Reply to reviewers' comments

We thank the two reviewers for their thorough reading of the manuscript and their valuable remarks that helped us to improve the manuscript. In the following, the original reviewer comments are given in italic and all line numbers and figure numbers refer to the original submitted version that was reviewed if not mentioned otherwise.

Reply to review of reviewer 2

The authors contend that copulas add unnecessary uncertainty to the estimation of compound event frequency. They therefore apply a non-parametric randomization test to find the spatial distribution of compound high river discharge -storm surge events across Northern Europe. Similar randomization techniques have been applied in past e.g., Svensson and Jones (2002) and Zhang et al. (2013), but the scale of this analysis arguably provides some novity. The compound events that impact most of the rivers on the German-Danish west coast are shown to be linked by common metrological drivers as characterized by weather types defined by the German Weather Service. In my opinion, the preprint has the bones of a good paper but more work is required before it is worthy of publication in Natural Hazards and Earth System Science Journal.

We thank reviewer 2 for the overall positive evaluation of our manuscript and very detailed comments.

The introduction contains a wealth of relevant information; however, it does not flow particularly well. I believe a major cause of this is the poor internal structure of certain paragraphs. The first sentence should summarize the main point being in the rest of the paragraph. The paragraph starting on line 31 is a prime example where this is not the case. It starts by stating "Several studies conducted over the last years have shown the importance and catastrophic nature of compound flood events for several locations" but no catastrophic compound flood events are subsequently discussed. Instead studies that explore the correlation between the flooding drivers are listed. The paragraph ends describing results from Hendry et al. (2019), a relevant reference, since this study goes on to investigate potential correlations between a river's catchment size and the number of compound flood events, but out of place here.

We changed the language at several places to avoid a mix of British and American English (e.g., changed randomized to randomised). In addition, we went through the manuscript and corrected a few typos as well as revised a few sentences for better readability.

The following changes were made to the introduction, containing also changes based on the specific comments:

Coastal flooding is one of the most frequent, expensive, and fatal natural disasters. In the U.S. alone, it dealt $199 billion in flood damages from 1988 to 2017 according to Davenport et al. (2021). For Europe, Vousdoukas et al. (2018) projected an increase of annual costs caused by coastal floods of up to $1 trillion in 2100 for representative concentration pathway (RCP) RCP8.5. Furthermore, more than 600 million people live in coastal areas that are less than 10 meters above sea level and less than 100 kilometres from the shore (United Nations, 2017; McGranahan et al., 2007). Drivers for floods are storm surges, waves, tides, precipitation, and high river discharge (Paprotny et al., 2020). Additionally, floods can also be the result of failures of critical infrastructure like hydropower dams or flood defences (ECHO, 2021).

The IPCC Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX) defined compound events as “(1) two or more extreme events occurring simultaneously or successively, (2) combinations of extreme events with underlying conditions that amplify the impact of the events, or (3) combinations of events that are not themselves extremes but lead to an extreme event or impact when combined. The contributing events can be of similar (clustered multiple events) or different type(s)” (Seneviratne et al., 2012). A more general definition was proposed by Leonard et al. (2014), who defined it as "an extreme impact that depends on multiple statistically dependent variables or events". Compound flood events occur when large runoff from, e.g., heavy precipitation, leading to extreme river discharge, is combined with high sea level (storm surge).

The occurrence of extreme flood events either simultaneously or in close succession can lead to severe damage, which greatly exceeds the damage those events would cause separately (de Ruiter et al., 2020, Xue et al., 2022). Several studies conducted over the last years have shown the importance and catastrophic nature of compound flood events for several locations. Examples for these are compound inland floods of 2013 and 2016 in Germany (Thieken et al., 2022), the Siret Basin flood in northeastern Romania in 2010 (Romanescu et al., 2018), and the compound flooding in Houston-Galveston Bay during Hurricane Harvey (Valle-Levinson et al., 2020). Moreover, several studies have been conducted on a larger spatial scale in Europe, which cover North-western Europe (Ganguli and Merz, 2019; Ganguli et al., 2020), Norway (Poschlod et al., 2020), eastern Britain (Svensson and Jones, 2002), and entire Europe (Bevacqua et al., 2019; Paprotny et al., 2018b, 2020). Those studies investigated possible dependencies, mostly between surge and discharge. All of them found that the assumption of independence between drivers leads to a huge underestimation of the occurrence rate of compound events.

In addition to the large-scale studies mentioned above, a large number of studies exist that focus on smaller regions. Examples are the studies of van den Hurk et al. (2015) and Santos et al. (2021a), which both analysed a near flood event in the Netherlands in January 2012, which was caused by a combination of extreme weather conditions. Studies have been conducted worldwide, with examples being the Zengwen River basin in Taiwan by Chen and Liu (2014), Shoalhaven River in Australia by Kumbier et al. (2018), Fuzhou in China by Lian et al. (2013), Dickinson Bayou watershed in Texas by Kim et al. (2022), and other locations in various countries.

