General comments:

The manuscript "Partitioning the uncertainty contributions of dependent offshore forcing conditions in the probabilistic assessment of future coastal flooding at a macrotidal site" by Rohmer et al. presents an interesting framework to assess uncertainties in future flood risk projections at a site in France, which undoubtedly falls under the scope of NHESS.

While it is clear that a lot of work has gone in this study and manuscript preparation, there are some improvements that are needed before the manuscript should be accepted for publication. These should include some clarifications with respect to the methodology, an update in the structure of the paper and some language improvements. The methodology and results are quite interesting, but due to the complexity of the paper, the current structure is not optimal. For example, Section 4 named "Application" has a mixture of methods and results. Subsections 4.1 and 4.2 could be incorporated in section 3. Sections 2 and 3 could be merged and reorganized as well, starting with a presentation of the study site and available data, then the multivariate analysis, then describing the hydrodynamic model, then the GP meta-model and so on. A nice diagram with the methodological framework and references to the respective sections could really help the reader. This could be placed at the end of the introduction sections or in the beginning of the methodological section. Moreover, the figures are a lot and some of them (Figures 9-13) could go in the Appendix without any problem with the flow of the paper.

While the dependency of the offshore parameters seems to be one of the main highlights of the paper, the presentation of this part and especially Figure 4, are not that clear. Some extra dependency indicators and improvements in the figure could help to clarify this (see my comments bellow for more specific information). Moreover, I would expect that the duration of an offshore event is an important parameter when assessing flood risk. However, the authors use a fixed duration of 20 minutes with uniform conditions. I would appreciate if the authors justify this choice.

Additionally, I feel that for such a complex methodological framework the discussion (including limitations sections) is rather short. Maybe section 5 could actually go in the discussion section (I feel like it can belong there since it discusses the assumptions used), especially if the figures that accompany it, are moved to the appendix. Other limitations and assumptions should be discussed as well; the way the dependency structure is modelled relatively to other methods available like e.g., copulas; the use of GP as a meta-model versus other statistical techniques etc.

In general, the language could improve as well, as I noticed there were some grammar issues and typos here and there.

Specific comments:

Line 54: Athanasiou et al. 2020 applied GSA as well for coastal erosion projections at the European scale

Athanasiou, P., van Dongeren, A., Giardino, A., Vousdoukas, M.I., Ranasinghe, R., Kwadijk, J., 2020. Uncertainties in projections of sandy beach erosion due to sea level rise: an analysis at the European scale. Sci. Rep. 10, 11895. https://doi.org/10.1038/s41598-020-68576-0

Line 271: How are the forcing conditions defined? Which is the time interval? Do you apply a peak over threshold to identify extremes? If yes, why don't you use the event duration as one of the offshore parameters, but rather assume the same duration for all events? I would expect that events with larger wave heights will have larger duration thus more flooding.

Figure 4: This an important figure and I think that some improvements are needed:

- First of all, while you mention in the text that the density difference is due to the threshold used, I don't see why this should be the case. I would expect the lower values of the pairs to

have higher densities, since lower wave conditions would be more frequent. It would generally help to plot histograms for each variable at the x and y axis.

- The grey dots are the simulated samples. Why don't you plot the observed data as well? This would be critical to see if the simulated samples follow the structure of the observed data.
- The yellow dots (training data) should be plotted on top of the simulated points, since now sometimes they are not visible. Additionally, the training data do not seem to sample well the parameter space. This should not be the case if the maximum dissimilarity algorithm (MDA) was used.
- Some correlation statistics (Pearson or tail dependency) would really provide some insights on the dependencies between the parameters, which is one of the main points of the paper. Along with point 2, it would be nice to compare these indicators between observed and simulated pairs.
- In the caption you refer to the next section. This type of referring to parts of the manuscript that come afterwards happens in various points. Consider positioning the figures at a point were all things presented in the figure have been discussed already.

Lines 294-296 and Line 299: The way you select the extra cases with high SWL is not clear. How are these cases defined? Are they based on simulated conditions using an offset for SWL? This should be clarified in the text.

Table 1: Here, you present the performance indicators for the meta-models you present in Section 5, so it is not clear what they are about yet. Additionally, while you use a 10-fold validation I see only one value? Is this the average of the performance indicators? Shouldn't the range be included as well?

Figure 5: See my previous comment. Do you present all validations in the 10-fold validation here? Sometimes Q^2 is presented in decimal and other on percentage, try to be consistent. It would be interesting to see some validation with the actual flood volume instead of the logarithm as well.

Figure 6: What does this figure show exactly? Is this for the median SLR projections of RCP4.5? Which stochastics procedures are included here? What do the uncertainty bands describe exactly? Is this the total uncertainty of the projections of flood risk? Then one could question why the decomposition of the uncertainty is important if the uncertainty itself is that small.

Figure 11: For some of the light blue bars the widths used are different than that of the white bars (H_s, T_p, D_p). I imagine this can change the count that is plotted?

Line 498: "By 2100, the threshold…", from the graph the contribution of RCP seems minimal, while the DEM one is even larger than in 2050, so I am not sure why you mention the RCPs here.

Technical corrections:

Line 33: "…, flood severity is…"

Line 63: "… and to probabilistic assessments…", to is not needed here

Figure 1: Caption needs to be rephrased. Consider having a general title and then describing the panels. Moreover, there are things in the figure that are not described like the star and point P.

Line 120: Here, there is a reference to Fig.3 which has a caption where the next section is referred (Section 3.1). It would make more sense to show the figure when everything about it has already been described.

Line 123: Consider clarifying in the figure caption that the 50 random series are a subset of a larger number or realization that have been used to get the actual confidence bounds.

Line 197: If I am not mistaken Q^2 is commonly referred to as skill score. Maybe use that phrasing?

Line 202: “Let use first focus on the presentation by considering Y to…”, what do the authors mean here?

Line 241: “In our study, the Shapley effect cannot be directly applied because the variable of interest Y is here not a scalar, but is binary and related to the flooding probability as defined in Eq.1”. I thought that *Y* is actually a scalar representing the water volume. Maybe rephrase?

Line 278: Here you express *u* as a probability while in eq. 2 it is a continues value if I am not mistaken.

Line 285: “conditions (if they were independent any structure would have been noticed), hence” the parenthesis does not make sense. Maybe you actually want to say the opposite?

Line 324-326: Consider mentioning already for which Y_c these values refer to.

Line 335: Here the authors jump to Fig.9 while the last figure was Fig.6.

Line 483: “The minor (with Y_C= 50m3), respectively very large (with Y_C=15,000m^3),…” this is not clear.

Figure 14: Why here there are no upper and lower bounds like in the other figures?