(The authors' responses to RC#1 and RC#2 are included in this document)

Authors' response to RC#1:

"Recent Trends in Frequency and Duration of Global Floods" by Nasser Najibi and Naresh Devineni

Dear Reviewer

Thank you very much for reviewing our manuscript and providing feedback. We have made efforts to revise the manuscript, applied the suggested revisions to it, and have responded to the comments. Since we have re-written parts of the introduction for clarity, we combined our response to few language related comments.

Review of the manuscript "Recent Trends in Frequency and Duration of Global Floods" by Najibi and Devineni for publication in the journal Earth System Dynamics.

Dear Editor,

I have just read the abstract and the introduction section of the above-referenced manuscript. In these two sections (three pages total), I find the writing of the manuscript is of poor quality, and it is difficult to understand the sentences well. I have observed many grammatical inconsistencies in the first two pages of the manuscript (see some specific comments below), and there I stopped making further suggestions on grammar/professional writing.

With these many comments in just three pages, I think it is in the best interest of the potential readers to return the paper for English writing corrections. I strongly suggest the authors to send the manuscript to an English writing professional for modifications and corrections before sending it for a review, to this journal or elsewhere.

Comments related to writing

L1: "The" Dartmouth.

L13: Remove "also".

L13-14: 'This analysis provides framework for understanding changing climate and socioeconomic conditions' is a very broad statement; I suggest removing this sentence and replacing it with one that is more relevant to the present analysis.

L14: use comma after 'conditions'.

L23: put ', respectively,' before 'showed that'.

L23: replace "had" with "have"

L24-27: Sentence may be improved for clarity. Use appropriate punctuation at appropriate positions (e.g., I would put a comma before 'mainly'). The term 'flood events' shows two times in the sentence, consider rephrasing.

L31: Replace 'was' with 'is'.

L30-35: This sentence can be improved for conciseness and clarity.

L50-51: This is not professional writing. Please revise the sentence.

L37-39: Put a comma before 'and the physical'. Put 'events' after 'rainfall'.

L44: Replace 'a' with 'the'.

L45-47: Try more direct sentences. In the sentence, you have 'utilization', 'information', 'inundation', all are nouns.

L50-51: This is not professional writing. Please revise the sentence.

L50-56: These sentences are not properly structured. There should be a smooth transition from one sentence to the next, right now this difficult to see. See also the paragraph from L66-76.

L57: Replace 'these' with 'the following'.

L66: Remove 'tried'.

L77-78: This is not professional writing.

L79: Replace 'using a' with 'using the'.

We thank the reviewer for detailed comments on the writing. We have re-written the manuscript with these in mind. We believe we have improved the presentation by revising the writing.

L50: No satellite is mentioned before this sentence, which ones are being referred here?

This paragraph starts by discussing the 'satellite microwave sensors'. We have re-written the sentence.

L23: Vogel et al., 2013 reference is not complete (see L694-695)

Thank you. We used LaTex v3.6 for the initial submission. We upgraded the revised manuscript to LaTex v4.6 (the latest version provided by Copernicus). It should now show up as complete reference.

L34: Mention the year(s) of flood(s).

Thank you, we added the years.

L35: 'This research' is this specific type of research? The cited references suggest otherwise.

We rephrased this sentence.

Correct the website address in the reference (see L563).

Thank you for pointing this. We fixed it.

L58-62: As flood duration has a probability distribution, annual frequency also has a corresponding distribution. Is there a reason that only one distribution is mentioned?

The metric of annual flood frequency (occurrences) is simply considered here as the total number of flood events that occurred in a specified year at a determined spatial scale. For example, 293 flood events occurred across the global in year 2003 according to the DFO database. Regarding the flood duration, since each flood event has a specific value (ranging from 1 day to n-days), in a given year, a set of values are available. Thus, the metric of flood duration is a probability distribution for durations of flood events. We were interested in exploring whether the probability distribution of the duration each year had a significant change or not. In this regard, we explored the moments estimated for each year from the pdf of the flood durations in that year.

L67: What does "probability distribution of floods' mean?

For each spatial scale, for each year, we first obtain the duration of the floods for all the events. We considered this as data coming from a probability distribution and estimated the trend of the moments of this distribution.

