We thank the reviewers for their valuable comments. We have replied the comments one by one below, and revised the manuscript accordingly. The replies are highlighted in blue color and the modified texts (in the revised manuscript) are shown italic.

Reviewer #1

This study quantitative assessed the human impact and climate change impact on streamflow in continental China. The simulations streamflow was used from six global hydrological models driven by three meteorological forcings. The research is very interesting and significative. However, there are a few issues that the authors need to address before the manuscript can be accepted. I recommend most of the issues I raise below just need clarification or justification.
Reply: Thanks for the positive comment. We have replied the comments below and revised manuscript accordingly.

1. The simulated results need to be verified further with observed streamflow, maybe, QQPLOT, NSE etc. method can be used.
Reply: We have calculated NSE for the 44 stations, and added a sentence describing the result with a table (new added Table S2) in the revised supplementary information.

Revision in the manuscript (Subsection 3.1, the second paragraph (new added)):
The Nash-Sutcliffe coefficients calculated for the multimodel median and observed monthly streamflow at each station (see Table S2) show that the multimodel medians have better performance in the southern basins.

Table S2. The Nash-Sutcliffe coefficients (NSE) for the simulated monthly streamflow from VARSOC experiment and observed monthly streamflow (m^3 s^{-1}) at the 44 stations over the 1971-2000 period. The observed mean annual streamflow (MAF, m^3 s^{-1}) averaged over the period is also shown for each station.

Number	Station Name	MAF	NSE	River name	Number	Station Name	MAF	NSE	River name
1	Guchengzi	151.26	-0.27	Songhua River	23	Xixian	117.87	0.31	Huai River
2	Fuyu	449.68	0.53	Songhua River	24	Fuyang	117.74	0.63	Huai River
3	Tonghe	1444.43	0.81	Songhua River	25	Lutaizi	639	0.80	Huai River
4	Kuerbin	26.94	0.004	Songhua River	26	Bengbu	800.63	0.81	Huai River
5	Chaoyang	18	-0.88	Liao River	27	Shishang	1968.28	0.93	Yangtze River
6	Chifeng	7.76	-0.25	Liao River	28	Changyang	431.02	0.75	Yangtze River
7	Tieling	84.36	<-1.0	Liao River	29	Pingshan	4546.38	0.77	Yangtze River
8	Liaozhong	101.43	0.50	Liao River	30	Sinan	910.98	0.79	Yangtze River
9	Changmapu	29.3	-0.37	Northwest Rivers	31	Cuntan	10747.92	0.68	Yangtze River
10	Yingluoxia	51.06	-0.26	Northwest Rivers	32	Datong	28460.19	0.78	Yangtze River
11	Zhamashenke	22.7	0.09	Northwest Rivers	33	Quzhou	207.66	0.80	Southeast Rivers
12	Sandaohezi	16.45	<-1.0	Hai River	34	Zhuji	40.24	0.58	Southeast Rivers
13	Panjiakou	60.87	0.01	Hai River	35	Zhuqi	1721.14	0.91	Southeast Rivers
14	Luanxian	96.09	0.71	Hai River	36	Yangkou	442.85	0.72	Southeast Rivers
15	Xiapu	4.58	<-1.0	Hai River	37	Daojieba	1746.97	0.12	Southwest Rivers
16	Huangbizhuang	32.14	-0.07	Hai River	38	Gulaohe	96.63	0.22	Southwest Rivers
17	Cetian	4.78	-0.01	Hai River	39	Manhao	310.84	0.82	Southwest Rivers
18	Lanzhou	976.8	0.53	Yellow River	40	Jiangbianjie	194.96	0.68	Pearl River
19	Shizuishan	867.25	0.45	Yellow River	41	Duanzhan	2005.11	0.88	Pearl River
20	Longmen	803.67	-0.47	Yellow River	42	Xiayan	449.63	0.82	Pearl River
21	Huayuankou	1103.51	0.09	Yellow River	43	Wuxuan	4130.25	0.81	Pearl River
22	Xianyang	107.26	0.63	Yellow River	44	Boluo	782.04	0.80	Pearl River

2. The simulated results are very bad in some basins, such as NW, SW, HA. These simulated streamflow need be post-processed, and then be used to analyzed the impact of human and climate change.

Reply: We recognized the poor performance of the simulations, especially in the northern basins (see above Table S2). A post-processing on the simulations could reduce the deviation in simulated streamflow from observations and narrow the spread across models (e.g., Yin et al., 2017). However, the post-processing can be affected by the distribution and the number of the stations. In this study, though we have collected hydrological observations from 44 stations in China, they may be not representative enough for all basins. For example, there are only three stations in the Northwest Rivers (NW) which cover small areas. More stations are also needed for the basins like the Southwest Rivers where the streamflow changes greatly from upstream to downstream. Therefore, we think post-processing is not appropriate to and not necessarily improve the streamflow simulations in this study based on multimodel simulations. Furthermore, we tend to focus on the multimodel uncertainty in the model results in the evaluation section. We have added a caution for the limited representative of the observations in the evaluation result to remind readers to treat it carefully.

