This manuscript by Lyddon et al. clearly sets out an approach to identify tidal and river flood hazard thresholds for an estuary to determine the dominant or compound drivers of flooding. This understanding is important for local hazard forecasting and response planning. The methodology to identify and explore the impact of compound flooding uses national monitoring sources and freely available numerical models allowing application to other estuaries. The results presented demonstrated the application using a flashy catchment, representing the most vulnerable estuary setting in the UK. The results show the spatial variability in inundation area in response to different flood driver combinations, while the discussion provides insights to improve hazard warning and investigate future uncertainty in such catchment types.

The authors use of impact information (e.g., bus cancellations) is an innovative way to develop a historic record of flood events in the absence of documented records. The figures are clearly presented, and all are thoroughly discussed in the text to illustrate key messages/findings. The contour plots Figs 6-8 are a clear way to illustrate compound hazard impact and compare results, which could be applied to different studies to explore multiple hazard drivers.

Thank you for these comments, and we are pleased that the application of the method and novel approaches have been recognised here. We provide responses to each comment below. The comments help to improve the clarity and readability of the figures, and the queries have provided an opportunity to consider the usability of the results for local authorities and identify future research questions.

Minor queries/comments:

The hazard thresholds use sea level and river discharge. I assume the elevation and discharge choice was to reflect the long-term monitoring available, so the results can be used by authorities in future assessments using available data. A question for consideration in the reply to reviewers is: would there be any value in using two elevation metrics in the analysis?

As the reviewer identifies, sea level (m) and river discharge (m^3/s) are used as flood hazard drivers in the analysis and numerical model due to availability of data. River discharge (m^3/s) is an appropriate metric for two reasons: (1) as a boundary condition for the model it ensures that the correct volume of freshwater is inputted (should there be any discrepancies between modelled and measured river cross-sectional area); and (2) discharge is a transferable metric to other systems and the outputs are therefore meaningful across catchments.

Elevation would also be a practical metric in the context of the river levels; local authorities could more easily assess and review rising river levels against a stage / river level threshold for flooding, as opposed to a river discharge.

In the abstract it should be clarified the time lag considered is 3hrs. Sensitivity to a range of time-lags is not considered.

Yes, this is a fair point, and the abstract has been edited:

L22: and sensitivity (~7%) due to a 3-hour time-lag between the drivers.

When mentioning locations in the introduction (e.g., Lancaster and Humber) they should be located as NW or NE England.

Yes, this is a good point, this has been added to the manuscript.

L110, clarify if the record length is determined by the monitoring duration or the start of this study (using available data when the simulations were performed).

The instrumental record length (from the tide gauge at Llandudno and river gauge at Llanwrst) is determined by the monitoring duration, up to the point when the analysis was performed. Instrumental data was only included in the analysis when paired TWL and Q observations were recorded at the gauge at the same time.

The following text has been added to the manuscript:

L120: The record length used in the analysis here is determined by the monitoring and modelling duration.

L130, 22:00 is pm the day before the peak river discharge. It's not clear if the flooding occurred on the falling max tide or the incoming tide.

The recorded flood event occurring on 22 November 1980, as described on Line 130, is an interesting event to explore in more detail. The Natural Resource Wales database of Recorded Flood Extents identifies the start date of this event as 22 November 1980, and an end date 23 November 1980. The Qmax is 433 m^3/s, and peak river discharge coincides with low tide on 22 November 1980, 05:00. The record of historic flooding describes this event as a fluvial flood, from the main river, caused by overtopping of defences. The location of flooding is described as 'Conwy Valley, Cae'r Groes, Llanrwst'.

The record does not provide enough detail to identify if the flooding happened on the falling tide or rising tide, but this may not be an important consideration in this event as it is considered a fluvial event.

The manuscript has been edited and the following text added to reflect that this is described as a fluvial event, and there is a lack of information to identify if compound flooding occurred as the exact timing of the flooding is not provided.

L130: Figures 1c and 1e show the 22 November 1980 flood event where Qmax was recorded as 428 m^3/s at 03:45 am. TWLmax was 4.5 m at 22:00 am (which included a 0.25 m skew surge) however lack of exact information on the timing of the flooding mean it is difficult to say if TWLmax contributed to flooding, and this is a compound flood. The NRW catalogue notes that there was widespread fluvial flooding in the Conwy Valley at this time, although since this was the pre-internet era there are no further online records.

Figure 2. Can Caesar-Lisflood provide information about the currents within the estuary to understand when slack tide occurs relative to the river flow? This would help understand the processes involved in holding the water within the estuary domain. The time lag is currently presented between Qmax and TWLmax. There could be asymmetries that prevent slack water occurring at TWLmax. It would be interesting to see the Qmax lag relative to slack tide (turning from flood to ebb). It would be worth checking the colour bar doesn't constrain the lag so it is relative to the TWLmax tide. The compound flooding could occur on the next tide (more than +6hrs might represent flood tide again). If there are events that occur more than +/- 6hrs either side of TWLmax can they be identified (two panels maybe)?