What are the reasons of performing trend analysis (or changes in other flood statistics) at global scale? There is a large body of literature that has studied the changes at local and regional scales, would it be easier to combine all studies (kind of a Review Paper) and see the global and latitudinal changes?

Here is our motivation for the study.

We wanted to understand global floods from the context of the ocean-atmospheric dynamics and socioeconomic factors. It is documented in climate literature that anomalous atmospheric flow patterns (e.g., atmospheric blocking and quasi-stationary Rossby waves) are affecting large spatial scales (i.e. entire extratropics) with similar weather proxies for several days to weeks [Perlwitz et al. 2017]. We wanted to check if the trends in the floods at these scales are related to the climate patterns. Understanding any specific monotonic patterns in global flood data can help in quantifying the flood-related hydrometeorological processes, atmospheric dynamics and climatic drivers. Ultimately, this can provide us with more comprehensive information to quantify the essential causes and drivers of floods' attributes (e.g. flood duration and timing). Besides, the predictions of future planetary flood events would be more robust using the inference based on the atmospheric teleconnections as the drivers widely spaced in distance and/or time. In addition, realization of specific trend and shifts can inform the impact of floods on the global economy through supply chains in order to reduce vulnerability to risks at different scales [e.g., Levermann, 2014; Haraguchi and Lall, 2015].

Regarding the second comment, although it is feasible to combine all the local and regional flood-related studies, one of our motivations here is to have an up-to-date investigation (1985−2015) into all global flood events that were recorded by one data provider/platform.

I understand that a coherent global scale analysis will be useful for disentangling different relationships; however, a strong motivation and justification for doing this is lacking here. Based on this and the above comments, the entire Introduction section needs to a complete revamp.

Thank you for your constructive suggestions. We have improved the manuscript and the presentation of the results.

References for "Authors' Response to RC#1"

Haraguchi, M. and Lall, U., 2015. Flood risks and impacts: A case study of Thailand's floods in 2011 and research questions for supply chain decision making. International Journal of Disaster Risk Reduction, 14, pp.256-272.

Levermann, A., 2014. Make supply chains climate-smart. Nature, 506(7486), p.27.

Perlwitz, K., T. Knutson, and J.P. Kossin, 2017: Large-scale circulation and climate variability. In: Climate Science Special Report: A Sustained Assessment Activity of the U.S. Global Change Research Program [Wuebbles, D.J., D.W. Fahey, K.A. Hibbard, D.J. Dokken, B.C. Stewart, and T.K. Maycock (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, pp. 228-266.

Authors' response to RC#2:

"Recent Trends in Frequency and Duration of Global Floods" by Nasser Najibi and Naresh Devineni

Dear Reviewer

We sincerely thank you for the constructive comments and thorough evaluation. We took all of them into consideration while revising the paper. Our point-by-point responses to your comments are given below.

Review of Najibi and Devineni (2017) in ESD

General Comment

Najibi and Devineni (2017) presents an analysis of trends in global flood frequency and duration between 1985-2015 based on flood data from the Dartmouth Flood Observatory's (DFO) global flood database. Using standard trend and change point analysis they find an increase in both the frequency and duration of floods in many regions (here using latitudinal belts), with a change point centred around the year(s) ~2000. A further step attempts to attribute these changes to large-scale drivers such as ENSO/PWC/GPH using a GLM framework. While the pursuit of understanding how floods are changing at a global scale is indeed a worthwhile topic, I have some serious concerns about the way the paper uses the DFO dataset as a single source of information on flood frequency and duration to go about this. I list a few of my major concerns about the use of this dataset, the methodology applied, and attempt to offer some suggestions to move towards a stronger manuscript. Until these major concerns are addressed I won't provide a full detailed review.

