Authors' Responses to Reviewers and Editor Decision:
'Recent Trends in Frequency and Duration of Global Floods' by Nasser Najibi and Naresh Devineni

Comments to the Author:

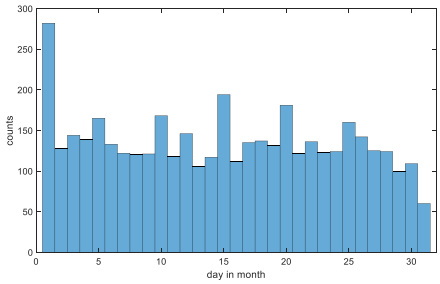
Dear Authors,
Thank you for submitting a revised version of your manuscript, which incorporated most of the referees comments. Following the latest reviews provided by the referees and the referee reports in the previous round of the discussion, it becomes clear that the data source used in this analysis is still a major concern. Particularly the comments provided by the Anonymous Referee #2 raise important points in this regard. For example, if in the data no date is mentioned the middle of the month is used. This certainly affects the calculation of the flood duration and any subsequent analysis of the time series. Both referees recommend that the uncertainties need to be better explained and highlighted and that results obtained should at least be partly validated with additional data sets such as the one from the GRDC. Therefore, to improve the confidence in the analysis, the authors should validate their results using a different data set. Regions for validation should be selected with particular focus on where the data quality in the DFO is perceived to be low. Moreover, please take into account all the other points highlighted by the two referees, as they both raise valid concerns. Additionally, the authors mention several times the changes before and after ~ 2000. This is an interesting finding; however, it also raises again concerns with the data homogeneity/reliability as the authors mention themselves that only consistent information in the data exists particularly since 1999. This fact needs to be further investigated and discussed in detail in the discussion section. The abstract strongly highlights the results about trend results. Given that the DFO data is used (which has known data quality issues and all the concerns raised by the referees) a statement regarding the uncertainties associated with this data needs to be included. Additionally, in the data section 2.1 a more detailed discussion on data quality is needed so that the reader is aware of the potential issues BEFORE the data is analyzed and the results are presented. Section 4.2.1. More detail to the results need to be added by linking the statistical results obtained to the actual physical mechanisms and how these interconnections interact with the floods and why there is the difference in their effects to different indices analyzed.

Dear Editor,

Thank you for handling our manuscript and providing the feedback. We have revised the manuscript based on your comments and the reviewers' suggestions. The revised version of our manuscript has incorporated both the reviewers' comments and the editorial remarks. Our responses to your overall remarks are provided here point by point.

1. *Particularly the comments provided by the Anonymous Referee #2 raise important points in this regard. For example, if in the data no date is mentioned the middle of the month is used. This certainly affects the calculation of the flood duration and any subsequent analysis of the time series.*

We investigated the distribution of the beginning date of the flood and found that less than 5% (194 out of 4311 events) have a middle of the month start date. Even if we consider the 1st of the month and 15th of the month, the number of events are 282+194 = 476 out of 4311 ~11%. As is evident from the adjacent figure, there is a wide distribution of the timing of the flood happening in different days of the month.



2. *Both referees recommend that the uncertainties need to be better explained and highlighted and that results obtained should at least be partly validated with additional data sets such as the one from the GRDC. Therefore, to improve the confidence in the analysis, the authors should validate their results using a different data set. Regions for validation should be selected with particular focus on where the data quality in the DFO is perceived to be low.*

We now compare the DFO dataset with GRDC dataset for locations with matching time periods. Much of the matching locations were in the mid-latitude and sub-tropics north and were concentrated in the USA, Europe and South Africa. We added an appendix to show these comparisons. Uncertainties in this context are also discussed.

3. *Moreover, please take into account all the other points highlighted by the two referees, as they both raise valid concerns.*

We considered and/or addressed all the points raised by the referees.