We did not explicitly simulate the events that are presented in Figure 2. The data presented in Figure 2 is an analysis of the instrumental data available from Llandudno tide gauge and Cwmlanerch river gauge and it is not possible to classify the lag time relative to slack water, as we don't know the time of slack water for these observed events. That is not captured in the instrumental records.

Events with a lag time greater than +/- 6 hours are not included in Figure 2; as the reviewer explains, it would be hard to assign a Qmax to a specific event.

Despite this, the comment is an interesting one and would be an interesting metric to consider when analysing compound events. Therefore we have added a comment to the discussion to reflect that the research could be developed to explore the lag time between Qmax lag relative to slack tide (turning from flood to ebb) using a 3D baroclinic model. Caesar LISFLOOD can provide information about flow velocity in fluvial and estuary settings however it is two dimensional (depth averaged) and may lack the process representation to fully resolve meaning a 3D hydrodynamic model may be better.

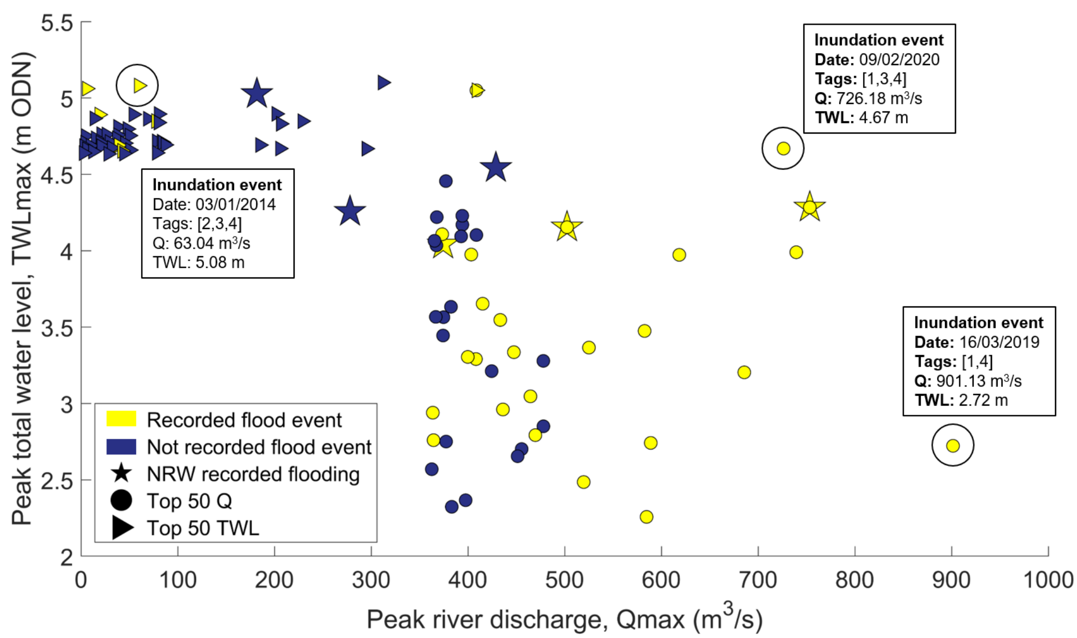
L683: The lag time is currently presented as between Qmax and TWLmax, however there could be asymmetries within the estuary that prevent tidal slack water occurring at TWLmax. The Qmax lag relative to slack tide (e.g. turning from flood to ebb) could be explored, however significant 3D lateral flows in the Conwy Estuary (e.g. Robins et al., 2012; Howlett et al., 2015) would mean that identifying location and timing of slack water would require a 3D baroclinic model.

