Response to Reviewer 3

We would like to thank all reviewers and the editor for taking the time to assess the manuscript. We will address each comment in the revised manuscript, as according to our responses below.

Referee comments are highlighted in bold with our response in normal font.

This paper investigates whether the differences between positive/negative and neutral phases of various climate indices in the tropical Pacific and Atlantic Oceans affect flood characteristics in the Amazon basin for the period 1979-2015. This statistical study do not imply much explanation (which is not the purpose of the work). It is an original work as it addresses not only flood magnitude but also the timing and the duration of the floods. It is interesting as it covers the whole Amazon basin. River data are both observed and simulated (GloFAS 2.1), the indices are the usual one, but attention is paid to the differences between the impact of Central and Eastern Pacific events. For these reasons, this paper is interesting but it is not very original in its form and methods.

We first would like to thank the reviewer for taking the time to review the paper in detail and for providing valuable feedback which will improve the paper. We would like to highlight that all of the comments raised, including the major ones, are all addressable without any significant undertaking of any new analysis or methods. The only major point raised which we do not agree with is that this study is "not very original study in its form and methods". This is because, as the reviewer has highlighted above, it is original as it addresses not only flood magnitude, but also the timing and duration of floods, in which analysis is limited for the Amazon basin.

The importance of how climate variability can affect flood timing has already been shown by Ficchí and Stephens (2019) for Africa, whilst Langill and Abizaid (2019) have highlighted the importance of the timing and duration of floods for different types of flood events in the Peruvian Amazon. In addition, Towner et al. (2020) concluded the need to look at other indictors (e.g., variations of ENSO, MEI, CP, and EP) to account for possible variations in results which we have applied across the entire basin, with many studies focusing on less indices, consider only a specific region of the basin, and or for one particular flood year.

Ficchì, A., & Stephens, L. (2019). Climate variability alters flood timing across Africa. *Geophysical Research Letters*, *46*(15), 8809-8819.

Langill, J. C., & Abizaid, C. (2019). What is a bad flood? Local perspectives of extreme floods in the Peruvian Amazon. *Ambio*, 1-14.

Towner, J., Cloke, H. L., Lavado, W., Santini, W., Bazo, J., Coughlan de Perez, E., & Stephens, E. M. (2020). Attribution of Amazon floods to modes of climate variability: A review. *Meteorological Applications*, *27*(5), e1949.

In addition, it suffers some imperfections that are listed underneath. The major ones are the following:

- You do not define precisely what the flood duration is

We intend to clarify our definition further by saying:

For each gauging station/grid point we consider the duration of flooding as the number of days (not consecutive) spent above the 95th percentile of the climatology. On page 7, line 22.

- You do not pay attention to the significance of the correlations. This cannot be accepted.

In all plots the significance level is clearly denoted, but we agree with the reviewer that it would be useful to include more discussion on the significance of correlation results within the main text. We intend to include more descriptive statements in terms of the number or percentage of stations for both datasets that are significant. We do feel it is important to note that focusing on the P value, whilst very important, is only one way of looking at things. If the P value does not reach an arbitrary threshold, but a whole region is presenting the same behaviour then this is important, likewise if the P value does reach that threshold but agreement in that region is poor we might think that it is just a reflection of the measure being used.

- You do not always comment all your results. See various remarks on this topic, below.

We agree with the reviewer that more of the results could be brought out into the results and discussion section. The reason not every single result is highlighted is purely to prevent the paper from becoming too long. We intend to bring out more of the results highlighted by the reviewer, particularly surrounding the results of GloFAS 2.1.

- in section 3.4.1, you comment three topics. Two are thematic; one is related to the data. How? Why did you chose these topics?

The reasons why we choose to focus on these particular topics are as follows:

Response in the Peruvian Amazon – we chose to focus on this region as extreme floods there have been consistently linked to cooler SSTs in the tropical Pacific (i.e. La Niña) (e.g. see Espinoza et al., 2013). However, we do not find this in our results and so wanted to explore why this might be further.

Response in the north-eastern Amazon – we chose to focus on this region as the results for both datasets stand out for the negative phase of ENSO, with a significant number of stations showing increased flood magnitude and duration. However, this was the region where the results differed particularly between the central and eastern Pacific indices and we wanted to explore why this might be by examining rainfall data.

Observations vs GloFAS 2.1 – we feel this section would be particularly beneficial to readers interested more on the hydrological modelling side and could help us understand why certain periods of modelled data perform better than others relative to the observed dataset.

We intend to make this clearer at the start of Sect. 3.4, including statements on the benefits of comparing the modelled data against the observation in more detail.

Espinoza, J. C., Ronchail, J., Frappart, F., Lavado, W., Santini, W., & Guyot, J. L. (2013). The major floods in the Amazonas River and tributaries (Western Amazon basin) during the 1970–2012 period: A focus on the 2012 flood. *Journal of Hydrometeorology*, *14*(3), 1000-1008.

Detailed comments:

Page 2, line 17: ". . .the largest rainfall anomaly" Largest in mm? in length ?