It is impossible to compare the results of different publications directly, since every study uses different approaches and measurements. There are currently no established standards for detecting extreme events. For example, the thresholds for extreme events were calculated by utilising the return period (Bevacqua et al., 2019), a certain number of events per year (Hendry et al., 2019; Ganguli et al., 2020), or utilizing a percentile approach (Paprotny et al., 2018a). Other studies chose block maxima to detect extreme events (Engeland et al., 2004). The exact parameters are chosen nearly arbitrarily by the authors, with the only common goal being a low number of events so that they can be declared as ‘extreme’. Nonetheless, there have been some studies that investigated the sensitivity of their results regarding different methods of identifying extreme events (Zheng et al., 2014) and changes to the model set-up (Jane et al., 2022). Additionally, there are studies like Ghanbari et al. (2021) that compare their results to those of other studies, e.g., to Ward et al. (2018). Basically, all studies found a correlation between drivers to a certain extent.

The influence of climate change on the frequency of compound flood events in Europe has been investigated by different studies. The increasing sea level due to climate change and higher occurrence of strong precipitation pose an increasing threat to important economic centres around the world and the people living there (Müller and Sacco, 2021). Feyen et al. (2020) projected that in case of a high emissions scenario, the damages caused by floods will represent a considerable proportion of some country's national gross domestic product (GDP) at the end of the century. Studies that investigated the effect of climate change on compound flood events focused on various regions of interest, for example, Bevacqua et al. (2019) on entire Europe, Poschlod et al. (2020) on Norway, Bermúdez et al. (2021) on the rivers Mandeo and Mendo in Spain, and Ganguli et al. (2020) on northwestern Europe. Bevacqua et al. (2019, 2020) reported a strong increase in the occurrence rate of compound flooding events for the future, especially for northern Europe, mainly due to the stronger precipitation as the result of a warmer atmosphere carrying more moisture. Contrarily, Ganguli et al. (2020) reported a lower risk of compound flooding due to a lower dependence between surges and river discharge peaks.

Many studies utilised copulas to describe the data distribution of two time series and investigate the dependence between extreme events. They were introduced in Sklar (1959). The Upper Tail Dependence Coefficient (UTDC) is used to estimate the probability of two extreme events occurring at the same time and to derive their level of dependence. It can be calculated, for example, with the Capéraá–Fougéres–Genest estimator (Capéraà et al., 2000). However, copulas introduce an additional amount of uncertainty (Heffernan and Tawn, 2004) and may require a large number of data points for robust tail dependence calculation (Moran, 1957). In climate research, the amount of available data points is often too small for this kind of analysis with many studies operating at merely 30 extreme events. Serinaldi et al. (2015) therefore concluded that those results are "highly questionable and should be carefully reconsidered". Using a thorough mathematical analysis, Frahm et al. (2005) showed that the calculation of the Upper Tail Dependence Coefficient is based on the assumption that an actual correlation between the datasets exists and therefore will always have a strong bias. This goes to the point that Joe (2014) stated that "the empirical measure of tail dependence for data doesn’t really exist". This raises the question of how to quantify the number of compound flood events and the resulting dependence between their drivers. Consequently, we chose to study compound flood events by using a methodology that does not utilise copulas.

In the present study, we analyse compound flood events by focusing on the question whether they occur more often than by pure coincidence. Utilising several available large-scale data sets allowed us to conduct this analysis for Northern and Central Europe, instead of focusing on a single river. Furthermore, we wanted to investigate if spatial patterns occur and if they are caused by one common meteorological driver. To achieve this, we implemented a simple statistical method that avoids the application of copulas. For this, we randomised our datasets and investigated if the resulting random distributions show comparable amounts of compound extreme events. Similar studies were so far only carried out by van den Hurk et al. (2015) for the Lauwersmeer in the Netherlands and by Poschlod et al. (2020) for Norway (in this case covering rain on snow events). To our knowledge, this will be the first recent publication investigating compound flooding in northern Europe without the use of copulas. For this, we utilised discharge and sea level datasets that were simulated based on reanalysis and hindcast data. Moreover, we investigated the robustness of the spatial patterns in our results by modifying various parameters of our method, like the thresholds for determining extreme events. Additionally, we investigated potential correlations between a river’s catchment size and the number of compound flood events that occur. Finally, we examined possible drivers that could cause the occurrence of compound flood events.

Furthermore, the reference to line 131:

For example, Hendry et al. (2019) found that the compound events on the west coast of Great Britain have a different meteorological background than those on the east coast.

Finally, paragraphs should never start with linking words such as "therefore" or "Next".

The following paragraphs were changed:

Line 71: In the present study, we analyse compound flood events by focusing ….

Line 247: To test the robustness for different time periods, we split the ECOSMO-coastDat3 and HD5-EOBS data into two 30 year sections, from 1960 to 1989 (Fig. 3b) and from 1990 to 2019 (Fig. 3c).

The term "potential compound flood events" needs to be defined. It appears the term "potential" is adopted in other studies as they only concentrate on the drivers of flooding and not the pathway or receptors component of the flooding process.

The idea was with this specific phrasing to take into account that the identified events might not be considered "extreme" due to local properties like topography and flood protection. We removed "potential" since the literature generally refers to them as compound flood events.
Additionally, we removed the sentence starting from line 28:
In the following text we will note "potential compound flood events" as "compound flood events" […].

The introduction (L62-64) contains some very strong statements regarding the suitability of copulas in climate research. These statements must be either backed up with references or toned down (see specific comments).

We added two additional references to line 62 and applied the suggested change from the specific comments:

However, copulas introduce an additional amount of uncertainty (Heffernan and Tawn, 2004) and may require a large number of data points for robust tail dependence calculation (Moran, 1957).

Furthermore, we toned down the statement in lines 63-64:

In climate research, the amount of available data points is often too small for this kind of analysis with many studies operating at merely 30 extreme events.

