questions, which are our objectives as well, are whether irrigation can explain the discharge overestimation in Xi et al. (2018) and what is the impact of dam operations on the river streamflow. Obviously, increasing spatial resolution is not helpful to interpret the mismatch. We agree with the referee's comments that high-resolution forcing is compulsory for accurate simulations. But before that, all important mechanisms should be implemented in the model.
In fact, the GSWP3 forcing has been corrected by a suite of ground-based observations (`http://hydro.iis.u-tokyo.ac.jp/GSWP3/exp1.html#boundary-conditions`). For instance, its precipitation assimilates with the GPCC (Global Climatology Centre) precipitation dataset that includes numerous gauges intensively distributed over China (Fig. R1, Becker et al. (2013)). Long-term (1982–2014) in-situ ET measurements (eddy covariance) that are still rare over China, particularly in the YRB (Chen et al., 2014; Lian et al., 2018). Although uncertainties exist in global ET products, they are able to reflect monthly ET magnitude and inter-annual variations (Pan et al., 2020). Nevertheless, our previous study (Yin et al., 2018) validated ORCHIDEE-simulated soil moisture (which indirectly reflects ET dynamics) over China by in-situ measurements, which shows a good agreement (median correlation coefficient 0.53 and RMSE $0.07\,\mathrm{m}^3.\mathrm{m}^{-3}$).

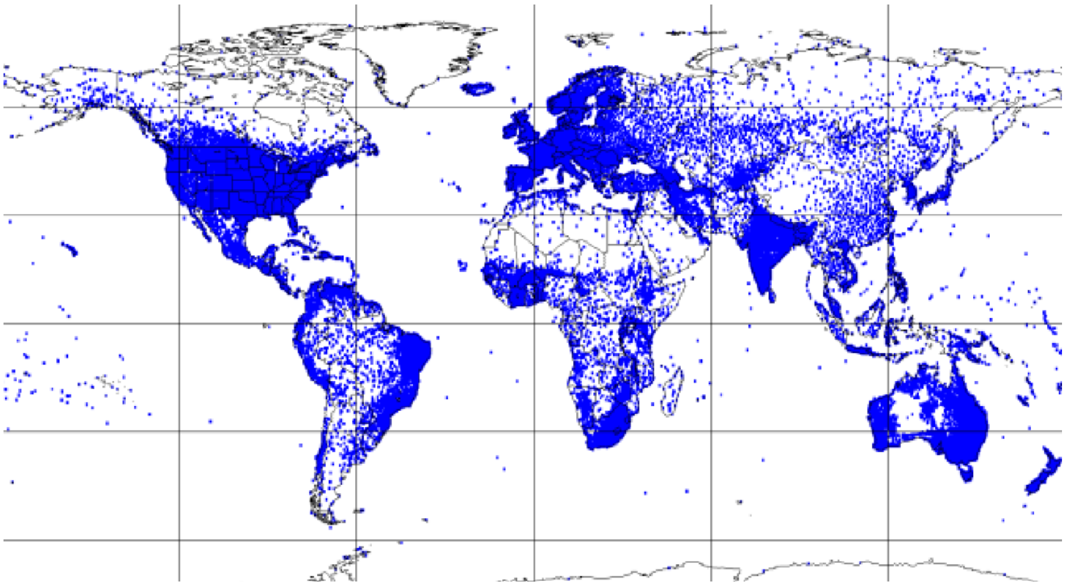

Figure R1: The map of 67,200 gauging stations used for the GPCC precipitation data production (from Becker et al. (2013)).

3 **"2. Extensive calibrations should be performed before using the model for quantifying the anthropogenic impacts. Authors argue that streamflow fluctuations have not been well examined in previous studies. but in figure 5-6, the model shows rather poor performance in simulating the seasonality and the peak streamflow, even with consideration of irrigation and dams."**

A: We agree that model calibration is necessary before utilization for scientific research. Previous studies demonstrate that our model performs well in simulating soil moisture dynamics (Yin et al., 2018), naturalized river streamflows (Table S1 in Xi et al. (2018)), leaf area index (Section S2 in Xi et al. (2018)), amount and trend of irrigation withdrawals (Yin et al., 2020), trends of total water storage (Section 3.4 in Yin et al. (2020)), and ET (Table S1 in online supplement) over China and in the YRB.