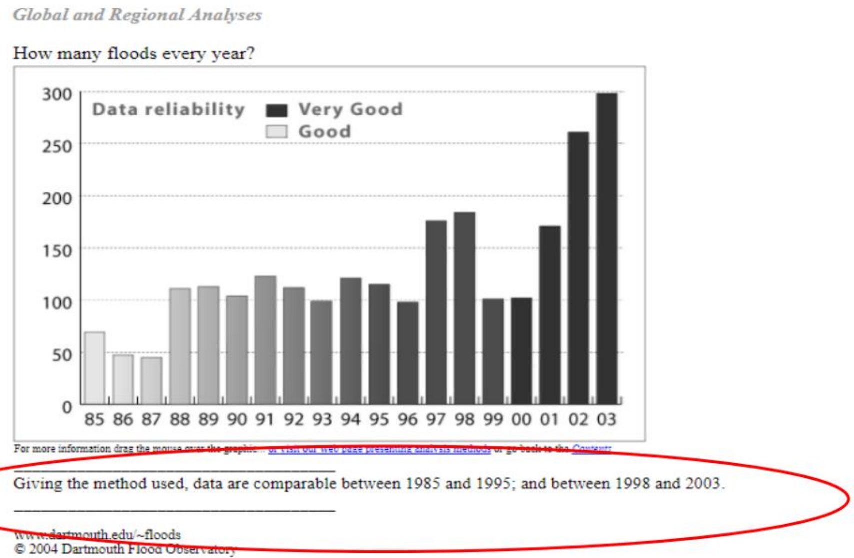
We have improved the presentation of our manuscript based on your comments. Regarding the DFO dataset as a single source of information employed here, we should emphasize that to our knowledge, this is the first analysis of "global flood events" that focuses exclusively on the variability of "flood duration" derived from DFO dataset over the last three decades. The database has recorded flood inundation events using satellite sources and media verification since 1985, and currently has over 4200 entries with approximate location of the center of the area flooded, the dates and duration of flooding and notes as to societal impacts. This is the only global data set of this kind and we believe that an analysis of the trends provides value. Much of the prior studies either focused on rainfall-based datasets or model-based river flow data. In this regard, our study adds a new dimension to the flood literature (especially the understanding of the floods that last for longer time) at a global scale.

Specific Responses:

Specific Comments

1) I am not convinced it is appropriate to statistically test for trends/change points using a database such as the DFO, and then assume any changes are physically driven. I do not have an issue with the DFO archive in general, and in fact believe it is a very useful database of reported floods and impacts, however I don't think you can treat such a database like an observed flood event series derived from e.g. gauged river flow stations. While the authors acknowledge some uncertainties with the dataset (e.g. Section 4.1), they really do not address some key known errors and uncertainties of the use of such flood report databases. Here are a few:

a) How can you know that the increase in floods is not because of an increased level of the media reporting floods/access archives or changing quality of information entering the archive? Here is a graph taken directly from the DFO website (http://www.dartmouth.edu/~floods/archiveatlas/floodnumber85_03.htm) showing the changing reliability of the data in the archive over time. I also highlight in red the sentence "… data are comparable between 1985 and 1995; and between 1998 and 2003" [that analysis was done in 2004] – is there something within the DFO methodology that means these two periods are not comparable?



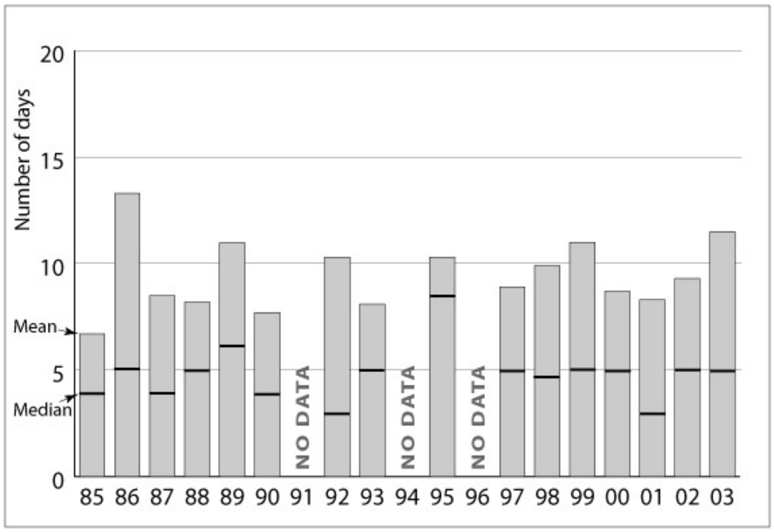
We would like to emphasize that, while we are identifying the potential contributor(s) associated with the observed trends in the floods (as shown in DFO) from the climate and atmospheric patterns, there are unexplained factor(s). Some of these will certainly be technological advances and our ability to monitor floods better. The improvements in the level of media reporting or information quality have increased over the years (likewise the improvements in the quality of the in-situ measurements, advances in the satellite or ground-based sensors, data storage and transfer facilities, etc.). Part of our attempt was to tease out the natural cycle from these unexplained factors. We acknowledge that from a hypothesis standpoint, ascertaining the truth (causal factor) to a trend may be a methodological challenge in itself.