Please note that we provided further statements on the underlying assumptions like dependence (Frahm et al., 2005) and results (Serinaldi et al., 2015) already in the original text.

The discussion of the limitation of the method could be further developed. For example, imposing a separation criterion of four days means events may not be independent in some of Europe's longest rivers (I agree this problem is not unique to this study). Another limitation is that not fitting a parametric model prevents the estimation of useful engineering quantities such as design events.

We added the following lines to line 340:

Furthermore, the de-clustering time of 4 days might be too short for some of the longest rivers that may contain very long extreme events. The lack of a parametric model impedes the possibility of deriving engineering quantities for design events to test flood protection structures.

Certain Figures can be improved. The maps look distorted and not all rivers are shown.

It is correct that not all rivers are shown with their river network. Showing all rivers in a detailed way at this zoom level would make the figures harder to read. The relevant location are the rivers in the figures indicated by the large circles. We considered only rivers where discharge observation were available for the evaluation of the HD model discharge (see Hagemann and Stacke 2022)

Any locations including rivers, seas and oceans mentioned in the text need to be identified in the Figures.

As suggested, we added the following image to the end of the "Methods" section.

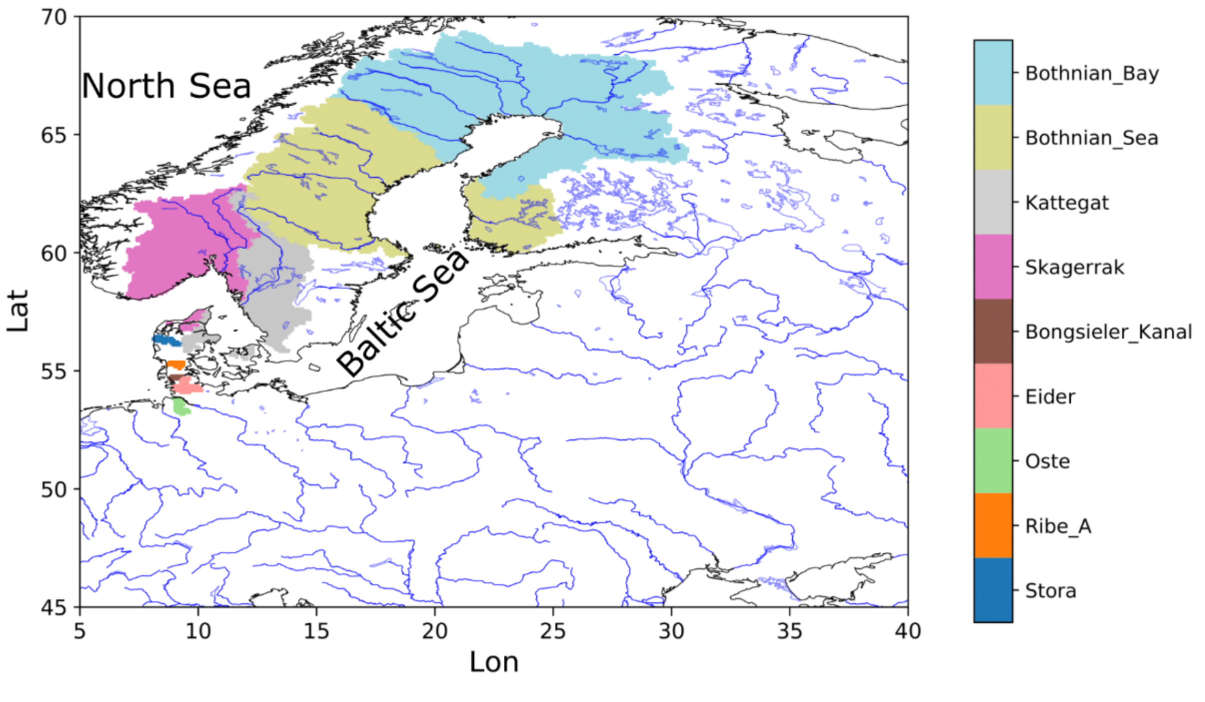


Figure 1. This figure contains the catchments, regions, and seas that are mentioned by name throughout the study. The first five entries in the colorbar contain maritime zones with highlighted catchment areas of rivers that discharge into them. The last five entries show the catchment area of five rivers on the German-Danish western coast.

We added the following after Line 157:

The domains of all catchments, regions, and seas that we mention by name for various reasons in this study can be seen in Fig.1.

The following lines were changed:

Line 230: The east and south facing coasts of the Bothnian Bay and Bothnian Sea in the Baltic Sea, as well as the eastern British coast, and Skagerrak show the lowest frequency of compound flood events.

Line 278: In the western facing coast of the Bothnian Sea, Cyclonic Westerly remained the predominant Großwetterlage.

A figure showing the proportion of each weather type responsible for the compound events at each location either as mini pie or bar charts or Figure 7 in Camus et al. (2022) would add value. This would provide evidence for statements such as "For Ireland, a distinct Großwetterlage could not be identified as a driver of compound flood events. We speculate that this might be because it offers a wide angle of attack for storm surges."

We added after Line 264:

The coordinates of those regions are available in Table 2.

