Dear Editor and Reviewers,

Thank you again for all the suggestions that help us to improve our study. We have revised our manuscript thoroughly according to the comments. Please see the replies to the reviewers' comments below. In the response, the blue texts are the comments and the green texts are the quotes of the manuscript.

On behalf of all co-authors,
Wencong Yang

Reviewer #1 Comment 1: (hereafter referred to as R1C1, R1C2…) *The authors have made many improvements to the paper and have addressed most of my concerns. My remaining concern relates to the amount of underlying variability that is used to estimate the panel regression coefficients. The authors should communicate the underlying variability (or lack of) more clearly, so that the reader is able to properly evaluate how the results may or may not apply to other regions or catchments. Below, I have explained my concern in detail and how I believe it could be addressed in a straightforward manner.*

A: Thank you for your constructive comments. We have carefully considered your suggestions and revised the paper. In the revision, we compare the characteristics (*Urban*, Δ*Urban*, *RI*, and Δ*RI*) of catchment samples used for the regression and the flood change attribution and find no substantial difference. Therefore, we infer that the regression coefficients can be applied on a national scale. In addition, using bootstrapping to derive coefficient intervals has already accounted for the sampling uncertainties to a certain extent. Nonetheless, we admit that the catchment samples used for the regression and the flood change attribution may not be completely homogeneous in terms of all catchment characteristics (e.g., climate, topography). Therefore, we added a statement in the discussion part for readers to transfer the results to other regions with caution. Details of the revision are presented in the following answers.

R1C2: *Panel regressions estimate the effect of a within-catchment change by taking advantage of panel datasets, which have variation in both space and time. This requires that there are temporal variations in urban area, RI, and crop area within catchments. As I suspected, and as shown in Table 1, a relatively small portion of catchments experience changes in RI over 1992-2017. For example, it appears that less than half of 207 watersheds (or <103 catchments) had changes in RI. Thus, the* *estimated effect of RI on Q will be almost entirely based on the effects within this small subset of catchments. My concern is that this subset of catchments may not be a representative subset of the overall 1,625 catchments (in terms of size, geographic distribution, topography, climate, etc).*

A: Thank you very much for your critical comment. We agree with you that the regression coefficients rely heavily on the catchment samples with changes in *Urban*, *Crop*, and *RI*. Therefore, in the revision, we compare the frequency distributions of *Urban*, Δ*Urban*, *RI*, and Δ*RI* for catchments used for the regression (Fig. 3) and the flood change attribution (Fig. 6 and Fig. 7). The comparisons are presented in Fig. A2 and Fig. A3 of the appendix and also as follows. We only focus on the catchments with Δ*Urban*>0 and Δ*RI*>0. According to these new figures, the catchments used for the regression are not substantially different from the catchments for the flood change attribution. Therefore, the catchments used for the regression are representative of a national-scale study in terms of the human factors that influence floods. In the revised manuscript, we added the contents about the comparisons above in Line 244-248 as "To avoid sample heterogeneity between these 1625 catchments and the 757 catchments used for regression, we compared the frequency distribution of $Urban$ and $\Delta Urban$ for catchments with $\Delta Urban > 0$ in the two samples of catchments in Fig. A2. Since there is no substantial difference between the two distributions in Fig. A2, the sensitivity of Q to $Urban$, which is derived from the 757 catchments, can be used to infer the changes in Q due to $\Delta Urban$ in those 1625 catchments." and Line 263-267 as "Similar to Fig. 6, to avoid sample heterogeneity between these 536 catchments and the 757 catchments used for regression, we compared the frequency distribution of RI and ΔRI for catchments with $\Delta RI > 0$ in the two samples of catchments in Fig. A3. Since there is no substantial difference between the two distributions in Fig. A3, the sensitivity of Q to RI, which is derived from the 757 catchments, can be used to infer the changes in Q due to ΔRI in those 536 catchments."