The chart from DFO (figure above) indicates the difference in number of flood events in two separate periods, 1985-1995 with ‘good’ and 1999-2003 with ‘very good’ data reliability. As mentioned before and by Brakenridge et al. 2012, Brakenridge et al 2003 and Brakenridge et al, 2005, the data is recorded after frequent temporal sampling of satellite-based observations cross-calibrated/verified with media reports. In this light, although the data prior to 1995 may be less reliable in comparison, this is still the only database of this kind and worthy of the exploratory data analysis.

A chart provided in DFO website given below (also available online from http://www.dartmouth.edu/~floods/archiveatlas/duration.htm) points out the interannual variability of flood duration since 1985. It also indicates the highest flood median year (between 1985 and 2003) was the year 1995 - consistent with Figure 3 in our manuscript. Their chart is based on data from 1985 to 2003.

Global and Regional Analyses

Interannual Evolution of Flood Duration
(since 1985)



For more information drag the mouse over the graphic... or visit our web page presenting analysis methods or go back to the *Contents*

The average duration (mean) of the floods recorded since 1985 is about 9.5 days; the median duration is 5 days. There is no definite evolution in terms of flood duration since 1985. Among these data, some years seem more interesting: 1986 mean duration is influenced by a couple of long floods - On the other hand 1995 have had a proportionally high number of long flood (high median value), but the data reliability for this year needs to be improved..

---

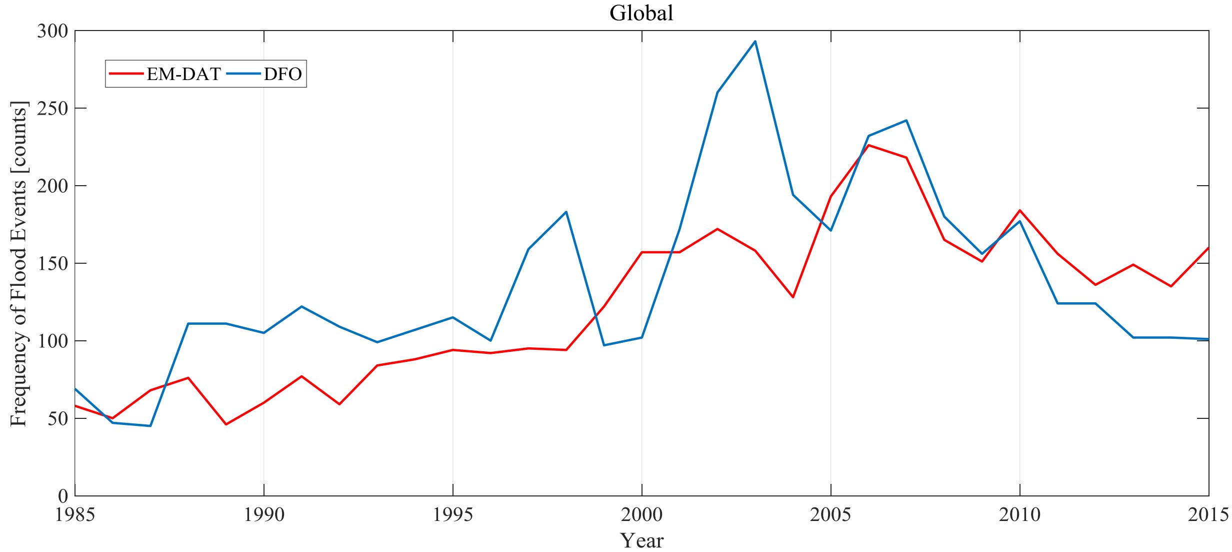
www.dartmouth.edu/~floods

---

b) L88-93: A change in the information DFO uses is known to have changed majorly after 1999, by introduction of MODIS. Surely this would have a large impact on the temporal homogeneity of the frequency and duration of floods and thus could be a likely explanation for why so many change points were detected in the time around the 2000s, and hence the increasing trend?

This may be one possibility. As we mentioned in the paper, we approached this in an exploratory sense and attempted to associate these changes to possible factors. We considered two possibilities that go against the MODIS factor.

