Do larger catchments respond different to forest cover change? Re-analysing a global data set.

R. Willem Vervoort[*,a,1], Eliana Nervi[b], Jimena Alonso[**,c]

[a] *School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia*
[b] *INIA, Uruguay*
[c] *Institute of Fluid Mechanics and Environmental Engineering, School of Engineering, Universidad de la República, 11200 Montevideo, Departamento de Montevideo, Uruguay*

Abstract

This is the abstract.

It consists of two paragraphs.

Introduction

Introduction

There has been an long and on-going discussion in the hydrologcal literature around the impact of forests on streamflow (Andréassian, 2004; Brown et al., 2013, 2005; Jackson et al., 2005; Zhang et al., 2017). The historic work highlights provides a general consensus that if forest areas increase, streamflow decreases and vice-versa. The most dramatic result in relation to this, is Figure 5 in Zhang et al. (2011) indicating (or Australian watersheds) a 100% decrease in stream flow for watersheds with 100% forest cover. However, on the other end of the spectrum, in a series of French watersheds (Cosandey et al., 2005), there was no change in streamflow characteristics in 2 of the three watersheds studied in relation to deforestation.

There have been several review papers aiming to summarize different studies across the globe, in relation to paired watershed studies (Bosch and Hewlett, 1982; Brown et al., 2005) and more generally (Jackson et al., 2005; Zhang et al., 2017). These studies are aiming to generalize the individual findings and to identify if there are global trends or relationships that can be developed. The most recent review (Zhang et al., 2017) developed an impressive database of watershed studies in relation to changes in streamflow due to changes in forest cover based on a global data set. This dataset, which covers over 250 studies are described in terms of the change in streamflow as a result of the change in forest cover, where studies related to both forestation (increase in forest cover) and deforestation (decrease in forest cover) were included.

The conclusions of the paper (Zhang et al., 2017) suggest that there is a distinct difference in the change in flow as a result of forestation or deforestation between small watersheds, defined as < 1000 km^2 and large watersheds > 1000 km^2. While for small watersheds there was no real change in runoff with changes

* Corresponding Author
** Equal contribution
Email addresses: willem.vervoort@sydney.edu.au (R. Willem Vervoort), eliananervi@gmail.com (Eliana Nervi), jalonso@fing.edu.uy (Jimena Alonso)

in cover, for large watersheds there was a clear trend showing a decrease in runoff with and increase in forest cover. Their main conclusion was that the response in annual runoff to forest cover was scale dependent and appeared to be more sensitive to forest cover change in water limited watersheds relative to energy limited watershed (Zhang et al., 2017).

Encouraged by the work presented by Zhang et al. (2017) and the fantastic database of studies presented by these authors, we believe we can add to the discussion by presenting further analysis of the data and by adding further watersheds and enhancements to the data base.

In particular, the main method in the work by Zhang et al. (2017) is using simple linear regression. And the main assumption is that the threshold at 1000 km^2 is a distinct separation between "small" and "large" watersheds. Given the fantastic data set collected, the analysis can be easily expanded to look at interactions between the terms and to test the assumption of a distinct threshold at 1000 km^2.

In particular, the objective of this paper is to 1) enhance the data set from Zhang et al. (2017) with further watersheds and spatial coordinates and 2) to analyse the possibility of non-linear and partial effects of the different factors and variables in the data base using generalised linear (GLM) and generalised additive models (GAM Wood (2006)). Finally we hope to point to further research that can expand our work and that outlined Zhang et al. (2017) to better understand the impact of forest cover change on streamflow.

Methods

The original data set

The starting point of this paper is the data base of studies which were included in Zhang et al. (2017) as supplementary material. The columns in this data set are the watershed number, the watershed name, the Area in km^2, the annual average precipitation (Pa) in mm, the forest type, hydrological regime, and climate type, the change in forest cover in % (ΔF%) and the change in streamflow in % (ΔQf(%), based on equation 1 in Zhang et al. (2017)), the precipitation data type, the assessment technique, and the source of the info, which is a citation. Several of these columns contain abbreviations to describe the different variables, which are summarised in Table 1.