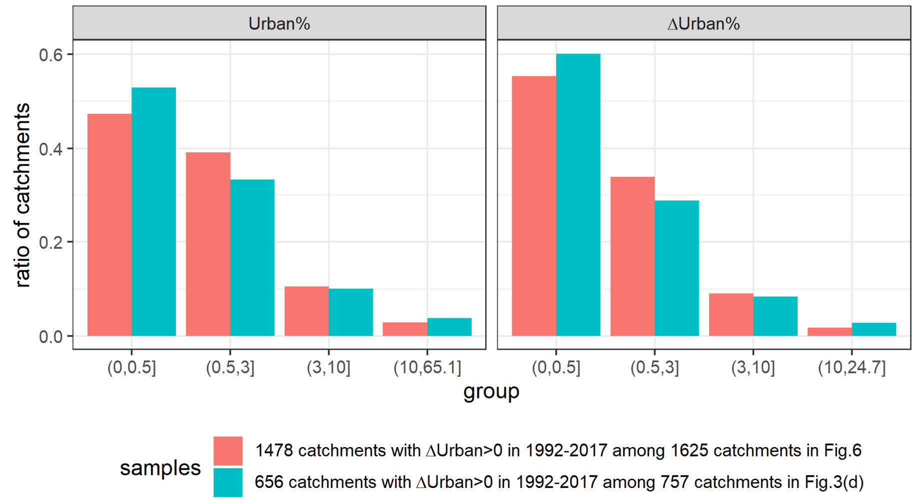

Figure A2. Frequency distribution of $Urban$ and $\Delta Urban$ in catchments with $\Delta Urban > 0$ from two catchment sets: the one used for the regression in Fig. 3(d) (green) and the one used for the flood change attribution in Fig. 6 (red).

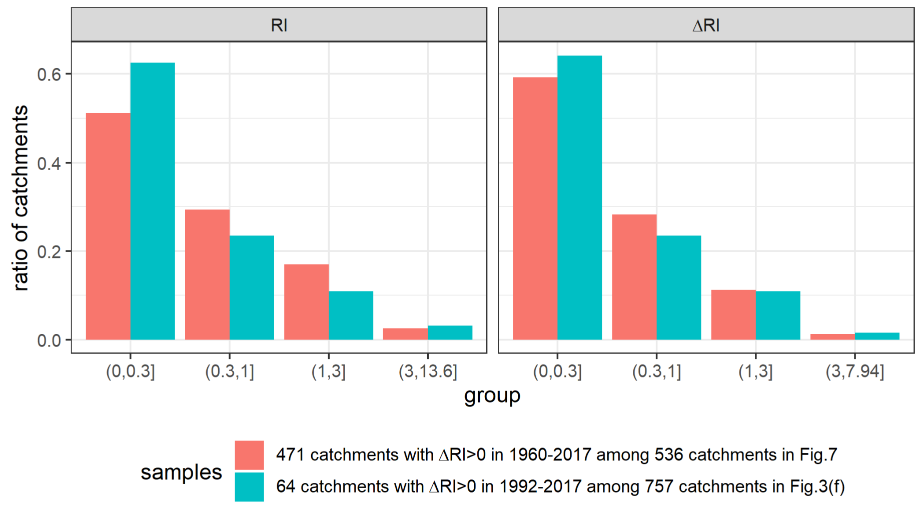

Figure A3. Frequency distribution of RI and ΔRI in catchments with $\Delta RI > 0$ from two catchment sets: the one used for the regression in Fig. 3(f) (green) and the one used for the flood change attribution in Fig. 7 (red).

