Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-609-RC2, 2021

Hydrology and Earth System Sciences Discussions
Open Access

HESSD

Interactive comment

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

Marc F. Muller (Referee)

mmuller1@nd.edu

Received and published: 21 January 2021

The study applies a panel research design to estimate the causal effect of three hypothesized human-related drivers (urban extent, cropland extent and reservoir regulation) of annual flood peaks in China. While the methodological contributions of the study are (in my view) limited compared to recent other studies using panel regressions in a similar context (e.g., Blum 2020 and Davenport 2020 cited in the study), the study is nonetheless valuable in that it provides important insights on how these process operate in conjunction, using a very large dataset in China. The study is in my view appropriate for publication in HESS, provided the author address the following major concerns that I have.

1. Potentially misleading map figures. To be clear, the panel approach does *not* allow to estimate heterogeneous treatment effects. It allows to estimate one average effect of (say) urban expansion on flow peaks (i.e. one single value of beta, if g() is linear) across the whole sample. It does *not* allow to say that urban expansion has a larger effect on flood peaks in some regions than in others. Yet the maps in figures 6 and 7 (and their discussion throughout the paper) appear to suggest exactly that, which I find misleading. The spatial variability in the "effect" of crop/urban on floods represented in these maps only emerges because changes crop and urban cover are themselves varying across regions. Figure 6 is nothing more than a map of urban cover change, scaled by a constant factor (the estimated beta) representing the linear effect it has on flood peaks. This point is important to clarify throughout the text, at the very least by specifying the estimated value of beta and theta in the captions of Figures 6 and 7 (see minor comments for other suggestions).

2. Fixed Effects. I am wondering why you use "regions" as space fixed effects, and not the individual basins themselves. For Blum et al., this approach made sense because they interact the treatment (X) with covariates (e.g., soil permeability, etc) but I don't really see the point of doing that here. I am concerned that it might introduce a bias associated with varying confounding factors within the regions (e.g., basin altitude can vary within regions and affect both the treatment âĂŤ crop or urban cover âĂŤ and flood magnitude). Adding a specification with basin-level fixed effect (i.e. setting k==number of basins) as robustness check might help alleviate my concern.

3. Heterogeneous treatment effect: I am wondering if your results are affected by heterogenous treatment effects in the sense that most basins of the sample likely have little impervious surface cover. (By the way, please add a table with descriptive statistics for the reader to assess that). If the deviates (even slightly) from the three arbitrary functional forms that you impute to g(), this may potentially bias your average estimates. A way to control for this (perhaps) would be to do a robustness check by running the analysis to a subset of highly (lowly) impervious basin to see how sensitive the effect

is.

4. Nestedness: Finally, the ordinary least square estimator that (I assume) you are using only provides an unbiased estimate of standard errors if residuals (epsilon) are independant. In your case, I am concerned that many of your observations might be nested (i.e. taken along different reaches of a same river), which might introduce a correlation in the epsilon. For instance a time- and space- specific shock on flow peaks observed in a headwater catchment will likely affect flow peaks observed at several gauges along that river. The fact that errors congregate around specific basins in Figure 8 is actually a strong indication of that effect! This effect might lead you to underestimate the standard errors on your regression coefficient and find a significant effect where there is none. A way to address that would be to use the topology of your river network to specify the structure of your variance-covariance matrix (see, e.g., Muller and Thompson 2015) which you can then incorporate in your estimation via Generalized Least Square or Restricted Maximum Likelihood. Alternatively, you could do a robustness check where you run your OLS estimation on multiple subsets of your full sample, for which you made sure that all observations are from different catchments âĂŤ hopefully the results will be similar.

Minor Comments.

The first sentence of the abstract is awkward (“because the knowledge and observations toward the effects are limited”). Please reformulate.

L79: middleâĂŤ> medium ?

L94: It took me a while to realize that you *defined* your regions such that climate is homogeneous within them (as oppose to assuming that climate is homogeneous within a bunch of predetermined regions). Maybe clarify that here?

Eqn 6: My understanding is that ΔQ varies of space but not time: if so, how to you “average over” the time index in the middle expression. Also, this would be an ideal place to clarify that ΔQ varies in space only because ΔX varies in space. Your estimation of g() is constant in space and time.

Fig 2: I agree with the other reviewer that p-values are an odd criteria for model selection. Either justify it, or use goodness of fit metric.

L155-160 and Fig 8. I find it a good idea to analyze the spatial distribution of model deviations (i.e. locations where variations in Q are not explained by the modeled drivers), but I find the approach chosen to identify these locations odd/arbitrary and challenging to understand. Wouldn't it be more straightforward to simply map the temporal variance of the residuals (i.e. Var_i(Eps_it))?

L294 "Coefficient of Variation" can be understood as the ratio between the standard deviation and the mean. I don't think that's what you mean here, so please reformulate.

L294. You provide a good illustrative example of the models inability to capture heterogeneous treatment in time, but here would also be a good opportunity to give an example of a heterogeneous treatment in space (i.e. a scenario where cropland might persistently have a stronger effect on flow peaks in some locations than in other). That would contribute alleviating my first major concern, above.

SI. Please add a descriptive statistics table with key stats on all the considered variable across your sample.

Marc Muller

References Blum, A. Et al (2020) Causal Effect of Impervious Cover on Annual Flood Magnitude in the United States, GRL

Davenport et al. 2020 Flood Size Increases Nonlinearly Across the Western United States in Response to Lower Snow-Precipitation Ratios, WRR

Muller, M.F. and Thompson, S.E. (2015) "TopREML: a topological restricted maximum likelihood approach to regionalize trended runoff signatures in stream networks", HESS

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

Printer-friendly version

Discussion paper