However, we cannot fully agree that our model performances are poor in simulating streamflow fluctuations based on Figure 5–6. First, after considering irrigation and dams, the bias of annual discharge and seasonality is substantially reduced (SB and SDSD reduce dramatically in Fig. 7a). Second, our study provides the comparison of simulated and observed water storage change of the LongYangXia and LiuJiaXia reservoirs **for the first time**. The correlation coefficient is 0.9, which, in our opinion, is quite good given the lack of information of the operation rules. Third, although natural discharge simulations with NSE=0.9 in a small sub-basin of the Yellow River is cited in our study, the NSE of them is incomparable to that of our simulations to conclude that our simulations are poor. A simple proof is given in our reply to the comment 13 from the second referee.

It is true that mismatches still exist between simulations and observations. However,

how to treat these mismatches depends on your goal. If the model services for short- or mid-term streamflow prediction, it is necessary to calibrate the parameters in the model to make the simulated streamflows fit the observations as well as possible regardless the detailed physical processes and other linked variables (e.g., surface energy balances, carbon cycles, vegetation dynamics, etc). However, such approach is probably not conducive to fundamental model improvements in terms of projecting streamflow variations under climate change, because some important missing mechanisms may be obscured by extensive calibrations. For instance, a study highlighted by HESS currently questioned why some well-calibrated models cannot perform well in forecasting river discharges under climate change (Duethmann et al., 2020). Through zooming in to a catchment in Austria, they revealed that "the importance of considering interrelations between changes in climate, vegetation and hydrology for hydrological modelling in a transient climate."
On the other hand, which is our case, if the model is used to demonstrate interactive mechanisms among climate, water resources, and human activities, these mismatches should be well investigated rather than be directly calibrated. For instance, we find that our model underestimates the annual discharge at LanZhou in the period 2000–2002 (Fig. 3b), during which $\hat{Q}_{\mathrm{IR}}$ was almost negatively correlated to the Q_{obs} (Fig. 5a). From China Water Resources Bulletin (2000-2002, `http://www.mwr.gov.cn/sj/tjgb/szygb/`), we find that to avoid discharge cutoff ($Q < 1\,\mathrm{m}^3.\mathrm{s}^{-1}$) irrigation and hydropower are strictly restricted. It suggests that integrated catchment management plays an important role in river flow variation, especially for extreme years. Obviously, models are not able to reproduce this special reaction by over calibration, if the related mechanisms are missing.
Moreover, from these mismatches, we also reveal other possible missing factors and mechanisms: 1) the Hetao Plateau withdraws $50 \times 10^8\,\mathrm{m}^3$ water from the Yellow River, which is neglected in most models because there is no large dam but multiple small reservoirs and complicated channel networks. It may lead to the overestimation of peakflows in Fig. 5; 2) The souring sediment is a special operation target of the XiaoLangDi dam, which release water one month ahead resulting in the delay of simulated water storage change (right panel of Fig. 6). All in all, as the famous statistician George Box said, "All models are wrong, but some are useful" (Box, G. E. P. 1976), if the "wrong" thing in the simulation can help us to discover important missing mechanisms rather than cover them by over calibration, I think the work is "useful". The discussion here are summarized in Sect. 4 of the revised manuscript.

4 **"3. In the irrigation scheme, irrigation water requirement is met only by the available stream water. How is the water availability defined? How does the model perform in simulating irrigation water use, compared to census data?"**
A: Thanks. It should be "available water resources", which has been corrected in the revised version. The available water resources include three water reservoirs in ORCHIDEE: 1) stream reservoir (streamflow); 2) fast reservoir (surface runoff); and 3) slow reservoir (deep drainage). Detailed introduction has been added in Section 2.1.1.