4. *Additionally, the authors mention several times the changes before and after ~ 2000. This is an interesting finding; however, it also raises again concerns with the data homogeneity/reliability as the authors mention themselves that only consistent information in the data exists particularly since 1999. This fact needs to be further investigated and discussed in detail in the discussion section. Section 4.1. Please discuss the possible effects on the trends obtained and the uncertainties associated with the in-homogeneity of the data pre- and post- inclusion of the MODIS product.*

We added discussion on this in Section 4.1.

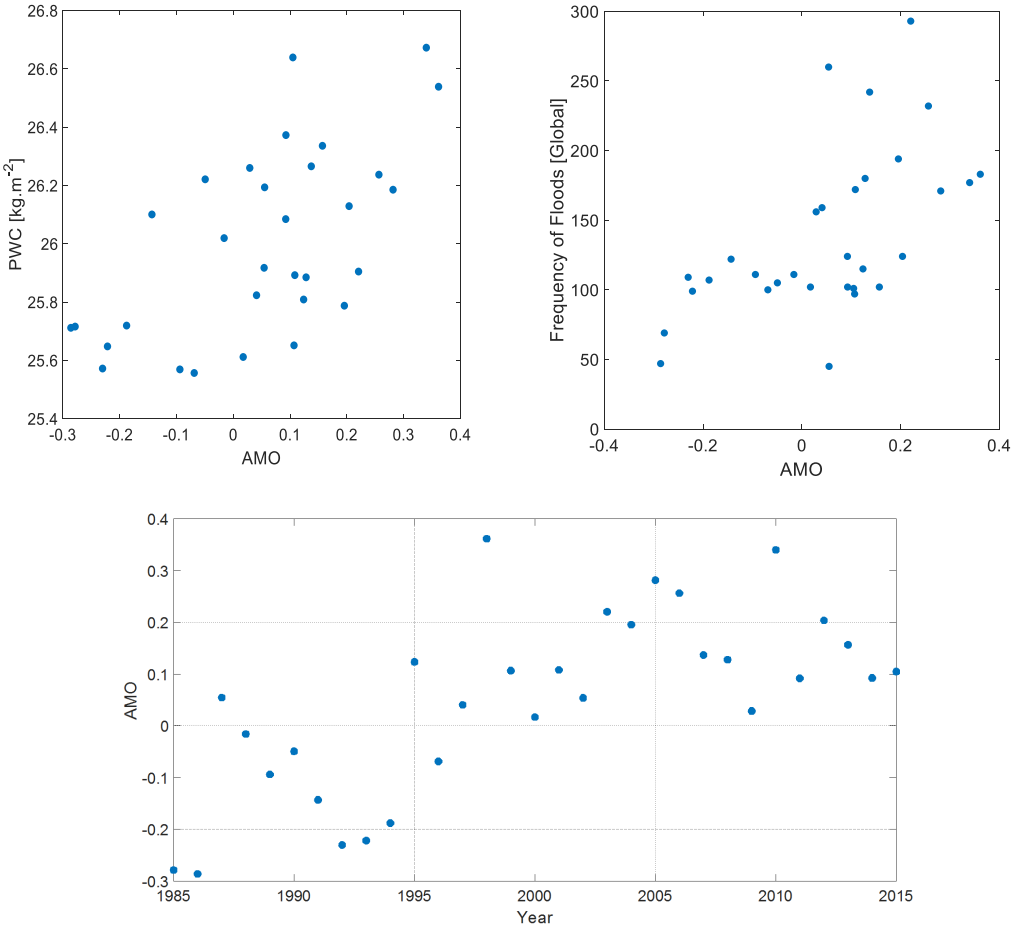
5. *The abstract strongly highlights the results about trend results. Given that the DFO data is used (which has known data quality issues and all the concerns raised by the referees) a statement regarding the uncertainties associated with this data needs to be included. Additionally, in the data section 2.1 a more detailed discussion on data quality is needed so that the reader is aware of the potential issues BEFORE the data is analyzed and the results are presented.*

We updated the abstract and added the discussion about data in both sections 2.1 and 4.1.