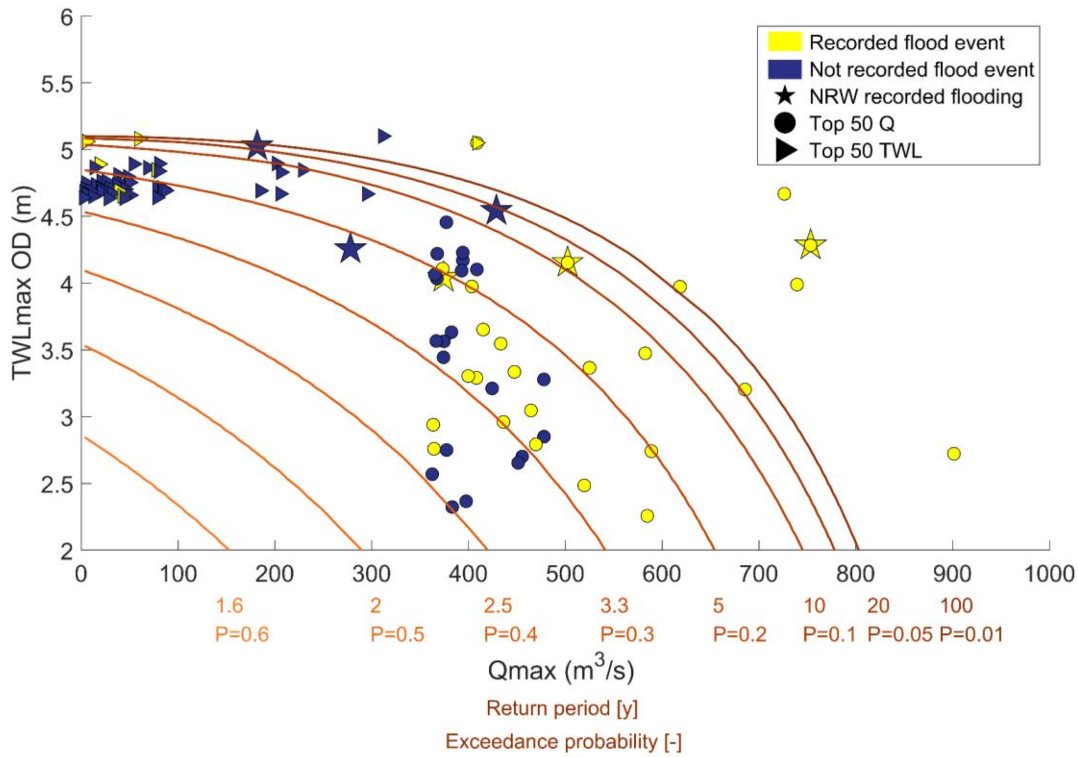
Howlett, E.R., Bowers, D.G., Malarkey, J., Jago, C.F. (2015) Stratification in the presence of an axial convergent front: Causes and implications. *Estuarine, Coastal and Shelf Science.* 161, 1–10.

Robins, P.E., Neill, S.P., Giménez, L. (2012) A numerical study of marine larval dispersal in the presence of an axial convergent front. *Estuarine, Coastal and Shelf Science*. 100, 172–185.

Figure 3 and 11. Using a different shape to explain the colour coding in legends (e.g., squares rather than circles) would help quickly separate this information from the circles used to indicate the top 50 Q.

Yes, this is a good suggestion. The legend in Figure 3 and 11 have been updated to include a panel of colour to represent recorded and not recorded flood event, as shown below:





L366, the flood extents aren't clearly assigned to the results in Figure 9. How were the three scenarios in each subplot selected? I assume it was to illustrate a similar FloodArea as defined on L365. In Figure 9 the distance up estuary could be marked (or the place names added) to show the locations of sensitivity in Figure 10. To help show how the icon boxes (Fig 9) link to the results the grid could be plotted on the TWLmax and Qmax axis, with the grid boxes labelled as tide dominant, river dominant, compounded and moderately compounded. Putting the grid onto Figure 11, would help show the probability of occurrence of the different flood drivers.

Thank you for highlighting this. The lateral flood extents in Figure 10 shows 12 model scenarios, with three selected to show sensitivity of FloodArea to small changes in Qmax and TWLmax for TWL, Q, moderate compound, and extreme compound scenarios. The three scenarios in each subplot were selected to be representative of FloodArea across the estuary for each type of scenario. For example, Figure 10a demonstrates that FloodArea is sensitive to Qmax when TWLmax remains the same; higher Qmax creates a greater FloodArea.

Different scenarios could be presented here (e.g. $Q^{3\text{-}4}TWL^{12}$), and the sensitivity illustrated in Figure 10 would have been the same. These scenarios were selected because they most clearly demonstrate the sensitivity of FloodArea to the drivers.

The four scenarios presented in Figure 9 are selected to represent FloodArea produced by different combinations of drivers.