Table 1 Summary of abbreviations of factors used in the Zhang et al. (2017) data set

Factor	Abbreviation	Definition
forest type	CF	coniferous forest
	BF	broadleaf forest
	MF	mixed forest
hydrological regime	RD	rain dominated
	SD	snow dominated
climate type	EL	energy limited

Factor	Abbreviation	Definition
	WL	water limited
	EQ	equitant
precipitation data type	OB	observed
	SG	spatial gridded
	MD	modelled
assessment technique	PWE	paired watershed experiment
	QPW	quasi-paired watershed experiment
	HM	hydrological modelling
	EA	elastictity analysis
	SH	combined use of statistical methods and hydrographs

While the Zhang et al. (2017) use the dryness index in their analysis, potential or reference evapotranspiration is not part of the data set. We combined the tables for small (< 1000 km^2) and large ($>= 1000$ km^2) watreshed data sets in our analysis. Some small naming errors and citations for some of the data sets for some of the small watersheds were fixed as we were familiar with the studies. But overall the original data set was not changed.

Additional data collection

To enhance the existing data set, this study added additional variables. The first variables added were the latitude and longitude for the center of the watershed as an approximation of its spatial location. Using this information annual average reference evapotranspiration (E0) was extracted from **XXXXX** if value of E0 were not available from the original papers. For large water watersheds, this value, similar to annual average rainfall, is an approximation of the climate at the location.

The length of the study can be a variable influencing the change in flow (e.g. Jackson et al., 2005) and therefore, the length of the different studies was extracted from the references provided by Zhang et al. (2017).

Several additional data points from watershed studies were extracted from Zhang et al. (2011), Zhao et al. (2010), Borg et al. (1988), Thornton et al. (2007), Zhou et al. (2010), Rodriguez et al. (2010), Ruprecht et al. (1991) and Peña-Arancibia et al. (2012), and these were checked against the existing studies to prevent overlap. In the citation column in the data set, in general the main reference for the calculated change in streamflow was used, because sometimes the original study did not provide the quantification of the change in streamflow (i.e. Table 6 in Zhang et al. (2011))

The final column in the improved data set is a “notes” column, which is not further used in the analysis, but gives context to some of the data for future research.

Statistical modelling

To estimate how the change in streamflow is affected by the change in forest cover while considering the effects of the other variables, we applied generalised additive modelling (GAM) (Wood, 2006). In particular we fitted the following initial model to replicate the variables in the analysis from Zhang et al. (2017):

$$\begin{aligned}\Delta\%Q \sim\ & \Delta\% forest + s(Pa) + s(Area) + forest\ type + \\ & climate\ type + assessment\ type + hydrologic\ regime + \varepsilon\end{aligned} \tag{1}$$

However, the overall skewed distribution of the predictant ($\Delta\%Q$) is problematic, and this results in a skewed distribution of the GAM model residuals. As a result we transformed $\Delta\%Q$ and $\Delta\% forest$ back to fractions (0 - 1) and log transformed using $log10(x+1)$, where x is either ΔQ or $\Delta forest$. This means that the model residuals are $\sim N(0, \sigma^2)$ results in the following equation:

$$\begin{aligned}log10(\Delta Q) \sim\ & log10(\Delta forest) + s(Pa, k=3) + s(Area, k=3) + forest\ type + \\ & climate\ type + assessment\ type + hydrologic\ regime + \varepsilon\end{aligned} \tag{2}$$

In this model, the assumption is that all continuous variables (such as Pa) can have a linear or non-linear relationship with $log10(\Delta Q)$. This means that a smooth function $s()$ is applied to the variable. To restrict the smoothness of the fit, the smoothness factor k is restricted to a value of 3 (Wood, 2006). This restriction was applied to smooth variables throughout this paper and we have dropped this from the subsequent equations.

For the model in equation 2, we only used the data from Zhang et al. (2017) to make sure that the additional watersheds added to the data set did not influence the analysis. Given that in Zhang et al. (2017), dryness ($\frac{E0}{Pa}$) is used to look at variations in the change in flow, we also fitted the following model:

$$\begin{aligned}log10(\Delta Q) \sim\ & log10(\Delta forest) + s(\frac{E0}{Pa}) + s(Area) + forest\ type + \\ & climate\ type + assessment\ type + hydrologic\ regime + \varepsilon\end{aligned} \tag{3}$$

Subsequently, using the full data set, including the additional watersheds and the additional variables the following two models were fitted:

$$\begin{aligned}log10(\Delta Q) \sim\ & log10(\Delta forest) + s(Pa) + s(Area) + s(Latitude) + s(Longitude) + \\ & s(years) + forest\ type + climate\ type + assessment\ type + \\ & hydrologic\ regime + \varepsilon\end{aligned} \tag{4}$$

$$
\begin{aligned}
log10(\Delta Q) \sim\ & log10(\Delta forest) + s(\frac{E0}{Pa}) + s(Area) + s(Latitude) + s(Longitude) + \\
& s(years) + forest\ type + climate\ type + assessment\ type + \\
& hydrologic\ regime + \varepsilon
\end{aligned}
\tag{5}
$$

