Dear reviewer,

We would like to thank the reviewers for their detailed reviews of our manuscript and for the relevant points that they have raised.

In the context of our response, we have conducted some further analysis and addressed all the points that were raised by the reviewers. In response to the suggestions, the main changes in the revised version of the manuscript are:

1) We have included a literature review on the wave contribution to WLs.
2) We have added goodness-of-fit estimates for the model validation and modified table 3.
3) We have elaborated in more detail on the scenario choice and dependence analysis in the methods and discussion.
4) We have conducted additional scenarios.
5) We have updated the graphical representation of figure 5.

We believe that the manuscript has benefited substantially from the revisions and hope that the reviewers agree with the changes that we have implemented.

Please find below the answers to the individual comments and suggestions (in blue bold font). The answers to the comments are placed below the comments (in black font) and changes made in the manuscript are included in italic black font.

As the main focus/novelty is the contribution of waves to WLs, more overview of literature and their findings on that matter would be of benefit to a reader.

We agree that a literature overview on the contribution of waves to WLs further highlights the focus of the manuscript and provides relevant background information to the reader. We have therefore added a paragraph from line 52 in the introduction as follows:

"Waves can raise water levels at the coast in terms of wave set-up, which is described in detail by Dodet et al. (2019). Tanaka and Tinh (2008) have shown that in a shallow and narrow estuary entrance, wave set-up can be up to 14% of the offshore wave height. For South Africa, Marcos et al. (2019) have shown a dependency of extreme water levels and waves, and according to Melet et al. (2018) and Theron et al. (2010) waves constitute the most important components of coastal flooding for the country. Large destructive swells are generated by cold fronts, cut-off lows, and cyclones (Guastella and Rossouw, 2012). These low-pressure systems cause additional heavy rainfalls, leading to immense fluvial flash floods (Pyle and Jacobs, 2016; Molekwa, 2013). Thus, a dependency between both drivers is likely. However, no published regional to local compound flood probability analyses exist for South Africa…."

Also, a discussion and comparative analysis with the authors' findings would be of use in the discussion section.

We understand and agree that a more extensive discussion and a comparative analysis with other findings on the wave contribution to water levels during compound flooding would provide further understanding to the reader. Unfortunately, we find a results comparison difficult as there is no published literature quantifying compound flooding from waves, river discharge and tides in an estuary.

The separate analyses of RMSE for neap, spring tides and low and high water would be useful in the process of validation.

Thank you for this comment. Following your suggestion and the suggestion of the second reviewer, we have estimated the goodness-of-fit for WL peaks of each individual event used for model validation. The results of the additional statistical analysis are presented in the following table. We moved this table to Appendix C, as we included the additionally calculated goodness-of-fit estimates for the peak values and each event separately in Figure 3.

Table C1. Goodness-of-fit estimates of model validation runs compared to observations. Columns 2-4 show goodness-of-fit estimates for each tidal event of flood peaks only. Column 5 shows the goodness-of-fit for tide-only conditions (entire time series) and column 6 for tides including high river discharge (entire time series).

Goodness-of-fit	Average	Neap	Spring	Spring + high Q	Spring-, neap-, average tide	Spring-, neap-, average tide + high Q
RMSE	0.25	0.07	0.14	0.11	0.21 m	0.23 m
R^2	0.52	0.79	0.78	0.69	0.8	0.94
r	0.96	0.91	0.85	0.81	0.9	0.91