Region	Coordinates
West coast of the Baltic states	54.52° N–59.00° N × 20.00° E–24.80° E
West coast of Great Britain	50.79° N–55.99° N × 4.85° W–2.50° W
German-Danish west coast	53.81° N–56.46° N × 8.02° E–9.12° E
Western facing coast in the Bothnian Sea	61.12° N–62.46° N × 21.18° E–21.80° E
West coast of Ireland	52.48° N–54.72° N × 9.30° W–7.90° W
Western facing coast of Sweden	55.37° N–59.37° N × 10.90° E–13.20° E

Table 2. Regions and their corresponding coordinates sorted in alphabetical order. They are used for the analysis in Sect. 4.3. These regions are also utilised in the visualisation of the results in Fig. 6.

Furthermore, we added the following image as requested:

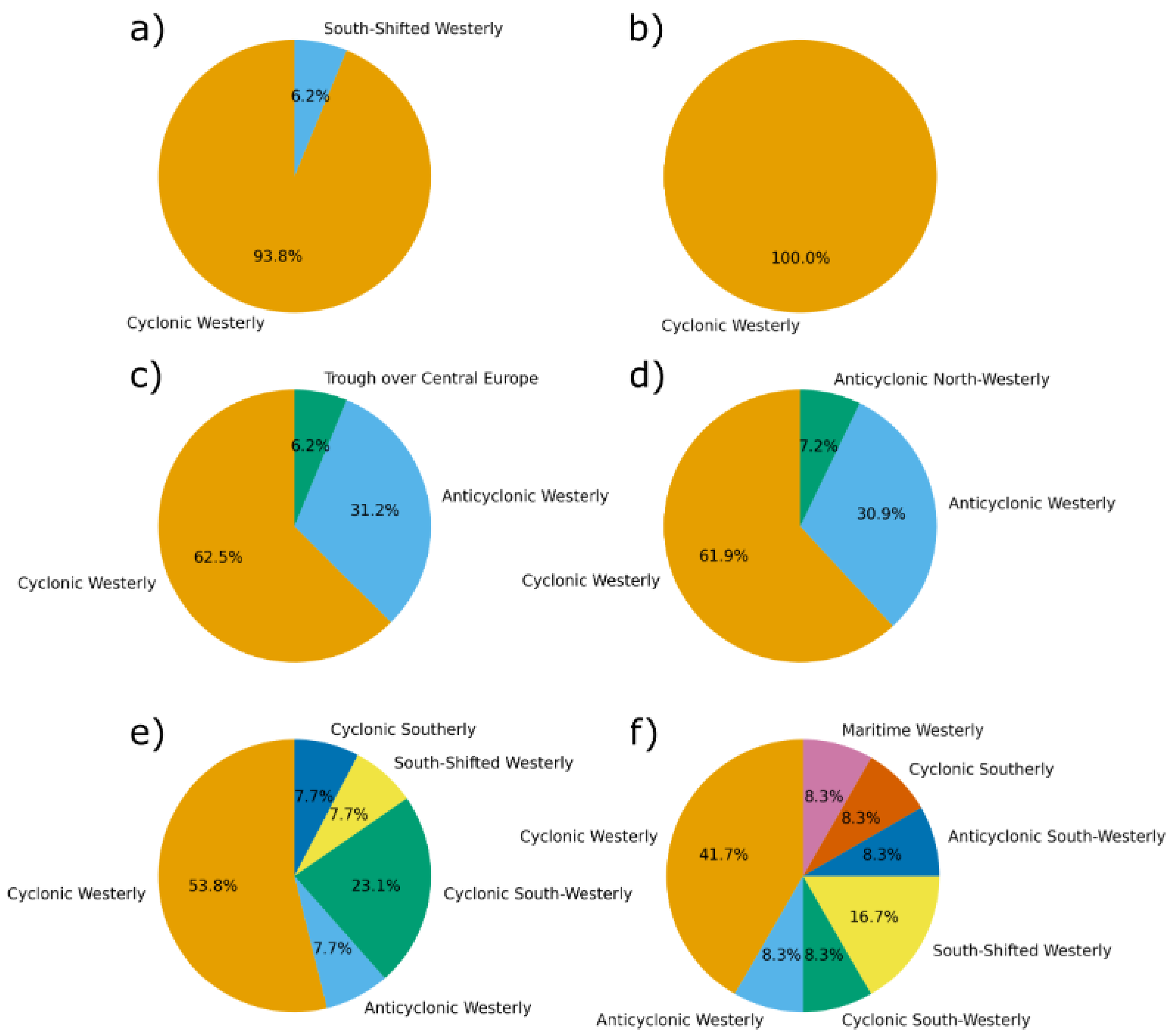


Figure 6. Distribution of Großwetterlagen that occurred during compound flood events in Europe. The following regions were analysed: a) German-Danish west coast b) Western facing coast of Sweden. c) Western facing coast in the Bothnian Sea d) West coast of the Baltic states e) West coast of Great Britain f) West coast of Ireland. Coordinates of those regions are given in Table 2.

We additionally inserted cross references to Fig. 6 where appropriate.

In my option the discussion section is very good. However, I suggest considering whether readability would be improved if the discussion and conclusion were disentangled. The conclusion would only need to be a few short sentences addressing the objectives set out at the end of the introduction.

We prefer to keep discussion and conclusion in one piece. We appreciate the suggestion and will consider it in the future.

* *L9 and elsewhere: Consider changing the text to make clear that "common driver" refers to a large scale metrological driver rather than a direct flooding driver like storm surge.*

Changes:
Line 9: Finally, we analyse if the observed compound extreme events had a common large-scale meteorological driver.