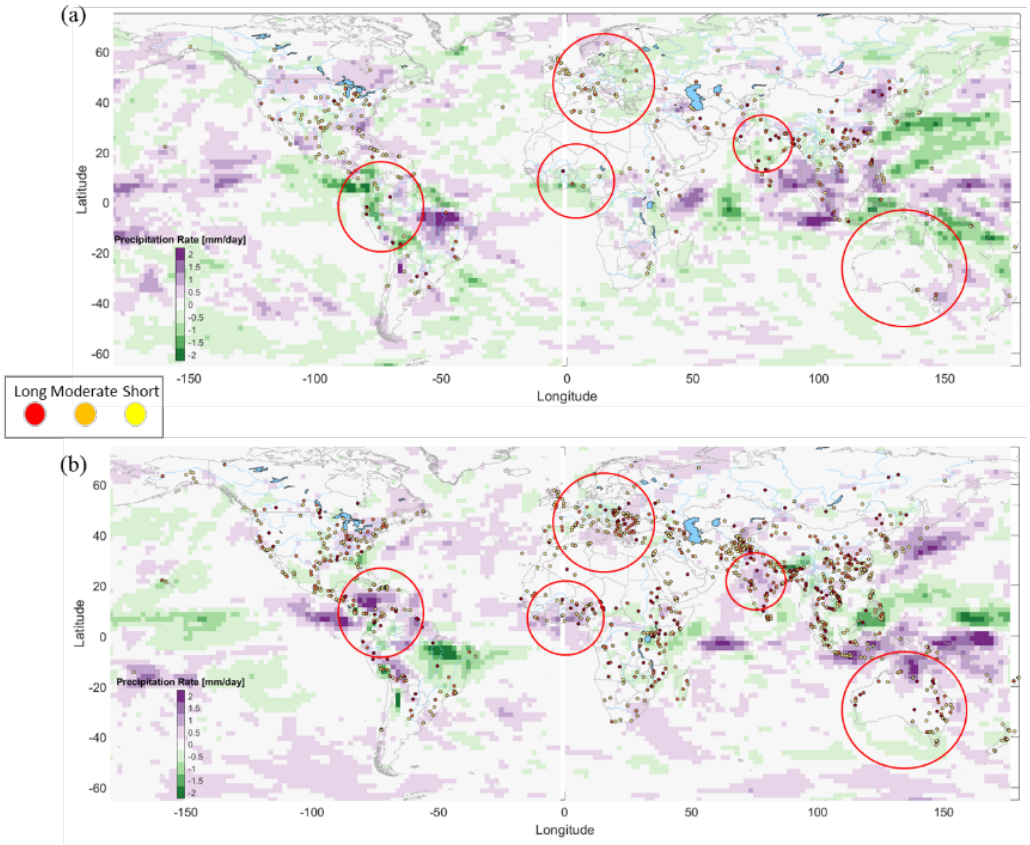
6. *Section 4.2.1. More detail to the results need to be added by linking the statistical results obtained to the actual physical mechanisms and how these interconnections interact with the floods and why there is the difference in their effects to different indices analyzed.*

To interpret the statistical results (e.g., AMO climate index in GLM) using the actual physical mechanisms and causation of the climate/atmospheric variables, we should acknowledge that this matter needs a separate comprehensive study to properly address several existing questions. These may include investigating the preceding hydroclimatological conditions including the antecedent moisture content (in soil and atmosphere), preceding rainfall events/intensities, and regional geomorphological conditions such as the river networks. However, we feel that such comprehensive study cannot be presented in this manuscript given the length and scope of this paper. For your reference, we conducted some preliminary investigation and presented it here.

First we assessed the bivariate correlation of precipitable water content (PWC) (left) and frequency of floods (right) with the AMO index at the global scale. The Pearson correlation yields a coefficient of r=0.59 and 0.62 respectively with p-value=0.0005 and 0.0002 (it is significant at 5% significance level).