Revision in the manuscript (Subsection 3.1, the second paragraph (new added)):

It should be noted that the stations are located at different reaches of individual basins. Thus, the station-averaged estimates are largely dominated by those with large streamflow (e.g., at the lower reaches). *Additionally, the coverage of stations used is relatively small (due to data availability), especially in hydrologically variable regions like in the Northwest Rivers, leading to not necessarily representative evaluation of the performance of the GHMs in the whole basin.*

3. The authors need add some explanation of ISIMIP2a about how to simulate water withdrawals.

Reply: We have added some description for the simulated water withdrawals in section 2.1 Simulation data.

Revision in the manuscript (Subsection 2.1, the second paragraph (new added)):

Human impact considered in the VARSOC experiment (see the maps in Figure S1 and Table S1 for more details) includes the time-varying areas for both irrigated and rainfed cropland (Fader et al., 2010; Portmann et al., 2010) and reservoirs (dams) from the Global Reservoir and Dam (GRanD) Database (Lehner et al., 2011) including their commissioning year (see Figure S1 and Table S1 for more detail). Reservoir regulation was considered in the VARSOC experiment, which often reduces high streamflow in high-flow seasons and increases streamflow in dry seasons (Masaki, et al., 2017). Inter-basin water transfer was not considered in any of the model runs. The simulations of water withdrawals are different between the GHMs with respect to water use requirements and water withdrawal sources which are shown in Table S1. The sources of water withdrawals, depending on models, may include river channel, reservoirs, groundwater and lakes, and their fractions can be determined from reported statistics (e.g., Siebert et al., 2010) or estimated in models (Wada et al., 2014). In addition to the irrigation water requirement which is usually estimated by coupling crop models, most

GHMs considered the requirements for domestic and industrial water use which were prescribed in H08 (Hanasaki et al., 2008), LPJmL and MATSIRO (Pokhrel et al., 2015) or were estimated according to the population, socioeconomic and technological development in PCR-GLOBWB (Wada et al., 2014) and the population, thermal electricity production, gross added value, and technological change in WaterGAP (Flörke et al., 2013). Water use requirement for livestock was also prescribed in the LPJmL model, and estimated according to livestock densities in PCR-GLOBWB and WaterGAP2.

Reference:

Fader, M., S. Rost, C. Müller, A. Bondeau, and D. Gerten (2010), Virtual water content of temperate cereals and maize: Present and potential future patterns, J. Hydrol., 384(3–4), 218-231, doi: 10.1016/j.jhydrol.2009.12.011.

Flörke, M., E. Kynast, I. Bärlund, S. Eisner, F. Wimmer, and J. Alcamo (2013), Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study, Global Environ. Change, 23(1), 144-156, doi: 10.1016/j.gloenvcha.2012.10.018.

Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, N. Shirakawa, Y. Shen, and K. Tanaka (2008), An integrated model for the assessment of global water resources – Part 1: Model description and input meteorological forcing, Hydrol. Earth Syst. Sci., 12(4), 1007-1025, doi: 10.5194/hess-12-1007-2008.

Lehner, B., et al. (2011), High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Frontiers in Ecology and the Environment, 9(9), 494-502, doi: 10.1890/100125.

Masaki, Y., N. Hanasaki, H. Biemans, H. Müller Schmied, Q. Tang, Y. Wada, S. N. Gosling, K. Takahashi, and Y. Hijioka (2017), Intercomparison of global river discharge simulations focusing on dam operation—multiple models analysis in two case-study river basins, Missouri–Mississippi and Green–Colorado, Environ. Res. Lett., 12(5), 055002.

Pokhrel, Y. N., S. Koirala, P. J. F. Yeh, N. Hanasaki, L. Longuevergne, S. Kanae, and T. Oki (2015), Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts, Water Resour. Res., 51(1), 78-96, doi: 10.1002/2014WR015602.

Portmann, F. T., S. Siebert, and P. Döll (2010), MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, Global Biogeochem. Cycles, 24(1), GB1011, doi: 10.1029/2008GB003435.

Siebert, S., J. Burke, J. M. Faures, K. Frenken, J. Hoogeveen, P. Döll, and F. T. Portmann (2010), Groundwater use for irrigation – a global inventory, Hydrol. Earth Syst. Sci., 14(10), 1863-1880, doi: 10.5194/hess-14-1863-2010.

Yin, Y., Q. Tang, X. Liu, and X. Zhang (2017), Water scarcity under various socio-economic pathways and its potential effects on food production in the Yellow River basin, Hydrol. Earth Syst. Sci., 21(2), 791-804, doi: 10.5194/hess-21-791-2017.

Wada, Y., D. Wisser, and M. F. P. Bierkens (2014), Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dynam., 5(1), 15-40, doi: 10.5194/esd-5-15-2014.