Line 17: ~~Common~~ Drivers for floods are storm surges, waves, tides, precipitation, and high river discharge (Paprotny et al., 2020).

Line 73: Furthermore, we wanted to investigate if spatial patterns occur and if they are caused by one common meteorological driver.

Line 130: Rivers with this behaviour might indicate a common large-scale driver that causes extreme discharge and sea level at the same time.

Line 262: A common meteorological driver for compound flood events

Line 263: To see if the regions with a higher than expected number of compound flood events have a common large-scale meteorological driver we analysed the meteorological situation during these events.

Line 266: Our goal was to scrutinise whether large-scale compound flood events in these rivers have a specific Großwetterlage as their common meteorological driver.

Line 320: We identified the Großwetterlage Cyclonic Westerly as the common meteorological driver for the occurrence of large-scale compound flood events in North and Baltic Sea regions.

* *L10: The "than expected" phrase is a little ambiguous. I believe you mean more potential compound events are expected compared to simple random chance but it could be interpreted as meaning more events expected based on similar previous studies.*

Changed Line 10:

The results of our investigation show that rivers along the western facing coasts of Europe experienced a higher amount of compound flood events than expected by pure chance.

* *L20: The sentence starting on this line requires a reference.*

We added the citation to line 20:

The occurrence of extreme flood events either simultaneously or in close succession can lead to severe damage, which greatly exceeds the damage those events would cause separately (de Ruiter et al., 2020, Xue et al., 2022).

* *L21: "Zscheischler et al. (2018) described in further detail why it is essential to consider compound events for risk assessment." This is not very insightful.*

Removed line 21 as suggested.

* *L27: "Potential compound flood events occur when large run-off from, e.g., heavy precipitation, leading to extreme river discharge, is combined with high sea level (storm surge)." So what type of compound event is this in terms of the Seneviratne et al., (2012) definition?*

As stated above we removed the term "potential" to avoid confusion. Therefore, the compound event described by us is type 1 in terms of the definition by Seneviratne et al. (2012).

L41: A lot of the listed studies focus on Asia/Oceana whereas much of the introduction concerns the U.S. and Europe. Kim et al. (2022) is a recent U.S. study that could be described here.

We added the reference as suggested:

Studies have been conducted worldwide, with examples being the Zengwen River basin in Taiwan by Chen and Liu (2014), Shoalhaven River in Australia by Kumbier et al. (2018), Fuzhou in China by Lian et al. (2013), Dickinson Bayou watershed in Texas by Kim et al. (2022), and other locations in various countries.

L43: I agree there is no established standards for detecting extreme events, however there are studies that compare the results obtained using different methods to identify extremes (e.g., Zheng et al. 2014) and changes in model set-up (e.g., Jane et al. 2022). There are also comparisons between findings from different studies (Ward et al. 2018; Ghanbari et al. 2021).

We added to the end of Line 48:

Nonetheless, there have been some studies that investigated the sensitivity of their results regarding different methods of identifying extreme events (Zheng et al., 2014) and changes to the model set-up (Jane et al., 2022). Additionally, there are studies like Ghanbari et al. (2021) that compare their results to those of other studies, e.g., to Ward et al. (2018).

L44: The items listed here are not equivalent. The percentile approach (inverse of return period) and events per year criteria are really two techniques amongst many for choosing a threshold in the peaks over threshold approach. Univariate extremes are typically identified using either a peak over threshold or block maxima approach. In terms of a bivariate sample, univariate extremes of a variable can be paired with concurrent or near concurrent values of a second variable in a procedure referred to as one-way conditional sampling (Moftakhari et al. 2019). In two-way conditional sampling, the procedure is repeated conditioning on the other variable producing two conditioned samples (Ward et al. 2018). Zheng et al. (2014) discusses the advantages and limitations of three methods of identifying bivariate extremes.

We thank the reviewer for their detailed information on bivariate sampling. We believe that this additional information would rather confuse the reader if added to the publication, since it is not used.

However, we modified Line 45:

For example, the thresholds for extreme events were calculated by utilising the return period (Bevacqua et al., 2019), a certain number of events per year (Hendry et al., 2019; Ganguli et al., 2020), or utilizing a percentile approach (Paprotny et al., 2018a). Other studies chose block maxima to detect extreme events (Engeland et al., 2004).

* *L51: Bevacqua et al. (2020) has some predictions regarding the high emissions scenario for Europe. Also, you need to state that flood damage amounting to a considerable proportion of some countries GPD by the end of the century is a prediction at this stage!*

We added the reference as suggested to Line 55:

Bevacqua et al. (2019, 2020) reported a strong increase in the occurrence rate of compound flooding events for the future, especially for northern Europe, mainly due to the stronger precipitation as the result of a warmer atmosphere carrying more moisture.

In addition, we modified line 51:

Feyen et al. (2020) projected that in case of a high emissions scenario, the damages caused by floods will represent a considerable proportion of some country's national gross domestic product (GDP) at the end of the century.

* *L53: Perhaps discuss the results of the Norway study and remove the sentence about different spatial scales as this feels a bit like repetition. Bermúdez et al. (2021) is a regional study which could be cited here.*

We believe that it is not target-oriented to discuss the results of the Norway study by Poschlod because it focuses on rain on snow and rain on saturated soil.

