Reply to RC3

We thank the referee for his/her review and constructive comments. Original review comments are shown in **black** while our replies are provided in blue.

The authors present an event-based analysis of flood drivers on a 344 km2 catchment using 50 years of concurrent daily rainfall and continuous streamflow data. The main conclusions are that heavy 4-day precipitation is the primary high flow driver and this, when combined with wet antecedent conditions, provides a stronger indication of flood likelihood than extreme daily precipitation alone.

Overall, I think the evidence presented provides reasonable support for the conclusions drawn, but this evidence could be strengthed and clarified. The authors have selected an interesting topic and a very worthwhile case study. I have three major comments and some minor ones, as detailed below.

"Major comments"

1. **The floods considered are based on a small number of factors (daily and 4-day precipitation occurrences occurring in the highest 1% and 5% of wet-day events, API, and various joint combinations). While in concept these are reasonable surrogates for the underlying flood processes of most relevance, it is a little surprising that no attempt appears to have been made to select factors of specific relevance to the catchment. For example, rather than adopt an arbitrary API, the decay factors of an API function could be fitted to the selected flood maxima and then used in the event-by-event analysis. Alternatively, a simple daily soil-moisture accounting function could be derived that implicitly allows for the influence of rainfall sequencing and evaporation; even without fitting to any observed data such an approach would appear to have greater efficacy than the adopted indicator of wetness.**

__**Answer:**__ Regarding the API and the exclusion of evaporative processes utilized for evaluating initial catchment conditions, the simple 30-day before an event effective rainfall (precipitation minus potential evaporation from the Maastricht station) was calculated instead of the simple API, and Fig. 5 from the manuscript was reproduced. Kindly see our reply to RC1.

Similarly, a simple correlation analysis could be used to justify the number of days adopted for the multi-day precipitation index, as at present no discussion is provided to justify the "critical" duration adopted. Such analyses would strengthen the physical reasoning used to assess the relative importance of the different flood drivers and may reveal greater insights about the nature of the interactions involved.

__**Answer:**__ The selection of the number of days (i.e. four) adopted for the multi-day precipitation (P_{MD}) index is partly discussed in lines 143-149. A four-day duration is selected considering the hydrological functioning of the catchment (L144). As we define extreme precipitation (P_{99}; 24-hour precipitation exceeding the 99$^{\mathrm{th}}$ percentile of rainy days) using two days (precipitation on the same or the previous day of the flood event) for P_{MD} we need more than 3 days. The 99$^{\mathrm{th}}$ threshold for the P_{99} is extracted only from wet days, while the 95$^{\mathrm{th}}$ for the P_{MD} is taken from all 4-day accumulations (rolling sum). This is done to ensure that the 95$^{\mathrm{th}}$ percentile of multi-day rainfall is larger than the 99$^{\mathrm{th}}$ percentile of single-day rainfall so as to make the distinction between P_{99} and P_{MD} clearer. This is achieved with longer accumulation durations for P_{MD}.

The sensitivity of P_{MD} to different precipitation durations is investigated in subsection 3.1.1. The mean relative probability of P_{MD} is highest at the 4-day accumulation period (confirming the usefulness of the selected period), while for the remaining indices, the results remain relatively stable regardless of the selected duration. As we are using the 95th percentile of all k-day accumulated (rolling sum) precipitation to define P_{MD} and we have "daily" values, usually this threshold is exceeded in prolonged events irrespective of the selected duration. This indicates that we have prolonged (multi-day) heavy events (larger than the 95th percentile of the selected k-day accumulations), however not so extreme as the 24-hour P_{99}, which helps us examine the relative contributions of extreme precipitation and prolonged heavy rainfall in generating high flows. We will consider including this analysis earlier in the revised version of the manuscript.

Overall, in the revised version of the manuscript, we aim to enhance the physical reasoning used to assess the relative importance of the employed drivers.

2. **Most of the analyses focus on the sample of events where it is known that conditions have resulted in floods. However, concentrating on the sample of 870 multi-day precipitation events (noted in Table 2) and examining the moderating factors which led to 50 annual maxima events should provide more insight about the processes leading to floods than does focusing on the much smaller sample of known flood maxima. For example, the analysis of these 870 events using similar diagnostics to that used in Fig 5 would make it clearer what combinations of factors lead to major flooding and which don't. It may be found that the combinations of conditions that are associated with floods may in some (or many) cases not lead to flooding, and this may highlight the influence of an additional factor that has not been considered. The "reverse" analysis described in the paper thus needs more focus and attention.**

Answer: We agree with the reviewer that it would be more thorough to expand the reverse analysis. This will be done in the revised version of the manuscript.