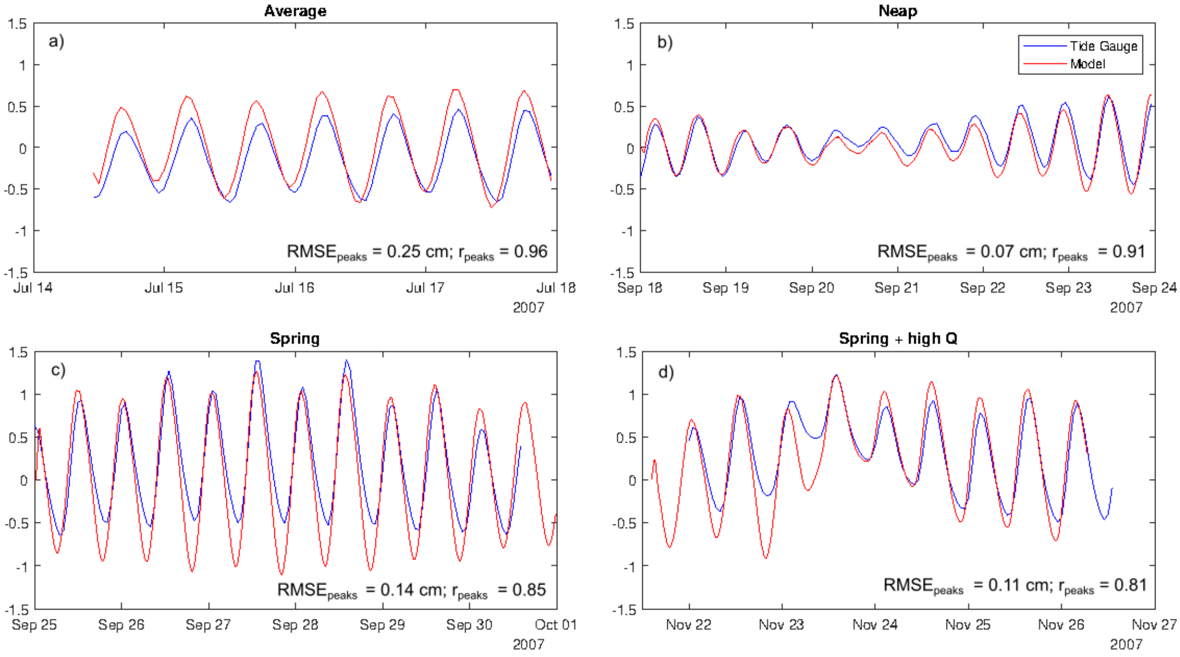

Figure 3. WLs of the model validation runs (red curve) at the tide gauge station, compared to observed WLs from the tide gauge (blue curve). (a) shows WLs of the average tide event, (b) the neap tide event, (c) the spring tide event, and (d) the high river discharge event, coinciding with the spring high tide. All panels include goodness-of-fit estimates for peak values of each event.

We hope that this analysis of RMSE for peak values brings more clarity regarding the source of the error and highlights model performance at peak values. Moreover, we would like to particularly thank the reviewer for this comment as, by recalculating the statistical measures, we detected an error in the previous goodness-of-fit estimates for the "Spring-, neap-, average tide + high Q" analysis. We included the corrected values in the table, which leads to large improvements in the model results. We also corrected the RMSE of all peak values, being now 0.15 cm.

The discrepancies could be due to the number of factors including spatial resolution of FES2014 tidal constituents and a number of constituents being simulated, as well as not including the non-tidal signals. Nesting or downscaling the boundary conditions from the regional model would be more appropriate, however I understand that due to the computational effort and the timescales required here this may not be possible.

Thank you for raising this interesting point. We must confirm that a nesting approach would require large computational effort and is therefore too challenging to realize within this short time period. However, we agree to add this point to the discussion. We have therefore added a sentence to the discussion in line 421:

"One way to overcome this limitation would be to downscale the tides from the model towards the location of the open boundary, which, however, is beyond the scope of this study."

Further, errors may arise from the digitized bathymetry, as Basson et al. (2017), who generated the bathymetry in 2017, could not provide detailed information. Moreover, the inner estuarine area is a highly dynamic area containing many sandbanks. These may have been in different positions in 2007 (the year we extracted the validation events) compared to 2017, affecting the tidal inflow and constituents. Thus, we added a paragraph to the discussion, elaborating on this limitation in line 424:

"Moreover, permanently opened estuaries are highly dynamic areas due to a constant influence of sediment deposition by river inflow and sediment removal due to floods (Moore et al., 2009; Whitfield et al., 2012). The sand bars and sand banks at the timing of the validation runs (covering events in 2007) were therefore likely in a different position than at the time when the input bathymetry was generated (Basson et al., 2017). This can have a high impact on water levels at the location of the tide gauge (Lanzoni and Seminara, 1998; Wang et al., 2019)."

Setting the north and south boundary as the no-flow wall boundary may play a role too.

We would like to point out that our validation shows that the model is generally underestimating what ebb tidal levels. Opening of the north and south boundary would lead to more water flowing out of the model domain. Therefore, we do not expect that open north and south downstream boundaries would improve model performance.

Also, it is not clear whether a 3D version of the model is used; the 3D-mode would be a better choice as the baroclinic conditions could be important in the region due to seasonal development of stratification.