We added the reference to line 53 as requested:

Studies that investigated the effect of climate change on compound flood events focused on various regions of interest, for example, Bevacqua et al. (2019) on entire Europe, Poschlod et al. (2020) on Norway, Bermúdez et al. (2021) on the rivers Mandeo and Mendo in Spain, and Ganguli et al. (2020) on northwestern Europe.

* *L63: The text states that copulas "rely on a huge amount of data points", given that copulas are widely applied in climate studies the following sentence appears to be a contradiction where it states that "the amount of available data points is always insufficient for this kind of analysis with many studies operating at merely 30 extreme events". Changing the first sentence to state that a large number of data points may be required for the copula fit to be robust will remove the apparent contradiction. Jane et al. (2022) explores the sensitivity of copula family to sample size.*

The requested changes were made and can be seen in the general comments section above.

* *L85: A few automatic threshold techniques could be cited here.*

We added citations as requested:
There are ways to use automatic threshold approaches for detecting extreme events, like goodness of fit p-value (Solari et al., 2017) or the characteristics of extrapolated significant wave heights (Liang et al., 2019), but they struggle due to the diverse characteristics in the time series of drivers that cause coastal floods (Camus et al., 2021).

* *L86: I would be really interested to read studies where local properties, like flood protection or elevation of the surrounding area play a role in threshold selection.*

We agree but we are also not aware of such studies.

* *L94: Compound flooding studies such as Ward et al. (2018) where a percentile threshold is applied over a large spatial scale should be noted here.*

We added the reference as requested:
We, therefore, chose the peaks-over-threshold (Pickands III, 1975) method to select extreme events by using percentiles, like in the works of Rantanen et al. (2021), Fang et al. (2021), Lai et al. (2021), Ward et al. (2018), and Ridder et al. (2018).

* *L97: I suggest requirement for independent events be discussed before describing the variation in the number of (independent) extreme events with river characteristics.*

As suggested, we moved the paragraph starting in line 104 to line 95.

* *L106: The definition of separated is not clear here. Is it separated by four consecutive days containing no threshold exceedances?*

This is correct. We improved the wording to:
This means that two events are considered to be separate if the threshold is not exceeded for four consecutive days, like in Haigh et al. (2016).

* *L108-109: I do not understand the point being made in this very short paragraph. "In order to enable a good comparison between different rivers, the number of extreme discharge and sea level data points should be the same for all of them." I don't think the number of extreme discharge and sea level points necessarily need to be the same for a comparison to be valid. Furthermore, with this rational is there not a danger of choosing a threshold that permits events that are not extreme so a comparison can be made?*

To make our point clearer we modified the text as follows:

In order to enable a good comparison between different rivers, the number of extreme discharge and sea level data points should be the same for all of them. Then again, extreme events should be rare by definition, regardless of the river size, therefore only occurring scarcely throughout the year. This especially prevents the accidental analysis of events that are normally not considered as extreme. To test the influence of the extreme event definition on possible patterns, […].

* *L114: Please define more clearly which events have an "average return period of 0.5 years for extreme events".*

We changed line 114:

Moreover, the threshold tuning results in an average return period of 0.5 years for extreme discharge and sea level events since the return period can be defined as […].

* *L130: Previous studies which implement a similar procedure should be cited e.g., Svensson and Jones (2002), Zhang et al. (2013) and Nasr et al. (2021). Couasnon et al. (2020) presents a similar approach to the one proposed in this paper that uses the binomial distribution to assess the number of annual maxima storm surge – river discharge co-occurrences expected under the independence assumption.*

As suggested, we added the following to line 130:

Other studies in the past also utilised data permutation, see, for example, Svensson and Jones (2002), Zheng et al. (2013), and Nasr et al. (2021).

* *L136: Couasnon et al. (2020) speaks to this.*

We added the following to line 138:

This was similarly stated by Couasnon et al. (2020).

* *L132: "east coast of Great Britain exhibits a low number of compound flood events" This is repetition of the previous sentence.*

We removed the repetitive part from line 231 as suggested:
The east and south facing coasts of the Bothnian Bay, Bothnian Sea and the Skagerrak show the lowest frequency of compound flood events. Similarly, the east coast of Great Britain exhibits a low number of compound flood events, in contrast to the west coast.

* *L161: Water levels at the mouth of a river are often heavily influenced by sea level (e.g., Moftakhari et al. 2019), and therefore are not desirable in this context since they are not a reliable measure of the river discharge caused by inland rainfall.*

Our results are based on simulated discharges at the river mouth, not on observed discharges. The HD model does not consider sea level impact; therefore, it does not cause problems in our analysis.