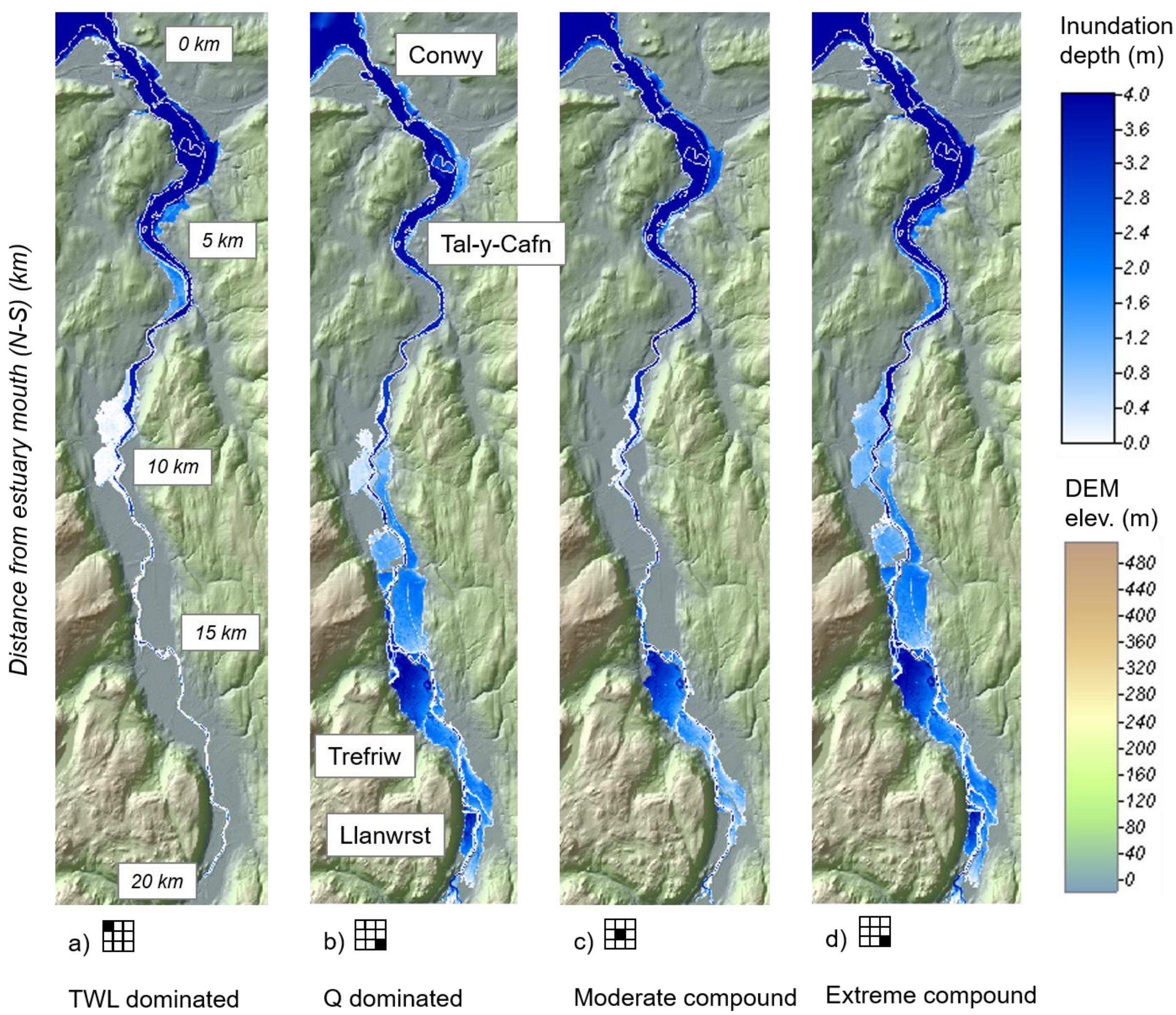
The following changes have been made to the manuscript:

L368: For each case (i-iv), three simulations were presented to illustrate sensitivity of FloodArea for scenarios with similar driver magnitude: (i) TWL dominated, 3.1 - 6.5 km^2, (ii) Q dominated, 11.13 - 11.8 km^2, (iii) moderate compound, 5.4 - 8.3 km^2, and (iv) extreme compound, 8.8 - 9.1 km^2.

Distance along the estuary (N-S) follows the channel, rather than as the crow flies. Therefore individual distance markers have been added to Figure 9a along the channel, rather than a N-S axis next to the figure.

Place names have been added to Figure 9b, that relate to the places named in Figure 10.

Plotting the grid on the TWLmax and Qmax axis, representing the complete 13 x 14 grid, makes the figure look busy. This is why we have kept the simplified 3 x 3 grid on Figure 9. The suggestion to label the grid boxes is a good one. This was stated in the Figure caption, but has now also been added to Figure 9. Please see the updated Figure below.



P20/21, check “FloodArea” is always in italics.

Checked and edited where necessary

Figure 11, should P=0.9 in the caption be 0.6 to match the contours plotted? For the joint probability it needs to be clarified that co-concurrence means Qmax must occur in +/-6hrs of TWLmax (for a flashy estuary), they do not have to occur at exactly the same time.

Thank you for highlighting this. Yes, the caption for Figure 11 has been edited to show that P=0.6 is the lowest probability displayed.

L684, have not.

This has been edited.

Author contributions, should CN be NC?

This has been edited.