

Replies to reviewer JoH

Abstract

This document records the replies to the reviewer for the first submission of “Factors determining how catchments respond to forest cover change. Re-analysing global data sets” 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. We need to rewrite the introduction to do this better.

0.1. 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.

We politely disagree with this, as we outline in our reply to the AE. If the statistical modelling aimed at developing the best predictive model, then this would be the right approach. However, in this case the statistical modelling is aimed at hypothesis testing and explanation of variance in the data set.

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.

Thank you for the suggestion of the Levy et al. (2018) paper, which is an excellent example of a careful statistical analysis taking into account possible variations in climate and dynamic landuse change. This is exactly the kind of statistical analysis that we believe needs to be undertaken to better understand how forest cover impacts The paper also provides the rainfall and runoff

data that was used in the study, but regrettably does not provide the landuse data. Rather than deriving this ourselves following the methods described in the paper, we decided that it would be better to use Levy et al. (2018) as a “verification” of our analysis.

Impact of forest cover on streamflow: it is not that simple. While easily hypothesised, the research clearly shows that the causal relationship between change in forest cover and streamflow is complex and not as straight forward as shown in earlier literature. In terms of the “novelty” of our work, we provide two key insights.

- While global databases can be interesting, we need to be careful with drawing major conclusions (as in Zhang et al. (2017), Filoso et al. (2017), Zhou et al. (2015) and Jackson et al. (2005)) based on basic regression analysis. In many cases statistical assumptions are violated and confounding factors can hide or strengthen assumed relationships. This is without considering the number of errors that existed in the data. This is particularly important, since results from these global analyses are used to build further models to analyse impacts (e.g. Hoek van Dijke et al. 2022), leading to possible wrong policy or management responses.

- Cumulative and average values of change can be misleading, especially when extracted

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.

0.2. 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. A first point that arises from this is that we clarify better what the aim of our statistical modelling is.

There are in essence two “modes” in relation statistical modelling. Generally a model is developed to be used in predictive mode: using a model to predict unknown values, either within or beyond the current data set (forecasting). In this case the model should be reduced to its most efficient version that minimises the bias - variance trade-off. In this case, automatic variable selection and potentially validation on independent data are important, as the aim is to develop the most robust model for prediction.

However, a second reason for fitting a statistical model is purely to explain the variance in the data. In this case, it is important to develop a a-priori hypothesis about the causal relationships in the data. This is subsequently followed by a step by step analysis to test the different causal relationship, either as single variables (as in Zhang et al. (2017)) or jointly (as in our approach). In this case there is no attempt to find the best “predictive model”, the focus is on the additional amount of explained variance from adding each variable.

Include a section in the paper to explain this clearly

Either way, understanding the diagnostic plots and the residual distribution is important, which is why this was included in the manuscript. Including such diagnostic plots for single variable regression is often omitted. For example, both Zhang et al. (2017) and Filoso et al. (2017) do not present any diagnostics for their relationships and simply looking at their regressions qualitatively indicates issues with the residuals.

In our case, we clearly indicate the steps we have taken to improve the quality of the regressions, such as transforming some of the variables.

Because our modelling is focused on explaining the maximum variance, excluding values that are outliers is probably not a good idea. However, identifying outliers (particularly values with high leverage), and investigating why these might be different from the overall distribution is a great suggestion.

Check Venables and Ripley p119 “Outliers are sample values that cause surprise in relation to the majority of the sample. This is not a pejorative term; outliers may be correct, but they should always be checked for transcription errors.”

However, another careful review of the data identified many further errors in the data, which were all originally in Zhang et al. (2017). A particular problem was that many catchments had the wrong sign for the change in forest cover. There are many catchments with reported positive change in cover and a large increase in flow. These were all checked and corrected:

- 124, D3, Amatya and Skaggs, 2008: The originally recorded 250% change by Zhang et al. (2017) is clearly wrong. The paper says on page 7: Both of these outflow ratios (0.64 and 0.50) were higher than the calculated expected values of 0.55 for 2003 and 0.44 for 2005, respectively. So value should be $0.64/0.55 * 100 - 100$ or $0.5/0.44 * 100 - 100$: 16% or 13%. corrected to 16% - 3, Baker Creek, Zhang and Wei, 2012. The original recorded 201.1% change by Zhang et al. (2017) is also wrong. Original paper says on page 2031: Annual

mean flow has been increased by 47.6%. corrected

- 67, April rd, which is incorrectly attributed to Ruprecht and Schofield (1991) in Zhang et al. (2017). This is actually from Ruprecht and Schofield (1989) and the original paper clearly indicates “clearfelling”. As a result the change in forest cover was changed to -100% rather than +100%. - 210, March rd, 100, 147.6. Same problem as 67, Bari et al. (1996) clearly state that the catchment was cleared, so therefor the change in forest cover changed to -100%. - 213, 214 and 215, Monda 1, 2 and 3. These catchments are tricky. The original paper (O’Shaughnessy et al. 1979) only reports on the control period and indicates that the catchments will be cleared. The later summary paper (Watson et al. 2001) shows the timeseries of the flow change, but does not report a single value, so the values in the database must have been estimated from the timeseries. The further complication is that the treatment included clearing and reseedling and regrowth. This suggest that the records should be removed from the database, or only the first few years of the experiment used. In any case, if the values are kept, the sign of the change in forest cover needs to be changed to negative (Clearing).