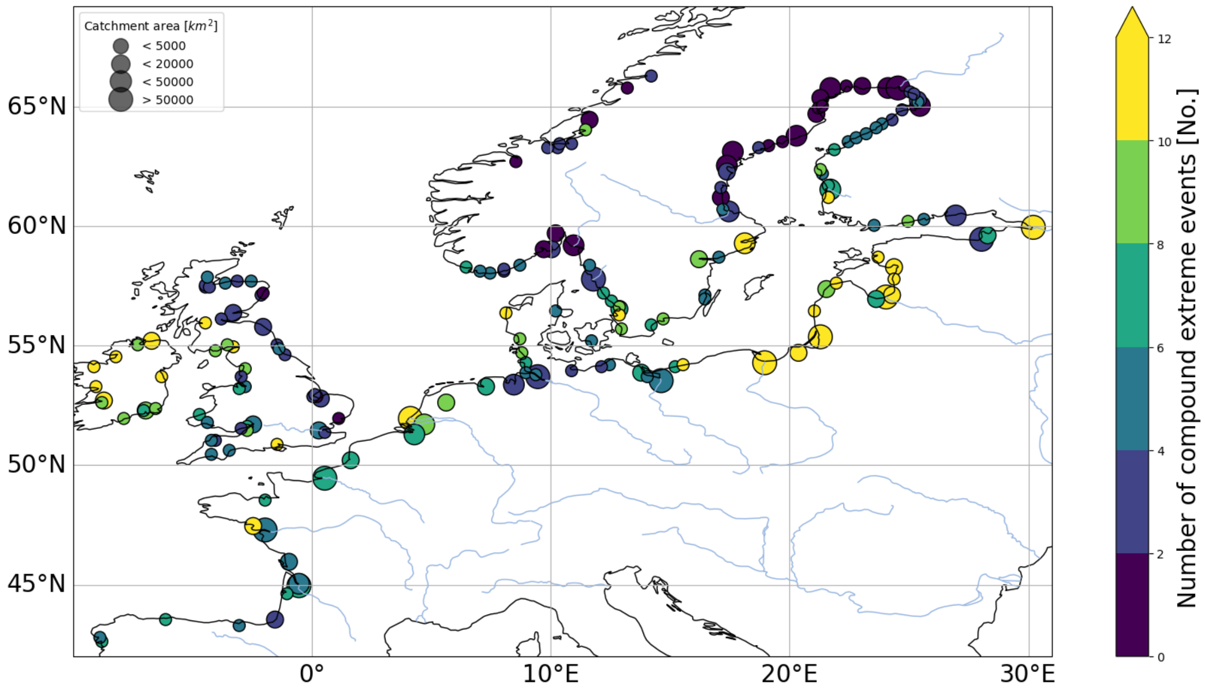
* *L228: What time period?*

We added information on the time period to line 228:

A total of 26% of the rivers along the coasts had eight or more compound flood events during the time period 1995-2018.

* *Figure 1: The choice of color bar makes this figure difficult to interpret. Consider using a cold to hot color scale to more easily identify hotspots (and cold spots).*

Changed colormap of Figure 1 to viridis:



* *L241: Could add "and are discussed in the following subsections" to the end of this paragraph to guide the reader.*

Added as suggested to line 241:

Some images for these tests are in Appendix A for the sake of readability, and they are discussed in the following subsections.

* *L248: What pattern are you referring to here?*

We modified line 248:

The pattern of western facing coasts having a higher number of compound flood events than expected by random sampling remained persistent throughout different time periods, even though it is somewhat more pronounced in the more recent one.

* *L260: Here and elsewhere it would be interesting to explore possible causes of these differences.*

We added the following information to line 329:

There was a certain amount of variation in the pattern, which can be attributed to randomness and the different setups. Due to the limited number of compound flood events, even small variations to their definition, like changes in the allowed lag, have a minor influence on the results.

* *L284: So what weather systems were responsible for the rest of the compound events in Great Britain?*

We added the following information to line 284:

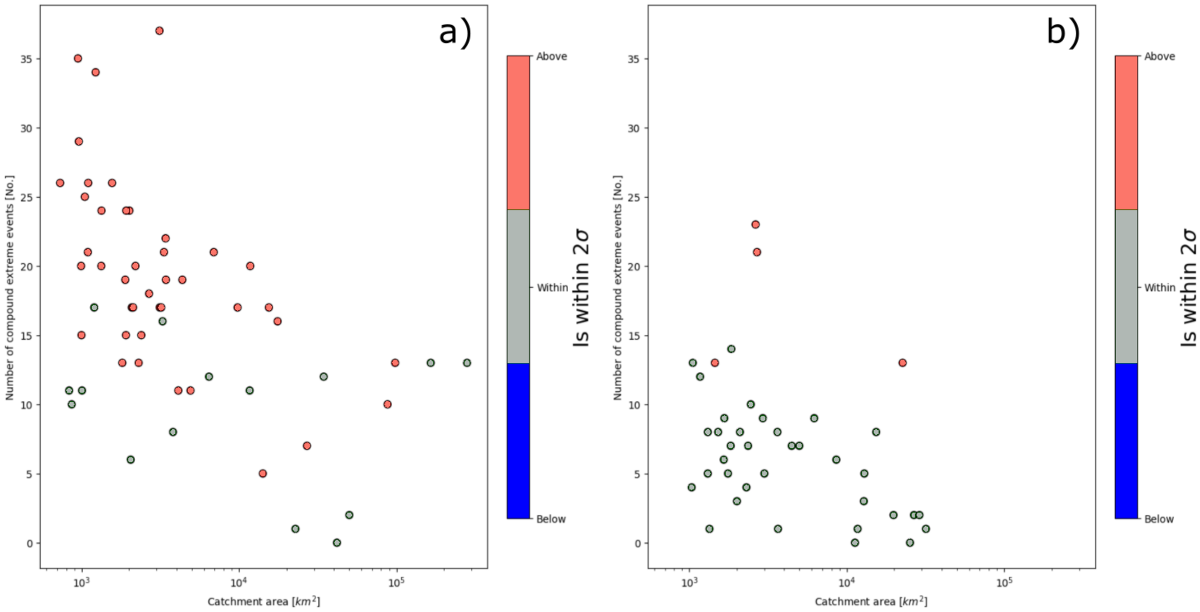
For the western facing coast of Great Britain, we found that half of the compound flood events happen during Cyclonic Westerly, a quarter of the events during Cyclonic South-Westerly and the remaining during other Großwetterlagen (Fig. 6e).