First, we corroborated global DFO data with the flood frequency data at global scale from the EM-DAT database (http://www.emdat.be/database) during the same timeframe (1985 to 2015). Below you can find the original EM-DAT flood frequency time-series (which is based on reporting information) and how it compares with the DFO data (which is based on both satellite observations and reporting information). It should be noted that for a disaster to be recorded in the EM-DAT database, at least one of the following criteria must be satisfied: 1) 10 or more people reported killed, 2) 100 or more people reported affected, 3) there was the declaration of a state of emergency, and 4) there was a call for international assistance.



We also see similar trend in EM-DAT data, indicating potential increase in floods due to various causes. It can be also inferred that DFO is collecting more flood information, especially, those events that are occurring in the regions with zero access to reporting facilities.

Second, specifying three flood duration classes and calculating the number of flood events for each class per year, we see that there is no significant trend in the frequency of ‘short duration’ events across all latitudinal scales (Table 7), but a significant trend can be seen across the Tropics for ‘moderate’ and ‘long duration’ flood events. The introduction of improved satellite products would have increases the chances of detecting more short duration floods (small events) along with providing better resolution for longer floods. We think that it is not possible to see the systematic contributions of such products into only one specified type(s) of flood duration. Hence, we continued our investigation to explain the trend from other factors also.

c) DFO Duration data. The start date/end date within the DFO appears to be very uncertain and dependent on news reporting time-scales. This is very different from actual physical flood duration/inundation. There is a well-known issue also with separating multiple-events in close succession within these such databases that tend to aggregate many smaller events (Chapter 3 in Kundzewicz (2012) might be useful here). While I recognise some flood inundations can take weeks or even months to recede, many large floods recede in a much shorter time-frame (within a few days). How likely are the floods reported to last e.g. 21 or longer to be actual inundation of an affected populated area, or an artefact of news reporting etc.? Overall, more careful consideration of what is meant by flood duration and a discussion of uncertainties is needed.

We agree that there may be several uncertainties in the calculation of flood duration (a combination of satellite imagery and reporting information) and we acknowledged this in section 2.1 where we introduce DFO data and in section 4.1 discussion. We also want to point out that the use of DFO's definition of > 21days floods is only done to cross validate the trends seen in the 90th percentile of the annual distribution. For all other purposes, we use the full duration data.

DFO also shows here (http://www.dartmouth.edu/~floods/archiveatlas/floodmonthdurationgraph.htm) that the floods with >21 days make up around 10-15% of the total events. These uncertainties are also ackowdged in the previous studies that use duration data for calculating the flood magnitude (i.e. Log {duration × severity × affected area}) (Kundzewicz et al. [2013], Kundzewicz et al. [2017], and Halgamuge and Nirmalathas [2017]).

d) A record of 31 years, no matter if using observations or DFO data, is very short for a trend analysis. I acknowledge that the spatial availability of observed gauged flow data is limited for many countries, but it would be important to know if trends even in the overlap period, for regions that have data, are similar to those shown for the DFO data. In those cases it would also be useful to know if the pattern seen in the 1985-2015 record is part of a longer-term trend or just short-term climate variability. Even though you use e.g. ENSO within the GLM framework to assess *drivers, other drivers that operate at a longer time-scale could be important (e.g. AMO and floods, see Hodgkins et al., 2017, especially Fig. 10b). You will see from their paper that the period of ~1985-2015 really is on the short side. It is my strong opinion that there needs to be some corroboration with the DFO data and observed flood frequency datasets that could be extracted from e.g. Global Runoff Data Centre (GRDC) records (and for 'duration' metrics if enough data exists?) or at least for a select number of countries with good data in the overlap period. Do floods also increase in a similar manner, is the inter-annual variability/trends the same between different flood datasets over 1985-2015?*

We agree with your comments. 31 years of data are short to make robust conclusions on trends, especially since the trends can also be a manifestation of climate cycles. For this reason, we do offer any extravagant claims about increases, but just report their existence in the seen data and point out the potential cyclical/atmospheric linkages. As per your recommendation, we have now included AMO and PDO data as potential predictors. As Hodgkins et al., 2017 reported, we do find AMO to be a significant predictor in explaining the flood frequency and durations. In most cases, we now find that the trend in the residuals in insignificant, indicating that much of what we observe might actually relate to cyclical variability. Still, given that the dataset is only 31 years, we reserve ourselves from making the causality claims.

