Supplementary Information part 2: Testing the improved data sets

R. Willem Vervoort??, Eliana Nervi, Jimena Alonso

Abstract

This supplementary material file compares whether the inclusion of additional catchments generates fundamentally different results as the original (but improved) data. Single variable regressions on the smaller (original) dataset are compared with the extended data set.

1. Introduction

In this document we tested whether the fundamental conclusions in the single variable regressions with the improved data base differed from the original conclusions in Zhang et al. [1]. This is to check how much influence the changes to the data set and the additional data might have changed the original conclusions.

2. Methods

First we will read in the data

We will combine the different tables, but will keep an indicator to see where the data are from.

```
Zhang_small$From <- as.numeric(Zhang_small$From)
Zhang_small$To <- as.numeric(Zhang_small$To)
Zhang_all <- bind_rows(Zhang_large,Zhang_small) %>%
  mutate(dataset = "original Zhang et al data")
new_data <- new_data %>%
  mutate(dataset = "new data")
All_data <- bind_rows(Zhang_all, new_data)
```

2.1. Implementing the changes to the overall data

The following code implements the changes described in the Supplementary data part 1. However, many of the changes were implemented manually into the data set. These are simply the remaining changes not implemented manually.

1. removing the duplicates.

*Corresponding author

Email addresses: willem.vervoort@sydney.edu.au (R. Willem Vervoort), eliananervif@gmail.com (Eliana Nervi), jalonso@fing.edu.uy (Jimena Alonso)

[1]Corresponding Author

```
All_data <- All_data %>%
  mutate(`Possible duplicate` =
           ifelse(is.na(`Possible duplicate`)==T,0,`Possible duplicate`),
         `Possible duplicate` = as.numeric(`Possible duplicate`)) %>%
  filter(`Possible duplicate` != 1)
```

2. calculating the dryness

```
# calculate dryness index
All_data <- All_data %>%
  mutate(Dryness = E0/Pa_mm)
```

3. remove watershed 1 (the Amazon) from the analysis

```
All_data <- All_data %>%
  filter(`Watershed #` != 1)
```

4. remove data set 188 and 254 Kamakia and Sambret

```
All_data <- All_data %>%
  filter(`Watershed #` != 188) %>%
  filter(`Watershed #` != 254)
```

5. add a column that indicates forst loss of forest gain

```
All_data <- All_data %>%
  mutate(forest_sign = ifelse(DeltaF_perc < 0, "Forest Cover Loss", "Forest Cover Gain"))
```

2.2. Approach and analyses

The approach is similar to Zhang et al. [1]. We run single variable regressions separating large (> 1000 km^2) and small catchments ($<= 1000$ km^2).

The paper by Zhang et al. [1] calculates the sensitivity of runoff as a function of runoff as:

$\Delta Q_f = 100 \times \frac{\Delta Q_{f,mm}}{Q}$

This first equation is superfluous in this case as the data (as extracted from Zhang et al. [1]) is already defined in terms of ΔQ_f.

$S_f = |\frac{\Delta Q_f}{\Delta F}|$

```
All_data <- All_data %>%
  filter(is.na(DeltaF_perc) == F) %>%
  mutate(S_f = abs(DeltaQf_perc/DeltaF_perc))
```

In sequence we analyse:

- the relationship between forest cover change and streamflow change for small and large catchments (i.e. Figure 2 in Zhang et al. [1]);

- the relationship between catchment size and the sensitivity to runoff change (i.e. Figure 3 in Zhang et al. [1]); and

- the sensitivity to forest loss as a function of dryness (i.e. Figure 4 in Zhang et al. [1]).

3. Results

3.1. The change in stream flow as a function of change in forest cover

Figure 1 highlights that the overall relationship in the updated dataset is the same as in Zhang et al. [1]. This means that while the modifications have cleaned up the transcription errors in the data, they have not fundamentally changed the conclusions in the original paper.

The next figure (Figure 2) is the same analysis, but this includes the new data that we identified in papers. Again, this figure highlights that the new datasets have not fundamentally changed the relationships found in Zhang et al. [1].