We used the depth-averaged mode the model validation, as well as scenario runs for computational reasons. The depth-averaged mode has been successfully applied in other flood modelling studies in estuaries, using Delft3D, such as Kumbier et al. (2018) as well as other hydrodynamic models used by Skinner et al. (2015). In the method section, in which we describe the model set-up we point out in line 166:

"We performed simulations […] in Delft3D-FLOW […] in a depth averaged mode for the model validation, as well as scenario runs. The 2D mode has been successfully applied in numerous hydrodynamic flood modelling studies (Kumbier et al. 2018; Skinner et al. 2015; Olbert et al. 2017)."

In the results section, the four selected scenarios represent only a combination of the most extreme (100-year RP) conditions. As such, this is a very conservative approach and does not explore all flood conditions and their probabilities of occurrence.

We agree that accounting for extreme scenarios only does not explore a wide range of flood conditions, which would be a limitation when performing a risk assessment. However, we have decided on the extreme scenarios, for the following reasons: First, we follow recommendations of previous flood assessment studies for South Africa. Most have previously used 50-year return periods, which is now shifting towards using 100-year return periods (Theron and Rossouw 2008; Basson et al., 2017). For computational reasons we did not explore a wider range of scenarios, as lower wave return periods (20- and 50-year) do not significantly differ from the 100-year return level and higher return levels were not available. As our results clearly show that the estuary is Q-dominated, we did not consider higher return periods for river discharge. We added a sentence in section 3.3, line 246 to explain our choice of scenarios:

"Due to computational constraints and data limitations, we have employed the 100-year return period for waves and Q, as this was also recommended by previous flood assessment studies for South Africa (e.g. Theron and Rossouw, 2008; Basson et al., 2017)."

As we have shown that Breede Estuary is, even during compound flooding, a Q-dominated estuary, we developed an additional scenario with extreme wave (100-year, all directions) and moderate (20-year) Q conditions. However, the results did not significantly differ from the 100-year RP compound flood scenario, which is why we decided not to include the additional scenario in the manuscript.

We elaborate on the scenario decision in the discussion from line 433. Please find the additional paragraph in the comment below.

The readers would benefit from dependence analysis between Q and WLs…I appreciate however, that such statistical analyses would require a large effort and lead to a substantially different paper. I would suggest to discuss the interactions and dependencies in context of future work.

We strongly agree that dependency analysis between waves and Q would give an overview on the probability of co-occurring extreme conditions. However, we agree with the reviewer that focussing on a statistical dependency analysis would change the focus of the paper. We have indeed performed a preliminary dependence analysis between Q and waves by estimating the Kendall's Tau coefficient (see Ward et al., 2018; Couasnon et al., 2019). However, we decided not to present these results, since we used data from two different sources, as wave observations were not available. To assess dependence however, data used should come from the same source (Marcos et al., 2019). We used observed Q and modelled offshore waves from the CSIRO hindcast, produced from NCEP-NCAR reanalysis data (Cox and Swail, 2001). Moreover, we did not include a dependence analysis in the manuscript, since independence or dependence does not affect our results, showing driver interaction. It would only provide an overview on the likelihood of occurrence of our chosen results, which is relevant for a risk assessment. However, we do agree that discussing dependencies in the context of future work would be useful. We therefore added a paragraph to the discussion in line 433:

"In this study we focus on the effects of driver interactions on flood characteristics but have not considered dependence and joint probabilities of waves and Q. This information becomes relevant when assessing risk from compound flooding, which is beyond the scope of this study and should be considered in future work. For such a risk assessment, a wider range of return periods should be explored."

Figure 5, right panel. The colour scale could be refined to better reflect the differences. The interval of 0.3 is substantial in the flood context. The negative values in the range are not used.

We thank you for this advice. We removed the negative values from the legend and decreased the interval to 0.1, as shown in the figure below:

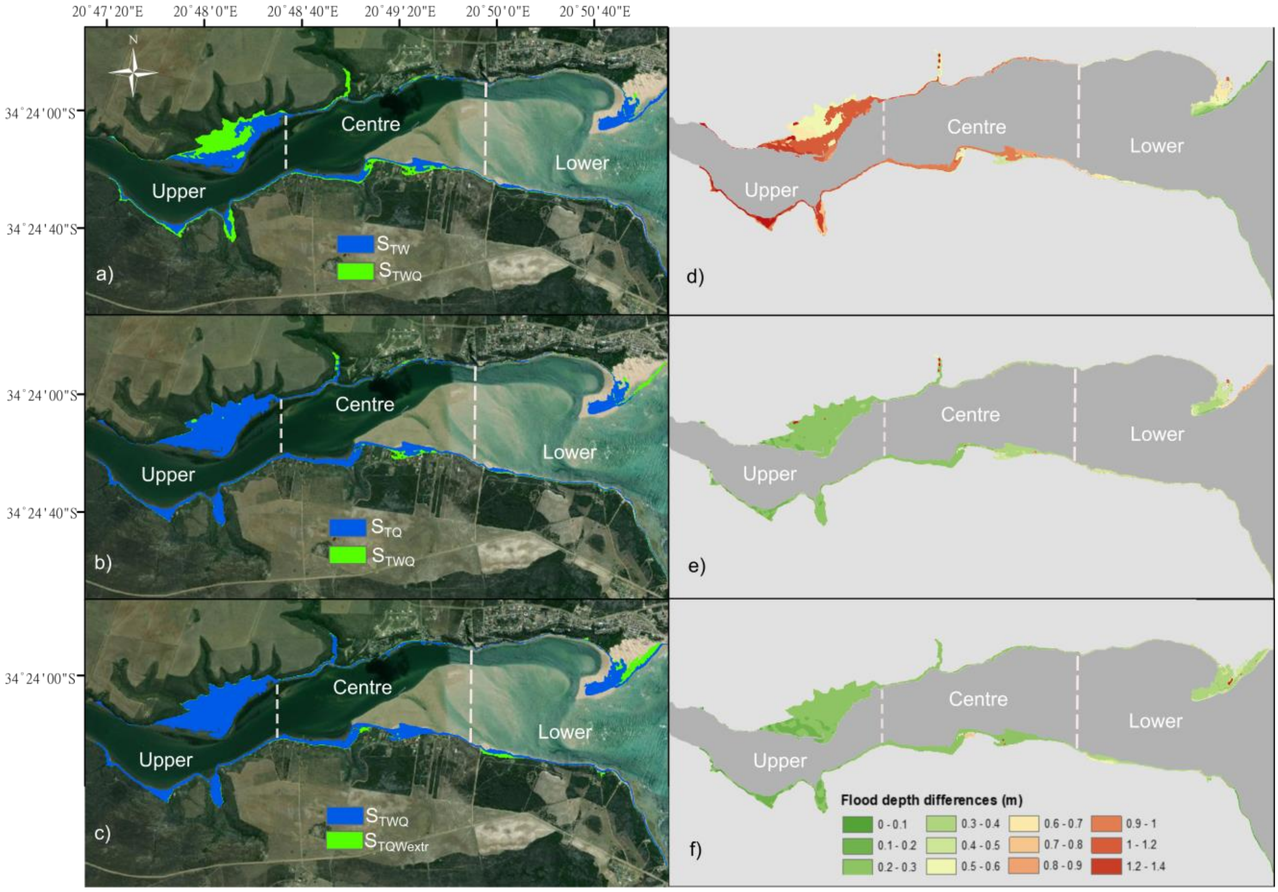

Figure 5. Comparison of flood extents of compound and excluding driver scenarios (left panel, a), b) and c)) and differences in flood depths (right panel). Panel d) shows the flood depths of S_{TWQ} - S_{TW}, e) shows S_{TWQ} - S_{TQ} and f) $S_{TQWextr}$ - S_{TWQ}.

Moreover, there is a number of typos (line 37- semicolon etc) which should be removed.

We have gone through the manuscript and have corrected all the errors we identified.

Additional references:

Cox, A. T., and Swail, V. R.: A global wave hindcast over the period 1958–1997: validation and climate assessment. *Journal of Geophysical Research*, **106**(C2), 2313–2329, https://doi.org/10.1029/2001JC000301, 2001

Marcos, M., Rohmer, J., Vousoukas, M.I., Mentaschi, L., Le Cozannet, G. and Amores A.: Increased Extreme Coastal Walter Levels Due to the Combined Action of Storm Surges and Wind Waves. Geophys Research Letters 46 (8), https://doi.org/10.1029/2019GL082599, 2019

Tanaka H., Tinh, N.X. and Nagabayashi H.: Wave setup at different river entrance morphologies, Coastal Engineering 2008 – 31st International Conference, http://doi.org/10.1142/9789814277426_0082, 2009