

As the authors rightly point out (lines 97-103) there is very little understanding on how changes in antecedent moisture conditions are modulating flooding under the assumption of non-stationarity. The authors not only demonstrate the modulating effect of changing antecedence on the 95th and 99th percentile of stream flow but also present results demonstrating the effect of catchment non-stationarity (e.g. urbanization). I enjoyed reading this manuscript. I find the manuscript to be novel and I recommend publication. A few minor comments are below.

Thank you for these positive comments about our work.

General comments: By chance of the three references I checked in the text two were missing from the bibliography e.g. Prein et al., 2016 and Bloschl et al., 2016. Please do check the referencing.

We have now included all references and checked the whole list.

Line 210-217: Does quantile regression use all the data? If so, how is the 99th per-centile of this comparable to the 99th percentile of the POT analysis where you only end up with a handful events per year?

Quantile regression only models the evolution of the 95th or 99th percentiles, not data above (or below). It is a convenient method to evaluate the change in different parts of the distribution and not only the mean.

Unlike the POT sampling, where all the values above the 95th or 99th percentiles are extracted, but then the consecutive threshold exceedances are de-clustered (i.e. only the maximum is kept) in order to avoid introducing autocorrelation.

Section 4.1: I get the impression that Figure 2 might not have used the Mann-Kendall test or quantile regression described in the methods but possibly a different technique? I am not sure. But this can be easily clarified but inserting a sentence at the start of this section.

We modified the caption of figure 2:

“Significant annual trends at the 10% level (Mann Kendall test) between 1958 and 2018 in precipitation, rainfall, temperature, soil moisture, actual evapotranspiration (AE) and reference evapotranspiration (ET0).”

We also clarified the text in the section as recommended:

“For each basin, the annual trends in precipitation, rainfall, temperature, soil moisture, actual and reference evapotranspiration have been analyzed with the Mann Kendall test.”

Line 330: You say “was not regionally significant” does that mean that all the figures have significance only tested on a site by site basis even though you said in the methods you use a FDR? This needs to be clarified. There a couple small grammatical things like line 48 “These results imply . . .” and Line166:

“As a very common . . .” but these are an easy fix that can be addressed at the editorial stag

We tested the significance of the trends at two levels:

- 1- At local scale, by simply reporting for each basin if the trend for a given variable is significant, at the 10% significance level.
- 2- At the regional scale, since the repetition of local tests and the possible cross-correlations between the different basin’s data could induce artificially more significant local values. This is why that FDR approach has been implemented, regional (or field) significance would be declared by this method if at least one local null hypothesis is rejected.

We better explain this point in the revised manuscript:

“This FDR method is applied to the MK test results to check if the trends are regionally significant. The detected trends are regionally significant if at least one local null hypothesis is rejected according to the global (or regional) significance level, α_{global} (Wilks, 2016). For consistency with the local trend analysis, the global significance level is also set to 10% in the FDR procedure.”

Line by line:

Line 70: A global study might make the point of high spatial variability better e.g. <http://dx.doi.org/10.1029/2011GL048426>; <https://doi.org/10.1002/2016GL071354> But this is at the authors discretion as it may be that they were referring to variability on a smaller spatial scale (not sure because the reference was missing).

We added the reference of Wasko et al 2016 in this section.

Line 106, 138, 482 I probably prefer “e.g” rather than “...” at the end of the examples. But again at authors discretion.

Changed in the text

Line 120: Number of rain days or mean rainfall or both?

As written, line 120, it is the number of rain days (usually days with precip > 1 mm).

Line 217: Remove “but preliminary tests” and just write “and”. This will sound more robust.

Changed

Line 274: Can you add the “with precipitation below 1mm” to the figure legend also please

Added

Line 325: I didn't think the trends were low? Actually they seemed quite large given the number of events per year?

Indeed that is a quite subjective statement. According to the average number of events per year, the maximum trend would correspond to a about -20% decrease in the number of events. I would not say it is large, but rather moderate. We changed the text to "moderate".

Line 400: Is the 30 and 365 day averages also lagged or is it just the coincident month/year that is averaged?

We agree this was not clear. We rephrased to =
At the monthly time scale representing the seasonal variability, the covariates have been averaged for the 30 days preceding the events. For the annual time scale the covariates have been averaged for 365 days preceding the events.

Line 401: remove "rather"

Removed

Line 403: opposite -> "other"

Changed

Line 421: "mean altitude" – this typo made me chuckle!

Indeed a funny spelling error! Thanks for noticing. We replaced attitude with elevation.

Line 427: This section needs rewording I think. You say "R values up to 0.6" for small basins but to counter this you say values "about 0.1 to 0.2" for large basins. One statistic is a maximum and the other is more related to the mean. It may be larger basin values also have R values up to 0.6 but I wouldn't know? Picking a more consistent statistic would give me more confidence in these results.

We modified the text as follows:

"for small basins (less than 500km²) event soil moisture and precipitation are good predictors for the time variations of the 95th and the 99th percentiles, with R^1 values up to 0.6, when for larger basins the R^1 values are much lower, reaching the maximum of 0.2 for some basins"

Figures: Can it be clarified in the manuscript text and on (every) figure legend that only statistically significant sites are shown (and at what level)?

We modified figure captions to be more precise in each case

Figure 2: At least on this figure (but preferably on all the figures) the axes should be labelled "lat/lon" on at least one panel.

We added lon/lat on all figures

Figure 2 caption: Add “rainfall” to the list of variables.

Added

Figure 3: Scale for triangles?

As requested by the other reviewers, we added on the plots the magnitude of the trends and in addition a table summarizing the trend testing results.

Figure 5: Are all sites presented here or just statistically significant ones

Thanks for asking this, because it helped us notice that we made a mistake by including on this figure 5 all the basins and not only those with a significant slope. We modified the figure. Now the figure is in agreement with its description in section 4.2 second paragraph.