This is in terms of accumulated rainfall anomaly in mm. We will include this on Page 2, line 17.

Page 2, line 26: the reference Bazo 2018 is a web page where nothing is told about the duration of the inundation. Please give a more precise reference or avoid it.

We will remove this reference from the paper.

Page 3, line 3: Apart from Tomasella, see also https://halshs.archives-ouvertes.fr/hal02987838/ on the topic of the timing of peak river flows.

We thank the reviewer for sending the link to this interesting paper. We intend to cite this paper also when referring to the magnitude of the travelling flood wave from coinciding tributaries, page 3, line 3.

Page 3, line 19: the prediction of the end of the wet season could also be useful for population. See https://www.sciencedirect.com/science/article/pii/S2214581817303543

We intend to cite this paper when discussing the how the timing of the wet season could provide skill in determining seasonal rainfall totals (see page 3, line 19).

Page 6, lines 3-4: why do you justify not using the MJO ? While you do not justify not using for instance the PDO, that combining with ENSO can promote extreme events (see the authors you mention in Towner et al 2020).

We felt it was important to give an example of why certain indices which we have described in a previous paper (i.e. Towner et a., 2020) that affect Amazon flooding were not used (i.e. because they do not fit into the tercile methodology). We plan to extend the reasoning to state why the Pacific Decadal Oscillation and the Atlantic Multidecadal Oscillation were also excluded to page 6, line 3.

Towner, J., Cloke, H. L., Lavado, W., Santini, W., Bazo, J., Coughlan de Perez, E., & Stephens, E. M. (2020). Attribution of Amazon floods to modes of climate variability: A review. *Meteorological Applications*, *27*(5), e1949.

Page 7, section 2.7: Are the days above the 95th always consecutive during an annual cycle? The answer is no. There can be a sequence of days above 95th, then a few days below and again days above 95th. How did you determine the duration of the flood? Did you count the days above 95th? Did you include in this count the intermediate days? Do the duration of the flood and the date of the beginning of the flood allow to know the date of the end of the flood? Or not?

We did not state anywhere in the manuscript that the number of days above the 95th percentile had to be consecutive. However, we will add in a statement as previously highlighted to remove any uncertainty and state that the days do not need to be consecutive. Only days above the 95th percentile are counted.

Page 8, line 16: what is the relationship between the fact that the annual maximum of GloFAS 2.1 occurs earlier than observations and dams? Explain, please. What about the role of the large floodplains in the Upper Madeira basin (Llanos de Mojos) ?

This is because dams can affect the navigation and timing of water passing through and GloFAS 2.1 does not currently include all dams into the model. Thus, in real life (i.e. observations) the peak at the gauge downstream of a dam might be later than it otherwise would and GloFAS 2.1 would not account for this delay. The role of large flood plains is an interesting point and could also affect not only the timing of the annual maximum but also the duration of flooding. In GloFAS, river-floodplain simulation works by incorporating a simple loss function to mimic river-floodplain interaction, however, this is considered one of the main sources of errors within the model and could lead to

earlier peaks than in the observed data (Alfieri et al., 2013). We intend to add statements surrounding these uncertainties when explaining the results in Sect. 3.

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013). GloFAS–global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, *17*(3), 1161-1175.

Page 8, lines 19-22: Neighboring stations are also subject to ITCZ migration. Why would only stations 45 and 51 behave differently?

This is a good point raised by the reviewer. We plan on removing this statement.

Page 8, 3.1.1: You show on the figures the significant values, but do not say in the text how many (or what %) of them are significant. This is essential, as non-significant values are not very interesting! This remark is true for all the indices.

We plan to add comments regarding the percentage of stations that are significant for all indices and flood characteristics, though non-significant values, where regional patterns emerge we believe have value as they describe a large location where you could then extend analysis to look at the atmospheric circulation in these regions.

Page 9, line 17: Northwestern or northeastern?

Northwestern, here we are referring to the Peruvian Amazon region that have previously seen floods associated with La Niña conditions (e.g. Espinoza et al., 2013). We do find this pattern also in the northeastern Amazon but here we are not referring to this.

Espinoza, J. C., Ronchail, J., Frappart, F., Lavado, W., Santini, W., & Guyot, J. L. (2013). The major floods in the Amazonas River and tributaries (Western Amazon basin) during the 1970–2012 period: A focus on the 2012 flood. *Journal of Hydrometeorology*, *14*(3), 1000-1008.

Page 10, lines 1-10: are these signals consistent with what is found for rainfall?

Yes, the same wet signal in the north-eastern Amazon is identified in the composite rainfall plots (Figure 12) for EN3.4 and the CP index but not for the EP. We intend to state this in the paragraph.

Pages 8-10: Why the signals are stronger with GloFAS 2.1? Why don't you comment that in the text?

In some cases, the signals are stronger as GloFAS 2.1 fails to pick up on smaller scale processes and does not represent the floodplain interactions well. This means the signal is likely due from the increase or decrease of rainfall associated with climate variability without any smaller scale interactions involved. We intend to insert comments on these factors in the results, though the results are not always stronger in GloFAS as can be seen in Figs. 3c, 4c, and 9a for instance.

