The paper describes a multimodel assessment of the relative impacts of human activities and climate on mean annual streamflow over the past 4 decades in China. This study shows that unlike previous assessments, the climate impact signal is much more pronounced than the human impact signal in 88% of river segments in China. The study also quantifies the impact of humans across basins and discusses regional differences. In general the paper is publishable after some moderate revisions.

- The use of the term 'climate change' in the title and throughout the manuscript is somewhat confusing and misleading because it gives the impression that the paper will be forward looking in time and over the coming several decades (e.g., 2050, 2100). A more appropriate term is 'climate impacts'
- The 3rd paragraph of the introduction makes the argument that "This is the first study to perform such a quantitative assessment for all rivers of China with comparable modeling experiments." Being aware of the ISIMIP publications (https://www.isimip.org/outcomes/publications/) in this space with global assessments including many of the authors on this paper, I find this argument to be an exaggeration. I think the last sentence of that paragraph is a key novelty of this work, and as such linking back to the content of the second paragraph in the introduction to make the case would be my suggestion. I do agree that focusing on China is somewhat unique about this study. So one suggestion is to tweak the noted sentence as follow "This is the first study to focus on performing such a quantitative assessment for all rivers of China with comparable modeling experiments."
    - Schewe et al.: Multimodel assessment of water scarcity under climate change. PNAS, 2014.
    - Haddeland et al.: Global water resources affected by human interventions and climate change, P. Natl. Acad. Sci. USA, 111, 3251–3256, https://doi.org/10.1073/pnas.1222475110, 2014.
    - Veldkamp et al.: Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century, Nature Commun., 8, 15697, https://doi.org/10.1038/ncomms15697, 2017.
    - Wada et al.: Human–water interface in hydrological modelling: current status and future directions, Hydrol. Earth Syst. Sci., 21, 4169-4193.
- P5, L2: I would suggest omitting 'preliminary'
- P7, L14-17: Showing the individual models in figure S2 makes the figure too busy to read. Why not use the same format as in figure 2 by showing a band around the median. Also, it would be useful to show the same type of figure as figure 2 but for streamflow.
- P7, L18-24: I realize that given the large departures in water withdrawal estimates, matching streamflow guage observations might be a challenge, unless the authors believe that simulated water withdrawals might be equally or even more reliable than the statistically collected data, which have their own challenges.
- P7, L18-24: Are water withdrawals taken from surface water sources or also groundwater sources? What about return flows? Also, I am assuming that glacier melting, which contributes to streamflow, is simulated in these models, but that region is not included in the analysis. I realize that some of these were mentioned in the results, but incorporating some of these

details briefly when discussing the method or the results from the evaluation exercise would suffice.

- P10, L23-29: how does the model specify how much water is taken from surface water vs groundwater sources? Are the small pockets of increased MAF due to human impacts (Fig 5c) attributed to technological change (e.g., irrigation efficiency), or return flow from groundwater pumping, or something else?
- P11, L7: I would suggest omitting the sentence about the US and Canada. It breaks the flow of the paragraph which is talking specifically about China.
- P13, L14-30: To me this, this is a key contribution of this study. Yes, I agree that the results are not necessarily comparable in term magnitudes due to the highlighted reasons by the authors. But a missing discussion point is to why they fundamentally differ in their findings. I don't agree that either one of these two approaches (small scale using statistical approaches vs large scale modeling similar to this study) is necessarily superior. Each approach has its own pros and cons. So articulating why this approach differs from earlier findings is critical.