- 230, Oleolega catchment. The paper describes a removal of forest up to 85%. changed Delta_F_perc to -85 from 90.

- 312, Yerraminup South. The original publication for this catchment is a Western Australian Water Authority report from 1987, which is hard to find, but we have added a copy in the “Papers” folder on github. In this report, in Table 2 on page 11, for the catchment a “Crown cover” decrease of 60% is given. Changed the sign of the change in forest cover: -60%.

- 100 Coachwood, 103 Corkwood, and 117 Bollygum, as cited by Cornish (1993) and Cornish and Vertessy (2001). In the database from Zhang et al. (2017), the forest change for all these catchments is positive. However, the paper highlights that these catchments were all logged and either naturally regenerated or were planted with a plantation species. So, similar to the earlier mentioned Monda catchments, the reported change probably only refers to the first couple of years after clearing (before regrowth). In any case, the reported change in forest cover should be negative (clearing) rather than positive. Corrected for all three catchments.

- 78, Black Spur 1, the treatments and effects are only reported in a conference paper (Jayasuriya and O’Shaughnessy 1988) and once again indicated clearing, meaning that the change in forest cover should be negative rather than positive (as reported in Zhang et al. (2017)). Corrected. Similar to other paired watershed experiments, only the first couple of years can be linked to the effect as later regrowth cancels out part of the increase in flow.

A VIF analysis is also a great suggestion as this will indicate variables which are highly correlated within the dataset, such as the Precipitation and Dryness (which we already indicate).

```
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```

## Warning: package 'tidyverse' was built under R version 4.1.3

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.6      v purrr  0.3.4
## v tibble  3.1.6      v dplyr  1.0.9
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.1      v forcats 0.5.1

## Warning: package 'ggplot2' was built under R version 4.1.3

## Warning: package 'tidyr' was built under R version 4.1.3

## Warning: package 'dplyr' was built under R version 4.1.3

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

require(lubridate)

## Loading required package: lubridate

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union

require(mgcv)

## Loading required package: mgcv

## Warning: package 'mgcv' was built under R version 4.1.3

## Loading required package: nlme

##
## Attaching package: 'nlme'

## The following object is masked from 'package:dplyr':
##
##   collapse

## This is mgcv 1.8-40. For overview type 'help("mgcv-package")'.

```

```

require(car)

## Loading required package: car
## Loading required package: carData
##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##      recode

## The following object is masked from 'package:purrr':
##
##      some

All_data <- read_csv("All_data.csv")

## Rows: 327 Columns: 24
## -- Column specification -----
## Delimiter: ","
## chr (10): Watershed name, Forest_type, Hydrological_regime, Precip_data_type...
## dbl (14): Watershed #, Area_km2, Pa_mm, DeltaF_perc, DeltaQf_perc, E0, Latit...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

All_data2 <- read_csv("All_data2.csv")

## Rows: 327 Columns: 26
## -- Column specification -----
## Delimiter: ","
## chr (11): Watershed name, Forest_type, Hydrological_regime, Precip_data_type...
## dbl (15): Watershed #, Area_km2, Pa_mm, DeltaF_perc, DeltaQf_perc, E0, Latit...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

All_data2 <- All_data2 %>%
  mutate(length = To - From,
         mid_year = From + (To - From)/2)

#Maybe first simply
# https://stackoverflow.com/questions/52554336/plot-the-equivalent-of-correlation-matrix-for
library(ggcorrplot)

```

```
## Warning: package 'ggcorrplot' was built under R version 4.1.3
```

```
png("testcorr_plot.png", width = 960, height = 960)
model.matrix(~0+DeltaF_perc_pos +
             Forest_Sign +
             log10(Area_km2) +
             Dryness +
             length +
             Precip_data_type + Assessment_technique +
             Forest_type +
             Hydrological_regime, data=All_data2) %>%
  cor(use="pairwise.complete.obs") %>%
  ggcorrplot(show.diag = F, type="lower", lab=TRUE, lab_size=4)
dev.off()
```

```
## pdf
## 2
```

```
# model6_all <- gam(DeltaQf_perc ~ DeltaF_perc_pos +
#                  Forest_Sign +
#                  log10(Area_km2) +
#                  Dryness +
#                  length + # s(From, k = 40, bs = "ts") +
#                  Precip_data_type + Assessment_technique +
#                  Forest_type +
#                  Hydrological_regime
#                  , data = All_data2)
# summary(model6_all)
# gam.check(model6_all)
# #plot(model6_all)
# inf1 = influence.gam(model6_all)
# hist(inf1)
# hist(cooks.distance(model6_all))
# which.max(cooks.distance(model6_all))
#
# regclass::VIF(model6_all)
#
# # Can only do car::vif on linear model and only on numerical values
# car::vif(gam(DeltaQf_perc ~ DeltaF_perc_pos + Forest_Sign +
#              log10(Area_km2) +
#              Dryness +
#              length, data = All_data2))
```

So the VIF really does not work, only works for numerical values and not for factor values. So this suggestion is not really relevant either.