However, the catchments used for the regression may not be representative if we consider more catchment characteristics, such as climate, topography, and so on. It is beyond the scope of this study to determine whether two catchment samples are homogeneous in terms of a comprehensive set of characteristics. Therefore, in the revision, we state the risk of transferring the results of this study to other regions in the discussion as "Caution is required to interpret the flood changes attributed to urbanization and dam constructions on a national scale because the sensitivity of floods to these factors is derived from a subset of catchments. Although the catchments used for sensitivity calculation and the ones used for flood change attribution have similar frequency distributions of urban areas and reservoir indexes in Fig. A2 and A3, these different sets of catchments may not be completely homogeneous in terms of all characteristics (topography, climate, etc.). Moreover, one should also be cautious to apply the sensitivity results to other regions such as catchments in other countries." (Please see Line 370-375)

R1C3: *Similarly, from Table 1, it appears that more than 75% of the 757 catchments had changes in urban area of less than 1%, and more than 50% of catchments had changes in crop area of less than +/– 1%. The few watersheds that had larger changes in urban area or crop area would have a larger influence on the estimated coefficients. Overall, this adds uncertainty for drawing conclusions about effects of RI, crop area, and urban area across the full region, if the regression coefficients are heavily influenced by a few catchments.*

A: Thank you for your important comment. Please see the answer to R1C2.

R1C4: *I agree with the authors' choice to use bootstrapping to estimate regression uncertainty. I think there are two additional aspects needed to properly convey the underlying variability and the implications for the analysis:*

1. Clarify the number of catchments with variability in each of the causal factors during the 1992-2017 period, as well as the distribution and characteristics of these catchments compared to the overall region. This could be accomplished with an additional table (or a revised version of Table 1), as well as some clarifications in the text. Fig. 3 is helpful, and it would also be helpful to see the spatial distribution of ΔRI, Δ%Urban, and Δ%Crop. If there are large differences between the subset of catchments and the full set of 1,625 (or 2,739) catchments, the authors should add a caveat about applying results from the subset to the full set.

A: Thank you very much for your valuable comment. Following your suggestions, in the revision, we clarify the number of catchments with variability in each of the causal factors in Line 195-197 as "Catchments with changes in $Urban$, $Crop$, and RI have large impacts on estimating regression coefficients. The numbers of catchments with $\Delta Urban > 0$, $\Delta Crop > 0$, and $\Delta RI > 0$ are 656, 351, and 64, respectively.". In addition, we add three subfigures to the original Fig. 3 to show $\Delta Urban$, $\Delta Crop$, and ΔRI.