While we did not explicitly get the GRDC data to conduct the verification, we would like to point out Do et al. 2017 who have conducted trend analysis of annual maximum flow on some river basins (from GRDC) in USA and Europe. They see both increasing and decreasing trends. A complete comparative analysis is required in this regard, especially to first identify DFO locations with river basins and then analysis the trends in those river basins. We believe that this involves developing a separate study in the future.

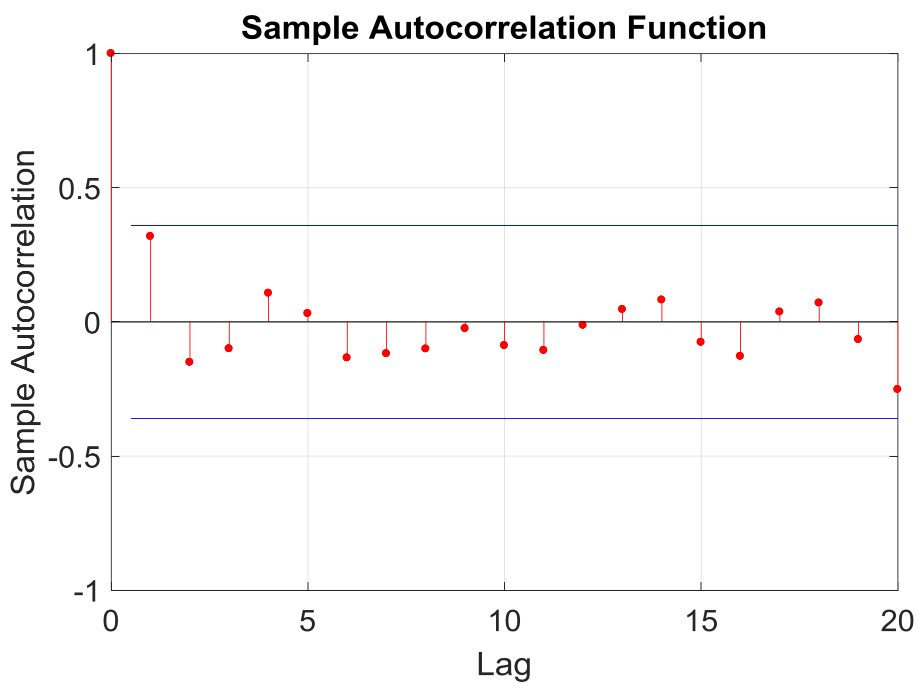
2) A note about the methods. The paper would benefit to have all the methods described within the methods section. E.g. The methods description (and results!) from the GLM analysis is given in the discussion section (L396-439) instead.

Given we are presenting the trend (which is the main focus) and then explaining some factors that could lead to trend, we preferred to keep GLM description when it is introduced. All other methods are moved to "methods" section.

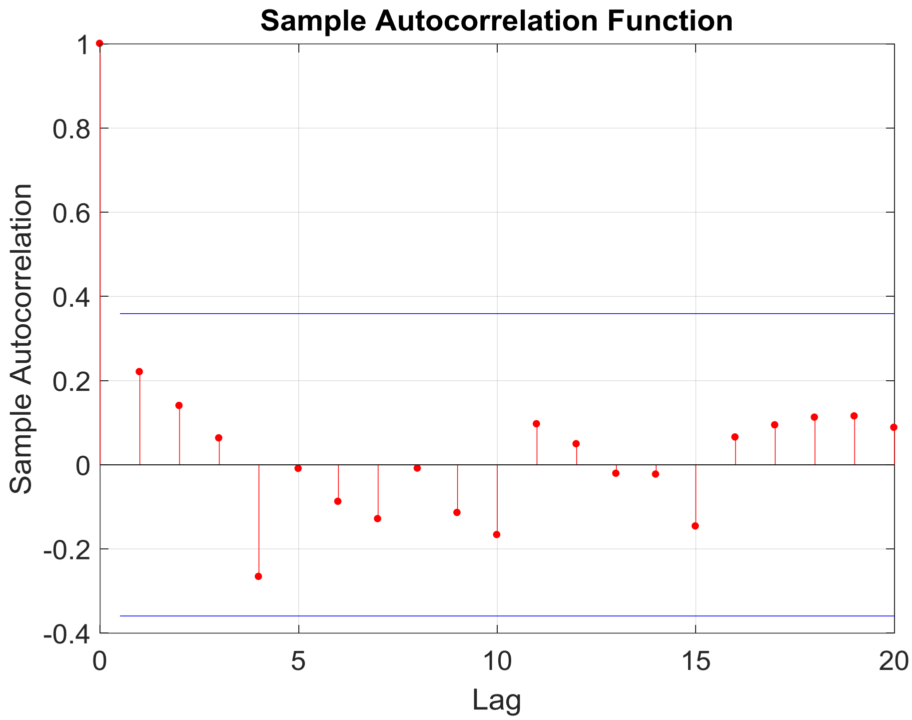
3) While the use of the Mann-Kendall and Pettitt tests are useful for working with environmental data as less assumptions are needed (non-parametric), they do still need to satisfy the assumption of independence in statistical hypothesis testing. So at a minimum the series should be tested for serial correlation and dealt with appropriately if detected (e.g. block bootstrapping perhaps).