* *L298: No need to start a new paragraph here.*

Removed new paragraph as suggested.

* *Figure 5: Spacing between (a) and (b) needs to be increased.*

We increased the spacing between (a) and (b) as suggested.



* *L345: Check grammar of sentence starting "Additionally, …".*

Rewrote the sentence to:

Additionally, future studies could focus on locations in close spatial proximity along the western facing coasts for which long time series of daily sea level and discharge data are available.

* *L350: Remove "very interesting topic" this is subjective.*

We modified line 350 as suggested:

In order to support future risk assessments, it will be important to analyse how compound events will change under different climate scenarios and sea level rises (Zscheischler et al., 2018).

List of added references:

Bevacqua, E., Vousdoukas, M.I., Zappa, G. et al. (2020). More meteorological events that drive compound coastal flooding are projected under climate change. Communications Earth Environ 1, 47. https://doi.org/10.1038/s43247-020-00044-z

Engeland, K., Hisdal, H., and Frigessi, A.: Practical extreme value modelling of hydrological floods and droughts: a case study, Extremes, 7,5–30, https://doi.org/10.1007/s10687-004-4727-5, 2004

Ghanbari, M., Arabi, M., Kao, S. -C., Obeysekera, J., and Sweet, W. (2021). Climate change and changes in compound coastal-riverine flooding hazard along the U.S. coasts. Earth's Future 9: e2021EF002055. https://doi.org/10.1029/2021EF002055

Heffernan, J. E. and Tawn, J. A.: A conditional approach for multivariate extreme values (with discussion), Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66, 497–546, https://doi.org/10.1111/j.1467-9868.2004.02050.x, 2004.

Jane, R. Wahl, T., Santos, V.M., Misra, S.K., and White, K.D. (2022) Assessing the Potential for Compound Storm Surge and Extreme River Discharge Events at the Catchment Scale with Statistical Models: Sensitivity Analysis and Recommendations for Best Practice. Journal of Hydrologic Engineering, 27, 3, 04022001. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002154

Kim, H., Villarini, G., Jane, R., Wahl, T., Misra, S., and Michalek, A. (2022). On the generation of high-resolution probabilistic design events capturing the joint occurrence of rainfall and storm surge in coastal basins. International Journal of Climatology, 1– 11. https://doi.org/10.1002/joc.7825, 2022

Lai, Y., Li, J., Gu, X., Liu, C., and Chen, Y. D.: Global compound floods from precipitation and storm surge: hazards and the roles of cyclones, Journal of Climate, 34, 8319–8339, https://doi.org/10.1175/JCLI-D-21-0050.1, 2021

Liang, B., Shao, Z., Li, H., Shao, M., and Lee, D.: An automated threshold selection method based on the characteristic of extrapolated significant wave heights, Coastal Engineering, 144, 22–32, https://doi.org/10.1016/j.coastaleng.2018.12.001, 2019.

Moran, P.: The statistical treatment of flood flows, Eos, Transactions American Geophysical Union, 38, 519–523, https://doi.org/10.1029/TR038i004p00519, 1957.

Romanescu, G., Mihu-Pintilie, A., Stoleriu, C. C., Carboni, D., Paveluc, L. E., and Cimpianu, C. I.: A comparative analysis of ex-ceptional flood events in the context of heavy rains in the summer of 2010: Siret Basin (NE Romania) case study, Water, 10, 216, https://doi.org/10.3390/w10020216, 2018.

Ruiter, Marleen C. de, Anaïs Couasnon, Marc J.C. van den Homberg, James E. Daniell, Joel C. Gill, and Philip J. Ward. 2020. "Why We Can No Longer Ignore Consecutive Disasters." Earth's Future. John Wiley and Sons Inc. https://doi.org/10.1029/2019EF001425.

Solari, S., Egüen, M., Polo, M. J., and Losada, M. A.: Peaks Over Threshold (POT): A methodology for automatic threshold estimation using goodness of fit p-value, Water Resources Research, 53, 2833–2849, 2017.

Thieken, A. H., Samprogna Mohor, G., Kreibich, H., and Müller, M.: Compound inland flood events: different pathways, different impacts and different coping options, Natural Hazards and Earth System Sciences, 22, 165–185, https://doi.org/10.5194/nhess-22-165-2022, 2022.

Valle-Levinson, A., Olabarrieta, M., and Heilman, L.: Compound flooding in Houston-Galveston Bay during Hurricane Harvey, Science of the Total Environment, 747, 141 272, https://doi.org/10.1016/j.scitotenv.2020.141272, 2020.

Xu, H., Tian, Z., Sun, L., Ye, Q., Ragno, E., Bricker, J., Mao, G., Tan, J., Wang, J., Ke, Q., et al.: Compound flood impact of water level and rainfall during tropical cyclone periods in a coastal city: the case of Shanghai, Natural Hazards and Earth System Sciences, 22, 2347–2358, https://doi.org/10.5194/nhess-22-2347-2022, 2022

Zheng, F., Westra, S., Leonard, M., and Sisson, S. A. (2014). Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk. Water Resources Research, 50 (3), 2050–2071. https://doi.org/10.1002/2013WR014616