Replies to reviewer JoH

Willem Vervoort/Eliana Nervi/Jimena Alonso

2022-06-13

This document records the replies to the reviewer for the first submission to Journal of Hydrology which was rejected.

Reviewer comments in blue responses in normal text

The most important point from the reviews is the fact that we have not communicated the need for this paper.

Editor comments

The first is to streamline the statistics. As suggested by the AE, a formal model selection process, followed by using only the selected model(s) to evaluate change, would be a suitable approach. In doing so, please carefully note the AEs' concerns about some of the statistical methodology - concerns which range from interpretation to the use of appropriate performance metrics across models of varying structure, to the selection of the appropriate metrics of forest cover change for analysis. The second - hopefully supported by a streamlining of the statistics - is to simplify and improve the coherence of the argument. Framing - as the AE states - a "systematic" exploration of the importance of parameters, will improve the readability and interpretability of the work. At present, I think because the statistical analysis itself is convoluted - so too the thread of the argument and clarity of the messages are hard to follow. The final - and perhaps in truth the most problematic issue that may stand in the way of the MS being published - relates to novelty. The AE highlights several recent works with similar questions, approaches and findings at global scales. I could add to that list with more regional studies (e.g. Levy et al 2018). So carefully identifying the knowledge gap being addressed, with respect to these recent studies, and making the case for the present study being "needed" will also be essential.

These are substantial changes and go beyond a major revision. For this reason, we're rejecting the MS at present. If the authors are able to address the 3 issues above in a substantial revision of the MS, we would be pleased to look at it again.

We would, however, consider as a new submission for review a substantially revised version of this paper that addresses all of the reviewers' comments. Should you choose to submit such a revised manuscript please refer to the present manuscript number, provide a detailed point-by-point reply to all of the reviewers' comments, and state how the revised manuscript addresses these.

Associate Editor:

The manuscript considers an enhanced dataset of streamflow and forest cover, to explore how deforestation/afforestation alters catchment water yields. The manuscript is potentially of interest of the JoH readership, but it is not ready for review in present form.

The main aspects that need to be addressed before the manuscript can be evaluated by experts in the field are listed here. - As apparent from the diagnostic plots, the model assumptions may be violated in many cases. This can make the results of the fitting (and hence the manuscript conclusions) incorrect. I urge the

authors to double check if this is indeed the case and consider ways to address the problem. It is also good practice to check the relevance of outliers (of data with high VIF) and set them aside before model fitting. It is also not correct to comment on models as if working better or worse in certain ranges, based on the residuals (P 29), because the residuals are the results of the data and fitted model, and the fitted model depends on all datapoints.

Thank you for raising these important points in relation to the validity of the statistical model

- The manuscript presents a number of alternative statistical models, differing by candidate explanatory variables. Each model is designed considering the key shortcomings of the previous one. The end result of such an approach is a complex and somewhat non systematic exploration of predictors and their explanatory power, where it is easy to get lost. I suggest restructuring the manuscript around a well-designed and robustly formalized model selection. One way to proceed could be to start with the most complex model suggested by the extant understanding of the processes at play, and then proceed with a model simplification, according to some consistent criteria (AIC, dropping non significant terms, or similar; high r2 is not a good criterion because it does not consider the number of parameters). A full blown model selection would also allow to retain or discard the interaction terms, which could be important (as also recognized by the authors; Section 4.5) and should not be discarded a priori. Doing a proper model selection and presenting the results only for the best model (according to a clearly specified criterion) would be less subjective and allow to drastically reduce the number of figures and tables, allowing the reader to focus their attention on the key message.

- The novelty of this work needs to emerge more clearly in the introduction. As it looks now, the manuscript could be easily considered somewhat confirmatory, with respect to most data, approaches and conclusions reached by Zhang et al 2017 and Filoso et al 2017. Furthermore, the introduction needs to be rearranged, starting with a clear statement of the problem, what we know about that based on previous results, what is missing/how these previous analyses can be improved, and, stemming from these knowledge gaps and/or our understanding of the mechanisms, the questions addressed in the work or the hypotheses tested.

There are also some typos and unfinished sentences (e.g., L 142, L 298). Some units are missing (for example those of length of the experiment in the figures) and symbols are not defined at their first appearance (E0/Pa in L 99; Dryness Index). Also: how is Table 1 used? These are not big issues per se but are nonetheless distracting.

I would also like to provide the authors with a couple of suggestions regarding the statistical model and their interpretation. - The models used in the manuscript consider the absolute value of the forest cover change and then its sign, but this choice is not well justified. It implicitly assumes that the status corresponding to no change distinguishes two 'realms'. Yet, I would expect (and it is also hinted at at some point in the manuscript) that what really matters is the %forested area (possibly in relation to the climatic conditions) and how it changes. So, I would suggest the authors to consider whether a model nearer to our understanding of the phenomena at play would be one including, for the forest part, %change in forested area (with sign) and %forested area, with the latter possibly as random effect, if not of interested. - The fact that the explanatory power is low (low r2) does not necessarily make the results uninteresting (against conclusion on L 530), simply it suggests there are other factors, not included in the model, which have a large effect, and that the model presented cannot be used in a predictive mode. While it is important to present also the r2, even a model with low r2 square we learn which factors significantly affect the change in streamflow and which do not do so.