We agree. We discussed this in section 4. The time series may have an autocorrelation structure that can manifest as trend in the data. However, in a limited dataset, at the annual scale, influence of exogenous variables (especially if they are persistent, like in the case of AMO) might also manifest as auto-correlated data. Hence, we preferred to first fit the model (GLM) and then analyze the auto-correlation structure in the residuals. Example acf plots from the residuals at the global scale are shown below. There is no auto-correlation in the residuals.

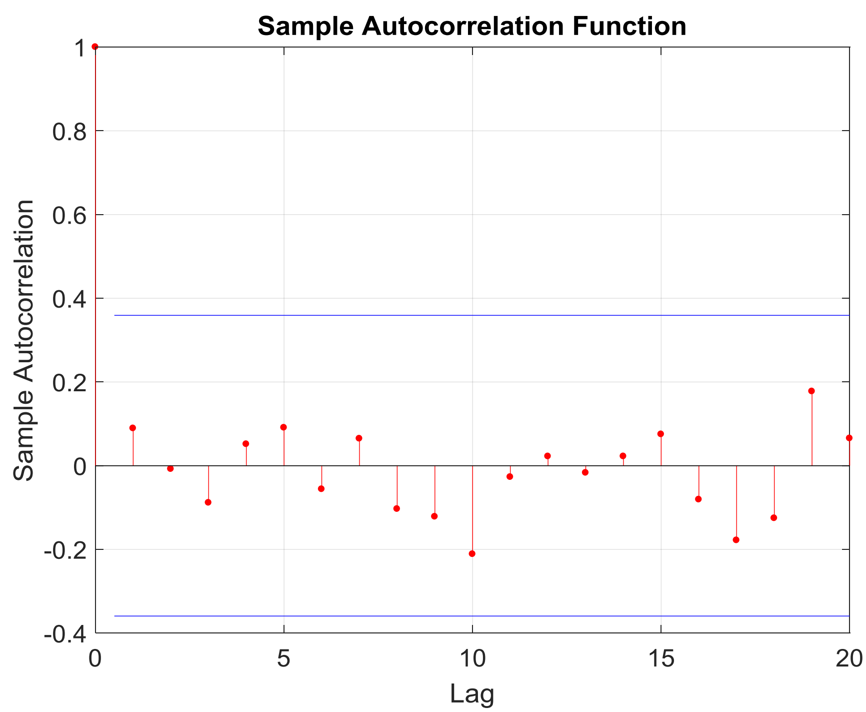
Fc



Fd (median)



Fd (90^{th})



4) Page 24 Fig. 2: The bottom right panel (i.e. Mid-latitudes (S)) have many zeros. This is an issue when using the Mann-Kendall test as it introduces many 'ties' that can impact the variance of the test. Does the Mann-Kendall version you use deal with ties? Or have you thought about using other methods often used for frequency or count data such as logistic regression? If not, perhaps *there are too few data points within the Mid-latitude (S) region, in particular, to use the MannKendall/Pettitt tests reliably.*

Yes, the employed Mann-Kendall version here handles ties. We use the enhanced version of Mann-Kendall test in Matlab (https://www.mathworks.com/matlabcentral/fileexchange/11190-mann-kendall-tau-b-with-sen-s-method--enhanced-?focused=3795005&tab=function) that is written to handle ties in the data.