Page 10, lines 30-31, last sentence: this is also true for TNA in Western Amazon.

We will state that both results found in the western Amazon and along the Madeira River are replicated in GloFAS 2.1 for the positive phase of TNA.

Page 11, line 21: not only in the Jurua, but also in the Madeira, Purus, Negro

We will state that early peaks are also found in the Madeira, Purus, and Negro rivers.

Page 12, line 2: You could explain that the extra days in Obidos may be related to the extra days along the Negro ad Branco rivers.

We thank the reviewer for raising this point. It is worth noting that the results identified at Óbidos could be related to increased flood duration from northern tributaries such as the Negro and Branco rivers and we plan to add statements to discuss this.

Page 12, line 4: What is DOT ?

DOT stands for days over threshold. We will define this.

Page 12, section 3.3.1: Figure 9h shows some consistent positive anomalies along the Purus and the Madeira. You do not comment them. Why?

We will comment on this in Sect 3.3. The results are similar to what is found in the observations, although statistically significant.

Section 3.4: why do you focus on these 3 topics. You should explain that in the introduction of this section.

We will include an explanation in the introduction to this section as mentioned in reply to the major comment.

Section 3.4.1: you do not mention that Espinoza et al 2013 tell that the intensity of floods is more likely related to an early La Nina event, as observed during the 2011– 12, early rainfall and simultaneous peaks of both tributaries of the Amazonas Rivers.

We will add this information when discussing the response to SSTs in the Peruvian Amazon (Sect. 3.4.1). This is an interesting point to discuss and it could be that the timing (i.e. onset and ending) of a La Niña event is more important with regards to flooding in this region than whether there is a La Niña event or not.

You should consider this in yours reflexions.

Page 13, line 24: show these years in figure 11.

We will add text denoting the years 1998, 1999 and 2007 in Fig. 11.

Page 14-15, section 3.4.2: why the differences between EP and CP? Rapidly. This has been commented in the literature (see your review paper).

The differences found in the results between the EP and CP indices are described in Sect. 3.4.2 for the north-eastern Amazon. In negative EN3.4 and CP years a positive rainfall anomaly is identified in the north-east, whereas a deficit in rainfall is identified for negative EP years (see Fig. 12).

Page 15, line 20: again, tell whether the correlation are significant or not.

We will add a statement regarding the median number of significant values between the observed data and GloFAS 2.1.

Page 17, lines 10-13: Can combining indices be a perspective?

This is an interesting point which we want to include into the conclusions when discussing future work. Previous studies have found links between flood events in the Amazon and simultaneous SST anomalies in different ocean basins. For instance, the 2012 Peruvian Amazon flood has been linked to both La Niña and a warm TSA ocean (Espinoza et al., 2013), while the 2014 flood in the Madeira

basin has been associated with warm conditions in the western Pacific-Indian Ocean and exceptionally warm SST conditions in the sub-tropical south Atlantic (Espinoza et al., 2014). Therefore, it would be important to consider a similar analysis in which the combined influence of indices are considered.

Espinoza, J. C., Ronchail, J., Frappart, F., Lavado, W., Santini, W., & Guyot, J. L. (2013). The major floods in the Amazonas River and tributaries (Western Amazon basin) during the 1970–2012 period: A focus on the 2012 flood. *Journal of Hydrometeorology*, *14*(3), 1000-1008.

Espinoza, J. C., Marengo, J. A., Ronchail, J., Carpio, J. M., Flores, L. N., & Guyot, J. L. (2014). The extreme 2014 flood in south-western Amazon basin: the role of tropical-subtropical South Atlantic SST gradient. *Environmental Research Letters*, *9*(12), 124007.

Tables and figures

Table 1: which are the significant values?

We will highlight these in bold font in the revised manuscript.

Figure 1:

1) **Are "mean annual maximum river flows" monthly values? Extreme annual values? Tell it in the caption.**

   Mean annual maximum river flows take the average of all the yearly maximum values within the analysis period. This is already described in the title above the plots (see Fig. 1a, d) and in the figure caption.

2) **Give more information in the caption about the index that is represented in 1c and 1f. Are all the figure essential in the main text? Some figures could be supplementary.**

   We will define the interannual variability coefficient, r, more clearly by stating that it refers to how consistent the date of the annual maximum is over time (i.e. values closer to 1 have a consistent flood regime as described on page 7, line 14.

Figure 1S: Fazenda Vista Alegre is not at the confluence of the Madeira and Solimões. Check it.

We thank the reviewer for highlighting this mistake. Apologies for this. The coordinates used to map this station were incorrect. We will correct this as it should be further upstream along the Madeira River.

Table S1: List the stations in number order and not alphabetical order. It will be easier to find the stations mentioned by their numbers in the text.

We thank the reviewer for this suggestion. We plan to change the layout of Table S1 to order stations by gauging number.