Second, we derived the composite maps of anomalous precipitation rates for two sets of flood events corresponding to two AMO phases (i.e., AMO<-0.2 and AMO>+0.2). In DFO database, there are 324 flood events with AMO index < -0.2 (in the years of 1985, 1986, 1992, 1993) (a), and 1056 flood events with AMO index > +0.2 (1998, 2003, 2005, 2006, 2010) (b). The figure below indicates this comparison. In addition, there are 576 events with positive precipitation rates/anomalies (the nearest grid value) when AMO>+0.2 compared to 190 events in AMO<-0.2 phase, clearly showing a significant influence of positive phase of AMO on enhancing the available moisture contents and precipitation rates at the global scale. The hollow red-colored circles point out the corresponding comparison of flood events in a specific location (e.g., Europe, Australia, West Africa, India, and Central America).



Additional comments:

P1L14-15: Add reference to this statement.
Done.

P2L 22: Unclear. Please elaborate more how 'temporal trend' and 'regime like behavior' are related.
We added more precise wordings in order to make these clear.

P3L6: Add sentence on the country scale analysis.
Done.

P4L15: Add already a sentence in this section why these 4 countries were selected.
Done.

Section 3.2: Please add some elaboration to each sub section what these changes mean for the floods in a physical way. E.g. what does it mean when 'the asymmetrical/symmetrical behavior of the distribution … changes from 5 to ~8.
Thank you. We added some explanations to these sections in order to link the statistics to physical interpretations.

Section 3.4: Much detail on the methods presented in this section should be moved to the methods section.
We moved this as Section 2.5.

P11L19-20: Please provide more detail on this data on damages (e.g. how was it derived in the original data set).
We added to the revised manuscript, the information provided by DFO relating to the flood damage.

Section 3.4: Please add a paragraph summarizing the most important findings from this section. Currently this is not clear.
Added.

Figure 8a) The figure is difficult to read due to the thin line width of the gray points. Please increase the quality of the figure.
We enlarged its dimension in the revised manuscript (the fitted lines here are of interest and discussion). Also, we improved the presentation of bar-plots in the Figure 8c using the log-scale approach on the horizontal axis.

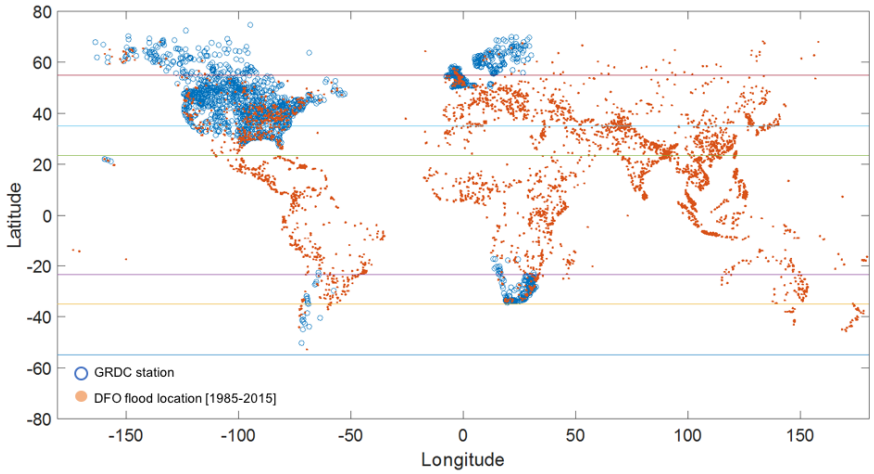
Anonymous Referee #2

I have re-reviewed the paper by Najibi and Devineni (2018) for ESD. My original major concern was the use of the DFO dataset as a single source of information on flood frequency and duration in which the conclusions of the paper are founded. I acknowledge that the authors have answered most of my original comments to a satisfactory degree, and indeed made many positive changes to the manuscript. This includes addressing issues around assumptions for statistical testing and including more information about the limitations of the DFO dataset.