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

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

3.2. The relationship between the area of the catchment and the sensitivity of streamflow to the change in forest cover.

This analysis replicates Figure 3 in Zhang et al. [1], which investigates for large and small catchments the sensitivity to runoff change from change in forest cover as a function of area. Note that in the original figure, the x-axis is on a log scale. In the original paper, the analysis is presented for all catchments as well as for large and small catchments. Here we only analyse the small and large catchments.

We can see from Figure 3 that again the updated database for the original dataset results in little change in the relationships for both large and small catchments. However, when the additional new catchments are added to the database (Figure 4), the relationships clearly change. In particular, for small catchments gaining forest cover, the sensitivity appears positively correlated with the logarithm of the size of the catchments.

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

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

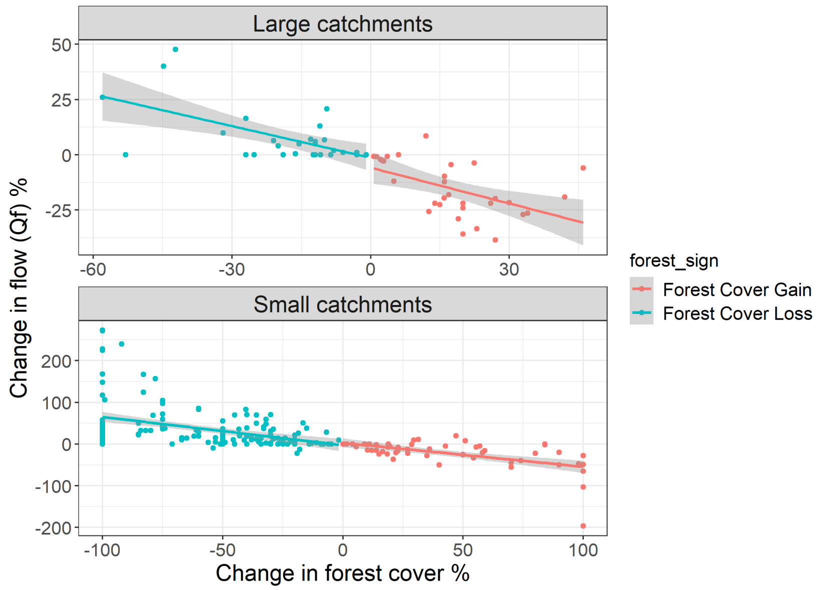

Figure 1: Changes in flow based on the catchments from the original data set

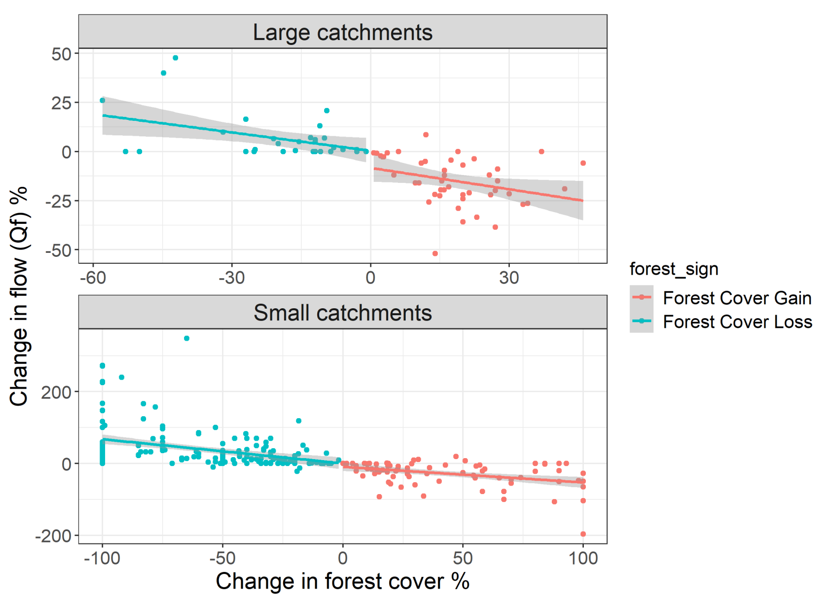

Figure 2: Changes in flow based on the catchments from the extended data set

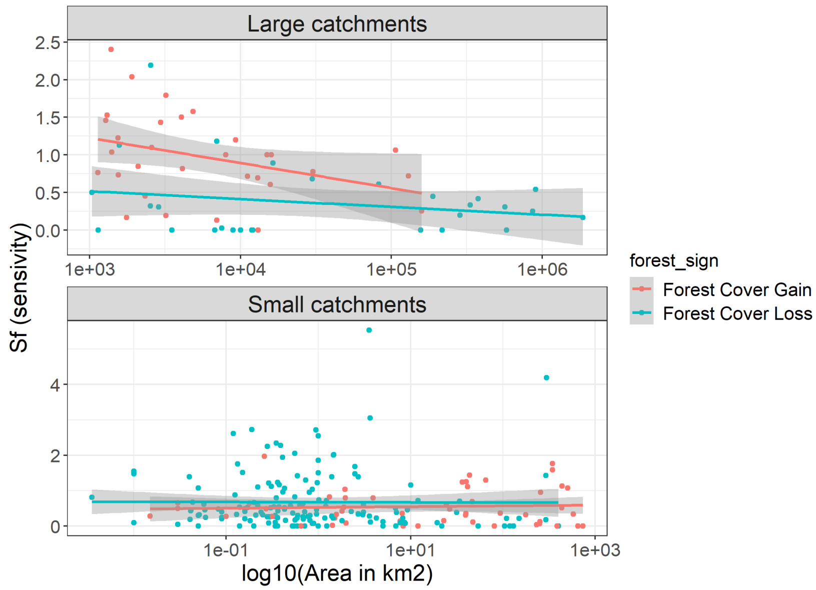

Figure 3: Changes in flow based on the catchments from the original data set

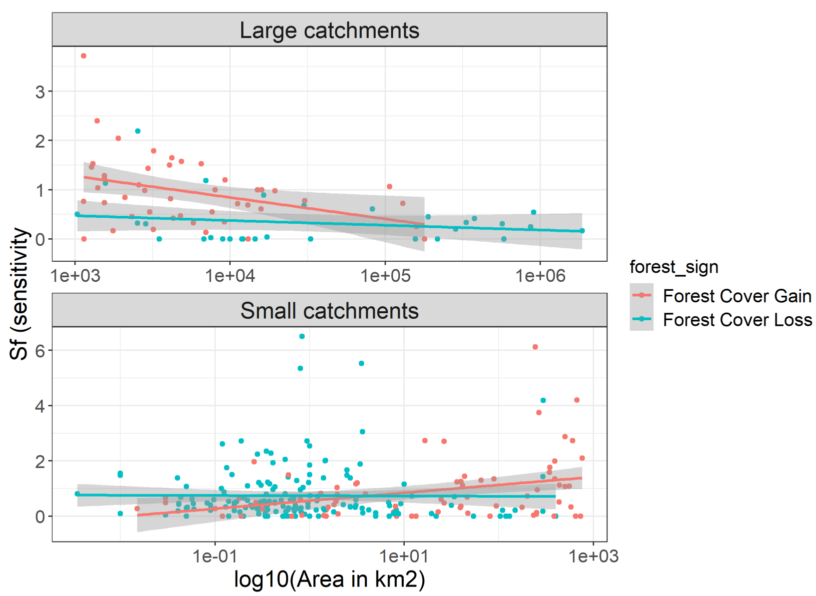

Figure 4: Changes in flow based on the catchments from the extended data set

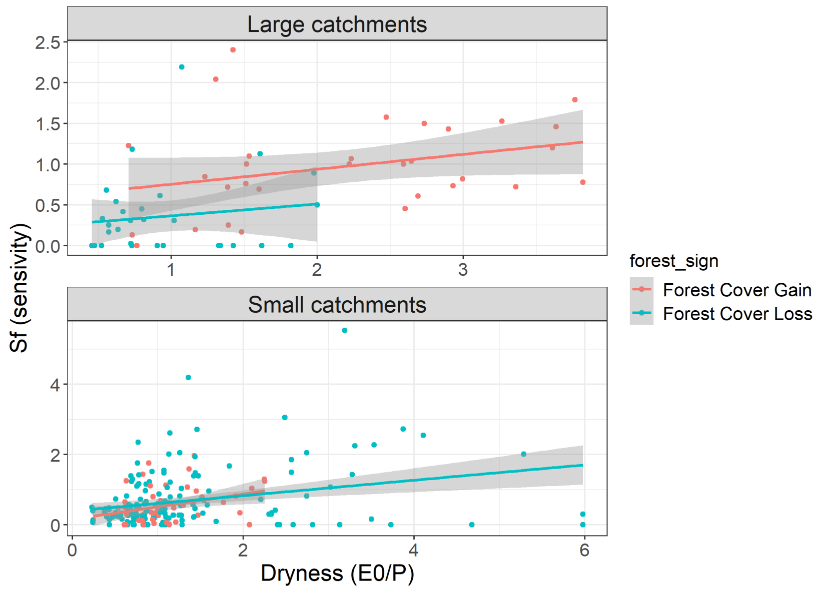

Figure 5: Changes in flow based on the catchments from the original data set

3.3. *The sensitivity to forest loss as a function of dryness*

The final analysis that we retest here is the relationship in Figure 4 in the original Zhang et al. [1] paper, which highlights the sensitivity to forest loss as a function of dryness. We are again showing just the for the small and large catchments, similar to the original paper.