5) The section from L326-333 does not belong in the results section. Further, the whole text in Sect. 4.3 is not formally being tested so does not deserve a separate section. It is useful to mention throughout the discussion in passing, but to aim of the paper is not to test "Have the exposures of residential/industrial sectors to flood events increased recently?"

Thank you for pointing it out. We removed L326-333 from the results section. We also completely removed section 4.3 and moved this with 4.1, the discussion on DFO data's accuracy.

6) Page 23 Fig. 1: Can the method differentiate floods spatially within a country? For example, if a flood was reported in 'Australia' how does the method determine if the flood occurs within the 'Subtropics (S)' or 'Mid-latitudes (S)' belts?

Based on the reported centroid for the flood event, we assign the event to whichever latitudinal belt it falls under. In this regards, a flood in Australia, depending on the latitude, can be assigned to either Subtropics (S), Mid-latitude (S) or Tropics.

7) A key strength of the paper is the effort to explore the drivers of flood variability and change using large-scale predictors (e.g. ENSO). This does not feature in the original hypotheses tested in L58-65 + Sect. 2.4 + Table 1. In my opinion, if the paper was to re-focus and expand upon this element, and/or strengthen the trend analysis components with other methods and datasets, it would lead to a much stronger analysis.

Thank you for mentioning this. In the revised version, we included this as H4, the hypothesis where we try to connect the trends to possible factors.

I'm happy to give more specific and detailed comments if my above concerns are addressed in a future iteration of the manuscript. I do of course see merit in this work on such an important and challenging topic, but I feel additional effort is needed with regards to corroborating the DFO dataset and strengthening the methods in the first instance.

Thank you very much. We look forward to receiving your more comments.

References

Hodgkins, G. A., Whitfield, P. H., Burn, D. H., Hannaford, J., Renard, B., Stahl, K., Fleig, A. K., Madsen,H., Mediero, L., Korhonen, J., Murphy, C. and Wilson, D.: Climate-driven variability in the occurrence of major floods across North America and Europe, J. Hydrol., 552, 704–717, doi:10.1016/j.jhydrol.2017.07.027, 2017.

Kundzewicz, Z. W.: Changes in Flood Risk in Europe. [online] Available from: http://www.crcpress.com/product/isbn/9780415621892 (Accessed 20 August 2014), 2012.

References for "Authors' Response to RC#2"

Brakenridge, G.R., E. Anderson, S.V. Nghiem, S. Caquard, T.B. Shabaneh. Flood Warnings, Flood Disaster Assessments, and Flood Hazard Reduction: The Roles of Orbital Remote Sensing, 30th International Symposium on Remote Sensing of Environment, Honolulu, HI, November 10-14, 2003. (available online from: https://trs.jpl.nasa.gov/handle/2014/38308).

Brakenridge, G.R., Nghiem, S.V., Anderson, E. and Chien, S., 2005. Space-based measurement of river runoff. EOS, Transactions American Geophysical Union, 86(19), pp.185-188.

Brakenridge, G.R., Cohen, S., Kettner, A.J., De Groeve, T., Nghiem, S.V., Syvitski, J.P. and Fekete, B.M., 2012. Calibration of satellite measurements of river discharge using a global hydrology model. Journal of Hydrology, 475, pp.123-136.

Do, H.X., Westra, S. and Leonard, M., 2017. A global-scale investigation of trends in annual maximum streamflow. Journal of Hydrology, 552, pp. 28-43.

Halgamuge, M.N. and Nirmalathas, T., 2017. Analysis of Large Flood Events: Based on Flood Data During 1985–2016 in Australia and India. International Journal of Disaster Risk Reduction.

Kundzewicz, Z.W., Pinskwar, I., and Brakenridge, G.R., 2013. Large floods in Europe, 1985–2009. Hydrological Sciences Journal, 58 (1), 1–7.

Perlwitz, K., T. Knutson, and J.P. Kossin, 2017: Large-scale circulation and climate variability. In: Climate Science Special Report: A Sustained Assessment Activity of the U.S. Global Change Research Program [Wuebbles, D.J., D.W. Fahey, K.A. Hibbard, D.J. Dokken, B.C. Stewart, and T.K. Maycock (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, pp. 228-266.