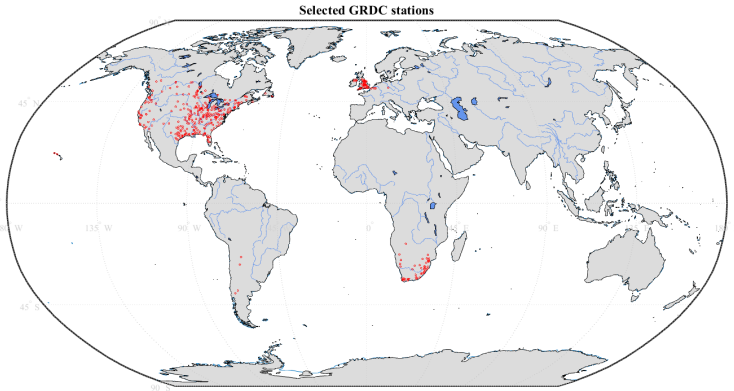
However, after reading the revised manuscript the same concerns about the sole use of the DFO dataset still remain. Here is the author response to this concern: "Regarding the DFO dataset as a single source of information employed here, we should emphasize that to our knowledge, this is the first analysis 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".

My view is that just because this study is the first to analyse an event based dataset at the global scale is not the only criteria for scientific publication. The conclusions made are profound and could be misleading if they are not supported by other event-based and/or physical flood datasets. For example, there are five key conclusions outlined on Pg. 16 in the revised manuscript: "The frequency of flood events has increased"; "there is a statistically significant trend in the moments of flood duration at the global scale"; "The yearly number of moderate and long duration flood occurrences increased"; "there was no monotonic trend observed in the frequencies of short duration floods", and "the increase in frequency of long duration floods during recent years can be related to the persistent patterns in the low-frequency climate induces". There is a chance that these conclusions could be an artefact of the known changing quality and/or changing sources of information ingested into the DFO dataset since 1985, and not to mention the relatively short period of record used for a trend analysis (without trying to place shorter term patterns into their longer term context).

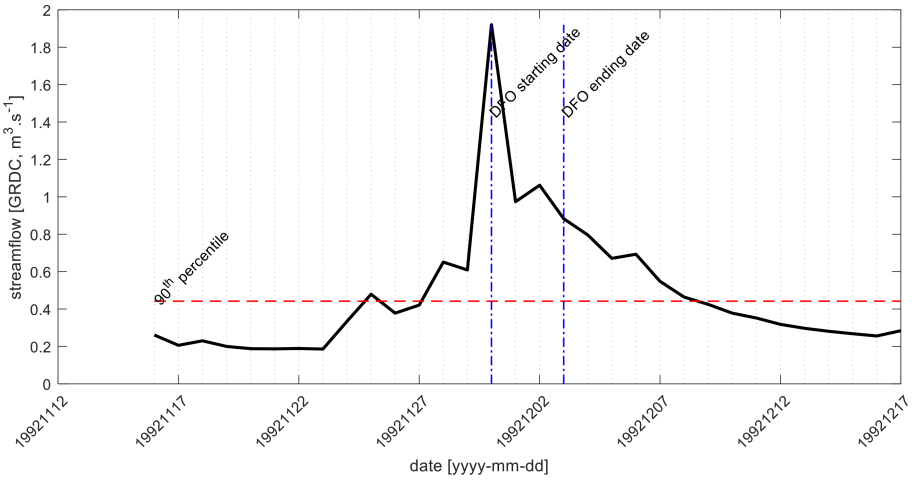
We now verify the DFO flood database with GRDC flood database. To do so, first we took into account those GRDC stations with available data (at least) between 1985 to 2015. The spatial distribution of these stations is shown below:



Second, within this set of GRDC stations (i.e., 1849 stations), we used the streamflow time-series of those stations which are located within one-degree radial distance of the DFO flood locations. This resulted in 517 stations from the GRDC database that we used to verify the DFO's flood statistics. The geographical distribution of these stations is as follows:

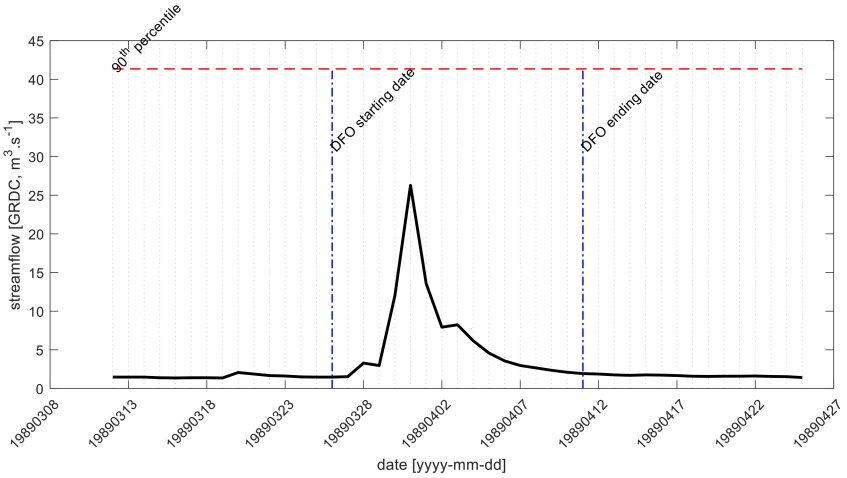


Third, as we discussed in details in Appendix B, the corresponding flood duration to DFO's duration is calculated as following based on the GRDC discharge observation:



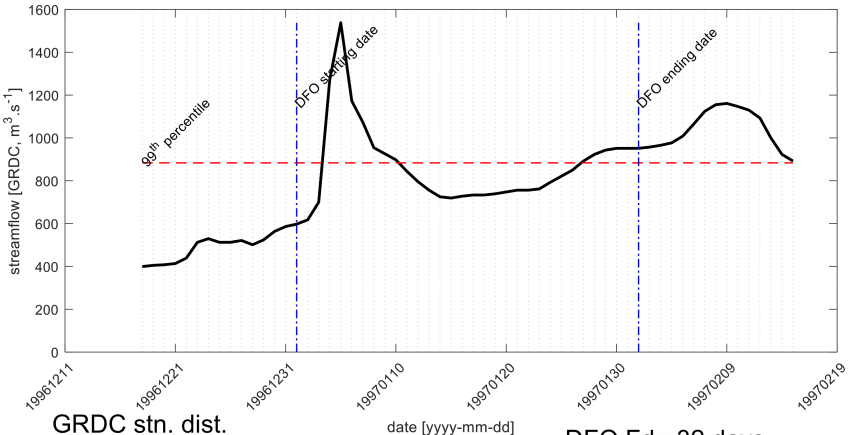
GRDC stn. dist.
0.39 km

DFO Fd= 4 days
GRDC Fd= 4 days



GRDC stn. dist.
0.35 km

DFO Fd= 17 days
GRDC Fd= 0 days

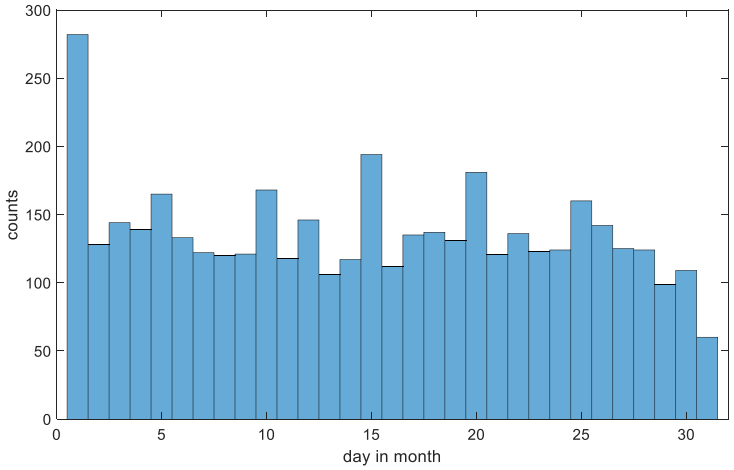


GRDC stn. dist.
0.62 km

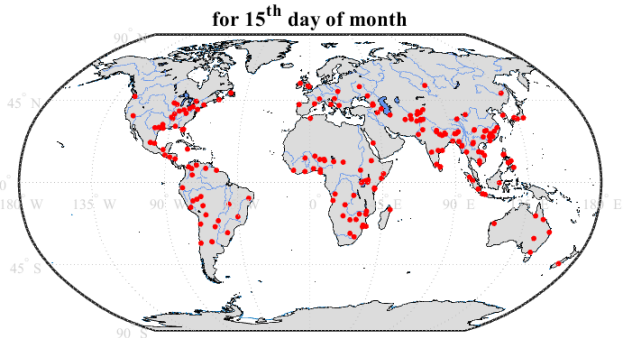
DFO Fd= 32 days
GRDC Fd= 13 days

We find that much of the events have a relatively small deviation in terms of flood manifestation. Further, we added the discussion of comparison with EM-DAT into the appendix along with GRDC. We also revised the conclusions by acknowledging how much of the trend can be placed within the long-term climate variability context. Trend in the frequency and the extreme duration can in fact be a result of decadal climate variability, and we state that now. Trends in the median flood duration however still remain unexplained.

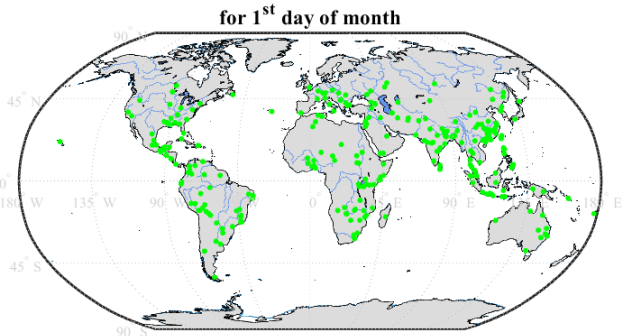
Of particular concern is how the 'flood duration' variable in the DFO dataset relates to physical flood inundation extent/duration of inundation, or just an artefact of news reporting. According to the DFO website (http://floodobservatory.colorado.edu/Archives/ArchiveNotes.html), flood duration is established from the reported flood start and end dates, and from this website: "Occasionally there is no specific beginning date mentioned in news reports, only a month; in that case the DFO date will be the middle of that month. Ending dates are often harder to determine - sometimes the news will note when the floods start to recede. We make an estimate based on a qualitative judgement concerning the flood event". To what degree of certainty does the flood duration variable used in the DFO, and hence a key aspect of the paper, relate to physical flood duration? Is this qualitative uncertainty mentioned on the DFO website only for a small number of floods in some lower income countries, or is the use of such qualitative estimates of flood duration common within the DFO dataset as a whole?



We investigated the DFO reported flood events from 1985 to 2015 in terms of the distribution of the "flood beginning date" within each month. It turned out less than 5% (194 events) out of 4311 events have been reported with the flood beginning date as the middle of the month. The following figure presents the distribution of flood beginning dates within each month at global scale from 1985 to 2015.



Even if we consider the first of the month as a potential judgment based estimate, these total to 282, making it a total of 476 out of 4311 events (~11%). The sample is small to have skewed the overall results. The following maps are the locations of these 194 and 282 events over the 31 years. It appears that they are randomly occurring across the globe and the timing distribution



also indicates a uniform spread across the month. We mentioned these in the data section.

I acknowledge that the authors corroborate flood frequency from DFO with the EM-DAT dataset in their response to my original review, why did this not feature in the main analysis in the revised manuscript/or at least supplementary information? Are both datasets using similar sources of news reporting? Does the temporal pattern of flood frequency match physical flood frequency as observed with river flow gauges over the same time period, notwithstanding the known issue of low density gauges in many regions?

Thank you for pointing this out. We added this comparison as supplementary material in Appendix B, which also has a new companion of DFO data with GRDC data for matching locations and time periods.