The irrigation module has been introduced and validated in Yin et al. (2020), which shows a good agreement of spatial distribution with census data. In Section 1, we added: "In a study focusing on China (Yin et al., 2020), ORCHIDEE estimated irrigation withdrawal coincided well with census data (provincial-based spatial correlations are ≈0.68), and successfully explained the decline of total water storage in the YRB."

5 **"4. In the abstract, 'Irrigation is found to be the dominant factor leading to 63.7% reduction of the annual discharges'. Is streamflow reduction caused by anthropogenic factors only? How about the effects of changing climate? Authors need to show the relative contribution of each factor (including irrigation) to streamflow changes in the abstract and conclusion sections."**

A: As industry and urban water consumptions are not taken into account in this study, we turn to report the amount of irrigation consumption instead of percentage of annual discharge. It is revised as: "Irrigation is found to substantially reduce the river streamflow by consuming approximately $242.8 \pm 27.8 \times 10^8\ \mathrm{m}^3.\mathrm{yr}^{-1}$ in line with the census data ($231.4 \pm 31.6 \times 10^8\ \mathrm{m}^3.\mathrm{yr}^{-1}$)." The stream reduction here means the difference between mean annual natural discharge and mean annual observed discharge due to irrigation (call it R1), not the impact of irrigation on the long-term decreasing trend of observed discharge (call it R2, if significant trend exists).

The streamflow reduction (R1) is mainly caused by anthropogenic factors (e.g., water consumption, reservoir surface evaporation, etc). However, the trend of streamflow reduction (R2) is not only caused by anthropogenic factors. Indeed, climate change is the primary driver of trends of the Yellow River streamflows, which has been demonstrated in our previous attribution study including climate change, CO_2 rise, land use change, and human activities (Xi et al., 2018). As this study concentrates on possible impacts of simulating anthropogenic factors on R1, we did not perform the similar analysis shown in Xi et al. (2018). Nevertheless, we demonstrate that climate change, at least the change of precipitation, has little effect on the change of streamflow seasonality (Section. 3.2 and Figure S4).

Bibliography

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth System Science Data, 5, 71–99, https://doi.org/10.5194/essd-5-71-2013, URL http://www.earth-syst-sci-data.net/5/71/2013/, 2013.

Chen, X., Su, Z., Ma, Y., Liu, S., Yu, Q., and Xu, Z.: Development of a 10-year (2001-2010) 0.1° data set of land-surface energy balance for mainland China, Atmospheric Chemistry and Physics, 14, 13 097–13 117, https://doi.org/10.5194/acp-14-13097-2014, URL http://www.atmos-chem-phys.net/14/13097/2014/, 2014.

Chen, Y. Y., Yang, K., He, J., Qin, J., Shi, J. C., Du, J. Y., and He, Q.: Improving land surface temperature modeling for dry land of China, Journal of Geophysical Research, 116, D20 104, https://doi.org/10.1029/2011JD015921, URL http://www.tandfonline.com/doi/abs/10.1080/10643380902800034http://doi.wiley.com/10.1029/2011JD015921, 2011.

Duethmann, D., Blöschl, G., and Parajka, J.: Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change?, Hydrology and Earth System Sciences, 24, 3493–3511, https://doi.org/10.5194/hess-24-3493-2020, URL https://hess.copernicus.org/articles/24/3493/2020/, 2020.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z. D., Wada, Y., and Wisser, D.: Global water resources affected by human interventions and climate change, Proceedings of the National Academy of Sciences, 111, 3251–3256, https://doi.org/10.1073/pnas.1222475110, URL http://www.pnas.org/lookup/doi/10.1073/pnas.1222475110, 2014.

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A global hydrological simulation to specify the sources of water used by humans, Hydrology and Earth System Sciences, 22, 789–817, https://doi.org/10.5194/hess-22-789-2018, URL https://www.hydrol-earth-syst-sci.net/22/789/2018/, 2018.