Similar to earlier analyses in this document Figure 5 show that the updated database for the original dataset results in little change in the relationships for both large and small catchments. However, when the additional new catchments are added to the database (Figure 6), the relationships clearly change. In particular, for small catchments both for forest gains and losses the relationship changes and appears stronger.

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

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

References

[1] Mingfang Zhang, Ning Liu, Richard Harper, Qiang Li, Kuan Liu, Xiaohua Wei, Dingyuan Ning, Yiping Hou, and Shirong Liu. A global review on hydrological responses to forest change across multiple spatial scales:

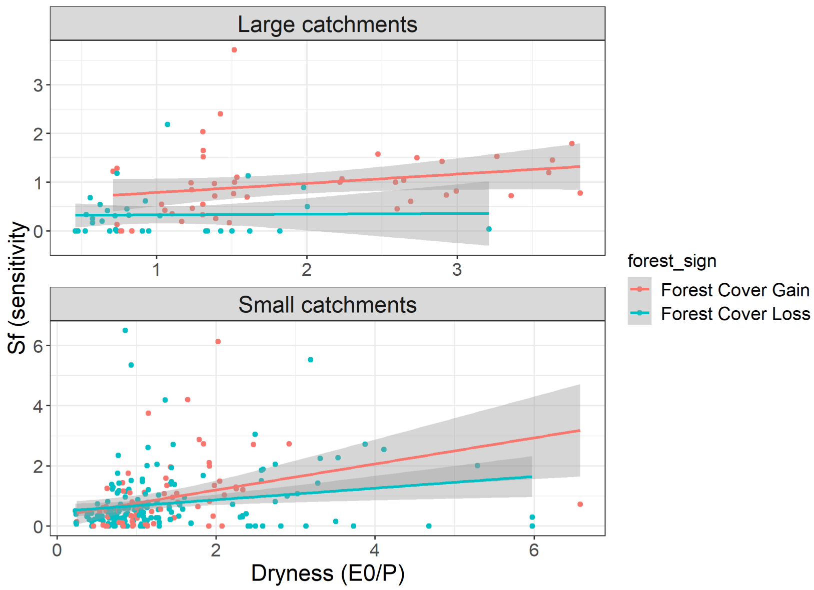

Figure 6: Changes in flow based on the catchments from the extended data set

Importance of scale, climate, forest type and hydrological regime. *Journal of Hydrology*, 546:44–59, 2017. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2016.12.040. URL `http://www.sciencedirect.com/science/article/pii/S0022169416308307`.