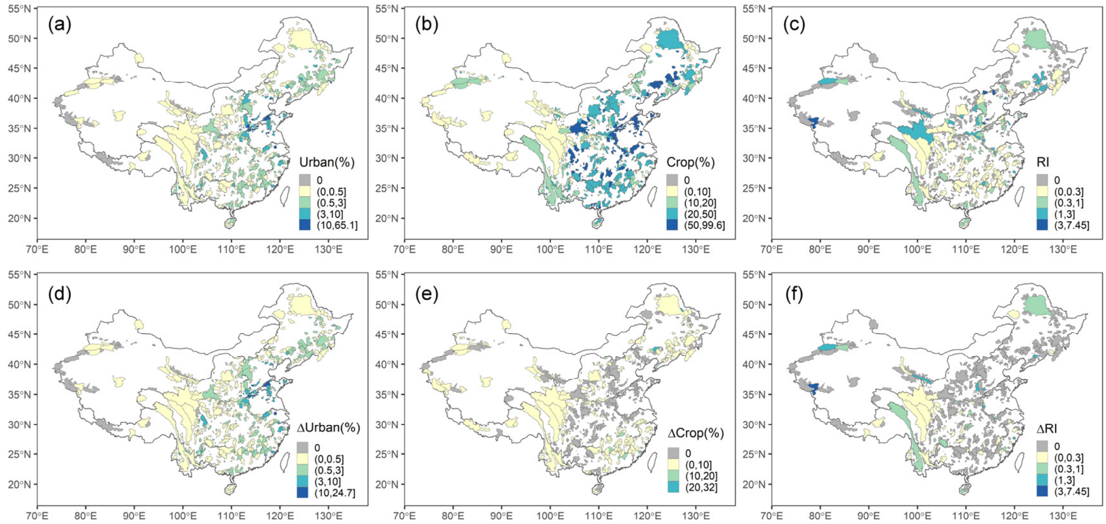

Figure 3. Spatial distribution of catchment characteristics in 757 independent catchments. (a) Urban percentages ($\boldsymbol{Urban}$), (b) cropland percentages ($\boldsymbol{Crop}$), and (c) reservoir indexes ($\boldsymbol{RI}$) in their last years with available flood data. The changes of (d) urban areas ($\boldsymbol{\Delta Urban}$), (e) cropland areas ($\boldsymbol{\Delta Crop}$), and (f) reservoir indexes ($\boldsymbol{\Delta RI}$) in 1992-2017.

As we mentioned in the answer to R1C2, Fig. A2 suggests no substantial difference in the frequency distributions of *Urban* and Δ*Urban* between the 757 independent catchments and the 1625 catchments used for detecting *Urban*-induced flood changes. Nonetheless, we add a caveat in the discussion (Line 370-375) for rigorousness, as we stated in the answer to R1C2.

R1C5: *2. In the case of RI (where different time periods are used to estimate the regression vs. cumulative effects), the authors should compare the amount of within-catchment variability during the period used for fitting the regression (1992-2017) to the RI changes during the period used to estimate cumulative causal effects (1960-2017). This could be accomplished in a table, or through clarification in the text. It is possible that changes in RI over this longer period are much larger than over the shorter period, and thus there may be higher uncertainty when extrapolating to these larger changes. If this is the case, I think it deserves a statement or caveat in the text.*

A: Thank you very much for your valuable comment. As we mentioned in the answer to R1C2, Fig. A3 suggests no substantial difference in the frequency distributions of

RI and Δ*RI* between the 757 independent catchments in 1992-2017 and the 536 catchments in 1960-2017. Nonetheless, we add a caveat in the discussion (Line 370-375) for rigorousness, as we stated in the answer to R1C2.

Additional Comments:

R1C6: *L164: "1. There are no other important time-varying sub-regional variables that significantly affect both human factors and floods". I think the assumption would more accurately be stated as: There are no other time-varying sub-regional variables that are correlated with human factors and affect floods.*

A: Thank you very much for your comment. We have changed the sentence to "There are no other time-varying sub-regional variables that correlate with both human factors and floods". (Please see Line 141-142)

R1C7: *Figure 8: I think this analysis should be restricted to the common period when both dam and land type data are available (1991-2017). Otherwise, potential effects of Δ%Urban in the early part of the period are unknown. Additionally, I am confused by the authors' choice to only examine basins where RI=0 and %Urban=0 at the beginning of the time period. Any catchments that have no change in RI and %Urban would be free of effects from these factors over the time period used to calculate trends in Q (regardless of initial RI or %Urban values). I don't think that the current analysis is wrong, per se, but a more comprehensive version of Fig. 8 would include all catchments where ΔRI and Δ%Urban yield expected changes in Q of less than 10% (regardless of initial values), and is something the authors should consider.*

A: Thank you very much. The purpose of Fig. 8 is to examine the unexplained flood changes in a long period (≥30 years). Long-term changes in floods dated back to the 1960s provide valuable information for studies related to the impacts of climate change or other factors on floods. Since we do not have land cover data before 1992, the strategy to select catchments is to filter all catchments that are potential to have large Δ*Urban* values in 1960-2017. In other words, the catchments with limited values of *Urban* are a subset of catchments with limited values of Δ*Urban*. Therefore, the selection guarantees the catchments are free from the impact of *Urban*, but does not guarantee all catchments that are free from the impact of *Urban* in 1960-2017 will be included. You indeed provide a good idea to redraw Fig. 8, but we aim to show the flood changes in a longer period and therefore we prefer to keep this figure.