Lian, X., Piao, S. L., Huntingford, C., Li, Y., Zeng, Z. Z., Wang, X. H., Ciais, P., McVicar, T. R., Peng, S. S., Ottlé, C., Yang, H., Yang, Y., Zhang, Y., and Wang, T.: Partitioning global land evapotranspiration using CMIP5 models constrained by observations, Nature Climate Change, 8, 640–646, https://doi.org/10.1038/s41558-018-0207-9, URL http://www.nature.com/articles/s41558-018-0207-9, 2018.

Pan, S. F., Pan, N. Q., Tian, H. Q., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Ottlé, C., Poulter, B., Zaehle, S., and Running, S. W.: Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling, Hydrology and Earth System Sciences, 24, 1485–1509, https://doi.org/10.5194/hess-24-1485-2020, URL `https://hess.copernicus.org/articles/24/1485/2020/`, 2020.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth System Dynamics, 5, 15–40, https://doi.org/10.5194/esd-5-15-2014, URL `http://www.earth-syst-dynam.net/5/15/2014/`, 2014.

Wada, Y., de Graaf, I. E. M., and van Beek, L. P. H.: High-resolution modeling of human and climate impacts on global water resources, Journal of Advances in Modeling Earth Systems, 8, 735–763, https://doi.org/10.1002/2015MS000618, URL `http://doi.wiley.com/10.1002/2015MS000618`, 2016.

Wang, X. H.: Impacts of environmental change on rice ecosystems in China: development, optimization and application of ORCHIDEE-CROP model, Ph.D. thesis, Peking University, 2016.

Wang, X. H., Ciais, P., Li, L., Ruget, F., Vuichard, N., Viovy, N., Zhou, F., Chang, J. F., Wu, X. C., Zhao, H. F., and Piao, S. L.: Management outweighs climate change on affecting length of rice growing period for early rice and single rice in China during 1991–2012, Agricultural and Forest Meteorology, 233, 1–11, https://doi.org/10.1016/j.agrformet.2016.10.016, URL `http://linkinghub.elsevier.com/retrieve/pii/S0168192316304087`, 2017.

Xi, Y., Peng, S. S., Ciais, P., Guimberteau, M., Li, Y., Piao, S. L., Wang, X. H., Polcher, J., Yu, J. S., Zhang, X. Z., Zhou, F., Bo, Y., Ottle, C., and Yin, Z.: Contributions of Climate Change, CO 2 , Land-Use Change, and Human Activities to Changes in River Flow across 10 Chinese Basins, Journal of Hydrometeorology, 19, 1899–1914, https://doi.org/10.1175/JHM-D-18-0005.1, URL `http://journals.ametsoc.org/doi/10.1175/JHM-D-18-0005.1`, 2018.

Yin, Z., Ottlé, C., Ciais, P., Guimberteau, M., Wang, X. H., Zhu, D., Maignan, F., Peng, S. S., Piao, S. L., Polcher, J., Zhou, F., and Kim, H.: Evaluation of ORCHIDEE-MICT-simulated soil moisture over China and impacts of different atmospheric forcing data, Hydrology and Earth System Sciences, 22, 5463–5484, https://doi.org/10.5194/hess-22-5463-2018, URL `https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-699/https://www.hydrol-earth-syst-sci.net/22/5463/2018/`, 2018.

Yin, Z., Wang, X. H., Ottlé, C., Zhou, F., Guimberteau, M., Polcher, J., Peng, S. S., Piao, S. L., Li, L., Bo, Y., Chen, X. L., Zhou, X. D., Kim, H., and Ciais, P.: Improvement of the Irrigation Scheme in the ORCHIDEE Land Surface Model and

Impacts of Irrigation on Regional Water Budgets Over China, Journal of Advances in Modeling Earth Systems, 12, 1–20, https://doi.org/10.1029/2019MS001770, URL `https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001770`, 2020.