The results were analysed to identify:

1. the significance of the different variables
2. the direction of the categorical or shape of the smooth variables

References

Andréassian, V., 2004. Waters and forests: From historical controversy to scientific debate. Journal of Hydrology 291, 1–27. doi:https://doi.org/10.1016/j.jhydrol.2003.12.015

Borg, H., Bell, R.W., Loh, I.C., 1988. Streamflow and stream salinity in a small water supply catchment in southwest western australia after reforestation. Journal of Hydrology 103, 323–333. doi:https://doi.org/10.1016/0022-1694(88)90141-2

Bosch, J.M., Hewlett, J.D., 1982. A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. Journal of Hydrology 55, 3–23.

Brown, A.E., Western, A.W., McMahon, T.A., Zhang, L., 2013. Impact of forest cover changes on annual streamflow and flow duration curves. Journal of Hydrology 483, 39–50. doi:http://dx.doi.org/10.1016/j.jhydrol.2012.12.031

Brown, A.E., Zhang, L., McMahon, T.A., Western, A.W., Vertessy, R.A., 2005. A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. Journal of Hydrology 310, 28–61.

Cosandey, C., Andréassian, V., Martin, C., Didon-Lescot, J.F., Lavabre, J., Folton, N., Mathys, N., Richard, D., 2005. The hydrological impact of the mediterranean forest: A review of french research. Journal of Hydrology 301, 235–249. doi:https://doi.org/10.1016/j.jhydrol.2004.06.040

Jackson, R.B., Jobbagy, E.G., Avissar, R., Roy, S.B., Barrett, D.J., Cook, C.W., Farley, K.A., Maitre, D.C. le, McCarl, B.A., Murray, B.C., 2005. Trading water for carbon with biological carbon sequestration. Science 310, 1944–1947. doi:10.1126/science.1119282

Peña-Arancibia, J.L., Dijk, A.I.J.M. van, Guerschman, J.P., Mulligan, M., Bruijnzeel, L.A., McVicar, T.R., 2012. Detecting changes in streamflow after partial woodland clearing in two large catchments in the seasonal tropics. Journal of Hydrology 416-417, 60–71. doi:https://doi.org/10.1016/j.jhydrol.2011.11.036

Rodriguez, D.A., Tomasella, J., Linhares, C., 2010. Is the forest conversion to pasture affecting the hydrological response of amazonian catchments? Signals in the ji-paraná basin. Hydrological Processes 24, 1254–1269. doi:https://doi.org/10.1002/hyp.7586

Ruprecht, J.K., Schofield, N.J., Crombie, D.S., Vertessy, R.A., Stoneman, G.L., 1991. Early hydrological response to intense forest thinning in southwestern australia. Journal of Hydrology 127, 261–277. doi:https://doi.org/10.1016/0022-1694(91)90118-2

Thornton, C.M., Cowie, B.A., Freebairn, D.M., Playford, C.L., 2007. The brigalow catchment study: II*. Clearing brigalow (acacia harpophylla) for cropping or pasture increases runoff. Australian Journal of Soil Research 45, 496–511. doi:doi:10.1071/SR07064

Wood, S., 2006. Generalized additive models: An introduction with r. CRC Press, Boca Raton, FL.

Zhang, L., Zhao, F., Chen, Y., Dixon, R.N.M., 2011. Estimating effects of plantation expansion and climate variability on streamflow for catchments in australia. Water Resources Research 47, W12539. doi:10.1029/2011wr010711

Zhang, M., Liu, N., Harper, R., Li, Q., Liu, K., Wei, X., Ning, D., Hou, Y., Liu, S., 2017. A global review on hydrological responses to forest change across multiple spatial scales: Importance of scale, climate, forest type and hydrological regime. Journal of Hydrology 546, 44–59. doi:https://doi.org/10.1016/j.jhydrol.2016.12.040

Zhao, F., Zhang, L., Xu, Z., Scott, D.F., 2010. Evaluation of methods for estimating the effects of vegetation change and climate variability on streamflow. Water Resources Research 46, W03505. doi:10.1029/2009wr007702

Zhou, G., Wei, X., Luo, Y., Zhang, M., Li, Y., Qiao, Y., Liu, H., Wang, C., 2010. Forest recovery and river discharge at the regional scale of guangdong province, china. Water Resources Research 46. doi:https://doi.org/10.1029/2009WR008829