Overall, I do not feel confident accepting these conclusions based only on the DFO dataset. This is unfortunate as I think the authors have performed a nice analysis with very interesting and extremely worthwhile hypotheses, but at this time, in my opinion, the underlying dataset by itself does not provide a strong enough foundation for the application of these hypotheses and hence conclusions. A possible suggestion for moving this study forward would be to perform a validation of reported flood frequency and duration based on in-situ observations, at least for regions with available overlapping data (as suggested in my original review). I understand this will be more straightforward for flood frequency than for flood duration (as flood inundation is not directly measured at a flow gauging station), but the only other option would be for a consistent satellite product to be available across the full 1985-2015 study period, which I do not believe exists(?). Nonetheless, some effort towards increased corroboration and/or validation even for a range of case study regions/basins would provide a path towards supporting the conclusions. I hope the authors understand my concern and can find the time to provide more validation.

Thank you so much for these constructive comments. We have added a new section to the supplementary part as Appendix B in which we present the validation of reported flood information by DFO with respect to the available in-situ streamflow observations in GRDC and previous EM-DAT comparison. We have tried to incorporate and address the comments here to our best ability.

Other comments;

Pg 3, Line 6: There has been several authors (e.g. Merz et al., 2012) who highlight the lack of scientific rigour in attribution studies in hydrology. The terms "attribute/attributed" should be avoided if not performing a rigorous formal attribution framework.

We considered the indicated reference (for section 2.7) and modified this sentence accordingly in the revised version.

Pg 4, Line 2 and 3: Need to mention that 'flood duration' within DFO is simply calculated from the flood beginning and end date, and need to expand on to what degree the start and end date is from news reporting and/or from satellite images, and over which time periods (i.e. MODIS only started in 1999)?

We now add that definition and the description of the fraction reported based on middle of the month judgment.

Pg 7, Lines 26-29 & pg 8 Lines 1-7: This is methodological detail about trend tests and would be best placed in section 2.

Thank you for your suggestion. We moved this part to Section 2.6.

Pg 10, Lines 2-3: Delete/rephrase sentence on "found an abrupt shift" as no longer assessing change points.

Done.

Pg 11, Line 24: Should "30 and 40%" not be "20 and 30%" if my interpretation of Fig. 7 is correct?

Thank you for pointing this. We have modified these numbers in the revised manuscript accordingly.

References

Merz, B., Vorogushyn, S., Uhlemann, S., Delgado, J., and Hundecha, Y.: HESS Opinions "More efforts and scientific rigour are needed to attribute trends in flood time series", Hydrol. Earth Syst. Sci., 16, 1379-1387, https://doi.org/10.5194/hess-16-1379-2012, 2012.

Anonymous Referee #3

I read the original and revised versions of esd-2017-59, Recent Trends in Frequency and Duration of Global Floods by Nasser Najibi and Naresh Devineni. I also looked at the previous reviews of the manuscript. No doubt that the paper is improved. I believe the paper is now well framed and well written. In addition, reordering the paper really helps following up the arguments throughout the paper.

While I find the paper technically tight, I am not really convinced if DFO is the best source of data for performing a global flood analysis when GRDC is available. The uncertainty in data is therefore should be highlighted better and it should be mentioned that this analysis should be validated using ground-based data at least in some representative regions. Having said that, DFO can provide a valuable information with which the flood events and corresponding exposure can be studied together and therefore can provide a basis for inter- and/or cross-disciplinary studies. As a result, there is a need to better highlight the pros-and-cons of the DFO data in comparison with alternative data such as GRDC in conducting such a global analysis.

I believe after this minor revision, the paper would be ready for publication, which makes a good read for the community. Good luck!

Dear Reviewer,

Thank you so much for evaluating our paper. We have added to the manuscript, the analysis that compares DFO data with GRDC ground-based measurement data for matching locations and time periods. We have added it as Appendix B in the revised manuscript.

Regards, -The Authors