

Interactive comment on “Causal effects of dams and land cover changes on flood changes in mainland China” by Wencong Yang et al.

Anonymous Referee #1

Received and published: 13 January 2021

In the paper, the authors estimate the effects of urbanization, cropland area, and dam regulation on the magnitude of annual maximum streamflow by analyzing historical data from 2739 gaged catchments in China. The authors use panel regression methods to identify these effects, and find that urbanization increases annual maximum streamflow, whereas dam regulation decreases annual maximum streamflow. Overall, the paper addresses an important question and provides new understanding about the factors leading to changes in annual floods. The paper is well-organized and clear, with ample references to previous studies, and the methods are appropriate for the questions studied. There are a few issues that could be addressed with minor revisions, which I have detailed below. First, more detail about the underlying data for the regression is necessary to fully interpret the results. Second, I would also advise against

C1

comparing the magnitude of p-values as a method to select between two different models. I have also noted some areas of the text needing clarification and included one suggestion for additional analysis that I think could be of interest to readers.

Main Comments

1. It would be helpful if the authors included more detail about the underlying distributions (or ranges) of the causal factors studies. For example, the authors state that the effect of urbanization is stable (i.e. linear), but it is unclear over what range of urbanization values this was tested, and if the effect could be increasing/decreasing outside of this range. Also, the panel regression specifically models within-unit (catchment) effects. This requires that there are within-catchment variations in the explanatory variables, but it is unclear how some of the variables, like dam regulation, are distributed in space and time. What percentage of catchments have changes in reservoir index, urbanization, and/or cropland area over the time period? Suggestions for ways to add this information include reporting these statistics within the text and/or adding a figure(s) of single or joint probability distributions or a time series of the causal factors.
2. Fig. 2 indicates that the lowest p-value is used to choose between a quadratic term (“increasing effect”) or square root term (“decreasing effect”) in cases where both terms have a p-value less than 0.01. P-values are not intended to be used to determine which model form is correct. It would be more appropriate to use a criterion designed for model selection, such as the AIC or BIC. Alternatively, in cases where both an increasing and decreasing effect are plausible and there is not a model form that clearly fits better, the authors could report both models.
3. In Fig. 8, the authors provide a map showing stations that had relatively small changes in urbanization or reservoir index but did have significant changes (>10%) in annual maximum streamflow. It would be interesting to extend this type of comparison for the stations shown in Fig. 6 and 7. In other words, it would be interesting to compare the observed streamflow trends to the trends predicted by changes in urban-

C2

ization and RI shown in Fig 6 and 7. This comparison could identify regions where additional causal factors are involved (and thus could identify interesting areas for future research). I do not believe this analysis is necessary for publication, but would likely be of interest to readers. Comparing observed trends with trends predicted by causal factors would also be a relatively new addition to the panel regression literature within hydrology.

Line-by-line comments:

Line 31: “whether a factor affects floods?” would read more clearly as “does a factor affect floods?”

Lines 47-64: In this paragraph, I find the descriptions of the existing methodological approaches to be unclear, particularly for the first method. Is the first “model-based” method referring to an empirical model, a physically-based model, or some combination?

Line 116: “the time-varying constant effects” should just be “the constant effects”

Lines 119-121: “Although $\delta\hat{S}EI$ may correlate with $\delta\hat{S}Lr_{\text{urban}}$ and Crop the effects of dams and land cover on floods can be derived independently since we have controlled their common drivers (Pearl and Mackenzie, 2020) in each equation, i.e., the regional time-varying term and the individual time-invariant term”. I’m not sure this is correct - it is possible that RI could be temporally correlated (within-watershed) with Urban or Crop, in which case the effects could be confounded if they are modeled separately. The results don’t seem to indicate that the variables are confounded, but this could also be checked by calculating the within-unit correlations.

Line 138: “Since the pooling samples were sufficient for statistical inference, . . .”. It is unclear what is meant by this statement, so it should be clarified.

Lines 145-147: “1. No other important. . .” would read more clearly as “1. There are no other important. . .”. Likewise, “2. No interactions between. . .” would read more clearly

C3

as “2. There are no interactions between. . .”

Line 155: different exponential formatting is used between Equations 6 and 7 (exp vs e). Equation 7 is also basically the same as Equation 6.

Line 158: “ $|\exp(g(X_i, t_2) - 0) - 1| < 10\%$ ”. I believe the “-0” should be replaced with $g(X_i, t_1)$?

Lines 158-160: Are the trends in annual maximum streamflow calculated using $\log(Q)$? This is what I expect based on the presentation of the results, but should be clarified in the text.

Line 191 (and elsewhere): Suggest changing “large and middle dams” to “large and medium dams” or “medium and large dams”.

Line 199: “The number of catchment groups $\delta\hat{S}Y$ in Section 2.2 had no optimal value.” Was there a method used to test for an optimal value?

Line 209: “with only one exceptional type of effect for Urban and Crop”. This phrasing is confusing, so I suggest rephrasing to something like: “except for two cases (Urban effect when $k=10$, and Crop effect when $k=40$)”.

Line 210: “percentage increase” – I believe this should be a percentage point increase? (For example, an increase from 10% to 11% is a 10 percent increase, or a 1 percentage point increase.)

Line 238 and 348: “more than 10% of increases in Q”. If I understand correctly, this should be written as “increases in Q of more than 10%”.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2020-609>, 2020.

C4