3. **The results are consistent with physical reasoning though in places I had to work quite hard to follow the logic of the narrative and the specific details of the results. It would thus be useful if the authors tightened up the narrative and provided additional discussion. For example:**

    1. **the information presented in Table 3 needs further explanation as the supporting discussion on this was not particularly helpful.**

Answer: Additional clarification will be incorporated in the revised version.

    2. **While the information presented in Figure 4 is broadly clear, I do not understand how the relative frequencies are calculated and why selected combinations of them don't add up to 100%.**

Answer: As reported in the figure's label, the relative frequencies are the "count of a driver leading to Q_{max} in the Q_{max} cases divided by the total number of cases". We simply count how many times a driver is observed in the total number of the Q_{max} events (49 cases). A single Q_{max} event can be preceded at the same time by more than one flood indicator, e.g. most of the P_{99} events that led to Q_{max} were also P_{MD} events (see Fig. 5 and also our answer to the following comment).

    3. **Fig 5 provides is a useful analysis as it differentiates between floods of different magnitude, yet it is not entirely clear what the different symbols are in Figure 5 denote - they appear to differ from the indicators listed in Table 1? It would perhaps be useful to examine such correlations for all selected indicators, allowing for timing lags as needed?**

Answer: Fig. 5 offers important insights by showing that wetter conditions (indicated by higher API values) are expected to have different effects on high flows, particularly in compound events, while it reveals the very weak correlation between Q_{max} and the event precipitation. These aspects were not as apparent in the analyses presented earlier (Fig. 3 and 4).

The symbol/marker of each Q_{max} event indicates its preceding indicator as defined in Table 1 and is connected with Fig. 4, however, it should be interpreted slightly differently. In Fig. 5, we can also see the specific drivers preceding Q_{max}, and subsequently the overlapping mentioned in the previous reply and partly the distinction between P_{99} and P_{MD} (that's why we present also the "only P_{99}" and "only P_{MD}" events).

The P_{99} markers (both in green and purple) in "Wet" classification (i.e., API > 1.5) indicate Compound I, the P_{MD} markers (both in orange and purple) in "Wet" classification indicate Compound II, and the P_{99} & P_{MD} purple markers in "Wet" classification indicate Compound III. All P_{MD} markers (both orange and purple markers), irrespective of their wetness (API), are classified as P_{MD} and thus used to calculate the relative frequencies of Q_{max} being preceded by P_{MD} in Fig.4b (e.g. 37/49 * 100% = 76 % relative frequency at Maastricht station), and so on. For example, the 1970 flood event in Fig. 5a at Maastricht is classified as P_{MD}, P_{WAC}, and Compound II. In this event, no P_{99} is observed. Or, for example, the 1987 event at Vaals (Fig. 5c) is classified only as P_{WAC}.

Correlations among all chosen indicators will be analyzed, and additional discussion will be included to provide further clarification on the meaning of Fig. 5.

4. **Further efforts should be made to strengthen the narrative thread throughout the paper as in many places I found myself going back and forth within the current and previous paragraphs to make sure I was following the intended logic. For example in Section 2.3.3 the discussion around the logic of the selected indicators commences before they are clearly defined two paragraphs later.**

Answer: We will work on improving the narrative flow throughout the paper and make it easier to follow.

"Minor comments"

- **Figure 2(b) – x-axis label is incorrect (it is not a rate, but rather the proportion of time that the given flows are exceeded)**
- **Line 219-220 – why is it the serial correlation of the precipitation time series assumed and not simply calculated?**
- **Line 249-250 – the justification for the last sentence of this paragraph is not clear**
- **Line 256 – should 86% be 83.7%?**
- **Line 397 – clearer justification is required for the 3rd sentence in this para regarding the cause for the rise in severe precipitation**
- **There are numerous small errors with the use of prepositions and other minor grammatical problems, and these should be reviewed and corrected.**

Answer: We thank the reviewer for the several minor comments. These will be addressed in a point-by-point response whilst preparing a revised version of our manuscript.