However, being too focused on just the residual distribution can also lead to over fitting or over smoothing of the variables. This means that important characteristics of the data might be lost.

- 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 r^2 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.

As outlined in the last part of the reply, the focus of the statistical modelling in the paper is on understanding the different covariates that can explain the variation in the data

- 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.

This is a good point and we have rewritten the paper to strengthen this point.

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 r^2) 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 r^2 , even a model with low r^2 square we learn which factors significantly affect the change in streamflow and which do not do so.

- Cornish, P. M. 1993. "The Effects of Logging and Forest Regeneration on Water Yields in a Moist Eucalypt Forest in New South Wales, Australia." Journal Article. *Journal of Hydrology* 150 (2-4): 301–22. <http://www.sciencedirect.com/science/article/B6V6C-487D3Y2-9J/2/73c981ba76284d9d629f6b221d6fd6c6> .
- Cornish, P. M., and R. A. Vertessy. 2001. "Forest Age-Induced Changes in Evapotranspiration and Water Yield in a Eucalypt Forest." Journal Article. *Journal of Hydrology* 242 (1-2): 43–63. <http://www.sciencedirect.com/science/article/B6V6C-429910G-3/2/0158b1f89ff436f338a9e688a47f06c4> .
- Filoso, Solange, Máira Ometto Bezerra, Katherine C. B. Weiss, and Margaret A. Palmer. 2017. "Impacts of Forest Restoration on Water Yield: A Systematic Review." Journal Article. *PLOS ONE* 12 (8): e0183210. <https://doi.org/10.1371/journal.pone.0183210>.
- Hoek van Dijke, Anne J., Martin Herold, Kaniska Mallick, Imme Benedict, Miriam Machwitz, Martin Schlerf, Agnes Pranindita, Jolanda J. E. Theeuwens, Jean-François Bastin, and Adriaan J. Teuling. 2022. "Shifts in Regional Water Availability Due to Global Tree Restoration." Journal Article. *Nature Geoscience* 15 (5): 363–68. <https://doi.org/10.1038/s41561-022-00935-0>.
- Jackson, Robert B., Esteban G. Jobbagy, Roni Avissar, Somnath Baidya Roy, Damian J. Barrett, Charles W. Cook, Kathleen A. Farley, David C. le Maitre, Bruce A. McCarl, and Brian C. Murray. 2005. "Trading Water for Carbon with Biological Carbon Sequestration." Journal Article. *Science* 310 (5756): 1944–47. <https://doi.org/10.1126/science.1119282>.
- Jayasuriya, M. D. A., and P. J. O'Shaughnessy. 1988. "The Use of Mathematical Models in Evaluating Forest Treatment Effects on Streamflow." Book Section. In *Hydrology and Water Resources Symposium 1988: Preprints of Papers*, 135–39. Hydrology; Water Resources Symposium. <https://doi.org/10.3316/informit.692214289455295>.
- Levy, M. C., A. V. Lopes, A. Cohn, L. G. Larsen, and S. E. Thompson. 2018. "Land Use Change Increases Streamflow Across the Arc of Deforestation in Brazil." Journal Article. *Geophysical Research Letters* 45 (8): 3520–30. <https://doi.org/https://doi.org/10.1002/2017GL076526>.
- O'Shaughnessy, P. J., K. J. Langford, H. P. Duncan, and R. J. Moran. 1979. "Catchment Experiments in Mountain Ash Forests at North Maroondah." Journal Article. *Australian Forestry* 42 (3): 150–60. <https://doi.org/10.1080/00049158.1979.10674220>.

- Watson, Fred, Rob Vertessy, Tom McMahon, Bruce Rhodes, and Ian Watson. 2001. "Improved Methods to Assess Water Yield Changes from Paired-Catchment Studies: Application to the Maroondah Catchments." Journal Article. *Forest Ecology and Management* 143 (1): 189–204. [https://doi.org/https://doi.org/10.1016/S0378-1127\(00\)00517-X](https://doi.org/https://doi.org/10.1016/S0378-1127(00)00517-X).
- Zhang, Mingfang, Ning Liu, Richard Harper, Qiang Li, Kuan Liu, Xiaohua Wei, Dingyuan Ning, Yiping Hou, and Shirong Liu. 2017. "A Global Review on Hydrological Responses to Forest Change Across Multiple Spatial Scales: Importance of Scale, Climate, Forest Type and Hydrological Regime." Journal Article. *Journal of Hydrology* 546: 44–59. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2016.12.040>.
- Zhou, Guoyi, Xiaohua Wei, Xiuzhi Chen, Ping Zhou, Xiaodong Liu, Yin Xiao, Ge Sun, et al. 2015. "Global Pattern for the Effect of Climate and Land Cover on Water Yield." Journal Article. *Nature Communications* 6 (1): 5918. <https://doi.org/10.1038/ncomms6918>.